Chapter 3: Acquiring Spatial Data

3.1 Introduction

There is no cheap and easy way of getting spatial data into a GIS. In many cases the *data capture* phase of a GIS-based project is the longest and most expensive of all phases. Bernhardsen (1999) suggests that the collection and maintenance of data accounts for 60 to 80% of the total cost, in terms of both money and time, of a fully operational GIS. Other authors give higher figures. Siebert (2000), for example, gives an honest account of the time taken and problems encountered building his historical GIS for Tokyo.

There are two basic sources of spatial data: *primary sources* where the data can be captured directly into the GIS, for example through the use of *Global Positioning Systems* (GPS); or *remote sensing* from satellites. More common for historians are secondary sources capture, where data from paper maps are converted into digital form. There are two ways of doing this: *scanning* the maps to produce *raster data*; or *digitising* the maps where points or line features are traced either directly from paper maps or from scanned images of the maps. This produces *vector data*. It is worth noting that the GIS definitions of primary and secondary data sources are different from the definitions used by historians. From a GIS perspective, primary data are data captured directly from the real world while secondary data are captured from abstracted sources such as maps. A map produced in the 1850s is a primary source from a historian's point of view, but a secondary source from a GIS perspective.

An alternative to capturing data yourself is to acquire digital data from someone else. This is less time consuming and less risky but spatial data are often expensive, may have serious copyright restrictions placed on their use, and need to be fit for the purpose that you require. In the context of historical research, many datasets that may be required have not been digitised owing to a lack of demand for them, so there is frequently no alternative but to capture them yourself.

The first, and probably most important, decision to take when acquiring digital spatial data is what the source should be. This applies whether the data are to be captured in-house or acquired from others. The limitations of the original source inherently limit any subsequent use of those data. The scale of the source is of particular importance here. In general, larger scale data are more flexible than smaller scale but will be more expensive, either in terms of purchase price or the time taken to capture them, and will have more redundant information which can lead to problems with file sizes. A second issue to consider is whether the reasons for the production of the original maps are compatible with the objectives of any digital representation.

3.2 Scanning maps to produce raster data

Scanning is a relatively straightforward process in which paper maps are placed on a scanner and a raster copy is produced. Smaller scanners are often relatively inexpensive; however, larger ones suitable for large map sheets are still expensive. The spatial resolution of the scanner, usually expressed in dots per square inch (dpi), and its spectral resolution (the number of colours it can distinguish), must be borne in mind, as this will affect the accuracy of the resulting data.

Converting from the resulting scan to the type of raster data described in Chapter 2 will require a certain amount of post-processing. If the source maps are relatively simple and focused on a single theme, such as land-use or soil type, this may be quite straightforward, whereas converting from more complex sources can be very time consuming.

3.3 Digitising maps to produce vector data

While scanning produces a copy of the source map, digitising extracts certain features from the source and creates *point*, *line*, and/or *polygon layers* from them. From an early series of Ordnance Survey (OS) inch-to-the-mile maps, a user might simply want to extract a layer of points representing the locations of churches. Other layers, such as the road and rail networks, administrative boundaries, and so on may also be extracted. However, it is not possible to create a direct copy of the source map in the way that scanning does.

Points are digitised by clicking on the features that are required. Lines are digitised by tracing along the lines and clicking at points where there is a significant change in the line's direction. Polygons are created by creating a topological structure on top of line data, as described in Chapter 2. Digitising can be done either directly from the paper map, using a *digitising table* or *tablet*, or by first scanning the map and digitising on screen, known as *head-up digitising*. Digitising tables and tablets (tables are usually larger), often simply known as *digitisers*, consist of a flat surface on which the map is firmly stuck down. The tablet or table's surface has a fine mesh of copper wires underneath it. There is also a *puck*, a hand-held device with a fine cross-hair and one or more buttons. To capture a point the cross-hairs are placed over the feature of interest and a button is pressed. The tablet or table is able to calculate the exact location of the cross-hairs from this and the coordinate is passed to the GIS software. Head-up digitising does not require a digitiser as a scan of the map is already in digital form. Instead a cursor is placed over the feature on the screen using the computer's mouse and a button is clicked to determine its exact location. Head-up digitising has the advantage that it creates a scanned copy of the source and also the vector data extracted from it. This both preserves a copy of the source, and can be used as a *backcloth* that provides context for the extracted vector features.

Digitising accuracy is extremely important especially where topology is to be created. A node will often have to be digitised two or more times to represent the end point of one line segment and the start point of another. Most GIS software provides tolerances that mean that if two nodes are within a set distance they will be 'snapped' together to form a single node. If the tolerance is set too high inappropriate features will be snapped. If it is too low then gaps will appear, known as *dangling nodes*, where polygons will not close properly. This will lead to corrupted topology.

3.4 Geo-referencing

Whether data have been scanned or digitised, their underlying coordinate scheme at this stage of the data capture process will usually be in inches or centimetres measured from the bottom left-hand corner of the scanner or digitiser. *Geo-referencing* is the process by which these coordinates are converted into real-world coordinates on a *projection system*. This allows distances and areas to be calculated and data from different sources to be integrated.

Map projections are intricate and complicated. In a country such as Britain where the use of the National Grid is almost universal, a detailed understanding of projection systems is rarely necessary. In this guide only the briefest description will be given, further details can be found in works listed in the bibliography. The earth is a globe and locations on that globe are described using *latitude*, the number of degrees north or south of the equator, and *longitude*, the number of degrees east or west of the Greenwich meridian. Maps are flat sheets of paper. Projections are the translations used to convert from a curved earth to a flat map surface. Doing this involves distorting one or more of distances, angles, areas or shapes. A projection will also often convert from degrees of longitude and latitude to miles or kilometres from a particular location with longitude becoming the x-coordinate and latitude becoming the y-coordinate.

The British National Grid is a Transverse Mercator projection. The origin of its ellipsoid runs north-south at 2^{o} west of the Greenwich meridian, approximating to the central spine of Britain. As one moves east or west from this line, distances in particular become increasingly distorted. Fortunately, as Britain is a long, thin country running approximately north-south these distortions are rarely significant at the kind of scales at which historians operate. Traditionally, Britain was subdivided into grid squares with sides of 100km. Each square was given a two-letter identifier and locations were expressed as distances in kilometres or metres from the south-west corner of the grid square. For example, NN is a grid square in the southern highlands of Scotland. Within a computer, this use of letter codes is clumsy. Instead a false origin is given for the whole country. This is a point south-west of the Scilly Isles that allows all locations on mainland Britain to be expressed to the nearest metre as non-negative integers of no more than six figures. Location 253000, 720000 is 253km east of the false origin and 720km north and is in the southern highlands. This structure allows easy calculations of the distances between any two points expressed using six-figure National Grid references.

Most GIS software packages make the process of geo-referencing appear quite straightforward. Using the source map the user finds the real world coordinates of a number of *reference points*, usually four. These are also called *tic points* and are often the corners of the map sheet. These points are then digitised. The user is then prompted to type in the real-world coordinates of the reference points and the software uses these four points to convert every coordinate in the layer to real-world coordinates. Frequently the software will also prompt for a projection system at this stage and convert the layer accordingly. This means that the locations of all points on the layer will be expressed in National Grid coordinates or whatever other coordinate system has been selected, and that all distances measured on the layer will be in metres, kilometres, or whatever referencing unit is used.

While this is relatively easy using modern maps where coordinate grids are shown and projection information is readily available, it can be difficult using older maps. Where no coordinate grid is provided, this can be worked around by finding reference points on the source map that are also mapped on modern maps (appropriate features may include churches, lighthouses, trig-points or railway stations). The coordinates of these points can then be found using modern maps that do include a coordinate grid. For the sake of accuracy, the modern maps used should preferably be larger scale than the source map. If no information is provided on map projections then books such as Delano-Smith and Kain (1999), Harley (1975), Oliver (1993), Owen and Pilbeam (1992) are good sources of information. For work on local study areas, the impact of projections may be so small that they might not be relevant.

3.5 Error and accuracy

As was previously stated, a digital representation of a paper map is at best of equal quality to the original map, but it will almost inevitably accrue some additional error or inaccuracy. It is important to distinguish between the different types of error and inaccuracy. Unwin (1995) uses a six-way classification as follows: *error* is the difference between reality and the digital representation of it; *blunders* are simply mistakes; *accuracy* is the closeness of measurements, results, computations or estimates to values accepted as true; *precision* is the number of decimal places given in a measurement which is usually far more than it can support; *quality* is the fitness for purpose of the data; and *uncertainty* measures the degree of doubt or distrust when using the data.

Scanning, *digitising* and *geo-referencing* are particular sources of locational error. To digitise a map it must first be placed completely flat on the digitiser or scanner. Even this is not always as easy as it sounds: maps may have been stored folded, paper warps over time, and so on. The accuracy of the scanning or digitising equipment itself is the next possible source of error, although if specialised (normally expensive) hardware is used this is usually only minimal. If head-up digitising is used there will be cumulative error as there is the error introduced by the scanning, and then error introduced by the digitising.

The next source of error comes from the user's involvement with the data capture process. When digitising, the person capturing the data has to place the puck or cursor over the point to be captured. Even a highly motivated and alert person will make minor positional errors. When digitising is done for many hours by low-paid staff, the potential for both inaccuracy and blunders is increased. It is not always possible to capture the exact location accurately. Point symbols are not always a precisely defined point: for example railway stations on an OS 1:50,000 map are represented by a circle that is nearly 2mm in diameter. Line features are even more difficult. Digitising a line relies on the operator capturing each point at which the line changes direction. For gentle curves, such as those on roads, rivers or contour lines, this is inevitably a subjective choice and no two operators digitising a line of this type will ever digitise exactly the same points to describe it.

A final source of error is the geo-referencing process. Coordinates measured from a map will have a certain amount of error in them. The locations of the reference points will also have some error. This means that the placement of every location on the layer will be slightly distorted. Most software packages will provide a measure of the error expressed as *Root Mean Square* (RMS) error. This is often expressed in both digitiser units and also real-world units. It is recommended that the RMS error should not exceed 0.003 digitiser inches (ESRI 1994a). Even this standard is not always possible with historic maps, and it is important that minimum standards of accuracy are established as part of the data capture process and that these are documented (see Chapter 9). The key point to this section is that there are many issues associated with error and uncertainty in spatial data that may not be familiar to the historian. The problems these cause may seem daunting; however by facing up to the issues and being aware of their implications these should not cause serious difficulties to the research process.

3.6 Digitising attribute data

This is usually a more straightforward process than capturing spatial data. Attribute data are usually captured by scanning and Optical Character Recognition (OCR), or by typing. These can either be typed in directly against spatial features in the GIS software, or can be captured separately and joined to the spatial data using a relational join. For more information on good practice in capturing attribute data see Townsend *et al*. (1999).

3.7 Raster-to-vector and vector-to-raster data conversion

The holy grail of spatial data capture is simply to be able to scan a map and then automatically extract point, line and polygon features from the scanned data. This is called *raster-to-vector conversion*, but it has not proved easy to implement. Some systems claim to be able to do this but the process usually requires a large amount of user interaction. Even then it can be inaccurate: for example, lines can pick up a stepped effect as a result of following the underlying raster grid pattern. *Vector-to-raster conversion* can be more successful, but care still needs to be taken. Many software packages include routines that allow both of these conversions to take place, but many also make exaggerated claims about the potential accuracy of the routines. Before committing to a major purchase, or major investment in time, it is well worth checking the results first rather than believing the sales pitch provided by the software company.

3.8 Primary data sources

Two main primary data sources are data from *satellite imagery*, and *Global Positioning Systems* (GPS). Satellite imagery is a form of *raster data* in which each pixel represents a part of the earth's surface. The exact dimensions of the pixel depends on both the type of scanner used and the post-processing applied, but is usually a square with sides of from 1 metre to around 100 metres. For each pixel data are provided about the light that was reflected back from that part of the earth's surface when the image was captured. Again, the exact details depend on the type of scanner and the post-processing applied. From this basic information, more sophisticated knowledge can be developed concerning, for example, types of land-use, health of vegetation, and so on. It is not the intention to describe remotely sensed satellite imagery here, as many good guides are available, as listed in the further reading at the end of the chapter. Obviously, satellite imagery has only become available in relatively recent times but it may provide useful information for historians; for example, in determining the location of certain features in remote areas, or to attempt to determine past land-uses. Aerial photographs may be an alternative (secondary) source of data on land-use in the past.

GPS receivers are an easy way of primary spatial data capture. The simplest form of GPS receivers give the coordinate for the current location of the user. This is calculated from a network of satellites launched by the United States military. Initially the accuracy of these locations was deliberately degraded for non-military users to around 100m. This has now stopped and accuracies of only a few metres are available to all users. More sophisticated systems allow the user to capture multiple points as the receiver is moved and download them directly into a computer. If more accuracy is required *differential GPS* (DGPS) is used. This requires the use of two receivers, one of which is kept stationary at a known location to assist in measuring the location of the roving receiver. This can potentially produce sub-metre accuracy.

GPS can be used to provide a relatively quick method of surveying and capturing features on the earth's surface. A good example of this is Lowe's (2002) description of surveying surviving earthworks and other features in Spotsylvania National Military Park in the United States, the site of four Civil War battles.

3.9 Buying data or acquiring it free

For academics in the United Kingdom there are a variety of organisations that provide GIS data either free or at low cost. For historians the most useful of these is likely to be AHDS History (formerly the History Data Service) <http://www.ahds.ac.uk/history/index.htm>, part of the Arts and Humanities Data Service <http://www.ahds.ac.uk/>. EDINA at the University of Edinburgh <http://edina.ac.uk/> provides a suite of services including UKBORDERS, which disseminates contemporary and historical boundary data for the UK, and Digimap, which provides Ordnance Survey (OS) digital data. The MIMAS service at the University of

Manchester <http://www.mimas.ac.uk/> provides access to a variety of datasets, including Bartholomew's map data and a variety of satellite imagery. They also host modern census data through their CASWEB interface and have a variety of GIS and visualisation tools. The National Archives (formerly the Public Record Office) <http://www.nationalarchive.gov.uk/> is increasingly becoming involved in electronic publishing, although only a limited amount of data are currently available online.

Other sources of free or low-cost data include the Geography Network <http://www.geographynetwork.com/>, and the Digital Chart of the World <http://www.maproom.psu.edu/dcw/> which provide data on a variety of different countries; and the United States Geological Survey (USGS) <http://www.usgs.gov/>, which provides a variety of products often at relatively cheap rates. The Electronic Cultural Atlas Initiative (ECAI) <http://www.ecai.org/> produces online e-publications of historical datasets, and also incorporates a metadata clearing-house that allows users to search for historical GIS datasets.

A variety of commercial companies sell spatially referenced data. The most obvious source of this in Britain is the Ordnance Survey <http://www.ordsvy.gov.uk/> which sells digital versions of many of their map products. Other commercial sources include the AA <http://www.theaa.com/aboutaa/data_sales.html> and Bartholomew's <http://www.bartholomewmaps.com/>. Purchasing these data can be expensive and copyright limitations may be placed on their use, but buying the data does provide high-quality products quickly and without the risks involved in capturing it yourself.

Whatever the source of data it is important to bear two things in mind: the first is the limitations of the original source material and its fitness for the purpose that you want to use it for. The second is any limitations that are placed on the use of the data through copyright or other restrictions. Misuse of this may result in civil or criminal action and is also likely to lead to people and organisations being less willing to disseminate data.

3.10 Conclusions

Acquiring spatial data is usually either time consuming or expensive or both. It is not to be entered into lightly. Once spatial data are available in digital form, however, they become a powerful resource with uses that may go far beyond the original reason why they were captured. It is important to remember that digital spatial data have all the limitations of the original source data, be they maps or other sources, plus some additional limitations and errors introduced by the data capture process. As will be described in Chapter 9, good documentation and *metadata* are essential elements of the data capture process, as this allows other users to evaluate the fitness of the data for their purpose. This ensures that data are widely used but only for suitable purposes.

Further reading from chapter 3:

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

Scanning and digitising

All good introductory texts on GIS will cover this (see Chapter 2).

Bernhardsen, T., 1999. Choosing a GIS. 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, 589-600

Calkins, H.W., 1990. Creating large digital files from mapped data. In: D.J. Peuquet and D.F. Marble, eds. *Introductory readings in Geographic Information Systems*. London: Taylor & Francis, 1990, 209-214

Chrisman, N.R., 1990. Efficient digitizing through the combination of appropriate hardware and software for error detection and editing. In: D.J. Peuquet and D.F. Marble, eds. *Introductory readings in Geographic Information Systems*. London: Taylor & Francis, 1990, 233-246

Dangermond, J., 1990. A review of digital data commonly available and some of the practical problems of entering them into a GIS. In: D.J. Peuquet and D.F. Marble, eds. *Introductory readings in Geographic Information Systems*. London: Taylor & Francis, 1990, 222-232

Environmental Systems Research Institute, 1994a. *Arc/Info Data management: concepts, data models, database design, and storage*. Redlands, CA.: Environmental Systems Research Institute Inc.

Peuquet, D.J. and Boyle, A.R., 1990. Interactions between the cartographic document and the digitizing process. In: D.J. Peuquet and D.F. Marble, eds. *Introductory readings in Geographic Information Systems*. London: Taylor & Francis, 1990, 215-221

Error in spatial data

Chrisman, N.R., 1991. The error component in spatial data. In: D.J. Maguire, M.F. Goodchild and D.W. Rhind, eds. *Geographical Information Systems: principles and applications. Volume 1: principles.* Longman: Harlow, 1991, 165-174. Available online at: http://www.wiley.co.uk/wileychi/gis/volumes.html

Fisher, P.F., 1991. Spatial data sources and data problems. In: D.J. Maguire, M.F. Goodchild and D.W. Rhind, eds. *Geographical Information Systems: principles and applications. Volume 1: principles.* Longman: Harlow, 1991, 175-189. Available online at: http://www.wiley.co.uk/wileychi/gis/volumes.html

Fisher, P.F., 1999. Models of uncertainty in spatial data. 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, 191-205

Goodchild, M.F. and Gopal, S., 1989, eds. *The accuracy of spatial databases.* London: Taylor & Francis.

Unwin, D.J., 1995. Geographic Information Systems and the problem of 'error and uncertainty'. *Progress in human geography*, 19, 549-558

Veregin, H., 1999. Data quality parameters. 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, 177-189

Map Projections

Bugayevskiy, L.M., and Snyder, J.P., 1995. *Map projections: a reference manual*. London: Taylor & Francis.

Environmental Systems Research Institute, 1994b. *Map projections: geo-referencing spatial data*. Redlands, CA: Environmental Systems Research Institute Inc.

Maling, D.H., 1991. Coordinate systems and map projection for GIS. In: D.J. Maguire, M.F. Goodchild and D.W. Rhind, eds. *Geographical Information Systems: principles and applications. Volume 1: principles*. Longman: Harlow, 1991, 135-146. Available online at: http://www.wiley.co.uk/wileychi/gis/volumes.html

Monmonier, M., 1996. *How to lie with maps.* Chicago: University of Chicago Press.

Richardus, P., and Adler, R.K., 1972. *Map projections for geodesists, cartographers and geographers.* London: North Holland Publishing.

Robinson, A.H., Morrison, J.L., Muehrcke, P.C., Kimerling, A.J. and Guptill, S.C., 1995. *Elements of cartography*. 6th ed. Chichester: John Wiley & Sons.

Capturing attribute data

Harvey, C. and Press, J., 1996. *Databases in historical research: theory, methods and applications*. London: Palgrave Macmillan.

Townsend, S., Chappell, C. and Struijve, O., 1999. *Digitising history: a guide to creating digital resources from historical documents.* Oxford: Oxbow Books.

Introductions to remote sensing

Barnsley, M., 1999. Digital remotely-sensed data and their characteristics. 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, 451-466

Campbell, J., 1987. *Introduction to remote sensing*. New York: Guildford Press.

Curran, P., 1985. *Principals of remote sensing*. Harlow, Essex: Longman.

Lillesand, T. and Kiefer, R., 1994. *Remote sensing and image interpretation*. New York: John Wiley & Sons.

Areal photography and GIS

Wolf, P.R., 2000. *Elements of photogrammetry: with applications in GIS*. 3rd edition. London: McGraw-Hill.

Global Positioning Systems

Hofmann-Wellenhof, B., Lichtenegger, H., and Collins, J., 1994. *Global Positioning Systems: theory and practice*. 4th edition. Wien: Springer.

Kennedy, M., 1996. *The Global Positioning System and GIS: an introduction*. Chelsea, Mich.: Ann Arbor Press.

Sources of information about historical maps of Britain

Delano-Smith, C. and Kain, R.J.P., 1999, *English maps: a history*. London: British Library.

Harley, J.B., 1975, *Ordnance Survey maps: a descriptive manual*. Southampton: Ordnance Survey.

Oliver, R.R., 1993, *Ordnance Survey maps: a concise guide for historians*. London: Charles Close Society.

Owen, T. and Pilbeam, E., 1992. *Ordnance Survey: map makers to Britain since 1791*. London: HMSO.

The Charles Close Society: http://www.charlesclosesociety.org.uk/

Data capture and historical research

Healey, R.G. and Stamp, T.R., 2000. Historical GIS as a foundation for the analysis of regional economic growth: theoretical, methodological, and practical issues. *Social science history*, 24, 575-612. See also Chapter 4.

Knowles, A.K., 2000. Introduction. *Social science history*, 24, 451-470

Lowe, D.W., 2002. Telling Civil War battlefield stories with GIS. In: A.K. Knowles, ed. *Past time, past place: GIS for history.* Redlands, CA: ESRI Press, 2002, 51-63

Siebert, L., 2000. Using GIS to document, visualize, and interpret Tokyo's spatial history. *Social science history*, 24, 537-574. See also Chapter 2.

Sources of data

Australian Centre of the Asian Spatial Information and Analysis Network (ACASIAN): http://www.asian.gu.edu.au/
Bartholomew's: http://www.bartholomewmaps.com/

Digital Elevation Data catalogue: http://www.geo.ed.ac.uk/home/ded.html

Old Maps.co.uk: http://www.old-maps.co.uk/

The National Archives: http://www.nationalarchives.gov.uk/

The AA: http://www.theaa.com/services/bookshop_home.html

The Arts and Humanities Data Service: http://www.ahds.ac.uk/

The Arts and Humanities Data Service, History (was the History Data Service): http://www.ahds.ac.uk/history/

The David Rumsey Map Collection: http://www.davidrumsey.com/

The Digital Chart of the World: http://www.maproom.psu.edu/dcw/

EDINA(University of Edinburgh): http://edina.ed.ac.uk/

The Electronic Cultural Atlas Initiative: http://www.ecai.org

The MIMAS service: http://www.mimas.ac.uk/

The Ordnance Survey: http://www.ordsvy.gov.uk/

UK Data Archive: http://www.data-archive.ac.uk

UK Economic and Social Data Service: http://www.esds.ac.uk/

United States Geological Survey (USGS) Earth Explorer service: http://edcsns17.cr.usgs.gov/EarthExplorer/