

Chapter 8: Qualitative Data in GIS

8.1 Introduction

Chapter 7 described some well-developed quantitative techniques that could be used by a historian from a social science tradition. The use of qualitative data of the kind more likely to be more relevant to a researcher from the humanities end of the discipline is less well developed. The basic principles are the same: the GIS imposes a structure where data have spatial and attribute components and are organised in a series of layers, with each layer representing a different theme. This structure allows the user to explore the geographical relationships within the data. For historians working in social science this is likely to involve thematic mapping and analysis of statistical data in ways similar to those that have been well explored in other disciplines. For humanities scholars the data used are likely to be more complicated, comprising textual descriptions, images, photographs, maps, sound and video. Although these are different from statistical data, if they can be geo-referenced using points, lines, polygons or pixels, then GIS can be used to explore them. Mapping will again be an important part of this but so will querying the data to investigate the spatial relationships within and between layers of data and perhaps the calculation of summary statistics. Because far less research has been done in this area, this chapter is significantly shorter than some others. This reflects the fact that historians working with qualitative sources are defining the research agenda as they go rather than applying techniques that originated elsewhere.

8.2 Types of qualitative data in GIS

Many qualitative projects follow a data-led approach. Often a large amount of disparate information about a place or places is gathered and integrated using the locational characteristics of the data. By assembling the data in this way it becomes possible to explore them to form new insights. Qualitative information is often point-based, but can also involve lines or polygons. Examples of the types of qualitative attribute data that may be used in GIS include information on the presence or absence of features such as roads and buildings, documents concerning events that occurred at a place, pictures or photographs relating to places, and audio or video that refer to a specific location. Textual descriptions of places from historical manuscripts can also provide data for a qualitative GIS. Historical maps or plans could also be part of the qualitative attribute data, particularly if they have relevance to an area but cannot be geo-referenced in the manner of modern-style maps. Rather than store all the information

about a place in a single GIS, the GIS attribute data could consist of hyperlinks to websites that contain information about specific places or themes. All that is required to turn these into GIS data is a coordinate-based location.

Traditionally, the locational component of qualitative information is less likely to fit the precisely defined locations insisted on by GIS software. In some ways, this is less of a problem than it can be for quantitative data because the types of analyses that are performed are less likely to be as demanding of the spatial component than, for example, areal interpolation operations using quantitative data. In spite of this, the accuracy of the locational data, particularly data derived from different sources, still needs to be carefully considered.

The use of qualitative sources to date often represents a more sophisticated use of the *metadata* about an object. For example, the catalogue for a collection of scanned photographs is likely to have included some information about the place that the photographs refer to, usually place names. By converting these place names into point data we have the ability to create a layer of scanned photographs. As with other GIS applications, this allows us both to find more information about the collection of photographs themselves, for instance, by comparing images of locations that are near to each other; and it also allows us to integrate the collection of photographs with other collections of photographs or other types of information.

8.3 Case studies

A basic ability of GIS is to provide structure to a single layer of data. Core to this is the ability to query features based on their location. The *International Dunhuang Project* <<http://idp.bl.uk/>> provides a good example of this. It is an electronic archive of over 100,000 documents, document fragments and artefacts from sites along the Silk Road. Its catalogue includes facilities to either allow users to search the catalogue in the conventional way and then use maps to see where the artefact is from, or can be searched through a map interface such that the user progressively zooms into an area and is then able to select specific features from the map and explore their attribute data.

GIS can also be used to integrate data from different sources through the use of multiple layers. The Perseus Project <<http://perseus.csad.ox.ac.uk/>> is a good example of this (Smith *et al.* 2000). It attempts to integrate various electronic libraries and archives from around the world and again uses locational information as a key component of this. One of the libraries they were concerned with was Edwin C. Bolles' collection of the history of London <<http://perseus.csad.ox.ac.uk/cgi-bin/perscoll?collection=Bolles>>.

compiled in the late 19th century. It includes: printed sources, some of which are unique; folio descriptions of the city from limited print runs; contemporary 19th century maps; and illustrations and prints from the 17th to the 19th centuries. Again, using GIS allows the user to access these data efficiently and flexibly based on location. In the case of the Bolles' collection, it allows data from different sources to be brought together and explored. This involves having digital maps from different dates linked together, having place names identified in source texts with hyperlinks to other references to the same place, and having coordinate-based locations associated with images. The Sydney TimeMap project (Wilson, 2001) is a system with a similar architecture that is based in the Museum of Sydney and allows users to explore information about Sydney through a map interface.

In the projects described above the GIS is used primarily as a structuring and retrieval tool. The following two examples show it being used more directly as an exploratory analytic tool. One example focuses on the Salem witch-trials in Massachusetts in 1692. Ray describes an archive that includes complete transcriptions of contemporary court documents, transcriptions of rare books written about the trials, contemporary maps of the village, historical maps relating to the trial, and information including transcriptions, scans of documents and catalogue information from a wide variety of archives in different places (Ray 2002). Access to all of these is available through a single website: Salem Witch Trials - Documentary Archive and Transcription Project <<http://etext.virginia.edu/salem/witchcraft/home.html>>. The GIS allows the spatial aspects of the trial to be queried and understood. In particular, it allows the 300 people mentioned in the court records to be put into their actual households to provide a better understanding of the property disputes that many historians believe to have underlain the accusations. This includes information on the age and gender of the people concerned, the frequency of the accusations made by or against them, their family relationships, and the relative wealth of the accused and their accusers. Queriable maps and animations are available to allow the user to get a better understanding of the geographic nature of these relationships.

The Valley of the Shadow project (Sheehan-Dean, 2002; Thomas & Ayers, 2003) follows a similar approach <<http://jefferson.village.virginia.edu/vshadow2/contents.html>>. Here the interest is to compare the experience of two communities two hundred miles apart and on different sides of the American Civil War. These communities are based in Franklin County, Pennsylvania and Augusta County, Virginia. The aim of the project is to track the lives of

soldiers and civilians from these two communities during the war and beyond. Available sources include census records, tax records, soldiers' dossiers, letters and diaries. A major difficulty was how to recreate these communities at a level localised enough to provide a context to individual lives. GIS provides one method. In 1870 the 'Hotchkiss' map of Augusta County was produced. This map was localised enough to name over 2,000 individual dwellings, many of which were private residences. The map was scanned and layers digitised from it, including points such as dwellings, churches, schools, mines and mills; lines such as roads and rivers; and polygons such as electoral districts that were used to map census data. Individuals could then be allocated to their houses and other sources about the individuals linked to these. Doing this 'allows us to locate people *within* the county and not simply treat them as undifferentiated residents *of* the county' (Ayers *et al.* <<http://jefferson.village.virginia.edu/vshadow2/ecai/present1.html>>).

8.4 Conclusions

The role of GIS with qualitative data has yet to be fully developed; however, early studies suggest that there is considerable potential here. This potential falls into two main areas. One area is as a catalogue tool that allows electronic archives to be queried and explored in a geographical manner. This is the approach followed by the IDP, Sydney TimeMap and Perseus projects. The second area is that it allows the historian to research the spatial side of a problem in a more integrated and usable way than would traditionally have been possible. This is demonstrated by the Salem witch-trials and the Valley of the Shadow projects' use of GIS to re-create local communities in order to gain a better understanding of events in the past. The use of GIS in qualitative research is not yet well developed, however, but the examples given here show that its power as an integrating and exploration technology and approach has considerable potential in this area, in addition to its more traditional role in humanities research.

Further reading from chapter 8:

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

Case studies

Lang, L., 1995. GIS supports Holocaust survivors video archive. *GIS World*, 8(10), 42-45

Ray, B.C., 2002. Teaching the Salem Witch Trials. In: A.K. Knowles, ed. *Past time, past place: GIS for history*. Redlands, CA: ESRI Press, 2002, 19-33

Sheehan-Dean, A.C., 2002. Similarity and differences in the Antebellum North and South. In: A.K. Knowles, ed. *Past time, past place: GIS for history*. Redlands, CA: ESRI Press, 2002, 35-49

Smith, D.M., Crane, G. and Rydberg-Cox, J., 2000. The Perseus Project: a digital library for the humanities. *Literary and linguistic computing*, 15, 15-25

Thomas, W.G. and Ayers, E.L., 2003. An overview: the differences slavery made: a close analysis of two American communities. *American history review*, 108, 1298-1307.

Wilson, A., 2001. Sydney TimeMap: integrating historical resources using GIS. *History and computing*, 13, pp. 45-69

The Perseus Project website: <http://www.perseus.tufts.edu/>

The Salem Witchcraft website: <http://www.iath.virginia.edu/salem/>

The International Dunhuang Project website: <http://idp.bl.uk/>

The Valley of the Shadow website:
<http://jefferson.village.virginia.edu/vshadow2/contents.html>