

ALISS Quarterly Vol. 7 Number 3

April 2012

© The authors

Each article is copyrighted to its author(s) and all rights are reserved. No part of this publication may be reprinted or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or any information storage or data retrieval system without prior written permission from the author(s) and full acknowledgement to *ALISS*.

Editor: Heather Dawson

h.dawson@lse.ac.uk

Published by ALISS.

ALISS Quarterly
Volume 7 no. 3 April 2012

Open Doors:
Gateways to Social Science Information Online

Editorial

Heather Dawson

Portals

The Third Sector Knowledge Portal

Jennie Grimshaw (Lead Curator of Social Policy and Official Publications, British Library) and Pete Lambert (Knowledge Portal Manager, Third Sector Research Centre)

Connected Histories: Building Sources for British History, 1500-1900

Jane Winters

Head of Publications & IHR Digital, Institute of Historical Research

Re-imagining Booth: The Construction of a Guided Walk with Reference to the Charles Booth Archive in the LSE's On-line Library

Sean Patterson

The Digital Education Resource Archive (DERA) one year on – lessons learned, plans for the future...

Rozz Evans

Institute of Education Library, University of London

Open Access Publishing

OAPEN-UK: exploring open access scholarly monographs in the humanities and social sciences

Caren Milloy

Head of Projects, JISC Collections

Huddersfield Open Access Publishing

Graham Stone

Information Resources Manager, University of Huddersfield

Repositories

Bringing a Buzz to NECTAR

Miggie Pickton

Research Support Specialist, University of Northampton

LSHTM Research Online: Release the Beast!

Andrew Gray, Manager LSHTM Research Online

London School of Hygiene and Tropical Medicine

Editorial

Welcome to the latest edition of ALISS Quarterly. It has been published by ALISS (Association of Librarians and Information Professionals in the Social Sciences). This special issue is based upon our Christmas 2011 event which was entitled Open Doors – Gateways to Social Science Information online. It was held at the British Library in December 2011 and featured presentations from three major projects:

Peter Lambert

Knowledge Portal Manager, Third Sector Research Centre

Jennie Grimshaw, Curator, Social Policy Collections, British Library

Introduced the Third Sector Knowledge Portal which was launched in October 2011 and aims to bring together research and information on the voluntary and community sector <http://www.tsrc.ac.uk/Research/KnowledgePortal/tabid/840/Default.aspx>

Rozz Evans

Collection Development Librarian Institute of Education, University of London spoke about the development, content and future plans of Digital Education Resource Archive (DERA) <http://dera.ioe.ac.uk/>

Dr Jane Winters Head of Publications & IHR Digital, Institute of Historical

Research was involved in the development of the Connected Histories: Sources for Building British History, 1500-1900 website <http://www.connectedhistories.org/> which was launched in 2011 and offers a federated search of 15 major databases of primary source economic, social and political history texts and images. She offered insight into the development and content of the site.

All of these presentations are discussed more fully in this issue. The original papers can be downloaded from slideshare at <http://www.slideshare.net/heatherdawson/> The issue also contains discussion of other innovations which are opening up social science information. They include open access publishing and the development of open access repositories on university websites.

However, before we move onto the main body of the issue, here are our recommended resources for keeping up to date with information on open access.

Key organisations

Use these websites for current awareness. All are key movers in the field.

Eprints.Org

<http://www.eprints.org>

Site developed by the School of Electronics and Computer Science, University of Southampton to support Eprints open source software. It provides free access to a wealth

of resources including: technical support documents, FAQs on self-archiving for authors and an extensive directory of useful links. Also available is the EPrints' ROARMAP service: <http://roarmap.eprints.org> a searchable register of mandates and policies on open access publishing from institutions and research funders.

JISC

<http://www.jisc.ac.uk>

JISC supports the innovative use of new technology in the UK HE and FE sector using funding received from the UK government. It has been heavily involved in promoting the management and development of institutional repositories. Its website is a key starting point for tracing both historic and current programmes. Topics covered include: technical infrastructure, digital preservation and legal concerns.

SHERPA

<http://www.sherpa.ac.uk>

Consortium led by the University of Nottingham which is investigating issues in the future of scholarly communication, focussing in particular on open access repositories. Key resources offered include:

RoMEO <http://www.sherpa.ac.uk/romeo.php> - Searchable listing of journals and publishers providing information on their policies regarding self-archiving and open access.

JULIET <http://www.sherpa.ac.uk/juliet/index.php> - Summary of policies given by various research funders as part of their grant awards.

SPARC

<http://www.arl.org/sparc/openaccess/index.shtml>

International consortium of research libraries which seeks to influence models of future scholarly communication to increase access and reduce financial pressure. A key member is Peter Suber who has produced the open access newsletter and currently maintains an excellent collection of bookmarks to new reports and sites via Connotea. See his website for further details: <http://www.arl.org/sparc/about/staff/suber.shtml>

Open Access Directory (OAD)

http://oad.simmons.edu/oadwiki/Main_Page

Wiki about open access (OA) hosted by the Graduate School of Library and Information Science at Simmons College and supervised by an independent editorial board. It relies on the OA community to help maintain and keep the links up to date. A key feature is The Open Access Bibliography, http://oad.simmons.edu/oadwiki/Bibliography_of_open_access which is based on Charles W. Bailey, Jr., *Open Access Bibliography: Liberating Scholarly Literature with E-Prints and Open Access Journals*, Washington, DC: Association of Research Libraries, 2005. It seeks to list and link to articles and reports covering all aspects of open access. Topics include- mandates, copyright, institutional repositories and interviews.

OpCit Project

<http://opcit.eprints.org/oacitation-biblio.html>

Based at the Department of Electronics & Computer Science, University of Southampton provides free access to a bibliography of articles and reports that focus upon the impact

of open access publishing upon citations. Aims to show that making articles open access increases the number of citations.

Key resources

These are directories of open access resources. Useful for searching and finding the full text.

OpenDOAR

<http://www.opendoar.org/>

Comprehensive list of institutional and subject-based repositories worldwide.

Directory of Open Access Journals

<http://www.doaj.org>

Maintained by Lund University Libraries. Get links to thousands of titles. Regularly updated.

Online ebooks Collection

<http://onlinebooks.library.upenn.edu>

Excellent index to all types of free ebooks maintained by Penn University. Highlights individual titles and large ebook directories. Over 1 million titles currently listed

OER Commons

<http://www.oercommons.org/>

Seeks to provide information and access to free online educational content. This can include: lecture notes, teaching materials and lesson plans. Also useful is the OpenCourseware consortium

OpenCourseware consortium

<http://www.ocwconsortium.org>

An organization of universities committed to advancing open access learning courses. Get news and links to courses worldwide.

Networked library of Digital Theses and Dissertations.

<http://www.ndltd.org/>

Useful starting point for tracing repositories of online doctoral and masters theses worldwide. Includes details on national projects.

We hope you enjoy the issue!

Keep up to date with our website at <http://www.alissnet.org.uk>

We also have a new twitter channel where you can keep up to date with our latest activities. http://twitter.com/aliss_info. We are using it to highlight weekly listings of new social science websites and new UK government publications online.

Remember that you can also keep up to date with ALISS news by subscribing to our free electronic mailing list LIS_SOCIAL SCIENCE at <http://www.jiscmail.ac.uk/lists/LIS-SOCIALSCIENCE.html>, or consulting our website at: <http://www.alissnet.org.uk>

Heather Dawson.

ALISS Secretary

The Third Sector Knowledge Portal

Jennie Grimshaw (*Lead Curator of Social Policy and Official Publications, British Library*) and Pete Lambert (*Knowledge Portal Manager, Third Sector Research Centre*)

Introduction

The term 'third sector' is used to describe "non-governmental organisations which are value-driven and which principally reinvest their surpluses to further social, environmental or cultural objectives. It includes voluntary and community organisations, charities, faith groups, social enterprises, cooperatives and mutuals." (Office of the Third Sector, May 2006).

The Third Sector Knowledge Portal is a new database of evidence, research and analysis, built around a unique thesaurus. It has been developed by the Third Sector Research Centre (TSRC) at the University of Birmingham in partnership with the British Library. The site is free to use, and has been financed for its first three years as part of the Big Lottery Fund's 'replication and innovation' work stream.

The Portal provides a user friendly, clearly signposted, gateway to the vast range of information available online and offline about civil society and the role, activities, organisation and impact of the Third Sector in the UK and worldwide. It can provide essential support to small charities seeking evidence to demonstrate the need for their services to funders, as well as enabling academic researchers to efficiently review the existing literature, and disseminating best practice in organisation administration and development to managers and leaders.

Launching the Portal at the British Library in October 2011, Big Lottery Fund Chief Executive Peter Wanless explained:

"There is no shortage of information on the work of voluntary and community organisations, but it is often fragmented and hard to access. Organisations re-invent wheels and start from scratch because they are not aware of resources already available."

The Portal has been designed to bring together a diverse range of information published within the sector itself, by the academic community, and by government agencies, in an easy to use 'one stop shop' bridging the information divides that separate these communities. It includes citations to research and practice reports, journal articles, government documents and academic monographs, both current and historic, with links to full text where available.

All records in the database are manually indexed using a thesaurus of over 800 descriptors which has been constructed specifically to meet the information needs of the sector on the basis of an analysis of the literature and in consultation with the research community.

Technical infrastructure and staffing

The Knowledge Portal database and search interface were created using Heritage Library Management Software. This was chosen in part because it was already in use for the Health Services Management Centre Library at Birmingham University and was supported by the IT Department. It also offers a user friendly interface for record creation accessible

to persons other than trained cataloguers, a flexible search interface for end users which could be easily customised and adapted in response to feedback, and an impressive range of facilities for generating usage reports and statistics to delight our funders and demonstrate our impact.

New records are created and added to the database daily by a small but enthusiastic group of volunteers recruited from within the University and the local third sector, under the supervision of the Portal manager. A full time volunteer has recently moved on to paid employment, with three new part time volunteers now taking her place. Their commitment and knowledge of the sector has enabled the database to be rapidly developed and expanded from zero to 900 records in advance of its launch in October 2011, and at the time of writing four months later this figure has doubled to 1800. Where permission can be negotiated, and especially when organisations are at risk of closure, we download and ingest documents into our own server to ensure their preservation. Back-up copies are also held in the British Library's digital store. Where permission to download has not been negotiated, or where access is fee-based, we link out to material on publishers' websites.

Terminology and the thesaurus

A constantly shifting specialist terminology can prove a significant barrier to the flow of useful information between researchers, policymakers, practitioners, commissioners and managers. The terminology applied to the sector and its activities is unstable and prone to changes of meaning and overlaps, as new governments, developing managerial processes, and shifting analytical approaches continually move its goal posts. The term 'third sector' itself exemplifies this, as it is often used interchangeably with a range of terms such as: voluntary and community sector, civil society, not-for-profit sector. However, each of these overlapping terms has subtle differences in definition and connotation which change over time. Moreover, as the sector itself evolves and policies change, new terms are constantly emerging, rising to prominence and disappearing. Thus the 'Big Society' agenda promoted by David Cameron has superseded the 'Third Way' policies espoused by Tony Blair.

The Knowledge Portal thesaurus has the ambitious aim of organising and defining the terms used in the literature and mapping the relationships between them using the normal techniques of synonym control, and linking broader, narrower and related terms. Its unique feature is its inclusion of detailed definitions of its terms. This has proved a significant challenge and has generated lively debate as researchers frequently disagree amongst themselves about meanings!

The thesaurus is 'locked down' as a database within the Heritage LMS, so that cataloguers cannot create new descriptors at will, but it is updated weekly by the Portal manager to ensure it captures new concepts and adapts to the developing research agenda.

Periodic reviews will be held to prune the thesaurus of unused terms, or to combine terms which have become effectively synonymous.

When users enter a term in the basic search box Heritage produces a list of all catalogue records containing that term, and additionally offers a 'See Also' list of alternative terms

they might like to try based upon the links to broader, narrower and related terms in the thesaurus.

For example, a search for 'grants' will, in addition to showing users all records containing that term, offer a list of related search terms:

'Income generation' - The broader term above 'grants' in the Funding and Finance section of the thesaurus hierarchy

'National Lottery' - A narrower term below 'grants' in the hierarchy

In addition to the hierarchical links, we also specify related terms in other hierarchies within the thesaurus. In the case of 'grants' we have additionally defined links to the keywords:

'Funding applications' - from a different branch of the Funding and Finance hierarchy

'Grant-making foundations' - from a different hierarchy, for types of organisations

We can access detailed reports on the search terms users are employing. For example, any searches made which do not find content will be highlighted. Such information will help us to develop the thesaurus and prioritise new content to enable us to better support the information needs of our users.

Content

There is no shortage of potential content for the Portal, and the development of a flexible system for prioritising items for inclusion is essential to its ongoing growth. Inevitably there must be experience-based judgements on a case by case basis but we prioritise the inclusion of works which are:

- research produced by TSRC and other reputable organisations in the sector itself,
- articles accepted for publication in peer reviewed academic journals,
- government agencies' research and guidance,
- part of an established collection which can be absorbed into the system,
- representative of 'knowledge at risk' as other print or online collections face cutbacks or closure.

The Portal includes international works which are of relevance to the UK, and its content reflects the differing agendas of the devolved governments in Northern Ireland, Scotland and Wales.

Users cannot upload their own content directly but they are actively encouraged to submit recommendations of publications for inclusion. TSRC and the British Library are building partnership arrangements with a range of third sector organisations which publish their own research, or have built collections of useful information which can be preserved in the Portal, as part of our commitment to preserving knowledge at risk. With third sector projects often supported only by short term finance, valuable information resources are often lost as websites close or key project officers move on. Case studies and project reviews from the recent past can be of use to organisations developing their work in similar fields, and through the Portal's reliable index and in-house servers access

to these works can be maintained beyond their original web presence.

Partner organisations' collections can be imported into the Portal in their entirety, and this will soon be taking place with the Evidence Library of the Scottish Council of Voluntary Organisations, an index of over 2000 government, academic and third sector publications of relevance to Scotland and the whole of the UK.

Depending on the format and media type of each item, Portal records provide users with:

- a free PDF (or any other electronic format) download from the publisher's own site or the Portal's secure server,
- a direct link to the British Library's online catalogue entry for the item, from which reference collection viewings or document supply options can be explored,
- a direct link to online journal articles, which may require a subscription to Athens or a similar system for full access, but provide free abstracts and related information.

We can also catalogue and provide downloads of databases, audio/video files, and links to other useful online resources – a recent example being a BBC Radio 4 documentary on the development of David Cameron's Big Society agenda.

Try it out

You can access the Portal through a link from the Third Sector Research Centre's home page at <http://www.tsrc.ac.uk/>. We welcome all comments (especially complimentary ones!) suggestions and feedback.

The future

Early statistics on use of the Portal have been very positive, with users making 2115 online visits and 6000 searches over these first four months. An optional entry questionnaire has been completed 575 times, with 49% of users identifying themselves as from the third sector, 28% academics/researchers, and 23% government or other.

Long term information preservation is a core objective of the Knowledge Portal project, and to this end we are preparing for a variety of potential scenarios once its initial funding period ends in December 2013. These include TSRC continuing to host and update the Portal with alternative financial support, its absorption into the British Library's own systems, or its transfer to another academic or third sector body.

The level of use the Portal is already obtaining shows clearly that it is needed, and we look forward to continuing to build and develop the distinctive service it offers to people working in, for, and with the third sector.

Connected Histories: Building Sources for British History, 1500-1900

Jane Winters, Head of Publications & IHR Digital, Institute of Historical Research

Connected Histories (<http://www.connectedhistories.org.uk>) was launched at the Institute of Historical Research, University of London (IHR) on 31 March 2011. At the time of writing it brings together 15 major sources for the study of British history from the early modern period to the 19th century with a single federated search. The diverse range of sources indexed by Connected Histories includes image databases (the John Johnson Collection of Printed Ephemera and British Museum Images); OCR'd texts (the 17th and 18th Century Burney Newspaper Collection, the House of Commons Parliamentary Papers and the JSTOR Nineteenth Century British Pamphlets); rekeyed texts (London Lives 1600-1900, the Proceedings of the Old Bailey Online, John Strype's Survey of London, British History Online and John Foxe's The Acts and Monuments Online); and structured databases/datasets (the Clergy of the Church of England Database 1540-1835, the Charles Booth Archive, Origins.net, Cause Papers in the Diocesan Courts of the Archbishopric of York 1300-1858 and the Convict Transportation Registers Database). Some of these sources are freely available, some are free at the point of use to researchers based in UK Higher Education Institutions and some are subscription only. However, even for those which require some form of payment, the snippet view provided in Connected Histories search results allows users to see enough information to get a good sense of the source material.

The diversity of technical approaches and methods of access described above provided the impetus for the Connected Histories project. Historians are among those to have benefited most from the large-scale digitisation of the past 10-15 years, and early modern British history in particular has been very well served. However, the digital landscape remains fractured, and even users within well-resourced universities have to familiarise themselves with a range of search interfaces, using different (and often unexplained) criteria to determine search results. While searching for information about a particular person, place or subject is, of course, exponentially faster when using a digitised resource, the process remains complex and ultimately time-consuming. Outside academia, this picture is further complicated by the restrictions of institutionally focused subscription models. A recognition of these problems lay behind the 'Clustering and enhancing' funding made available by the Joint Information Systems Committee (JISC), which supported Connected Histories. The project essentially aims to simplify access to a wide range of resources, thereby encouraging usage and allowing researchers to concentrate on developing new research questions rather than on learning new systems.

Search is key to Connected Histories, as developing a meaningful browse facility for such a large and disparate set of material is more or less impossible (and certainly beyond the scope of the project). Connected Histories supports full text searching, but Natural Language Processing (NLP) has also been applied to all of the sources to identify places, people and dates that appear in the texts and image metadata. Searching for 'Sheffield' as a place, for example, produces very different results from a general keyword search for 'Sheffield' (44,212 results as opposed to 55,551); Robert Sheffield, a Lincolnshire rector and vicar, no longer appears. Similarly, searching for Hatfield as a surname excludes the

majority of instances of Hatfield the place. I say majority, because the NLP can never be completely accurate, particularly when applied to OCR'd texts. However, when dealing with such a large corpus of material (around 10 billion words at present), the NLP is an invaluable tool for narrowing and refining searches. In order to preserve the integrity of the underlying resources, and to ensure that findings are not decontextualised, search results are presented in 'channels'. The Sheffield place-name search mentioned above, for example, returns results from 14 of the 15 sources included in Connected Histories, and the user can see at a glance that 27 of these are from the Convict Transportation Registers Database, while 1,077 are from British History Online and 33,410 from the House of Commons Parliamentary Papers. Results can also be filtered by 'Document type' (for example 'Newspapers' or 'Images and Maps'), by 'Date' (for example '1600-1699'), by 'Availability' ('Free' or 'Subscription'), or by any combination of these. Users may further choose to limit their search by excluding individual sources from the results.

Bringing together a range of resources focused on a particular place and period of time has clear benefits for the researcher. An obvious example of this is the option to search the Proceedings of the Old Bailey Online alongside the Convict Transportation Registers Database. The ultimate fate of those sentenced to transportation at the Old Bailey is revealed by a single search, as in the example of Francis and Charles Wiltshire, sentenced to seven years' transportation in 1833 for the theft of a handkerchief worth three shillings. They were among 260 individuals transported to Van Diemen's Land (modern day Tasmania) on 3 August in that year. The co-location of these sources also reveals unexpected patterns and strengths. For example, the House of Commons Parliamentary Papers is the richest source for the terms 'Indian summer' and 'Halloween' – the former in the context of expedition reports from North America, the latter in relation to festivities in Irish and Scottish 'lunatic asylums'. What might be viewed as primarily a resource for the administrative and political history of Britain is revealed to be of great value for social and cultural history too.

Registration is not required for Connected Histories, but registered users can access a range of additional options from their personal workspace. It is possible to save whole searches, to save single or multiple individual search results, and ultimately to create connections between those results. The facility to 'connect' particular results allows users to build up their own collections of related material, which they may choose to keep private or share with other users of the site. The connections that people have chosen to share so far reveal a wide range of research interests, and in some cases suggest unexpected uses for the source material. For example, there is a large collection of material relating to early film and cinema-going, which might otherwise have remained obscured. Other users have chosen to focus on gin, snuff, playing cards and even hot air balloons. Topical recent collections include one for St Valentine's Day, and one for Scottish Union, Home Rule and Independence.

In order to help users navigate their way through Connected Histories, we have produced detailed descriptions of all of the resources searched, with information about the technical methods used in their creation, strengths and weaknesses and so on. An understanding of the different ways in which resources have been produced, and the editorial and

technical decisions taken, is essential in understanding the results generated by a Connected Histories research. Short research guides also discuss relevant content and give suggestions for search strategies and further reading. Subjects covered include 'Crime and justice', 'Family history', 'Poverty and poor relief', 'Local history' and 'Searching for images'.

Connected Histories is not a static resource, and new sources will be added regularly. Among those to be indexed in the next few months are the Witches of Early Modern England project, the History of Parliament Online, the Victoria County History and Lane's Masonic Records of England and Wales. Further additions are planned, and it is hoped to include two major updates annually, at least for the next couple of years.

Connected Histories is a JISC-funded project, and a collaboration between the IHR and the Universities of Sheffield and Hertfordshire. The Directors of the project are Professor Tim Hitchcock, Professor Bob Shoemaker and Dr Jane Winters.

Re-imagining Booth

The Construction of a Guided Walk with Reference to the Charles Booth Archive in the LSE's On-line Library

Sean Patterson sean@west367.demon.co.uk

Introduction

The Discovering Britain website <http://www.discoveringbritain.org/> was created by the Royal Geographical Society in association with educational charity IBG. It aims to create a portal of over 200 walks for the public which bring to life the geography and history of Britain.

This paper examines the construction of a guided psychogeographic walk in Deptford, south-east London called *District 45*. <http://www.discoveringbritain.org/walks/region/greater-london/london-deptford.html> It will include a description of the finished walk and will mark through the process of creation from its inception with specific reference to the use of the LSE's Charles Booth archive which played a vital role both in formulating ideas and in the fully realised product.

Background

The walk was originally conceived as the centrepiece of a dissertation for an MA in London Studies at Queen Mary and Westfield University, a dissertation which aimed to prove that certain walking practices are a valid tool to help cultural geographers examine notions of *space* and *place* and it was influenced by the differing walking practices of geographers such as Tim Brennan, Janet Cardiff and Toby Butler, all of whose work has gained general acceptance in the academic field.

Inception

All good walking tours are mental maps. They are a physical picture of a linear narrative carried within the mind of the guide and presented as an unfolding story to the walkers. There are many factors to be considered in choosing an area or subject and once chosen, the area and subject must support and inform one another. Psychogeographers such as Ian Sinclair and Tim Brennan have tended towards areas such as the East End of London because of the layers of history and experience available and I initially chose Deptford because I understood it to be virgin territory with multiple layers of human history, lines of communication and a degree of mystery contained in events such as the death of Christopher Marlowe. For my purposes it also had the advantage of relatively few grand and obvious sights to distract from my particular narrative.

That narrative needed a concept, something to mark it out from a traditional, bullet-pointed walk, that would ultimately be a *route with meaning* in itself. Psychogeographers have used techniques such as placing a map of Paris over a map of London to find an eccentric route, they have tossed coins at junctions to decide which way to go and have even recommended having a small child decide the route, but it was a chance encounter with the LSE's Booth archive that provided the missing element, the *route with meaning* that I was seeking.

Mapping

Like many people I was aware of Charles Booth's famous socio-economic mapping of late Victorian London only from the finished coloured maps and I had little idea of how they were researched. I searched for Booth on the internet originally only to make a simple comparison between Deptford's past and present socio-economic standing but the chance discovery on the LSE's online Booth library of a hand drawn map of the route walked by one of his researchers provided the answer to the *route with meaning* problem that I had been trying to solve.

What the archive contains is scans of the original hand written notebooks compiled whilst walking around the area with local policemen, supplemented by original sketched maps and linked by modern maps and key-words. These notes offer a snapshot of exactly what was seen on every street on a particular day in Deptford in 1899 and crucially they are completely unconcerned with bullet points such as grand buildings.

Now I had two layers: the route would be Booth's and the stopping points my own.

The ability to swap between modern maps and the Booth maps at the click of a mouse was a great time saver because so much of the topography of Deptford has been altered by slum clearances and bombing leaving some parts of the route un-walkable. My plan was to walk as much of the area of Deptford covered in the notes as possible (which in 1899 took three days and was designated *District 45*), and then try to mesh the most interesting bits of text and most interesting streets with the bullet-point sites I wanted to arrive at. In the final product I would be able to recite the notes to walkers so that they could compare a particular street directly with its past.

Unfortunately the LSE library does not have a transcription for most of the Booth archive and the task of transcribing the *District 45* notes from the scribbled text was laborious. Much of it is actually quite tedious and much simply difficult to read (for a whole day I pursued the exciting notion that there were 'gallows' at Deptford Broadway in 1899, only to realise that I had miss-read the word 'barrows'). Ironically South East London is the one region fully transcribed in Jess Steele's book *The Streets of London: The Booth Notebooks* and had I contacted the LSE online librarian earlier I could have been saved a lot of trouble and been directed to her book. It is worth noting, however, that Steel's book does not contain the sketched maps that alerted me to the possibility of using the archive in the first place, a fact that makes a compelling argument for the availability of primary sources, in this case scans of the original notebooks. The point being that the Booth archive is comprehensive and has not been arranged according to an inherent assumption about how it will be used by researchers.

The finished product

Around three months work resulted in a three hour circular walk consisting of three layers. The first layer is the route provided by Booth and the leader reads the archived notes at the exact locations they were originally made. This route takes the walkers to locations or bullet points that have historic or present interest or resonance, and information about these locations comprises the second layer. Finally the walkers themselves are given readings to recite to the group, readings from sources such as

historical diaries, a coroners report, poetry, an estate agent's report, a bombing report and a sea-shanty. This last layer is often oblique and tangential and essentially forms the psychogeographic element, appealing to senses and memories and inhabiting the space between what the walkers can see now, and what Booth's assistant saw over a hundred years ago. That space is what cultural geographers would call a *sense of place*. It is worth mentioning the often maligned Google search engine as a useful tool with regard to this final layer because it is exactly its 'scattergun' approach to sifting information that can lead to the discovery of hidden and unlooked for connections and facts.

Conclusion

District 45 can be seen simply as an interesting and amusing diversion from the more traditional methods of socio-economic study and general cultural geographic practice, but the intention is both to provide a deeper, sensory experience of what *place* means, and by extension to establish a mindset for discovering other areas in a similar way without the help of a guide. The LSE's online Booth library proved not only a pivotal point in the creation process by providing the inspiration of the sketched map, but ultimately became an integral part of a project that required a kind of glue to bind it together.

Bibliography

Brennan, Tim (1999) *GUIDEBOOK: three manoeuvres by Tim Brennan in London E1/E2* (Camerawork).

Butler, Toby (2005) *Memoryscape* (Royal Holloway) www.memoryscape.org.uk/.

Cardiff, Janet (1999) *The missing voice (case study b)* (London: Artangel).

LSE Booth Archive <http://booth.lse.ac.uk/>

Patterson, Sean (2010) *District 45: A walking tour of Deptford* (charlesboothwalks.com)

Sinclair, Iain and Lichtenstein, Rachel (2000) *Rodinsky's Room* (London: Granta).

Steele, Jess (1997) *The Streets of London: The Booth Notebooks* (Deptford Forum Publishing Ltd).

The Digital Education Resource Archive (DERA) one year on – lessons learned, plans for the future...

Rozz Evans

Institute of Education Library, University of London

It seems incredible to me as I'm writing this that it is already a year since the launch of the Digital Education Resource Archive (DERA). I wrote an article in the October 2011 ALISS Quarterly describing how we set it up and my colleague Bernard Scaife wrote an article on the more technical aspects of the database in the July 2011 edition of Ariadne. The slides from my talk at the ALISS 2011 Open Gateways event in December are also available on slideshare: <http://www.slideshare.net/heatherdawson/presentations>. I don't intend to repeat the content of those as I'm keen to share lessons learnt since we launched, progress we've made and plans for the future.

As a reminder: the Digital Education Resource Archive (DERA) is a permanent digital archive of largely born digital document-based publications from government and related (publicly funded) bodies selected for their relevance to the subject of Education. It is not a website archive. We developed it using a customised version of EPrints and it is now firmly embedded in our collection. It is not a website archive.

DERA's first year

When we launched in February 2011 there were just over 1000 documents in the archive. This was made up of electronic publications published in 2010 plus a deposit of almost 300 publications material from BECTA (the British Educational Communications and Technology Agency), a quango which was formerly funded by the Department for Children, Schools and Families (DCSF).

When it was announced that BECTA was to be abolished in May 2010 we saw an opportunity to demonstrate the potential for DERA as a permanent home for born-digital publications. Since then we have also been entrusted with large deposits of electronic publications from the General Teaching Council for England (GTCE) and the Learning and Skills Council (LSC). The former is currently winding up and closes for good on March 31st 2012. The latter was split into two new organisations (the Skills Funding Agency and the Young Peoples Learning Agency) in 2010. We have also been given approximately 300 selected resources from ESCalate, the HEA Education Subject Centre which closed in December 2011. Their publications are still accessible at the moment via their website, but a selected range will be added to DERA for permanent preservation in the near future.

To date there are over 11000 publications in DERA, from almost 100 different organizations. The archive received more than 10600 visitors this month, with over 50000 page views. There is evidence of use from all over the world but the top 5 countries are: United Kingdom (a massive 85% as you'd expect), followed by the United States, Australia, India and Canada.

This is really encouraging and clearly demonstrates that people are finding and using the material. The vast majority of traffic is coming via Google searches. Google Analytics provides us with a large amount of information but appears to show some disparities with

the EPrints own statistics package. We are not entirely sure why this is but we have asked the EPrints community for their thoughts.

We have now established an efficient workflow which means that the selection of new material, the uploading of the publication and any copyright issues are dealt with by our Collection Development Services team and all the editing of records is done by colleagues in Cataloguing and Acquisitions teams. We add approximately 260 new current publications per month, selected by our Official Publications librarian.

The other major source is the deposit of digital publications archives by organisations as mentioned earlier.

This alone doesn't account for our leap from 1000 documents in February 2011 to over 11000 in February 2012 though, however prolific government departments and agencies are!

The POPE project

The reason we have been able to add such a huge number of documents in DERA in the last few months is thanks to funding from JISC for our POPE (Preserving Official Publications in Education) project. This project enabled us to add an additional 6000 documents retrospectively from 1990 – 2010.

These were records which used to be in the library catalogue and were added before 2010. We did have the resource to upload the metadata (in bulk) from our LMS (Symphony) but not to link to the documents themselves which had to be a manual process. Our funding from JISC meant we could carry out this work. We anticipated some problems with the tracing of full text documents from dead links and were expecting to have to digitise possibly hundreds of documents from our printed-out backup copies as a result. In the end there were a lot fewer than we expected that had disappeared totally but it took sustained and expert searching by Antony, our POPE project staff member to track down many, and certainly required skills beyond the standard that you'd expect most researchers to possess. POPE finished at the end of February and we are very happy with what we've achieved.

We do have some outstanding issues. The two main ones are around copyright and format.

1. Copyright. We expected this may be problematic and it was (and is). We still have several hundred items in review (i.e. in DERA but not visible or accessible to the public) as we are still dealing with this. The main issue is the surprising number of publications published by an official body where the copyright remains with the authors of the report, or their institution.
2. Format. We have so far only added PDF documents to DERA. There was a period in the late 1990s/early 2000s when a number of organisations issued publications in html format, so that while they look like 'normal' electronic publications, they are actually websites.

What next?

We will be upgrading the EPrints software to version 3.3.6. soon. There are some enhancements here, in particular with relation to preservation tools.

We need to find out a lot more about how people use it. It is clear that the vast majority of users come to us via Google. We have no idea (aside from anecdotal evidence) how useful people find it, whether they find what they want, or what they are using the material for, so we need to work on that.

And, although we have done some promotion of DERA via our blog, among our own student and staff body, attendance at Umbrella 2011, and a few articles we probably need to increase this. We would be interested in talking to other libraries who might want to do something similar in other subject fields as this may mean that we can scale up this sort of solution across the HE sector to solve the exact same problems with materials in other subject areas.

We are considering whether it is feasible or sustainable to expand the scope of DERA to include other types of organisation which are not either government or official but now publish exclusively online. We have to consider this in terms of sustainability of the digital archive based on resources (technical, financial and staff).

Digital Education Resource Archive (DERA) <http://dera.ioe.ac.uk/>

Evans, R (2011) Preserving access to Government and Official Publications in Education – the Digital Education Resource Archive (DERA). ALISS Quarterly, Volume 7, No. 1, October 2011

Scaife, B (2011) From Link Rot to Web Sanctuary: Creating the Digital Educational Resource Archive (DERA) <http://www.ariadne.ac.uk/issue67/scaife/>

Digital Education Resource Archive (DERA) – ALISS 2011 Open Gateways <http://www.slideshare.net/heatherdawson/december2011evans>

General Teaching Council for England - <http://www.gtce.org.uk/>

Young People's Learning Agency - <http://www.ypla.gov.uk/>

Escalate - <http://escalate.ac.uk/>

JISC - <http://www.jisc.ac.uk/>

POPE project blog <http://ioepope.wordpress.com/>

Newsam News (the IOE Library blog) <http://newsamnews.ioe.ac.uk/>

OAPEN-UK: exploring open access scholarly monographs in the humanities and social sciences

Caren Milloy, Head of Projects, JISC Collections



OAPEN-UK¹ is a 4 year research project gathering evidence to help stakeholders take informed decisions about a move towards an open access model for scholarly monograph publishing in the humanities and social sciences (HSS). It is an extremely timely project for a number of key reasons:

1. Print sales of scholarly monographs are in decline - in the US, between 1980 and 2000, "a monograph's average library sales plummeted from around 2,000 copies in 1980, to 1,000 in the late 1980s, to 500 in the 1990s, to a little more than 200 in the early years of this century."²
2. Library spending on print books is in decline - in 2010, the Research Information Network's Trends in the finances of UK higher education libraries: 1999 – 2009 reported that "between 2004 and 2009, expenditure on print books fell in real terms by 13% across the sector."³
3. Library monograph budgets are under pressure to support journal expenditure
4. Publishers and university presses are collaborating to offer electronic subscription based access to scholarly monographs – in the last 12 months a number of new initiatives have been launched, OUP's University Press Scholarship Online, CUP's University Publishing Online, Project MUSE and Books at JSTOR
5. A recent OAPEN-UK publisher focus group⁴ estimated sales of e-monographs currently only account for around 6% of publishers' revenue.

The scholarly monograph, the gold standard for academic excellence in many HSS disciplines, is under threat. The effective dissemination of the monograph is critical to fostering a world class, vibrant and innovative research environment, but with print sales declining at a higher rate than e-sales are increasing and new models continuing to rely on dwindling library budgets, are we at risk of stifling HSS scholars?

Open Access Publishing

Relatively little research has been undertaken with regards to the impacts of OA in HSS and even less in relation to digital publishing of monographs. To enable us to really assess the prospect of an OA model for scholarly monographs, and to take account of our UK specific policies and processes, we need to gather a body of evidence.

Open access (OA) is a model that could help to increase dissemination and reduce the burden on libraries. OA does not limit access to those that pay - subscribers, particular institutions or one country. OA publishing is international; therefore once a title is made available online under a Creative Commons licence it can be downloaded and placed on a multitude of platforms, thus increasing the potential for wider dissemination and readership of a title to all users across the world.

OAPEN-UK, follows on from the recent European project, Open Access Publishing in European Networks (OAPEN)⁵ and is taking an evidence based approach to gathering data, investigating attitudes and perceptions and gathering impacts.

Introducing OAPEN-UK

OAPEN-UK is collaborating with publishers, authors, research funders, academics and institutions to gather evidence to help stakeholders make some informed decisions about how they may work towards an OA model for HSS monographs, if indeed it is feasible. It is piloting the OAPEN model, where a PDF version of the title is made available in OA using a Creative Commons licence, with the publisher also making available the print version for sale to help recoup costs. This hybrid model recognises that a move to OA will not happen overnight and that while a transition takes place, publishers will continue to publish print – just as academics will continue to read print and want print versions for reward and recognition.

Palgrave Macmillan, Taylor & Francis, Berg Publishers, Liverpool University Press and University Wales Press are all participating in OAPEN-UK and have had HSS scholarly monographs selected for inclusion in the pilot⁶, which commenced on the 1 September 2011. For each monograph title submitted, the publisher had to submit a matched title to make a pair. The titles were matched as closely as possible on publication date, subject area, age to print sales ratio, price and format history. A total of 60 HSS scholarly monographs were selected, making 30 pairs.

A title from each pair has been randomly placed in either the Experimental Group or the Control Group. The Experimental Group titles are available on the OAPEN Library under a Creative Commons licence. In addition, the titles can be made discoverable via the publisher's own website, institutional repositories, the authors own website and via a 100% view in Google Book Search.

The Control Group titles are available as ebooks for sale under the publishers normal licensing and route to market models. The titles are discoverable as per the publisher's standard processes, but only a 10% preview in Google Book Search.

For both the Experimental Group titles and the Control Group titles, libraries and readers will be able to purchase print copies and where the publisher is making them available, ebook device friendly editions such as ePub.

The pilot will allow sales and usage data to be gathered for the experimental and control group titles and to evaluate differences between the groups.

OAPEN-UK Research Plan

A large programme of research⁷ is being undertaken as part of the project, this includes annual benchmarking surveys, focus groups, interviews and surveys with key stakeholders. By combining the quantitative data from the pilot with sales and usage data, three key questions will be addressed:

1. How might policies, processes and mechanisms need to change to enable OA publication of monographs?

- a. What is the most appropriate business model for OA monographs?
 - b. What is the impact on organisational policies?
 - c. What technical changes may be required in order to move to OA?
2. What are the measurable effects of a move to OA monographs?
 - a. On readership/usage?
 - b. On sales?
 - c. On citations?
 3. How do perceptions of OA monograph publication change among participants during the project?
 - a. Perceived risks and benefits to various stakeholders of publishing in OA format?
 - b. Perceived quality of OA monographs?

Year 1 of the plan includes an initial literature review, an annual benchmarking survey, focus groups and initial scoping interviews / surveys.

Current Findings

The first annual benchmarking survey (which will track the changing opinions of all project participants - authors, publishers, AHRC, JISC, members of the steering group towards OA monograph publication) and the first round of focus groups have now been completed.

Focus groups have been held with publishers, researchers (as both authors and readers), institutional representatives (including librarians, IR staff and research managers), learned societies, ebook aggregators and research funders. Key themes that arose included metadata, versioning, preservation, delivery and the user experience, quality and prestige, business models and risks and consistency. The findings⁸ are being used to define the next set of activities – surveys and in-depth interviews.

The OAPEN-UK researcher survey⁹ (running until 16 April 2012) explores many of the themes from the focus groups in more depth. In particular it looks at researcher's attitudes towards OA, the value they place on being published and their publishers, the viability of business models and their reading behaviours.

In summer 2012 OAPEN-UK will make available the analysis of the sales, usage and citation data alongside the full write up of the focus groups, survey results and interviews. As the project progresses, all the evidence gathered will provide a comprehensive examination of OA as a potential model and enable stakeholders to make informed decisions about the future of the scholarly monograph.

Further information

A guide for librarians¹⁰ and IR managers¹¹ includes a full title list, instructions to download MARC records, author affiliation and how to can get involved.

Details on the project, the full research plan, events and latest news are available on the OAPEN-UK website¹. The project can be followed on twitter: @oapenuk #oapenuk or contact c.milloy@jisc-collections.ac.uk for further information.

References

1. OAPEN-UK: <http://oapen-uk.jiscebooks.org/>
2. Willinsky, J., (2010) Towards the Design of an open Monograph Press. *Journal of Electronic Publishing*, 12, 1, doi: 10.3998/3336451.0012.103
3. Research Information Network (2010) Trends in the finances of the UK higher education libraries: 1999–2009. <http://www.rin.ac.uk/uk-libraries-trends>
4. OAPEN-UK publisher focus group: <http://oapen-uk.jiscebooks.org/research-findings/y1-initial-focus-groups/publishers/>
5. OAPEN: <http://project.oapen.org/>
6. OAPEN-UK pilot: <http://oapen-uk.jiscebooks.org/pilot>
7. OAPEN-UK research plan: <http://oapen-uk.jiscebooks.org/overview/research-plan/>
8. OAPEN-UK focus groups: <http://oapen-uk.jiscebooks.org/research-findings/y1-initial-focus-groups/>
9. OAPEN-UK researcher survey: <http://www.surveymonkey.com/s/oapenukresearcher>
10. OAPEN-UK guide for librarians: <http://oapen-uk.jiscebooks.org/pilot/info-for-librarians/>
11. OAPEN-UK guide for IR managers: <http://oapen-uk.jiscebooks.org/pilot/info-for-repository-managers/>

Huddersfield Open Access Publishing

Graham Stone - Information Resources Manager, University of Huddersfield

The Huddersfield Open Access Publishing (HOAP) Project¹ aimed to develop a low cost sustainable Open Access (OA) journal publishing platform using EPrints Institutional Repository software in order to disseminate University of Huddersfield Press journals. The project was funded by JISC and led by Computing and Library Services at the University of Huddersfield, in conjunction with the School of Education and Professional Development and the Research and Enterprise Directorate.

The HOAP platform

The principle aim of the project was to develop a platform to convert the peer reviewed University journal, *Teaching in Lifelong Learning*, from a print subscription model to an OA e-journal. After discussions with EPrints, a specific front end interface was created for the journal, with content being archived in the University Repository. The journal landing pages² and contents pages follow a specific template. This enabled title specific branding for future journals launched by the University Press to be created easily with no or little additional development costs.

The creation of the journal landing pages and the volume/issue pages is fully automated. The articles are uploaded into the Repository using the existing workflows. The first article of a new issue automatically creates a new entry on the landing pages and a new issue page by referencing the ISSN, year, volume, issue and page numbers in the articles. Each subsequent article deposited is then listed on the journal pages. The efficient workflow means that an entire issue can be uploaded in around 30 minutes. The articles themselves maintain the standard Repository branding linking back to the journal landing pages on the platform. This simplified process aids discovery, e.g. the article only has one instance in the Repository and can be discovered through the Repository, landing pages Google (Scholar), and in due course, the Directory of Open Access Journals (DOAJ)³ and thus retrievable from resources discovery systems such as Summon.

The IRStats package from EPrints is used to monitor usage for all articles. This also allows authors to see their own statistics immediately. In addition the 'impact' of a particular volume can be measured over time using the reporting feature on the IRStats administration pages, e.g. Volume 3 (2) has had papers downloaded from 14 countries around the world within 5 months of publication.

As part of the project, the Press joined CrossRef⁴; this entailed creating DOIs for 42 articles in the journals back run and the annotation of around 200 references. It was also agreed that the Press become a member of the Committee on Publishing Ethics (COPE)⁵, a forum for editors and publishers of peer-reviewed journals to discuss publication ethics.

The move from a subscription model to an OA model required a complete revision of the notes for contributors. A new 'Licence to Publish' has been written to replace the previous copyright transfer agreement, allowing the author to retain the copyright under a Creative Commons Attribution 3.0 Unported License (cc-by)⁶, which permits unrestricted use, distribution and reproduction, provided the original work is correctly cited.

Huddersfield Research Review

The second objective of the HOAP project was to assess the feasibility of a new title, *Huddersfield Research Review*. This title aims to showcase the most significant research at the University of Huddersfield by including interviews with the authors of the most cited and/or downloaded articles in the Repository together with an editorial overview by a senior researcher who will locate Huddersfield research within the broader national and international literature in the relevant fields and disciplines.

Other journals

An audit of the University's journals resulted in the addition of two further titles during the project, *Mental Health and Learning Disabilities Research and Practice* and *RADAR*. These titles have now been assigned DOIs. 5 other titles could also be hosted on the HOAP platform, including 2 new peer-reviewed OA titles, it is intended to launch these titles during 2012 using guidelines inspired by a presentation given by Damien Short at the launch event a sister JISC project, SAS Open Journals⁷.

Toolkit

In order to disseminate its output, the project developed a toolkit⁸ for other institutions, and features sections on:

- Moving to Open Access
- Setting up the landing pages using EPrints
- Adding content
- Dissemination
- Workflows
- Setting up a new journal
- Setting up a best of research title
- Appendices

In addition, the notes for contributors, Licence to Publish and notes for reviewers are reproduced for others to use, the project hopes the toolkit will inspire other institutions to investigate OA journal publishing.

Social Media

Computing and Library Services at Huddersfield has been experimenting with social media and web 2.0 tools and technologies since 2005⁹ and the project was very keen to encourage reader comments and ratings and social tagging as part of the publication process. Although this has been partly achieved through the bookmarks and sharing features of the existing Repository, RSS feeds and automated tweets for new articles, and through the project blogs and Tweets, the project team wanted to go one step further by encouraging authors and readers to use social media based on the recommendations of the RIN report, 'If you build it, will they come? How researchers perceive and use web 2.0'¹⁰. The Repository will be implementing the SNEEP (Social Networking Extensions for EPrints)¹¹ suite of social networking extensions as part of the next release of EPrints. This will allow readers of the journal (as well as all other Repository content) to comment, tag and make notes once they log in.

Recommendations

The final project report will be available later in 2012 and will discuss a number of recommendations for the wider community, including best practice guidelines, further development of the PIRUS2¹² project, a role for the JISC Journals Usage Statistics Portal (JUSP)¹³ to run custom reports for OA journals and the suggestion of an audit of other UK universities with a view to depositing the metadata in a central repository; such as the Knowledge Base+¹⁴ project?

There is still further work to be done, particularly at the national level, however, as stated by the Bloomsbury Qatar Foundation Journals White Paper in its opening paragraph, "Open access publishing has arrived"¹⁵.

Acknowledgements

HOAP would like to thank all those that have assisted during the project, particularly Josh Brown, Programme Manager for the Campus-based publishing strand, for his support and advice and the JISC for funding the project, Adam Field at EPrints for his work on the design of the pages and his insightful comments and suggestions, The University of Glasgow for the original inspiration for this project and the project team, Sue White, Dr Ian Pitchford, Dr Denise Robinson and Cherry Edmunds.

References

1. Huddersfield Open Access Publishing blog: <http://library.hud.ac.uk/blogs/projects/hoap/>
2. Teaching in Lifelong Learning: http://eprints.hud.ac.uk/journal_till/
3. Directory of Open Access Journals: <http://www.doaj.org>
4. CrossRef: <http://www.crossref.org/>
5. Committee On Publishing Ethics (COPE): <http://publicationethics.org/>
6. Creative Commons: <http://www.creativecommons.org.uk/>
7. SAS Open Journals: <http://sasopenjournals.blogspot.com/>
8. HOAP toolkit: <http://eprints.hud.ac.uk/12239/>
9. Stone, G. Social Media in Computing and Library Services at the University of Huddersfield. *Illuminea* 7, (2011). <http://eprints.hud.ac.uk/11010/>
10. Research Information Network, If you build it, will they come? How researchers perceive and use web 2.0, Research Information Network, London, 2010. <http://www.rin.ac.uk/our-work/communicating-and-disseminating-research/use-and-relevance-web-20-researchers>
11. Social Networking Extensions for EPrints (SNEEP): <http://www.jisc.ac.uk/whatwedo/programmes/reppres/sue/sneep.aspx>
12. PIRUS2 project: <http://www.cranfieldlibrary.cranfield.ac.uk/pirus2/tiki-index.php>
13. JISC Journals Usage Statistics Portal (JUSP): <http://jusp.mimas.ac.uk/>
14. Knowledge Base +: <http://www.jisc-collections.ac.uk/knowledgebaseplus/>
15. Bloomsbury Qatar Foundation Journals, Open Access, BQFJ White Paper, London (2011). <http://www.qscience.com/userimages/ContentEditor/1322148977938/oa-whitepaper.pdf>

Bringing a Buzz to NECTAR

Miggie Pickton, Research Support Specialist, University of Northampton

Introduction

NECTAR (the Northampton Electronic Collection of Theses and Research) is the University of Northampton's institutional repository for research outputs. First launched in 2008 (Pickton, 2008), the repository showed a steady growth and by the end of 2010 was recognised as a successful tool for describing and reporting on university research activity.

But it could not really be described as an 'open access' repository.

Despite agreement that

"Entry of bibliographic details (metadata) only will be accepted in NECTAR if it is legally or technically impossible to include the full work (e.g. if copyright prevents submission or if it is artwork or a performance). Full text (or other complete content) is preferred."

(The University of Northampton, 2007)

we found that very little full content had found its way into NECTAR.

The repository was meeting the need for a university research reporting system but not its primary purpose (as far as the NECTAR Steering Group were concerned) of providing open access to research outputs. It was time for a re-think.

The project

At the end of 2010 the JISC put out a call for projects for their 'Repositories: take-up and embedding' programme (JISC, 2011). The purpose of this call was to fund projects which would use existing repository applications, tools and good practice to further embed repositories in their institutional workflows.

The timing was just right for us. We had identified some areas where improvements could be made to NECTAR and we had found some tools and services that we believed would help. So we designed a project to cover a range of procedural, technical and promotional activities. We hoped that the result of these activities would be a sleeker, more efficient NECTAR workflow; greater engagement with the research community and, ultimately, an increase in the numbers of full text items in the repository. Our bid for funding was successful.

Re-branding and the Kultur plug-in

We started by re-branding NECTAR to match the new university website. By visually embedding the repository we were making a clear statement that NECTAR was integral to the university website and as important to the university as any other core service.

Next, with the help of Eprints Services, we implemented the 'Kultur' plug-in to Eprints (Brody and Gray, 2009). Originally developed by a project team from the University of Southampton, the University for the Creative Arts, the University of the Arts London and

the Visual Arts Data Service, the Kultur code enhances the way in which images, audio and video outputs are managed and displayed in the repository (Kultur, 2009). It was felt that these improvements would encourage researchers in the creative arts to engage more positively with the repository.

Procedural change

Alongside these essentially cosmetic changes, we introduced a significant amendment to the procedure for depositing research outputs in NECTAR. When NECTAR was first launched, each School had a NECTAR Administrator who was responsible for entering that School's research details into NECTAR. The data were then checked by metadata specialists in the library and moved by them to the 'live' repository (Figure 1.) This process worked fairly well – especially in gathering comprehensive data from the Schools – but bottlenecks were common.

It was felt that if researchers could enter details of their own work directly into NECTAR then several benefits would ensue:

- More comprehensive metadata would be entered – while previously a researcher would simply have passed a bibliographic reference to the School NECTAR Administrator, now they could enter an abstract, key words, relevant URLs, and a host of other information;
- More full content would be uploaded – the opportunity to upload a file is presented as a routine step in the NECTAR workflow;
- School NECTAR Administrators would spend less time typing in metadata and more time checking it for accuracy, consistency and completeness; this in turn would help the library's metadata specialists in their quality control.

A proposal for this change in procedure was taken to the university's Research Committee and, after some discussion, was approved. The committee were keen that the School NECTAR administrators should continue to perform a 'gatekeeping' role, ensuring that the appropriate research outputs were showcased in NECTAR.

Before:



After:



Figure 1: Change in procedure for submission of research outputs to NECTAR

Electronic theses

Early in NECTAR's lifetime the university's Research Degrees Committee (RDC) had approved a change to the university's thesis submission procedure to require research students to submit an electronic copy of their thesis for subsequent deposit in NECTAR. By January 2011 the first students to have been subject to this mandate were due to complete their theses.

Alongside some work with the Graduate School to review and gain RDC approval for a revised thesis submission process, a sub-project was also started to develop a Northampton theses collection in the repository. The University of Northampton only received its research degree awarding powers in 2005 so it was felt that this would be an achievable goal. To this end each printed thesis in the library had its abstract scanned and a metadata record created in NECTAR. At the same time, software changes were made to ensure that the fields held for theses matched those required by the British Library's EThOS service (British Library Board, n.d.). We then invited EThOS to harvest our thesis metadata.

Added value services

In order to increase levels of interest in NECTAR we were keen to provide services that added value to the basic NECTAR functionality and to find new ways of attracting full content.

As members of the UK Council of Research Repositories (UKCoRR) mailing list (JISCmail, 2011) and as regular attendees of repository related events, we had a large network of repository colleagues to call upon:

- William Nixon at the University of Glasgow gave us the idea of automatically generating emails to authors each time their work was made live in NECTAR;
- Sarah Hall at the University of the Arts London advised us to make a point of asking each researcher we trained to provide a file to upload;
- Karen Bates' work at the University of Salford showed us the benefits to be gained from sharing usage statistics with university colleagues. We installed the 'IRStats' statistics package (University of Southampton, 2011; Field et al. 2012) to help with this;
- From various contributors to the mailing list we were given ideas for exporting repository bibliographies to researcher web pages.

It was the latter that generated the greatest impact. We implemented two feeds:

- Latest NECTAR entries were listed for each School on its research home page (e.g. The University of Northampton, 2011a). Each list drew one item from each Division within the School, thereby displaying research outputs from a representative range of authors and subject areas (Figure 2).
- For each researcher a list of all their NECTAR outputs was exported to their staff profile page. If the staff member already had a list of publications on their profile then the NECTAR list was hidden behind a link at the top of the page, if no list already existed then the NECTAR list was automatically expanded (Figure 3).

Recent publications

Latest School of Science and Technology output added to the University's institutional repository, NECTAR

Standardised Radon Index (SRI): a normalisation of radon data-sets in terms of standard normal variables

Crockett, R. and Holt, C. (2011) Standardised Radon Index (SRI) - a normalisation of radon data-sets in terms of standard normal variables. *Natural Hazards and Earth System Science*. 11(7), pp. 1839-1844. 1684-9981.

Morphometric analysis of aeolian bedforms in the Namib Sand Sea using ASTER data

Bullard, J., White, K. and Livingstone, I. (2011) Morphometric analysis of aeolian bedforms in the Namib Sand Sea using ASTER data. *Earth Surface Processes and Landforms*. 36(11), pp. 1534-1549. 0197-9337.

Predicting leather handle like an expert by artificial neural networks

Wang, Y., Picton, P., Turner, S. and Attenburrow, G. (2011) Predicting leather handle like an expert by artificial neural networks. *Applied Artificial Intelligence*. 25(2), pp. 180-192. 0893-9514.

Investigation of morphological and electrical properties of the PMMA coating upon exposure to UV irradiation based on AFM studies

Darowicki, K., Szociński, M., Schaefer, K. and Mills, D. (2011) Investigation of morphological and electrical properties of the PMMA coating upon exposure to UV irradiation based on AFM studies. *Progress in Organic Coatings*. 71(1), pp. 65-71. 0300-9440.

All School of Science and Technology NECTAR publications

Figure 2: Export of publications from NECTAR to School research web page

Figure 3: Export of publications from NECTAR to staff profile

Advocacy and promotion

The project activities were supported throughout by a programme of advocacy. This took the form of training sessions for researchers; presentations to School research groups; talks to professional colleagues; and a variety of written advice and guidelines.

Our greatest effort though, was focused on a major event to celebrate Open Access Week 2011. We gave advance notice of the event as early as June, and sent out personal invitations (Figure 4) to researchers from September onwards. We invited prestigious external guest speakers and even our Vice Chancellor was keen to attend. Unfortunately, the timing of the event (in October) proved to be inconvenient for the majority of staff and uptake was so low that we eventually postponed the event.

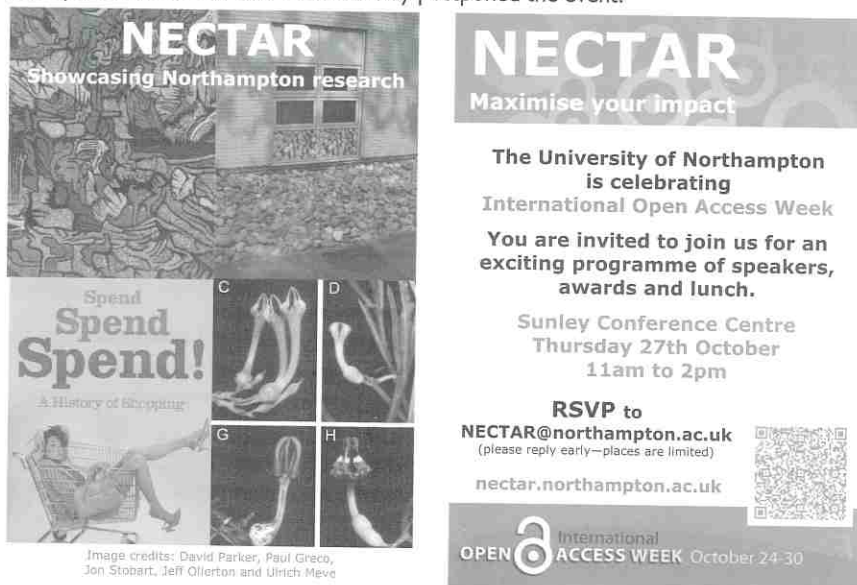


Figure 4: Postcard style invitations to our Open Access Week event

Project outcomes

By the end of the project we had seen a 50% increase in the number of live NECTAR items displaying open access content. Even more items were awaiting processing (e.g. copyright checking, author permissions and coversheets) in the review area. These will eventually translate into further full text availability.

With the exception of the quieter summer period, NECTAR visitor statistics had shown steady growth throughout the year, as had the number of full text downloads. Two of the university's six Deans had put in requests for NECTAR usage statistics and exploratory discussions had been held with the Research and Strategic Bidding Office regarding the reporting of NECTAR activity.

By January 2012, almost 300 researchers were registered NECTAR users and one third of these had recorded details of their work in the repository.

Strong positive feedback had been received from researchers following the addition of NECTAR bibliographies to staff profiles and, as a result of this new visibility, there had been a significant increase in the number of queries from staff regarding their work in NECTAR.

In terms of both researcher engagement and full text availability the 'Bringing a buzz to NECTAR' project was therefore a success.

The way forward

One element of the project is yet to be completed. Eprints Services are working on new software which will enable NECTAR (and other Eprints repositories) to assemble and submit the necessary publication information to the Research Excellence Framework (REF). When implemented, this will further reinforce the value of NECTAR to the university's research community.

As a result of a recent restructuring of professional services at Northampton, the Research Support team has gained a new member of staff. This person will spend at least half of their time supporting NECTAR, in particular focusing on the processing of full content and the development of new repository services.

The University's Vice Chancellor is a powerful advocate of open access (Petford, 2011) and under his leadership we hope to continue to advance the open access agenda. As well as the repository, we already have a (small) collection of open access journals (The University of Northampton, 2011b) and a research data policy that promotes open access to research data (The University of Northampton, 2011c). With such high level support we are optimistic that the future of open access at The University of Northampton is assured.

References

1. British Library Board (n.d.) *ETHOS – Beta. Electronic theses online service* [online]. London: British Library. URL: <http://ethos.bl.uk> [accessed 02.02.12].
2. Brody, T. And Gray, A. (2009) Kultur Eprints Extension [online]. URL: <http://files.eprints.org/466/> [accessed 02.02.12].
3. Field, A., Brody, T., François, S. and Tarrant, D. (2012) IRStats 1.0.3. [online]. URL: <http://files.eprints.org/722/> [accessed 03.02.12].
4. JISC (2011) *Repositories: take-up and embedding* [online]. URL: <http://www.jisc.ac.uk/whatwedo/programmes/infl1/reptakeup.aspx> [accessed 30.01.12].
5. JISCmail (2011) *UKCORR-DISCUSSION* (Closed list). URL: <http://www.jiscmail.ac.uk/> [accessed 02.02.12].
6. Kultur (2009) *Welcome to the Kultur Project* [online]. URL: <http://kultur.eprints.org/> [accessed 02.02.12].
7. Petford, N. (2011) A Vice Chancellor's view of Open Access [online]. URL: <http://www.northampton.ac.uk/info/20137/research/806/open-research-at-the-university-of-northampton/2> [accessed 03.02.12].
8. Pickton, M. (2008) Gathering NECTAR at The University of Northampton. *ALISS Quarterly*, 3(4), pp. 33-38. Also available from URL:
9. The University of Northampton (2007) NECTAR: the Northampton Electronic Collection of Theses and Research [online]. URL: http://nectar.northampton.ac.uk/NECTAR_briefing_sheet_200607.pdf [accessed 30.01.12].
10. The University of Northampton (2011a) *Research in the School of Science and Technology* [online]. URL: <http://www.northampton.ac.uk/info/20521/research-in-the-school-of-science-and-technology> [accessed 02.02.12].
11. The University of Northampton (2011b) *Northampton open journals* [online]. URL: <http://journals.northampton.ac.uk/> [accessed 03.02.12].
12. The University of Northampton (2011c) *Research data policy* [online]. URL: <https://www.northampton.ac.uk/info/20283/academic-research/1606/research-data-policy> [accessed 03.02.12].
13. University of Southampton (2011) IRStats [online]. URL: <http://wiki.eprints.org/w/IRStats> [accessed 03.02.12].

LSHTM Research Online: Release the Beast!

Andrew Gray, Manager LSHTM Research Online.

London School of Hygiene and Tropical Medicine.

<http://researchonline.lshtm.ac.uk/> <http://lshtmresearchonline.blogspot.com/>



Welcome to LSHTM Research Online



Atom RSS 1.0 RSS 2.0

If you have any questions or comments about the repository please email researchonline@lshtm.ac.uk

This year in January, the 19th to be exact we had a 'soft' launch of LSHTM Research Online <http://researchonline.lshtm.ac.uk/> to the School. I had obviously wisely chosen this day since it was one of the days I did not work. After some email and phone call panics our staff had a chance to finally see what we had been working upon for the last year and what was obviously the highlight of their new year! So what reactions did we have? All were pretty good, though of course I have a sneaking feeling that for many staff and researchers their minds were elsewhere. However I think we were all pleased that we managed to launch our repository with over 19,000 bibliographic records of which 1,800 were full text. While nearly 10% is not an enormous percentage those who have experiences of repositories will know that it is still a pretty good level of full text but we are not smug just yet.

For those who don't know about LSHTM I will quote our website....

"The London School of Hygiene & Tropical Medicine is a world-leading centre for research and postgraduate education in public and global health, with 4000 students and more than 1300 staff working in over 100 countries. The School is one of the highest-rated research institutions in the UK, and was recently cited as one of the world's top universities for collaborative research."

While in terms of student figures and staff we could possibly be termed a small institution our research output belies this and we are a much bigger institution. I think one of defining characteristics is our global engagement. With over '100 members of staff living and working outside the UK' http://www.lshtm.ac.uk/research/about/overseas_staff_map.

html and 'close partnership with institutions in low- and middle-income countries' including a 'number of long-term, well-established major collaborations with institutions http://www.lshtm.ac.uk/research/about/collaborating_map.html where we have had a senior member of LSHTM staff based for more than five years'

So having such a global institution not only with many international students but researchers working and collaborating in many areas of the world an institutional repository should help to further remove barriers of distance, geography and finance from world class health research and that is still the aim.

I suddenly realise it's been 16 months since I've been here and we started developing and building the repository, so I will attempt a quick recap to give a flavour of what happened. The initial few months were given over to familiarising myself with the structure and workings of LSHTM, meeting the various departments who would have early input and continued input into the development of the repository and establishing a steering group and a technical support group drawn from the members of these departments. Alongside this I began to review the various different repository software packages and options around for hosting. I'd be a cheeky barefaced liar if I did not mention the fantastic resources and information that I used from the RSP website for this. We invited a number of repository software companies to give us demonstrations of their products including Islandora (<http://islandora.ca/>); Digital Commons (<http://digitalcommons.bepress.com/>); Eprints (<http://www.eprints.org/>); University of London Computer Centre (<http://www.ulcc.ac.uk/>); Open Repository (<http://www.openrepository.com/>) and ContentDM (<http://www.contentdm.org/>). Those companies based overseas gave us online demonstrations while those based within the UK were happy to visit, provide a demo and discuss our requirements. Without much ado I'd like to announce that we decided to use the Eprints software and commission University Computer Centre London to build and host our repository. The reasons for our choice of Eprints were various, I happily admit that my previous repository experience working on the Kultur project (<http://kultur.eprints.org/>) UAL Research Online (<http://ualresearchonline.arts.ac.uk/>) and GALA (<http://gala.gre.ac.uk/>) were all with Eprints repositories so this may have created a certain bias but I truly felt that after reviewing the other providers that Eprints best suited our needs. We are fairly 'late' to creating a repository and believed that there was a certain safety in numbers regarding the software and Eprints is the most widely used within the UK, it also seemed to be most responsive to the changing needs of the academic research community with new developments and releases imminent and their own engagement with requirements for the upcoming Research Evaluation Framework. Deciding between Eprints Services and University Computer Centre London was a harder decision but ULCC had either built or hosted a number of repositories within the University of London, of which we are a member; and again speaking with these institutions and getting good feedback from those who used them gave them the edge. Of course the actual physical proximity of ULCC's offices (across the road) meant that we could meet face to face when we want to and they would be able to attend meetings for questions, praise and scorn. ULCC's experience in hosting repositories and LSHTM's lack of experience in repository software along with the knowledge that IT and the webteam's focus for the immediate future was the new school website meant that we opted for the hosted option.

One of the interesting situations we had was the fact that LSHTM already had their own publications database, which was built in house and maintained by the fantastic Andy Reid. This had great advantages in that there was in house knowledge and experience of harvesting information from external databases such as Web of Science and was already doing this and tackling issues around author authentication. However this role had grown each year and as Andy said to me at the beginning 'if I'm knocked over by a bus then no one will know how to work the publications database'. Sadly.....

I am not a very technical person but I'm pretty good at getting people to talk to each other. A lot of work was done on ensuring that our repository could harvest data from our publications list and then this was in turn matched against data from UK PubMed Central (<http://ukpmc.ac.uk/>) to check for any full text publications. Any full text publications that were open access and part of a specific subset we pulled into the repository. Our initial excitement at finding nearly 3000 full text items in UKPMC was cooled by the dawning realisation that even though the full text material was within UKPMC and could be downloaded, not all of it could be hosted within our repository due to pesky copyright restrictions. As previously mentioned we pulled in 1,800 full text but now we need to go through the other 1,200 full text articles in UKPMC to determine whether or not the license attached to that particular piece of full text is actually fully open access and will allow us to host it within LSHTM Research Online.

Other fantastic problems included preventing duplicates appearing in our mass harvest of records and then to continue detecting them. I would like to say that there are absolutely no duplicates but sometimes I do find I'm zapping a couple before I do a demo for researchers! Like everyone else we also had the issue of unique identifiers! We don't have truly unique identifiers for staff. Our initial hope to rectify this within the school hit a brick wall when we confronted the immense task in driving this through the School but it is something that we all hope for one day. So we stuck with the staff ID's that are derived from Personnel and already used by the Publications database. Of course I would now like a ban on anyone marrying, changing their name, leaving or returning to the School.

For those looking at our repository you may notice that we do not have subject headings!! This was not the initial plan and it took quite a bit of time and effort to decide not to have them. We knew that Library of Congress headings would not work for our disciplines so looked around and after speaking to Liverpool School of Tropical Medicine decided to look at National Library of Medicine classification; the very kind librarian Sarah Lewis Newton from Liverpool sent through to me their own sweat driven taxonomy which matched their disciplines with NLM. We then decided to look at how our own subject headings system Barnard Classification matched that of NLM. This took forever! When we discovered that so much was absent or mismatched e.g. NHS was in NLM classified as a 'function of the state'. We then looked at the research headings used by faculty on the website and thought maybe this could be an answer, admittedly uncontrolled but an answer. Our steering group quickly pointed out that a lot of those headings didn't even begin to cover the areas of research and that they were out of date. As I wandered the lonely low-lit pages of other institutional repositories I became more aware of how certain institutions did not even bother with subject headings! Suddenly this seemed to make sense, saved time and provided relief! (Sorry librarians I tried)

We have been lucky in comparison to other new repositories in having a considerable amount of full text available from the launch. Open Access publishers such as PLoS (<http://www.plos.org/>) and Bio Med Central (<http://www.biomedcentral.com/>) of course have contributed greatly to this. We are also in receipt of a Wellcome Trust grant to pay for Open Access, so together they provide a good and steady amount of full text for the repository. However this does not cover all the research we publish, so at present we are in the process of convincing researchers to send us their final peer reviewed version. This of course has brought up lots of questions and misunderstandings. To gauge the engagement of our researchers and to help clarify a procedure for deposit we began a trial with the Medical Statistics department facilitated by one of our steering group. We initially had 5 requests for the department

1. Look at their list of publications (we sent them the url)
2. Indicate if a publication belonged to a research centre/group
3. Provide keywords
4. Indicate the funder for any publication
5. Provide us with 'author accepted manuscript' for 2011 publications

There were approximately 30 people in this department and we initially ran the trial for a month. After a month we had only received 2 responses. It appeared that we had asked for too much information and people misunderstood the phrase 'author accepted manuscript' and we were sent the publisher pdf's. We extended the trial for another couple of weeks and this time just asked for the full text but phrased as 'the submitted version, after peer review without publisher pagination and typesetting'. This resulted in another 2 replies but we did get all the papers. While this is not a huge response it was not entirely unexpected and it was run over the Christmas period and unfortunately coincided with the School running a dummy REF exercise. However we learnt that we need to be clearer about what we want, the benefits, offer as much help and continue to get 'out there' to explain why they should submit their full text.

So the advocacy continues and slowly we are having more engagement. We have individual researchers contacting us, research centres suddenly interested in being able to feed lists of their publications to their own sites. We are continuing to go around to each department and to now demonstrate the site with its lovely new interface and the fact that we also have IR Stats and can provide some admittedly small stats about downloads has perked up more interest. We have an official launch with the Director Peter Piot speaking on 18th April and we will use this as another springboard for penetrating the minds of all!

There is of course plenty more we want to do to the repository and sometimes it seems every week we discover another 'glitch' but we carry on. Our plan is to wheedle our way into the hearts of the administrators in these departments and convince them that the repository can help them in monitoring the research produced by their departments, provide more online information, have drop-in sessions, use our popular copyright course to promote the repository, postcards, individual emailing, display of latest publications on the electronic notice board in reception, replacing all publication links on staff profiles from PubMed to that of the repository, working with the new school Marketing team

so that they can pull out research to present on their newsletters and in coordination with new events and of course bribery, blackmail, threats and sulking. Hopefully the combination of these will convince but come back to me this time next year.

Finally I'd like to say thank you to the many people who have worked extremely hard on this including but not limited to Emma Golding, Vivienne Eades, Andy Reid, Rory McNicholls and Jose Martin. Without you I'd be somewhere else.