GEOGRAPHIC INFORMATION SYSTEMS: AN OVERVIEW

Tang Yang Tze
School of Computing and Information Technology,
INTI College Subang Jaya, Selangor, Malaysia.
(tangyt@inti.edu.my)

ABSTRACT

Geographic Information System (GIS) is an integrated information system that provides timely and quality spatial information for specific geographical applications. GIS ranges from a small piece of software programme to a highly complex and sophisticated integrated communications system. GIS is itself an integration of various specific subsystems. It also serves as an integration technology as it integrates knowledge from different disciplines and functions of various information technologies into one single system. GIS has been extensively used in government and private sectors for a number of scientific and non-scientific applications. Open GIS is the concept that provides inter-operability, extensibility and flexibility to this specific type of information system. This paper presents a generic definition of GIS from a broad perspective, the principal features and functions of GIS, and the application areas of GIS. The paper also includes a brief discussion on the integration technology concepts of GIS.

INTRODUCTION

Information systems have become an integral part of the information society today. Information systems support the information needs for the operations, management and decision making functions in an organisation, which include collecting, processing, storing, transmitting and displaying information in an effective and efficient way. Information systems are widely used in various modernised organisations for a wide range of applications. The types of information systems that are frequently used to support the operations, management and decision-making functions include transaction processing systems, knowledge work systems, decision support systems and management information systems. However, this paper is concerned with the specific information systems called "Geographic Information Systems."

"Geographic Information Systems" is a collective, general term for describing information systems manipulating geographic data. Geographic information systems vary significantly from one to another, depending on the application domain. However, most geographic information systems have one thing in common: they were designed to provide timely and quality spatial information for specific geographical applications.

DEFINITIONS OF GIS

Generally, there is no universal definition for GIS. Different researchers presented definitions with different emphases on various aspects of GIS. In the broadest possible terms, geographic information systems are tools that are designed for processing geographic data into useful and meaningful information.

The United States Geological Survey (USGS) defines a GIS as "a computer system capable of capturing, storing, analysing, and displaying geographically referenced information; that is, data identified according to location" (Anon., 2003).

Gianfranco defines GIS as "an application software (user interface, general and specific functions) that allows the input, visualisation, and analysis of geographic data (geo-references) or remote-sensing images." (Scrinzi *et al.*, 2000) DeMers defines GIS as "tools that allow for the processing of spatial data into information, generally information tied explicitly to, and used to make decisions about some portion of the earth." (DeMers, 2000).

Jeffrey defines GIS as "an information system that is designed to work with data referenced by spatial or geographic coordinates. In other words, a GIS is both a database system with specific capabilities for spatially-reference data, as well as a set of operations for working with data." (Star and Estes, 1990).

Densham used the term spatial decision support system (SDSS) to describe a system that "... normally is implemented for a limited problem domain. The database integrates a variety of spatial and non-spatial data and facilitates the use of analytical and statistical modelling techniques. A graphical interface conveys information, including the results of analyses, to decision makers in a variety of forms. Finally, the system adapts to decision makers' style of problem solving and is easily modified to include new capabilities" (Densham, 1991). Most of the commercial GIS, particularly desktop GIS, fall into the definition of SDSS.

CONCEPT OF GIS

A Geographic Information System (GIS) is a special-purpose information system, which is designed to serve specific geographical applications. GIS supports the input, visualisation and analysis of geographic data (geo-references) or spatial information. GIS associates the graphical representation of each territorial element with its respective classification in reference to a specific chosen theme. The geographic data in a GIS are described and organised in layers, each representing a thematic characteristic with respect to a specific spatial region.

A comprehensive GIS is equipped with an input function that is capable of capturing geographic data from various sources (maps, aerial photos, satellite images and surveys), a storage function that is capable of managing the data (storage, retrieval and query), a processing function that is capable of manipulating the geographic data (transformation, analysis and modelling), and an output function that is capable of visualising the data in various forms (reports, maps, plans, spatial statistics, etc.).

GISs are often related to other database applications, but with a major difference where information in a GIS is linked to spatial references. This means that the database of a GIS uses geo-references as the primary means of storing and accessing information. A database application may use locational information (such as street, address or zip code) for manipulations, but a GIS database uses coordinates for manipulations.

GISs should be perceived as an integrated solution rather than a software or hardware. This is because GISs are often used for analysing situations and making decisions. The way in which data are entered, stored and analysed within a GIS must reflect the way spatial information will be used for a specific research or decision-making task.

CONCEPTS OF INTEGRATING TECHNOLOGY

GIS is a special-purpose information system that requires various specific data capturing and processing technologies. Hence, the development of GIS is more toward integrating the existing technologies into the system, rather than developing the entire system, which could be really costly and time-consuming.

The concept of integrating technology is important in the development of GIS. GIS integrates a number of discrete technologies into a whole, rather than being completely new. Therefore, GISs differ significantly from one to another, depending on the information requirements and the types of applications. GIS may range from a small piece of application software such as Desktop GIS (ArcView and MapInfo) to a large, complex information system such as Global Positioning System (GPS).

GIS is often an integration of a number of new and existing discipline-specific tools into a single system, depending on the problem domain. Each of the discipline-specific tools is with a specific application domain. For example, the development of an airway navigation GIS could be an integration of the computer-assisted cartographic drafting tool, photogrammetric tool, spatial databases, spatial statistical analysis and modelling tool, and a coordination system.

Besides, GIS is also an integration of knowledge from different disciplines into a single application. The integration is often determined by the desired application domain. For example, the development of a natural resource management application may require adequate knowledge in geography, cartography, geodesy, geology, statistics, engineering, and computer science.

These integrating concepts (technologies and knowledge) of the GIS provide flexibility and extensibility to the development life cycle. The integrating concepts enable new technologies to be integrated into an existing GIS, rather than rebuild the system. Thus, it is more flexible and cost effective. Besides, the integrating concepts make it possible to customise GIS for similar or different problem domains.

CONCEPTUAL MODEL OF GIS

Current GISs range from loosely coupled systems to integrated systems. The first GISs were dedicated systems, which are also known as application-specific systems. The lack of ability to develop new applications and the lack of general database functionalities such as query languages, along with the lack of functional extensibility, are among the limitations of such architectures. Today, the available systems (from the second and third generation) rely on two systems of architectures: loosely coupled systems architecture (e.g. ArcInfo from ESRI and Tigris from Intergraph) and integrated systems architecture (e.g. ArcView from ESRI and Smallworld from Smallworld, Ltd.) (Philippe *et al.*, 2002).

Several conceptual models have been proposed by researchers for presenting the unique features and characteristics of GISs. However, these models are generally application-specific and lack details of integrating concepts. For example, the conceptual model proposed by Mennecke (1997) for presenting the features in GIS shown in Figure 1.

The integrating concept should always be a major factor in the designing of GIS. When designing a GIS, the system developer needs to consider what requirements and functionalities are to be incorporated into the design.

A conceptual model of the GIS proposed by the author is presented in Figure 2 for describing the concept of integrating technologies. In this model, GIS consists of a core system that serves as an integrator or coordinator for the other subsystems. The subsystems include digitizing or data capturing devices (input subsystem), modelling or imaging devices (output subsystem), spatial DBMS (spatial data storage subsystem), relational DBMS (general storage subsystem), and information processing and knowledge management facilities (processing subsystem). Each of these subsystems could be an existing technology that has its unique characteristics and behaviours.

A Spatial Database Management System (SDBMS) is a software module that can work with an underlying database management system, for example, Object-Relational DBMS or Object-Oriented DBMS. Unlike the relational database, in which every attribute of an entity consists of exactly one atomic value, the attributes in a spatial database may contain a

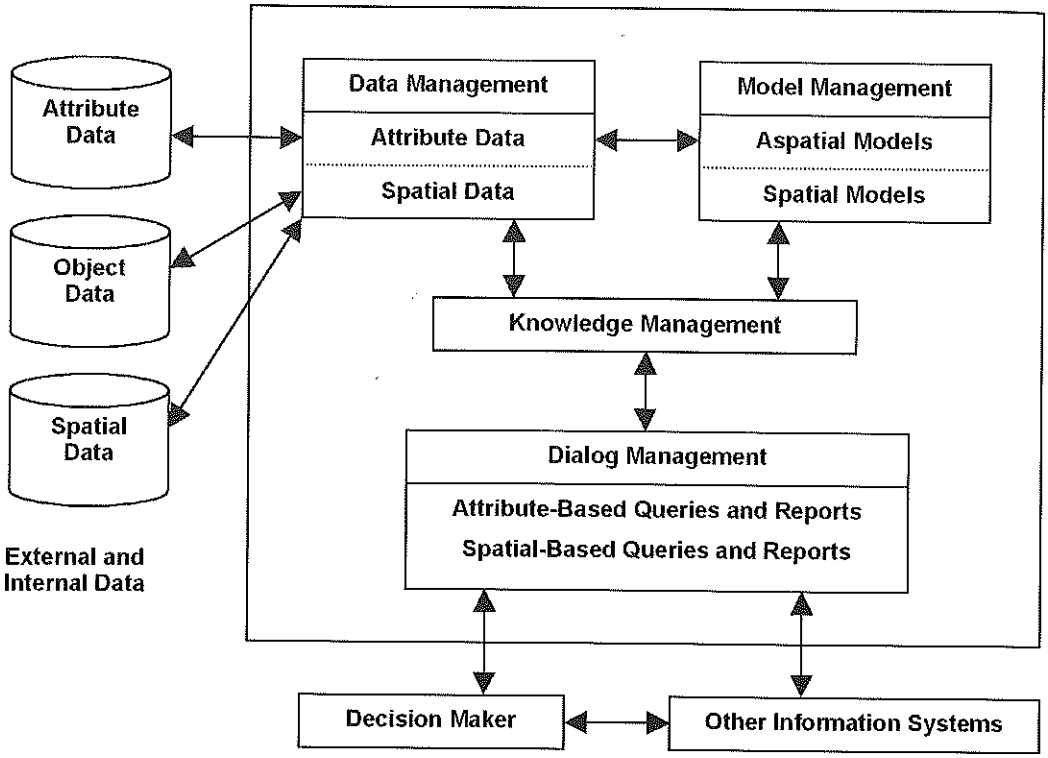

Figure 1. A conceptual model of a GIS used for decision support.

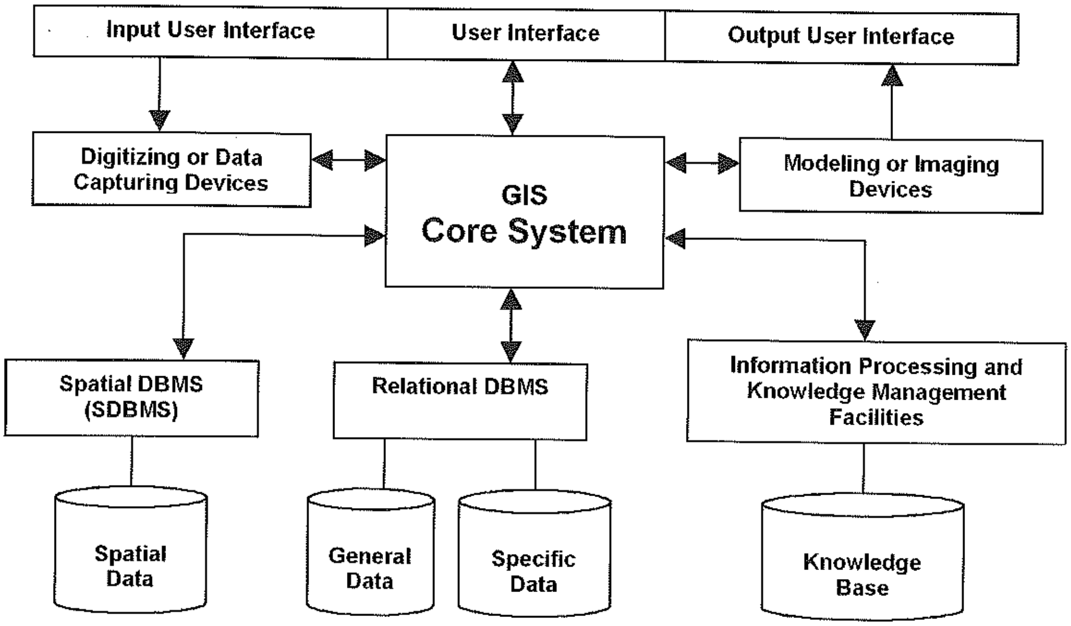

Figure 2. Conceptual model of GIS

set of spatial data (such as coordinates and geometric data). The SDBMS supports multiple spatial data models, commensurate spatial abstract data types (ADTs), and a query language from which these ADTs are callable (Shashi and Sanjay, 2003).

IMPLEMENTATION OF GIS

As with other information systems, GIS must be managed properly to be used effectively. Business School researchers possess a rich literature that can be applied to studying the use of GIS in organisations (Moore, 1993). For instance, an important concern facing management when adopting a new technology is whether the technology will be accepted and used by members of the organisation. Besides, when designing a GIS, it is not necessarily talking about the actual software design, although this is an important part of the process. System designers have many non-technical issues to consider as well as some broader technical issues when implementing GIS into organisations.

One of the most important issues for a successful GIS implementation is the accuracy of the identification of user requirements. This issue has many subdivisions as user requirements cover a broad range of things. Designers will need all information on user requirements if the system is to be successful. Proper identification of user requirements can avoid a concept known as 'design creep', which occurs in the absence of organisational learning. DeMers addressed the importance of accurate information that initiate the system design process in such a way that "at one extreme, is a system with more functionality than necessary, and at the other, is a system having an incorrect functionality" (DeMers, 2000).

Another important issue for consideration is the adaptability of the system. Whether it is just a piece of software system or a large GIS project, the more specialised it is, the less adaptable it is for new tasks. It is important that GIS is designed in a way that it can be altered when needed. Therefore, when organisations discover new needs, the system can be modified efficiently and at a reasonable cost.

The projection and registration of the digital map is probably another important concern for the use of GIS. Projection is a fundamental component of mapmaking. A projection is a mathematical means of transferring information from the Earth's three-dimensional, curved surface to a two-dimensional medium. Different types of maps use different projections because each projection is particularly appropriate for certain uses. For example, a projection that accurately represents the shapes of the continents will distort their relative sizes.

Map information in a GIS must be manipulated so that it registers, or fits, with information gathered from other maps. The digital data may have to undergo projection conversion before they can be analysed. Since much of the information in a GIS comes from existing maps, digital information gathered from different sources with different projections, need to be transformed into a common projection.

POTENTIAL USES OF GIS

GIS has the ability of relating information from different sources in a spatial context and reaching a conclusion about this relationship. Most of the real-world information contains a location reference, which relates that information at some point on the globe, using a location reference system (longitude and latitude, and perhaps, elevation).

By relating and comparing the information from different sources, scientists (or decision makers) can identify the interactivity of the resources based on specific locations. Therefore, GIS can reveal important new information that leads to better decision making.

Many databases that can be directly entered into a GIS are being produced by many different organisations such as government agencies,

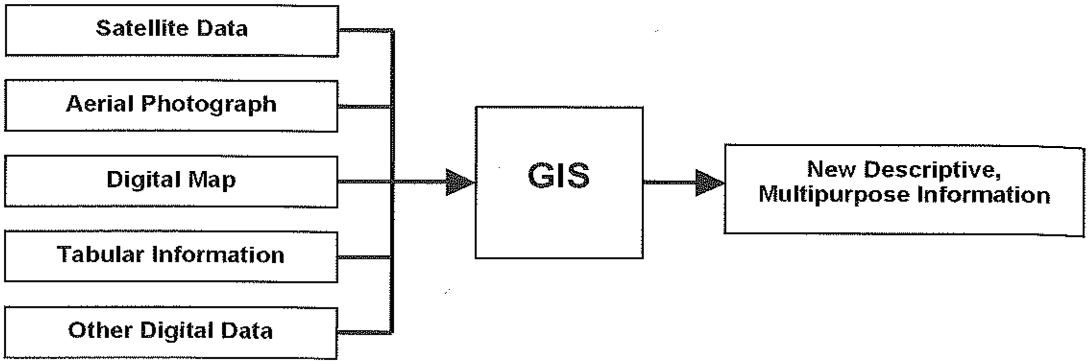

Figure 3. Concept of data integration with GIS.

private companies, academic institutions, and nonprofit organisations. Different kinds of data in map form can be entered into a GIS, and then converted into forms it can recognize and use. For example, digital satellite images can be analysed to produce a map of digital information. Likewise, tabular data can be converted to a map form and serve as layers of thematic information in a GIS.

This concept of relating and converting information of different types into useful new information is known as the data integration. The concept of data integration is illustrated in Figure 3.

Some analytical information could be difficult to associate or relate. But, GIS makes it possible to link or integrate them. GIS can use combinations of mapped variables to build and analyse new variables. For example, by using GIS technology, it is possible to combine industrial records (such as the amount of waste water produced by a parcel of industrial land) with hydrography data to determine which streams will carry certain levels of water pollution. The GIS can be used to predict the level of water pollution in each stream by locating these parcels and intersecting them with streams. Then as streams converge, the total loads can be calculated downstream where the stream enters a river, or the sea.

It is extremely difficult to collect data over every square metre of the Earth's surface. Therefore, samples are taken at discrete locations. A GIS can be used to depict two- and three-dimensional characteristics of the Earth's surface, subsurface, and atmosphere from points where samples have been collected. For example, a GIS can quickly generate a map with isolines that indicate the pH of soil from test points. Such a map can be thought of as a soil pH contour map. The two- and three-dimensional contour maps created from the surface modelling of sample points from pH measurements can be analysed together with any other map in a GIS covering the area.

APPLICATIONS OF GIS

Decades ago, the applications of GIS were mainly focused on scientific researches. Business organisations considered such systems as very costly and complex, with very limited use. However, the rapid enhancement in GIS technology and its integrating ability has altered this perception. Today, GIS has been widely used in government and private sectors for both scientific applications and business applications.

The research and development institutions use GIS for a wide range of scientific applications. These applications include natural resource management, emergency response planning, environmental effects simulation, chemical contamination evaluation, wildlife evaluation, computer-assisted cartography or mapmaking, and land management.

The use of GIS incorporated with satellite systems has been a very effective approach for scientists and geologists to study the surface of Earth. The GIS technology has provided a new way of looking at geographic data. It has dramatically changed the speed at which geographic information can be produced, updated and distributed; for example, the use of NAVSTAR's Global Positioning System (GPS), a sophisticated satellite system, for acquiring locational information with reference to a highly accurate geodetic framework. Many analyses can be performed with the assurance of a high degree of measurement accuracy and precision (DeMers, 2000).

Natural disasters such as earthquakes, hurricanes and storms are unpredictable in many countries, and have caused the loss of many lives and properties. Scientists are always looking for better monitoring systems so that these disasters can be predicted. GIS has been extensively used in such research areas to improve the studies of these natural disasters. The ability of integrating geographic data from different sources has made it possible for scientists to study the correlations between geographic factors that contribute to the occurrences of natural disasters. The result is encouraging with the assistance of GIS.

Landis (1993) suggests that most business organisations use information systems for one or more of the five applications: transaction processing, operations, inventory control, planning and decision making, and internal management and control. GIS can be used for these functions because this technology possesses capabilities that are common to traditional aspatial information systems. The GIS business applications include spatial data collection and automated mapping, facility management, demographic and market analysis, transportation and logistics, strategic planning and decision-making, design and engineering.

Mennecke (1997) proposed a conceptual model that portrays four GIS functions and

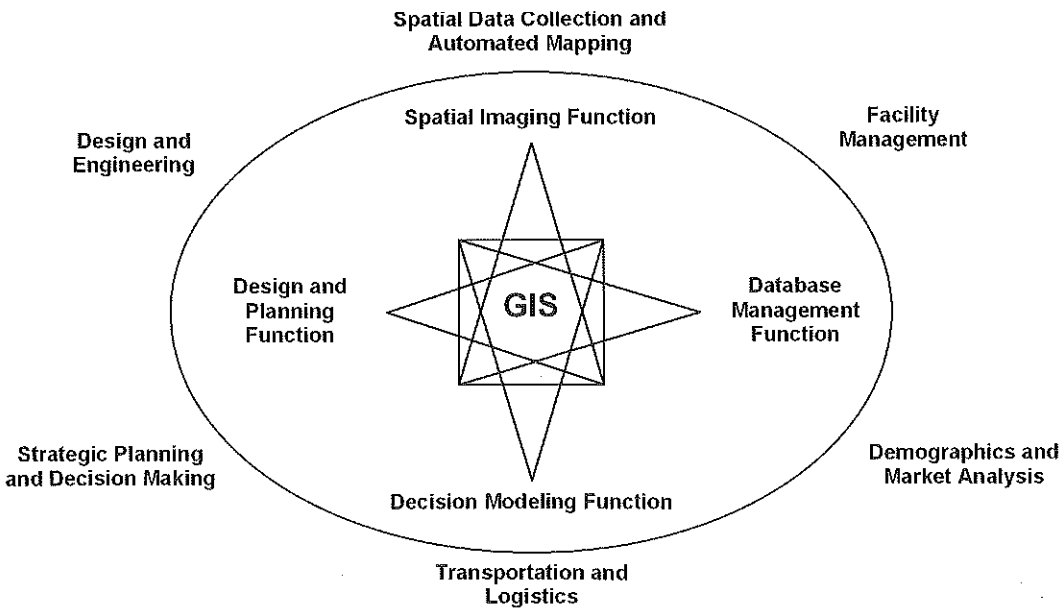

Figure 4. Conceptual model of GIS functions and business related applications

related applications in business (Figure 4). The four functions are derived from four unique activities for which GIS can be used to address the needs of business. These functions are spatial visualisation/imaging, database management, decision modelling, and design and planning. The business related applications are spatial data collection and automated mapping, facility management, demographics and market analysis, transportation and logistics, strategic planning and decision making, and design and engineering.

The use of GIS has compressed the time and distance needed by businesses for getting quality information from a large geographical area. The GIS technology has provided an evolutionary way of improving the responsiveness and competitiveness of businesses and decision makers. As a result, this causes strategic impacts on businesses, which help businesses to gain competitive advantages over the competitors.

Besides, the GIS technology is also providing functionality for effective direct marketing via the spatial point pattern analysis. The analysis uses the kernel density estimation to establish a rough description of the target area based on the customer data. This approach shall provide a tentative hint for the selective employment of geodemographic data in order to decrease the respective expenditures (Adam, 2002).

OPEN GIS

The progress of Open Source Software (OSS) development is becoming the new focus in information systems development. As organisations are always looking toward a more cost effective way of systems development, OSS has become an alternative solution for some of the small- and medium-scale companies. The main idea of OSS is to encourage the development of software by providing the source codes to the users.

Some of the commercial GISs are self-contained within a closed system. This has created a visible barrier on the use of the systems. The objective of Open GIS is to improve the interoperability between GISs, and with other existing systems, or new systems. The idea is to move GIS away from a monolithic system toward a modular system that would encompass different softwares (Philippe *et al.*, 2002).

The Open GIS Consortium (OGC) is a consensus-based association of public and private sector organisations to the creation and management of an industry-wide architecture for inter-operable geoprocessing. OGC works with the government, private industry, and academia to create open and extensible software application programming interfaces for GIS and other mainstream technologies. With a significant amount of commercial GIS software in use by the government, the issue of inter-operability has become very important.

Other major concerns of Open GIS are the extensibility and flexibility of the system. These are the qualitative factors that influence the quality of an information system. Most organisations now operate in a rapidly changing environment. So, when the application domain has been extended due to new information requirements, Open GIS should be able to cope with changes in the new functional requirements.

The flexibility of an information system means different things to people with different perspectives. From the business perspective, flexibility refers to the adaptability of the system to the changing business environment. From an end-user point of view, flexibility probably refers to intuitive and tailorable user interfaces.

LIMITATIONS OF GIS

First, the unavailability of specific GIS functions often makes it impossible to actually perform spatial statistical analysis within the software. In most cases, the geo-coded observations must be exported to another statistical software to calculate the Search and Rescue (SAR) statistics.

Second, some GIS programmes require all data points to be calculated individually rather than using a matrix of distances. This requirement takes more time and computer processing power. This is especially true for high-definition spatial models, which involve a large amount of spatial data.

Third, GIS software is not particularly user-friendly. The integrated working environment for GIS applications often requires a wide range of supporting tools, which require specialised knowledge. Hence, it is rather difficult to make GIS as user-friendly as general application softwares such as a word processor.

The last limitation is the availability of data. Although this limitation is diminishing over time, many GIS data sets have yet to be created for public use. Consequently, researchers must create the datasets themselves. For example, U.S. Census data are available in GIS format only for 1990 and beyond, a limitation that becomes particularly relevant when research using earlier data is undertaken. Moreover, older versions of current maps often do not exist, making time series data difficult to use (Mathis *et al.*, 2003).

SUGGESTIONS FOR FURTHER RESEARCH

One of the major concerns in the implementation of GIS technology is the uncertainty in spatial data and GIS-based analyses. GIS technology has permitted spatial scientists to acquire, process and analyse geographic data, and access multiple spatial databases through networks. All spatial data introduced into a GIS infrastructure carry a certain amount of uncertainty with it.

Without an awareness of the implications of data uncertainty, there is a real danger of making inappropriate decisions from GIS-based analyses. To address this concern, geographic information scientists have proposed three major areas of research: (1) to investigate the origins of data uncertainty and how it propagates through GIS and spatial data infrastructures, (2) to develop techniques for the detection, visualisation, quantification, and prediction of data uncertainty, and (3) to establish a coordinated and systematic approach to study the effects of uncertainty and provide expertise to the public for data management.

Another relatively new and rapidly developed area for further research is spatial data mining. Data mining is the process of discovering potentially interesting and useful patterns of information embedded in large databases. Spatial data mining is a niche area within data mining for spatial analysis, which can potentially influence major scientific challenges, including the study of global climate change (Shekhar and Chawla, 2003).

CONCLUDING REMARKS

The awareness of the importance of GIS has slowly shifted from scientific research to business. This is because most business problems include significant spatial components, and GIS enables decision makers to utilise their resources more effectively. Due to the increasing use of GIS, there is a concern that such systems are designed accurately according to the specifications. Besides meeting user requirements, adaptability and integration are important issues to be considered in order to implement GIS successfully.

GIS technology has enhanced the efficiency and analytic power of traditional mapping. Through visualisation, a GIS can be used to produce images that allow researchers to view their subjects in ways that have never been seen before. GIS technology enables analyses with multiple variables between different regions to be achieved through the availability of digital data on regional and global scales.

An active GIS market has resulted in lower costs and continual improvements in the hardware and software of GIS. These developments will result in a much wider application of the GIS technology throughout government, business and industry.

REFERENCES

DeMers, M.N. (2000). *Fundamentals of Geographic Information Systems. 2nd Edition.* New York : John Wiley & Sons Inc.

Densham, P.J. (1991). Spatial Decision Support Systems. In: *Geographic Information Systems: Principles and Applications*, Vol.1. Maguire, D.J. M.F. Goodchild, and D.W. Rhind (eds), pp. 403-412. London: Longman Scientific and Technical Public.

Landis J.D. (1993). *GIS Capabilities, Uses, and Organizational Issues*. In: *Profiting From A Geographic Information System.* Edited by G.H. Castle. Fort Collins, M.F. Goodchild, and D.W. Rhind, pp. 403-412. London: Longman Scientific and Technical Public.

Mathis M., Fawcett A. and Konda L. (2003). *Valuing Nature: A Survey of the Non-Market Literature.* (VNT-03-01). Houston Advanced Research Center, Woodlands, Texas. (http://www.harc.edu/harc/Projects/Nature/Reports/LitReview-GeoSpatial.pdf)

Mennecke B. E. (1997). Understanding the Role of Geographic Information Technology in Business: Applications and Research Directions. *Journal of Geographic Information and Decision Analysis*, Ontario, Canada. 1 (1): 44-68.

Moore G.C. (1993). Implications From MIS Research for the Study of GIS Diffusion: Some Initial Evidence. In: *Diffusion and Use of Geographic Information Technology*. Masser, I and Onsrud, H.J. (eds.), pp. 77-94. Dordrecht, Netherlands.

Rigaux P, Scholl M. and Voisard A. (2002). *Spatial Databases: With Application to GIS*. Morgan Kaufmann Publishers, USA.

Scrinzi G., Tabacchi G., Tosi V., De Natale F. (2000). Geographic Information Systems (GIS) in Forestry: From Thermatic Representation to Territorial Variable Analysis. Italy. (http://www.isafa.it/scientifica/gis/eng_gispersito.htm)

Shekhar S. and Chawla S. (2003). *Spatial Databases: A Tour*. Prentice Hall. Upper Saddle River, N.J.

Sliwinki A. (2002). Spatial Point Pattern Analysis for Targeting Prospective New Customers: Bringing GIS Functionality into Direct Marketing. *Journal of Geographic Information and Decision Analysis*, Ontario, Canada. 6 (1): 31-48.

Star J. and Estes J. (1990). *Geographic Information Systems*. Prentice Hall, Englewood Cliffs, N.J.

United States Geology Survey (USGS) – Science for a changing world. *Geographic Information Systems*. http://www.usgs.gov/isb/pubs/gis_poster/ (03/12/2003)