

What's new ...

Digital Preservation Coalition | Digital Curation Centre Issue 24, 1st April 2010

In this issue:

- **What's on and What's New**
- **Editorial:** Ending the news-cycle that never ends - data management and the universe that turned to goo, William Kilbride, DPC
- **Who's who:** Sixty second interview with Anusha Ranganathan, Oxford University Library Services
- **One World:** Inge Angevaare, Netherlands Coalition for Digital Preservation
- **Your View:** commentary, questions and debate from readers

What's on:

JISC Conference 2010

12-13 April

Retrieved 31 March 2010 from: <http://www.jisc.ac.uk/events/2010/04/jisc10.aspx>

The theme 'Technology: at the heart of education and research' confirms JISC's view that universities and colleges must integrate technology into all aspects of their strategic planning.

Strategic Content Alliance Brunch Briefing on Business Modelling and Sustainability

14 April

Retrieved 31 March 2010 from: <http://www.jisc.ac.uk/events/2010/04/scabbbs.aspx>

This briefing aims to improve your professional knowledge and understanding of sustainability planning for digital projects and will encourage delegates to consider in more depth the range of business strategies open to the digital project. Recommendations are taken from recent, ground-breaking research which offers an insight into real world case studies with financial measurements in order to inform strategic planning.

Digital Preservation - The Planets Way

19-21 April

Retrieved 31 March 2010 from: <http://www.planets-project.eu/events/rome-2010/>

The final in the series of three-day training events, 'Digital Preservation - The Planets Way', will take place in Rome on 19 - 21 April 2010.

Blue Ribbon Task Force on Sustainable Digital Preservation and Access

6 May

Retrieved 31 March 2010 from: <http://www.jisc.ac.uk/events/2010/05/brtf.aspx>

The Blue Ribbon Task Force will be presenting their final report Sustainable Economics for a Digital Planet: Ensuring Long-term Access to Digital Information at this one day free symposium. The report provides an economic framework and practical recommendations for ensuring the future sustainability of digital preservation and access. The symposium will allow UK and European stakeholders to respond to this report through a mixture of panel and audience discussions and presentations from the senior figures.

CETIS meeting: Repositories and the Open Web

19 April

Retrieved 31 March 2010 from: http://wiki.cetis.ac.uk/Repositories_and_the_Open_Web

The theme of the meeting is looking web 2.0 / social sharing sites compare to repositories when used for the management and dissemination of learning materials.

EGEE User Forum

12-15 April

Retrieved 31 March 2010 from: <http://indico.cern.ch/conferenceDisplay.py?confid=69338>

Enabling Grids for E-science (EGEE) is Europe's leading grid computing project, providing a computing support infrastructure for over 13,000 researchers world-wide, from fields as diverse as high energy physics, earth and life sciences. The 5th EGEE User Forum will be held in collaboration with EGI and NDGF in Uppsala, Sweden, April 12-15, 2010, hosted by SNIC, UPPMAX and PDC.

What's New:

ESRC FRAMEWORK FOR RESEARCH ETHICS (FRE) 2010

Retrieved 31 March 2010 from:

http://www.esrcsocietytoday.ac.uk/ESRCINFOCENTRE/OPPORTUNITIES/research_ethics_framework/

The Research Ethics Framework (REF) for social science research was formally introduced by ESRC in January 2006. The dignity, rights and welfare of research participants are core criteria. Since that date, we have only funded research where consideration has been given to ethical implications, and in those institutions where appropriate arrangements are in place. The Framework therefore has implications for applicants to ESRC, research ethics committees within HEIs and for those assessing research proposals. ESRC have always made it clear that the ethics framework should be a 'living' document which we are committed to developing over time. Updates to the document will take into account the experiences of the ESRC and Research Organisations as well as future challenges arising from the wider ethics environment. The Revised Framework for Research Ethics (FRE) is now available.

IDCC10– Call for papers

Retrieved 31 March 2010 from: <http://www.dcc.ac.uk/events/conferences/6th-international-digital-curation-conference/call-papers>

IDCC10 will be presented jointly by the Digital Curation Centre, UK and the Graduate School of Library and Information Science, the University of Illinois at Urbana-Champaign, and in partnership with the Coalition for Networked Information (CNI). The Programme Committee invites the submission of papers that reflect developing trends in curation and address the issues of growing the curation community to meet the challenges of the next decade.

JORUM Newsletter

Retrieved 31 March 2010 from: <http://www.jorum.ac.uk/news/newsletter.html>

The March edition of the Jorum Update newsletter is now available. This packed issue highlights these activities and features an article on using Twitter to help promote your resources in JorumOpen.

Planetarium – the ninth newsletter from Planets

Retrieved 31 March 2010 from: http://planets-project.eu/docs/newsletters/Planets_newsletter9_March2010.pdf

The Planets project has just released its 9th issue of Planetarium. This issue features a spotlight on Preservation Planning: Protecting Files For The Future, presents research in the impact of transition from analogue to digital communication on researchers and civil servants, and also contain an article about how the Planets research partners develop tools and services and put Planets into practice. In addition you will find an article about Planets and the OAIS. The newsletter also contains a News Round Up and a list of Planets publications and presentations at events.

Assessing the Future Landscape of Scholarly Communication: An Exploration of Faculty Values and Needs in Seven Disciplines

Retrieved 31 March 2010 from: http://escholarship.org/uc/cshe_fsc

Since 2005, the Center for Studies in Higher Education (CSHE), with funding from the Andrew W. Mellon Foundation, has been conducting research to understand the needs and practices of faculty for in-progress scholarly communication (i.e., forms of communication employed as research is being executed) as well as archival publication. The complete results of this work can be found at the Future of Scholarly Communication's project website. This report brings together the responses of 160 interviewees across 45, mostly elite, research institutions to closely examine scholarly needs and values in seven selected academic fields: archaeology, astrophysics, biology, economics, history, music, and political science.

Knowledge Exchange: Successful Reuse Of Research Data Requires Integrated Approach

Retrieved 31 March 2010 from: <http://www.knowledge-exchange.info/Default.aspx?ID=62&M=News&NewsID=77>

On 23-24 September 2009 an international discussion workshop was held in Berlin, prepared and organised by Knowledge Exchange. The main focus of the workshop was on the benefits, challenges and obstacles of re-using data from a researcher's perspective. The presentations are now available to view online.

Ending the news cycle that never ends – data management and the universe that turned to goo (William Kilbride, DPC)

There's an old joke that being brainy isn't as much fun as it used to be. In the good old days boffins furnished us lesser mortals with such useful things as gravity, electricity, King Tut, laughing gas, the triple condensing boiler, clean water and salvation as the limit of evil. All good things as Sellar and Yeatman would say. In our generation they've discovered the H-bomb, semiotics and cryptosporidiosis. The company of smart people, we may whimsically conclude, is not the boon it used to be.

Looking at the press it's certainly possible to put empirical evidence behind this fanciful reckoning. Headlines like the ones quoted below from the last six months give a deeper and more concerning echo of how scientists and science are sometimes presented, perceived and abused in popular imagining. They show an unhappy band of researchers trapped for good reasons and bad in a news cycle that never ends:

- **'Callous, unethical and dishonest'** (The Sunday Times, 31/01/2010, reporting on Dr Andrew Wakefield's research on the MMR jab);
- **'Fatuous, dangerous and utterly irresponsible – the Nutty professor who's distorting the truth about drugs'** (The Daily Mail, 04/11/2009, welcoming the removal of the UK Government's chief advisor on drugs)
- **'Disgraced cloning expert convicted of fraud in South Korea'** (New York Times, 26/10/2009, reporting the trial of Korean geneticist Hwang Woo-suk)

These are not the same and no comparison is intended. The point is simpler: who would be a researcher? Galileo you will remember got a tough time of it but at least he was pardoned (eventually). And, when our main view of research is through the prism of the media, with its righteous pomposity and fascinations with frailty, vanity and celebrity, how can a reasonable and circumspect citizen trust what they are told? The well-intended *'Large Hadron Collider will not turn world to goo promise scientists'* (The Times, 06/09/2008) gives incomplete solace.

A recent nadir has been the inexorable indignity of 'Climategate' (or **'Climategate Prof raked in \$22.6 million in grants'** as you would know it from The Detroit News on 29/11/2009). There's nothing good about this story. Prof Phil Jones and colleagues at the University of East Anglia have for several years been pursuing research on climate change using data gathered from multiple meteorological sources. This data provided the basis for some important conclusions which in turn became influential in the climate change debate. Prof Jones's view is that 'we face enormous challenges ahead if we are to continue to live on this planet.' But research invites scrutiny and scepticism which, in Prof Jones's case, meant numerous Freedom of Information requests to see data which was not always his to provide. Lively and at times colloquial discussion about these requests formed the basis of email exchanges between colleagues at UEA. In November, curiously close to the start of the Copenhagen climate summit, these emails were taken illegally from a server at UEA and published. These emails 'do not read well' and have been used to question the Prof Jones's conclusions. They have also lead to an extended, unpleasant and often-times ill-informed furore about the value, virtue and independence of research. Prof Jones has stood aside as director of his research unit, he has had to defend his research in front of a parliamentary committee and there are now two independent reviews of the centre in progress.

I've seen this story spun many different ways since November but I've still not seen a positive angle - from any commentator. A favourite approach has been to focus on process rather than the substantive issues, deflecting and distorting the substantive issues that need thorough and measured analysis. It is rueful comfort that the parliamentarians who questioned Prof Jones have had their own information management problems recently too: and that leaked documents about moat cleaning and floating duck islands provide more lurid headlines than data from Chinese weather stations. Iain Gibson, Norwich North's erstwhile MP, former chair of Commons Science and Technology Select Committee, and former Dean of Biology at UEA would likely agree.

Information management is not trivial and it's not new: but dependence on digital sources for research and a massive increase in data volumes and complexity means that researchers face new challenges. When research intersects with government, when researchers engage with the cut and thrust of political debate, and when data is used to inform policy, then we can be certain that conclusions *and the data sources on which they are*



based will be questioned. Confident data management, short and long term, will be essential if we are to keep our focus – and the media's focus – away from process and on the really important issues. Transparency, clarity and a new kind of professionalism with information management will benefit parliamentarians and researchers alike.

Let's not underestimate what's at stake here. It's difficult to argue with an opponent who buys ink by the barrel and speaking truth to power can be deadly. In the digital age it's all too easy to make fanciful statements on partial and inadequate evidence cheaply downloaded, fabricated, altered and shared. We can prove almost anything that way. The academy's only safeguard is the integrity of research, demonstrated by scrupulous documentation of source and process.

Researchers and politicians might not be able to end the self-propelled data management news cycles that have caught them out, but we can at least prevent ourselves getting drawn into new ones. A highlight of my first year at DPC was an invitation to help launch a new and major programme from JISC – dedicated to improving research data management. The proposition for the first strand of work is remarkably simple: to identify what researchers need to do to improve data management; then to pilot the implementation of infrastructures that will deliver these improvements. In the short term this will be good news for the projects funded, and in the medium term it will be good news for the wider research community. But the lessons are of wider relevance. Some of the tools embedded in this programme – like the Data Audit Framework, AIDA and DRAMBORA – have been widely trailed in the digital preservation community.

Thorough and heterogeneous assessment of these tools and the refinements that follow could bring benefits to diverse organisations in many sectors: and a very good reason to hang around with brainy people again.

Who's who:

In this section we invite a DPC member working on digital preservation to introduce themselves and their work. It will help us build connections between staff working on similar issues, will help promote the work of members, and will help identify priorities for the development of shared tools or services This time we hear from ...

Anusha Ranganathan, Oxford University Library Services



Where do you work and what's your job title?

I work at the Bodleian Libraries of University of Oxford as a DAMS (Digital Asset Management System) software engineer. We are a small team and work on many different projects, dealing with all aspects of digital data.

Tell us a bit about your organization

The Oxford libraries are home to 400 years of research not only housing special collections of rare books, manuscripts, archives, maps, music, and other primary research collections, but also digital books of the 20th century and current research output (thesis, publications, books, journals) generated by the Oxford community. Being a legal deposit library, this collection is constantly growing, along with a lot of effort put into digitizing traditional paper collections. Our aim is to develop a system for the life-cycle management of all these digital data, enabling continued future access.

How did you end up in digital preservation?

Digital preservation was a natural extension to the work I was doing as a software test engineer, where the need was to maintain software and manage test data and results. So I jumped at the opportunity to work on digital preservation when it was offered to me by Oxford University. It gives me the opportunity to be at the bleeding edge of software technology.

What projects are you working on at the moment?

I am working on a couple of projects. Blockbooks is a digitization project. Books printed from woodblocks and metal cut prints from the 15th and 16th century, taken from Bodleian's special collections are being digitized. They represent early European printing of the 15th century, combining words and images in elaborate but easily-reproduced series. These high quality digital images along with all the metadata are being made available online and preserved in our digital repository. Explore the special collections at <http://www.odl.ox.ac.uk/digitalimagelibrary/index.html>.

The second project, Building the Research Information Infrastructure (BRII) enables efficient sharing of research activity data (i.e. data about research) at the University of Oxford. It involves harvesting data from various websites and open data stores using semantic web technologies, to not only re-use the data in new

ways and exploit the powers of collating data, but also archive the research activity data for future, long past the availability of local systems. This registry of information can be searched and is made available to web managers to extract in a number of different formats using simple APIs. More information can be obtained from <http://brii.bodleian.ox.ac.uk/>

What are the challenges of digital preservation for digital object repositories?

Digital preservation is all about preserving digital objects for posterity. We not only have to deal with availability of resources (or the lack of it), but need to ensure the digital object remains useful. This means continued access and meaningful information about the object is made available along with today's challenge in deciding the worthiness of a digital object for preservation.

What projects would you like to work on in the future?

My family is home to three accomplished authors, each specializing in different fields and from a different era. We have the joy of being caretakers to their wonderful and vast hybrid libraries, along with several original manuscripts. It would be a great achievement if we could get down to actively working on preserving them.

What sort of partnerships would you like to develop?

Collaborations leading in developments in programming applications for digital curation, specializing in semantic web technologies is something I would like to build on.

If just one tool or standard could be brought into existence that would make your job easier, what would it be?

I have slowly come to realize the truth in this statement - working on the technology to be used to manage the life-cycle of digital preservation, though complex by itself, is an easier task at hand. The difficult task still is advocating digital preservation, stakeholder buy in and getting people to realize it costs to preserve digital data.

If you could save for perpetuity just one digital file, what would it be?

The works of the renowned classical Sanskrit writer Kalidasa from the period between the 4th and 6th Century.

Anything else you want to share with us?

Life in the internet world is a great boon, but it has its needles. Ownership of data published online is something that concerns me, as we seem to forget about privacy and are willing to share our life experiences, but are not equally concerned about the rights to our data.

Finally, where can we contact you or find out about your work?

To find out more about the libraries at the University of Oxford visit <http://www.bodleian.ox.ac.uk>. The best way to get in touch with me is by emailing me at anusha.ranganathan@bodleian.ox.ac.uk

One World:

In this section we invite a partner or colleague to update us about major work on their home country that will interest readers. In this issue we hear from...

Inge Angevaere,

Netherlands Coalition for Digital Preservation

Faced with the same digital challenges as the UK, it was the very example of the DPC that inspired ten Dutch public sector organisations to establish their own *Netherlands Coalition for Digital Preservation* (NCDD) in 2008, with the express and very ambitious aim of bringing about a national infrastructure for permanent access to digital information in the public sector. A national digital preservation survey was to lay the groundwork for a national policy, and an English-language summary of the report, [*A future for our national memory*](#), was published in the Summer of 2009.



The report concluded that a number of pioneering organisations are now able to secure preservation at bit level (see the report for details), but that large categories and quantities of digital information are still at risk, not so much because technical obstacles cannot be overcome, but rather because these data simply never reach digital archives with proper facilities. Producers of digital information are either unaware of the risks their information runs or their interest is chiefly in short-term use and they are not motivated to put in the extra work involved in long-term preservation. The pattern which emerges is that long-term access is relatively secure in information chains that have changed little since the onset of the digital age. Publishers of scholarly information are a case in point. Where the world has changed more drastically, such as on the internet with its billion suppliers of digital input, the situation is not even close to being addressed. In the same vein, digital

copies of printed collections are relatively secure, as they were produced in the context of institutions with long-term remits, while born-digital information is at far greater risk. All in all, as yet there seems to be no direct correlation between the value and/or uniqueness of digital collections and the degree to which their access has been secured in the long run.

Whoever tries to address this discrepancy is immediately faced with the enormity of the digital 'phenomenon'. All too often, stakeholders are overwhelmed to such an extent that they put off decisions – interminably. The Dutch advocate a pragmatic attitude to ward off this paralysis. Users are the designated community of long-term access, but as many of them have not yet been born they can hardly be relied upon to take the initiative. Producers are similarly unlikely candidates, as their actions are determined by short-term goals. Therefore, the initiative must lie with archiving institutions – but those of a 'digital' frame of mind, i.e., keenly aware of their position in information *chains* and their dependency on both producers and users.

The strategic agenda the NCDD is now drafting is based on the premise that digital preservation is too enormous a challenge for single institutions or even sectors to take on. Therefore, cooperation and collaboration are called for. But at what level? The NCDD distinguishes four sectors (government/archives, scholarly communications, audiovisual media and cultural heritage) and the national survey has shown that in many respects these sectors have very distinct dynamics. Consequently, a substantial amount of work on standards, facilities, protocols and knowledge should be organised at sector-level. As digital preservation is, furthermore, a discipline requiring substantial financial investments, manpower, and specialised knowledge and expertise, the NCDD decided to designate a 'prime mover' organisation within each sector to facilitate the development of an infrastructure for digital preservation. Four such 'prime mover' organisations have now been named: the [Koninklijke Bibliotheek](#) (KB), national library of the Netherlands, for scholarly publications; [Data Archiving and Networked Services](#) (DANS) for research data, the [Netherlands Institute for Sound and Vision](#) for audiovisual media and the [National Archives](#) for public records. In the cultural heritage sector no 'natural' prime mover could be identified, but the Ministry of Culture has acknowledged the need for such an organisation and is actively involved in finding one.

The NCDD is the platform where the four/five prime movers meet to discuss national policies, the division of labour between sectors, and matters of mutual interest, such as mass storage. Also it is the arena where all other custodians of digital objects show their commitment to the long-term cause by their membership and by sharing knowledge and expertise at a national level.

The Ministry of Education, Culture and Science has expressed its support for the NCDD approach to digital preservation by matching membership dues for the NCDD for the next four years, but funding for infrastructural works in the sectors has been delayed by budget cuts caused by the credit crunch.

Nevertheless the NCDD partners will do what they can within their present budgets to further a national infrastructure. To name just a few national initiatives: the KB e-Depot harvests the institutional repositories of the universities and is developing a national web archive; the National Archives have reached agreement in principle with all stakeholders to establish a national authority for permanent access to all government records; DANS has presented plans to evolve from a data archive for the humanities and social sciences into a discipline-independent data archive, and Sound and Vision is further developing a knowledge bank for audiovisual objects and a hosting service for audiovisual archives from third parties.

The NCDD provides the overall context – and keeps beating the drum for the need for long-term preservation in The Hague where the Dutch government resides.



A coalition of coalitions meets in October 200

Left to right: Inge Angevaare of the Netherlands Coalition for Digital Preservation (NCD) Abi Potter of Library of Congress NDIIP, Sabine Schrimpf of NESTOR, Martha Anderson of Library of Congress and William Kilbride of the DPC.

Your View: commentary commentary, questions and debate from readers

Get involved! Comment online at: <http://www.dpconline.org/newsroom/what-s-new/>

Credits

What's New is a joint publication of DPC and DCC and is compiled by Najla Rettberg. Thanks to contributors. If you would like to feature in What's New or have a contribution to the news or events, then please contact us.

For more information on the DPC see <http://www.dpconline.org/>

For more information on the DCC see <http://www.dcc.ac.uk/>