

making mapping real

experiences & thoughts from

National Library of Scotland

gill hamilton
digital access manager
g.hamilton@nls.uk

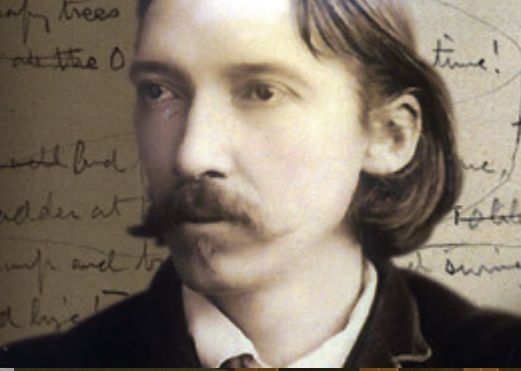
and introducing ...
Le Chef

agenda

- introduction(s)
- background to the DODLOD@NLS
- *The Big Issue* : mapping local instances to the global graph
- practical techniques
- discussion



Mon frere mon beau frere étant par la permission
de Dieu pour mes pechez comme ie croy venir
me refter entre les bras de ceste Roynie ma
confiance en ray en beaucoup deuans et passe
pres de vingt ans ie suis entrin pareille et sus
estete condampnee a la mort et ayant demandé
mes papiers juy eulx ostez a ceste fin de feure
mon testament ie nay peu veue recevoir d'un
curist ny obtenir chose den faire un libre
ny quapris ma mort mon corps fust transporté
de l'on mon deuy en votre royaume ou ray en
honneur de estre royne vostre sœur et ancienne



Bibliographic information organization in the semantic web

Mirna Willer and Gordon Dunsire



my glamorous
assistant

LE CHEF

For lossless triples, create local element set based on RDBMS structure



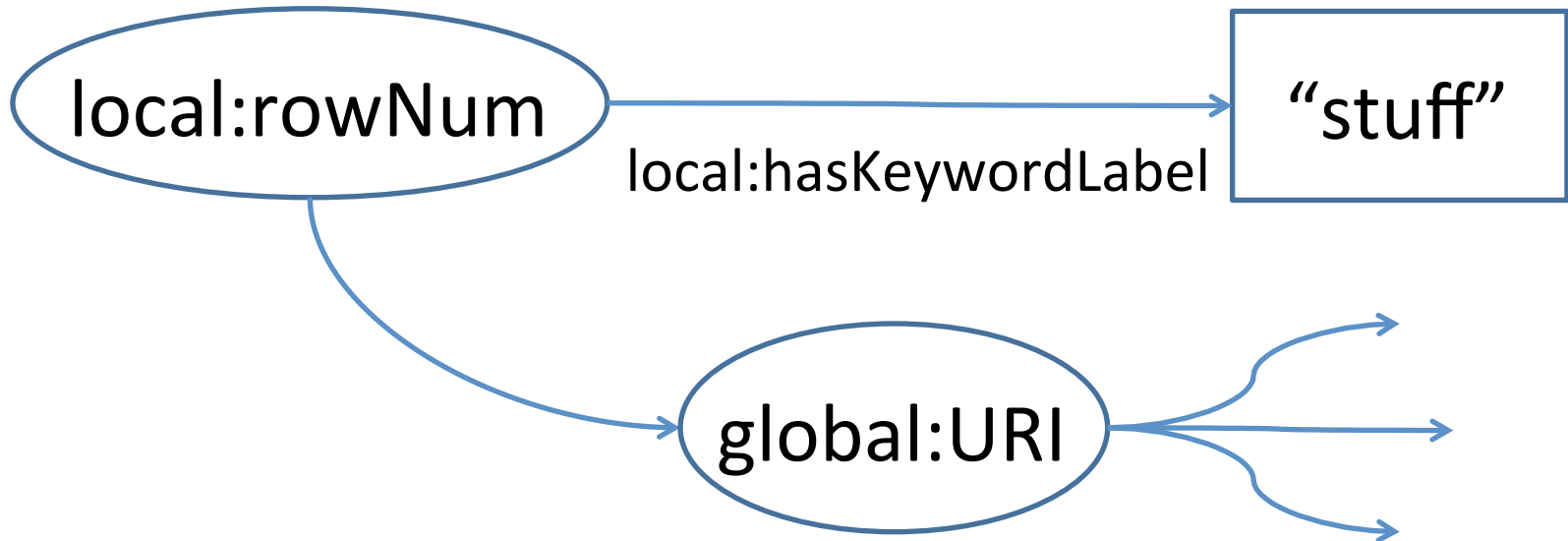
Element Sets: Show detail for National Library of Scotland Digital Object Database elements

Detail	Elements	History	Maintainers			
Label	Type	URI	Status	Updated	Last Updated by	Actions
Resource	class	.../dod/elements/Resource	Published	2012-12-30 12:21	Gill Hamilton	
Place	class	.../dod/elements/Place	Published	2012-12-30 12:24	Gill Hamilton	
Date or Event	class	.../dod/elements/DateOrEvent	Published	2012-12-30 12:28	Gordon Dunsire	
Keyword	class	.../dod/elements/Keyword	Published	2012-12-30 12:30	Gill Hamilton	
Who	class	.../dod/elements/Who	Published	2012-12-30 12:31	Gill Hamilton	
Form and Genre	class	.../dod/elements/FormAndGenre	Published	2012-12-30 12:33	Gill Hamilton	
has title	property	.../elements/tDescription-title	Published	2012-12-30 13:50	Gill Hamilton	
has description	property	.../tDescription-description	Published	2012-12-30 13:52	Gill Hamilton	
has parent	property	.../elements/hasResourceParent	New-Proposed	2012-07-15 10:16	Gill Hamilton	
has shelfmark	property	.../elements/hasResourceShelfmark	New-Proposed	2012-07-15 10:16	Gill Hamilton	
has Voyager bib id	property	.../hasResourceVoyagerBibId	New-Proposed	2012-07-15 10:17	Gill Hamilton	
has form or genre	property	.../dod/elements/hasFormOrGenre	Published	2012-12-30 13:04	Gill Hamilton	
has who	property	.../dod/elements/hasWho	Published	2012-12-30 13:12	Gill Hamilton	
has who depicted	property	.../elements/tWhoType-Depicted	New-Proposed	2013-01-24 7:32	Gill Hamilton	

Extracting triples

- Extract a collection from the operational database (SQL Server)
- Upload into MySQL database
- Run PHP script against MySQL
 - Create URIs using design templates
 - Base domain + entity/authority type + local/table record/row number
- Script outputs triples in TTL format

“Authority” strings assigned local things (URIs)



Local things can then be matched to global things

External mappings

- Mappings from local to global things stored as turtle files
 - Separate from extracted local triples
- New mappings can be added ad hoc
 - When resources available
 - Prioritised by local usage/focus/context
 - E.g. “Haig, Douglas, Sir, 1861-1928”
 - E.g. “Ypres, 3rd Battle of, Ieper, Belgium, 1917”
 - dbPedia “Battle of Passchendaele”



What IS the URI for "Spud" anyway?



for NLS

(and everyone else probably)

is the mapping of *historical strings* to their modern day *things* or URIs ... that we would have used if we were starting now.

LoC^(h) Getty

not Loch Gelly

THE BIG
ISSUE



Name Authority File (names)

[go](#)

Subject Authority File (keyword)

[go](#)

Thesaurus for Graphic Materials (keyword)

[go](#)



Thesaurus of Geographic Names (place)

[go](#)

Art & Architecture Thesaurus (keyword)

[go](#)

D74548618

Keyword

D1790



D1790

Keyword-keyword


“Kilts”

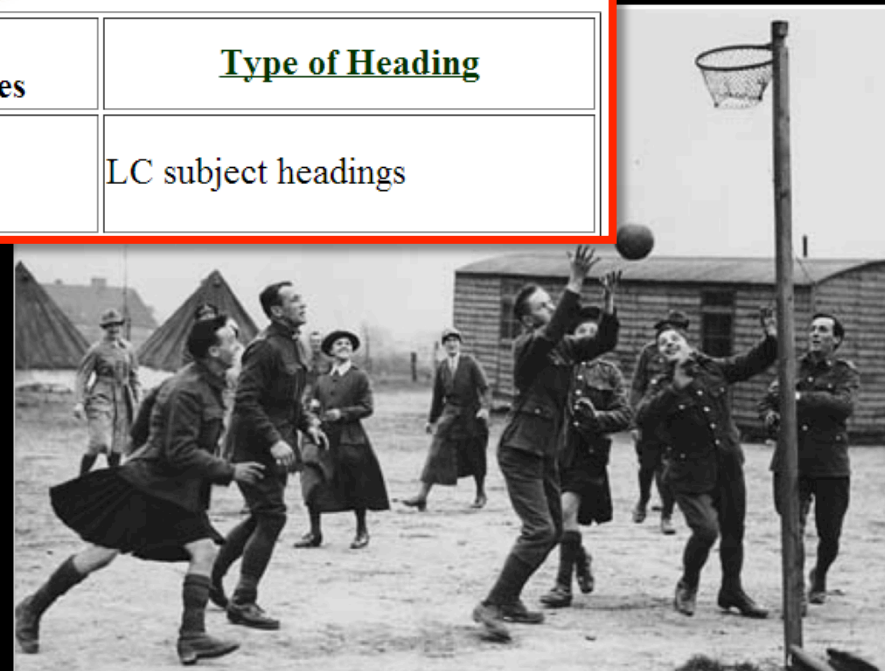


D1790

exactMatch

sh85072341 ?

#	Bib Records	<i>select icon in first column to...</i> View Authority Headings/References	<u>Type of Heading</u>
 1	1	Kilts	LC subject headings



D74548618 tWhoType-Depicted

D9886



D9886 tWho-who


“Great Britain. Army.
Women’s Army
Auxiliary Corps”

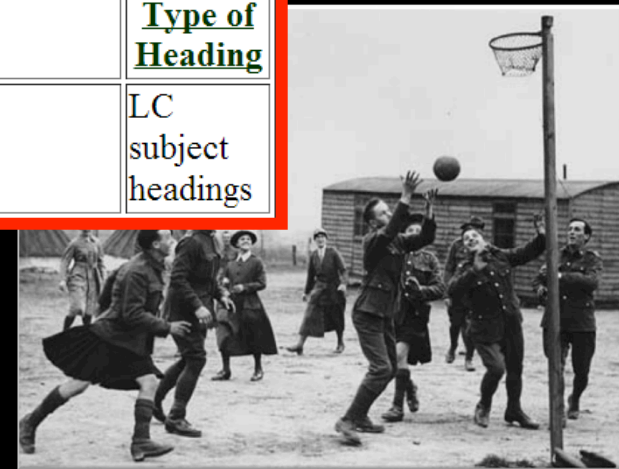


D9886 exactMatch

no2006000034



#	Bib Records	<i>select icon in first column to...</i> View Authority Headings/References	Type of Heading
 1	0	Great Britain. Army. Women's Army Auxiliary Corps	LC subject headings



D74549224 tPlaceType-Placedepicted

D575



D575 tPlace-place

“Cambrai”



D575 exactMatch



ID: 7008456

Record Type: administrative

 Cambrai (inhabited place)

Coordinates:

Lat: 50 10 00 N *degrees minutes* Lat: 50.1667 *decimal degrees*

Long: 003 14 00 E *degrees minutes* Long: 3.2333 *decimal degrees*

Note: Was a Roman center on Escaut river; was a Frankish capital by 445 CE; obtained rights from bishop in 1227; was longtime pawn of contending rulers; league against Venice was formed here 1508; site of signing of Peace of Cambrai 1529; ceded to France 1678.





art & architecture

THE BIG ISSUE

D74546504 Keyword

D1234




D1234 Keyword-keyword "War photography"



D1234 exactMatch



ID: 300157838

 **war photography** (photography, <photography and photographic processes and techniques>)







Note: Use with regard to photography in war for the purposes of historical or documentary photography, such as for mapping or training, use "military photography."

Terms:

- war photography** (preferred, C,U,LC,English-P,D,U,U)
- photography, war** (C,U,English,UF,U,U)
- oorlogsfotografie** (C,U,Dutch-P,D,U,U)
- fotografía de guerra** (C,U,Spanish-P,D,U,U)

Facet/Hierarchy Code: K.KT

Hierarchical Position:

-  Activities Facet
-  Processes and Techniques (G)
-  <processes and techniques> (G)
-  <processes and techniques by specific type> (G)
-  <image-making processes and techniques> (G)
-  <photography and photographic processes and techniques> (G)



LoC_(h) Getty

**THE BIG
ISSUE**



Name Authority File

[go](#)

Subject Authority File

[go](#)

Thesaurus for Graphic Materials

[go](#)



Thesaurus of Geographic Names

[go](#)

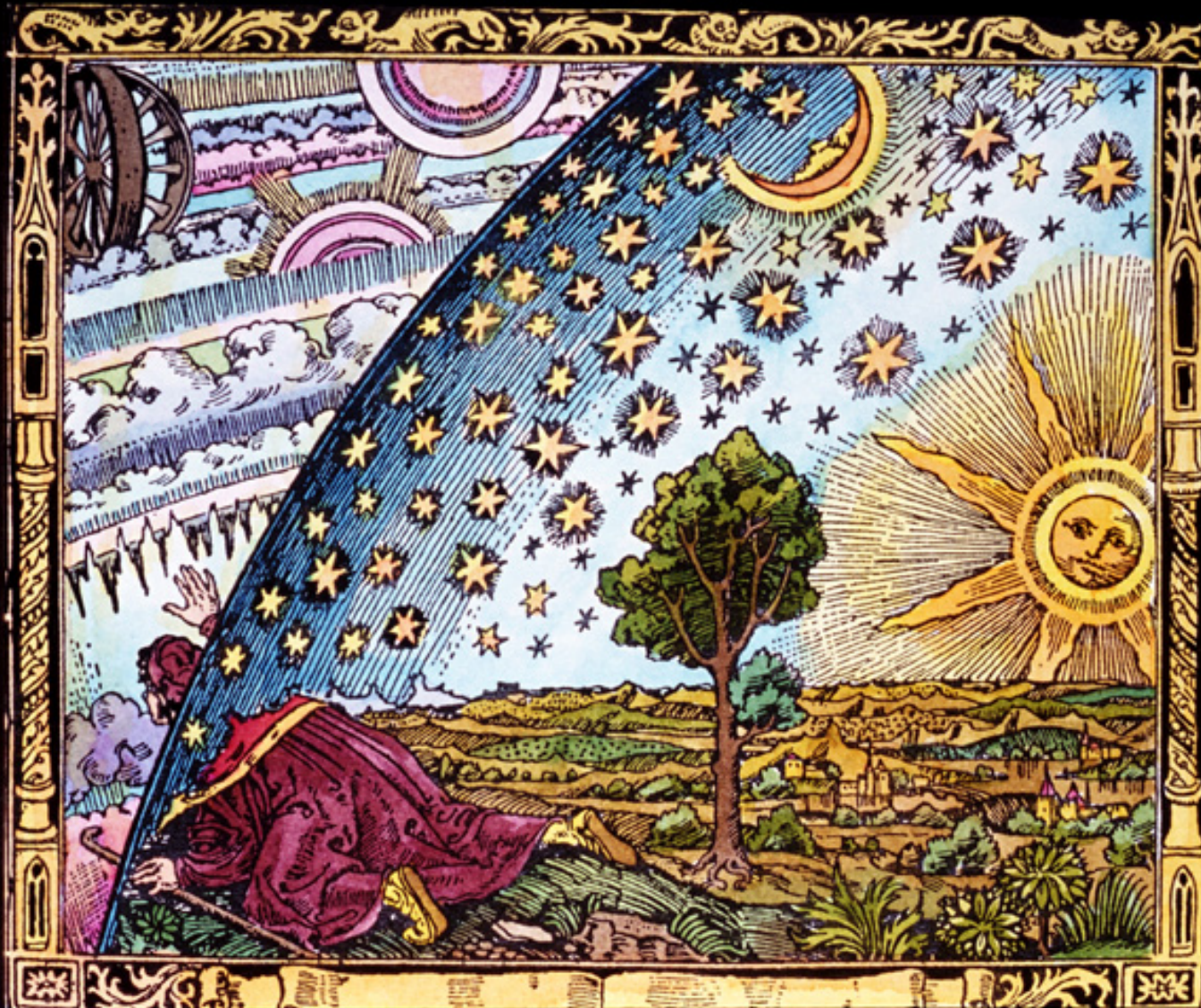
The Art and Architecture Thesaurus

[go](#)

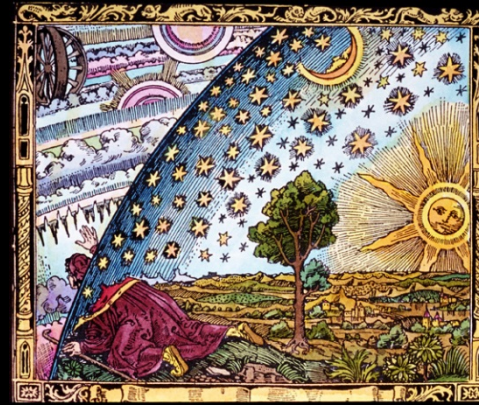


the extended graph

**THE BIG
ISSUE**



once for all time!



THE BIG
ISSUE

LCSH:sh85072341

exactMatch productontology.org/doc/Kilt

exactMatch dbPedia:Kilt

DOD:D575

exactMatch

[geoNames:3029030](https://www.geonames.org/3029030)



matching local to global

statistical

literals

humans

influences

statistical

the law of large numbers

statistical matching

DDC to LCSH

Monsieur Le Chef will explain

(he's a mathematical physicist too dontcha know)

MAGIC

literals

String matching Yeah baby!

What do we want?

UNIQUE MATCHES!

When do we want them?

NOW! NOW! NOW!

OH NO!!! hang on a minute



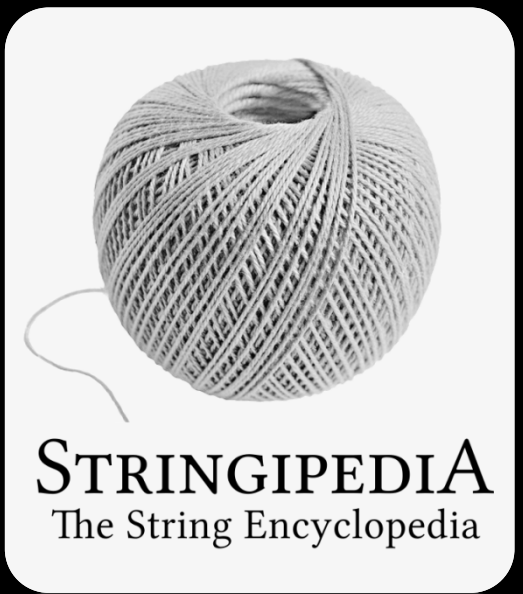
literals

String matching **W****A****O****N****H****O****O**!!

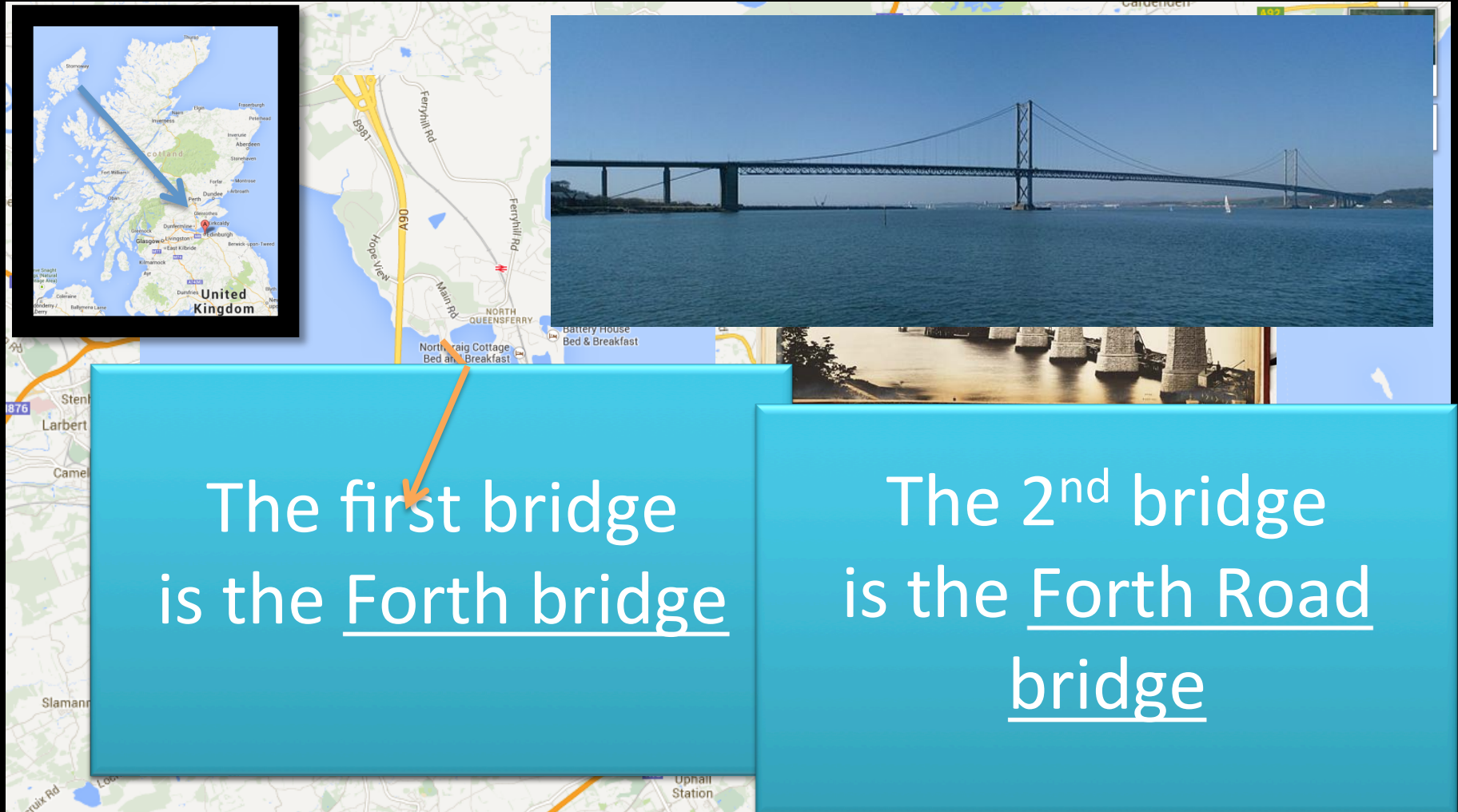
EXACT MATCH

means we've gotta URI
need resolution

but can you really
REALLY trust it?



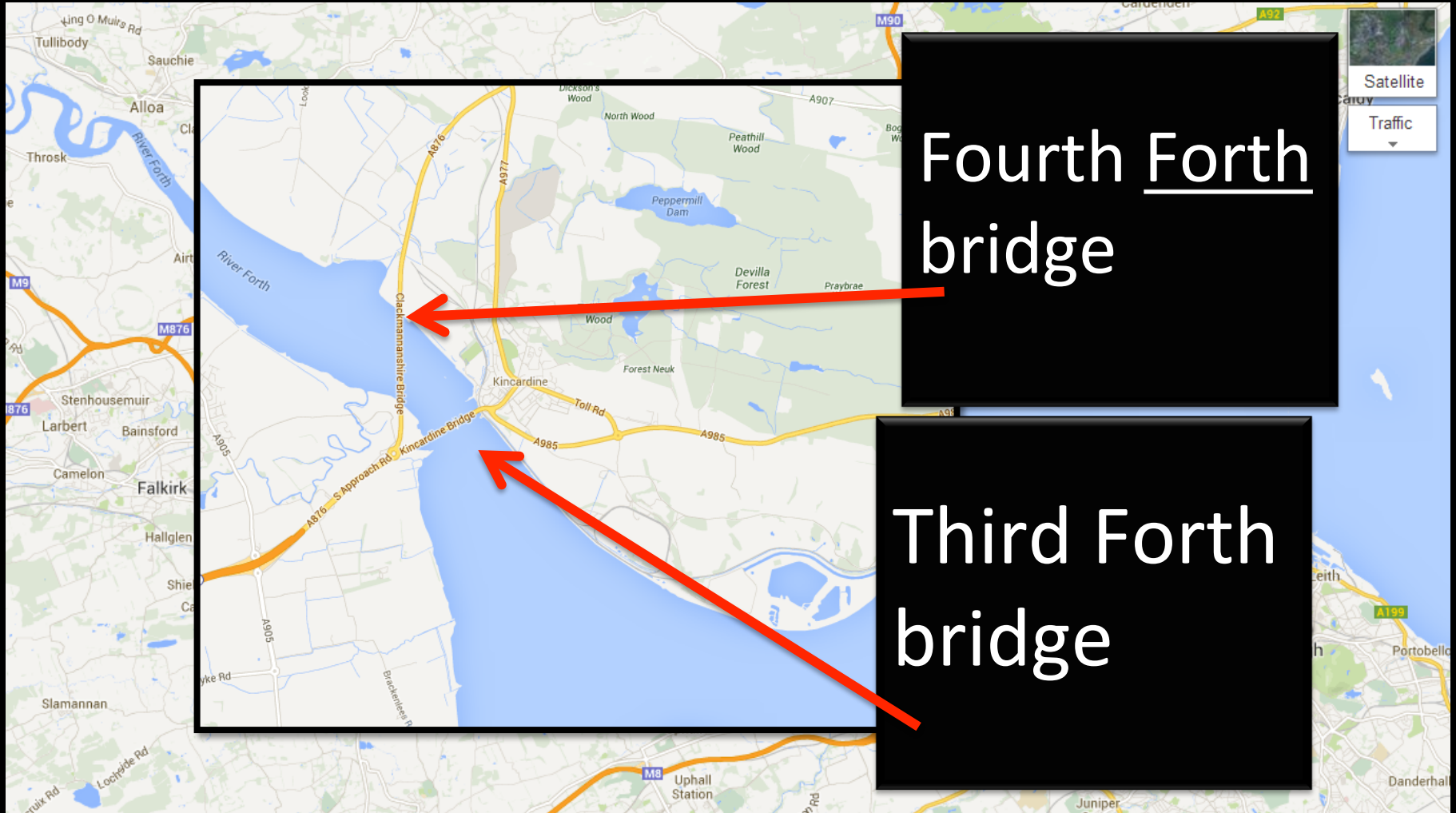
an aside the innumerate Scots



The first bridge
is the Forth bridge

The 2nd bridge
is the Forth Road
bridge

an aside the innumerate Scots



an aside

the innumerate Scots

The First Bridge on the Firth of Forth is the Forth bridge

There's a third Forth Bridge on the Firth

The 2nd Bridge on the Firth of Forth is the Forth Road bridge

And Finally, there's the Fifth Forth bridge on the Firth of Forth

And on the Firth of Forth there's a Fourth bridge but it's not the Forth bridge

humans

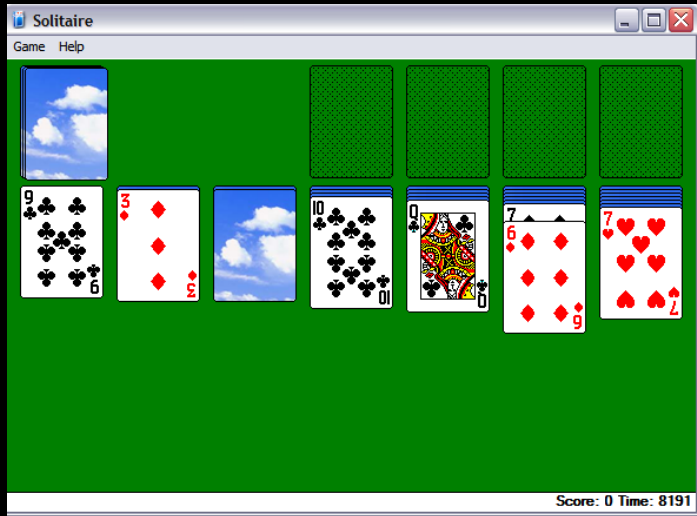


individuals

groups

crowds

individuals



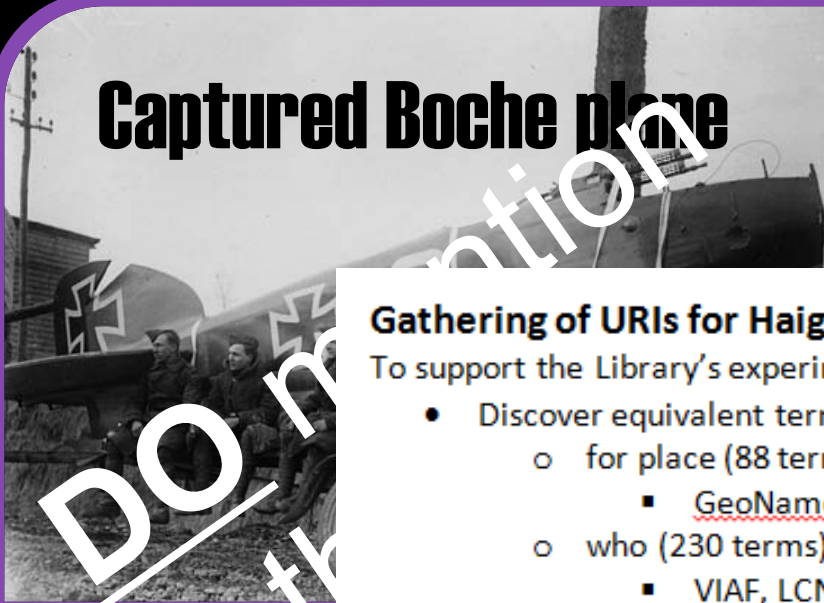
versus

Names	Surname	Forename	Middle name	Person/Org	Not sure	More than one
Anthony Stuart	Stuart	Anthony		P		
Alan Christopher Oliver	Oliver	Alan	Christopher	P		
Lukas van der Graaf				P	x	
Hamilton, Gill	Hamilton	Gill		P		
St Ola of Kirkwall				X	x	
John Smith & Alison Jones				X		x
British Broadcasting Corporation				O		

groups



Captured Boche plane



Gathering of URIs for Haig collection

To support the Library's experiment and learning with linked open data.

- Discover equivalent terms in other vocabularies
 - for place (88 terms)
 - [GeoNames](#), [dbPedia?](#), others?
 - who (230 terms)
 - VIAF, LCNAF, others?
 - keyword (1235 terms)
 - LCSH, DDC, TGM, others?
 - Form/genre (9 terms)
 - Date or event (19 terms)
- record URIs
- record SKOS mapping (close match, broad match [etc](#))

German Guest

Basil Fawly: O

German Guest: or course.

Basil Fawly: Ah, wonderful! WUNDERBAR! Ahh! Please allow me to introduce myself, I am the owner of Fawly Towers. And may I welcome your war... your war... you wall... you all... you all, and hope that your stay will be a happy one. Now, would you like to eat first, or would you like a drink before the war... AHH! Er... trespassers will be tied up with piano wire... SORRY, SORRY!

http://upload.wikimedia.org/wikipedia/en/f/fb/Basil_Fawly.jpg

groups



keyword	LCSH	Match	Wikipedia	Match	keyword Authority	DODid	keywordID
Earls			http://en.wikipedia.org/wiki/Earls	exact	AAT	74465029	1048
Earth (soil)	sh85124396	close	http://en.wikipedia.org/wiki/Soil	close	AAT	74549258	3723
Earthworks (engineering works)	sh85040505	exact	http://en.wikipedia.org/wiki/Earthworks_(engineering)	exact	AAT	74546044	4844
Eating & drinking					TGMI	74548320	1055
Editors	http://id.loc.gov/				AAT	74546674	4352
Keyword	LCSH	Match	Wikipedia		Match	Originating voc	
Eating & drinking						TGMI	
Editors	sh85040976	close	http://en.wikipedia.org/wiki/Editing		broad	AAT	
Edwardian			http://en.wikipedia.org/wiki/Edwardian		close	AAT	
Egg	sh85041248	close	http://en.wikipedia.org/wiki/Egg_(biology)		close	AAT	
Elderly	tgm007221	close	http://en.wikipedia.org/wiki/Elderly		close	AAT	
Electricity	sh85042065	close	http://en.wikipedia.org/wiki/Electricity		close	AAT	
Embankments	sh85042664	exact	http://en.wikipedia.org/wiki/Embankments		exact	AAT	
Entertainment	sh96009616	broad	http://en.wikipedia.org/wiki/Entertainers	exact	TGMI	74548150	5191
Entrances	?		http://en.wikipedia.org/wiki/Entrance	exact	AAT	74549258	1117
Epaulets	-		http://en.wikipedia.org/wiki/Epaulette	exact	AAT	74546740	1118
Equestrians	sh85062154	close	http://en.wikipedia.org/wiki/Equestrianism	broad	AAT	74549594	1123
Equipment	sh85085299 ?	broad	http://en.wikipedia.org/wiki/Military_equipment	close	AAT	74545814	1124
Equipment & supplies	sh85085299	close	http://en.wikipedia.org/wiki/Military_equipment	broad	TGMI	74549442	3782
Ethnic groups	sh85045172	exact	http://en.wikipedia.org/wiki/Ethnic_group	exact	AAT	74548678	1139
Events	sh96009616	close	http://en.wikipedia.org/wiki/Competition	close	AAT	74547864	1148
Excavation (process)	sh85046104	broad	http://en.wikipedia.org/wiki/Excavation_(archaeology)	related	AAT	74549618	4850
Exhibiting	sh85046354	close	http://en.wikipedia.org/wiki/Exhibition	broad	AAT	74548718	5038
Exhibitions (events)	sh85046354	close	http://en.wikipedia.org/wiki/Exhibition	exact	AAT	74546188	1163
Explosions	sh85046465	exact	http://en.wikipedia.org/wiki/Explosion	close	AAT	74549252	3750

crowds

Which person is Old Fox?



Is Glencoe here?



Order to Capt. Campbell by Maj. Duncanson

You are hereby ordered to fall upon the rebels, the McDonalds of Glencoe, and put all to the sword under seventy. you are to have a speciall care that the old Fox and his sones doe upon no account escape your hands

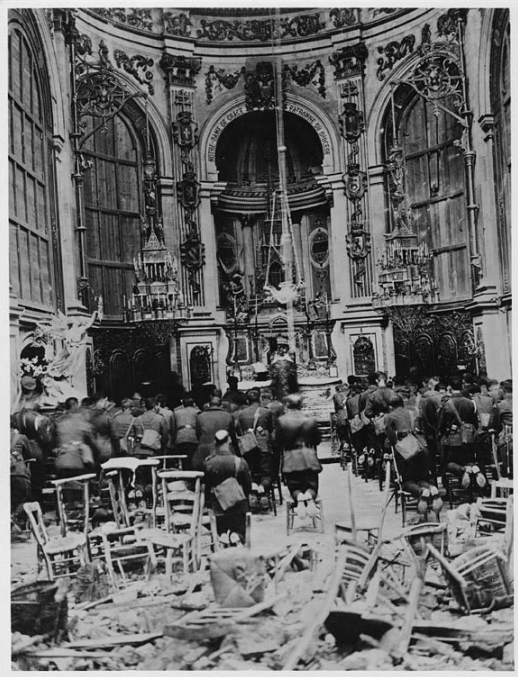
You are hereby ordered to fall upon the Rebels, the
all Donalds of Glencoe, and putt all to the sword under
seventy. you are to have a speciall care that the old
Fox and his sones doe upon no account escape your
hands. you are to have all the care that you
can take. This you are to putt in execution
with a strong party: if I doe not come to you
all sones, you are not to tary for me, butt to fall on.
This is by the Kings speciall command, for the good
safety of the Country, that thos miscreants be putt
off root and branch. See that this be putt in execution
without fauld or falshood, also you may expect to
deall with us one not to tary for the King nor Government,
nor a man fitt to ease Commissioners in the Kings
service. Expecting you will not faill in this
filling thereof, as you tell your selfe, I subscribe
these with my hand at Salisbury Feb. 12, 1692
For the Kings service
Robert Campbell
of Glencoe.

crowds & geonames



We think this is
Cambrai ...

Do you think
this Cambrai is
here?



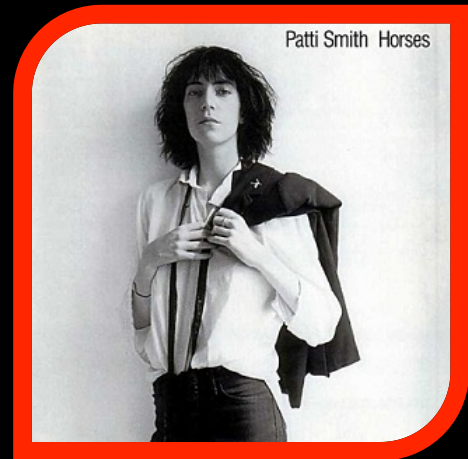
Or do you think
it's here?



crowds & dbPedia



Is the horse in
this picture ...



none of
these

crowds & LCSH



Would you describe this horse in any of these ways?

- Show jumpers (horses)
- Horses in motion pictures
- Toy Horses
- Horses
- War horses
- Travel with horses



none of these

influences - when 1 loch becomes 2

Loch of Harray

The **Loch of Stenness** together with the **Loch of Harray** are the two largest freshwater lochs of **Mainland, Orkney**. Their names are *Steinnesvatn* and *Heraðvatn*, the nearby parishes of **Harray** and **Stenness**.

From Wikipedia, the free encyclopedia

Coordinates:  59°1′24″N 3°13′40″W﻿ / ﻿﻿ / ﻿

The **Loch of Harray** is the largest loch of **Mainland Orkney**^[1] and is named for the nearby parish of **Harray**. It lies immediately to the north of the **Loch of Stenness** and is close to the **World Heritage** neolithic sites of the **Stones of Stenness** and **Ring of Brodgar**.^[3] In **Old Norse** its name was *Heraðvatn*.^[4]

Hydrology [edit source | edit beta]

The loch was surveyed^[1] on 21 August 1903 by **Sir John Murray** and later charted^[5] as part of the **Bathymetrical Survey of Freshwater Lochs of Scotland 1897-1909**. Murray observed that **Loch of Harray** is a freshwater loch, the largest in all Orkney with an area of

Loch of Harray

Harray loch



Loch of Harray



Loch of Stenness

From Wikipedia, the free encyclopedia

Coordinates:  59°00′N 3°15′W﻿ / ﻿﻿ / ﻿

The **Loch of Stenness** is a large **brackish** loch on **Mainland, Orkney**^[2] and is named for the parish of **Stenness**. It is 2 miles (3.2 km) northeast of the town of **Stromness**, lies immediately to the south of the **Loch of Harray** and is close to the **World Heritage** neolithic sites of the **Stones of Stenness** and **Ring of Brodgar**.^[3] In **Old Norse** its name was *Steinnesvatn*.^[4]

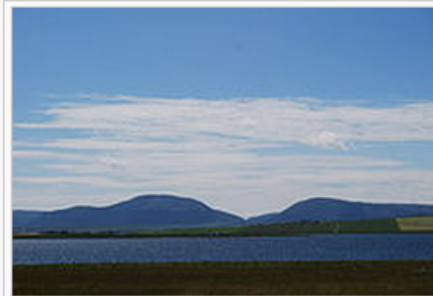
Contents [show]

Hydrography [edit source | edit beta]

The **Loch of Stenness** is a sea loch and is the deepest loch on the **Mainland**, it is slightly smaller in area and volume than the **Loch of Harray**. It is located at 3.25 miles (5.23 km) northeast of

Loch of Stenness

Stenness loch



Over Loch of Stenness towards Hoy - geograph.org.uk - 1385582



The Elusive Loch of Shandon



discussion

- using a voc without URIs?
 - should we change?
- are there good ways to string match?
 - are they trustworthy?
- are crowds helpful?
- what vocs are mapped to what other vocs?
 - can/should we help map vocs beyond our domain?
- should we alter the global to match the local?