Using Open Standards To Build A Web Based Gis For Public Use

Ian Turton[1], Humphrey Southall[2]

[1]Centre for Computational Geography, University of Leeds, Leeds, LS2 9JT.
Phone: +44 113 343 3392,
Fax: +44 113 343 3308,
Email: ian@geog.leeds.ac.uk

[2]Great Britain Historical GIS Project, Department of Geography,
University of Portsmouth, Portsmouth PO1 3HE,
Email: Humphrey.Southall@port.ac.uk

Biography

Ian Turton is director of the Centre for Computational Geography and the GeoComputation Research Fellow in the School of Geography at the University of Leeds. Research Interests include democratisation of geographic and statistical information, development of open web mapping standards, and novel methods of visualisation of geographic data.

Introduction

This paper discusses the development of the geographic front end to the Great Britain Historical GIS. The system is built around a series of open standards which allow easy extensibility and external use of the data contained in it.

Existing 'historical' GIS systems fall into two categories. Firstly, those which hold historical data but lack a time dimension, covering just one date. Most archaeological systems record the locations of 'finds' rather than reconstruct past landscapes. Secondly, systems that incorporate time by using mainstream GIS technology in novel ways. The original Great Britain Historical GIS (Gregory and Southall, 1998) was a good example of this, implemented within ArcInfo using extensive AML to manipulate dates. This worked well on a small scale but achieving high data integrity proved hard once the system was scaled up to record changing boundaries across 15,000 parishes. Like other national historical GIS systems, it was created primarily to support academic research, primarily demographic.

New National Lottery funding has created both the opportunity and the need for a new system, providing a more consistent framework for both our existing statistics and digital boundaries, and new material including descriptive text and historical maps. This system is designed as a national historical GIS not just **of** Britain, covering the whole country, but **for** Britain, freely accessible over the web to the general public. The original system emerged from historical demography, supported by the Office of National Statistics. This new system involves collaboration with English Heritage, the British Library, and the network of local archives.

Expectations of a national historical GIS

Our system is a **historical** GIS, a framework for **documentary** information from the past, not the objects recorded by archaeologists. We are mainly concerned with relatively large administrative units or vaguely defined 'places', not precisely located archaeological remains. We need to be able to store information about a few administrative units whose locations are unknown. Conversely, our dates are often precise. It also means we primarily liaise with libraries and archivists, not museums and archaeologists.

Archivists are not our ultimate customers. 'Life-long learners' include all English nine year olds, 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', using information on 'education; population movement; treatment of the poor and care of the sick …' (Quality and Curriculum Authority, 1999). They also include the millions researching local and family history. The result of building for this target audience is that the information has to be presented in a manner which is easily accessible to users with little or no background knowledge of the system. Users may also be unsure of what questions they would like answered. The system is therefore designed to allow the exploration of the data via a series of interactive maps and pre-designed graphs of the variables. As in all census activities it is vital that the correct denominator is chosen for each variable before converting them to rates. The database holds information as to which is the right denominator for each variable allowing an appropriate graph to be presented to the user.

Supporting web sites and web services via open standards

The central issue is how to provide information from the GIS to this large and very diverse user base. While the core data structure is monolithic – the only information held outside Oracle will be geo-referenced image scans of historic maps – and tightly integrated, web content will be generated via a network of middleware servers with very 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 via web services; for example, by converting a set of place names into co-ordinates, then creating a map superimposing them on a suitable historic base. This communication must be open standards based.

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 lacking a historical dimension (EDINA, 2001). The Thesaurus of Geographical Names (Getty Information Institute, 2000) adopts librarians' familiar thesaurus structure, 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 Gazetteer Content Standard (2002) offers a happier medium and is easily supported by our core database. The required fields are place-name, an arbitrary ID, a 'feature type' and a location, but extensions include alternative names, chronology and polygonal 'footprints'. We are also implementing the International Council on Archives' standard for Corporate Bodies (2002), as each of our administrative units was a corporate body as well as a polygon.

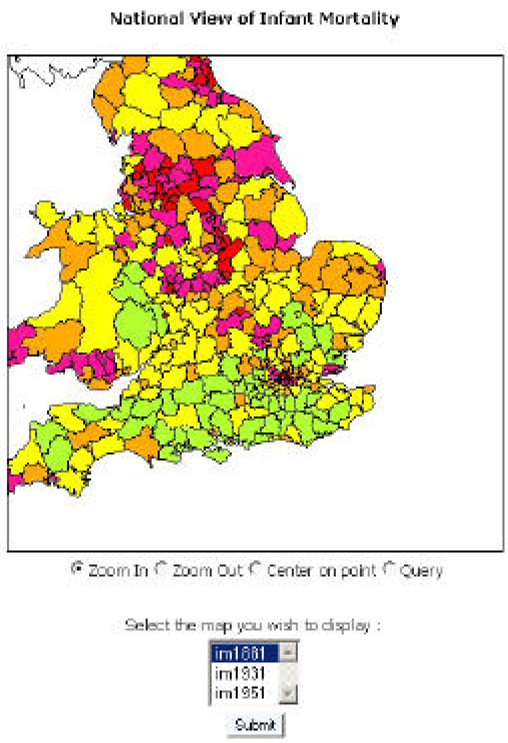

Figure 1: Interactive Mapping using the GeoTools WMS, users can zoom, pan and query the map, as well as changing the year displayed

Once users and client systems obtain a location, they want geographical visualisation of the location and of associated attribute data. These are provided by two separate OGC-compliant web map servers (OGC, 2002). The first, using MapServer (http://mapserver.gis.umn.edu) holds geo-referenced scanned images of 19th and 20th century 1" maps. The second, based on GeoTools (http://www.geotools.org) software running within a Tomcat server, generates thematic maps from statistics and polygons (see figure 1). A Web Map Service (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 (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 clients can be simply a web browser. Clients are generally smarter than this, but the simplest requests can still be submitted via a simple html form to a server program which generates the new URL in the form and returns that to the browser. This allows a WMS to be accessed by any browser capable of displaying images and forms.

When two or more maps are produced with the same Bounding Box, SRS, and output size, the results can be accurately 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 only be the steward of its own data collection. This stands in contrast to vertically-integrated web mapping sites that gather in one place all of the data to be made accessible by their own private interface.

It is this simplicity and standardised interoperability that allows us to make use of two WMS providers to spread the load across machines. By providing the two services separately we can allow others to make use of the maps in their own mapping on different web sites and on their desktop if they have an OGC compliant GIS. For example, if a web site requires a map of infant mortality in 1881 for the parishes of a county. Instead of copying the boundaries and statistical data to their server it is possible to use a simple link to the web map server inserted in the page, thus if the data are updated the map will be updated automatically. Alternatively if a historical raster map for a locality is required as a backdrop for a project it can be easily imported into the GIS from the WMS and then locally held data can be drawn over it. This saves the user downloading large amounts of raster data for the area of interest and (nearly always) the surrounding tiles as well.

The end result will be a system which is both highly scalable, to meet the needs of very large number of users, and very flexible, able to present many different faces to the world and collaborate in the rapidly evolving new world of web (geo)services.

References

Alexandria Digital Library (2002), Gazetteer Content Standard, `http://alexandria.sdc.ucsb.edu/~lhill/adlgaz`

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

National Council on Archives (1997) Rules for the construction of place names, `http://www.hmc.gov.uk/nca/rules3.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`

NKOS (Networked Knowledge Organization Systems), ‘Digital gazetteers: Integration into distributed digital library services’, JCDL2002 Workshop,

Portland, Oregon, http://nkos.slis.kent.edu/DL02workshop.htm