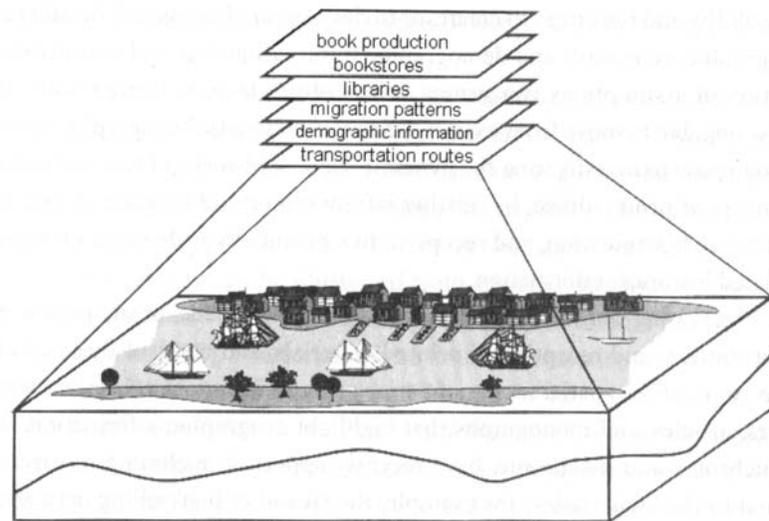


Chapter 2: The World as Viewed through a GIS

2.1 Introduction

This chapter gives a detailed description of the data models that GIS software packages use to store their data. Little of what is said here is directly historical; however, it is important to present a brief overview of the way in which GIS models the world. The chapter also establishes the terminology used throughout this book, as this varies between different authors and different software packages. There are many good introductions to GIS that are appropriate for historians. In 1991 Maguire, Goodchild and Rhind edited a large collection of papers on GIS in two volumes that rapidly became the standard reference work in GIS (Maguire *et al.* 1991). In 1999 a completely re-written second edition of this book was published, edited by Longley, Goodchild, Maguire and Rhind (Longley *et al.* 1999). Most of the first edition of the ‘big book’, as it became known, was then made available online. In 2001 the four authors responsible for the second edition produced a monograph on GIS that provides a wide-ranging introduction to GIS at a more affordable price than the edited works (Longley *et al.* 2001). Other good introductions include Chrisman (2002), DeMers (2000), Heywood *et al.* (2002) and Martin (1996a).

As described in Chapter 1, GIS software uses a conventional database, termed *attribute data* (sometimes termed ‘non-spatial data’). Each record in the attribute data is given a spatial reference using coordinates: these coordinates are termed the *spatial data*. There are two main types of GIS data model: in *vector data models* the spatial data consist of *points*, *lines* or *polygons*; while in *raster data models* the study area is usually sub-divided into square *pixels*, although other regular *tessellations* such as hexagons or triangles can also be used. In a vector model, therefore, space is sub-divided into discrete features, while the raster model attempts to represent space as a continuous *surface*. There are also other ways of sub-dividing the world into surfaces such as *triangular irregular networks* (TINs), often used to represent relief (the altitude of the earth’s surface) through *digital terrain models* (DTMs, also known as digital elevation models or DEMs).



[Abstracting the real world into layers for the study of book history]

Rather than store all their information about a study area in a single structure, the GIS separates the world into different *layers* (also known as coverages or themes) with each layer representing a different type of information. Most GIS texts give examples such as data on relief being one layer, data on the road network another, data on settlements another, data on rivers another, and so on. MacDonald and Black (2000) give an historical example. They are interested in exploring the development of print culture in the 19th century. They argue that the spread of books was partly due to factors directly related to books themselves, including sources of production such as printers and sources of supply such as libraries and bookshops. In addition the spread of books was also influenced by broader factors such as the transport network and demographic factors such as migration. Figure 2.1 shows how they can abstract the real world into layers in order to create a GIS that will allow them to explore book culture.

Understanding that GIS models the world through spatial and attribute data, that spatial data is made up of a small number of crude *graphic primitives*, and that data on different themes are brought together using layers is critical to an understanding of GIS. This basic model fundamentally defines what can and cannot be done within a GIS and the strengths and weaknesses of all GIS operations. There is an alternative approach to modelling the world in GIS using an *object-orientated* approach. Here features are not subdivided into layers but are instead grouped into classes and hierarchies. Although there are theoretical advantages to doing this, few commercially available GIS software packages have fully implemented this approach. For this reason object-orientated models will not be discussed in this guide.

2.2 Attribute data

Attribute data are data in the form that most people understand by the term. Many GIS software packages include their own attribute databases but allow the user to link to external *database management systems (DBMS)* or spreadsheets. Examples of this include *MapInfo*, allowing the user to link to data in Microsoft Excel, *ArcView* to DBase, or *ArcInfo* to Oracle. Attribute data are frequently either statistical or textual, however they can be in virtually any format that the DBMS used to store them can support. Increasingly this includes image formats, animations, hyperlinks, multimedia, and so on.

Most of the DBMSs used in GIS are *relational database management systems*. This means that two or more tables can be joined together based on a common field known as a *key*. With historical data this is often either a place name or an ID number (note that place names are not considered to be spatial data: to be spatial the data must have a coordinate-based location), and it allows data from various sources to be integrated without requiring spatial data. For example, a user has a table of Poor Law data organised by Poor Law Unions (these were a type of administrative unit used in England and Wales in the 19th and early 20th centuries to administer relief of the poor), a table of voting statistics organised by parliamentary constituency, and some employment statistics based on towns. A *relation join* will join all three tables together and all of the data for 'Bristol', for example, will appear on a single row.

There are three main problems with doing this: firstly, the join has no knowledge that the entity referred to as 'Bristol' may be a different entity in each table. Secondly, problems will occur with names such as 'Whitchurch', which is not unique and appears in both Hampshire and Shropshire. Different software will handle this in different ways, the most common (and theoretically sound) being to duplicate rows of data. One way round this is to use more than one column as the key, for example, place name and county. The third problem with using place names is that their spellings must be identical to produce a match. Even minor differences in the use of hyphens or apostrophes will cause a non-match. This can be worked around using gazetteers that standardise all possible spellings and create a single spelling from an authority list, or through the use of ID numbers. Creating these can be time consuming.

Many attribute databases use *Structured Query Language (SQL)* to allow flexible querying and joining. This is often implemented though a Graphical User Interface but follows the basic structure:

```
select <column names>
from <table names>
where <condition>
```

So, for example, we have two tables: ‘unemp’ that contains data on unemployment rates, and ‘inf_mort’ that contains data on infant mortality. These have the following fields:

Unemp		Inf_mort	
Field Name	Contents	Field Name	Contents
Name	The area’s place name	Name	The area’s place name
U_rate	The area’s unemployment rate	Im_rate	The area’s infant mortality rate
Tot_pop	The area’s total population	Births	The number of births in the area

Table 2.1: Sample tables of attribute data

The SQL query:

```
select name, u_rate, tot_pop
from unemp
where u_rate>10.0
```

will select the names, unemployment rates, and total populations from the table unemp for places with an unemployment rate of over 10% as is shown in Table 2.2.

Name	U_rate	Tot_pop
Bolton	13.2	10,000
Oldham	12.1	7,500
Rochdale	10.9	8,000

Table 2.2: Sample data returned by the query above.

Relational joins are also implemented in this way. For example the query:

```
select unemp.name, unemp.u_rate, inf_mort.im_rate
from unemp, inf_mort
where unemp.name=inf_mort.name
```

will select the unemployment rates from `unemp` and infant mortality rates from `inf_mort` where the values in the name fields in both tables are identical.

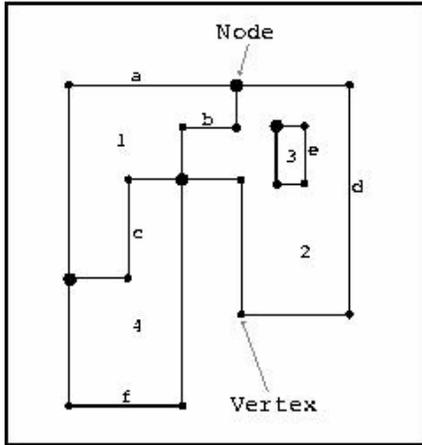
Unemp.name	Unemp. U_rate	Inf_mort. im_rate
Bolton	13.2	120
Oldham	12.1	115
Rochdale	10.9	106
Burnley	9.4	98
Colne	9.1	89

Table 2.3: Sample data returned by the query above

While relational databases and SQL are not fundamental to an understanding of GIS, knowledge of them can be useful in order to enhance an understanding of GIS data and GIS software. Many guides to the use of SQL and relational databases are available: see the bibliography for further information.

2.3 Vector systems

In vector systems spatial data are represented either by *points*, *lines*, or *polygons*. A point is represented using a single *coordinate pair*. A line (or arc or segment) is represented by a string of coordinate pairs giving the start and end point of the line and the coordinates of all points where the line changes direction. A line's start and end points are often referred to as *nodes*. Polygons are created by completely enclosing an area using one or more lines. A basic polygon data model is shown in Figure 2.2. The method of connecting lines to create this model is known as *topology*. Both polygons and lines need topological information. A polygon needs to know the ID numbers (or letters in the case of figure 2.2) of all the line segments that make up its boundaries, while a line segment usually knows which polygons are to its left and its right. For example in Figure 2.2, polygon 1 is bounded by line segments *a*, *b* and *c* while polygon 3 is only bounded by line segment *e*. Line segment *c* has polygon 1 to its left and polygon 4 to its right.



Polygons - line segments

- 1: a, b, c
- 2: b, d, e
- 3: e
- 4: c, f

Lines - co-ordinates

- a: (1.36, 4.38) (1.36, 8.62) (5.12, 8.62)
- b: (5.12, 8.62) (5.12, 7.65) (4.03, 7.65)
- ...
- f: (1.36, 4.38) (1.36, 1.55) (4.03, 1.55)

[Simplified structure of polygons in a vector GIS]

Topology is required if polygons are used, as without it boundaries would have to be stored twice: for example, line segment *b* would have to be stored as part of polygon 1's boundary and also as part of polygon 2's. It is important to note that the topological data model means that polygons cannot overlap and that every location in the study area can only belong to one polygon. Topology can also be used to turn a collection of lines into a *network*. This is usually based on each node knowing which line segments it is connected to, so in other words the nodes represent junctions. This can be extremely useful as it means that transport and other networks can be used, for example, to determine the shortest path between two points.

Spatial data:

Attributed data:

	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="text-align: center;">Polygon id</th> <th style="text-align: center;">Parcel no.</th> <th style="text-align: center;">Owner</th> <th style="text-align: center;">Value</th> </tr> </thead> <tbody> <tr> <td style="text-align: center;">1</td> <td style="text-align: center;">437</td> <td style="text-align: center;">A. Elliot</td> <td style="text-align: center;">2201</td> </tr> <tr> <td style="text-align: center;">2</td> <td style="text-align: center;">617</td> <td style="text-align: center;">J. Sutcliffe</td> <td style="text-align: center;">1432</td> </tr> <tr> <td style="text-align: center;">3</td> <td style="text-align: center;">133</td> <td style="text-align: center;">A. Elliot</td> <td style="text-align: center;">5025</td> </tr> <tr> <td style="text-align: center;">4</td> <td style="text-align: center;">991</td> <td style="text-align: center;">A. Smith</td> <td style="text-align: center;">1692</td> </tr> </tbody> </table>	Polygon id	Parcel no.	Owner	Value	1	437	A. Elliot	2201	2	617	J. Sutcliffe	1432	3	133	A. Elliot	5025	4	991	A. Smith	1692
Polygon id	Parcel no.	Owner	Value																		
1	437	A. Elliot	2201																		
2	617	J. Sutcliffe	1432																		
3	133	A. Elliot	5025																		
4	991	A. Smith	1692																		

[Linking spatial and attribute data]

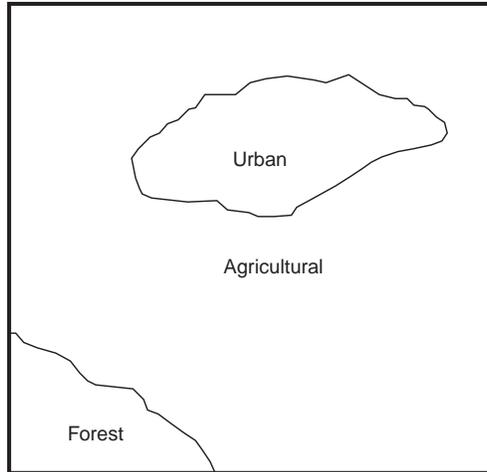
Attribute data is linked to points, lines, and polygons using a relational join on ID numbers as is shown in Figure 2.3. In modern software the creation and maintenance of topology, its associated ID numbers, and the links between the spatial and attribute data are kept hidden from the user. Nevertheless it is important to have some understanding of how the data are assembled.

Some features, such as towns or buildings, can be represented by either a polygon or a point. The choice of which representation to use is dictated by the purpose they are to be used for and, in particular, the scale of use. Polygons can be represented by points if required, usually by using the centre point of the polygon, termed the *centroid*.

The Great Britain Historical GIS (Gregory *et al.* 2002; Gregory and Southall 2000), SECOS (Gatley and Ell 2000), the historical GIS of Belgian territorial structure (De Moor and Wiedemann 2001) and the United States National Historical GIS all have attribute databases containing a variety of census and other data. These are linked to polygon representations of the administrative units used to publish the data such as registration districts in England and Wales, and counties in the United States. In this way the data are put on the map and the map is populated with data.

2.4 Raster systems

The *raster data model* sub-divides *space* into square *pixels* or other regular *tessellations* to provide a continuous representation of the study area rather than subdividing it into discrete points, lines, or polygons. For example, to represent relief each pixel might have its height above sea-level as an attribute; to represent land-use each pixel would have a land-use class attached to it, and so on. Although less likely to be of use to historians, *satellite images* are a more complex form of raster data. On the image the earth's surface is sub-divided into pixels with each pixel storing information about the amount and type of light being reflected by that part of the earth's surface, for example, how much green, how much blue, how much red, and how much infra-red.



1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	2	2	2	2	1	1	1	1
1	1	1	1	1	2	2	2	2	2	2	2	2	1	1
1	1	1	1	2	2	2	2	2	2	2	2	2	1	1
1	1	1	1	2	2	2	2	2	2	1	1	1	1	1
1	1	1	1	1	1	1	2	2	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3	3	1	1	1	1	1	1	1	1	1	1	1	1	1
3	3	3	3	1	1	1	1	1	1	1	1	1	1	1
3	3	3	3	3	1	1	1	1	1	1	1	1	1	1
3	3	3	3	3	3	1	1	1	1	1	1	1	1	1

[The raster data model]

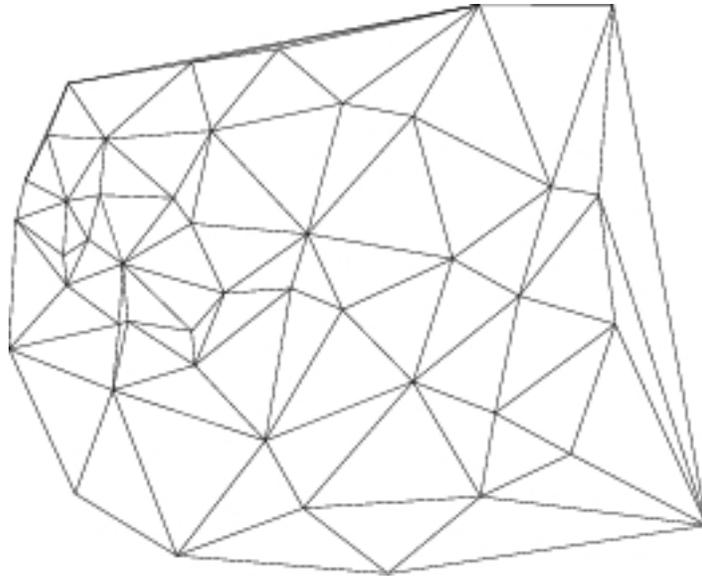
The simplest raster file formats are two-dimensional arrays in which each value corresponds to a pixel on the grid that the array is modelling. The header information provides data on, for example, pixel size and the location of the bottom left-hand corner of the grid, while the remainder of the file simply consists of the values for each cell. Figure 2.4 shows an example of this. The original map shows three types of land-use: agriculture, urban and forest. The raster sub-divides the study area into pixels with each pixel being allocated a numeric value representing land-use class. The pixel is allocated to the type of land-use that covers most of its area. Obviously the choice of pixel size is very important to the model. Too large a size will lead to a poor representation of the features, too small a size will lead to unwieldy file sizes. More sophisticated methods, such as *run length encoding* and *quadtrees*, can be used to compress the file sizes, but the

basic model remains the same. For more information on the details of different raster structures see, for example, chapter 3 of Heywood *et al.* (2002).

In general, raster data are more suited to environmental applications while vector data are more suited to human activity. Raster systems model complex spatial patterns with limited attributes, such as land-use patterns, very well, while the vector data model is better for more clearly defined space with complex attributes such as census data. There are exceptions to this: Martin (1996b) uses derived raster surfaces to model 1981 and 1991 census data and compare change between the two, claiming that the raster model provides a more realistic model of the underlying population distribution. A good example of a raster system in a historical context is provided by Bartley and Campbell (1997). They examined the Inquisitions Post-Mortem of the 14th century and used these to create a raster GIS of medieval land-use that, they claim, is potentially the most detailed survey possible until the 19th century tithe surveys. A raster system was used because it provides a complete coverage of the land area, it provides a more realistic representation of land-use than polygons, and because it handles the inaccuracies of the sources better than polygons.

2.5 Other systems: terrain modelling with TINs

A *triangular irregular network* (TIN) can be created from point data that, in addition to x and y coordinate values, also have a z value that usually represents altitude. Spot-heights are the most common source of data in this form. The first stage in generating a TIN is to plot all the x,y coordinates as points. Next a straight line is drawn between each point and its nearest neighbours. This is shown in Figure 2.5. As the heights of all the points are known, the gradient of the line of these lines is also known. If data are available from enough points, these can be used to generate a surprisingly realistic, apparently three-dimensional, representation of the shape of the landscape, known as a *digital terrain model* (DTM). Often raster and vector data are *draped* over the DTM to provide features on the landscape. Models such as these are often shown on news programmes; for example, the siege of Sarajevo was regularly pictured this way.



[A Triangulated Irregular Network (TIN)]

One use of DTMs in historical research is to recreate historical landscapes. This has been done for a number of reasons including exploring how a landscape would have appeared to people in the past (Harris 2002), recreating historical battlefields to investigate visibility and decision making during battles (Knowles, 2004), or to bring old maps to life (Rumsey and Williams 2002).

2.6 Bringing it all together with layers

Each separate theme of information is represented as a layer in the GIS. To build up a realistic representation of the study area a variety of layers are commonly used. For example, a DTM may be used to model the relief of an area, a raster surface to model its land-use, a point layer used to represent buildings of interest, a line layer to represent rivers, a network layer to represent the transport system, and polygon layers to represent field patterns and administrative boundaries. All of these contain the relevant attributes. Exactly when to use a separate layer is often a matter of choice. For example, if we have three types of buildings – churches, hospitals, and museums – these could be stored as three separate layers, or they could be stored on a single layer with attributes to say what type of building each point represents.

Most common GIS software packages can handle most of these types of data although the majority will concentrate on one. *MapInfo* and *ArcGIS* are both primarily vector systems that have limited raster and terrain modelling functionality. *Spans* and *Idrisi* on the other hand are primarily raster based.

A good example of a study area modelled through a GIS is provided by Pearson and Collier's work on the parish of Newport in Pembrokeshire (Pearson and Collier 1998). They were interested in land ownership and agricultural productivity in the mid-19th century. To investigate this they needed to combine environmental data such as soils, slope, aspect and altitude, with historical statistics such as census data, and information from the tithe survey, a detailed and comprehensive survey of land ownership and land-use carried out in England and Wales during the early 19th century. Information from the tithe surveys was recorded as polygons. Each field was represented as a polygon with attribute information such as the owner, the occupant, the field name, the state of cultivation, the acreage and the tithe rent-charge. The environmental data started as altitudes for a regular grid of points 50m apart covering an area of 10km by 10km. This was used to create both a raster grid for which each cell had an altitude, and a terrain model that allowed slope and aspects to be calculated. Bringing the data together in this way allowed a detailed reconstruction of the various factors affecting agricultural productivity in the mid-19th century, and provided a starting point for more sophisticated analyses.

A second example is provided by Siebert (2000). He was interested in reconstructing the history of Tokyo's development with a particular emphasis on physical features such as the shoreline and rivers, the administrative areas, the population distribution and the rail network. To do this he created a variety of layers for each type of feature: the rail network was represented by line layers; the shoreline and the rivers were represented either using lines (for small rivers) or polygons (for larger ones); and the administrative areas were clearly represented as polygons, as were the census data. As with Pearson and Collier's work, bringing the data together in this way allows a comprehensive picture of the study area to be built up.

2.7 Conclusions

The core idea behind GIS is that it attempts to represent features that are, or were, on the earth's surface. The first stage in doing this is to take real-world themes that are relevant to the research and abstract them as layers, with each theme being represented by a separate layer. Features on each layer are represented using a combination of spatial and attribute data with each layer's spatial data consisting of one of four types of graphic primitives: points, lines, polygons or pixels. The resulting data are therefore highly abstracted from the real world. Where the data can be effectively represented using this model, GIS provides a powerful research tool because of its unique ability to handle

location and attribute simultaneously. However, this structure fundamentally restricts the way that GIS can be used and where features cannot be effectively represented in this way the use of GIS may not be appropriate.

The implementation of most of what has been described in this chapter is performed automatically by modern GIS software and is well hidden from the user, but it is important to understand the way in which GIS represents the world. Many datasets simply do not fit this model comfortably. If they don't, it is unlikely to be worth investing much effort in GIS. Where they do, this model represents a powerful way of integrating space into an analysis or exploration of a particular location, be it an area of only a few square metres, a whole country, or even the whole world. In addition to the primitive method of representing space, many other issues also need to be borne in mind. These include scale, accuracy and uncertainty. The GIS research community expends large amounts of energy worrying that GIS users will not understand these limitations. They should, however, be familiar to historians as they are the limitations of the original source plus the additional limitations created by converting the source into digital form. The problem is that for spatially referenced data these are more complicated than they may appear at first sight. The issues involved will be discussed in more detail in subsequent chapters.

Further reading from chapter 2:

References giving in **bold** are key references.

Good introductions to GIS

Burrough, P.A. and McDonnell, R., 1998. *Principals of Geographical Information Systems for land resources assessment. 2nd edition.* Oxford: Clarendon Press.

Chrisman, N.R., 2002. *Exploring Geographic Information Systems. 2nd edition.* Chichester: Wiley.

Clarke, K.C., 1997. *Getting started with Geographic Information Systems.* New Jersey: Prentice Hall.

DeMers, M.N., 2000. *Fundamentals of Geographic Information Systems. 2nd edition.* Chichester: Wiley.

Heywood, D.I., Cornelius, S. and Carver, S., 2002. *An introduction to Geographic Information Systems. 2nd edition.* Harlow: Longman.

Longley, P.A., Goodchild, M.F., Maguire, D.J. and Rhind, D.W., 1999, eds. *Geographical Information Systems: principals, techniques, management and applications. 2nd edition.* Chichester: John Wiley.

Longley, P.A., Goodchild, M.F., Maguire, D.J. and Rhind, D.W., 2001. *Geographic Information Systems and Science.* Chichester: John Wiley.

Jones, C. 1997. *Geographic information systems and computer cartography.* Harlow: Addison Wesley Longman.

Maguire, D.J., Goodchild, M.F. and Rhind, D.W., 1991, eds. *Geographical Information Systems: Principles and Applications.* London: Longman Scientific and Technical. Available online at: <http://www.wiley.co.uk/wileychi/gis/volumes.html>

Martin, D., 1996a. *Geographic Information Systems and their socio-economic applications. 2nd edition.* Hampshire: Routledge.

GIS on the web:

GIS World Wide Web resource list:
<http://www.geo.ed.ac.uk/home/giswww.html>

The Association for Geographic Information (AGI): <http://www.agi.org.uk/>

The Center for Spatially Integrated Social Science: <http://www.csiss.org>

The ESRI virtual campus: <http://campus.esri.com/>

The Geography Network: <http://www.geographynetwork.com/>

Relational database management systems and SQL

Date, C.J., 1995. *An introduction to database systems.* Reading, MA: Addison-Wesley.

Healey, R.G., 1991. Database management systems. In: D.J. Maguire, M.F. Goodchild and D.W. Rhind, eds. *Geographical Information Systems: principles and applications. Volume 1: principles*. Harlow: Longman, 1991, 251-267. Available online at: <http://www.wiley.co.uk/wileychi/gis/volumes.html>

Oxborrow, E.A., 1989. *Databases and database systems: concepts and issues*. 2nd edition. Bromley: Chartwell-Bratt.

Worboys, M.F., 1995. *GIS: A computing perspective*. London: Taylor & Francis. Chapter 2.

Worboys, M.F., 1999. Relational databases and beyond. In: P.A. Longley, M.F. Goodchild, D.J. Maguire and D.W. Rhind, eds. *Geographical Information Systems: principals, techniques, management and applications*. 2nd edition. Chichester: John Wiley, 1999, 373-384

GIS data models

DeMers, M.N., 2002. *GIS modelling in raster*. Chichester: Wiley.

Egenhofer, M.J. and Herring, J.R., 1991. High-level spatial data structures for GIS. In: D.J. Maguire, M.F. Goodchild and D.W. Rhind, eds. *Geographical Information Systems: principles and applications. Volume 1: principles*. Harlow: Longman, 1991, 227-237. Available online at: <http://www.wiley.co.uk/wileychi/gis/volumes.html>

Peuquet, D.J., 1990. A conceptual framework and comparison of spatial data models. In: D.J. Peuquet and D.F. Marble, eds. *Introductory readings in Geographic Information Systems*. London: Taylor & Francis, 1990, 250-285

Weibel, R., and Heller, M., 1991. Digital terrain modelling. In: D.J. Maguire, M.F. Goodchild and D.W. Rhind, eds. *Geographical Information Systems: principles and applications. Volume 1: principles*. Harlow: Longman, 1991, 269-297

See also all the introductions to GIS listed above

Case studies

Bartley, K. and Campbell, B.M.S., 1997. Inquisitiones Post Mortem, GIS, and the creation of a land-use map of medieval England. *Transactions in GIS*, 2, 333-346. See also Chapter 7.

Gatley, D.A. and Ell, P.S., 2000. *Counting heads: an introduction to the census, poor law union data and vital registration*. York: Statistics for Education. See also chapter 6.

Gregory, I.N. and Southall, H.R., 2000. Spatial frameworks for historical censuses – the Great Britain Historical GIS. In: P.K. Hall, R. McCaa and G. Thorvaldsen, eds. *Handbook of historical microdata for population research*. Minneapolis: Minnesota Population Center, 2000, 319-333. See also Chapter 5.

Gregory, I.N., Bennett, C., Gilham, V.L. and Southall H.R., 2002. The Great Britain Historical GIS: From maps to changing human geography. *The Cartographic Journal*, 39, 37-49. See also Chapter 5.

Harris, T.M., 2002. GIS in archaeology. In: A.K. Knowles, ed. *Past time, past place: GIS for history*. Redlands, CA: ESRI Press, 2002, 131-143. See also Chapter 6.

Knowles, A.K., 2004. *Visualising Gettysburg: problems in historical terrain modelling*. Paper presented to the 29th Annual Meeting of the Social Science History Association

MacDonald, B.M., and Black, F.A., 2000. Using GIS for spatial and temporal analysis in print culture studies: some opportunities and challenges. *Social science history*, 24, 505-536

Martin, D., 1996b. Depicting changing distributions through surface estimations. In: P. Longley and M. Batty, eds. *Spatial analysis: modelling in a GIS environment*. Cambridge: GeoInformation International, 1996, 105-122. See also Chapter 6.

Miller, D. and Modell, J., 1988. Teaching United States history with the Great American History Machine. *Historical methods*, 21, 121-134

Pearson, A. and Collier, P., 1998. The integration and analysis of historical and environmental data using a Geographical Information System: landownership and agricultural productivity in Pembrokeshire c. 1850. *Agricultural history review*, 46, 162-176

Rumsey, D. and Williams, M., 2002. Historical maps in GIS. In: A.K. Knowles, ed. *Past time, past place: GIS for history*. Redlands, CA: ESRI Press, 2002, 1-18.

Siebert, L., 2000. Using GIS to document, visualize, and interpret Tokyo's spatial history. *Social science history*, 24, 537-574

The China Historical GIS project: <http://www.people.fas.harvard.edu/~chgis/>

The Great Britain Historical GIS project: <http://www.gbhgis.org/>

The United States Historical GIS project: <http://www.nhgis.org>