Geographical Information and historical research: Current progress and future directions

Ian N. Gregory, Karen Kemp, Ruth Mostern

Gregory, I., K.K. Kemp and R. Mostern, 2003, in press. Geographical Information and Historical Research: Current progress and future directions. *Humanities and Computing* 13:7-22.

Abstract

Information about places is ubiquitous in the study of history, events occur, people live, and phenomena flourish in more or less identifiable locations. To a greater or lesser extent, all historians make use of geographical information. Potentially, there is a great potential for the use of Geographical Information Systems (GIS) in the discipline. However, GIS with origins in the earth sciences and the way that those disciplines model the world, is still based largely on the traditions and requirements of these fields. This means that while there are many advantages to using GIS in historical research, these technologies and analytical approaches must be applied with caution, based on the characteristics of historical data and the traditions of historical scholarship. In this paper we define GIS and explain why it is relevant to historical research. We then use a wide variety of examples to illustrate the ways in which historians are beginning to use GIS. Finally, we discuss how existing software and methods for working with geographic information need to be improved to make it more applicable to historical research. Our aim is to demonstrate that GIS, if properly used, is not only applicable to quantitative, scientific historical paradigms, but is also extensible to the more humanities-driven, qualitative areas of the discipline.

1. Introduction

Geographic Information Systems (GIS) originated in the 1960s.[1] Created to assist in land management, these systems evolved to manage data from two related but distinct perspectives:they provided graphical depictions of the locations of features on plotted maps or map-like displays, and they were also capable of linking sophisticated conventional database and statistical functionality to mapped information in order to allow more sophisticated interpretation of the characteristics of locations. The ability to combine conventional approaches to data

management with the capacity to explicitly handle geography created a new approach to representing and understanding the world.

Through the 1970s and 1980s, GIS was applied to a wide range of problems, and during the 1990s, with the advent of relatively cheap and user-friendly computers and GIS software packages, GIS became a well-established part of academic geography. Nevertheless, the spread of GIS was heavily influenced by its origins in earth sciences. Many early academic studies using GIS were conducted by land planners, earth scientists and physical geographers. Gradually other quantitative geographers and social scientists, including historians, became involved in adapting models and analytical approaches that had been developed in their own domains. The use of GIS continues to spread into new areas. Increasingly it is being used not only as a quantitative tool, but also as a means for sharing and comparing historical maps, linking texts and images to the places that they describe, representing historical boundaries and trade routes, and, in other ways, is being explored as a part of a new humanities computing methodology that mixes quantitative and qualitative approaches. In this latter context, however, its use is still at an early stage.

This paper examines the emerging role of GIS in historical research. While many historians are still not aware of the potential of GIS as a methodology, there is a strong quantitative geographical tradition in some parts of the discipline, for example K.D.M. Snell and Paul Ell's[2] work on the geography of religion in Britain in the mid-nineteenth century ,Robert Woods and Nicola Shelton's[3] atlas of mortality in Victorian England and Wales, or William Skinner's work on regional systems analysis in late imperial China.[1] In other parts of the discipline using geography simply means determining the location of an event such as a battle or a birthplace, or the diffusion of a cultural practice. Through GIS, hypotheses concerning the influence of place and representations of the ways that features are arranged on the earth's surface can be explicitly incorporated into historical research in a far more systematic and analytically powerful manner than was previously possible. GIS has the potential to significantly enhance the use of spatially referenced information in all parts of the discipline, not just those with a strong geographical tradition.

[1] See for instance William Skinner, "Marketing and Social Structure in Rural China," Parts 1 and 2 *Journal of Asian Studies* 24.1,2 (1964-5). For recent applications using GIS, see http://www.people.fas.harvard.edu/~chgis/data/skinner/ [August 2, 2002]

There are many ways in which tools for structuring, visualizing and analysing space, spatial behaviour, and spatial change can benefit historical study, and it is impossible in this brief paper to explore them all in detail. In the paper we introduce basic GIS concepts and provide some examples of the application of GIS to demonstrate the range of methods and results produced by a selection of existing historical studies. The paper then goes on to discuss what is required to make GIS an effective part of the historian's toolbox. The key point here is not that GIS needs to be seen as a revolutionary new approach, but that it adds the capacity for sophisticated spatial thinking to existing historical scholarship through both its ability to portray locations and their attributes in innovative ways and to foster original questions about the relationship of locations to one another. While this paper demonstrates some advantages of using GIS, we are also hopeful that we demonstrate the need to apply it in ways that are appropriate to both the limitations of geographical information and to the historical problems to which GIS is applied.[4]

2. Geographical Information and Geographical Information Systems

Any information with a reference to a location on the earth's surface can be regarded as geographical information. This most obviously includes information shown on topographical maps such as elevation contours, rivers and landforms, or information associated with regions on thematic maps, such as those showing density of population or rates of mortality. However, maps need not be Cartesian—Chinese maps representing physiographic features using a methodology that highlights the cosmic energy forces of *fengshui*, or Australian Aboriginal maps depicting journeys drawn according to the shapes of totem animals, also render geographical information to those trained to read them.

Geographical information comes in forms other than maps. Tabular data is geographical information[5] if it includes columns, or data fields, that provide information about places. This type of data is in widespread use in history and related disciplines. Census tables are a good examples: they link each item of data on statistical themes such as household structure or employment patterns with a spatial reference that is often either a place name or ID that refers to a location on the earth's surface. A large proportion of statistical sources similarly link their data to locations.

However, a great deal of geographical information is less formally structured and, as a result, spatial aspects of many datasets may be overlooked. Telephone directories have spatial information in the form of addresses; descriptions of places that include their postal codes likewise have spatial references. If the location represented in a photograph can be determined, then the photograph itself is a form of geographical information. Images and texts, stock in trade for most historians, often contain a wealth of spatial information, though frequently it will be descriptive, requiring some interpretation before locations can be precisely defined. Where a text references a place name, this is obvious geographic information. Given some caveats discussed later in this article, virtually any information with a geographic reference can be fodder for GIS. Thus, the ubiquity of geographic information and the range of ways of reflecting on it in the kinds of sources that historians conventionally use and produce suggest the potential for GIS in historical scholarship.

Unlike most historians, who primarily read and write texts, people trained in GIS typically subdivide their geographical information into two components. *Attribute data* provides information about each object or location: the characteristics of features shown on a map, the statistical data associated with a census tract, the nature of images or texts about places. Each item of attribute data is linked to its spatial component, often called *spatial data*, which provides the information about where on the earth's surface the data object is located. Spatial data can be stored most directly using geographic co-ordinates, although place names and location identifiers are also commonly used. GIS explicitly brings together the spatial and attribute components of geographical information.[6] This integration of attribute and spatial data distinguishes GIS as a methodology for history from traditional historical texts describing places, and from other forms of historical geographical investigation.

Including spatial data in historical investigations allows historians and others to ask new kinds of questions. There are at least three ways that location, once it is formally specified within a GIS, encourages new perspectives:

- Location provides a universal means of integrating diverse kinds of data,
- Data arranged geographically have geometric properties,
- Data about places collected within a single dataset are spatially dependant.

Integrating data. Since geographical locations can be uniquely and precisely described with respect to coordinates on the surface of the earth, such as latitude and longitude, any

information about a specific location thus described can be immediately related to any other information for the same location. Thus, dots on a map representing, say, shrine sites can be integrated with geographically referenced lines representing rivers or trade routes. Of course, the ability to collate disparate datasets by geographical location does not automatically imply a causational spatial relationship between the data. Additional techniques need to be applied to test any relationship as discussed in the following sections of this paper. Using maps, however, actually adds a layer of complexity. Cartographers convert from locations on the earth's surface to locations on a piece of paper through the use of map projections and scale, each chosen specifically for the map's purpose. Map projections convert the locations on the curved globe to locations on a flat sheet of paper, while scale reduces the size of the area being mapped to a size that will fit on a page. Once drawn on a piece of paper, projections and scale are fixed so it is very difficult to compare information mapped at different projections and scales. When locations are stored in a computer as opposed to on a map, projections and scales are easily changed, thus allowing datasets that would previously have been incompatible as maps to be directly integrated.[7] This integration can simply be done visually using a mapping interface or may be done more formally at a database level. Location can therefore be regarded as a universal integrator when the spatial component of data can be specified.

Taking advantage of geometric properties. In addition to location, geographic information also has geometric properties: a line has length and a polygon[8] has an area, perimeter and shape. This means that quantities such as distance and area can be computed directly from the stored data. Thus, it becomes easy to determine which sacred site is further from the coastline or which ancient walled city covered more territory. One aspect of geometry that GIS is particularly well suited to handle is topology, or the way in which spatial elements are connected to one another. For example, topological information on a network of lines that represents a rail network allows the shortest route from one station to another to be calculated. Topological information on a set of polygons will reveal whether one polygon is adjacent to the next. Using topological information, historians can locate events within a geographical framework and ask questions about the relationship between events and such factors as population density, transportation routes, and topography. Thus one could ask questions such as "Which counties border on those that contain the main trading cities?"

Spatial dependence. Many of the most commonly used statistical procedures, including regression, have a basic assumption that the dataset to be analysed consists of independent observations. Tobler's First Law of Geography states, "Everything is related to everything else, but near things are more related than those far apart".[9] This means that the value of an observation at a specific location is similar to those at other locations, with closer locations being more alike. Since this violates the independence assumption, potentially spurious or invalid results may be found when independence-based statistical techniques are applied to spatially dependent datasets. Thus, running a regression analysis on a set of census data where the values of variables in individual census districts is likely to be related to those of their neighbours absolutely violates the fundamental assumption of the technique.

The common solution to this problem is to simply ignore it, a potentially dangerous blunder. A better understanding of the characteristics of geographical information will encourage researchers to use statistical techniques that explicitly incorporate the spatial component of their data into the analysis. This ensures the use of appropriate techniques and may, in fact, bring more information into play.

3. Uses of geographical information in historical research

When historians collect and publish geographical information, regardless of whether they are working in an electronic or paper medium, they often lose, ignore, or over-aggregate its spatial component. Census data is a good example of this. The population is enumerated using very small, clearly defined, geographical zones thus the data contains extremely detailed information about how demographic indicators vary spatially. However, printed reports of census data, and database tables constructed from them, remove almost all of the geographical information, leaving only an implied geography based on place names and perhaps administrative hierarchies. By themselves, these are geographically meaningless, requiring either local knowledge on the part of the researcher, or reference to separate maps, to re-introduce the geographical component. Without the geographical component, researchers cannot make spatial queries, either about that data itself, or about its relationship to other geographical data from the same place and different times.

However, many historians have begun to understand the value of the often hidden or ignored geographic information available to them. Amongst these historians, GIS is generally used in three ways:

- to discover, manage and integrate historical research resources
- to graphically display research results in the form of maps and other visualisations
- to perform spatial analysis on historical data

This shows that GIS is far more than computer cartography as, in addition to its ability to map data, its also provides both a spatial database and spatial analytical functionality.[10] The following sections illustrates each of these uses within historical scholarship.

a. A tool for scholarship: Discovering and managing scholarly sources

Geographical information can be a powerful way to index digital archives, allowing historians and other researchers to use location, and in some cases, a map interface, to discover all the available sources associated with a particular place. The Perseus Project is a pioneering endeavour in this regard.[11] This project integrates several digital libraries and archives. For instance, they have digitised and geo-referenced the Edwin C. Bolles archive of the history of London, collected in the late nineteenth century. The archive includes printed documents, some of which are unique, folio descriptions of the city from limited print runs, contemporary nineteenth century maps, and illustrations and prints of places in London dating from the seventeenth to the nineteenth centuries. To index and integrate this resource, the Perseus Project created a gazetteer[12] with a map interface and hyperlinked all place names referenced in all the archival materials to it. This means that a researcher reading the digitised version of Daniel Defoe's *Journal of the Plague Year,* an account of the start of the plague in London in 1664, who encounters a reference to Drury Lane, can navigate immediately to a list of all other occurrences of Drury Lane in the archived texts, to historical images of Drury Lane, and to various historical maps showing its location.

Another example of the use of geography to integrate historical resources is the International Dunhuang Project (IDP) at the British Library.[13] This digital library is based on a catalogue of over 40,000 paper manuscripts, printed documents and fragments, and other artefacts, primarily on Buddhist topics, dating from the fifth to the eleventh centuries. They represent the contents of a monastic library in northwest China, sealed on the eve of an invasion in the early eleventh century and later forgotten. The caves were excavated by archaeologists

from several countries in the early twentieth century, including Sir Aurel Stein from Britain, and their contents have been dispersed among institutions in China, Britain, France, Japan, and Russia. The artefacts, many in poor condition, cannot be made generally available to the public, and so the IDP has created an integrated digital archive of them. The database of manuscripts aims not only to replace the printed catalogues and microfilms, but also to become a scholarly tool in its own right. In fact, the digital version of the collection has even permitted fragments of the same manuscript that were excavated from separate cave sites by different teams and stored in dispersed locations to be reunited in cyberspace.

Taking advantage of the geographical information inherent in all these resources, the IDP catalogue includes mapping functionality that allows users to search for manuscripts by spatial location and proximity. In addition to the manuscripts, the project is also digitising and geo-referencing Aurel Stein's notes, pictures and expedition maps. By geo-referencing individual items in the collection, users of the collection are able to reconstruct the original composition of the sites. In the same way that fragments of a single document can be digitally recombined, here artefacts collected by three separate expeditions and now stored in four widely dispersed institutions can be virtually reunited and made available to a global audience.

The use of geography to structure data collection and manage research notes is common in archaeology and can also be used effectively in historical research. The Cambodian Genocide project is collecting data on the mass killing of 1.7 million people by Pol Pot's regime between 1975 and 1979.[14] Based on archival documents and interviews, the project is creating a spatial database with information about burial sites, prisons and memorials to the victims. These sites, often located in areas that are highly inaccessible, are rapidly deteriorating as a result of the warm, wet climate, destroying evidence that might be used for both historical research and war crimes advocacy. Using the geographical information they have collected, researchers working on this project travel to the sites and use GPS (Global Positioning Systems) to record their exact locations. At the same time, they collect other data about the site and gather detailed interviews from anyone who witnessed the events that took place there. Thus GIS is a tool for the researchers collecting data in the field, and also allows the scholars involved to move seamlessly from notes, to interpretation, to dissemination.

b. A medium for publishing: Using geography to conceptualise and visualize history

Through the integrating power of GIS, vastly different historical resources can be brought together to create comprehensive and rich contexts for historical discovery. A groundbreaking example of this approach is the Valley of the Shadow project at the University of Virginia.[15] Based on an archive of thousands of sources, this project seeks to compare the experiences of the individuals living in two valley counties before and during the American Civil War: Franklin County, Pennsylvania and Augusta County, Virginia. While only two hundred miles apart and with similar climate and terrain, these counties were on different sides during the war.

The challenge for the project team was to use traditional sources such as agricultural and population census, tax and church records, soldiers dossiers, letters, diaries, and photographs in such a way as to make them relevant at an individual level. This was accomplished for Augusta County through the discovery and use of the 1870 "Hotchkiss" map which specifically locates and names over 2,000 individual dwellings. These locations are then linked through family names to information about individuals listed in the other sources. Once details about individuals are linked to geography, they can become a part of the rich context of what is known about their neighbours and surrounding communities. In this way, the project is able to "…locate people *within* the county and not simply treat then as undifferentiated residents *of* the county".[16]

All of the historical data is not indexed only by the spatial component—time is also an important element to be catalogued and examined. By bringing time and space together, the context for discovery becomes even richer. Interactive, animated maps, showing the movements of various locally-formed military regiments during the period of the Civil War provides yet more understanding of the lives of these soldiers and their families. [17] Based on the resources compiled and spatially referenced for this project, the authors have been able to argue that there are no significant factors internal to the social geography of the counties themselves that could have predicted their membership in opposing regimes during the war.

An even more localized example of putting historical data into its geographical context is provided by Benjamin Ray's research on the seventeenth century witchcraft trials in and around Salem, Massachusetts.[18] Ray's animation of the locations of significant events during the month when the accusations reached their peak allows a detailed examination of the spatial and temporal patterns associated with the lives of the women accused of witchcraft and the people who accused them. This research is driven in part by an effort to explore the assertion that the

accusations and subsequent trials were in fact the result of local disputes over land ownership. The relationships between where the women and their accusers lived are clearly critical in exploring this interpretation.

Much effort is now being directed in many fields to use GIS and related technologies to build visualizations of reconstructed past landscapes and built environments. One example of this is provided by Trevor Harris's work on Moundsville, West Virginia.[19] An ancient burial mound in the centre of a modern town was virtually reconstructed in its original environment by placing it in a GIS generated 3-D model of the landscape and reconstructing upon that landscape what is known of the native vegetation. By visualizing what might be seen of the ancient landscape from the top and environs of this mound, researchers gain some understanding of what the builders might have been attempting to achieve through their construction.

c. An analytical methodology: Using GIS to understand and analyse spatial relationships

Beyond simply providing tools for the visual analysis of spatial relationships, GIS also provides tools for much more complex spatial analytical techniques. Bertrum MacDonald and Fiona Black's work on the spread of print culture demonstrates the importance of spatial analysis in historical research.[20] Their argument is that print culture spread across continents and oceans as a result of the interaction of a complex set of spatially distributed variables, and that the diffusion of new reading practices can be modelled using spatial analysis. Literacy rates and the distribution of printers, libraries and booksellers can be mapped. However, to understand their spread these must be integrated with other kinds of geographical information like migration, developments in transport and changes in education. To model the relationship between such a wide range of factors requires a complex database able to handle variables that change through both time and space for a wide variety of disparate types of data. Since GIS provides a means to construct a geographical framework with which all these variables can be integrated, the researchers hope to be able to examine the relationships between the growth of print culture and broader developments in society. Without a powerful tool for handling the multi-dimensional complexities of their data, and in particular its spatial and temporal elements, they would be forced to simplify aspects of their data and thus reduce the potential for gaining understanding from it.

Once data such as that collected by MacDonald and Black is organized in a GIS, spatial analysis can be performed. Spatial analysis (also known as geographical data analysis) refers to a

set of statistical and other computational techniques that differ from conventional quantitative analysis techniques in that location is incorporated as an explicit part of the analysis. This means that the results of a spatial analysis will change if the location of the objects analysed changes. Examples of spatial analysis techniques include testing spatial patterns to determine if clusters exist (for example, finding cancer "hot-spots"), interpolating data from one set of spatial units to another (for example, estimating how many people counted in census tracts live in overlapping but not coincident school districts), identifying shortest routes from one location to another on a network, and regression techniques that allow the parameters in the model to vary across the space, rather than simply producing a single regression model for the entire study area.[21]

So far, only limited uses of spatial statistical techniques have been made by historians. Ken Bartley and Bruce Campbell[22] used such techniques to produce a land use classification for medieval England based on 6,000 records in the *Inquisitiones Post Mortem* (IPM), detailed descriptions of estates produced on the death of landholders. They began by allocating the records to individual points on a map of England using a gazetteer of place names. Since many of these points overlapped and potentially represented large areas that may have contained a variety of land uses, they then used a spatial statistical technique known as kernel estimation to distribute the values from these points onto a grid of square cells overlaid on a map of England (a raster). Using other techniques, they were then able to assign primary and secondary land use types to each cell and create a map of medieval land use in England.

Ian Gregory, Daniel. Dorling and Humphrey Southall[23] analyse changing patterns of poverty in England and Wales through the twentieth century as recorded by the census. They argue that conducting long-term comparisons is not normally possible as it is difficult to distinguish changes in patterns caused by real change from changes in patterns caused by the changing geometry of the administrative units used to collect and publish census data. To work around this, they interpolate data published for a variety of mosaics of spatial units onto a single set of administrative units. Once on a common geography, it is possible to compare patterns on different dates using a combination of simple statistics and sophisticated visualizations.

4. Making GIS work for history and historians

The examples shown above demonstrate how GIS can be used to discover and organize the texts, notes and statistics that form the basis of historical research, how it can be used to

visually display results and illustrate hypotheses about historical experiences that are spatially dependent, and how it can be used as a tool for spatial analysis. However, because the GIS research agenda has been driven by earth scientists and, more recently by social scientists, the tools offered by GIS software and the approaches used by GIS practitioners are heavily rooted in the traditions of these disciplines. While the resulting framework has led to applications of relevance to a wide range of historical problems, there are areas that still need research to enhance the capabilities of GIS for use in historical scholarship. At the same time, historians need to conduct and disseminate research that is dependent on GIS in order to establish disciplinarily-appropriate best practices for using the existing technology and methodology of GIS in a way that does not compromise proven methods of historical research and documentation.

Given this, several observations can be made about historical methodology and data that suggest areas in which historians, other humanities scholars, and geographic information scientists can usefully collaborate to extend both the technology and the historical applications of it.

- Information about time, place and attributes in historical sources is often ambiguous, incomplete or contradictory. It is inevitably associated with particular ideologies or political positions that may not be completely clear to modern scholars. GIS is poorly suited to handle spatial, let alone temporal, ambiguity.
- The basis of the study of history as a discipline is often to describe change over time — how some particular aspect of the human experience was transformed, and why. Spatial temporal visualization and analysis in GIS is rudimentary at best.
- As a discipline, history has both quantitative and qualitative traditions but a significant gulf has developed between the two approaches. GIS allows data from different sources to be integrated at a technical level. Being able to integrate data more easily should allow progress to be made in integrating quantitative and qualitative approaches. This is not simply a technical issue, the real challenges will be in adapting and devising appropriate methodologies.

We expand on these observations in the following sections.

a. Incorporating uncertainty into geographical information

Unlike many of the disciplines that use GIS, historians are accustomed to dealing with sources that are incomplete and inaccurate and may contain ambiguity, contradiction and bias. Indeed, many historians develop research topics by seeking loci of silence or conflict in the texts they read. For example, the twelfth century frontier between the Song and Xi Xia empires in what is now northwest China was contended between the two regimes, and the information produced by each side is not consistent. Many maps were compiled at the time, but they were based on surveys conducted rapidly, by poorly trained generalists operating under wartime conditions, using pre-industrial techniques that we would now consider rudimentary. Even in the twelfth century, officials complained that the maps were inadequate. Moreover, all of those maps are now lost, so the only information available to historians about the location of the border is elliptical and fragmentary texts that describe the maps. From the point of view of historians' best practices, "getting the border right" does not mean drawing a very precise line, but rather documenting as precisely as possible two different kinds of ambiguity: the ambiguity vividly described in the documentary sources of the time, and the conundrum faced by modern scholars who have to adjudicate between contradictory sources and elisions in the record. For historians working with topics other than political history that have an explicit spatial characteristic—say, the spread of the English language over time and place, its transformations and variations, and its coexistence with other linguistic practices—these issues are even more complex.

Location is obviously a key component of geographical information. Unfortunately, in a historical context, location is often described simply as a place name, an ill-defined region such as "the southwest US," or descriptively, as in "close to the banks of the river". These all represent various forms of uncertainty. The simplistic approach to handling these in GIS is to attempt to infer the precise geographical co-ordinates and features required by GIS software. Thus, the place name is redefined as a point location taken from a modern map, the region is given precise boundaries and "close to the river bank" is taken as being within an exact and entirely arbitrary distance from a line representing the modern course of the river. These solutions are at best limited and at worst entirely inappropriate.

What is required instead is an ability to explicitly incorporate uncertainty into geographical information. This is an area that has already attracted a considerable amount of research by the GIS community[24] and where a number of lessons have been learned. However,

commercial GIS software has yet to incorporate functionality for easily managing uncertainty. This leaves the responsibility with the researcher to first of all learn about how uncertainty has been handled in different areas of GIS, and then to adapt these lessons to historical problems.

The handling of uncertainty can be achieved through several means. *Data models* can be developed that allow, for instance, several versions of the geographical location of a particular object to coexist, with certainty ratings assigned to each. There are emerging standards for *digital gazetteers* which can link named places to more than one specification of their locations and allow the inclusion of information about the different names used for the same place, changes in the location and extent of the place over time and the sources of knowledge about place locations. Continued efforts are needed in *cartography and other areas of geo-visualization* to develop robust methods for displaying various forms of uncertainty. In addition, new *analytical techniques* need to be developed that take into account all the various types of uncertainty across space and time.

b. Better functionality to handle change over time

The way in which GIS models the world is highly effective at representing certain types of locational information; however, its ability to represent time as well are extremely limited.[25] Progress is being made but it is still in its early stages. Commercial software vendors have been slow to incorporate temporal functionality into GIS, as they have not seen a strong demand for it. Recently they have started to include certain types of temporal functionality such as vehicle tracking or methods for recording changes in ownership and shape of land parcels. Academic projects may be of more relevance to historians. For example, the TimeMap Project at the University of Sydney, described in detail elsewhere in this issue, is designed to allow historians to work with temporal as well as spatial data.[26] It includes a GIS viewer linked to a metadata clearinghouse using customized metadata extensions that allow objects to be filtered and mapped in time as well as space. Using TimeMap, once objects in the database are assigned both a spatial coordinate and a time instance or time range, a timeline slider bar and an animation tool allow simple temporal change to be explored visually.

Progress has also been made in devising GIS data models suitable for use with data that change over time. So far this has been focused on enabling the construction of databases of census and similar datasets published at different times for spatial units that have changed over

time.[27] Gregory, Dorling and Southall show how controlling the spatial component of the data while varying time allows a better understanding of changing patterns of poverty.

Showing change over time has long been a challenge for traditional cartographers. New methods of geo-visualization and animation in the digital environment show great promise. The Salem witchcraft and Valley of the Shadow projects described above give examples of how animation can be used to develop a better understanding of spatio-temporal change.

But just as historical documents are often frustratingly silent about location at the level of precision that contemporary GIS is best suited for, the same is true when it comes to time. The Ricci Institute is developing a spatial database of historical Christian missions in China largely based on a single census taken in a particular year.[28] The missions, obviously, did not exist only at the moment that the census recognized them, but it would take a monumental amount of research in order to determine precisely when each one of hundreds of missions was founded or dissolved. Short of such an effort, documenting historical temporal information is an uncertainty issue. Scholars need to use information external to the dataset itself to determine the earliest possible date that each mission could have existed, the date by which it seems extremely likely that it existed, and how the one instance of its documented existence relates to the plausible time range of its existence. Thus there is much work to be done to model the relationship between what we do know about places in time and the actual way places have evolved over time.

c. Integrating quantitative and qualitative scholarship

GIS's ability to integrate information from a wide variety of seemingly incompatible sources through location has great potential to help to integrate quantitative and qualitative scholarship. Traditionally these two branches of historical study have used very different sources. Quantitative scholarship is largely based on the use of statistical data, while qualitative scholarship uses diverse sources such as narrative texts, images, interviews and archaeological data. There is no reason why these approaches should be mutually exclusive, but the difficulties in integrating the different types of data used by the two may be one reason. Technically, GIS has solved most of the problems in doing this. If one has a description of a place such as Drury Lane which can be converted to a coordinate based location, then historians can easily determine, for example, which parish or census district this lay in at certain dates and can also bring descriptive and statistical information together. To adhere to scholarly practice, a quantitative database like this must somehow be able to incorporate rigorous information about sources, and

free or structured text evaluating the quality of those sources and the information in them. Some of this documentation may be handled in metadata.[29] However, very few existing metadata standards apart from those in the GIS domain include temporal and spatial information, and none have been developed to accommodate documentation in the tradition of history and the humanities. Developments in this area will need to involve the standards community as well as the GIS and historian communities.

In addition to putting information about sources into databases and metadata, GIS implementations for historians, if they are to be credible within existing scholarly communities, must allow access to the historical documents from which spatial and attribute information was derived, along with historians' commentaries about them. Texts that describe places, and images that depict them, can be hyperlinked to geographical data about those places. Some historical maps can be geo-referenced and incorporated into a GIS; others that are not in a recognizable coordinate system can simply be hyperlinked as images. More elaborate semantic and interpretive questions such as what a time-enabled map of, say, *War and Peace*, with its multiple geographies and large cast of characters ought to look like, are on the research frontier of collaboration between humanists and geographic information scientists.

Comparing different datasets, each with different collection methodologies and assumptions, presents many problems not unique to historians l. It relies on rigorous and informed scholarship that shows a thorough understanding of the limitations of the different types of data, and a commitment to linking that scholarship to the GIS. Historical scholarship in the age of GIS will have to adhere to all of the traditional good practices, and develop new ones as well.

5. Conclusions

GIS use among historians is still in its early stages and there is still a long road to travel with respect to technology, content and the development of practices and methodologies before it becomes a widely accepted and adopted component of historical scholarship. There is, however, a clear and significant potential for using GIS. This paper demonstrates the breadth of applications for which GIS can be used. These fall into three basic categories: geographically referenced databases that allow data to be integrated through location, alternatives to traditionally published monographs for visualizing and disseminating geographically referenced

data related to a topic area and hypotheses about it, and tools for analysing data spatially. The common thread is that GIS, allows researchers to effectively handle spatial information. This in turn allows geography and human behaviour in space to be better understood.

The geographical component of historical information has traditionally been under-utilized. There will be a learning curve to using GIS and geographical information appropriately. The software and techniques associated with GIS come from academic disciplines with strong scientific traditions. Many areas of history are not engaged with this tradition, and the adoption of GIS does not mean that they should begin to be. Rather than simply impose the techniques used by other disciplines, historians must use GIS in a manner that builds on the traditions of historical scholarship. There are technical issues in integrating GIS into historical research, but even more than this, a key concern must be the substantive issues of how best to frame and answer scholarly research questions within the limitations of the available data and tools. How to deal with uncertainty in historical data and how to integrate quantitative and qualitative scholarship are examples of this. While these are long-standing issues within historical research, GIS technology allows greater demands to be made on the data than was previously possible, and thus throws these issues into sharper focus.

When GIS started to be widely adopted among geographers it was accompanied by over-optimistic claims that it would provide a cohesive scientific framework that would re-unite an academic discipline frequently regarded as fragmented and lacking a central common thread.[30] This produced a backlash among other geographers who regarded the quantitative and scientific approach as an anathema to the way they approached geographical research.[31] As a result of this debate there was, subsequently, a more temperate examination of what GIS had to offer to academic geography, and GIS has now become a widely used and commonly accepted part of many geographers' toolboxes. It is hoped that by ensuring that GIS becomes more widely but appropriately used among historians, both over-optimistic and over-hostile attitudes towards GIS can be avoided. This can be achieved as long as historians demand that the tools and techniques offered by GIS are appropriate to their research questions and data, and that they use the technology in a careful, informed manner that yields both sound scholarship and interesting methodological challenges. It is not a contradiction to hope that GIS use can help keep the humanity in the humanities.

[1] T. W. Foresman, ed., The history of geographic information systems: perspectives from the pioneers. (Upper Saddle River, NJ, 1998)

[2] K. D. M. Snell and P. S. Ell, *Rival Jerusalems: the geography of Victorian religion* (Cambridge, 2000)

[3] R. Woods and N. Shelton, *An atlas of Victorian mortality* (Liverpool, 1997)

[4] The parallel can be drawn between GIS and statistics. In both cases the advent of easy-to-use software means that there is the potential for both the use and the abuse of complex techniques.

[5] Note that we do not differentiate here between the terms "data" (strictly, raw facts) and "information" (facts plus interpretation) since one person's information is often another person's data. These terms are used somewhat interchangeably in GIS.

[6] Good introductory texts on GIS include: N. R. Chrisman, Exploring Geographical Information Systems, 2nd ed. (New York, 2002); M. N. DeMers, *Fundamentals of Geographic Information Systems*, 2nd ed. (New York, 2000); D. Martin, *Geographic Information Systems and their socio-economic applications*, 2nd ed. (Hampshire, 1996).

[7] It must be noted that although software will allow projections and scales to be changed at will it is up to the user to ensure that this is implemented within the limitations of the source data.

[8] *Polygon* is GIS terminology for an area or region.

[9] W. Tobler, 'A Computer Movie Simulating Urban Growth in the Detroit Region', *Economic Geography*, 46 (1970), pp. 234-240.

[10] D. J. Cowen, 'GIS versus CAD versus DBMS: what are the differences?', in D. J. Peuquet and D. F. Marble, eds., *Introductory readings in Geographic Information Systems* (London, 1990)

[11] The Perseus Project website is at: <URL: http://perseus.csad.ox.ac.uk/> [22 Feb 2002]. See D. M. Smith, G. Crane and J. Rydberg-Cox, 'The Perseus Project: A digital library for the humanities.', *Literary and linguistic computing*, 15 (2000), 15-25.

[12] The term gazetteer is used slightly differently in GIS than in history. In GIS it refers to a document (digital or printed) that relates place names to their location. Additional information such as the type of place may also be included.

[13] <URL: http://idp.bl.uk/> [22 Feb 2002]

[14] <URL: http://www.yale.edu/cgp/databases/> [22 Feb 2002]

[15] <URL: http://jefferson.village.virginia.edu/vshadow2/> [22 Feb 2002]. See also A. Sheehan-Dean, 'Similarity and difference in the antebellum north and south', in A. K. Knowles, ed., *Past time, past place: GIS for history* (Redlands, California, 2002), 35-50.

[16] Quoted from E. L. Ayers, A. S. Rubin. and W. G. Thomas <URL: http://jefferson.village.virginia.edu/vshadow2/ecai/present1.html> [22 Feb 2002]

[17] <URL: http://jefferson.village.virginia.edu/vshadow2/MAPDEMO/theaterintro.html> [22 Feb 2002]

[18] <URL: http://jefferson.village.virginia.edu/%7Ebcr/salem/salem.html> [22 Feb 2002]. See also B. C. Ray, 'Teaching the Salem witch trials', in A. K. Knowles, ed., *Past time, past place: GIS for history* (Redlands, California, 2002), 19-34.

[19] T. M. Harris, 'Moving GIS: exploring movement in prehistoric cultural landscapes using GIS', in G.R. Lock, ed., *Beyond the map: archaeology and spatial technologies* (Oxford, 2000), 116-123 talks about the potential of these techniques.

[20] B. M. MacDonald and F. A. Black, 'Using GIS for spatial and temporal analysis in print culture studies: Some opportunities and challenges', *Social Science History*, 24 (2000), 505-536

[21] A good introduction to spatial analysis is provided by chapter 12 of P. A. Longley, M. F. Goodchild, D. J. Maguire, and D. W. Rhind, *Geographic Information Systems and Science* (Chichester, 2001). More statistical discussions include A. S. Fotheringham, C. Brunsdon and M. E. Charlton, *Quantitative geography: perspectives on spatial data analysis*. (London, 2000) and T. C. Bailey and A. C. Gatrell, *Interactive spatial data analysis.* (Harlow, 1995).

[22] K. Bartley and B. M. S. Campbell, 'Inquisitiones Post Mortem, GIS, and the creation of a land-use map of medieval England', *Transactions in GIS*, 2 (1995), 333-346.

[23] I. N. Gregory, D. Dorling and H. R. Southall, 'A century of inequality in England and Wales using standardised geographical units', *Area*, 33 (2001), 297-311.

[24] D. J. Unwin, 'Geographical Information and the problem of "error and uncertainty"', *Progress in human geography*, 19 (1995), pp. 549-558 provides a review of this.

[25] See G. Langran, *Time in Geographical Information Systems*. (London, 1992) and D. J. Peuquet, 'Its about time: a conceptual framework for the representation of temporal dynamics in Geographic Information Systems', *Annals of the Association of American Geographers*, 84 (1994), 441-461

[26] See A. Wilson, this volume, and <URL: http://www.timemap.net> [22 Feb 2002]

[27] See M. DeMoor and T. Wiedemann, this volume for an example of this. Broader discussions of the data models required for historical temporal data are provided by I. N. Gregory 'Time-variant databases of changing historical administrative boundaries: a European comparison', *Transactions in GIS,* 6 (2002), 161-178 and F. Swiaczny and T. Ott, *Time-integrative Geographic Information Systems. Management and analysis of spatio-temporal data.* (Berlin, 2001).

[28] See <URL: www.usfca.edu/ricci/institutions/public/index.html> [02 April, 2002]. This is a database of all Christian institutions in Chinese history. It includes extensive spatial information, but it has not been linked to a gazetteer, and the GIS version of it exists only in a prototype that is not available publicly.

[29] Metadata is data that provides additional information about the sources and characteristics of the data.

[30] See in particular S. Openshaw 'A view on the GIS crisis in geography, or putting Humpty-Dumpty back together again', *Environment and planning A*, 23 (1991), 621-628 and S. Openshaw 'The truth about Ground Truth', *Transactions in GIS*, 2 (1997), 2-24.

[31] See in particular M. Curry 'GIS and the inevitable ethical inconsistency', in J. Pickles, ed., *Ground truth: the social implications of Geographical Information Systems*. (Oxford, 1995) 68-87 and J. Pickles 'Representations in and electronic age: geography, GIS and democracy' in the same volume.