Alternative ways of indexing by geography

Many archives index their collection geographically to make records more accessible for local history. However, a recent survey by the UK Archives Discovery Network found great variation in working practices between archives. Some follow the recommendations of the National Council on Archives' *Rules for the Construction of Personal, Place and Corporate Names* (1997) and index by the most relevant administrative unit, generally a historic parish; some index by less formal place; and some use a geographical coordinate or location. What are the pros and cons?

I created the Great Britain Historical GIS (Geographical Information System), and the web site *A Vision of Britain through Time* which accesses it. Our lottery funding was obtained with strong support from the National Register of Archives, and a central aim was to computerise the name authorities identified in the *NCA Rules*. http://www.ncaonline.org.uk/materials/namingrules.pdf

Uniquely, our database supports systematic cross-walking between administrative units, places and locations.

Administrative units

The *NCA Rules* emphasised indexing by parish, and identified Youngs '*Guide to the Local Administrative Units of England* (1979 and 1991). Richards*' Welsh Administrative Territorial Units* (1969) and the *Index of Scottish Place Names from 1971 Census* as authorities. We obtained permission to computerise Youngs and Richards, and also R. Cheffins' *Parliamentary*
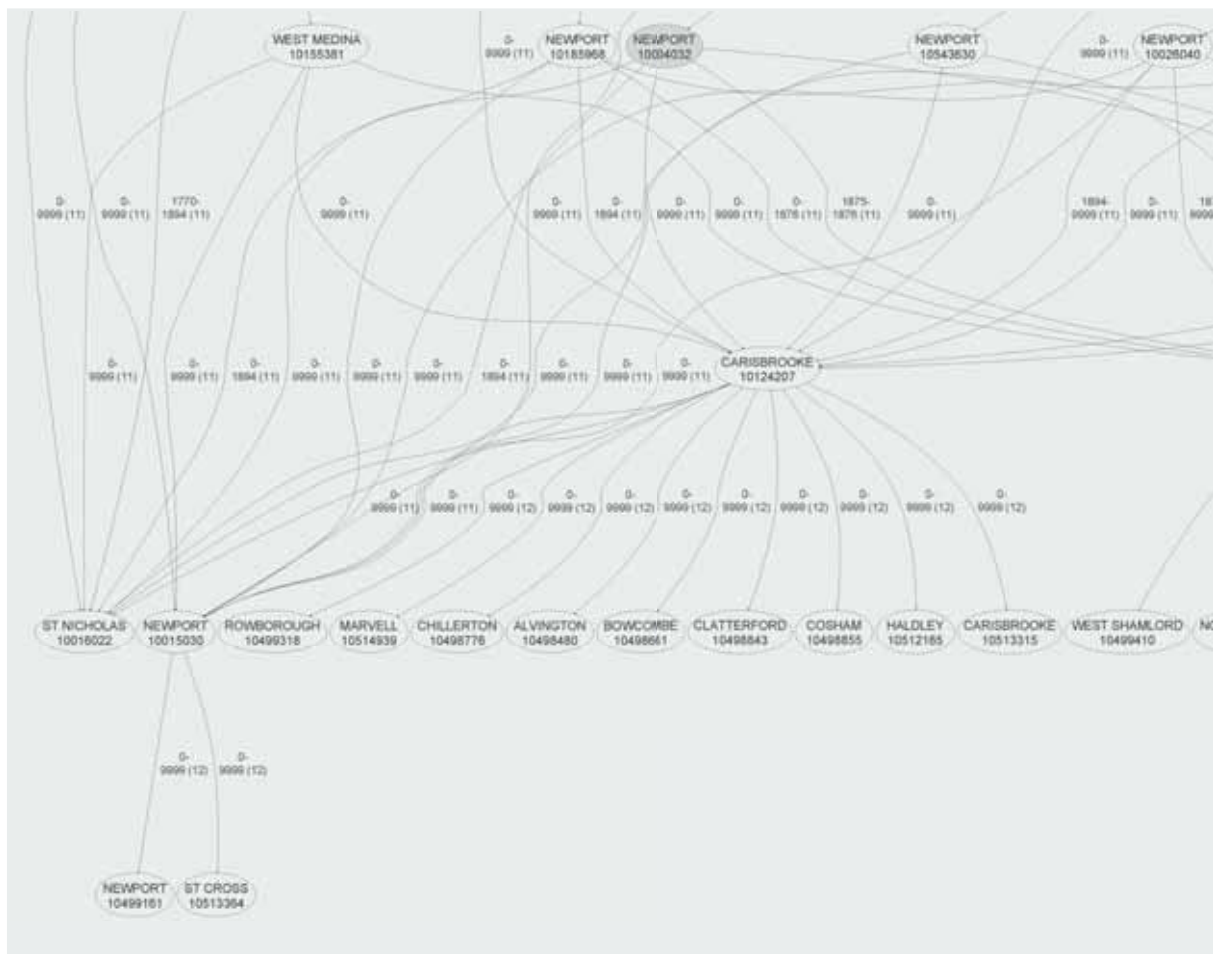
A visualisation of the 'IsPartOf' relationships, with dates, linking the parish of Carisbrooke, Isle of Wight, to higher-level districts, to its component manors, and to the component chapelries of Newport and St. Nicholas, as held in the GB Historical GIS based on information in Youngs' Local Administrative Units and the Manorial Documents Register. NB this shows just 20 units out of 78,470. Thanks to Vojtech Kupca for the graphic.

Although Carisbrooke the place is a village west of Newport, Isle of Wight, the parish of Carisbrooke at one time included most of Newport and a large surrounding area. 'Historic parishes' were not set in stone: Carisbrooke's boundaries were revised in 1882, 1894 and 1933, and the map shows these changes. Boundary lines were researched from published administrative area maps and unpublished maps in the National Archives, and are shown here overlain on a 1940s New Popular Edition one-inch map.

Constituencies And Their Registers 1832-1996 (2005). We also drew on the gazetteer of counties, parishes and burghs created by the Scottish Archives Network. This information is best searched via our expert search interface: http://www.visionofbritain.org.uk/units

The advantages of indexing by administrative units are firstly that they are corporate bodies which create records, and secondly that they are areas not points, so a limited set of terms indexes all possible locations and brings related records together. Indexing by parish provides a relatively detailed geographical partitioning, and the system of parishes has been much more stable over time than the many different systems of districts.

Unfortunately, the names of administrative units will not always be familiar to users, and familiar names may mislead. Further, while parishes have been more stable than districts there have been substantial changes over time. Although parishes do partition up the country, archivists have not always had access to the necessary maps (*A Vision of Britain* can identify an ancient parish, or a registration district, from a postcode typed into the home page).

Places

While administrative units are defined in law, 'places' are simply locations people give names to. Their great advantage for a user interface is that they are how most people think about geography. Identifying a satisfactory controlled vocabulary of places is problematic, but technology now makes it easy to offer cross-walking from places to administrative units or locations. The Ordnance Survey's *1:50 000 Gazetteer* has 259,080 entries, but the only sense in which it

includes variants names is that some 'names' actually include 'OR'. The Getty Information Institute's *Thesaurus of Geographical Names* includes variants, but misses substantial settlements. The crowd-sourced *GeoNames* includes variant names and initially seems vast, but is stronger on the names of hotels in tourist areas than on British villages.

All these sources lack provenance, and include duplicate entries. We have worked hard to limit these problems in our own list of places. It is far smaller than the OS50K or GeoNames but our 16,000 or so 'places' include every place that gave its name to an Ancient Parish, to a Civil Parish in 1911, to a constituency or to most kinds of district, from ancient hundreds to modern local authorities. Over 80,000 place-names are linked to these places, each with an attribution.

Another problem with 'places' is that it is hard to define them as anything other than points. Many records relate to a whole county or even the whole country, which in a place-based system may make it very difficult to separate the records of a County Council from the local history of the county town.

Locations

Indexing by geographical coordinates would be useless in a traditional card catalogue. However, in a computerised system coordinates offer precision, and easy identification of geographically related records. The downside is that the accuracy may be spurious, in two ways. Firstly, archival documents rarely contain coordinates and old maps will often be topographically inaccurate. The geography in historical documents consists of place-names, and associating these with

coordinates often requires detailed knowledge of both geography and the historical forms of names. Coordinates are therefore contestable interpretations.

Secondly, holding coordinates is likely to mean holding points. It is relatively easy to modify a standard records management system to hold points as pairs of numbers, and equally easy to programme in functions that find available records within a defined rectangle or within a defined radius from some point. The problem is that records, which relate to a large area, such as a district or county, have to be indexed via a single point. This problem is avoided if the computer system holds polygons as well as points, but adding that functionality to existing management systems is much harder. GIS software is about exactly this kind of functionality but can be very expensive and of course lacks all the other functionality archivists need.

Conclusion

Mainstream authorities on geographical information, without experience of historical sources or the requirements of records management, would almost certainly recommend storing coordinates not names. The same is probably even truer of your corporate people who have picked up a smattering of GIS concepts, but the result will often be catastrophically bad resource discovery.

All approaches to geographical indexing have significant problems, and your best approach will depend on the collection type. I am sure of three points. Firstly, if you decide to index historical documents by location, you must treat the coordinates as interpretations and also retain the placenames. Secondly, you should distinguish between what you hold in your index and what you allow your users to search by, and the latter should include informal places. Thirdly, there are limits to how much geo-spatial functionality can be included in mainstream records management software, and efficient cross-walking between users' search terms and the actual index, or to link collections, needs more specialised capabilities available as web services. We are very interested in discussing how our system could be used in this way.

Humphrey Southall
Department of Geography, University of Portsmouth

Tracks in Time

The Leeds Tithe Map Project

The tithe maps and associated apportionment records cover the modern City. They provide details of land ownership, land occupancy and land use between 1838 & 1861 and they represent the earliest systematic, large-scale cartographic record of the area. In an effort to improve access to Leeds' tithe map collection, with funding from the Heritage Lottery Fund, the *West Yorkshire Archives Service* has undertaken a project remove physical, financial and intellectual barriers to these important resources.

Although generally in good condition, several of the plans had become fragile and suffered peripheral damage following many years of use. Some had accumulated a layer of coal dust due to previous exposure to poor storage conditions. The collection has been conserved and captured digitally to provide free online access at www.tracksintime.wyjs.org.uk

High-resolution scans have yielded top-quality digital surrogates of the maps, while data from the apportionments has been transcribed by volunteers and linked to the relevant tithe plots and townships. Researchers can now interrogate the resource and obtain answers and display results to suit individual requirements. The user may query the database to find a specific landowner, for instance, and be taken to the portion of the map once held by that person. Or they might assess land usage and highlight all orchards, woodland and other land types within a given area.

For further information please email: tracksintime@wyjs.org.uk

Stefanie Davidson
West Yorkshire Archive Service

Use of RD/RT/10 with thanks to West Yorkshire Archive Service, Leeds