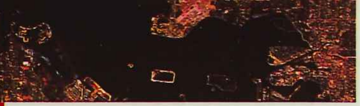
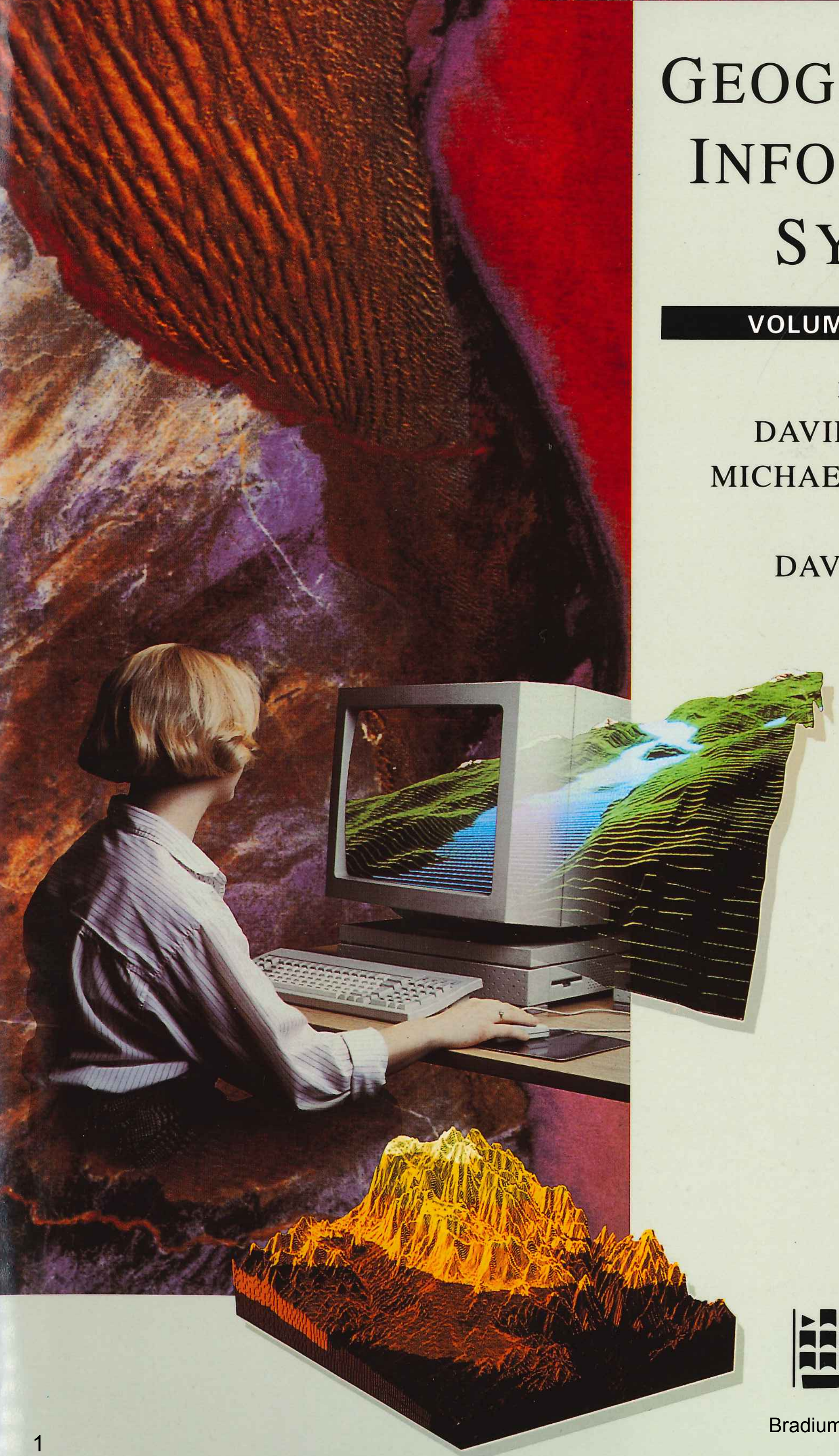
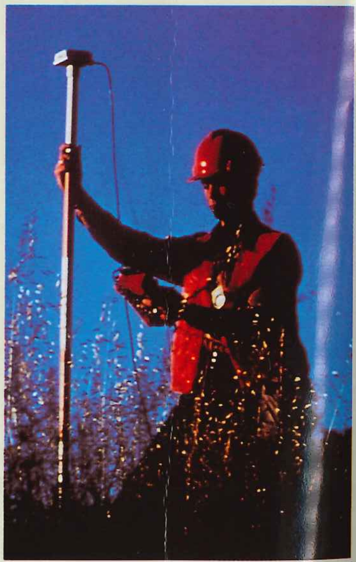




GEOGRAPHICAL INFORMATION SYSTEMS

VOLUME 1 : PRINCIPLES

EDITED BY
DAVID J MAGUIRE,
MICHAEL F GOODCHILD
AND
DAVID W RHIND



Longman
Scientific &
Technical

Exhibit 2006
Bradium Technologies LLC - patent owner
Microsoft Corporation - petitioner
IPR2016-00448

GEOGRAPHICAL INFORMATION SYSTEMS

The Editors

David J Maguire is Lecturer in Geographical Information Systems at the University of Leicester.

Michael F Goodchild is currently co-Director of the NSF-Funded National Center for Geographic Information and Analysis, University of California, Santa Barbara.

David W Rhind is Professor of Geography at Birkbeck College, University of London.

Cover illustrations courtesy © CNES 1991; Earth Satellite Corporation, Geopic; ESRI UK; Intera-Tydac Technologies Inc. ; Laser-Scan Limited; Trimble Navigation Ltd. Paint box facility supplied by IQ Videographics.

GEOGRAPHICAL INFORMATION SYSTEMS

VOLUME 1 : PRINCIPLES

GEOGRAPHICAL INFORMATION SYSTEMS

PRINCIPLES AND APPLICATIONS

EDITED BY
DAVID J MAGUIRE,
MICHAEL F GOODCHILD
AND
DAVID W RHIND

Longman
Scientific &
Technical
Copublished in the United States and Canada with
John Wiley & Sons, Inc., New York

Longman Scientific and Technical,
Longman Group UK Ltd
Longman House, Burnt Mill, Harlow,
Essex CM20 2JE, England
and Associated Companies throughout the world.

copublished in the United States and Canada with
John Wiley & Sons, Inc., 605 Third Avenue, New York,
NY 10158

First published 1991

British Library Cataloguing in Publication Data
Maguire, David J.
Geographical information systems: Principles and applications
I. Title II. Goodchild, Michael F.
III. Rhind, David W.
910.901

ISBN 0–582–05661–6

Library of Congress Cataloging-in-Publication Data
Maguire, D. J. (David J.)
Geographical information ssyyyystems / by D. J. Maguire,
Michael F. Goodchild, and David W. Rhind.
p. cm.
Includes bibliographical references and index.
Contents: v. 1. Principles – v. 2. Applications.
ISBN 0–470–21789–8
1. Geographical information systems.
I. Goodchild, Michael F. II. Rhind.
David. III. Title.
G70.2.M354 1991
910′.285 – dc20 91–3724
CIP

Set in Great Britain

Printed and Bound in Great Britain at the Bath Press, Avon

Dedicated to the memory of

DAVID S SIMONETT

1926–90

David Simonett was born in Australia in 1926. After earning a Doctorate at the University of Sydney, he became a leading pioneer in the field of Remote Sensing, holding faculty positions at the University of Kansas, the University of Sydney and the University of California, Santa Barbara. He was director of land use applications at Earth Satellite Corp from 1972 to 1975.

As Chair at Santa Barbara from 1975, he was able to build one of the foremost Geography programs in the US, culminating in 1988 with the establishment of the National Center for Geographic Information and Analysis. The Santa Barbara site of the Center was renamed the David Simonett Center for Spatial Analysis in 1990 in recognition of his role in its creation. He received the Honours Award from the Association of American Geographers and the Victoria Medal from the Royal Geographical Society.

David Simonett lost a courageous fight against cancer on December 22, 1990 in the course of the preparation of his contribution to this book. The editors dedicate this book to his memory and to the outstanding role he has played in the development of the field of Geographical Information Systems.

Section I Overview

Section II Principles

(a) Nature of spatial data

(b) Digital representation

(c) Functional issues

(d) Display issues

(e) Operational issues

VOLUME 2 : APPLICATIONS

Section III Applications

(a) National and international GIS programmes

(b) Socio-economic applications

(c) Environmental applications

(d) Management applications

Section IV Epilogue

2

THE HISTORY OF GIS

J T COPPOCK AND D W RHIND

Computer-based GIS have been used since at least the late 1960s: their manual predecessors were in use perhaps 100 years earlier. Acknowledging the paucity of well-documented evidence, this chapter describes the background to the development of such systems, stressing the context in which such development took place, the role of organizations and individuals where this can be ascertained, and the applications which the systems were intended to meet. A broad definition is taken of GIS so as not to exclude any significant developments; computer mapping systems of all types (including those with line-printer graphics, the forerunners of contemporary raster systems) are included.

It is demonstrated that most, but by no means all, of the early developments originated in North America. The roles of key organizations such as the US Bureau of the Census, the US Geological Survey, the Harvard Laboratory for Computer Graphics and the Experimental Cartography Unit are described and the activities of the commercial sector are exemplified by a case study of Environmental Systems Research Institute. Reasons are suggested for significant international differences in the development of GIS, such as the attitudes to ownership of data and the perceived role of the state. It is concluded that several stages of evolution of GIS can be defined. These overlap in time and occur at different moments in different parts of the world. The first, or pioneering age, extended from the early 1960s to about 1975; in this, individual personalities were of critical importance in determining what was achieved. The second phase, approximately from 1973 until the early 1980s, saw a regularization of experiment and practice within and fostered by national agencies; local experiment and action continued untrammelled and duplication of effort was common. The third phase, running from about 1982 until the late 1980s, was that of commercial dominance. The fourth (and current) phase is one of user dominance, facilitated by competition among vendors, embryonic standardization on open systems and increasing agreement on the user's perception of what a GIS should do and look like.

INTRODUCTION

A variety of information indicates that the field of GIS has expanded rapidly in recent years (see Maguire 1991 in this volume). From where did all this business and the resulting jobs arise? Unhappily, we scarcely know. GIS is a field in which history is little more than anecdotal. To rectify this, a search through the archives of government departments and agencies would certainly help. As yet, however, few organizations have given any thought to formalizing the history of their involvement in GIS and at least one major player (Ordnance Survey; see Finch 1987) has refused to let its detailed records be examined by external researchers. Less certainly, the records of computer hardware and software companies could also be a source of relevant information but no such material has been uncovered. Unfortunately for those writing the history of GIS, neither staff of commercial companies nor government officials have a tradition of writing books or papers on their experience of an emerging technology. Research staff in government or private sector research organizations are exceptions to this rule but, even for them, writing papers for the benefit of the scientific community at large has a relatively low priority. As far as is known, the only official attempt anywhere to provide a broad overview of the field as a whole is that given by the Report of the Committee of Inquiry into the Handling of Geographic Information (Department of the Environment 1987; Rhind and Mounsey 1989).

The main source of information, with all the risks of partisan bias, remains researchers in the academic community. In reality, however, even the numbers of academics working in this field were quite small until the expansion of the last decade. Moreover, as Chrisman (1988) and Rhind (1988) both testify, those active in universities in this field in the early stages of the development of GIS were often outside the formal academic career structure and were so heavily involved in project work that they had little time or inclination to write papers. In any case, at the beginning there were no obvious outlets for publication in a topic that was seen as marginal to a large number of interests; Rhind's (1976) report, for instance, may well be the first example of a record of GIS conference papers which were described as such in a mainstream academic publication. While the advent of specialist GIS conferences (often disguised by use of other titles such as AUTOCARTO) provided one publishing mechanism from 1974 onwards, the early conference proceedings were intermittent and were not easily accessible to those who had not attended the gatherings. We do not believe this postulated paucity of recorded history represents incompetence on our part: a correspondence prompted by the editor of *Photogrammetric Engineering and Remote Sensing*, for example (Marble 1989; Tomlinson 1989), generated great controversy and revealed a lack of documentation on the first use of GIS in the refereed literature.

Finally and most crucially, the content of any history of GIS depends in large measure on the definition adopted. A strict definition, as a computer-based system for *analysing* spatially referenced data, would greatly restrict the field because, with the major exception of the Canada Geographic Information System (Tomlinson 1967), this was not a common feature until the 1980s. A more general interpretation, as any system for handling geographical data, would greatly widen the field and hence enlarge the number of contributors. Such a definition would embrace, not only the whole field of automation in cartography (which was often the precursor to any involvement in GIS and provided, in terms of computer-generated graphics, the most common form of output for most early systems), but also many general-purpose statistical and database packages capable of handling x,y,z point data. Formal definitions of GIS are not, therefore, of much help and relatively little reliance is placed on them in this book as a whole. In any event, the field evolved not from some *ex cathedra* definition of the subject but through sets of interactions. The main backgrounds of those involved have been cartography, computer science, geography, surveying, remote sensing, commercial data processing, mathematics and statistics. The purposes to which the systems have been put include environmental protection, urban and regional planning, land management, property ownership and taxation, resource management, the management of utilities, site location, military intelligence and tactics, and many others – as later chapters in this volume testify. The field has developed, then, from a melting pot of concepts, ideas, practice, terminology and prejudice brought together by people from many different backgrounds, interacting with each other often on a chance and bilateral basis in the early days and normally proceeding in blissful ignorance of what was going on elsewhere. The essence of GIS is thus its multidisciplinary character, with some at least of those involved in developing this technology having little previous involvement, or even interest, in the handling of geographical data as such (see Maguire, 1991 in this volume for further discussion of the definition of GIS).

This review of the history of GIS is inevitably a consequence of the authors' accidental exposure to early developments and their own set of value-judgements; different views certainly exist, such as that manifested in Cooke's portrayal of the genealogical structure of geoprocessing systems in general (Fig. 2.1). In particular, it is suspected that the role of those who did not contribute to the formal literature has been underplayed, especially those working in the military. While regrettable, this is probably unavoidable: history very often consists solely of what has been written down.

THE GRASS ROOTS EVOLUTION OF GIS

What seems clear is that there were many initiatives, usually occurring independently and often in ignorance of each other, concerned with different facets of the field and frequently originating in the interests, often disparate, of particular individuals. Like the reality (as opposed to the reporting) of scientific research, there was no strictly logical progression towards the development and implementation of GIS, but rather a mixture of failures, set-backs, diversions and successes. Inevitably, more is known about the successes than about the failures which, according to both Dangermond and Smith (1988) and Tomlinson (1988), have been numerous and often attributable to bad advice, ignorance and a determination to go it alone. This is unfortunate because failures are often as illuminating as successes, if not more so (Giles 1987). What also seems clear is that particular individuals and institutions played key roles, acting as examples or as sources of expertise, advice and often skilled personnel; since these contributions are now better recorded than is the generality of progress, this account will tend to emphasize them, particularly those of Howard Fisher in the Harvard Laboratory for Computer Graphics (LCG), Roger Tomlinson in the Canada Geographic Information System (CGIS) and Jack Dangermond in the Environmental Systems Research Institute (ESRI) in North America, and David P. Bickmore at the Experimental Cartography Unit (ECU) in the United Kingdom. Many others played significant parts (e.g. Tobler 1959; Nordbeck 1962; Cook 1966; Hagerstrand 1967; Diello, Kirk and Callander 1969 and Boyle (see Rhind 1988)), but these four have been the subject of particular articles in a special and invaluable issue of *The American Cartographer* (Tomlinson and Petchenik 1988). Fortunately, these individuals seem to typify the interests, attitudes and commitments of those working in the vintage era of GIS from the late 1950s to the end of the 1970s.

The motivations for developing GIS or components of such systems have varied very widely. They have ranged from academic curiosity or challenge when faced with the possibility of using new sources of data or techniques, through the desire for greater speed or efficiency in the conduct of operations on spatially referenced data, to the realization that desirable tasks could be undertaken in no other way. The last was undoubtedly a powerful motive in two key developments which are discussed in more detail below – the Oxford System of Automated Cartography and the Canada Geographic Information System. It was the experience of publishing the *Atlas of Great Britain and Northern Ireland* (Bickmore and Shaw 1963) and the criticisms this attracted of being out of date and unwieldy that convinced D. P. Bickmore, probably in 1958 but certainly no later than 1960, that only the computer could provide a cost-effective mechanism to check, edit and classify data, to model situations and to facilitate experiments in graphic display (Rhind 1988). Similarly, it was the impossibility of analysing maps of East Africa at an acceptable cost that first led R. Tomlinson (1988) to think of a digital approach. A calculation made in 1965 indicated the need for some $Can 8 million in 1965 prices and a requirement for 556 technicians for three years in order to overlay the 1 : 50 000 scale maps of the Canada Land Inventory; this unacceptable level of resources acted as an incentive to develop a more automated approach.

It was, of course, the advent of the digital computer and the order-of-magnitude decrease in computing costs every six years over a 30-year period (Simonett 1988) that made such alternative digitally based approaches viable. It is interesting to note, however, that not all early work used the digital computer. Thus perhaps the earliest attempt to automate map production, the preparation of the *Atlas of the British Flora*, employed a modified punch card tabulator to produce maps on pre-printed paper from cards on which had been punched the grid references of recorded occurrences (Perring and Walters 1962). Although this approach was not repeated and Perring (1964) later recognized that the analysis of voluminous data could more easily be undertaken by computer, it anticipated the widespread mapping in the late 1960s by line printer. It is also interesting to note that Perring was a botanist, with no training in cartography, who was faced with the task of providing 2000 maps from data that had been

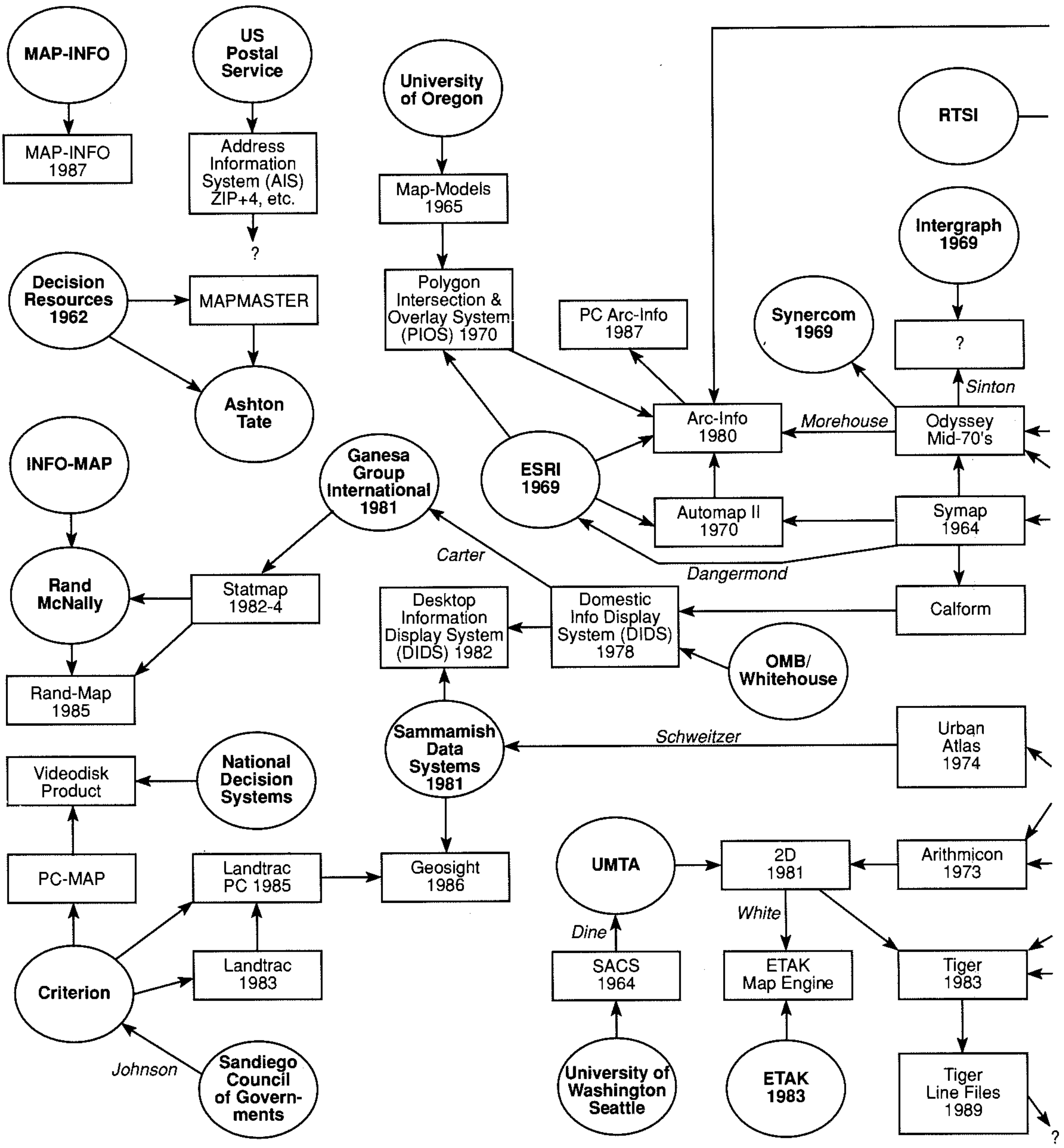
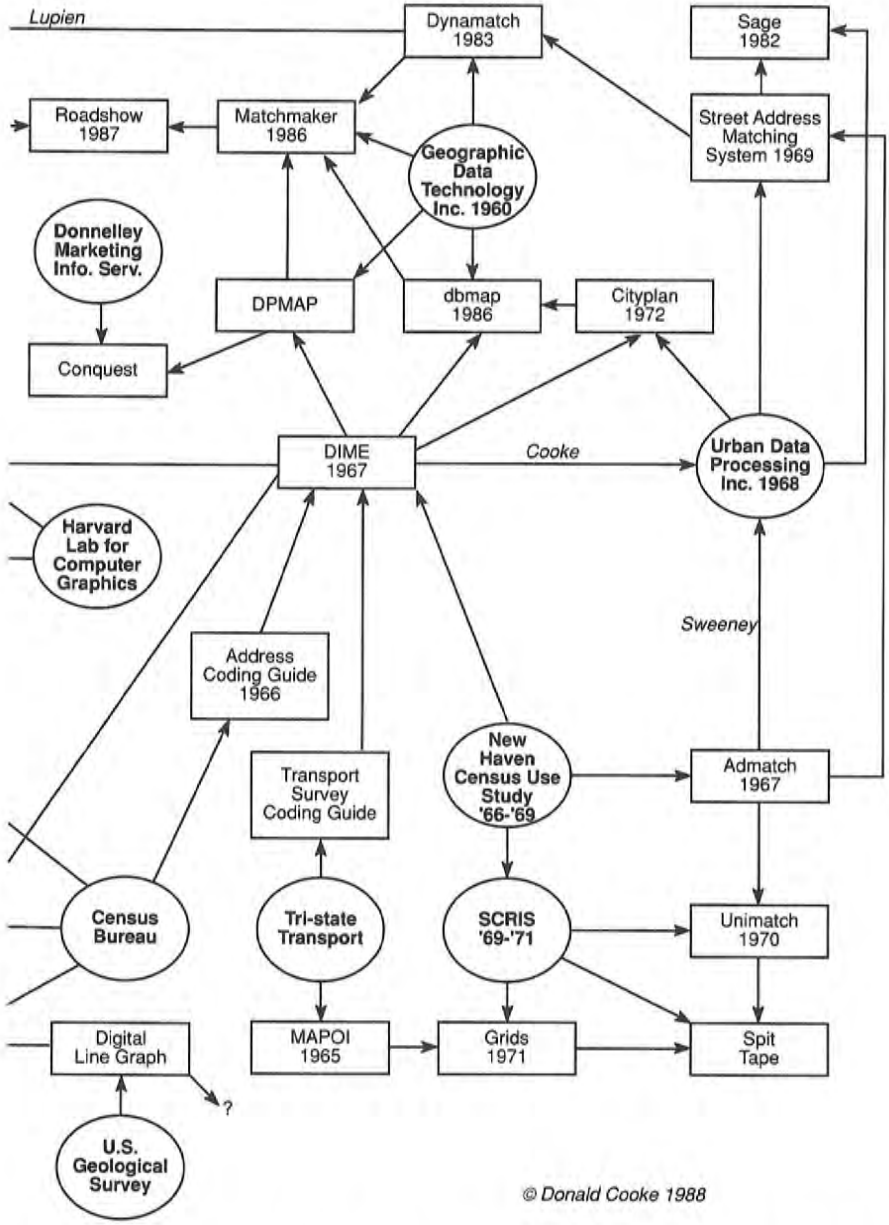


Fig. 2.1 An individual perception of the genealogy of geoprocessing in the United States (Pers. Comm. Don Cooke, 1990). Circles are 'places', i.e. companies, government agencies, universities, etc.; rectangles are ideas or concepts, often embodied in a software package or database; directed lines show direct or indirect migration or influence in a number of different ways. Examples of flows or lack of expected ones include:



© Donald Cooke 1988

- Harvard Labs influence on GIS vendors (Morehouse to ESRI, Sinton to Intergraph; Odyssey to Synercom)
- DIME was independent from the SACS (Small Area Census Studies)
- the diagram suggests that the USGS and the US Postal Service had very little influence on most developments.

recorded on punch cards. His initiative also illustrates an aspect to be repeated in many later projects where the application of technology was driven by an urgent need of the users, that such a task would have to take advantage of the best available technology – whatever its limitations – rather than await the ideal solution; it was also similar to many later applications in that it was a 'one-off' development which, having served its purpose, was not taken any further. Slightly later work (around 1967) by Bertin in Paris involved the modification of IBM 'golfball' typewriters driven directly by punch card readers to produce proportional symbol maps.

It is also clear that it was in North America that most of the significant early developments in, and applications of, GIS and related technology were made. By the early 1980s, Tomlinson (1985) estimated that there were probably more than 1000 systems in North America, a figure that must have represented a very high proportion of the systems then existing in the world as a whole. The bulk of this account will accordingly focus on North America, with later references to the United Kingdom and other European countries and to developments elsewhere in the developed world. It is only in the late 1980s that any significant developments have occurred in developing countries and then often through the aid and encouragement of developed countries (see Taylor 1991 in this volume).

THE NORTH AMERICAN SCENE

Aangeenbrug (pers. comm. 1990) has argued that the earliest antecedents of GIS in the United States can be traced back to the University of Washington. In the 1950s, both geographers (notably Garrison) and transportation engineers (notably Horwood) developed quantitative methods in transportation studies. Garrison's colleagues and students included Berry, Tobler and Marble; Horwood's included Barb and Dueker (see Dueker's important 1974 paper). Much of the original leadership of the Urban and Regional Information Systems Association (founded in 1963) and that of other key bodies was derived from or directly influenced by this group.

By the early 1960s, at least in North America, large mainframe computers were becoming widely available. In 1964, IBM introduced its 360/65 computer, with a processing speed 400 times faster and a memory 32 times as great as its predecessor, the IBM 1401 (Tomlinson 1985). These machines were employed primarily for one of two very different purposes: for routine administrative and data management tasks in business and government (such as pay-roll, stock control and record keeping of various kinds) and for scientific applications involving extensive computations, notably in chemistry, mathematics and physics. There was inevitably a good deal of discussion in government departments and agencies about the possibility of applying computer technology to handle numerical data, especially where these were already in machine-readable form, as with many censuses, where punch-card technology was widely used. In 1965 the US Bureau of the Budget compiled an inventory of automatic data processing in the Federal Government, in which it noted the significant use of computers to handle land use and land title data (Cook and Kennedy 1966). The following year, a conference on a comprehensive unified land system at the University of Cincinnati was advised that a system must be designed such that it obtained the maximum benefit from electronic data processing equipment (Cook 1966). The conference also heard that the District of Columbia already had a property data bank, which could be searched, updated and retrieved, and that Nassau County in New York would be the first to provide fully-automated access to records of land ownership.

The significance of the developments at the US Bureau of the Census, stemming directly from its need for automated address matching, is difficult to overemphasize. This need arose from the predominantly mail out/mail back nature of the US census and the requirement to produce area based tabulations from records whose only geographical reference was the postal address. An early advisory committee on small area data included Garrison (see above), who urged a development project to test automated data linkage procedures. A director hired to run the test, Caby Smith, recruited a team which included Corbett, Cooke, Maxfield, White, Farnsworth, Jaro, Broome and others who appear elsewhere in these pages. The first demonstrations of address matching, computer mapping and small area data analysis were provided through the 1967

New Haven Census Use Study (USBC 1969–73). Subsequent studies elsewhere in the United States, the launch of the DIME workshops in 1970 and the development and widespread distribution of ADMATCH (address matching software) all had major impacts upon government and academia in the United States. Indeed, the Census Use Study also sponsored the First International DIME Colloquium in 1972, leading to the creation of the Segment (later re-named as the Spatially) Orientated Referencing Systems Association (or SORSA), an organization which still holds international conferences.

Increasing availability of computers in universities was undoubtedly instrumental in the development of the quantitative revolution in academic geography in the early 1960s (James and Martin 1978; Hudson 1979), particularly in the field of spatial analysis (a term which was in general use by the late 1960s – see Berry and Marble 1968), with its emphasis on the statistical treatment of geographical data and on modelling. However, these applications, despite their potential relevance to handling geographical data, had little interaction with computer mapping, primarily because the statistical methodology was largely aspatial. One exception is a paper in an edited collection on computers in geography which related modelling to a crude cartography using the line printer (Rushton 1969). It is only in the middle and late 1980s that successful attempts have been made to develop closely coupled spatial statistics and 'geographical' displays.

Computers in the 1960s had, in general, no explicitly graphical facilities, usually operated in batch mode and were very expensive by today's standards. Despite this, Tobler (1959) had early recognized their potential for automating cartography, as had Nordbeck (1962) in Sweden. There were, indeed, developments in automating cartography in several national agencies concerned with mapping and in military establishments which could afford equipment that was prohibitively expensive to others. The US National Ocean Survey was creating charts on a Gerber plotter for the production of 'figure fields' or matrices of depth values and such organizations as the Aeronautical Charting and Information Center at St Louis, the Rome Air Development Center and the Central Intelligence Agency were active in aspects of this field (Diello, Kirk and Callender 1968; Tomlinson 1972). By the end of the 1960s, map production assisted by computer appears to have become widespread; for example, the Canadian Hydrographic Survey had automated display facilities in operation and Surveys and Mapping had embarked on a programme to apply automated cartography to the 1 : 50 000 series in Canada. In the main, however, the aim in computer applications in national mapping agencies was to mimic manual methods of production and so to produce maps that were virtually indistinguishable from their manual counterparts. Little information appears to be available on the extent to which these methods were cost effective, although Tomlinson (1985) suggests that the high cost of hardware placed them at a disadvantage in competition with manual systems: continuing evaluations of costs by the Ordnance Survey in Britain, for example, did not find automated approaches to map production as a whole to be cost effective until the 1980s. Unlike the situation in Britain, where a digitizing production line was in operation from 1973, the Topographic Division of the United States Geological Survey did not implement plans to automate the production of topographic maps until the start of the 1980s – a severe handicap to the development of many geographically-based information systems in the United States.

An entirely different approach to the automation of cartography was adopted elsewhere, notably in the universities, using the standard line printer as a mapping device. In cartographic terms, the results were crude, but this was not the point; the aim was to produce maps quickly and cheaply so as to display the characteristics of the data (especially statistical data for census tracts and the like) and to undertake simple analyses of such data by relating different parameters. It was here that Howard Fisher made a significant contribution and this approach found ready applications in landscape design, in urban and regional planning and, to a lesser extent, in resource management.

The Harvard Laboratory for Computer Graphics

Fisher was not a cartographer but trained and practised as an architect. He had begun work on a computer mapping system in 1963 while at the North Western Technical Institute (Schmidt and

Zafft 1975). On his retirement, he succeeded in obtaining a grant from the Ford Foundation to develop this work and, after making unsuccessful approaches to Chicago and Northwestern Universities (both strongholds of non-spatial computer applications to the analysis of geographical data), established the Laboratory for Computer Graphics (a title subsequently lengthened by the addition of 'and Spatial Analysis') in 1965 in the Graduate School of Design at Harvard University – from which he himself had graduated. There he built up a team of programmers and others to create a mapping package (SYMAP) which used the line printer as a mapping device and was capable of producing isoline, choropleth and proximal (Thiessen polygon or Dirichlet tessellation) maps. The package was easy to use by the standards of the day, particularly in relation to data for census tracts, incorporated default options when nothing was specified by users and was widely distributed. In addition to many pirated copies, over 500 institutions acquired SYMAP (Schmidt and Zafft 1975; Chrisman 1988); half of these were in universities, with the remainder equally divided between government agencies and private institutions. Copies were acquired not only in North America but also in Europe and elsewhere and the manual was translated into several languages, including Japanese. A subsequent program, CALFORM, which produced higher quality choropleth maps by pen plotter and reflected the increasing (if still sparse) availability of these plotters, seems to have had less success although it too was a pioneering effort. SYMAP was important as the first widely distributed computer package for handling geographical data. It introduced large numbers of users to the possibilities of computer mapping; it was the precursor, and possibly the progenitor, of a large number of other programs using the line printer; and it found a wide range of applications particularly through the connection between the Harvard Laboratory and landscape architects in the Graduate School of Design, notably C. Steinitz and his associates – one of whom, D. Sinton, produced a cell-based program (GRID) which permitted multiple overlays of data. Somewhat surprisingly, the appointment of a theoretical geographer, W. Warntz, to succeed Fisher as Professor of Theoretical Geography and Planning and head of the Laboratory in 1969, had little effect on the work and apparently stimulated little interaction between quantitative geography and computer mapping.

The Laboratory generated a wide range of contracts which, after the expiry of its grant from the Ford Foundation, became the main source of finance, along with income generated by the sale of mapping packages. It never developed a teaching programme (which might have prolonged its life) and thus only directly added a few new professionals to the field, although it did organize a highly significant symposium on topological data structures in 1977 and hosted influential Harvard Computer Graphics Weeks between 1978 and 1981. It also attracted at various times talented individuals who contributed in many ways to the development of computer mapping and, by extension, to geographical information systems. Among these are N. Chrisman, J. Dangermond, G. Dutton, S. Morehouse, T. Peucker and D. Sinton, several of whom contributed to the design and construction of ODYSSEY, arguably the prototype of contemporary vector GIS (Chrisman 1988). Unhappily, the subsequent history of this system was characterized by a series of unsuccessful marriages between the Laboratory and commercial enterprises and the departure of key staff from Harvard. As a consequence, numbers of staff declined and the Laboratory finally closed in the late 1980s. Overall, probably its most important contributions were in sparking creative thinking on GIS, creating a widespread awareness of the possibilities of handling and (to a lesser extent) analysing spatial data, and in stimulating programmes elsewhere which have contributed to the longer term development of GIS.

The Canada Geographic Information System

At about the same time as Fisher was developing his ideas on computer mapping at Harvard, R. Tomlinson (Tomlinson 1988) was involved in creating possibly the first true GIS – and certainly the first to be so entitled. Tomlinson can be thought of as the father of GIS through his role in persuading the Canadian Government that the creation of the Canada Geographic Information System (or CGIS, as it became known) in 1966, was a worthwhile investment. The origins of this, however, go back to 1960 when he was working for an air survey company, Spartan Air Services, which was undertaking a forest survey in East Africa. The firm had been asked to analyse all available map sources to identify locations for new plantations and for a new mill. The estimated costs of doing this manually were so high that the proposal was rejected. Tomlinson had argued that such analyses could be undertaken by computer and was given the opportunity to develop a digital methodology. None of the computer companies he approached was interested, although a subsequent chance encounter led to an expression of interest by IBM, which was already involved in digitizing air photographs. Another chance encounter on an internal flight found him sitting next to Lee Pratt, an administrator in the Department of Agriculture, which was then planning a Canada Land Inventory (CLI) involving the production of many maps of land capability for the whole of settled Canada; the analysis of these maps was expected to throw light on the agricultural rehabilitation of marginal farms. Tomlinson again expressed his belief that computer-based techniques would perform such analyses both faster and more cheaply. He clearly succeeded in impressing Pratt, whose subsequent support was critical to the development of the system. A contract was awarded to Spartan Air Services to undertake a feasibility study of a computer mapping system for the CLI. With the help of computer expertise from the staff of IBM, Tomlinson compiled a report which was accepted by the Department of Agriculture and he was then invited to direct its development within the Canadian Agricultural Rehabilitation and Development Administration (ARDA).

This development involved a large number of people both within ARDA and in IBM, and led to several significant developments for the future of GIS – among them, the creation of a drum scanner for the rapid digitization of maps (based on earlier IBM work of digitizing aerial photographs), of a data indexing scheme (the Morton Index 1966, which was subsequently widely emulated) and of a topological coding of boundaries involving the first known use of the link/node concept of encoding lines. The drum scanner, together with digitizing tables, provided the input to the system which was then based on an IBM 360/65 mainframe computer; output could be by line printer for numerical results and by ponderous Gerber plotter if a graphical output was required. It is interesting to note that there was minimal contact between CGIS and other bodies engaged in automated cartography and quantitative geography.

Tomlinson left the CGIS project in 1969, by which time Pratt had also left the Department of Agriculture. Although the various capabilities of the CGIS had been successfully demonstrated by this time, it was not until 1971 that the system was fully operational and subsequently it was reorganized and simplified. It now contains a digital archive of some 10 000 maps on more than 100 different topics. From the outside, it is difficult to evaluate the success of CGIS. Excluding those systems based on remote sensing data and the much more recent TIGER system (see Rhind 1991 in this volume), it may still be the largest GIS in operation and the only one to cover an area of continental extent in such detail; but its use seems to have been limited, in part no doubt because it took much time to build up the database as maps became available and because it came successively under four different departments, being given a different remit on each occasion. No doubt the facts that 'land' is a provincial responsibility in Canada and that CGIS was, for most of its existence, a passive organization, administered by technicians waiting for users to seek its services were also contributory factors. Its location in Ottawa, the lack of computer networking at the time and the prior availability of easily distributed printed maps of land capability – which users elsewhere in Canada may have regarded as more accessible – were other possible factors in its limited operational (as opposed to technical) success.

Early government activities in the United States

Tomlinson departed in 1969 to become a private consultant within the GIS field, one of an initially rare but increasingly necessary breed as proprietary systems multiplied and salespeople sought to persuade clients of the desirability of their systems. In addition, he continued to play an important role as Chairman, for the first 12 years of its existence, of the International Geographical Union's Commission on Geographical Data Sensing and Processing, which had been established in 1968 and which sponsored two major international conferences in Ottawa in 1970 and 1972, the first such conferences to be specifically identified with GIS. This emphasis helped to give currency to the name which the CGIS had pioneered, to promote contacts between delegates from a wide range of disciplines and locations, and to provide an overview of developments in the early 1970s (Tomlinson 1970, 1972). Subsequently, the Commission undertook an evaluation of the handling of digital spatial data within the US Geological Survey which, in 1976, had more than 15 information system activities concerned with the gathering and handling of spatial information in the fields of geology, geography, topography and water resources. The Commission also published, under the auspices of the UNESCO Natural Resources Research Series, a major monograph on the computer handling of geographical data (Tomlinson, Calkins and Marble, 1976).

These investigations also revealed that by 1976 there were at least 285 items of computer software handling spatial data which had been developed outside the USGS; by 1980, when a revision of the findings of this project was published, the number had risen to over 500 (Marble 1980). These figures are one measure of the rapid progress in this field in the late 1960s and the 1970s. They also illustrate one feature of that development which the CGIS had earlier demonstrated, that there was relatively little contact between the developers of such software. As a result, there was also considerable duplication, with programs being developed independently by different agencies to perform the same function. In part this was a consequence of a growing awareness of the possibilities provided by computers for handling spatial data and for displaying results of analyses through automated cartography. Within the universities, too, individuals were writing programs, often to fulfil a contractual obligation with a locally based agency that lacked the expertise to do so itself.

Tomlinson (1988) described the 1970s as a period of lateral diffusion rather than of innovation and there is considerable piecemeal evidence to show an increasing interest among a variety of agencies at all levels of government – federal, state, county and city. These include military and security establishments (such as the Central Intelligence Agency, which developed a world data bank for its own purposes and then made it available in the public domain (see Anderson, Angel and Gurney 1978); land management agencies such as the US Forest Service (Shumway 1986); conservation bodies such as the Fish and Wildlife Service (Christianson 1986); the Department of Housing and Urban Development (Goldstein, Wertz and Sweet 1969); the Bureau of the Census (see earlier description and USBC 1969–73)) and others at federal level; states such as California, Maryland, Minnesota, New York and Oregon; counties such as Fairfax in Virginia (Lay 1975) and cities such as Kansas City and Oakland. Some took existing software, as in the application of SYMAP in the Oakland Planning Information System (Goldstein, Wertz and Sweet 1969); others developed their own software in-house.

No comprehensive record exists of these various, local approaches but those adopted in the Bureau of the Census and in the USGS may be taken to represent the federal level. Initially, they developed quite separately, but have partially come together in the late 1980s to link digital cover from the USGS 1 : 100 000 topographic map series with census tracts for the 1990 population census, developments that will provide the basis for a variety of GIS initiatives throughout the United States (Callahan and Broome 1984).

As already indicated, the Bureau's substantial involvement in geographical data processing began with the New Haven Census Use Study in 1967 (USBC 1969–73) which led to the Dual Independent Map Encoding (DIME) scheme as the standard method for encoding data for census areas and (later) the preparation of experimental computer-generated maps of census data (Schweitzer 1973). The essence of DIME was a method of describing the urban structure through recording the topological relationships of streets; the earliest DIME files contained no coordinates. The advantage of proceeding thus was to provide an automated method of checking the completeness of areas built up from street boundaries – of particular importance since the US Census is substantially a mail out/mail back operation and the descriptions of the geography were assembled in many Census offices across the country (Dewdney and Rhind 1986). During 1972, the Bureau decided to embark on the creation of atlases for the major metropolitan areas (Schweitzer 1973), a task with which the Harvard Laboratory became involved in 1975 (Chrisman 1988). The Urban Atlas Project required the digitizing of maps of some 35 000 census tracts in the metropolitan areas and demonstrated the cost effectiveness of such an approach. It also required the development of software for handling this large quantity of data. Particularly significant in all these developments in the Bureau was a small group of mathematicians led by James P. Corbett and including Marvin White. Although it appeared subsequent to the Loomis (1965) paper, Corbett's definitive paper on the topological principles underlying cartography and GIS appeared in 1975 and a readily obtainable version of it was later published by the Bureau itself (Corbett 1979); it is clear that much of the credit for defining how topology theory is applied in the field of GIS is due to this group and to others working in applications areas in the Census at the time, such as Don Cooke. From this beginning came the subsequent extensions to DIME, the development of ARITHMICON and, ultimately, the creation of TIGER (see Rhind 1991 in this volume), possibly the largest and most all-embracing civilian GIS project yet and on which the success of the 1990 US Census critically hinged.

The practical involvement of the USGS in the GIS field is exemplified by the development of a system to handle and analyse land resource data, the Geographical Information Retrieval and Analysis System (GIRAS), developed from 1973 onwards to handle the increasingly large data sets becoming available (Mitchell *et al.* 1977). This was developed specifically to handle information on land use and land cover, held on manually produced maps at a scale of 1 : 250 000 derived from aerial photography, although these are subsequently derived/updated directly from imagery without manual intervention, using data from the Landsat series of satellites. Such maps had first to be digitized in polygon format to provide the input to the system which had been designed to store, manipulate and analyse these data, together with others on political and administrative subdivisions and public land ownership. GIRAS was developed initially as a batch processing system (GIRAS 1) but an interactive version was subsequently developed (GIRAS 2), with access via remote terminals. Output could be produced either in statistical or graphical form, the latter through display on a CRT screen or as a plot from a Calcomp drum plotter. The system also had a capacity to convert polygonal data into gridded data, an attribute of increasing importance with the use of remotely sensed data. The USGS has since made a myriad of other contributions to the development of GIS, not least in converting many data from paper map or stereo photo to digital form. Starr and Anderson (1991 in this volume) has summarized its recent activities and future plans.

State systems may be exemplified by the Minnesota Land Management Information System (MLMIS) which appears to be one of the more successful of such systems and illustrates the transition from a university research facility to an operational system within a state agency. It is described in more detail by Robinette (1991 in this volume) but it illustrates the difficulty of finding a secure financial base for such a system and the weaknesses deriving from an early decision to collect data on a rather coarse grid. MLMIS was started in 1976 as a research project located in the Center for Urban and Regional Analysis in the University of Minnesota, where the emphasis was on pilot projects but where some limited production work was carried out (which proved to be unsuited to a university environment). The system was based upon a digital land use map of the state, prepared from aerial photography. It was subsequently taken over by the state and established as a service bureau within the state planning agency, where it operated on a 'fee for service' basis, an approach which nevertheless required that the system operation and management be subsidized. It appears that very few users were at that time willing to pay for database development and the service found that it had to take on an increasing number of projects that were marginal to its main purpose in order to remain viable. This requirement to pay its way also led to raised fees and a consequent reduction in use. Nevertheless, it has undertaken several hundred successful GIS projects during its lifetime.

The commercial sector: the example of ESRI

MLMIS is interesting as one of the systems that had a continuing existence. Many others came into being, often created by university groups under contract to local or national agencies, and subsequently disappeared for lack of funding or because a key member of the team left. Little is known of the many equivalent developments in the commercial sector although the later history of the Harvard Laboratory provides one illustration of a commercial system that failed to get off the ground. After reaching a low point in the 1970s, following the exhaustion of the original Ford grant and the withdrawal of support by the Harvard landscape architects, the Laboratory had grown again, developing software and applications. One of its central activities was the development of the vector-based GIS system, ODYSSEY. A working version of this system was in operation by 1979 and a 'hazy deal' was struck with ISSCO, a software firm involved in computer graphics, to market it. The firm advertised the software but then withdrew, leaving the Laboratory with heavy debts which left it unable to recover as a major innovator in this field.

A happier example, also with roots in the Laboratory, is represented by the success of the Environmental Systems Research Institute (ESRI), founded in 1969 by J. Dangermond, a landscape architect who had gone to Harvard in 1968 to complete a master's degree and had then returned to his native California. ESRI was not, of course, the only firm operating in this field. Intergraph (again involving a product of the Harvard Laboratory, D. Sinton, and led by James Meadlock), ComputerVision and Synercom were other major players even in the 1970s. Most of these – apart from ESRI – came into GIS from the CAD/CAM area. But, in the light of published knowledge and because it is a highly successful enterprise, ESRI must serve as an exemplar for them all.

ESRI began as a non-profit organization engaged in the field of environment consultancy, although a brochure published in 1970 identified computer graphics as one of the professional services provided (Dangermond and Smith 1988). It used and developed the cell-based package GRID as its main applications package until the launch of ARC/INFO in 1982, and also developed a three-dimensional version called GRID TOPO; in the mid/late 1970s, it developed and sold a vector-based system, the Planning Information Overlay System (PIOS). A few years after its launch, it became clear that ESRI would not succeed in raising the necessary finance for growth as a non-profit organization and it consequently became a with-profit enterprise. The firm initially used the University of California mainframe computer but, with falling costs of hardware and increasing computer use, found it more convenient and cost effective to acquire its own minicomputer. By the mid-1970s it was also advertising its competence in GIS and by the early 1980s was providing a turnkey GIS. This proved very popular and a large and growing number of such systems have been installed. ESRI's ability to make its ARC/INFO system function across computer platforms ranging from personal computers, through workstations and minicomputers up to the largest mainframes has clearly been beneficial to the company.

Initially, most of ESRI's project work was on relatively small applications, relating to site or location analyses, but it became increasingly involved in environmental questions, reflecting the growing recognition of environmental problems in the United States. In 1973, it began work on its first state-wide system designed for mapping environmental suitability, the Maryland Automatic Geographic Information (MAGI) system, which became a model for other state systems. It had earlier participated in several applications in town planning in the United States, Australia, Canada, France, Japan and Venezuela. Other projects were undertaken in wastewater management, biological conservation, land reclamation, floodplain management, recreational planning and other topics.

Throughout the 1970s and early 1980s, staff of ESRI undertook a great deal of the project work themselves in the absence of appropriate expertise in commissioning agencies. It is unclear how far this widening range and increasing number of applications was due to a growing awareness of ESRI's capabilities and its own efforts to make these known, and how far it was due to an increasing need by potential clients to find efficient ways of handling large quantities of data. Dangermond and Smith (1988) have suggested that, in the 1970s and early 1980s, it was a matter of pressing GIS solutions on unaware and unwilling potential users, involving constant selling and subsequent support. Nevertheless, the fact that ESRI staff were heavily involved in the projects meant that they identified any flaws in their own software at an early stage and had a strong project-oriented approach, a fact that helped to build confidence in the firm. In contrast to the 'selling job' of the 1970s, the 1980s were characterized by an increasing and accelerating trend towards acceptance of GIS, with increasing numbers of requests for information and advice (Dangermond and Smith 1988). In the circumstances where this has been the case for over a decade, users can undertake projects with little outside advice or help although ESRI seeks, through its ARCNEWS and user conferences in different parts of the world, to provide continuing support. Of all GIS vendors, ESRI has probably been the most successful in the 1980s; much of its success can be attributed to ARC/INFO. Many other factors played a role, including the personality of the ESRI founder and the forging of close links with users in education and other sectors. By the end of the 1980s, more than 2000 systems of GIS software for use on personal computers were being sold each year and ESRI had expanded from a staff of 15 in the early 1970s (though already operating throughout the United States and overseas) to one with over 350 staff and operating in a global market.

Spreading the word

Two other aspects are worthy of note, the development of teaching (initially in computer cartography and then, in the 1980s, in GIS itself) and the growing communication between workers in these fields. Many of those who had attended the early SYMAP conferences at Harvard began to develop teaching applications and, as early as 1972, a monograph on computer cartography (on which a whole chapter was devoted to SYMAP) was available in the Association of American Geographers' Resource Papers Series (Peucker 1972). Among those developing competence in this field were the University of New York at Buffalo, Simon Fraser University and the University of Saskatchewan in Canada, although Tomlinson (1988) has argued that in the 1970s there was probably more on-the-job training in commercial and government agencies than in universities.

The roles of conferences and publications in this field have already been noted. One of the first groups to publish a newsletter was the Urban and Regional Information Systems Association (URISA), founded in 1963 and holding annual meetings thereafter. Like many maturing organizations, it has eventually found the need to create its own journal and the founding issue in 1989 was largely devoted to GIS. The roles of the Harvard Graphics Weeks and the two Ottawa conferences under the auspices of the IGU Commission have already been noted. The most significant other development was the AUTOCARTO series begun (although not under that name initially) in 1974 as an International Symposium on Computer-Assisted Cartography held at Reston, Virginia (Chrisman 1988). It is interesting to note that, whereas some 40 people attended the first Ottawa conference in 1970, 300 attended the second and some 500 the first AUTOCARTO meeting (Tomlinson 1970, 1974; Chrisman 1988). Such meetings have been held at biennial or shorter intervals since that time. Such conferences became a fruitful outlet for publications by those involved in this field, although other papers were being prepared in a wide variety of professional journals (notably in the *International Journal of Geographical Information Systems*, co-edited in the United Kingdom and United States). Of course, developments in related fields must not be forgotten. Relevant papers appeared in the conferences of the Association for Computer Machinery, in publications devoted to computer-aided design, and in computing and engineering journals, especially those of the latter related to the utility companies.

By the late 1980s, then, GIS can be said to have become widely accepted in North America. The numbers of systems, courses, conferences, projects and facilities continue to multiply. Central, regional, state and local governments are increasingly involved, as are those in retailing and service delivery (see Beaumont 1991 in this volume) and in asset management (see Mahoney 1991 in this volume). The field has also acquired a degree of scientific recognition in the establishment of a National Center for Geographic Information and Analysis (NCGIA), funded by the National Science Foundation, as a cooperative venture between the Universities of California, Maine and New York (Abler 1987; see also Morrison 1991 in this volume). What is particularly interesting to an outsider is the speed with which acceptance of GIS has accelerated, to a stage where GIS is now a 'buzz word', and the extent to which the development has largely happened outside the political process, at least at a federal level. There seems, at a national level, to have been no official declaration of policy that this is a desirable path to follow. Initiatives and investments seem largely to have been effected through the bureaucratic system, although the USGS has been recognized by other federal agencies as having a coordinating role in this field at the federal level. As far as can be seen, the decision of the USGS and the US Bureau of the Census to develop what became the TIGER files and the associated line graphs for the whole of the United States was largely an internal, inter-agency agreement, unprompted by wider political considerations.

THE BRITISH EXPERIENCE

The British experience in the GIS field is in marked contrast in this regard, although there are some similarities – notably in the influential role of key personalities at an early stage in the development of GIS. Perhaps because there is little publicly owned land in the United Kingdom (and hence little direct responsibility for land management) and because the small size of the country and the highly centralized government have made for limited appreciation of spatial contrasts, the official pressures to develop computer-based methods of handling spatial data have been more limited. Moreover, local government, where such systems might usefully be applied, has experienced increasing financial constraints and a reduction in several of its key functions throughout the late 1970s and 1980s. These differences, have, however, been turned to advantage.

As in North America, the first beginnings can be seen in the 1960s in a number of areas, in the national mapping agency, in local planning and in the universities and polytechnics. A proto-GIS was proposed to the short-lived Natural Resources Advisory Committee in 1964 but, with the demise of that body, no further action was taken (Coppock 1988). A growing perception that computers provided a potential aid to land-use planning led in the late 1960s to the formation of joint working parties of officials from both central and local government and to the publication in 1972 of a report entitled General Information Systems for Planning (or GISP). This report (DoE 1972) outlined an approach that local authorities might follow, although there was never much encouragement by central government to do so nor any indication that it would adopt such an approach in its own agencies.

Bickmore, the ECU and the Ordnance Survey

Again, as in North America, a driving force in the automation of cartography was a research group, established almost entirely through the persistence and persuasiveness of one individual, D. P. Bickmore (Rhind 1988). He had been the cartographic editor of the Clarendon Press and one of the two editors of what was in effect a national atlas (Bickmore and Shaw 1963). This experience led him to the conclusion that the computer offered the only possibility of undertaking such a project reasonably expeditiously with up-to-date data. In collaboration with R. Boyle, later to be a major player in the development of automated cartography in North America, he secured funds from the press to develop the Oxford System of Automated Cartography and, although this led to the manufacture of the world's first free-cursor digitizer and possibly the first map-making using a photohead on a high precision plotting table, no complete system was ever produced. However, Bickmore was successful in persuading the newly formed Natural Environmental Research Council (NERC) to fund a research unit in automated cartography. The Experimental Cartography Unit (ECU) became fully operational in 1967–68 at the Royal College of Art in London and focused its attention initially on the computer-assisted production of high-quality printed maps (see, for instance, Experimental Cartography Unit 1971; Rhind 1971). From 1973 onwards, it largely concerned itself with GIS issues. In 1975, the ECU was absorbed into the Natural Environment Research Council headquarters at Swindon and later became the NERC Unit for Thematic Information Systems at Reading University.

Like the Harvard Laboratory in its later years, the work of the Unit was highly project based, in collaboration with the Ordnance Survey (the national mapping agency), national agencies for geology, soils and oceanography, and planning agencies. Software was developed for changing projections, editing, data compression, automated contouring and so on, and experimental maps produced for the cooperating agencies. No marketable software analogous to SYMAP or other Harvard Laboratory programs was developed although plans were laid in 1971–72 to market the Unit's interactive editing software which ran on a DEC PDP 15 computer. But like the Laboratory, it exercised an important influence on thinking about automation and Rhind (1988) has argued that, without the Unit, the Ordnance Survey (OS) would not have begun its investigations into digital mapping or the production of digital maps until several years later at least.

The path initially followed by the ECU was not directly concerned with GIS. Nor, indeed, did the OS involvement in automated map making at first have any direct bearing on GIS. The approach adopted by the OS from 1971 onwards (see Sowton 1991 in this volume; Finch 1987) was to simulate manual production of its maps as closely as possible, an approach which facilitated map production but made it impossible to use the resulting digital data in an information system because they are unstructured. By 1973, the OS had established a production line to digitize and plot automatically versions of its large-scale plans as they came up for revision, a process that was necessarily random in its spatial occurrence. This approach did not become cost effective until the late 1980s, but it gradually came to be appreciated (both within the OS and outside) that a digital map framework could provide the base through which digital records of soil, geology, vegetation, assets and plant and the like could be related. An attempt was, therefore, made as early as 1974–75 to restructure the digital data through the development of appropriate software. The approach to preparing such a digital base through the process of map revision was not only random but also slow and the OS has come under increasing pressure to accelerate production of digital maps, at least for the bulk of the populated areas (which now seems likely to be completed not later than 1995). If this is achieved, OS will have well in excess of 200 000 maps in computer form – a total unmatched anywhere else in the world. In the British situation, where a consistent and large-scale series of maps exists, compiled to a common standard and datum and updated continuously, this coverage provides a highly valuable spatial framework on which other data can be assembled and linked (see Sowton 1991 in this volume).

Finally, it may not be appreciated by readers outside the United Kingdom that there are two Ordnance Survey departments within the nation state: one for Great Britain and the other for Northern Ireland. The latter became committed to a database approach rather than a map reproduction one earlier than did the former (see, for instance, Brand 1986). The Northern Ireland OS operates in a quite different – and generally more favourable – government and financial environment from that responsible for mapping the rest of the United Kingdom, but its task is much smaller and less onerous.

Other British developments of GIS

A variety of other historical developments has influenced the current GIS scene in the United Kingdom. For example, research staff in the Department of the Environment (DoE) had developed a mapping system for use in the Department and in central government as early as 1969 (Gaits 1969), but it seems to have had only limited application there, presumably because its usefulness was not recognized by administrators and planners. The Scottish Development Department, the planning ministry in Scotland, did fund a pilot rural information system, the Rural Land Use Information System (RLUIS), involving a large number of national agencies and two local authority districts in Fife, but this work was discontinued when the originating body, the Standing Committee on Rural Land Use, was abolished (Lyall 1980). Probably the greatest interest in GIS was expressed by the agencies concerned with environmental matters but these were small, with limited budgets, and were largely dependent on data collected by others, so that until the 1980s only limited progress was made. In Scotland, however, a consortium of three agencies (the Countryside Commission for Scotland, the Forestry Commission and the Scottish Tourist Board) did commission a university research unit, the Tourism and Recreation Research Unit, to construct and operate a simple GIS, the Tourism and Recreation Information Package (TRIP), to assist them in planning and policy making (Duffield and Coppock 1975).

Initiatives taken within planning authorities in local government to establish information systems following the 1972 GISP Report varied greatly, although in total they can only be described as disappointing until the late 1980s (Rhind 1987). The National Gazetteer Pilot Study in the Metropolitan County of Tyne and Wear was the first of a number of property-based systems, probably taken furthest in the now-abolished Greater London Council and Merseyside Metropolitan County. The experience of several such systems, with particular reference to Berkshire, Nottinghamshire and Warwickshire, has been reviewed by Grimshaw (1988), who suggests that, probably for a variety of institutional reasons, the use of such systems has declined. Certainly the encouragement offered by the GISP Report led to only limited development, in part because of the determination and ability of central government to restrain expenditure by local authorities. There appears to have been an increasing interest in the use of automated map production in planning departments, initially by line printer maps using programs such as SYMAP; a study sponsored by the DoE and the Scottish Development Department to develop a proto-information system for planning in Scotland led instead to guidance through workshops and consultancies on the ways in which such mapping could assist the planning process (Coppock and Barritt 1978). There was, too, an increasing attention to management information systems in local government, as exemplified by the innovative Local Authority Management Information System (LAMIS) developed by International Computers Ltd (Rhind and Hudson 1980), but this appears to have had relatively little impact on GIS. The potential significance of applications of computer-based handling of data was, however, increasingly recognized. The Royal Town Planning Institute established a committee on the topic in the early 1970s, a British equivalent of the Urban and Regional Information Systems Association (BURISA) was established in 1970 and the National Computing Centre attempted to promote interest in this field. Nevertheless, the practical results of such interests were small, although these developments no doubt assisted the explosion of interest which occurred in the late 1980s.

As in North America, initiatives by central and local government were sometimes undertaken through the agency of university researchers, with the Universities of Durham, Edinburgh and London among the main foci of such work. The staff of such universities were also involved in a number of initiatives in computer mapping; in particular, T. C. Waugh (1980) developed his widely used GIMMS package at Edinburgh from 1970 onwards. Other initiatives were undertaken by researchers in the Institute of Terrestrial Ecology (ITE), where national databases of environmental and land use data were developed (Brown and Norris 1988).

Despite the late start, the single most important influence on GIS in the United Kingdom in recent years has probably been the growth in use by utilities of digital mapping data in conjunction with records of their own plant. Indeed, they are now the main driving force behind the acceleration of the digitizing of the coverage of OS digital maps. Since this is very recent, currently involves no major methodological advance in GIS and is described by Sowton (1991) and by Mahoney (1991) in this volume, no more will be said here except to stress its importance in providing data which are also useful for other purposes.

Official inquiries relating to GIS

A distinguishing characteristic of the United Kingdom in the decade between 1978 and 1987 is the series of inquiries carried out by the British Government or Parliament into geographical data and its use (see Chorley and Buxton 1991 in this volume). These demonstrate a shift in emphasis in successive inquiries. The first of these was the Ordnance Survey Review Committee, appointed in 1978 to advise on long-term policy for the OS. It paid considerable attention to the adoption of digital mapping, recommending that – if it was proved (as the Committee expected it would be) that digital mapping was cost effective – the OS should accelerate conversion from manual methods to provide digital coverage at the 1 : 50 000 scale by 1982–83 and at the basic (i.e. primary mapping) scales (of 1 : 1250, 1 : 2500 and 1 : 10 000) by 1992–93; it made no reference to GIS as such. This report, then, considered OS very much in terms of its role as a traditional map producer, but producing the paper maps in future by digital means.

The Report of the House of Lords Select Committee on Science and Technology (House of Lords 1984; see also Rhind 1986) investigated both digital mapping and remote sensing. It was much more appreciative of the potential benefits of digital data as such and, among its many recommendations were several on the need to accelerate the OS digitizing programme and the manner in which this should be done. In addition, it recognized the pervasive nature of geographical data and the many interrelationships and dependencies being built up, principally with regard to OS data. Therefore, it recommended the establishment of a Committee of Enquiry into the handling of geographical data. Such a Committee was established under the chairmanship of Lord Chorley and found widespread interest in such data; a large part of its

Report (DoE 1987; Rhind and Mounsey 1989; Chorley and Buxton 1991 in this volume) was devoted to the role of the OS in providing a digital map framework. In addition, however, it strongly emphasized the role of GIS, examined the value of standards for spatial referencing, noted that the lack of 'awareness' of what is possible and what is already going on in other areas was a major impediment to the widespread use of GIS, and recommended the establishment of a centre for coordination and advice on the handling of geographic data. Some of the recommendations were acted upon speedily: thus considerable resources have since been put into GIS research in the United Kingdom by the Natural Environment and the Economic and Social Research Councils (Goodchild and Rhind 1990). More significantly, there has been a move away from maps as products in themselves towards the view that they are only one of many sources providing geographical data for use in GIS. In essence, a database approach is now firmly established in British government thinking and, because of the highly centralized nature of the state in comparison to that of the United States and the existence of a ubiquitously used and country wide spatial referencing framework (the National Grid), this approach may be expected to permeate most organizations quite quickly.

Yet, despite this initiative to study what should be done in the national interest, the commitment by government departments to making geographical data generally available has been less than had been hoped. This is significant because data are not 'in the public domain' in the United Kingdom; government sees itself as entirely justified in charging for its data and, indeed, OS receives over one-third of its budget from cost recovery – a figure which will certainly rise (Rhind 1990). The military have marketed (through OS) Digital Elevation Models at 50 m resolution derived from the OS 1 : 50 000 scale contours and the Census offices have agreed to code each 1991 Census household response with a postcode (1.3 million of which exist for the country, so a multiplicity of small area data sets could be constructed for particular purposes). A significant factor is that income generated by sales of data by government departments does not generally accrue to them but reverts to the Treasury; since no part of their budget is provided for making data available and since to do so demands resources, little incentive exists for data to be provided. However, recent relaxations in rules as some departments are converted into executive agencies may alter the situation.

If activity in the United Kingdom in GIS was disappointingly small in the 1970s and early 1980s after the pioneer work of the ECU and OS, it expanded greatly in the late 1980s, stimulated by the publication of the Chorley Report. As already pointed out, much of the funding to stimulate action has come from the utilities, which need large-scale digital databases for the efficient management of the networks they control (NJUG 1987); their pressure has been influential in simplifying the specification for digitizing OS maps and in accelerating that programme. Local authorities, too, are showing a keener interest and, along with the utilities, have been major actors in the establishment in 1989 of the Association for Geographic Information (AGI), a national umbrella organization which involves vendors, chartered surveyors, geographers, educationalists, users in commerce and industry, software houses, learned societies and many others: its primary roles are to ensure the dissemination of knowledge, promote standards and advance the field. Commercial interest, which had been notably lacking in the 1970s also rapidly developed, particularly in relation to market research (see Beaumont 1991 in this volume), plant location, local planning and traffic guidance systems, and is exemplified by the formation of Pinpoint Analysis Ltd which, in collaboration with the OS, has prepared a national database of centroid references of all properties in Britain (Rhind 1988). Finally, an exceptional contribution has been the creation of GIS demonstrators (Green 1987) and tutors (Raper and Green 1989) which have served to instruct and inform a world-wide audience.

DEVELOPMENTS ELSEWHERE IN THE WORLD

It is obvious that developments elsewhere cannot be covered in the same detail as for North America and the United Kingdom. None the less, several important distinctions need to be made. Developments elsewhere in Europe appear, apart from any military interest, to be associated primarily with national mapping agencies and with the maintenance of cadastral records of property. In respect of the latter, one of the most interesting developments has been that of the Swedish Land Databank System (SLDS), initiated in the early 1970s by a decision of the Swedish Parliament to replace the earlier manual system of property and land registration by an electronic system. An agency, the Central Bureau for Real Estate Data (CBRED), was created to establish and maintain this system. A pilot system was introduced in Uppsala County and operated under legal authority from 1976, and the system was progressively implemented, half the country having been covered by 1986 (Andersson 1987). It had a much wider application than solely in the field of property (see Dale 1991 in this volume; also Ottoson and Rystedt 1991 in this volume) and is being combined with statistics on housing and population (routinely available in Sweden on the basis of individual properties) as an input to urban and regional planning and to throw light on policy issues of various kinds, such as nuclear emergencies, civil defence and second homes.

Many applications elsewhere in the world are more recent than those in the United States. In Australia, initiatives have come primarily from two sources, cadastral mapping (a state responsibility) and applied scientific research (O'Callaghan and Garner 1991 in this volume). The handling of land records is now well established in all Australian states, having begun very early in South Australia, and is indeed internationally regarded as a major Australian success. GIS in the science area has been developed mostly through the Commonwealth Scientific and Industrial Research Organization (Cocks, Walker and Parvey 1988), where a continental-scale GIS, the Australia Resources Information System (ARIS), was begun in the late 1970s. It arose from an initial need seen in the mid-1970s for the production of maps of local government data and was stimulated by the availability of digital files of local government boundaries from the Division of National Mapping. By 1982, a wide variety of natural and socio-economic data were available. This system is essentially a research tool which has been used in a number of applications, from the location of new cities, through the identification of areas that should be withdrawn from pastoral use and of other ungrazed areas that might be devoted to this purpose, to the representativeness of National Parks. It remains a prototype, produced on a limited budget, which might serve as a model for an operational system at some future date.

Despite its immense commitment to the electronics industry, Japan showed little interest in the 1960s and 1970s in GIS, although here too interest in GIS developed very rapidly in the 1980s (Kubo 1987; see also Kubo 1991 in this volume). The only related activity in the 1970s appears to have been the production of digital land data by the Geographical Survey Institute using a 1 km grid; but no indigenous software was available to process the data and their use was accordingly very limited. The recent surge of interest has apparently been encouraged by the commercial survey and computing industries, which are very keen to acquire new business, and by central government agencies wishing to extend their control.

It is unclear from the account by Koshkariov, Tikunov and Trofimov (1989) at what stage interest in GIS occurred in the Soviet Union, although the immense challenges which that country faces in the management of natural resources throughout its vast territory seem to cry out for use of such systems; the impression is that not much had happened before 1980. By 1983, the Institute of Geography in the USSR Academy of Sciences was holding a conference on the problems of GIS science and, in the same year, the first of a series of schools and seminars for young scientists was held in the Far Eastern Research Centre on cartographic modelling and GIS. The main emphasis generally seems to have been on developing systems of automated mapping and on the preparation of cartographic databases, rather than in GIS *per se*. In this regard, the Soviets seem to be following the same path undertaken by the United States and the United Kingdom a decade or more earlier.

The history of GIS in developing countries is, in the main, similarly restricted to the 1980s when, partly through the initiatives of aid groups, a number of systems was established (see Taylor 1991 in this volume). For example, the Jamaica GIS (JAMGIS) was begun in 1981, funded by the US Agency for International Development using Michigan State University's Comprehensive Resource Inventory and Evaluation System (CRIES) as the basis (Eyre 1989) and a Land Titling Project was begun in Thailand in 1985 through a World Bank loan and technical assistance from the

Australian International Development Assistance Bureau (Angus-Leppan 1989). The latter has highlighted the importance of the staff of the host country being intimately involved in the work, of using familiar concepts and terms wherever possible, of providing comprehensive training and education, and of the role of persuasion by adequately briefed advisers.

Not all developments in such countries, however, are dependent on outside support. In the People's Republic of China, where work on digital mapping had begun in 1972 and tapes of satellite imagery had been acquired by 1975, a conference at the Academia Sinica in 1980 led to the establishment of a working group on GIS and to a number of regional initiatives. Chen Shupeng (1987) has summarized the numerous developments since then, mostly related to environmental hazard prediction and management, and carried out on microcomputers.

For completeness, reference should be made to the attempts at multinational and global GIS. Of course, the problems of developing and implementing GIS across national boundaries reflect, in exaggerated form, experiences within each country (notoriously those within the United States) of a lack of correspondence between mapping systems, data collection and the like. The European Commission (EC) has sought to develop a coordinated system of environmental mapping for the whole community. The main emphasis to date has been on the collation and evaluation of comparable data for the constituent countries and on software to handle them (Wiggins *et al.* 1987). At a global level, the availability of data from satellites, an increasing concern for the global environment and experience with such cartographic databases as World DataBank II have led to increasing interest in the possibility of world wide systems, a view that has led to initiatives to develop a topographic framework for such databases, initially through the activities of D. P. Bickmore, Chairman of a Commission of the International Cartographic Association (ICA) on a World Digital Database for Environmental Science from 1987 until 1990. Mounsey and Tomlinson (1988) have chronicled the progress of GIS in managing and exploiting global databases; it is clear from their book that such developments are still embryonic but are developing rapidly (see Clark, Hastings and Kinneman 1991 in this volume; also Townshend 1991 in this volume).

CONCLUSIONS

A history of GIS is necessarily piecemeal and partial. Inevitably, events are duplicated in different countries at different times. But, despite the unsatisfactory nature of the evidence and the fact that such conclusions are necessarily approximations to reality, four overlapping phases may be distinguished in the development of GIS in the more advanced countries. The first is the pioneer or 'research frontier' period, from the 1950s to about 1975 in the United States and the United Kingdom. This was characterized by individual – even idiosyncratic – developments, limited international contacts, little data in machine-readable form and ambitions which far out-ran the computing resources of the day. Individual personalities greatly influenced events. The second phase was that in which formal experiment and government-funded research was the norm, stretching from about 1973 to the early 1980s; the role of individuals was diminished somewhat in the international and national arenas except for strong-minded heads of national mapping agencies, but at the local level the effect of individuals persisted strongly. Rapidly replacing this phase was the commercial phase commencing *circa* 1982 which, in the light of strong competition among vendors, is now giving way to a phase of user dominance. The last two phases can also be characterized as ones in which systems handling individual data sets on isolated machines (latterly workstations) gave way to those dealing with corporate and distributed databases, accessed across networks and increasingly integrated into the other non-spatial databases of the organization. A vital characteristic of both the latter phases is that these activities became routine: in earlier phases, skilled 'fixers' were required to be on hand to cope with problems in the software, data or hardware.

What particularly emerges from this chapter is the dominant contribution of North America to the development and implementation of GIS up to the mid- and late-1980s, a function of the persuasive power of key individual pioneers, the size of the internal market, the leading role of the United States in the development of computer hardware and software and – above all – an increasing appreciation by many North American users of the need for efficient, speedy and cost-effective means of handling large quantities of geographical data. It is that perception of need which led potential users to seek GIS solutions and has encouraged commercial providers to develop and offer turnkey systems to convert that perceived need into a reality. What is not clear from the piecemeal evidence, however, is the ratio of failures to successes or how many operational systems are fully used and living up to their promises. A federal system of government, where large bureaucracies have considerable powers to take initiatives on their own account and where states are often as large as many independent countries, are no doubt important features, as is the large area of public land to be managed directly by federal and state agencies. Being continental in scale faces both Canada and the United States with particular problems, but it also helps to create an awareness of the importance of GIS to policy. Even so, CGIS remains unique in its scale, comprehensiveness and ambition at a time of inadequate technology.

Developments elsewhere in GIS were more limited until late in the 1980s, although those in Japan, the United Kingdom and several other countries in mainland Europe seem in rapid evolution. Land registration promises to make GIS a globally used technology from the 'bottom up' while earth monitoring from satellites promises to achieve global use 'top down'. It is a reasonable expectation that routine (and often boring, if valuable) use of GIS will be nearly ubiquitous over the next 20 years. This is the end of the beginning of GIS.

REFERENCES

Abler R F (1987) The National Science Foundation National Center for Geographic Information and Analysis. *International Journal of Geographical Information Systems* **1** (4): 303–26

Anderson D E, Angel J C, Gurney A J (1978) World Data Bank II. In: Dutton G (ed.) *Harvard Papers on Geographical Information Systems 2*. Laboratory for Computer Graphics and Spatial Analysis, Harvard University, Cambridge Massachusetts

Andersson S (1987) The Swedish Land Data Bank. *International Journal of Geographical Information Systems* **1** (3): 253–63

Angus-Leppan P (1989) The Thailand Land Titling Project: first steps in a parcel-based LIS. *International Journal of Geographical Information Systems* **3** (1): 59–68

Beaumont J R (1991) GIS and market analysis. In: Maguire D J, Goodchild M F, Rhind D W (eds.) *Geographical Information Systems*. Longman, London, pp. 139–51, Vol 2

Berry J L, Marble D F (1968) *Spatial Analysis: a reader in statistical geography*. Prentice Hall, Englewood Cliffs New Jersey

Bickmore D P, Shaw M A (1963) *Atlas of Great Britain and Northern Ireland*. Clarendon Press, Oxford

Brand M J D (1986) The foundation of a geographical information system for Northern Ireland. In: Blakemore M J (ed.) *Proceedings of AUTOCARTO London*. Royal Institution of Chartered Surveyors, London, pp. 4–9

Brown M J, Norris D A (1988) Early applications of geographical information systems at the Institute of Terrestrial Ecology. *International Journal of Geographical Information Systems* **2** (2): 153–60

Callahan M, Broome F R (1984) The joint development of a national 1 : 100 000 scale digital cartographic database. *Proceedings of the Annual Conference of the American Congress on Surveying and Mapping*, Washington DC, pp. 246–53

Chen Shupeng (1987) Geographical data handling and GIS in China. *International Journal of Geographical Information Systems* **1** (3): 219–28

Chorley R, Buxton R (1991) The government setting of GIS in the United Kingdom. In: Maguire D J, Goodchild M F, Rhind D W (eds.) *Geographical Information Systems: principles and applications*. Longman, London, pp. 67–79, Vol 1

Chrisman N (1988) The risks of software innovation: a case study of the Harvard Lab. *The American Cartographer* **15** (3): 291–300

Christianson C J (1986) Geoprocessing activities in the Fish and Wildlife Survey. *Proceedings of a Geographical Information Systems Workshop*, American Society of Photogrammetry and Remote Sensing, Falls Church Virginia, pp. 43–6

Clark D M, Hastings D A, Kineman J J (1991) Global databases and their implications for GIS. In: Maguire D J, Goodchild M F, Rhind D W (eds.) *Geographical Information Systems: principles and applications*. Longman, London, pp. 217–31, Vol 2

Cocks K D, Walker P A, Parvey C A (1988) Evolution of a continental-scale geographic information system. *International Journal of Geographical Information Systems* **2** (3): 263–80

Commission on Geographical Data Sensing and Processing (1976) *Second interim report on digital spatial data handling in the US Geological Survey*. International Geographical Union, Ottawa

Cook R N (1966) The CULDATA system. In: Cook R N, Kennedy J L (eds.) *Proceedings of a Tri-State Conference on a Comprehensive Unified Land Data System (CULDATA)*. College of Law, University of Cincinnati, pp. 53–7

Cook R N, Kennedy J L (1966) (eds.) *Proceedings of a Tri-State Conference on a Comprehensive Unified Land Data*

System (CULDATA). College of Law, University of Cincinatti
Coppock J T (1988) The analogue to digital revolution: a view from an unreconstructed geographer. *The American Cartographer* **15** (3): 263–75
Coppock J T, Barritt M (1978) *Application of digital techniques to information systems for planning*. Consultants' report to the Scottish Development Department, Edinburgh
Corbett J P (1979) Topological principles in cartography. Technical Paper 48, US Bureau of the Census, Suitland (also published in *Proceedings of AUTOCARTO 4* 1975, pp. 22–33. American Congress on Survey and Mapping/ American Society for Photogrammetry, Washington DC)

Dale P F (1991) Land information systems. In: Maguire D J, Goodchild M F, Rhind D W (eds.) *Geographical Information Systems: principles and applications*. Longman, London, pp. 85–99, Vol 2
Dangermond J, Smith K L (1988) Geographic Information Systems and the revolution in cartography: the nature of the role played by a commercial organization. *The American Cartographer* **15** (3): 301–10
Diello J, Kirk K, Callender J (1968) The development of an automatic cartographic system. *Cartographic Journal* **6** (1): 9–17
Department of the Environment (DoE) (1972) *General Information Systems for Planning*, Department of the Environment, London
Department of the Environment (DoE) (1987) *Handling geographic information: the report of the Committee of Inquiry*. Her Majesty's Stationery Office, London
Dewdney J C, Rhind D W (1986) The British and United States Censuses of Population. In: Pacione M (ed.) *Population Geography: progress and prospects*. Croom Helm, London, pp. 35–57
Diello J, Kirk K, Callander J (1969) The development of an automated cartographic system. *Cartographic Journal* **6**: 9–17
Dueker K (1974) Urban geocoding. *Annals of the Association of American Geographers* **64**: 318–25
Duffield B S, Coppock J T (1975) The delineation of recreational landscapes: the role of computer-based information systems. *Transactions of the Institute of British Geographers* **66**: 141–8

Experimental Cartography Unit (1971) *Automatic Cartography and Planning*. Architectural Press, London
Eyre L A (1989) JAMGIS, the first Jamaican Government comprehensive multi-data geographical information system: achievements and problems. *International Journal of Geographical Information Systems* **3** (4): 363–71

Finch S (1987) *Towards a national digital topographic data base*. Unpublished PhD thesis, University of London.

Gaits G M (1969) Thematic mapping by computer. *Cartographic Journal* **6**: 50–68
Giles R H (1987) The creation, uses and demise of a Virginia USA Geographical Information System. *Proceedings of the International Geographical Union GIS Workshop, Beijing, China*: pp. 507–24
Goldstein H, Wertz J, Sweet D (1969) *Computer Mapping: a tool for urban planners*. Battelle Memorial Institute, Cleveland
Goodchild M F, Rhind D W (1990) The US National Center for Geographic Information and Analysis: some comparisons with the Regional Research Laboratories. In: Foster M J, Shand P J (eds.) *The Association for Geographic Information Yearbook 1990*. Taylor & Francis, London, pp. 226–32
Green N P A (1987) Teach yourself geographical information systems: the design, creation and use of demonstrators and tutors. *International Journal of Geographical Information Systems* **1** (3): 279–90
Grimshaw D J (1988) Land and property information systems. *International Journal of Geographical Information Systems* **2** (1): 67–79

Hagerstrand T (1967) The computer and the geographer. *Transactions of the Institute of British Geographers* **42**: 1–20
House of Lords (1984) *Remote sensing and digital mapping*. Report 98 of the House of Lords Select Committee on Science and Technology, Her Majesty's Stationery Office, London
Hudson J C (1979) (ed.) Seventy-five years of American geography. *Annals of the Association of American Geographers* **69** (1): 185pp.

James P E, Martin G J (1978) *The Association of American Geographers: the first seventy-five years 1904–1979*. Association of American Geographers, Washington DC

Koshkariov A V, Tikunov V S, Trofimov A M (1989) The current state and main trends in the development of geographical information systems in the USSR. *International Journal of Geographical Information Systems* **3** (3): 257–72
Kubo S (1987) The development of geographical information systems in Japan. *International Journal of Geographical Information Systems* **1** (3): 243–52
Kubo S (1991) The development of GIS in Japan. In: Maguire D J, Goodchild M F, Rhind D W (eds.) *Geographical Information Systems: principles and applications*. Longman, London, pp. 47–56, Vol 2

Lay J C (1975) Mapping services in Fairfax County, Va. *Proceedings of AUTOCARTO 4*. American Congress on Survey and Mapping/American Society for Photogrammetry, Washington DC, pp. 143–7
Loomis R G (1965) Boundary networks. *Communications, Association for Computing Machinery* **8**: 44–8
Lyall G A (1980) Planning and land assessment in Scotland – the role of the Rural Land Use Information Systems Working Party. In: Thomas M F, Coppock J T (eds.) *Land Assessment in Scotland*. Aberdeen University Press, Aberdeen, pp. 107–17

Maguire D J (1991) An overview and definition of GIS. In:

Maguire D J, Goodchild M F, Rhind D W (eds.) *Geographical Information Systems: principles and applications*. Longman, London, pp. 9–20, Vol 1

Mahoney R P (1991) GIS and utilities. In: Maguire D J, Goodchild M F, Rhind D W (eds.) *Geographical Information Systems: principles and applications*. Longman, London, pp. 101–14, Vol 2

Marble D F (1980) (ed.) *Computer Software for Spatial Data Handling*, 3 volumes. Commission on Geographical Data Sensing and Processing/International Geographical Union, Ottawa Canada

Marble D F (1989) Letter to PERS. *Photogrammetric Engineering and Remote Sensing* **55** (4): 434–5

Mitchell W B, Guptill S C, Anderson E A, Fegeas R G, Hallam C A (1977) GIRAS – a Geographic Information Retrieval and Analysis System for handling land use and land cover data. *Professional Paper 1059*, USGS Reston Virginia

Morrison J L (1991) The organizational home for GIS in the scientific professional community. In: Maguire D J, Goodchild M F, Rhind D W (eds.) *Geographical Information Systems: principles and applications*. Longman, London, pp. 91–100, Vol 1

Morton G M (1966) *A Computer Oriented Geodetic Data Base and New Technique in File Sequencing*. IBN Ltd, Ottawa Canada

Morton Index (1966) In: Tomlinson R F (1972) (ed.) *Geographic Data Handling*. Commission on Geographical Data Sensing and Processing. International Geographical Union, Ottawa Canada

Mounsey H M (1988) (ed.) *Building Databases for Global Science*. Taylor & Francis, London

National Joint Utilities Group (NJUG) (1987) *The NJUG specification for digital map data*. Working Note 12, NJUG, c/o Electricity Council, London

Nordbeck S (1962) Location of areal data for computer processing. *Lund Studies in Geography, Series C, General, Mathematical and Regional Geography No. 2*, Lund University Sweden

O'Callaghan J F, Garner B J (1991) Land and Geographical Information Systems in Australia. In: Maguire D J, Goodchild M F, Rhind D W (eds.) *Geographical Information Systems: principles and applications*. Longman. London, pp. 57–70, Vol 2

Ottoson L, Rystedt B (1991) GIS in Sweden. In: Maguire D J, Goodchild M F, Rhind D (eds.) *Geographical Information Systems: principles and applications*. Longman, London, pp. 39–46, Vol 2

Perring F H (1964) Contribution to Session 6, The mapping of vegetation flora and fauna. In: Bickmore D P (ed.) *Experimental Cartography – Report of the Oxford Symposium*. Oxford University Press, Oxford

Perring F H, Walters S M (1962) *Atlas of the British Flora*. Nelson, London

Peucker T K (1972) Computer cartography. *Commission on College Geography, Resource Paper No.17*. Association of American Geographers, Washington DC

Raper J F, Green N P A (1989) Development of a hypertext-based tutor for geographical information systems. *British Journal of Educational Technology* **3**: 164–72

Rhind D W (1971) The production of a multi-colour geological map by automated means. *Nachr. aus den Karten und Vermessungswesen* Heft Nr. **52**: 47–51

Rhind D W (1976) Geographical Information Systems. *Area* **8** (1): 46

Rhind D W (1986) Remote sensing, digital mapping and Geographical Information Systems: the creation of government policy in the UK. *Environment and Planning C: Government and Policy* **4**: 91–102

Rhind D W (1987) Recent developments in geographic information systems in the UK. *International Journal of Geographical Information Systems* **1** (3): 229–41

Rhind D W (1988) Personality as a factor in the development of a new discipline: the case of computer-assisted cartography. *The American Cartographer* **15** (3): 277–89

Rhind D W (1990) Topographic databases derived from small scale maps and the future of Ordnance Survey. In: Foster M J, Shand P J (eds.) *The Association for Geographic Information Yearbook 1990*. Taylor & Francis, London, pp. 87–96

Rhind D W (1991) Counting the people: the role of GIS. In: Maguire D J, Goodchild M F, Rhind D W (eds.) *Geographical Information Systems: principles and applications*. Longman, London, pp. 127–37, Vol 2

Rhind D W, Hudson R (1980) *Land Use*. Methuen, London

Rhind D W, Mounsey H M (1989) The Chorley Committee and 'Handling Geographic Information'. *Environment and Planning A* **21**: 571–85

Robinette A (1991) Land management applications of GIS in the State of Minnesota. In: Maguire D J, Goodchild M F, Rhind D W (eds.) *Geographical Information Systems: principles and applications*. Longman, London, pp. 275–83, Vol 2

Rushton G (1969) A comprehensive model for the study of agricultural land use patterns. *Computer Assisted Instruction in Geography. Commission on College Geography, Technical Paper No. 2*. Association of American Geographers, Washington DC, pp. 141–50

Schmidt A H, Zafft W A (1975) Progress of the Harvard University Laboratory for Computer Graphics and Spatial Analysis. In: Davis J C, McCullagh M J (eds.) *Display and analysis of spatial data*. Wiley, London, pp. 231–43

Schweitzer R H (1973) *Mapping Urban America with Automated Cartography*. Bureau of the Census US Department of Commerce, Suitland Maryland

Shumway C (1986) Summary of the US Forest Service Geographical Information Systems activities. *Proceedings of Geographical Information Systems Workshop*. American Society for Photogrammetry and Remote Sensing, Falls Church, pp. 49–52

Simonett D S (1988) Considerations on integrating remote sensing and Geographic Information Systems. In:

Mounsey H M, Tomlinson R F (eds.) *Building Databases for Global Science*. Taylor & Francis, London pp. 105–28

Sowton M. (1991) Development of GIS-related activities at the Ordnance Survey. In: Maguire D J, Goodchild M F, Rhind D W (eds.) *Geographical Information Systems: principles and applications*. Longman, London pp. 23–38, Vol 2

Starr L E, Anderson K E (1991) A USGS perspective on GIS. In: Maguire D J, Goodchild M F, Rhind D W (eds.) *Geographical Information Systems: principles and applications*. Longman, London, pp. 11–22, Vol 2

Taylor D R F (1991) GIS and developing nations. In: Maguire D J, Goodchild M F, Rhind D W (eds.) *Geographical Information Systems: principles and applications*. Longman, London, pp. 71–84, Vol 2

Tobler W R (1959) Automation and cartography. *Geographical Review* **49**: 526–34

Tomlinson R F (1967) *An Introduction to the Geographic Information System of the Canada Land Inventory*. Department of Forestry and Rural Development, Ottawa Canada

Tomlinson R F (1970) (ed.) *Environment Information Systems*. Commission on Geographical Data Sensing and Processing. International Geographical Union, Ottawa Canada

Tomlinson R F (1972) (ed.) *Geographic Data Handling*. Commission on Geographical Data Sensing and Processing. International Geographical Union, Ottawa Canada

Tomlinson R F (1974) *The application of electronic computing methods to the storage, compilation and assessment of mapped data*. Unpublished PhD thesis, University of London

Tomlinson R F (1985) Geographic Information Systems – the new frontier. *The Operational Geographer* **5**: 31–6

Tomlinson R F (1988) The impact of the transition from analogue to digital cartographic representation. *The American Cartographer* **15** (3): 249–62

Tomlinson R F (1989) Letter to PERS. *Photogrammetric Engineering and Remote Sensing* **55** (4): 434–5

Tomlinson R F, Calkins H W, Marble D F (1976) *Computer Handling of Geographical Data*. Natural Resources Research Series XIII, UNESCO Press, Paris

Tomlinson R F, Petchenik B B (1988) (eds.) Reflections on the revolution: the transition from analogue to digital representations of space, 1958–1988. *The American Cartographer* **15** (3): 243–334

Townshend J R G (1991) Environmental databases and GIS. In: Maguire D J, Goodchild M F, Rhind D W (eds.) *Geographical Information Systems: principles and applications*. Longman, London, pp. 201–16, Vol 2

USBC (1969–73) *Census Use Study Reports 1 to 12*. US Bureau of Census, Washington DC

Waugh T C (1980) The development of the GIMMS computer mapping system. In: Taylor D R F (ed.) *The Computer in Contemporary Cartography*. Wiley, London, pp. 219–34

Wiggins J C, Hartley R P, Higgins M J, Whittaker R J (1987) Computing aspects of a large geographical information system for the European Community. *International Journal of Geographical Information Systems* **1** (1): 77–87

THE COMMERCIAL SETTING OF GIS

J DANGERMOND

This chapter deals with the commercial setting of GIS, in contrast with historical, technological, government or academic settings. After an introduction and a brief indication of the present scope of the GIS business, GIS firms are distinguished from those in closely related technologies, and a typical commercial GIS is briefly described. Then, in the chapter's major section, the wide variety of commercial GIS firms are categorized, and each category briefly described; a related section describes how commercial firms support the GIS systems cycle. Then the contribution of commercial firms to the field, their problems, some of their frustrations, the GIS trends of which they are a part, and the future of the commercial sector are each discussed in turn. Finally, some brief concluding remarks are provided.

INTRODUCTION

GIS technology provides a framework for all forms of spatial data storage, retrieval, analysis, display and modelling. It provides the 'front end' technology for multimedia spatial databases including video, CD-ROM, tabular, and other forms of data.

Market forces and healthy commercial competition are a primary driving force causing thinkers to forge ahead with new ideas, concepts and techniques in any technological field. It was apparent, 20 years ago, that the GIS field would grow rapidly only if large commercial organizations could be enticed to enter it.

Development in the GIS field demanded many elements: hardware, to provide the fundamental enabling capabilities required; construction of a sound theoretical basis for geographical relationships and a model of how geographical reality could be abstracted for data processing; engineered software products which would encapsulate the scientific notions of spatial analysis and geographical data processing; creation of demand for spatial information in order to address complex problems about geography; creation of an industry which could manufacture and distribute GIS technology; and creation of a research environment with all its competitive mechanisms, for ensuring advances in methods and techniques. Each of these elements required appropriate institutional settings and the development of people who would create and drive these institutions.

Over the last two decades, these elements have come into being. The GIS field is now coming of age with: the founding of the US National Center for Geographic Information and Analysis (NCGIA) and its equivalent in other countries; the selection by IBM of GIS as one of its five strategic markets for the 1990s; the adoption of GIS by virtually all US national agencies; the massive emerging general interest in cartography and geography; environmental crises at local, national and global levels; resource shortages; the decay as well as rapid growth of cities; and the need to manage natural resources better.

This chapter describes what GIS in the commercial sector presently supplies, how it goes about its business, what problems it faces, and where present trends are leading. The chapter concentrates on those aspects of commercial operations which are influenced directly by doing business in the GIS field. The perspective is that of a major firm in the field, which now supplies a broad range of GIS technology, but which recalls vividly what a 'start-up mode' of operation was like.

THE LITERATURE

In comparison with the literature of the computer field, the literature of the GIS field is still in a rather undeveloped state; this is especially so with regard to the strictly commercial aspects of the field. Nevertheless, there are a few sources which identify vendors of GIS hardware, software, and services. Often company profiles are also provided.

Surveys of the commercial sector in the United Kingdom are included in the AGI Yearbooks (Shand and Moore 1989; Foster and Shand 1990). The most thorough survey of the US industry is that by Walker and Miller (1990). An annually updated sourcebook for the United States (*GISWorld* 1990) also provides useful up-to-date information.

A number of privately prepared (and quite costly) survey reports aim to describe the commercial GIS field. These provide dollar estimates of total revenues, market share, sales by sector, estimates of growth rates by sector and similar kinds of data. Some also include company profiles and technical accounts of recent developments. These include, for example, those of Dataquest, Inc. (San Jose, California, a subsidiary of Dun & Bradstreet) and Daratech, Inc. (Cambridge, Massachusetts).

SCOPE OF COMMERCIAL GIS BUSINESS

The GIS business began about 20 to 25 years ago, was relatively small for the first decade of its existence, and has been growing very rapidly in the last decade. While estimates vary, growth rates around 25–35 per cent per year over the next five years seem reasonable (see Maguire 1991 in this volume). The number of firms which claim to provide at least some GIS-related goods or services is now approaching 1 000. Annual gross sales are difficult to determine, in part because of varying definitions of what the GIS business includes (e.g. including or excluding sales of general-purpose computing machinery which is used for GIS applications). Using a broad definition, annual sales of GIS technology are probably about 500 million US dollars at the present time. Software sales may represent a quarter of that amount.

Commercial GIS and closely related technologies

Commercial GIS are commonly based on both vector and raster computer technologies; however, the commercial GIS field is also related to a number of other spatial information technologies, usually because of interchanging data with them, or using them in managing or displaying data. Many commercial firms supply these other technologies along with GIS-related products. Many commercial organizations which perform photointerpretation, remote sensing, photogrammetry, or image processing now offer some GIS-related products, such as creating automated geographical databases which include the products of their work; many organizations which supply computer aided drafting or cartographic services also offer GIS-related products. A number of computer hardware manufacturers have begun to offer GIS software as well. A large number of firms which have specialized in services to a particular industry (e.g. forestry, engineering) have begun to offer GIS services and consulting in support of their project work. General purpose computer software suppliers are now creating some software (such as DBMS software) useful in GIS.

'TYPICAL' COMMERCIAL GIS

A minimal GIS might consist of a computer with accompanying memory, some sort of GIS software, geographical data, a person to operate the system, and a set of procedures which are to be followed in its use. More complete GIS commonly include digitizing tablets, colour graphics terminals, and hardcopy devices such as plotters, electrostatic printer/plotters, or the like. Larger GIS can be created by adding more components; using larger capacity or more sophisticated devices; connecting components together; and increasing staffing, data and organizational complexity. Because GIS often perform compute-intensive operations (such as overlaying), many user organizations have found it best to dedicate a computer system entirely to GIS operations rather than running GIS as one of several applications. Even then, graphic displays are often slow unless a very powerful graphics processor is included in the system. Because large databases are usually involved, GIS often require very large memory resources. At present, commercial GIS can be based on mainframe computers, minicomputers, workstations, or personal computers, and GIS and their component hardware can be networked together in various ways. The accompanying figures (Figs. 4.1–4.5) show several such GIS configurations and their approximate costs (in US dollars). At present, sales of workstation and personal computer based systems are growing more rapidly than those of other configurations.

CATEGORIES OF COMMERCIAL GIS

Primary GIS businesses

Hardware technology

Hardware development continues to be a major driving force in the development of GIS. The major trend in this area is towards less expensive, faster and smaller computers; in the early 1990s these are workstations of the 20–50 MIPS variety. By the mid-1990s, these are expected to become 50–250 MIPS workstations at a cost affordable for personal computing. These computers, together with file servers and very large capacity storage devices, are rapidly being networked and, in the future, they will operate as nodes within a vast computer and data management environment (for more discussion of the technological setting of GIS see Goodchild 1991 in this volume).

Most hardware development is of general-purpose computing machinery, which is then applied to GIS. Only a few hardware devices are completely GIS specific, but some devices, such as digitizers, plotters and scanners, find especially important applications in the GIS field. The marketing of computer hardware for GIS use has become increasingly oriented to an 'open system' approach in which buyers are purchasing a variety of independent system devices and avoiding the overhead costs of solutions in which a single vendor supplies all the hardware for a complete GIS.

A discussion of the diversity of the commercial hardware business, even as it applies to GIS, is far beyond the scope of this chapter.

Software technology

Software used in GIS may be GIS specific, although some software, such as DBMS software, may be a general purpose product applied to GIS. The development of GIS-specific software may represent the 'purest' primary GIS business, since the product has no other application than GIS and is essential to the existence of a GIS.

Because of intense competition and rapid technical developments, the most critical factor in a firm's commercial success in GIS software development is probably a programme of continuous research and development, together with the creation of popular user applications. Reputation among users, the number of installed systems, the quality of software documentation and user training, and the firm's ability to deal with its customers/users are also quite critical to commercial success. To meet these demands, a successful GIS software developer probably requires staff trained in computer science, geography, cartography and a range of related disciplines. As the software becomes more sophisticated, staff specializing in the particular fields to which the GIS software is to be applied also become valuable.

Some commercial GIS software systems are hardware specific, others run on a variety of hardware systems. Similarly, some systems interface with a variety of DBMS, graphics software, and so on, while others offer less flexibility. A variety of strategies are used in coping with the problems which portability, or lack of it, presents.

The organization and functions of a commercial firm supplying GIS software are similar to firms supplying other kinds of software. These include functions like software development, quality assurance, documentation, installation, training, field support, marketing, and the like. Sales of software for major GIS often involve considerable consulting, often by third parties. Sales and distribution seem to be most effective when potential users receive actual demonstrations of the

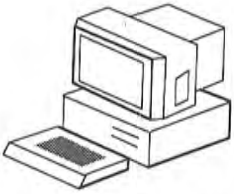


Fig. 4.1 Personal computer ($5 000–$25 000)

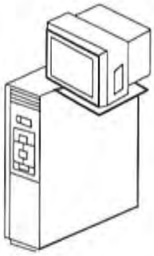


Fig. 4.2 Single user workstation ($10 000–$50 000)

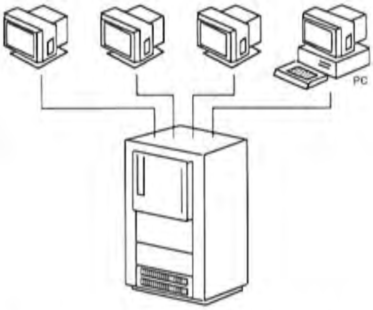


Fig. 4.3 Minicomputer using VMS, AOS/VS, X-OS, or similar operating systems ($80 000–$1 million)

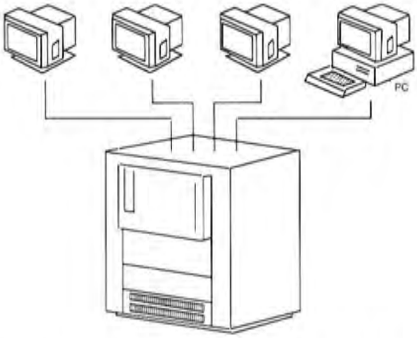


Fig. 4.4 Mainframe using VM/CMS, MVS, or similar operating systems ($2 million–$15 million)

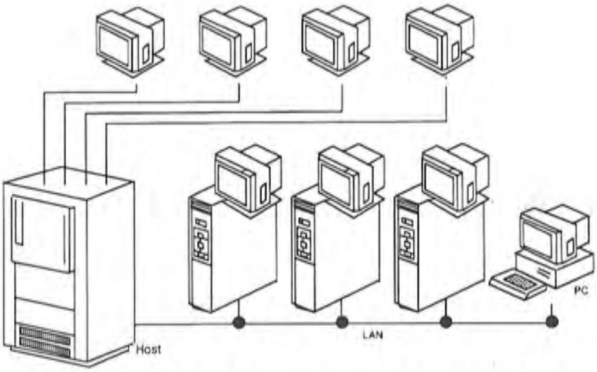


Fig. 4.5 Network with server(s) using UNIX and mainframe or minicomputer operating systems ($100 000–virtually unlimited)

Five different hardware platform approaches to supporting geographical information systems and ranges of typical hardware cost. (Cost ranges assume purchase of new equipment) capabilities of the software, although with the wider distribution of GIS, more potential buyers are familiar with GIS capabilities before they contact software vendors. Programming support and hotline support are important in the GIS field because the technology is relatively new, because it is just beginning to be taught extensively in colleges and universities, and because the applications of the technology are rapidly diversifying. User groups are still useful because of these same factors, but they are already specializing by application type, hastened by the rapid increase in the number of GIS users in the world.

Sales of GIS software for personal computers is already a much simpler process than that for GIS software on other computer platforms: following receipt of a written or phone request, and subsequent payment for the product, the user receives the user installable software, written documentation and a training package (perhaps including videotaped instruction and a training database). As personal computers continue their evolution into higher performance workstations, these lower end focused application products will play the valuable role of attracting new users into the GIS field.

Applications programming

General purpose GIS software must often be adapted to particular applications (urban planning, forestry, etc.; see also the chapters in Section III of this volume). This application may be accomplished by the user, by the original software vendor who offers prepackaged applications, by using 'macro' languages (which make it easier for users to write their own applications programs), or by a custom programming service. A major trend in recent years has been the development of third party organizations which provide application assistance to users. Many persons who gained experience in the GIS field in the late 1970s and early 1980s are now providing such 'value added' services to clients who have GIS systems. Some of these firms are also becoming GIS software distributors for the major GIS software vendors. They then add value by providing either their applications software packages or custom programming support.

Database development

The largest cost in most GIS continues to be the database. For the last 20 years automation of geographical data has been performed chiefly by digitizing and key entry. In recent years, scanning and conversion from existing automated files have also become important. Nevertheless, no present technology permits easy and inexpensive capture of previously mapped data, let alone spatial data in other non-digital forms.

For this and other reasons, there is an enormous backlog of geographical data which various organizations would like to put into digital form. A growing number of firms is providing complete assistance in the processes of database design, data capture, data conversion, data automation, data editing, database creation and related services. Many of these firms specialize, some working chiefly with natural resource data, others with urban data, and so on.

Database creation is an exceedingly complex task (see, for example, Dangermond and Harnden 1990), involves many steps, and requires great care, skill and experience if the result is to be satisfactory. Some of the firms perform their work parameter by parameter, perhaps integrating the data after automation; others integrate and standardize data before they are automated.

Some GIS are old enough that one or more cycles of complete updating by commercial firms have now been undertaken (as in rapidly growing urban areas).

Many GIS are now being designed for continuous updating through transactions, presenting problems in GIS administration which are now under intense study.

Commercial organizations applying GIS technology

The growth in the number of commercial firms which apply GIS technology in particular fields has been explosive in the last few years. Such organizations work in the fields of surveying, photointerpretation and photogrammetry, image processing, urban and regional planning, real estate, vehicle navigation, utilities, energy, mining and minerals, market research, landscape architecture, architecture, development, forestry, coastal planning, ecology, environmental planning and regulation, parks and recreation, land management, agriculture, military/defence, cartography and national mapping, water resources, flood control, civil engineering, transportation engineering, sanitary engineering, transportation planning, communications, and many others. (In part because of the employment opportunities being created in such firms, a growing number of colleges and universities are acquiring GIS for use in teaching.)

Consulting

A significant number of organizations and individuals now offer consulting on various aspects of GIS. The services offered include assessment of possible applications of GIS, GIS feasibility studies, system and database design, project design, assistance in designing specifications and Requests for Proposals (RFPs), project supervision, benchmark testing, advice on special problems in GIS, and so on. Some of the consultants also offer the capability to execute the plans they devise, including complete implementation of a system; others are third-party providers of services only.

General-purpose GIS technology firms

A few firms offer to provide a very wide range of GIS-related goods and services. These are often firms which have been in the field for more than a decade, are relatively large in size, and have accumulated a good deal of experience. These firms typically offer a wide range of consulting services, turnkey GIS, database creation, custom software programming, and complete project support services, from concept to final working system. Another approach is for a firm to offer organizational management and financing for the GIS, and then contract for all the required elements of the final GIS.

GIS-related businesses

Training/education

As contrasted with colleges and universities which may deal with education, these firms usually offer highly specific training about GIS and its applications. Some of these firms offer a broad range of related services as well. A few concentrate on seminars and courses which introduce GIS in a more academic setting. A number of commercial organizations now offer services in organizing GIS-related meetings, conferences, symposia and the like. These seem to function in ways somewhat similar to professional societies.

Publishing

A number of commercial organizations are now heavily involved in providing GIS-related journals, newsletters, market reports, research reports, and the like. A much larger number of such organizations deal with the field more peripherally by running occasional GIS-related advertisements, articles, reviews, application descriptions and similar materials. As well as potential sources of revenue, these offer a useful means of information dissemination and advertising.

Database publishers

A few commercial firms exist which might be said to be 'database publishers'. These firms provide previously created digital databases of various kinds of geographical information in some exchange format. An example is the firms providing road centreline files, initially driven by the requirements of road navigation systems. As rights in data are clarified, arrangements for data sharing with governmental agencies are worked out, and the demand for quick availability of geographical data increases, these firms may find an important niche in the GIS field, much like national government's publication of digital files.

Others

In addition to the firms mentioned above with involvement in photogrammetry, image processing, there are other firms, in a wide range of data capture and data processing areas, which simply supply their products to others for automation. These firms, nevertheless, play an absolutely essential role in the creation and support of GIS technology: their development in the last 20 years, along with developments in computer hardware technology, are probably the two major driving forces in the growth of the GIS field itself.

SYSTEM CYCLE SUPPORT

While some experienced GIS user organizations continue to buy selected GIS goods and services, the rapid growth of GIS use has meant that a growing number of customers want complete support for their GIS, from initial concept to a complete functioning system. The complete GIS development cycle, as it is supported by some commercial organizations, will be sketched out in this section. The basic steps in the development cycle are indicated in Figure 4.6 (see also Clarke 1991 in this volume).

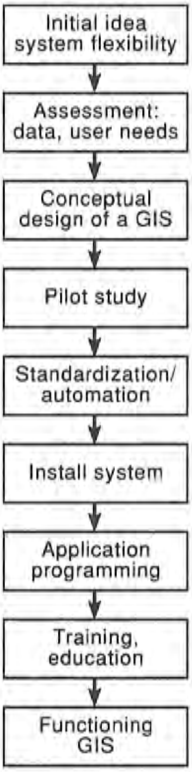


Fig. 4.6 A simplified view of the GIS system cycle

The process (described here only in skeleton outline) begins with formulation of needs and a concept; this is best done under a single director, using hard working, highly motivated 'teams of two' persons (one technically oriented, the other management oriented). Then a thorough study is made of user needs, of existing data resources, and of how the GIS will affect decision making. A cost/benefit or feasibility analysis of the proposed GIS may be performed at this point. The hardware, software, database, staff, organization, and other resources required are identified and then a design for the system and an implementation plan are created and reviewed with the user organization. If approved, the system implementation is begun. Vendors are contacted, site visits made, necessary consultants employed, applications identified, alternatives considered. A request for proposal may be drafted, including specifications, performance criteria, requirements for a technology selection test, and other requirements. Vendor selection follows. The data to be used are gathered, organized, reclassified as necessary, standardized for the new database, updated as necessary, integrated, prepared for automation, automated, edited, displayed for comparison with the source data, prepared for storage, and included in the growing database. Hardware is acquired, installed and tested, and systems software installed. Software is acquired and installed. Users are trained, and then their training reinforced by work on a pilot project which tests the entire GIS.

For a specific project, the system is used to prepare atlas maps of the project site, models are prepared for analysing the project data, outputs of various kinds are obtained, and, based on these, a plan, recommendations, decisions, or similar outputs are provided.

For a GIS providing continuing support to an organization's decision-making process, the initial database is gradually expanded, updated and enhanced, perhaps through the transaction process. Over the years the system cycle may have to be repeated, providing for updated, enlarged, or additional capabilities. Reorganizations, changes in mission or mandate, and other changes, may affect this process. Experience suggests that many GIS will have to be expanded as soon as users discover what they are capable of doing. Upgrading of skills, introduction to new techniques and methods, and similar ongoing education and training are required throughout the system's life.

Commercial organizations may also assist in integration of a GIS with an existing information system, GIS technology transfer, software conversion to new hardware platforms, creation of specific GIS products (such as atlases, special maps, etc.), and other, more specialized, services.

COMMERCIAL FIRMS' CONTRIBUTIONS TO THE GIS FIELD

The commercial sector makes a variety of contributions to the GIS field as a whole. A good deal of research and development is undertaken, in order to bring new products to the market; pilot studies of actual applications are often an important part of the process. Commercial firms are often the first actually to provide delivery of new technology to users in the form of hardware, software, methods and techniques, applications, training and so on. Having made it available, commercial firms continue its support through hotlines, technical consulting, technical literature, user groups and the like. Commercial firms may provide partial or complete subsidies of the development of new technology by other sectors, especially educational institutions and individual researchers; they may also make the technology (e.g. hardware, software, project support) available at low cost or no cost to educational institutions and special classes of users, such as non-profit organizations, international organizations dealing with the underdeveloped world, or the like. Donations of new and used equipment may be made to institutions which could not otherwise afford to acquire it. Commercial organizations usually provide major sponsorship of technical meetings and conferences, technical journals and technical research. Advertising by such firms performs an educational function for potential users, making them more aware of the technology and what it can do; advertising revenues are a major source of funds supporting the existence of professional journals. Representatives of individual firms or industry groups often take the lead in organizing industry committees and working groups or joint committees involving all the sectors of the GIS field, such as those dealing with standards; industry representatives often provide certain kinds of public policy advice to government agencies at nominal cost. Industry is a major employer of persons who practise GIS, probably exceeded by government, but, today, ahead of academia in this regard. Commercial firms often conduct extensive education and training efforts for their own employees, either formally or informally, and create a major reservoir of trained and experienced professionals and technicians which eventually enriches all sectors of the GIS field.

PROBLEMS OF COMMERCIAL GIS BUSINESS

While its success has been remarkable in recent years, the commercial sector of the GIS field must work hard to solve various problems if it is to continue to flourish; a few of these problems are mentioned below.

Problems of perception and understanding of the technology

The vast majority of people who could usefully employ GIS technology still know little or nothing about it. The message about its usefulness is still not being heard in the places where the technology is most needed, such as in the Third World, in the vast majority of municipal governments throughout the world, on the farms of the world, and by average citizens, whose access to the technology (often paid for by public funds) is difficult or impossible. Present educational efforts in colleges and universities, and even the expanding use of the technology, will only improve this situation over a period measured in decades (see Unwin 1991 in this volume for further details).

Potential buyers and users of GIS technology continue to be confused and, in some cases, turned away from GIS use by various advertising and sales practices of some commercial GIS firms. The number of commercial firms claiming to have GIS is rapidly expanding, but the vast majority of these software systems have only rudimentary GIS capabilities.

Companies continue to announce GIS software long in advance of the earliest possible delivery dates, and many such programs are never delivered to users at all. Many GIS continue to perform so badly that users, and those who listen to them, are convinced that the 'promise' of GIS technology is no more than a deception.

Many of these problems would be eliminated if potential users were far better informed and educated about the technology, if professional standards of practice were observed, and if more users would make use of objective tests, such as benchmarks, to evaluate the competing claims of vendors. But beyond this, all sectors of the GIS field need to work harder at explaining what GIS technology is, what it can do, and how it can be most effectively and inexpensively employed.

Problems with GIS technology

There are numerous technical problems to resolve at the present time, but, perhaps paradoxically, these are likely to be resolved much sooner than the problems mentioned above. These technical problems, discussed elsewhere in this volume, include the difficulties in connecting hardware and software from different vendors; the difficulties in interconverting data created in different ways; the data automation problem; and so on.

Problems between the sectors of the GIS field

Finally, commercial organizations continue to find themselves in competition with government and sometimes academia in providing goods and services to users. As in other technical fields, this is a source of concern for all parties; it may also be a source of creative tension which benefits all.

SOME FRUSTRATIONS OF THE COMMERCIAL SECTOR

Some problems are so intractable and persistent that they colour every aspect of commercial operations in the GIS field. While opinions will differ as to what these intractable problems are, here is one short list:

- Given the speed with which the field is developing, and the difficulty commercial organizations are having in keeping up with that growth, it may seem paradoxical to suggest that the slowness with which the technology has been accepted is frustrating; but, given what the technology is capable of, and the rapid march of the global problems which it could help to alleviate, the relative snail's pace of its development is extremely frustrating.
- Though costs are falling rapidly, the continuing high cost of the technology is also frustrating.
- It is frustrating to GIS professionals to deliver such an effective and powerful technology and then see it either underused, misused, or abandoned by users, for reasons which have nothing to do with the technology itself. A related frustration is commercial firms' inability, thus far, to deliver GIS technology to those who need it most in the world, the people in underdeveloped countries. Both these frustrations may be related to a third, the difficulty in transferring this technology to users who lack a rather rich educational and, perhaps, cultural background (this issue is explored further in Taylor 1991 in this volume).
- Of the GIS technical problems, the most frustrating continues to be the collection and automation of data, still probably the chief technical barrier to wider use of GIS (see Jackson and Woodsford 1991 in this volume).

TRENDS OF THE COMMERCIAL SECTOR

Many changes have occurred in the business of GIS in the 25 or so years of its existence. Only a few can be mentioned here. Figure 4.7 also indicates some of these.

Project oriented GIS uses → Database oriented GIS uses
Mainframe-based GIS → Minicomputer-based GIS
Minicomputer-based GIS → Workstation-based GIS
Minicomputer-based GIS → Personal computer-based GIS
Minicomputer-based GIS → Network-based GIS
Small area databases → Global databases
Isolated users → Networked users

Fig. 4.7 Some of the trends in the commercial sector

The dominant trend in the commercial sector is unquestionably the rapid growth in the sales of GIS. This certainly reflects, in turn, steeply rising user acceptance of, and user demand for, the technology.

An underlying trend, which has probably fuelled these rapid increases, has been the rapid increase in the performance/cost ratio of general-purpose computer hardware over the entire history of GIS. This permitted commercial firms to break into the GIS turnkey system business in the 1970s by being able to support GIS on minicomputers instead of just on mainframes; in the 1980s, GIS could be based on personal computers. Now performance is being enhanced through the use of workstations and by networking hardware components.

One reflection of these changes has been a clear trend from GIS being used to perform the work of single, isolated projects, to the sale of the technology to users who create databases and systems which they intend to use continuously, over a long period, for a series of applications. This trend has made the industry what it is today. A decade ago the largest part of the business was in services, usually performance of complete projects for users. Now the largest part is sales of hardware, software, training and support, to users.

As users have become increasingly responsible for their own systems, and the cost for entry-level GIS has fallen, the diversity of users (Fig. 4.8) has increased; so also has the number of systems serving many different purposes.

	Computer-aided drafting	Remote sensing	Raster GIS	Vector GIS	Network analysis	Coordinate geometry	3-D modelling	Laser disk storage
National development	▨	▨	▨	▨	▨	▨	▨	▨
Urban planning	▨	▨	▨	▨	▨	▨	▨	▨
Renewable resources		▨	▨	▨		▨	▨	
Utilities	▨	▨	▨	▨	▨	▨		▨
Transportation	▨	▨	▨	▨	▨	▨		▨
Environment		▨	▨	▨	▨	▨		
Agriculture		▨	▨	▨		▨		

Fig. 4.8 Spatial data technology and applications

The practical upper limit in the size and complexity of GIS databases is growing rapidly; the first true global GIS have only been around for a few years (see Clark, Hastings and Kineman 1991 in this volume). As database size has expanded, increased efforts have been devoted to data capture and automation. As the field has aged, increasing efforts have been devoted to database maintenance and updating.

If present trends in falling cost, increasing ease of use, and rising user interest continue, GIS may become as commonly used as computer graphics.

FUTURE COMMERCIAL DEVELOPMENT OF NEW TECHNOLOGIES

Many factors influence whether and when a particular development in GIS technology becomes 'commercial'. These include cost, potentially useful applications, ease of use or 'user friendliness', concept demonstration through suitable pilot applications, the willingness of one or more firms to invest in bringing the technology to market, availability of necessary supporting technologies, and so on. Sometimes what is most necessary, or most lacking, is a 'champion': a person or an organization to push the technology until it is accepted.

An example of a technology which seems to offer considerable promise, but which is not presently significant in GIS-related sales, is artificial intelligence/expert systems (see Smith and Ye Jiang 1991 in this volume). Some GIS applications have been made, but have not yet been widely accepted by users.

Scanning is finding increasing acceptance for data capture, although it continues to have major technical limitations in dealing with many kinds of mapped data.

CONCLUSIONS

The commercial sector of the GIS field is just beginning its period of most rapid growth. Like the rest of the field, the commercial sector is just emerging from its 'pioneering' phase. As information technology improves and diversifies, users will increasingly be able to mix freely GIS, CADD (computer-aided design and drafting), image processing, and other spatial information technologies. On the one hand, the decade or two just ahead may see commercial GIS further emerge as a recognizable industry; on the other hand, the technology may become so pervasive that it 'disappears', becoming transparent to users in the same way as that of the telephone, the computer and computer graphics. At present, the former course seems the more likely.

The commercial sector of the GIS field is increasingly recognized as a major player in the field as a whole. It provides competitive and market mechanisms and creative forces that can be channelled to make great progress, if parochialism, protectionism, nationalism, and unfair forms of competition can be avoided, and open, global markets for GIS technology can be created.

REFERENCES

Clark D M, Hastings D A, Kineman J J (1991) Global databases and their implication for GIS. In: Maguire D J, Goodchild M F, Rhind D W (eds.) *Geographical Information Systems: principles and applications*. Longman, London, pp. 217–31, Vol 2

Clarke A L (1991) GIS specification, evaluation and implementation. In: Maguire D J, Goodchild M F, Rhind D W (eds.) *Geographical Information Systems: principles and applications*. Longman, London, pp. 477–88, Vol 1

Dangermond J, Harnden E (1990) *Map Data Standardization*. Environmental Systems Research Institute, Redlands California

Foster M J, Shand P J (eds.) (1990) *The Association for Geographic Information Yearbook 1990*. Taylor & Francis and Miles Arnold, London

GISWorld (1990) *GIS Technology '90: results of the 1990 GISWorld geographic information systems survey*. GISWorld, Fort Collins Colorado, 16 pp.

Goodchild M F (1991) The technological setting of GIS. In: Maguire D J, Goodchild M F, Rhind D W (eds.) *Geographical Information Systems: principles and applications*. Longman, London, pp. 45–54, Vol 1

Jackson M J, Woodsford P A (1991) GIS data capture hardware and software. In: Maguire D J, Goodchild M F, Rhind D W (eds.) *Geographical Information Systems: principles and applications*. Longman, London, pp. 239–49, Vol 1

Maguire D J (1991) An overview and definition of GIS. In: Maguire D J, Goodchild M F, Rhind D W (eds.) *Geographical Information Systems: principles and applications*. Longman, London, pp. 9–20, Vol 1

Shand P, Moore R (1989) *The Association for Geographic Information Yearbook 1989*. Taylor & Francis and Miles Arnold, London

Smith T R, Je Yiang (1991) Knowledge-based approaches in GIS. In: Maguire D J, Goodchild M F, Rhind D W (eds.) *Geographical Information Systems: principles and applications*. Longman, London, pp. 413–25, Vol 1

Taylor D R F (1991) GIS and developing nations. In: Maguire D J, Goodchild M F, Rhind D W (eds.) *Geographical Information Systems: principles and applications*. Longman, London, pp. 71–84, Vol 2

Unwin D J (1991) The academic setting of GIS. In: Maguire D J, Goodchild M F, Rhind D W (eds.) *Geographical Information Systems: principles and applications*. Longman, London, pp. 81–90, Vol 1

Walker T C, Miller R K (1990) *Geographic Information Systems – an assessment of technology, applications, and products*. SEAI Technical Publications, Madison Georgia USA

20

THREE-DIMENSIONAL GIS

J F RAPER AND B KELK

This chapter introduces and profiles three-dimensional (3-D) GIS which are differentiated from computer-aided design systems by the ability to represent complex geoscientific objects and apply volumetric spatial functions. Such 3-D GIS have grown rapidly to suit the needs of earth, atmospheric and ocean sciences, and are capable of using 2-D and 3-D spatially referenced data in a heterogeneous representation scheme. New forms of representation have emerged, based on 3-D vector and raster data structures, which can index spatial form and process, and support complex 3-D queries. In the future, the success of this new form of modelling depends on the quality of the model on which it is based, and the availability of 3-D data.

INTRODUCTION

For many GIS applications, a key assumption is that all spatial data handled are referenced to a 2-D Cartesian coordinate system. This convention restricts the scope for mapping the 'vertical' dimension over terrains or within the earth, oceans or atmosphere, since this third dimension must be converted to an attribute and expressed in 2-D as a line or zone with a constant value. Hence, following the considerable growth in GIS to meet the needs of digital mapping and of spatial database development (GIS World 1990), attention is now focusing on the design and implementation of 3-D GIS in a range of geoscientific application areas. Recent developments have been reported in the fields of oil exploration (Youngmann 1989), mining (Bak and Mill 1989), meteorology (Slingerland and Keen 1990), hydrogeology (Turner 1989), geological modelling (Kelk 1991), environmental monitoring (Smith and Paradis 1989), civil engineering (Petrie and Kennie 1990) and landscape architecture (Batten 1989).

Many of the recent developments referred to above have pioneered the use of new data structures to overcome the limitations of earlier (generally non-geoscientific) approaches to 3-D modelling. While many 3-D modelling systems have been developed for high quality computer-aided design (CAD) (Requicha 1980) and graphic rendering of solid objects (Pixar 1988), these systems have limitations for geoscientific applications. Many such graphic modelling systems can only generate high quality visualizations of the features under study (Salmon and Slater 1987), which cannot be analysed or interrogated, while CAD systems have limited facilities for the management of complex geo-objects. However, with the simultaneous improvement in the price/performance ratio for hardware, the rapid development of software for realistic graphic display, and recent developments in the theory of 3-D spatial data structuring, many geoscientific modelling problems have become increasingly tractable using these systems.

The primary impetus behind the rapid improvement of hardware performance has been the development of new processor architectures (Goodchild 1991 in this volume; Franklin 1991 in this volume), in particular the Reduced Instruction Set Chip (RISC). This has brought compute performance to a point where the very large number of individual elements in a 3-D model can be analysed. For example, Bak and Mill (1989) showed the development of a mine model with more than

800 000 nodes which, following the construction of an index to the model, could easily be interactively manipulated on a workstation.

In addition to the developments of general purpose hardware, graphics accelerator chips have been developed to boost the performance of colour rendition and the floating point operations required to carry out interactive transformations in 3-D space. Since hundreds of thousands of 3-D transformations per second are used in 3-D applications (Flynn 1990), each involving hundreds of floating point calculations, the performance required is equal to tens of MFlops (millions of floating point operations per second). General purpose hardware must be augmented with special graphics processors to make this procedure interactive. Flynn (1990) estimates that the computational requirements for fully realistic 3-D graphic animation are 200 000 times more demanding than 2-D static graphics. Other hardware developments include immediate mode graphics to avoid the limitations of display list techniques, additional device buffers to allow refreshment of the buffer as the screen is built, and fast screen clears. Parallel hardware architectures also offer substantial performance improvements, although this hardware has more specific applications and requires associated software engineering.

Software developments have also helped to create the conditions for an expansion in 3-D modelling. Firstly, the movement towards UNIX as an operating system for high performance workstations has helped the process of standardization and reinforced the market development of fast hardware. Secondly, the creation of high quality computer graphics systems has been a key development (Salmon and Slater 1987). The growth of computer graphics led initially to 2-D graphics standards such as the Graphical Kernal System (GKS). This 2-D system (which is accepted as an ISO standard – ISO 7942) provides a library of machine-independent operations on screen representations. The X-Windows system developed at MIT is also widely used as a window management system on workstations, but is currently limited to 2-D. Basic 3-D modelling systems have developed through the definitions of standards such as the Programmers Hierarchical Integrated Graphics System (PHIGS); while PHIGS and PHIGS+ support the management of 3-D objects, they are based on (slower) display list graphics. However, PHIGS and X-Windows may be merged to form the PHIGS-Extended-to-X or PEX standard by the early 1990s. Currently, the state-of-the-art 3-D modelling systems are based on standards defined by commercially led consortia such as Renderman (Pixar 1988) and Silicon Graphics/IBM (Flynn 1990).

Other key developments in the field of graphic display include the routines for the visualization of 3-D models using perspective, hidden surface removal and depth cueing (see Fig. 20.1), along with the transformations needed for model manipulation through interactive viewpoint change (McLaren and Kennie 1989). Software for the realistic rendering of colour and shading also now permits the use of 24-bit colour schemes with a palette of 16.8 million colours as well as anti-aliasing techniques and lighting effects.

Fig. 20.1 Surface model for terrain in the Telford area, England.

Finally, it is important to note that the methodology of 3-D modelling has developed separately in a variety of different fields. A significant role has been played by the development of CAD applications, such as the EMS software from Intergraph (Kelk and Challen 1989), but solid modelling has also been pioneered in cinematic animation, using systems such as Renderman (Pixar 1988). Other developments have originated in 'scientific visualization' within crystallography, high energy physics, medicine (Gargantine 1991) and fluid dynamics (Harig 1990) where images of objects such as body organs have been built using the well known characteristics of these objects. Solid modelling has also become strongly developed in architectural planning and landscape design (Turnbull, McAulay and McLaren 1990).

However, typical algorithms for the visualization of solids assume valid, spatially unique and unambiguous solids. This kind of representation is ideal for visualizing molecular structures, engineering parts or architecture since establishing the primitives, solid geometry or bounding edges is usually straightforward. However, the geometrical, structural and resolution complexity of geoscientific data sets usually makes this approach difficult to apply in the modelling of geology, geomorphology, ocean or atmosphere (see Plate 20.1). These data are not easily modelled using simple primitives or algorithms and have required extensions to spatial theory (Raper 1990; Frank & Buyong 1991).

This chapter is composed of four main sections following this introduction. The next section discusses the dimensionality of spatial data and this is followed by sections discussing the role of surfaces in 3-D modelling, solid modelling in the geosciences, and the process of 3-D model development.

DIMENSIONALITY OF SPATIAL DATA

Geoscientific spatial data can be represented in two clearly distinct Euclidean dimensional contexts:

- 2-D: a spatial object or region which is defined in 2-D space by measurements on axes x, y;
- 3-D: a spatial object or domain extending through 3-D space defined by axes x, y, z.

The use of a 2-D representation has generally been to delineate 'objects in the plane' or 'fields of observations' (Goodchild 1990), specifically the mapping of spatial pattern and extent. Types of 2-D spatial objects or fields and the operations which can be carried out on them are considered at length in texts such as Burrough (1986) and Aronoff (1989) and need no elaboration here. Typical examples of such spatial objects would be land parcels, coastlines or fire hydrants which are readily handled in a GIS with an x,y coordinate system. It should be noted that most commercially available GIS are only designed to handle 2-D spatial data: some systems achieve limited 3-D capabilities for surface modelling by assigning an attribute for z values (such as elevation) to a set of x,y locations. Here only the x,y locations are stored within the spatial indexing system and the z value is defined as a pseudo-attribute. This involves making the assumption that it is only necessary to store a single z value for each x,y location, that is, the surface defined is not overfolded (Weibel and Heller 1991 in this volume). In practice this has been an acceptable limitation on the representation of surfaces.

However, many forms of geoscientific analysis seek to collect data about spatial objects and domains such as features of the solid earth (e.g. aquifers), oceans (e.g. currents) or atmosphere (e.g. weather fronts), which fill or enclose 3-D space. A complete representation of these phenomena requires the definition of each location known within an x,y,z coordinate system. This fully 3-D system allows a direct analogy between the real space and a simulated space to be established in the model. However, it also requires 3-D forms of spatial indexing which are much more difficult to create and manage than the existing 2-D systems (Raper 1989a).

Note that such spatial objects or domains can be described by any number of attributes. However, these are always expressed as descriptions of x,y,z locations, since these are the cardinal spatial dimensions. Time can be considered a further dimension, but must be considered qualitatively different, as Hazelton, Leahy and Williamson (1990) show, since time is not generally measured in the same units. Typically, therefore, evolving systems are represented by a sequence of 3-D models, although Hazelton, Leaky and Williamson (1990) show how a 4-D representation of space-time can be constructed where all units are light-seconds. In general, 3-D geoscientific models are also based on precise (though not necessarily accurate) representations rather than fuzzy ones. Due to the difficulty of visualizing fuzzy 3-D models (x-ish, y-ish, z-ish coordinates), a series of models reflecting different estimates is usually constructed, although some fuzziness can be displayed by the use of transparency to blur the edges of geo-objects (Flynn 1990).

However, these two classes of spatial data *representation* (2-D, 3-D) are also associated with sets of operations which can be carried out on each set of data types supported. For example, volume cannot be established in all situations where solid objects are represented in 2-D, due to the inability of the representation to handle multiple *z* values for a single *x*,*y* location in the plane. At present 3-D spatial operations are poorly understood, as there are few implemented systems which work with 3-D representations, although considerable research into appropriate algorithms is presently under way. Hence, at present many applications are being moved to 3-D representations in order to exploit the availability of these new 3-D operations.

One of the major problems for the development of true 3-D models is that generally they must still be imaged on semi-flat cathode ray tube screens. Although Welch (1990) reports the use of a 3-D stereo visualization system based on 120 Hz circularly polarized images, the majority of 3-D models are still portrayed in 2-D. This has given rise to a variety of forms of *visualization* for 2-D and 3-D representations which are discussed below.

2-D visualization

A 2-D visualization is a graph or raster where the *z*-value defining a surface is projected on to a 2-D plane and 0-D, 1-D and 2-D objects can be displayed (Fig. 20.2). Since *z* is usually a continuously varying value on the ratio scale, the value of *z* is usually grouped into a class and the class boundaries shown. This can be achieved by shading *z*-value classes or labelling the isolines which divide them. This is the preferred technique for the display of a 3-D spatial object such as a terrain on a 2-D map. Multiple *z*-values for a given location cannot be handled: if this situation occurs in some locations the isolines are often simply omitted.

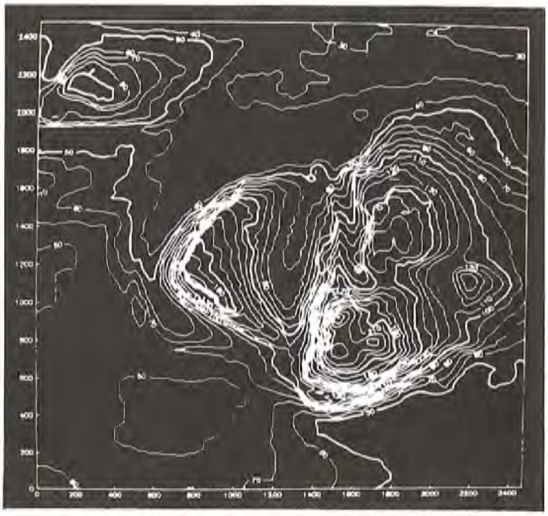


Fig. 20.2 A 2-D isoline map of Arthur's Seat, Edinburgh, Scotland.

2.5-D visualization

A 2.5-D visualization is an isometric model where the *z* attribute associated with an *x*,*y* location is projected onto an *x*,*y*,*z* coordinate reference system and all three axes displayed (Fig. 20.3). This operation transforms the map of *z* attributes for an *x*,*y* position, so that each *z* attribute defines a position on the *z* axis, creating a surface with no thickness visualized within 3-D space. This approach simulates the view that would be seen by a human observer from a point within the 3-D space. However, note that a 2.5-D *visualization* of a 2-D *representation* is only the display of a single-valued surface: multiple *z*-values cannot be handled for a single *x*,*y* point except by stacking a number of surfaces in the same 3-D space. This kind of visualization only contains information about the described surface, the state of a 3-D spatial object at a plane or interface or for a single observation cycle. Hence, a 2.5-D visualization is still limited by the basis of a 2-D representation.

3-D visualization

A true 3-D visualization is a full 3-D solid model where many *x*,*y*,*z* observations are structured into a solid structure and visualized in perspective view, complete with multiple occurrences of *z*. This kind of view is a precise analogue for the physical space inhabited by human observers, and allows the full specification of 3-D operations on the observed

Fig. 20.3 A 2.5-D isometric model of Arthur's Seat, Edinburgh, Scotland.

phenomena within the limits of the geometrical model employed (Plate 20.2).

Within contemporary geoscientific analysis only 2-D and 2.5-D *visualizations* of 3-D *representations* of reality are in common use, specifically, surfaces visualized in 3-D space. For such visualizations, 2.5-D techniques offer the opportunity to view the model, and are popular since they are calculated easily and match certain application requirements well. For example, surfaces can often be considered to be approximate analogues for sedimentary strata, since they frequently have a high area to thickness ratio, and the compression of the third dimension to zero may not cause the loss of significant information. Display of surfaces in 2.5-D may also be the appropriate visualization for interfaces such as water tables. However, these objects can only be handled where there are no multiple occurrences of z-values for any given x,y: in practice, therefore, 2.5-D visualization is limited to spatial objects with a planar continuity, although this assumption holds true for almost all terrains. It is also conventional for the z axis to be parallel to the action of gravitational forces, although this axis may be arbitrarily defined.

Note, however, that although several adjacent surfaces can be used to show multiple values of z for given locations in 2.5-D visualizations, this situation has several shortcomings for analysis. In particular, if the surfaces intersect they must be topologically connected, and it is still difficult to extract the true 3-D properties of solid spatial objects. Thus, to overcome these problems and to handle solid 3-D objects will require a 3-D visualization and a true 3-D representation scheme. This is particularly appropriate for the visualization of complex geo-objects where solid thresholding of internal property variance may be required, and where 3-D spatial operations are needed to characterize 3-D spatial objects.

SURFACE MODELLING IN 3-D

The creation of surfaces is a widespread form of modelling in the geosciences, and a wide range of application software is available to create 2.5-D visualizations. As discussed in the previous section the 2.5-D surface display can only be used to visualize a 3-D representation of reality within certain limits. For example, multiple z-values cannot be displayed. However, geometrically, surfaces may be thought of as complex spatial objects whose spatial configuration is not limited by the constraints of a 2.5-D visualization. Surfaces are developed by spatially structuring point- or line-based z-value data using raster (grid based) or vector (triangulated) techniques: therefore, surfaces can be used in a 3-D representation as a means of bounding 3-D space. This can be achieved by using limited 2.5-D visualization data or by developing new ways to represent surfaces in 3-D space. In this chapter the study of surfaces is confined to their use in 3-D modelling: the generation of 2.5-D surfaces and the storage of digital elevation models (DEMs) is considered extensively in Weibel and Heller (1991 in this volume).

While 2.5-D surface visualizations cannot handle multiple z-values, and can only partition space and not enclose it, they fulfil a useful role as basic building blocks or constraints in solid models. These links can be implied as in the case of plotting together multiple surfaces within a common spatial frame: this can allow the computation of simple solid characteristics such as the volume vertically between any two surfaces. These links can also be made by the geometrical connection of surfaces (Fig. 20.4): Christiansen and Sederberg (1978) described an algorithm for connecting isolines describing a surface together to form a solid.

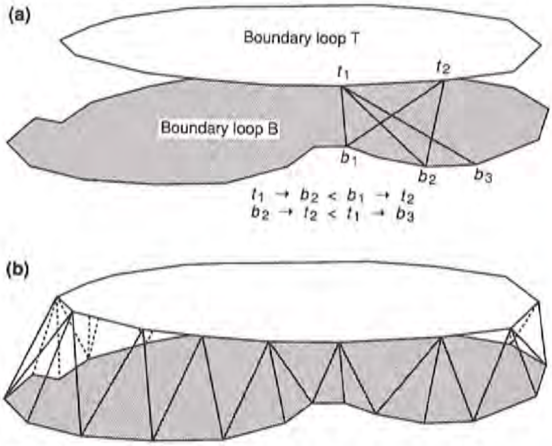


Fig. 20.4 An example of two planar, parallel isoline loops connected to form a triangulated solid: (a) the shortest span method of surface triangulation; (b) surface triangulation of two boundary layers.

However, the development of new forms of representation for surfaces has produced new approaches unconstrained by 2.5-D limitations. Such an approach described by Fisher and Wales (1991) has been the use of non-uniform rational B-splines (NURBS) to generate surfaces in hydrocarbon exploration. NURBS are piecewise parametric polynomial functions which can be used to model complex surfaces. These functions fit the data exactly, but create a smooth surface form, subject to local control by the manipulation of the poles associated with each segment of the surface (Fig. 20.5). NURBS can handle overfolds in a surface: however, solids must still be assembled from constituent surfaces. The procedure to model a geological surface described by Fisher and Wales involved picking the tops of geological formations (identified by the geologist) and interpolating 2-D isoline maps honouring all the data points using NURBS techniques. These isoline maps were then converted into a set of surface planes incorporating structural information such as faults, which were then connected to form a solid.

Kelk and Challen (1989) describe an application of the NURBS technique using Intergraph's EMS package to connect two limbs of an overfold in South Wales Coal Measures. In this case the use of NURBS allowed them to model the form of the fold connection from knowledge of the regional structure. The process was interactive allowing them to view four screen windows, one each for x,y and z perspectives and one 2.5-D view. Much of the work in this case was eased by rotating the surfaces to look along the strike of the folds when making the actual connection. Once connected, the complete new surface was incorporated into the structure of the two original folds (Fig. 20.6).

The value of surfaces in 3-D models is also demonstrated when they are used as estimators of general spatial process behaviour. Surfaces can be used to filter or transmit spatial processes: thus, one set of surface characteristics such as terrain, can be used to characterize other process surfaces such as runoff characteristics. 'Soft' data (indicative of behaviour) expressed in surface z_1 can be used as a template or predictor of specific local behaviour in surface z_2: surface z_1 can then be discarded. A more sophisticated procedure capable of relating complex behaviour of one surface to another is co-kriging (Leenaers, Burrough and Okx 1989; Burrough 1991 in this volume). When one variable is spatially dependent on another they are described as co-regionalized. By describing the form of the relationship and its variance it is possible to predict more accurately the value of the spatially dependent property at locations where it has not been measured. Leenaers, Burrough and Okx (1989) applied this technique to the modelling of lead deposition across a floodplain where the lead concentrations were expensive to sample but the floodplain elevations were easily surveyed. The co-regionalization of these two variables was found to be dependent on floodplain inundation frequency.

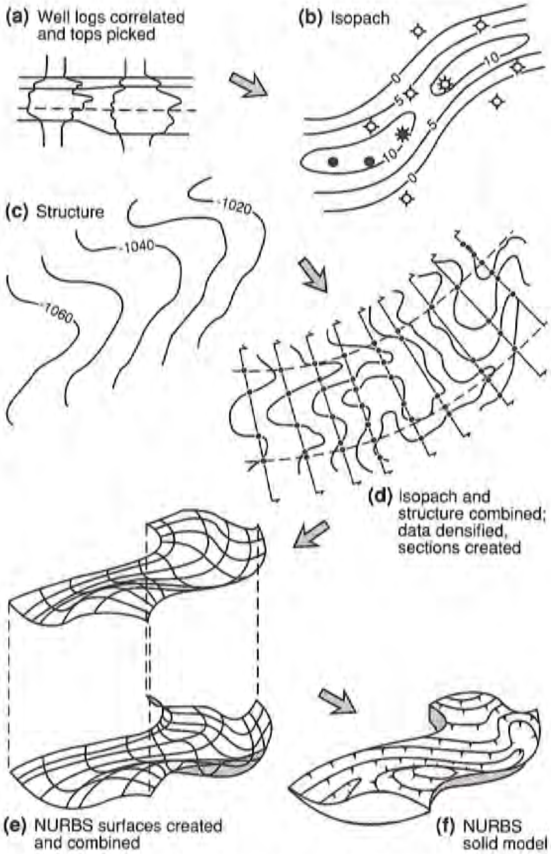


Fig. 20.5 The procedure for the creation of a NURBS solid model.

The development of surfaces can also be made completely interactive as described by Schaeben (1989) and Auerbach and Schaeben (1990). In this study, scattered points are triangulated under controlled conditions with the user placing the initial vertices of a Delaunay triangulation and inserting polylines which define the axis of any

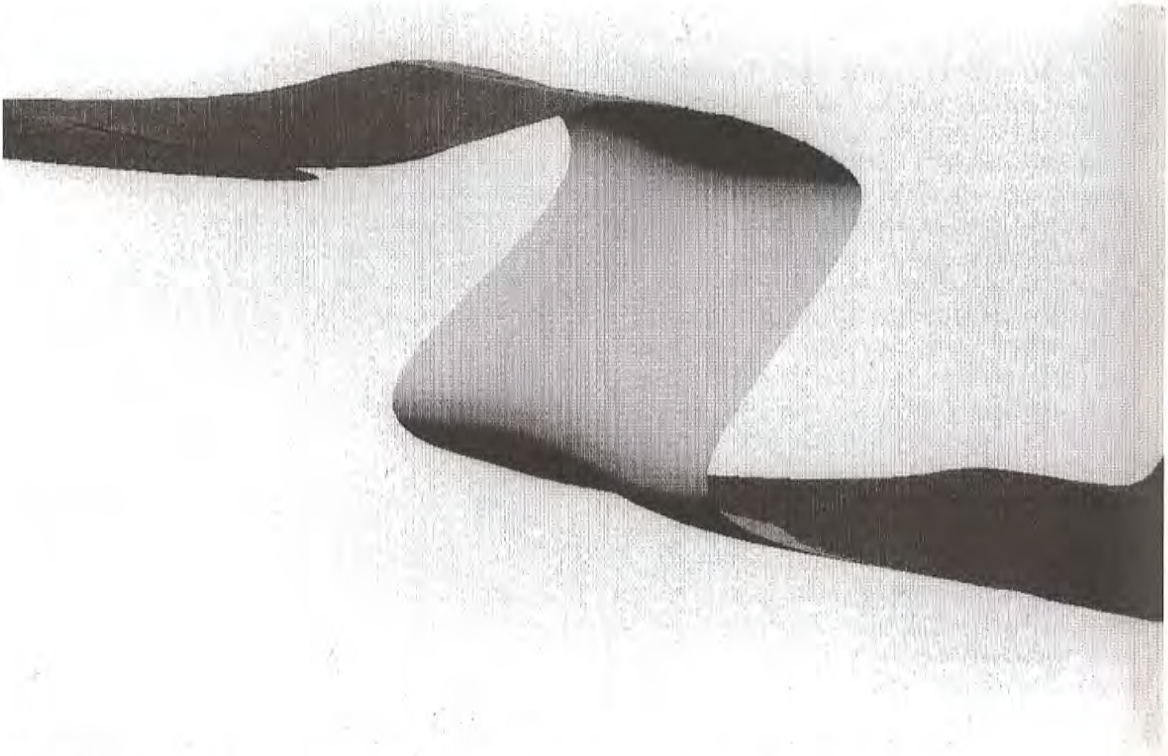


Fig. 20.6 A view of the two original limbs of the fold structure now connected by the overfold.

geological features influencing the structure of the study area. The surface is then constructed using bivariate quadratic simplicial B-splines to create a smooth surface, although directional constraints can be defined. In order to make the surface representation modifiable the coefficients of the splines are visualized as geometrical points which can be manipulated to control the shape of the surface. This procedure can be used to create multiple surfaces which are then fastened together to create a full solid model.

Interactive surface design is also available within the GOCAD system (Mallet 1991) which uses a triangulated irregular network (TIN) to model a surface. GOCAD allows the creation of triangles interactively in a CAD system or algorithmically by conversion of surfaces interpolated from a grid into a triangulated form. The GOCAD TIN model can handle overfolds since the triangulation is constructed by iteratively examining a 3-D grid of cells for the entrance and exit points of triangular facets defining the boundary conditions for the surface. GOCAD can also incorporate fuzzy constraints into the specification of nodes and vectors in the TIN. The user interface to the GOCAD system is based on the use of 'cameras' which view the 3-D domain from any angle: transformation of the bounding surfaces can take place within these windows.

One particular characteristic of geological surfaces is the likely presence of faults which may lead to 'steps' in surfaces. The steps may be associated with vertical and lateral movement which may juxtapose different materials across the fault. Hence when creating surfaces to model solid geology it is necessary to prevent interpolation across a fault and to control interpolation around the ends of faults where they may plunge below the surface. Rüber (1989) has developed a surface interpolation system to support such constraints based on a series of modules: BILDED to support data capture from digitizers and grid files; SEISCO to convert depth to seismic travel time for heterogeneous seismic and geological data sets; DFAULT to allow the interactive definition of fault patterns and plane angles; FLINT to interpolate the fault-constrained surface; and INSECT to calculate intersections between interpolated surfaces and fault planes.

Finally, it should be noted that surfaces can also be generated as the result of a spatial query on a solid when defined either in terms of form, orientation or position in the model. Extractions of surfaces from solids can be useful when attempting to locate an interface which is important as a boundary of some kind. Hence surfaces can be seen to be important geometrical constructs whether 2.5-D or full 3-D techniques are used.

SOLID MODELLING IN 3-D GIS

Much of the early experience of solid modelling has been gained in the computer-aided drawing (CAD) field as described by Requicha (1980) and Meier (1986) who identified several distinct groups of 3-D representation techniques:

1. *Sweep representation* (*SR*): the sweep technique represents an object by sweeping a defined area or volume along a defined trajectory.
2. *Primitive instancing* (*PI*): this represents an object by a set of pre-defined shapes, or mathematical primitives, which are positioned in 3-D space without intersection. An instance of a primitive is defined by a set of numeric values where each value is a parameter in the mathematical equation describing the primitive shape.
3. *Constructive solid geometry* (*CSG*): this technique represents an object by combining primitive point sets using Boolean operations (union, intersection and difference).
4. *Boundary representation* (*BR*): this technique defines an object by its bounding surface. The latter can be represented as a set of coordinates and their connectivity.
5. *Spatial occupancy enumeration* (*SOE*): this represents an object by the union of a set of cells where the cell is a primitive shape which can be either regular or irregular. Cells are adjacent, connected and do not intersect.

These representations form a set of techniques to create 3-D models: Fig. 20.7 shows a typology of the representations commonly used in the geosciences. The suitability of one or other of these representations depends on the characteristics of the data set, the operations which it is desirable to carry out and the specific form of spatial indexing employed.

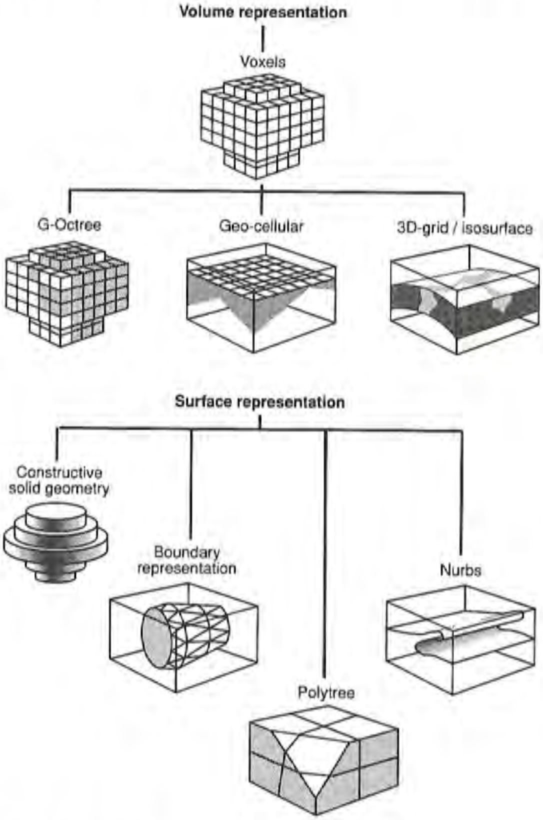


Fig. 20.7 3-D representations used in the geosciences.

Sampling

The first step in the solid modelling process is collection of appropriate measurements in the study domain consisting of attribute data for x,y,z locations which meet a specified set of criteria. The collection of these measurements is, however, a sampling exercise and in the subsequent modelling these measurements should be considered as the spatial representation of the probability that geo-object(s) exist in a particular location and configuration. Since many geo-objects have poorly known characteristics, the correct sampling frame (random or structured) and the sampling density can often not be determined in advance.

Often the type of sampling adopted is governed by economic or project-related, rather than geoscientific, criteria. For example, the sampling of ground conditions for a new road is often constrained to the centreline of the carriageway, and the automated estimation of surface heights in a photogrammetric survey generates a regular grid of values. Typical methods of obtaining geoscientific samples are seismic profiling, overflights, weather balloons, probes, systematic survey and boreholes. However, there are several drawbacks associated with point sampling of a geo-object using vertical sampling lines through a domain. First, it can be very difficult to distinguish between the multiple occurrence of a similar feature down a line and the repeated occurrence of a single object through folding or faulting. Secondly, the continuity between any two occurrences of the boundary of a geo-object taken from two neighbouring point samples cannot be assumed to be simple or predictable.

Another sampling difficulty associated with 3-D geo-objects arises in the case of dynamic phenomena, such as water movement in aquifers (Turner 1989) and oil flow through hydrocarbon reservoirs, or in evolutionary systems such as coastal sedimentary environments. It may be necessary to sample these data sets repeatedly over time which will generate multiple values of attributes for an x,y,z position. Burns (1988, 1990) described a system of 'structured vector fields' to manage x,y,z,t data (where t is time), and reported an application of the technique in the development of a geothermal production field northeast of San Francisco, California. The values of attributes in an x,y,z,t framework may also change over time according to positive process feedback. For example, the greater accumulations on a coastal salt marsh near a major creek (French 1989), or the drawdown of a piezometric surface near a well, may provide further sampling difficulties.

The results of the sampling process are crucial determinants of the forms of representation which are available. If the sampling procedure produced a regular grid of x,y,z point values or series of parallel lines, then using a regular tessellation of volume elements (voxels) involves no further operations. If the sampling has produced non-regular patterns of x,y,z points or polylines then some form of connection must be made among these elements, and a form of structuring must be selected. In this case objects must be formed from the points and are usually constructed using lines and polygons (to create volygons).

Representation and structuring

The approaches to 3-D representation and structuring of geo-objects can be categorized as raster, vector and function based: Jones (1989) gives a good summary of the main data structures available. The raster solutions to 3-D data representation are mostly based around the voxel as a basic unit. Optimization of this form of representation for complex geo-objects requires the use of spatial indexing systems to eliminate redundant storage when necessary. The simplest form of indexing for such data is a complete 3-D layer- and row-ordered raster, with each voxel stored explicitly and associated with attribute values. These models have been built by many oil exploration and reservoir management companies, for example by Shell (the MONARCH system) and by Exxon (the GEOSET system: Jones (1988)). These models can be displayed for any range of attribute values and offer high quality visualizations. However, their performance in queries and spatial analysis, such as connectivity tracing, can be poor.

These 3-D sets of voxels can be indexed effectively in a variety of ways, including 3-D run encoding (Mark and Cebrian 1986) where all the voxels are visited by, for example the Morton Order, although this technique can develop huge demands for storage. A more sophisticated technique is the 3-D equivalent of the quadtree called the octree (Kavouras and Masry 1987; Bak and Mill 1989) which recursively divides space into eight until any part of the subdivision is empty (outside the object) or full (inside the object), with the process continuing to a pre-determined level of resolution. The advantage of this form of indexing is the very efficient conduct of Boolean operations on geo-objects.

A modified version of the octree called the polytree has been proposed by Carlbom (1987) which identifies the logical content of each voxel, that is whether full or empty, or a vertex, edge or surface cell of the geo-object. This scheme has the advantage of a solid representation amenable to rapid Boolean operations for 3-D spatial operations, while also identifying the nature of the individual voxel, giving a pseudo object representation.

Various hybrid forms of representation utilizing both raster and vector concepts have recently been proposed which use or convert observations in *x*,*y*,*z* coordinate space to a 3-D raster and then form a vector-type representation from it. Smith and Paradis (1989) outlined an approach which uses a minimum tension interpolation algorithm to calculate the values of a 3-D phenomenon (sampled non-regularly) at the intersections of a gridded bounding block. The system then forms 'iso-surfaces' using a triangulated vector data structure (see Fig. 20.10), which are used to partition the 3-D grid by attribute value. Other design elements such as faulting and structural features with a known form (such as a salt dome) can also be added to the model. This design incorporates elements of both raster- and vector-based structuring, and the system can choose whichever structuring is most appropriate for a particular spatial query operation. This system forms the basis for the commercially available Interactive Volume Modelling (IVM) system from Dynamic Graphics.

Denver and Phillips (1990) show how known geological structure can be used to define bedding planes and faults, which can each be represented by 2.5-D grid mesh surfaces, where the cells in the mesh are shaded for geological composition. Composition shaded cells on the surface can then be used to define solid voxels which may fill each geological unit if desired. This system is the basis of the commercially available Stratamodel system. The advantage of these two hybrid systems is that they offer the facility to convert between space-bounding vector techniques and space-filling raster methods.

In the field of vector data structuring for geo-objects the systems used most commonly employ 3-D boundary representations ('B-reps') for the indexing of geometrical data. Attribute data can then be linked to this structure using an appropriate geo-relational system, although the processing overheads can be high (Molenaar 1990; Fritsch 1990). This kind of structure also requires the acceptance of planar enforcement as an organizing principle, that is, no domains or objects can overlap.

The simplest 3-D B-rep is a triangulated irregular network (TIN). Carlson (1987) described the theoretical basis for a 3-D TIN system based on simplicial complexes: a 0-simplex is equivalent to a point, 1-simplex to a line, 2-simplex to an area and 3-simplex to a volume. This structure is relatively easy to transform and visualize, but is very difficult to create. This is because 3-D geometry is poorly understood, and if the problem is reduced to the connection of 2-D nodes by projecting the points on to a plane, multiple values are likely. Assumptions necessary for 3-D triangulations include the need to ensure that volumes do not intersect except at edges, and that all the objects defined are tetrahedrons which are mutually disjunct.

Few, if any, systems using this full 3-D form of representation have been implemented. Most implemented systems fasten together surfaces defined in *x*,*y* or *x*,*z* space to avoid the problems of 3-D polygon structuring. The Lynx mine modelling system uses this technique to build 3-D objects by connecting geological sections to make volumetric 'components'. In order to define the 'components' each section is projected to a mid-plane, where the geological transitions are resolved. Note that Lynx also allows the creation of a voxel model for all the 3-D components identified oriented at any angle appropriate to the 3-D structure of the ore body (Clark, Houlding and Stoakes 1990).

Another 3-D system which uses a constructive framework for 3-D vector modelling is IREX (Lasseter 1990). IREX allows the display and inter-connection of boreholes and their lithological characteristics within a visualization composed of a 3-D cuboid. The system can also show segments of 2.5-D surface and associated faults along with seismic sections available within the study area. IREX offers an environment where geoscientists can assemble structures from diverse forms of data using CAD-like tools: however, the tools embrace many of the complex data types which the geoscientist uses.

Function-based representations are generally based on piecewise parametric polynomial functions, such as NURBS (Fisher and Wales 1991), which have robust properties. These functions can structure points or primitive geometrical forms into a single exact 3-D model by assembly of surface components. Transformations and analysis of this representation are rapid and efficient, however, as a representation it enforces a continuity of curvature between known points: this is an acceptable assumption for many, but not all, geoscientific applications.

The choice of data structures will constrain the scope of the modelling by defining the 3-D functionality available. It is common for the coordinates of a geo-object of interest to be determined and the visualization created each time a model is made, since true 3-D representation of geoscience data has not hitherto been possible. However, now that 3-D spatial structuring is becoming more common, the alternatives of re-computing a new model or storing a 3-D representation will arise. As processing power becomes cheaper per unit, the choice between the alternatives of re-computing and full 3-D representation and structuring of the data will be determined by the application involved and the storage space available. Thus, representation is likely to become the rule when the resources consumed in the model creation process justify the retention of the model and its addition to a 3-D database.

Spatial functions

The keys to efficient access and use of 3-D geoscientific data, however structured, are the spatial functions or queries supported. It is possible to define sets of 3-D spatial functions in relation to several different systems of user conceptualization and system representation. Previous work on the definition of spatial functions has originated in diverse fields. Work in cognitive science by Johnson (1987), discussed by Mark (1989), has led to the definition of a class of spatial functions based around human metaphors for space known as 'image schemata'. Other work in linguistics (such as Talmy 1988) has emphasized the role of spatial prepositions, topology, viewpoint and distribution of attention as expressed through the structure of language used in spatial description. These approaches have confirmed the role of perception and understanding in the conceptualization of spatial relations in 2-D and 3-D, and hence the framing of spatial queries by the user of the system.

Work on spatial representation in 2-D has generated a more formal range of spatial functions, such as simple visualization properties (Freeman 1975), and complete sets of topological relationships between objects (Pullar and Egenhofer 1988). A set of functions for 3-D modelling of geoscientific data was proposed by Raper (1990) to formalize spatial functions without defining representation. Figure 20.8 shows a series of simple illustrations of the operation of these functions on a single object, and Table 20.1 gives an indication of the relative performance of raster and vector data structures across all these functions.

Table 20.1 shows in outline terms how raster and vector spatial structuring affects the speed of operation of the 3-D spatial functions. The relative merits of integral structuring and re-computation of the model also vary according to the type of data structure used. Note also that the accuracy of the representation must not exceed the real world determinacy of the data set. It is, therefore, suggested that a complete analysis of the available data and expected queries is the optimum way to decide on the type of representation which is appropriate for each model. The functionality of existing commercial systems can also be evaluated using this approach. Few if any of these 'solid' functions have yet been fully implemented for real geoscience data sets with all their inherent complexity, and so only the theoretical performance of these queries is considered in the establishment of Table 20.1 (although see Bak and Mill 1989 for some estimates of octree performance in spatial functions).

MODEL DEVELOPMENT IN 3-D

The application of these techniques of representation is crucially dependent on the identification or identity of the geo-objects of interest. While in many modelling environments the data and their spatial configuration are defined by external factors or customary practice, the process of conceptualization which precedes the representation of a model is a crucial stage in 3-D model development (Goodchild 1991). Raper (1990) suggested that the process of discretization of the perceived reality leads to the creation of a 'conceptual set' in 3-D: the structuring of the set of tuples x,y,z,p (where p is an attribute at time t)

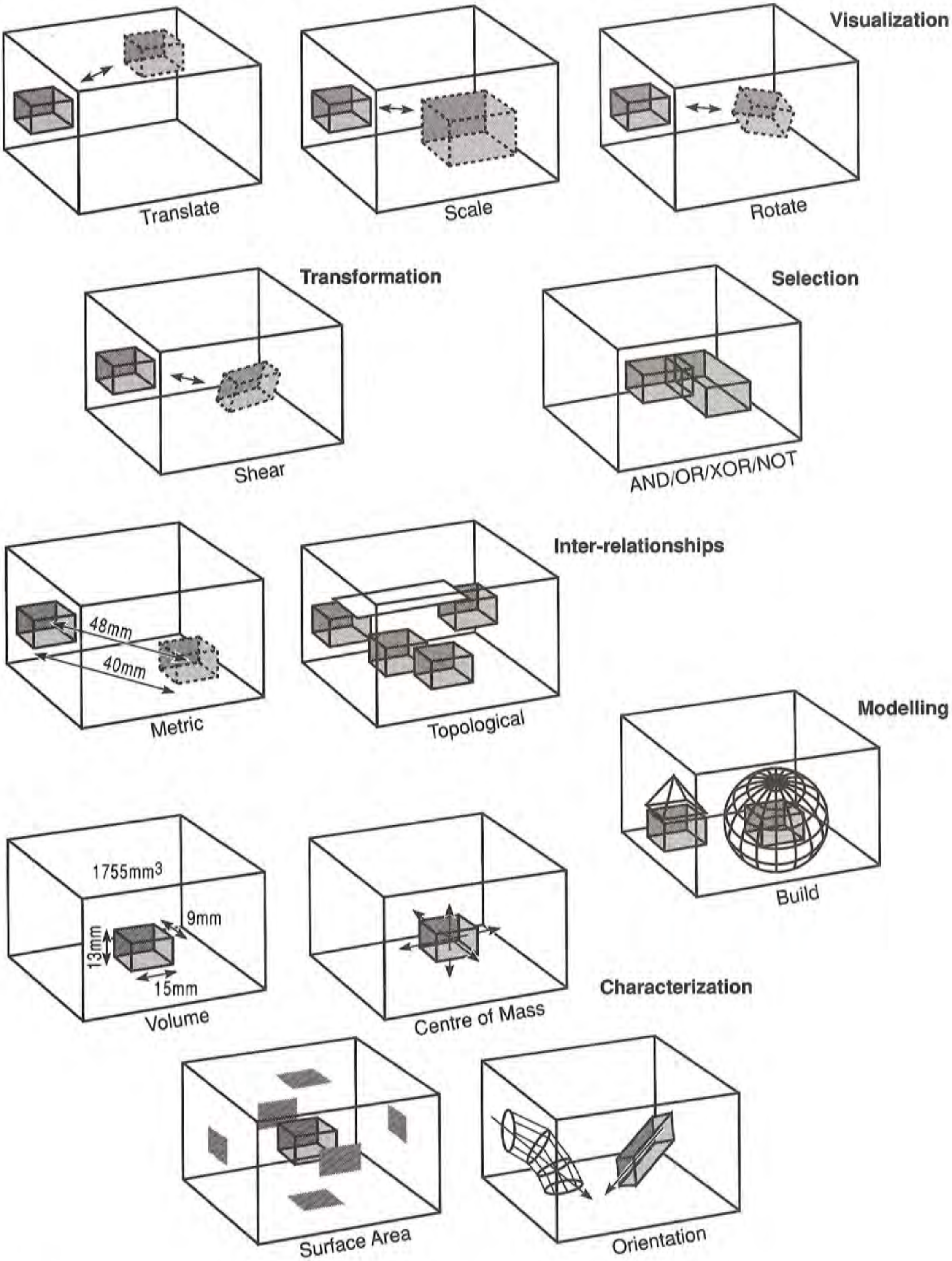


Fig. 20.8 Key 3-D spatial functions illustrated for a simple surface representation.

Table 20.1 Generic spatial functions in 3-D spatial modelling and their effectiveness under different 3-D data structures.

Spatial Function	Raster	Vector
Visualization		
Translate	slower	fast
Rotate	slower	fast
Scale	slower	fast
Reflect	slower	fast
Transformation		
Shear	slower	fast
Characterization		
Volume	fast	slower
Surface area	fast	fast
Centre of mass	fast	slower
Orientation	slow	fast
Selection		
AND	fast	slower
OR	fast	slower
XOR	fast	slower
NOT	fast	slower
Topological relations		
Separation	fast	fast
Adjacency	fast	fast
Modelling		
Build (axial or bounding)	fast	slow

associated with the conceptual set has a vital role in the representation choices made later.

The process of discretization (or data modelling) begins with the identification of structure in reality and is deeply embedded in practice and 'experience'. This step is normally followed by measurement of a number of parameters associated with the 'structure'. The observations on the 'suggestive' parameters are usually made with reference to a 3-D coordinate system which acts as a geometrical frame for the located tuples x,y,z,p. Perhaps two main kinds of approaches can be identified:

1. *Device exploratory*: in this case the measurement technology defines the geometrical arrangement of the observed tuples, and there is no search for an *a priori* object. Thus, boreholes impose a linear structure on the measurements, characterizing a regular sequence of points downhole, and a photogrammetric survey of a terrain or section will usually generate a grid of measurements over the visual field. However, a survey of the positions of tracer pebbles over a channel bar will be governed by hydraulic processes and recovery factors, and will generate a spatially non-regular set of observed tuples. In this process the tuples recorded only need have the means of collection or selection in common.
2. *Object exploratory*: in this case the search for an *a priori* object defines the geometrical arrangement of the observed tuples. The located tuples identified by the combination of 'suggestive' parameters form a spatial cluster of arbitrary configuration. In this process the tuples recorded may have a distinct spatial structure.

When the 'object' can be identified by a single key parameter and is known to exist in a discrete form from knowledge of the domain (e.g. mine access, destructive examination) then the spatial object can be termed 'sampling limited' (Raper 1989b). In this case the discretization proceeds by using the sampling theorem to create a parsimonious description of the object from selective observations. An example would be a perched aquifer, or a salt dome or a fault-limited block (Fig. 20.9).

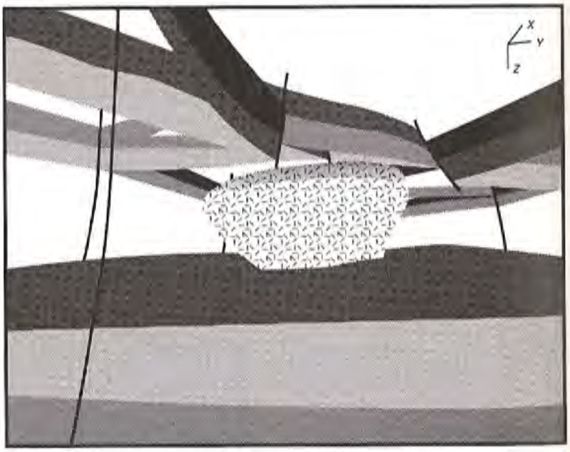


Fig. 20.9 A sampling limited geo-object: a salt dome and associated faulting modelled with IREX.

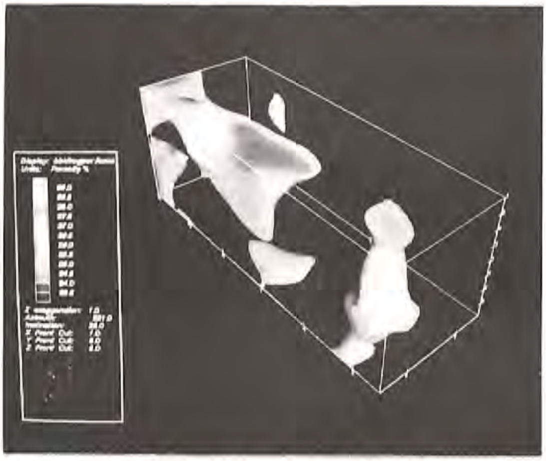


Fig. 20.10 Definition-limited objects: an IVM model of high porosity zones in a destructively tested limestone block.

When the object is transient or part of a continuum (e.g. a temperature field), or exists only as a spatially clustered set of observed tuples defined by a group of 'suggestive' parameters, then the spatial object can be described as 'definition limited (Raper 1989b). In this case the discretization proceeds by assembling a set of 'suggestive' parameters and searching for locations matching this *a priori* description. An example of a definition-limited object would be a plume of pollutants in the atmosphere or ocean defined by a physical threshold, or a sedimentary facies. Often the boundaries or conditions for selection of objects are 'picks' made on the basis of an interpretation of raw data (Fig. 20.10).

Although the set of tuples selected during the data collection stage can be structured in a variety of ways, there are probably only two major strategies:

1. *Domain partition*: use of the selected tuples to subdivide the whole 3-D domain into regular or non-regular constituent units.
2. *Entity construction*: use of the selected tuples to define entities within 3-D space using basic geometrical elements.

These two structuring strategies control the set of spatial functions available for the analysis of a model: they do not necessarily define the form of representation that should be chosen. In the main, however, domain partition leads to a raster type representation and entity construction leads to the use of vector or function type representations.

With the relative paucity of information often available to carry out such definition limited 3-D spatial modelling, the importance of the conceptual data model used in the analysis increases. This is because the geo-object itself is usually defined by the sampling or selection of parameters established by the data model. This data modelling process ought to be supported by knowledge about the geoscientific domain and associated processes (Raper 1988). For example, Kelk (1991) identified a number of characteristic geological data constraints which provided support for this exercise:

- Geological discontinuities (e.g. fault types)
- Regional structure (e.g. dip/strike and fold patterns)
- Sedimentation (e.g. fan systems)
- Volcanic (e.g. plutonic history)
- Process environment (e.g. glacial weathering).

The use of vertical profile models in sedimentary geology by such workers as Bouma and Allen has demonstrated that there is significant information content within sedimentary sequences (Turner 1989). This may include Markovian models of succession and cycling which reflect the behaviour of the physical system. This knowledge can be used to establish a data model for a sparse data set by providing guidance on the spatial dimensions or nature of a particular geo-object: one example would be the Bridge and Leeder (1979) model of meander belt migration behaviour. Conversely, by the use of simulation using Markov techniques, definition-limited models can be created under known conditions. This procedure can be used to form comparisons with other geo-objects which are established by the structuring of sparse data.

Specifically, the data model for a stratum believed to exist in the subsurface may be defined by lithological and structural parameters in a specific combination. The 3-D spatial identity of the geo-object is then established by searching the population of selected characteristics for the boundaries of the defining conditions and recording the x,y,z coordinates. Note, therefore, that by altering the contents of the data model and iterating the search process, a new set of x,y,z coordinates defining the object can be created. In the case of a study attempting to establish the overall architecture of a sedimentary sequence, the basic spatial arrangement of geo-objects can often be defined in different ways (and may overlap), depending on the contents of the data model for each element of the sequence defined (Raper 1988). Reconstruction of the structural development of basins, or dynamic behaviour of current systems or fluid flow regimes illustrates other evolutionary systems where the data model will change rapidly.

It is clear, therefore, that establishing the spatial identity of definition limited geo-objects in the subsurface is highly sensitive to the contents of the data model. The essential point is that the errors or bias inherent in the process of defining this data model can be as great as, or even greater than, those introduced in the spatial sampling of the parameters defining the geo-object or in the process of its visualization. Finally, it may be necessary to edit the data to select the values to be used in the 3-D spatial modelling. For example, it may be necessary to parse or validate the raw data, and subsequently to parameterize or regionalize the values before further analysis. These operations may impose their own indeterminacy on the identity of the geo-object under consideration.

Using the techniques and approaches described above it is possible to generate complex and realistic 3-D models visualized in 3-D. However, a key component in the modelling process is the spatial database and its design at the level of data types permitted. Any modelling process must also operate within the constraints of a database environment (see Fig. 20.11). Few specific database designs for the 3-D environment have been developed: Molenaar (1990), Fritsch (1990) and Hazelton, Leahy and Williams (1990) are the most complete attempts to date, all using relational concepts. One particular problem in a 3-D environment is the lineage of the model development, and the management of model versions (Newell, Theriault and Easterfield 1990). In a modelling environment it is also important to be able to incorporate data from different database sources (Turner, Kolm and Downey 1990), although few prospects for structured data transfer exist at present.

One well developed strategy for creating a 3-D model database involves spatial clustering of the elements defining the geo-object by a geometrical attribute in a geoscience database. Work by Schek and Waterfeld (1986) and Horn *et al.* (1989) has led to the establishment of a prototype database (the 'Geokernel'). This implements the storage of 3-D vector data items which are fully integrated with the associated attribute data. The storage structure used is based on hierarchical subdivision of the primary data attributes into sub-objects. The 3-D geometry is represented within this structure as a set of points, polylines or cuboids. These are spatially clustered by a regular or hierarchical subdivision of the 3-D space enclosing the objects. The usual mode of access would be by specifying (in system terminology) a 'clip' or 'compose' spatial query followed by a 'test' for the relevant non-spatial attributes. The chief advantage of this system is that it binds together the spatial and non-spatial components in a form of object orientation. It also allows the database to be set up for any data configuration. There are two main disadvantages. First, it requires a complex database operation to set up and bind each new data description to the database. Secondly, it does not use an explicit data structure to store 3-D geometry, but subdivides space around the objects identified in each project.

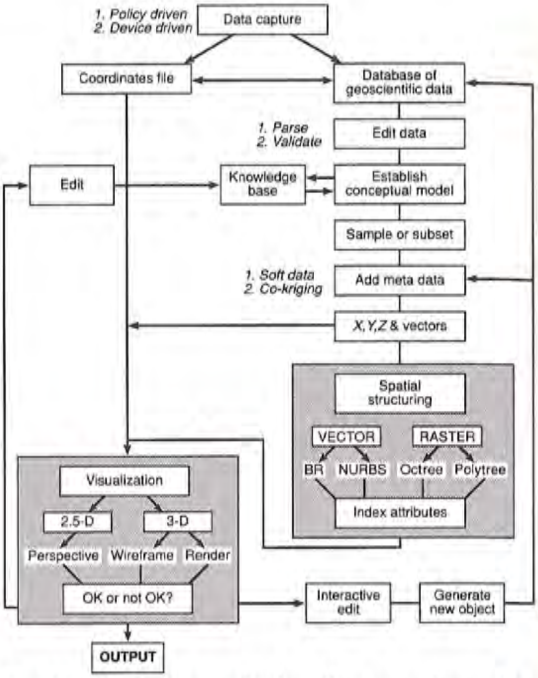


Fig. 20.11 A strategy for handling geoscience data and database input/output.

Three-dimensional visualizations cannot often be interactively interrogated. Hence, one ultimate design objective in the development of a 3-D modelling system is the production of high quality images which can also be interrogated by graphic interaction with the model on the screen, perhaps using a 3-D cursor such as the 'Dataglove' demonstrated by the Media Lab at MIT (Aldersey-Williams 1989). The rapid development of 'virtual reality' may also provide tools for interactive experiential manipulations of the model.

Dynamic 3-D models are also becoming a reality within the geoscience field. Tetzlaff and Harbaugh (1989) have recently described how clastic sedimentation can be simulated for a variety of processes, including river channel, deltaic and alluvial fan environments. These models are made in the SEDSIM environment which simulates clastic transport and sedimentation on a grain-by-grain basis. SEDSIM produces 3-D fence diagrams, shaded for grain size, which illustrate the geometrical outcome for a specific sedimentological model.

Experience with 3-D modelling also opens a new opportunity to accumulate knowledge of 3-D object metrics and behaviour which can be placed in knowledge bases. Dawson (1989) describes the development and structure of such a knowledge base at BP referred to as SEDMAC. Although this knowledge base is non-spatial in form, the potential to link to 3-D models exists. Finally, it should be noted that 3-D data can also be input into process driven models of 3-D behaviour, such as those used in hydrogeology (Turner 1989). The scope for this type of application is increasing as more systems are developed.

CONCLUSIONS

This chapter has reviewed and interpreted the developments within the field of 3-D GIS at a time of particularly rapid development in theory and applications. This rapid growth indicates that the development of 2-D GIS has not provided the techniques of spatial data handling required by geoscientists. As a consequence, further developments in this field can be anticipated. The great advantages of the new 3-D modelling techniques are clear for all geoscientists to see: many existing geoscience questions can be more effectively answered in a 3-D modelling domain and since completely new questions can be posed, this form of analysis will be continuously extended.

However, the key control over the long-term development and refinement of tools for 3-D GIS will undoubtedly be the ability of the geoscientist to respond to the challenge to produce more *complete* models of geoscientific processes and their physical expression. This response is called for since these new modelling techniques require a greater quantification of assumptions and more complete data to establish geometrical form. In this respect, 3-D GIS has the potential to fuel a revolution in geoscientific data handling which will increase the speed at which models can be made and improved. These developments also form part of the revolution in scientific visualization whose developments will greatly improve the geoscientists' ability to communicate with their peers and the public.

One final consideration for 3-D GIS is the issue of how to manage the output of these new modelling processes, that is how to organize the database of models. Past and present practice in the geosciences has been to enshrine the results of a particular study or model in a paper map, perhaps as part of a definitive series. However, within a 3-D GIS database, models can be continuously replaced by new models, each of which may be capable of many visualizations. This suggests that a major challenge for the management of this powerful new technology is how to store and make available for future use the new insights which will emerge, in particular, how to ensure that the data and their interpretation are archived together.

REFERENCES

Aldersey-Williams H (1989) A Bauhaus for the media age. *New Scientist* **1655**: 54–60
Aronoff S (1989) *Geographic Information Systems: a management perspective*. WDL Publications, Ottawa
Auerbach S, Schaeben H (1990) Computer-aided geometric design of geologic surfaces and bodies. *Mathematical Geology* **22**: 723–42

Bak P, Mill A (1989) Three dimensional representation in a Geoscientific Resource Management System for the minerals industry. In: Raper J F (ed.) *Three Dimensional*

Applications in Geographical Information Systems. Taylor & Francis, London

Batten L G (1989) National capital urban planning project: development of a 3-D GIS. *Proceedings of GIS/LIS '89*. ACSM/ASPRS, Falls Church Virginia, pp. 781–6

Bridge J S, Leeder M R (1979) A simulation model of alluvial stratigraphy. *Sedimentology* **26**: 617–44

Burns K L (1988) Lithologic topology and structural vector fields applied to subsurface prediction in geology. *Proceedings of GIS/LIS '88*. ACSM/ASPRS, Falls Church Virginia, pp. 26–34

Burns K L (1990) Three dimensional modelling and geothermal process simulation. *Proceedings of Symposium on Three Dimensional Computer Graphics in Modelling Geologic Structures and Simulating Processes*. Freiburger Geowissenschafliche Beitrage **2**: 10–12

Burrough P A (1986) *Principles of Geographical Information Systems for Land Resources Assessment*. Clarendon, Oxford

Burrough P A (1991) Soil information systems. In: Maguire D J, Goodchild M F, Rhind D W (eds.) *Geographical Information Systems: principles and applications*. Longman, London, pp. 153–69, Vol 2

Carlbom I (1987) An algorithm for geometric set operations using cellular subdivision techniques. *IEEE Computer Graphics and Applications*. May: 45–55

Carlson E (1987) Three dimensional conceptual modelling of subsurface structures. *Proceedings of AUTOCARTO 8*. ACSM/ASPRS, Falls Church Virginia, pp. 336–45

Christiansen H N, Sederberg T W (1978) Conversion of complex contour line definitions into polygonal element mosaics. *ACM Computer Graphics* **12** (3): 187–92

Clark I, Houlding S, Stoakes M (1990) Direct geostatistical estimation of irregular 3-D volumes. *Proceedings of Symposium on Three Dimensional Computer Graphics in Modelling Geologic Structures and Simulating Processes*. Freiburger Geowissenschafliche Beitrage **2**: 13–15

Dawson, M (1989) Developing a sedimentological database on an Apple Macintosh II. *BP International Information Systems Services paper*

Denver L E, Phillips D C (1990) Stratigraphic geocellular modelling. *Geobyte* February: 45–7

Fisher T, Wales R Q (1991) 3-D solid modelling of geo-objects using non-uniform rational B-splines (NURBS). In: Turner A K (ed.) *Three Dimensional Modelling with Geoscientific Information Systems*. Kluwer, Dordrecht

Flynn J J (1990) 3-D computing geosciences update. *Geobyte* February: 33–5

Frank A U, Buyong T B (1991) Geometry for 3D GIS in geoscientific applications. In: Turner A K (ed.) *Three Dimensional Modelling with Geoscientific Information Systems*. Kluver, Dordrecht

Franklin Wm R (1991) Computer systems and low-level data structures for GIS. In: Maguire D J, Goodchild M F, Rhind D W (eds.) *Geographical Information Systems: principles and applications*. Longman, London, pp. 215–25, Vol 1

Freeman J (1975) The modelling of spatial relations. *Computer Graphics and Image Processing* **4**: 156–71

French J R (1989) Hydrodynamics and sedimentation in a macrotidal salt marsh, north Norfolk, England. Unpublished PhD Thesis, University of Cambridge, England

Fritsch D (1990) Towards three dimensional data structures in geographic information systems. *Proceedings of the EGIS '90*, pp. 335–45

Gargantini I (1991) Modelling natural objects via octrees. In: Turner A K (ed.) *Three Dimensional Modelling with Geoscientific Information Systems*. Kluwer, Dortrecht

GIS World (1990) *GIS Software Survey 1990*. GIS World Inc., Fort Collins Colorado

Goodchild M F (1991) Geographical data modelling. *Computers and Geosciences* **17**

Goodchild M F (1991) The technological setting of GIS. In: Maguire D J, Goodchild M F, Rhind D W (eds.) *Geographical Information Systems: principles and applications*. Longman, London, pp. 45–54, Vol 1

Harig J (1990) Visualisation of 3-D finite element solutions of Navier–Stokes equations. *Proceedings of Symposium on Three Dimensional Computer Graphics in Modelling Geologic Structures and Simulating Processes*. Freiburger Geowissenschafliche Beitrage **2**: 36–8

Hazelton N W J, Leahy F J, Williamson I P (1990) On the design of a temporally-referenced, 3-D Geographical Information System: development of a four dimensional GIS. *Proceedings of GIS/LIS '90*. AAG/ACSM/AMFM/ASPRS/URISA, Bethesda Maryland, pp. 357–372

Horn D *et al.* (1989) Spatial access paths and physical clustering in a low level geo-database system. *Geologisches Jahrbuch* **A 104** (Construction and display of geoscientific maps derived from databases)

Johnson M (1987) *The Body in the Mind: the bodily basis of meaning and reason*. University of Chicago Press, Chicago

Jones C B (1989) Data structures for 3-D spatial information systems. *International Journal of Geographical Information Systems* **3**: 15–32

Jones T A (1988) Modeling geology in 3 dimensions. *Geobyte* February: 14–20

Kavouras M, Masry S (1987) An information system for geosciences: design considerations. *Proceedings of AUTOCARTO 8*. ACSM/ASPRS, Falls Church Virginia, pp. 336–45

Kelk B (1991) 3-D GIS for the geosciences. *Computers and Geosciences* **17**

Kelk B, Challen K (1989) Experiments with a CAD package for spatial modelling of geoscientific data. *International Colloquium on 'Digital maps in the Geosciences', Würzburg Germany*

Lasseter T (1990) An interactive 3-D modelling system for integrated interpretation in hydrocarbon reservoir exploration and production. *Proceedings of Symposium on Three Dimensional Computer Graphics in Modelling Geologic Structures and Simulating Processes*. Freiburger Geowissenschafliche Beitrage **2**: 45–6

Leenaers H, Burrough P A, Okx J (1989) Efficient mapping of heavy metal pollution on floodplains by co-kriging from elevation data. In: Raper J F (ed.) *Three*

Dimensional Applications in Geographical Information Systems. Taylor & Francis, London, pp. 37–50

Mallet J-L (1991) GOCAD: a computer-aided design program for geological applications. In: Turner A K (ed.) *Three Dimensional Modelling with Geoscientific Information*. Kluwer, Dortrecht

McLaren R A, Kennie T (1989) Visualisation of digital terrain models: techniques and applications. In: Raper J F (ed.) *Three Dimensional Applications in Geographical Information Systems*. Taylor & Francis, London, pp. 79–98

Mark D M (1989) Cognitive image schemata for geographic information: relations to user views and GIS interfaces. *Proceedings of GIS/LIS '89*. ACSM/ASPRS, Falls Church Virginia, pp. 551–60

Mark D M, Cebrian J A (1986) Octrees: a useful method for the processing of topographic and subsurface data. *Proceedings of ACSM–ASPRS Annual Convention*, Volume 1. ACSM/ASPRS, Falls Church Virginia, pp. 104–113

Meier A (1986) Applying relational database techniques to solid modelling. *Computer Aided Design* **18**: 319–26

Molenaar M (1990) A formal data structure for three dimensional vector maps. *Proceedings of EGIS '90*, pp. 770–81

Newell R, Theriault D, Easterfield M (1990) Temporal GIS – modelling the evolution of spatial data in time. *Proceedings of GIS Design Models Conference*. Midlands Regional Research Laboratory, Leicester

Petrie G, Kennie T (eds.) (1990) *Terrain Modelling in Surveying and Civil Engineering*. Whittles, Latheronwheel

Pixar Inc. (1988) *Renderman Interface*. Pixar, San Rafael California

Pullar D, Egenhofer M J (1988) Towards formal definitions of topological relations amongst spatial objects. *Proceedings of the 3rd International Symposium on Spatial Data Handling*. International Geographical Union, Columbus Ohio, pp. 225–42

Raper J F (1988) A methodology for the investigation of landform-sediment relationships in British glaciated valleys. Unpublished PhD Thesis Queen Mary College, University of London

Raper J F (ed.) (1989a) *Three Dimensional Applications in Geographical Information Systems*. Taylor & Francis, London

Raper J F (1989b) The geoscientific mapping and modelling system: a conceptual design. In: Raper J F (ed.) *Three Dimensional Applications in Geographical Information Systems*. Taylor & Francis, London, pp. 11–20

Raper J F (1990) An atlas of 3-D functions. *Proceedings of Symposium on Three Dimensional Computer Graphics in Modelling Geologic Structures and Simulating Processes*. Freiburger Geowissenschafliche Beitrage **2**: 74–5

Requicha A A G (1980) Representations for rigid solids: theory, methods, and systems. *ACM Computing Surveys* **12** (4): 437–64

Rüber O (1989) Interactive design of faulted geological surfaces. *Geologisches Jahrbuch* **A 104** (Construction and display of geoscientific maps derived from databases)

Salmon R, Slater M (1987) *Computer Graphics*. Addison-Wesley, Reading Massachusetts

Schaeben H (1989) Improving the geological significance of computed surfaces by CADG methods, *Geologisches Jahrbuch* **A 104** (Construction and display of geoscientific maps derived from databases)

Schek H-J, Waterfeld W (1986) A database kernel system for geoscientific applications. *Proceedings of the 2nd International Symposium on Spatial Data Handling*. International Geographical Union, Columbus Ohio, pp. 273–88

Slingerland R, Keen T R (1990) A numerical study of storm driven circulation and 'event bed' genesis. *Proceedings of Symposium on Three Dimensional Computer Graphics in Modelling Geologic Structures and Simulating Processes*. Freiburger Geowissenschafliche Beitrage **2**: 97–9

Smith D R, Paradis A R (1989) Three-dimensional GIS for the earth sciences. In: Raper J F (ed.) *Three Dimensional Applications in Geographical Information Systems*. Taylor & Francis, London, pp. 149–54

Talmy L (1988) How language structures space. In: Mark D M (ed.) *Cognitive and linguistic aspects of geographical space*. National Center for Geographic Information and Analysis Publication. NCGIA, Santa Barbara California

Tetzloff D M, Harbaugh J W (1989) *Simulating Plastic Sedimentation*. Van Nostrand Reinhold, New York

Turnbull M, McAulay I, McLaren R A (1990) The role of terrain modelling in computer aided landscape design. In: Petrie G, Kennie T (eds.) *Terrain Modelling in Surveying and Civil Engineering*. Whittles, Latheronwheel, pp. 262–75

Turner A K (1989) The role of 3-D GIS in subsurface characterisation for hydrogeological applications. In: Raper J F (ed.) *Three Dimensional Applications in Geographical Information Systems*. Taylor & Francis, London, pp. 115–28

Turner A K, Kolm K, Downey J (1990) Potential applications of geoscientific information systems (GSIS) for regional ground water flow systems. *Proceedings of Symposium on Three Dimensional Computer Graphics in Modelling Geologic Structures and Simulating Processes*. Freiburger Geowissenschafliche Beitrage **2**: 108–10

Weibel R, Heller M (1991) Digital terrain modelling. In: Maguire D J, Goodchild M F, Rhind D W (eds.) *Geographical Information Systems: principles and applications*. Longman, London, 269–97, Vol 1

Welch R A (1990) 3-D terrain modelling for GIS applications. *GIS World*, October/November: 26–30

Youngmann C (1989) Spatial data structures for modelling subsurface features. In: Raper J F (ed.) *Three Dimensional Applications in Geographical Information Systems*. Taylor & Francis, London, pp. 129–36

22

INFORMATION INTEGRATION AND GIS

I D H SHEPHERD

The role of the GIS as an 'information integrator' is examined, and the various approaches taken in current systems to achieve information integration are described. It is shown that considerable effort must be expended to create a consistent geographical database before GIS can successfully integrate diverse information. Alternative approaches to information integration found in information technology are also discussed, in order to identify more effective approaches to integration within a GIS framework.

INFORMATION INTEGRATION AND GIS

> 'The benefits of a geographical information system depend on linking different data sets together.' (DoE 1987: 2).

> 'A GIS brings information together, it unifies and integrates that information. It makes available information to which no one had access before, and places old information in a new context. It often brings together information which either was not or could not be brought together previously.'
> (Dangermond 1989: 25).

GIS as an integrating technology

One of the most persistent and pervasive buzzwords in the field of GIS is 'integration'. Indeed, the ability of GIS to integrate diverse information is frequently cited as its major defining attribute, and as its major source of power and flexibility in meeting user needs (Maguire 1991 in this volume).

There are strong similarities in the way that GIS and the discipline of geography have been promoted as 'integrating' mechanisms. Geography, it is argued, is an integrating discipline because it unites the study of society with that of the physical environment. GIS, it is claimed, is an integrating technology because of the way in which it links together diverse types of information drawn from a variety of sources. By integrating information, users can take a unified view of their data and large organizations can establish a single, coherent, corporate information system.

The benefits of information integration

The benefits that follow from the integration of diverse information within a GIS are widely recognised:

- A broader range of operations can be performed on integrated information than on disparate sets of data.
- By linking data sets together, spatial consistency is imposed on them. This adds value to existing data, making them both a more effective and a more marketable commodity.
- Through the integration of data which were previously the domain of individual disciplinary specialists, an interdisciplinary perspective to geographical problem solving is encouraged.
- Users benefit from the perception that they have access to a seamless information environment, uncomplicated by the need to

consider differences in data sources, information types, storage devices, computer platforms, etc.

Further advantages accrue if several organizations pool their individual data into a single integrated database (Bracken and Higgs 1990):

- Data acquisition costs are reduced, because of the elimination of duplicate data collection and conversion activities.
- Organizations can draw on a broader base of information than hitherto, and are thus able to address issues that were previously beyond their individual data resources.
- Organizations can cooperate with one another within the context of shared information, and thereby make more effective management decisions.

What is 'Information Integration'?

What is meant by 'information integration', and how is it achieved? The dictionary definition of integration, 'the combination of separate parts into a whole', provides a poor starting point for a discussion of this concept. The GIS literature can be equally unhelpful, for integration is frequently ascribed several distinct meanings, including:

- The bringing together of spatial data from a number of sources, including maps, field survey equipment, photogrammetry, and remote sensors, within a single system (e.g. Aybet 1990).
- The creation of a geometrical description of the earth's surface within a consistent topological framework (e.g. Marx 1986).
- The inter-conversion of raster (i.e. image) and vector (i.e. map) models of the world within a single software environment (e.g. Jackson and Mason 1986).
- The provision of a comprehensive set of geographical information handling functions within a unified software framework (e.g. Dangermond 1986).
- The interlinking of both spatial and attribute data within a single, coherent representation or model (e.g. ESRI 1990a).
- The synthesis of diverse spatial information by means of fundamental geographical operations such as spatial search and overlay (e.g. Cowen 1988).

It is also perhaps worth noting that over and above these multiple meanings of integration, there is further terminological confusion in the use of terms such as 'integrate', 'link', 'relate', 'combine' and 'match'. These are sometimes used interchangeably, but are sometimes given distinct meanings (e.g. in Rhind *et al.* 1984).

It may be useful, therefore, to provide a working definition that clarifies these issues, and which serves as a framework for subsequent discussion. In this review, *information integration* is seen as the synthesis of geographical information in a computer system which depends for its effectiveness on *information linkage* (i.e. of spatial and attribute data) within a coherent data model. This involves bringing together diverse information from a variety of sources (*information interchange*), requires the effective *matching* of supposedly similar entities in these sources, and demands *information consistency* across the source data sets.

THE CLASSIC APPROACH TO INFORMATION INTEGRATION IN GIS

The integration of spatial and attribute data

In commercial information systems, non-spatial data are typically integrated by storing them non-redundantly in a single database engine. In computer mapping systems, spatial data of various kinds are frequently integrated within a single graphical database, in either vector or raster format. Within each of these systems, the relationship between items of consistently recorded data can be used to answer a variety of questions. However, the range of questions that can be asked must be of a spatial or non-spatial nature.

Within GIS, information integration goes one step further, involving the linkage of non-spatial or attribute data to spatial information describing real world features. By performing operations across the two sets of information in tandem, a far richer set of questions may be asked, and a far broader range of problems can be solved than in those systems that handle just attribute or spatial data alone. For example:

1. Users can interrogate geographical features displayed on a computer map and retrieve associated attribute information, for display or further analysis.
2. Maps can be constructed by querying or analysing attribute information in a database.
3. New sets of information can be generated by performing spatial operations (such as polygon overlay) on the integrated database.
4. Different items of attribute data can be associated with one another through a shared locational code.

Two approaches, or models, have been widely adopted for achieving the linkage between spatial and attribute information within GIS: the composite map model and the geo-relational model (Aronson 1987; Bracken and Webster 1989; Healey 1991 in this volume). Each of these approaches to data integration is based on a particular type of spatial data model: the composite map model is usually based on a tessellated representation of space (typically the grid-cell tessellation), while the geo-relational model is usually associated with a vector representation of space (and particularly the arc/node topological model). For a fuller discussion of the relative merits of these and other spatial data models, see Peuquet (1984), Burrough (1986), Davis and Simonett (1991 in this volume) and Egenhofer and Herring (1991 in this volume).

The key element in both of these models is that links are established between attribute information and spatial features. The precise techniques used to create these links vary from GIS to GIS but, in general, involve establishing a pointer between each spatial feature in the database and its associated attribute information. In the composite map model, the links are implicit in the way that specific attributes are assigned to individual map layers, and in the process of assigning specific attribute values to the spatial entities (i.e. cells) on each layer. In the geo-relational model, by contrast, the links are established by arranging for each spatial feature's unique identifier (or ID) to be recorded in a key field of the appropriate database table(s) that store its attribute information.

For each data model, an appropriate set of operations is available that integrates across (or synthesizes) the two kinds of data in response to user queries. In the composite map model, the principal integration mechanism is a 'map modelling' capability that relates together two or more grid cell layers to produce new layers, on which further modelling operations may be performed (Berry 1985, 1987; Tomlin 1990, 1991 in this volume). In the geo-relational model, integration is accomplished by undertaking operations such as spatial search and overlay (ESRI 1990b).

The composite map model

The first method of linking spatial and attribute information is by means of a multi-layered map (Berry 1985; Tomlin 1990). This model is essentially a computerized version of the technique of filter or sieve mapping developed for landscape planning and resource management (Steinitz, Parker and Jordan 1976), but it is also closely associated with more recent developments in image processing.

In the composite map model, attribute information is referenced to artificial units of space (typically square grid cells) that form a regular grid. These units are usually constant in size, shape and orientation, and rarely have intrinsic meaning. Grids of cells are usually 'sliced' horizontally into a number of layers or planes, with information about specific variables or attributes stored on individual layers (see Fig. 22.1).

Within the context of this model, information integration proceeds by combining attribute values for cells that lie above or below one another in a 'stack' of superimposed layers. A quantitative model is generally used to derive secondary values from the source attribute values. This can be done in one of two ways: either the model is applied to each cell independently, or else the model operates on cell values within a layer that are within a given range or window. These are sometimes termed Type 1 and Type 2 operations (Heuvelink, Burrough and Stein 1990), and are illustrated in Fig. 22.1.

The grid-cell model is a relatively simple approach to data integration, both conceptually and operationally, and it has therefore been popular

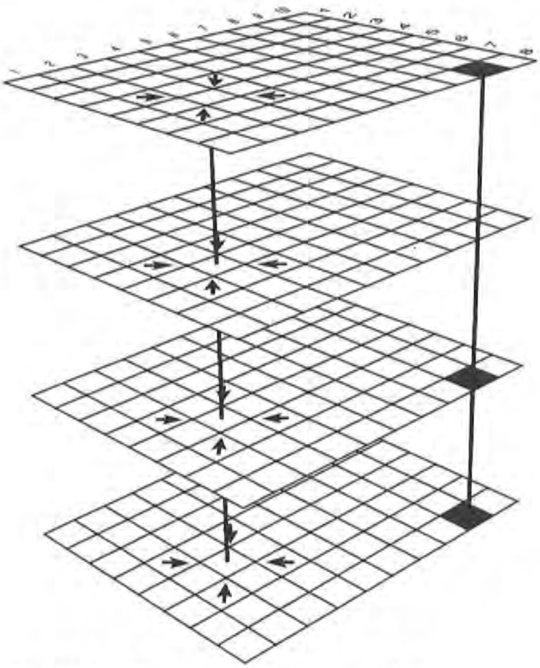
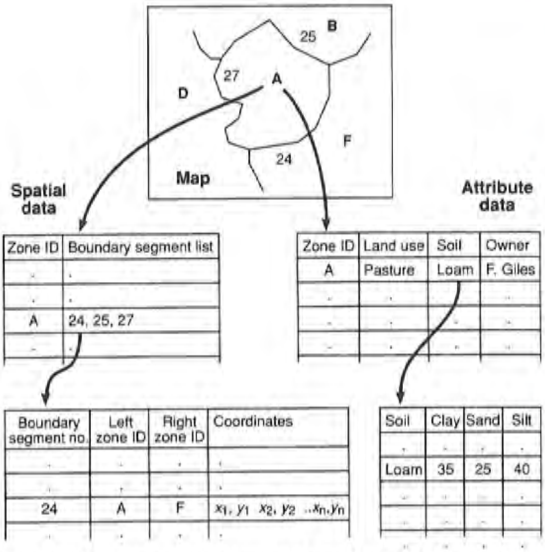


Fig. 22.1 The composite map model of information integration in which both spatial and attribute data are combined in the same layer.



Zone ID	Boundary segment list
.	.
.	.
A	24, 25, 27
.	.

Zone ID	Land use	Soil	Owner
A	Pasture	Loam	F. Giles
.	.	.	.
.	.	.	.
.	.	.	.

Boundary segment no.	Left zone ID	Right zone ID	Coordinates
.	.	.	.
.	.	.	.
24	A	F	$x_1, y_1\ x_2, y_2\ ..x_n, y_n$
.	.	.	.

Soil	Clay	Sand	Silt
.	.	.	.
Loam	35	25	40
.	.	.	.
.	.	.	.

Fig. 22.2 The geo-relational model of information integration in which spatial and attribute data are stored separately.

since the earliest days of GIS development. The model is currently implemented in a large number of 'cartographic modelling' and raster-based image processing systems.

The Geo-relational Model

In the second, and more recently developed, approach to information integration, attribute information is associated with point, line, zonal or other spatial entities that describe features occurring in the real world. Thus, for example, a point located feature such as a house may have associated with it such items of information as its local tax rate, its current market price, its owner, its occupants, and so on. Similarly, a linear feature such as a river might have associated with it such information as mean discharge, pollution levels, fishing rights, etc. In like manner, a zonal feature such as a field might be linked to information describing its owner, its current land use, its soil type, and so forth.

In this approach to data integration, spatial entities are usually linked with their associated attribute data by means of a common spatial key (commonly a unique identifier or ID assigned to each spatial feature). Different sets of attribute information are stored in different attribute tables, and the relevant information for a given set of spatial features is accumulated by relating (or joining) two or more tables of information (see Fig. 22.2). The collation of information from the attribute tables may be carried out in several ways: by exact matching, by hierarchical matching, or by fuzzy matching (ESRI 1990b; Flowerdew 1991 in this volume).

The geo-relational model has two important characteristics. The first is that the join works both ways; thus, attribute information can be found by selecting spatial features, and spatial features can be found by querying attribute information. Secondly, all attribute information within the model is associated with one or more spatial features, even though it may require a chain of joins between numerous tables in order to establish this relationship. Attribute information that is not linked (either directly or indirectly) to a spatial feature is normally not included in the attribute tables.

Most contemporary, vector-oriented GIS adopt the geo-relational model to conceptualize the links between spatial and attribute data. In addition, most usually adopt a dual or hybrid data storage strategy to implement the model, with spatial data held separately from the attribute data, and the system maintaining links between the software modules that handle each information type (Duecker 1985; Morehouse 1985). Locational data are typically represented using a topological (or arc–node) spatial data model, and thematic data are usually stored in the tables of a standard relational database. Thus, for example, in the case of ARC/INFO, the archetypical example of this approach, the spatial model is implemented in the ARC program, and the attribute data model is implemented in INFO, a fourth generation database system. Users of the system are able to address each software component separately, by means of the ARC macro language and INFO commands respectively. However, the tight coupling between ARC and INFO ensures the integrity of the relationships between spatial entities and their attributes.

Similarly, the GDS/AMS system consists of a graphics database (containing land feature data) and a textual database (containing assigned data) which are connected by an interface consisting of a lookup table. SICAD, too, combines a geographical database for storing topologically structured spatial information with a relational database for storing textual information (Wagner 1990). Many other operational GIS (including TerraSoft, Strings, and ILWIS) adopt this hybrid approach.

Because of the relative simplicity of this hybrid approach to implementing the geo-relational model, many LIS and AM/FM systems have been developed by end users linking together a CAD system (to store the spatial data) and a database program (to hold the attribute data) (Cowen 1986; Shepherd 1990a, 1990b, 1990c). In contrast to commercial GIS that adopt the hybrid data model, the component items of software in such 'do-it-yourself' GIS are usually loosely coupled, with intermediate files being used to exchange data about the links between the two types of information.

Limitations of current integration models

GIS do not (and maybe cannot) always provide an entirely satisfactory approach to the integration of geographical information. In particular, the overlay mechanism at the heart of most GIS software is not able of itself to guarantee error-free integration of spatial and attribute data. Errors introduced by data capture and conversion, together with residual errors not resolved by data consistency analysis, can combine with errors generated by the overlay process itself to provide output maps that have a probabilistically uncertain level of error.

In vector GIS, for example, where the accuracy of points, lines and areas stored in the geographical database is subject to small margins of error, the overlay process will propagate these errors when producing resultant maps (Rhind *et al.* 1984). Experiments in geographical sensitivity analysis clearly reveal how perturbations or errors in input maps can change the appearance of output maps (Lodwick, Monson and Svoboda 1990). The integration process can itself also generate errors, a well-known effect being the appearance of spurious sliver polygons as a result of vector overlay operations (Goodchild 1976). Chrisman (1984, 1991 in this volume) argues that earlier assessments of the errors associated with the overlay process need to be revised in the light of modern software, but acknowledges that the lack of appropriately accurate base map sources remains a prime source of integration errors.

Similar problems arise with integration operations in grid-cell GIS. For example, the validity of the layer combination process may be compromised if attribute information has been erroneously allocated to cells by the process of cell aggregation used in creating a consistent spatial base (Walker and Hutton 1986). Additionally, where uncertainties associated with coefficients used in the map combination model are unknown, then the relative contribution of the model to the total error in the output map will also be unknown (Heuvelink *et al.* 1990). Finally, where spatial units and/or data attributes are fuzzy rather than discrete, the standard set theoretic approach implicit in the overlay process may break down entirely (Wang, Hall and Subaryono 1990; Burrough 1989). For a fuller discussion of these issues, see Chrisman (1991 in this volume) and Tomlin (1991 in this volume).

Preparing for information integration

Organizations that are new to geographical information handling frequently make the mistake of assuming that the GIS is some kind of magical data integrator, an automatic 'information melting pot'. In reality, the rewards that may be reaped by adopting a GIS to perform integrated information analysis (as outlined above) are contingent upon the amount of effort expended in creating a consistent database in the first instance. The adoption of one or other of the currently fashionable models of information integration does not automatically create harmony among previously incompatible data sets. Information integration is in fact a two-step process: the first step involves the removal of inconsistencies between disparate data sets, to ensure their compatibility; the second involves the use of GIS software to interrelate consistent data in ways that meet the needs of particular applications (see Flowerdew 1991 in this volume).

The following sections examine the causes of inconsistent data, and the various methods available for, and some of the broad approaches taken in, creating consistent geographical databases.

The data inconsistency problem

Geographical data sets are difficult to integrate where there are inconsistencies between them. These inconsistencies may affect the spatial and attribute characteristics of data, and necessitate the use of various corrective measures. Inconsistencies result from a variety of causes:

- differences in recording or measurement techniques (e.g. observation times or periods, data gathering equipment, data categories, human observers/interviewers);
- errors in measurement or survey methods (e.g. malfunctioning data loggers, data recording errors by human observers);
- variations in the resolution (spatial, temporal or attribute) of the data gathered;
- vagueness or imprecision in definitions (spatial or attribute);
- fuzziness in spatial objects (e.g. soil boundaries);
- variations in the use of terminology and nomenclature.

Inconsistencies tend to be greater where:

- the study area involves several administrative or governmental data units, and/or where several organizations are responsible for gathering information;
- the information is drawn from multiple sources, at several scales;
- the information consists of several types (map, image, text, numeric, sound, etc.);
- the information is available at several points in time;
- the information is stored on diverse media, and on more than one computer system.

Problems of inconsistent data are frequently greater in developing countries than in the developed nations of the world (Casley and Lury 1981; Robinson *et al.* 1989). However, considerable problems arise even in developed countries when attempts are made to integrate information across national boundaries. An example is provided by the CORINE project, which sought to demonstrate the viability and utility of an integrated (vector) GIS for soil and land use planning in the European Community (Mounsey 1991 in this volume). This project revealed that problems of data inconsistency are frequently far more significant than those of data availability, accuracy, geographical coverage, large data volumes, or lack of adequate data structuring. Moreover, despite the use of a variety of techniques for overcoming problems of spatial inconsistency, it was nevertheless concluded that where a GIS involves a large surface area (in this case, 2.26 million km^2), and a number of political entities (in this case 12 nation states), 'the problem of data inconsistency is not easy to tackle and cannot wholly be resolved' (Briggs and Mounsey 1989: 16).

Techniques for removing inconsistencies in data

The creation of a consistent geographical database requires effort at a number of distinct levels: the resolution of differences in the meaning and significance ascribed to data by cooperating organizations; the integration of incompatible database schemata (Frank 1986; Nyerges 1989b); and the conversion of information between contrasting data models, data structures, and data media. Ideally, the removal of inconsistencies should proceed in a top-down manner, dealing first with differences in the meanings of data, before addressing issues of data representation and structure (Nyerges 1989b).

The specific techniques that may be used to create consistent data can be classified on the basis of the information being handled (see Table 22.1). Fuller accounts of these techniques are provided elsewhere (e.g. Rhind and Tannenbaum 1983; Rhind *et al.* 1984; Flowerdew 1991 in this volume).

Table 22.1 Methods for creating consistent databases (see Flowerdew 1991 in this volume for further details).

Spatial data

Map projection standardization, rectification of local geometrical distortions, coordinate registration, coordinate density equalization, feature generalization, edge matching, scale conversion, image-to-image registration, co-registration of maps and images, conversion between vector and raster data models, creation of multi-image mosaics, allocation of one set of zones to another, etc.

Attribute data

Aggregation of data classes, reclassification of raw data, reduction in numerical precision of data, reduction of levels of measurement to a common level, address matching, nominal record linkage, schema integration analysis, conversion from network or hierarchical database to relational database, greyscale normalization, etc.

In some cases, the task of ensuring consistency in spatial data can be undertaken separately from that of ensuring consistency in attribute data. Considerable benefits accrue, however, if these efforts proceed in tandem (an example is address/location matching). In particular, there are many occasions when the removal of inconsistencies in spatial data sets should follow their removal in attribute data. This is because a consistent attribute data set can often provide valuable guidance to the process of removing inconsistencies in spatial data (Nyerges 1989b). Similarly, the creation of a topologically consistent map base is a valuable precursor to ensuring geometrical consistency in a spatial database.

Sometimes, as is the case with map conflation (Saalfeld 1988), incompatibilities are best resolved by handling items of source information one pair at a time. In other operations, such as database schema integration (Nyerges 1989b), several sources of information are considered together.

A number of broad strategies may be adopted in creating consistency among disparate data sets. One approach is to permit inconsistent data to coexist, with explicit information being provided on the nature of the inconsistencies (Briggs and Mounsey 1989). Another approach is to convert all source data sets into a single target version, as in the integration of multiple database schemata (Nyerges 1989b). A third approach is to convert data stored in one data model to another data model, as is the case in preparing data for vector-only or raster-only GIS (see Peuquet 1981a, 1981b; Piwowar, LeDrew and Dudycha 1990). In many applications, data consistency is often achieved by reducing diverse spatial or attribute data to some lowest common denominator representation. For example, all attribute data may be reduced to nominal scale measurements, or all spatial data may be reduced to a relatively coarse grid-cell representation.

These approaches are also found in many mainstream commercial systems. Some integrated office packages, for example, store different types of data in separate files, and/or maintain different data models for each of the individual functions provided. Information is exchanged between the various components by automatic or manual file swapping. By contrast, some data handling packages use a single data model (e.g. the spreadsheet), and provide the necessary internal functions to store and structure the various types of user information within this single model.

The 'single model' approach is perhaps taken to its extreme in Document Image Processing (DIP) systems, in which various source materials (e.g. text, drawings, maps, photographs) are reduced to a single form: the Digital Image Document (DID). This is a standard scanned image, and large collections of such images may be indexed, stored, managed, accessed and displayed using tools available in a DIP system. As in raster GIS, the availability of optical storage devices (such as CD-ROMs and WORMs) with huge storage capacities has greatly accelerated the adoption of this approach.

Typically, a variety of techniques is applied in a given application to ensure consistent data. In the

CORINE project, for example, several techniques were adopted to remove or reduce spatial incompatibilities between source data incorporated into a geo-relational GIS (Briggs and Mounsey 1989). These included: conversion of data to a single reference map projection; removal of local geometrical distortions by rubber sheeting; adoption of two standard scales (1 : 1 million and 1 : 3 million) at which to store data; and generalization of larger scale data down to these standard scales.

The creation of a consistent database is still essentially a manual operation. It is frequently labour intensive, and usually requires the participation of those who are knowledgeable both in the characteristics of the source data and also in the application area for which the data are being assembled. More recently, however, many of the tasks undertaken to create consistent databases have been supported by software tools. These are available either as modules in a general-purpose GIS, or as special-purpose utility programs. Examples of the latter include software for rubber sheeting (Bedell 1988), zipping (Beard and Chrisman 1986), map conflation (Saalfeld 1988), and vector-to-raster conversion (Steneker and Bonham-Carter 1988).

Unfortunately, not all data inconsistencies are resolvable simply by using the battery of tools available. For example, the use of an interactive graphics editor to displace polygon boundary coordinates (what Dangermond (1988) calls 'graphic fudging') may be constrained or prevented by legal definitions of one of the boundaries in question. Similarly, the use of distance-based or point-in-polygon techniques to allocate one set of zones to another may lead to unacceptably high error rates. In other cases, it may not be possible or acceptable to reduce the variability in source information by creating or adopting a single, 'clean' data set. In some applications, it may be desirable to let mutually inconsistent versions of data to coexist in a controlled fashion within a GIS (Newell, Theriault and Easterfield 1990). Finally, many inconsistencies may be 'hidden' in apparently consistent data, and these may have to be flagged rather than corrected (Briggs and Mounsey 1989).

The operations needed to create a consistent geographical database can be implemented at various stages in the data conversion process (Dangermond 1988):

- Before data conversion: use of standard base maps (e.g. Integrated Terrain Units) to record field information, etc.
- During data conversion: templating, on-line transformation, automatic text annotation, automatic snapping, on-the-fly topological construction, etc.
- After data conversion: manual interactive editing, rubber sheeting, conflation, line snapping, schema integration analysis, attribute consistency checking, etc.

The amount of effort needed to ensure data consistency tends to increase at each stage of the conversion process. The dictum 'prevention is better than cure' applies just as much to the creation of a consistent database as it does in matters of personal health and safety. It is far cheaper to 'design out' possible inconsistency at an early stage in the development of a geographical database. Unfortunately, many geographical databases are created by merging existing digital data sets, and incompatibilities have, therefore, to be removed after data conversion has taken place.

Integrating contrasting spatial data models

One of the long-standing problems of operational GIS has been the separation of information derived from maps and information derived from images. The former has typically been the preserve of vector-oriented GIS, the latter has been the preserve of image processing systems.

Recently, however, major efforts have been made to integrate remote sensing data with cartographic data (e.g. Zobrist 1983; Jackson 1987; Logan and Bryant 1988; Goodenough 1988; Annoni *et al*. 1990). Designers of remote sensing systems have attempted to bridge the gap either by adding mapping capabilities to existing image processing functions (e.g. the drawing of vector ground reference features such as coastlines, rivers, and national boundaries), or by providing users with on-line access to GIS information (e.g. to assist with the process of image classification). Some raster GIS provide vector tracing or annotation functions for raster images (e.g. ERDAS and SPANS).

For their part, developers of vector GIS have tried to add a variety of raster/image processing

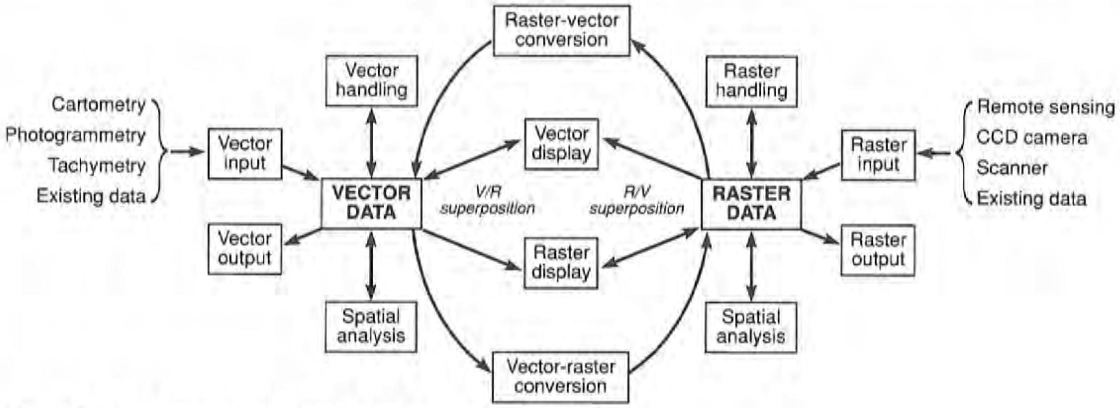


Fig. 22.3 The integration of vector and raster data in a GIS.

capabilities: displaying raster backgrounds beneath vector maps, vectorizing scanned images, and linking images to vector coordinate systems (Pevsner 1989; Annoni *et al.* 1990). A prime example of this trend is ARC/INFO, which enables vector information to be overlaid on raster images by means of its Image Integrator module. An increasing number of systems provide utilities to convert spatial data from the alternative model into its own native model, or to convert between the vector and raster data models.

Figure 22.3 illustrates the internal structure of a fully integrated GIS which attempts to combine both vector and raster information. (This figure is derived from a system diagram of the SICAD–HYGRIS system, and from Laser-Scan's idealized view of an integrated GIS architecture.) Several systems are now available that claim to be able to combine vector data and functionality with raster data and functionality, and these are sometimes labelled 'Integrated GIS' (or IGIS) to distinguish them from those that are limited to a single spatial data model (Jackson and Mason 1986; Gorte, Liem and Wind 1988).

Resolving inconsistent data formats and media

Beyond inconsistencies in the content of geographical data lie incompatibilities in the format or medium used to record the data. Format and media incompatibilities can arise wherever data are gathered by different methods or equipment, where data are stored on two or more computer systems, or where data are handled by two or more computer programs. Such incompatibilities require a further set of operations to enable successful data integration within a GIS (Aybet 1990).

The most common solution to this problem lies in the use of data translators to inter-convert a wide variety of cartographic, image and attribute data (numeric as well as text). Most translators are special purpose, working only with specific data formats, but general purpose data translation tools, driven by relational database technology, are also available (Waugh and Healey 1986; Pascoe and Penny 1990; Guptill 1991 in this volume). Once again, these translators are available either as stand-alone utility programs, or as modules within general-purpose GIS. By way of example, the data conversions available within the ARC/INFO system are illustrated in Fig. 22.4. In mainstream computer applications, utility programs are widely available for converting between text, database, spreadsheet and graphics data formats.

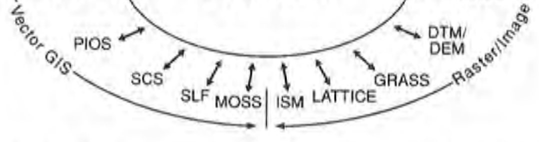


Fig. 22.4 Data conversion facilities in ARC/INFO. Reprinted courtesy of Environmental Systems Research Institute, Inc. Copyright © 1991 Environmental Systems Research Institute, Inc.

Two approaches are generally taken to the inter-conversion of incompatible data formats and media: one-to-one conversion using direct translators (necessitating n^2n-n translators for n different formats); and translation to and from a universal intermediate format (requiring 2_n translators for n different formats). Clearly, any strategy that reduces the number of data formats in general use simplifies this conversion problem. In the wider world of data processing, a number of *de facto* standards has emerged that simplify the exchange of attribute and spatial data (Table 22.2). The emergence of national and international standards for cartographic data (e.g. Haywood 1986; Zarzycki and Jiwani 1986; DCDSTF 1988) is an equally welcome step in reducing the problems associated with data exchange.

For the future, data exchange mechanisms other than the commonly used external file are likely to assume greater importance. One such mechanism is the 'live link', which enables items of information in separate application programs to be inter-connected. (In the personal computer environment, this facility is currently available in the form of the Dynamic Data Exchange (DDE) protocol, and the Inter-Application Communications (IAC) mechanism.)

By linking information directly together, some of the problems of data loss associated with conventional transfer may be avoided. In addition, the 'live link' approach makes possible several forms of data interchange: data snapshots may be passed automatically from one program to another; changes in the data held in one program (e.g. a spreadsheet) may be reflected in automatic changes to information held in another program (e.g. a desktop mapping program); real-time data (e.g. from data loggers) may be accessed automatically by data analysis and display software; and data queries may be made across multiple linked documents (Vose 1990). In addition, live data links may be established among programs running on a single multitasking computer, or on several different computers connected across a network.

Table 22.2 Standard formats for data exchange.

Structured data

Database formats: dBASE (DBF).

Spreadsheet formats: DIF, WK1 and WKS, SYLK.

ASCII (the lowest common denominator; this contains the data and, perhaps, data definition information, but none of the internal data structuring information specific to a particular data handling program) – common formats are CDF and SDF.

Text

Word processing formats: Word Perfect, Word, Wordstar, DCA/RFT.

ASCII (the lowest common denominator; this contains the text but none of the display and formatting characters found in standard word processing files).

Line Drawings

Line art and business graphics program formats: HPGL, CGM, Lotus PIC, Windows Metafile, General Parametric Videoshow, Postscript.

CAD program formats: DXF, IGES, SIF, Set, Step/PDES.

Maps

DIGEST, VPT, NTF, DLG, etc.

Painted pictures

Painting program formats: PCX, GEM, IMG, TGA.

Raster images

Scanned image format: TIFF, GIF.

FAX format: CCITT Group 3 and 4.

Presentation

NAPLPS, EPS.

Documents

ODA, CDA, DCA, DDIF, SGML, Edifact.

Linking existing databases

A major problem in establishing an integrated spatial database is that many organizations already have a massive installed base of computer-readable, geo-referenced data. In many cases, this information is handled by application-specific software; in others, it is handled by a standard database management system. Increasingly, the component data sets of a potential geographical database are distributed over several computers, often with incompatible operating systems. Frequently, attribute data are stored in one place and spatial data in another.

The ability to integrate data sets across several items of software and several computer systems is, therefore, a key factor in the success or failure of GIS technology within modern organizations. McLaren (1990) sees four approaches to the integration of existing information in such circumstances, each providing a different level of effectiveness:

1. *Data absorption*: information is copied from existing data handling software, either on a temporary or a permanent basis, into a GIS for analysis.
2. *Querying of external data sets*: features in a GIS are linked to sets of information held in external databases. The links are usually *ad hoc*, and the items of software are only loosely coupled.
3. *DBMS interfacing*: spatial objects modelled in a GIS are linked through a generic interface to attribute data held in an external DBMS.
4. *Full integration*: information on spatial objects is managed by a single, heterogeneous, distributed DBMS (or DDBMS).

The first three approaches are now commonplace, frequently being adopted where it is economically or operationally infeasible to transfer existing information to an entirely new system.

Several authors (e.g. Webster 1988; McLaren 1990) argue that the full potential of GIS will only be realized if a corporate approach is taken to information management, and particularly one that addresses the issue of distributed data. They also argue that conventional approaches to data integration which do not use mainstream DBMS technology to handle *both* spatial and attribute data are unlikely to allow the creation of integrated corporate databases. In this scenario, it is the coherent management of geographically distributed data, by means of a DDBMS, and not just the interlinkage of data items in diverse software systems, that is crucial to the success of an integrated geographical information system. Thus far, there have been only a few implementations of this approach (e.g. Green 1990; Nicholson 1990).

Organizational factors in data integration

The significance of organizational factors in the creation of fully integrated geographical databases cannot be overemphasized. For example, because different organizations collect data for different purposes, they may ascribe contrasting meanings to their data. These purposes, therefore, need to be taken into account when decoding the meanings of specific data items. This conclusion emerges clearly from work undertaken to create multipurpose land information systems in North America (Liley 1985; Nyerges 1989a).

Organizational factors are also critically important when establishing operational data exchange arrangements. This is illustrated by the development of a prototype GIS for inter-organizational decision making in mid-Wales (Bracken and Higgs 1990). A major conclusion derived from this exercise was that the benefits of an integrated database will only accrue if there is consensus among the participants (individuals, departments or organizations) involved in creating it, and that this consensus will only emerge if the data users see a clear value in the integrated database. An important contributory factor in achieving success is that the database users should have a tradition of information sharing.

ALTERNATIVE APPROACHES TO INFORMATION INTEGRATION

Limitations of current approaches to information integration in GIS

Several assumptions, both explicit and implicit, are made by current GIS vendors in their approach to information integration. For example, the range of information types currently integrated in most GIS is largely dictated by the *data models* they implement. In particular, the information they accept is limited to that which easily fits the database model of attribute space (i.e. numeric data and structured alphanumeric strings), and that which fits the map or image model of geographical space (i.e. vector or raster representations of space). Similar restrictions are imposed on the ability to integrate data available on different media.

Similarly, the way that information is integrated, and therefore the uses to which it can be put, are imposed by the *integration model* adopted by a particular GIS. This has two major implications, one concerning the use of precise locational referencing, the other concerning the cognitive model adopted for the integration process.

The traditional mechanism for integrating data within most GIS is the overlay process. This implies the use of some form of locational reference or geocode to link attribute data to their respective spatial features. The assignment of a unique spatial key to all data entities enables the software to store, retrieve, analyse and display diverse data in a precisely coordinated fashion, and ensures that all information for a given location can be brought together as required.

However, one of the dangers of using an unambiguous and precise location code as the principal data key or linkage mechanism is that information which cannot be explicitly or precisely tied into a particular locational referencing scheme runs the risk of being ignored, or relegated to an ancillary role. This has occurred with several types of information in the past, including: non-vertical aerial imagery (ground-based and low oblique aerial photographs); free text which describes broad geographical areas; environmental sounds; moving images; etc. However, it is just such information, with its richness, ambiguity, fuzziness and subtlety that many users want to be able to access in their information systems. Although steps have been taken during the past few years to accommodate such information within GIS (Dangermond 1988), this has typically been accomplished within the framework of an existing model of data integration.

Current thinking and research in GIS tends to ignore the information richness of the real world. In a recent volume on the design of databases for global science (Mounsey 1988), for example, it was assumed by most contributors that the structure of world-wide databases would depend primarily on just two major types/sources of information: the cartographic and the remotely sensed. Elsewhere, too, the discussion of geographical data in GIS is dominated by consideration of those models (notably the vector and the raster) which are best able to accommodate these two types of information.

A number of GIS workers do recognize that 'research is needed on ... techniques for integrating heterogeneous data' (Abler 1987: 308). However, the alternatives that are proposed are often conditioned by the prevailing view of geographical information (i.e. consisting of highly structured data), and by existing models of data integration (i.e. those based on precise locational referencing of spatial objects). These views are also heavily conditioned by an analytic form of thinking among geoscientists, and lead to the continued development of GIS which concentrate on the computer processing of precisely coordinated spatial data sets. However, in its overemphasis of the role of the GIS as a machine for data handling and analysis, this approach de-emphasizes the role of the end-user as an information integrator, interpreter and analyst.

It is clear from the use made of multiple window displays on modern workstations, for example, that users can benefit from being able to perform their own visual integrations of diverse information displayed on screen. Users also frequently integrate multiple sources of information in a relatively loose way outside a GIS. The use of desktop publishing software to produce environmental reports by merging information of many kinds is an example. In certain applications, therefore, user-centred methods of information integration complement the machine's algorithmic approach.

For these reasons, three alternative approaches to the integration of multiple types of information that have emerged from other fields in recent years will be considered further here. Each represents an alternative, even competitive, model of data integration to that provided by current GIS technology. They merit attention, because they suggest directions in which GIS might move in order to integrate information about the real world more effectively.

Multimedia databases

Recent developments in computer technology have led to a rapid expansion in the number of information types that can be routinely handled by computers. Most notable have been developments in optical storage technology, hardware for combining images from digital and video sources, digital sound sampling, and hi-fi sound output. Of greater importance in the present context, is the convergence of previously separate information handling technologies (e.g. video and audio, digital and analogue, and television and computing) (Press 1990), and the development of desktop 'multi media' computer systems that permit users to access many types of information (e.g. text, graphics, animation, sound and motion video) from a number of complementary sources (e.g. videodisk, CD-ROM, scanner, video camera, audio tape, digital sound sampler and data logger). Figure 22.5 illustrates a typical PC-based multimedia system.

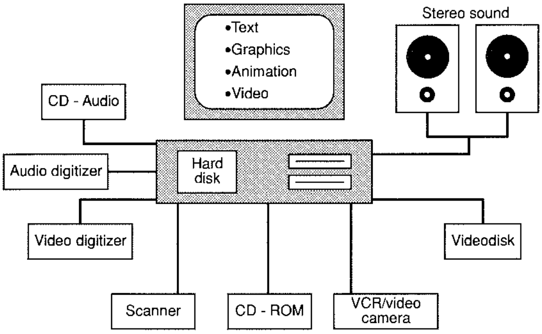


Fig. 22.5 The components of a desktop multimedia workstation.

This section considers how non-conventional information has been embraced by conventional database technology, and what implications this might have for GIS. Some DBMS have addressed the problem of non-standard data items by introducing the 'long record' or the 'binary large object' (BLOB). BLOBs consist of large volume data items (e.g. documents, software, faxes, graphs, images and voice information) that do not fit neatly into the standard database management system framework. One system that has embraced this approach is Informix, a standard RDBMS whose derivative product, Informix-OnLine, can handle not only the usual structured information, but also large-volume items with little or no standard structure (Shetler 1990). In this system, the two kinds of BLOB (text and byte) are stored separately from the regular data, perhaps on separate storage devices, but they are handled entirely by the standard database mechanisms. BLOBs can be selected, updated and inserted using standard SQL commands, and displayed to the database user on request. (How information is presented to the user depends on the output devices and device drivers available.) However, searching for BLOBs, or relating BLOBs to one another, on the basis of their content, would require extensions to standard SQL.

A number of multimedia database systems are also available in the PC environment. These usually involve relatively simple extensions to a standard database or information retrieval program so that database records can be linked to scanned or drawn images.

The multimedia database appears to be the current development route for several GIS. ESRI, for example, has identified eight related geographical data technologies which need to be interlinked in a GIS framework (ESRI 1989) (Fig. 22.6). Two of these, the vector GIS and the DBMS, are at the core of ARC/INFO operations, and are commonly found in many other GIS. The other technologies may be connected to the system by one or more of the following generic integration techniques:

1. *Image overlay*: Superposition of vector data on a raster image (e.g. network lines drawn on a scanned base map), and/or superposition of a raster image on a vector display (e.g. a satellite image draped over a DTM). In either case, no logical connection is made between the two data sets.
2. *Data conversion*: Conversion between vector and raster data representations, including data exchange between mapping and CAD systems with differing data formats.
3. *DBMS interfacing*: Linking cartographic feature data with attribute data managed by an external DBMS.
4. *Feature and attribute association*: Creation of an index link between a GIS feature and a displayable item (e.g. a raster document, a video image, an environmental sound, an animation sequence, an engineering drawing). These items may be stored on a separate storage medium (e.g. optical disk).

Each of these techniques is now increasingly used in commercial GIS. The last mentioned has

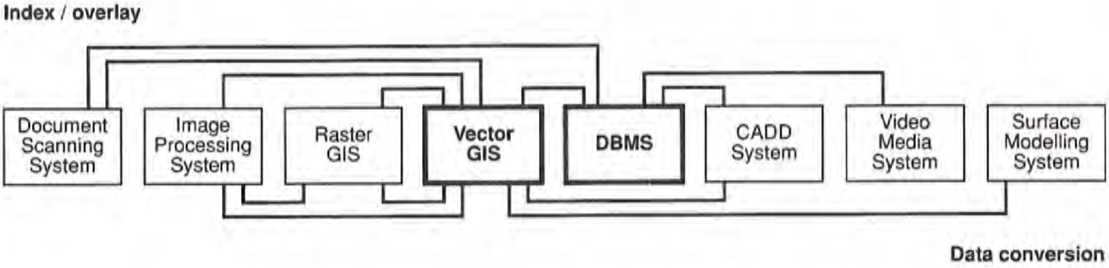


Fig. 22.6 Linkages between different information types in ARC/INFO. Reprinted courtesy of Environmental Systems Research Institute, Inc. Copyright © 1991 Environmental Systems Research Institute, Inc.

become particularly popular with the advent of optical mass-storage devices. For example, in a system that is used by Paris police to monitor events in the city as they happen, the operator can zoom down to street level on a vector map, and bring up a video image of key buildings from an optical disk system (Anon 1990).

However, several issues remain to be resolved before multimedia databases achieve their full potential in a GIS context. Among these are: how to handle spatial as well as linear multimedia objects (e.g. images and sound); how to handle the transfer of multimedia objects from the database, when some may require 'persistent' presentation and others 'non-persistent' presentation; how to control the capture and/or presentation of multimedia data once the data transfer has begun; and how to ensure efficient storage and transfer of multimedia information (Woelk and Kim 1987).

Some idea of the complexities involved in integrating non-conventional information in a GIS can be provided by briefly considering visual images. In GIS that associate images with database records, the link is usually made by assigning a field in the appropriate attribute data table in which to store a pointer (such as a file name) to an image stored elsewhere. The problem lies in deciding which spatial feature(s) in the database are to be associated with which image(s).

The simplest approach is to link an image to a point-located feature. Thus, for example, in a road network GIS, the user might point to a road intersection marker and bring up on screen a photograph of that intersection; or in a property database, clicking on a property symbol might bring up a photograph of a house. However, a single image may be linked to several types of feature (point, line or zone), in which case the spatial extent of the photograph will not match the geometry of one or more of these features.

This type of linkage provides only a weak form of integration of the spatial extent and content of an image with that of the other (conventional) information in the database. In order to associate the image with other spatial data in the database fully, some attempt needs to be made to rectify its geometry, perhaps using appropriate photogrammetric techniques. However, the space covered by a single image may be geo-referenced in several different ways, each appropriate for specific applications. It may, therefore, be necessary to store additional geometrical parameters in the database to suit these (Table 22.3).

Table 22.3 Geometrical parameters which can be added to databases to assist rectification.

Geometrical parameter	Application
Centre point	Retrieval
Viewer's eyepoint	Retrieval
Camera location, etc.	Photomontage
View direction	Environmental impact analysis, etc.
View envelope	Environmental impact analysis, etc.

If a particular geographical database is to serve multiple applications, then individual images will have to be described in each of these different ways. Moreover, different types of image (vertical, oblique, ground level, etc.), may require particular types of spatial referencing, and image sets and moving images will require yet more complex approaches to ensure effective linkage between the images and the database (see, for example, Baker 1990).

Another aspect of the image integration problem concerns the most appropriate means of representing the non-spatial content of images. One approach that holds promise for the future is the automatic scene interpretation techniques developed for industrial robots and autonomous vehicles. For the moment, however, manual methods of extracting and encoding the thematic content of images are more commonly used, if they are used at all.

The Community Disk of the BBC Domesday interactive video system provides an example of the way in which both systematic spatial referencing and a looser form of geographical association can coexist (Openshaw and Mounsey 1987; Rhind and Openshaw 1987; Maguire 1989). Users are able to search for a location by 'map walking' (i.e. zooming and panning visually across a seamless map of the United Kingdom), or by entering a place name which is checked against a gazetteer of some 250 000 place names. When an appropriate map is found, the user may then display photographs and pages of text associated with some part of the area covered by that map. Alternatively, users can search for textual material by entering appropriate keywords, and the associated photographs and map are retrieved and displayed by the system. Photographs may also be searched in this way, and the associated text and maps recovered for inspection (see Fig. 22.7).

The system is able to perform this linkage by means of a thesaurus of terms describing the contents of the text and photographs stored on the disk. By means of this mechanism, users can move rapidly from one type of information to the other.

This three-way cross-indexing of contrasting information is extremely powerful, and suggests a model for multimedia indexing of geographically locatable information. However, the Domesday model needs to be generalized, to provide *n*-way cross-indexing between a larger number of information types, to permit an arbitrary number of photographs to be linked to a map and text, and also to ensure that information about the links is not 'hard-wired' into the system.

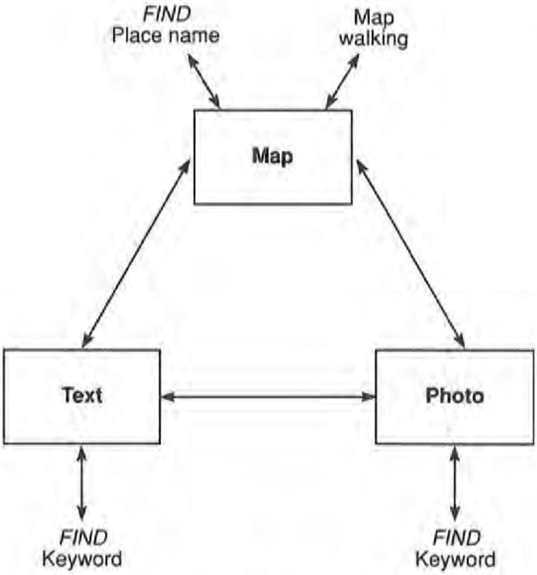


Fig. 22.7 Three-way information linkages in the BBC Domesday system.

One example of the direction this research is taking is the attempt by Ruggles (1990) to construct a generic solution to the problem of describing the links between images, so that standard query methods can be used to access them, and move from one to another. The problem is a complex one, because of the need to accommodate images with different geometries (vertical, oblique, ground level), individual images as well as sets of images, whole images as well as parts of images, single images as well as hierarchical collections of images, and tiled images as well as overlapping images. Nevertheless, systems such as Domesday demonstrate that highly effective multimedia GIS can be developed even where the locational links connecting items of information may not be precise, and where the geographical content of much of the information may be relatively low.

The Domesday approach to information integration also needs to be extended in other directions if it is to be more widely used as a basis for multimedia GIS. For example, the methods adopted for indexing the content of text and photographs need to be considerably more sophisticated, perhaps making use of artificial intelligence techniques that are being incorporated into commercial information retrieval systems. The indexing and linking of other types of real world information present additional problems. For this reason, the design of multimedia GIS will remain a fertile area for further research in the 1990s.

Interactive hypermedia systems

Another response to the emergence of multimedia technology that appeared during the 1980s is interactive multimedia (IMM). IMM is based on a combination of a personal computer and a mass digital storage device (commonly an optical disk such as a CD-ROM or WORM drive). The system captures, stores and displays any type of information that can be represented in digital form: sound, images, text, video, speech, maps, film, etc. (Ambron and Hooper 1988). As with earlier interactive video systems from which they derive many of their characteristics (Bayard-White 1985; Duke 1983; Parsloe 1983; Laurillard 1989), a major feature of IMM systems is that they handle large volumes of data. A critical factor in their success has, therefore, been the availability of an appropriate software environment which enables non-expert users to browse interactively through these data.

The currently preferred browsing model for multimedia information in such systems is represented by hypertext software (Conklin 1987; McAleese 1989; Nielsen 1990b; Hall and Papadopoulos 1990). Initially suggested by Bush in the 1940s (Bush 1945), and developed by Ted Nelson in his Xanadu project in the 1970s (Nelson 1981, 1987), hypertext was popularized by the HyperCard software distributed free with Apple Macintosh computers from the mid 1980s. Hypertext software enables users to follow chains of linked items of information, by pointing on screen to such graphical devices as 'buttons' and 'hot spots'. What hypertext does with items of text, hypermedia does with multimedia information: graphics, animation, sound, video, etc.

The key notion underlying hypertext and hypermedia is that all information and knowledge is inter-connected, and that users should therefore be able to wander freely through a universe of stored information in order to create 'islands of meaning' by making the appropriate associations. The user, rather than the computer, 'integrates' the various units of information in a hypertext or hypermedia system, by mentally linking together information retrieved on a particular theme. In many hypermedia systems, the user need not follow the links provided by the system designer; tools are provided for the user to modify existing links and to forge entirely new links between 'chunks' of information. In this way, the user can create personal pathways or networks in the stored database. This is sometimes referred to as the 'reader as author' approach to information browsing.

There are several major differences in the way information is handled in conventional GIS and hypermedia systems:

- Any type of information may be handled in hypermedia systems, drawn from a variety of media.
- Hypermedia treats any suitable 'chunk' of information as a basic data unit, whether it be a single word, an image, a multi-layered table of numeric data, a sequence of sounds, or a motion video extract. This contrasts strongly with the database model commonly used in conventional GIS, which demands highly structured attribute data, arranged in standard tables, and typically uses pointers to refer to non-standard information.
- The content of individual items of information (i.e. the nodes) is usually fixed in hypermedia systems. This is in contrast to GIS which can reconstitute information on the basis of (say) data aggregation or spatial overlay.
- In hypermedia systems, there is no explicit model of what the stored information means. In a database schema, some attempt is made to encode the meaning of data stored in the database.
- In a conventional GIS, the various sets of attribute data are systematically tied to a spatial data model using a common geo-referencing system. In a hypermedia system, such precise spatial referencing is less commonly found. Although IMM systems have been developed which permit point-located features (such as banks, shops or properties) to be displayed on a map, a more relaxed form of spatial referencing is more typically adopted. Thus, for example, a photograph may be related to a map because they both show a particular street; a recording of a bird's call may be linked to a map of a large expanse of salt marsh which forms that bird's habitat; or a computer simulation of the

hydrological cycle may be linked to a map of an entire drainage basin.

- Hypermedia systems permit data units to be linked together in any way felt appropriate by the system designer and/or the system user. For example, the links between units of information in multimedia systems are often arbitrary, subjective and/or tenuous, and mostly serve a non-analytical purpose. (This idea goes far beyond the notion of multiple schemata in database technology.)
- In a hypermedia system, the user plays a major role in making sense of stored information, by switching rapidly from one unit of information to another, following any number of associative links to piece together a collage of information that serves a particular purpose, and by re-configuring links between data units as needs dictate (see Fig. 22.8). Essentially, therefore, hypermedia provides a browsing model for GIS (Marchionini and Schneiderman 1988). This contrasts with conventional GIS in which the software provides a toolbox of ready-made analytical techniques, most of which are designed to store, retrieve, link and compare spatially coordinated data in a systematic way.

Typical geographical applications for hypermedia technology include education and training, technical documentation, information browsing, and product catalogues (e.g. a 'browser' catalogue of GIS software or data sets). A training example is the GIST system (Raper and Green 1989), which was designed as an interactive demonstrator for basic GIS concepts. An educational example is provided by Ecodisk (McCormick and Bratt 1988), an optical disk which contains a wealth of information on a nature reserve in southern England. Finally, the use of hypermedia technology to provide flexible user navigation facilities for a distributed environmental database is provided by the Whole Earth Decision Support System (WEDSS) prototype (Davis and Deter 1990).

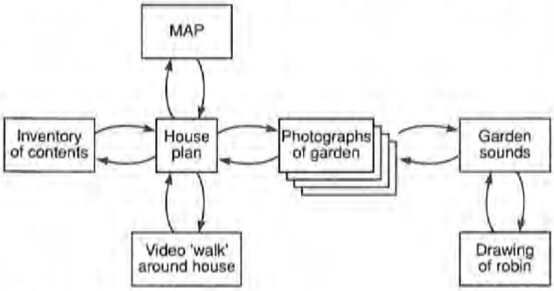


Fig. 22.8 Browsing through a multimedia GIS.

The challenge ahead is to combine the hypermedia approach to handling multi-source information with mainstream GIS. The rudimentary hypermedia ideas embedded in the Domesday system have already been described. Another approach, which aims to combine hypermedia techniques with the more traditional map-based approach of the GIS, is the 'geomatic hypermap' (Laurini and Milleret-Raffort 1989) in which a map is used as a graphical point-of-entry or front-end to information organized as hyperdocuments. A similar approach is taken by Wallin (1990), who suggests that the map might take the role of the 'home hyperdocument' in a hypermedia GIS.

Virtual reality systems

During the first two decades or so of GIS development, the user has been presented with (confronted by?) a highly unnatural representation of the real world involving a number of stylizations or abstractions. There is a strong emphasis on the two-dimensional, on a vertical or bird's eye view, highly symbolized feature representation, well structured data, and static time slices. Moreover, tight rules are usually imposed on how the user can interact with the computer model, and on the methods that are available to browse, navigate, query and manipulate the stored information.

Modern techniques of information gathering, particularly in the field of remote sensing, have greatly expanded the range of human sensory perception. It is ironic, therefore, that this sensory amplification has been accompanied by a corresponding sensory deprivation within GIS. In the GIS model, the direct human experience of the world, involving all of the senses, is replaced by an experience that involves the use of only a limited range of the senses – frequently, only the sense of sight.

Moreover, the informational richness of the real world is shoe-horned into a small number of information types within GIS, most notably numeric, textual and visual. Thus, for example, sounds are not heard, but are measured as point-located decibel readings and displayed as contours; colours are conventionalized or falsified; gradual change in the environment is replaced by widely spaced time slices; living conditions in an inner-city slum are summarized in numerical indexes of multiple deprivation; and the surrogate map walker expends no physical effort, and feels no physical strain, when following a footpath up hill and down dale on a computer screen. This is clearly not the fault of GIS alone, but it is perhaps a reflection of the way in which GIS have tended to reflect the approaches to information and information handling adopted in contemporary science.

Because of the limited sensory bandwidth of current GIS representations of the world, only a limited number of user responses are invoked, with a bias towards the cognitive rather than the affective and, more particularly, a bias towards visual information processing of a relatively abstract kind. This raises two questions: first, whether it is possible to provide users with a fuller sensory experience of the world, from their own desktop; and secondly, whether it is appropriate and useful to do so.

On the technical front, several developments have occurred during the 1980s, in a range of different fields, which suggest that it will soon be technically feasible to provide users with a multi-sensory representation of the world, involving a far wider range of information than hitherto. Most of these developments have been aimed at producing realistic computer graphic displays of real-world objects (e.g. Schacter 1983; Magnenat-Thalman and Thalman 1987; Muller *et al*. 1988). Others have focused on the development of multi-sensory interfaces to computer models (e.g. Foley 1987; Brooks 1987; Ware 1990; Ware and Osborne 1990). Among the relevant emerging technologies and their major application areas are: data visualization (scientific analysis), computer photomontage (civil engineering), building 'walk throughs' (architecture), environmental 'fly overs' (flight and battlefield training), surrogate walks (travel, education and training), hypertext/media (education), and multisensory simulations (arcade games and flight simulation).

Not only do many of these technologies integrate a far richer set of information than hitherto, and generate photo-realistic images, but some also permit users to interact directly with the computer representation of the world, using a fuller range of senses and faculties to do so. In the recently developed technologies associated with 'cyberspace' and 'alternative realities' (Gibson 1984; CADalyst 1989), prosthetic devices such as data gloves and head-mounted visors are used to manipulate the virtual world directly, and users (or 'participants') receive feedback in the form of continuous motion stereoscopic display, tactile response, and even whole body motion. This takes the concept of the 'direct manipulation interface', previously coined in the context of graphical user interfaces (Shneiderman 1983), to its logical conclusion.

The emerging technologies make it possible, perhaps for the first time, for GIS developers to consider integrating a complete range of information about the world in their systems and, moreover, to provide this information in a form that addresses the same senses that would be used by the observer or analyst when confronting the real world at first hand. Unlike current GIS, the emerging technologies integrate information in a more comprehensive manner, and attempt to present an environment for action that is fully three dimensional, is represented naturalistically, and in which changes occur in real time. Moreover, the user's role is considerably extended: users may interact directly with the environment, modify the environment, and receive multisensory information on the state of the environment, as well as multisensory feedback as a result of taking certain actions.

Of course, a considerable amount of research and development is needed before such facilities can be successfully provided for the purpose of geographical problem solving. Little work has been done, for example, on how to involve the human senses of touch, taste or smell within computer information systems. A wide-ranging discussion is also needed on the advisability of adopting such technologies for real world problem solving, as opposed to game playing in a simulated environment.

This is not the place to argue the relative merits of naturalistic representations of the world and more abstract and symbolized models. Nor is it appropriate to launch a discussion on the ethical or moral issues involved in providing users with surrogate multisensory experiences of the kind anticipated by Aldous Huxley's 'feelies'.

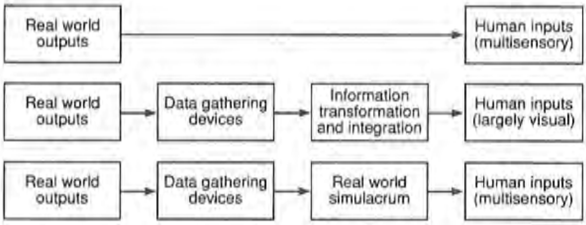


Fig. 22.9 Information flows between the real world and the geoscientist.

Nevertheless, it is important to consider in what circumstances, and for which applications, the heavily stylized model of the world provided by contemporary GIS should be retained, and those for which the more naturalistic, multisensory representation would be more appropriate. The alternatives are summarized in Fig. 22.9. Evidence is already available that in applications such as travel previews, estate agency demonstrations and emergency training, the provision of surrogate walks on an interactive video system, with their more naturalistic representation of space, is preferable to (and, more importantly, more useful than) the more abstract representation provided by the map. In work carried out on vehicle navigation (Mark 1987), it has been similarly found that a map-based display is not the most effective means of communicating information about tasks such as way finding.

Perhaps GIS developers should sidestep this decision altogether, and leave it to the user to decide. If both types of representation are made available in GIS, together with a simple means of switching from one to the other, then users can judge for themselves when it is most appropriate to use each model of the world.

Towards the 'omni-informational' GIS

It is possible that none of the alternative approaches discussed earlier will provide an appropriate model for information integration within GIS. Critics of hypermedia, for example, point to the fact that many users of hypertext and hypermedia systems become 'lost in hyperspace' when they begin to explore vast quantities of information. Much recent research has been aimed at refining this browsing model, either by constraining the freedom of users to ramble aimlessly through stored information, or else by providing orientation and structuring aids such as indexes, 'maps', tours, and quizzes (Weyer and Borning 1985; Hammond 1989; Hammond and Allinson 1987, 1989; Nielsen 1990a).

In the meantime, other developments may yet provide the impetus for more effective information integration in GIS. For example, a greater degree of integration than that afforded by the RDBMS environment may be provided by object-oriented database management systems (OODBMS). In one such system, for example, not only are multimedia objects (images, audio messages, etc.) defined as a hierarchy of classes, but so too are the devices (microphone, camera, screen, etc.) which capture, store and present those objects (Woelk and Kim 1987). In this system, a message-passing protocol is also established for interaction between these classes, and both the class hierarchies and the message protocol may be modified and/or extended by the user. In the 1990s, the hypermedia systems may well converge with GIS that adopt the object-oriented approach.

Alternatively, it may be necessary to look to developments in the field of electronic messaging and document handling, which have led to the creation of a number of standards that define the structure of multimedia messages and documents. The office document architecture (ODA), for example, is an ISO standard that defines the information content and structure of documents (IOS 1986). This currently permits a document to contain text, and geometrical and raster images, but an extension to the standard also allows colour, spreadsheets and sound. Yeorgaroudakis (1990) argues that future GIS software will need to adapt more effectively to office environments where information sharing is the norm, and this might imply adoption of emerging office information standards such as ODA, and the embracing of a far broader range of non-geographical information than hitherto.

Finally, although it is now generally accepted that information integration is best ensured by adopting a single, unifying data model, McKeown and Lai (1987) suggest that multiple data representations should be adopted by all spatial information systems. They argue that several internal representations should be adopted for spatial databases, each applied where it is most appropriate, and that a coordinating item of software should be developed to provide users with a uniform interface to these internal representations.

CONCLUSIONS

Over the past 25 years or so, considerable progress has been made in enabling GIS to integrate diverse information from a variety of sources. Currently, integration is usually achieved by adopting a coherent data model (typically the composite map model or the geo-relational model), which ensures that both spatial and attribute information are synthesized in any problem-solving or decision-making operation carried out by users.

This approach to integration provides major rewards to the user, but it is not without its problems. For example, it is clear that integration is not a magical by-product of pouring diverse data into a GIS; it is usually the result of considerable effort in overcoming the mutual intransigence of data drawn from various sources. Before a GIS can be used to interrelate diverse information, that information has first to be made comparable, compatible and consistent, and this usually involves considerable human effort.

There are also limitations on the extent to which current GIS can integrate information. Most current GIS betray their cartographic origins, and are consequently myopic in respect of the information they are able to integrate. In particular, most GIS are poor at integrating information through the three dimensions of the real world, across time, and for information about the world (such as environmental sounds) that have not been easy to handle using traditional computer facilities. Other systems reflect their underlying database technology, and are therefore highly selective in the types of information they handle. Most are designed to handle highly structured information, and link information by means of unambiguous locational references. Few are designed to handle unconventional information, that is only loosely geo-referenced, and is associated with other information by non-spatial links.

There are several alternative ways in which diverse information can be integrated in a computer-based information system. The traditional GIS approach has been to use a spatial location code as the primary key to relate all data. However, other methods which involve a looser, non-locational approach to information cross-referencing are also proving to be effective. It has also been shown how alternative user interfaces can considerably broaden the scope for user-centred information integration. Whether these approaches will lead to the appearance of several distinct types of GIS (e.g. the multimedia GIS, the hypermedia GIS and the multisensory GIS) is an open question.

It is certainly the case, however, that different approaches to data integration will be found to be more appropriate for specific applications and users. Thus, for example, spatial analysts might require a consistent spatial data representation, or facilities for converting between alternative spatial representations; those engaged in environmental monitoring, by contrast, might value facilities for the automatic update of linked information in real time; while environmental management trainees and the general public might be better served by interactive facilities for information navigation and browsing, and by tools that enable the arbitrary association of loosely connected information.

In the final analysis, information integration is not simply a function of smart algorithms, standard locational referencing schemes, unifying data models, powerful computers and multimedia storage devices. It is, perhaps above all else, the result of empowering users to explore alternative ways of linking the previously unlinked, and of enabling them to do this by a process of conjecture and refutation. Successful information integration is thus an essentially human activity, which may be carried out in private by the lone analyst, or across networks by communities of decision makers. Effective information integration thus hinges crucially on the ability of spatial analysts and decision makers to interact with one another, through the medium of computer-based information systems, as part of the process of making sense of the world about them.

REFERENCES

Abler R F (1987) The National Science Foundation Center for Geographic Information and Analysis. *International Journal of Geographical Information Systems* **1** (4): 303–26

Ambron S, Hooper C (1988) *Interactive Multi-media.* Microsoft Press, Redmond, WA

Anon (1990) The French Revolution of 1989. *Mapping Awareness* **4** (9): 48–9

Annoni A, Ventura A D, Mozzi E, Schettini R (1990) Towards the integration of remote sensing images within a cartographic system. *Computer Aided Design* **22** (3): 160–6

Aronson P (1987) Attribute handling for geographic information systems. *Proceedings of AUTOCARTO 8.* ASPRS/ACSM, Falls Church Virginia, pp. 346–55

Aybet J (1990) Integrated mapping systems – data conversion and integration. *Mapping Awareness* **4** (6): 18–23

Baker H H (1990) Scene structure from a moving camera. In: Blake A, Troscianko T (eds.) *AI and the Eye.* John Wiley, Chichester England, pp. 229–60

Bayard-White C (1985) *An Introduction to Interactive Video.* National Interactive Video Centre, and Council for Educational Technology, London

Beard M K, Chrisman N R (1986) Zipping: new software for merging map sheets. *Proceedings ACSM (vol I)* ACSM, Falls Church Virginia **1**: 153

Bedell R (1988) *WARP: a program to warp computer drawings, maps and plans.* Terra Investigations and Imaging Ltd, Guildford Surrey

Berry J K (1985) Computer-assisted map analysis: fundamental techniques. *Proceedings of the 6th Annual NCGA Conference*, Volume 2, pp. 369–86

Berry J K (1987) Fundamental operations in computer-assisted map analysis. *International Journal of Geographical Information Systems* **1** (2): 119–36

Bracken I, Higgs G (1990) The role of GIS in data integration for rural environments. *Mapping Awareness* **4** (8): 51–6

Bracken I, Webster C (1989) Towards a typology of geographical information systems. *International Journal of Geographical Information Systems* **3** (2): 137–51

Briggs D J, Mounsey H M (1989) Integrating land resource data into a European geographical information system. *Applied Geography* **9** (1): 5–20

Brooks F P (1987) Grasping reality through illusion: interactive graphics serving science. CHI '90. *SIGCHI Bulletin* (special issue): 1–11

Burrough P A (1986) *Principles of Geographical Information Systems for Land Resources Assessment.* Clarendon Press, Oxford

Burrough P A (1989) Fuzzy mathematical methods for soil survey and land classification. *Journal of Soil Science* **40**: 477–92

Bush V (1945) As we may think. *Atlantic Monthly* **176**: 101–8

CADalyst (1989) When reality is not enough. *CADalyst* December: 40–53

Casley D J, Lury D A (1981) *Data Collection in Developing Countries.* Clarendon Press, Oxford

Chrisman N R (1984) The role of quality information in the long-term functioning of a geographical information system. *Cartographica* **21**: 79–87

Chrisman N R (1991) The error component in spatial data. In: Maguire D J, Goodchild M F, Rhind D W (eds.) *Geographical Information Systems: principles and applications.* Longman, London, pp. 165–74, Vol 1

Conklin J (1987) Hypertext: an introduction and survey. *IEEE Computer* **20**: 17–41

Cowen D J (1986) PC-CAD manages geographical data. *Computer Graphics World* **9** (7): 38–41

Cowen D J (1988) GIS versus CAD versus DBMS: what are the differences? *Photogrammetric Engineering and Remote Sensing* **54**: 1551–5

Dangermond J (1986) The software toolbox approach to meeting the user's needs for GIS analysis. *Proceedings of the GIS Workshop, Atlanta, Georgia, 1–4 April 1986*, pp. 66–75

Dangermond J (1988) A review of digital data commonly available and some of the practical problems of entering them into a GIS. *Proceedings of ACSM–ASPRS St Louis.* ACSM/ASPRS, Falls Church Virginia

Dangermond J (1989) The organizational impact of GIS technology. *ARC News* Summer: 25–6

Davis F W, Simonett D S (1991) GIS and remote sensing. In: Maguire D J, Goodchild M F, Rhind D W (eds.) *Geographical Information Systems: principles and applications.* Longman, London, pp. 191–213, Vol 1

Davis J S, Deter R S (1990) Hypermedia application: Whole Earth Decision Support System. *Information and Software Technology* **32** (7): 491–6

Digital Cartographic Data Standards Task Force (DCDSTF) (1988) The proposed standard for digital cartographic data. *The American Cartographer* **15**: 24–8

DoE (Department of the Environment) (1987) *Handling Geographic Information. Report of the Committee of Enquiry chaired by Lord Chorley.* HMSO, London

Duecker K J (1985) Geographic information systems: towards a georelational structure. *Proceedings of AUTOCARTO 7.* ASPRS/ACSM, Falls Church, Virginia, 172–75

Duke J (1983) Interactive video: implications for education and training. *CET Working Paper 22.* Council for Educational Technology, London

Egenhofer M J, Herring J R (1991) High-level spatial data structures for GIS. In: Maguire D J, Goodchild M F, Rhind D W (eds.) *Geographical Information Systems: principles and applications.* Longman, London, pp. 227–37, Vol 1

ESRI (1989) Integration of geographic information technologies. *ARC News* Winter: 24–5

ESRI (1990a) ESRI, IBM, and HTE develop interactive link between ARC/INFO and IBM AS/400 Parcel Management System. *ARC News* Winter: 30

ESRI (1990b) *Understanding GIS: the ARC/INFO method.* ESRI Inc, Redlands California

Flowerdew R (1991) Spatial data integration. In: Maguire D J, Goodchild M F, Rhind D W (eds.) *Geographical Information Systems: principles and applications.* Longman, London, pp. 375–87, Vol 1
Foley J D (1987) Interfaces for advanced computing. *Scientific American* **257** (4): 83–90
Frank A U (1986) Integrating mechanisms for storage and retrieval of land data. *Surveying and Mapping* **46**: 107–21

Gibson W (1984) *Neuromancer* Ace Science Fiction.
Goodchild M F (1976) Statistical aspects of the polygon overlay problem. In: Dutton G (ed.) *Harvard Papers on Geographical Information Systems*, Volume 6. Addison-Wesley, Reading Massachusetts, pp. 1–22
Goodenough D G (1988) The integration of remote sensing and geographic information systems. In: Damen M C J, Smit, G S, Verstappen (eds.) *Symposium on Remote Sensing for Resource Development and Environmental Management, Balkema, Rotterdam*, pp. 1015–28
Gorte B, Liem R, Wind J (1988) The ILWIS software kernel. *ITC Journal* **19**: 15–22
Green N P (1990) Towards truly distributed GIS. *Proceedings of the GIS Design Models and Functionality Conference.* Midlands Regional Research Laboratory, University of Leicester, 8pp.
Guptill S C (1991) Spatial data exchange and standardization. In: Maguire D J, Goodchild M F, Rhind D W (eds.) *Geographical Information Systems: principles and applications.* Longman, London, pp. 515–30, Vol 1

Hall P A V, Papadopoulos S (1990) Hypertext systems and applications. *Information and Software Technology* **32** (7): 477–90
Hammond N (1989) Hypermedia and learning: who guides whom? In: Maurer H (ed.) *Computer Assisted Learning.* Springer-Verlag, Berlin
Hammond N, Allinson L (1987) The travel metaphor as design principle and training aid for navigating around complex systems. In: Diaper D, Winder R (eds.) *People and Computers III.* Cambridge University Press, Cambridge, pp. 75–90
Hammond N, Allinson L (1989) Extending hypertext for learning: an investigation of access and guidance tools. In: Sutcliffe A, Macaulay L (eds.) *People and Computers V.* Cambridge University Press, Cambridge, pp. 293–304
Haywood P (1986) National transfer standards for Great Britain. *Land and Minerals Surveying* **4** (11): 569–78
Healey R G (1991) Database management systems. In: Maguire D J, Goodchild M F, Rhind D W (eds.) *Geographical Information Systems: principles and applications.* Longman, London, pp. 251–67, Vol 1
Heuvelink G B M, Burrough P A, Stein A (1990) Propagation of errors in spatial modelling with GIS. *International Journal of Geographical Information Systems* **3** (4): 303–22

IOS (International Organisation for Standardisation) (1986) *Office Document Architecture (ODA).* ISO 8613. ISO, Geneva

Jackson M J (1987) Digital cartography, image analysis, and remote sensing: towards an integrated approach. *Interdisciplinary Science Reviews* **12**: 33–44
Jackson M J, Mason D C (1986) The development of integrated geo-information systems. *International Journal of Remote Sensing* **7**: 723–40

Laurillard D (ed.) (1989) *Interactive Media: working methods and practical applications.* Ellis Horwood, Chichester England
Laurini R, Milleret-Raffort F (1989) Principles of geomatic hypermaps *Ekistics* **56** (338–39): 312–17
Liley R (1985) Integration – the big pay-off for geobased municipal systems. *Papers of the Urban and Regional Information Systems Association – URISA '85.* URISA, Ottawa, Canada **2**: 11–27
Lodwick W A, Monson W, Svoboda L (1990) Attribute error and sensitivity analysis of map operations in geographical information systems suitability analysis. *International Journal of Geographical Information Systems* **4** (4): 413–28
Logan T L, Bryant N A (1987) Spatial data software integration: merging CAD/CAM mapping with GIS and image processing. *Photogrammetric Engineering and Remote Sensing* **53** (10): 1391–5

McAleese R (ed.) (1989) *Hypertext: theory into practice.* Intellect Books, Oxford
McCormick S, Bratt P (1988) Some issues relating to the design and development of an interactive video disc. *Computers in Education* **12** (1): 257–60
McKeown D M, Lai R C T (1987) Integrating multiple data representations for spatial databases. *Proceedings of AUTOCARTO 8.* ACSM/ASPRS, Falls Church Virginia, pp. 754–63
McLaren R (1990) Establishing a corporate GIS from component data sets – the database issues. *Mapping Awareness* **4** (2): 52–8
Magnenat-Thalman N, Thalman D (1987) An indexed bibliography on image synthesis. *IEEE Computer Graphics and Applications* **7** (8): 27–37
Maguire D J (1989) The Domesday interactive videodisc system in geography teaching. *Journal of Geography in Higher Education* **13** (1): 55–68
Maguire D J (1991) An overview and definition of GIS. In: Maguire D J, Goodchild M F, Rhind D W (eds.) *Geographical Information Systems: principles and applications.* Longman, London, pp. 9–20, Vol 1
Marchionini G, Schneiderman B (1988) Finding facts versus browsing knowledge in hypertext systems. *Computer* **3** (1): 70–80
Mark D M (1987) On giving and receiving directions: cartographic and cognitive issues. *Proceedings of AUTOCARTO 8.* ACSM/ASPRS, Falls Church Virginia, pp. 562–71
Marx R W (1986) The TIGER system: automating the

geographic structure of the United States census. *Government Publications Review* **13**: 181–201

Morehouse S (1985) ARC/INFO: a geo-relational model for spatial information. *Proceedings of AUTOCARTO 7*. ACSM/ASPRS, Falls Church Virginia, pp. 388–97

Mounsey H M (ed.) (1988) *Building Databases for Global Science*. Taylor & Francis, London

Mounsey H M (1991) Multisource, multinational environmental GIS: lessons learnt from CORINE. In: Maguire D J, Goodchild M F, Rhind D W (eds.) *Geographical Information Systems: principles and applications*. Longman, London, pp. 185–200, Vol 2

Muller J-C (1977) Map gridding and cartographic errors: a recurrent argument. *Canadian Cartographer* **14**: 152–67

Muller J-P, Day T, Kolbusz J, Dalton M, Richards S, Pearson J (1988) Visualisation of topographic data using video animation. In: Muller, J-P (ed.) *Digital Image Processing in Remote Sensing*. Taylor & Francis, London, pp. 21–38

Nelson T (1981) *Literary Machines*. (2nd edn) Theodore Holm Nelson, Swarthmore

Nelson T (1987) *Computer Lib*. Microsoft Press, Redmond, WA

Newell R G, Theriault D, Easterfield M (1990) Temporal GIS – modelling the evolution of spatial data. *Proceedings of the GIS Design Models and Functionality Conference*. Midlands Regional Research Laboratory, Leicester, 10pp.

Nicholson R (1990) Public access to spatial information: the use of value added networks in the UK. *Proceedings of EGIS '90*, Volume 2. EGIS Foundation, Utrecht, pp. 782–8

Nielsen J (1990a) The art of navigating through hypertext. *Communications of the Association of Computing Machinery* **33** (3): 296–310

Nielsen J (1990b) *Hypertext and Hypermedia*. Academic Press, San Diego California

Nyerges T L (1989a) Information integration for multipurpose land information systems. *URISA Journal* **1** (1): 27–38

Nyerges T L (1989b) Schema integration analysis for the development of GIS databases. *International Journal of Geographical Information Systems* **3** (2): 153–83

Openshaw S, Mounsey H M (1987) Geographic information systems and the BBC's Domesday interactive videodisk. *International Journal of Geographical Information Systems* **1** (2): 173–9

Parsloe E (ed.) (1983) *Interactive Video*. John Wiley, Chichester England

Pascoe R T, Penny J P (1990) Construction of interfaces for the exchange of geographic data. *International Journal of Geographical Information Systems* **4** (2): 147–56

Peuquet D J (1981a) An examination of techniques for reformatting digital cartographic data. Part 1: The raster-to-vector process. *Cartographica* **18**: 34–48

Peuquet D J (1981b) An examination of techniques for reformatting digital cartographic data. Part 2: The vector-to-raster process. *Cartographica* **18**: 21–33

Peuquet D J (1984) A conceptual framework and comparison of spatial data models. *Cartographica* **21**: 66–113

Pevsner S (1989) Image processing in a GIS environment. In: Barrett E C, Brown K A (eds.) *Remote Sensing for Operational Applications*. The Remote Sensing Society, Nottingham England, pp. 323–30

Piwowar J M, Le Drew E F, Dudycha D J (1990) Integration of spatial data in vector and raster formats in a geographic information system. *International Journal of Geographical Information Systems 4* (4): 429–44

Press L (1990) Compuvision or teleputer? *Communications of the Association for Computing Machinery* **33** (9): 29–36

Raper J R, Green N P A (1989) GIST: an object-oriented approach to a GIS tutor. *Proceedings of AUTOCARTO 9*. ACSM/ASPRS, Falls Church Virginia, pp. 610–19

Rhind D W, Openshaw S (1987) The BBC Domesday system: a nation-wide GIS for $4448. *Proceedings of AUTOCARTO 8*. ACSM/ASPRS, Falls Church Virginia, pp. 595–603

Rhind D W, Tannenbaum E (1983) Linking census and other data. In: Rhind D W (ed.) *A Census User's Handbook*. Methuen, London, pp. 287–300

Rhind D W, Green N P, Mounsey H M, Wiggins J S (1984) The integration of geographical data. *Proceedings of Austra Carto Perth*, Volume 1. Australian Institute of Cartographers, Perth, pp. 237–53

Robinson G M, Gray D A, Healey R G, Furley P A (1989) Developing a geographical information system (GIS) for agricultural development in Belize, Central America. *International Journal of Geographical Information Systems* **9** (2): 81–94

Ruggles C L N (1990) An abstract model for the structuring of a spatially indexed set of images. *Proceedings of EGIS '90*, Volume 2. EGIS Foundation, Utrecht, pp. 948–57

Saalfeld A (1988) Conflation: automated map compilation. *International Journal of Geographical Information Systems* **2** (3): 217–28

Schacter B J (1983) (ed.) *Computer Image Generation*. John Wiley, New York

Shepherd I D H (1990a) Mapping with desktop CAD: a critical review. *Computer Aided Design* **22** (3): 136–50

Shepherd I D H (1990b) Computer mapping: 21 roles for AutoCAD. *Bulletin of the Society of University Cartographers* **23** (2): 1–15

Shepherd I D H (1990c) Build your own desktop GIS? *Land and Minerals Surveying* **8** (4): 176–83

Shetler T (1990) Birth of the BLOB. *BYTE* **15** (2): 221–6

Steinitz C F, Parker P, Jordan L (1976) Hand-drawn overlays: their history and prospective uses. *Landscape Architecture* **66**: 444–55

Steneker M, Bonham-Carter G F (1988) *Computer Program for Converting Arc-Node Vector Data to Raster Format*. Geological Survey of Canada, K1Z 8R7, 300pp.

Tomlin C D (1990) *GIS and Cartographic Modelling*. Prentice-Hall, Englewood Cliffs New Jersey

Tomlin C D (1991) Cartographic modelling. In: Maguire D J, Goodchild M F, Rhind D W (eds.) *Geographical Information Systems: principles and applications*. Longman, London, pp. 361–74, Vol 1

Vose M (1990) Hot links to go. *BYTE* **15** (12): 373–7

Wagner G (1990) SICAD: profile of a raster indexed topological vector GIS. *Proceedings of the Conference on GIS Models and Functionality*. Midlands Regional Research Laboratory, University of Leicester

Walker P A, Hutton P G (1986) Grid cell representation of soil maps: an Australian example. *Australian Geographical Studies* **24** (2): 210–21

Wallin E (1990) The map as hypertext: on knowledge support systems for the territorial concern. *Proceedings of EGIS '90*, Volume 2. EGIS Foundation, Utrecht, pp. 1125–34

Wang F, Hall G B, Subaryono (1990) Fuzzy information representation and processing in conventional GIS software: database design and application. *International Journal of Geographical Information Systems* **4** (3): 261–83

Ware C (1990) Using hand position for virtual object placement. *The Visual Computer* **6** (5): 245–53

Ware C, Osborne S (1990) Exploration and virtual camera control in virtual three dimensional environments. *1990 Symposium on Interactive 3D Graphics: Computer Control* (special issue): in press

Waugh T C, Healey R G (1986) The GEOLINK system, interfacing large systems. In: Blakemore M J (ed.) *Proceedings of AUTOCARTO London*, Volume 1. Royal Institution of Chartered Surveyors, London, 76–85

Webster C (1988) Disaggregated GIS architecture: lessons from recent developments in multi-site database management systems. *International Journal of Geographical Information Systems* **2** (1): 67–79

Weyer S A, Borning A H (1985) A prototype electronic encyclopedia. *ACM Transactions on Office Information Systems* **31** (1): 63–88

Woelk D, Kim W (1987) Multimedia information management in an object oriented database system. In: Stocker P M, Kent W (eds.) *Proceedings of the 13th Very Large Databases Conference, Brighton*, pp. 319–29

Yeorgaroudakis Y (1990) The GIS of the future. *Proceedings of EGIS '90*, Volume 2. EGIS Foundation, Utrecht, pp. 1188–99

Zarzycki J M, Jiwani Z (1986) Canadian standards for exchange of digital topographic data. *Proceedings of the XVIII FIG Congress, Commission V*: 171–181

Zobrist A L (1983) Integration of Landsat image data with geographic databases. In: Peuquet D J, O'Callaghan J (eds.) *Proceedings of the United States/Australia Workshop on Design and Implementation of Computer-based Geographic Information Systems*. IGU Commission on Geographical Data Sensing and Processing, Amherst New York, pp. 51–63