INT. J. GEOGRAPHICAL INFORMATION SYSTEMS, 1991, VOL. 5, NO. 3, 341–351

Taxonomy for surveying the use and value of geographical information†

HUGH W. CALKINS

National Center for Geographic Information and Analysis, State University of New York at Buffalo, 301 Wilkeson, Amherst, New York 14261, U.S.A.

and NANCY J. OBERMEYER

Department of Geography and Geology, Indiana State University, Terre Haute, Indiana 47809, U.S.A.

Abstract. Understanding the use and value of geographical information in decision making has been identified as crucial to the further development of geographical information systems and related spatial analysis techniques. Research into the use of geographical information will be through surveys and case studies. A taxonomy for investigating the use of geographical information and its associated value is needed to structure the surveys and case studies. This paper presents an initial attempt at such a taxonomy, intended to support continuing research into the use and value of geographical information. The survey methodology to be supported will be mostly telephone and/or mail questionnaires, and this has influenced the structure of the taxonomy. The goal of this work is to create a taxonomy suitable for use in surveys that will enhance the understanding of the use and value of geographical information in decision making.

1. Introduction

As the technology of geographical information systems (GIS) develops and gains wider acceptance and increasing rates of adoption, there is a need to develop a system for classifying and characterizing the uses of geographical information. The sharing of individual experiences via a commonly accepted set of definitions will facilitate a faster and better understanding of the use of such information. Although the purpose of this taxonomy is to consider the use of geographical information, no matter what technology is used, the vehicle for observing the use will be through GIS. This is necessary as GIS are more rigorously structured and documented than the general use of geographical information from other sources (maps, diagrams, photos). However, the broad objective is to understand the use of geographical information and geographical analysis in decision making.

† This paper reports work that is related to research being carried out at the National Center for Geographic Information and Analysis, supported by a grant from the National Science Foundation (SES-88-10917). The Government has certain rights in this material; support by the NSF is gratefully acknowledged.

Interest in GIS can be traced to the efforts of the U.S. Bureau of the Census (in the late 1950s) and to the large metropolitan land use on a transportation studies conducted in the late 1950s and 1960s. In both instances, funding was provided by the Federal government, either for federal agencies or through planning grants to metropolitan areas. Although the computer technology of that time was very rudimentary, many of the concepts and ideas incorporated into the commercial GIS packages of today were used in the early systems. The history of GIS is generally one of many individuals and agencies creating limited purpose GIS to meet specific needs. Early efforts to understand and describe these were conducted as case studies (Tomlinson *et al.* 1976, Calkins and Marble 1980). Case studies of systems and applications documented the level of activity at that time. These studies were limited for several reasons: (1) each GIS was 'home-grown' and therefore unique; (2) widely accepted definitions of GIS and their constituent parts did not exist; and (3) the GIS community was not well defined. The most comprehensive picture of activity in GIS was a survey conducted by Calkins and Marble (1980). Other surveys attempting to determine the scope and uses of GIS were generally ineffectual.

The focus on GIS became dormant in the late 1960s and 1970s as federal programmes and funds for the development of GIS were severely cut. Isolated user groups, fledgling professional associations (e.g. the Urban and Regional Information Systems Association) and a small number of GIS vendors (e.g. the Environmental Systems Research Institute) kept the ideas alive during the late 1970s.

The use of GIS began to grow very rapidly in the early 1980s, concurrent with the introduction of the microcomputer. Although most GIS covering a large area cannot effectively use microcomputers because of limitations on data storage, computer literacy and awareness of the potential for computing contributed to increased interest in GIS. Along with the increasing interest, the need for the more effective communication of ideas, experiences and potential applications to a large number of possible GIS users has become essential. As long as the GIS professional community remained relatively small, formalized structures for communication were not really needed. However, given the truly explosive growth of interest and activity in GIS experienced in the 1980s, communication relying on more precise definitions and formalized structures has become a necessity.

One of the first systematic scientific activities in the study of any group of related objects is to develop meaningful taxonomies. The purpose of taxonomies is to classify objects to facilitate comparison between these objects. Next, generalization about the objects can allow the discovery of significant characteristics which are scientifically meaningful. 'The theory of the use and value of geographic information in decision-making is poorly developed... The role of information in decision-making, the mechanisms and processes by which information is defined and utilized, and the economic theory of information must all be addressed' (NCGIA 1988, 53). Research focusing on the use and value of geographical information should determine if current theories adequately explain the observed use and value, or if expanded or new theories are needed. The initial research will focus on observing and measuring the uses of geographical information. A taxonomy will provide the necessary means to structure, classify and generalize the data collected through literature reviews, surveys and case studies.

The remainder of this paper is devoted to an introduction to the characteristics of a taxonomy and to a taxonomy suitable for investigating the use of geographical information.

2. Definitions and distinctions in taxonomy

Before proceeding with a taxonomy (or taxonomies) for the use of geographical information, it is useful to clarify the terms used in taxonomies. The definitions below follow the work of Sneath and Sokal (1973) and Simpson (1961):

Systematics: the scientific study of the kinds and diversity of organisms and of any and all relationships among them

Classification: the ordering of organisms into groups (or sets) on the basis of their relationships. Such relationships may be one of several types: (1) genetic; (2) phylogenetic (evolutionary); or (3) phenetic (referring to similarities of phenotype)

Taxonomy: The theory and practice of classifying organisms; the theoretical study of classification, including its bases, principles, procedures and rules

Identification: the allocation or assignment of additional unidentified objects to the correct class once a classification has been established

The objective of the taxonomy described in this paper is to provide a framework for the study of the use of geographical information and a subsequent attempt to identify the benefits of such use (preferably in a price-based form compatible with economic evaluation methods). Taxonomies attempt to classify objects on the basis of one or more characteristics (or attributes) of the object for identification and description in unambiguous terms. Taxonomies, or classifications, can be useful in a systematic study of the objects. The classifications should allow each instance of the object under study to be observed and measured in a manner to facilitate comparison between observations and generalization to recognize significant patterns exhibited by the objects.

Taxonomy is neither a simple nor a neutral task. As Gould (1983, 72) argues:

> Taxonomy is often regarded as the dullest of subjects, fit only for mindless ordering and sometimes denigrated within science as mere "stamp collecting"... If systems of classification were neutral hat racks for hinging the facts of the world, this disdain might be justified. But classifications both reflect and direct our thinking [about the objects under observation]. The way we order represents the way we think. Historical changes in classification are the fossilized indicators of conceptual revolution.

Perhaps the greatest difficulty inherent in the development of a taxonomy is subjectivity. In paragraphs subsequent to his argument that 'classifications both reflect and direct our thinking', Gould (1983, 72) describes and discusses Michel Foucault's challenge to the French mid-seventeenth century practice of imprisoning the poor, the unemployed and the insane. Foucault argued that this practice reflected a way of viewing the world and its citizens that saw a common characteristic of individual members of these three groups of people: idleness. That there were what we would regard as legitimate extenuating circumstances (such as mental illness and lack of opportunity) to explain the idleness of many of these individuals was irrelevant.

Numerical taxonomists hold that, '... classification can be neither right nor wrong. It is not a theory, but merely a way of summarizing information in an intelligible form. One assesses its value by considering its usefulness to other [scientists]' (Dunn and Everitt 1982). Once in place, systems of classification have the capacity to influence our thinking. Such systems select a limited number of characteristics that are formulated into criteria by which we categorize and evaluate objects or phenomena. In addition to

subjectivity in the development of classes, subjectivity is also possible in identifying objects or phenomena as members of a specific class. Ideally, a system of classification would allocate every object to one, and only one, class and all objects would be assigned to a class. Although a clear and unambiguous classification will be sought for each instance of an object, this may be difficult to achieve. Differences in the concepts, definitions and terminology previously used in a field tend to influence the way we observe objects. These factors interact with the process of observation and may bias the resulting classifications. Further, contradictions and conflicts may be discovered, suggesting either modifications to the taxonomic structure or, more fundamentally, a re-thinking of concepts and definitions within the field. The taxonomy presented in this paper is a first attempt to develop a structure suitable for the investigation of the use and value of geographical information in decision making. It will certainly require modification and change resulting from both review by experts in the field and lessons learned during subsequent surveys of the users of geographical information.

3. Classification strategies for the use of geographical information

A taxonomy should reflect the purposes or objectives of the activity for which the taxonomy is created. In this instance the objective is to characterize the use of geographical information in decision making. A second objective is to identify and quantify the value, or benefits, of the use of geographical information to one or more parties.

A concept or model of the process of decision making using geographical information is needed for the basis of the taxonomy. This model identifies the components of the decision making system which need to be represented in the taxonomy (figure 1). The model shows a complete set of steps, from observing entities and their spatial characteristics to GIS produced information, which provides information for decision making. The model consists of a number of stages through which the geographical information passes as it is prepared for use in decision making. Each stage involves the transformation of the geographical information which may: (1)

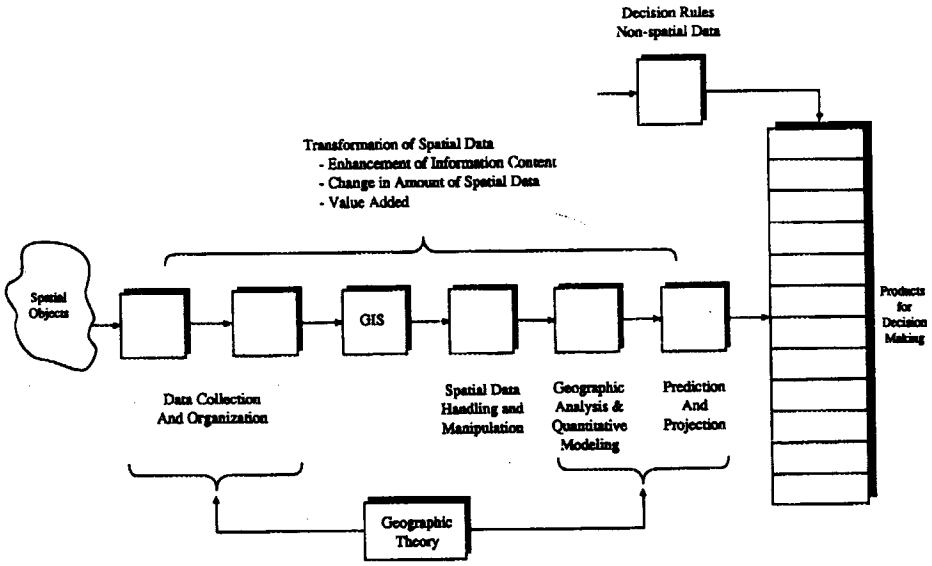

Figure 1. Organizational use of geographical information for decision making.

enhance or deplete the information content; (2) change the amount of data or information; and/or (3) add or subtract value to/from the data/information. The GIS is represented as a stage in the middle of the overall model. At some point, possibly the GIS stage, data become information as a result of the transformations required for a specific instance of decision making. The view that information improves decision making assumes that the information content is enhanced, the amount of data is decreased (summarization, abstraction), and that value is always added by each successive transformation. The other extreme is also possible: information can be lost, the volume of data increased, and the value decreased with each transformation. Decisions, in the context of this paper, can be either very complex, involving a large number of transformations, or relatively simple. A simple decision would be the use of a map type product to determine location by executing a simple spatial query. On the other hand, a complex decision to determine the optimum location for a facility might require a spatial decision support structure to analyse many possible alternative locations. In each instance, however, selected transformations will occur between the GIS and the end-user of the information. The objective of the taxonomy presented is to provide a structure for studying such transformations. The transformations from entities in space to the GIS are not of particular interest in this study as these have already been adequately described. If research using the proposed taxonomy indicates that the pre-GIS transformations independently affect use and value, these transformations will have to be re-examined.

The taxonomies presented here are defined in terms of the set of questions for which answers are sought to further the understanding of the use of geographical information and its value (or benefits). Each question, or part, of the overall taxonomy represents a single view of the factors thought to affect the use of geographical information. These views not only reflect what we would like to know at the present time, but also recognize the current state of the development of systematic procedures for the use of geographical information and the constraints imposed by the current level of development. The taxonomies presented here are not representative of any complete taxonomy of the use of geographical information or GIS, but reflect the questions which can be answered at the present time, within the objectives of this study. Other taxonomies designed for different objectives could be constructed and would be valid for a broader understanding of the role of GIS and geographical information in a societal context.

The components of the taxonomy are based on three major questions reflecting what we want to know about the use of geographical information.

(1) What successful uses are being made of geographical information, successful being defined as use that creates benefits? Specifically:
   Who uses geographical information?
   What is the level of decision making of the user?
   What is the purpose for using geographical information?
   What are the steps of the decision making process where geographical information is used?
   What application areas are using geographical information?
(2) How effective is the use of geographical information? Specifically:
   Who receives the benefits?
   What is the level of need for the geographical information?
   What is the effect of the use of geographical information?

(3) What are the benefits attributable to the successful use of geographical information? Specifically:
What are the types of benefits?
How are these benefits measured?

Additionally, there is a second category of questions that are of interest here. These questions are related to the way in which geographical information is processed and made available to users within organizations.

(1) What characteristics of geographical information processing are associated with successful use?
What data are associated with successful use?
What geographical units are used?
How are the data represented (individual observations, summarized)?
What are the primary functional capabilities required in a GIS to support uses?
What forms of spatial analysis are used/useful?
What is the form or format of the geographical product?
Why is the display of geographical information useful?
What amount of geographical information is used?
What is the response time of the GIS?
(2) What organizational factors are associated with successful use?
How is the task of handling spatial data organized?
What type of organization is using the geographical information?
What is the role of the organization in the handling of geographical information?

4. Taxonomy for investigating the use and value of geographical Information

The development of a taxonomy of geographical information and its use has gained interest as research on geographical information systems turns increasingly towards institutional issues. It is a significant research effort within Initiative 4 (titled the Use and Value of Geographic Information in Decision-making) of the National Center for Geographic Information and Analysis. The basis of the taxonomy presented here is 'phenetic' as this approach relies on the presence of observable features for classification purposes (for a fuller discussion on different classification strategies and the reason for using phenetic in this effort, see Obermeyer 1989).

4.1. *Characteristics of successful uses of geographical information*

(1) Who uses the geographical data or information?
a. professional analyst
b. middle management
c. upper management (decision maker)
d. general public
e. organized public opinion groups
(2) What is the level of decision making of the user?
a. policy
b. management
c. operations
(3) What is the purpose for using geographical information?
a. to build or maintain an inventory of significant data (possibly responding to a mandate)

b. to monitor a process, such as the issuance of permits where the objective is to insure compliance with appropriate regulations or procedures; to keep track of a series of planned events over a predictable time period;
c. to browse through a geographical database
   (i) structured (goal oriented) browsing, e.g. environmental monitoring
   (ii) unstructured, e.g. spatial exploratory data analysis
   (iii) looking up specific facts
d. to summarize geographical data
   (i) simple counting
   (ii) measurement (i.e. area or perimeter of a polygon)
e. to analyse mapped or spatial data, using the geographical domain as the organizing principle where the problem is well defined
   (i) map overlay analysis
   (ii) network analysis (route finding, districting, optimal site selection using network distance, etc)
   (iii) terrain analysis models
   (iv) spatial interaction models (based on a zonal structure)
   (v) flow analysis through a network (stream network models)
f. spatial decision support systems to assist in reaching a decision in situations where the problem is ill-defined

(4) What are the steps in the decision making process where geographical information is used?
   a. identification of problems
   b. development of alternatives
   c. selection of solution
   d. programming of implementation
   e. monitoring of results

(5) What are the application areas? (e.g. urban planning, forestry, water resource management, facilities management)

4.2. *How effective is the use of geographical information?*

(6) Who receives the benefits (to whom do benefits accrue?)
   a. agency using GIS (direct)
   b. another (possibly higher level) agency (indirect)
   c. general public
   d. individual

(7) Information is:
   a. needed—decision cannot be made without
   b. desirable—decision may be made with/without
   c. extraneous—decision will be made anyway

(8) Effect of information
   a. assumptions questioned
   b. additional alternatives evaluated
   c. decision is modified or changed
   d. no effect

4.3. *What are the benefits of the use of geographical information?*

(9) What are the benefits (Type 1)?
   a. current use
   b. anticipated use (future)
   c. unanticipated use (future)

(10) What are the benefits (Type 2)?
   a. savings in staff time
   b. increased productivity
   c. avoidance of costs
   d. reduction of risk (decision making)
   e. new applications, formerly not possible

4.4. *Measurement of benefits*

(11) How are benefits measured or identified?
   a. price based (benefits in terms of cost including savings in staff time converted to cost)
   b. intangible benefits

(12) What is the appropriate period for accrual of benefits?
   a. short, immediate (costs and benefits occur within the same budget cycle)
   b. long term (benefits accrue during and after the cost budget cycle)

4.5. *Characteristics of geographical data and spatial analysis*

(13) What data are associated with successful use?
   a. Themes or layers, such as land use, vegetation

(14) What geographical units are used?
   a. point
   b. line
   c. polygon
   d. regular tessellation (e.g. grid cells)

(15) How are the data represented?
   a. sample data
   b. network topological information
   c. areal unit information
   d. surface information

(16) What are the primary functional capabilities required in GIS to support uses?
   a. display of graphic data
   b. query and display
   c. counting and direct measurement operations
   d. map overlay (map combinations)
   e. network algorithms
   f. spatial models

(17) What forms of spatial analysis are used/useful (from Brandeau and Chiu 1989)?
   a. objective (optimizing; non-optimizing)
   b. decision variables (server/facility location; service area/dispatch priorities; server capacity/volume; type of goods produced by each server; routing/flows of server or goods to demand points; queue capacity; other)

c. system parameters (topological structure; travel metric; travel time/cost; demand; number of servers; number of service facilities; number of commodities, server location; server capacity; service area and dispatch priorities; service discipline; queue capacity)

(18) What is the form or format of the geographical product?
   a. maps
   b. other graphics
   c. tabular listing

(19) Why is the display of geographical information useful?
   a. general base map for orientation
   b. visualization of complex data and spatial relationships
      (i) results of analysis (structured analysis or spatial decision support)
      (ii) volume of data is very large
   c. display not required

(20) What amount of geographical information is used? Actual value measured in one of several ways
   a. requirements for data storage
   b. number of geographical entities
   c. amount of attribute data
   d. amount of image data
   e. combination of these

(21) Response time requirements
   a. perceived need
   b. actual

Response time can be measured in seconds, minutes, hours, etc. Most user perceptions of GIS are, however, conditioned by the high technology image of computing. Consequently, expectations normally run much higher than are needed, or than can be supported by current technology.

4.6. *Organizational factors*

(22) How is the task of spatial data handling organized?
   a. task level (specific use of a procedure by a user to complete an individual task)
   b. applications—a set of tasks that have similar characteristics, both procedural and substantive
   c. module—a related set of application procedures organized as a single unit and presented to the user through a standard interface.

(23) What type of organization is using the geographical information?
   a. public
   b. private
   c. non-profit making

(24) What is the role of the organization in the handling of geographical information?
   a. user only of geographical information
   b. provider of geographical information (no use by providing department or agency)
   c. both provider and user
   d. maintains and uses its own geographical data

5. Discussion

This taxonomy has identified 24 questions, loosely grouped into six categories. The categories represent what we would like to know about the use of GIS and the role of geographical information in decision making.

Although each component of the taxonomy has been presented as a single question, it is expected that the resulting surveys using the taxonomy will identify dependencies between components. In fact, identifying these dependencies is one of the objectives of investigations into use and value of geographical information. For example, the purpose for using geographical information may correlate with the kind of geographical data, the geographical units, the data representations, the GIS functional capabilities and the kind of spatial analysis. If such correlations are found, a grouping of some or all components should be possible to define a high level use or application of geographical information. Then the subsequent identification of the benefits and value of such an application can be determined.

The use of GIS as a starting point for investigating these questions is necessary to gain an initial overview of the use and verification of the taxonomy. Testing the taxonomy will involve three activities: (1) building a suitable survey instrument; (2) verifying the ability of users to supply answers to the survey questions; and (3) determining the appropriate universe for continuing surveys.

A programme of surveys (both broad, general surveys and selected, in-depth surveys) can then be defined to investigate the following questions:

(1) Who are the users of geographical information in decision making and for what purpose are they using the information?
(2) Who comprise the set of similar individuals or institutions who are not using geographical information for similar purposes?
(3) What classes of decisions are supported by GIS and spatial analysis?
(4) What benefits are reported by the users of geographical information, focusing on questions of identification, quantification and verification of the benefits?
(5) What personal, organizational and institutional factors are related to the successful use of geographical information and what benefits are associated with such use?

6. Conclusions

The taxonomy presented aims to answer several questions and to foster further discussion on the use and value of geographical information. The taxonomy will serve as an initial structure to provide the basis for surveys (by telephone or mail). The taxonomy should evolve in two separate ways. It may be difficult or impossible to frame an appropriate set of answers for selected questions or respondents may have difficulty with the defined classes. If this is so, the taxonomy will need revision accordingly. Secondly, it may take more questions and more detailed answers to develop a complete picture of the uses and benefits of geographical information. The development and testing of survey instruments and the analysis of preliminary results may reveal the need for additional questions. The provision of geographical information in the past, either by map products or geographical information systems, has been an expensive undertaking. The continuing costs of providing this information, especially if new technology is to be used, need to be justified both economically and in terms of overall value to society. This taxonomy has been presented as a tool to assist in determining the value of this information.

References

BRANDEAU, M. L., and CHIU, S. S., 1989, An overview of representative problems in location research. *Management Science*, **35** (6), 645–674.

CALKINS, H., and MARBLE, D., 1980, *Computer Software for Spatial data Handling: Full Geographic Information Systems* (Ottawa, Canada: International Geographical Union Commission on Geographical Data Sensing and Processing).

DUNN, G., and EVERITT, B., 1982, *An Introduction to Numerical Taxonomy* (Cambridge: Cambridge University Press).

GOULD, S. J. (editor), 1983, The titular bishop of Titiopoulos. In: *Hen's Teeth and Horses Toes* (New York: W. W. Norton).

NCGIA, 1988, Technical Paper # 88-2: Proposal of the National Science Foundation, The University of California at Santa Barbara, the State University of New York at Buffalo, and the University of Maine at Orono.

OBERMEYER, N., 1989, A systematic approach to the taxonomy of geographic information use, *Proceedings of GIS/LIS89, Orlando, Florida, U.S.A.* (Bethesda, MD: American Society for Photogrammetry and Remote Sensing; Washington. D.C.: Association of American Geographers; Washington, D.C.: Urban and Regional Information Systems Association; Englewood, CO: AM/FM International.)

TOMLINSON, R. F., CALKINS, H. W., and MARBLE, D. F., 1976, *Computer Handling of Geographical Data* (Paris: UNESCO).

SIMPSON, G. G., 1981, *Principles of Animal Taxonomy* (New York: Columbia University Press).

SNEATH, P. H. A., and SOKAL, R. R., 1973, *Numerical Taxonomy* (San Francisco: Freeman).