CHALICE USE CASE

CLERGY OF THE CHURCH OF ENGLAND

Introduction

The Clergy of the Church of England Database (CCEd) began primarily as a prosopographical reference resource listing the names of all clerics ordained in the Church of England between 1540 and 1835 (a relatively short period in the context of placename history). It provides a list of names, and began with the relatively straightforward ambition of documenting the name of every ordained person between those dates, drawing on a wide variety of historical sources. The main aim was to provide a ready reference resource to historians of subjects to which clerical lists are of some degree of relevance. The technical research and development of CCEd is the responsibility of the Centre for Computing in the Humanities (CCH) at KCL. Like so many other online historical resources, CCED's main motivation is to bring things together, in this case information about the CofE clergy between 1540 and 1835, just after which predecessors to the Crockford directory began to appear. There is, however, a certain divergence between what CCED what Crockford (simply a list of names of all clergy) provides. CCEd seeks to explicitly contextualize the names with certain historical parameters. CCEd's sources vary greatly from diocese to diocese, but include such archival resources as parish records, clergy lists, and church records of letters testimonial, presentations, subscriptions and resignations. For an overview of the sources, see http://theclergydatabase.org.uk/about/database/sourceessays.html.

In the execution of the project, it swiftly became apparent that the list of names in and of itself would be of only limited use in paper form; that it would be unwieldy and cumbersome to use. Therefore, a digital approach was needed to cope with the sheer quantity of information involved. CD-ROMS were mooted at first; although it was subsequently decided that the resource should exist as an online database. In this sense, there are close parallels between the CCEd and EPNS: both are what might be termed 'strategic content'. Although they have intrinsic value for domains which are highly specific, perhaps even niche, (clerical prosopography and English placename research), both sets of content act (or rather are able to be used as) pointers along information pathways about persons, places and events, and the documentation of these through time. For this to be realized however, they need to be available as online reference resources. This is certainly the original spirit in which CCEd was conceived, and is key to the value identified by CHALICE in EPNS. CCEd staff also decided at an early stage that a facility to build queries around location would be critical to

the use historians make of the resource. There is therefore clearly scope for considering how Chalice and CCED might complement one another (of which more below).

Usage

Even more importantly however, some of the issues which CCED have come up against in terms of structure have a direct bearing on CHALICE's ambitions. What is most interesting from the point of view of a linked data resource that could link to CCEd was the great complexity which the geographic component contains; and also the variable nature of the source material. It is important to note that there was no definitive list of English ecclesiastical parish names prior to the CCED (crucially, what was needed, was a list which also followed through the history of parishes – e.g. dates of creation, dissolution, merging, etc.), and this is a key thing that CCED provides, and is and of itself of great benefit to the wider community. In this sense, it is the very comprehensiveness of EPNS that acts as a barrier to linking to other resources. Since it is the EPNS is the main scholarly source for the placename material, in many cases there is no corresponding material to link to beyond what is already represented and documented in EPNS.

Location

In general terms, an interesting aspect that emerges when the EPNS/CHALICE resource is compared with CCEd is the evident tension which the latter, by its nature, places on location. CCEd employs a rudimentary feature type thesaurus to deal with the different kinds of entity (see http://www.theclergydatabase.org.uk/about/database/stype001.html). These are: Parishes, Parochial Chapels, Non-parochial Chapels, Schools, Diocesan offices, Domestic chaplaincies, Locations in Cathedrals and Collegiate Churches, The Chapel Royal, Secular institutions and semi-ecclesiastical institutions. This is in parallel to the units used by EPNS, and loosely defined by CHALICE – county, parish, and sub-parish units. Many other digital scholarly resources have similar feature types, and it makes little sense building semantic links to placenames from any system that does not in some way reflect these. References to location in CCED, as distinct from types of feature, is dealt with in two ways: jurisdictional and geographical. The jurisdictional element is particularly significant in that it ties location data to the idiosyncratic hierarchy of the Church of England. CCEd uses a typology to associate its records with place. Contrary to popular opinion, which tends to perceive a well-ordered *cursus honorum* descending from bishop to archdeacon to deacon to incumbent to curate etc, ecclesiastical hierarchies can be very complex. For example, a vicar might be geographically located within a diocese, and yet not report to the bishop responsible for that diocese ('peculiar' jurisdictions). This tends to follow the EPNS's hierarchical structure, although in EPNS it is less clearly defined.

In the geographic sense, location is dealt with in two distinct ways – according to civil geographical areas, such as counties, and according to what might be described as a 'popular understanding' of religious geography, treating a diocese as a single geographic unit. Where known, each parish name has a date associated with it, and for the most part this remains constant throughout the period, although where a name has changed there are multiple records (a similar principle to the attestation value of Chalice names, but a rather different approach in terms of structure).

Sub-parish units are a major issue for CCED, and there are interesting comparisons in the issues this throws up for EPNS. Chapelries are a key example: these existed for certain, and are contained with CCED, but it is not always possible to assign them to a geographical footprint (the notion of footprinting is severely challenged by the kind of pre-Ordnance Survey data CCEd deals with, and this is where close coordination with, and avoidance of replication of, projects such as Vision of Britain is essential: http://www.visionofbritain.org.uk). Almost by definition, each record – i.e. each clergyman - will be been associated with a particular building, which is associated with a particular place, however that place is not necessarily known. This, of course, creates problems in terms of identifying which location is in question. One example comes from East Greenwich, where there is a record of a curate being appointed, but there is no record of where the chapel is or was, and no visible trace of it today or in the known archaeological record. A linked resource whith CHALICE's precision – which is known – should, when linked to a resource such as CCEd, be careful to avoid imposing any 'false accuracy' or 'enforced crispness'[1]. Boundaries are particularly problematic. The phenomenon of 'beating the bounds' around parishes only occurred where there was an economic or social interest in doing this, e.g. when there was an issue of which jurisdiction tithes or other taxes should be paid to. Other factors in determining these boundaries were folk memories, and the memories of the oldest people in the settlement. However, it is the case that, for a significant minority of parishes at least, pre Ordnance Survey there was very little formal/mapped conception of parish boundaries. For this reason, many researchers consider that mapping based on points is more useful that boundaries, with an appropriate system of associating certainty and hierarch among points (CHALICE has broadly adopted this approach by associating sub-parish units with the point value for each parish). This links in with some early discussions the CHALICE project had with the Peiades project (http://pleiades.stoa.org) on mereology of place and the 'nesting' of spatial footprints over time. An exception is where boundaries followed natural features such as rivers. This is an important issue for Chalice to consider in its discussion about capturing and marking up natural features: where and how have these featured in the assignation and georeferencing of placenames, and when?

A similar issue is the development of urban centres in the late 18th and 19th centuries: in most cases these underwent rapid changes; and a system of 'implied boundaries' reflects the situation then more accurately than hard and fast geolocations. Ecclesiastic jurisdictions changed with these changes, and any link with a source as geographically consistent as CHALICE (which is accurate, if not precise) would need to reflect this. It could be argued that the EPNS concept of geographical location is directly related to the source the variant place-name is taken from; an analysis of which would offer interesting and varying definitions of space. Change over time is of course a major issue more generally: within CCEd time is constrained; but it this is not always the case.

Despite these considerations, CCED reflects the formal structured entities of the parish lists. Its search facilities are excellent if you wish to search for information about specific parishes whose name(s) you know, but, for example, it would be very difficult to search for 'parishes in the Thames Valley'; or (another example given in the meeting), to define all parishes within one day's horse riding distance of Jane Austen's home, thus allowing the user to explore the clerical circles she would have come into contact with but without knowing the names of the parishes involved.

At sub-parish level, structured information is lacking. For example, there remains no definitive list of chapelries. CCED has 'created' chapelries, where the records indicate that one is apparent (the East Greenwich example above is an instance of this). In such cases, a link with Chalice and/or Victoria County History (VCH: http://www.victoriacountyhistory.ac.uk) could help establish/verify such conjectured associations.

When one dips below even the imperfect georeferencing of parishes, there are non-geographic, or semi-geographic, exceptions which need to be dealt with: chaplains of naval vessels are one example; as are cathedrals, which sit outside the system, and indeed maintain their own systems and hierarchies. Currently EPNS does not match such structures. In such cases, it is better to pinpoint the things that can be pinpointed, and leave it to the researcher to build their own interpretations around the resulting layers of fuzziness. One simple point layer that could be added to Chalice, for example, is data from Ordnance Survey's describing the locations churches: a set of simple points which would associate the names of a parish with a particular location, not worrying too much about the amorphous parish boundaries, and yet eminently connectible to the structure of a resource such as CCED.

---

[1]   See Kronenfeld, B. J., Triangulation of gradient polygons: a spatial data model for categorical fields, COSIT'07 *Proceedings of the 8th international conference on Spatial information Theory*

In the main, the interests that CCEd share with Chalice are ones of structural association with geography. Currently, Chalice relies on point based grid georeferencing, where that has been provided by county editors for the English Place Name Survey. However, the story is clearly far more complex than this. If placename history is also landscape history, one must also accept that it is also intimately linked to Church history; since the Church exerted so much influence of all areas of life of so much of the period of history in question.

Overlap between CCEd and CHALICE

To explore the overlap between EPNS and CCEd we defined a relatively simple matching method between the two sets of records. For the samples taken for the CHALICE project, we found that there was very considerable overlap between the CCEd coverage and that of EPNS. Of 450 CCEd records for Cheshire, only 7 had no match to EPNS while 17 CCEd records had a partial match to EPNS. The reasons for imperfect matching are usually either because the EPNS extraction has missed some data or because our simple definition of matching is not sufficiently sophisticated. The example of places with 'Peover' in their name illustrates both problems. CCEd contains 'Over Peover', 'Lower Peover' and 'Peover Superior' while from EPNS we have extracted records for

- Lower Peover Hall
- Lower Peover Mill
- Nether Peover
- Nether Peover Chapelry
- Over Peover Bridge
- Over Peover Chapelry
- Peover
- Peover Cottage
- Peover Eye Bridge
- Peover Hall
- Peover Heath
- Peover Inferior
- Peover Mill
- Peover Superior

A proper match is achieved for 'Peover Superior' but not for the other two CCEd records. For 'Over Peover' the match should be to 'Over Peover Chapelry' but 'Chapelry' wasn't stripped out of the EPNS record and the matching wasn't configured to ignore the 'Chapelry'. For 'Lower Peover' there was no match because although EPNS gives this as an alternative name for 'Peover Inferior' our EPNS processing did not preserve it. The entries for all the 'Peover' places in EPNS are very complex: they are spread across three separate hundreds with, for instance, 'Over Peover Chapelry' occurring part in Macclesfield Hundred and part in Bucklow Hundred. In cases such as this, it is not even possible for a human annotator to be able to tell which match is correct. 5 mismatches were attributable to typographic/data inputting errors in CCEd (Longendale (x3), Sotckport, Chorlton, Medlock).

141 CCEd records contain more than one place name which match EPNS; and 121 distinct EPNS places matched to CCEd (there is a lot of repetition in CCEd).

A closer examination of the matches across CCEd place types allows certain observations to be drawn. There are two CCEd records for 'Hospital', although both of these occur in Chester. All the CCEd instances of *Moiety* appear in EPNS, but only two have variants listed (it should be noted at this point that a deep analysis of how many historical variants each

Placename matches between EPNS Cheshire volume and the CCEd

parish name has is not possible at this stage, since the EPNS records for attribution are linked to township rather than parish. Attribution is therefore via the township contained within the parish). The general pattern within CCEd is that parish names are by far the most common location element, and this is the case in Cheshire, where around half (220) are parish church names. 22% of these appear not to have historical variants given in EPNS, although there is a total overlap of 98%. Parish schools tell a slightly different story: with a total Cheshire population of 53 in CCEd, 3 are not listed in EPNS, and 16 have no variants. The CCEd's article on the treatment of location makes it clear that there are many instances where entries for schools are subordinate to the parishes in which they resided, suggesting that they are likely to exist at a more granular level, i.e. at one which is less likely to be picked up by the EPNS's records at this level (which would typically be more concerned with units such as field and street names). This may also help explain the 24% of 'School' records that do not have an EPNS equivalent. It should be noted in this context that outside urban areas, EPNS tends to be less concerned with buildings, except where the building has an especial interest for, or relevance to, placenames. One important class of such exceptions is manors (see VCH use case).

In our exploratory matching we have matched names but have not constrained the match by type, raising the question of whether location entities identified in EPNS cut across the CCEd database's data types. For example, we successfully match the name 'Burwardsley' across the two collections but the 'Burwardsley' in CCEd is listed as 'Parish (church)' while

the 'Burwardsley' in EPNS is a township in the parish of Bunbury and EPNS does not have an entry for 'Burwardsley' as a parish. In order to benefit from linking the two collections it will be necessary to investigate the origins of this kind of misalignment and to implement a much more sophisticated matching process.

Conclusion

The CCEd is 'strategic content' in that its value is enhanced if it is linked to other things: its original purpose was to support a broader gamut of historical research. The technical challenges of using CHALICE/EPNS as a cross-search facility are relatively few. CCEd is contained within an online MySQL database, and the structure of the CHALICE resource should ensure that, from a purely technical point of view, linking is quite possible. The challenges arise in the content and the structure. CCEd uses a bespoke feature type system, without which the placename information in it is not useful. The CHALICE resource should be customisable to deal with such structures, otherwise the utility of linking placenames is lost. For CHALICE, CCEd raises three issues:

1. What visual interface/structure would work best to display complex layers of information? This would have to include timelines, information structures and geographic mapping, but searching on one of these facets must not be contingent on the outcomes of searches on others.

2. How can the existing (limited) georeferencing of EPNS be enhanced by linking to it, and how could this linking enhance CCEd? CCEd researchers told us that even having georeferences attached to CCEd records would be very valuable. However, CCEd deals with location in a highly bespoke way: linking with a consistent and unique resource such as ENPS risks exposing tensions in both resources' data structures.

3. EPNS is a unique resource, and while this fact underpins its utility as 'strategic content' it also restricts the level at which linking is possible. The sources underlying the EPNS material is gathered by EPNS scholarship: it must be remembered therefore that CHALICE links EPNS scholarship. It is not a quantitative gazetteer such as the Ordnance Survey and, as this study (among many others) has shown, such kinds of gazetteer are conceptually unsuited to dealing with fuzzy historical data.

It is clear, however, that the association of (EPNS, placename, church, CCEd, VCH) could allow historians to construct the kind of queries they have not been able to construct before.