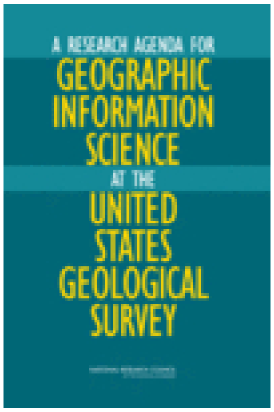


A Research Agenda for Geographic Information Science at the United States Geological Survey

Committee on Research Priorities for the USGS Center of Excellence for Geospatial Information Science, Mapping Science Committee, National Research Council

ISBN: 0-309-11155-2, 156 pages, 6 x 9, (2007)

This free PDF was downloaded from:
http://www.nap.edu/catalog/12004.html

Visit the National Academies Press online, the authoritative source for all books from the National Academy of Sciences, the National Academy of Engineering, the Institute of Medicine, and the National Research Council:

- Download hundreds of free books in PDF
- Read thousands of books online for free
- Purchase printed books and PDF files
- Explore our innovative research tools – try the Research Dashboard now
- Sign up to be notified when new books are published

Thank you for downloading this free PDF. If you have comments, questions or want more information about the books published by the National Academies Press, you may contact our customer service department toll-free at 888-624-8373, visit us online, or send an email to comments@nap.edu.

This book plus thousands more are available at www.nap.edu.

Copyright © National Academy of Sciences. All rights reserved.
Unless otherwise indicated, all materials in this PDF file are copyrighted by the National Academy of Sciences. Distribution or copying is strictly prohibited without permission of the National Academies Press <http://www.nap.edu/permissions/>. Permission is granted for this material to be posted on a secure password-protected Web site. The content may not be posted on a public Web site.

THE NATIONAL ACADEMIES
Advisers to the Nation on Science, Engineering, and Medicine

A RESEARCH AGENDA FOR GEOGRAPHIC INFORMATION SCIENCE AT THE UNITED STATES GEOLOGICAL SURVEY

Committee on Research Priorities for the USGS Center of Excellence for Geospatial Information Science

Mapping Science Committee

Board on Earth Sciences and Resources

Division on Earth and Life Studies

NATIONAL RESEARCH COUNCIL
OF THE NATIONAL ACADEMIES

THE NATIONAL ACADEMIES PRESS
Washington, D.C.
www.nap.edu

THE NATIONAL ACADEMIES PRESS • 500 Fifth Street, N.W. • Washington, DC 20001

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

The opinions, findings, and conclusions or recommendations contained in this document are those of the authors and do not necessarily reflect the views of the agency that provided support for the project. Mention of trade names or commercial products does not constitute their endorsement by the U.S. government. The study was supported by a U.S. Geological Survey Grant, No. 04HQAG0138.

International Standard Book Number-13: 978-0-309-11154-6
International Standard Book Number-10: 0-309-11154-4

Additional copies of this report are available from the National Academies Press, 500 Fifth Street, N.W., Lockbox 285, Washington, DC 20055; (800) 624-6242 or (202) 334-3313 (in the Washington metropolitan area); Internet http://www.nap.edu

Cover: Designed by Van Nguyen.

Printed in the United States of America.

THE NATIONAL ACADEMIES
Advisers to the Nation on Science, Engineering, and Medicine

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Ralph J. Cicerone is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. Charles M. Vest is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Harvey V. Fineberg is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Ralph J. Cicerone and Dr. Charles M. Vest are chair and vice chair, respectively, of the National Research Council.

www.national-academies.org

Copyright © National Academy of Sciences. All rights reserved.

COMMITTEE ON RESEARCH PRIORITIES FOR THE USGS CENTER OF EXCELLENCE FOR GEOSPATIAL INFORMATION SCIENCE

ROBERT P. DENARO, *Chair*, NAVTEQ Corporation, Chicago, Illinois
KATE BEARD-TISDALE, University of Maine, Orono
CYNTHIA A. BREWER, The Pennsylvania State University, University Park
MICHAEL DOMARATZ, Michael Baker Jr. Inc., Alexandria, Virginia
PENG GONG, University of California, Berkeley
ROBERT B. MCMASTER, University of Minnesota, Minneapolis
MING-HSIANG TSOU, San Diego State University, California
JOHN P. WILSON, University of Southern California, Los Angeles

National Research Council Staff
PAUL M. CUTLER, Study Director
ANN G. FRAZIER, Program Officer
CAETLIN M. OFIESH, Associate Program Officer
NICHOLAS D. ROGERS, Research Associate

MAPPING SCIENCE COMMITTEE

KEITH C. CLARKE, *Chair*, University of California, Santa Barbara
ROBERT P. DENARO, NAVTEQ Corporation, Chicago, Illinois
SHOREH ELHAMI, Delaware County Auditor's Office, Delaware, Ohio
HON. JAMES E. GERINGER, Former Governor of Wyoming, ESRI, Wheatland, Wyoming
GEORGE F. HEPNER, University of Utah, Salt Lake City
JOHN R. JENSEN, University of South Carolina, Columbia
NINA S.-N. LAM, Louisiana State University, Baton Rouge
LEWIS A. LAPINE, South Carolina Geodetic Survey, Columbia
MARY L. LARSGAARD, University of California, Santa Barbara
XAVIER R. LOPEZ, Oracle Corporation, Nashua, New Hampshire
ROBERT B. MCMASTER, University of Minnesota, Minneapolis
SHASHI SHEKHAR, University of Minnesota, Minneapolis
NANCY TOSTA, Ross and Associates Environmental Consulting, Ltd., Seattle, Washington
EUGENE TROBIA, Arizona State Land Department, Phoenix

National Research Council Staff
MARIA E. UHLE, Program Officer
JARED P. ENO, Research Associate

BOARD ON EARTH SCIENCES AND RESOURCES

GEORGE M. HORNBERGER, *Chair*, University of Virginia, Charlottesville
GREGORY B. BAECHER, University of Maryland, College Park
STEVEN R. BOHLEN, Joint Oceanographic Institutions, Washington, D.C.
KEITH C. CLARKE, University of California, Santa Barbara
DAVID COWEN, University of South Carolina, Columbia
WILLIAM E. DIETRICH, University of California, Berkeley
ROGER M. DOWNS, The Pennsylvania State University, University Park
JEFF DOZIER, University of California, Santa Barbara
KATHERINE H. FREEMAN, The Pennsylvania State University, University Park
RHEA L. GRAHAM, Pueblo of Sandia, Bernalillo, New Mexico
RUSSELL J. HEMLEY, Carnegie Institute of Washington, Washington, D.C.
MURRAY W. HITZMAN, Colorado School of Mines, Golden
V. RAMA MURTHY, University of Minnesota, Minneapolis
CLAYTON NICHOLS, Idaho National Engineering and Environmental Laboratory *(retired)*, Standpoint
RAYMOND A. PRICE, Queen's University, Ontario, Canada
BARBARA A. ROMANOWICZ, University of California, Berkeley
JOAQUIN RUIZ, University of Arizona, Tucson
MARK SCHAEFER, Global Environment and Technology Foundation, Arlington, Virginia
WILLIAM W. SHILTS, Illinois State Geological Survey, Champaign
RUSSELL STANDS-OVER-BULL, Anadarko Petroleum Corporation, Billings, Montana
TERRY C. WALLACE, JR., Los Alamos National Laboratory, New Mexico
THOMAS J. WILBANKS, Oak Ridge National Laboratory, Tennessee

National Research Council Staff
ANTHONY R. DE SOUZA, Director
PAUL M. CUTLER, Senior Program Officer
ELIZABETH EIDE, Senior Program Officer
DAVID A. FEARY, Senior Program Officer
ANNE M. LINN, Senior Program Officer
ANN G. FRAZIER, Program Officer
SAMMANTHA L. MAGSINO, Program Officer

CAETLIN M. OFIESH, Associate Program Officer
RONALD F. ABLER, Senior Scholar
JENNIFER T. ESTEP, Financial Associate
VERNA J. BOWEN, Financial and Administrative Associate
JARED P. ENO, Research Associate
NICHOLAS D. ROGERS, Research Associate
TONYA E. FONG YEE, Program Assistant

ACKNOWLEDGMENTS

This report was greatly enhanced by input from participants at the workshop and public committee meetings held as part of this study (see Appendix B). These presentations and discussions helped set the stage for the committee's fruitful discussions in the sessions that followed.

This report has been reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise, in accordance with procedures approved by the National Research Council (NRC) Report Review Committee. The purpose of this independent review is to provide candid and critical comments that will assist the institution in making its published report as sound as possible and to ensure that the report meets institutional standards for objectivity, evidence, and responsiveness to the study charge. The review comments and draft manuscript remain confidential to protect the integrity of the deliberative process. We wish to thank the following individuals for their participation in the review of this report:

Kenneth Boyko, U.S. Geological Survey *(retired)*, Newburg, Missouri
Barbara Buttenfield, University of Colorado, Boulder
Keith Clarke, University of California, Santa Barbara
Thomas Cova, University of Utah, Salt Lake City
Mark DeMulder, National Geospatial-Intelligence Agency, Reston, Virginia
Liping Di, George Mason University, Greenbelt, Maryland
William Gail, Microsoft Corporation, Boulder, Colorado
May Yuan, University of Oklahoma, Norman

Although the reviewers listed above have provided many constructive comments and suggestions, they were not asked to endorse the conclusions or recommendations nor did they see the final draft of the report before its release. The review of this report was overseen by Michael Goodchild, University of California, Santa Barbara. Appointed by the NRC, he was responsible for making certain that an independent examination of this report was carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final content of this report rests entirely with the authoring committee and the institution.

CONTENTS

APPENDIXES

Summary

Comprehensive and authoritative baseline geospatial data content is crucial to the nation and to the U.S. Geological Survey (USGS). To maintain its prominence, realize its potential, and fulfill its mission to develop and distribute these national data assets in a fast-moving information technology environment, USGS needs a coordinated geographic information science (GIScience) research presence that provides the scientific underpinning for these operations and exerts leadership in research that is critical to serving USGS's unique role. USGS founded its Center of Excellence for Geospatial Information Science (CEGIS) in 2006 to perform this task.

BACKGROUND

The USGS consists of major programs in the areas or "disciplines" of biology, geography, geology, water, and geospatial information. CEGIS resides in the National Geospatial Program Office (NGPO) within the Geospatial Information Office of the USGS. CEGIS is currently small, with three full-time equivalent (FTE) positions, a budget of $1.2 million per year to support an additional two to three USGS researchers and four to six FTE equivalent support staff, and a further $1 million for seven multicollaborator projects that were funded in FY 2007. The center's mission is to "conduct, support, and collaborate in research to address critical geographic information science questions of importance to the USGS and to the broader geospatial community" and "as an outgrowth of and complement to this research program, CEGIS will support and collaborate in technological innovations that further the implementation of the National Spatial Data Infrastructure" (CEGIS, 2006).

STATEMENT OF TASK

Given CEGIS's size and resources, it needs focus among the huge list of possible research topics encompassed by its mission. Consequently, USGS asked the National Research Council's Mapping Science Committee to convene a study panel charged to

1. Identify current and future USGS needs for GIScience capabilities;
2. Assess current capabilities in GIScience research at the USGS and recommend strategies for strengthening these capabilities and for collaborating with others to maximize research productivity; and
3. Using knowledge of the current state of the art in GIScience, make recommendations regarding the most effective research areas for CEGIS to pursue.

Over the course of 11 months, the committee met three times and received input from many sources (Appendix B). Given the short time frame, USGS urged the committee to focus on "what should we do and how should we do it"—primarily the second and third tasks—and to rely on published material to address the first task and the first part of the second task.

RECOMMENDATIONS

With a focused agenda as the key goal, initial attention is needed to research that will improve the capabilities of *The National Map,* which includes map layer databases, web map servers, and *The National Map* viewers. This USGS product, which was first envisioned and implemented in 2001, is "a database of continuously maintained base geographic information for the United States and its territories that will serve as the Nation's topographic map for the 21st century" (USGS, 2001). *The National Map* is the USGS vehicle for providing authoritative data content that has broad application within and beyond USGS. Success with *The National Map* is the prerequisite for any additional GIScience research at CEGIS. This success will then return fundamental, visible benefits to CEGIS, to NGPO in which CEGIS is located, and to the USGS disciplines of biology, geography, geology, and water, while tackling some of the most significant GIScience topics confronting the geospatial community.

RECOMMENDATION 1. CEGIS should initially focus on research that will improve the capabilities of *The National Map*.

Even with a focus on *The National Map*, the list of potential CEGIS research topics is large. Consequently, the committee drew on expert testimony and its own experience to develop eight prioritization criteria: CEGIS research should (1) be important to *The National Map*; (2) be important to USGS disciplines; (3) be relevant to society; (4) solve a problem and target a customer; (5) be foundational, understandable, and generalizable; (6) enable multidisciplinary integration; (7) focus on data content; and (8) show potential for early, visible success. These prioritization criteria led to the committee's recommendation for CEGIS's initial general areas of research and guided subsequent recommendations for detailed research within these areas.

RECOMMENDATION 2: The three priority research areas for CEGIS should be (1) information access and dissemination, (2) integration of data from multiple sources, and (3) data models and knowledge organization systems.

1. *Investigating New Methods for Information Access and Dissemination*. Access to information content by users is a key success factor at many levels for *The National Map*. The USGS disciplines need effective data access to carry out their missions. Other federal and state agencies need effective interfaces to *The National Map* content so that their organizations can maximize productivity when working with national and local data. This priority also supports society in general, since citizens need access to a trusted, up-to-date source of geospatial data for the nation that is easy and flexible to use. This diversity of users, from government agency experts to ordinary citizens, represents a significant challenge for effective information access and dissemination. In addition, this is an area with potential for visible early success enabling interim milestones in CEGIS's longer-term research agenda.

2. *Supporting Integration of Data from Multiple Sources*. Given the diversity of source data from state and local agencies as well as many add-on themes and the desire for multidisciplinary research across USGS, achieving efficient and accurate data integration is fundamental to the effectiveness of *The National Map* and will be a unique feature of *The National Map* relative to other online geospatial data sources. Within USGS, researchers in the biology, geography, geology, and water disciplines will need to find common reference data in *The National Map* and be able to load and share their thematic layers. Furthermore, the types of models and forms of spatial analysis that are increasingly needed to solve social and environmental problems will require that spatial data sets can be integrated in real time. CEGIS will have to find solutions to integrating data with widely varied quality, scale, and resolution.

3. *Developing Data Models and Knowledge Organization Systems*. To support society in general, *The National Map* will need both the semantic flexibility of a well-designed framework and models that enable a variety of user queries. This research will transform *The National Map* database into a comprehensive geographic knowledge base enabling knowledge discovery and analysis far beyond the typical mapping portal. This objective will likely require the most research effort, but will deliver enormous power to *The National Map* application and lead to its clear differentiation from other web-based products.

Under each of these three general research areas, the committee recommends a series of narrower research topics and associated research questions and suggests whether they can be answered in the short term (one to four years) or the longer term (four to eight years; Table S.1). Two topics under each research area require initial attention by CEGIS. All six research topics are derived from a critique of the present *National Map* and focus on the most important areas for improvement. These topics are summarized in the following paragraphs.

RECOMMENDATION 3: The two priority research topics within the area of information access and dissemination should be to reinvent topographic maps in an electronic environment and to investigate user-centered design for *The National Map* web services.

A well-designed and user-friendly map browser is essential for effective use of USGS data and map products. More importantly, however, topographic maps are among the essential *brands* of the USGS and a basis for *The National Map*. In the digital mapping age, CEGIS has the opportunity to conduct research that will transform well-designed traditional paper topographic maps into an electronic, web-based, multipurpose utility. Immediate attention is required to innovative formats and designs to reinvent topographic maps in the electronic environment. Also, to accomplish effective usability of *The National Map*, user-centered design (UCD) will result in the best solution for its users. UCD for *The National Map* web services will improve the usability of map viewers, web mapping services, and data themes by accommodating different needs for cartographic display and GIS functionality. Substantial results should be achievable in a short time on both research topics.

RECOMMENDATION 4: The two priority research topics for CEGIS within the area of data integration should be generalization and fusion.

Integrating spatial data sets from a wide range of sources presents a fundamental research challenge for CEGIS. Data integration may require many types of operations or techniques. For example, data fusion includes merging and linking information elements such as map features and images that could be disparate in scale, resolution, and quality. Generalization reduces the information content of maps due to scale change, map purpose, intended audience, and/or technical constraints. *The National Map* will require CEGIS researchers to develop unique generalization operations that can be automated for the many possible data types and map scales. Despite more than a decade of research on the topic, fusing disparate data sources together is still a significant challenge to be confronted by CEGIS.

RECOMMENDATION 5: The two priority research topics in the area of data models and knowledge organization systems should be developing geographic feature ontologies and building the associated feature data models and gazetteers.

Transformation of *The National Map* database into a comprehensive geographic knowledgebase can bring new dimensions to topographic information delivery and revitalize the role of the USGS as provider of geographic information and a valuable geospatial integration framework. *The National Map* cannot respond to simple queries such as “where is Canyon X” because it simply does not know what a canyon is. The use of geographic feature ontologies can formally define a set of geographic features to enable knowledge discovery through such queries.

Table S.1 lists the research topics and questions generated by the committee under each of the three research areas identified in Recommendation 2.

TABLE S.1 Summary and Time Lines of Recommended CEGIS Research Areas, Topics, and Questions

Research Area	Research Topic (Bold = Priority Topic)	Research Questions	Time Range
Information Access and Dissemination	**Innovative formats and designs to reinvent topographic maps in an electronic environment**	1. What is the widest range of scales that can be mapped only by adjusting map symbols combined with selectively removing feature types?	Short term
		2. What is the minimum amount of change to map symbols and content that provides the maximum scale range maintaining topographic map usability?	Short term
		3. What is the stability of topographic map design (with the goal of establishing a coherent set of designs that function from coarse to fine resolutions through scale change)?	Short term
		4. What should be the visual hierarchies for the base *National Map* layers?	Short term
		5. How should USGS select a subset of automated and manual approaches to visual hierarchies to provide a tool that effectively serves the largest number and variety of *National Map* users seeking to answer geographical questions that are not served by commercial point-to-point navigation tools (e.g., Google Maps, MapQuest, Yahoo!)?	Short term
		6. What is the optimal combination of types and number of symbols for an inexperienced user to create an effective topographic map and accommodate a data overlay on a topic of interest using web tools?	Short term

User-centered design (UCD) for implementation of *The National Map* web services	1.	With the goal of updating and evaluating *The National Map* viewer user interface, (a) what types of user interfaces are appropriate for *The National Map* viewers, (b) does *The National Map* need different viewers for different users and map contents or is a single one appropriate, and (c) what kinds of communication methods are effective for disseminating geospatial information through web browsers?	Short term
	2.	Will new web mapping technologies, such as vector-compression algorithms, AJAX, and Adobe Flex, improve the usability and system performance of *The National Map* servers and general web mapping applications?	Short term
	3.	What is an appropriate standardized user testing and evaluation method for assessing and improving the effectiveness of *National Map* products?	Short term
Open Geospatial Consortium (OGC) Standard Profiles for *The National Map* web mapping services and map layer design	1.	How should USGS create OGC standard profiles (which are a subset of standard specifications and customized standard content) to bring layers in *The National Map* databases into conformance with OGC standards?	Short term
	2.	How can USGS overlay well-positioned labels with clear categories and hierarchies on top of symbolized features dynamically set to foreground and background depending on user interests?	Short term

Integration of Data from Multiple Sources	**Generalization**	1.	What are the specific new generalization operations and algorithms that will be needed for *The National Map*?	Short term
		2.	What feature-based generalization is needed for *The National Map* (the focus would be on a specific feature, such as a stream, and approaches needed for stream generalization) and how can that be accomplished?	Long term
		3.	What new kinds of measurements will be needed to determine locational conflicts between USGS features?	Short term
		4.	What are the effective scale ranges for fusing two layers together, and how does generalization affect fusion?	Long term
	Data fusion	1.	What are the data quality issues related to spatial data integration and fusion?	Short term
		2.	How can areal interpolation—as a key method for fusing aspatial data with spatial data—be applied in *The National Map?*	Long term
Data Models and Knowledge Organization Systems	**Geographic feature ontologies**	1.	What are the key sets of topographic features portrayed within *The National Map* layers that should be explicitly represented in ontologies (these might align with the set of features already identified within the Spatial Data Transfer Standard; USGS, 1994)?	Short term
		2.	What are the formal operational definitions for these features, their parts and structures, and their relationships to other features?	Short term
		3.	What automated feature extraction methods are derivable from these operational definitions?	Short term

Ontology driven data models and gazetteers	1.	How does a geographic feature ontology operationally support a *National Map* feature database?	Long term
	2.	How can the collection, validation, modeling, and management of vernacular names be facilitated?	Short term
	3.	How can the creation of more detailed or smart feature footprints be automated?	Short term
	4.	How can the implications of fuzzy footprints in gazetteers be managed?	Short term
Quality-aware data models	1.	How can sampling distributions of complex objects be defined and managed (e.g., reduce them to points in some *N*-dimensional shape space)?	Long term
Data models for time and change	1.	What can be learned from spatiotemporal use cases for advancing spatiotemporal models for *The National Map?*	Long term
	2.	How is change effectively represented in spatial data sets?	Long term
	3.	How can process-based models be used to improve data quality or quality awareness in *The National Map*?	Long term
Transaction Processing	1.	What is the transaction processing logic for complex spatiotemporal transactions among *National Map* and other USGS databases?	Long term

To address these research topics, CEGIS needs a sustainable research management process involving a portfolio of collaborative research that balances short and long-term goals. This will require an in-house staff of Ph.D.-level scientists (including postdoctoral fellows and visiting scientists) working in small teams on each research topic. CEGIS should fund external research on these same topics through directed or competed grants (for shorter-term topics) and support of university centers (for longer-term investigations). To develop future research directions driven by user requirements fed through the National Geospatial Technical Operations Center, USGS disciplines, and *The National Map* design team, the center should host specialist meetings, consult an advisory board, and track developments and lead discussions within the GIScience community. The committee makes the following recommendations in this regard:

RECOMMENDATION 6: CEGIS should initially comprise six to eight Ph.D.-level scientists working in teams of at least two on the high-priority topics identified in Recommendations 3 to 5. Each team would comprise a mix of USGS scientists and visiting scientists and/or postdoctoral fellow(s) as appropriate to the topic. Their location should not be constrained to USGS facilities if the most efficient progress could be made in another setting (e.g., an academic center of excellence).

Placing USGS researchers at university centers of excellence in GIScience would potentially be of great benefit to cultivating CEGIS's GIScience leadership. These centers would extend the research reach of CEGIS and focus on longer-term projects.

RECOMMENDATION 7: CEGIS should establish and/or support one to two centers of excellence in GIScience at universities with relevant GIScience focus and capabilities that address its longer-term research challenges.

Since not all required research can be conducted with a small internal team, to extend the scope of research CEGIS could utilize contracting mechanisms to accomplish part of its research agenda. This outsourced research could be performed with academic or industrial entities.

RECOMMENDATION 8: CEGIS should supplement the work of its core research teams with Broad Area Announcements, Cooperative Research and Development Act agreements, and targeted contracts on high-priority research topics.

By liaising with other agencies with GIScience research activities and coordinating on topics of common interest, CEGIS could piece together a national geospatial research agenda. In addition, significant work is being performed in private-sector firms and professional societies. Relationships with these activities are critical to CEGIS. Collaboration with other agencies and organizations doing GIScience research is crucial to realizing a national need to integrate these activities and is also crucial to establishing the leadership of CEGIS and the USGS in GIScience for the nation. For example, the Open Geospatial Consortium (OGC) is a leader in generating interoperability standards for web mapping applications. There are many OGC standards closely related to the development of *The National Map* and other USGS products, such as Web Map Services (WMS), Web Feature Services (WFS), Web Coverage Services (WCS), and Catalog Service for Web (CSW). Reenhancing strong connections between USGS and OGC will ensure interoperable web mapping services in USGS products and support their broader usage and accessibility.

RECOMMENDATION 9: To reestablish USGS's leadership role in GIScience, maximize efficiency, and share in the cost of addressing common challenges, CEGIS should forge connections with other federal agencies, professional societies, and private-sector firms that conduct, support, and/or promote GIScience research.

A major tenet of *The National Map* is its aggregation of highly relevant and timely local data from states and counties. Some of these entities have made significant progress in the application of GIScience, and mutual benefits would accrue to each with explicit collaboration.

RECOMMENDATION 10: Because of USGS's core role in integrating data from local sources for *The National Map*, CEGIS should establish collaborative activities with state and local agencies that have progressive activities in GIScience.

Visibility in the geospatial community will be vital to USGS's reemergence as the nation's GIScience leader. CEGIS could participate in and lead key events and sessions to establish its role in the nation's GIScience activities.

RECOMMENDATION 11: CEGIS should use specialist meetings, perhaps in conjunction with the University Consortium for Geographic Information Science winter meeting or summer assembly, to advance its state of knowledge and plans for addressing emerging research challenges.

An advisory board that includes the other disciplines of USGS would foster channels of connection and communication with these disciplines, as well as providing a path for defining requirements and research needs. The disciplines would benefit by having influence on the research agenda of CEGIS to suit their needs.

RECOMMENDATION 12: To provide broad-based input, review, and critique of CEGIS plans, activities, and progress and to institutionalize CEGIS's connection to the USGS disciplines, the National Geospatial Program Office should establish an advisory board for CEGIS that includes members from each of the USGS disciplines as well as non-USGS GIScience experts.

With these actions, and a focus on the research areas recommended in this report, CEGIS could become a nimble, dynamic, cutting-edge research unit that emerges as the critical research engine underpinning USGS's capability to supply the nation's authoritative geospatial base content. Although the committee's charge was to suggest an agenda for this early phase of CEGIS, it is assumed that as CEGIS grows in resources and expertise, it will expand to encompass research into broader areas of GIScience. The committee believes that in the future, not only could CEGIS provide the structure to conduct research—its approach could establish GIScience leadership for the USGS.

1

Introduction

The U.S. Geological Survey (USGS) has a long history in the development of geospatial data, starting with the topographic mapping of the nation that began in the late nineteenth century and continued throughout the twentieth century. The USGS developed topographic maps on several scales that have been used over the past 100 years by professionals and citizens alike. To carry out its mapping mission the USGS gained considerable expertise in cartography and, with the advent of computer technology, was one of the leaders in the development of techniques and standards in the field of digital cartography. Even further advances in technology have pushed the USGS to strive toward a fully electronic implementation of the topographic maps and related geospatial data, which has been named *The National Map*.

The challenges of developing *The National Map* differ greatly from those faced by cartographers even 20 years ago. Meanwhile, geographic information system (GIS) technology has become ubiquitous, with digital mapping sources to be found anywhere from specialized government agency sites, to state and local government web pages, to commercial sites that have caught the interest of the general public. Within this fast-changing environment, the USGS realized the need to assess the focus of its research in geographic information science (GIScience) to determine how it could best meet the needs of *The National Map*, the USGS, and the nation.

THE EMERGENCE AND FOCUS OF CEGIS

The idea of a Center of Excellence for Geospatial Information Science (CEGIS) was first proposed by McMahon et al. (2005) in a report that describes a science strategy for geographic research, including GIScience, at the USGS between 2005 and 2015. CEGIS was initiated by the Associate Director for Geospatial Information in 2006.

CEGIS is housed within the National Geospatial Program Office (NGPO). The NGPO was created in 2004 when the USGS reorganized its geospatial information programs to better invest in technology and partnerships aimed at modernizing its collection, management, processing, updating, and delivery of geospatial information.[1] The major elements of USGS's geospatial programs and services unified under NGPO include *The National Map*, the National Atlas of the United States of America®, the Federal Geographic Data Committee secretariat, Geospatial One-Stop, and other geospatial program elements (Figure 1.1). Geospatial information is one of five disciplines within USGS, the others being water, geology, geography, and biology (Figure 1.1).

Among the NGPO's responsibilities are defining the overall GIScience (see Box 1.1) research agenda and championing GIScience research as a component of USGS's science portfolio. CEGIS undertakes these GIScience research responsibilities. The USGS's vision for CEGIS is to "conduct, lead, and influence the research and innovative solutions required by the National Spatial Data Infrastructure (NSDI)" (CEGIS, 2006).[2] CEGIS's mission is to "conduct, support, and collaborate in research to address critical Geographic Information Science questions of importance to the USGS and to the broader geospatial community" and "as an outgrowth of and complement to this research program, CEGIS will support and collaborate in technological innovations that further the implementation of the NSDI" (CEGIS, 2006).

---

[1] See *http://www.usgs.gov/newsroom/article.asp?ID=80*.

[2] The NSDI is the means to assemble geographic information that describes the arrangement and attributes of features and phenomena on the Earth. The infrastructure includes the materials, technology, and people necessary to acquire, process, store, and distribute such information to meet a variety of needs (NRC, 1993).

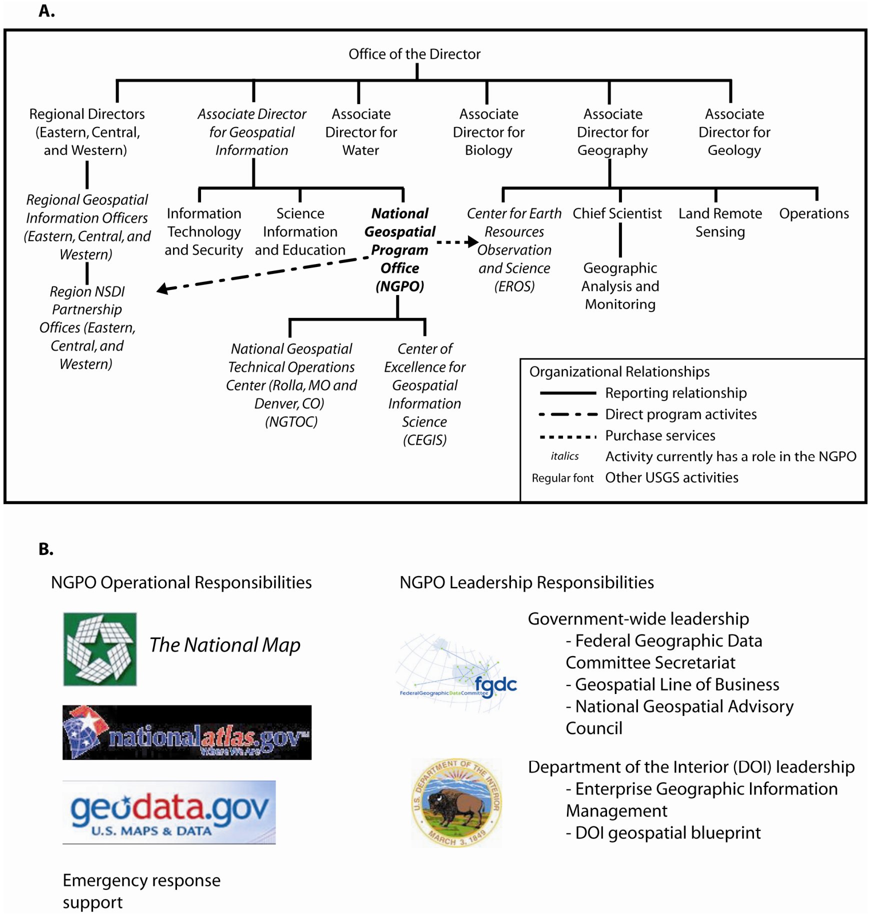


FIGURE 1.1 Organizational chart that emphasizes (in italics) (A) organizations that carry out geospatial responsibilities of the Associate Director for Geospatial Information and (B) the major program activities through which these responsibilities are carried out. SOURCE: USGS.

BOX 1.1
Definition of Geographic Information Science

The individual who coined the term GIScience defined it as “a multidisciplinary research enterprise that addresses the nature of geographic information and the application of geospatial technologies to basic scientific questions” (Goodchild, 1992).

CEGIS staff recently identified the reach of GIScience as including “the traditional mapping disciplines of surveying, aerial photographic interpretation, photogrammetry, remote sensing, and cartography. It also encompasses a broader scope of issues related to the modeling and representation of geographic phenomena, data, and processes; human cognition of geographic information; the analysis, depiction, and use of uncertainty information; spatial analysis and modeling, including geographic information systems (GIS); scale sensitivities; geographic ontologies; visualization; and other similar topics” (CEGIS, 2006). GIScience relies on expertise from many allied fields and has intimate ties to geospatial technology and applications.

As stated earlier, the starting point for planning CEGIS’s GIScience research activities is a study by McMahon et al. (2005) that describes a science strategy for geographic research, including GIScience, at USGS between 2005 and 2015. The recommendations of the McMahon report were, in fact, quite broad and include the needs of the other USGS disciplines. Furthermore, many of the authors, including McMahon, were from disciplines within USGS other than geography, or were from outside of the USGS. The McMahon report recommended that USGS establish CEGIS to lead USGS GIScience research (for details on the McMahon report’s ideas for CEGIS, see Appendix C). Proposed areas of focus within CEGIS are drawn from goals 8, 9, and 5 in McMahon et al. (2005), namely:

- Provide timely, efficient, and intelligent access to new and archived USGS geographic data needed to conduct science and support policy decisions (Goal 8).
- Develop innovative methods of modeling and information synthesis, fusion, and visualization to improve our ability to explore geographic data and create new knowledge (Goal 9).
- Develop credible and accessible geographic research, tools, and methods to support decision making related to the human and environmental consequences of land change (Goal 5).

While the call to action in the McMahon report is the primary reason that CEGIS is being established now, the agency feels a need to retain key talent with a critical mass of researchers for which a center of excellence would be helpful (Steve Guptill, USGS, personal communication, 2006). The McMahon report suggests that such a center would build, nurture, and maintain a core of GIScience researchers and provide a focal point and sense of identity for these researchers. Lastly, CEGIS will build a science role within NGPO in addition to its operational and leadership roles.

CEGIS TODAY

The CEGIS budget covers three full-time equivalent (FTE) employees who are considered CEGIS staff—one in Reston, Virginia, and two at Rolla, Missouri. In addition, CEGIS funds GIScience-related projects conducted by two to three other USGS FTEs and a support staff of four to six FTEs.[3] The USGS also funds a CEGIS-affiliated postdoctoral research position managed by the University Consortium for Geographic Information Science (UCGIS) (CEGIS, 2006). In late 2007, three additional postdoctoral positions in GIScience will be added through USGS's participation in a National Research Council (NRC)-administered postdoctoral program.

The FY 2007 budget has two components. The first is $1.2 million to cover the activities listed above, which are located not only at Reston, but also at the Center for Earth Resources Observations and Science (EROS) in Sioux Falls, South Dakota, and in Rolla, Missouri (to support all of the FTEs mentioned above). The second component of the CEGIS budget is an FY 2007 call for proposals, which resulted in approximately $1 million being awarded to seven interdisciplinary projects that involve people across USGS in collaboration with non-USGS partners.

CEGIS's research portfolio is thus a mix of (1) preexisting GIScience-related projects already funded within USGS and placed under the management of CEGIS when it was formed in 2006 and (2) the seven new projects funded under the FY 2007 call for proposals.

FOCUS OF THIS REPORT

Recognizing the need to develop a set of research goals and priorities for CEGIS that will best meet its needs for future capabilities in GIScience, the USGS approached the NRC through its Mapping Science Committee (MSC) and asked the MSC to form a committee to develop these research goals and priorities (Appendix A). Using knowledge of the current state of the art in GIScience and information from USGS on current and future needs and capabilities, the committee was asked to determine which areas of research would be most effective for the CEGIS to pursue. The three primary tasks follow:

---

[3] The ranges arise because CEGIS provides partial support to a number of people whose combined contribution is equivalent to between eight and nine FTE positions.

1. Identify current and future USGS needs for GIScience capabilities.
2. Assess current capabilities in GIScience research at the USGS and recommend strategies for strengthening these capabilities and for collaborating with others to maximize research productivity.
3. Using knowledge of the current state of the art in GIScience, make recommendations regarding the most effective research areas for the CEGIS to pursue.

To complete its task, the committee met three times in person—twice in Washington, D.C., and once in Irvine, California—and numerous times by phone. The committee benefited from input from a range of experts (Appendix B) and drew from a broad range of documents listed in the references.

Given the short time frame and potentially broad scope of its task, the committee chair and study director met with the USGS Associate Director for Geospatial Information, the CEGIS Director, and the Associate Director's Chief Scientist (Karen Siderelis, Steve Guptill, and Anne Frondorf, respectively) prior to the first full committee meeting to discuss the sponsor's expectations from the study and to gain insights into their priorities among the items in the committee's task. As the primary audience for the report, these senior USGS staff indicated that they were most interested in the committee's insights on item 3 and the second half of item 2 of its task. These tasks were summarized as, What should CEGIS focus on and how can this be achieved? The committee's discussion of and recommendations on these tasks are covered in Chapters 3 and 4, respectively.

In addition to emphasizing a desire for the committee to focus on tasks 2 and 3, USGS steered the committee away from primary research on the first task and the first half of the second task. Instead, the USGS urged the committee to draw on recent reports such as McMahon et al. (2005) and on responses to the call for proposals for this information.[4] These tasks are covered in Chapters 1 and 2.

In addition to the USGS's guidance on emphasis among the committee's three tasks, it encouraged the committee to shape the CEGIS research portfolio based on its need for GIScience research across USGS (i.e., not solely within NGPO). However, the committee is keenly aware of the importance of *The National Map* to the mission of NGPO and to the USGS as a whole, not to mention

---

[4] The McMahon et al. report was written by a team whose primary objective was to develop a strategy for USGS geography science activities from 2005 to 2015 by analyzing regional, national, and global scientific issues and needs. The 12-person team that conducted the work spanned all the disciplines of USGS and also included the chairs of NRC panels that earlier had reviewed the geography program and the vision for *The National Map*. The group heard from 175 people from all facets of government, academia, industry, and nongovernmental organizations. That effort was supported by almost an order of magnitude more funds than this NRC study, and USGS sponsors of this study urged the committee to draw from that resource with respect to USGS needs for GIScience capabilities.

the USGS's topographic mapping mission responsibilities. In 2002, the NRC stated that "developing *The National Map* is the most important single initiative in the Geography Discipline at the USGS" (NRC, 2002), while recognizing that "*The National Map* as a database product and an information base is an attainable goal by 2010, but some of the basic knowledge needed to create it (and other spatial data products) is not yet available," and that "present knowledge, methods, and tools are inadequate to create *The National Map* . . . ". In the committee chair's meeting with Karen Siderelis, Steve Guptill, and Anne Frondorf, they confirmed that supporting *The National Map* is of highest priority and that there were critical research needs for accomplishing that objective.

The USGS urged the committee not to constrain the scope of research based on current CEGIS resources and to think in terms of the next decade of research. Even with this guidance, the committee concurred with the USGS sponsors that the range of needs for GIScience research (as described in USGS, 2001; NRC, 2002; McMahon et al., 2005, and responses to the FY 2007 call for proposals, for example) would readily exceed even the most optimistic expectations of resource availability for CEGIS. Consequently, the committee focused its view of CEGIS's role on applied, technical aspects of GIScience and away from software engineering, product development, and nontechnical aspects (e.g., institutional issues, digital rights management challenges) of supporting the National Spatial Data Infrastructure (while recognizing the need for these roles elsewhere in USGS). The committee's process of prioritization of research tasks is explained in Chapter 3.

GISCIENCE CAPABILITIES AT USGS

The USGS employs a small cadre of GIScience professionals. In addition, it uses several mechanisms that bridge to external GIScience expertise.

Internal GIScience Resources

Much of USGS's GIScience expertise has already been identified and linked to CEGIS. As one indication of the GIScience capabilities already tapped by CEGIS, the projects funded in FY 2007 (excluding those funded through the bureau-wide call for proposals) focus on the following:

- Automated data integration ($280,000)
- Generalization for *The National Map* ($190,000)
- Building an ontology for *The National Map* ($250,000)
- Multiresolution raster data for *The National Map* ($200,000)
- LIDAR-derived elevation technology assessment ($80,000)

- Elevation feature extraction ($127,000)
- Fractal and variogram analysis of scale and resolution effects in geospatial data ($100,000)

Appendix D includes more information on each of these projects, and Chapter 3 describes how these projects fit with the committee's suggested research topics.

Another indication of GIScience capabilities across the bureau arises from the topical focus of the proposals received in response to CEGIS's FY 2007 USGS-wide call for proposals. This call by CEGIS leadership had a goal to identify "hidden" GIScience talent within the disciplines at USGS that might be a valuable resource for the center (Steve Guptill, USGS, personal communication, 2006). Although each research team must be led by a USGS researcher, the team is encouraged to be multidisciplinary and include non-USGS expertise. Consequently, an additional result from this call is that it reveals the broader network of GIScience capabilities to which USGS experts are already connected.

Of the 69 proposals received in response to the call, 23 were submitted from the water discipline, 20 from geography, 15 from biology, 10 from geology, and 1 from NGPO. In all disciplines there was a broad range of proposed topics that spanned monitoring and data capture through data integration, analysis, and error propagation to modeling and decision support. CEGIS leadership now has a better sense of the distribution and range of GIScience and related capabilities across USGS. The seven projects funded through this call for proposals are:

1. Scaling, Extrapolation, and Uncertainty of Vegetation, Topographic, and Ecologic Properties in the Mojave Desert ($73,000);
2. A Landscape Indicator Approach to the Identification and Articulation of the Ecological Consequences of Land Cover Change in the Chesapeake Bay Watershed, 1970 – 2000 ($132,000);
3. Assessing Local Uncertainty in Non-Stationary Scale-Variant Geospatial Data ($117,000);
4. Methods to Quantify Error Propagation and Prediction Uncertainty for USGS Raster Processing ($135,000);
5. The Geoscience of Harmful Invasive Species: Integrating LANDFIRE (Landscape Fire and Resource Management Planning Tools Project) and Invasive Species Data for Dynamic and Seamless Integration of Raster and Vector Data to Meet Management Needs at Multiple Scales ($150,000);
6. Mapping Inundation at USGS Stream Gage Sites: A Proof of Concept Investigation ($150,000); and

7. GEOLEM:[5] Improving the Integration of Geographic Information in Environmental Modeling through Semantic Interoperability ($150,000).

See Appendix D for more information on these projects.

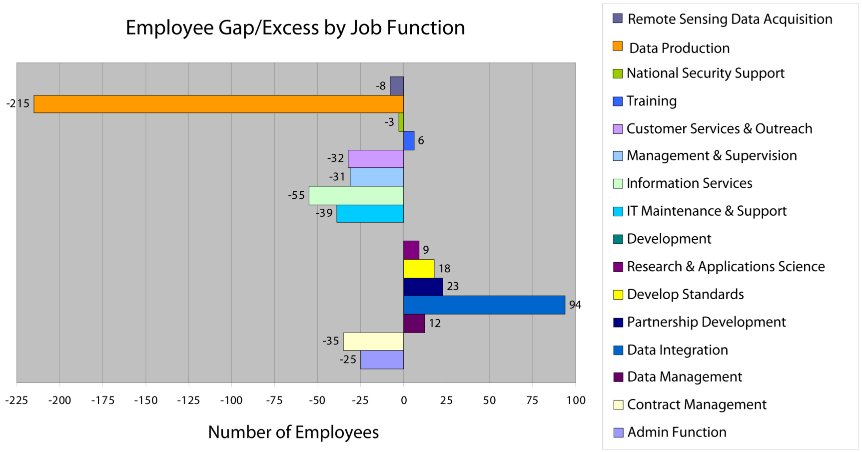


FIGURE 1.2 Difference between employee functions in 2003 and projected future geography discipline employee functions. Negative numbers represent an excess of employees in the function; positive numbers represent a current shortage in the function. Largest proposed growth areas (positive numbers) are in data integration and partnership development. Because this figure was generated before the USGS map production functions moved to NGPO, it is at best a general indication of planned changes that are not yet complete. SOURCE: USGS, 2003.

GIScience research activities funded within and by CEGIS are occurring in the midst of USGS's shift in emphasis of staff expertise away from paper map production and toward integration of digital data (Figure 1.2). An assessment of resources against the current mission shows an excess of staff engaged in data production and a shortfall of staff skilled in data integration—a function that will be important for assimilating data from other sources as the USGS moves out of internal production and into working with partners.

---

[5] Geospatial Object Library for Environmental Modeling.

Leveraging External Resources

USGS uses Cooperative Research and Development Act (CRADA) agreements as one means of connecting with external GIScience expertise. An example of such an agreement for geospatial activities is with Microsoft Corporation on the development of Terraserver.[6] In addition to CRADAs, USGS has leveraged external GIScience expertise by arranging a series of visiting academic GIScientists who have been based at Reston, Virginia during sabbatical leave; participating with the National Geospatial-Intelligence Agency (NGA) in the solicitation and review of NGA University Research Initiative proposals; conducting a graduate school training program in which more than a dozen USGS employees pursued GIScience studies at universities including Ohio State University, State University of New York Buffalo, University of California at Santa Barbara, and University of South Carolina; organizing research meetings (e.g., the Public Health Colloquiums) with expert participants invited through UCGIS; funding postdoctoral positions in GIScience at USGS facilities; and sponsoring GIScience professional meetings (Steve Guptill, USGS, personal communication, 2007).

FUTURE USGS RESEARCH NEEDS AND CEGIS

Geography, GIScience, and mapping will be increasingly important to the USGS's water, geology, and biology disciplines. With the new roles of map and information integration, CEGIS will face increasing demands for solutions to complex geospatial data processing challenges as well as automation of those functions so that USGS researchers can handle large amounts of dissimilar and nonconforming data with frequent updates. In addition, USGS's major role in analyzing land change over time will require new GIScience-derived methods.

In 2007, the USGS released its report outlining a 10-year science strategy for the agency. *Facing Tomorrow's Challenges: USGS Science in the Coming Decade* (USGS, 2007) sets the bar high early in the introduction: "The USGS is the Nation's and the world's leading natural science and information agency . . . [whose efforts] . . . allow the USGS to map and understand land use/land change trends across the Nation." The 2007 report highlights the GIScience needs that could be fulfilled by CEGIS. These needs are described in Chapter 2.

---

[6] See *http://nationalmap.gov/gio/viewonline.html.*

2

USGS Needs for GIScience Capabilities

This chapter addresses the first of the committee's tasks—to identify current and future U.S. Geological Survey (USGS) needs for geographic information science (GIScience) capabilities. The chapter begins with USGS-wide needs, then describes National Geospatial Program Office (NGPO) needs, and ends with the needs of *The National Map*. Given the strategic importance of *The National Map* to NGPO and its potential value across USGS and to the nation, the majority of this chapter is devoted to understanding the drivers of potential priorities among the many GIScience research needs of *The National Map.* Following USGS's suggestion that the committee expend the majority of its effort on the other two items in its task, the committee relies heavily on summarizing existing information in this chapter.

BUREAU-WIDE GISCIENCE RESEARCH NEEDS

Multidisciplinary research is a central goal of the USGS science strategy (NRC, 2001; USGS, 2007). GIScience research underpins this goal by addressing such challenges as data integration across the disciplines. As the Center of Excellence for Geospatial Information Sciences (CEGIS)'s USGS-wide FY 2007 call for proposals states, "The role of the [Geographic Information Office] in providing a geospatial framework for integrating information among all the USGS science disciplines is an important element of GIScience at the USGS." Access to and mining of geographic base data are also critical components of the USGS research infrastructure (McMahon et al., 2005), and GIScience research supports these capabilities.

The USGS's science strategy outlines an inspirational science agenda for the agency. Every one of the six interdisciplinary "science directions"[1] discussed in the report has a profound reliance on geospatial databases and methodologies. This broad and deep reliance on geospatial content and solutions points to a rich opportunity for CEGIS to contribute to the other USGS disciplines, especially through an improved *National Map*. Examples are cited in the following paragraphs.

For the ecosystem direction, the report states that "USGS and partners will develop new products, including standardized national maps of ecosystems in the United States and regularly updated status and trends assessments" The human health direction underscores its dependence on geospatial technology with the statement that "USGS capabilities in environmental monitoring and mapping are important components for understanding environmental relations to human health and evaluating probable outcomes of future human health risk factors." However this initiative goes further in defining a strategic action for geospatial technology to "develop an online atlas of potential environmental health threats . . . [and] develop and implement a national-scale, real-time, environmental health threat warning system that combines biological, water-quality, and geologic information with GIS [geographic information system] decision-support tools." Indications that there is more work to be done in GIScience also appear in relation to the water census direction: "The Water Census will also require improvements in the mapping and characterization of the geologic and geomorphic framework of the Nation's principle aquifers and watershed systems."

These objectives will necessitate enhancements to current databases, platforms, and tools—in particular, advances in scale handling, temporal analysis, real-time access to data, standards, and multidisciplinary analysis. Requirements for the first three are captured in the vision of the ecosystem direction that "comprehensive, multi-scaled, online digital maps of the Nation's ecosystems and their physical and biological components are routinely used for management, education, and portrayal of change over time. Real-time ecological data, images, maps, and research findings are available to the public on interactive USGS websites."

Scale Adjustment

The ecosystem direction aims to "coordinate, develop, and regularly update a standardized national map of ecosystems and their physical and biological components, at scales appropriate to land-manager needs, to facilitate the ability to assess, monitor, manage, and restore ecosystems." This range of scales is unique to USGS: "The position of USGS as a non-regulatory agency, with capabilities in environmental

---

[1] Water census, human health, hazards, ecosystems, climate, and energy.

monitoring and mapping at all scales from national to local and the ability to understand environmental and ecological processes, is found nowhere else in the Federal Government."

Temporal Analysis

Temporal analysis and real-time access expand the roles for a geospatial database. The temporal analysis requirement also underpins the key question in the climate direction: What links between climate, land use, and hydrology influence the temporal and spatial characteristics of water resources?

Access to Real-Time Data

Access to real-time geospatial and other data is a rapidly growing requirement for managing the nation's natural resources. The human health direction identifies a strategic action requiring real-time analysis and challenges the USGS to "develop an online atlas of potential environmental health threats, which consolidates USGS data and information and provides real-time data for researchers and public-health agencies to enhance the Nation's ability to respond quickly to current threats and anticipate potential future threats."

Management of natural hazards is also highly dependent on timely data dissemination. The USGS's vision in terms of hazards expects that "by 2017, the USGS will . . . significantly expand urban hazard mapping throughout the Nation . . . and we will have the models, metrics, decision-support tools, and portals that provide intelligent access to remotely sensed data and geospatial information for cost-effective risk-reduction, response, and recovery efforts . . . [G]eographic methods and tools need to be developed . . . and intelligent access [provided]." In addition, the energy direction anticipates the need to share information as well as "maintain and update the geological and geophysical databases and geochemical baselines used to develop national and global resource assessments . . . [and] . . . assure the data are accessible both internally and externally." Lastly, the vision for the hazard direction is to have all seismically active areas served whereby "associated maps of shaking level, population density, and susceptibility to landslides will be posted on the Internet within minutes of the determination of the earthquake location and magnitude."

The implications of information dissemination are global and multinational—as summarized in USGS's final vision statement: "An international consensus is developing on the need to leverage recent advances in computer science and related technologies to create a next-generation, integrated science computing and collaboration platform that will be as transformational as the Internet."

Standards

Standards are a catalyst for the success of GIScience technology at USGS. The final vision of the science strategy refers to the importance of standards: "The use of open standards and of tools has minimized the difficulty of merging or comparing data sets; searches on a location or topic of interest quickly yield comprehensive research data."

Multidisciplinary Analysis

The USGS's multidisciplinary composition is unique among federal agencies. "The USGS is the only Federal agency that combines scientific expertise in biology, hydrology, geology and geography." This creates unique capabilities such as in hazard forecasting, for example: "These accomplishments [hazard forecasting, monitoring, diagnosis] are all possible because the USGS is able to bring a unique combination of disciplines—biology, geology, hydrology, geography, and geospatial information technology—to bear on all these hazards." The key ingredient for success is data integration, captured in the strategy for the climate direction: "Our breadth of multidisciplinary scientific expertise . . . enables us to deliver uniquely integrated information." In fact, this capability has the potential to be world class, because, as the human health direction states, "USGS databases constitute one of the most comprehensive and high-quality arrays of national, regional, and local biologic, organic and inorganic analyses available from any single source." Yet the integration of diverse data sets is still a great challenge, as discussed in the next section.

Challenges

Although the USGS has a unique set of databases and disciplines, there is substantial work ahead because, as discussed in the health direction, "the many data sources are scattered across the USGS and not easily available to most users. If [catalogued,] this array of environmental data and information could provide our partners and customers with unified spatially and temporally referenced sources of information. An important step in the overall goal of protecting public health is to integrate existing USGS databases [to identify] potential environmental health threats and provide the underlying framework for USGS environmental health studies."

The solution to these and other data integration needs is a geospatial platform comprising spatial and semantic reference systems (Kuhn, 2005) that will support spatial and semantic integration. *The National Map* is such a geospatial platform, but significant research and development is needed for it to meet

USGS's science needs over the next decade. In addition, it is both conspicuous and of concern to the committee that throughout USGS's inspirational science strategy—with its numerous references to mapping databases, capabilities, and tools—*The National Map* is never mentioned. However, this can be considered an opportunity for NGPO and CEGIS. As shown above, for the USGS disciplines to successfully implement their science strategy, there are numerous challenges that are inherent to GIScience, including issues with scale, temporal analysis of geospatial data, improving access to real-time data, and developing the standards and data integration techniques that will allow true multidisciplinary analysis. All of these are challenges that must be resolved for successful implementation of *The National Map*, therefore it became clear to the committee that the GIScience needs of the USGS as a whole and those of *The National Map* are highly complementary.

GISCIENCE RESEARCH SUPPORTING NGPO'S COORDINATION RESPONSIBILITIES FOR THE DEPARTMENT OF THE INTERIOR

The Department of the Interior leads federal activities to improve the use of geospatial data through the development of the National Spatial Data Infrastructure (NSDI) and the coordination of geospatial activities across the federal government. NGPO's four federal coordination activities are the Federal Geographic Data Committee (FGDC), Geospatial One-Stop, Geospatial Line of Business, and the National Geospatial Advisory Committee. These activities have not established research needs, which raises the question of whether these coordination activities might generate a need for specific research. Early work in relation to federal coordination focused on technology to reduce barriers to data sharing, but emphasis has now shifted to identifying common agency needs and aligning business practices and investments. The remaining technology component focuses on improved interoperability or the ability of dissimilar networked computer systems to exchange data and instructions to provide computing services. Of particular interest are approaches that improve interoperability but minimize the need to change or retrofit participating organizations' existing technology implementations. Nonetheless, while the success of NGPO's coordination activities has a technical component, participants in NGPO "listening sessions" have noted that the main challenges are organizational, not technical, in nature. Therefore, the committee has assumed that any research needs of these geospatial activities would be adequately met by the research agenda developed to meet the criteria developed for this report.

NATIONAL MAPPING GISCIENCE RESEARCH NEEDS

The list of potential topics in a CEGIS research agenda is daunting. An all-encompassing list would have to address the needs of the NGPO, the USGS disciplines, other federal agencies, state and local agencies, and the public (commercial firms and private citizens). In addition, a wealth of potential research topics has already been identified in a number of documents (e.g., USGS, 2001; NRC, 2002; NRC, 2003; McMahon et al., 2005; CEGIS, 2006; and DiBiase, et al., 2006). In light of CEGIS's lean start with two full-time researchers and a director there is no doubt that the center needs to prioritize among all these demands.

As discussed in Chapter 1, there is wide consensus on the importance of *The National Map*. This consensus began to build with the conclusion by the National Research Council (NRC, 2002) that developing *The National Map* was the single most important initiative for the geography discipline. McMahon et al. (2005) reiterated this sentiment. Also in an interview for this study, Karen Siderelis, Associate Director for Geospatial Information at the USGS, said that NGPO must get *The National Map* right as its first priority (Karen Siderelis, USGS, personal communication, 2006). An effective *National Map* is the most critical success factor to CEGIS and the NGPO. If *The National Map* is to be the primary source of USGS geospatial products to customers outside of the USGS, then ultimately focusing on the research needs of *The National Map* will address the needs of these external users.

Yet what about the additional research needs of the NGPO and the USGS disciplines? As shown earlier in this chapter, the committee consensus is that topics relevant to *National Map* implementation cover a wide range of the GIScience research agenda that is broadly relevant to USGS (e.g., integration of data from diverse sources, effective display and processing, temporal processing of spatial data). Given the time constraints of this study and the sponsor's request to focus on tasks 2 and 3, it was difficult for the committee to assess the unique GIScience needs of the USGS disciplines above and beyond those that meet the overall USGS science strategy. However, the needs of the USGS disciplines of water, geology, biology, and geography can be directly assessed by CEGIS and included in the research agenda to ensure that *The National Map* will meet their important needs with appropriate priority. Consequently, a dedicated initial focus by CEGIS on *The National Map* will achieve three broad USGS goals:

1. *The National Map* is primed for success, its use is increased, and USGS regains a leadership role in GIScience.
2. *The National Map* becomes an essential platform serving the geospatial needs of the USGS disciplines and supports execution of the 2007 USGS science strategy.

3. The initial focus provides a concrete and necessary early target for CEGIS that yields a visible and measurable outcome of the research.

It is important to note that *The National Map* is not just a data viewer. It includes the data, the processes for obtaining and maintaining the data, the national standard-certified nature of those data, the data models and relationships, and the data knowledge systems and ontologies that make up the vision of *The National Map*. In fact, there may be multiple viewers for *The National Map*, government, academic and commercial, as these organizations seek to exploit the richness and value of the data in *The National Map.* Success with *The National Map* is the prerequisite for any additional GIScience research at CEGIS. Since CEGIS must prioritize its research agenda to optimize the effectiveness of limited resources, the committee feels that an initial focus on the GIScience needs of *The National Map* would provide the most visible benefits to CEGIS, NGPO, and the USGS disciplines while tackling some of the most significant GIScience topics confronting the geospatial community. As the resources of CEGIS grow and early successes with *The National Map* are achieved, the research agenda can and should broaden to more diverse areas of research.

RECOMMENDATION 1. CEGIS should initially focus on research that will improve the capabilities of *The National Map*.

Within *The National Map* focus, the delivery of topographic maps deserves special mention. Topographic maps were the virtual brand of the USGS during its formative years. *The National Map* is often thought of as the electronic evolution of the topographic map.[2] However, the current *National Map* does not deliver fully on the concept of the traditional topographic map nor has it fully embraced a broader concept of topographic information representation and delivery in other than map form. Despite the impressive data access and manipulation features of *The National Map*, the topographic map "brand" has been diluted. Also, because paper maps are still favored by many users (particularly in the emergency management community), USGS will have to consider bringing its branded product back into play in *The National Map*.

USGS's focus would need to remain on topographic information but with an expanded range of products and services. One thrust involves revitalization of the traditional topographic map as an effective cartographic product in light of new digital design research as well as research on user interaction and effective interactive cartographic products. Another thrust involves ontology development and new data models for *The National Map* that move it beyond traditional and current forms of topographic information delivery. Investigation of "intelligent"

---

[2] See *http://nationalmap.gov/report/national_map_report_final.pdf.*

topographic features and access to fundamental topographic information in other than map form open new research directions and enhanced capabilities for *The National Map* that support queries for information not easily or fully served by current methods.

The remainder of this section describes critical details of *The National Map* (including its possible evolution, mentioned in the previous paragraph) upon which the committee builds its discussion of research priorities in the next chapter.

The Current *National Map*

The National Map is "a database of continuously maintained base geographic information for the United States and its territories that will serve as the Nation's topographic map for the 21st century" (USGS, 2001). This new topographic map is an electronic realization of the original paper maps, and with that structure comes greatly increased flexibility and opportunity. The term "national map" encompasses the underlying data, data management approach, data access methods, supporting partnerships, and models of data use (Box 2.1). USGS describes the relationships among the *The National Map*, Geospatial One-Stop*,* and the NSDI in the following way: "NSDI is a *concept* defined as including all aspects of geospatial data, Geospatial One-Stop is a communications *portal* for geospatial information content and related information, and *The National Map* is geospatial information *content* in the form of data and applications" (USGS, 2003).

The relationships among *The National Map*, Geospatial One-Stop, and the FGDC can be visually expressed in the architecture of *The National Map* (Figure 2.1). In addition to the architectural elements, the figure is divided into three levels that represent general topics under which research can be organized: data integration, data access, and data use (Chapter 3).

1. *Data Integration*. In the data integration level, all eight national data themes defined by the USGS are integrated into *The National Map* database (as the data management component). *The National Map* databases are accessed and used by *The National Map* web server, the National Atlas database, and the Geospatial One-Stop portal.
2. *Data Access*. At the data access level, *The National Map* web server provides the graphics and the data streams for *The National Map* viewers (web browsers) to display maps. The end users can query and combine multiple layers provided by the map server and customize their final mapping design and layout by using *The National Map* viewers. The Geospatial One-Stop portal provides data catalog services and web mapping catalog services to the end users (who might be GIS professionals or scientists). End users can search, request, and download the data they need. In addition, the National Atlas server provides mapping functions for the end users through the National Atlas viewers.

3. *Data Use*. At the data use level, users create end products of *The National Map*, including paper maps or graphic maps displayed on a computer screen or embedded in a program. Two major groups of users are USGS scientists in the various disciplines and the general public. Scientists can download data from Geospatial One-Stop or directly from *The National Map* databases for their domain-specific research or spatial analysis. These applications will create new GIS data sets with added value. These new data sets are added back to the Geospatial One-Stop Portal or *The National Map* web server for future applications or use. Public users generate maps for activities such as hiking, school projects, and map reading.

BOX 2.1

What Is *The National Map?*

The USGS's vision of *The National Map* targets improvements in data characteristics such as currentness, seamlessness, consistent classification, variable resolution, completeness, consistency, variable positional accuracy, spatial reference systems, standardized content, metadata, and temporal dimensions (USGS, 2001). The goals of *The National Map* are ambitious in attempting to integrate data from a large number of sources, many with widely varying specifications.

Here, the committee describes the components of *The National Map* and then illustrates how the components fit together in the cases of two layers within *The National Map*: the National Hydrography Dataset and high-resolution imagery over urban areas.

Data: The data themes in *The National Map* are orthoimagery, elevation, hydrography, geographic names, land cover, transportation, structures, and boundaries of governmental units (e.g., states, counties) and publicly owned lands (e.g., national forests, state parks). These themes were selected in large part because USGS is authorized to provide them if no other source is available, and they typically comprise the information portrayed on USGS topographic maps. USGS's original intent was that the data would retain the characteristics appreciated by users of USGS topographic maps, including consistent feature identification and classification and comprehensive national coverage. In addition, the data would be seamless (not interrupted by arbitrary edges introduced by the data production process), would be more current, and would have positional accuracy equivalent to or better than that of USGS topographic maps.

Data management: USGS's data management approaches support two methods through which partners participate in *The National Map*. These methods generally parallel the "blanket" and "quilt" approaches described by the National Research Council (2003). In the blanket method, the USGS maintains databases of national data coverage. Partners help maintain these databases by sending data snapshots or transactions that are incorporated into a national coverage. In the quilt method, USGS maintains a web-accessible database that catalogs web services maintained and hosted by partners. The former approach offers improved data integration and utility to customers, especially for advanced geocoding and modeling applications; the latter offers the potential for faster availability of more current data.

Data access: USGS offers several methods to access data in *The National Map*. For users seeking to view a map of the data, a map viewer is available (see *http://nmviewogc.cr.usgs.gov/viewer.htm*) and provides a "print" function to create a page-size hardcopy. For those developing their own viewer or other applications, the USGS offers an application programming interface (API) and related web service through which users can use the catalog to access inventoried services. Metadata entries in the catalog are "harvested" into Geospatial One-Stop and so can be discovered through that portal. For those interested in retrieving copies of data, especially those in national coverages, USGS offers interactive and preprocessed methods to select and retrieve data through the web or on media. Agencies have negotiated service-level agreements with USGS for more advanced web-based access to national databases.

Partnerships: Perhaps the most fundamental change represented in *The National Map* approach is the transition from USGS relying on internal resources to collect new data to relying on partners to provide new data. These partnerships are based on an exchange of value between USGS and partners. In exchange for partners' data, USGS has provided value in the forms of funding, data, data models, data collection software tools, and access to contracts and related management and quality assurance processes, information technology, web and other data management services, and expertise.

Data use: The "build once, use many" approach of *The National Map* supports up to three models of data use. The first is visualization of the data in the form of a map graphic, in which the data encode the location and a basic description of a feature. The second is support for geocoding, in which the data support the assignment of a position based on another reference system (such as street addresses or stream reach codes). The third supports more advanced modeling, in which the data encode spatial information in addition to position, such as the direction of flow along streams. The three use models are a progression from lower value but higher volumes of use (and lower costs of development) to higher value but lower volumes of use (and higher costs of development).

The National Hydrography Dataset, the hydrography component of *The National Map*, provides an example of the partnership activities mentioned above. Federal agencies, including USGS, the U.S. Environmental Protection Agency, and the U.S. Forest Service, provided the initial data model, supplied software tools and related training to implement the model and maintain the data, and identified common business needs among the agencies and with state partners. The agencies and more than 35 state and other partners pooled funding and data, and used the software tools to develop the data within their organizations or through contractors. USGS reviewed the resulting data, incorporated them into the national coverage, and made them available. USGS has begun to accept updates from partners in the form of transactions from partner-maintained databases. These data provide the "blue lines" for portrayal of hydrography on graphics, the basis for geocoding observations of water quality, water quantity, habitat, and other characteristics relative to the hydrographic network, and the network for modeling the flow of water. In addition to uses in USGS, the data provide a geospatial basis for the U.S. Environmental Protection Agency's Watershed Assessment, Tracking, and Environmental Results project and for the exchange of geospatially referenced environmental information between the agency and its state partners, the U.S. Forest Service's Natural Resource Information System water module, the Census Bureau's Topologically Integrated Geographic Encoding and Referencing (TIGER) system modernization program, and the Federal Emergency Management Agency's flood map modernization program.

In the example of high-resolution imagery over urban areas, USGS pools its funds with those of intelligence and homeland security agencies, and reaches out to state and local governments to identify opportunities to share costs of data collection. In cases where partners are available, the federal and other organizations' funds are pooled and the USGS or partners let contracts to collect data. In cases where partners already have data that meet or exceed federal requirements, the organizations can provide the data to *The National Map* through web mapping services registered in the catalog. They also provide the data for incorporation into USGS national data holdings, although the distribution of such data might be restricted to selected federal agencies. This approach spreads the costs of data development over more uses and users than would the initial "contingency" application of disaster response, in which data would likely be used only in case of an emergency. It also helps ensure that local and federal responders would be using the same base for a "common operating picture" when there is an emergency.

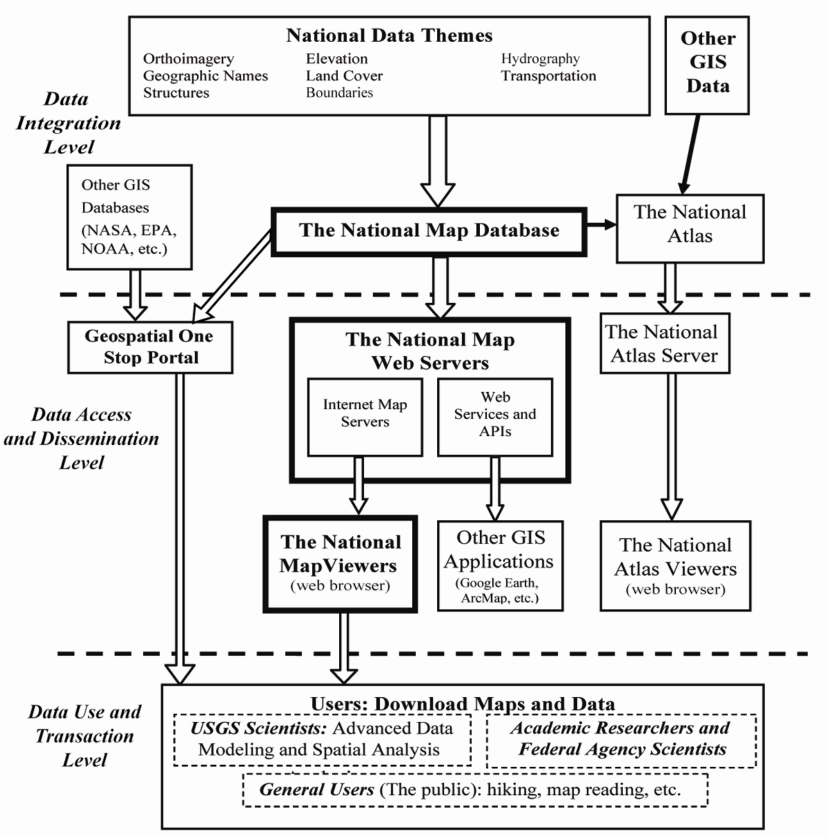


FIGURE 2.1 The current framework of *The National Map*. NOTE: API = application programming interfaces. EPA = Environmental Protection Agency; NASA = National Aeronautics and Space Administration; NOAA = National Oceanic and Atmospheric Administration.

It was beyond the scope of this committee to do a detailed analysis or critique of *The National Map*; however, based on testimony from invited experts (Appendix B), USGS science strategy documents and other published material, and informal testing by the committee, a general consensus about its current state was reached. While the concept of *The National Map*, drawing its content from local, reliable sources, and integrating those data for consistent access by all, is an excellent process for serving the nation's geospatial needs, implementation of *The National Map* requires improvement in three key areas. First, access to and presentation of the data in the current system make usage cumbersome and could benefit from user interface design and display methodologies. Secondly, the enormous benefit realized by combining data from reliable local sources brings with it difficult challenges in integrating and fusing those data with respect to scale, resolution, and quality. The current system has not automated these functions for the most part because, the technology needed is not available today. To deal with these issues effectively, the committee believes specific GIScience research is needed that will contribute not only to *The National Map* but to the field in general. Finally, with this foundation of superb data and integration, the committee believes that *The National Map* can take mapping to a new level of knowledge access and understanding, even beyond simple mapping. This requires significant restructuring of the data models in *The National Map* but it could be the key to being a world-class resource that integrates data not only over broad geographies but across varying semantics and time as well. Based on these and an analysis of the future requirements for *The National Map* described in the next section, the committee could begin to develop a list of priority research needs.

Requirements for *The National Map*

Requirements for *National Map* products should drive the design characteristics and content (Jack Dangermond, ESRI, personal communication, 2006). From the desired characteristics and content flow the practical challenges of delivering on these requirements. The challenges, in turn, determine the areas in which GIScience research could be most helpful. Based on discussions with USGS participants at the committee's first meeting and a range of non-USGS participants at its second meeting (Appendix B), and through the committee's general experience, it is clear that there is still great diversity in perceptions of the requirements, as well as the characteristics and content, of *The National Map* five years since it was first envisioned by USGS.

To the committee's knowledge, the principal requirements assessment for *The National Map* was conducted during the development of *The National Map* vision (USGS, 2001).[3] Several high-level (i.e., nonspecific) needs emerged from this assessment (USGS, 2003):

---

[3] USGS interviewed "key customers and individuals familiar with the development and use of geospatial data from five federal, two state, and two regional government agencies; nine private sector organizations; one educational organization; four professional organizations; and fifteen leaders and scientists in the USGS. A draft version of the report was posted for public comment, and this elicited 122 responses. A second, directed review was also conducted" (USGS, 2003).

- There continues to be an unmet need for a common set of basic spatial data.
- For some places, large amounts of data are available; for others, very little.
- A standing collection of basic spatial data is needed.
- These data need to be current and useful for any arbitrarily defined geographic area.
- Both digital and paper forms of basic spatial data are needed.
- Federal leadership and commitment are needed to ensure that basic spatial data are available to support federal agencies in accomplishing their missions.
- As the nation's civilian mapping agency, the USGS mission is to lead in the development and maintenance of this common set of basic spatial data.

Requirements feature heavily in the initial phase of the USGS's implementation plan for *The National Map* (USGS, 2003) (which was projected to end in FY 2005). This implementation plan proposes that requirements be collected by and fed through an advisory board. On a USGS web page dated March 2005[4] that describes the needs for *The National Map,* this entity is mentioned in the future tense. To the committee's knowledge, specific requirements have not been published.

In 2004, USGS published a cost-benefit analysis for *The National Map* (Halsing et al., 2004). The framework for estimating the benefits was "based on expected improvements in processing information to perform any of the possible applications of spatial data." Consequently, the results are, at best, only an indirect indicator of requirements.[5] Taken at face value (since the committee is not aware of an independent evaluation of this analysis), results from the "most likely" estimates of model parameters and data inputs indicate that over its 30-year projected life span, *The National Map* would bring a net present value of benefits of $2.05 billion in 2001 dollars, with a 95 percent confidence interval of approximately ± $1 billion (Halsing et al., 2004). In

---

[4] *http://nationalmap.gov/nmabout.html* (accessed August 29, 2007).

[5] The analysis did not attempt to determine the benefits and costs of performing spatial data-driven applications. Rather, "it estimates the change in the differences between those benefits and costs with *The National Map* and the current [2004] situation without it. The estimates of total costs and benefits of *The National Map* were based on the projected implementation time, development and maintenance costs, rates of data inclusion and integration, expected usage levels over time, and a benefits estimation model."

this scenario, the average time until the initial investments are recovered is fourteen years.[6]

Given the general ways in which requirements for *The National Map* have been expressed to date, the committee explored additional ways of exposing drivers for prioritization of CEGIS's research. One such approach is to consider what differentiates (or could differentiate) *The National Map* from the many other geospatial resources now available.

Unique Features of *The National Map*

The National Map resides in an environment with a large number of other electronic mapping products and services from government agencies, academia, and private industry. In particular, the emergence of Google Earth and Microsoft Virtual Earth has captured the interest of the public as well as professional users. To develop a clear understanding of the added value and niche of *The National Map*, the committee considered this crowded environment and focused on what differentiates (or could differentiate) *The National Map* from other electronic map services. Although the government does not compete with other entities in a classic business sense, it has a responsibility to provide value from its investments and to serve its customers with the highest-quality products designed precisely to meet their needs.

Who are USGS's customers? They span many disciplines, organizations, levels of geospatial knowledge, individual experiences, education, and expectations—and their needs for spatial data vary from recreational to social to professional. *The National Map* must be the trusted geospatial information source for all of these constituencies and applications. The measure of the success of *The National Map* will be the extent to which these diverse users embrace and depend on the product. That said, it is impossible to be all things to all users right from the start (NRC, 2003) and identifying what differentiates *The National Map* in the crowded geospatial product field can point toward the

---

[6] Because of many uncertainties in model assumptions and input data, the analysis considered 60 different scenarios to obtain an indication of the effects of these uncertainties. This sensitivity analysis shows that the baseline results are robust to large changes in one or more of the input values (several model parameters could be doubled, halved, or eliminated without causing substantial changes). However, the results were quite sensitive to changes in average change in the net benefit of an application implementation as a result of data from *The National Map*, rate of innovation of new applications, and amount of cumulative diffusion of data from *The National Map* and of the new and existing applications those data can inform. A few scenarios resulted in slightly negative outcomes, whereas the majority clustered around the "most likely" value.

high-impact initial research areas that help *The National Map* serve many different users.

The USGS mission and the purpose of *The National Map* establish five requirements that in combination would make *The National Map* the most trusted source of quality data and knowledge about the geography of the United States:

1. **High quality**: Data incorporated in *The National Map* need to be reliable and accurate. Critical applications including disaster response and homeland security must be able to depend on the quality of the data offered by *The National Map*.
2. **National coverage**: *The National Map* will have to represent geospatial data across the United States to serve federal, state, local, and tribal needs in as consistent a manner as possible.
3. **Accessible**: *The National Map* will need to be readily accessible to a wide constituency, from private citizens to professionals within government, academia, and industry.
4. **Continuously updated**: While the original goal of *The National Map* was to have all data current to seven days, it is more practical to implement a continuously updatable database with processes to automate data ingestion, validation, and integration. The frequency of updates should reflect data latency, which may vary according to type of data, sources, and complexity of validation. Previous versions should be retained under a temporal model from which the history of changes can be obtained.
5. **Standardized**: *The National Map* will need to reflect standards in classification schemes, metadata documentation, naming, and other characteristics so that data are easily shareable among different uses and users.

Achieving this vision for *The National Map* is partially a matter of research and partially a matter of resolve and decision making. NGPO has an opportunity to define the personality and structure behind *The National Map*. With thoughtful but clear choices, this definition will establish a nationally recognized geospatial resource.

After initial decisions to position *The National Map*, it will undoubtedly evolve, perhaps into different configurations and with specific access tools for varying constituencies. This is a benefit of an electronic product and will need to be exploited. For the important purpose of foreseeing what new areas of research would have to be conducted by CEGIS to support the evolution of *The National Map*, imagine what *The National Map* might become—a next generation that evolves from USGS's initial vision from 2001 and reflects changes in geospatial capabilities since that time (Box 2.2). The committee's vision is a means of highlighting what could be the priority areas for CEGIS research, since it will need to support a version of *The National Map* that reflects today's capabilities and their evolution. This vision is intended as a point of departure for an ongoing discussion led by USGS that should be linked to requirements for *The National Map* when these are defined.

BOX 2.2
An Updated Vision for *The National Map*

It is now 2015 and *The National Map* is one of the most visible products of the USGS—within the agency, to other federal agencies, to state and local geospatial offices, to commercial industry, and to the public. Users and organizations build their systems and models on *The National Map*. Because it is the only national trusted source of high-quality and comprehensive geospatial data, most organizations mandate the use of *The National Map* for all geography-based activities. At the public level, hikers, homebuilders, and others depend on *The National Map* for comprehensive topographic information just as in the mid-twentieth century, but now with the benefit of interactive geospatial services as a front end to a powerful and semantically rich database. Just as the paper topographic map has been distinguished by the quality of the data it represents, *The National Map* is distinguished by the quality of its database. The quality of a database is not only a function of the quality of the data it contains but of the underlying data models and the capabilities of the database management system. Thus, these have become focal points for setting a new *National Map* brand of excellence.

Capabilities

The National Map has become an intelligent topographic knowledge base capable of responding to wide ranging public needs for geographic information. This knowledge base is

- Knowledgeable about geographic features as they exist in the world, not just their cartographic representations;
- Beginning to accumulate process knowledge about how these features interact and how they change over time;
- Capable of responding to a range of different queries that might be posed by users seeking geographic information;
- Capable of delivering not just current information but information on past states and histories of features;
- In certain cases capable of making projections of future states; and
- Aware of the quality of its information and capable of communicating variation in quality and uncertainty to users.

Topographic maps and their digital counterparts play a key role in the spatial integration of information by providing a common spatial reference framework. *The National Map* has expanded its role as an integrating mechanism by establishing itself as a semantic reference system (Kuhn, 2005). The explicit semantics of geographic features imparted through ontologies associated with *The National Map* provide this mechanism.

An ontology formally defines the concepts of a domain and relationships among these concepts such that the concepts can be understood and shared by both humans and computer systems. *The National Map* is distinguished by a supporting geographic feature ontology that provides an authoritative specification of topographic features for the geospatial community. *The National Map* is further supported by a *National Map* gazetteer that is the most comprehensive source for geographic features names, place names, and feature footprints in the nation and by a geographic feature thesaurus. The gazetteer and feature thesaurus have been embedded in *The National Map* as critical pieces for its information integration role. The gazetteer provides the essential information to translate between the heterogeneous location representations (place names, addresses, geocodes, spatial footprints) used by different agencies and information providers. The geographic feature thesaurus provides the essential information to translate between equivalent or synonymous terms—in this case equivalent or similar feature classes (e.g., “cove” as equivalent to “bay”). These knowledge sources not only assist human interactions with *The National Map* but provide the basis for the expanding machine to machine and agent interactions with *The National Map*.

Information Demands

Information demands on the new *National Map* range from simple fact-based queries (what is the source of the Ohio River?), data requests across USGS data assets (streamflow data from stations on first- and second-order streams in the Potomac River watershed), complex collections of data for process model input (temporally specific fuel loadings by slope and aspect face for fire models), requests for previous states or the history of a feature (e.g., flood stages of the Missouri River over the last two years), user-defined maps (e.g., a map of the hiking trail for Mt. Katahdin, including streams, points of interest, and elevation profile), to requests for traditional topographic maps.

To respond nimbly to such a range of information requests, *The National Map* has an intelligent spatiotemporal data model based on formal specification of geographic feature concepts in the form of ontologies. *The National Map* has an individual feature database as well as the more traditional layer databases associated with the original version. The set of features (streams, lakes, canyons, watersheds, mountain ranges) and the feature database schema derive from specification of geographic feature ontologies (e.g., hydrographic feature ontology, terrain feature ontology). The central role of the individual feature database is to manage multiple spatial scale and temporal versions of the same geographic feature that result from observations of the feature at different levels of detail or resolution, from different sensor sources (aerial photography, light detection and ranging [LIDAR], ground based, or human), and from different times. The feature database is complimented by a layer database that organizes features seamlessly into their expected connectivity with other features.

Differentiation

USGS topographic maps have always maintained recognized standards of quality, and the updated *National Map* maintains the highest-quality data as embodied in standards for classification schemes (e.g., land use classifications, stream orders), official place names, relevant web services, (e.g., Open Geospatial Consortium standards), and cartographic display templates. In addition, the updated *National Map* is newly distinguished by quality-aware features that rely on the underlying feature ontologies and on the accumulation of several observations (versions) of the same feature over time. The management of multiple spatial versions of features and multiple temporal or time-stamped values for feature properties in *The National Map* database creates empirical distributions for a feature’s attributes, shapes, and positions. Such distributions provide a basis for identifying outliers and change, supporting temporal synchronization of feature versions and properties as well as other quality attributes. New versions of features being submitted to *The National Map* database can be compared against these distributions for quality assessment and recommended inclusion in or exclusion from *The National Map* database.

The updated *National Map* also offers new supporting functions to respond to user needs. It employs local validation tests in which citizens can examine *National Map* data to verify the correctness of place names, feature classes, feature attributes, features shapes and locations, and expected relationships among features (e.g., “the courthouse is north of Main Street”). By encouraging citizen participation in local validation, the reliability of *The National Map* has grown along with user satisfaction.

Also in the spirit of responsiveness to user needs, *The National Map* accommodates various map display settings from high-resolution computer screens (e.g., 5000 x3000 resolution) to small portable devices such as pocket PCs or cell phone displays (with 300 x240 resolution) and supports various mapping tasks.

Hard-Copy Products

The topographic map is a very visible feature of the updated *National Map*. Although there are an enormous number of functions and data inclusion possibilities for the power user, *The National Map* viewer includes a prominent "Topo Map" button that provides a wizard to guide the user through construction of a topographic map in the area of interest, at several selectable scales, and with a few key data choices for enhanced utility. The power user can manipulate the final product many more ways, and citizens and organizational users have ventured into these additional functions to produce customized topographic maps for their specific needs.

Performance Evaluation

The National Map is now designed for better performance evaluation. *National Map* web services track usage statistics including which features, features types, geographic regions, and information products (maps on demand, topographic maps, fact queries) are requested most frequently. The resulting statistics are available to evaluate usage patterns and help in the allocation of resources to higher-use, higher-demand needs.

In addition, *The National Map* tracks demands for new data. For example, if users in a search for data (e.g., orthophotoquads for northwestern Tennessee) find them unavailable, they can log data requests. Accumulation and evaluation of these user logged data requests highlight potentially overlapping data demands and help USGS target and prioritize new data collection initiatives.

Local Ownership

To build a productive sense of local ownership, *The National Map* now encourages local submissions of geographic information (e.g., from local names and places and feature of interest to spatial locations and alignments). Every individual or autonomous sensor with a Global Positioning System receiver or other tracking device is a potential data collector; thus, a human and sensor data collection "workforce" is being tapped to contribute generally to *The National Map* feature database (in the vein *The National Map* corps suggested for the original *National Map*). For example, a *National Map* hiking trail program has solicited and supported submission of hiking trail alignments, sights, wash-outs, and other trail features captured in real time by hikers as they are hiking.

Furthermore, *The National Map* is now rich with detailed, accurate and very fresh data from every state and virtually every county in the United States that collects and manages such data. The USGS has been successful in demonstrating to these local authorities that regular submission and updating of their data, in fact management of their data, in *The National Map* has enormous benefit to them in terms of standardized tools with available training; standardized formats and symbology, joint management of national emergencies such as earthquakes, tornadoes, and fires with national emergency response organizations; and access to seamless data in neighboring counties and states. With local authorities and organizations depending on *The National Map* for their operations, they now have a stake in its accuracy and freshness that makes *The National Map* the single most authoritative and useful source of geospatial data in the nation.

The vision for the next generation of *The National Map* outlined in Box 2.2 includes a number of additional differentiators to those listed earlier for the current *National Map* (quality, accuracy, national coverage, standardization, and continuous updates), including the following:

- Authoritative geographic knowledge base of topographic features based on a geographic feature ontology;
- Comprehensive database of official geographic feature names, and local, regional, and historic variants in *The National Map* gazetteer;
- Enhanced spatial-temporal integration framework for organization and synchronization with other USGS data collections;
- Geographic semantic reference system;
- Multiple levels of spatial detail;
- Feature histories (for spatial locations, attributes, names);
- User-supported local validation;
- Flexible product generation (e.g., response to fact queries, process model data packages, maps on demand, traditional topographic maps); and
- Smart adjustment of maps or other visual display settings for different devices.

This is a challenging but necessary and achievable vision for CEGIS, NGPO, and USGS. Its realization will require focused work and resources, but a dedicated effort has the potential to ensure a continued role for USGS as the preeminent distributor of topographic information. The new knowledge base model for *The National Map*, enabling widely diverse queries over time as well as over geography, along with highly flexible product generation means that *The National Map* has gone far beyond being just a map.

SUMMARY

CEGIS is not lacking in potential GIScience research topics. The hard decisions facing CEGIS leadership are, (1) Which among the myriad potentially useful research topics will provide the most benefit in advancing the goals of *The National Map*, NGPO, and the other disciplines? and (2) What is the right research portfolio with respect to the balance among serving *The National Map*, NGPO, and the other disciplines? On the second question, the committee favors an initial emphasis on serving *The National Map* because of its strategic importance to NGPO and USGS as a whole. On the first question, and in the absence of detailed published requirements for *The National Map*, the committee explored the distinguishing traits of *The National Map*—present and future—as a guide to what traits are most worth acquiring or developing through research.

3

Research Priorities

This chapter addresses the committee's third task—to make recommendations regarding the most effective research areas for the Center of Excellence for Geospatial Information Sciences (CEGIS) to pursue. The need for prioritization is the clear driver for this study—for, as noted earlier, there are many more research challenges than even the most optimistic assessment of CEGIS's future resources can support. The committee has already established the need for, and recommended an initial focus on, research to support *The National Map* (Chapter 2). This chapter describes and recommends research priorities under that overarching theme. Although other research topics such as visualization, cognition, and land use or land cover change are very important, the committee feels that enhancing *The National Map* will optimize initial efforts while leaving open the possibility of expanding to other topics mentioned by McMahon et al. (2005) in due course as resources allow.

The chapter has two parts. The first part describes the committee's approach to determining priorities and applies the resulting prioritization criteria to yield an initial set of priority *research areas* for CEGIS. The second part delves more deeply into priority *research topics* that fit within each of the three general research areas and demonstrates how these priorities are interrelated within *The National Map*. In the long run, this set of priorities will have to adapt to changing U.S. Geological Survey (USGS) needs and resources.

PRIORITY RESEARCH AREAS

This section defines and applies criteria for the prioritization of CEGIS research. The committee deliberated on candidate criteria based on information from meeting participants, interviews, and other inputs (Appendix B). Not only do the criteria help define broad research areas, they point to more specific priorities among focused topics within these areas. Consequently, the criteria are used again later in this chapter. The committee's eight prioritization criteria for CEGIS research follow:

Prioritization Criteria for CEGIS Research

1. *Importance to* The National Map. *The National Map* is a critical product and service of the USGS and, in particular, of the National Geospatial Program Office (NGPO). Consequently, an initial research emphasis on serving the needs of *The National Map* is a high priority. Furthermore, if applied to enhancing *The National Map*, the results will be a visible and high-profile measure of the success of such research.
2. *Importance to USGS disciplines*. After serving the needs defined by *The National Map*, the most important constituencies for CEGIS are the USGS disciplines. discipline needs and *The National Map* needs are not mutually exclusive. New capabilities for *The National Map* described in Chapter 2 are envisioned to serve the disciplines and multidisciplinary interactions.
3. *Relevance to society*. CEGIS serves not only USGS but also the nation. Its research projects will have to demonstrate high relevance to society.
4. *Solves a problem and targets a customer*. At this early stage in CEGIS's evolution and with limited resources, CEGIS will have to focus on applied research with measurable payoff. Solving key customers' problems should receive high priority.
5. *Foundational, understandable, and generalizable*. CEGIS's most important projects will be those that solve problems in geographic information science (GIScience) that have general applicability to the field and are easily comprehensible by users and customers. A measure of success in this criterion would be acceptance of CEGIS research results in a peer-reviewed publication.
6. *Enables multidisciplinary integration*. Due to the wide variety of users of CEGIS's research, the most effective research will be that which serves the widest breadth of users and supports an "enterprise solution."
7. *Focus on content*. Content is the defining ingredient provided by the USGS—whether from *The National Map* or elsewhere. CEGIS's research will need to focus on content-related issues. CEGIS may at times do

conceptual design of tools, but tool development is considered part of development engineering.

8. *Potential for early, visible success*. As with any organization, CEGIS has strongest prospects for longevity and value to USGS if it achieves and builds on early successes. CEGIS will need to target programs with this in mind.

It is important to note that these criteria are intended only as a starting point for CEGIS. From here it is essential that CEGIS continue to review this prioritization as well as take it to the next level of detail to resolve further trade-offs on what to do first within the available resource pool. These criteria for prioritizing CEGIS research point toward a program of research areas with underlying focused topics that supports users of *The National Map* data content and produces visible results in a short period of time.

Research Areas

Three broad research areas emerged from the committee's deliberations on the eight prioritization criteria:

1. *Investigating New Methods for Information Access and Dissemination*. Access to information content is a key success factor at many levels for *The National Map*. The USGS disciplines need effective data access to carry out their missions. Other federal and state agencies need effective interfaces to *The National Map* content so that their organizations can maximize productivity when working with national and local data. This priority also supports society in general because citizens need a trusted, up-to-date source of geospatial data for the nation that is flexible and easy to use. In addition, this is an area with potential for visible early success enabling interim milestones in CEGIS's longer-term research agenda.

2. *Supporting Integration of Data from Multiple Sources*. Given the diversity of source data from state and local agencies as well as many add-on themes and the desire for multidisciplinary research across USGS, achieving efficient and accurate data integration is fundamental to the effectiveness of *The National Map.* Within USGS, researchers in the various disciplines will need to find common reference data in *The National Map* and be able to load and share their data. Furthermore, the types of models and forms of spatial analysis that are increasingly needed to solve social and environmental problems will require that spatial data sets can be integrated on the fly. CEGIS will need to find solutions to integrating data with different semantics and widely varying quality, scale, and spatial and temporal granularities and resolution.

3. *Developing Data Models and Knowledge Organization Systems*. To support society in general, *The National Map* will need both the semantic flexibility of a

well-designed framework and models that enable a variety of user requests for information and information products. This objective will likely require the most research effort, but it will deliver enormous power to *The National Map* applications and lead to its clear differentiation from other web-based products.

Because all three research areas are core geographic information science research areas that are of general interest to the broad GIScience community in addition to USGS (see, e.g., DiBiase et al., 2006), CEGIS will be able to leverage ongoing research activities in this broader community. The arguments presented above lead to the following recommendation:

RECOMMENDATION 2: The three priority research areas for CEGIS should be (1) information access and dissemination, (2) integration of data from multiple sources, and (3) data models and knowledge organization systems.

PRIORITY RESEARCH TOPICS

Authorities were notified early yesterday of a fire raging in the hills around San Diego. The local fire district office immediately accessed The National Map *and displayed a topographic map of the area, including known fire trails in the hills and water resources. Given the terrain, fuel supply and impending weather, the team realized that it had a very difficult challenge on its hands and team members would be depending on technology, as well as the hard work of their crews, to deal with the crisis.*

To bring the discussion of a GIScience research agenda to life, we have woven into the remainder of this chapter firefighting and management examples in the form of a scenario of the use of *The National Map* to manage and fight a wildfire in San Diego, California. Geospatial information and tools are useful in wildfire risk assessment, modeling, monitoring, and firefighting, emissions modeling, and burn scar mapping (Rothermel, 1972; Radke, 1995; Clinton et al., 2006; Gong et al., 2006). Firefighting can benefit from accurate static geospatial data (e.g., topography) as well as dynamic information (e.g., fuel and weather) viewed in a spatial context. Improved data access, data integration, and data modeling and knowledge organization are all key to an enhanced *National Map* that can more effectively serve fire management applications as well as many others. (Note that these scenarios are intended for illustration purposes only and are not intended to reflect actual current or planned capabilities).

Each of the recommended broad research areas from the previous section encompass a range of focused research topics. These also need to be prioritized for CEGIS's research portfolio. The following three subsections describe in detail these research topics and, drawing again on the prioritization criteria listed earlier, recommend the two highest-priority topics under each research area. The order of these three subsections is driven by which research areas will likely result in early “wins” for CEGIS. Consequently, the subsections progress from near-term toward the longer-term and more challenging research. All of the research topics identified could span a broad range from basic to applied research. To provide context for the state of the art, the discussion generally begins with a description of the basic nature of the topic and lists references to relevant research. However, the recommended research questions are focused on applied research since they are motivated by the goals of *The National Map* and therefore are aimed specifically at how this research will advance the capabilities of *The National Map.* Of course, the application of this applied research does not stop with *The National Map* and will serve the other USGS disciplines as well as other agencies and users in the field. In this way, the leadership of the USGS and NGPO is demonstrated not only by the creation of a powerful *National Map*, but also by the far-reaching influence and value of the applied research the agency conducts.

Each subsection provides a general explanation of the problem; describes the relationship of the focused research topics to the USGS context (its relevance to *The National Map*, NGPO, and/or any of the USGS disciplines); and describes the maturity of the problem, approximate time frame to complete the research (near term or longer term) and in which organizations the research center of gravity resides. Although the committee did not evaluate the potential duration of research projects in great detail, in general short term is considered to be one to four years, and long term four to eight years.

Three presentation tools are utilized in this section to help clarify the main points and tie the material together. First, the specific research questions offered under each topic as starting points for CEGIS research are collected in a summary table in the final section of the chapter. Second, the aforementioned scenarios of wildfire management and operations are revisited in each subsection to illustrate how the proposed research relates to an operational application. Third, the committee uses Figure 3.1 to illustrate how the research areas and topics are linked in the context of *The National Map*.

Figure 3.1 illustrates the relationship of the recommended research topics to the overall framework of *The National Map* introduced in Figure 2.1, addressing most of its components. Colored boxes are research topics that would add a new capability or feature to *The National Map* and the three colors relate to the three research areas discussed in this chapter. The pink boxes and arrows indicate research topics covered in the section on Information Access and Dissemination. The blue boxes and arrows indicate research topics covered in the section on Integration of Data from Multiple Sources. Research topics in the yellow boxes and arrows are covered in the section on Data Models and Knowledge Organization Systems. The committee’s six recommended priority research topics for CEGIS are bolded in Figure 3.1. Box 3.1 describes how research in these areas would enhance the capabilities and functionality of *The National Map.*

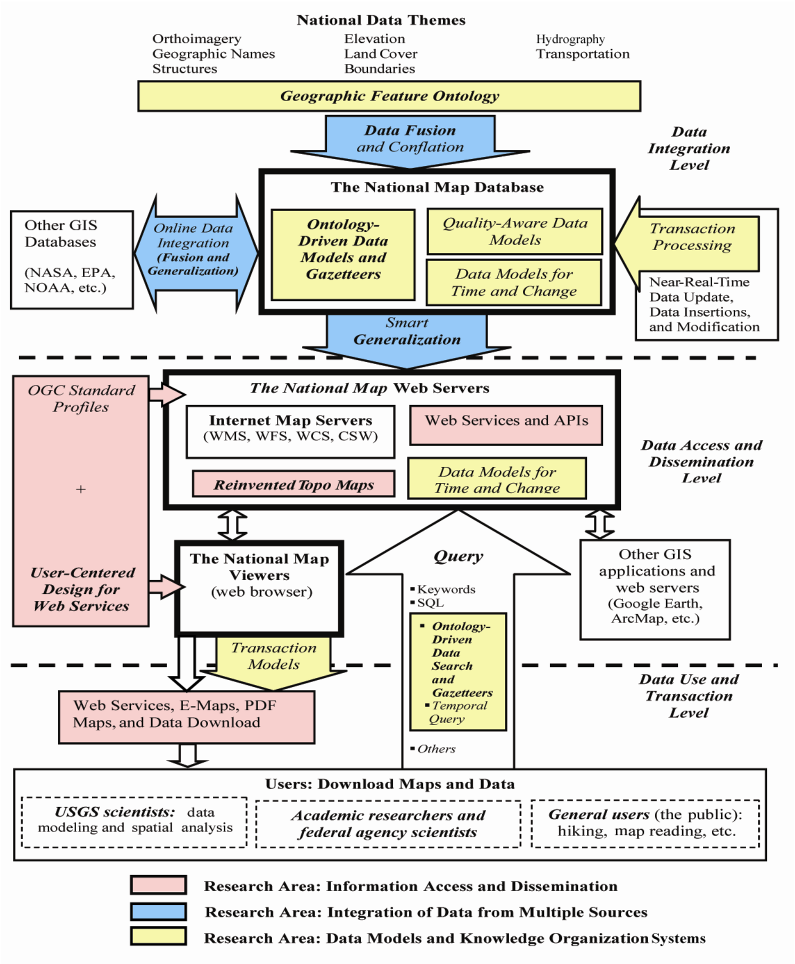


FIGURE 3.1 A potential framework for *The National Map* of the future and areas of GIScience research for CEGIS that will fuel its evolution. Recommended priority research topics are in bold within the colored boxes and arrows. This framework is adapted from that in Figure 2.1 and emphasizes *The National Map* aspects of the diagram—not those relating to the National Atlas. NOTE: API = application programming interfaces. CSW = Catalog Service for Web; EPA = Environmental Protection Agency; NASA = National Aeronautics and Space Administration; NOAA = National Oceanic and Atmospheric Administration; OGC = Open Geospatial Consortium; WCS = Web Coverage Services; WFS = Web Feature Services; WMS = Web Map Services.

Box 3.1
Benefits of the Research Topics to *The National Map*

Information Access and Dissemination

Reinvented Topographic Maps

- Provide easy public access to a valuable USGS product

User-Centered Design

- Improves usability of the human interface
- Easy access to high-quality maps in various media
- High-quality printing for all users

OGC Standard Profiles

- Facilitate a systematic framework for a distributed *National Map* computing system

Integration of Data from Multiple Sources

Data Fusion

- Integration of dissimilar data types enriches *The National Map* database
- Facilitates integration of local data with various scales, types, etc.

Generalization

- Allows automatic scaling of output to user's needs

Data Models and Knowledge Organization Systems

Geographic Feature Ontologies

- Specify feature semantics for richer data models

Ontology Driven Data Models and Gazetteers

- Organize data to support queries by place name, feature types, feature parts and multple representations

Quality-Aware Data Models

- Add ability to automatically assess quality of diverse input data

Data Models for Time and Change

- Analyze and track land feature changes

Transaction Processing

- Supports frequent data updates from distributed sources

The current version of *The National Map* has created an excellent field test of a prototype or beta version from which to build. The current *National Map* implementation reveals its strengths as well as its limitations. In fact, it is probably true that the only way to understand the highly complex information system design needs of *The National Map* is to field a prototype and measure its good and bad points. To break through the technology barriers that stand between the current *National Map* and the way it is envisioned in Figure 3.1, a thorough review of the system design is warranted. The committee has suggested one possible scenario (Box 2.2) based on current capabilities and trends, but in the long term *The National Map* system design team within USGS would feed requirements-based research challenges to CEGIS that would undoubtedly result in adjustments to the set of priorities listed in this chapter.

The National Map of the future is envisioned to be a highly dynamic and flexible transactional information system. Those transactions occur on both the input and output sides of the system, with the powerful concept of seamlessly integrating local feature level granularity data into the database in real time. A wide diversity of users access data, use tools to geospatially and temporally analyze data, and construct products and tools on top of *The National Map* in sessions with application programming interfaces on the output side. The magnitude and breadth of these transactions define the potential value of *The National Map,* but also create an information system design challenge.

In addition to the influence of evolving capabilities in GIScience research and technology on the potential framework of *The National Map*, broader trends on the web[1] will inevitably affect its architecture because the web is the delivery platform for *The National Map*. These trends will, for example, push an enhanced *National Map* toward a service-oriented rather than system-oriented approach, a collective intelligence rather than a single knowledge base, data as the driving force, lightweight user interfaces and development models for fast and reliable system performance, mapping software that supports multiple devices (e.g., personal digital assistant [PDA], cellular phone), and direct feedback opportunities that support rich user experiences and user participation.

Information Access and Dissemination

Wildfires are spreading rapidly across a San Diego mountainside. Firefighters have deployed with two-way radios and Global Positioning Systems (GPS). In the command center, the new three-dimensional topographic maps overlaid with near-real-time airborne color-infrared thermal imagery, real-time GPS wireless sensor data, and National Weather Service maps of wind direction, precipitation potential, and temperature displayed on the computers allow the command center team to tell the firefighters where the wildfire boundaries are and help them estimate the likely fire spread directions and speed in the next two hours The operators at the command center find it intuitive to toggle between the various layers of data to analyze the situation and can select different combinations to produce PDF files for fast printing to distribute to the crews. Meanwhile, the GPS and wireless communication enable the transmission of the position of the crew back to the command center, which has a large screen to display the overview maps with current positions of all firefighters and current fire perimeters. With comprehensive geographic information system (GIS) modeling technology and the information provided from The National Map *(topography, slope, aspect, weather, soil moisture, vegetation, etc.), the command and control center calculates potential dangers for firefighters and immediately distributes a warning to the crews on the west side of the mountain to relocate 300 m farther west. Based on information from the overview maps, the center also dispatches another crew to*

---

[1] See *http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html.*

the highest-risk zone and moves two more toward that zone. Their earlier participation in design phases is paying off in powerful but easy to use geospatial tools in a frantic and hostile environment.

A well-designed and user-friendly web mapping service is essential for effective use of USGS data and map products. The design of web-based USGS mapping applications is a great challenge because users can change the contents immediately by manipulating map browsers in such simple functions as zooming in, zooming out, or changing layers. The communication mechanism between map production and map users has never been as immediate and important as it is now in the rapidly expanding web mapping environment. The three GIScience research topics described in this section contribute to improved web services and map display—with the ultimate goal of improving accessibility and usability of USGS products. These topics, listed in the committee's recommended priority order based on the criteria presented earlier and beginning with the highest priority, are:

1. Innovative formats and designs to reinvent topographic maps in an electronic environment;
2. User-centered design (UCD) for implementation of *The National Map* web services; and
3. Open Geospatial Consortium (OGC) Standard Profiles for *The National Map* web mapping services and map layer design.

This subsection covers each of these foci in the order presented above. The first two topics need immediate action by CEGIS because of their fundamental value to users of *The National Map* and the potential for near-term, visible success.

RECOMMENDATION 3: The two priority research topics within the area of information access and dissemination should be to reinvent topographic maps in an electronic environment and to investigate user-centered design for *The National Map* web services.

Priority CEGIS Research Topic: Innovative Formats and Designs to Reinvent Topographic Maps in an Electronic Environment

Topographic maps are the one of the most important products of the USGS and *The National Map*. They were established in the nineteenth century (Thompson, 1988) and are the USGS's most recognized and popular map product. In the digital mapping age, CEGIS has the opportunity to conduct research that will transform the well-designed traditional paper topographic maps into an electronic, web-based, multipurpose utility. Effective delivery of topographic maps will serve both society and professionals who use this information as a base layer for analyses. This work requires immediate attention by CEGIS and can be accomplished in the short term (one to four years)—drawing in particular on the expertise of USGS's many well-trained cartographers in collaboration with software vendors with established technologies for map display. As illustrated in the vignette at the start of this section, well-designed three-dimensional electronic topographic maps will become a critical source of information in such applications as wildfire spread predictions and emergency response.

Two research foci are of particular and immediate value to the cartographic display of *The National Map*: (1) development of PDF topographic maps for wide distribution and (2) development of foreground and background data layers for control of visual hierarchies in each of the eight data layers for which USGS has responsibility in *The National Map*. These two foci arise because the available methods for creating online topographic maps using *The National Map* viewers are fairly complicated for public use—users must often select layers and symbols from among hundreds of choices; alternatively, they are confronted with a map made with all themes as strong high-contrast symbols—sometimes with confusing color choices (Figure 3.2).

PDF Topographic Maps. In the simplest case, CEGIS could develop PDF topographic maps with an associated specialized map viewer. PDF is preferred because it retains the resolution needed in print products, has wide distribution, and would accommodate viewing, saving, and printing maps by users with the most minimal computing capabilities. All topographic map symbols and layer contents are predefined by USGS cartographers. Topographic map symbol colors, widths, textures, shapes, and sizes could mimic the existing map style if scale is restricted to 1:24,000, for example.

CEGIS research needs to address design changes that accommodate changes in scale and resolution. Existing point, line, and area elements can be used at a range of scales and resolutions with minor adjustment to symbols and selection of features (Brewer and Buttenfield, 2007). Changes in symbol size and shape, line width, use of outlines, color, transparency, and texture all extend the readability of map data without requiring geometric changes through generalization.

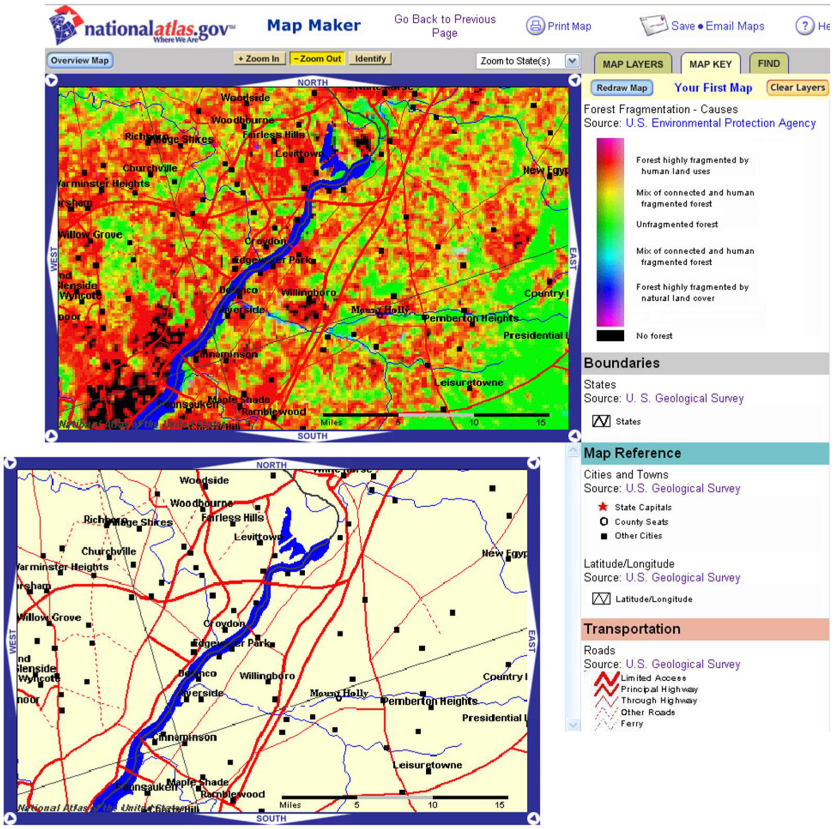


FIGURE 3.2 Multicolor forest fragmentation theme (upper pane) that interferes with the base information overlaid on it (shown separately in the lower pane). Red, blue, and black symbols in the forest theme are the same color as the base elements, making them ineffective location cues despite simple and consistent display choices. SOURCE: USGS (*http://nationalatlas.gov*).

> **Research Question:** *What is the widest range of scales that can be mapped only by adjusting map symbols combined with selectively removing feature types?* (short term)

> **Research Question:** *What is the minimum amount of change to map symbols and content that provides the maximum scale range maintaining topographic map usability?* (short term)

Advanced use of the PDF to deliver topographic maps could make use of Optional Content Groups (OCG) (Adobe Systems Incorporated, 2004). OCG allow groups of graphics to be set to "visible" or "invisible" by viewers. These are dynamic changes that may be used to mimic GIS map layer visibility settings or symbol redesign to suit smaller scales and coarse-resolution viewing. For example, a subset of layers may be set to invisible when zoomed out, or larger labels may be shown for readability at smaller scales within one PDF file delivered to the user. USGS cartographers have accumulated substantial cartographic knowledge on the design of topographic maps from their past work with paper topographic maps. This knowledge can be adapted and revised for the design of PDF topographic maps.

The reinvented USGS electronic topographic map will also need to focus on integrating GPS and accommodating various display devices. For example, the new topographic maps could include a version designed for display on portable computers and personal navigation devices (with screen resolutions of 300 x240) for general hiking purposes. The same electronic map might be converted to a design suited to a very large high-definition screen (3000 x2000) for conference and meeting presentation purposes.

> **Research Question:** *What is the stability of topographic map design (with the goal of establishing a coherent set of designs that function from coarse to fine resolutions through scale change)?* (short term)

CEGIS's goal could be to automate these adjustments so that map users are not faced with many symbol options for every device and scale at which they request topographic maps. CEGIS does not have any research in this area at this time, but it could call on USGS's long experience in paper map design as a starting point.

Visual Hierarchy. Even without enumerating all possible users and uses for spatial data, CEGIS can help reinvent the topographic map by investigating how to offer the base layers for which USGS has responsibility (elevation, hydrography, transportation, boundaries, orthoimagery, land cover, structures, and names) as foreground and as background information to allow flexible, user-determined combinations. Point, line, and area elements within each theme that form a map can be the emphasis of mapmaking (foreground) or they can be background information on which other information is overlaid. Either map elements can automatically adjust to being foreground or background or the user can be given the capability to choose between the two options. Without the ability to change the prominence of elements, map users will not be able to make readable or presentable maps without having GIS or graphics software and understanding how to download, open, and re-symbolize features. This requirement leaves behind a majority of users.

The combination of forest fragmentation information with a road and stream base map shown in Figure 3.2 demonstrates this need for controlling symbol design combinations. The National Atlas viewer offers the base information in a seemingly obvious set of symbols (Figure 3.2, lower pane). One colorful overlay belies the utility of these choices. The red for roads and blue for hydrography are used in the forest symbols (represented with a rainbow of colors common in mapping scientific data). Any color and line weight choice limits the options for representing other data, so basic style sets that can be selected to produce a usable appearance are needed. This is not a matter of refining the beauty of the display; it is a distinction between offering information or an unusable mass of data. One example of a research application in this area among the 2007 CEGIS-funded projects is provision of Internet flood mapping and flood warning layers. These detailed hydrographic data would be foreground with locational information provided as background information (base data).

Developing a visual hierarchy avoids the complexity of offering a design suited to each application, such as hiking, bus routing, or voter districting. There are, in fact, more than 100 GIS uses at a county level (Halsing et al., 2004). Customized template designs for each use would produce a cumbersome web interface. In addition, emphasis on foreground designs for each base layer offers a structured approach to custom design.

> **Research Question**: *What should be the visual hierarchies for the base* National Map *layers?* (short term)

Visual hierarchy conventions are already being developed for road maps with the private sector leading the way on implementation. Contemporary examples of symbolization schemes for roads include those from Google Maps (established earlier in print in the National Geographic Road Atlas [previously GeoSystems Road Atlas] and European topographic mapping). These designs use color and line width redundantly and a hierarchy of oranges and yellows for wider main roads and gray or white narrow lines for local roads (and offer a sound implementation of visual variable principles long established in cartographic design) (Bertin, 1983; MacEachren, 1995; Brewer, 2005). Orange-yellow-white is a series of adjacent hues that also change in lightness and saturation, making simultaneous use of these three color dimensions to build maximum contrast between symbols for enhanced readability. This combination, shown in Figure 3.3, is becoming a widely understood and clearly organized symbol set. In the Google Maps road symbol set, hydrography in blue and parks in green are examples of background information presented in desaturated low-contrast colors that push them into the background of the map. Brown urban areas and grayish-blue water in Figure 3.3 are similarly background information.

FIGURE 3.3 Example map display with detailed hierarchy of roads shown in different widths and a range of colors from orange through yellow to white. SOURCE: Original graphic made from National Atlas data sources.

Because so much work has already been done on visual hierarchies for roads, the transportation layer of *The National Map* is a likely candidate for early development of visual hierarchies to be used in *The National Map* viewer and products. The other major data themes will have to be reviewed to determine which can also build on existing work.

No automated procedures exist for designating the most relevant themes for a map purpose and allocating them to foreground and background through symbol design. Expert users could make these distinctions by manually setting priorities for each layer. They could set an order of importance for themes or a legend order, or choose among starting templates that include some initial priorities and then add additional themes to the template. Alternatively, for the nonspecialist, the web viewer could analyze the requested combination of information and select visual priorities based on a likely mixture so that, for example, a wetlands theme would prompt detail in hydrography and background representation of boundaries while population themes would prompt hydrography to be displayed with background symbolization. User profiles could also be included in procedures for setting symbolization hierarchies.

> **Research Question:** *How should USGS select a subset of automated and manual approaches to visual hierarchies to provide a tool that effectively serves the largest number and variety of* National Map *users seeking to answer geographical questions that are not served by commercial point-to-point navigation tools (e.g., Google Maps, MapQuest, and Yahoo!)?* (short term)

USGS's mapping challenge goes beyond route navigation and point location because it is responsible for multiple themes. The USGS viewer or server needs to flexibly alter symbols to bring them into the foreground or background based on map purpose and visual combination with other data that have been requested. Implementing this flexibility will accommodate the majority of user needs and will aid more adept users in getting to a useful starting symbol set when these prioritized styles are exported with downloads to be used further with GIS software.

> **Research Question:** *What is the optimal combination of types and number of symbols for an inexperienced user to create an effective topographic map and accommodate a data overlay on a topic of interest using web tools?* (short term)

This knowledge will guide further refinement of topographic mapping web resources for the public.

Priority CEGIS Research Topic: User-Centered Design for Implementation of *The National Map* Web Services

User-centered design of *The National Map* web services and viewers will improve usability by accommodating different needs for display and functionality. UCD is an interactive process of system development with user participation and evaluation (Box 3.2) and is a major research area in computer science and human-computer interaction (Nielsen, 1993, 1999; Shneiderman, 1998; Garrett, 2002). Web services are an advanced technology framework for web applications. The framework can provide high-level integration of multiple data process functions and information services hosted on different machines (web sites) (Tu and Abdelguerfi, 2006). Web services are very important for the future development of *The National Map* because they will extend *The National Map* from generic mapping functions to advanced geospatial analysis and modeling tasks. The adoption of UCD and web services approaches for *The National Map* will facilitate the integration of distributed geospatial information in USGS and improve the usability of *The National Map*.

The UCD approaches developed in computer science are ready to be adapted for *The National Map*. In the vignette at the start of this section, UCD could provide a better mapping tool to help firefighters obtain easy access to *The National Map*.

To fulfill different user needs among the USGS disciplines for displaying *National Map* layers, CEGIS research in UCD could address the design of multiple cartographic presentation methods. In fact, two of the 2007

CEGIS projects already relate to this topic. One project will create maps of ecohydrologic properties that can be used by cooperating projects and agencies. The other proposes data layers featuring 100 invasive plant species that are continuously updated with new field observations for use by the public and scientists. CEGIS guidance on how to format and represent new data layers to work with USGS base information (collected in *The National Map*) advances the scientific mission of all USGS disciplines.

Web services are interoperable and self-describing applications that can communicate with each other over the web services platform. There are many online GIS applications utilizing web services for web mapping or geocoding functions (Peng and Tsou, 2003). For example, popular Google Map application programming interfaces (APIs) and U.S. Census Bureau's geocoding services (converting U.S. streets addresses into *x,y*-coordinates) are lightweight web service examples for GIS applications.

An important concept in the development of web services is Service Oriented Architecture (SOA). SOA can allow multiple applications running on heterogeneous platforms to connect to each other and create a chain of web services for different users and applications. Interoperability and openness are the two key advantages for the development of web services. The openness of web service specifications encourages software developers to create flexible and customizable web applications based on web service standards. Interoperable web services can allow end users or service consumers to combine multiple functions and operations into a single web document for their own needs.

The OGC envisions that web services will allow future web applications to be assembled from multiple geoprocessing and location services (OGC, 2004). The adoption of web services in *The National Map* can integrate various GIS functions, maps, and data servers into a systematic web service framework rather than create scattered Internet GIS applications.

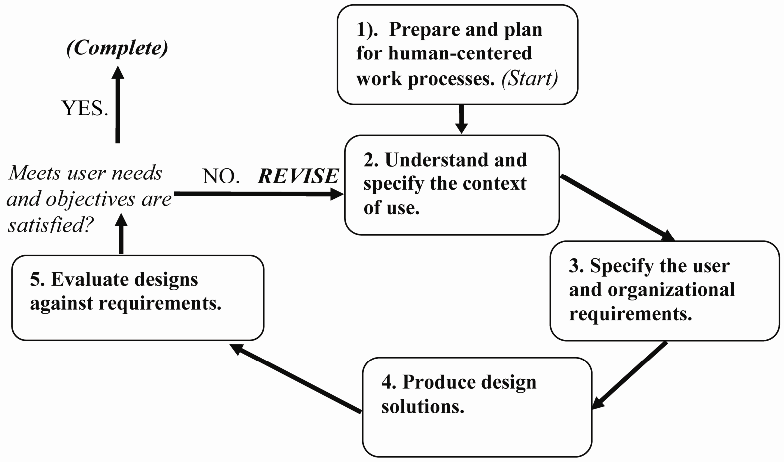
There are three major UCD-related tasks for the implementation of *National Map* web services: (1) user interface design for map viewers, (2) functional analysis for map servers, and (3) user testing and evaluation methods for *National Map* products. These short-term tasks can be accomplished by CEGIS within one to four years because the computer science community has developed comprehensive and effective UCD approaches that are ready to be used for the implementation and evaluation of *National Map* web services. Nonetheless, adopting UCD for *The National Map* and other USGS products will be challenging because it is difficult to classify multiple user groups, and what types of map content, symbols, user interfaces, and system functions are appropriate for different user groups.

BOX 3.2
International Organization for Standardization (ISO) Standard 13407: Human-Centered Design for Interactive Systems

The ISO 13407 standard includes five steps for the implementation of UCD applications:

1. Prepare and plan for human-centered work processes,
2. Understand and specify the context of use,
3. Specify the user and organizational requirements,
4. Produce design solutions, and
5. Evaluate designs against requirements.

One important aspect of the ISO 13407 standard is iterative looping through these steps (see figure below). Results from the fifth step can be applied to the product from the second step in an iterative loop of user feedback and revision (ISO, 1999) that continues until the objectives and user needs are satisfied.



The ISO 13407 human-centered design processes for interactive systems. SOURCE: Modified from the ISO 13407 Model Overview, *http://www.iso.org*.

User Interface Design for* National Map *Viewers. User Interface Design (UID) develops computer software or hardware that can be used to access and operate information stored in a computer or other storage device. In GIS, design issues often focus on the graphical user interfaces offering icons, menus, and windows with which users manipulate and display geospatial information.

CEGIS's UID focus for *The National Map* viewer will allow customization for domain-oriented applications in USGS disciplines. For example, a currently funded CEGIS research project seeks to build a "user-friendly GIS tool" for local uncertainty analysis in watershed management decisions. Confirmation of a user-friendly result requires user testing. Another current CEGIS project seeks to combine GIS, an invasive species database, and statistical capabilities for decision support using web tools, and this goal will require a user interface evaluation. A third project allows users to generate flood warnings for specific locations for end users, again requiring examination of the user interface.[2] Developing and propagating user interface evaluation techniques to the USGS disciplines developing these tools will ensure the researchers effectively meet their stated goals.

The communication methods of *The National Map* viewers are also a very important issue for user interface design. Currently, Simple Object Access Protocol (SOAP), Representational State Transfer (REST), and JavaScripts are popular methods for creating web-based mapping applications and viewers. CEGIS needs to analyze the advantages and disadvantages of various technology frameworks and methods for creating *The National Map* viewers and the user interfaces.

The user interface designs of Google Earth and MapQuest provide good examples for CEGIS on how to improve *The National Map* viewer. However, this viewer will also have unique characteristics and will need to establish and promote the USGS topographic map brand and provide access to the full depth of USGS geospatial data. This viewer will need to move beyond the point and route functionality of popular map tools to smart combinations of features with names and attributes, including networks, areal data, model output, and spatial analysis results suited to environmental and social decision making.

> **Research Question:** *With the goal of updating and evaluating* The National Map *viewer user interface, (a) what types of user interfaces are appropriate for* The National Map *viewers, (b) does* The National Map *need different viewers for different users and map contents or is a single one appropriate, and (c) what kinds of communication methods are effective for disseminating geospatial information through web browsers?* (short term)

---

[2] See *http://cegis.usgs.gov/proposals.html.*

Once these basic questions are answered, potential longer-term research topics for CEGIS in UID include investigating novel interface approaches, such as voice commands with natural language input, touch-table navigation tools, and augmented reality displays. Topics to be addressed will have to be based on what users demand—fed to CEGIS through *The National Map* system design team, contact with users in the USGS disciplines, and other forums (discussed further in Chapter 4).

Functional Analysis for* The National Map *Server. Commercial geospatial data servers (such as the Environmental Systems Research Institute's [ESRI] ArcIMS) and open source solutions (such as MapServer and GeoServer) can both provide a range of mapping functions and GIS capabilities. A major area of research is to determine which type of web mapping server is appropriate for *The National Map*. CEGIS's research would have to draw on comprehensive user needs analysis and user feedback to help select appropriate map servers.

In general, open source GIS servers can provide flexible functions and customizable user interfaces, but the developers of web applications require advanced programming skills and knowledge. On the other hand, commercial web mapping packages are easy to implement and can provide advanced mapping functions with out-of-the-box tools and user interfaces. Open source programs can provide flexibility and adaptability for complicated projects, while commercial programs come with paid support that might be better for certain applications.

CEGIS application developers will have to choose software packages that can provide enough functionality to fulfill the needs identified in functional specifications and mapping service objectives. These evaluations will include both the characteristics of mapping formats (such as image-based engines or stream-vector data based engines) and customizable GIS functions (identification, buffering, changing symbols and colors, etc.) (Tsou, 2004). Research could also assess web technologies, such as vector-compression algorithms, Asynchronous JavaScript and XML (AJAX), and Adobe Flex, that can improve the performance of web map servers. In addition, some software development platforms such as Java or NET can combine different web mapping technologies; these could also be evaluated. Selected and refined web map servers would be connected with *The National Map* database (through collaboration between CEGIS and the National Geospatial Technical Operations Center [NGTOC]).

> **Research Question:** *Will new web mapping technologies, such as vector-compression algorithms, AJAX, and Adobe Flex, improve the usability and system performance of* The National Map *servers and general web mapping applications?* (short term)

User Testing and Evaluation Methods for* National Map *Products. Iterative processes of prototyping and user evaluation are central to UCD (Box 3.2). After the implementation of a web server and map viewers, the next stage is to carry out comprehensive user testing and evaluation procedures for assessing the effectiveness of the web mapping services in the context of map use. This is an excellent way to determine the usefulness and functionality of a new system or application (Schneiderman, 1998). To help in planning future revisions, evaluation methods focus on usability and functional aspects of the prototype. Usability problems are design flaws encountered while working with a web mapping tool with poor or diminished user control, flexibility, efficiency, legibility, understandability, feedback, error prevention, visibility, ease of use, consistency, conformance to standards, and accessibility. These design flaws, if not attended to, can distract users from the overall purpose and potential of the prototype. There are many user testing and evaluation methods available for web-based GIS applications, such as expert review, questionnaires, videotaping, and web-log analysis (Schneiderman, 1998). These techniques need to be evaluated to determine which ones should be adopted for the web-based *National Map* products. The most appropriate procedures for user testing and evaluation have to be selected, and then the most useful statistical methods for analyzing the test results need to be established—again, with CEGIS collaborating closely with NGTOC and *The National Map* system design team.

> **Research Question**: *What is an appropriate standardized user testing and evaluation method for assessing and improving the effectiveness of* National Map *products?* (short term)

Open Geospatial Consortium Standard Profiles for Mapping

Interoperable web mapping services allow map users to combine multiple web-based map layers from different map servers in a single map viewer. For example, users would be able to combine any layers from *The National Map* with data layers provided by the National Aeronautics and Space Administration (NASA), the Environmental Protection Agency (EPA), the Census Bureau, or local government map servers to seamlessly integrate mapping services. In the earlier vignette, interoperability standards ensure that *National Map* data can be overlaid in real time with airborne imagery and National Weather Service maps.

The OGC is a leader in generating interoperability standards.[3] By customizing OGC Standard Profiles for *The National Map* web mapping services and

---

[3] OGC is an international industry consortium of 335 companies, government agencies, and universities participating in a consensus process to develop publicly available interface specifications. See *http://www.opengeospatial.org/ogc*.

map layer design, USGS will ensure interoperable web mapping services in all of its products and support their broader usage and accessibility. The major research challenge is to create standard profiles that are customized for USGS products yet strike a balance between OGC standards and proprietary data formats and protocols. This topic is a short-term task that can be accomplished in one to four years. Work on this topic will serve as an efficient step forward for USGS because many of its products are already built on OGC standards. There are many OGC standards associated with Web mapping applications, including Web Map Services (WMS), Web Feature Services (WFS), and Web Coverage Services (WCS).

> **Research Question**: *How should USGS create OGC standard profiles (which are a subset of standard specifications and customized standard content) to bring layers in* The National Map *databases into conformance with OGC standards?* (short term)

The committee offers three examples of where it would be useful for CEGIS to work on adding or adapting OGC standard profiles.

1. A hydrological layer in *The National Map* using WCS (as opposed to the current WMS) standards would allow inclusion of temporary information (such as pictures). This augmentation would, for example, help in recording flood danger and damage.
2. Customized metadata formats for particular USGS map layers would help meet the needs of different USGS users. Lengthy metadata with a one-size-fits-all format for all elements obstruct readability and interpretation by lay users.
3. Work on OGC styled layer descriptors (SLDs) would improve the quality and clarity of USGS graphic products. USGS cartographers are currently constrained by the limited specifications of SLDs, which contribute to the coarse character of WMS products.

OGC’s SLD extends WMS by allowing users and servers to set symbols and colors. The options for point, line, polygon, and text elements include color, width or size, opacity, texture, rotation, and halos. Lines may be outlines of polygons or centerlines and can be dashed. Text options include font family, style, weight, size, offset or displacement, and label anchor points. These are basic specifications that are limited to setting only two font characteristics—weight and size—with no dynamic variation in anchor position based on nearby features. Other cartographic needs, such as line spacing for stacked labels, curved labels, and character spacing are not specifically addressed in the OGC SLD and symbology encoding documentation.

Much current work on automated label placement focuses on point label placement. For example, Kameda and Imai (2003) extend a slider algorithm, Ebner et al. (2003) develop a force-based simulated annealing algorithm, and Stadler et al. (2006) apply a two-step approach that combines algorithms. In contrast to this automated placement work that can be applied for on-the-fly labeling, WMS such as Google Maps look good because they are built from predesigned tiles with careful label placement and line joins made in advance rather than on the fly. By studying how to combine pre-placed elements in on-the-fly combinations, CEGIS can maintain USGS's mapping prominence, enhance the combination of multiple data sets, and contribute to the quality of public data display for more map purposes than covered by Google Maps and other popular WMS. For example, the type of road and point location reference information shown over remotely sensed images on Google Maps need not be limited to transportation features for USGS mapping. Hydrographic features such as streams and springs, physiographic features such as ridges and valleys, cultural features such as post offices and landmarks, and commercial features such as gravel pits and orchards could all be annotated on imagery depending on user interests.

It is unlikely that online map users will want to download the spatial data files for an area of interest and combine them with labels in high-end graphics software. CEGIS research on automatic label placement could offer the advantage of USGS design skills for the reference labels over dynamic selections of spatial data instead of the limited WMS SLD labels that plague current viewers.

> **Research Question**: *How can USGS overlay well-positioned labels with clear categories and hierarchies on top of symbolized features dynamically set to foreground and background depending on user interests*? (short term)

Through such research, nonspecialist users would have the benefit of USGS label placement skills (see, e.g., Figure 3.4 and the stark contrast in quality between map viewer and printable PDF maps for the National Atlas in Figure 3.5) and be able to produce ready-to-read and ready-to-share mapping to support their localized decisions.

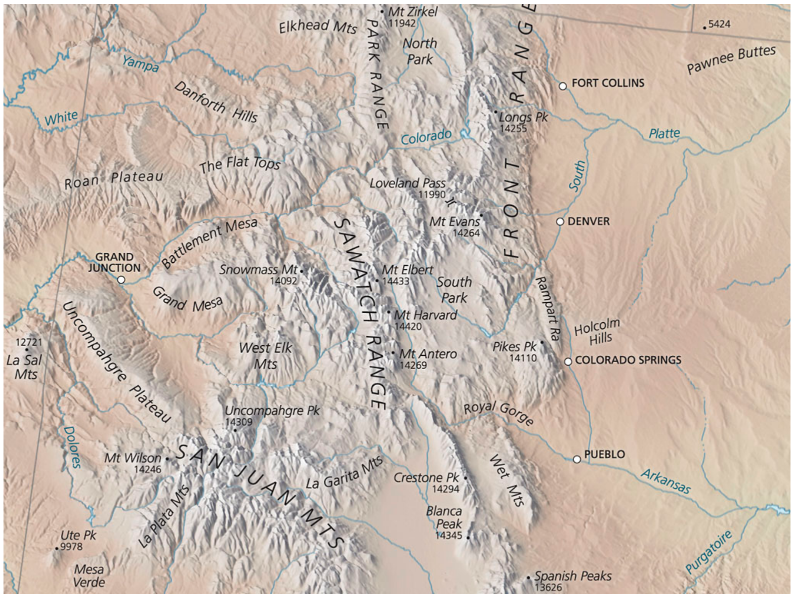


FIGURE 3.4 An illustration of the quality of label placement USGS is able to supply to annotate a landforms map by Tom Patterson, National Park Service. This label quality is not offered by Web Map Servers. SOURCE: Tom Patterson, *http://www.nacis.org/data/us_physical/gallery/10_colorado.jpg*.

To provide the most usable mapping tools, the major challenge in WMS research is to select a balance between the OGC standards and proprietary data formats and protocols (such as ArcIMS AXL or GoogleEarth's Keyhole Markup Language). To develop OGC standard profiles, CEGIS may work closely with OGC, ISO/TC 211, and NASA's Geoscience Interoperability Office. USGS has already established the brand and look of U.S. topographic maps and can build on this background to adjust and update the look for current display media. These may be delivered by OGC SLD, ArcMap style files, Illustrator templates and styles, or other style mechanisms.

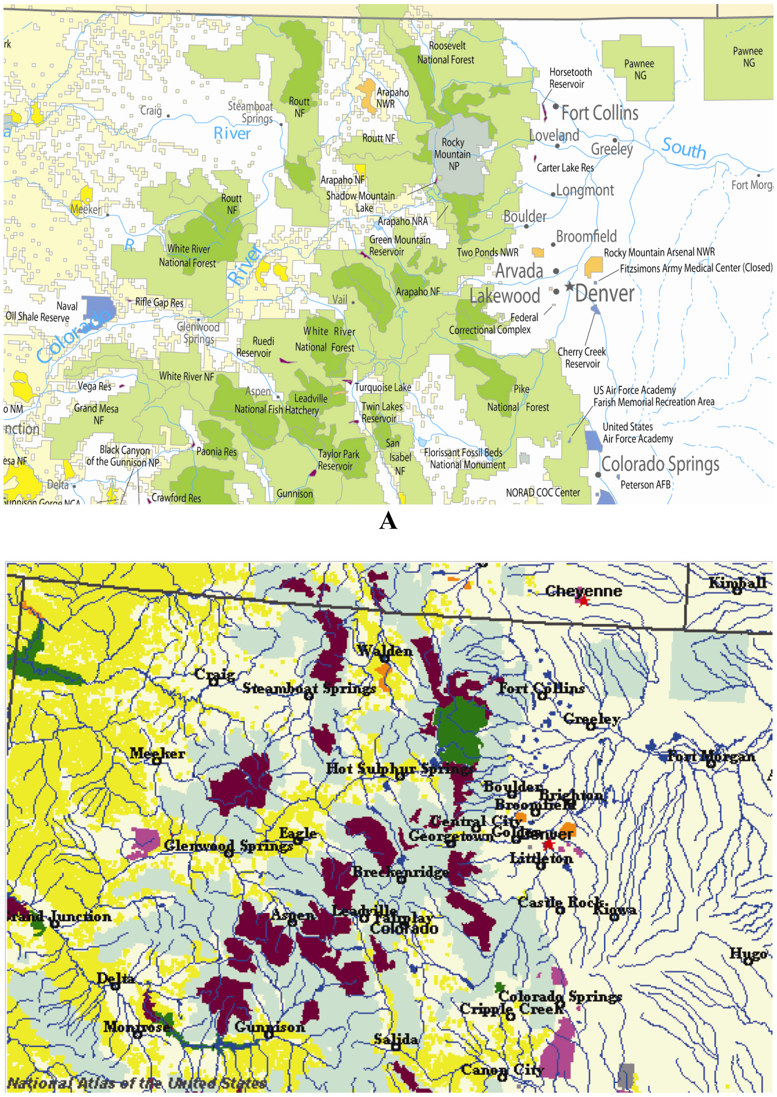


FIGURE 3.5 Differences in map design and label quality between a prepared PDF map file and dynamically generated web map both offered at the USGS National Atlas web site: (A) a portion of the Colorado Printable Map of Federal Lands and Indian Reservations. SOURCE: USGS National Atlas. *http://www.nationalatlas.gov/printable/images/pdf/fedlands/co.pdf* (B) Federal Lands map produced in the National Atlas Map Maker, comparable in purpose and extent to (A). SOURCE: USGS National Atlas, *http://www.nationalatlas.gov/natlas/Natlasstart.asp*

Integration of Data from Multiple Sources

The San Diego fire is not yet contained. The crew assesses the current boundary of the fire, overlaid on the topographic map, which explains the difficulty of containing the spread upslope. However, there is still the unexplained spread to the east. The crew accesses the National Weather Service wind forecast, which is provided at a scale of 1:125,000 compared to the topographic map at 1:24,000. The crew invokes a tool for generalization of the topographic map to the smaller-scale weather data, and a trend emerges. To determine high-priority targets, the crew calls up an address directory and uses simple controls to geocode the addresses spatially on the fire map, showing location of structures in the fire's path. To understand possible paths to fire sites, another layer with roads and another with trails are spatially matched (conflated) with the generalized map of topography. Finally, a remote sensing image with vegetation types is fused with the other layers to determine potential fuel loads for the fire path.

Integrating spatial data sets from a wide range of sources presents a fundamental research challenge for CEGIS. Spatial data sets at disparate scales, resolutions, and quality are difficult to fuse or merge, and there is a series of issues in bringing these disparate data together for spatial analysis and decision making. The most basic challenge involves the compatibility of the geometry. For example, when the original topographic data are not available can the stream network layers and contour layers be aligned? Do the streets align with the census tract boundaries? Due to myriad decisions in the creation of the original data set, including scale, resolution, data quality, and feature selection, many or most features will not exactly align. A second challenge involves the "semantics" of integration. Is a "swamp," for instance, categorized identically by two different agencies? Do two different counties classify a county road in a similar way? This involves semantic interoperability, where fusion also must occur at the attribute description level.

As an example, integration was a central challenge when comparing changes over time in census tract boundaries in the National Historical Geographic Information System (NHGIS). This system compiled all tract boundaries for the United States back to 1910 when the eastern cities were first tracted. Figure 3.6 illustrates the spatial mismatch of the 1990 Topologically Integrated Geographic Encoding and Referencing (TIGER) boundaries and the 1980 tract boundaries. When integrating these two data sets, the less accurate 1980 tract boundaries would be snapped to the more accurate 1990 TIGER boundaries.

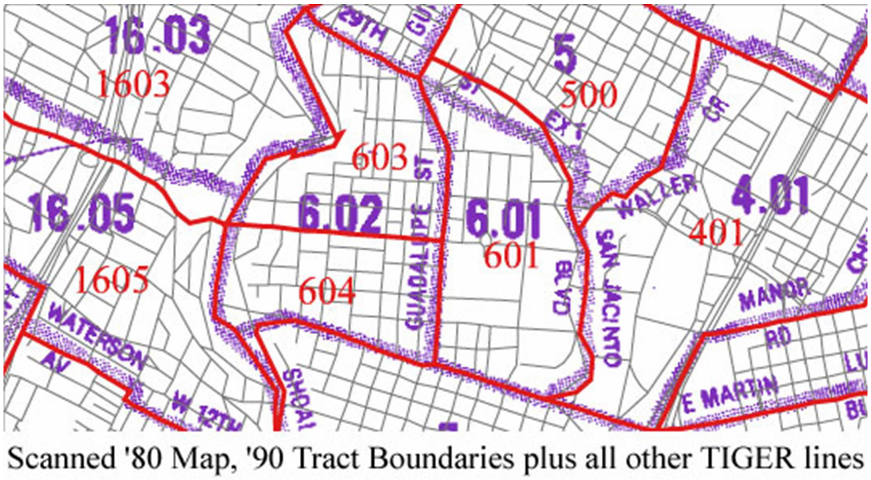


FIGURE 3.6 Mismatch between 1990 tract boundaries (red lines) and 1980 boundaries (purple lines). SOURCE: Van Riper (2003); used with permission of the author.

Given the many data sets that might be integrated into *The National Map*, including terrain and contours, vegetation, hydrography, transportation, and cultural features for instance, the success of *The National Map* will depend on the ability to integrate or harmonize these disparate sources spatially and with the accompanying attributes. Two research topics are at the heart of data integration and must be addressed early on. These are generalization and fusion (Box 3.3).

RECOMMENDATION 4: The two priority research topics for CEGIS within the area of data integration should be generalization and fusion.

Scale is a fundamental consideration in spatial data integration, and differences in scale between two data sources influence the level of the integration challenge. There are direct and strong relationships among scale, information content, and generalization. Geographical features are often scale-dependent, and appear differently depending on the scale at which they are portrayed. For example, Figure 3.7 shows same section of a stream taken from topographic maps at four different scales (1:24,000, 1:50,000, 1:100,000, and 1:250,000). The top half of the figure depicts the line segments at the original map scale, and the bottom half enlarges three so they are all represented at 1:24,000. The detail on each line segment is dependent on the scale (which influences the amount of space possible to represent the line segment at this scale). If, for instance, two hydrographic data sets were being fused at two different scales (i.e., the 1:50,000 and the 1:100,000), the 1:50,000 scale version would need to be generalized to the 1:100,000 scale version for spatial compatibility; that is, the level of spatial detail needs to be commensurate. An additional problem involves the selection of the “optimal” scale that will be needed for the given problem, such as hydrologic modeling or depicting the dispersion of West Nile Virus in a region. This relates to the problem of “scale hierarchy,” which argues there is a certain natural range of scales at which geographical processes operate—the operational scale (McMaster and Sheppard, 2004).

BOX 3.3
Data Integration, Data Fusion, and Generalization

Integration. Data integration is the assembly of information from different sources such that they work together as a whole. It is the combination of complementary information physically and logically such that applications can be written to make use of all relevant data (Jhingran et al., 2002).

Fusion. Data fusion has the connotation of physically merging data sets and is often associated with merging images and other forms of sensor data. Llinas and Hall (1998) describe data fusion as the techniques for combining data from multiple sensors to achieve improved accuracy and more specific inferences than could be achieved by the use of a single sensor. Data fusion has also been broadly defined as "the process of organizing, merging and linking disparate information elements (e.g., map features, images, text reports, video, etc.) to produce a consistent and understandable representation of an actual or hypothetical set of objects and/or events in space and time" (OGC, 2000). For the purposes of *The National Map,* data fusion will be required, for instance, to merge remote sensing images at several resolutions (e.g., 30 m Thematic Mapper ™ imagery for year 1 with 10 m SPOT (Systeme Pour l'Observation de la Terre) imagery for year 2); to merge two layers generated at different scales (e.g., 1:24,000 road layer and a 1:100,000 vegetation layer); or to bring together two shoreline data sets from different time periods. Nearly all modeling and spatial analysis routines assume that data are harmonious in terms of scale, resolution, and quality.

Generalization. Generalization reduces the information content of maps due to scale change, map purpose, intended audience, and/or technical constraints. For example, when reducing a 1:24,000 topographic map (large scale) to 1:125,000 (small scale), some of the geographical features must be either eliminated or modified since the amount of map space is significantly reduced. All maps are, to some degree, generalizations as it is impossible to represent all features from the real world on a map, no matter what the scale.

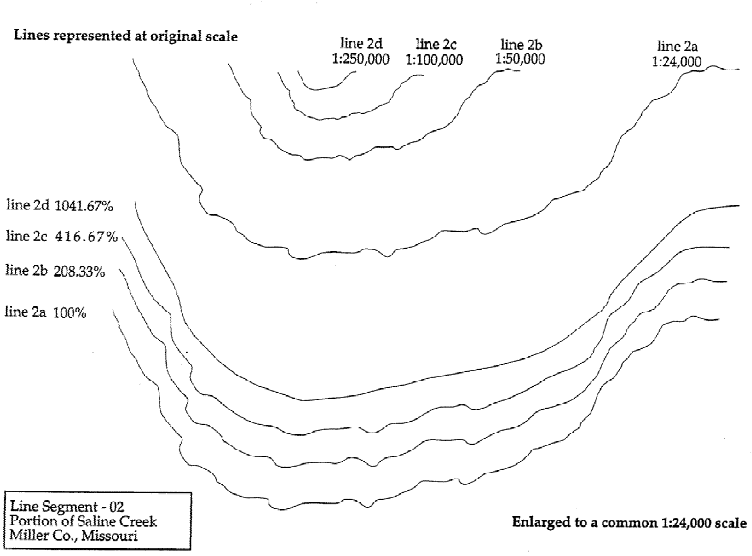


FIGURE 3.7 Four lines representing the same section of a stream taken from topographic map sheets drawn at four different scales. SOURCE: Thibault (2001); used with permission of the author.

With the goal of building *The National Map* from multiple (multiscale) data sources, methods will be needed to generalize and represent data at other scales—even down to neighborhood-level information on topics such as environmental quality or social conditions (e.g., position of the first sighting of a bird in spring, addresses of boarded-up housing, location of noxious odors) (Ghose and Huxhold, 2002). Generalization is thus required for scale harmonization before fusion. Two data sets must be basically at the same level of detail—or granularity—before they can effectively be merged.

Automated generalization and fusion have proven to be difficult research areas to move forward in a practical manner; therefore it is a challenge to identify particular research questions that meet the priority for "early wins." The following section covers generalization and then fusion research needs at CEGIS that the committee feels can best be addressed in a range of time frames.

Priority CEGIS Research Topic: Generalization

Generalization research has progressed over the last 20 years, including on algorithmic design (Regnauld and McMaster, 2007), database requirements (Mustiere and van Smaalen, 2007), and fundamental understanding of the process (Harrie and Weibel, 2007). Much of the current research in generalization is reported in a recently published book by the International Cartographic Association (Mackaness, et al., 2007). Many of the leading generalization researchers are in Europe. The Conception Objet et Généralisation de l'Information Topographique (COGIT) Laboratory at the Institute Géographique National (IGN) has worked in many areas of cartographic generalization, including data modeling to support generalization, algorithmic design and testing, and the application of agent-based methods to enable intelligent generalization. The latter project involved a collaboration among the IGN, University of Edinburgh, University of Zurich, and Laser-Scan (now called 1-Spatial)—a GIS company based in Cambridge, England. Significant work on generalization is also taking place at the University of Hannover in Germany and at European National Mapping Agencies.

In the United States, the National Science Foundation-funded National Historical Geographic Information System housed at the University of Minnesota is working to develop a multiple-scale database for census information. In addition, ESRI has developed several generalization procedures and has a small generalization team in place to improve ARC GIS's capabilities. The field is mature in the development of generalization tools, such as simplification and smoothing routines, but the research community is not as far along in understanding the importance of scale and generalization and identifying optimal scales for certain geographical processes. The USGS already has several ongoing projects related to generalization, including "Generalization for *The National Map*" as detailed in Appendix D.

Figure 3.8 shows the importance of robust approaches to generalization. In this example, Figure 3.8(a) represents raw TIGER data and the problems in coastal areas with a high density of coordinate information. Figure 3.8(b) depicts the results of the NHGIS generalization of this coastal area, while Figure 3.8(c) shows the Census generalization. Figure 3.9 shows the generalization of a piece of the Florida coastline at two different scales (1:150,000 and 1:400,000). The NHGIS generalization eliminates some larger inlets and simplifies the boundaries for an appropriate representation at the desired scale.

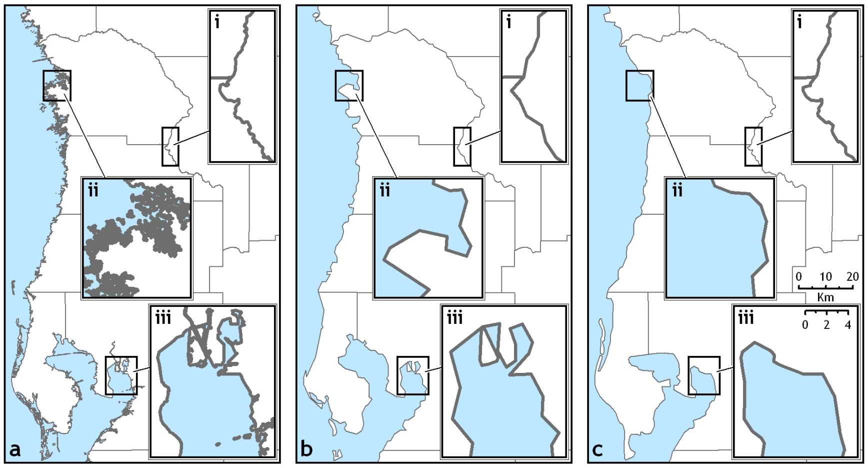


FIGURE 3.8 County boundaries along the Florida Gulf Coast drawn at 1:2,000,000: (a) base data from the Census TIGER files, with inland water extensions clipped, (b) NHGIS generalization for a 1:2,000,000 target scale, and (c) the Census cartographic boundary files. SOURCE: With permission from the NHGIS and Jonathan Schroeder.

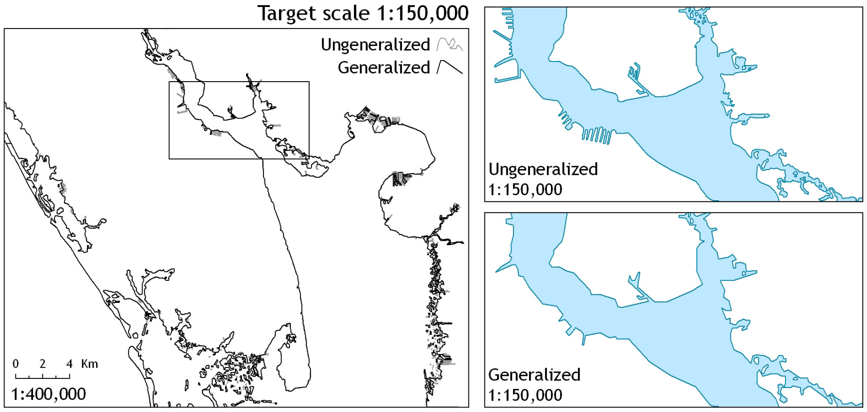


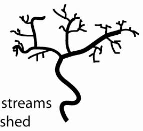
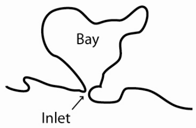
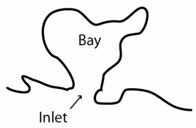
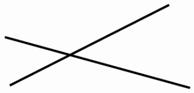
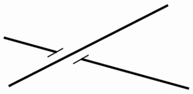
FIGURE 3.9 The Florida Gulf Coast produced a scale of 1:150,000. SOURCE: With permission from the NHGIS and Jonathan Schroeder.

In an automated environment the generalization process is complex and mathematically based. Whereas "human" cartographers have been generalizing maps for hundreds of years through the application of geographic logic, computers require exact instructions, or algorithms. The latest generalization processes involve applying one or a series of "operations." There are many such operations, and the five that CEGIS would most likely have to consider are simplification, smoothing, refinement, exaggeration, and enhancement (Box 3.4).

BOX 3.4
Types of Generalization Operations

Five possible generalization operators that CEGIS would need are discussed below. These are only a subset of a much larger collection of operations that have been developed. The five are: simplification, smoothing, refinement, exaggeration, and enhancement (see figure below).

1. **Simplification** is the most commonly used generalization operator. This involves, at its most basic level, a "weeding" of unnecessary coordinate data. The goal is to retain as much of the geometry of the feature as possible, while eliminating the maximum number of coordinates. Most simplification routines utilize complex geometrical criteria (distance and angular measurements) in selecting significant or critical points.
2. **Smoothing** (not to be confused with simplification) shifts the position of points to improve the appearance of the feature. Smoothing algorithms both relocate points in an attempt to plane away small perturbations and capture only the most significant trends of the line, or they can add points using splining routines (McMaster and Shea, 1992). As with simplification, there are many approaches to the process. Some of these operate at the local level while others process the entire line at once. Careful integration of simplification and smoothing routines can produce a simplified, yet aesthetically acceptable, result (McMaster, 1989).
3. **Refinement** is another form of resymbolization that involves reducing a multiple set of features such as roads, buildings, and other types of urban structures to a simplified representation. The concept with refinement is that such complex geometries are resymbolized to a simpler form that represents a "typification" of the objects. The example of refinement shown in the figure below is the selection of a stream network to depict the major characteristics of the distribution in a simplified form. This might be accomplished, for instance, by eliminating streams of order 4 or higher based on the attribute field for the stream.
4. **Exaggeration** is one of the more commonly applied generalization operations. Often it is necessary to amplify a specific part of an object to maintain clarity in scale reduction. The example in the figure below depicts the exaggeration of the mouth of a bay that would close under scale reduction, as would occur with San Francisco Bay or New York Harbor.
5. **Enhancement** involves a symbolization change to emphasize the importance of a particular object. For instance, the delineation of a bridge under an existing road is often portrayed as a series of cased lines that assist in emphasizing one feature over another. Enhancement also involves the enlargement of certain symbols—such as buildings—as scale is reduced and the minimum size of the object becomes too small.

Spatial Operator	Representation in:	
	Original Map	Generalized Map
Simplification Selectively reducing the number of points required to represent an object		
Smoothing Reducing angularity of angles between lines		
Refinement Selecting specific portions of an object to represent the entire object	 **All streams in watershed**	 Only major streams in watershed
Exaggeration To amplify a specific portion of an object	 	
Enhancement To elevate the message imparted by the object	 Roads cross	 Roads cross; one bridges the other

Fundamental generalization operations. SOURCE: Slocum et al. (2005); used with permission.

Work on *The National Map* will require CEGIS researchers to develop unique generalization operations that can be automated for the many possible data types and map scales. The research questions in this area vary, with short- and long-term time lines. As a first step, CEGIS will have to complete a needs assessment for generalization to prioritize work in this area. This should yield the specific generalization challenges for *The National Map*. Further refinement of this would specify those features of most relevance to *The National Map* to limit the scope of this research. In generalization and scale manipulation, locational conflicts between objects are possible, and research should address this issue. When maps are generalized and fusion is attempted, the loss of information can affect the success of the fusion process. Research is needed to define the boundaries of scale ranges for such operations. Some key questions that need to be addressed in this prioritization include the following:

> **Research Question:** *What are the specific new generalization operations and algorithms that will be needed for* The National Map*?* (short term)
>
> **Research Question:** *What feature-based generalization is needed for* The National Map *(the focus would be on a specific feature, such as a stream, and approaches needed for stream generalization) and how can that be accomplished?* (long term)
>
> **Research Question:** *What new kinds of measurements will be needed to determine locational conflicts between USGS features?* (short term)
>
> **Research Question:** *What are the effective scale ranges for fusing two layers together, and how does generalization affect fusion?* (long term)

After the assessment of needs addressed by these questions, priority will have to be given to determining which algorithms should be applied to the major feature classes. For example, this study might focus on the generalization of road networks using simplification, smoothing, and displacement. CEGIS research on generalization will have to address challenges beyond integrating *data*. These challenges include generalization for cartographic *display*, which fortunately shares many of the same approaches. Generalization is also needed to create multiscale-multiresolution databases. These databases are increasingly required by cartographers and other geographic information scientists. This approach assumes that from a master database one can generate additional versions at a variety of scales (Box 3.5).

The scope of such multiscale databases is driven by the user. For example, when mapping census data at the county level, a user might wish to have significant detail in the boundaries. Alternatively, when using the same boundary files at the state level, less detail is needed. Since the generation of digital spatial data is extremely expensive and time consuming, one master version of the database is often created and smaller-scale versions are generated from this master scale. USGS will need to carefully consider multiscale multiresolution databases in the context of *The National Map* and their relationship to geospatial data and processes. CEGIS has one project focused on generalization (see Appendix D) from which further research can be built.

BOX 3.5
Research on Multiscale Databases

Although the European literature contains a number of conceptual frameworks for automated map generalization, few have had as significant influence on U.S. researchers as the models of Sarjakoski (2007) and Brassel and Weibel (1988). The German ATKIS (the Official Authoritative Topographical Cartographic Information System; Vickus, 1995) and Kilpelainen (1997) developed innovative frameworks for the representation of multiscale databases. Assuming a master cartographic database, called the Digital Landscape Model (DLM), this research proposed a series of methods for generating smaller-scale Digital Cartographic Models (DCMs). The master DLM is the largest-scale, most precise database possible, whereas secondary DLMs are generated for smaller-scale applications. DCMs, on the other hand, are the actual graphical representations, derived through a generalization-symbolization of the DLM. The master DLM is used to generate smaller-scale DLMs (model generalization), which are then used to generate a DCM at that level. The assumption is that DCMs are generated on an as-needed basis (cartographic generalization). An additional complexity may be that the boundaries change at given times, such as decadal change for the Census Bureau's TIGER data. For many human-social and environmental databases, research is needed to develop and update multiscale versions.

Priority CEGIS Research Topic: Data Fusion

Despite more than a decade of research on this topic, fusing disparate data sources is still a significant challenge to be confronted by CEGIS. The most significant challenge is the fusion of spatial data sets that are generated from different sources, such as a Department of Transportation road layer, a USGS hydrographic layer, and a set of census tract boundaries. It is also a significant issue when considering fusing the diverse scientific data sets of the USGS into *The National Map*. Each could be at a slightly or significantly different scale, of different quality, and in different data models that must be converted. Harmonizing such disparate data sets, through the application of map conflation and edge matching, will need to be an early and high priority for CEGIS. Other related, but perhaps longer-term and lower-priority, topics include fusion across time and fusing spatial with spatial data. These approaches are described in more detail below.

The success of CEGIS and *The National Map* will require that the capability for data integration and fusion is in place at an early stage. Indeed, CEGIS recognizes that this is a significant issue and has an active research project titled Automated Data Integration (Appendix D). This project is looking at integration from both a layer-based and feature-based approach.

Conflation and Edge Matching. Map conflation and edge matching are related approaches that are often utilized to fuse data. Conflation involves first identifying features within one reference map that are accurate locations of real-world objects that need to be combined with one or more target maps (DeMers, 2003). A set of control points is identified, and their locations are reconciled with the selected objects. Next, and often in an iterative process, the features are shifted to obtain the best possible alignments. Edge matching involves “zipping” together features along map sheet edges. This process in particular will be critical for creating the seamless quality desired for layers in *The National Map*. Typically, conflation and edge-matching alignments are performed manually by identifying a set of control point pairs across different layers and then using “rubber-sheeting” techniques to align the features (Saalfeld, 1987). This approach is slow and labor intensive and usually generates a single solution across the entire layer(s) that need to be aligned, that is, the current methods do not usually allow for specific location-based solutions.

Knoblock and Shahabi (2007) developed techniques for accurately and automatically integrating vector data—representing space with points, lines, and areas—with high-resolution color imagery. Their approach utilizes automated localized image-processing techniques to find control point pairs in the pair of data sets being fused. In addition, they developed novel filtering techniques to remove inaccurate control points. Their approach does not need to locate all of the intersection points to accurately align the vector data with the imagery, and current implementations of their methods rely on triangulation and rubber-sheeting algorithms to align the remaining lines and points in the two data sets using the control point pairs (Figure 3.10). This approach, while promising, generates almost as many questions as it does answers. First and foremost, a much larger and more versatile set of tests is needed to demonstrate the efficacy of these new techniques and whether or not they can be used to integrate multiple data themes and feature types across a variety of land surfaces and cover types. These are important subtleties, and early test results (Chen et al., 2003, 2004) along with those from the standard manual techniques suggest that transportation data—the only type examined so far—are among the most likely to generate favorable results and that this outcome is especially likely when the method is applied to areas with flat terrain.

FIGURE 3.10 Automatic conflation of road vector data with imagery. This illustration depicts the conflation process applied to an orthophoto and street centerline file. After the identification of key registration points (street intersections) by the circles, the street line file is spatially aligned with the orthophoto. The right-hand illustration shows the corrected images. SOURCE: Knoblock and Shahabi (2007); used with permission.

Although theoretically *The National Map* will be a seamless database not requiring true edge matching, many other projects that integrate USGS data with other data from state and local levels will require edge-matching capability. Another process that will be needed is "splicing," sometimes called coordinate inlay, in which a part of one map (database) is spliced into a larger map such as a newly designed freeway interchange into an existing transportation layer. The challenges of this type of fusion are illustrated in the case of integrating hydrologic networks derived from topographic data (often from digital elevation models) with transportation data in which the roads often follow the hydrology. Given the different processes that generated these layers, they will not exactly match and will have to be harmonized. Conflation and edge matching are typically not automated and are therefore time consuming.

> **Research Question**: *What are the data quality issues related to spatial data integration and fusion?* (short term)

CEGIS would benefit from coordinating its existing fusion research with related activities in the National Geospatial-Intelligence Agency (NGA) geospatial research program. In that program, academic researchers are looking at topics such as Multi-Sensor Data Fusion, Analysis, and Visualization, Seamless Integration of Geospatial Data from Water to Land, Conflation Research in Support of Gazetteers, Spatial Uncertainty Models to Automatic and Enhance Fusion, and Spatiotemporal Data Fusion (NGA, 2006).

Fusion Across Time. The need for fusing data layers across time becomes apparent when considering coastal processes that are being analyzed using data such as shoreline position that change through time. Fusion methods for harmonizing such temporal data sets will be essential to realize the full utilization of *The National Map*. Although research on data fusion has developed a series of approaches over the past few decades, such as the efforts in map conflation and rubber sheeting, there is no comprehensive "fuse-layers" button to push. There remains significant work first in harmonizing the data sets in terms of scale and quality and then in approaches to spatially fine-tune the integration on specific parts of the map. The upcoming section on data models includes a related discussion on research on spatiotemporal models.

Fusing Spatial and Aspatial Data. To maximize its value, *The National Map* will likely move beyond solely fusing the traditional USGS natural resource data types and be sufficiently flexible to integrate multiple forms of social data, including census, economic, and agricultural data—many of which are aspatial.

> **Research Question**: *How can areal interpolation—as a key method for fusing aspatial data with spatial data—be applied in* The National Map*?* (long term)

Significant work on areal interpolation has been completed by Tobler (1979), with his pycnophylactic approaches; by Gregory (2002), working with the United Kingdom's historical GIS project; and in research at the NHGIS (Van Riper, 2003). Much of this earlier work has focused on areal interpolation across disparate boundaries such as census tract boundaries. Many census tract boundaries have shifted over time due to enumeration splits, and fusions will result in noncomparable statistical units (e.g., census tract 101 in 1980 may be split into 101A and 101B in 1990). The result is that a direct comparison of the population statistics, such as median housing value, are not meaningful unless the data are reaggregated. There are several procedures for dealing with such spatially noncoincident boundary files. One involves spatial interpolation, where estimates may be made by computing areally-weighted estimates *from* one area *for* another area. Using the example above, if one third of the 1980 tract 101 area was assigned to 101A in 1990 and two thirds of the area was assigned to 101B, then by overlaying the two boundary files the percentage of the area could be interpolated and used as a weight in a statistical comparison. Areal interpolation is a fairly mature research area—especially with census information, although there has been less work in integrating census data with other layers such as land use or land cover information.

Data Models and Knowledge Organization Systems

A California regional dispatch operator gets a call about a new fire that has just been spotted in Sycamore Canyon. The caller further indicates that the fire is moving quickly up the west face of the canyon. The dispatcher does not know

Sycamore Canyon or its location. Using a local geographic region profile to search the online National Map*, the dispatcher enters Sycamore Canyon and obtains a coordinate footprint of the canyon from* The National Map *gazetteer. Using the returned footprint, the dispatch system zooms to the canyon's location. The dispatcher selects an option within* The National Map *portal that uses the canyon footprint to automatically query geospatial databases housed in several different locations to obtain information on roads, streams, land cover, houses, and fire hydrants within the canyon. In addition, the dispatcher is able to select a three-dimensional image of the canyon terrain that is offered as part of the initial query results. The dispatcher clicks the west wall of the canyon to select it and adds annotation that the fire was sighted moving rapidly up this face.* The National Map *portal seamlessly integrates the retrieved streams, roads, houses, and land cover onto the three-dimensional display and the dispatcher sends the assembled data set to the fire control and command center. With this information in hand, an emergency response team departs only minutes after the call was received.*

Three important elements in this scenario are the immediate access to information based on a common place name, explicit representation of a landform feature (canyon) as a queriable object in the database, and explicit definition and representation of landform feature parts as objects (canyon wall). The feasibility of these capabilities and of rapid response in such a scenario is in large part a function of the underlying data models (Box 3.6) and supporting geographic knowledge organization systems (Hodge, 2000).

New data models and associated knowledge organization systems for *The National Map* can translate traditional topographic information into a flexible spatiotemporal knowledge base that can serve many different application areas. Transformation of *The National Map* database into a comprehensive geographic knowledge base can bring new dimensions to topographic information delivery and revitalize the role of the USGS as provider of geographic information and a valuable geospatial integration framework.

CEGIS, as a research unit of the USGS, carries with it authority for research on topographic information and hence can play a critical formative role in the development of ontologies or knowledge bases for topographic information. This section prioritizes and describes five research areas on which CEGIS could focus. These topics, listed in the committee's priority order based on the criteria presented earlier, and beginning with the highest priority, follow:

1. Geographic feature ontologies (e.g., hydrographic, terrain, or coastal feature ontologies)
2. Geographic feature data models based on these ontologies, and an associated gazetteer as an extension of the Geographic Names Information System (GNIS)
3. Quality-aware data models

4. Data models for time and change
5. Semantics-driven transaction processing

BOX 3.6
Data Models, Knowledge Organization Systems, and Ontologies

A *data model* is an abstract description of real-world entities for representation in a database management or information system. Data models are critical for effective representation and retrieval of information. They specify what entities are explicitly represented, their attributes, and the relationships among entities. The model determines the capability and flexibility to access information, manage multiple versions of information, and update information effectively. *Knowledge organization systems* are an important complement to data models. They are formalized specifications of domain knowledge that include taxonomies, thesauri, gazetteers, and ontologies. They provide important authoritative or community-sanctioned domain knowledge in forms that are explicit and shareable by both humans and computational systems.

Ontologies are one of a number of knowledge organization systems (Soergel, 2000) that help to define and organize the information resources for a domain, discipline, or institution. Ontologies specify the kinds of concepts or entities that exist or may exist in some domain or subject area and relationships among them (Sowa, 1998). An ontology, through specifications, provides formalized definitions of concepts and expected relationships among concepts for consumption by humans and computers. Geographic feature ontologies are important because they provide standardized definitions for use within the community, a basis for operational definitions of features that can support automated feature extraction, and a semantic reference framework for matching and exchanging information across communities. Geographic feature ontologies that formally specify topographic features, their parts and structures, and their relationships to other features additionally *provide the conceptual foundation for enhanced data models organized around features rather than map layers and around the semantic structure of features as opposed to abstract geometric and topological structures*.

The first two components are critical infrastructure for the geospatial and broader communities and are therefore vital initial areas of CEGIS research.

RECOMMENDATION 5: The two priority research topics in the area of data models and knowledge organization systems should be developing geographic feature ontologies and building the associated feature data models and gazetteers.

The USGS science strategy (USGS, 2007) stresses the importance of data integration and these two components in Recommendation 5 are critical to a data integration strategy. Integration and interoperability at the syntactic and structural levels have been addressed in large part by OGC standards, but integration at the semantic level needs attention. Effective semantic integration of information relies on knowledge supplied by ontologies and other knowledge organization systems. In the broader context of the web and multiple online distributed repositories of information, geographic feature ontologies and an enhanced gazetteer become essential for effective geospatial information access, retrieval, and exchange. The geographic feature ontologies are, in particular, a top research priority because as they build the conceptual foundation for the subsequent data model research components. For example, data quality-aware features and smart transactions on features depend on having a clear and formal specification of these features.

Priority CEGIS Research Topic: Geographic Feature Ontologies

Information not made explicit by a data model and knowledge sources is information not directly and easily accessible to users. While the current *National Map* contains much geographic information, it is not able to respond to many types of requests for geographic information because of the lack of explicit representation of certain features, part of features, and feature-feature relationships.

Canyons for example are implicit within a terrain model (some may have labels in the GNIS), but generally they are not explicitly modeled features for which one can query a database. Other terrain and coastal features such as bays, coves, peninsulas, gulfs, and sounds and some hydrologic features (e.g., oxbows) appear in *The National Map* layers, but they are not explicitly modeled. The current *National Map* model supports geometric and topological relationships among features but not the semantic relationships commonly used by people. Rivers and streams, for example, are connected geometrically and topologically into abstract networks, but these features are not additionally modeled and identified according to common semantic parts (e.g. mouth, source, and tributaries) as understood and likely to be requested by people. Thus, *The National Map* cannot respond directly to a simple geographic information query such as, Where are the mouth and tributaries of the Kennebec River?

In another fire example, suppose a user wishes to retrieve all canyons in California involved in fires over the last five years. *The National Map* currently has no ability to respond to such a query. One issue of course is that there is no information on fire events in *The National Map* database, but the critical point is that *The National Map* cannot respond to any query on canyons (a basic topographic feature) because it simply does not know what a canyon is. A canyon represents one basic topographic feature among many about which *The National Map* has no knowledge. If *The National Map* is expected to be more than a source of maps and become a source of comprehensive geographic information, then it must be made to “know” about basic geographic features. The point of the geographic feature ontology is to do just that—create formal specifications of canyons and other topographic features and their parts and structures such that *The National Map* is able to perform automated computations and respond to queries on geographic features.

The USGS specification of the digital line graph represented an early example of an essentially ontological specification of cartographic features. A new research increment on this work is the development of semantic specification of key geographic features. Geographic feature ontologies can formally define a set of key geographic features of interest to NGPO and more generally USGS. Recognizing the importance of research on this topic, NGPO has already embarked on exploration of ontologies for *The National Map* and funded a prospectus project on this topic in the first round of CEGIS projects (Appendix D). This is a good start, but efforts should be coordinated in such a way as to create a systematic focus on foundation topographic features. Just as the topographic map served the central role for integrating map information, an ontology that specifies the semantics of geographic features can provide a new and critical integration framework at the semantic level. The United Kingdom (U.K.) Ordnance Survey has also embarked on research on ontologies (Goodwin, 2005), and coordinated development and shared knowledge between CEGIS and the U.K. Ordnance Survey could be mutually beneficial, as well as coordination with U.S. government agencies involved in the use of geospatial data such as NGA, Department of Homeland Security (DHS), EPA, National Oceanic and Atmospheric Administration (NOAA), and U.S. Department of Agriculture (USDA). In addition, the biology community has made substantial progress in the development of ontologies (e.g., the gene ontology) and lessons from its work could be examined.

Research on ontologies is relatively recent, and ontology specification languages have only recently become readily available as open source and commercial products. A concentration of active researchers on geographic ontologies resides at the University of Buffalo. Relevant research for USGS on geographic ontologies ranges from development of upper-level ontologies (Smith and Mark, 2001; Fonseca et al., 2002; Grenon and Smith, 2004) to domain-specific ontologies (Feng et al., 2004; Sorokine et al., 2004; Sorokine and Bittner, 2005). Specification of ontologies of features implies specification of objects—and objects are assumed to have complete and closed boundaries. However, a common characteristic of geographic features is indeterminate or ambiguous boundaries (Burrough and Frank, 1996). Ambiguous boundaries create a challenge for topographic feature ontology development (Mark and Smith, 2004), and the difficulties of delimiting many topographic features are likely the reason many have remained only implicitly represented in geospatial databases. The goal of feature definition should not be to have one ultimate boundary but rather to allow the possibility of many boundaries. Just as features can have multiple names they can have multiple boundary representations. The existence of multiple boundaries for features in a database is in fact a logical expectation and outcome of observations by different sensors at different scales and different times and under different interpretation and processing. The suggested research questions in this area could be addressed in the short term.

> **Research Question:** W*hat are the key sets of topographic features portrayed within* The National Map *layers that should be explicitly represented in ontologies (these might align with the set of features already identified within the Spatial Data Transfer Standard; USGS 1994)?* (short term)

Research Question: *What are the formal operational definitions for these features, their parts and structures, and their relationships to other features?* (short term)

Research Question: *What automated feature extraction methods are derivable from these operational definitions?* (short term)

An important consideration in framing some feature definitions is the potential to translate them to operational definitions useful for automated or semiautomated feature extraction. Consider for example the definition of a bay. It can be defined as a body of water partially enclosed by land but with a wide mouth, affording access to the sea. Automatic segmentation and extraction of bays from existing digital shoreline files or image sources however would require more detailed operational specifications. Most research in this area relates to landform classification based on digital terrain models. There is less work related to identification and extraction of individual landform features from terrain models (Saux et al., 2004)—particularly on the basis of a landform ontology—although this is the objective of Mark and Smith (2004). Ontological specification of geographic features offers an effective formal basis for such work.

Priority CEGIS Research Topic: Ontology Driven Data Models and Gazetteers

Ontologies provide a framework for structuring information system content and clarify the things one wishes to model. An ontology can synthesize collective knowledge of things that exist, their properties, and relationships among them. Such synthesized knowledge of "what a canyon is" or "what a stream is," for example, can provide a framework for organizing and integrating myriad information gathered on such features by different observers or sensors, at different scales, resolutions, and time periods.

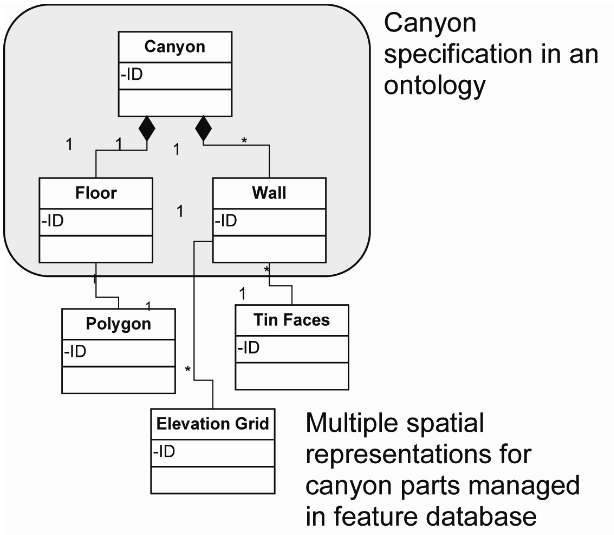
To illustrate the clarity of ontology, consider how one defines a canyon. A canyon will have a certain structure and behavior that can be specified in an ontology. Given this specification, all observations on canyons would be expected to be consistent with it. Multiple observations on any one canyon however can generate multiple versions of its location, size, shape, spatial relations to other features, and many versions of its nonspatial attributes, creating potential organizational and representational complications. The ontological clarity is that there is one underlying canyon with some invariant structure and different observations are simply multiple views of this entity. The invariant structure of a feature type should be captured by the ontology. The invariant structure of a canyon is that it has a floor and walls. The expectation would be that the feature database schema would inherit the invariant structure of the feature type as specified in the ontology and link this structure with the multiple representations obtained for each instance of a feature. A canyon and its invariant parts (walls and floor) could be associated with many spatial representations that might include different structure and resolution terrain representations or different temporal versions depicting different temporal states (e.g., pre-landslide, post-landslide).

An ontology thus provides a conceptual organizing framework for masses of heterogeneous data and information. Development of a feature based data model is a logical follow-on from geographic feature ontologies, and so is a longer-term research question. Also, there are open research questions about the nature of this association: If the ontology is modified, how are modifications propagated to the database? Do queries to the ontology propagate to the database?

> **Research Question:** *How does a geographic feature ontology operationally support a* National Map *feature database?* (long term)

Figure 3.11 depicts a canyon and its parts as specified in an ontology (gray box). A canyon can have one floor and many walls. The associated feature database includes multiple spatial representations for canyon parts that can be assembled to represent a canyon. A canyon floor can be represented with one or more polygons representing different levels of detail, and its walls can be represented with one or more sets of triangulated irregular network (TIN) faces or elevation grids.

A feature ontology provides a conceptual foundation not only for a geographic feature database but also for an enhanced gazetteer for *The National Map*. A gazetteer is a knowledge organization source that manages and translates between heterogeneous location representation forms (text [place names], geocodes, coordinate footprints) for geographic features. A place name is an implicit location reference but an important human-centered one. The gazetteer serves the essential role of connecting place names to other forms of location representation. By providing the connection across different location representations, a gazetteer can significantly expand human and machine search capabilities for geographic information and is thus another vital component of a semantic integration framework. A three-way integrated model construct that includes geographic feature ontology, gazetteer, and geographic feature database could form the basis for a comprehensive geographic information integration framework. To illustrate the potential, consider a case in which a USGS researcher wants to investigate the relationship of fire damage and landslide incidence. An example extract of a fire event database from the dispatch office is as shown in the upper part of Figure 3.12. Location is a text field that includes place names. A landslide database compiled by extraction of slide scars from satellite imagery has the format shown in the lower part of Figure 3.12. Locations are given by coordinates defining polygons. Given these two particular database configurations, no direct spatial connection between fire and landslide events is possible.



Canyon specification in an ontology

Multiple spatial representations for canyon parts managed in feature database

FIGURE 3.11 Depiction of an ontology of a canyon and multiple spatial representations as stored in a feature database.

Fire_event_ID	*Start_location*	*Start_time*	*Duration*	*Area Burned* *(acres)*
18452005	Sycamore Canyon, west face	10/13/2005	16 hours	234
19332005	Highway 262, Milpitas	9/01/2005	3 hours	89
23432005	Corral Canyon, Malibu	9/22/2004	22 hours	439

Landslide_ID	*Date_time*	*Polygon*
23	2/26/2004	1023445
24	3/2/2006	2203556

FIGURE 3.12 Example of incompatible event databases. The fire event database (upper box) has only a place name location representation, while the landslide scar database (lower box) has a coordinate location representation. Without the mechanism to translate between these different location representations (as supported by a gazetteer) there is no basis to determine collocation or other spatial relationships among these different events.

Knowledge supplied by a gazetteer and geographic feature ontology linked to a geographic feature database can make possible the automated integration of the two event databases. The gazetteer provides the translation between place names and coordinate footprints (Hill, 2006). The specification of a canyon in the ontology can define canyon structure and parts. Support for the above scenario requires further research on gazetteers that could be addressed in the short term. New data models for the gazetteer can align with ontology development and be consistent with ongoing work on digital gazetteer content standards developed as part of the Alexandria Digital Library Project (Hill, 2000; Hill and Goodchild, 2000; ADL, 2004) and OGC web gazetteer services.

> **Research Question:** *How can the collection, validation, modeling, and management of vernacular names be facilitated?* (short term)
>
> **Research Question:** *How can the creation of more detailed or smart feature footprints be automated?* (short term)
>
> **Research Question:** *How can the implications of fuzzy footprints in gazetteers be managed?* (short term)

The current GNIS lacks entries for many important named features whose locations and extents are fuzzy or incompletely defined such as mountain ranges, valleys, plains, basins, gulfs, bays, and harbors. These are limitations that may be corrected as an outcome of the ontology development. Research on gazetteer development is not extensive, but relevant research in this area includes work of the Spatially-Aware Information Retrieval on the Internet (SPIRIT) project (Jones et al., 2003, 2004), work on fuzzy features and footprints (Alani et al., 2001; Wilson et al., 2004; Fisher et al., 2004; Purves et al., 2005), and work on ontologies and gazetteers (Lutz and Klien, 2006). While there is potentially more expressive query power in having more explicitly defined features, there are also associated uncertainty issues and implications with fuzzy footprints. For example, while a query such as how many streams originate in the Appalachian Mountains? or how many islands there are in Casco Bay? may be enabled, the attendant uncertainty must be addressed.

The NGPO, with the responsibility for GNIS, is in a key position to direct and coordinate gazetteer research and development. This is a short-term research initiative that could make a substantial contribution to enhancing the spatial and semantic integration role for *The National Map*.

Development of feature ontologies, ontology driven data models, and the gazetteer has to be coordinated since the hard work of defining feature classes (e.g., canyons, mountains, bays) and their boundaries is necessary for the specification of feature footprints in the gazetteer. An important role for CEGIS in such an effort is the official sanction or standardization it can lend to this process.

Standardization of operational definitions for topographic features and hence feature footprints can lead to standardized algorithms for feature footprint extraction rather than many ad hoc approaches.

Quality-Aware Data Models

Uniformly high-quality data has been a signature characteristic of USGS topographic information. Given the changing environment in which many heterogeneous sources now contribute to the databases of *The National Map,* new challenges for data quality assessment and management arise. In the past, quality control was standards driven, and generally externally and globally applied. In other words, whole data sets were compared to independent sources of higher accuracy to assess compliance with the standard. In an environment where new data may be submitted in small increments as updates on individual features in a spatially and temporally ad hoc manner and from diverse sources, new methods to assess data quality need to be explored. An internally managed approach in which the database has built-in redundancy and benchmarks to assess quality offers some promise. A feature database that supports multiple spatial versions of the same features at different levels of detail or different temporal states creates the opportunity for such a quality-aware data model. Multiple versions of features create replicates and the potential for empirical distributions on features states.

Imagine many versions of the boundary of a lake collected over time and with different levels of detail from several different sources. As a database accumulates these versions, it begins to have the information to identify means, medians, and percentiles for feature attributes including locations. Such a strategy creates a trade-off in storage overhead for multiple versions but with the benefits of enhanced quality assessment. Tu et al. (2005) examine this problem of multiple-quality replica selection subject to an overall storage constraint. If each feature has a distribution of observed values for its various properties, the database can be designed to work with these distributions for various quality assessment tasks. These distributions can be used for example as a basis to evaluate and categorize incoming transactions from local cooperative partners. Suppose a partner submits a new GPS-generated road segment and suppose several versions of this road segment are stored in the database. The new submission can be compared with the existing set to see if it is an outlier (i.e., could represent a legitimate change, or an error) or falls “close” to the mean of the set. Such a concept raises a number of longer-term research questions around which a coherent research initiative could be built. In particular,

> **Research Question:** *How can sampling distributions of complex objects be defined and managed (e.g., reduce them to points in some N-dimensional shape space)?* (long term)

The expectation is that multiple spatial versions of features would follow a normal distribution, but statistical tests would require specification of means and variances for these complex objects. Work at the University of Maine has investigated Least Squares Collocation (LSC) and geostatistics as positional accuracy diagnostics tools (Agouris et al., 2001). The potential of simulated versus empirical sampling distribution could be explored. Various types of feature update transactions could then follow from distribution specifications. For example, new versions of a feature that are close to the mean might be rejected as redundant. Outlier versions might also be rejected as errors or alternatively checked and retained as important variants due to change or temporal state. This is a new area of research in terms of context (i.e., embedding quality assessment within database transactions for complex spatial objects), but it can bring to bear ongoing work on least squares, geostatistics and high-dimensional statistics, indexing, and dimension reduction.

Data Models for Time and Change

The National Map data model does not now support explicit representation of temporal states, change, and dynamic relationships among geographic features. Consequently, there is no framework for storage, retrieval, and access to previous states of geographic features, changes, and events. For example users cannot query for past states (e.g., "get flood states for the Kennebec River for the last five years"); query for feature states within a specified time interval and spatial range, and display returned states (e.g., "retrieve the state of lakes for April 1-15 for latitudes 44-45 degrees"); or ask for projected views for future states (what is the expected water level in Lake X for the month of April?). A range of physical process models including fire behavior models, ecosystem phonological modeling, and disease spread models (see McMahon et al., 2005, for more examples) require spatiotemporal inputs and could benefit substantially from the addition of the time dimension to *The National Map*.

The history of spatiotemporal models for GIS (Armstrong, 1988; Langran, 1992) begins in the late 1980s. Early work viewed the central unit of analysis as the spatial layer and change was conceived as modifying the fabric of a layer (Langran, 1992). More recent views include an object change view and an event view. Spatiotemporal information queries can then be done based on various spatiotemporal models (Yuan and McIntosh, 2002).

Object change-based data models (Abraham and Roddick, 1999; Worboys, 2005) can track changes in geographic features. In contrast, the event-based model focuses explicitly on tracking the change itself (Claramunt and Thériault, 1995; Peuquet and Duan, 1995; Worboys and Hornsby, 2004; Worboys, 2005; Beard, 2006). While the object-based model records the changes in the property of an object, the event model considers change as the explicit entity of interest.

For example, assume that the average pH of a lake changes over the course of a year. In the object-based model the primary object is the lake and one would track pH changes as non-spatial property changes to the lake. In the event view, a change itself, for example an abrupt drop in pH, is the specific entity of interest along with type specific event properties such as intensity, time of onset, duration, or cessation, and location.

The USGS science strategy strongly advocates development of methodologies for change analysis and *The National Map* has a role to play here. Data model enhancements that support time, change, and events are therefore central to the interdisciplinary science agenda of the USGS.

While *The National Map* has adopted the role of the topographic map as a framework for spatial information integration, it may be further investigated as a framework for spatiotemporal information integration. Records of events and processes are a basis for understanding dynamic behaviors, and USGS is already collecting and accumulating event data (seismic events, landslides, etc.). Environmental monitoring by other agencies and emerging sensor networks are creating repositories of information with high temporal resolution that support the analysis of change. Additionally, physical process-based spatiotemporal models that produce spatial snapshots in time at regular intervals (e.g., hourly, daily, annually) are used widely in such fields as hydrology, ecology, and biogeography. These models make use of geographic information layers as input but are not well accommodated by traditional GIS databases. It is important to have data models that are both spatially and temporally explicit. This is particularly the case in USGS where hydrologists, ecologists, and geographers are adopting more quantitative modeling tools and considering using geospatial data to calibrate models or vice versa. Therefore, research in this area could have great benefits to other USGS disciplines and other scientific agencies in developing new techniques for combining and analyzing spatiotemporal data.

A role for *The National Map* in such a setting is to provide the appropriate temporal as well spatial contexts in which to analyze change or event data and support spatiotemporal process models. As an example, suppose researchers have detailed spatial records of burn scars for a set of historic wildfire events and they wish to run a fire model to examine how well the model can replicate such events. Assume the fire model runs over a detailed landscape-terrain representation that include roads, structures, and land cover. The researchers want to assemble the landscape settings that are most temporally consistent with each fire date. Ideally the researchers should be able to search *The National Map* database for the spatial and temporal location of each fire event and retrieve temporal versions of the terrain, roads, structures, and land cover most consistent with the fire date. Such a scenario illustrates one potential spatiotemporal support role for *The National Map* by CEGIS that could be addressed in the long term.

Research Question: *What can be learned from spatiotemporal use cases for advancing spatiotemporal models for* The National Map*?* (long term)

Research Question: *How is change effectively represented in spatial data sets?* (long term)

Research Question: *How can process-based models be used to improve data quality or quality awareness in* The National Map*?* (long term)

Semantics-Driven Transaction Processing

The contribution of data to *The National Map* and other USGS databases from multiple local data sources and partners has a real benefit in improving the update cycle and easing the burden of centralized data collection. Indeed, distributed, locally based geographic data collection stands to substantially help the USGS maintain current, locally verified, comprehensive databases of geographic information. However, such an approach can create a substantial new burden for transaction processing (insertions, modifications, and metadata management) on these databases. Insertion transactions are likely to become much more frequent (e.g., as sensors generate near-continuous data streams), pertain more to individual features, and generate more complex metadata records given that data sources may include many different heterogeneous technologies with potentially quite different accuracy or quality characteristics. Transaction processing also becomes more complex in the more complex data model environments described above.

Revisiting the fire example, let us assume that the firefighters collected information on portable computers or from deployed sensors in the field as they were fighting the fire. At the end of the day, the goal is to distribute this information as updates to appropriate databases. Suppose the information collected by the firefighters includes estimates and extents of burned areas and an inventory of burned structures. Several long-term questions arise on what the transaction processing logic is for updating *National Map* or other USGS databases. The firefighters are not expected to be database experts and so need support for simply uploading the data. There is, however, complex transaction processing logic that stands behind uploading these data to the correct databases. For example the records of burned areas could be added to a fire events database. In addition, it may be appropriate to update a set of land cover databases and associated products. *The National Map* includes several land cover and associated products, and the transaction logic would have to consider whether some or all of these should be subject to fire updates. The transaction logic might be such that only those products in which the resolution or granularity of cover classes matches the extent of the burn area are subject to updates. Coarse land cover products might be immune to small burn area updates. On the other end of the granularity spectrum, if the fire data are sufficiently detailed to indicate differential burn on different cover types, the transaction logic might be that differential burn damage information is applied to different land cover classes. The information on destroyed structures requires similar sets of transaction decisions. Presumably the structures database should be updated with information that structures X, Y, and Z were destroyed by fire on the given date. A follow-on question could be what additional databases and sources should be updated? For example, should any high-resolution images depicting these structures be updated?

> **Research Question**: *What is the transaction processing logic for complex spatiotemporal transactions among* National Map *and other USGS databases*? (long term)

Research in this area resides predominantly in the database research community. Transaction processing generally is a mature field, but spatial and temporal transaction processing and transaction processing in distributed database contexts are still new. The OGC Transaction Web Feature Server is addressing the ability to create, update, and delete geographic features in a distributed computing environment and CEGIS may consider some collaborations with OGC in developing distributed *National Map* transaction processing. Open research issues remain with respect to spatial transaction processing on multiresolution databases. Kafeza et al. (1996) and Rigaux and Scholl (1995) describe approaches to transaction processing in multiscale and multiresolution environments, and there is relevant work on transaction processing for mobile systems (Hampe and Sester, 2004). Hampe and Intas (2006) have recently proposed extensions to OGC Web Feature Services standards to support transactions on multiple representation databases. Transactions in spatiotemporal databases must address issues of when or how frequently new versions or states of feature properties are updated. Some update transactions may be event driven as in the case of the fire event described above. CEGIS might investigate what USGS or other agency databases record events (e.g., earthquakes, landslides, floods) and consider the automation of *National Map* database update transactions in response to such events.

SUMMARY

This chapter has laid out a recommended research priority structure for CEGIS, with three priority research areas, each broken down into research topics. The two highest-priority topics are recommended for immediate action in each area, with other important topics described as well for the purposes of longer-term research planning. Specific research questions for each topic are also suggested as potential starting points. Table 3.1 summarizes this research structure and is organized to show the broad research areas, the recommended research topics within those areas, and the committee's suggested initial research questions. Building on this foundation as resources allow and requirements evolve, CEGIS can expand its research portfolio to address a broader range of key GIScience issues of national relevance.

TABLE 3.1 Summary and Time lines of Recommended CEGIS Research Areas, Topics, and Questions

Research Area	Research Topic (Bold = Priority Topic)	Research Questions	Time Range
Information Access and Dissemination	**Innovative formats and designs to reinvent topographic maps in an electronic environment**	1. What is the widest range of scales that can be mapped only by adjusting map symbols combined with selectively removing feature types?	Short term
		2. What is the minimum amount of change to map symbols and content that provides the maximum scale range maintaining topographic map usability?	Short term
		3. What is the stability of topographic map design (with the goal of establishing a coherent set of designs that function from coarse to fine resolutions through scale change)?	Short term
		4. What should be the visual hierarchies for the base *National Map* layers?	Short term
		5. How should USGS select a subset of automated and manual approaches to visual hierarchies to provide a tool that effectively serves the largest number and variety of *National Map* users seeking to answer geographical questions that are not served by commercial point-to-point navigation tools (e.g., Google Maps, MapQuest, and Yahoo!)?	Short term
		6. What is the optimal combination of types and number of symbols for an inexperienced user to create an effective topographic map and accommodate a data overlay on a topic of interest using web tools?	Short term

User-centered design for implementation of *The National Map* web services	1. With the goal of updating and evaluating *The National Map* viewer user interface, (a) what types of user interfaces are appropriate for *The National Map* viewers, (b) does *The National Map* need different viewers for different users and map contents or is a single one appropriate, and (c) what kinds of communication methods are effective for disseminating geospatial information through web browsers?	Short term
	2. Will new web mapping technologies, such as vector-compression algorithms, AJAX, and Adobe Flex, improve the usability and system performance of *The National Map* servers and general web mapping applications?	Short term
	3. What is an appropriate standardized user testing and evaluation method for assessing and improving the effectiveness of *National Map* products?	Short term
Open Geospatial Consortium (OGC) Standard Profiles for *The National Map* web mapping services and map layer design	1. How should USGS create OGC standard profiles (which are a subset of standard specifications and customized standard content) to bring layers in *The National Map* databases into conformance with OGC standards?	Short term
	2. How can USGS overlay well-positioned labels with clear categories and hierarchies on top of symbolized features dynamically set to foreground and background depending on user interests?	Short term

Integration of Data from Multiple Sources	**Generalization**	1.	What are the specific new generalization operations and algorithms that will be needed for *The National Map*?	Short term
		2.	What feature-based generalization is needed for *The National Map* (the focus would be on a specific feature, such as a stream, and approaches needed for stream generalization) and how can that be accomplished?	Long term
		3.	What new kinds of measurements will be needed to determine locational conflicts between USGS features?	Short term
		4.	What are the effective scale ranges for fusing two layers together, and how does generalization affect fusion?	Long term
	Data fusion	1.	What are the data quality issues related to spatial data integration and fusion?	Short term
		2.	How can areal interpolation—as a key method for fusing aspatial data with spatial data—be applied in *The National Map?*	Long term
Data Models and Knowledge Organization Systems	**Geographic feature ontologies**	1.	What are the key sets of topographic features portrayed within *The National Map* layers that should be explicitly represented in ontologies (these might align with the set of features already identified within the Spatial Data Transfer Standard; USGS, 1994)?	Short term
		2.	What are the formal operational definitions for these features, their parts and structures, and their relationships to other features?	Short term
		3.	What automated feature extraction methods are derivable from these operational definitions?	Short term

Ontology driven data models and gazetteers	1.	How does a geographic feature ontology operationally support a *National Map* feature database?	Long term
	2.	How can the collection, validation, modeling, and management of vernacular names be facilitated?	Short term
	3.	How can the creation of more detailed or smart feature footprints be automated?	Short term
	4.	How can the implications of fuzzy footprints in gazetteers be managed?	Short term
Quality-aware data models	1.	How can sampling distributions of complex objects be defined and managed (e.g., reduce them to points in some *N*-dimensional shape space)?	Long term
Data models for time and change	1.	What can be learned from spatiotemporal use cases for advancing spatiotemporal models for *The National Map?*	Long term
	2.	How is change effectively represented in spatial data sets?	Long term
	3.	How can process-based models be used to improve data quality or quality awareness in *The National Map*?	Long term
Transaction processing	1.	What is the transaction processing logic for complex spatiotemporal transactions among *National Map* and other USGS databases?	Long term

4

Realizing USGS's Vision for CEGIS

The success of the U.S. Geological Survey's (USGS) Center of Excellence in Geospatial Information Science (CEGIS) will depend on the relationships and network it develops to conduct its research and establish national leadership. From a small and lean beginning, and through such a network and relationships, CEGIS is envisioned to "conduct, lead, and influence the research and innovative solutions required by the National Spatial Data Infrastructure (NSDI)" by conducting, supporting, and collaborating in "research to address critical Geographic Information Science questions of importance to the USGS and to the broader geospatial community" (CEGIS, 2006). This is a grand vision toward which USGS and CEGIS can build. This chapter presents ideas on CEGIS's first steps toward fulfilling its vision.

LEADERSHIP

In creating CEGIS, USGS recognizes that it needs a research "hub" that identifies and prioritizes geographic information science (GIScience) research challenges, finds answers available in the geospatial community, and sponsors new research projects to develop new answers and field solutions. Such a center provides an opportunity to develop a research culture, a sense of identity, a place in which to mentor young researchers, and a proactive intellectual community (McMahon et al., 2005). In addition, CEGIS has the opportunity to lead the way in showing how such a center can provide an enterprise solution to geospatial challenges confronting an entire organization. "CEGIS should be a model for the world of how to organize a GIScience research agenda" (Michael Goodchild, University of California, Santa Barbara, personal communication, 2006). Not only would CEGIS provide the structure to conduct research—its approach could establish GIScience leadership for the USGS.

Indeed, in proposing that CEGIS be established, McMahon et al. (2005) stated, "The USGS must redevelop and reassert its leadership role in GIScience." Until the 1980s, the USGS was the nation's leader in collecting, processing, producing, and distributing spatial data. USGS researchers drove the national research agenda, and several of the major research conferences, including the early Auto-Carto events, were cosponsored by the USGS. The geospatial community looked to the USGS as a source of high-quality spatial data and maps, for coordination of access to spatial data, and as a place where some of the world's best ideas emerged on spatial processing, projections, uncertainty, and visualization.

In recent years, however, the USGS has undertaken several transitions that have weakened its national leadership in these aspects of GIScience. Although other entities such as the University Consortium for Geographic Information Science (UCGIS) have stepped in to coordinate national-level geospatial research, they have even more limited resources than USGS to do so. A dynamic, nimble, cutting-edge research unit at CEGIS could lead a nationally important GIScience research agenda that more effectively focuses and harnesses the nation's GIScience assets in academia, industry, government, and nongovernmental organizations. Also, even though *The National Map* will be the initial impetus for identifying research challenges (Chapter 3), solutions to these problems will be of broader interest and be applicable throughout the geospatial community.

MODELS FROM WHICH CEGIS CAN LEARN

No two mapping agencies are identical because each has different demands and constraints. Nonetheless, the committee received a range of information on how research is conducted within mapping organizations around the world and within the United States, and drew upon the lessons of these organizations as it deliberated on ideas for recommending an approach that CEGIS might follow. Case studies from three such organizations are presented in this section. The Conception Objet et Généralisation de l'Information Topographique (COGIT) Laboratory in France (Box 4.1) and the research unit within the Ordnance Survey (OS) in the United Kingdom (Box 4.2) are both successful research laboratories that support foreign national-level mapping capabilities. The Basic and Applied Research Office of the National Geospatial-Intelligence Agency (NGA) (Box 4.3) supports U.S. geospatial intelligence provision capabilities.

BOX 4.1
Institut Géographique National COGIT Laboratory

The COGIT Laboratory is the research unit of France's Institut Géographique National (IGN). Over the past 15 years, the laboratory has tackled some of the thorniest problems in GIScience including automated generalization—a topic in which it is viewed as the international leader. Its research, although basic in nature, also supports the IGN production units. Its reputation throughout the European research community is one of excellence.

COGIT is supervised by IGN and funded directly from France's Ministry of Research. It has a staff of approximately 20, including the director, 7 Ph.D.-level scientists, 7 Ph.D. students, and several engineers who mostly come from the IGN school. The laboratory also can sponsor visiting faculty members and has brought in international researchers to work on key projects, often for periods of one to three weeks. Some of COGIT's research projects are developed in collaboration with other European agencies and the private sector.

Every five years the laboratory creates a research plan that is approved by the IGN through a series of consultative meetings. Research teams of two to five people focus on problems identified in the plan. Currently, five such research teams are in place, working on research issues including (1) helping access to geographic information; (2) colors and legends; (3) automation of generalization; (4) integration and multiple representation; and (5) spatial analysis. Increasingly, COGIT researchers are presenting their work at international conferences such as the International Cartographic Association, and they are evaluated, in part, on the quality of their publications.

SOURCE: Anne Ruas, COGIT.

BOX 4.2
Research at the United Kingdom Ordnance Survey

The research facility at the U.K. Ordnance Survey (OS) provides an engine of innovation and insights. It is a knowledge store for the organization and positions the OS as a thought leader. The facility functions as a radar screen for new technologies that will impact the organization and its partners. Its research is internally focused on the needs of the organization, which generates all of its operating revenue from licensing its information products and services.

The research unit comprises 30 researchers and support staff. The majority of staff consists of graduate students and postdoctoral fellows. A shift toward a higher percentage of postdoctoral fellows over graduate students is under consideration because of the likelihood of a tighter research focus and speedier return on investment in the case of postdoctoral scientists.

The annual budget of the research unit is approximately $4 million (approximately 2 percent of OS's revenue), of which one-quarter goes to research contracts with universities that are primarily located in the United Kingdom. The unit collaborates on research with other U.K. government agencies and with other European mapping agencies. In addition, it has joint industry research projects.

Having previously followed a product development approach to managing its research needs, the facility has, in the last two years, adopted a "portfolio" approach. This approach includes short-, medium-, and long-term goals that are generated internally within the research unit and taken to the OS governing council, which weighs their value to the OS business. The approved goals then define a series of research priorities, and the balance of investment on each topic is influenced by the likely level of success.

As part of its goal development process and to cement its role as a thought leader, the research unit hosts "Terrafuture"—an annual conference that focuses on societal challenges over the next 10 to 15 years and how they could affect research. The current foci of research within the unit are on data capture, data modeling, and semantic technologies.

SOURCE: Duncan Shiell and Ed Parsons, U.K. Ordnance Survey.

BOX 4.3
Basic and Applied Research Office at the National Geospatial-Intelligence Agency

The Basic and Applied Research Office (BARO) "executes cutting-edge scientific research, providing the foundation for . . . solutions to NGA's most difficult technical challenges." BARO addresses NGA's "unsolvable" problems, investigates the scientific issues and basic phenomena surrounding "hard" problems, and demonstrates rapid proof of concept based on basic and applied science. The office draws no hard line between basic and applied research—it prefers the term "use-inspired research."

Research priorities within BARO are initially considered by its senior scientists. These priorities are then reviewed by senior agency officials and subsequently scrutinized by the broader intelligence community (the ultimate users of NGA's geospatial intelligence products) when NGA presents its budget case. Priorities emerge in three areas: (1) technical and methodological capabilities that might lead to breakthroughs, (2) exploiting new data sources, and (3) pressing concerns among geospatial intelligence analysts. Because NGA is operations driven, item 3 tends to carry the greatest weight. In general, the research priorities are influenced by where investments will make the biggest difference.

NGA performs research in many ways (NRC, 2006). Most research is contracted out, and the in-house technical staff consists predominantly of program managers. BARO's academic connections focus on funding research, attracting talent, and training.

With respect to research funding, BARO supports NGA University Research Initiatives, Historical Black College and University-Minority Institution Research Initiatives, and Intelligence Community Postdoctoral Fellowships. Recipients of awards under all three initiatives gather annually to report on progress and share ideas.

To attract talent, BARO utilizes a visiting scientist program that brings young scientists from undergraduate to postdoctoral level to work in a classified environment at NGA facilities. This helps NGA to enhance technologies, tools, and methods; leverage academics specializing in NGA's scientific areas of interest; build long-term relationships with top universities; build recruitment opportunities for NGA; and obtain experienced Ph.D.-level scientists on a temporary basis to augment in-house expertise.

As part of its efforts to strengthen the academic base through training, NGA supports three academic centers of excellence, four service academies (e.g., U.S Air Force Academy), and ten intelligence community academic centers of excellence.

In addition to its academic partnerships, BARO collaborates with other government entities such as the Defense Advanced Research Project agency (DARPA), the National Science Foundation (NSF), and Department of Energy laboratories. NGA works with USGS on domestic emergency response in an operations context, but does not currently collaborate on research.
SOURCE: Beth Driver, NGA.

CONSIDERATIONS FOR BUILDING AND OPERATING CEGIS

Karen Siderelis (Associate Director for Geospatial Information, USGS) encouraged the committee to think in terms of what is needed for CEGIS to succeed and, while being practical, not to be constrained by the current dimensions and budget of the center.

In pondering how CEGIS might achieve its vision, the committee considered a series of points, many of which were raised during discussions with presenters (Appendix B) at the first two meetings. At the general level, CEGIS is essentially a new start. This presents a rare opportunity for those framing its future, and their initial actions will set a tone for CEGIS's business practices, performance expectations, and accountability. Other points fall under one of two categories: research portfolio or research management.

Research Portfolio

Early "wins" for CEGIS will be important for establishing confidence in the activity, developing a track record, experiencing the full "life cycle" of a research project, and obtaining feedback. A research portfolio that balances this early-win research with longer-term, and perhaps more risky, research works well for the U.K. Ordnance Survey (Box 4.2). Such long-term research may require different approaches than the shorter-term research.

The research portfolio will need to concentrate most resources on operational research projects that have research objectives with performance outcomes and milestones. This will provide a baseline for realistic management decisions about budget (what can CEGIS afford?), staffing (who should CEGIS hire, and can it afford them?), and opportunities to work with other organizations inside and outside USGS (with whom should CEGIS work, and why?). Developing a research portfolio will also provide CEGIS and other program management personnel with "practice" in setting priorities and working together.

A tight focus of CEGIS's initial research activities on supporting *The National Map* would provide a purpose that reinforces, and does not distract from, other program activities. It also provides access to problems of partners and customers that generate new CEGIS research activities and demonstrates the need for the activity. One such partner and customer for CEGIS is the National Geospatial Technical Operations Center (NGTOC). Operational challenges faced by NGTOC would need to be presented to CEGIS in the form of research questions on which to develop solutions. NGTOC would then implement those solutions into the structure and operations of *The National Map*. In addition, CEGIS staff would be involved in the system design team for *The National Map*, and this process would also provide research questions for CEGIS. Other partners and customers that may develop research questions for CEGIS include the water, geology, geography, and biology disciplines of the USGS and, as the agency forms relationships with other agencies that come to depend heavily on *The National Map*, potentially the Federal Emergency Management Agency (FEMA), the U.S. Department of Agriculture (USDA), and Department of Homeland Security (DHS) and others, these agencies could also provide input to the research agenda of CEGIS.

Research Management

The Office of Management and Budget provides criteria for managing research. These include relevance, quality, and performance (Box 4.4). To fully benefit from the wealth of GIScience research in science, technology, and operations, CEGIS will need to follow investments made by others and look for opportunities that could be adapted with CEGIS involvement to support USGS operations. This approach has an added advantage of allowing CEGIS managers to learn best practices for managing a larger portfolio of technology and scientific research as resources permit and needs require. In relation to managing resources in particular, opportunities to reduce overhead costs and maximize the allocation of resources to research activities will maximize the benefit of the initial budget.

To ensure that the external research component is managed appropriately, CEGIS needs to plan for a ratio of internal to external resources of no more than 3:1. In addition to practical experience of this being an effective balance in the context of a 6- to 20-person GIScience lab, the rationale behind this ratio is the desire to achieve continuity across multiple projects (milestones) and to advance in-house capabilities to remain at the leading edge with respect to designing, managing, and conducting new and innovative research. Such in-house capabilities could be further enhanced if CEGIS scientists are given “space” to become established by not requiring them to pursue outside funding.

BOX 4.4
Office of Management and Budget Management Criteria for Research and Development Activities

- **Relevance**: “Investments must have clear plans, must be relevant to national priorities, agency missions, relevant fields, and ‘customer’ needs, and must justify their claim on taxpayer resources . . . Review committees should assess program objectives and goals on their relevance to national needs, ‘customer’ needs, agency missions, and the field(s) of study the program strives to address.”
- **Quality**: “Programs should maximize the quality of the research and development (R&D) they fund through the use of a clearly stated, defensible method for awarding a significant majority of their funding. A customary method for promoting R&D quality is the use of a competitive, merit-based process.”
- **Performance**: “R&D programs should maintain a set of high priority, multi-year R&D objectives with annual performance outputs and milestones that show how one or more outcomes will be reached. Metrics should be defined not only to encourage individual program performance but also to promote, as appropriate, broader goals, such as innovation, cooperation, education, and dissemination of knowledge, applications, or tools . . . Programs must demonstrate an ability to manage in a manner that produces identifiable results. At the same time, taking risks and working toward difficult-to-attain goals are important aspects of good research management, especially for basic research.”

SOURCE: Espinosa, (2006)

REALIZING CEGIS’S POTENTIAL

The committee envisions a small core of world-class, dedicated GIScience researchers within CEGIS coupled to a well-orchestrated network of researchers and resources at university centers; geospatial centers of excellence at other federal,[1] state, and local agencies (perhaps with unique foci relevant to their specific missions); and professional societies with their extensive constituencies of the best and brightest in the field (Figure 4.1). Collaboration with other agencies and organizations doing GIScience research is crucial to realizing a need of the nation to integrate these activities and is also crucial to establishing the leadership of CEGIS and the USGS in GIScience for the nation.

CEGIS's external network serves two purposes. First, it enables CEGIS to retain a broad perspective across the GIScience field so that it can identify and exploit relevant technology advances that may not necessarily happen within the core team at CEGIS. Second, by sponsoring and participating in these external research activities, CEGIS and the USGS will provide the cohesive leadership needed in GIScience in the United States and will, in turn, be recognized for this important contribution, thereby regaining that leadership role. The following sections describe the potential components of CEGIS's network in more detail. These include the CEGIS core within USGS and connections to academia, government, and industry. Some types of external connection are not new to USGS and CEGIS (see Chapter 1), and in these cases, CEGIS can carry forward and build on the experiences and lessons learned.

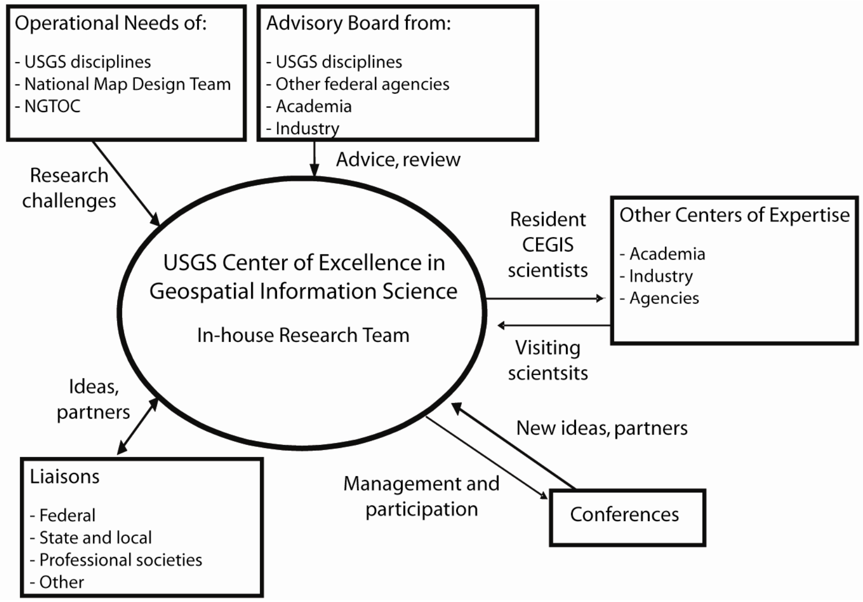


FIGURE 4.1 Conceptual diagram of CEGIS's key relationships

---

[1] For example the NGA, DHS, Census Bureau, Department of Transportation, USDA, and National Aeronautics and Space Administration (NASA).

The Core Research of CEGIS

The scientific core of CEGIS will initially consist of a group of approximately six to eight Ph.D.-level scientists. These scientists would be a mix of full-time USGS employees, visiting professors and other visitors (as with COGIT—Box 4.1), and/or postdoctoral researchers. The visitors would provide cross-fertilization with work outside USGS (Box 4.3; Michael Goodchild, University of California, Santa Barbara, personal communication, 2006). Each of CEGIS's research foci would be addressed by teams built from this initial group of scientists. The teams would be tightly coordinated and they would balance a focus on the small number of research topics with regular dialogue across these topics to raise ideas and share efforts so that the entire center is committed and dedicated to overall execution and results.

RECOMMENDATION 6: CEGIS should initially comprise six to eight Ph.D.-level scientists working in teams of at least two on the high-priority topics identified in Recommendations 3 to 5. Each team would comprise a mix of USGS scientists and visiting scientists and/or postdoctoral fellow(s) as appropriate to the topic. Their location should not be constrained to USGS facilities if the most efficient progress could be made in another setting (e.g., an academic center of excellence).

The core science group will, as momentum builds and demands expand (and as budgets allow), grow to between 12 and 20 scientists over the next five years. The initial range of six to eight is the minimum critical mass to produce significant work on the handful of high-priority topics discussed in the previous chapter. The larger range of 12 to 20 would bring CEGIS's dimensions in line with what is needed to address the broader range of research identified in Chapter 3 and is comparable to that of research groups supporting other mapping agencies (e.g., Boxes 4.1 and 4.2). Details on how such a program would be led are for the USGS to decide, but the committee sees the need for two leadership positions: (1) a director who focuses on administering the center and providing national leadership, inspiration, and enthusiasm for CEGIS programs and their development as well as maintaining a liaison role with outside cooperative agreements, and (2) a deputy director who manages the research portfolio (and who would be one of the Ph.D.-level scientists mentioned earlier). These leaders would, as is already the case, be supported by administrative personnel (e.g., administrative assistants, computer support staff).

Connections with Academia

CEGIS can foster fundamental research in academia. Three types of connection to academia are envisaged: establishing and supporting academic centers of GIScience excellence that focus on USGS-related research (e.g., Box 4.3), placing USGS researchers within these academic institutions, and providing grant support through competitive or sole source means (e.g., Box 4.3). The visiting professors and postdoctoral researchers envisaged as critical members of the CEGIS science core could (but need not) be associated with these centers. The majority of support for such centers (and personnel therein) would be in addition to support for the temporary visitors within the CEGIS core science team.

The handful of CEGIS-sponsored university GIScience centers of excellence, perhaps one to two start and eventually growing to two to four, would be selected by matching their expertise to the critical GIScience research areas confronting CEGIS. Such centers could focus on long-term and perhaps more risky research topics in CEGIS's portfolio or simply areas of particular expertise and innovation that serve CEGIS needs. The infusion of graduate students into CEGIS's research activities through such centers would bring human resources at affordable rates and potentially attract future research employees to USGS (e.g., Box 4.3). The number of centers could increase as their success generates justification for additional resource funding.

RECOMMENDATION 7: CEGIS should establish and/or support one to two centers of excellence in GIScience at universities with relevant GIScience focus and capabilities that address its longer-term research challenges.

Placing USGS researchers at university centers of excellence in GIScience would potentially be of great benefit to cultivating CEGIS's GIScience leadership (Mike Goodchild, University of California, Santa Barbara, personal communication, 2006). In this model, a CEGIS research team leader could be embedded for several weeks or months with top researchers in an academic setting to work on one of CEGIS's high-priority research topics. As indicated in Figure 4.1, the model of embedding a CEGIS scientist elsewhere could also be applied within another USGS discipline, agency, or organization. In addition to funding the university centers, and as a more dynamic adjunct to them (Jack Dangermond, ESRI, personal communication, 2006), CEGIS could manage a program of university grants through sole source selections when the choice is obvious (e.g., when there is a clear leader on a topic of particular interest to CEGIS) or otherwise, through open competition (i.e., requests for proposals) (e.g., Box 4.3).

Connections with Industry

Partnerships among industry and mapping agencies are commonplace because of the wealth of knowledge and experience in the private sector. NGA, for example, engages industry partners through a number of mechanisms (NRC, 2006). The Ordnance Survey's partnership with Oracle, USGS's own work with Microsoft on Terraserver, and many government agencies' participation in Open Geospatial Consortium (OGC) activities are other examples. With the widespread use of the Global Positioning System (GPS), field data collection devices, desktop geographic information system (GIS) software, online mapping services, and consumer navigation, the geospatial industry has created impressive applications and infrastructure for the collection, processing, distribution, and display of geospatial data. In addition, a sub-industry of technical support for geospatial products has emerged to support commercial products. These activities can potentially contribute to the national geospatial research agenda. Consequently, provision of resources by CEGIS to tap into and catalyze industry's work on topics under CEGIS's research portfolio is an important element of the overall CEGIS research strategy and national leadership role. Directing an industry research agenda establishes CEGIS's prominence within industry just as funding of university centers establishes a prominent position for CEGIS within the academic community. In addition, funding industry research on aspects of *The National Map*, such as infrastructure development, human-computer interface, or applications built on top of *The National Map* application programming interfaces for specific users, would help establish *The National Map* as the preeminent source of quality national geospatial data and services.

Because industry research is typically focused on product development or projects in response to customer requirements, connecting into these research capabilities is best accomplished by contractual instruments such as Broad Area Announcements or Cooperative Research and Development Act agreements. In addition, targeted contracts may be most effective in instances of narrowly defined research needs.

RECOMMENDATION 8: CEGIS should supplement the work of its core research teams with Broad Area Announcements, Cooperative Research and Development Act agreements, and targeted contracts on high-priority research topics.

Government Agency Liaisons

By liaising with other agencies with GIScience research activities, CEGIS could "piece together the National Agenda" (Jack Dangermond, ESRI, personal communication, 2006). Candidate federal agencies with significant programs or research in GIScience include NGA, the National Aeronautics and Space Administration (NASA), NSF, the National Oceanic and Atmospheric Administration (NOAA), FEMA, the Census Bureau, the Central Intelligence Agency, Defense Advanced Research Projects Agency (DARPA), the Army Corps of Engineering's Topographical Engineering Center, and DHS. The GIScience at NGA, FEMA, and DHS is particularly relevant to USGS and provides fertile opportunities for collaboration on a comprehensive research agenda. Properly executed, this collaborative research agenda would allow each agency to focus on its unique needs, such as imagery analysis at the NGA or floodplain mapping at FEMA, while depending on the USGS for a consistent, current, high-quality base map including the themes that the USGS collects. This eliminates duplication and allows each agency to focus on the data with which it has the most expertise. There is already a desire for better coordination of federal geospatial research funding and the possibility of NSF-USGS partnerships on targeted research topics (Maria Zemankova, NSF, personal communication, 2006). Even within the USGS, CEGIS could establish a strong liaison with the disciplines of water, geology, geography, and biology with co-located professionals to work on research to address common application challenges that involve core geospatial data.

RECOMMENDATION 9: To reestablish USGS's leadership role in GIScience, maximize efficiency, and share in the cost of addressing common challenges, CEGIS should forge connections with other federal agencies, professional societies, and private-sector firms that conduct, support, and/or promote GIScience research.

In addition to federal agencies, state and local government organizations are a potentially important source of collaboration for CEGIS. Some of these entities conduct geospatial research and apply this to operations in their jurisdiction. Many of the progressive local agencies are also key contributors of data to *The National Map* (Ivan DeLoatch, USGS, personal communication, 2006). Where expert groups exist at the state or county level,[2] CEGIS could work with USGS's

---

[2] Examples of groups working on topics relevant to *The National Map* are (1) Missouri Spatial Data Information Service—a spatial data retrieval and archival system at the Geographic Resources Center at the University of Missouri; (2) New Hampshire Geographically Referenced Analysis and Information Transfer System (NH GRANIT)—a cooperative project to create, maintain, and make available a statewide geographic database serving the information needs of state, regional, and local decision makers, managed at Complex Systems Research Center, University of New Hampshire; (3) Kansas Data Access and Support Center (DASC)—a central delivery and distribution center for core GIS databases essential to ensure the effective and efficient development and implementation of GIS technology in state government; and (4) Utah Automated Geographic Reference Center (AGRC), which provides a wide range of GIS support to the State of Utah and stewardship of the State Geographic Information Database.

existing state liaisons to identify a few key programs and support research collaborations with these entities. Such collaboration would have additional benefits of promoting use of *The National Map* at all levels and supporting the establishment of semantic, thematic, and other standardization at the source, thereby reducing the challenge of automated techniques over time and enhancing data integration.

RECOMMENDATION 10: Because of USGS's core role in integrating data from local sources for *The National Map*, CEGIS should establish collaborative activities with state and local agencies that have progressive activities in GIScience.

Multilevel collaborations among government agencies must negotiate many political and other complexities. One example of success in this regard is DHS's standardization of 17 layers of the National Critical Infrastructure (DHS, 2004). When federal, state, and local agencies use these common databases for national events and disasters, coordination and collaboration are greatly facilitated.

Conferences and Workshops

By hosting specialist conferences and workshops, CEGIS could raise its national and international profile while concurrently serving its research interests (e.g., Box 4.2). The National Center for Geographic Information and Analysis model of specialist meetings on focused topics could be particularly effective in this regard (Jack Dangermond, ESRI, personal communication, 2006). CEGIS could tailor the agenda of each meeting to its most pressing research need. For example, specialist workshops might coincide with the launch or conclusion of directed research at one of CEGIS's university centers or with industry research thrusts. This would focus national attention on the research agenda of CEGIS.

Such meetings could be coordinated through the UCGIS in conjunction with its winter assembly government research and summer assembly academic research meetings. Both meetings are already established, and with USGS support they could emerge as major venues for scientific exchange in GIScience.

RECOMMENDATION 11: CEGIS should use specialist meetings, perhaps in conjunction with the University Consortium for Geographic Information Science winter meeting or summer assembly, to advance its state of knowledge and plans for addressing emerging research challenges.

On a broader scale, a liaison from CEGIS to the many professional societies involved with GIScience would enhance awareness of CEGIS activities, maintain a strong connection to the latest developments in the field that might be of relevance to CEGIS research foci, and yield special opportunities for collaboration and dissemination of research findings. Such societies include the University Consortium for Geographic Information Science, Cartography and Geographic Information Society, American Society for Photogrammetry and Remote Sensing, Urban and Regional Information Systems Association, North American Cartographic Information Society, Geospatial Information Technology Association, and International Cartographic Association.

CEGIS Advisory Board

The advisory board would have two purposes: (1) to provide broad-based input, review, and critique of CEGIS plans, activities, and progress; and (2) to institutionalize the liaison with USGS's geology, biology, hydrology, and geography disciplines (see Box 4.3 for a description of a somewhat similar approach at NGA). The flow of ideas between the disciplines and CEGIS about research needs, opportunities, and resources is essential to CEGIS's success (Marty Goldhaber, USGS, personal communication, 2006), as is the flow of research needs from *The National Map* design team and NGTOC. To be most efficient, CEGIS's advisory board would have to be limited in size (e.g., less than ten members) and meet regularly (e.g., two to three times per year). To be most effective and well-rounded to capture insights and experience from all facets of CEGIS's broader set of relationships outside USGS, the board would include non-USGS membership from relevant agencies, academia, and industry. CEGIS leadership would need to define the role and processes of the advisory board carefully to ensure it provides benefits without becoming cumbersome or hampering productivity.

RECOMMENDATION 12: To provide broad-based input, review, and critique of CEGIS plans, activities, and progress and to institutionalize CEGIS's connection to the USGS disciplines, the National Geospatial Program Office should establish an advisory board for CEGIS that includes members from each of the USGS disciplines as well as non-USGS GIScience experts.

SUMMARY

In this report, the committee has set forth a vision for the future of the National Geospatial Program Office (NGPO) in GIScience—a vision not only of leadership within the USGS, but across the federal government anywhere that geospatial data are critical to operations; across the state and local offices that rely on consistent, up-to-date geospatial data; and even across academia and industry where many of the difficult research problems in GIScience will be solved.

Central to successful execution of this charge is focus. CEGIS has lean resources. It has the potential to fill the role of leadership in GIScience in the United States, but it must focus on a few critical projects and execute them flawlessly. With early successes and a focus on the research priorities described in this report, the committee is confident that the USGS will grow and emerge as strong as in the days of the paper topographic maps and to eventually encompass a much broader research agenda than discussed here.

In this report, the committee endeavored to prioritize the many possible avenues of research into a solid core of interrelated research topics that provide early and visible results that are of importance to the nation's need for accurate and accessible geospatial information. This was a compact study, in both time and resources. It is possible that some aspects of CEGIS's mission were missed; therefore, this report should be considered a starting point for refinement and final prioritization of the CEGIS research agenda. However, it is the committee's hope that this report will provide a fertile beginning for achieving focus and ultimately successful execution by CEGIS and NGPO and that it will be the start of the reemergence of the USGS as the leader for GIScience research for the nation.

References

Abraham, T., and J. Roddick. 1999. Incremental meta-mining from large temporal data sets. Pp. 4154 in Y. Kambayashi, D. Lee, E.-P. Lim, M. Mohania, and Y. Masunaga (eds.), Advances in Database Technologies: *Proceedings of the First International Workshop on Data Warehousing and Data Mining*. Berlin: Springer Verlag.

ADL (Alexandria Digital Library). 2004. *Guide to the ADL Gazetteer Content Standard Version 3.2.* Santa Barbara, Calif.: Map and Imagery Lab, Davidson Library, University of California, Santa Barbara. Available online at http://www.alexandria.ucsb.edu/gazetteer/ContentStandard/version3.2/GCS3.2-guide.htm. Accessed July 3, 2007.

Adobe Systems Incorporated. 2004. *PDF Reference Fifth Edition: Adobe® Portable Document Format Version 1.6*. Berkeley, Calif.: Adobe Press. 1248 pp.

Agouris, P., A. Stefanidis, and S. Gyftakis. 2001. Quality-aware deformable models for change detection. *Proceeding of the 2001 International Conference on Image Processing* 2: 805-808.

Alani, H., C. Jones, and D. Tudehope. 2001. Vornoi-based region approximation for geographical information retrieval with gazetteers. *International Journal of Geographical information Science* 15:287-306.

Armstrong, M. 1988. Temporality in spatial databases. Pp. 880-889 in *GIS/LIS 88 Proceedings: Accessing the World, Volume II*. Falls Church, Va.: American Society for Photogrammetry and Remote Sensing.

Beard, K. 2006. Modeling change in space and time: An event based approach. In J. Drummond, R. Billen, D. Forrest, and E. João (eds.), *Innovations in GIS: Dynamic & Mobile GIS: Investigating Change in Space and Time.* Boca Raton, Fla.: CRC Press. 310 pp

Bertin, J. 1983. *Semiology of Graphics: Diagrams, Networks, Maps.* Madison: University of Wisconsin Press [French edition 1967].

Brassel, K., and R. Weibel. 1988: A review and conceptual framework of automated map generalization. *International Journal of Geographical Information Systems* 2(3):229-244

Brewer, C. 2005. *Designing Better Maps: A Guide for GIS Users.* Redlands, Calif.: ESRI Press. 203 pp.

Brewer, C., and B. Buttenfield. 2007. Framing guidelines for multi-scale map design using databases at multiple resolutions. *Cartography and Geographic Information Science* 34(1):3-15.

Burrough, P., and A. Frank. 1996. Pp. 109-121 in *Geographic Objects with Indeterminate Boundaries*. London: Taylor and Francis.

CEGIS (Center of Excellence for Geospatial Information Science), 2006. Call for Proposals for Fiscal Year 2007. Available online at http://cegis.usgs.gov/pdf/FY2007_CEGIS_Prospectus_Call_for_Proposals.pdf. Accessed August 9, 2007.

Chen, C., S. Thakkar, C. Knoblock, and C. Shahabi. 2003. Automatically annotating and integrating spatial datasets.. *Proceedings of the Eighth International Symposium on Spatial and Temporal Databases* (SSTD 2003), Santorini Island, Greece

Chen, C., C. Knoblock, C. Shahabi, S. Thakkar, and Y. Chiang. 2004. Automatically and accurately conflating orthoimagery and street maps. *Proceedings of the Twelfth ACM International Symposium on Geographic Information Systems* (ACM-GIS '04), Washington, D.C.

Claramunt, C., and M. Theriault. 1995. Managing time in GIS: An event-oriented approach. Pp. 23-42 in J. Cliffford and A. Tizhilin (eds.), *Recent Advances in Temporal Databases*. Berlin: Springer Verlag..

Clinton, N., P. Gong, and K. Scott. 2006. Quantification of pollutants emitted from very large wildland fires in Southern California, U.S.A. *Atmospheric Environment* 40(20):3686-3695.

DeMers, M. 2003. *Fundamentals of Geographic Information Systems,* 3rd ed. New York: John Wiley & Sons. 467 pp.

DHS (Department of Homeland Security). 2004. The National Plan for Research and Development in Support of Critical Infrastructure Protection. April. Available online at http://www.dhs.gov/xlibrary/assets/ST_2004_NCIP_RD_PlanFINALApr05.pdf. Accessed July 3, 2007.

DiBiase, D., M. DeMers, A. Johnson, K. Kemp, A Luck, B. Plewe, and E. Wentz. 2006. *Geographic Information Science and Technology Body of Knowledge,* 1st ed. University Consortium for Geographic Information Science. Washington, D.C.: Association of American Geographers.

Ebner, D., G. Klau, and R. Weiskircher. 2003. Force-Based Label Number Maximization. Technical Report. Vienna: Institute of Computer Graphics and Algorithms, Vienna University of Technology.

Espinosa, D. 2006. Research and development program investment criteria (Appendix C). Memo to OMB Program Associate Directors, OMB Program Deputy Associate Directors, Agency Budget and Performance Integration Leads, and Agency Program Assessment Rating Tool Contacts*, 2006* (March 10), regarding guidance for completing 2006 PARTs (Program Assessment Rating Tool Guidance No. 2006-02): Washington D.C.: U.S. Executive Office of the President, Office of Management and Budget, 92 pp. Available online at http://www.whitehouse.gov/omb/part/fy2006/2006_guidance_final.pdf. Accessed November 20, 2006.

Feng, C., T. Bittner, and D. Flewelling. 2004. Modeling surface hydrology concepts with endurance and perdurance. Pp. 67-80 in M. Egenhofer, C. Freksa, and H. Miller (eds.), *Proceedings of Giscience*, LNCS 3234. Berlin, Heidelberg: Springer-Verlag.

Fisher, P., J. Wood, and T. Cheng. 2004. Where is Helvelyn? Multiscale morphometry and the mountains of the English Lake District. *Transactions of the Institute of British Geographers* 29:106-128.

Fonseca, F., M. Egenhofer, C. Davis, and G. Camara. 2002. Semantic granularity in ontology-driven geographic information systems. *Annals of Mathematics and Artificial Intelligence* 36(1-2):121-151.

Garrett, J. 2002. *The Elements of User Experiences: User-Centered Design for the Web.* New York: American Institute of Graphic Arts.

Ghose, R., and W. Huxhold. 2002. Role of multi-scalar GIS-based indicators studies in formulating neighborhood planning policies. *URISA Journal* 14(2):5-17.

Gong, P., R. Pu, Z.Q. Li, J. Scarborough, N. Clinton, and L. Levien. 2006. An integrated approach to burned area mapping in California with NOAA AVHRR data. *Photogrammetric Engineering and Remote Sensing* 72(2):139-150.

Goodchild, M. 1992. Geographical information science. *International Journal of Geographical Information Systems* 6(1):31-45.

Goodwin, J. 2005. What have ontologies ever done for us—potential applications at a national mapping agency. In OWL: Experiences and Directions. November 11-12. Galway, Ireland. Available online at http://www.mindswap.org/2005/OWLWorkshop/sub18.pdf. Accessed July 3, 2007.

Gregory, I. 2002. The accuracy of areal interpolation techniques: Standardizing 19th and 20th century census data to allow long-term comparisons. *Computers, Environment, and Urban Systems* 26:293-314.

Grenon, P., and B. Smith. 2004. SNAP and SPAN: Prolegomenon to geodynamic ontology. *Spatial Cognition and Computation* 4(1):69-104.

Halsing, D., K. Theissen, and R. Bernknopf. 2004. A Cost-Benefit Analysis of *The National Map*. USGS Circular 1271. Reston, Va.: U.S. Geological Survey. Available online at: http://pubs.usgs.gov/circ/2004/1271/c1271.pdf. Accessed July 3, 2007.

Hampe, M., and M. Sester. 2004. Generating and using a multi-resolution database (MRDB) for mobile applications. ICA Research Workshop on Generalisation and Multiple Representation, Leicester, U.K., August 20-21.

Hampe, M., and S. Intas. 2006. Extension of the OGC Web Feature Service Standard for Multiple Representation Data. ISPRS Technical Commission II Symposium, Vienna, July 12-14.

Harrie, L., and R. Weibel. 2007. Modeling the overall process of generalisation. Pp. 67-88 in W. Mackaness, A. Ruas, and T. Sarjakoski (eds.), *The Generalization of Geographic Information: Cartographic Modelling and Applications.* Amsterdam: Elsevier.

Hill, L. 2000. Core elements of digital gazetteers: Placenames, categories, and footprints. Pp. 280-290 in J. Borbinha and T. Baker (eds.), Research and Advanced Technology for Digital Libraries: *Proceedings of the 4th European Conference, ECDL 2000,* Lisbon, Portugal, September 18-20.

Hill, L. 2006. *Georeferencing. The Geographic Associations of Information*. Cambridge, Mass.: MIT Press.

Hill, L., and M. Goodchild. 2000. Digital Gazetteer Information Exchange (DGIE). Final Report of Workshop Held October 12-14, 1999. Available online at http://www.alexandria.ucsb.edu/~lhill/dgie/DGIE_website/DGIE%20final%20report.htm. Accessed July 3, 2007.

Hodge, G. 2000. Systems of Knowledge Organization for Digital Libraries: Beyond Traditional Authority Files. Council on Library and Information Resources. Available online at http://www.clir.org/pubs/abstract/pub91abst.html. Accessed July 3, 2007.

ISO (International Organization for Standardization). 1999. ISO 13407: Human Centered Design for Interactive Systems. Geneva, Switzerland: International Organization for Standardisation.

Jhingran, A., N. Mendonca Mattos, and H. Pirahesh. 2002. Information integration: A research agenda. *IBM Systems Journal* 41(4):555-562.

Jones, C., A. Abdelmoty, and G. Fu. 2003. Maintaining Ontologies for Geographical Information Retrieval on the Web. *Proceedings of OTM Confederated International Conferences CoopIS, DOA, and OOBASE* 934-951. Available online at http://www.geo-spirit.org/publications/SPIRIT_maintaining_ontologies.pdf. Accessed 3 July 2007.

Jones, C., A. Abdelmoty, D. Finch, G. Fu, and S. Vaid. 2004. The SPIRIT spatial search engine: Architecture, ontologies and spatial indexing. Pp. 125-139 in *Proceedings of Third International Conference on GIScience*, Lecture Notes in Computer Science. Berlin: Springer.

Kafeza, E., V. Delis, and T. Hadzilacos. 1996. Transaction synchronization in multiresolution geographic databases.. Pp. 113-120 in *Proceedings of the 4th ACM International Workshop on Advances in Geographic Information Systems.* New York: ACM Press

Kameda, T., and K. Imai. 2003. Map label placement for points and curves. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences* E86-A(4):835-840.

Kilpelainen, T. 1997. Multiple representations and generalizations of geo-databases for topographical maps. Ph.D. thesis. Kirkkonummi: *Publications of the Finnish Geodetic Institute.* 229 pp.

Knoblock, C., and C. Shahabi. 2007. Geospatial data integration. In J. Wilson and A. Fotheringham (eds), *Handbook of Geographic Information Science*. Oxford, U.K.: Blackwell.

Kuhn, W. 2005. Geospatial Semantics: Why, of what, and how? Special Issue on Semantic-Based Geographical Information Systems, *Journal on Data Semantics III, Lecture Notes in Computer Science* 3534:1-24.

Langran, G. 1992. *Time in Geographic Information Systems*. London: Taylor and Francis. 189 pp.

Llinas, J., and D. Hall. 1998. An introduction to multi-sensor data fusion. *Proceedings of the International Symposium on Circuits and Systems* 6(June 1):541-544.

Lutz, M., and E. Klien. 2006. Ontology-based retrieval of geographic information. *International Journal of Geographic Information Science* 20(3):233-260.

MacEachren, A. 1995. *How Maps Work: Representation, Visualization, and Design,* Hampshire, U.K.: Guilford Press. 513 pp.

Mackaness, W., A. Ruas, and T. Sarjakoski (eds.). 2007. *The Generalization of Geographic Information: Cartographic Modelling and Applications.* Amsterdam: Elsevier Press. 370 pp.

Mark, D., and B. Smith. 2004. A science of topography: From qualitative ontology to digital representations. Chapter 3 in M.P. Bishop and J.F. Shroder (eds.), *Geographic Information Science and Mountain Geomorphology.* Chichester, U.K.: Springer-Praxis. 486 pp.

McMahon, G., S. Benjamin, K. Clarke, J. Findley, R. Fisher, W. Graf, L. Gundersen, J. Jones, T. Loveland, K. Roth, L. Usery, and N. Wood. 2005. Geography in a Changing World: A Science Strategy for the Geographic Research of the U.S. Geological Survey. USGS Circular 1281. Sioux Falls, S.D.: U.S. Geological Survey. 74 pp.

McMaster, R. 1989. The integration of simplification and smoothing foutines in line generalization. *Cartographica* 26 (1):101-121.

McMaster, R., and K. Shea. 1992. *Generalization in Digital Cartography.* Resource Publications in Geography. Washington, D.C.: Association of America Geographers. 134 pp.

McMaster, R., and E. Sheppard. 2004. Introduction: Scale and geographic inquiry. Pp. 1-15 in: E. Sheppard, and R. McMaster (eds.), *Scale and Geographic Inquiry*. Oxford, U.K.: Blackwell. 288 pp.

Mustiere, S., and J. van Smaalen. 2007. Database requirements for generalisation and multiple representation. Pp. 113-136 in W. Mackaness, A. Ruas, and T. Sarjakoski (eds.), *The Generalization of Geographic Information: Cartographic Modelling and Applications.* Amsterdam: Elsevier Press. 386 pp.

Nielsen, J. 1993. *Usability Engineering*. Boston, Mass.: Academic Press.

Nielsen, J. 1999. *Designing Web Usability: The Practice of Simplicity*. Indianapolis, Ind.: Peachpit Press. 432 pp.

NGA (National Geospatial-Intelligence Agency). 2006. NGA Academic Research Program Symposium Agenda. Washington, D.C. September 13-15.

NRC (National Research Council). 1993. *Toward a Coordinated Spatial Data Infrastructure for the Nation*. Washington, D.C.: National Academy Press.

NRC. 2001. *Future Roles and Opportunities for the U.S. Geological Survey*. Washington, D.C.: National Academy Press.

NRC. 2002. *Research Opportunities in Geography at the U.S. Geological Survey*. Washington, D.C.: National Academies Press.

NRC. 2003. *Weaving a National Map: A Review of the U.S. Geological Survey Concept of The National Map*. Washington, D.C.: National Academies Press.

NRC. 2006. *Priorities for GEOINT Research at the National Geospatial-Intelligence Agency*. Washington, D.C.: National Academies Press.

OGC (Open Geospatial Consortium). 2000. Open GIS Consortium Web Mapping Testbed Public Page. Wayburn, Mass.

OGC. 2004. OGC Web Services Initiative, Phase 2. Wayland Mass.: Open Geospatial Consortium Inc. Available online at http://www.opengeospatial.org/projects/initiatives/ows-2#rfq. Accessed August 9, 2007.

Peng, Z., and M. Tsou. 2003. *Internet GIS—Distributed Geographic Information Services for the Internet and Wireless Networks.* Hoboken, N.J.: Wiley.

Peuquet, D., and N. Duan. 1995. An event based spatio-temproal data model (ESTDM) for temporal analysis of geographical data. *International Journal of Geographic Information Systems* 9:7-24.

Purves, R., P. Clough, and H. Joho. 2005. Identifying imprecise regions for geographic information retrieval using the web. Pp. 313-318 in *Proceedings of GIS Research U.K. 13th Annual Conference*, Glasgow, U.K.

Radke, J. 1995. Modeling urban/wildland interface fire hazards within a geographic information system, *Geographic Information Sciences* 1(1):7-20.

Regnauld, N. and R. McMaster. 2007. A synoptic view of generalisation operators. Pp. 37-66 in William Mackaness, Ann Ruas, and Tiina

Sarjakoski (eds), *The Generalization of Geographic Information: Cartographic Modelling and Applications.* Amsterdam: Elsevier Press..

Rigaux, P., and M. Scholl. 1995. Multiscale partitions: Applications to spatial and statistical databases. Pp. 170-183 in Advances in Spatial Databases. *Lecture Notes in Computer Science 951*. Berlin: Springer..

Rothermel, R.C. 1972. A Mathematical Model for Predicting Fire Spread in Wildland Fuels. Researc Paper INT-115. Ogden, Utah: U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station. 40 pp.

Saalfeld, A. 1987. Conflation: Automated Map Compilation. Statistical Research Division Report Series. RR-87/24. Washington, D.C.: U.S. Bureau of the Census. Available online at http://www.census.gov/srd/papers/pdf/rr87-24.pdf. Accessed July 24, 2007.

Sarjakoski, L. 2007. Conceptual models of generalisation and multiple representation. Pp. 11-36 in W. Mackaness, A. Ruas, and T. Sarjakoski (eds.), 2007. *The Generalization of Geographic Information: Cartographic Modelling and Applications.* Amsterdam: Elsevier.

Saux, E., R. Thibaud, K.J. Li, and M.H. Kim. 2004. A new approach for a topographic feature-based characterization of digital elevation data. Pp 73-81 in *Proceedings of the 12th annual ACM international workshop on Geographic Information Systems***.** New York: Association for Computing Machinery.

Shneiderman, B. 1998. *Designing the User Interface: Strategies for Effective Human-Computer Interaction*, 3rd. ed. Boston: Addison-Wesley. 639 pp.

Slocum, T., R. McMaster, F. Kessler, and H. Howard. 2005. *Thematic Cartography and Geographic Visualization,* 2nd ed. Upper Saddle River, N.J.: Pearson Prentice Hall. 575 pp.

Smith, B., and D. Mark. 2001, Geographical categories: An ontological investigation. *International Journal of Geographic Information Science* 15(7):592-612.

Soergel, D. 2000. Taxonomy of Knowledge Organization Structures/Systems. Kent, Ohio: Kent State University. Available online at http://nkos.slis.kent.edu/KOS_taxonomy.htm. Accessed July 24, 2007.

Sorokine, A., T. Bittner, and C. Renschler. 2004. Ontological investigation of ecosystem hierarchies and formal theory for multiscale ecosystem classification. *Geoinformatica* 10(3):313-335.

Sorokine, A. and T. Bittner. 2005. Understanding taxonomies of ecosystems: a case study. Pp. 559-572 in P. Fisher (ed.), *Developments in Spatial Data Handling*. Berlin: Springer-Verlag.

Sowa, J.F. 1998. *Knowledge Representations Logical, Philosophical and Computational Foundations*. Boston, Mass.: PWS Publishing Co.

Stadler, G., T. Steiner, and J. Beiglböck. 2006. A practical map labeling algorithm utilizing image processing and force-directed methods. *Cartography and Geographic Information Science* 33(3):207-215.

Thibault, P. 2001. Cartographic generalization of fluvial features. Ph.D. dissertation. Minneapolis: University of Minnesota.

Thompson, M. 1988. *Maps for America,* 3rd ed. Washington, D.C.: U.S. Geological Survey.

Tobler, W. 1979. Smooth pycnophylactic interpolation for geographical regions. *Journal of the American Statistical Association* 74(367):519-535.

Tsou, M. 2004. Integrating web-based GIS and on-line remote sensing facilities for environmental monitoring and Management. In Special Issue on the Potential of Web-Based GIS, *Journal of Geographical Systems* 6(2):155-174.

Tu, S., and M. Abdelguerfi. 2006. Web services for geographic information systems. *IEEE Internet Computing* 10(5):13-15.

Tu, Y-C., J. Yan, and S. Prabhakar. 2005. Quality Aware Replication of Multimedia Data. Pp 240-249 in *Database and Expert Systems Applications*. Berlin: Springer-Verlag..

U.S. Executive Office of the President, Office of Management and Budget. 2002. August 19. Coordination of Geographic Information and Related Spatial Data Activities. Circular A-16. Available Online at http://www.whitehouse.gov/omb/circulars/a016/a016_rev.html.. Accessed November 19, 2006.

USGS (U.S. Geological Survey). 1994. Spatial Data Transfer Standard. FIPS Publication 173. Gaithersburg, Md.: National Institute of Standards.

USGS. 2001. *The National Map*: Topographic Mapping for the 21st Century. Reston, Va.: U.S. Geological Survey, 38 pp. Available Online at http://nationalmap.gov/report/national_map_report_final.pdf. Accessed November 19, 2006.

USGS. 2003. Implementation Plan for *The National Map*. October 18, Version 1.0. Available online at http://nationalmap.gov/report/Implementation_Plan_ver_1.0.pdf. Accessed July 24, 2007.

USGS. 2007. Facing tomorrow's challenges—U.S. Geological Survey science in the decade 2007-2017. Circular 1309. Reston, Va.: U.S. Geological Survey. 70 pp.

Van Riper, D. 2003. Assessing the accuracy of areal interpolation methods for the National Historical Geographic Information System. Master's thesis. University of Minnesota-Minneapolis.

Vickus, G. 1995. Strategies for ATKIS-related cartographic products. Pp. 246-252 in J. Muller, J. Lagrange, and R. Weibel (eds.), *GIS and Generalization: Methodology and Practice.* Bristol, Pa.: Taylor & Francis. 257 pp.

Wilson, J., C. Lam, and D. Holmes-Wong. 2004. A new method for the specification of geographic footprints in digital gazetteers. *Cartography and Geographic Information Systems* 31(4):195-207.

Worboys, M. 2005. Event-oriented approaches to geographic phenomena. *International Journal of Geographical Information Science* 19(1):1-28.

Worboys, M., and K. Hornsby. 2004. From objects to events: GEM, the geospatial event model. Pp. 327-344 in M. Egenhofer, C. Freksa, H. Miller (eds.), *Third International Conference on GIScience—Lecture Notes in Computer Science* 3234. Berlin: Springer Verlag.

Xu, J. 2006. Modeling the scale right: Mitigating wildfire in the wildland-urban interface. Ph.D. dissertation. Department of Landscape Architecture, University of California, Berkeley. 156 pp.

Yuan, M., and J. McIntosh. 2002. A typology of spatiotemporal information queries. Pp. 63-82 in K. Shaw, R. Ladner, and M. Abdelguerfi (eds.), *Mining Spatiotemporal Information Systems*. Norwell, Mass.: Kluwer Academic Publishers.

APPENDIXES

A

Committee and Staff Biographies

Robert P. Denaro, *Chair*, is vice president of NAVTEQ, a corporation that specializes in digital road maps for navigation, and is leading NAVTEQ's new thrust into Advanced Driver Assistance Systems. He holds a master's degree in electrical engineering from the U.S. Air Force Institute of Technology, a master's degree in systems management from the University of Southern California, and a bachelor's degree in engineering sciences (Astronautics) from the U.S. Air Force Academy. Mr. Denaro joined NAVTEQ from Rand McNally & Co. where he was senior vice president and general manager of Global Business Solutions, responsible for business-to-business applications and consumer technology products and services in mapping and routing. Prior to joining Rand McNally, Mr. Denaro was vice president and director of Motorola's Consumer Telematics Products, a division he launched after heading the company's Global Positioning System (GPS) business for five years. Earlier in his career, Mr. Denaro launched Trimble Navigation's Fleet Management and Vehicle Tracking Division and was co-founder of TAU Corporation, producer of the first commercial differential GPS systems. He started his career in the U.S. Air Force, where he served for nine years, initially working on research, development, and flight testing of the first cockpit digital map displays, and ultimately carrying out research and development as a captain at the Navstar Global Positioning System Joint Program Office. Mr. Denaro is a member of the National Research Council's Mapping Science Committee, a former member of the Board of Directors of the Intelligent Transportation Society of America (ITS America) and served as a Policy Board Director of the 511 National Traveler Information Number Deployment Coalition. He is a past vice president of the Institute of Navigation, past vice chairman of the U.S. GPS Industry Council, and was a North Atlantic Treaty

Organization (NATO) Advisory Group for Aerospace Research & Development (AGARD) lecturer.

Kate Beard-Tisdale is a professor in the Department of Spatial Information Science Engineering at the University of Maine. She is director of the National Center for Geographic Information and Analysis (NCGIA). She holds an M.S. (1984) and a Ph.D. (1988) from the Institute for Environmental Studies, Land Resources Program, University of Wisconsin-Madison where she specialized in geographic information systems. Her research interests cover multiple representations and cartographic generalization; investigations and visualization of data quality and uncertainty; metadata services, representation, and visualization; digital libraries; the integration of geospatial data and imagery; and gazetteer development. Her recent research addresses modeling, analysis, and visualization of space-time events. She is participating in a new research project that involves collaboration with oceanographers to develop an ontology of ocean-related events, to detect oceanographic events from multiple ocean observing sensors, and to develop methods for exploration of event patterns. Dr. Beard also collaborates on other applications-oriented projects in areas such as water resources and bioinformatics. She serves on the Editorial Board of URISA Journal and is a member of the American Congress on Surveying and Mapping, the Urban and Regional Information Systems Association, the Geospatial Information Technology Association, and the Association for American Geographers. She has been a member of the NRC U.S. National Committee for CODATA since 2003.

Cynthia A. Brewer is a professor in the Department of Geography at The Pennsylvania State University. She has M.A. and Ph.D. degrees in geography with emphasis in cartography from Michigan State University. Her research interests are in map design, color theory applications in cartography, multiscale and multirepresentation cartography, hypothesis generation in visualization, choropleth classification for maps in series, and atlas mapping. Dr. Brewer is a consultant to Environmental Systems Research Institute, Inc. (ESRI), doing cartographic critique of ArcGIS 9.0 and multiscale map design, and to the Population Division of the U.S. Census Bureau, where she planned and produced a diversity atlas of the 2000 Census and did design and analysis consulting for the comprehensive second atlas. She is currently chair of the U.S. National Committee to the International Cartographic Association, was president of the North American Cartographic Information Society in 1998-1999, and has been a member of the Editorial Board of *Cartography and Geographic Information Science* since 2000. She is the author of *Designing Better Maps: A Guide for GIS Users,* ESRI Press, 2005.

Michael Domaratz is a senior technical consultant with the Geospatial Information Technology practice of Michael Baker Jr., Inc. His duties include supporting geospatial data coordination activities of the Federal Emergency Management Agency's Flood Map Modernization program. He recently retired from the U.S. Geological Survey (USGS), where his responsibilities included implementing *The National Map*, a plan to provide current and accurate digital map data for the United States. He also had responsibilities for coordinating with the Bureau of the Census and internal USGS activities related to transportation data development. Mr. Domaratz co-chaired the Homeland Security Working Group of the Federal Geographic Data Committee (FGDC). Previous positions included serving as a liaison in the office of the Assistant Secretary for Water and Science in the Department of the Interior and as the executive secretary, metadata coordinator, and framework coordinator for the FGDC. He also held research, procurement, and production positions in USGS. He is a member of the American Congress on Surveying and Mapping and the Association for Computing Machinery. He has received the U.S. Department of the Interior Distinguished Service Award (highest service award) and an Outstanding Service Award from the National States Geographic Information Council. Mr. Domaratz has a B.A. in geography from the State University of New York at Buffalo and did graduate work in geography at the Ohio State University.

Peng Gong is a professor in the Department of Environmental Science, Policy and Management at the University of California-Berkeley and director of the Center for Assessment and Monitoring of Forest and Environmental Resources. He received a B.S. and M.S. in geography from Nanjing University, China, and a Ph.D. in geography from the University of Waterloo, Canada, in 1990. His research involves using remote sensing and geographic information systems (GIS) to monitor and map natural resources and human settlement. Specific areas of research include feature extraction, land use and land cover mapping, change detection, uncertainty modeling and error analysis, photo-ecometrics, and automated map generalization. Projects include the development of algorithms for forest fire monitoring and mapping; soil diversity studies; invasive species monitoring; and the development of an image-based analysis system for precise ecological measurements. He has written numerous books in Chinese and one in English on remote sensing, GIS, and land use-land cover. He is the director of International Institute for Earth System Science, Nanjing University, and director of the State Key Laboratory of Remote Sensing Science, jointly sponsored by the Institute of Remote Sensing Applications, Chinese Academy of Sciences and Beijing Normal University, China. Dr. Gong is editor-in-chief of *Geographic Information Sciences*, editor of the *International Journal of Remote Sensing*, and a member of the editorial board of the *Journal of Remote Sensing*.

Robert B. McMaster is professor and chair of the Department of Geography at the University of Minnesota. He received a B.A. (cum laude) from Syracuse University in 1978 and a Ph.D. in geography and meteorology from the University of Kansas in 1983. His research interests include automated generalization (including algorithmic development and testing, the development of conceptual models, and interface design), environmental risk assessment (including assessing environmental injustice to hazardous materials, the development of new spatial methodologies for environmental justice, and the development of risk assessment models), and the history of U.S. academic cartography. He has coauthored several books including *Map Generalization: Making Rules for Knowledge Representation* (Longman Publishing Group, 1991), *Generalization in Digital Cartography* (Association of American Geographers 1992), *Thematic Cartography and Geographic Visualization* (Prentice Hall, 2003), *A Research Agenda for Geographic Information Science, and Scale and Geographic Inquiry* (with E. Sheppard, CRC Press, 2004). Dr. McMaster has served as editor of the journal on *Cartography and Geographic Information Systems* from 1990-1996, and is an editor of the Association of American Geographers' Resource Publications in Geography. He served as a member of the U.S. National Committee for the International Cartographic Association, president of the U.S. Cartography and Geographic Information Society, and is a University Consortium for Geographic Information Science (UCGIS) board member; he chairs the UCGIS Research Committee. In 1999, he was elected as a vice president of the International Cartographic Association, and was reelected in 2003. He is a member of the NRC Mapping Science Committee.

Ming-Hsiang (Ming) Tsou is an associate professor in the Department of Geography, San Diego State University. He received a B.S. from National Taiwan University in 1991, an M.A. from the State University of New York at Buffalo in 1996, and a Ph.D. from the University of Colorado at Boulder in 2001, all in Geography. His research interests are in Internet mapping and distributed GIS applications, mobile GIS and wireless communication, multimedia cartography and user interface design, and software agents and distributed computing technologies. He has applied his research interests in applications such as wildfire mapping, environmental monitoring and management, habitat conservation, and border security. He is coauthor of the book, *Internet GIS: Distributed Geographic Information Services for the Internet and Wireless Networks* (John Wiley & Sons, 2003). Dr. Tsou has been the cochair of the National Aeronautics and Space Admininistration (NASA) Earth Science Enterprise Data System Working Group (ESEDWG) Standard Process Group (SPG) from 2004 to present. He is a member of the Association of American Geographers and the American Congress on Surveying and Mapping. He received the 2004 Outstanding Faculty Award at San Diego State University and was recently elected as the 2006-2007 vice chair of the Cartographic Specialty Group in the Association of American Geographers.

John P. Wilson is professor of geography at the University of Southern California, where he directs the GIS Research Laboratory. He studies the development of new terrain analysis techniques; modeling of soil erosion, vegetation, and water quality processes and problems; modeling of spatial patterns of urban growth and habitat change and impact of land use change, urban growth, and conservation policies on these patterns; description of environmental and socioeconomic characteristics and their impacts on selected health and quality-of-life outcomes; and the development of web-based map and gazetteer services for digital libraries and archives. He has held several visiting appointments in environmental studies, geography, and planning at the Australian National University, the University of Utrecht in the Netherlands, and the University of Waikato in New Zealand. Dr. Wilson founded the journal *Transactions in GIS* (Blackwell Publishers) in 1996 and has served as editor-in-chief since its inception. He has served on the editorial board of *Applied Geography* (1992-2001) and has just started a four-year term on the Editorial Board of the *Annals of the Association of American Geographers*. He has chaired the Applied Geography Specialty Group of the Association of American Geographers (1989-1991) and the Research Committee of the University Consortium for Geographic Information Science (2002-2005). He is currently president of the University Consortium of Geographic Information Science and an active participant in the UNIGIS International Network, a worldwide consortium of more than 20 institutions that collaborate on the development and delivery of online geographic information science academic programs. He has published numerous books and articles on these topics, including two edited volumes, *Terrain Analysis: Principles and Applications* (John Wiley and Sons, 2000) and the *Handbook of Geographic Information Science* (Blackwell Publishers, 2006).

National Research Council Staff

Paul M. Cutler is a senior program officer with the Board on Earth Sciences and Resources of the National Academies. His interests are in surficial processes, hydrology, glaciology, global change, mapping science, and geographical science. Earlier work at the National Academies was with the Polar Research Board and the Board on Atmospheric Science and Climate. Prior to joining the Academies, Dr. Cutler was an assistant scientist and lecturer in geology and geophysics at the University of Wisconsin. He holds a Ph.D. in geology (University of Minnesota), an M.Sc. in geography (University of Toronto) and a B.Sc. in geography (Manchester University, England). In addition to postdoctoral work on numerical modeling of the Laurentide and Scandinavian ice sheets funded by the National Science Foundation (NSF), he has carried out fieldwork in Alaska, Antarctica, Arctic Sweden, the Canadian Rockies, the Swiss Alps, and the Karakoram Mountains of Pakistan.

B

Meeting Participants and Presenters

Nadine Alameh, NASA
Clint Brown, ESRI
Bill Carswell, USGS
Jack Dangermond, ESRI
Ivan Deloatch, USGS
Beth Driver, NGA
Anne Frondorf, USGS
Marty Goldhaber, USGS
Mike Goodchild, University of California, Santa Barbara
Steve Guptill, USGS
Geoff Jacquez, BioMedware
Bruce Jones, USGS
Michael Jones, Google
Xavier Lopez, Oracle
Jerry McMahon, USGS
Joel Morrison, Ohio State University, retired
Ed Parsons, U.K. Ordnance Survey
George Percivall, Open Geospatial Consortium
Duncan Scheill, U.K. Ordnance Survey
Karen Siderelis, USGS
Chris Tucker, Ionic Enterprises
Maria Zemenkova, NSF

C

CEGIS as Envisioned by McMahon et al. (2005)

The following series of quotes from McMahon et al. (2005) gives a sense of the dimensions and process of the Center of Excellence in Geospatial Information Science (CEGIS) envisioned to be needed to support geographic information science (GIScience) needs within the geography discipline and across the U.S. Geological Survey (USGS).

"The GIScience Center of Excellence should lead in the planning and implementation of the research associated with goals 7, 8, and 9[1]. The USGS must invest in personnel, resources, and infrastructure to establish a center of excellence focused on GIScience that builds, nurtures, and maintains a core of GIScience researchers to further these goals and actions."

McMahon et al. note that additional master's and doctoral researchers will be needed. They suggest the need for collaboration with the University Consortium for Geographic Information Science (UCGIS) and academic departments.

---

[1] Goal 7: Observe the Earth at all scales using remote sensing to understand the human and environmental dynamics of land change.
Goal 8: Provide timely, intelligent access to new and archived USGS geographic data needed to conduct science and support policy decisions.
Goal 9: Develop innovative methods of modeling and information synthesis, fusion, and visualization to improve our ability to explore geographic data and create new knowledge.

"Through collaboration with other major research organizations . . . the USGS can establish major research objectives in GIScience that meet the needs of all USGS disciplines and *The National Map*." "Formal exchanges with other Federal agencies involved with spatial data (e.g. National Geospatial-Intelligence Agency [NGA], Census) are highly desirable." "USGS Geographers must evaluate linkages with data user communities so that data specifications and analytical capabilities are based on both strong peer-reviewed science and USGS needs."

McMahon et al. recommend using postdoctoral fellows, internships, and close collaboration with USGS scientists and university researchers. "Direct support of university based GIScience research critical to the USGS mission will be necessary." "Linkages to private industry for research and development in GIScience will also be fruitful."

"Scientists affiliated with these centers [CEGIS and two other centers proposed by McMahon et al.] need not all work in a single geographic location, although it is desirable that a core group of staff be collocated. Core researchers include senior scientists who provide leadership and guidance for junior investigators. USGS scientists who are not directly affiliated with the centers but have an active interest in the themes addressed by a center can expect center scientists to serve as an important part of their extended intellectual community and act as collaborators and direct colleagues on some projects."

McMahon et al. suggest the following performance measures for CEGIS:

- Establish CEGIS within two years.
- Staff the center initially with 10 Ph.D. scientists with support staff (within two years) and expand to 20 scientists within five years.
- Prepare a science plan addressing key topics from goals 7, 8, 9 that are needed to meet the overall goals of this plan within two years.
- Establish postdoctoral, internship and visiting scholar relationships with universities with at least 20 scholars in residence within three years and provide continual rotation to maintain 20 per year.

D

Details of CEGIS-Funded Activities in Fiscal Year 2007

The Center of Excellence in Geospatial Information Science (CEGIS) funds two sets of research activities. The first set includes activities that are considered "in-house" within CEGIS. The second set includes activities that are funded through CEGIS but led by U.S. Geological Survey (USGS) researchers outside CEGIS, in some cases in collaboration with non-USGS colleagues.

In-house Activities

The information for this section was provided by Steve Guptill, USGS.

Automated Data Integration

Data integration is a significant problem for *The National Map*. This project will examine data integration from a layer-based approach, developing a conceptual framework based on resolution, geometric accuracy, and topological consistency, and apply it to five of *The National Map* data layers—digital orthoimages, elevations, land cover, hydrography, and transportation. From the experience with the layered approach and the data developed, the project will examine a feature approach to integration based on a model previously developed and implemented as a feature library. The team anticipates significant results leading to an automated approach based on the conceptual framework, the empirical results, and the use of these in metadata to drive an automated process.

- Study Sites

  - Atlanta, Georgia
  - St. Louis, Missouri
- Publications and Reports
  - Implementation of *The National Map* Road Database—from the American Congress on Surveying & Mapping (ACSM) Annual Conference, Nashville, Tennessee, April 2004
  - Integration of *The National Map*: Data Layers and Features—from the American Society for Photogrammetry and Remote Sensing (ASPRS) Annual Conference, Denver, Colorado, May 2004
  - Integration of *The National Map*—from the XXth Congress of the International Society of Photogrammetry and Remote Sensing, Istanbul, Turkey, July 2004
  - Integrating Data Layers to Support *The National Map* of the United States—from the International Cartographic Conference, Coruña, Spain, July 2005

Generalization for **The National Map**

To meet the goals of *The National Map* the USGS must accept high-resolution data from local, state, and other sources and merge these data into a consistent framework at an appropriate resolution. To the extent possible, this process should be automated, transparent to users, and occur in real time as part of *The National Map* viewer or the data delivery system. Part of this process will require spatial data generalization.

- Publications and Reports
  - Generalization for *The National Map* with emphasis on the NHD—Abstract from the 25th Environmental Systems Research Institute (ESRI) International User Conference, July 25-29, 2005
  - Estimation of Accumulated Upstream Drainage Values in Braided Streams Using Augmented Directed Graphs—Paper from the *Auto-Carto 2006, A Cartography and Geographic Information Society Research Symposium*, Vancouver, Washington. June 25-28, 2006

Multiresolution Raster Data for **The National Map**

As science moves toward regional and global analyses with models of climate and human-induced change, methods are needed to project raster data accurately. The approach for this research theme is to use the theoretical and empirical base of knowledge to (1) design a new projection method accounting for raster cell size and latitude effects on accuracy, (2) systematically analyze the error effects and develop error correction procedures, and (3) develop raster resampling algorithms that use the error analysis and correct for inaccuracies. The project will also leverage results from previous USGS and academic research on projecting raster data to establish the necessary knowledge base for the decision support system and the error correction procedures.

- Publications and Reports
    - Open-File Report 01-181—Methods To Achieve Accurate Projection of Regional and Global Raster Databases
    - Open-File Report 01-383—Methods To Achieve Accurate Projection of Regional and Global Raster Databases
    - Projecting Global Raster Databases—Abstract from the Geoinformatics for Global Change Studies and Sustainable Development Conference, Nanjing, China, June 2002
    - Projecting Global Raster Databases—Paper from the International Symposium on Geospatial Theory, Processing, and Applications Conference, Ottawa, Canada, July 2002
    - A Comparison of Equal-Area Map Projections for Regional and Global Raster Data
    - Projecting Global Datasets to Achieve Equal Areas—Peer-Reviewed Paper from the Cartography and Geographic Information Science Journal, Vol. 30, Issue 1, Jan 2003
    - User's Guide to the Decision Support System for Map Projections
    - Accurate Projection of Small-Scale Raster Datasets—Paper from the 21st International Cartographic Conference, 10 . 16 Aug 2003, Durban, South Africa
    - Open-File Report 03-433—Users Guide for the MapImage Reprojection Software Package
    - Open-File Report 2004-1394 User's Guide for the MapImage Reprojection Software Package, Version 1.01
    - Scientific Investigations Report (SIR) 2004-5297—A Decision Support System for Map Projections of Small Scale Data
    - Re-projecting Raster Data of Global Extent—Abstract from Auto-Carto 2005: A Research Symposium, March 21-23, 2005, Las Vegas, Nevada

***Building an Ontology for* The National Map**

The current evolving standards for the various themes of *The National Map* and the historic developments of Digital Line Graph-Enhanced (DLG-E), Digital Line Graph-Feature-Based (DLG-F), and National Hydrography Dataset (NHD) formal specifications provide a cohesive basis for a new ontology that can support *The National Map*. The existing standards must be cast into the new environment of multiscale representation, near-real-time and web access, and on-demand product generation. This can only be accomplished with a complete ontology of all features at all possible representation scales as the basis for feature and information retrieval from the multiple databases that comprise *The National Map.* This project (which starts in 2007) will be the initial step in building such a comprehensive ontology and will use current geographic information science (GIScience) methodologies developed in the ontology of geographic information that have evolved over the last five years.

Fractal and Variogram Analysis of Scale and Resolution Effects in Geospatial Data

Fractals and variograms are established methods to determine effects of scale and resolution in geospatial phenomena and processes. This project (which starts in 2007) will use these methods to examine the impacts of scale and resolution on data integration and generalization for *The National Map* and the National Spatial Data Infrastructure (NSDI).

Elevation Technology Assessment

The evolving private sector capabilities in the elevation technology area, primarily LIDAR (light detection and ranging), are growing at a rapid rate. Multiple states are now in the process of acquiring state-wide coverage and planning for a broad range of statewide applications including floodplain mapping, hydrologic mapping, watershed characterization, vegetation characterization, and structure analysis. In order to incorporate new technological capabilities into the elevation theme's National Elevation Dataset and non-bare earth elevation features, time must be spent evaluating new developments in the areas of bathemetric LIDAR, full-waveform LIDAR, bare earth processing algorithms, software packages, ground-based LIDAR as an approach to accuracy assessment, and requirements studies. References on this work are listed at *http://lidar.cr.usgs.gov*.

National-Scale Elevation Feature Extraction

This project will conduct the research necessary to build a strategy for the nationwide extraction of important elevation derivatives. As the National Elevation Dataset moves more and more to a LIDAR base, and support mounts for a nationwide collection, there is intense interest in the potential high-quality, high-resolution elevation parameters that can improve the flood, fire, landslide and debris flow, storm surge, and water quality and quantity modeling processes. A national approach will unite federal agencies and other partners in the systematic development of this information.

Much of this work will be based on existing work with elevation derivatives embodied in the Elevation Derivatives for National Applications (EDNA) database. EDNA is a multilayered database derived from a version of the National Elevation Dataset (NED, documented at *http://ned.usgs.gov*) has been hydrologically conditioned for improved hydrologic flow representation. The seamless EDNA database provides 30 meters resolution raster and vector data layers including aspect, contours, filled digital elevation models (DEM), flow accumulation, flow direction, reach catchment seedpoints, reach catchments, shaded relief, sinks, slope, and synthetic streamlines.

Hydrologically conditioned elevation data, systematically and consistently processed to create hydrologic derivatives, can be useful in many topologically based visualization and investigative applications. Drainage areas upstream or downstream from any location can be traced accurately facilitating flood analysis investigations, pollution studies, and hydroelectric power generation projects. For further information including publications, conference proceedings and downloadable posters, see *http://edna.usgs.gov.*

CEGIS-Funded USGS-wide Research

CEGIS supports interdisciplinary teams of scientists to address GIScience research issues. The CEGIS Research Prospectus supports cross-bureau research on issues that are a high priority for the USGS. The objectives are to support the use of diverse scientific data from multiple sources and to provide new insights to address complex issues; to apply established science tools and techniques to unique and challenging questions; and to foster opportunities to conduct and report collaborative research that can increase the impact of the science CEGIS does. The projects funded in FY 2007 are the following:

- ***Scaling, Extrapolation, and Uncertainty of Vegetation, Topographic, and Ecologic Properties in the Mojave Desert***
  Principal Investigator (PI)—David R. **Bedford**; Co-PIs – Leila Gass, Sue Phillips, Jayne Belnap; Collaborator—David M. Miller (Southwest Surficial Processes and Mapping) ($73,000)
- ***A Landscape Indicator Approach to the Identification and Articulation of the Ecological Consequences of Land Cover Change in the Chesapeake Bay Watershed, 1970-2000***
  PI—Peter **Claggett**; Co-PIs—Janet S. Tilley, E. Terrence Slonecker (EPA); Collaborator—Bill Jenkins (EPA) ($132,000)
- ***Assessing Local Uncertainty in Non-stationary Scale-Variant Geospatial Data***
  PI—Susan **Colarullo (**$117,000)
- ***Methods to Quantify Error Propagation and Prediction Uncertainty for USGS Raster Processing***
  PI—John **Gurdak**; Co-PI—Sharon Qi ($134,000)
- ***The Geoscience of Harmful Invasive Species: Integrating LANDFIRE and Invasive Species Data for Dynamic and Seamless***

Integration of Raster and Vector Data to Meet Management Needs at Multiple Scales
PI—Thomas J. **Stohlgren**; Co-PIs—Zhi-Liang Zhu, Catherine Jarnevich; Collaborators—Tracy R. Davern, Robert K. Peet (University of North Carolina), James F. Quinn (University of California-Davis), James J. Graham (Colorado State University), Gregory J. Newman (Colorado State University), Kathryn Thomas ($150,000)

- ***Mapping Inundation at USGS Stream Gage Sites: A Proof of Concept Investigation***
  PI—James P. **Verdin**; Co-PIs—Kwabena O. Asante (Science Applications International Corporation [SAIC]), Jerad Bales; Collaborators—Jodie Smith (SAIC), Kristine Verdin (SAIC), Silvia Terziotti ($150,000)
- ***GEOLEM: Improving the Integration of Geographic Information in Environmental Modeling through Semantic Interoperability***
  PI—Roland **Viger**; Barbara Buttenfield (University of Colorado); Co-PIs—Olaf David (Colorado State University/U.S. Department of Agriculture), Charles O'Hara (Mississippi State University); Collaborators––Frank Geter (Natural Resources Conservation Service), Jeff Hamerlinck (University of Wyoming ($150,000).

E

Acronyms and Abbreviations

AJAX	Asynchronous JavaScript and XML
API	application programming interface
BARO	Basic and Applied Research Office
CEGIS	Center of Excellence for Geospatial Information Science
COGIT	Conception Objet et Généralisation de l'Information Topographique
CRADA	Cooperative Research and Development Act
CSW	Catalog Service for Web
DARPA	Defense Advanced Research Projects Agency
DCM	Digital Cartographic Model
DHS	Department of Homeland Security
DLM	Digital Landscape Model
DOE	Department of Energy
DOI	Department of the Interior
EPA	U.S. Environmental Protection Agency
EROS	Center for Earth Resources Observation and Science
ESRI	Environmental Systems Research Institute
FEMA	Federal Emergency Management Agency
FGDC	Federal Geographic Data Committee
FTE	full-time equivalent
GEOLEM	Geospatial Object Library for Environmental Modeling
GIS	geographic information system

GIScience	geographic information science
GNIS	Geographic Names Information System
GOS	Geospatial One-Stop
GPS	Global Positioning System
IGN	Institute Géographique National
ISO	International Organization for Standardization
LCS	Least Squares Collocation
LIDAR	light detection and ranging
MSC	Mapping Science Committee
NCGIA	National Center for Geographic Information and Analysis
NGA	National Geospatial-Intelligence Agency
NGAC	National Geospatial Advisory Committee
NGPO	National Geospatial Program Office
NGTOC	National Geospatial Technical Operations Center
NHGIS	National Historical Geographic Information System
NOAA	National Oceanic and Atmospheric Administration
NRC	National Research Council
NSDI	National Spatial Data Infrastructure
NSF	National Science Foundation
OCG	Optional Content Group
OGC	Open Geospatial Consortium
OS	Ordnance Survey
PDF	Portable document format
R&D	research and development
REST	Representational State Transfer
SLD	styled layer descriptor
SOA	Service Oriented Architecture
SOAP	Simple Object Access Protocol
SPIRIT	Spatially-Aware Information Retrieval on the Internet
TIGER	Topologically Integrated Geographic Encoding and Referencing
TIN	triangulated irregular network
TNM	The National Map
UCD	user-centered design
UCGIS	University Consortium for Geographic Information Science

UID	User Interface Design
USDA	U.S. Department of Agriculture
USGS	U.S. Geological Survey
WCS	Web Coverage Services
WFS	Web Feature Services
WMS	Web Map Services