Rapid Impact Analysis

The Impact and Embedding of an Established Resource: British History Online as a Case Study

Grant 7/10. JISC e-Content and Digitisation Programmes: Impact and Embedding of Digitised Resources

Jonathan Blaney and Peter Webster, Institute of Historical Research, November 2010

Overview

In 2010 the JISC funded the Institute of Historical Research to conduct a rapid impact analysis of British History Online (BHO), the digital library of core sources for the medieval and modern history of Great Britain. This report details the findings of that analysis. Since the project's inception in 2002, the BHO team have conducted successive enquiries into the impact of digital scholarly resources in general, and BHO in particular. Three interview-based enquiries have been undertaken: in 2002, 2005, and in 2009-10 as part of the JISC-funded Connected Histories project. These successive longitudinal studies have identified clear shifts in research practice in history, and shown the impact of BHO in certain fields. They were, however, explicitly limited to research practice, and were not concerned with the use of BHO in teaching and learning. That gap is remedied by this report, which provides answers to the following questions:

(i) how is BHO currently used in university-sector teaching in the UK? Is it used for teaching to a greater or lesser extent than it is for research?

(ii) which new functions, or improvements to existing ones, would be most welcomed by teachers and learners, and thus aid the greater embedding of BHO in teaching practice?

(iii) which new functions would be most welcomed by university librarians?

(iv) can the impact of BHO in research, about which the project team already knows a great deal, be demonstrated in a clearer and more statistically grounded way?

Methodology

In accordance with the Toolkit for the Impact of Digitised Scholarly Resources (TIDSR), the following activities were undertaken. For reasons of space, we will give a very brief summary of the outcome of each in this methodology section, reserving general comment for later in the report. The quantitative measures were adopted since BHO, as a mature project, has a comprehensive set of data dating back several years. Amongst the qualitative measures, the team exploited the unrivalled connections and national reputation of the IHR to gain access to participants for both telephone and face-to-face interviews, for the two focus groups and for the survey of librarians.

Qualitative	Quantitative
Focus groups	Webometrics
User surveys	Analytics
Interviews	Log files
Site feedback	Bibliometrics
Referrer analysis	Content analysis

Findings

Qualitative activities

(i) two **focus groups**: one with university teachers, and one with graduate research students. Both groups expressed strong preferences in favour of enhanced tagging facilities on the site; the graduate group also expressed interest in screencasts.

(ii) an online **survey** of users, which received 969 responses. Of those respondents who were university-based, just short of half were from the UK, with 20% from north America and 10% from Europe. A third were academic staff, 40% graduate students and 12% undergraduates. As well as this rich data on our users and their preferences, the use of 10 different collectors on the survey has enabled us to track our most efficient means of communicating with our users.

(iii) an online **survey** of university history librarians. There were 38 responses to this survey, which as a percentage of UK History subject librarians is significant.

(iv) a series of **stakeholder interviews** with practising historians, adopting the same format as the last iteration of the benchmarking study of 2009-10, but extended to include teaching. We carried out seven in-depth interviews with academics with teaching experience and a range of seniority and career stages; we were able to extend our previous benchmarking studies and add a new section on teaching.

(v) an analysis of the historical data from the site **user feedback** function, summarised below:

Total feedback counted	2012	%
General historical query	498	25%
Subscriptions	384	19%
Site content	303	15%
Site structure or functions	237	12%
Non-subscription access	215	11%
Reproduction rights	132	7%
Other	256	11%

(vi) an analysis of inbound referrers, **referrer analysis**, summarised below (*source: Google Analytics*):

Academic Referrers	Evidence of course use	Academic library references
(top) 500	22	181

This represents the academic referrers for the period October 2009 to September 2010, for a minimum of three referrals. Strings indicating VLE software (eg Moodle) provided evidence of use in courses.

Quantitative activities

(i) **webometrics** *(source: LexiURL).* BHO performs very well in comparison with sites containing similar content. Old Bailey Online has very similar figures, but BHO comfortably outperforms the others on all measures listed.

Site	URLs	Domains	Sites	STLDs	TLDs
British History Online	998	841	805	47	35
Old Bailey Online	988	843	788	47	32
Parliamentary Papers	395	276	234	31	21
Clergy Database	315	248	230	29	20
England's Past for Everyone	297	238	238	20	15
TNA Documents Online	621	512	512	42	32
Tanner Ritchie	92	79	74	15	13
Colonial Papers	39	28	28	8	7

(ii) **analytics** *(source: Google Analytics)*

Period	Pages	Visits	% referrals from search	% referrals other sites	% referrals direct	%UK	%Non-UK
1-10-2009 - 30-9-2010	17,191,624	3,955,260	81.2	11.4	7.4	70.59	29.41
1-10-2008 - 30-9-2019	15,849,717	3,757,259	83.5	10	6.5	69.70	30.30
1-10-2007 - 30-9-2018	15,041,729	3,462,509	84.8	9.8	5.4	70.69	29.31

Further investigation into analytics, including keywords, will be included in the final report.

(iii) **log file** analysis *(source: IIS extended format logs).* Here we compared usage across groups, in an effort to learn more about user interest in content. We calculated a standard deviation on the per-volume usage in each group, to see if that would point to underused resources. The idea was that groups with a high standard deviation would point to low-usage within the same subject area as high-usage volumes, and qualitative analysis could then be applied to obtain a more finely nuanced view of user preferences; this would then inform better content selection in future. Again, the final report will provide deeper analysis of this issue. As an example, here is the usage of the volumes in the group 'Monastic and Cathedral Records':

Publication title	URLs	Total views	Average by URL
Additional material for the history of the Grey Friars, London	10	2983	98
Annales Cestrienses	12	4648	125

The Grey Friars of London	32	9458	96
London and Middlesex Chantry Certificate 1548	8	3739	151
Register & Records of Holm Cultram	82	31480	123
Registrum Statutorum et Consuetudinum Ecclesiae Cathedralis Sancti Pauli Londiniensis	82	9070	36
Staffordshire Historical Collections, vol. 4	37	19725	173
Staffordshire Historical Collections, vol. 5 part 1	22	8490	126
Staffordshire Historical Collections, vol. 6 part 1	32	9616	97
Staffordshire Historical Collections, vol. 11	50	12874	84
The Cartulary of Holy Trinity, Aldgate	99	7723	25
The Ledger Book of Vale Royal Abbey	17	6415	123
Westminster Abbey Charters, 1066 - c.1214	28	5128	59

(iv) **bibliometrics:** Scopus and Google Scholar were consulted for journal citations of British History Online in 2010. Scopus returned 14 results and Google Scholar returned 17 results. Of these, seven were common to both sources. The final report will extend the search period to cover citations from 2008 forwards. Scopus may be omitted.

(v) **content analysis**: we tracked references to British History Online, alongside the comparator sites, in blog posts for the period June to November 2010 (*Source: Nielsen BlogPulse*)

Resource	URL	Recent blog mentions
British History Online	www.british-history.ac.uk	84
Old Bailey Online	www.oldbaileyonline.org	38
Clergy Database	www.theclergydatabase.org.uk	5
England's Past for Everyone	www.englandspastforeveryone.org.uk	0
Ancestry.co.uk	www.ancestry.co.uk	55
Tanner Ritchie	www.nationalarchives.gov.uk/documentsonline	20
The National Archives	www.tannerritchie.com	0
Parliamentary Papers	parlipapers.chadwyck.co.uk	1
Colonial Paper	colonial.chadwyck.com	0
Early English Books Online	eebo.chadwyck.com	11

It is notable here that that the second highest number of mentions is for Ancestry.co.uk, the biggest genealogical site in the UK, but that BHO is nevertheless attracting almost double the number of mentions as it.

Impact on research

It is the team's intention to publish, under separate cover, an article detailing the findings of a longitudinal study of the impact of digital resources on historical research in general, stretching back to the inception of BHO in 2002. As such, a fuller analysis of the impact of BHO on research as revealed by the qualitative aspects of the present study will be incorporated into that publication. Consequently, only some summary findings are presented here.

Some 26% of respondents to the general survey said that BHO was "very important" to their research, and 27% "quite important". If undergraduate students and users outside the professional academic context are excluded, nearly two-thirds describe the service as very or quite important, with only 2% thinking it not at all important.

There were also some indications of the particular uses to which BHO is being put. Amongst the academic users, the spread of tasks was wide: from resource discovery activity at the beginning of a research project ('finding new works') to the consultation of known works and tasks associated with writing ('checking references.') Over a third used the site most often for general searching: a trend towards theme- rather than source-driven and search-enabled research practice that will be drawn out further in the later study.

Part of the user survey asked respondents to describe whether and how BHO had changed their research. Exactly half considered that BHO had indeed changed the way they did their research, and the responses as to why matched very closely the findings of the previous longitudinal study. They concerned not only gains in speed and efficiency, but also shifts in the manner in which research is carried out.

Most predictably, many users welcomed the fact that sources were available on their desktop, yielding an obvious gain in travel time and expense. Several users also reported that they now had access to sources, the printed versions of which their home libraries did not possess; this was particularly the case for users located overseas.

The search function also led many users to approach sources in a different way. Users reported being able to save considerable time in searching resources that previously necessitated a more lengthy reading, often looking for a limited number of scattered references. This was particularly the case in sources which were originally printed without adequate indexes.

It was also the case that searches across the whole site now produced 'unexpected treasures'; whole sources not previously known to the user, or unanticipated results in known sources. The findings for standard deviations within groups did indicate that some resources may be underutilised. Further investigation of these figures will be included in the final report.

There were also some indications that online access with search was beginning to change the order in which research was carried out, and the relative weight given to sources. One user thought that they were probably now prioritising the digitised source over the printed one. Another thought that it was now possible, at or near the writing stage, quickly to check sources for supplementary references; sources that would not have been consulted at all previously. Another, preparing an edition of correspondence, was more likely to make the attempt to fill in more of the detail surrounding that edition, given that speculative enquiries could now be carried out without a great investment of time.

In general, the responses were enthusiastic about BHO; one particular comment is perhaps worth quoting in full:

> British History Online is my favorite and first source for primary sources in British history. As a student of history, librarian, and writer, I return again and again. Even when I'm not researching, I often visit BHO for the sheer fun of what I might learn and discover. The site is easy to navigate, convenient, and its offerings thorough and accessible. Where else online can I find such a bounty of Britain's heritage? It is a generous endeavor and an absolute goldmine.

This comment, with its mention of useful functionality is mirrored by some of the site feedback, such as the following two:

> Just to thank you for a superb site - I found the answers to questions I'd been trying to answer for a long time. Very clear, super indexing and the 'highlight' function is fantastic.

I have just copied a paragraph from one of your pages and found that the citation is automatically added. FANTASTIC. This saves so much time, and is likely to be more accurate.

Impact in teaching

It had been the impression of the team for some time that whilst BHO has made a very considerable impact on research practice, the service has not been comparably widely used in teaching. By and large, the results of this survey would confirm that impression.

Amongst the respondents to the user survey, only 27% did any teaching in any case. This is in part to be expected: of the 969 respondents, only 196 (less than 20%) classed themselves as academic staff, contract lecturers or tutors or research postgraduates: the classes most likely to do any teaching.

More surprising was the fact that, of those who did teach (the majority of whom presumably taught courses on British history), only 36% used BHO in that teaching in any way. Only 13% thought that BHO was central to their teaching (some 11 respondents), whilst 51% thought it supplementary to their work rather than core.

Also surprising was the level at which that teaching was done. Whilst teachers at undergraduate level formed the majority (55%), more than a quarter (27%) were teaching at foundation or further education level. BHO has hitherto tended to assume an academic audience of at least undergraduate level. See the **Solutions** section for some implications of this.

Respondents were also asked about how they used BHO in teaching. 93% set BHO resources as independent reading for classes and assignments, whilst a considerably smaller proportion (35%) used BHO material 'live in class'.

Quite what this class teaching involves was further illuminated by both the teachers group and the interviews. Overwhelmingly, this 'live' use of BHO materials was in fact from printed copies, and in some cases with a screenshot projected using PowerPoint. Very few interviewees gave live demonstrations of online sources in the classroom, and none at all when asked had ever set a class working collaboratively using multiple devices; neither did they think that the infrastructure was in place to allow such use should they have wanted it.

Use by librarians

The survey of librarians served to interpret the referrer analysis, where almost 40% of academic links came from library pages. Some 82% of the librarians surveyed had been aware of BHO before receiving the survey, and over two-thirds recommend the site to students in general terms. Just over one quarter were from libraries subscribing to the premium content.

The state of affairs regarding more systematic integration of systems was more mixed. Whilst a small minority of library catalogues provided integration at an individual title level, the majority either listed BHO as a single electronic resource in their catalogue or in other general guides to resources. This corroborates the lack of title-level referrers from library catalogues in the quantitative data.

Challenges

The toolkit case studies are small by comparison with the huge amount of data generated in the rapid analysis (for example, the thousands of referrers available to BHO examined under referrer analysis). Larger datasets offer more opportunity for segmentation, but the toolkit does not offer much information on how this might be done.

There were also challenges involved in the focus groups, particularly that involving teachers. Despite the centrality, physically and metaphorically, of the IHR in the historical profession in London and the south-

east, it was difficult to secure sufficient participants. This may be due to the fact that this phase of the project coincided exactly with the opening of a new academic year: certainly, this was the response of at least one interviewee.

The site feedback was voluminous and time-consuming to categorise but did not lead to much insight for this particular project. This was because there was no discernible pattern in the suggestions for site functions, other than the two dominant topics mentioned above.

The nomenclature used in the toolkit, and consequently this report, such at "STLDs" is appropriate for a technical audience, but if the results are to be useful to a more general audience then much terminology will need to be glossed or explained.

Solutions

Prior to beginning the project, the BHO team formulated an indicative list of the broad areas of potential developments of site functions. The two focus groups discussed these at length, and online survey participants were asked to rate them for usefulness. When the user survey results were filtered to include only staff and students in university contexts, with librarians, support for the new tools was as follows:

Potential Development	Users	Librarians
Extra citation formats and download options	65%	95%
Stable and simple URIs	64%	89%
Screencasts covering content, searching and browsing	39%	75%
Personalisation facilities, such as personal workspace	36%	68%
A feed of URIs recently bookmarked or mentioned (by others)	35%	20%
Curriculum-based learning modules	31%	40%
Alternative 'tag cloud' views on taxonomy pages based on relevance	28%	45%
Add and share your own tags	26%	40%

The potential importance of the cool URIs in teaching and research was highlighted by an interviewee, who said that he would never cite URIs unless (like the Oxford Dictionary of National Biography) they have transparent citations, with a clear reference to the underlying material; he makes his students change URL references to, for example, State Papers Online to the archival SP references. This may be the case with academics preferring not to cite BHO, and we would expect this to be reflected in the bibliometric results. Perhaps the format of cool URIs should as closely match the prevailing bibliographic conventions among historians as possible.

The potential of additional download options was welcomed by users as an efficient way of bibliographic management. It was also the case that this would be welcomed by librarians, with 95% thinking them either very or fairly useful. One respondent specifically requested the provision of MARC records, or metadata that might be mapped to MARC, to aid integration into their library catalogue at item level. This may help neglected resources by automating their import into library systems.

It is conceivably the case that these two options were highest rated by reason of the fact that they are most familiar and easy to imagine. This would appear to be backed up by an apparent mismatch between this data and that derived from the two focus groups. Whilst both groups agreed that cool URIs and better

citation options were useful 'low-hanging fruit', there was more time available during the sessions to explain the possibilities of other options. Once the possibilities were expounded, there was very considerable interest in the potential of folksonomies and of screencasts.

The possibility of tagging resources, whether for private use, for global sharing or for use by tutors and students in dedicated course-specific ways, was enthusiastically welcomed by both focus groups. The graduate students welcomed the possibility of sharing their tags with and learning from other users; the teachers saw particular possibilities for setting seminar groups collaborative assignments involving the building of private group folksonomies. This might be reflected in analytics as an increase in the number of pages per visit.

Both groups saw screencasts as a useful introduction to particular clusters of content on BHO, or to particular sources. In addition, survey responses for both research and teaching pointed to the same clusters of resources as the most heavily used (for example Victoria County History volumes). The need for this facility was strengthened by the unexpected presence amongst the users of a high proportion of teachers in further education; a sector in which there would a greater need for introductory materials. Screencasts were also ranked more highly in the librarians' survey than in the general one. The impact of screencasts could be ascertained from log file analysis.

There was however a strong feeling against providing more discursive learning materials (the 'learning modules'), since tutors were most likely still to devise their own bespoke materials; few used any other such materials from third parties in existing courses.

We therefore conclude that we ought to proceed to develop in the following areas with impact measurable using the method(s) listed:

Development	Effect measurable via
Cool URIs	Bibliometrics
Extra citation and metadata download formats	Bibliometrics, content analysis
Folksonomies and shared tagging facilities	Log files, user feedback
Screencasts introducing content and individual sources	Referrer analysis, analytics

More generally, we would hope to see increased usage of BHO in teaching and learning, as reflected across all indicators.