

## **Report on the DPC Forum held at the British Library Conference Centre, Wednesday 23 June 2004.**

The 8<sup>th</sup> DPC Forum attracted the biggest audience to date for a DPC Forum. Around 100 delegates were kept interested and informed by a very rich programme with presentations from several experts from the U.S and Europe. One key theme running throughout the day was the need for active collaboration at every level and across sectoral and geographic boundaries. Speaker after speaker illustrated how this collaboration was essential. Other consistent messages were the importance of trust (between partners and stakeholders in the emerging technical infrastructure), the need to find effective mechanisms to distribute responsibilities, developing standards and tools and above all, the need to develop and share practical experience.



Speakers from the 8th DPC Forum, Digital Preservation: the global context  
L to R Taylor Surface, OCLC; Robin Dale, RLG; Nancy McGovern, Cornell; Peter Burnhill, DCC; Seamus Ross, HATII;  
Eileen Fenton, Ithaka; David Seaman, DLF; Vicky Reich, LOCKSS; Tony Hey, eSCP; Laura Campbell, Library of Congress

Delegates received a sense of the broad range of activities going on, the progress that has been made, and the increasingly compelling need to accelerate progress. Feedback from the Forum, both formal and informal, has been overwhelmingly positive and is indicative of the consistently high quality of the presentations and a stimulating and thought provoking programme.

On the evening before the Forum, there was the presentation of the annual Conservation Awards, which included the inaugural [DPC Award for Digital Preservation](#). This was won by the National Archives, for their

[Digital Archive](#)

. The

[CAMiLEON](#)

had received a special commendation. This event was regarded as another stepping-stone on the way to raising the profile of digital preservation. The Forum was equally important, bringing together people from all over the world, recognising the need for international collaboration, noting that no one can do this on their own.

Lynne Brindley, Director of the British Library and Chair of the DPC Board, chaired the day and noted in her welcome the importance the DPC placed on international links and the need to ensure that digital preservation issues are increasingly on the political and policy agendas. The DPC is committed to making practical progress and to sharing best practice through its membership.

To open PDFs you will need [Adobe Reader](#)

Ms Brindley introduced the first speaker, Laura Campbell, Associate Director for Strategic Initiatives at the Library of Congress, who provided an up-to-date picture of what the Library of Congress was doing through their [NDIIPP](#) (PDF 772KB) (National Digital Information Infrastructure and Preservation Program) program. Ms Campbell described NDIIPP, which developed as a result of a report commissioned by the Library of Congress to assess whether it was prepared for the 21<sup>st</sup> century. Much experience in digital technology had come from building their Digital Library and they had learned the power of digital surrogates as well as their vulnerability to loss.

The \$US175m plan consisted of \$5m approved by Congress to produce a plan, \$20m upon approval of the plan, and a further \$75m which would be contingent on obtaining matching funds. Scenario planning helped show how a distributed effort might operate.

Key lessons and messages of NDIIPP to date include the belief that there will never be a single right way of doing things, so the architecture needs to be sufficiently modular and flexible to take account of this, the need for a distributed and decentralised approach and the need for new tools and technologies. NDIIPP needs to build partnerships and networks and then create a technical infrastructure to support the partners. Partnerships already forged included an alliance with the DPC, helping to establish the International Internet Preservation Consortium (IIPC), business model partnerships such as subscription and archiving services for e-journals, and technical partnerships, taking full advantage of the skilful technical talent which exists.

The next stage of NDIIPP would include testing architectures to support archive ingest and handling. In summing up, Ms Campbell indicated that during the next five years, the intention was to form a range of formal partnerships, encourage standards for digital preservation, establish a governing body, and make recommendations to Congress for funding.

Seamus Ross, Director of HATII (Humanities advanced Technology and Information Institute), provided an introduction to [ERPANET](#) (PDF 1087KB) (Electronic Resource Preservation and Network), the European Commission funded project which has brought together partners from Italy, the Netherlands, U.K and Switzerland. ERPANET has created a number of resources, organised seminars on several key topics, carried out an analysis of relevant literature and developed other tools, such as business cases for digital preservation and off-the-shelf policy statements. It was stressed that a lot of expertise already exists but there is a pressing need to bring it together and to work together.

Lessons learned were that the digital preservation community needs practical case studies and reports of "real world" experience. Simple tools for costing digital preservation exist but much more work needs to be done here. Guidance on digital repository design is also needed. ERPA E-prints (a repository of digital preservation papers and reports) is growing very slowly and needed to be marketed better. ERPANET have negotiated with the Swiss National Archives to preserve material held in this repository in perpetuity. In summing up, Dr Ross emphasised the great need for knowledge sharing so ERPANET events and DPC Forums were extremely important in helping to raise the level of awareness and understanding.

Presentations from David Seaman (Director of the Digital Library Federation); Robin Dale (Program Officer for the Research Libraries Group); and Taylor Surface (OCLC), described the work their organisations are doing in developing practical, collaborative tools, all of which will play a role in increasing trust in the developing infrastructure for digital preservation.

David Seaman's presentation, ' [Towards a Global File Format Registry](#) ' (PDF 67KB) described the developing global file format registry, which is responding to an immediate need. The importance and value of linking to other relevant work, such as the National Archives' PRONOM system and the DCC (Digital Curation Centre) in the UK, was also stressed.

The title of Robin Dale's presentation, ' [The Devil's in the Details- working towards global](#)

[consensus for digital repository certification](#)

' (PDF

77KB), aptly summarised the challenge of articulating and reaching broad consensus on what elements and what process can be put in place to certify digital repositories against a commonly understood standard.

Taylor Surface described the work of OCLC's Digital Collections and Preservation Services in '[the OCLC Registry of Digital Masters](#)

' (PDF 464KB), which arose from a DLF Steering Committee recommendation. Taylor described how the registry linked to the OCLC's WorldCat service to provide enhanced discovery , encourage use of standards and limit duplication of effort of digitisation initiatives.

During the lunch break, Lynne Brindley and Laura Campbell signed an agreement between the DPC and the Library of Congress. A poster session on the Digital Curation Centre gave delegates the opportunities to ask specific questions before the afternoon presentations.

The afternoon session began with Tony Hey, Director of the e-Science Core programme. Tony's presentation was '[e-Science - preserving the data deluge](#)' (PDF 543KB) . The e-science grid (or cyberinfrastructure as it's known in the U.S) has the vision enunciated by Licklider, of being able to bring together all material throughout the world and build a truly global, collaborative environment which enabled researchers to work together regardless of geographic location. Describing the impetus for the development of the DCC, Dr Hey said that over the next 5 years, e-science will produce more scientific data than has been collected in the whole of human history. The goal is to bring together the digital library community with the scientific community so that each can learn from the other.

Peter Burnhill, Interim Director of the DCC, described '[The Digital Curation Centre](#)' (PDF 146KB) which has received funding of £1.3m p.a. from JISC and the e-Science Core programme. The DCC was not a digital repository, he said, but would provide services and research for the community involved in digital preservation. It is still very early days, in that the DCC has only been operational for a few months but progress has been made. A website has been launched, an e-journal is planned and focus groups would help to articulate who the user community for the DCC is and what their needs are. It was anticipated that a permanent Director would be in post by the official launch, scheduled for early November.

Nancy McGovern spoke of '[The Cornell Digital Preservation Online Tutorial and Workshop](#)' (PDF 419KB). This is yet another illustration of the pressing need to develop practical support for those already involved in, or about to embark on, digital preservation programmes. It was also another example of the strength of collaboration, as the curriculum had been developed collaboratively and Cornell looked forward to working closely with the DPC, who have been inspired by Cornell's work to develop a similar programme geared towards the U.K. Nancy

described the five organisational stages of digital preservation which are: Acknowledge; Act; Consolidate; Institutionalise; Externalise. Nancy noted that none of these stages can be skipped and it was essential to realise that there is no on/off switch for digital preservation, it is something which needs to build over time. Cornell has now run four workshops which have received very positive feedback from participants. All have been oversubscribed, which illustrates the need for intensive training which provides a toolkit to enable participants to take practical short-term strategies appropriate to their own institutional settings. A fifth workshop is planned for November 2004.

The final session of the day provided an opportunity to hear two very different approaches to preserving e-journals. Vicky Reich described the ' [LOCKSS Program approach](#) ' (PDF 804KB), applicable to any content available through http protocol, and which enables libraries to collect and preserve content in the same way as they do for print. Vicky stressed that LOCKSS preserves the content, not the services publishers provide (e.g. search buttons). LOCKSS has established contact with several publishers and it is essential to have the cooperation of publishers to allow LOCKSS crawlers to gather their content. Trust was also an issue here - publishers needed to trust that libraries would gather content they have purchased under licence. Key advantages of LOCKSS are its inbuilt redundancy and ease and cheapness to install. Vicky stressed that some institutions needed to have large, central repositories as well but this need not preclude the use of LOCKSS.

The final speaker of the day was Eileen Fenton, on ' [Preserving e-journals, the JSTOR model](#) ' (PDF 58KB). The Electronic Archiving Initiative has involved working with publishers and is focused on preserving the source files. Archiving e-journals requires a significant investment in the development of organisational and technological infrastructure, it was not either/or. Eileen also described Ithaka, a not-for-profit company, supported by Mellon, Hewlett and Niarchos funding. This has the goal of filling gaps not being supplied by the free market. Both Eileen and Vicky agreed that at this nascent stage of development, the community needs multiple approaches.

In closing the Forum, Lynne Brindley thanked all of the speakers for the significant contribution that had made to the success of the Forum. The next DPC Forum will be a joint DPC/CURL event and will be held on Tuesday 19 October 2004. Further details will be available in the coming months.