OLD MAPS ONLINE: FINDING AND REFERENCING HISTORICAL MAPPING AS A PLATFORM FOR RESEARCH AND TEACHING

PAULA AUCOTT, HUMPHREY SOUTHALL & PETR PRIDAL

GRANT HOLDER: UNIVERSITY OF PORTSMOUTH

PARTNERS: KLOKAN TECHNOLOGIES GMBH, THE BRITISH LIBRARY &

THE NATIONAL LIBRARY OF SCOTLAND

JISC CONTENT PROGRAMME 2011-13, STRAND C: CLUSTERING DIGITAL CONTENT

WWW.OLDMAPSONLINE.ORG

SYNOPSIS

Old Maps Online is an online search tool for historical maps. It uniquely draws together map collections from libraries around the world into a single, easy to use search interface. Using cutting edge technology the user searches for old maps by location, narrowed by date, rather than through more traditional catalogue listings. Following a successful early launch in February 2012 its popularity has grown continuously. We have delivered large benefits to end users, but this project, and the two conferences it organized, have also succeeded in raising awareness of the importance of spatial metadata for historical maps within the libraries world.

INTRODUCTION

The project aimed to create a single search interface allowing users to find old maps in multiple library collections around the world without having to know either which library held the map or any of the conventional catalogue information such as title, author or publisher, which have traditionally been used when searching for maps. Instead users search using geography, either by typing in a place name or by utilizing the interactive map view, to find maps showing the place they are interested in. They receive instant results as thumbnails with accompanying metadata which each provide a direct link to a full size interactive scan of that map on the website of the host institution.

During the project there has been a steady stream of interest from map librarians wishing to become involved. We launched with just five collections but by February 2013 we provided access to nineteen, and were working to resolve metadata issues from several more.

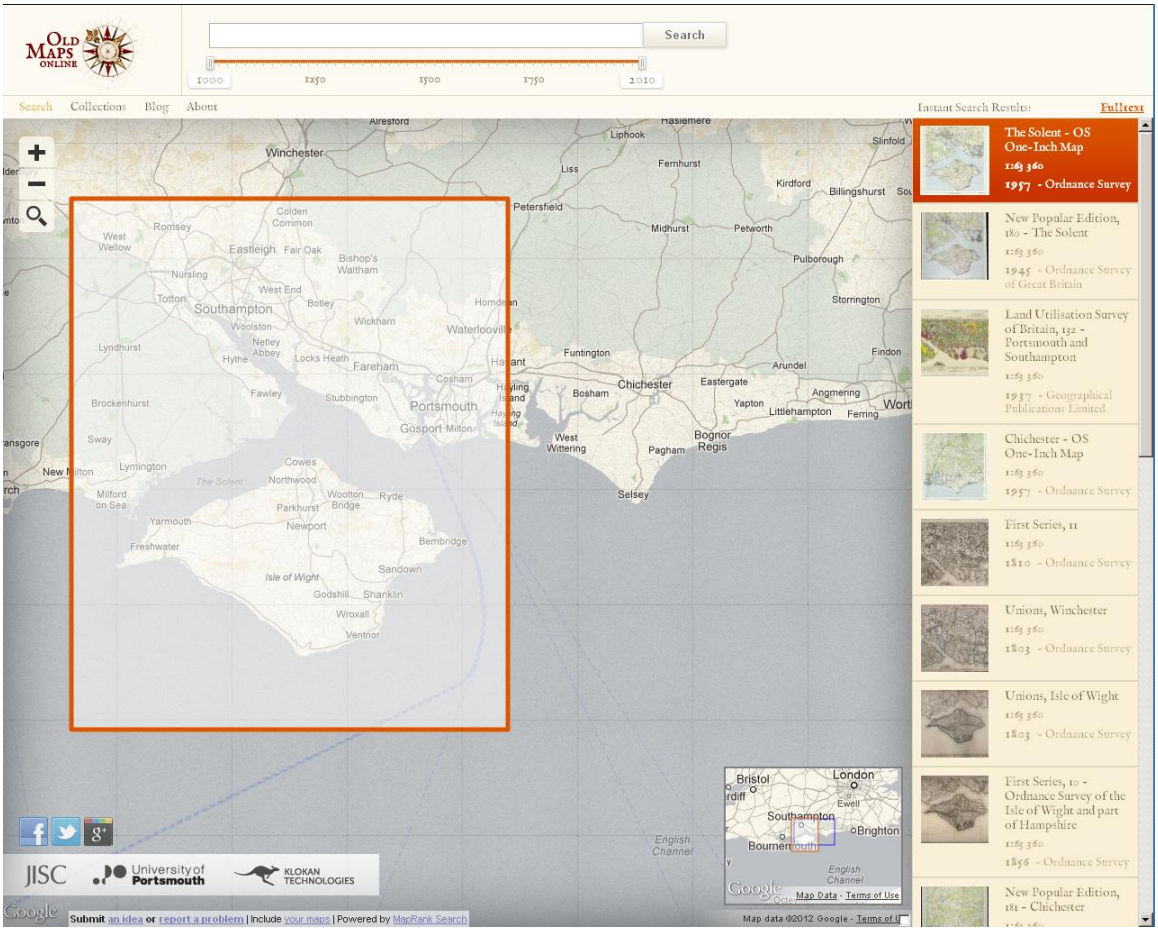

Figure 1: Old Maps Online Homepage, the search interface

As well as creating the Old Maps Online search interface, we aimed to raise awareness of the usefulness of firstly adding spatial metadata to conventional library metadata; and secondly of the need for stable and consciously designed URIs (Uniform Resource Identifiers) rather than URLs as web addresses. Using URIs raises citation rates, increases the readability of the address for humans, and grants longevity to web addresses by removing dependencies on both server software and organizational structures.

CONTEXT AND CHALLENGE

Old maps both have immense intrinsic appeal to users as cultural artifacts, and contain a mass of information on past environments and societies; they arguably tell us more about particular periods in the relatively recent past than, say, old censuses and are available for substantially earlier dates. However, most historical researchers are based far from major map libraries. Widening community access, along with reducing wear and tear on irreplaceable originals, has justified projects in many libraries to digitize old maps and make them available online, but further problems then emerge:

- Online users know little about which particular libraries have maps of where they are interested in, so requiring them to know which online catalogue they should access is a major barrier.
- Traditional map library catalogues enable historians of cartography to find maps by author, publisher or title, but are of little use to the generality of users who want maps which show the places they are interested in. In practice, most users probably relied far more on personal advice from map librarians, which helps explain why libraries seem to have given a particularly low priority to computerizing map catalogues. Unfortunately, online map users obviously lack direct contact with librarians, so have been stuck with the fairly unhelpful catalogues. They need to find maps mainly by **where** they show.
- Most map libraries are embedded within larger libraries and have to rely on web sites, online finding aids etc. designed mainly for books, or for drawings. Few can afford specialized geospatial tools.

The need for a new kind of finding aid for historical maps, itself map-based and spanning multiple collections, is an obvious one and Old Maps Online was not the first such project: other examples are DIGMAP (http://portal.digmap.eu), Cartomundi (http://www.cartomundi.fr) and OpenGeoportal (http://opengeoportal.org), but the first two have failed to attract much content while the last is designed mainly for born-digital geospatial content. The need to actively recruit map librarians to the project, and then work with them to add their content to our system, lay behind our crucial decision to launch our service only four months into a fifteen month project, in February 2012.

That decision necessitated building the first version on existing software, MapRank Search from our partners, Klokan Technologies GmbH; but that software was already far more user friendly than DIGMAP, whose interface required users to type in numerical coordinates rather than simply click on a map, and by the end of our project we were using an enhanced version of MapRank Search scaled up to work with potentially millions of maps in any number of collections. Another key decision was to make contributing metadata as easy as possible: unlike Cartomundi we offered an entirely free service to map libraries, and we worked with them to make their existing metadata fit into our system.

We asked libraries for their metadata, not their map scans. This meant we did not have to provide a large amount of file store as the only images we hold are thumbnails and that libraries retained full control over their images, and how users see them. In many cases the libraries metadata is available under creative commons or is in the public domain, so legal agreements were unnecessary, and where legal agreements were necessary we kept them very simple and did not require libraries to make a long term commitment. Our insistence on bounding box data did limit the amount of metadata available to us, as only a small fraction of the maps that have been scanned have been geo-referenced, but we believe the project has helped make many librarians realize the desirability of this relatively small and cheap extra step for raising usage.

PARTNERSHIPS AND STAKEHOLDERS MANAGEMENT

THE TEAM

The project linked staff at the University of Portsmouth with a software developer in Switzerland, a designer in the Czech Republic, and map librarians all over the world. Overall the relationship between the developer and the academic team has been positive; we funded four visits by the main developer to the UK plus his travel costs to the launch event in New York, and otherwise relied on email and Skype. The Portsmouth team included a separate part-time developer focused on improvements to our own existing map library.

THE CONTRIBUTORS

Encouraging multiple libraries to participate meant there were inevitably issues with ensuring all metadata was of a similar standard. Initially we produced a requirements sheet and provided an example spreadsheet to show what we needed. As the collections contacted us we refined the requirements and discussed recurrent problems, such as 'how do you define a map held within an atlas?' Decisions made for one collection informed the others. Some libraries offered us their metadata; others simply stated an interest in future involvement once their metadata had a spatial element; for two UK libraries, the National Library of Wales and Manchester University, we **did** the geo-referencing.

Often libraries did not realize the funding limitations on the project and felt the short duration was unhelpful as it would take significantly longer to add a spatial element to their data, even though they appreciated the value of our system. Another drawback of the short time scale for libraries was the implied restriction on updating previously submitted metadata, either by adding new records or removing redundant links. To some extent we hope to mitigate both these shortcomings by continuing to offer to incorporate new material into the website after the funding ends. However this will be on an *ad-hoc* basis and there will be little time for the detailed metadata checking that has been occurring during the project period. Our ability to extend or amend the metadata depends on the specialist Portsmouth team being sustained with other funding, but we have a good history of this.

Figure 2: Collections visible in Old Maps Online February 2013

CREATING AND RELEASING

PLATFORM

Although we launched using existing software, we later re-architected MapRank Search to handle millions of maps. It is unique in how it returns results ordered by geographical relevance, and it is computationally very efficient: Old Maps Online runs on the cheapest virtual server available from Portsmouth Information Services, so we were able to cover most of our sustainability obligations by paying in advance for five years' use of the server. One way we limit the computational load on the server is by off-loading as much as possible onto Google services: all pages in the Old Maps Online site except the search interface itself run on Google Sites or on Blogger; and a large part of what users see in the search interface is background mapping from Google Maps. This is possibly a long-term weakness, but we are far from alone in such dependence.

STANDARDS

Old Maps Online organizes metadata following the Dublin Core standard, using the DC.box element for bounding box coordinates. We consciously decided not to base the project around automated harvesting; most metadata arrived as Excel files, or similar, sent as e-mail attachments. Although a few libraries make map metadata available for harvesting using the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH), the metadata so exposed mostly lacks bounding box coordinates, even where we know the library has coordinates for the maps. This may be because the OAI-PMH facility is managed by part of the library organizationally remote from the map library. In the longer term, we hope libraries will properly implement OAI-PMH; but with our limited resources we do not regret our manual approach.

Promoting use of Uniform Resource Identifiers for maps was a significant focus in our public presentations, but implementation within our own map library, and the consequent guide to good practice, were delayed because we needed specialized help on URL rewriting from Portsmouth Information Services; much of the *Vision of Britain* site is now using true URIs, and much of the guide exists as a first draft. Our two *Working digitally* conferences did much to promote good practices more generally.

LICENSING

We quickly discovered that many libraries were very reluctant to sign binding legal agreements. This was especially true of non-UK libraries, where there were issues both of which jurisdiction should apply and of non-native English speakers being baffled by legal jargon. As a result, most of the metadata we hold is not covered by any formal agreement, and the written agreement we developed allows libraries to withdraw content at a month's notice. However, no library has shown any sign for wanting to withdraw content. We see the ideal solution not as a more formal agreement just with us, but as all libraries exposing their metadata for automated harvesting under a machine-readable Creative Commons license. This is some way off.

DISSEMINATION

Launching early meant dissemination work was spread over eleven months. We had two linked launch events: our one-day workshop on "Working digitally with historical maps" on 26 February 2012, in the New York Public Library's main lecture theatre and promoted as a special session within the Association of American Geographers Annual Meeting; and part of the National Archives' Gerald Aylmer seminar, at Senate House London on 29 February. Further presentations were made at the UK Archives Discovery Forum at Kew; a JISC sponsored workshop on geo-spatial engagement at Kings College London; the European Social Science History

Conference in Glasgow; the LIBER Groupe des Cartothécaires meeting and the International Cartographic Association's workshop on 'Digital approaches to Cartographic Heritage', both in Barcelona; EDINA Geoforum in York; Google IO in San Francisco, where the Google Maps team themselves cited Old Maps Online as a "cool" use of their technology; the World History Association congress in Albuquerque; the BCS Map Curators Workshop in Leeds; and the AHRC Digital Transformations Moot in London. Our second "Working digitally with historical maps" conference was held at the National Library of Scotland on 13 December.

All this activity both publicised the web site to end-users and recruited more libraries as metadata contributors. Our initial launch went viral on Twitter, triggering much coverage in technology blogs and slightly random items in traditional media: e.g. a TV news item in Barcelona and an article in *La Stampa* (Milan). We suspect few readers of *Ars Technica* or *Gizmodo* have an enduring interest in old maps; recommendations on map library web sites, and from local and family history blogs probably contribute more to long-term usage.

Our impact on map librarians is more easily traced. We knew that our partners already held most of the geo-referenced on-line historical maps in Britain. Organizing the New York meeting and having a major presence at the Barcelona conferences rapidly gave us a high profile internationally; but it then usually took some months to turn expressions of interest into actual metadata working within our system. Other impacts are that eleven papers from New York were published as a double issue of the *Journal of Map and Geography Libraries* (http://www.tandfonline.com/toc/wmgl20/9/1-2), while our Barcelona presentation is now a paper in e-Perimetron (http://www.e-perimetron.org/Vol_7_2/Southall_Pridal.pdf).

SUSTAINABILITY

Some earlier map portal projects arguably failed partly because of over-ambitious sustainability models. Operation of Old Maps Online until January 2018 is assured simply by an advance payment for hosting, and by including software maintenance until then in our contract with Klokan Technologies. Google Adsense income from our Vision of Britain site now covers hosting and software maintenance indefinitely. Adsense income from Old Maps Online is lower than user numbers predict, but we hope to experiment further.

Greater Adsense income is not needed to keep the site running, except post-2018, but would make it easier to commit resources to updating content and adding more collections. We are committed to doing one major update in 2013, intended to include metadata from the Library of Congress, US Geological Survey and the Australian National Library, all of which we already hold in preliminary form. While our specialist staff are needed to extend or update the system, a simple interface has been included enabling non-specialists to "turn off" specified collections, and so honor requests from libraries to withdraw their content.

Really long term we hope our system will be replaced by one using automated harvesting to gather metadata exposed by libraries to all-comers. However, we believe our project is a necessary preliminary step, changing library practices both on what metadata is created and on whether it is then made available to others, as well as delivering large immediate benefits to end-users.

LESSONS FOR FUTURE PROJECTS

Our single largest issue was agreeing licenses with contributors. Although we began discussing these early in the project, unresolved issues here still led to delays in receiving some metadata: we spent considerable time on a formal contract before opting instead for a simple permission letter from libraries: because this is clearly revocable at fairly short notice, most clauses of the proposed contract become unnecessary. For future

content aggregation projects involving international partners we strongly recommend this letter-based approach, provided the loss of any one source of content would not be fatal to the project.

Both the original launch of the Vision of Britain site in 2004 and its re-launch in 2009 were very slightly after funding ended. Launching a limited site early in the project and then enhancing it greatly reduces overall pressure on the project team and enables far more effective dissemination, but of course is possible only if major software development can be avoided.

That past experience also enabled us to plan the launch more carefully; for example, Portsmouth Information Services made a second identical virtual server available to us just for the launch, with users shared transparently over both servers via a load balancer. However, use of cloud services creates new risks: as well as Google services we were using a third party service that enabled us to include place interactive 'like' links with Twitter, Facebook and Google+; but the sudden popularity of our site at launch caused the service to blacklist us as malicious, and interaction between their embedded Javascript and MapRank Search caused our site to lock up for most users. New risks created by relying on cloud services need careful evaluation; if available information is insufficient for this, don't use it.

IMPACT

THE SITE

Over 200,000 different users visited Old Maps Online in March 2012, but by June this had dwindled to under 20,000. However, a much more positive story appears when we exclude initial fleeting visits by the Twiterati:

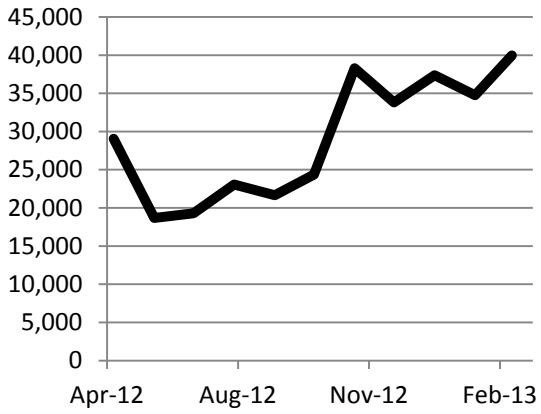

Figure 3: Unique users per month May 2012-Mar 2013

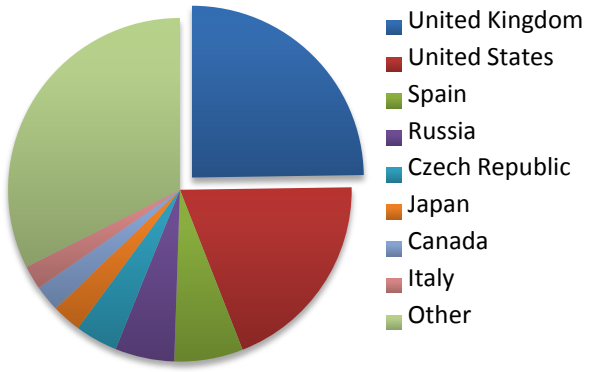

Figure 4: Visitor Origins March 2013

The graph shows the end of the launch peak in April 2012, but then a fairly steady rise to current levels of just under 40,000 unique visitors per month. The pie chart shows that we have a genuinely international user base with 23.6% of visitors from the UK, 20.3% from the US but the majority from other countries. It may seem unfair for the UK higher education sector to be funding this global service, but as noted above the service runs on the cheapest possible server, so restricting access would not cut costs. Further, only two UK universities are contributing content and libraries outside the sector would not contribute to a narrowly academic service; so everyone benefits. For example, a user at University College London needing maps of nineteenth century London found "the three maps which will be of most use … on the oldmapsonline.org website, … provided by Harvard University Library".

Usage data also show that most visits to the site are brief, but our aim is to forward users on as quickly as possible to the library site holding the map they need.

CONTRIBUTING COLLECTIONS

Collection	No. of Maps
Biblioteca Nacional de Colombia	66
British Library, Map Library	2,225
Cartography Associates, the David Rumsey Map Collection (US)	17,242
Charles University in Prague, The Map collection (Czech Republic)	7,892
Dutch National Archives	693
Geo-spatial.org, eHarta Project (Romania)	1,612
Harvard Library, Harvard Map Collection (US)	1,263
Institut Cartogràfic de Catalunya, Cartoteca Digital (Spain)	717
Land Survey Office Czech Republic	53,051
Moravian Library, Mollova mapová sbírka (Czech Republic)	4,030
National Library of Scotland, Maps of Scotland	42,258
National Library of Wales	*277
New York Public Library, Map Division (US)	5,176
Norman B. Leventhal Map Center at the Boston Public Library (US)	845
North Carolina Digital Heritage Center (US)	35
Retromap.ru (Russia)	146
Thuringian State and University Library (ThULB), Jena, Historische Bestände (Germany)	9
University of Manchester Map Library (UK)	*115
A Vision of Britain through Time, Historical Map Library (UK)	2,042
TOTAL	**139,694**

Table 1: Number of Maps from each contributing Collection in the site in February 2013 (* = geo-referenced by Old Maps Online team)

More important is to measure our impact on usage of library sites. The video includes comments from three of the contributing collections. Even those that contributed only a few maps have seen increased usage. The clearest data are from the Boston Public Library, which had 618 maps in our system when this usage data was harvested, but with 1.72 referrals per map we are now their top referrer beating Google, Wikipedia and their own institutional website:

Monthly referral traffic from http://www.oldmapsonline.org as per cent of all referrals:

Submitted first batch of 122 maps late July, 2012

August 2012 – 13.3%

September 2012 – 10.1%

October 2012 – 7.0%

Submitted second batch of 496 maps late October, 2012

November 2012 – 28.1%

December 2012 – 24.8%

MAP LIBRARIES GENERALLY

The fact we were sought out by the National Library of Columbia, wanting to be included, shows how wide our impact has been: there are relatively few specialist map librarians in the world, but most have now heard of us. The New York workshop contributed substantially to this and it cost the project little.

We have promoted discussion of three issues:

- **How should maps and other geographical content be found**? In April 2013 Old Maps Online was listed among "100 Websites which [curators] judge will be essential reading for future generations researching our life and times": "this type of visual search is revolutionising the way in which we conduct historical research and will be invaluable for future researchers wishing to dive through layers of history geographically" (http://www.bl.uk/100websites/top100.html).
- **How should maps be referenced**? Can we get rid of URLs like http://visionofbritain.org.uk/iipmooviewer/view.html?map=gsgs_2957%252FGSGS_2957_19&source=National%2520Library%2520of%2520Scotland&website=www.nls.uk&publisher=British%2520War%2520Office&series=AMS%25201202&sheet=19%2520-%2520Central%2520Europe&x=43&y=42?
- **Can digital map libraries be better than paper?** Many librarians still see digital services as a worse *surrogate*, but with better finding aids and viewers, they can be easier to use as well as easier to get to.

CONCLUSIONS

We have built a well designed website which does what it says on the tin: **it finds old maps which are online.** Growing usage reflects the advantages of a site which is open access, easy to use and instantly returns relevant results. We have persuaded many libraries spread over four continents to provide metadata, and we hope to go on adding more maps and more collections to extend the resource. Our contributors have been impressed with the results. They like the improved findability the service offers for their collections, doing more with existing metadata, and therefore with minimal effort by them. Map librarians can use Old Maps Online to show their colleagues what can be done if catalogues include specifically spatial metadata.