

What's new ...

Digital Preservation Coalition | Digital Curation Centre Issue 22, 1st February 2010

In this issue:

- **What's on and What's New**
- **Discussion:** Winning Formulas and the Enduring Web (William Kilbride, DPC)
- **Who's who:** sixty second interview with Iain Flemming, Public Record Office of Northern Ireland.
- **One World:** The Library of Congress National Digital Information Infrastructure and Preservation Programme

What's on:

Digital Futures Academy

1-5 February 2010

Retrieved 1 February 2010 from: <http://www.digitalconsultancy.net/digifutures/>

Led by experts of international renown, Digital Futures focuses on the creation, delivery and preservation of digital resources from cultural and memory institutions. Lasting five days, Digital Futures is aimed at managers and other practitioners from the library, museum, heritage and cultural sectors looking to understand the strategic and management issues of developing digital resources from digitisation to delivery.

Digital Preservation - The Planets Way

9-11 February 2010

Retrieved 1 February 2010 from: <http://www.planets-project.eu/events/london-2010/>

Planets (Preservation and Long-term Access through NETworked Services) will host the fourth in its series of three-day training events at The British Computer Society, London on 9-11 February 2010. Digital Preservation – The Planets Way will examine the need to preserve digital content, the action that needs to be taken and the Planets approach to addressing these issues.

Research Data Management Forum: Dealing with Sensitive Data: Managing Ethics, Security and Trust 10-11 March 2010

Retrieved 1 February 2010 from: <http://www.dcc.ac.uk/events/data-forum-2010-march/>

This will be the fourth meeting of the Research Data Management Forum (RDMF). Aimed at researchers, digital repository managers, staff from library, information and research organisations, data curators, data centre managers, data scientists, research funding organisations and research networks, the event will address the topic "Dealing with Sensitive Data: managing ethics, security and trust."

Digital Preservation Training Programme, London

29-31 March 2010

Retrieved 1 February 2010 from: <http://www.dptp.org/about/dptp-london-soas-march-2010/>

The DPTP is a modular training programme, built around themed sessions that have been developed to assist you in designing and implementing an approach to preservation that will *work for your institution*. The course content is aimed at multiple levels of attendee: people from technical and archival professions come together, to learn the same standards and methods for digital preservation. The overall aim of the course is *critical thinking*, assessing ways of acting and planning at an organisational level.

What's New:

The January/February 2010 issue of D-Lib Magazine is now available

Retrieved 1 February 2010 from: <http://www.dlib.org/>

This issue contains eight articles, two conference reports, the 'In Brief' column, excerpts from recent press releases, and news of upcoming conferences and other items of interest in 'Clips and Pointers'. This month, D-

Lib features The Swingle Plant Anatomy Reference Collection, a historical collection of plant anatomical microscope slides, courtesy of University of Miami Libraries.

UKSG and NISO Release First KBART Recommendations for Improved OpenURL Data Supply

Retrieved 1 February 2010 from: <http://www.uksg.org/kbart>

UKSG and NISO are pleased to announce the first report by the KBART (Knowledge Bases and Related Tools) Working Group, a joint initiative that is exploring data problems within the OpenURL supply chain. The KBART Recommended Practice (NISO RP-9-2010) contains practical recommendations for the timely exchange of accurate metadata between content providers and knowledge base developers.

The changing role of Jorum: sharing using JorumOpen and JorumUK

Retrieved 1 February 2010 from: <http://jorumnews.blogspot.com/2010/01/changing-role-of-jorum-sharing-using.html>

As Open Educational Resources (OER) increasingly become part of the landscape of learning and teaching resources, Jorum has now opened up for the sharing of resources, created under Creative Commons (CC) licences, through its new collection JorumOpen.

Data Dimensions: Disciplinary Differences in Research Data Sharing, Reuse and Long term Viability

Retrieved 1 February 2010 from: <http://www.dcc.ac.uk/scarp/>

The Digital Curation Centre is pleased to announce the report "Data Dimensions: Disciplinary Differences in Research Data Sharing, Reuse and Long term Viability" by Key Perspectives, as one of the final outputs of the DCC SCARP project. The project investigated attitudes and approaches to data deposit, sharing and reuse, curation and preservation, over a range of research fields in differing disciplines. The synthesis report (which drew on the SCARP case studies plus a number of others, identified in the Appendix), identifies factors that help understand how curation practices in research groups differ in disciplinary terms. This provides a backdrop to different digital curation approaches.

Library of Congress Digital Preservation Newsletter

Retrieved 1 February 2010 from: <http://www.digitalpreservation.gov/news/newsletter/201001.pdf>

The January 2010 issue of the Library of Congress Digital Preservation Newsletter is now available.

BioMed Central Joins CLOCKSS to Archive its Open Access Articles

Retrieved 1 February 2010 from: <http://www.clockss.org/clockss/Home>

CLOCKSS is pleased to announce that it has partnered with BioMed Central, the world's largest open access publisher, to preserve over 60,000 articles in CLOCKSS's geographically and geopolitically distributed network of redundant archive nodes, located at 12 major research libraries around the world.

First issue of SCA's Digital Content Quarterly

Retrieved 1 February 2010 from: <http://sca.jiscinvolve.org/2010/01/10/digital-content-quarterly-watch-the-video-interviews/>

In this fast-paced, ever-changing environment the Strategic Content Alliance's Digital Content Quarterly (DCQ) provides a news round-up of digital content issues from around the world, thought-provoking features highlighting key debates in the field and regular columns from experts in areas that have most traction in terms of digital content provision: intellectual property rights and business modelling and sustainability.

Collaborative Business Model Changes Funding Structure

Retrieved 1 February 2010 from: <http://arxiv.org/help/support/whitepaper>

In a move to expand support for sustaining arXiv, Cornell University Library is broadening the funding base for the online scientific repository. Nearly 600,000 e-prints - research articles published online in physics, mathematics, statistics, computer science and related disciplines - now reside in arXiv, which is an open information source for hundreds of thousands of scientific researchers. arXiv will remain free for readers and submitters, but the Library has established a voluntary, collaborative business model to engage institutions that benefit most from arXiv.

Information-handling Training for Researchers - Towards a More Cohesive Environment

Retrieved 1 February 2010 from: <http://www.rin.ac.uk/our-work/researcher-development-and-skills/information-handling-training-researchers-towards-more-co>

The RIN is taking forward the conclusions from its 2008 report, Mind the skills gap: Information-handling training for researchers, to develop information competency among researchers. The RIN has set up a working group, comprised of representatives from organisations with an interest in improving the information-handling training environment. The group's first task was to frame a joint response to the Vitae consultation to the Researcher Development Framework, submitted in December 2009. During 2010 the group will identify examples of good practice across the UK in information-handling training.

SCARP Synthesis Study

Retrieved 1 February 2010 from: <http://www.dcc.ac.uk/scarp/>

Anyone involved in curating research data knows well that disciplines do things differently. Whether we see research domains as a thousand blooming flowers or a tower of babel, that diversity presents challenges for data curation. Through a series of seven case studies in different disciplines, the Digital Curation Centre SCARP project aimed to help understand these challenges and build on current digital data management and preservation practices in each case. One of this JISC-funded project's final outputs is a synthesis study. DCC is pleased to announce the report "Data Dimensions: Disciplinary Differences in Research Data Sharing, Reuse and Long term Viability" by Key Perspectives. The synthesis report draws on the SCARP case studies plus a number of others (identified in the Appendix), and identifies factors that help understand how curation practices in research groups differ in disciplinary terms. This provides a backdrop to different digital curation approaches.

JISC TechDis Launches New Free Tool To Improve Internet Accessibility

Retrieved 1 February 2010 from: <http://access.ecs.soton.ac.uk/ToolBar/>

The JISC TechDis toolbar has been created in partnership with the University of Southampton's School of Electronics and Computer Science. The free application, released as a beta version, provides a range of editing and speech tools to enable internet users to better access information and interact with web pages. This is particularly useful for researchers and learners in helping them to complete research and gather information for their studies.

JISC Briefing Paper: Assessing the Business Case for Standards

Retrieved 1 February 2010 from:

<http://www.jisc.ac.uk/publications/briefingpapers/2009/bpbusinesscaseforstandards.aspx>

A recent briefing provides advice to help people incorporate standards in their ICT-related business cases.

Learning Object Metadata (LOM) Standard

Retrieved 1 February 2010 from: <http://ariadne.cs.kuleuven.be/mediawiki2/index.php/LomWhatNext>

Five years after its finalization, the IEEE LTSC Learning Object Metadata (LOM) standard has been reaffirmed and corrigenda are under ballot. The LOM Working group is now considering how to evolve LOM. For this purpose, we request input and feedback from the community.

Discussion:

Winning formulas and the enduring web (William Kilbride, DPC)

I've been thinking a lot about web archiving just now.

It started in December when I was sorting through some bookshelves and came across 'The Whole Internet Guide and Catalogue'. This curio was most likely obsolete when I bought it. It offered a 200-page classified listing of (most of) the Internet with a synopsis of their content. If you wanted to know what the Internet was like in the summer of '94, then it's a good place to start.

It came to mind when the Department for Culture Media and Sport published its long-awaited consultation on changes to the UK's legal deposit regime. They seek views on how deposit libraries - in particular the British Library - should deal with the Web. Legal deposit works well for print. A seminal moment of my early career was when the British Library recovered and delivered an obscure journal of woodland studies from some remote stack: an apparently obsolete article in a discontinued journal of which there were only



five copies in the UK. It seems hard to believe now, but it was the final piece of a complicated jigsaw.

Legal deposit originates in the 16th Century. From that perspective we've not had much time to consider legal deposit for the Web. Our limited experience has led the Legal Deposit Advisory Panel to recommend that libraries should be empowered to take a copy of every website in the UK that is not protected with authentication. The current dispensation - which required agreement from the copyright holder before a copy could be made - was well-intentioned but not scalable. The alternative - doing nothing - would have relegated this part of the UK's cultural memory to unfunded proclivities in the private sector.

This is big news for digital preservation in the UK and is likely to have implications beyond the deposit libraries. But it's hard to imagine it's the last piece of the jigsaw. For one thing, the restriction that publications be accessed on the premises of the libraries will probably bewilder users. For another thing web technology has moved a long way since legal deposit was first mooted. Moreover the exclusion of authenticated content raises some questions. The business case of charged internet services depends on valuable, desirable content: so the collections from which posterity will benefit most will be out of scope. The value is perhaps the problem. There will be different perspectives and interests: the libraries, the publishers, the funders, the authors, the technologists and the users will all have view. Our hope must be to agree on the free stuff so we can move onto more complex discussions. In the meantime we rely on a piecemeal provision of commercial partnership and self help to secure the greater part of this generation's scientific legacy.

Closer collaboration in web archiving is the purpose of the Web Archiving and Preservation Task Force, to be launched in March. Recognising the need to collaborate for better policy development, and recognising the utility of engaging a wider community, the UK Web Archiving Consortium approached the DPC in November 2008 to develop a joint action plan. The new task force is a direct response to this. UKWAC was established to oversee the operation of the UK Web Archive. We are now well served with technology: policy has been slower to develop than the tools. Interest in web archiving and preservation has grown as more organisations have developed and implemented their own digital preservation policies. It is relatively to procure services – but what to harvest and what to do once it is harvested is a different story. The new group has identified collecting policies and access policies as the first stops on their roadmap for discussion. 2010 will be an interesting and eventful year for web archiving in the UK.

But web archiving has been on my mind for more mundane reasons. Regular visitors to the DPC website will have spotted some changes. We've embraced web 2.0 tools allowing much more frequent updating via Twitter, dynamic content and user commentary (though not yet fully deployed). It also allows us to update content more quickly. To do this we've physically moved our server and shifted the domain name with it. Anyone who has done this will know how tricky it can be. You need to make sure that all the old content is ported across to the new site, then you need to synchronise an elaborate string of changes, and finally you need to go back and check that all the links and re-directions work. Our process has been greatly aided by our patient, diligent and good humoured technology suppliers. The really galling thing is that the more successful you are the less people will notice.

Working for an organisation like the DPC can make even the most laid back developer slightly neurotic. 'Our digital memory available tomorrow' says the slogan on the home page: true but only if the redirects are accurate and the export has succeeded. We thought it wise to take precautions. A complete harvest of the old site was arranged with the UK Web Archives on the day before we threw the switch. But things never quite go to plan. The file transfer was executed a couple of weeks before we had loaded the slides and the report from our web archiving conference in July. And so, late on a December evening, I found myself raiding the web archive to download the presentations about web archiving, to load them onto the new site.

'Our digital memory available tomorrow' says the slogan on the home page: true, but only if the redirects are accurate and the export has succeeded

The redevelopment makes DPC more responsive to members and it includes this new edition of 'What's New'. What's new was originally intended as a quarterly round up of everything you might need to know about digital preservation. For 21 issues it has tracked the growth of the community and it has been of real utility to our members. But the growth in DP activities over the years has made it harder to compile and more opaque to the casual readers. Our new publications plan includes a commitment to lower the barriers to digital preservation - that it should be accessible to all. So 'What's new' is changing too. From this month it will become a monthly production done prepared jointly between DPC and DCC. A monthly news roundup, an

editorial, a profile of one of our members and a short overview from one of our partners overseas will hopefully restore the immediacy, retain the utility and improve our readability. What's New has been a success, and the old website served us well. Changing a winning formula to keep up with the times can be trickier than it appears, as those involved with Legal Deposit are all too aware.

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.

Sixty second interview with Iain Fleming, Public Records Office of Northern Ireland

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

I work in the Public Record Office of Northern Ireland (PRONI), based in Belfast, and my job title (in Civil Service speak) is 'Curatorial Grade E'. I work on 'all things' digital with my colleague Joy Ardill and also with colleagues from PRONI's Information Services section.

Tell us a bit about your organisation

PRONI is Northern Ireland's official archive. As such we aim to identify and preserve records of historical, social and cultural importance and make them available for the information, education and enjoyment of the public.

Under our primary legislation PRONI is the official place of deposit for public records in Northern Ireland. However, we have a dual role in that we also collect a wide range of archives from a huge variety of private sources. We also act as promoters of best practice in archive and records management in order to ensure that today's records will be available for future generations. Our dual role presents a particular challenge for digital preservation in that (while not always strictly true) in the world of public record keeping we can plan for the preservation of records from government Departments and non-Departmental public bodies, however, when it comes down to the private donor we may have to deal with...well who knows what?

What projects are you working on at the moment?

I am involved with two digital preservation projects at the moment. The first project, which is about to 'launch', is the Northern Ireland Web Archiving Project. In partnership with National Archives and the European Archive we are about to harvest and then preserve a number of Northern Ireland public sector websites. These sites will then be made available via PRONI's own website for our users. The second project, on a much larger scale, is to turn PRONI into a Trusted Digital Repository ...which looks so much more simple on paper!! As ever the reality is much, much different. PRONI as an organisation diverts most of its financial and staff resources towards managing paper records, whether private or public. Out of a staff of 80, there are two full time and one part time staff members working on all aspects of electronic records management – from implementing electronic disposal and retention in the Northern Ireland Civil Service (NICS) EDRMS, to ensuring digital continuity within all electronic record creating systems used by NICS and the wider public service, to planning how PRONI will be able to take in, maintain and then make available those electronic records selected for preservation.

We have engaged the help of consultants to create a business case arguing for the need to provide staff and resources to turn PRONI into a Trusted Digital Repository. Currently we are about half way through that process and hope to be able to take the findings of the business case to our Departmental Board for approval in the very near future. The next stage is implementing those recommendations which should be challenging, yet ultimately rewarding.

What are the challenges of Electronic Records Preservation?

Preserving a digital record for perpetuity. Keeping abreast of the public sector's uptake and subsequent use of electronic record creation systems. The cat has been well and truly let out of the bag in this respect and those who use these systems, or greenlit (greenlight? Perhaps clarify) their implementation, have given little or no thought to the long term management of the records they create within them.

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

At heart I am an historian. I studied history, ancient history and Byzantine history through school and then University. Understandably PRONI is a little thin on the ground when it comes to archives reflecting these



subjects but we do have some fantastic private collections that I would like to see digitized, made available on our website and preserved in our soon-to-exist digital archive.

What sort of partnerships would you like to develop?

I am keen for the Digital Preservation Coalition's Web Archiving Taskforce to be a real success and provide leadership and direction in what is rapidly developing into a distinct, very important, aspect of digital preservation. We have just completed our web archiving project but are keen to develop it into a more complete process capturing not just public sector websites but the whole range of web resources pertinent to Northern Ireland on the internet. I look forward to how the taskforce takes off over the next year and feel that it will be a real benefit to PRONI as we look to ways of developing our own web archiving programme.

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

This might not fall under the category of tool or standard but I would like to see a sudden, bolt of lightning from a clear blue sky realization of the importance of records & information management to business continuity and ultimately to digital preservation! This would make my job so much easier.

As I mentioned earlier my team consists of myself and my colleague Joy Ardill and we are responsible for all aspects of electronic records and information management for the Northern Ireland public sector. The situation in the public sector itself, whether Departmental or non-Departmental, is little better with small staffs and allocation of meager financial resources directed towards what is an essential function. I believe there is an appreciation of the benefits that electronic ways of working can bring but there is little attention paid toward the long term management of information created within these systems.

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

My iPod collection – it has become very precious to me! Although I am sure that really should count as lots of little digital files.

Anything else you want to share with us?

Although I am a firm advocate of all things digital and am deeply attached to my iPod, the idea of e-books leaves me cold. Until those who create applications like Kindle or Sony Books are able to successfully replicate the smell and feel of a brand new book I will remain a Luddite in this one instance...

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

I don't tweet, being Irish means that I could never get across what I am feeling in less than 140 characters, so you can visit PRONI's website at www.proni.gov.uk [Retrieved 1 February 2010] or email me at ian.fleming@dcalni.gov.uk for further information.

One World:

In this section we invite a partner or colleague to update us about a major digital preservation activity in their country. In this first issue we look to the US and hear from ...

William Lefurgy, The Library of Congress National Digital Information Infrastructure and Preservation Programme

The Library of Congress National Digital Information Infrastructure and Preservation Program is working to build on the accomplishments of its partnership network, which currently consists of more than 150 organizations from around the world. A key to this goal involves extending NDIIPP interaction within specific digital data domains. The Library has convened **meetings of experts to focus on public policy on the web, digital news and geospatial data**. Each of the meetings fostered a full discussion of information categories with value for preservation, as well as exploration of collaborative stewardship models. Participants suggested many ideas for next steps to improve preservation of and access to significant digital information assets that will inform future work.



NDIIPP is working to expand professional and public outreach activities. The program web site, digitalpreservation.gov, provides frequently updated news from the Library and its partners. A new feature, **On the Leading Edge**, retrieved 1 February 2010 from: focus on organizations from around the world; the series was recently initiated with an article about DigitalPreservationEurope. Future articles will examine institutions,

partnerships and projects that are breaking new ground with tools, services, policies and best practices. Popular parts of the digitalpreservation.gov, including the **Tools and Services** Inventory and **Sustainability of Digital Formats** are expanding. An associated web site documenting the **Federal Agencies Digitization Guidelines Initiative** is also presenting new information about guidelines, methods, and practices to digitize historical content in a sustainable manner.

Aiming to exploit multichannel marketing, the Library recently launched a Digital Preservation Video Series featuring short presentations aimed at professional practitioners and the general public. Running in parallel is a new **Conversations about Digital Preservation**, podcast series. The videos and podcasts are posted on digitalpreservation.gov as well on the Library iTunes U site. NDIIPP is also planning to launch a Facebook page in the near future.

The Library is continuing its close involvement with a number of pilot and demonstration projects. Work with DuraSpace will continue on the **DuraCloud** project, which provides “cloud-based” preservation support and access services for academic libraries, academic research centers, and other cultural heritage organizations. Substantial findings are also coming from the **Preserving State Government Information Initiative**. States participating in the project now total 35, and important work is underway to explore solutions for preservation of state and local government digital publications and records.

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/>

What’s New is online at: <http://www.dpconline.org/newsroom/what-s-new/>