Case Study: Embedding *A Vision of Britain through Time* as a resource for academic research and learning

JISC e-Content and Digitisation Programmes: Impact and Embedding of Digitised Resources

Paula Aucott, Richard Healey, Humphrey Southall

March 2011

1. Background

From 1994 onwards, the Great Britain Historical GIS established itself as a worldleading historical GIS project. Since receiving National Lottery funding in 2001, it has also developed as one of the UK's leading digitisation projects. Although the initial focus was on historical statistics from diverse sources, including the full run of the Census of Population 1801 to 2001; and mapping as vectors the changing boundaries of major statistical reporting units, it now also includes scanned historical maps and travel writing. The entire collection is tightly integrated by place and location. This data collection has been made freely available via the web site *A Vision of Britain through Time,* which was originally built with national lottery funding. The site was launched in 2004 and modified and re-launched with JISC funding in 2009. Usage has grown steadily and continuously, and in January 2011 it reached its highest recorded peak yet, with over 142,500 unique visitors.

The website effectively meets the needs of "life-long learners" interested in local history. Previously it has been impossible to exactly measure "academic use" as there are no access controls, but e-mails and other contacts show that many academics with primarily local interests benefit from our integration of data for their study area. The current JISC project is aimed at redressing some of the balance by providing a focussed facility for academic users to access our data directly, as they have frequently requested over the years. The other outcome to the project is a complete analysis of the existing usage of the site, conveyed via an impact report¹. This makes interesting reading and goes some way to supporting what we believe are our user requirements in the usage trends seen therein.

2. Findings

Outlined in the rapid impact report is the extensive usage being experienced by the website. Since its initial launch the website has continued to attract new users, compare reported figures of just under 40,000 in November 2004 with the 100,781 unique visitors we received in November 2010. Its popularity stems from its "findability" and its ease of use:

"The main site itself seems full of goodies, and I think I'm only just starting to scratch the surface, so keep up the good work!" (User, July 2009)

"Very good website... The easiest to search website I have ever been on." (User, May 2009)

"The well-designed site is easy to use and clearly labelled...A Vision of Britain won't help you find ancestors, but it will put where they lived in a historical context." (P. Christian, 'Site test: Back to our roots' in The Sunday Times, 15th January 2006)

The wealth of different information we have combined together in a single place;

¹ Aucott, P., Southall, H. & Healey, R. (2011) *Impact Report on 'A Vision of Britain through Time' 2004-10. Investigating the current use and impact of a popular digital resource for local history research* <u>http://www.jisc.ac.uk/media/documents/programmes/digitisation/EVoB%20Impact%20Report%20Feb%2020</u> <u>11.pdf</u>

"Well, what can I say but that the *A Vision of Britain through time* website is a wonderful creation, and with the datasets it includes, makes for most interesting viewings of various places." (User, October 2010)

"I greatly value your service. Without it, I would have had no way of finding the information I needed." (User, May 2009)

"I think that the website is the best I've seen in giving information on both the historic and modern administrative names for different places." (User, May 2005)

And digital access to specific historical material for the first time;

"I have been using the Vision of Britain website and finding it really useful, thanks for putting together such a useful source... Your website offers a lot of useful stats." (Liverpool John Moores University User, February 2011)

"I have been using your site recently and have found it excellent for statistical data from the census" (PhD Student, July 2009)

"I want to congratulate you on the splendid work you've done. This semester I'm teaching a course on the literature of the British countryside, and imagine my delight to find how easily my students can locate on-line excerpts from Defoe's Tour of Great Britain, Young's Annals of Agriculture, and Cobbett's Rural Rides. Without your work, this course would be much poorer in content." (USA University User, January 2009)

The steadily growing user base will generally enter the site through an internet search engine, usually based on a search for a place name. This shows the successful impact of our architecture and data ontology structure in taking users direct to the information they want. The data ontology is a framework to hold information about geographical entities and their relationships over time. A gradual growth in the repeat visitor numbers is visible, rising from just over 18% in 2007-8 to 22.13% in 2009-10. Also highlighted is the international audience the site caters to in addition to the British traffic, although most come from English speaking (en) countries.

Language	Visits	Visits	Visits
en-us	1,331,071	78.78%	
en-gb	276,243	16.35%	16.35%
en	30,943	1.83%	
de	8,982	0.53%	
fr	7,095	0.42%	
ni	4,131	0.24%	78.78%
es	3,619	0.21%	
it	3,125	0.18%	
📕 pl	2,933	0.17%	
🔲 ru	2,479	0.15%	

Figure 1: Top 10 languages: January 1st 2010 – February 28th 2011

Perhaps the most worrying aspect of the usage analysis was the lack of depth of visits being achieved. The vast majority of users only stay a very short time on the website. In theory this is a positive outcome as it may indicate users find the information they require in a single page-view resulting from improved data presentation and structure. However, it also suggests we are not reaching the academic audience we would like to and that they are not exploiting the data held within the system to its full potential.

Figure 2: Depth of visit: January 1st 2010 – February 28th 2011

www.visionofbritain.org.uk Depth of Visit

Jan 1, 2010 - Feb 28, 2011 Comparing to: Site

Pageviews in the visit	Visits with this many pageviews	Percentage of all visits
1 pageviews	760,128.00	50.03%
2 pageviews	176,961.00	11.65%
3 pageviews	180,293.00	11.87%
4 pageviews	66,750.00	4.39%
5 pageviews	77,435.00	5.10%
6 pageviews	37,369.00	2.48%
7 pageviews	42,277.00	2.78%
8 pageviews	22,876.00	1.51%
9 pageviews	25,465.00	1.68%
10 pageviews	15,326.00	1.01%
11 pageviews	16,427.00	1.08%
12 pageviews	10,753.00	0.71%
13 pageviews	11,376.00	0.75%
14 pageviews	7,545.00	0.50%
15 pageviews	7,974.00	0.52%
16 pageviews	5,554.00	0.37%
17 pageviews	5,943.00	0.39%
18 pageviews	4,298.00	0.28%
19 pageviews	4,370.00	0.29%
20+ pageviews	40,131.00	2.64%

Most visits tracked: 1 pageviews

The report also details the potential to generate self-sustaining income and the challenges faced in setting up new methods of income generation with their relative merits and weaknesses. We generate revenue through co-marketing partnerships, licensing of data to commercial companies and via Google Ads running on the website. Income raised goes directly towards maintaining the server running the site. By monitoring usage and income generation together we can focus on improving

areas of the website that are the most lucrative and thereby encouraging users to return and potentially increase revenue further. The following graph shows the general upward trend in our revenue since the dual monitoring was introduced.

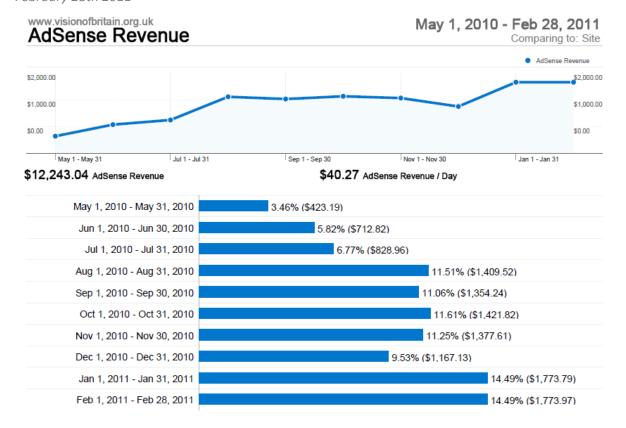


Figure 3: Revenue recorded since Google Analytics began recording AdSense data: May 1st 2010 – February 28th 2011

A lack of citations revealed in qualitative analysis raised concern, as did the apparent lack of prominence amongst academic circles highlighted in the academic user interviews. There is awareness of our data in certain academic communities. The evidence from UKBorders at Edina given in the impact report substantiates the view that many academics do know about, and in some cases judging by the download figures must use in teaching, our data. However, they are not relating this information to that held within *A Vision of Britain through Time* and using the data held there to the same extent. By providing a useful new resource specifically aimed at academic users and targeting our publicity at them, we hope to encourage greater use by this audience.

2.1 Content Analysis Findings

Content Analysis using Google News search found no articles which included hyperlinks to the Vision of Britain url and a search of the News Archive revealed very few (See Appendix 1). We do know some articles were not being picked up by this analysis, for example an Observer article by Juliette Jowit, published on the 24th October 2004: <u>http://www.guardian.co.uk/uk/2004/oct/24/britishidentity.population</u>. Links to different blogs identified through the Google blog search totalled 120. These

range from posts actually about the site, through mentions of the site in passing, to posts where just a link is used to illustrate a point made by the author. Notably, of those directly about the site, only 5 relate to the time around the re-launch and 2 to the initial launch. Altogether there were 3 regional radio and 1 national television interview plus 18 news stories covering the initial launch, and at the re-launch there were a further five radio interviews and a regional TV feature. The relative paucity of hits in the news archive is of concern considering the efforts that went into publicity around the launch. On the other hand it may be unsurprising considering the resource we offer is a data website, therefore we cannot expect mainstream media to be overly interested when it is updated as this is the expected norm for digital data.

A Search of Thomson Reuters Web of Knowledge, which acts as a platform to many scientific journal and citation references, revealed no direct citations for the website and only two references to "The Great Britain Historical GIS". This perhaps reflects a tendency to reference scholarly publications rather than websites in the material searched. Google Scholar was more forthcoming in references to the website. However, by searching for "The Great Britain Historical GIS" (i.e. the project behind the website) Google Scholar scored much better with a total of 105 references, of which 8 were citations.

Collectively the content analysis findings highlighted the lack of recognition the website has in our own field. By being better embedded within multiple disciplines we become more aligned to the priorities of our institution and thus increase the income generation opportunities to contribute to our long-term sustainability. Focused publicity in the academic sector should also help improve this situation. Altogether these changes should enable us to be at the forefront of work in the field of Historical GIS, rather than just having a select few recognize the importance of our output.

3. Resource Enhancement

By far the largest number of requests from users we have received over the years is for a download facility, whether it be for historical maps, statistics or boundaries. These requests show what academics actually use within the site and consequently directed our thinking on what to provide as part of our first download facility.

"I would also like to integrate data published in the Statistical Atlas 'A vision of Britain through time' (population, industry, social structure etc.)"... (PhD student, September 2010)

"Could you tell me if your population data is available in downloadable format please? We're doing some research into population shifts and economic geography, and so spatial population data from 1801 would be very helpful." (User, April 2010)

"It'd be really helpful, though, if the stuff that's available on the Vision of Britain site actually had a button to download the underlying dataset... plus it would widen access significantly." (University of Birmingham User, December 2009)

"I was wondering whether the Dudley Stamp map is available in digitised geo-registered GIS raster [rather] than just simply jpeg tiles (as appears in DIGIMAP)." (University of Reading User)

Offering a range of possible types of data download will, we hope, greatly improve accessibility to our material. Comments from a user survey run before the website re-launch also highlighted the failure to meet the standards of mapping expected by GIS professionals, academics and students. Although the user interface has been improved since that survey, the critical comparison in GIS terms is still relevant;

"The user interface for browsing around maps is pretty awful by comparison with similar features in other applications and websites" (User, May 2009)

"I find problematic the fact that where maps have historical and current boundaries, there's no way to tell which is which. A lot of my research benefits from knowing how boundaries have changed through history." (User, May 2009)

Providing a download facility is in itself a major achievement, but before this could happen significant work had to go into improving the background information that would be available via the download site. On occasion errors within the boundaries of units in the site are picked out by users;

"The changing boundaries of Meppershall Beds being not as I know them to have been [it's OK nobody who hasn't done in depth research knows them" (User, May 2009)

Corrections and improvements to this boundary data and in matching up unit information with the statistical data for the units have been made. This will manifest its usefulness in the new download and statistical search facilities and the redeposits of data at UKBorders and the UK Data Archive. It will also increase the overall efficiency of the statistical reporting in the site and, together with the improved boundary mapping, will lead to better statistical mapping being visible on the website as well.

3.1 User Survey 2011

In March 2011 we undertook a further user survey to assess current opinion of the revised website. We obtained opinions by directly requesting them from members of our mailing list (15 respondents, hereafter referred to as 'invited'). Separately we canvassed opinion through a general request for a short period on the website itself (304 respondents, hereafter referred to as 'spontaneous'). Appendix 2 gives the survey questions.

The majority of 'invited' users had been using the site for several years and mostly found it through directed advice, whereas most 'spontaneous' users were recent, arriving via a Google search. Despite this the three most important sections of the website used for both groups were the same; place searching, followed by historical maps then census reports. For the 'spontaneous' users learning resources came next and this suggests we should continue to develop this area to help them.

In the 'spontaneous' group a third of users classified their main interest as family history/genealogy followed by local history and they mainly use the website to research these topics. The 'invited' users had much more specific agendas, using the website to provide data, as background, as a teaching aid and for interest. Almost three quarters of both groups had not cited us and there was great variety in the responses of those that had. This indicates we should endeavour to make acknowledgement and copyright notices much clearer on the website. It is pertinent to note we already planned to automatically include these details as part of the data downloads.

Despite place searching being the most important feature for both groups the majority (60% 'invited'/77% 'spontaneous') had not noticed any changes recently, suggesting the improvements implemented had gone very smoothly. When asked about the facilities currently being created, over half of all respondents thought they would use boundary downloads. A smaller proportion said they would use map downloads and an even smaller group wanted the statistical downloads. One respondent actually described our design concept exactly; "Enter place and year information to find results and historical trends".

Comparison between our website and known equivalents weighed in our favour. Of the 'invited' group 13% use our site weekly, 27% monthly and 60% now and again. The 'spontaneous' group use it 4% daily, 10% weekly, 6% monthly and 37% now and again. Only Wikipedia achieved a higher daily usage response from the 'spontaneous' group. Overall the most regular response for all websites for both groups was "now and again". We also asked which other sites they used to complement our data. Perhaps not surprisingly the 'invited' group used mostly reference sites, such as Histpop, The National Archives Catalogue and Heritage Gateway whilst the 'spontaneous' group used Genealogical sites, followed by data, general and archive sites.

Type of Website	Count
Genealogy	139
Data	36
General	28
Archives	26
Maps	17
Local History	14
Search	11
Publications	8
General History	5
Images	5
Discussion	3
Libraries	3
Video	1
Grand Total	296

Figure 4: Complementary website references from 'spontaneous' respondents to the 2011 user survey

Less than half of either group were involved in teaching or classroom learning in any way. Of those that were, the 'invited' group mainly used the website in assignments and class, whilst the 'spontaneous' group used it mainly for background research. Only 7 people had ever used our data from the national data repositories, confirming the supposition that those academics using our data are mainly not using the website as well.

Suggestions for improvement included direct data and Shapefile download, adding more maps and data especially pre-1800 material and for smaller places. They also proposed incorporating new types of content, such as photographs, historical pictures and parish church information. Usability suggestions included a search box visible at all times, further links to other relevant websites, a means of sharing information and more multi-media. For example;

"...this is the kind of digital resource I see could be useful for school and older learners. More interactive and problem-solving/reflective learning resources are needed though. Currently the site is rather passive and more of an encyclopaedia (which is still, nevertheless, a good thing)."

Many were grateful for the great effort put into creating and maintaining the website, thankful for its free access and wanted the results as widely disseminated as possible;

"This is a useful site and I recommend it to students in HE as part of their library induction"

"This is probably the best site for local information that I have come across and I feel you should encourage people who are working on their family history to use it"

4. Challenges and Solutions

We wanted to give users, and in particular academics, the ability to download some of our data direct from the website. However, the sheer volume of traffic meant the number of users given this capability had to be limited in some way. The solution was the implementation of Shibboleth, a software access system designed to allow only members of UK Education Institutions access to data.

This task was originally assigned to the project team, but for project management reasons we later decided to outsource the specific technical task of the installation of Shibboleth to an external contractor. This decision was complicated by the fact the University already had Shibboleth installed in its identity mode, but not as a content provider. In other words, the University already operated services providing access to content held at other institutions via the Shibboleth log-in, but the software was not installed to enable it to provide locally sourced content. Although various individuals within the institution had already expressed an interest in using Shibboleth to become content providers, our project was the first to actually request it. This raised various issues, including the revelation that the identity provider software currently being used was provided by an external contractor. It took some time to determine

that we could in fact install the content provider software independently on a separate machine within the University and therefore we could use another contractor. In future any project requiring software installation will be assessed in terms of what the institution already provides and what contractual agreements exist prior to the commencement of a funded project.

The most challenging part of the impact report was getting specific feedback from academics. Although we contacted several academic users, only two responded positively to a request to discuss their use of the site. This led to a much reduced capacity to generalize on views held by academics about our resource.

5. Outcomes

By providing a facility that enables academic users to download their own copy of our scanned map collection and vector boundary files we believe the research community has been given a significant tool for progressing their personal research. By accompanying that with a specific end user license and relevant metadata we also make it easy for the user to know to whom and how they should make acknowledgements.

The statistical search facility provides a technically innovative tool which allows users for the first time to search for statistics by location, using their own specified parameters, rather than being constrained by the limits imposed by the original statistical table. Providing such a tool will enable us firstly to see how useful it is per se, but also to assess its likely usefulness to the wider community if it were to be released on a more general scale. Initial usage will allow us to assess what functionality needs to be altered, added or removed to improve the successfulness of the service being offered.

Incorporating the improved statistical mapping within the main body of the site, inserted as a byproduct of the vector boundary creation process, and the limited number of vector download files available to all mean the entirety of our user base will benefit from these improvements to the site. The re-deposit of our updated vector material in UKBorders at Edina will also improve our standing in the academic community as more academics become aware of our material.

6. Further Developments

The compilation of the impact report has highlighted areas for future development within the website. Despite plenty of references to the Great Britain Historical GIS Project, there are very few citation references to *A Vision of Britain through Time* itself. It would perhaps be advantageous to create a downloadable citation format, with different formats for different purposes. For example, one format for Wikipedia, others for more traditional formats like the Oxford, Harvard or Chicago citation styles. We suspect having clear stable citation information available will encourage teachers and researchers to use the site more.

Another area to develop will be to try to address the lack of recognition of the benefits of the website in academic circles by generating a more focused communication plan with the audience we wish to attract. As part of this strategy we are in the process of setting up a blog for the project which will operate in place of the current "news" item on the website. This will allow users to comment on adaptations we make as we announce them and they begin to be used.

7. Conclusion

By undertaking a full review of the impact of the website from its initial publication until October 2010 as discussed in the impact report, we have created a baseline analysis to provide easy comparison for the future. It allows us to monitor trends in usage of the website once the download facilities are in place. This will enable twofold tracking, giving us a better picture of academic use and revealing how useful the specific facilities actually are.

Implementing the new services will allow us the opportunity to focus on publicizing the benefits of using our website to the academic community. Thus hopefully the use of our material will increase and as a by-product of that, the number of citations to our work. Of course the traditional method of publishing our own articles in the discipline will also help in this respect and we are beginning to see some of these in press.

The download facility will provide access to our scanned images for home use for the first time. It will allow academics to access vector data through our website rather than Edina, although of course all the updated boundaries will also be available through them. For the first time we will offer vector downloads of county boundaries for all our users. The statistical search facility offers a way for academics to choose their geographical area of interest and select statistics solely for that area, rather than having to download whole national datasets and sift them for the relevant data. By enabling this kind of search and retrieval facility we are offering an improvement in efficiency of data selection. The statistical search facility is viewed as a proof of concept and initial version. Once it has been in put into operation we will be able to monitor it to assess its impact and success and investigate methods to improve its usefulness in the future.

Appendix 1: Content Analysis

The Guardian published two articles, one about the initial launch of the site in 2004 and another concerning the fight to free data from May 2006.

• Ye olde bloggers find plot in cyberspace by Polly Curtis, education correspondent, Tuesday 26 October 2004:

http://www.guardian.co.uk/technology/2004/oct/26/highereducation.elearning

• A Sidestep in the right direction by SA Mathieson, Thursday 11 May 2006:

http://www.guardian.co.uk/technology/2006/may/11/copyright.epublic

There are also two Guardian bookmarks to specific pages, one for Cardiff the other for Leeds, plus an article in the Liverpool Echo more recently.

• Chance access a slice of family history by Martin Rigby, Liverpool Echo, Feb 21 2009:

http://www.liverpoolecho.co.uk/views/liverpool-columnists/echocolumnists/2009/02/21/chance-access-a-slice-of-family-history-100252-22974331/

For the re-launch we focused our publicity strategy on informing networks of local historians through print and digital means. Perhaps most revealing is that despite the apparent lack of large-scale success in mainstream media suggested by the lack of digital news items, the site has still managed to increase its general popularity. This in part is due to targeted marketing, and linkage with other general websites such as Wikipedia and specific websites such as GenUKI, as well as the improved system architecture. In future we will try to target publicity at specific user groups, as we did during the re-launch, to inform relevant sub-groups as we develop elements of the website that suit their particular needs and encourage them to test the features and hopefully give feedback on the results.

Appendix 2: User Survey 2011 Questions

Questions marked * required a response in order to proceed.

1. * How would you describe your main role? (N.B. Please tick all that apply)

Lecturer – Further Education (FE)	Student – School		
Lecturer – Higher Education (HE)	Student – Other (as part of taught course)		
Teacher – School	Archivist		
Teacher/Tutor - Other	Librarian		
Researcher – Professional/HE/FE	Map Librarian/Cartographer		
Student - Research postgraduate	Non academic - Family History		
Student – Taught postgraduate	Non academic - Local History		
Student – Undergraduate	None of the above		
Student – Further Education/Foundation Education Level			

2. * Which academic discipline are you based in? (If 'Other' selected please define)

•	
Environmental science	Information Systems
Geography	Population Studies
Health care/Medical research	Other
History	l am not based in an academic discipline

3. * Where are you based geographically?UKAustraliaNew ZealandRepublic of IrelandCanadaUSAOther European Union (EU)/European EconomicRest of the WorldCommunity (EEC) CountryKealand

4. * Roughly how long have you been using A Vision of Britain through Time?

Less than 1 month	7-12 months	3-4 years
1-3 months	1-2 years	4-5 years
4-6 months	2-3 years	5+ years

5. * How did you first find A Vision of Britain through Time? (If 'Link from other website' selected please give name of other website (N.B. Please do not use full url addresses as these will be eliminated by the spam filter))

Electronic mailing list	Google Search
Newsletter	Other internet search engine search
Recommended by someone I know	Cannot remember
Link from other website	

6. * Which areas of A Vision of Britain through Time do you use? (N.B. Please tick all that apply)

Place searching	Learning resources	Expert search
Historical maps	Statistical atlas	Other
Census reports	Travel Writing	

7. * Which area(s) of the website are most important to you? (N.B. Please select up to 3)
Place searching Statistical atlas All
Historical maps Travel Writing None
Census reports Expert search
Learning resources Other

8. What is/are your main area(s) of research interest?

9. What do you use our website/data for?

10.* Have you cited our data in any of your own work? (If 'Yes' selected please describe how)

Yes / No

11.* Have you noticed any changes to place searching recently? *Yes / No*

12.* Would a facility to download historical boundary map files for your own research be useful?

Yes / Maybe / Don't know / No

- 13. What should a statistical search system in A Vision of Britain through Time let you do?
- 14.* Do you have any major issues with how the current site works? (If 'Yes' selected please describe)

Yes / No

15.* Do you have any suggestions on how the site might be improved? (If 'Yes' selected please describe)
Yes / No

Yes / No

TO. THOW ORIGINAL	<u> </u>		ing aignair	00001000	•		
	Daily	Weekly	Monthly	Now	Rarely	Single	Never
				and		Visit	
				Again			
A Vision of							
Britain through							
Time							
Histpop							
British History							
Online							
OldBailey							
Online							
EDINA							
services (UK							
HE)							
UK Data							
Archive							
services (UK							
HE)							
Genuki							
Wikipedia							

16.* How often do you use the following digital resources?

17.* Do you use digital resources differently from how you use original documents? (If 'Yes' selected please describe how)Yes / No

18. Please name any other digital resources you use to complement our data (N.B. Please do not use full url addresses as these will be eliminated by the spam filter)

19.* Do you do any teaching? Yes / No

- 20. If you could download scans of historic maps for use on your own pc, would this encourage you to use the website in your teaching/get your students to use it? Yes / Maybe / Don't know / No
- 21. If you could download historic boundary map files for use on your own pc, would this encourage you to use the website in your teaching/get your students to use it?

Yes / Maybe / Don't know / No

22. If you could download historical statistics for a specific geographical area for use on your own pc, would this encourage you to use the website in your teaching/get your students to use it?

Yes / Maybe / Don't know / No

- 23.* If you are involved in classroom work (as either the teacher or the student), is A Vision of Britain through Time used live in class, as part of background learning/research or in assignments for students? (N.B. Please tick all that apply) In class / Background learning or research / Assignments / None I am not involved in classroom work
- 24.* Have you ever downloaded any of our data (i.e. deposited by the Great Britain Historical GIS Project team) from the national repositories at UKBorders at EDINA or the UK Data Archive (formerly the History Data Service)? Both / UK Data Archive / UKBorders / Neither
- 25. If you have downloaded our data, which datasets did you have? Did you have any problems with the data itself?
- 26. Have you had to make any adjustments to our data? (If 'Yes' selected please describe)
- 27. If you have any other comments about the website please make them here