Toward Improved Geographic Information Services within a Digital Government

Organizing Committee

An organizing committee was formed to develop the agenda and work items for those who were invited to participate in this workshop. The committee also provided the National Center for Supercomputing Applications (NCSA) and Open GIS Consortium, Inc. (OGC) with advice and guidance regarding the findings, conclusions, and recommendations contained in this report.

Cochairs	*Organization*	*Email*
Louis Hecht	Open GIS Consortium, Inc.	lhecht@opengis.org
Barbara Kucera	National Computational Science Alliance, University of Illinois at Urbana-Champaign	bkucera@ncsa.uiuc.edu

Members	*Organization*	*Email*
Greg Buehler	Open GIS Consortium, Inc.	gbuehler@opengis.org
Max Egenhofer	University of Maine	max@spatial.maine.edu
Kenn Gardels	University of California–Berkeley	gardels@regis.berkeley.edu
Mike Goodchild	National Center for Geographic Information and Analysis, University of California-Santa Barbara	good@geog.ucsb.edu
Emil Horvath	U.S. Department of Agriculture	ehorvath@ftw.nrcs.usda.gov
Doug Johnston	University of Illinois at Urbana-Champaign	johnston@gis.uiuc.edu
Cliff Kottman	Open GIS Consortium, Inc.	ckottman@opengis.org
David Mark	SUNY–Buffalo	dmark@geog.buffalo.edu
John Moeller	Federal Geographic Data Committee	jmoeller@usgs.gov
Richard Muntz	University of California–Los Angeles	muntz@cs.ucla.edu
Hanan Samet	University of Maryland	hjs@umiacs.umd.edu

Toward Improved Geographic Information Services within a Digital Government

Report of the NSF Digital Government Initiative

Geographic Information Systems Workshop

Held July 1998

This material is based upon work supported in part by the National Science Foundation under Grant No. EIA-9818131. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

June 1999

Printed with soybean ink on recycled paper

Acronyms Used in This Report

API	application programming interface
BAA	Broad Area Announcement
CAP	Cooperative Agreement Program
CCIC	Committee on Computing, Information, and Communications
CRADA	Cooperative Research and Development Agreement
CS	Computer Science
DGI	Digital Government Initiative
EOSDIS	Earth Observing System Data and Information System
ESRI	Environmental Systems Research Institute, Inc.
FGDC	Federal Geographic Data Committee
GI	Geographic Information
GIS	Geographic Information Systems
GPRA	Government Performance Results Act
HUD	Housing and Urban Development
I/O	input-output
ISO	International Standards Organization
IT	Information Technology
MSA/PMSA	Metropolitan Statistical Area/Primary Metropolitan Statistical Area
NASA	National Aeronautics and Space Administration
NCSA	National Center for Supercomputing Applications
NIMA	National Imagery and Mapping Agency
NPR	National Performance Review
NSDI	National Spatial Data Infrastructure
NSF	National Science Foundation
OCG	Open GIS Consortium, Inc.
SQL	Structured Query Language
TIGER	Topologically Integrated Geographic Encoding and Referencing
UCGIS	University Consortium for Geographic Information Science
USCB	United States Census Bureau
USGS	United States Geological Survey

Table of Contents

Executive Summary

The world of geographical spatial data was once the exclusive domain of the mapping community. Color-coded maps, which summarized a wide variety of information in the context of geographic location, were updated every few years. Today, however, we live in a world described by data generated at rates of terabytes per day—image and sensor data from satellites, ground sensors, field studies, surveys, and census activities. Through huge federal programs these data are collected by planning offices in small towns as well as by government research and regulatory agencies, academic research programs, and the private sector. Time has become an important variable as we begin to understand the rapid rate at which our world is changing.

The purpose behind the large investment by federal, state, and local governments in geographical and geospatial data is increasingly clear. Through the past 50 years' revolution in information and computing technology, transportation, and engineering, rapid change can be initiated by public or private interests and can have a significant impact on national and global politics, economics, and the environment. Change needs to be understood and managed effectively for the nation and world to accommodate growth and sustain prosperity.

The worlds of computer and information sciences, statistics, geography, ecology, and cognitive science all converge in Geographical Information Systems (GIS) to provide glimpses into the world we live in. The ultimate application of any particular GIS view of the world can be as diverse as building a dam, planning a city, combating terrorism, or charting the destruction of a rain forest. While specific applications of GIS are diverse, the essential need is to share data across sources. In effect, any data from any source could be required for any GIS application regardless of whether the analysis is sponsored by government agency or private interest.

To address the need for cooperation across government, academia, and industry as well as among research disciplines, an invitational workshop was held to begin defining a research agenda for geographic information systems and technologies. This workshop was held in response to the Digital Government Initiative (DGI) of the Federal Information Services and Applications Council.

The recommendations from the workshop that are summarized in this report strongly support the need for cooperation in setting the national GIS research agenda and in linking innovative research with practice. The final recommendation specifically addresses the need for a national prototyping center that would not only bring together research results and experts but also provide a natural link among the communities focused on research, practice, and policy. All three of these communities are necessary to successfully define, develop, and fund a national GIS agenda.

Objectives

The Digital Government Program is managed by the National Science Foundation and supports projects that innovatively, effectively, and broadly address—through research—the potential improvement of agency, interagency, and intergovernmental operations as well as citizen-government interactions. Such research is expected to enable the creation and use of a continuous stream of advanced information technologies for early adoption and integration into the federal information systems community.

The GIS workshop was held July 8–9, 1998, at the Hilton Hotel in Arlington, Virginia. Objectives were:

- To identify research and demonstration needs
- To develop a common research agenda
- To build cooperation among the government, academic, and industrial communities that have research interests in geographic/spatial information technology, data, and services

Consequently this workshop began a process for shaping research that can be funded under the digital government program. The resulting research agenda—presented in this report—addresses the factors that, along with technology itself, *shape* the ability of government agencies to adopt advanced geospatial technology and information to deliver effective services to citizen, business, and government customers.

Recommendations

Recommendation 1. Advance research efforts directed toward the study of optimizing geographic query mechanisms and incorporating geometry and spatial relational operations.

Recommendation 2. Develop improved mechanisms for storing and representing time-varying geospatial data.

Recommendation 3. Support research on integrating spatial data fusion from multiple agencies, distributed data, and multiple collection devices.

Recommendation 4. Support research on multiple representations/interfaces focused on task-specific (procedural) workflow classes.

Recommendation 5. Support research in developing algorithms for knowledge discovery applied to very large, frequently updated, spatial datasets such as those derived from space-borne Earth-monitoring sensors.

Recommendation 6. Support research in the theory and methods of representing data with varying degrees of exactness and reliability.

Recommendation 7. Support research in the context of decision-making to improve the representation of diverse data and the dynamics of geographic phenomena.

Recommendation 8. Extend the promise of cognitive research to make geographic information technologies more accessible to inexperienced and disadvantaged users and also examine how government information policies affect access to and use of geospatial data for a broad spectrum of public and private sector stakeholders.

Recommendation 9. Support research to examine commerce's issues in geospatial information such as preserving privacy despite geographic locators and breaking potential bottlenecks in distributing geographic information services due to GIS's unique workflow processes.

Recommendation 10. Develop a Geospatial Digital Government Prototyping Center to create a network for testing and developing processes consistent with U.S. priorities for geographic information technologies and services in the government workplace.

Background

Numerous initiatives in the federal government are focused on improving both the delivery of services to the people and the population's ability to be aware of and participate in government decision-making. Vice President Gore's report on reinventing government has led to a great deal of consideration and rethinking of the means by which we can have "better government at less cost" (Gore 1993). At the same time, rapid advances in computing and communications suggest that information technology will provide a major vehicle for reaching the vice president's goals. This convergence led to the publication of *Towards the Digital Government of the 21st Century*, which lays out the following vision for employing information technology as an enabling mechanism to improve the delivery and execution of government services (Schorr and Salvatore 1997, 1–2):

> The disconnect between providers of private and public information services and the information and computer technologies research community is of great concern. Because information plays many roles in Government agencies (as a currency of exchange, as an important by-product of their missions, as the carrier of work flows and services), it is essential that as agencies restructure and reengineer to exploit these new efficiencies, the R&D community and other sectors form new relationships.
>
> The new network-centric world will present other societal challenges beyond the pragmatic concerns of efficient and affordable Government information services. . . . A Digital Government will allow public access to Government information and services, and group participation in discussions at any time and from anywhere on the globe with the required security and trust. . . . Realizing the Digital Government presents great technical and intellectual research challenges, but also promises great value and can provide valuable new insights and interesting new applied research problems, leading to deployable new systems. . . . Research topics with a special focus on the unique characteristics of the Federal information services domain might include the following:
>
> - Storing and Archiving Information
> - Finding and Accessing Information
> - Integrating Information from Multiple Sources
> - Mining and Knowledge Discovery
> - Universal Access to Information
> - Validating and Visualizing Information
> - Security, Privacy, and Electronic Commerce

Digital Government Vision

This workshop, part of a parallel series of workshops sponsored by NSF, was intended to define both research requirements for the geospatial information science community and ways to provide digital government services. The workshop was also a forum for advancing relationships among the GIS community, the federal agencies that increasingly rely on geographic information, and the private sector that is bringing standards and services to users. As federal agencies create strategic visions and operational philosophies for information technologies, they pursue the vision set forth in the digital government program announcement (NSF 1998, 1):

The Federal government is a major user of information technologies, a collector and maintainer of very large datasets, and a provider of critical and often unique information services to individuals, states, businesses, and other customers. . . . As society relies more and more on network technologies, it is essential that the Federal Government make the most effective use of these improvements. . . . Within this context, the objective of the Digital Government Program is to support projects that innovatively, effectively, and broadly address through research the potential improvement of agency, interagency, and intergovernmental operations and/or government/citizen interaction.

Special Case of Geographical Information

Geographic Information (GI) creates unique challenges and opportunities for realizing the vision of a digital government. GI—information explicitly linked to a location relative to the Earth's surface—has become nearly ubiquitous in many industries and government agencies. For instance, GI is a foundation of the U.S. Census through the TIGER system. The U.S. Geological Survey is a major provider of GI regarding the earth's physical systems. The Department of Defense relies heavily on accurate and frequent surveys of terrain, land use, weather, and other factors. NASA's air- and space-borne imaging systems record aspects of the Earth at frequent intervals. Other governmental agencies use GI for work such as emergency relief, crime analysis, natural resource management, and regulatory missions. Industry uses GI for such tasks as reducing costs of transport systems, improving the delivery of services, and managing natural resources.

Geographically referenced data are widely sought from both private and government providers. However, the diversity of data and user requests constrain the government's ability to provide access to GI-based data and services. While geographic data may be viewed as a subset that shares many of the general problems of information science and technology, several *unique characteristics* make providing geospatial services especially challenging.

Among the particular challenges associated with GI are the following as identified by the University Consortium for Geographic Information Science (UCGIS 1996, Introduction):

- Geographic information is rich and voluminous. Given that Earth's surface is infinitely complex, geographic information must always be an approximation.
- Geographic information is increasingly essential to many activities of modern society. The growth of international trade and globalization of economies require unprecedented knowledge of the diverse conditions in different parts of the planet. Geographic information is essential to understanding Earth's physical system as well as interrelationships among its components.
- Geographic information science is inherently multidisciplinary. No existing or traditional discipline can claim a singular role in solving problems of handling geographic information. Indeed, research into these issues has traditionally been divided among a number of disciplines that have often competed among themselves for available resources.
- Growth of geographic information technologies has already had profound and often unanticipated effects on society. For example, the GIS-enabled ability to link together digital street maps and telephone directories means that a house's telephone number can be identified by pointing to an image of the house on a computer screen. Marketing campaigns can now target the imputed socioeconomic status of each household. Such possibilities are the simple result of technological improvements, but implications for individual privacy are much more profound.

Other recent events involving the federal administration further point out the growing demand for geographic information and services. Vice President Al Gore has called for initiatives in the physical and social sciences that explicitly include geographic information sciences as a necessary component.

In the January 1998 speech announcing his plan for a Digital Earth, Vice President Gore called for advances in computational science, including high-performance computing for geospatial modeling and simulation as well as improved mass storage to accommodate the terabytes-per-day collection of GI from high-resolution remote sensing platforms, broadband networks, interoperability, metadata, data fusion, and other activities (Gore January 1998).

In an address at the Brookings Institution in Washington, DC, Vice President Gore called for stronger efforts nationwide to enhance the livability and economic competitiveness of American communities. He said the administration would expand its support for the use of GI systems technologies and would encourage increased public access to and sharing of geographic data to put "more control, more information, more decision-making power into the hands of families, communities, and regions" (Gore September 1998).

Initiatives among many federal programs (including but not limited to those undertaken by the Bureau of the Census, Department of Defense, Geological Survey, and Department of Interior) have geospatial data at the core of their information technology missions. Research and development in areas pertaining to the creation, analysis, and delivery of geographic data are important components of the Digital Government Initiative.

Perhaps because of the historically interdisciplinary development of geographic information technologies, there is a growing need to establish research and development priorities and to increase funding support of the basic underlying information science and engineering. These priorities must explicitly build a bridge between the communities developing methods, data, analyses, and applications and those communities developing the computer science foundations for storage, query, display, modeling, analysis, and other components of geographic information.

Participants' Perspectives

Approximately 50 people representing 38 academic, federal, and private sector organizations attended this invitational workshop. The complete list of participants is on page 28. The workshop's agenda is reproduced on the inside back cover of this report.

Participants were asked to review the following research topics from the perspectives of computer science and geographic information science and technology:

- Information storage, archiving, search and access
- Integration of information from multiple sources
- Knowledge mining and discovery
- Information validation and visualization
- Universal access to information
- Electronic commerce (security and trust)

Government Needs

The two-day workshop opened with a panel discussion of different views of digital government, much of which is included in the Backgrounds section of this report.

NSF's digital government report stresses that all aspects of government need to focus on information service technologies (Schorr and Stolfo 1997, 2–3):

> Research and development in computer and information science and technology, as well as research in statistics and the social sciences, is funded by many Federal agencies, universities, and businesses. What is necessary now is to define a new range of cross-disciplinary applied R&D objectives with the needs of the Federal information services community in mind, applying that R&D to the interesting problems and data within that arena.
>
> Although many Federal agencies understand the potential of an integrated and interoperable Internet and intranet approach to support their missions, most Federal information still resides in legacy vertically-integrated systems designed for specific applications, datasets, and mission areas. These systems have been maintained in stovepipe fashion and have grown individual cultures in the service community without attracting the interest of the R&D community. This disconnect between information services and R&D is by no means limited to the Federal sector, but is mirrored in the private sector and causes concerns there also. . . . To effect this move from vertically integrated legacy systems toward multidimensional integrated information sets is a significant research problem with very high potential payoff; it represents a National Challenge for Federal Information Systems.

Breakout Sessions

Getting input from government agencies, the private sector, and research communities was a significant aim of the workshop. To that end, participants broke out into groups that contained representatives from each type of commuity. The groups were to consider issues and priorities in accordance with topics identified in the digital government solicitation, to define research needs, and to explore opportunities for collaboration.

The breakout groups were specifically asked to consider the following points and how they relate to each of the digital government topics:

- Define a topic area in terms of contributions by geographic information and computer science.
- Identify impediments to working together such as interagency cooperation and joint research agendas.
- Identify unique aspects of government services that are dependent on geospatial information.
- Define a geographic information (GI) *services* vision for the year 2005.
- Define a computer science (CS) *technologies* vision for the year 2005.
- Discuss potentials for adding GI services in nontraditional agencies.
- Consider societal benefits that could result from maintaining data and system redundancies across government.
- Identify educational requirements needed to adjust for CS-centric GI.
- Discuss how Government Performance Results Act (GPRA) requirements and sound business approaches should be considered.

Workshop participants focused more on identifying unfulfilled research needs in the various dimensions of GIS than on addressing the various items listed above. Although many of these points may need to be considered, they were outside the scope of the workshop given the available time.

Notes from the breakout groups are provided in the Appendix.

Federal Agency Perspectives

Representatives from the federal agencies participating in the workshop discussed many of the basic research items and programs that they felt related to geospatial aspects of the Digital Government Initiative program.

National Imagery and Mapping Agency (NIMA) http://164.214.2.59

Currently NIMA funds ten basic research projects that are not classified. Eight different universities are working on these projects. Ten grants totaling 1.5 to 2 million dollars were funded for three years and can be extended if necessary. The areas of basic research that are involved were announced via a Broad Area Announcement (BAA) for:

- Intelligent agents for geospatial information discovery and retrieval
- Advanced Virtual Experts-knowledge-based expert systems
- Like-feature detection
- Cognitive workload—key information needs in high-stress environments
- Uncertain geospatial information representation, analysis, and decision support
- Cognitive aspects of geospatial information; formal models of spatial objects
- New methods in geodesy
- Models for information imagery and geospatial information
- Neuroscience-inspired target recognition

NIMA is interested in partnering and is looking in all possible areas. It is important to discover technology and services that overlap across agencies so that sponsored research attracts additional funding cooperators and research activities are not duplicated. NIMA will entertain basic and applied research proposals outside of normal BAA and funding vehicles.

Federal Geographic Data Committee (FGDC) http://www.fgdc.gov

The majority of the work done by FGDC is applied in the form of *demonstrations*. Many focus on sharing National Spatial Data Infrastructure (NSDI) with nontraditional agencies like HUD and Justice. Currently FGDC is looking for projects that involve crime, urban growth, quality of water or watershed, or emergency preparedness. They are also looking for matching funds. The DGI program is considered a vehicle for knitting together a working relationship in nontraditional spatial applications.

FGDC is concentrating on the following areas:

- Metadata
- Standards such as those of the International Standards Organization (ISO)
- NSDI clearinghouse nodes
- Data standards
- How to get implementation standards to work
- National spatial data framework

In addition to working on education, outreach, and training projects, FGDC is committed to working with agencies that have not traditionally included geospatial information services in their missions. A primary FGDC goal is to incorporate outcome statements in the agencies' strategic plans under the Government Performance Results Act.

National Aeronautics and Space Administration (NASA) http://www.nasa.gov

NASA is currently working on several projects. Of particular relevance to GSI are the Digital Library Project and the Digital Earth program.

The Digital Library Project relates to DGI's use of geospatial information. NASA has identified opportunities to conduct research for the digital government, including the massive amounts of data that are and can be made available for use in research. Already in the public domain, these data are ready for testbeds in all types of research and development scenarios.

The Digital Earth program may prove to be a parallel project that could help both define and solve some technology problems and also provide like or matching funds for digital government research projects. As the Digital Earth program gets up and running, funding is expected to be available for research in this area.

NASA called for workshop attendees to identify scenarios that illustrate the use of geospatial information to support government services. NASA has been using the Cooperative Agreement Program (CAP) for its research program. This process is generally satisfactory because it provides good synergy for interdisciplinary research, but NASA representatives also noted difficulties faced in meeting hard milestones.

United States Census Bureau (USCB) http://www.census.gov

The Census Bureau suggested several applications and tools that could contribute to the DGI project.

The TIGER® mapserver prototype (http://tiger.census.gov), which was developed by the Census Bureau and modeled after the Xerox PARC mapserver, now produces over 60,000 maps per day. This web service produces dynamic geographic map images from the Census Bureau's TIGER digital map database. Although current software services are somewhat inadequate, further development could support access and queries to large spatial datasets for high-volume map generation with simultaneous access to large database repositories of demographic, housing, economic, and survey data.

Another application that may prove useful to the DGI project is the ForNet MapServer package (http://www.gis.umn.edu/fornet/docs/MapServer/). Developed at the University of Minnesota, this application quickly and interactively generates maps with the shapefile format used in GIS software developed by Environmental Systems Research Institute, Inc. (ESRI). ForNet also offers advanced features such as the ability to make secondary database queries by intersection or by polygon. Census Lookup and U.S. Gazetteer (http://www.census.gov/cdrom/lookup and http://www.census.gov/cgi-bin/gazetteer, respectively) are database retrieval tools that contain thousands of data elements and tables for demographic and housing characteristics from the 1990 Decennial Census. The databases include geographic areas—nation, state, metropolitan data (MSA/PMSA), county, place, zip code, tract, block groups—for the entire country and outlying areas. The current Web-accessible facilities are written in C++, with back-end application programming interfaces (APIs) for remote query and acquisition of geography and data.

The MABLE/Geocorr Geographic Correspondence Engine (http://www.census.gov/plue) may prove to be another viable DGI tool. This software, which provides access to the MABLE geographic database and generates customized correlation lists, facilitates upward aggregation from city blocks to user-defined geographies, a step that is necessary for allocating and weighting data. Aggregation and correlation methodologies are key to comparing dissimilar geographies and to acquiring temporal, geospatial, and multidimensional data previously considered impossible to compare.

Census Bureau staff would like to expand the functionality of the MABLE tools by introducing more noncensus geographies (i.e., school districts, precincts, economic regions, telephone area codes, local planning areas, traffic zone areas, health service areas) to accommodate applications geared towards regional, state, and local governments. Adding temporal and geospatial inputs to the database creates multidimensional complexities, and the key to preserving these relationships is archiving and building an easily queried database that can resolve such issues as which areas intersect where and when. Beyond expanding the querying capabilities of the geo-aggregation tools, Census Bureau staff want to tightly couple these data to the mapping engine tools described above.
Additional funding is needed to update the MABLE tools, and the Census Bureau is looking for partners to further develop these technologies. The most challenging problem of relating GIS to census data is more completely describing conceptual approaches to supporting user-defined geographies, defining moving geographies, and identifying geographic changes over time (Meij and Blodgett 1997).

United States Geological Survey (USGS)

http://www.usgs.gov

The USGS representative discussed major themes crosscutting the agency: hazards, resources, environment, and information. The survey's DGI cooperation will probably be in the area of information. Instead of providing datafiles that cover the area of a topographic map, USGS is moving towards providing seamless databases with national coverage that will let any user chose the area of interest. Research priorities include:

- ensuring consistent information content across various spatial resolutions of geospatial data.
- developing means to distribute large-scale database and information (TerraServer is an early example of this technology).
- mining data and developing knowledge from the massive amounts of data collected by the Earth Observing System Data and Information System (EOSDIS).
- developing an effective user interface for land-based, multispectral, national, seamless database products.
- responding to emergencies (e.g., distributing and analyzing real-time volcano, earthquake, and stream gauge data).
- developing a framework in the areas of elevation and hydrography.

USGS is already working with industry on some of these research priorities through a number of Cooperative Research And Development Agreements (CRADAs). By easing institutional barriers to DGI participation, CRADAs may offer a good model for joint work between academia and industry. No funds were specifically allocated for external research in the FY99 budget, but partnering with universities would be possible.

Academic Research Priorities

Representatives from several academic communities central to GIS spoke about their respective research priorities. Academic communities involved in GIS research and applications cover nearly the entire breadth of a university:

- The root discipline of geography focuses on understanding geographic phenomena and developing methods to analyze them.

- Land use planning and natural resource management programs have been the historic drivers of GIS applications.
- GIS technology draws on database theory as well as computer and information science methodology, visualization, and other aspects.

University Consortium for Geographic Information Science (UCGIS)

http://www.ucgis.org

UCGIS is a nonprofit organization of universities and other research institutions dedicated to advancing the understanding of geographic processes and spatial relationships through improved theory, methods, technology, and data. Representatives gave a brief overview of the following ten UCGIS research priorities that had been identified through a process of white papers, committee reviews, and membership review and ranking.

Spatial Data Acquisition and Integration

With rapidly growing sources of geospatial data (such as satellite imagery, global positioning satellites, and other sensors) and a concurrent increase in the resolution and coverage of data collection, numerous problems need to be addressed. Issues include having frameworks that integrate data collected from multiple sources, accurate geographic referencing, and other problems related to blending data from multiple sources (also known as *data conflation)*. Also, the enormous volumes of data generated by automatic sensors require sophisticated algorithms for directing ground-based sampling that can recognize patterns and analyze data directly in the field.

Distributed and Mobile Computing

Concerted efforts are required to study implications of distributed computing with respect to geographic information (and vice versa). Research priorities include:

- Building models of GIS activities as collections of special services in distributed object environments to support their integration into much broader modeling frameworks. This will help promote the longer-term objective of making GIS services readily accessible in general computing environments of the future.
- Developing methods to efficiently use bandwidth when transmitting large volumes of geographic data (including progressive transmission and compression), to investigate the current status of such methods for raster data, and to research the use of parallel methods for vector data.
- Developing improved models (i.e., structure and format) of geographic metadata to facilitate sharing GIS data, to increase search and browse capabilities, and to allow users to evaluate appropriateness of using compilers to judge fitness of data for inclusion in GIS.

Geographic Representation

Current techniques for storing and accessing spatial data are not designed to handle the increased complexity and representational robustness needed to integrate diverse data across a wide range of applications and disciplines. Only the most rudimentary ability to represent and examine the dynamics of observed geographic phenomena is currently available within a GIS context. Problems that require analyzing change through time and patterns of change range from urban growth to agricultural impacts to global warming. Research in this area has particularly high priority because representational schemes must exist before databases can be built or analytical techniques based upon them can be developed. Given the rapidly increasing use of geographic information systems for policy analysis and decision-making, other urgent issues are how to represent data of varying exactness and degrees of reliability and then how to convey the additional information to a user.

Cognition of Geographic Information

Cognitive research promises both to make geographic information technologies more accessible to inexperienced and disadvantaged users and also to increase the power and effectiveness of these technologies in the hands of experienced users. Internet development has opened the possibility of systems that emulate the functions of map libraries. The future of such technologies depends on our ability to successfully reproduce map libraries' functions, including the assistance provided by library personnel to users with wide ranges of experience.

Interoperability of Geographic Information

Short-term goals are to more completely specify the formal semantics underlying models of geographic information and to use these models to intercommunicate among systems with significant geographical components. These goals can be addressed by developing geographical information systems with process-based models and a formal method of representing the semantics of a specific domain. There is a risk of identifying nonscalable solutions if the research focuses too much on narrow problems, but such projects are expected to develop largely in specific domain sciences such as hydrology, ecology, and regional science.

Scale

Primary issues are to better understand:

- how scale affects human perception.
- how to effectively and efficiently measure and characterize scale.
- how to use scale information in judging the fitness of data for a particular use.
- how to simultaneously automate change of scale and represent data on multiple scales.
- how scale and change in scale affect information content and analysis.

Spatial Analysis and Modeling in a GIS Environment

Analytical and computational methods must be devised that allow for solutions to problems conditioned both by GIS data models and by the nature of spatial and space-time inquiries. New forms of statistical analysis are needed to assess the relationships between variables in assorted spatial contexts. New theories must be devised to frame our understanding of relationships between variables at new levels of resolution and dimension.

Future of the Spatial Information Infrastructure

The principal activities underlying research and development in spatial information infrastructure can be assigned to the following four broad areas.

Information Policy. Research will identify optimal government information policies and practices to promote a robust spatial information infrastructure.

Access to Government's Spatial Information.

GIS and Society

Two concerns are overarching:

- How will GIS technology influence society's political, economic, legal, and institutional structures?
- How may these societal structures influence GIS development?

Computer Science Research Areas

Geographic information science has coevolved along with computer and information science. GIS data demands, visualization requirements, and unique search-and-query functions have challenged research in computer and information sciences. Research directions of faculty in computer and information sciences echo the nine research issues below, which were identified through formal UCGIS processes.

Optimize Spatial Queries

A mechanism based on the combination of spatial and nonspatial attribute values must be developed to optimize queries. For example, a number of methods respond to a query that seeks to find all cities with population greater than *x* within *y* miles of the Mississippi River. If *y* is small, then the spatially selected region is also small, and it is best to execute the spatial component of the query first. In contrast, if *x* is very large, then the nonspatial component is quite small (i.e., there are very few cities), and it is best to execute the query's nonspatial component first. Automatic techniques are needed to generate execution plans. This requires developing cost models for different operations as well as researching selectivity estimation techniques to sample the underlying data, especially spatial data. Computer vision techniques developed for probing may help respond to this challenge.

Incorporate Geometry into Database Queries

Conventional databases such as relational databases are accessed via selections, projections, and join operations while GIS databases are accessed via geometric operations. A useful area of investigation would be to incorporate geometric operations into standard database operations. For example, a recently developed operation called *ranking* lets a user find objects in the order of their distance from other objects with a spatial component. As a result, a user can browse (or *iterate*) through the database on the basis of a spatial condition such as distance. Other possible uses of ranking include connectivity (e.g., shortest path or route planning), gravity (for drainage), and frequency. Combining a semijoin with ranking results in the ability to perform clustering, which is equivalent to a discrete-space Voronoi diagram. The goal is to find analogs of other geometric operations. Any new algorithm must be efficient and should incorporate applicable results from computational geometry and databases.

Develop Techniques for Graphical Query Specification instead of SQL

Structured Query Language (SQL) has been the traditional means of interacting with a database. Unfortunately SQL is rather rigid and difficult to use to specify such spatial concepts as region boundaries. The natural way to interact with a spatial database is graphically. Thus graphical interaction techniques need to be developed along with methods for integrating them with SQL. Graphical query specification techniques could facilitate a user's interaction with GIS and could help extract answers, especially when GIS does not have the explicit tools and operations to answer a user's query.

Incorporate Time-Varying Data into GIS

Spatial data are not fixed but often vary with time. Resulting data are *spatio-temporal*. Research is needed into techniques for modeling such data and incorporating these data into GIS. Most work on temporal databases involves the concepts of valid time and transaction time. However, we cannot always look at data only via snapshots. Since spatial data are constantly changing, it would also be useful to represent the rates at which data are changing. Representing rates is one issue for investigation.

Incorporate Imagery

Data usually stored in conventional databases are discrete and nonspatial in nature, and conventional databases are specially designed to deal with such data. Spatial data, in contrast, have a continuous aspect or *extent* (e.g., lines, areas, surfaces, volumes), not just the locational component needed for point data. The most general spatial data is an image. There is a great need for techniques that will incorporate image data into GIS and will retrieve image data on the basis of content. This research area overlaps efforts towards content-based retrieval from image databases.

Develop Spatial Representations and Indexing Methods to Support Feature-Based Queries

Traditional work in spatial representations and indexing has focused on location-based queries (i.e., given location *x*, find all objects at *x*). Equally important are feature-based queries (i.e., find all locations where feature *f* is present). Such queries are increasingly important in areas such as mining spatial data. Thus, alternative indexing methods—which ideally will respond to both location- and feature-based queries—must be developed.

Investigate Using GIS on the World Wide Web

The dramatic rise in use of the Internet and World Wide Web has created an inevitable need to investigate issues that arise from using GIS on the Web. Important issues include transfering large amounts of data, considering system-like issues such as security, distributing work between client and server to best respond to a query, and processing data by determining the best ways to distribute data among different servers and algorithms.

Identify Issues in Knowledge Discovery and Spatial Data Mining

Research is needed to identify spatial data mining issues in addition to the index-related issues mentioned above. Operations based on spatial relations could be included in knowledge discovery and data mining techniques. Using GIS to support hypothesis formation and, more importantly, validation through intelligent access to a database needs consideration.

Increase Interoperability

Different geographical information systems must be able to communicate with each other. Even more importantly, systems need to work independently of the format of underlying data or in a way that makes underlying differences transparent to the user. Metadata issues also require research.

Industry Perspectives

Louis Hecht, vice president for Business Development at Open GIS Consortium, Inc. (OCG) (http://www.opengis.org), provided comments and perspectives from a consortium of companies and other organizations seeking to develop GIS industry standards for interoperability. He focused on the advantages of industry's involvement in setting research priorities and collaborating through the Digital Government Initiative program.

Companies gain market intelligence about digital government by participating in DGI discussions and watching the development of use cases. Market intelligence includes identifying potential near-term customers, identifying emerging markets and market trends, and evaluating the prospects of competitors. These activities are essential elements of the process that gives customers highly diversified, user-focused products and services.

When research/commercialization communication channels are rich, consensus develops around basic themes and general directions. A playing field takes shape and a set of common expectations arise that mold the collective technology effort as well as debates about policy issues such as copyrights for spatial data distributors, privacy, privatization of government functions, redundancy of data collection, and the role of government as producer or consumer. Policy and institutional arrangements evolve with the technology.

DGI provides an opportunity for industry to insert commercially critical research topics and priorities into the nation's digital government research agenda and architecture. This process helps shape a vision of digital government that harmonizes with industry's broader technology roadmap and increases the likelihood that DGI research will benefit other application domains.

OGC is a membership organization dedicated to developing open system approaches to geoprocessing. Through its activities to build consensus and develop technology, OGC affects the global community interested in geodata and geoprocessing standards. OGC promotes a vision of OpenGIS technologies that integrate geoprocessing with the distributed architectures of enterprise and Internet computing.

OGC was founded in response to widespread recognition of the following needs (http://www.opengis.org/info/brochure/index.htm):

- *The need to integrate geographic information contained in heterogeneous data stores with incompatible formats and data structures.* Incompatibility has limited the use of the technology in enterprise and Internet computing environments. The time, cost, and expertise required for data conversion have slowed adoption of geoprocessing.
- *The need to improve access to public and private geodata sources yet still preserve the data's semantics.* Agencies and vendors need to develop standardized approaches for specifying geoprocessing requirements for information system procurements. Industry needs to incorporate geodata and geoprocessing resources into national and enterprise information infrastructures so that these resources can be found and used as easily as any other network-resident data and processing resources. Users need to preserve the value of legacy geoprocessing systems and legacy geodata while incorporating new geoprocessing capabilities and geodata sources.

Workshop Findings and Recommendations

This section summarizes the major points that emerged from all workshop discussions and connects each point with the organizing committee's set of recommendations.

Information Storage, Archiving, Search, and Access

A primary goal of digital government research and development activities should be studying cost-effective designs of storage hierarchies that support the needs of sundry governmental applications focused on search-and-query processing tasks for geospatial data as an important and unique datatype.

Research is needed to develop algorithms and experimental systems that will facilitate the effective representation, storage, and access of various forms of geospatial data.

Research and development are needed into ontological representations of data, information sources and available expertise, and the means of integrating multiple query results into a comprehensible whole. Sharable ontologies that provide domain-specific assistance are needed to search sources, systems, practices and problem-solving capabilities of various agencies. Machine- and human-readable taxonomic descriptions of available information sources—as well as their meanings and relationships—would benefit from automated searching and dissemination of appropriate information. To help ensure quality of service, geospatial data objects should "know" about themselves, how they were collected and filtered, when they were collected, and what security/access control was used.

Recommendation 1. Advance research efforts directed toward the study of optimizing geographic query mechanisms and incorporating geometry and spatial relational operations.

Recommendation 2. Develop improved mechanisms for storing and representing time-varying geospatial data.

Integration of Information from Multiple Sources

Studies are needed to consider error propagation derived from variations in projections, collection accuracy and precision, bias, quality assessment, and multiple semantics.

Research is needed to span spatio-temporal frames that range from simple independent transactions to long sessions to ongoing problem-solving activities that draw upon local collections of persistent and semipersistent information structures.

Research is needed to focus on the integration of multiple platforms as well as information from various collection devices.

Additional integration research is needed to support a variety of views, with multiple engineering interfaces, that focus on specific classes of tasks.

Recommendation 3. Support research on integrating spatial data fusion from multiple agencies, distributed data, and multiple collection devices.

Recommendation 4. Support research on multiple representations/interfaces focused on task specific (procedural) workflow classes.

Knowledge Mining and Discovery

Information techniques and infrastructures are needed to analyze remote, inherently distributed data and information. New approaches need to be developed to effect search engines that feature geospatial analysis and modeling

A series of research opportunities relates to knowledge discovery and the efficiency of government operations. In addition to the challenges of long-term data archiving, the sheer size of many datasets needed to run in real time demands improved data transport and the development of scalable I/O architectures and interfaces. Research in very high algorithm complexity for advanced analysis includes developing algorithms for unstructured data, scaling algorithms to work with larger datasets, and achieving data reduction. Geospatial requirements such as spatial autocorrelation must be factored into this work.

Recommendation 5. Support research in developing algorithms for knowledge discovery applied to very large, frequently updated, spatial datasets such as those derived from space-borne Earth-monitoring sensors.

Information Validation and Visualization

A major research area concerns the evaluation of data, such as the error in estimating spatial relationships or in disaggregating spatially referenced data to smaller grids. Better information visualization technology is needed to support searching, browsing, navigating, and discovery.

A fundamental problem is developing guidelines and insights to which visual techniques are most effective in particular contexts, especially in the exploration of mechanisms for visualizing uncertainty in geographic data.

Advanced visualization capabilities, including immersive virtual reality, should be examined for their utility in three-dimensional spatial comprehension and analysis.

Recommendation 6. Support research in the theory and methods of representating data with varying degrees of exactness and reliability.

Recommendation 7. Support research in the context of decision-making to improve the representation of diverse data and the dynamics of geographic phenomena.

Universal Access to Information

The next generation of this work points to a valuable generalization of the cross-modal methods and standards already developed. Specific areas include networking protocols; multiple media session transcription; capture, search and replay methods; conversational abstracting technologies; and various agent and brokering services to support negotiations between user and system service capabilities. From a geospatial perspective, universal access could mean tactile maps and voice synthesis of geospatial data. These and other experimental themes of access need to evolve consistently with base/enabling technology developments.

Recommendation 8. Extend the promise of cognitive research to make geographic information technologies more accessible to inexperienced and disadvantaged users and also examine how government information policies affect access to and use of geospatial data for a broad spectrum of public and private sector stakeholders.

Electronic Commerce (Security and Trust)

Although issues of security and scale are paramount, they exceeded the scope of the workshop and were not addressed directly. However, the explicitly local nature of geographic information (by definition, describing the characteristics of a particular place) raises immediate concerns over privacy, protection of proprietary information, and other elements.

A key difficulty with geospatial information in electronic commerce is that we do not understand how to effectively introduce new technologies and new workflow processes in organizations and then migrate them out to citizen users. Highly desirable research will address workflow technologies that involve geospatial information so that bottlenecks and inefficient processes are identified and so that information processes become far more automated. This research needs to incorporate federal agencies, such as USGS or USCB, that deal extensively with spatial data but have not been the historic drivers of federal geospatial data and analysis.

Recommendation 9. Support research to examine commerce's issues in geospatial information such as preserving privacy despite geographic locators and breaking potential bottlenecks in distributing geographic information services due to GIS's unique workflow processes.

Role of Prototypes and Demonstrations

At some point the Digital Government Initiative program will need to coordinate sponsored developments and research in a setting where progress is demonstrated to agencies. The program may provide a forum for cooperation, dissemination, and evaluation among DGI participants. Developing a digital government prototyping facility will be a useful component. This concept could be either incorporated into the fabric of existing NSF-sponsored programs or run as an effort that is unique to DGI.

Testbeds and demonstration projects provide concrete examples that help people imagine digital government scenarios. DGI testbeds and demonstrations will uncover many unexpected issues that are unique to government agencies.

Capabilities should remain cutting-edge, to be used as a strong educational tool for demonstrating to governmental decision-makers what progress has been achieved to date, where the digital government is heading, and what steps should be taken to achieve the best results. A prototyping facility will also leverage academic, government, and industrial sectors to cooperatively produce cutting-edge tools, applications, and processes. Other federal agencies with laboratory programs should be invited to participate, thus creating a network for testing and developing processes that are consistent with U.S. priorities for information technologies in the workplace.

Recommendation 10. Develop a Geospatial Digital Government Prototyping Center to create a network for testing and developing processes consistent with U.S. priorities for geographic information technologies and services in the government workplace.

Conclusions

Within the last decade opportunities for electronic distribution of information and services have mushroomed along with the advances in computing and networking. These advances offer a tremendous opportunity to meet the needs of digital government, but only if concomitant research issues, such as those identified in this report, are addressed. Advancing technologies without advancing agencies' abilities to exploit those technologies would be a tragic loss of services and resources.

Geographic information and services constitute a large and growing part of the federal government's information requirements. Their use and development are pervasive through federal agencies such as defense, law enforcement, commerce, social services, land management, and emergency management. Geographic information faces several challenges over other information types due to the ubiquity of geographic space and the variety of mechanisms for querying, referencing, storing, and representing data. Tremendous improvements in the collection, management, analysis, and distribution of geographic information are possible—if given an appropriate investment in research and development.

At the same time, federal agencies are being called upon within their mission areas to provide increasing services in geographic information. However, the gap is growing between demands for services and our knowledge of how to address the myriad issues raised by the use of geographic information. An active program of GIS research and prototyping is critical to achieving the goals of the Digital Government Initiative.

This workshop was held to:

- Prioritize possible new federal research investments in computing, communications, and information sciences as well as engineering related to federal geographical information services.
- Identify opportunities for cooperation among federal agencies as they incorporate the DGI program and digital government issues in their strategic information services planning and as they address ongoing applied research initiatives within their agencies.
- Facilitate communications among federal agencies' R&D programs, university computer science and geographic information and technology departments, the National Science and Technology Council's Committee on Computing, Information, and Communications (CCIC), and other cross-agency activities and groups
- Establish geospatial priorities for research and demonstration projects of the federal Information Technology (IT) Innovation Fund.

Consequently this workshop merely *began* a process for shaping a research agenda that could be advanced under a DGI program. Clearly this advancement will be achieved only if resources are drawn from several thematic and disciplinary areas. Needs have been stated at all levels of government and across the public and private sectors. The Digital Government Initiative provides an important opportunity to take federally initiated geographic information services into the next decade.

References

References

Goodchild, Michael F., Max J. Egenhofer, and Robin Fegeas. December 5–6, 1997. Section 8: Research Agenda; Interoperating GISs: Report of a specialist meeting held under the auspices of the Varenius project panel on computational implementations of geographic concepts. (http://www.ncgia.ucsb.edu/conf/interop97/report.html#section8)

Gore, Al. January 31, 1998. The Digital Earth: Understanding our planet in the 21st century. Speech at the California Science Center, Los Angeles, CA. (http://digitalearth.gsfc.nasa.gov/VP19980131.html).

Gore, Al. September 2, 1998. Strengthening America's communities: Remarks at the Brookings Institution Policymakers Series, Washington, DC. (http://www1.whitehouse.gov/WH/EOP/OVP/speeches/sprawl.html).

Gore, Al. September 7, 1993. From red tape to results: Creating a government that works better and costs less. *Report of the National Performance Review.* (http://www.npr.gov/library/nprrpt/annrpt/redtpe93/).

Meij, Hendrick, and John Blodgett. June 24, 1997. MABLE+ (http://www.census.gov/plue/geocorr/doc/mablewp.html) and MABLE/Geocorr Tackling the Geographic Babel Problem (http://www.census.gov/plue/geocorr/doc/article.html). White paper and supporting document, respectively, in response to Call for White Papers: Workshop on Research and Development Opportunities in Federal Information Services (http://www.isi.edu/nsf/).

NSF. June 1, 1998. Digital Government Program Announcement–NSF98-121. National Science Foundation, Directorate for Computer and Information Science and Engineering, Division of Experimental and Integrative Activities. (http://www.nsf.gov/cgi-bin/getpub?nsf98121).

Schorr, Herbert, and Salvatore J. Stolfo. June 1997. Towards the digital government of the 21st century: A report from the workshop on research and development opportunities in federal information services, May 12–15, 1997. (http://www.isi.edu/nsf/final.html).

University Consortium for Geographic Information Science (UCGIS). 1996. Research priorities for geographic information science. *Cartography and geographic information systems,* vol. 23 no. 3. American Congress on Surveying and Mapping. (http://www.ncgia.ucsb.edu/other/ucgis/CAGIS.html).

Selected Research Funding Sources

Access America.: Government Information Technology Services (GiTS) of National Partnership for Reinventing Government. (http://www.gits.gov).

Federal Geographic Data Committee(FGDC). Program of metadata, clearinghouse, standards, framework and cooperative partnerships. (http://www.fgdc.gov).

National Aeronautics and Space Administration (NASA). NASA research opportunities. (http://www.earth.nasa.gov/nra).

National Imagery and Mapping Agency (NIMA). NIMA's solicitations and announcements. (http://164.214.2.59/poc/contracts/).

Open GIS Consortium, Inc. OpenGIS® abstract specification and implementation specifications. (http://www.opengis.org/techno/specs.htm).

Open GIS Consortium, Inc. OpenGIS® process documentation and technology development schedule. (http://www.opengis.org/techno/development.htm#Development Schedule).

United States Census Bureau (USCB). Geographic services and information. (http://www.census.gov/geo/www).

United States Census Bureau (USCB). Press release: Census Bureau, GeoResearch announce partnership. (http://www.census.gov/Press-Release/cb98-86.html).

United States Geological Survey (USGS). Contracts and grants information. (http://www.usgs.gov/contracts).

United States Geological Survey (USGS). National mapping information. (http://mapping.usgs.gov).

University Consortium for Geographic Information Science (UCGIS). Opportunities. (http://www.ucgis.org/f2oppor.html).

Appendix. Workshop Details

All breakout sessions discussed the same topics, then reported key discussion points to the plenary session.

Breakout Group A

Purpose of breakout group: Is it to guide NSF's proposal language and evaluation criteria? Or set the long-term vision? Or anticipate promising research opportunities?

Unique Requirements of Federal Government Agencies

Management of distributed information (Usselman)
Empowering citizens to know what personal info is being kept about them; to add to this information; and to protect their privacy (Behrens)
Economics / cost-benefit / policy aspects (Kottman/Herring/Usselman)
Search engines & indexing (Usselman)
Redundancies: linked vs. nonlinked (Herring)
Inventory of existing geo-information, as a sine qua non for improving efficiency (Herring/Hudson)
Fixing government (efficiency) vs. new opportunities (Tanimoto)
New opportunities: linked redundancies (Behrens)

Discussion of the Matrix

What's missing from the matrix? An infrastructure for research (Behrens)—e.g. framework data and "crosswalks" between ontologies (ISO/FGDC/etc.)

Information Storage, Archiving, etc.

Study of "as-is" vs. future technologies (Stephens)
Not so much Research as developing and prototyping and consensus-building (Kottman)
Demonstrate framework datasets that conform to public data and metadata standards (Behrens)
NSF-funded "infrastructure" work, e.g. building a lab (Stephens)

Info. Integration

Shared vs. different semantics (Kottman)
How to capture and reuse different people's processes for fusing datasets (Stephens)

Knowledge Mining and Discovery

Automatic information extraction from imagery (long-term) (Usselman)
Data mining: looking for a specific fact (Kottman); looking for statistical patterns (Tanimoto)
New additional geographic search and mining tools
No consensus on the meaning of *Knowledge*

Validation/Visualization

How to define and represent validity
Determine how many different dataset types are needed for a healthy digital government (Kottman)
Visually distinguishing the highly accurate data vs. the poor or even madeup data (Behrens)
"Commonsense GIS"—more or less geographic detail depending on (user or task) context (Stephens)
Not only lots of views, but a few standard views that we can all use (Tanimoto)

Universal Access

Rendering information in different ways for different users/clients (Lal)
Access to Universal Data vs. Universal Access to data (Hudson)
Adapt data delivery to client tools (push) vs. interfaces for information retrieval (pull)

E-commerce, Security, Trust

Protection and copyrights for spatial data distributors (Tanimoto)
How to use location as a catalyst for a geospatial marketplace, e.g. buying DEMs with satellite images, or looking at EPA maps of an area as part of a real estate transaction (Kottman)
Commerce may not be the key; rather, privacy protection
What are the limits on privacy? Each citizen should be able to set them! (Behrens)
What parts of government's current mission are amenable to e-commerce and could be privatized? (Stephens)

Architecture, Point Studies

E.g., the OpenGIS Web-mapping SIG—a pilot for the purpose of learning ("Deploy first, think later")

Consensus that this is a good idea

Other

NSF, OGC, FGDC: different constituencies, audiences, styles: need a business model! (Behrens)

Can users deal with the high complexity of all this additional info? (Tanimoto)

Tools should be handed not to end users, but to the professions (Kottman)

Growth of professionalism in the various geospatial disciplines (Kottman)

Educational component

Breakout Group B

Universal Access

What are agencies' responsibilities to make their data available?

Data files

Services on the data

Commercial value-added supplier or expert as an interface to data?

Focus discussion on intra-agency coordination/collaboration/datasharing—taxpayers/stakeholders will benefit as a result

Goals for the Digital Government

To make government operations more efficient in mission critical activities

To facilitate government's ability to support new and useful tasks based on their information

Vision

Access to all data through a common entry point and services

Achieve convergence of how geographic information is handled

Not data standards alone

Shared semantics and understanding

User asks any question and gets a useful answer

Interfaces Vision

Predominantly visual due to: human bandwidth, volume of input, implicit semantics

Direct manipulation—interact, increment, reverse

Augmented reality

3D flythrough

Fused data

Maps on demand

Research Issues—Semantics

How do you document semantics?

Community specific definitions

Sharable semantics/schema discovery

Translation and conflation of semantics

Mapping between queries, databases, and answers

Knowledge discovery, data mining

Adaptable query/query optimization

Open queries and intelligent agents

Research Issues-Geographic Information

Interface between MIS and GIS

Geospatial reference

Nonspatial data models

Gazeteers

Semantic translators

Culture change required to do this

Research Issues—Data

What data are available?

What spatial references do they have?

Are they redundant or represent different models, scales, representations?

How to handle redundancy?

Multiple accuracy, precision, data models, semantics

Breakout Group C

Citizen's View

Communication via new media, with exceptions handled by individuals in person
People in Gov't. have access to Gov't information
People in Gov't have power to make decisions

CS Perspective

Using natural interactions
Agent-based software systems

Impediments to Collaboration

Intellectual Property Rights
Internal politics within an agency
Politics across agencies

Unique Geospatial Aspects

Classified information
Need for aggregation
Privacy

GI in Nontraditional Agencies

GIS available to everyone but…not always called GIS

Redundancy Desired?

Back up	Yes
Duplication for mirroring	Yes
Acquisition	Mixed
Higher Confidence	
Data Integrity	
Maintenance	

Universal Access

ADA Compliance
Provide information, don't try to describe document
Verbal Tactile

Mining

Searching for interesting outlines or patterns
Searching for correlation
Finding redundancies
Needs-controlled vocabularies to facilitate discovery

Breakout Group D

Universal need to show results
Basic versus applied—agency mission
Adapt the Schorr report to geospatial
Strategy

Storage, Retrieval, Archiving, Searching

Large spatial databases, lack of scalability
OO vs. relational
Feature extraction, delivering higher-level products
Government as producer or consumer
New products of existing assets
Declassification
Multiagency
Representations, higher-level abstractions
Semantics
Metadata
Analysis
Data collection methods
What to archive?

Integration

Conflation, data fusion
A theory of geospatial data conflation
Across scales, time, disciplines

Knowledge Discovery and Data Mining

Search engines for geographic content
Find me places like this
Mining = feature extraction, web search, anomalies
Finding knowledge in government databases

Information Validation and Visualization

Validation = data quality
Role for government in validation?
Good housekeeping institute
Binary (accept/reject) or truth in labeling
Develop advanced tools
Visualization

Universal Access

ADA, visual impairment
Connectivity, geographically determined
Privacy, confidentiality
Analysis outside the firewall
Theory of privacy
Relative importance

Electronic Commerce

Traps, watermarking
Nothing special about spatial

Architecture Design Point Studies
Need to discuss further

Impediments
Lack of examples of good academic/government interaction

Government Services
Anything point-centered—private land, development—

Geoservices Vision
Max's house renovation permit
Maximizing the value of data assets
Emergency services

CS Technologies Vision
Solve these research problems

Potential in Nontraditional Agencies
Not FGDC founding agencies
Huge

Societal Benefits of Redundancy
Normalization
Research question, but is it geospatial?
How to measure redundancy, dollar cost?

Education
For researchers
For government cs/is specialists
Is the government set up to go digital

GPRA And Business Case
Change in agency strategic plan to include DG
Satisfied government users
d(LLPH)/dt

Breakout Group E

Goal
Digital government to increase self-governance (improve data transparency)

Vision
When we think of digital government we should think of facilitating government rather than the act of automating basic government processes.

Its identifying the activities that will facilitate government that is the objective of digital government initiative. This can best be facilitated through the development of topical/thematic white papers that are then used to drive tactical testbeds.

Focus on Process
Process includes 5 stages:

- Identification and clarification of topical areas.
- Commission white papers in topical areas.
- Panel review (workshops) to consider white papers and define parameters for testbeds.
- Testbeds (multiple) to exercise thematic area
- Development of use case scenarios to characterize the process

Minutes of Breakout Group E
DJ-Overview and Inventory of Topics

- Information storage, access, search archiving
- Integration of information from multiple sources
- Knowledge mining and discovery
- Info validation and visualization
- Universal access to information
- E-commerce
- Architecture Design/Point Studies (testbeds?)

Breakout AM—General Discussion
DJ-What is the DG initiative
What is its scope
DM—2 classes of agencies:
Those with a history of GIS and research
Those with little or no background in GIS and an aversion to basic research
Concern that DG initiative will be to applied for NSF and to research oriented for many agencies
EH—Thinks that last issue is relevant
DM—Concerned about the division that exists between CSEG and users, need to ensure that system design

JFER—Need to remember that many activities (e.g., planning) require both heterogeneous participation that x-cuts many groups and also that any system infrastructure must be flexible. Cites a continuum/cycle that exists from applied→operation and back to new applied
DJ—What are the "services" that should be provided
JFER—Basically need to balance systems
AM—What are the impacts of geospatial availability on decision-making
DM—Need to ascertain real cost-benefit from systems
EH—National programs have higher payoff
DM—Need to evaluate payback curves carefully
JF—Payback may include intangibles, that are hard to measure (e.g. better decisions)
DM—Need better methods to ascertain ROI
DJ—Do we know what initial impediments are in effect, is this a research topic for this group
AM—ROI may not be the criteria which drives this process or encourages NSF funding
EH—Who can place a cost-benefit on lower redundancy
DJ—If you develop better conversation plans based on the use of geodata the unit of measurement is the conservation plan not the geodata or geoprocessing capability
JFER—The cost-benefit is less a focal point than a background canvas against which the discussion occurs, at some point the focus needs to broader issues. Are we talking about automating government or improving range/scope of "services"?
Enabling self reporting through geoprocessing
Increased efficiency in government
LH—(to EH) Isn't that what you envision with the new farm plans. How we deliver service to the public is changing.
gov to get services and data as much.
JF—You're in transition from serving data to providing and packaging information
EH—We are now augmenting data with interpreted information and relying on new delivery mechanisms.
JFER—There is an element of giving the end-user access to primary data. If you enable this process of self-government then how do you enforce some form of regulation on the process to protect the "citizens". It's the problem of striking a balance between the individual and the government and creating unconstrained access/but regulating the process
DJ—If you use the permitting example you can either streamline the process or increase the speed of performance or both.
JFER—Perhaps better decisions are a result
AM—A related thing, there is the potential to pull together diverse information
JF—Need to remember that there is not one user but a range of users with a wide range of needs and requirements/expectations from data (raw→derived→synthetic) all of whom need to be supported and satisfied.
LH—What you are talking about is use case scenarios
DJ—So we're talking about an individual level orientation
EH—Much of this is very much in line with the FGDC-NSDI type of framework
MD—Points out the many aspects of the problem of scaling data and applications from the national level to the individual
EH—Need a vision for this discussion
MD—Is technology a .20 solution to a .05 problem

LH—A basic question, how do you get from legacy systems to object models
JF—Seamless example from Arkansas
DM—Data upon conversion don't match, need analytic objects that perform geoprocessing methods
JFER—Need to remember that these are not production items they are experimental in nature
DM—Emphasizes the experimental nature of objects
MD—Is OGC really providing a service or are they building what is currently possible?

Breakout PM—General Discussion of MS TerraServer

EH—Should we give some additional structure to the discussion

DJ—Have we established a sense of mission?

DM—Worries about the limitations associated with a few small pilot projects in the context of global datasets on an n dimensional sphere

General concern over the interest in agencies funding "research" projects

MD—It doesn't pay in government to be bleeding edge

AM—This is true in state and local government but not necessarily in fed

JF—Can we define what we mean by gov (fed->local) and research (applied, basic etc) and users and use these definitions as a framework for our recommendations. Does the Moore Chasm story have a use here.

DM—U of Minn has a program devoted to the study of spatial database's, how much can we really accomplish in 2–3 years, what can be realized.

AM—We don't need to consider only research which can be realized in 2–3 years. If the problem is worthwhile the projects should fly regardless of the time frame.

JG—How much of this (what we are talking about) could be accomplished with better hardware/software

DM—Expectations grow faster than data that grows faster than the computing infrastructures capacity

JG—We can talk about these in the abstract

strawman research topic

E-Commerce and integration of data from multiple sources are together interesting to the USGS, they would be interested in having input from a broad spectrum of industry/academia and others.

Is this an appropriate tact

DJ—Yes.

JG—The redundancy issue is important. Finding data is not hard, finding the right data is hard.

JF—These are 2 problems different measurements of the same phenomena (1) and the same measurement being applied to different phenomena (e.g. a product which has been morphed at different locations over time)

JFER—Currency in data is an issue.

LH—Parsing and evaluating information is a problem

EH—A research topic—*Centralization vs. redundancy*

DJ—Optimization of queries x-DB's is a potential research topic, if we are working towards a more individual notion of a user community what are the corresponding changes required in GIS

DM—Recall that for this round of funding NSF envisions limited funding focused on workshops etc. Perhaps we should create a program of thematic workshops targeting specific components of interest. Focus would be engineering bringing together experts in the field.

DM—Reiterates that the line agencies will be concerned with events that can addressed over the next several budget cycles

JF—Perhaps if the composition of the workshop groups includes good mix of private sector/agency/academic to develop a gant chart that measures needs and priorities by time to market and availability

EH—Need input from industry about what is possible, what NRCS is doing with reps from OGC community is to develop a network of testbeds across a range of environments and communities. What if FGDC does result in the development of a ubiquitous framework db, how is it served. The issue is the tremendous problem of education which is needed in the 2 communities (CS and natural resource management) how do we get them to understand and use. It's a Tech Transfer problem.

Testbeds are needed to help define the issues.

AM—It's not just a tech transfer/education issue. Fundamental cognitive research etc is needed. We've built a small interface . . . we were funded to do research in how domain specialists, not familiar with geospatial apps use and think about their data. Domain specialists in biostats etc used or avoided tools selectively resulting in a corresponding skew in analysis.

JG—Its somewhat of an artificial intelligence question, intelligent agents/domain agents can help to filter information and help to work to process data and return information

JFER—Research Topic—testbeds are a topic—need critical mass and the definition of area to focus on.

DM—The output from the panel review is a prioritization of tasks/technologies and a schedule

LH—Some concepts can't be dealt with at the abstract level. why hasn't anyone set up a spatial search engine yet, what's the industry doing about it? How do you package geographic information in XML.

Workshop Participants

Andersen, Ken	Lockheed Martin	ken.andersen@lmco.com
Anderson, Eric	City Manager, Des Moines	citymngr@mail.dwx.com
Arctur, David	LaserScan	arctur@lsiva.com
Backe, Kevin	U.S. Army TEC	kevin.backe@usace.army.mil
Beard, Kate	University of Maine	beard@spatial.maine.edu
Behrens, Cliff	Telcordia Technologies (*formerly* Bellcore)	cliff@bellcore.com
Brandt, Larry	National Science Foundation	lbrandt@nsf.gov
Buehler, Greg	Open GIS Consortium, Inc.	gbuehler@opengis.org
Caelli, Terry	The Ohio State University	caelli@cfm.ohio-state.edu
Connelly, David	University of Albany	dconnelly@ctg.albany.edu
Darling, Gary	Resources Agency of California	gary@ceres.ca.gov
Domaratz, Mike	USGS—National Mapping Division	mdomarat@usgs.gov
Egenhofer, Max	University of Maine	max@spatial.maine.edu
Evans, John	Massachusetts Institute of Technology	jdevans@mit.edu
Farley, Jim	Center for Advanced Spatial Technologies (CAST), University of Arkansas	jim@cast.uark.edu
Ferreira, Joe	Massachusetts Institute of Technology	jf@mit.edu
Gaines, Alan	National Science Foundation	againes@nsf.gov
Gardels, Kenn	University of California-Berkeley	gardels@regis.berkeley.edu
Goodchild, Mike	National Center for Geographic Information and Analysis (NCGIA)	good@geog.ucsb.edu
Gordon, Jed	U.S. Nuclear Regulatory Commission (NRC)	jgordon@nas.edu
Guptill, Steve	USGS—National Mapping Division	sguptill@usgs.gov
Hecht, Louis	Open GIS Consortium, Inc.	lhecht@opengis.org
Herring, John	Oracle Corporation	jrherrin@us.oracle.com
Horvath, Emil	U.S. Department of Agriculture (USDA)	ehorvath@ftw.nrcs.usda.gov
Hudson, Doug	Argonne National Laboratory	dlhudson@anl.gov
Johnston, Doug	University of Illinois at Urbana-Champaign	johnston@gis.uiuc.edu
Kemp, Karen	National Center for Geographic Information and Analysis (NCGIA)	kemp@ncgia.ucsb.edu
Kottman, Cliff A.	Open GIS Consortium, Inc.	ckottman@opengis.org
Kucera, Barbara A.	National Computational Science Alliance, NCSA	bkucera@ncsa.uiuc.edu
Lal, Nand	NASA—GSFC, National Space Science Data Center	nand@voyager.gsfc.nasa.gov
Lightfoot, James	Logicon Geodynamics	james.lightfoot@emu.logicon.com
Lyday, Win	National Association of Counties	wlyday@naco.org
MacEachren, Alan	The Pennsylvania State University	alan@essc.psu.edu
Marble, Duane	The Ohio State University	marble.1@osu.edu
Mark, David	SUNY-Buffalo	dmark@geog.buffalo.edu
Moeller, John	Federal Geographic Data Committee	jmoeller@usgs.gov
Muntz, Richard	UCLA	muntz@cs.ucla.edu
Nebert, Doug	Federal Geographic Data Committee	dnebert@usgs.gov
Radke, John	University of California-Berkeley	ratt@ced.berkeley.edu
Reed, Carl	Intergraph Corporation	creed@hq.pcmail.ingr.com
Saalfeld, Alan	The Ohio State University	saalfeld.1@osu.edu
Samet, Hanan	University of Maryland	hjs@umiacs.umd.edu
Senus, Walt	National Imagery and Mapping Agency	senusw@nima.mil
Shapiro, Stuart C.	SUNY-Buffalo	shapiro@cs.buffalo.edu
Stephens, Larry	National Imagery and Mapping Agency	stephenl@nima.mil
Stuber, Chris	Bureau of the Census	cstuber@census.gov
Tanimoto, Steve	University of Washington	tanimoto@cs.washington.edu
Thurston, Keith	U.S. General Services Administration (GSA)	keith.thurston@gsa.gov
Usselman, Tom	National Academy of Sciences	usselman@nas.edu

Workshop Agenda

July 8–9, 1998, Hilton Hotel in Arlington, Virginia

Wednesday, July 8

0830 **Welcome from NSF and the Organizing Committee**
Mel Ciment and John Cherniavsky
0900 **Panel—Different Views of Digital Government**
Horvath
Industry—Louis Hecht
Geographic Information—Max Egenhofer
Computer Science—Richard Muntz
1000 **Break**
1030 **Workshop Expectations**
Louis Hecht
1100 **Breakout Sessions**
1200 **Working Lunch**
1300 **Breakout Sessions continued**
1500 **Break**
1530 **Plenary Session—Breakout Reports**
1630 **Plenary Session—Crosscutting Themes and Unique Topics**
1700 **Close**

Thursday, July 9

0830 **Plenary Session—Current Technology Research Targets**
NIMA
FGDC
UCGIS
NASA
NSF-Federal IT Innovation Fund
1000 **Break**
1030 **Breakout Session—**Areas of collaboration between Government, Research, and Industry
1200 **Working Lunch**
1300 **Plenary Session—Breakout Reports**
1415 **Plenary Session—Discussion and Conference Wrapup**
1500 **Organizing Committee—Synthesis of Ideas**
1600 **Wrapup and Close**

This report is available on the EOT-PACI website:

- as a downloadable document http://www.eot.org/NE/DG-GIS.pdf
- as a fully functional Web document http://www.eot.org/NE/DG-GIS/index.html

For additional printed copies of this report, please contact:
phone: 217-244-4917
fax: 217-265-8022
tes@ncsa.uiuc.edu

National Computational Science Alliance
University of Illinois at Urbana-Champaign
605 East Springfield Avenue
Champaign IL 61820-5518