

In this issue:

- [What's On](#) - Forthcoming events from January 2012 onwards
- [What's New](#) - New reports and initiatives since the last issue
- [What's What](#) - The Next Big Thing - William Kilbride, DPC
- [Who's Who](#) - Sixty second interview with Lee Hibberd, National Library of Scotland
- [One World](#) - Digital Preservation in Spain - Alice Keefer and Miquel Térmens, University of Barcelona
- [Your View?](#) - Comments and views from readers

What's New is a joint publication of the DPC and [DCC](#)

---

## What's On

The DCC have a number of events coming up that may be of interest to you. For further details on any of these, please see our DCC events listings at <http://www.dcc.ac.uk/events/>. You can also browse through our DCC events calendar to see a more extensive list of both DCC and external events.

### **Biomedical Research Infrastructure Software Service (BRISKit) one day workshop**

*19 January 2012*

<http://briskit.eventbrite.com/>

BRISKit will design a national shared service brokered by JANET to host, implement and deploy biomedical research database applications that support the management and integration of tissue samples with clinical data and electronic patient records. This workshop will introduce participants to the service and its potential benefits.

### **'What I wish I knew before I started' DPC Student Conference**

*24 January 2012*

<http://www.dpconline.org/events/details/38-studentconference?xref=38>

The DPC and the Archives and Records Association are pleased to invite students and researchers in archives, records management and librarianship to a half day conference on practical workplace skills in digital preservation. Hosted by University College London, and organised in partnership with the University of Aberystwyth and the University of Dundee, this mini-conference will bring a select group of leading practitioners together with the next generation of archivists, records managers and librarians to discuss the challenges of digital collections management and digital preservation. In a lively set of presentations and discussions, each of the speakers will be invited to reflect on 'the things I wish I knew before I started' - giving students an advantage in their own career development, and helping those who frame the curriculum a chance to extend their student's readiness for the workplace.

### **Emulation in Digital Preservation**

*24-25 January 2012*

<http://keep-cardiff.eventbrite.com>

The EC KEEP project will facilitate a free-to-attend workshop at the Novotel Cardiff focusing on the technical and legal aspects of emulation in the context of digital preservation. It will show the results of two and half years' work by KEE, as well as considering the main issues relating to media transfer tools and the use of emulation as part of a digital preservation strategy.

### **Preservation Of Complex Objects Symposia (POCOS) on Preservation of Games and Virtual Worlds**

*26-27 January 2012*

<http://www.pocos.org/index.php/registration>

Preservation of video games and virtual worlds presents challenges on many fronts, including complex interdependencies between game elements and platforms; online, interactive and collaborative properties; and diversity in the technologies and practices used for development and curation. This exciting two-day symposium will provide a forum for participants to discuss these challenges, review and debate the latest developments in the field, witness real-life case studies, and engage in networking activities.

### **Trust and E-journals**

*31 January 2012*

<http://www.dpconline.org/events/details/39-trust?xref=39>

Perhaps the most advanced part of the digital preservation community, the E-Journal sector has growing experience in fixing technical challenges and is supported by a well-developed - if complicated and at times dysfunctional - value chain that connects authors, publishers, sellers, purchasers and consumers. A range of service providers and tools now aim to secure this supply chain with digital preservation. Outsourcing - specifically knowing how to trust services that claim to provide digital preservation - has been one of the key barriers to preservation being adopted more widely so the experience of the E-Journal community is of much wider relevance

than just the library and academic community.

### **Data-Intensive Computing in Biology**

*6th - 8th February 2012*

<https://eventbooking.stfc.ac.uk/news-events/data-intensive-computing-in-biology>

A major theme of modern biology is the quantity and wealth of data available: the data deluge. Biologists need to be able to handle very large datasets, and to extract useful information and derive knowledge. Example applications include sequencing, metagenomics, proteomics, imaging and neuroscience. This workshop will look at the computational challenges in data-driven biology. It will bring together computer hardware and infrastructure experts with scientists involved in challenging data-driven disciplines. We will cover scientific challenges, instrumentation, computational platforms, data standards, and scalability of current applications.

### **DCC East Midlands Data Management Roadshow**

*7-8 February 2012*

<http://www.dcc.ac.uk/events/data-management-roadshows/dcc-roadshow-loughborough>

The Digital Curation Centre is running regional roadshows to support HEIs with research data management. The East Midlands roadshow is being organised with the Department of Information Science at Loughborough University and will take place 7-8 February 2012 at Burleigh Court Hotel and Conference Centre. The programme includes an exciting range of speakers who will share examples of developing research data management services and infrastructure. A draft programme is available online.

### **Hackathon – A Practical Approach to Database Archiving**

*7-9 February 2012*

<http://opfdatabasearchiving.eventbrite.co.uk>

The large and growing volume of data held in an increasing variety of relational databases presents a huge challenge to the archiving community. In this practitioner and developer hackathon we will run requirements workshops and facilitate hands-on hack-sessions to address the issues of preserving databases. We will consider SIARD as a potential archiving standard and review its implementation at the Danish National Archives as well as discussing other existing practical solutions, for example, MIXED and RODA etc. During the event there will be demonstrations of existing tools as well as opportunities to discuss alternative solutions such as emulation.

### **CERIF and euroCRIS meetings in Bath**

*9-10 February 2012*

<http://www.ukoln.ac.uk/events/cerif-2012-02/>

Registration is now open for the CERIF Tutorial and UK Data Surgery organised by UKOLN, euroCRIS and JISC. The day will comprise a CERIF tutorial followed by a data surgery in which it will be possible to examine the use of CERIF in real life scenarios - so please bring your CERIF queries and data modelling/mapping issues for discussion with euroCRIS CERIF experts.

### **RSP Workshop: How embedded and integrated is your repository?**

*10 February 2012*

<http://www.rsp.ac.uk/events/embedding-and-integrating-repositories/>

This event will present the 'Embedding Repositories: A Guide and Self-Assessment Tool' which was launched in December 2011. The embedding repositories guide and self-assessment tool has been published by the Repositories Support Project (RSP) and it aims to help institutions achieve the best value from their repositories through integration with other university systems, particularly research management systems.

---

## **What's New**

For more information on any of the items below, please visit the DCC website at <http://www.dcc.ac.uk>

### **Quick Survey: Depositing Research Data using SWORD**

<http://swordapp.org/2011/12/survey-depositing-research-data-using-sword/>

The SWORD v2 project has been asked by the JISC to look into the applicability of the SWORD protocol for depositing Research Data. The SWORD protocol has always been agnostic about the type of resource it is depositing, however its initial development stemmed from a requirement for the deposit of scholarly communications outputs into repositories – these typically being small text-based items. In order to investigate how well SWORD and SWORD v2 would deal with Research Data, we need to know about the different types of research data that you are working with. This will allow us to discover some of the range of different data types in use, and the general and specific requirements of each.

### **TIMBUS Survey on Long-Term Aspects of Business Continuity**

<http://timbusproject.net/news/>

[https://intel.qualtrics.com/SE/?SID=SV\\_1LWzcs15h0TliAk](https://intel.qualtrics.com/SE/?SID=SV_1LWzcs15h0TliAk)

TIMBUS builds on and extends existing research and experience in risk management and business continuity management on the one hand, and digital preservation on the other, aligning these complementary approaches. It is a strategic project co-funded by the European Union and it brings together expertise from across Europe from industry partners, SMEs and research partners. In order to focus the project on the practical needs of the community, the project team would like to hear from anyone in the DP community about the occasions when you have needed to repeat, replay or exhume dormant, possibly interdependent, processes to make sense of process outputs, to repeat experiments or to diagnose the old process.

### **Repositories Support Project Embedding Guide and Toolkit**

<http://www.rsp.ac.uk/embeddingguide/>

The Repositories Support Project Embedding Guide and Toolkit has been published by the Repositories Support Project and will help institutions to get the best value from their institutional repositories through integration with other university systems, particularly research management systems. It is aimed at repository staff but will be of interest to other groups such as academic librarians and research management staff.

### **Paving the way to an open scientific information space: OpenAIREplus – linking peer-reviewed literature to associated data.**

<http://www.openaire.eu>

OpenAIREplus (2<sup>nd</sup> Generation of Open Access Infrastructure for Research in Europe) was launched in Pisa in early December. The 30 month project, funded by the EC 7<sup>th</sup>

Framework Programme, will work in tandem with OpenAIRE, extending the mission further to facilitate access to the entire Open Access scientific production of the European Research Area, providing cross-links from publications to data and funding schemes. This large-scale project brings together 41 pan-European partners, including three cross-disciplinary research communities.

### **Eight international research funders announce winners of 2011 Digging into Data challenge**

<http://www.jisc.ac.uk/whatwedo/programmes/digitisation/diggingintodata.aspx>

Analysing 600 years of music, drilling down into population databases, understanding social unrest through digitised newspapers – these are just some of the new lines of research that the winners of the second Digging into Data Challenge will now investigate.

## **NISO Releases Updated Draft of SERU: A Shared Electronic Resource Understanding for Public Comment**

[www.niso.org/workrooms/seru/](http://www.niso.org/workrooms/seru/)

SERU offers publishers and libraries the opportunity to save both the time and the costs associated with a negotiated and signed license agreement for e-resources by both content provider and customer agreeing to operate within a framework of shared understanding and good faith. The SERU framework provides a set of common understandings for parties to reference as an alternative to a formal license when conducting business. The draft updated SERU Recommended Practice and an online comment form are now available.

## **Consultation on DPC Tech Watch Report: 'Preservation, Trust and eJournal Content'**

[http://www.dpconline.org/component/docman/doc\\_download/715-twrejournalsooutlinedraftforcommentjan2012](http://www.dpconline.org/component/docman/doc_download/715-twrejournalsooutlinedraftforcommentjan2012) (login required)

DPC has opened a consultation among members on a new Technology Watch Report on the topic of 'Preservation, Trust and eJournal content'. This report will examine preservation of eJournals and the services that support it will consider how standards and workflows have been established in this sector and what this might mean for the wider community. Comments are welcome to the start of February. A draft outline of the report is available to members at:

## **Release of DPC Tech Watch Report: 'Preserving Email'**

[http://www.dpconline.org/component/docman/doc\\_download/714-twrpreservingemailpreviewdecember2011](http://www.dpconline.org/component/docman/doc_download/714-twrpreservingemailpreviewdecember2011) (login required)

DPC launches a preview of a new Technology Watch Report on Preserving Email by Chris Prom at the University of DPC Technology Watch Reports identify, delineate, monitor and address topics that have major bearing on ensuring our collected digital memory will be available tomorrow. 'Preserving Email' is the first Technology Watch Report to be published by the DPC in association with Charles Beagrie Ltd. Neil Beagrie, Director of Consultancy at Charles Beagrie Ltd, was commissioned to act as principal investigator and managing editor of the series in 2011. In spite of email's potential historical, legal and administrative value, few organizations have developed sustainable programmes that are dedicated to preserving it. Several factors, including perceived technological barriers and legal mandates favouring destruction, have led many organisations pursue policies that amount to little more than benign neglect. As a result, the end users of email systems frequently shoulder the ultimate responsibility for managing and preserving their own email, exposing important documentary records to needless and counterproductive risk of loss. The report is now available for preview by DPC members.

## **Launch of the SPRUCE Project**

<http://www.dpconline.org/advocacy/spruce>

Leeds University Library is delighted to announce the launch of the Sustainable PReservation Using Community Engagement (SPRUCE) project. SPRUCE will inspire, guide, support and enable HE, FE and cultural institutions to address digital preservation gaps; and to use the knowledge gathered from that activity to articulate a compelling business case for digitalMore details of the project will follow shortly at:

### **New DPC Strategic Plan**

[http://www.dpconline.org/component/docman/doc\\_download/713-dpcstrategicplan2012-15](http://www.dpconline.org/component/docman/doc_download/713-dpcstrategicplan2012-15)

DPC has launched a new strategic plan for 2012-Five main areas of work are identified which refine and update our previous strategic objectives: workforce development and capacity building; knowledge exchange; assurance and practice; advocacy; and partnership.

### **Project Manager for SPRUCE Project**

<http://t.co/xkM7ltmA>

Leeds University Library is looking to appoint a project manager to lead the SPRUCE project. A graduate with a good knowledge of relevant issues in Higher Education, you will be able to demonstrate your experience of working in a multi-disciplinary team, and with senior business and IT/IS partners. In particular you will have experience of working in one or more of the following areas: digital content creation and/or description; digital preservation strategies, tools or techniques; business case creation and/or cost/benefit analysis; management of a JISC funded project; institutional policy creation; stakeholder evidence capture and analysis; event organisation and facilitation; formal report writing. More details at:





## What's What - Editorial: The Next Big Thing

*William Kilbride, DPC*

There's a saying that technology changes more slowly than you might think in a year and more quickly than you might realise in a decade. I've had this in mind while writing the new [DPC Strategic Plan for 2012-2015](#)

, which came into effect on the 1<sup>st</sup>

January, because it marks our own tenth anniversary. We'll get to that: but let's remind ourselves of what technology looked like in 2002 and just how much has changed since then:

- In January 2002, Ilford Imaging employed 740 staff and competed with Kodak in the production of black and white film for cameras. The company collapsed and mass production stopped in 2004. A successor called Ilford Photo now supplies a dwindling niche market with 'manufacture on demand' products - a market which Kodak has long abandoned.
- In January 2002 the iPod was about six months old and there was no such thing as an iTunes Store. The music industry was in rude health and EMI had just signed Robbie Williams on a six album contract worth \$157 million (US). EMI is now split between Universal and Sony and the future of the [music industry does not seem so assured as record sales fall](#).
- Twitter, Facebook, Myspace, Flickr, YouTube, LinkedIn and Bebo did not exist in January



2002 and Wikipedia had only just been set up. Sharing files had become harder given that NAPSTER had just been closed down; so if you wanted a social network you could always try Geocities. Google wave has come and gone in the same time.

Digital preservation has had a similar journey over the decade. In 2002 two influential and linked projects CEDARS and CAMiLEON were in progress: the former publishing its final report and a useful guide to digital preservation strategies ( <http://bit.ly/s7TFmM> ), the latter examining the practicalities of emulation. My own employer at the time - The ADS - was experimenting to see if 'checksums' were any use in transferring data to the AHDS deep store at the University of Essex while the Public Records Office (sic) was working towards the launch of a prototype file format registry which they christened PRONOM. Fashionable reading for 2002 included the newly published Reference Model for an Open Archival Information System.

But the really exciting news from 2002 – and the origin of this editorial - was the foundation of the DPC. The idea of a coalition to work on mutual concerns around preservation was first mooted in 1999 and the proposal received practical and financial support in 2000 when Neil Beagrie, then employed by JISC, began to put the business plan together. A series of workshops and meetings helped build momentum and the DPC became a legal entity on 20th December 2001 when the articles of association and memorandum of association were adopted by a newly created board of directors. The coalition was [launched at the House of Commons in February 2002](#) . As 2012 is the year tenth anniversary of the launch, it's a good time to take stock and update our plans for the future.

The urgency of awareness-raising about digital preservation has lessened since the DPC was founded. It is now possible and necessary to articulate more subtle messages about how citizens, agencies and governments can work together to ensure a dependable digital legacy and the benefits that accrue from it.

The last decade has seen a proliferation of digital preservation tools and services that are of great benefit in ensuring the longevity of data sets that are themselves expanding in scale, complexity and importance. But the fragmentation and impenetrability of digital preservation research is now routinely identified as an impediment and disincentive to those who need solutions urgently. There are growing risks that tool developers will find it hard to reach a market meaning their solutions are under-deployed and their investment under-exploited; standards developers struggle to achieve consensus meaning their approaches are under-consulted and ignored; and problem owners find it hard to find solutions, increasing the short term risks of data loss and the long term costs of deployment. By bringing people together, the DPC reduces the

barriers to digital preservation. We enable higher quality research, better informed standards and faster deployment of solutions. So the DPC's mission has changed since 2002 and is certainly more nuanced and arguably trickier than it was then.

Being more subtle means we can talk about preservation as a capacity as well as a challenge. A determined effort to identify, document and retain data of enduring value means that the right data is available at the right time in a form that can be used, and that redundant data are relegated or eliminated confidently. Without digital preservation, agencies are unable to consolidate their data infrastructures and are forced to maintain and repair a profusion of redundant systems which add cost and reduce effectiveness. So, as you may have heard me say before, digital preservation is also about confident deletion. And we don't do digital preservation for the sake of the data: we do because of the competitiveness, competence and distinctiveness which follow.

You would be mistaken if you thought our job was done. The DPC's new strategic plan gives us a clear direction for the next three years. Five main areas of work are identified which refine and update our previous strategic objectives: workforce development and capacity building; knowledge exchange; assurance and practice; advocacy; and partnership.

These five headings mean the continuation and refinement of some familiar themes for the DPC, such as the popular expert briefings or the frequent updates and news through What's New. It includes the continuation and expansion of the 'Technology Watch Reports' - such as the latest issue by [Chris Prom on the topic of Preserving Email \(which is now available to members as a preview – login required\)](#) ). But the plan also includes a few new commitments such as developing a more active role as a reviewer and contributor to standards; and a promise to experiment with staff exchanges across the coalition, leading on from work being carried out as part of our commitment to the APARSEN network.

The DPC is owned by its members. The plan focusses on their needs and it comes from an extended consultation with our engaged, informed and active membership. The core membership has grown significantly since we were first established and it is sometimes hard to know how we can service their diverse needs with such a small core team. I am sure that, in its second decade, the DPC will continue to make some significant contributions to how data is created, preserved and used: and I have no doubt that the year ahead will be as busy as any so far.

2012 will be a good year to be member of the DPC and a great year to join us.

---



### **Who's Who: Sixty Second Interview with Lee Hibberd, Digitisation Manager, National Library of Scotland**

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

I work in the National Library of Scotland's Digital Collections team as the Digitisation Manager

#### **Tell us a bit about your organisation**

The National Library of Scotland has a history harking back to the early 1680s. Today NLS is Scotland's largest library and one of the major research libraries in Europe. Our world-class collections include more than 14 million printed items and as a legal deposit library we are entitled to a copy of any UK publication. We also have hundreds of thousands of maps, tens of thousands of hours of film and video, and miles of boxed manuscripts. Naturally we have an online presence at [www.nls.uk](http://www.nls.uk) where millions of digital objects are available through our use of

licensed resources and the digitisation of our collections

### **What projects are you working on at the moment?**

Most of my time is spent managing the content produced by the rest of the Digital Collections team who create digital versions of our collections. At the moment we're queuing up several digitisation projects so that we can release a project a month onto [www.nls.uk](http://www.nls.uk). The latest has been a major addition to the Medical History of British India website (<http://digital.nls.uk/indiapapers>

), supported by the Wellcome Trust, which focuses on veterinary medicine between 1864 and 1959. In spring we will be providing improved search tools for our Scottish Street Directories project (<http://www.nls.uk/family-history/directories/post-office>

). This will provide an invaluable resource for family historians tracking their relatives between the census years. On the Digital Preservation front I've been working closely with colleagues in our Scottish Screen Archive to refresh our Digital Preservation strategy. In the New Year I'll be working with our developers to increase the automation and efficiency of our file integrity checks that confirm whether or not our digital data has changed over time.

### **How did you end up in digital preservation?**

It wasn't long after joining National Library of Scotland in 2002 as Digitisation Officer that my director, Fred Guy (now at EDINA), pushed a pile of folders my way and suggested I go to Copenhagen to find out all about OAIS (Were you there?). That was my first deep immersion in digital preservation and the fascinating challenge that accompanies it. It was also my first experience of working in Europe and I have strong positive memories of it. Since then I have dipped my toe in and out of the subject, always thinking of how I should care for the growing collection of digitised resources under my management. With some luck I'll be plunging my feet, legs and body into more digital preservation work in the near future.

### **What are the challenges of digital preservation for an organisation such as yours?**

My feelings, founded on the great work that has taken place around the world for the last decade or so, suggest that digital preservation is doable but that it requires a substantial and sustainable investment of time, effort and money. In 50 years time I'm sure we can all sit back and relax as the preservation environment invisibly embeds itself into the workplace. In the meantime doing all of this new stuff alongside all of the old stuff without affecting the services we provide requires careful rebalancing. This has always struck me as being the most difficult challenge for NLS, and I know it's one faced by many others.

### **What sort of partnerships would you like to develop?**

Whenever I meet someone in the field I'm always hopeful that they have practical experience and can say to me, "Oh you want to do that. Well this is how you do it". This has happened many times over the years and can really help you to move rapidly forward. Looking at the bigger picture NLS will increasingly work with other information institutions, researchers and the private sector to solve our preservation problems. Scotland is often quoted to be ideally sized for collaboration and we are already working with the National Records of Scotland and the Registers of Scotland on the preservation of public records. To understand which partnerships will be successful requires a clear understanding of your current position and aims, a clear vision. The library's desire to refresh our digital preservation strategy will stimulate this and help enormously throughout 2012 and beyond.

**If we could invent one tool or service that would help you, what would it be?**

As I've already mentioned I'm working to improve our file integrity checking tools. Knowing whether files have changed or not over time has to be one of the digital preservation fundamentals. The "DPC-for-Lee tool" would be able to check 100s of TBs of data stored on disk and tapes once a month indicating whether it had changed and replacing it with an unchanged copy. Simple.

**And if you could give people one piece of advice about digital preservation ....?**

Get your hands dirty. This helps you to understand what you have and what you need and helps you to communicate with others who could help.

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

What a great question. I've been secretly turning my 2005 honeymoon video footage into a little movie for my wife's birthday present. It's been a challenge to do it without her noticing and I'd be mightily upset if it gets lost before the premiere. Afterwards, and after I'm long gone, it would be great to keep this for the kids (who would then be old and wrinkly themselves) so that they could remember the movements and motion their young parents made.

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

I'm absolutely hopeless at returning e-mails so the best way to ever get hold of me is to call. And you can always enjoy our website and digitised collections at <http://digital.nls.uk/>



## **One World: Digital Preservation in Spain**

***Alice Keefer and Miquel Térmens, Faculty of Information Sciences, University of Barcelona***

In comparison with the digital preservation activities of other countries, Spain can be considered to be in the middle of the pack. The results of the 2011 EC survey on national open access and preservation policies find that, in contrast to Spain's dedicated efforts regarding OA, "the national strategy on preservation ... has not been approached". Although the report specifically addresses scientific information, the same observation generally applies to other types of digital content as well. However, a number of noteworthy individual efforts make up for the absence of activities at a national level.

The preservation of web-based material has been the focus of attention of the Spanish National Library (BNE) and the national Library of Catalonia (BC). In 2005 the BC set up the *PADICAT* service with the aim of preserving the "Catalan web". The system has a triple approach for content acquisition: automatic harvesting of all material under the .cat domain; agreements with major political, cultural and social organizations for the authorized capture and curation of their web content; and programmed theme-based capture, especially of election-related material. For its part, the BNE in 2009 initiated a program to preserve the .es domain, by engaging Internet Archive to carry out periodic harvesting. The BNE and BC are both members of the International Internet Preservation Consortium (IIPC).

In July, 2011 a new law on legal deposit was approved that will facilitate the preservation of born digital material by Spain's various national libraries. For the first time non-tangible digital works are also subject to legal deposit. Although producers of web pages are exonerated from

this requirement, the law specifically permits the national libraries to “reproduce freely accessible web sites of interest for reasons of legal deposit, while respecting existing legislation on data protection and intellectual property”. Although a number of issues still remain to be worked out, it represents a good starting point. Regarding tangible digital material, the law requires that the deposited content be free of any obstacle to access, such as passwords; that it be accompanied by any software needed for purposes of research or conservation; and that the depositor assure the transferability of the data to media required for preservation.

On the university front, there have been various initiatives as well. In 2009, a Working Group of the Network of Academic Libraries (REBIUN) published a guide to digital preservation resources for its members. Special emphasis was placed on the need to ensure the preservation of digital material stemming from digitization projects. This is also a key concern of the Ministry of Culture that now requires the use of METS and PREMIS in all digitization projects funded by it. Individually, there have been several cases of academic libraries joining with international preservation initiatives. Two examples: the Universidad Complutense de Madrid, with the second largest library in Spain, is a member of Hathi Trust and the Catalan consortium of university libraries, CBUC, joined the MetaArchive Cooperative for ensuring the preservation of its joint repository of doctoral theses.

In archives, initiatives have been set up to meet new e-administration requirements. For instance, the Catalan and the Basque autonomous governments each have set up secure storage for e-documents --*iArxiu* and *Metaposta* respectively-- to meet official retention requirements. On the international level, Spanish archivists have participated in both Interpares2 and Interpares3.

---

{jcomments on}