

A New Architecture for the Great Britain Historical GIS

Humphrey Southall
Department of Geography
University of Portsmouth
Portsmouth PO1 3HE
Humphrey.Southall@port.ac.uk

Ian Turton
School of Geography
University of Leeds
Leeds LS2 9JT
i.turton@geog.leeds.ac.uk

Nick Burton
Department of Geography
University of Portsmouth
Portsmouth PO1 3HE
Nick.Burton@port.ac.uk

Neil Pearson
Department of Geography
University of Portsmouth
Portsmouth PO1 3HE
Neil.Pearson@port.ac.uk

Version 4: May 20th 2003

Introduction

'Historical' GIS systems can be grouped into four categories:

- (1) Systems holding historical data without a time dimension. Examples abound. Most archaeological GISs record the locations of 'finds' rather than reconstructing the past.
- (2) Systems including time as a series of cross-sectional coverages, with obvious applicability to working with historical censuses.
- (3) Systems implementing a time dimension within mainstream GIS technology. One example is the Belgian historical GIS which holds sequences of separate coverages for each province, then assembles a national map from these using a concordance table identifying the appropriate provincial coverage for any date (de Moor and Wiedemann, 2001). A purer implementation is the original Great Britain historical GIS, described below.
- (4) Systems going beyond temporal GIS to reflect the special characteristics of historical data, notably its incompleteness: holding information about entities whose locations and dates of existence are known with varying degrees of precision, and sometimes only indirectly. Perhaps the only two large systems of this type are the China Historical GIS, going back millennia (Berman, 2002), and the new Great Britain Historical GIS, which hopes one day to go back to Domesday.

The system described in this paper could hold statistical data for an entity called either Heckmondwith or Almondwith, somewhere in Yorkshire during the late middle ages. Obviously, these particular data would only show up in maps of county aggregates, but the same system can map parish-level data for all parishes with known boundaries. The system is implemented in an object-relational database, Oracle 9i.

The original Great Britain Historical GIS (Gregory and Southall, 1998) consists of a large statistical database held in Oracle 7 loosely linked to boundary data managed by ArcInfo 7. Boundary data are held as arcs, not polygons, and the arc attribute table holds manually inserted numbers recording the first and the last date each arc was a legal boundary. Polygon attribute tables hold label points for each statistical reporting unit, again with dates of creation and abolition. ArcInfo itself does not know that our numeric fields are year, month and day values, but a large suite of custom AML selects out arcs which were legal boundary lines at a specified date and constructs polygons around the label points of units which also existed for that date. A steadily growing collection of *ad hoc* look-up tables held in Oracle mapped the many different unit names in the statistical data to the single name held in the GIS; the latter was simply the name the unit had in the 1911 census report, or rather the name we typed into our original 1911 transcription.

This system was the first large scale historical GIS able to hold boundary changes accurate to the day, and worked well both for the c. 630 Registration Districts of England and Wales, and for the c. 1,800 Local Government Districts. However, scaling it up to 15,000+ parishes proved problematic, and the only way to create usable parish-level mapping has been to generate polygonal coverages for each census date, then manually resolve topological errors and name mis-matches. This has taken three years, but has permitted us to realise the analytic goals of the original project: by linking these static coverages to the parish tables from each census, we have created geography conversion tables that redistrict statistics for 19th century Registration Districts and 20th

century Local Government Districts to the 408 local authorities of modern Britain; and these redistricted data now forms the basis of a public web site which traces long-term change for each of these units, from 1801 to 2001:

<http://www.VisionOfBritain.org.uk>

However, retreating from a type 3 historical GIS to a type 2 is only an interim solution. Fortunately, new National Lottery funding made it both necessary and possible to design a new system providing a more consistent framework for our existing statistics and digital boundaries, and for new material including descriptive text and historical maps. This paper begins by discussing users needs for the new system, and then outlines four levels of architecture. Firstly, the ontology at the core of the system: a collection of statements of what entities and relationships exist; at this level the system contains no locational information. Secondly, the space-time architecture, associating entities with geography and chronology. Thirdly, the middleware architecture which mediates between users and the core resource, supporting multiple web sites and web services. None of these architectures defines how our resource appears to a particular user community, so our final ‘architecture’ is that of the primary web interface.

Expectations of a national historical GIS

Our new goal is to create not simply a national historical GIS, holding historical data for the whole country, but **the** national historical GIS for Britain, for everyone and not just academic researchers. ‘Historical’ implies not simply a GIS concerned with time and the past, but one providing a framework for documentary information, not the archaeologists’ finds. Our geographical entities are therefore relatively large administrative units or vaguely defined ‘places’, not precisely located archaeological remains. However, dates are often precise.

Many distinct user communities can be identified:

- **Academic researchers:** Demographers need primarily to create choropleth maps from historical censuses, but most historians want simply to locate place-names. Names change over time, so we have worked to develop a relationship with the community of place-name historians. Our system can associate any number of names with an entity, recording dates and language for each.
- **Public sector bodies:** ‘History’ started not long ago, ‘modern’ GIS data becoming available only in the early 1970s and many current processes can only be understood through data on longer periods. The project has been strongly supported by the Office of National Statistics. An example of policy-relevant research using the historical GIS linked the ONS Longitudinal Study (LS), recording the recent health of today’s elderly, with 1931 census data. This was possible because LS members are identified by their NHS numbers, originally issued by the National Registration of September 1939 and identifying the local authority of residence. Linkage to 1931 data was complicated by extensive boundary changes in the mid-1930s, but showed that being brought up in areas of high unemployment significantly worsened people’s health today, regardless of where they have since moved (Curtis *et al*, 2003). More recently, the Environment Agency have funded pilot work to integrate the 1930s Land Use survey into the GIS, for linkage with modern satellite imagery to assess long-run change.
- **‘Memory Institutions’:** Museums, galleries and especially, given our emphasis on the documentary record, archives and libraries. A major new justification for our

project is improving cataloguing of place-specific resources. In particular, UK local archives systematically catalogue by location. This means including geographical names in card catalogues, but which name to include: a street, a neighbourhood or the general district? General practice has been to use the name of the Ancient Parish containing the location but this poses two new problems: which parish covered a given location, and what *exactly* was its name? A set of standard authorities for names has been identified (National Council on Archives, 1997). Geo-referencing provides an alternative to place-name indexing, with many benefits, but must be polygon not point based. While heritage workers generally have limited experience with GIS, they are at least as computer literate as most academics, and are trained to organise information more rigorously than many GIS researchers. The detailed standards laid down by the funding body (New Opportunities Fund, 2002) require us to take metadata creation and standards support very seriously.

- **Schools:** All English children aged 8-11 are legally required via the National Curriculum to prepare ‘a study investigating how an aspect in the local area has changed over a long period of time ...’, examples of ‘aspects’ being ‘education; population movement; houses and housing; religious practices; treatment of the poor and care of the sick’ (Quality and Curriculum Authority, 1999). The QCA identify relevant web resources and, at the time of writing, for ‘finding out about history in the local area’ a site run by Sainsbury’s about their stores’ history ranks third.
- **‘Life-long learners’:** These are the principal audience for our lottery funding programme. The term includes everybody, but of particular concern are the hundreds of thousands, many retired, studying local and family history. Our primary interface is very simple: you type a post code into the home page, and immediately receive some basic information about the history of your area. Those with time will be able to follow link after link indefinitely, browsing from locality to locality, up to regions and down to parishes.

It is self-evident that any resource catering for such a wide audience must be network-accessible, and performance under heavy load is a major issue.

Core Architecture: GIS, thesauri and ontologies

So much of the system’s expected use is place-based that it is a gazetteer as much as a GIS as normally understood. Digital gazetteers are the focus of current activity linking GIS with digital libraries (Hill, 2000; NKOS, 2002). The JISC-funded Geo-X-Walk project is prototyping a UK gazetteer service, but lacks a historical dimension (EDINA, 2001). The *Thesaurus of Geographical Names* (Getty Information Institute, 2000) uses the thesaurus structure familiar to librarians, emphasising hierarchy not location. Conversely, ISO TC211’s gazetteer proposals borrow too heavily from GIS and can only handle change via vast redundancy. The Alexandria Digital Library’s (ADL) Gazetteer Content Standard (2002) offers a happier medium.

The heart of our new architecture is a set of relational tables containing no locational data: one key goal was to have a system which could hold information whose locations were unknown, or rather could only be inferred from relationships with other units. The clearest examples are administrative units associated with abandoned settlements, recorded only as rows in a medieval tax record, as parishes with names but no location; but we know which county and maybe which hundred they were in. However, we currently hold data on 1950s wards but mapping boundaries would need new funding.

The main sources used in constructing the core system are existing reference works which are among the most specialised, and maybe the hardest to use, bodies of ‘geographical information’ in existence. Figure 1 shows a page from Youngs’ *Local Administrative Units of England* (1979 and 1991) and we also have permission to computerise Melville Richards’ equivalent book for Wales (1969). For Scotland, we are extending an existing digital authority list of counties, burghs and parishes created by the Scottish Archives Network. The National Archives are supplying lists of English manors. Devising methods to computerise Youngs was difficult but the resulting structure has proved capable of holding data from other sources with minimal modification. Placing Youngs’ information into a database framework enables complete consistency checking as well as supporting web pages in which all cross-references are hyperlinks.

The result is neither a GIS, as it contains no locational information, nor a thesaurus, as relationships are not strictly hierarchical, but an *ontology*. The structure is not based on any existing standard precisely because it is designed to support multiple standards. For example, we support both the ADL Content Standard, which follows other librarians’ standards in distinguishing between *preferred* and *alternate* names for entities, and allowing only one *preferred* name for each entity, and the International Council on Archives ISAAR (CPF) standard (International Standard Archival Authority Record for Corporate Bodies, Persons and Families), which allows entities to have multiple consecutive official names. For example, Queen’s County in Ireland was renamed Laoighis post-independence.

Defining our entities is non-trivial:

- We record administrative units, not ‘places’. ‘Place-names’ are nicknames for vaguely defined areas, but administrative units are created, abolished, named and altered by clear legal processes. It is precisely because we define our entities as corporate bodies that we can allow them to change name, status and location. Of course, local authorities rarely change these drastically.
- The core system is a framework to which much locational data is tightly linked, so each unit has a fixed *type*. Some *types* are our own invention: ‘Ancient Districts’ include both Hundreds and Boroughs. The general aim is that selecting all the polygons for a given type at a given date provides a complete coverage of the country. Where two sets of boundaries, and therefore populations, existed simultaneously for what some would regard as the same unit, we require that two distinct units exist. Ancient, Registration and Administrative Counties of the same name are therefore distinct units of different types, which we link by *SucceededBy* relationships.
- Within each unchanging *type*, units can have many *status* values. These, for example, distinguish between the many Urban Districts and Rural Districts of the same name in the pre-1974 local government system. Youngs provides two separate entries for an Urban District that became a Municipal Borough, which not uncommon, but we treat it as a single entity with two consecutive status values.

Figure 2 shows the core set of relational tables holding our data, as distinct from the many associated metadata tables. The master list of all administrative units – counties, districts, parishes – that have existed is held in the **g_unit** table, with creation and abolition dates but without names or locations. Almost all unit characteristics can change: name (**g_name**), status (**g_status**), location and hierarchical position.

Locations are stored as polygons with beginning and end dates (**g_foot**), as further discussed in the next section. Hierarchic relationships are stored via a table (**g_rel**) which again includes dates and many non-exclusive relationships: e.g. parishes split between two districts. Theoretically, hierarchic relationships can be inferred from polygons but in practice this is limited by map accuracy. Any number of historical sources can document each item (**g_auth**).

Figure 2 also shows, in lighter grey, tables holding linked attribute data. Our statistical information is drawn from literally hundreds of transcribed tables in historical reports, but is all held in a single column of a single table (**g_data**). This simplifies the creation of maps and graphs from the data. The meaning of each value is defined entirely through metadata, not by the table and column it appears in. For now, we use a relatively simple structure which assigns each value to a 'variable' (e.g. the number of Catholics attending church), then to a 'group', (e.g. religious denominations in 1851) and finally to a theme (e.g. 'Religion'). This works well with simple frequency counts but not with complex cross-tabulations, such as age against gender against occupation. We are exploring approaches based on the Data Documentation Initiative's Aggregate/Tabular Data Extension (<http://www.icpsr.umich.edu/DDI>).

The system contains c. 5m. words of text from 19th century descriptive gazetteers: Goring's *Imperial Gazetteer of England and Wales* (1870-72, Groome's *Ordnance Gazetteer of Scotland* ((1882-5) and the first edition of Bartholomew's *Gazetteer of the British Isles* (1887). These data relate to 'places', including physical features like mountains, while entries for settlements may cover several distinct administrative units taking their name from it. The **g_text_gaz** table contains all the head-words in each gazetteer, while **g_text** holds the entries themselves, sometimes broken down into sub-entries for the administrative unit covered: for example, one for the parish of Ledbury, one for the sub-District and one for the District. **G_text_link** holds many-to-many relationships between gazetteer text and administrative units (Southall, 2003).

Space-Time Architecture

While the core tables contain no locational data at all, the system is designed to hold a very rich collection of date-stamped polygons defining historic boundaries: our own earlier work on the changing Civil Parishes of England and Wales since the 1870s; a separate record of Ancient Parishes created at Exeter University (Kain and Oliver, 2001), and converted into a GIS by us; and a new GIS we have built for Scotland. All this information is held within conventional database tables, using Oracle's spatial extension but also incorporating detailed chronological data.

Extensible database management systems are typically based on the relational model permitting easier and more natural data management by embedding abstract datatypes (ADTs) into the kernel while also supporting user-defined datatypes (UDTs). Object-Relational databases (ORDBMS) achieve extensibility by coupling the semantic richness of Object-Oriented with the operational efficiency of the Relational model (Stonebreaker & Brown 1999, Rumbaugh et. al. 1991; Date 2000). ORDBMS offer powerful modelling repositories for GIS (Worboys (1995;1999), Shekhar & Chawla (2002), Rigaux *et al* (2002)). Object support within the Oracle database began with version 8i, offering *object-types* that model complex real-world entities.

Object-types should not be confused with standard RDBMS objects such as tables, indexes and views. Object-types implement the object-oriented class concept, and can

be used in the same way as any other RDBMS data-types such as character and date, i.e. specified as the data-type of a column in a standard relational table, or declared as a variable in program code. Creating an object-type involves defining *attributes* and *methods*. Attributes hold the data for a given object, with each attribute itself assigned a data-type; these may be further object-types allowing objects to be nested. Methods are procedures defining the behaviour of an object-type and determine what it can do. Using Oracle object-types enables us to model the complex entities in our system, changing shape over time, and also permits inherently efficient query processing. For an introduction to Oracle Spatial see Sharma *et al* (2002).

Our system is both spatial and temporal, so we must be able to associate creation and abolition dates with any unit. We must also be able to assign dates to any unit attributes: names, status values, relationships and polygonal ‘footprints’. Further, these dates may be precise or just vague descriptions; two examples from Youngs are ‘early seventeenth century’ or ‘around the reign of Edward II’. We therefore define the DATE_T object-type:

```
GBHGIS.DATE_T(D_FULL_DATE  DATE,
              D_YEAR        NUMBER,
              D_DESC_STRING VARCHAR2(50))
```

If a full date is known, DD-MMM-YYYY information is stored in the D_FULL_DATE object attribute. If only the year is known, the D_YEAR attribute is used. Vague date references are stored in D_DESC_STRING. Mappings for the DATE_T object-type will enable us to scale, and therefore compare dates using boolean operators. Where only imprecise dates are given, object-type methods will use a look-up table to translate to conventional years; for example, ‘early seventeenth century’ would translate to the range 1600 to 1650, or to 1625. Crucially, the original imprecise term is stored in the system, and an alternative translation table could be used by researchers who disagreed with ours.

Recording the changing boundaries of administrative units combines boundary polygons with date values. We define a space-time object-type, or STOB, by combining two DATE_T instances, as already defined, with an instance of the standard Oracle Spatial SDO_GEOMETRY object-type:

```
GBHGIS.STOB_T(M_START_DATE  DATE_T,
              M_END_DATE    DATE_T,
              G_FOOT        MDSYS.SDO_GEOMETRY)
```

These nested object-types cross database schemas, and so the definition is qualified with the object owner. Populating this structure from the data we hold is complex, as the polygon coverages for single census years lack detailed chronology. We therefore first establish a crude chronology by comparing polygons for the same unit from adjacent censuses, discarding duplicate polygons and assigning a date range to those that remain. We then establish a more precise chronology from the boundary change data abstracted from Youngs and from census reports, which we separately load into the **g_rel** table and link to the master list of units.

While the core architecture enables us to hold information about units whose location and history is barely known, to some extent defining the agenda for future research, the space-time architecture enables us to hold detailed locational histories, representing each unit as a three-dimensional object.

Middleware Architecture

The core ontology plus its space-time extensions form a vast body of ‘local knowledge’, perhaps the largest resource for historical research and teaching ever assembled. The next question is how best to provide information from this structure to a large and very diverse user base. While the core data structure is monolithic – the only information held outside Oracle are geo-referenced image scans of historic maps – and tightly integrated, web content will be generated via a network of middleware servers with extensive caching: many user requests will be met without extracting new data from the core database. Further, while our immediate commitment is to provide a large but conventional web site, we also aim to provide a ‘historical geography’ server that other non-spatial sites can use to geographically enable their own content. This is further discussed below.

Our work is heavily based on standards developed by the Open GIS Consortium (www.opengis.org), especially their Web Map Server (WMS) standard. A WMS produces maps of geo-referenced data, a ‘map’ being defined as a visual representation of the geo-data, not the data itself. WMS clients may specify the information to be shown on the map as one or more ‘Layers’, possibly the ‘Styles’ of those layers, the portion of the Earth to be mapped (a ‘Bounding Box’), the geographic coordinate reference system to be used (the ‘Spatial Reference System’ or SRS), the output format and size, and background transparency and colour. WMS requests are sent as URLs, so all web browsers able to display images and forms work as clients. Clients are generally smarter, but the simplest requests can be submitted via an html form to a server program which generates the new URL from the form and returns that to the browser.

When two or more maps are produced with the same Bounding Box, SRS and output size, the results can be layered to produce a composite map. The use of image formats supporting transparent backgrounds allows the lower Layers to be visible. Furthermore, individual map Layers can be requested from different Servers. The WMS specification thus enables the creation of a network of distributed Map Servers from which Clients can build customized maps. A particular WMS provider in a distributed WMS network need manage only its own data collection, in contrast to vertically-integrated web mapping sites that must hold all the data they map.

Figure 4 identifies the main elements in the web site architecture. Web pages will be generated by Java Server Pages (JSPs) running within Tomcat. Each JSP interprets a request from a web browser, in the form of a URL, and converts it into requests for historical content from the database; the requests are turned into JavaBeans, small re-useable pieces of code, that execute an SQL (Structured Query Language) query against the database via the JDBC protocol. As well as querying the database, JavaBeans can also call upon specially tailored application services to manipulate and present the information returned, as plain text, a table, pie-chart or thematic map. Most of the work of our site will be carried out by three such applications or server processes:

- The generation of graphs is carried out by a custom servlet used to wrap JFreeChart (an open source Java graphing package). The servlet was developed in Leeds by members of the Centre for Computational Geography, and is tightly coupled with the underlying database and web site. In future it is hoped to make the servlet more general. JFreeChart is a reasonably standard graphing API. See

<http://www.jfree.org/jfreechart/index.html>

- Maps are generated by two distinct OGC web map servers. The first is a Java implementation of the specification using the GeoTools toolkit. While not fully compliant, this WMS creates maps from vector data styled using the OGC Styled Layer Description (SLD). This ability allows the production of thematic maps from statistical data, where the styling of the map is based upon the values of attributes associated with the polygons representing the mapped units. See:

<http://www.geotools.org>

- The project is constructing two complete raster coverages of Great Britain by geo-referencing image scans of Ordnance Survey one inch maps, the original 19th century *First Series* and the 1940s *New Popular Edition*. The second WMS provides access to this content, and is implemented using the open source MapServer. This does not currently implement the full OGC specification, but is particularly well suited to handling the large raster datasets used in the background maps required on the GBHGIS site. MapServer is a CGI program run directly by the Apache web server, and unlike all other parts of the system manages its own data rather than extracting it from Oracle. One significant problem with a set of images this large (each of the 400 maps is around 60 Mb) is how to manage them in order to serve them efficiently. Although the maps have been scanned at 24 bit for archiving purposes, the maps display adequately using 8-bit colour and MapServer is also more efficient in this mode. The maps have therefore been pre-processed to lower their colour depth. MapServer allows the rasters to be tiled and builds a spatial index over the tiles, so that only tiles for the excerpt currently being displayed need be sent to the browser; panning involves downloading only some additional tiles. This is particularly useful for the larger sheets of the *New Popular* series. Each tile covers an area of one Km². See:

<http://mapserver.gis.umn.edu>

- A fourth server process is being implemented to support the Alexandria Digital Library Gazetteer Service Protocol; the core database already supports the ADL content standard. This again involves Java code running within Tomcat. This facility will not contribute to our own web site, but will enable us to provide place-name searching facilities to other sites. We are also interested in enabling our system to function as a self-contained TimeMap Clearing House, supporting a standard for linking cultural web sites developed by the Electronic Cultural Atlas Initiative (ECAI). See:

<http://alexandria.sdc.ucsb.edu/~lhill/adlgaz>

<http://www.timemap.net/>

The great virtue of this architecture is that it modularises what would otherwise be a very large and sprawling web site, and creates a large potential for dividing the processing load between many distinct servers. The initial site is hosted on a single computer, but if and when usage grows we will be able to easily extend capacity, unlike the well-known recent problems of the Public Record Office's 1901 site. The final section discusses how we can also make use of this modularity to publish aspects of the site as web services.

User Architecture: Sites and Services

The middleware architecture gives us enormous flexibility in what facilities we offer our user communities: we can create an almost unlimited range of ‘web sites’ and we will also be publishing aspects of the resource as web services.

- Our current national lottery funding means our first priority is creating a site for ‘life-long learners’, with primarily local interests and accessing the site from computers in public libraries (‘The People’s Network’), record offices (the ‘New Archives Network’) and home. This is the sole concern of our Phase I system, which places 2001 census data for current local authorities in long-run perspective. Figure 3 shows its home page. We expect most users to select a locality by typing in their postcode. The statistical data are presented mainly as maps and graphs.
- Our second priority is to provide a more professional interface for librarians and archivists using our system as a placename authority. At least initially, we will cater to them via an ‘expert search’ option in our sidebar, which will permit more sophisticated searches by type of unit (wappentake, liberty or hundred), geographical area and part of the country. It will also provide fuller details of sources.
- It seems likely that at least one more web interface will be needed, tailored for schools. This would include alternative explanatory text, and it may be possible to offer facilities for teachers to prepare and store teaching materials which incorporate relevant parts of our content tailored to that particular school’s catchment.

Interest in web services has expanded rapidly, and we have already explained how aspects of our data can be made available using OGC and Alexandria Digital Library standards. Although these service implementations are of some technical interest, explaining to people in the UK heritage sector how they will be of practical use is a major task for the second half of the project; we have certainly gone beyond anything our ‘customers’ required. A relevant scenario is:

- Individual librarians and, especially, archivists cataloguing materials will of course be able to go to our site, rather than to a book like Youngs, to check the definitive spelling of a place-name. However, it would be much easier if their cataloguing software functions as an ADL client, automatically checking names they typed in. Of course, once this was done it would be fairly trivial to extend their system to hold either a grid reference or a unit ID number from our system.
- Holding polygons and creating maps requires spatial functionality which most database systems lack. However, most software can hold numeric values, and therefore hold geographical coordinates; if they can compute maximum and minimum values, they can construct bounding boxes from a set of points. A set of dots is not really a map, but if that set is superimposed on a historic scanned map it begins to provide real geographical display; and that is precisely what our Web Map Server implementations will permit.

We therefore hope that once museums and archives start using our system to improve their catalogues, they will begin to see how they can geographically-enable their digital holdings in new ways. Although we are operating in a sector which until recently had little experience with GIS, it is a sector where standards-based interoperability is widely understood. Our work is based heavily on open source software, so implementation costs can be quite modest. It seems quite possible that there will be more rapid take-up of ‘open GIS’ here than in sectors where GIS in general is longer established.

References

- Alexandria Digital Library** (2002), Gazetteer Content Standard,
<http://alexandria.sdc.ucsb.edu/~lhill/adlgaz>
- Berman, M.L.** (2002) 'CHGIS Database Model: Development and Implementation';
http://www.fas.harvard.edu/~chgis/work/docs/papers/lex_davis.pdf
- Curtis, S.E., Southall, H.R., Congdon, P. and Dodgeon, B.** (2003), 'Area effects on health variation over the life-course: analysis of the longitudinal study sample in England using new data on area of residence in childhood', *Social Science and Medicine*, in press.
- De Moor, M., and Wiedemann T.** (2001, in press) 'Reconstructing territorial units and hierarchies: A Belgian example', *History and Computing*, vol. 13(1).
- EDINA** (2001), Geo-Crosswalk - A Feasibility Study into the Provision of a British and Irish Gazetteer Service, <http://edina.ac.uk/projects/crosswalk.shtml>
- Getty Information Institute** (2000), Thesaurus of Geographical Names,
<http://www.getty.edu/research/tools/vocabulary/tgn>
- Gregory, I.N., and Southall, H.R.** (1998) Putting the Past in Its Place: the Great Britain Historical GIS, in Carver, S. (ed.) *Innovations in GIS 5*, Taylor & Francis), pp.210-21
- Hill, L.L.** (2000), Core elements of digital gazetteers: placenames, categories, and footprints, in Borbinha, J., and Baker, T. (eds.), *Research and Advanced Technology for Digital Libraries: Proceedings of the 4th European Conference, ECDL Lisbon, September 18-20, 2000*, Springer, pp. 280-290;
http://www.alexandria.ucsb.edu/~lhill/paper_drafts/ECDL2000_paperdraft7.pdf
- International Council on Archives** (2002), International Standard Archival Authority Record for Corporate Bodies, Persons and Families, Draft Second Edition
- Kain, R.J.P. and Oliver, R.R.** (2001), *Historic Parishes of England and Wales: An electronic Map of Boundaries before 1850 with a Gazetteer and Metadata*, History Data Service UK Data Archive.
- National Council on Archives** (1997) Rules for the construction of place names,
<http://www.hmc.gov.uk/nca/rules3.htm>
- New Opportunities Fund** (2002), NOF digitise technical standards and guidelines, Version Four,
http://www.peoplesnetwork.gov.uk/content/ts_index.asp
- NKOS** (Networked Knowledge Organization Systems), 'Digital gazetteers: Integration into distributed digital library services', JCDL2002 Workshop, Portland, Oregon,
<http://nkos.slis.kent.edu/DL02workshop.htm>
- Open GIS Consortium** (2002) Web Map Services Specification,
<http://www.opengis.org/techno/specs/01-068r3.pdf>
- Quality and Curriculum Authority** (1999), National Curriculum for England: History: Key Stages 1-3,
<http://www.nc.uk.net/download/cHi.pdf>
- Rigaux, P. Scholl, M. & Voisard, A.** (2002). *Spatial Databases With Application to GIS*. Morgan Kaufmann Publishers.
- Rumbaugh, J. Blaha, M. Premerlani, W. Eddy, F. & Lorensen, W.** (1991) *Object-Oriented Modelling and Design*. Prentice Hall.
- Sharma, J., Herring, J. & Ravada S.** (2002) *Oracle Spatial. An Oracle Technical White Paper*.
- Shekhar, S. & Chawla, S.** (2002) *Spatial Databases: A Tour*. Prentice Hall.
- Southall, H.R.** (2003) 'Defining and identifying the roles of geographic references within text: Examples from the Great Britain Historical GIS project' in Kornai, A. (ed.), [Papers presented to North American Chapter of the Association for Computational Linguistics' workshop on the analysis of Geographical Names] [NB interim title; workshop is in 2 weeks time but papers are already on the web: <http://www.kornai.com/NAACL/Conf/index.html>]

- Stonebraker, M. & Brown, P** (1999) *Object-Relational DBMS: Tracking the Next Great Wave*. Morgan Kaufmann.
- Worboys, M. F.** (1999) 'Relational Databases And Beyond'. In Longley, P. A., Goodchild, M. F., Maguire, D. J. & Rhind, D. W. (eds.) *Geographical Information Systems, Volume 1, Principals and Technical Issues*. Second Edition. Jon Wiley & Sons, Inc.
- Worboys, M. F.** (1995). *GIS: A Computing Perspective*. Taylor and Francis.
- Youngs, F.** (1979 and 1991) *Local Administrative Units of England*, 2 vols., Royal Historical Society

- Cheerwood EP.²¹⁶ Salf. RDn. Abol 1967 to help cr Lower Broughton St Clement with St Matthias EP.³⁰⁵
- LOWER BROUGHTON ST CLEMENT WITH ST MATTHIAS**
EP Cr 1967 by union Lower Broughton St Clement EP, pt Salford St Matthias with St Simon EP.³⁰⁵ Salf. RDn.
- BROUGHTON EAST**
Tp and chap (built 1745³²²) in Cartmel AP, sep CP 1866,⁴ sep EP 1748 as 'Field Broughton',¹² as such eccl refounded 1875,³⁰⁴ qv. *LG* Lonsd. Hd, Ulv. PLU, pt Grange USD, pt Ulv. RSD, Ulv. RD (1894—1960), N Lonsd. RD (1960—74). Civ bdry: 1884,⁵¹ 1884,⁵² 1884,⁵³ 1884,⁵⁵ 1884,⁵⁶ 1894 (loses the pt in the UD to help cr Grange CP),⁵⁸ 1914.⁵⁹ Transf 1974 to Cumbria.⁴⁰ *Parl* Seq 10.
- BROUGHTON ST JOHN [THE EVANGELIST]**
EP Cr 1878 from Broughton EP.³⁰³ Salf. RDn.
- BROUGHTON IN FURNESS**
Chap in Kirkby Ireleth AP, sep EP [St Mary Magdalene] 1870,³⁰⁶ sep CP 1866 as 'Broughton West',⁴ qv. Ulv. RDn. Eccl bdry: 1886.⁸ Abol 1956 to help cr Broughton in Furness with Woodland EP.²⁴⁴
- BROUGHTON IN FURNESS WITH WOODLAND**
EP Cr 1956 by union Broughton in Furness EP, Woodland EP.²⁴⁴ Ulv. RDn (1956—70), Furness RDn (1970—*).
- BROUGHTON WEST**
Chap in Kirkby Ireleth AP, sep CP 1866,⁴ sep EP 1870 [St Mary Magdalene] as 'Broughton in Furness',³⁰⁷ qv. *LG* Seq 14. Bdry: 1935 (exchanges pts with Millom Rural CP, Cumb).³⁰⁷ Transf 1974 to Cumbria.⁴⁰ *Parl* Seq 10.
- BRUNSHAW**
CP Cr 1894 from the pt of Burnley CP not in Burnley refounded 1854,¹⁹⁹ 'Burnage' a sep EP 1875 from Withington EP, Barlow Moor EP.¹⁵⁰ *LG* Salf. Hd, Chorlton PLU, Withington USD, UD (1894—1904), Manch CB (1904—10). Civ bdry: 1883.³¹ Abol civ 1910 ent to Manch CB (assoc with Lancs) and to South Manchester CP.³¹⁰ *Parl* S'rn Dv (1832—67), S-E Dv (1867—85), Stretf. Dv (1885—1918). *Eccl* Ardwick RDn (1875—1912), Heaton RDn (1912—*). Eccl bdry: 1900 (help cr Ladybarn EP),²⁰² 1905,¹⁹⁹ 1927,¹³⁵ 1929.²⁰⁴
- BURNAGE ST NICHOLAS**
EP Cr 1929 from Barlow Moor EP,¹⁵² Heaton RDn.
- BURNLEY**
The following have 'Burnley' in their names. Insofar as any existed at a given time: *LG* Black. Hd, Burnley PLU, pt Burnley USD, pt Burnley MB (1861—89 [enlarged pt 1871—89]), Burnley CB (pt 1889—94, ent 1894—1974), pt Burnley RSD. *Parl* N'rn Dv (1832—67), Burnley Parl Bor/BC (1867—*). *Eccl* Blackb RDn (until 1847), Whalley RDn (1847—72), Burnley RDn (1872—*).
- CP1/EP1—BURNLEY [ST PETER]**—Chap in Whalley AP, sep CP 1866,⁴ sep EP 1716 (incl tps Briercliffe with Extwistle, Cliviger, Habergham Eaves, Worsthorpe with Hurstwood; also incl 2 civ tp/eccl ex-par areas of Ightenhill Park [eccl, 'Ighton Hill' until eccl abol 1845 to help cr Habergham All Saints EP¹²⁴] and Reedley Hallows, Filly Close and New Laund Booth).¹² Civ bdry: 1894 (the pt not in the CB cr Brunshaw CP, and gains the pts in Burnley CB of Briercliffe with Extwistle CP, Habergham Eaves CP, Ightenhill Park CP; Reedley Hallows, Filly Close and New Laund Booth CP),²⁹⁵ 1911,³⁰⁸ 1926.³¹¹ Eccl bdry: 1741 (pt orig area tp Cliviger cr 'Holme in Cliviger' EP,¹² refounded 1843 as 'Holme St John'²⁹³), 1837 (orig area tp Worsthorpe with

Figure 1: Sample page from Youngs' *Local Administrative Units*

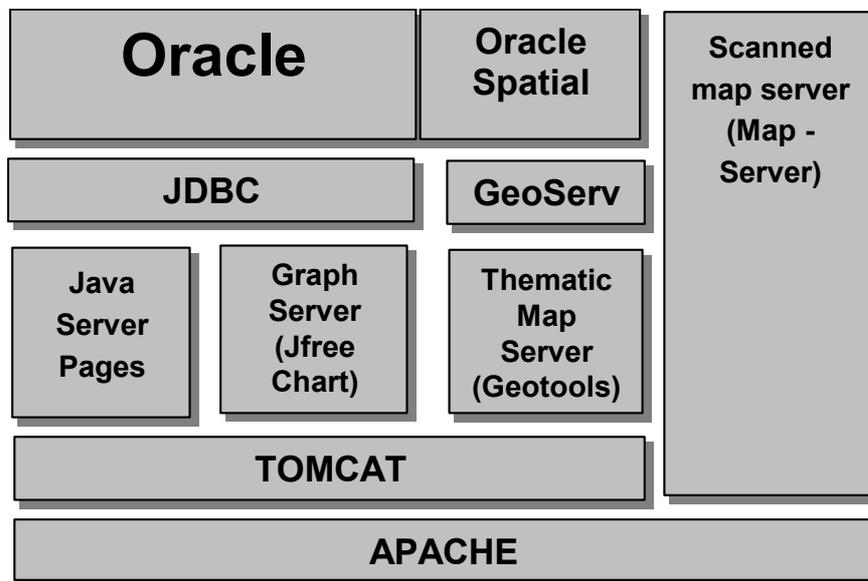


Figure 3: GBH GIS Middleware Architecture

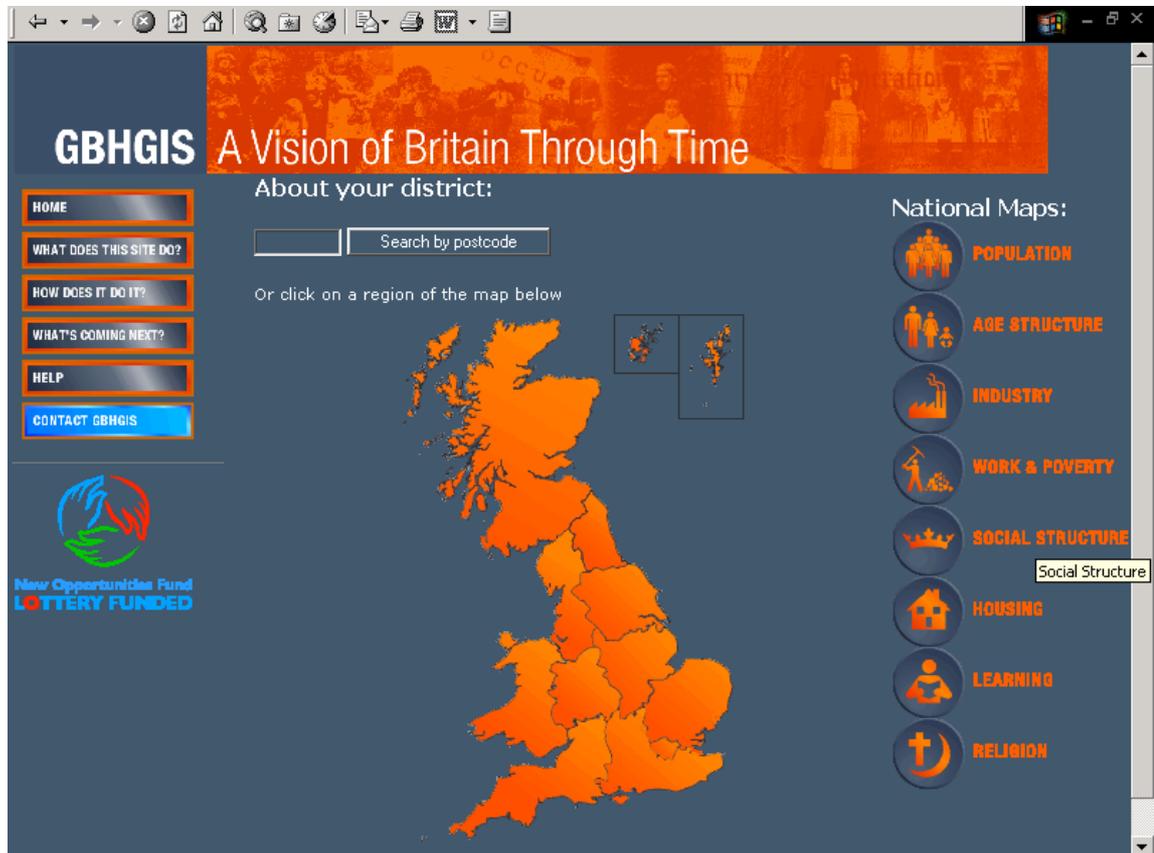


Figure 4: GBH GIS Phase I Home Page