The National Center for Geographic Information and Analysis

by Michael F. Goodchild

GIS merges an amazing array of diverse interests, and the fact that a number of interests can, with different levels of development, use a common technology is important. Three decades ago remote sensing hardly existed, and disciplines such as photogrammetry, cartography, geodesy and spatial analysis had their own unique technologies. With a common digital technology, many of the old disciplinary boundaries are breaking down. What is emerging may be the beginnings of a science of geographical information, with the ability to focus on generic issues rather than on the idiosyncracies of individual technologies.

The suggestion that science be organized around different classes of information makes sense in an age in which information is increasingly becoming the limiting factor in human activity. The extent to which spatial information, as distinct from other types, is valuable is open to debate, but it is clear that an enormous variety of human activities are organized spatially and that spatial information was largely ignored in the first three decades of the digital revolution. Problems in spatial statistics also are less tractable than the more conventional kind, despite their importance in understanding uncertainty in spatial data, and there has never been a spatial information theory to parallel theories in other nonspatial disciplines (Coombs 1964).

This growing sense of the commonality of issues prompted the U.S. National Science Foundation's (NSF) decision in 1988 to establish a National Center for Geographic Information and Analysis (NCGIA). The center consists of a consortium of three institutions: the University of California, with U.C. Santa Barbara as the lead institution, the State University of New York at Buffalo and the University of Maine. David S. Simonett and Michael F. Goodchild are the center's co-directors at Santa Barbara; Terence R. Smith, Ross D. MacKinnon and Andrew U. Frank are the associate directors at Santa Barbara, Buffalo and Maine respectively. General oversight is provided by a 17-member board of directors, chaired by John E. Estes. Abler has described the process of creating the center, which began with a proposal to NSF from Jerome Dobson of Oak Ridge National Laboratory in 1984 (Abler 1987). MSF funding for the center is at an initial level of $1.1 million per year for five years.

The center's primary purpose, as defined in the NSF solicitation document issued in late 1987, is to conduct "basic research on geographic analysis, utilizing geographic information systems." The emphasis is on applications, particularly scientific and policy-oriented applications, rather than on technical developments. On one hand, the statement of purpose implies that the technology offers significant potential for a wide range of geographically based analyses. On the other, it suggests that although GIS has been widely adopted as a technology in numerous fields, its applications to date often have been relatively unsophisticated. Openshaw (1987) described GIS as "20th century technology being used for 19th century purposes," implying dissatisfaction with the somewhat rudimentary nature of many applications. According to NSF's solicitation document, the center's goals are to: "Advance the theory, methods and techniques of geographic analysis based on GIS in the many disciplines involved in GIS research; augment the nation's supply of experts in GIS and geographic analysis in participating disciplines; promote the diffusion of analysis based on GIS throughout the scientific community; and provide a central clearinghouse and conduit for disseminating information regarding research, teaching and applications."

The first and third goals give the center a role in promoting GIS specifically as an enabling technology for science. The GIS community currently is dominated by applications in management, inventory and policy formation, and the potential scientific applications of GIS have attracted comparatively little interest to date from the scientific community. Scientific applications stress the power of GIS to place information in a spatial context, suggest relationships based on spatial proximity and explore the role of distance as a causal factor. For example, GIS can assist in superimposing spatially organized data such as maps from different sources, and one can envision an "exploratory spatial analysis (ESA)" tool analogous to the Exploratory Data Analysis (EDA) tools now common in statistics (Tukey,1977). GIS technology increases our ability to view data from different perspectives and under different forms of manipulation and summary. In fact, it seems that the power of spatial organization to suggest causes, explanations and relationships is significantly superior to other forms of data organization such as tables or graphs. Recently ESA tools have begun to offer the capability to view data through multiple windows (tubular, graphical and cartographic) with simultaneous updating, so that movement of a cursor within the cartographic window results in continuous and appropriate updating of the tabular window, for example.

Applications of GIS in the social sciences have lagged behind those in the earth and natural sciences for many reasons. Funding for technical tools is less readily available; there is a general suspicion of technical approaches in many areas of social science; data may be less reliable and harder to obtain; and spatial analysis has only recently become common in disciplines such as history and anthropology. The center's role as a promoter of scientific applications of GIS applies especially to the social sciences.

In addition to ESA, scientific applications require comparatively sophisticated capabilities for modeling and analysis, which are lacking in many GIS activities. The center therefore stresses not only the incorporation of spatial modeling and analysis techniques with GIS, but also the formal, theoretical basis for GIS and the development of an intellectual and conceptual core in the field.

A useful analogy is to compare GIS as a scientific tool and other widely distributed scientific software tools, such as the statistical packages including SAS, SPSS and BMD. These software packages have emerged in a similar timeframe as GIS, but almost entirely as products for the scientific market, whereas the scientific applications of GIS have had a relatively small role in driving and directing GIS' development. One can argue that GIS is a supporting tool for spatial analysis in the same way that SAS, for example, is a supporting tool for statistical analysis. Given the usefulness of spatial organization and the spatial context for data, one might contend that the long-term potential for GIS in this context is as large as it is for scientific software.

Based on these goals, the center's programs fall into the general areas of research, education and outreach.

Research Plan

The center's research agenda is based on the proposition that GIS technology has enormous potential in a variety of applications, particularly science, but that numerous impediments currently exist that constrain the full realization of that potential. The Research Plan (NCGIA 1989) described these impediments in detail under five general headings (in large part these match the five suggested areas of research contained in the NSF solicitation), so they will be merely summarized here:

- Spatial analysis and spatial statistics: Impediments exist in the lack of implementation of spatial analytic methods within GIS, and also in the lack of explicit treatment of data quality (accuracy and uncertainty) in current systems.
- Spatial relationships and database structures: The power of a GIS is constrained by the methods used to represent spatial data within its database. Current systems use a limited range of data models and structures, often derived from cartographic representations, and are based on an inadequate understanding of the nature of spatial relationships.
- Artificial intelligence (AI) and expert systems: Many GIS analysis and modeling requirements are poorly structured and could benefit from AI technologies, as could the complex process of data input and output.
- Visualization of spatial data: Electronic displays offer enormous potential for improved visualization methods, but current GIS technology largely fails to exploit the capabilities of the new medium.

• Social, economic and institutional issues: GIS technology raises numerous managerial, organizational and legal issues, and the adoption of GIS is currently impeded by the difficulties of accurately assessing its costs and benefits, and by inadequate understanding of its impact on organizations.

In formulating a specific agenda of research, the center looked for a mechanism that would allow it to operate as a three-institution, multidisciplinary, multi-investigator consortium. but at the same time focus attention on well-defined topics within the broader research plan. The center also realized that the plan defined a range of research topics that would be far too large for it to handle alone and that the center would need to encourage and stimulate as much outside research as possible, and provide effective two-way communication.

Research Initiatives

The research plan is implemented through a series of research initiatives, based on a model which has worked well at UCSB's Institute for Theoretical Physics. An initiative lasts for one or two years, with four or five initials running at any one time. An initiative begins with a specialist meeting, with 25 to 50 people from three constituencies: organizations with experience with the problem in question and its effects, researchers with interests in solving the problem, and representatives from the vendor community who can implement solutions. The purpose of the specialists' meeting is to lay out the specific research agenda, including tasks that can be accomplished by the center or affiliated groups and individuals in the timeframe of the initiative, but recognizing that the meeting will stimulate research by other groups and individuals as well. Subsequent research at the center is undertaken by faculty research assistants and visiting researchers, and the results are reported at a suitable occasion such as a national or international conference. In several cases it has been useful to establish a mailing network associated with an initiative to distribute current information about research activities during the research period.

The current program of initiatives is as follows, with the names of initiative leaders and dates of specialist meetings:

1. Accuracy of Spatial Databases (Michael F. Goodchild, Santa Barbara, December 1988). GISs are high-precision systems that process data as if the data were perfectly accurate. In reality, spatial data are subject to high levels of uncertainty and inaccuracy, which current GIS designs ignore. More than 50 people attended the specialist meeting for the first research initiative and a book has been published from the proceedings (Goodchild and Gopal 1989). Research has been under way for almost a year, and results are starting to appear in different forms. Specific research activities include a bibliography and taxonomy of spatial data errors to increase user awareness, fundamental work on the formulation of error models, methods for incorporating error information within spatial databases, analysis of error propagation through GIS processes, and development of finite resolution data structures.

2. Languages of Spatial Relations (David M. Mark, Buffalo; Andrew U. Frank, Maine; January 1989). GIS can be broadly seen as a technology for helping people work with spatial data and, more specifically, as a tool for learning and reasoning about space and spatial relationships. As such, the technology will be most useful when its data representations and operations emulate the learning and reasoning processes of users; yet current data models and structures fall far short of this ideal. Initiative 2 is conducting research to help improve the digital representation of spatial data and the design of GIS user interfaces. The research agenda includes the following topics: wayfinding, driving directions and processes of spatial knowledge acquisition; analysis of the structure, cross-cultural and cross-linguistic variations of driving directions, with potential applications to vehicle navigation aid systems; cross-linguistic analysis of locative expressions, and studies of linguistic variation in natural language terms for spatial relations; user interface design, including research on multi-media interfaces, metaphors for conveying and perceiving spatial information, and the visualization of spatial relations; and formalization of spatial relationships, the algebra of spatial relations and formal reasoning.

3. Multiple Representations (Barbara P. Buttenfield, Buffalo; February 1989). The representation of a geographical feature on a map depends on the map's scale, and the same feature can be represented in different ways at various scales. Within a spatial database it is attractive to imagine that a feature might be given a single representation, which would be generalized or simplified for display at different scales. Because this has proven difficult to do, it is common for databases to contain several representations of the same features. Moreover, GIS databases currently provide no explicit and fully satisfactory means of relating the various representations of a feature logically. Hierarchical data structures offer potential and are one of the research topics of this initiative. Other topics include definition of the rules required to automate the generalization process, systems for describing the ways features change with scale, and data structures that formalize the logical relationships of multiple representations. The center is developing a multi-agency, multi-scale database to be distributed as a standard for research work in this area.

4. Use and Value of Geographic Information (Harlan Onsrud, Maine; Hugh Calkins, Buffalo; May 1989). This is the first initiative to address the social, economic and institutional issues raised by the adoption of GIS technology. Three research themes emerged at the specialist meeting: the need for a taxonomy of geographic information and its uses, development of objective methods for measuring the value of geographic information, and empirical studies of the diffusion of GIS technology. The taxonomy must address questions such as what types of geographic information exist and how they relate to the variety of data models of spatial databases; whether certain types of geographic information are more or less suitable for handling in spatial databases; and what role geographic information plays in human activity, who uses this information, and for what purposes. Objective measurement of the value of geographic information is an essential component of any serious attempt to evaluate GIS benefits. Finally, research on the diffusion of GIS technology will determine the factors that control the rate of diffusion, and how these factors can be modified.

5. Design and Implementation of Large Spatial Databases (Terence R. Smith, Santa Barbara; Andrew U. Frank, Maine; July 1989). Two meetings were held on this topic in July 1989: a symposium with formal position papers, and a smaller workshop discussion to lay out the initiative's research agenda. The initiative will examine the technical problems that arise in handling the large spatial databases now being constructed, such as the U.S. Geological Survey's digital cartographic database.

6. Spatial Decision Support Systems (Paul J. Densham, Buffalo; Michael F. Goodchild, Santa Barbara; March 1990). The specialist meeting for this initiative was held in March 1990, and addressed the issues surrounding the development of decision support systems based on GIS technology (Densham and Goodchild 1989).

Future Initiatives

Future plans call for further initiatives that will extend the range of research topics, but at the same time build on the results of previous work. Initiatives 7 through 13 have been identified:

7. Visualization of the quality of spatial information,

8. Expert systems for cartographic design,

9. Institutions sharing spatial information,

10. Temporal relations in GIS,

11. Space-time statistical models in GIS,

12. Remote sensing and GIS, and

13. User interface design.

This list of initiatives represents a grouping and prioritization of the much larger set of impediments identified in the research plan. It is likely that the order will change and that the list will be extended in response to future changes in priorities.

Several specific issues are likely to affect future planning for research initiatives. First, although the methodology and subject matter for much of Initiatives 2, 4 and 9 belong in the social sciences, as well as many of the applications of Initiative 1, no initiative is directed explicitly at developing the social science applications of GIS. To rectify this, the center is planning a conference specifically on this topic in late 1990. Second, GIS technology has enormous potential in modeling the global environment and the interaction between human and physical systems at global scales. Yet no initiative is explicitly concerned with global scale GIS, although there is relevant research in Initiatives 1 and 5. Data input and conversion are significant for the GIS community but are not explicitly treated in the current agenda. It may be necessary to find ways of increasing the number of initiatives that can be supported at any one time, perhaps by a mechanism that would allow initiatives to be led at sites outside the center.

Education

The rapid development of the GIS field in the past few years has led to an acute shortage of adequately trained staff at all levels, particularly in those areas that require a moderate level of technical skills combined with an understanding of GIS application areas. The major center effort in education in the first year has been the Core Curriculum project, designed as a way to quickly increase the availability of GIS teaching materials so that courses could be introduced in new institutions and new disciplines.

The curriculum has been developed as a one-year sequence, as three quarter courses or two semester courses, with a total of 75 one-hour lectures. In the three quarter arrangement, the courses are:

1. Introduction to GIS (lectures 1-25),
2. Technical Issues in GIS (lectures 26-50), and
3. Application Issues in GIS (lectures 51-75).

The materials consist of lecture notes (typically six to eight pages for a one-hour lecture), supporting materials (handouts, slides and overheads), lab exercises for courses one and two, exam questions and discussion topics. Course one consists of 250 pages of material, a selection of slides and six diskettes containing data and text for the lab exercises.

Between August and December 1988, the proposed course outline was developed and circulated for comment, and presented at several conferences. Thirty-five experts from the GIS community were then invited to contribute the texts of specific lectures. These were edited and augmented by the center, and preliminary versions of the courses were completed by July 1989 (courses one and two) and October 1989 (course three). The project was discussed extensively at the GIS in Higher Education Conference at Ohio State University in June 1989. During the 1989/90 academic year 74 institutions agreed to evaluate and test the materials by incorporating them into their offerings. A large proportion of these institutions are geography departments, but the test sites also include marine science, geology, anthropology and engineering departments. Instructors will be providing feedback to the center on each lecture and lab, and comments will also be obtained directly from students. These will be used to develop a second, final version of the curriculum that will be ready for distribution in the summer of 1990.

The Core Curriculum project has generated significant interest, not only in the academic community but also in industry and in government agencies. At the same time the project addresses only part of the overall GIS education and training problem. Specifically, other areas the center would like to pursue in the future include short courses for Core Curriculum instructors; short courses in specialized topics (two such courses were developed in the first year as one-day workshops and presented at several conferences); a case course to study several well-documented case studies designed around the model used in many business schools; and training courses for users of specific systems, emphasizing analytic and modeling capabilities.

The Future for GIS

The high growth rates experienced recently in the GIS industry are exciting, but inevitably lead to concerns about the future including how long the GIS phenomenon can last, and whether GIS has been oversold. To some extent, growth has been sustained because new application fields have appeared, or new disciplines have become interested in GIS tools, but this cannot continue forever.

GIS is a loose collection of interests held together by common hardware and software solutions, whose long-term survival depends on the emergence of an intellectual core and on the symbols and institutions normally associated with a discipline. It is difficult to think of precedents in the form of disciplines that have been founded on tools, although one might argue that computer science emerged when computing found an intellectual foundation and developed a program of basic, fundamental research. The future seems to offer two alternative scenarios for GIS: a technology that failed to deliver on its promises, or a technology that blossomed into a geographical information science.

Many believe that GIS is driven by technology, rather than by any clearly understood set of objectives. Its drive is undisputable and shows little sign of diminishing. It raises a host of issues, many of which have been recognized for a long time in more established fields like cartography, but gives these issues new impetus. When the GIS phenomenon finally runs its course, and the acronym disappears from view, I hope we will be left with a heightened awareness of the common issues that underlie all of the disciplines that collect, compile and analyze geographical information, and a better knowledge of how to deal with these issues.

Michael F. Goodchild is co-director for the National Center for Geographic Information and Analysis, University of California, Santa Barbara, CA 93106.

Acknowledgment

The National Center for Geographic Information and Analysis is supported by the National Science Foundation, grant SES 88-10917.

Further Information

The center publishes a newsletter in June and December each year, and a technical papers series. Information on these and other publications, and the center's activities in research, education and outreach can be obtained from any of the three sites:

NCGIA
University of California
Santa Barbara, CA 93106
NCGIA
State University of New York
Buffalo, NY 14260
NCGIA
University of Maine
Orono, ME 04473

References

Abler, R.F. 1987. **The National Science Foundation's National Center for Geographic Information and Analysis**. *International Journal of Geographical Information Systems.* 1:303-26.

Coombs, C.H. 1964. *A Theory of Data.* Wiley & Sons. New York.

Densham, P.J. and M.F. Goodchild, 1989. **Spatial decision support systems: A research agenda.** *Proceedings, GIS/LIS '89.* ASPRS/ACSM. Bethesda, MD.

Goodchild, M.F. and S. Gopal, editors. 1989. *Accuracy of Spatial Databases.* Taylor and Francis. London

NCGIA. 1989. **The research plan of the National Center for Geographic Information and Analysis.** *International Journal of Geographical Information Systems.* 3:117-36.

Openshaw, S. 1987. **Guest editorial: An automated geographical analysis system.** *Environment and Planning.* 19:431-6.

Tukey, J.W. 1977. *Exploratory Data Analysis.* Addison-Wesley.