Research Project Plan

Project Title

- *The Places of Scotland: Assessing the Value of a Definitive Place-name Gazetteer to Scotland*

Description

Place-name gazetteers are a rising priority and key to connecting swathes of information held in a range of organisations (Gittings, 2009). Place-name gazetteers are important in linking cultural and other data which doesn’t have an explicit geo reference (Hill, 2006). Unlike other countries e.g. USA, Estonia and Ireland, the UK has yet to implement a definitive place-name gazetteer. This project overall aim is to assess the value of a definitive place-name gazetteer for Scotland. The importance for the study of current place-name gazetteer’s in Scotland has been underlined by the potential release of Ordnance Survey 1:50K gazetteer, which has been regarded as the standard until now and the INSPIRE directive which is being implemented in the UK from this year (Illert, 2009). From this context, this project’s first objective is to identify and survey the various gazetteers available in Scotland. The value of a consistent set of data which can provide a central look up and form a hub for connecting a range of different services has been demonstrated through CYGnus (Carter and Gittings, 2009). This projects second objective is to demonstrate through CYGnus the value of a consistent set of place-names for Scotland through connecting the range of services through place-name.
Background.

Importance of a definitive place-name Gazetteer

The growing use of GIS technology in conjunction with the World Wide Web since the 1980’s has seen a proliferation of both the creation and dissemination of geographical information (Gittings, 2009). Geographical Information is vital to governmental and organisational decision making as many decisions such as those concerning industrial accidents, environmental emergencies and national security are based on geography (Hill, 2006). It is estimated that 70% of all information is Geographical Information (GI). This information has traditionally been stored in large geographical directories known as gazetteers. As GIS has grown, so too has the awareness that much of this information in gazetteers is being duplicated, is of variable quality and is trapped in their respective ‘data silos’ (Masser, 1998). There are two possible solutions to these problems (both with their advantages and disadvantages), through either having a definitive spatial reference gazetteer or having a comprehensive place-name gazetteer. A definitive place-name gazetteer allows features and places to be referenced by place-name, however this approach dictates that all place-names must match and be unique. Such a standardised approach to national GI has been implemented in the USA through the enforcement of a National Spatial Data Infrastructure (UNSD, 2010). An estimated 50 other countries have followed suit including the National Place-names Register of Estonia (Kivisalu et al., 2009) and The Place-names Database of Ireland (Mac Giolla, 2009).

In summary, place-names are crucial to storing and linking geographical information as they allow non addressable features to be defined (Hill, 2006). A definitive gazetteer has thus been seen as key to ensuring unique and accurate geo referencing and allowing the connection of ‘information silos’ across organisations (Gittings, 2009). Currently, a definitive place-name gazetteer exists in many countries including the USA, Estonia and Ireland but an equivalent standard is not currently available in the UK. This project aims to assess the value of such a list to Scotland. The importance of this study has been underlined by the potential release of Ordnance Survey 1:50K gazetteer, which has been regarded as the standard until now and the INSPIRE directive which is being implemented in the UK from this year (Illert, 2009). From this context the current policy agenda on Scotland’s place-name gazetteers will be discussed below.
Literature Review

The Problem of Place-names in Scotland

As has been discussed above place-names, or toponyms - a comprehensive term for geographical names and extraterrestrial names, (Kadmon, 2000) – are the most common way we identify geographic features (Hill, 2006:91). When place-names are unambiguous they are the simplest way of georeferencing features. The importance of place-names was highlighted by the Scottish GI Strategy, which identified three forms of GI: physical and map based data (The Spaces of Scotland); address-based data (The Faces of Scotland) and place-name data (The Places of Scotland). Unique place-names are able to connect the rich background of geographical and historical information that exists in Scotland; including that of libraries, images and catalogues across government, education and heritage sectors (Scottish Executive, 2005). However Scotland and the United Kingdom at large have failed to link these resources as a consequence of the lack of a standardised place-name gazetteer (Gittings, 2009). This is despite the lack of a standardised approach to place-names being recognised as a real problem in the ‘One Scotland – One Geography’ publication (Scottish Executive, 2005). Too many people regard the name on Ordnance Survey maps as definitive; this is not the case. Indeed, currently we don’t even have a list of the various gazetteers which are available, never mind a standardised gazetteer. This problem has recently been illustrated through a recent survey of gazetteers drawn together for Tobhas and Dualchais in Scotland. The study compared place-names in the area to the Geonames and Ordnance Survey 50K gazetteer’s respectively, finding wide variations between different gazetteers (Unlocking Scottish Place-names, 2009). From this context, the first objective of the project will be to identify and survey the various gazetteers that contain place-name data of Scotland against one another.

Whereas unambiguous place-names can have huge benefits to governmental and organisational decision making ambiguous place-names can have the converse impact. The problem with place-names is that they are often not unique, each place may have many names, and many different places may have the same name (Caldwell and Shine, 2008). The term used to define an individual place-name that refers to many places is a homonym (Randall, 2001). There are numerous examples of toponymic homonyms in Scotland one of which is that of the three settlements called
‘Newbigging’ found nearby within Angus (Gittings, 2009). The term used to define multiple place-names that refer to the one place is allonyms (Kadmon, 2000). Toponymic allonyms are a particular issue in Scotland with regards to Gaelic place-names. This stems from the Ordnance Survey ‘Gaelic Name Policy’ (Ordnance Survey, 2003), which aimed to standardise and maintain Gaelic names across Scotland. The outcome of this policy has been a mass translation of names wherein there has been a lack of consistency with names being changed due to linguistic criteria rather than being based on geographic accuracy (Munro and Gittings, 2006). Alongside the OS Gaelic gazetteer the Scottish Parliament has its own gazetteer of English and Gaelic toponyms for every Scottish place-name (Mac an Tailleir, 2003) and recently there has been the creation of a Gaelic Place-names of Scotland gazetteer (Bord na Gaidhlig, 2007). Both the toponymic homonym and toponymic allonym problems endemic in Scottish placenames will have to be resolved in this project, when surveying the various gazetteers.

The Value of a Definitive Place-name Gazetteer for Scotland

It has been argued that the problems outlined above will only be fully resolved by having a Scottish Place-names Authority responsible for a definitive place-name gazetteer (Gittings, 2009). As already discussed above the benefit of such a database would not only solve toponymic problems but allow more efficient and effective linkages between different kinds of services. The need to share data was raised previously by the Chorley Report (1988), NGDF (Masser, 1998) and the Scottish Strategy (Scottish Executive, 2005). The value of such a consistent set of data which can provide a central look up and form a hub for connecting a range of different services has been demonstrated through CYGnus (Carter and Gittings, 2009). This projects second objective is to demonstrate the value of a definitive list of place-names for Scotland.

As previously mentioned in the background section, there are numerous countries where definitive place-name gazetteers have been implemented. The most famous example is that of the US Board on Geographic Names was established in its present form by public law in 1947 with the aim of maintaining uniform geographic name usage throughout the federal government (UNSD, 2010). The definitive place-name gazetteer contains domestic, foreign, Antarctica and undersea place-names. The creation of such a body stemmed from inconsistencies and contradictions among many names.
with definitive national place-name gazetteers include those in Ireland and Estonia. In Estonia the 2008 appointment of the Land Board of Estonia (Maa-amet) to be responsible for maintaining the National Place-names Register has seen the place-names register integrated into other systems dealing with address data, cadastral information, etc (see figure 1) (Kivisalu et al., 2009). The Placenames Database of Ireland, was officially launched in 2008 with its primary purpose to provide standardised Irish language versions of Ireland’s geographical names on line to customers (Mac Giolla, 2009). The site has proved hugely popular with over 2.5 million visitors to its website over a 6 month period. From looking at other case studies demonstrating the value of a definitive place-name gazetteer, this project will also consider links to other pre-existing services in Scotland and the potential impact of these services for Scotland.

Figure 1: Shows how the Place-names Register links to other Services. Source: Kivisalu et al., 2009: 3
Aim and Objectives

Aim

• Assess the value of a definitive place-name gazetteer to Scotland

Objectives

• Collect and survey the various place-name gazetteers available
  o Collect various gazetteers
  o Survey gazetteers based on their completeness, spatial and typological accuracy
  o Resolve any issues surrounding homonyms or allonyms

• Demonstrate the value of a definitive place-name gazetteers
  o Assess relative coverage
  o Examine links to different gazetteers
  o Examine the impact of a definitive place-name gazetteer on Scotland
Methods

The methodology used to achieve the aim and objectives is outlined below:

Collecting and surveying the various place-name gazetteers available

Collecting the gazetteers

This will involve communicating and networking with those already known to be involved in Scotland’s place-name gazetteers system as well as conducting research in the National Library of Scotland and seeing if there are any other relevant place-name directories available. Those place-name gazetteers selected for further research will be limited to those of national coverage.

Surveying the gazetteers

A systematic method will be adopted in surveying the various place-name gazetteers of which one of two options will be implemented. The first of these methods would be to write a Perl script that analyses every place-name in each gazetteer and compares it against the others. A second potential method would adopt the same approach to that set out by the Tobar and Dualchais place-name data analysis (Unlocking Scottish Place-names, 2009). This method involved the place-names from all records being extracted onto a spreadsheet from which the ‘Unlock Places Web Service’ was used to search for each place-name, this process was carried out for the various gazetteers. From either of these methods it would be possible to assess each gazetteer in terms of their completeness and spatial and typological accuracy against one another as well as identify resulting problems of homonyms and allonyms (see below).

Resolving Homonyms

Homonyms are multiple places that have the same name (Kadmon, 2000) e.g. NewBigging in Angus (see literature review). Along with allonyms these are one of the fundamental problems in using place-names to georeference toponyms. The approach adopted to deal with these homonyms will
be that of ‘disambiguating’ - the resolutions of a name to a specific feature and location is (Smith and Crane, 2001) - whereby additional geographical identifiers (e.g. postcode) will be added in order that the place-name is unique.

**Resolving Allonyms (The ‘Gaelic Problem’)**

Allonyms are the converse of homonyms in that they entail multiple name to the same place (Kadmon, 2000), a problem endemic of Scotland’s Gaelic place-names i.e. the ‘Gaelic Problem’ (see literature review). This project will initially take the approach of storing them all as a list of variants. It is hoped that these variants will later be researched further through consulting the local community and Gaelic historians in order to get an agreement on main Gaelic name. This Gaelic name and the main English name will then be stored as toponyms from which the various allonyms will be attached. Another issue will be that of accents whereby as far as the computer is concerned a letter with an accent is completely different from a letter without. There are various ways by which this can be handled; because of this a database system will have to be used that allows the search of place-names with or without accents e.g. Oracle.

**Demonstrating the value of a definitive place-name gazetteer**

**Coverage**

The value of a definitive place-name gazetteer will be demonstrated through comparing the number of places identified from a national archive document using Scotland’s places gazetteer to the number of places identified from the created ‘definitive’ place-name gazetteer (through moulding all the various gazetteers together). This will be done using the CYGnus system (Carter and Gittings, 2009)
Links to Historical allonyms

Historical allonyms are historical name variants (Kadmon, 2000). These will be used to demonstrate one of the advantages of having a definitive place-name gazetteer. Although there are various records of historical allonyms available this study would initially use that of the Ordnance Survey Name Books which can be obtained from the National Archives. In order to join all these allonyms to the current toponym CYGnus would have to be adapted.

Links to National Address Gazetteer

Another possible way in which the value of a definitive place-name gazetteer could be demonstrated is through showing the value in linking it with the Scotland’s definitive National Address Gazetteer (see literature review). This would initially involve communicating researching how other countries have linked their definitive place-name gazetteers to other systems (e.g. Estonia, Ireland and USA) before communicating with Iain Mackay.

Business Plan

The final stage in this project would be to produce a business place which would list the potential capabilities that a definitive place-name gazetteer could bring to Scotland. This would include different systems links and explicitly outlining the ways in which such a service would impact on public service efficiency and effectiveness. As with the previous stage above this would involve in depth research of the impact of an equivalent service in different countries (e.g. Estonia, Ireland) as well as networking people in government and business who have direct experience of such a service.
Ethics

The methods by which this project seeks to approach the aims and objectives outlined above have several ethical implications. The main issue is that of copyright whereby it is anticipated that the majority of gazetteers (including the National Address Gazetteer) will not allow their directories to be distributed. Therefore it is imperative that the place-name gazetteer created for the purpose of this study be used for research purposes only. There is not anticipated to be any costing or risks associated with this project.

Project Timetable

<table>
<thead>
<tr>
<th>Task</th>
<th>Mar</th>
<th>Apr</th>
<th>May</th>
<th>Jun</th>
<th>Jul</th>
<th>Aug</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gather Data</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Write Literature Review and Introduction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conduct analyses</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Write up methodology</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Final Write up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Review and Edits</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Gantt Chart
Bibliography


• Smith, D and Crane, G (2001) ‘Disambiguating Geographic Names in a Historical Digital Library’ In *Research and Advanced Technology for Digital Libraries*: 127-136. Heidelberg: Springer Berlin. Available at: [http://www.springerlink.com/content/h7em0v5803e6h7yb](http://www.springerlink.com/content/h7em0v5803e6h7yb)

• Unlocking Scottish Place-names (2009)