Gazetteer Services: A Scottish Perspective

Bruce M. Gittings
Edinburgh Earth Observatory
University of Edinburgh
bruce@geo.ed.ac.uk
Summary

• Scottish Context
• Importance of Gazetteers
• Definitive Address Gazetteer for Scotland
• Importance of Descriptive Geographical Information
• The Gazetteer for Scotland
• Historical Descriptive Gazetteers
• Problems with Place Names
Some Definitions

**Short-form gazetteers:**
- often used in conjunction with computer mapping and GIS systems, may simply contain a list of place-names together with their locations in latitude and longitude or other spatial referencing systems (eg. British National Grid reference). Short-form gazetteers appear as a place-name index in the rear of major published atlases.

**Long-form (or Descriptive) gazetteers:**
- may include lengthy textual descriptions of the places they contain, including explanation of industries, government, geography, together with historical perspectives, maps and / or photographs.

**Thematic gazetteers:**
- list places or geographical features by theme; for example fishing ports, nuclear power stations or historic buildings. Their common element is that the geographical location is an important attribute of the features listed.

**Address Gazetteers:**
- geo-referenced address lists, often used by local authorities for purposes such as local taxation
Scottish Context

• Some big problems:
  – Regularly changing administrative geographies
  – Multiple (nearby) places of the same name
  – Gaelic names (no agreed version)
  – No definitive place-name gazetteer

• But some advantages:
  – Lots of mapping versions (from Pont in 16th C.)
  – Detailed country-wide mapping from 1860s
  – Scottish place-names survey (1980s) – but not digital

• Some real demands for gazetteer services
  – Who is the custodian of place names?
  – Place is the connection between heritage projects (although they don't realise that!)
  – Efficient government
Gazetteers for Contextualising Places

- All manner of google, yahoo, microsoft are pushing forward with local search, integrated with mapping
- Imagery (vertical, oblique, street-frontage) and geographically-referenced texts become key to these services
- Gazetteers are vital to provide the linkages
- Many independent resources built with public money, which all include significant GI component (although they may not realise that either!)
- Could create a world-class resource
- In Scotland, these might include:
  - Local and Central Government
  - RCAHMS Canmore (archaeology)
  - National Archives
  - National Library of Scotland (maps and archives)
  - Scottish Cultural Resources Archive Network (SCRAM)
  - Historic Scotland
  - Great Britain Historic GIS
  - Gazetteer for Scotland
PART 1: National Address Gazetteer for Scotland
National Address Gazetteeer for Scotland

- Definitive National Address for Scotland Gazetteer (DNA-Scotland) will be aggregated from local corporate address gazetteers (CAGs) created and maintained by Local Authorities (Councils) and Property Tax Assessors
- May or may not be a postal delivery address (ie. in PAF)
- Effectively it will provide a national view of local data
- Aim is for it to become the trusted address source of choice within the Scottish Public Sector
- Seen as a standard means of connecting other databases together
- Based on the property, to which individual people can be connected
- Implementation accelerated due to the sudden need for Senior Citizen entitlement cards
- Is being resourced – will be maintained
Differing Stakeholder Views of a Property

- Property to be valued, classified and attributed [Assessors]
- Development to be controlled [LA Planning]
- Legal Title to be registered [RoS]
- Status of buildings and occupants [GROS]
- Property to be surveyed and attributed [OS]
- Property to deliver mail to [Royal Mail]
- Home address of 2 pupils [LA Education]
- Fire to be put out [Fire Brigade]
- Property to be protected [Police]
- Refuse to be collected [LA Cleansing]
- Council Tax to be collected [LA]
- Patient to be Collected [NHS Ambulance]
From BS 7666 to DNA Gazetteer

Thanks to Iain McKay (Programme Manager) from whom these slides have been adapted
Overview of National Infrastructure

32 Local Authorities

- CAG
- CAG
- CAG
- CAG

Application Server -> DNA Gazetteer

Application Server -> Property Database (Subset DNAG)

Application Server -> Citizen Database (Subset GRoS)

Person to Address Index

National Infrastructure
Gazetteer role in Shared Services Agenda

Shared Services (possible examples being considered)

- Revenue Collection
- e-Planning
- Land Terrier
- Asset Management
- Community Planning

Other Council Systems

- Individual Councils x 32
- Royal Mail
- Assessor
- OS
- ROS
Implications of DNA-Scotland

- Deals only with street address and post-town
- Does not deal with non-addressableable features (natural features etc. etc.)
- Also does not deal with urban or rural districts, which are so important in defining communities

- Not just addresses, people are very much linked in
- Very unlikely to be available to the private sector, so only a definitive address for some
- Even so, there are some real civil liberties implications
- The *entitlement card* is effectively a national identity card by the back door
- For the moment, saved only by the incompetence of government
PART 2: Textual GI and Descriptive Gazetteers
A Case for Textual GI?

- Are points, lines and polygons with attributes enough?
- Do these fully represent the world around us?
- How do we navigate from one place to another?
- Maps are often not terribly well understood by non-specialists
- Use features which are not currently part of maps or GI databases!
  - Landmarks
  - Descriptions or Textures of Buildings
  - Colour of Features
  - May be seasonally dependent
- Maybe we need richer data
- Building frontage data is being captured by videography (eg. Amazon A9, Microsoft’s intentions for local.live)
- Textual GI is one form of this richer data
So Descriptive Gazetteers are Important

- Descriptive GI is poor little brother of maps
- But this textual GI makes an important contribution
- This is exactly what google etc. are trying to search
- A traditional form of GI, which has recently been forgotten:
  
  "Next to a good dictionary, the most generally useful book is a good gazetteer"
  -- W.G. Blackie (1855)

- Actually, there was a solid tradition of gazetteer production in Scotland during the 19th C.
- Complements the map
- Potentially rather more comprehensible by the public
- The geographical equivalent of the geological map and monograph
- Helps to solve the navigation problem
The Gazetteer for Scotland

• Aims to be a definitive and detailed resource
• Began work in 1995; at least another 10 years to go!
• Originally envisaged as a book
• Statistics:
  – Around 13,500 entries
  – 1.2 million words of contemporary text
  – 2.25 million words of historical text
  – c.7500 photographs
  – Maps with every entry as mashups using:
    • OS get-a-map
    • Google Maps
    • Oldmaps.co.uk
  – Also direction-finding via multimap.com
  – Approaching 1 million hits / week
• Significant interest for tourism and genealogy
• There are places in the GfS which are not on OS maps
Welcome to the Gazetteer for Scotland

The Gazetteer for Scotland is a vast geographical encyclopaedia, featuring details of towns, villages, bens and glens from the Scottish Borders to the Northern Isles. The first comprehensive gazetteer produced for Scotland since 1895, it includes tourist attractions, industries and historical sites, together with histories of family names, biographies of famous people and descriptions of historical events associated with Scotland. The network of connections between all of these entries make this gazetteer unique. Visitors will also find we have included the text of Groome's 19th Century Ordnance Gazetteer of Scotland in our entries to provide a historical perspective.

With at least 13442 detailed entries, the Gazetteer is already the largest Scottish resource available on the web, yet it is growing constantly, with some 931 entries added or updated in the last week alone. Please look at 'Gaz Stats' for more information.

Browse the Gazetteer for Scotland using either:

Maps and Places OR History Time-Line OR Search Pages

Alternatively, look at our:

Featured Town: Milnathort
Featured Geographical Feature: Loch Portree

You can also use the interactive facilities in our free Member's Area to make maps and add your own comments to individual entries.

NEWS FLASH!
Just published: Scotland: An Encyclopedia of Landscape and Place, the most comprehensive reference on the geography of Scotland in more than 100 years, drawing on material from the Gazetteer from Scotland, but also including 32 pages of full-colour mapping and much more. Buy it from our bookshop in time for Christmas.

New for December 2006, more than 500 additional photographs for the West and Southwest of Scotland.
Details

• Main user interface designed for the general public
• Implemented as an Oracle database
• Makes full use of relational joins to link entities:
  – Settlements  -  Geofeatures
  – People  -  Families / Clans
  – Events  -  Associated biblio references
  – Media (photographs, sounds, video)
• Database is not queried in real-time; rather pages are pregenerated at 4-6 week intervals (except new)
• Compromise between performance and disk-space
• Actually a spatio-temporal database:
  – Dates are included where poss.
  – Linkages allow geographical inferences to be made (eg. people associated with various georeference places)
• Could easily form hub of a short-form gazetteer service
Temporality appears to public as timelines

The Gazetteer for Scotland History Time-Line

The Scottish History Time-Line provides an alternative route to access the Gazetteer for Scotland. The time-line incorporates thousands of Scottish historical events including the births and deaths of famous Scots, archaeological sites and historic buildings in Scotland, together with the dates of battles and other cultural events. Each event is colour coded, with a key given below. The time-line is fully linked into the vast store of information which forms the Gazetteer for Scotland database.

Choose a time-period from the list below:

- Until the 13th Century
- 14th and 15th Centuries
- 16th Century
- 17th Century
- 18th Century
- Early 19th Century; 1800-49
- Late 19th Century; 1850-99
- Early 20th Century; 1900-49
- Late 20th Century; 1950-99
- 21st Century

Note that, within a single year, events are not necessarily presented in chronological order. The time-lines only give direct access to a limited number of Gazetteer for Scotland entries, because only some can ever have dates associated with them.

KEY:  
- Airport, Bridges, Canals and Streets
- Archaeological Sites and Historic Monuments
- Births and Deaths
- Events
- Exhibition, Centre, Museum or Gallery
- Historic Buildings
- Industrial or Commercial Sites
- Institutions or Government Buildings
- Gardens or Parks
- Sports Venues
- Other Types
Port Logan

Use the tabs on the right of this page to see other parts of this entry
PART 3: Historical Descriptive Gazetteers
Issues of History

• Richness of textual description
• Text maintains the subtleties of history and details ephemeral knowledge much more easily and often more effectively than a map
• Unfortunately recently the trend has been towards a tourist guide rather than a systematic description of places which traditionally formed a gazetteer
• But there are problems:
  – Neither have we good tools to interrogate descriptive GI or make inferences from it
  – Data mining techniques represent only the beginnings of a solution
  – Nor good tools to generate it
  – We don't even do geo-parsing particularly well
  – But then how do we differentiate the three Newbiggins within a few miles of each other in Angus??
Muirhouse, a mansion in Cramond parish, Edinburghshire, on the shore of the Firth of Forth, 1½ mile NNE of Davidson's Mains and 4½ miles WNW of Edinburgh. Only two round towers remain of a previous mansion (circa 1670); and the present house is a picturesque Tudor edifice of about 1830, with a square battlemented tower and beautiful well-wooded grounds. Its drawing-room is adorned with several frescoes by Zephaniah Bell. Purchased by his ancestor in 1776, the estate belongs now to Thomas Davidson, Esq., F.R.S., F.I.S., F.G.S. (b. 1817; suc. 1865), who, residing in Edinburghshire, valued at £1216 per ann. in 1857. See John Small’s Castles of the Lothians (Edinb. 1883).

Muirhouse
City of Edinburgh

Lying to the north of Ferry Road in Edinburgh and to the west of Pilton, Muirhouse is a public housing estate begun in 1953. It takes its name from the Tudor-Gothic mansion built on Marine Drive in 1832 for the Davidson family, who were wealthy merchants trading in Rotterdam. The estate was built in mixed housing styles, include regular white boxes, with low-pitched roofs, and large tower blocks, mostly constructed in the 1960s. The area has suffered from significant social problems. From the mid-1980s Muirhouse, together with the neighbouring West Pilton have been undergoing significant renewal, with many of the tired and vandalised council-owned blocks being replaced by private housing. Muirhouse suffers unemployment rates of more than three times the average for Edinburgh, with a large proportion of long-term unemployment.
Roslin (Brit. ross, 'a point,' and lynn, 'a waterfall,' the name often and perhaps more correctly spelled Rosslyn), a quoad sacra parish containing a village, chapel, and castle of the same name, in the civil parish of Lasswade, in the county of Edinburgh. The village, which stands on high ground near the NW bank of the river North Esk, has in its neighbourhood three railway stations on different sections of the North British railway system, and each of them distant about 10 miles from Edinburgh. The nearest, Roslin, on the Edinburgh and Glencorse branch, is close to the village. Rosslyn Castle, on the Edinburgh and Penicuik branch, is distant about 1¼ mile, and Rosslynlee, on the Edinburgh and Peebles line, about 1¾ mile. By road the village is about 6½ miles S of Edinburgh; and from Polton station, 7 miles SSE of Edinburgh, a public footpath winds through the beautifully wooded glen of the North Esk to the village, the distance being about 2½ miles. About 1440, under the fostering protection of William St Clair, Prince of Orkney, Duke of Oldenburg, and having a string of other titles that it would weary even a Spaniard to repeat, the place is said to have stood third in Scotland for importance...

[ Groome, Ordnance Gazetteer of Scotland, 1885 ]
Case Study: Learning from Failure

- Wish to provide historical depth to the *Gazetteer for Scotland*
- Attempted to semi-automatically link Groome’s *Ordnance Gazetteer*
- Limited success (initially as low as 15-20% match)
- Lots of manual checking required to confirm
- Why?
  - settlements, country houses etc. disappeared
  - historical places not in contemporary gazetteer
  - relative importance of places (e.g. estates, laird's houses)
  - new places (airports, industrial facilities settlements etc.)
  - spelling change / inconsistency
  - old Gaelic renderings
  - inadequate description to distinguish similarly named entries (e.g. surprising number of Black Lochs)
Geographical-guided Soundex Matching

```sql
select distinct gr.seqno, gr.name, gr.type1, 
gr.county1, gr.county2, gr.county3, gf.seqno, 
gf.name, ca.name caname 
from groometemp gr, geofeatures gf, grcouca_rel, 
geoca_rel, council_areas ca 
where (soundex(gr.name) = soundex(gf.name) 
and gr.county1 = grcouca_rel.couname) 
and grcouca_rel.caseqno = geoca_rel.caseqno 
and geoca_rel.gseqno = gf.seqno 
and grcouca_rel.caseqno = ca.seqno

Gets a bit more complicated than this...
```
Generating Text from Databases

• As part of the same process of historical depth to the GoS, wished to create contemporary parish entries to which Groome text could be linked
• Contemporary entries needed to relate parishes to modern geographies
• Would be enormously tedious to manually research and write 1000+ of these entries
• Take digitised parishes for 1951 (out-of-copyright)
• Extract topology to give relationships one to another
• Encode the database with as many facts as possible
• Relate to historical *counties* and current *council areas*
• Try to make results interesting to read (non-formulaic)
• Not surprisingly this is rather difficult to achieve!
• But most people look at relatively few entries
Monikie

Monikie is a parish, which lies in Angus Council Area, some 5 miles (8 km) northwest of Carnoustie in Angus and 8 miles (12 km) northeast of Dundee in Dundee City.

It lies in the old county of Forfarshire, which disappeared following the Local Government (Scotland) Act 1974.

The parish has an area of 36.5 sq. km (22.6 sq. miles). Monikie has 5 neighbouring parishes; namely Barry, Carmyllie, Inverarity, Monifieth, Murree and Parbribe.

Forfar

Forfar is a parish, which lies in Angus Council Area, some 6 miles (9 km) southeast of Kirriemuir and 10 miles (17 km) southwest of Brechin in Angus.

It lies in the old county of Forfarshire, which disappeared following the Local Government (Scotland) Act 1974.

Longforgan

Longforgan is a coastal parish, which lies in Perthshire and Kinross Council Area, some 7 miles (11 km) west of Dundee in Dundee City and 8 miles (13 km) west of Newport-on-Tay in Fife.

It lies in the old county of Perthshire, which disappeared following the Local Government (Scotland) Act 1974.

The parish has an area of 34.5 sq. km (21.4 sq. miles). Longforgan has 7 neighbouring parishes; namely Abernyte, Carlill, Dundee, Fowlis Easter, Inchture, Kettins and Liff and Benvie. The boundaries of this parish were modified among significant changes recommended by the Boundary Commissioners after the Local Government (Scotland) Act 1889.

Arbroath and St Vigeans

Arbroath and St Vigeans is a coastal parish, which lies in Angus Council Area, some 7 miles (12 km) northeast of Carnoustie and 10 miles (16 km) south of Brechin in Angus, and includes a small offshore island.

It lies in the old county of Forfarshire, which disappeared following the Local Government (Scotland) Act 1974.

The parish has an area of 51.6 sq. km (32 sq. miles) and includes the settlement of Arbroath. Arbroath and St Vigeans has Arbrothot, Carmyllie and Inverkeilor as its neighbouring parishes. The boundaries of this parish were modified among significant changes recommended by the Boundary Commissioners after the Local Government (Scotland) Act 1889.
PART 4: Problems with Place-names
Defining Places

• Textual GI helps us define places
  – These three Newbiggings in Angus!

• Government, industry, communities need to be able to locate and describe places
  – Angus Fire Brigade were recently reported as saying that determining which place they were heading for was a significant issue! People don't know their coordinates

• The first step in community-building and social inclusion

• In Scotland have chaos in our place-names!

• But, did you know that the UK is one of relatively few countries in the world which doesn't have a place-names authority?
Some Examples

• Consider the OS 1:50,000 sheet vs the Collins Bartholomew 1:250,000 road atlas

• Four different renderings of 'Tolsta' on just two maps!

• Typically also problems between:
  – LA directions signs
  – Village name signs
  – Local Shop / PO
Consider Balivanich (Benbecula)

- **Balivanich** (42000) or **Ballivanich** (419) or **Balivanish** (34)
- In Gaelic: **Baile Mhanaich** (1780), **Baile a' Mhanaich** (1100), **Bhaile a' Mhanaich** (33), **Bailebhainich** (17) or **Bhaile a Mhanaich** (11)
- Four different renderings on *Comhairle nan Eilean Siar* web site
- Ordnance Survey Gazetteer, 1:50K and 1:25K maps, Collins and AA road atlases are all different, but also in flux!
- Plus historical versions
- OS differences possibly due to progressive implementation of their Gaelic names policy
Some More Confusion

Where was the folk-singer Calum Kennedy (1928 - 2006) born?

The *Guardian* and *Herald* obituaries gave "Orasay", the *Times* and the *Independent* gave "Orosay". With a little research, it turns out to be "Orinsay", which the current OS map portrays only in Gaelic as "Orasaigh", although the first edition (widely used for historical research) gives "Orosaigh".

How can we effectively GEOPARSE and GEOREFERENCE TEXTS, and DISAMBIGUATE PLACENAMES, when we have this sort of PROBLEM?
The Need for a Place Names Authority

- Resolution IV of the *UN Conference of Geographic Names* (1967) directed that standardisation of place names be achieved by establishing national place name authorities
- The USA has the US Board on Geographic Names
- We have the *Permanent Committee on Geographical Names for British Official Use*:
  
  "Principal function is to advise British government departments and agencies on policies and procedures for the proper writing, in official publications, of geographical names for places and features outside the United Kingdom, excluding those of the Antarctic"
- The Scottish Place Names Society have an academic interest in the historical aspects of toponomy
- Scotland desperately needs an Authority:
  - Advising local and central government
  - Tourism
  - Integrated transport
  - Citizen-focussed services
  - Informing Utilities
  - Ensuring the consistency of names
  - Promoting Scotland
  - Understanding and reinforcing our heritage
  - Joining up GI
  - Advising property developers (new streets etc.)
- Those involved in managing cultural resources are desperate for a means of joining up their resources
A GI Strategy for Scotland

• Connecting the *Spaces, Faces and Places* of Scotland through geography and geographic information
• Specifically mentions place-name data
• Includes the *non-addressables*
• States:
  - place-names are basic elements of library, museum and other archive catalogues (name-based archives)
  - The first element in searching for information about places with computer search tools
  - Such techniques have the potential to unlock Scotland's "hidden wealth" of archived information about its places and people
  - Currently limited by the lack of a standardised approach to place names in Scotland
  - Raised as a significant issue during consultation on the draft Strategy, particularly with regard to Gaelic place-names
Conclusions

• Descriptive gazetteers are enormously valuable
• Can include the history which gives authority
• Gazetteer services as vital because:
  – they can act as the glue to join up other services
  – they can provide definitive names, statistics and possibly even definitive descriptions
• Scotland (and I'm sure the rest of the UK) needs a Place-name Authority
• The Vision for Scotland:
  – A free-to-access definitive gazetteer service
  – Because we can hang all sorts of things from this
• I was going to have rant about Wikipedia, but I've run out of time...
The Wikipedia Effect

• We all tell our students not to use Wikipedia, because its information is potentially unreliable

• **However much more importantly:**
  • Wikipedia recycles information from one place on the web to another, with no added value, but detracting from the original source
  • Its playing hard-and-loose with copyright laws are a disincentive to professional reference editors putting material on the web
  • This discourages scholarship,
  • Its clones repeat the same bias / errors several times over; an intelligent geo-aggregator might regard facts as correct because they are repeated
  • Issues of political bias, peculiar agendas, or just plain wrong!
  • The Wikipedia model is a disastrous one for gazetteer services, but Wikipedia is already a geo-referenced authoritative gaz service whether we like it or not !!!
The three core elements of gazetteers—placenames, place categories, and geospatial locations—support the translation between informal georeferencing using placenames (e.g., Santa Barbara) and place categories (e.g., city) and the formal georeferencing of mathematical schemes (e.g., longitude and latitude coordinate systems). These elements plus explicit relationships between named geographic places and the identification of time frames for places and their characteristics are the fundamental components of digital gazetteers. Within the context of gazetteer services—such as support for enterprise georeferencing systems, geoparsing of text to derive spatial locations, navigation services, and support for geographic information retrieval (GIR)—the complexities of each of these components challenge the collection and use of gazetteer data. This session will explore such issues as:

* appropriate generalization of the geospatial location
* creation and sharing of category schemes for gazetteers
* Accommodation for the variations and repetitions of placenames on a worldwide basis
* effective treatment of the space-time linkages
* integration of gazetteer data from multiple sources
• Great to hear from you. Very happy to be involved in this, it is certainly an area of significant interest to me. I am continuing to develop the Gazetteer for Scotland (which now contains 3.5 million words of contemporary and historical text as 13,500 entries with a lot of images), and thus have strong interests in long-form (descriptive) gazetteers.

• The issue of placenames has become a 'hot' topic in Scotland. You may or may not be aware that the UK has no placename authority and with the now-devolved government in Scotland we are trying to change this. There is a UK authority which operates internationally, debating other people's names, and we are close to having a Definitive National Address Database for Scotland, which will deal with addressable properties (and hence to people) and post-towns, but not neighbourhoods or geo-features.

The issues you highlight are important ones: Gazetteer services need to be developed as a hub to enable the connection of other services. Here there are significant problems with large numbers of publicly-funded historical / cultural / archive projects which all have place as a common element, yet I have heard librarians seriously suggesting that Dewey-Decimal be extended to cope. Laughable, if it weren't a serious suggestion. I am also involved in issues of georeferencing historical texts - there are different approaches, isolating and disamiguation individual places is fraught with difficulty, but an approach of isolating geographical units (the parish of x in the county of y) allows aggregation of places at these levels, but rely on semi-structured text (eg. historical gazetteers, statistical accounts etc.). I am also involved in the supervision of a chap called Jochen Leidner, based in Informatics here, who has sig. experience of geoparsing contemporary texts. He is close to finishing and may also be worthy of invitation. There are an increasing range of services which auto-geoparse news and suchlike and put the results up as a google-map. I also have some comments on wikipedia, which is both a useful development but also dangerously unreliable and with a propensity to plagiarise. (I did write most of the 'Gazetteer' entry on Wikipedia). It doesn't represent place-oriented entries in a very useful way.

Interoperability is key; with the ability to be able to exchange / hub basic gazetteer information, while respecting the copyright of a value added / descriptive component.
Descriptive Gazetteers

- Needed to distinguish places (e.g., Newbigging)
- Effective description of places
Summary

- Scottish Context
- Importance of Gazetteers
- Definitive Address Gazetteer for Scotland
- Importance of Descriptive Geographical Information
- The Gazetteer for Scotland
- Historical Descriptive Gazetteers
- Problems with Place Names
Some Definitions

- **Short-form gazetteers:**
  - often used in conjunction with computer mapping and GIS systems,
  - may simply contain a list of place-names together with their locations in latitude and longitude or other spatial referencing systems (eg. British National Grid reference). Short-form gazetteers appear as a place-name index in the rear of major published atlases.

- **Long-form (or Descriptive) gazetteers:**
  - may include lengthy textual descriptions of the places they contain, including explanation of industries, government, geography, together with historical perspectives, maps and / or photographs.

- **Thematic gazetteers:**
  - list places or geographical features by theme; for example fishing ports, nuclear power stations or historic buildings. Their common element is that the geographical location is an important attribute of the features listed.

- **Address Gazetteers:**
  - geo-referenced address lists, often used by local authorities for purposes such as local taxation.
Scottish Context

- Some big problems:
  - Regularly changing administrative geographies
  - Multiple (nearby) places of the same name
  - Gaelic names (no agreed version)
  - No definitive place-name gazetteer
- But some advantages:
  - Lots of mapping versions (from Pont in 16th C.)
  - Detailed country-wide mapping from 1860s
  - Scottish place-names survey (1980s) – but not digital
- Some real demands for gazetteer services
  - Who is the custodian of place names?
  - Place is the connection between heritage projects (although they don't realise that!)
  - Efficient government
Gazetteers for Contextualising Places

- All manner of google, yahoo, microsoft are pushing forward with local search, integrated with mapping
- Imagery (vertical, oblique, street-frontage) and geographically-referenced texts become key to these services
- Gazetteers are vital to provide the linkages
- Many independent resources built with public money, which all include significant GI component (although they may not realise that either!)
- Could create a world-class resource
- In Scotland, these might include:
  - Local and Central Government
  - RCAHMS Canmore (archaeology)
  - National Archives
  - National Library of Scotland (maps and archives)
  - Scottish Cultural Resources Archive Network (SCRAM)
  - Historic Scotland
  - Great Britain Historic GIS
  - Gazetteer for Scotland
PART 1: National Address Gazetteer for Scotland

Click to add text
National Address Gazetteer for Scotland

• Definitive National Address for Scotland Gazetteer (DNA-Scotland) will be aggregated from local corporate address gazetteers (CAGs) created and maintained by Local Authorities (Councils) and Property Tax Assessors
• May or may not be a postal delivery address (ie. in PAF)
• Effectively it will provide a national view of local data
• Aim is for it to become the trusted address source of choice within the Scottish Public Sector
• Seen as a standard means of connecting other databases together
• Based on the property, to which individual people can be connected
• Implementation accelerated due to the sudden need for Senior Citizen entitlement cards
• Is being resourced – will be maintained
Differing Stakeholder Views of a Property

- Property to be valued, classified and attributed [Assessors]
- Development to be controlled [LA Planning]
- Legal Title to be registered [RoS]
- Status of buildings and occupants [GROS]
- Property to deliver mail to [Royal Mail]
- Home address of 2 pupils [LA Education]
- Fire to be put out [Fire Brigade]
- Property to be protected [Police]
- Refuse to be collected [LA Cleansing]
- Council Tax to be collected [LA]
- Patient to be Collected [NHS Ambulance]

Brief definition of what a CAG is.
From BS 7666 to DNA Gazetteer

History of the Programme to indicate the complexity and inter-relationships between its components

BS 7666 is an generic standard and requires “wrappers” built around it to deliver a functioning system.

The Strategic Framework Document provided the blueprint of what the programme was going to deliver 32 CAGs aggregated into a national dataset.

Step by step guide provide assistance to councils on governance as well as operational considerations

Tech requirements and Conventions absolutely essential if CAGs were to be joined up otherwise there would be 32 different “flavours”

National Stakeholder Requirements find out what was required if public sector was to adopt DNA-Scotland as definitive

BPM both “As Is” to baseline current (pre CAG) situation and “To Be” to define best practice and max benefits NO ONE GENERIC model though.

Quality is the fundamental to all of the programme – we will keep coming back to this ground-breaking work. Not been attempted anywhere else.
Description of how it is planned to work

Property Database held in Trading Community Architecture within Oracle e-Business suite

Originally Property Database only but realise the potential for DNA Gazetteer THE DNAG
The scary diagram same as previous but more impressive!!
Even more scary as it is in the process of being built

Brief Description

Pilot Councils: SSDN to centre: Web Services etc
This is the high level overview of a possible future scenario
Better business processes to other organisations
Potential for shared services
Increased usage for gazetteer
Red circle is local BOI – benefit will exceed resource costs
Implications of DNA-Scotland

- Deals only with street address and post-town
- Does *not* deal with non-addressable features (natural features etc. etc.)
- Also does *not* deal with urban or rural districts, which are so important in defining communities

- Not just addresses, people are very much linked in
- Very unlikely to be available to the private sector, so only a definitive address for some
- Even so, there are some real civil liberties implications
- The *entitlement card* is effectively a national identity card by the back door
- For the moment, saved only by the incompetence of government
PART 2: Textual GI and Descriptive Gazetteers

Click to add text
A Case for Textual GI?

- Are points, lines and polygons with attributes enough?
- Do these fully represent the world around us?
- How do we navigate from one place to another?
- Maps are often not terribly well understood by non-specialists
- Use features which are not currently part of maps or GI databases!
  - Landmarks
  - Descriptions or Textures of Buildings
  - Colour of Features
  - May be seasonally dependent
- Maybe we need richer data
- Building frontage data is being captured by videography (eg. Amazon A9, Microsoft’s intentions for local.live)
- Textual GI is one form of this richer data
So Descriptive Gazetteers are Important

- Descriptive GI is poor little brother of maps
- But this textual GI makes an important contribution
- This is exactly what google etc. are trying to search
- A traditional form of GI, which has recently been forgotten:
  
  "Next to a good dictionary, the most generally useful book is a good gazetteer"
  
  -- W.G. Blackie (1855)

- Actually, there was a solid tradition of gazetteer production in Scotland during the 19th C.
- Complements the map
- Potentially rather more comprehensible by the public
- The geographical equivalent of the geological map and monograph
- Helps to solve the navigation problem
The Gazetteer for Scotland

- Aims to be a definitive and detailed resource
- Began work in 1995; at least another 10 years to go!
- Originally envisaged as a book
- Statistics:
  - Around 13,500 entries
  - 1.2 million words of contemporary text
  - 2.25 million words of historical text
  - c.7500 photographs
  - Maps with every entry as mashups using:
    - OS get-a-map
    - Google Maps
    - Oldmaps.co.uk
  - Also direction-finding via multimap.com
  - Approaching 1 million hits / week
- Significant interest for tourism and genealogy
- There are places in the GfS which are not on OS maps
Welcome to the Gazetteer for Scotland

The Gazetteer for Scotland is a vast geographical encyclopaedia, featuring details of towns, villages, bars and goods from the Scottish borders to the northern Isles. The first comprehensive gazetteer produced for Scotland since 1884, it includes tourist attractions, industries and historical facts, together with histories of family names, biographies of famous people and descriptions of historical events associated with Scotland. The network of connections between all of these entries make the gazetteer unique. Visitors will also find we have included the text of Gordon's 19th Century Ordnance Gazetteer of Scotland in our entries to provide a historical perspective.

With at least 32442 detailed entries, the Gazetteer is already the largest Scottish resource available on the web, yet it is growing constantly, with some 991 entries added or updated in the last week alone. Please visit our FAQ for more information.

Browse the Gazetteer for Scotland using either:

Maps and Places OR History Time-Line OR Search Pages

Alternatively, look at our:

Featured Geographical Feature: Loch Dochart.

You can also use the interactive facilities in our free Member's area to make maps and add your own comments to individual areas.

NEWS FLASH!

Just published: Scotland: An Encyclopedia of Landscape and Race, the most comprehensive reference on the geography of Scotland in more than 100 years, drawing on material from the Gazetteer from Scotland, but also including 32 pages of full-colour mapping and much more: buy it from our bookshop in time for Christmas.

New for December 2006, more than 500 additional photographs for the West and Southwest of Scotland.
Details

- Main user interface designed for the general public
- Implemented as an Oracle database
- Makes full use of relational joins to link entities:
  - Settlements - Geofeatures
  - People - Families / Clans
  - Events - Associated biblio references
  - Media (photographs, sounds, video)
- Database is **not** queried in real-time; rather pages are pregenerated at 4-6 week intervals (except new)
- Compromise between performance and disk-space
- Actually a spatio-temporal database:
  - Dates are included where poss.
  - Linkages allow geographical inferences to be made (eg. people associated with various georeference places)
- Could easily form hub of a short-form gazetteer service
Temporality appears to public as timelines

The Gazetteer for Scotland History Time-Line

The Scottish History Time-Line provides an alternative route to access the Gazetteer for Scotland. The time-line incorporates thousands of Scottish historical events including the births and deaths of famous Scots, archaeological sites and historic buildings in Scotland, together with the dates of battles and other cultural events. Each event is colour coded, with a key given below. The time-line is fully linked into the vast store of information which forms the Gazetteer for Scotland database.

Choose a time-period from the list below:

- Until the 13th Century
- 14th and 15th Centuries
- 16th Century
- 17th Century
- Early 18th Century; 1700-49
- Late 18th Century; 1750-99
- Early 19th Century; 1800-49
- Late 19th Century; 1850-99
- 20th Century

Note that, within a single year, events are not necessarily presented in chronological order. The time-lines only give direct access to a limited number of Gazetteer for Scotland entries, because only some can ever have dates associated with them.

KEY:

- Airport, Bridges, Canals and Streets
- Archaelogical Sites and Historic Monuments
- Births and Deaths
- Events
- Exhibitions, Centres, Museum or Gallery
- Historic Buildings
- Industrial or Commercial Sites
- Institutions or Government Buildings
- Gardens or Parks
- Sports Venues
- Other Types
Port Logan

Use the tabs on the right of this page to see other parts of this entry.
PART 3: Historical Descriptive Gazetteers

Click to add text
Issues of History

- Richness of textual description
- Text maintains the subtleties of history and details ephemeral knowledge much more easily and often more effectively than a map
- Unfortunately recently the trend has been towards a tourist guide rather than a systematic description of places which traditionally formed a gazetteer
- But there are problems:
  - Neither have we good tools to interrogate descriptive GI or make inferences from it
  - Data mining techniques represent only the beginnings of a solution
  - Nor good tools to generate it
  - We don't even do geo-parsing particularly well
  - But then how do we differentiate the three Newbiggings within a few miles of each other in Angus??

LAST POINT: People Don't now their co-ordinates!
Muirhouse, a mansion in Craigmillar parish, Edinburghshire, on the shore of the Firth of Forth, 14 miles NNW of Edinburgh's Main and 4 miles WNW of Edinburgh. Only two round towers remain of a previous mansion (circa 1670); and the present house is a picturesque Tudor edifice of about 1880, with a square battlemented tower and beautiful well-wooded grounds. Its drawing-room is adorned with several frescoes by Zephaniah Bell. Purchased by his ancestor in 1779, the estate belongs now to Thomas Davidson, Esq., F.R.S., F.L.S., F.G.S. (b. 1717; suc. 1805), who lived at..

Muirhouse
City of Edinburgh

Lying to the north of Ferry Road in Edinburgh and to the west of Milton, Muirhouse is a public housing estate begun in 1933. It takes its name from the Tudor-Gothic mansion built on Marine Drive in 1832 for the Davidson family, who were wealthy merchants trading in Rotterdam. The estate was built in mixed housing styles, including regular white boxes with low-pitched roofs, and large tower blocks, mostly constructed in the 1940s. The area has suffered from significant social problems, especially from the mid-1980s Muirhouse, together with the neighbouring West Milton, have been undergoing significant renewal, with many of the tired and vandalised council-owned blocks being replaced by private housing. Muirhouse suffers unemployment rates of more than three times the average for Edinburgh, with a high proportion of long-term unemployment.
Roslin (Brit. ross, 'a point,' and lynn, 'a waterfall,' the name often and perhaps more correctly spelled Rosslyn), a quoad sacra parish containing a village, chapel, and castle of the same name, in the civil parish of Lasswade, in the county of Edinburgh. The village, which stands on high ground near the NW bank of the river North Esk, has in its neighbourhood three railway stations on different sections of the North British railway system, and each of them distant about 10 miles from Edinburgh. The nearest, Roslin, on the Edinburgh and Glencorse branch, is close to the village. Rosslyn Castle, on the Edinburgh and Penicuik branch, is distant about 1¼ mile, and Rosslynlee, on the Edinburgh and Peebles line, about 1¾ mile. By road the village is about 6½ miles S of Edinburgh; and from Polton station, 7 miles SSE of Edinburgh, a public footpath winds through the beautifully wooded glen of the North Esk to the village, the distance being about 2½ miles. About 1440, under the fostering protection of William St Clair, Prince of Orkney, Duke of Oldenburg, and having a string of other titles that it would weary even a Spaniard to repeat, the place is said to have stood third in Scotland for importance...

[ Groome, Ordnance Gazetteer of Scotland, 1885 ]
Case Study: Learning from Failure

- Wish to provide historical depth to the *Gazetteer for Scotland*
- Attempted to semi-automatically link Groome’s *Ordnance Gazetteer*
- Limited success (initially as low as 15-20% match)
- Lots of manual checking required to confirm
- Why?
  - settlements, country houses etc. disappeared
  - historical places not in contemporary gazetteer
  - relative importance of places (eg. estates, laird’s houses)
  - new places (airports, industrial facilities settlements etc.)
  - spelling change / inconsistency
  - old Gaelic renderings
  - inadequate description to distinguish similarly named entries (eg. surprising number of Black Lochs)
select distinct gr.seqno, gr.name, gr.type1, gr.county1, gr.county2, gr.county3, gf.seqno, gf.name, ca.name caname
from groometemp gr, geofeatures gf, grcouca_rel, geoca_rel, council_areas ca
where (soundex(gr.name) = soundex(gf.name)
and gr.county1 = grcouca_rel.couname)
and grcouca_rel.caseqno = geoca_rel.caseqno
and geoca_rel.gseqno = gf.seqno
and grcouca_rel.caseqno = ca.seqno

Gets a bit more complicated than this...
As part of the same process of historical depth to the GoS, wished to create contemporary parish entries to which Groome text could be linked.

Contemporary entries needed to relate parishes to modern geographies.

Would be enormously tedious to manually research and write 1000+ of these entries.

Take digitised parishes for 1951 (out-of-copyright).

Extract topology to give relationships one to another.

Encode the database with as many facts as possible.

Relate to historical counties and current council areas.

Try to make results interesting to read (non-formulaic).

Not surprisingly this is rather difficult to achieve!

But most people look at relatively few entries.
First Attempts at Database-Generated Text

Monikie

Monikie is a parish, which lies in Angus Council Area, some 5 miles (8 km) northwest of Carnoustie in Angus and 8 miles (12 km) northeast of Dundee in Dundee City.

It lies in the old county of Forfarshire, which disappeared following the Local Government (Scotland) Act 1974.

The parish has an area of 30.9 sq. km (22.6 sq. miles). Monikie has 6 neighbouring parishes; namely Barry, Camyllie, Inverarity, Monifieth, Munroes and For}\n
Forfar

Forfar is a parish, which lies in Angus Council Area, some 6 miles (9 km) southeast of Kirkcaldy and 10 miles (17 km) southwest of Brechin in Angus.

It lies in the old county of Forfarshire, which disappeared following the Local Government (Scotland) Act 1974.

The parish has an area of 34.3 sq. km (21.3 sq. miles) and includes the settlement of Forfar. Forfar has 6 neighbouring parishes; namely Currie, Carn, Inverarity, Monifieth, Monroes and For}\n
Longforgan

Longforgan is a coastal parish, which lies in Perthshire and Kinross Council Area, some 7 miles (11 km) west of Dundee in Dundee City and 9 miles (13 km) west of Newport-on-Tay in Fife.

It lies in the old county of Perthshire, which disappeared following the Local Government (Scotland) Act 1974.

The parish has an area of 34.5 sq. km (21.4 sq. miles). Longforgan has 7 neighbouring parishes; namely Aberdour, Carn, Dundee, Fowlis Easter, Inchture, Kintore and Liff and Bowrie. The boundaries of this Parish were modified among significant changes recommended by the Boundary Commissioners after the Local Government (Scotland) Act 1889.

Arbroath and St Vigeans

Arbroath and St Vigeans is a coastal parish, which lies in Angus Council Area, some 8 miles (12 km) northeast of Carnoustie and 10 miles (16 km) south of Brechin in Angus and includes a small offshore island.

It lies in the old county of Forfarshire, which disappeared following the Local Government (Scotland) Act 1974.

The parish has an area of 51.6 sq. km (32 sq. miles) and includes the settlement of Arbroath. Arbroath and St Vigeans has Arblastic, Camyllie and Invintakler as its neighbouring parishes. The boundaries of this parish were modified among significant changes recommended by the Boundary Commissioners after the Local Government (Scotland) Act 1889.
PART 4: Problems with Place-names

Click to add text
Defining Places

- Textual GI helps us define places
  - These three Newbiggings in Angus!
- Government, industry, communities need to be able to locate and describe places
  - Angus Fire Brigade were recently reported as saying that determining which place they were heading for was a significant issue! People don't know their coordinates
- The first step in community-building and social inclusion
- In Scotland have chaos in our place-names!
- But, did you know that the UK is one of relatively few countries in the world which doesn't have a place-names authority?
Some Examples

- Consider the OS 1:50,000 sheet vs the Collins Bartholomew 1:250,000 road atlas
- Four different renderings of 'Tolsta' on just two maps!
- Typically also problems between:
  - LA directions signs
  - Village name signs
  - Local Shop / PO
Consider Balivanich (Benbecula)

- **Balivanich** (42000) or **Ballivanich** (419) or **Balivanish** (34)

- In Gaelic: **Baile Mhanaich** (1780), **Baile a' Mhanaich** (1100), **Bhaile a' Mhanaich** (33), **Bailebhainich** (17) or **Bhaile a Mhanaich** (11)

- Four different renderings on Comhairle nan Eilean Siar web site

- Ordnance Survey Gazetteer, 1:50K and 1:25K maps, Collins and AA road atlases are all different, but also in flux!

- Plus historical versions

- OS differences possibly due to progressive implementation of their Gaelic names policy
**Some More Confusion**

**Where was the folk-singer Calum Kennedy (1928 - 2006) born?**

The *Guardian* and *Herald* obituaries gave "Orasay", the *Times* and the *Independent* gave "Orosay". With a little research, it turns out to be "Orinsay", which the current OS map portrays only in Gaelic as "Orasaigh", although the first edition (widely used for historical research) gives "Orosaigh".

How can we effectively GEOPARSE and GEOREFERENCE TEXTS, and DISAMBIGUATE PLACENAMES, when we have this sort of PROBLEM?
The Need for a Place Names Authority

- Resolution IV of the *UN Conference of Geographic Names* (1967) directed that standardisation of place names be achieved by establishing national place name authorities
- The USA has the US Board on Geographic Names
- We have the *Permanent Committee on Geographical Names for British Official Use*:
  
  "Principal function is to advise British government departments and agencies on policies and procedures for the proper writing, in official publications, of geographical names for places and features outside the United Kingdom, excluding those of the Antarctic"
- The Scottish Place Names Society have an academic interest in the historical aspects of toponomy
- Scotland desperately needs an Authority:
  - Advising local and central government
  - Tourism
  - Integrated transport
  - Citizen-focussed services
  - Informing Utilities
  - Ensuring the consistency of names
  - Promoting Scotland
  - Understanding and reinforcing our heritage
  - Joining up GI
  - Advising property developers (new streets etc.)
- Those involved in managing cultural resources are desperate for a means of joining up their resources
A GI Strategy for Scotland

- Connecting the *Spaces, Faces and Places* of Scotland through geography and geographic information
- Specifically mentions place-name data
- Includes the *non-addressables*
- States:
  - place-names are basic elements of library, museum and other archive catalogues (name-based archives)
  - The first element in searching for information about places with computer search tools
  - Such techniques have the potential to unlock Scotland's "hidden wealth" of archived information about its places and people
  - Currently limited by the lack of a standardised approach to place names in Scotland
  - Raised as a significant issue during consultation on the draft Strategy, particularly with regard to Gaelic place-names
Conclusions

• Descriptive gazetteers are enormously valuable
• Can include the history which gives authority
• Gazetteer services as vital because:
  – they can act as the glue to join up other services
  – they can provide definitive names, statistics and possibly even definitive descriptions
• Scotland (and I'm sure the rest of the UK) needs a Place-name Authority
• The Vision for Scotland:
  – A free-to-access definitive gazetteer service
  – Because we can hang all sorts of things from this

• I was going to have rant about Wikipedia, but I've run out of time...
The Wikipedia Effect

- We all tell our students not to use Wikipedia, because its information is potentially unreliable
- **However much more importantly:**
  - Wikipedia recycles information from one place on the web to another, with no added value, but detracting from the original source
  - Its playing hard-and-loose with copyright laws are a disincentive to professional reference editors putting material on the web
  - This discourages scholarship,
  - Its clones repeat the same bias / errors several times over; an intelligent geo-aggregator might regard facts as correct because they are repeated
  - Issues of political bias, peculiar agendas, or just plain wrong!
  - The Wikipedia model is a disastrous one for gazetteer services, but Wikipedia is already a geo-referenced authoritative gaz service whether we like it or not !!!
The three core elements of gazetteers - placenames, place categories, and geospatial locations - support the translation between informal georeferencing using placenames (e.g., Santa Barbara) and place categories (city) and the formal georeferencing of mathematical schemes (e.g., longitude and latitude coordinate systems). These elements plus explicit relationships between named geographic places and the identification of time frames for places and their characteristics are the fundamental components of digital gazetteers. Within the context of gazetteer services - such as support for enterprise georeferencing systems, geoparsing of text to derive spatial locations, navigation services, and support for geographic information retrieval (GIR) - the complexities of each of these components challenge the collection and use of gazetteer data. This session will explore such issues as:

* appropriate generalization of the geospatial location
* creation and sharing of category schemes for gazetteers
* Accommodation for the variations and repetitions of placenames on a worldwide basis
* effective treatment of the space-time linkages
* integration of gazetteer data from multiple sources
Great to hear from you. Very happy to be involved in this, it is certainly an area of significant interest to me. I am continuing to develop the Gazetteer for Scotland (which now contains 3.5 million words of contemporary and historical text as 13,500 entries with a lot of images), and thus have strong interests in long-form (descriptive) gazetteers.

The issue of placenames has become a ‘hot’ topic in Scotland. You may or may not be aware that the UK has no placename authority and with the now-devolved government in Scotland we are trying to change this. There is a UK authority which operates internationally, debating other people’s names, and we are close to having a Definitive National Address Database for Scotland, which will deal with addressable properties (and hence to people) and post-towns, but not neighbourhoods or geo-features.

The issues you highlight are important ones: Gazetteer services need to be developed as a hub to enable the connection of other services. Here there are significant problems with large numbers of publicly-funded historical / cultural / archive projects which all have place as a common element, yet I have heard librarians seriously suggesting that Dewey-Decimal be extended to cope. Laughable, if it weren’t a serious suggestion. I am also involved in issues of georeferencing historical texts - there are different approaches, isolating and disamiguating individual places is fraught with difficulty, but an approach of isolating geographical units (the parish of x in the county of y) allows aggregation of places at these levels, but rely on semi-structured text (eg. historical gazetteers, statistical accounts etc.). I am also involved in the supervision of a chap called Jochen Leidner, based in Informatics here, who has sig. experience of geoparsing contemporary texts. He is close to finishing and may also be worthy of invitation. There are an increasing range of services which auto-geoparse news and suchlike and put the results up as a google-map. I also have some comments on wikipedia, which is both a useful development but also dangerously unreliable and with a propensity to plagiarise. (I did write most of the ‘Gazetteer’ entry on Wikipedia). It doesn’t represent place-oriented entries in a very useful way.

Interoperability is key; with the ability to be able to exchange / hub basic gazetteer information, while respecting the copyright of a value added / descriptive component.
Click to add title

- Click to add an outline
Descriptive Gazetteers

- Needed to distinguish places (eg. Newbigging)
- Effective description of places