Introduction: From Historical GIS to Spatial Humanities: Deepening Scholarship and Broadening Technology

IAN N. GREGORY AND ALISTAIR GEDDES

When geographical information systems (GIS) first began to be used by academic geographers in the late 1980s, their use was nothing if not controversial. Proponents of the new field argued that it had the potential to reinvigorate geography as a discipline under a more computational paradigm. Opponents argued that it marked a lurch toward an unacceptable form of positivism with no epistemology or treatment of ethical or political issues. One thing on which they both agreed—or perhaps took for granted—was that GIS was a quantitative technology that was to be used in a social scientific manner (to its supporters) or a positivist way (to its antagonists).

When GIS first began to be used by historians it was not surprising that much of the early focus was also quantitative and social science based. It is no coincidence that the first special issue of a journal dedicated to historical GIS (HGIS), published in Social Science History, included essays on topics such as fertility, migration, urban history, and economic growth, all well suited to quantitative analysis. In 2008, eight years after this issue was published, a conference devoted to HGIS was held at the University of Essex. It attracted 125 delegates, with papers organized in 21 sessions. While some of these sessions were themed on topics that still had a strong quantitative bent—demography, urban history, environmental history, transport, and so on—there was also an increasing number of papers and sessions that concentrated on topics that were clearly qualitative and did not follow traditional social science paradigms. These topics included art, performance culture, literature, the Bible, and medieval and early modern history. What was happening
in the quantitative sessions was also interesting. Rather than concentrating on issues associated with database construction and potential applications, many of these papers had developed to focus on conducting applied works of history – studies that developed the historiography by answering applied research questions. This was an indication of two emerging trends within HGIS that have continued since: HGIS is deepening from an applied perspective, and it is broadening from a technical perspective. It is deepening in that it has reached a stage where researchers apply it to scholarship that develops new knowledge about the past. This must be the ultimate aim of the field, as it takes HGIS beyond a narrow technical specialism and makes it relevant to a much wider audience. HGIS is also broadening its technical scope in terms of the ever-widening potential for its application to both qualitative and quantitative sources. This means that GIS is thus able to expand beyond social science history – a fairly narrow field – to be applicable to the discipline more broadly, and beyond that to spread outside the disciplinary boundaries of history into other humanities disciplines.

**GIS AND HISTORICAL RESEARCH**

There are many different definitions of GIS and related terms such as GISc (Geographical Information Science). The emergence of new geospatial technologies such as Google Earth that do not fit traditional definitions only complicates these definitions. Originally, “GIS” was considered as the umbrella term for the field, and it is often still used in this way. The more recent trend, however, has been to use “GIS” to describe the tools offered, while “GISc” emphasizes the broader understanding of how these tools can be developed, used, and applied.

To take this further, GIS can be thought of as a type of software that provides a way of representing features on the Earth’s surface and a suite of operations that allow the researcher to query, manipulate, visualize, and analyze these representations. The representations, or data models, combine two types of data: *attribute data*, which were traditionally held in a table and tend – or perhaps tended – to be quantitative, and *spatial data*, which locate each item of data using a point, a line, a polygon (which represents an area or a zone), or a pixel. Points, lines, and poly-
gons are used to represent discrete features, and data in these formats are referred to as vector data, while pixels are used to represent continuous surfaces and are referred to as raster data. In this way the attribute data say what, while the spatial data say where. Thus, from this perspective GIS is a type of software that allows the user to store, retrieve, visualize, and analyze data that are georeferenced to a location on the Earth's surface. GIS allows researchers to ask questions about their topics or sources that stress the importance of location and thus geography. This emphasis on geography, combined with the tools to represent and explore georeferenced data, is what allows scholars to conduct their research in new ways.

This approach to defining GIS leads to the conclusion that a Geographic Information System is really a database for managing georeferenced data. Until recently this conclusion made a quantitative paradigm almost inevitable, as databases were, almost by definition, quantitative, holding either numbers or structured textual information such as is found in library catalogs. Recent developments in Information Technology (IT) mean that this paradigm is no longer the case. Increasingly, almost any type of data can be held within a computer system, including unstructured texts such as books and web pages, still images, moving images, and sounds. As long as a location can be found for these items, they can be held within a GIS-type structure. This means that the need for attribute data to be quantitative is increasingly disappearing.

While this is true of attribute data, it is not true of spatial data, which, although they tend to be represented graphically, are indisputably quantitative in nature. What appears to be a point on the map is actually a pair of numbers representing x and y or latitude and longitude. A line is a series of points joined together, and the boundaries of a polygon are made up of one or more lines. Despite the many critiques of maps, researchers are usually far more comfortable interpreting the crude quantitative abstraction of space created from spatial data than they have been interpreting the crude quantitative abstraction of society created from quantitative data. For example, many humanities researchers would be happy with a map showing the locations of certain events as points but would be suspicious of a scatterplot showing the relationship between two variables such as the unemployment rate and the number of crimes in tracts around a city. In reality, the two have much in common. They
both simply show dots whose location is determined using values of $x$ and $y$, values that are typically expressed to a far higher degree of precision than the accuracy of their measurement can really support. There are a number of possible reasons for this apparent inconsistency. These include the valid reason that measuring and interpreting statistics about characteristics of the population are more difficult than doing the same for locations on the Earth’s surface. Less justifiable are a misplaced confidence in the authority of maps and, from the opposite perspective, a misplaced suspicion of quantitative approaches among some researchers in the humanities.

Using GIS in humanities research presents the researcher with two major sets of challenges. The first is to get the data into a GIS. GIS databases are usually vector data, which require that every item of attribute data be located using a precisely defined point, line, or polygon. Techniques such as using raster surfaces or networks have been used effectively in HGIS research to represent imprecision in location, but even these approaches usually require a precise point-based location to be given initially and do not solve the inherent uncertainties within the spatial data. In some cases it may simply be impossible to get a source into a form suitable for GIS; however, as we will see, this does not mean that such a source cannot contribute to a study that is centered on a GIS database.

The second, and more important, challenge is to get information back from the GIS databases and turn it into new scholarship that advances our knowledge of the past. Once a GIS database has been created it is very easy to produce large numbers of maps, graphs, tables, and summaries. Going beyond this to produce new knowledge or an innovative narrative requires a different set of skills. Creating a GIS and analyzing the data that it contains requires technical GIS skills. Producing new scholarship requires the skills of the historian or other humanities scholar to turn the GIS output into a contribution to our understanding of the past. A key test of the effectiveness of historical research is that the work that it produces should be of interest not only to HGIS specialists but also to an audience of subject specialists who are more interested in the results of the research than in the methodology that was used to achieve these results.
These challenges show that there are certain principles that researchers interested in GIS must consider. Clearly, there must be a geographical dimension to whatever study is being undertaken, and, beyond this, the requirements and limitations of the vector and raster data models need to be understood. However, the fact that GIS research is ultimately based on a software tool does not, as is sometimes claimed, force a positivist approach onto the researcher. As described above, GIS merely provides a platform on which research can be conducted. It does not impose any approach other than the fact that the data within the GIS database have to be represented using attribute and spatial data and that the spatial data must be in the form of points, lines, polygons, or pixels. How scholars turn these data into information about the past and then to humanities scholarship is their decision. They would start from their discipline’s existing paradigms and add the more explicitly geographical. That said, as noted above, GIS does require locations that are usually expressed with high degrees of precision—usually far more precision than it makes sense to express them to. This is not, of itself, positivism, which is concerned with using statistical approaches to define relationships between variables so that empirical generalizations can be made. Early fears of GIS causing a return to “the very worst sort of positivism” were prompted by calls for GIS to reinvigorate the use of quantitative attribute data and statistical approaches rather than by fears about the accuracy of spatial data. It is also worth noting that representing the location of a mountaintop or a city using a point whose coordinates are given submillimetric precision makes no more sense in the earth sciences or social sciences than it does in the humanities. The limitations of the vector data model mean that locations are expressed with this spurious level of precision. It is up to the researcher to interpret this precision and its consequences sensibly, which is not always as easy as it may seem; indeed, here the humanities may have an advantage over more empirical approaches. While in other fields coordinates are often used as inputs into statistical approaches where their overly precise nature becomes lost in summary statistics, in the humanities—thewhere the emphasis is on close reading and careful interpretation—imprecision and ambiguity are actually easier to handle, as these limitations remain more transparent to the critical researcher.