

### DPC Forum with Industry

5th June, 2002. Prospect House York Road London SE1 7AW

Since its inception the DPC has aimed to gain industry awareness of its key messages and of the future needs and opportunities that lie ahead. This forum is part of that process. During the day representatives from the private and public sector will be speaking. They will address long-term trends and the research and development issues involved in the implementation of continuing access and preservation strategies by industry and government. Issues covered will include preserving TV and broadcast archives and research and development in the public and private sector.

### Meeting Report

Introducing the first Digital Preservation Coalition (DPC) Forum focussed on developing a dialogue with Industry, DPC Secretary **Neil Beagrie** welcomed guests and members of the DPC. He then placed the question of digital preservation and opportunities for industry participants, firmly within an international context. The US National Digital Information Infrastructure and Preservation Program (NDIIPP): a \$175m national programme; the National Science Foundation Cyberinfrastructure initiative; the Information Society Technologies programme of the EU 6<sup>th</sup> Framework; and, in the UK, the Research Grid and the work of the DPC, will be central to bringing political, technical and organisational impetus to bear on the challenges of digital preservation. With accelerating development of digital content, the issue of maintaining long-term access will be of concern to an increasing ranging of sectors and individuals.

Forging strategic alliances with industry in the context of preservation was a vital component of any initiative in this area, said Neil Beagrie. He outlined how the DPC's constitution was shaping its emerging links with industry. Two important principles governed the DPC's work: firstly, the DPC will support the development of standards and generic approaches to digital preservation, which can be implemented by a range of hardware, software and service vendors; in short, he continued, the goals of the DPC are vendor-neutral. The second principle is that the DPC is a coalition of not-for-profit organisations including industry associations; it is committed to promoting and disseminating information so that all can learn from the transferable lessons and outcomes. The DPC is actively interested in broadening its industry links including those to individual companies. There is a major potential role for industry, concluded Mr Beagrie; this event therefore marked an important step in realising that potential.

**Philip Lord**, formerly of SmithKlineGlaxo, and now a Digital Archiving consultant, spoke of his own experience in industry and of some of the major drivers for industry in the field of digital preservation. He particularly emphasised the importance of the US Federal Drug Agency (FDA)'s regulation 21CFR Part 11

on the maintenance of good electronic records for the pharmaceutical industry. Legal and regulatory issues were, he said, clearly of major importance, as were the various voluntary drivers, such as contractual and IP obligations, operational efficiency considerations and the need to preserve for future reuse. Rarely were records preserved, he said, for historical or sentimental reasons.

Mr Lord talked of the many challenges that faced industry in this area: the heterogeneity of data sources and systems, geographical dispersion (spanning legal and regulatory jurisdictions), the lack of suitable preservation systems and services, management issues, cost, and the lack of expertise in this area. Mr Lord reported that little progress had been made so far within the sector, although a few companies, he reported, were leading the way.

**David Ryan** from the Public Records Office reported that digital preservation was a core part of the PRO's e-business strategy. Research and development in this area was focussing, he said, on proprietary format migration, emulations and simulation, open format export and migration and product reviews. The importance to users of *interaction* with preserved digital information, rather than simply using that information as an historical record (as is more likely the case with printed data) was made, and this gave an added importance to the work of the PRO in its e-preservation activities. Mr Ryan concluded by saying that there was an overriding need for the archives community to establish credibility with the key stakeholders in e-preservation, including industry, the media and the general public. The community as a whole needed to prioritise so that the most urgent tasks and challenges were tackled first.

The film and sound archives of the BBC contain some 1.75m film and videotape items and around 800,000 radio recordings from the late 1940s onwards. A ten-year preservation strategy had, **Adrian Williams** from the BBC reported, recently been approved and would cost around £60m. A key part of this strategy was a programme of digitisation for both access and preservation. He reported on the European Commission Presto project, which involved 10 partners, lasted 24 months and cost around 4.8 million Euros. Findings from this project suggested, for example, that digitisation and mass storage is about 50% more expensive, but is expected to double usage of an asset; and moreover, that the value of an item must be four times the preservation cost to be financially viable. He concluded by suggesting that Europe requires a "dedicated preservation factory" given the scale of the task facing national broadcast archives. There was substantial audience interest in the approaches to cost and business

modelling described in the Presto project.

**Julian Jackson** noted in the headlong rush to put photographic images into digital form, little thought seemed to have been given to the problem of the longevity of digital files. There is an assumption that they will be lasting, but that is now under question. He addressed general issues surrounding preservation and obsolescence in digital images. He surveyed the techniques of refreshing, migration and emulation and emphasised the crucial role that metadata and metadata standards have to play in these preservation processes.

**Paul Wheatley** of the CAMiLEON project spoke of some of the practicalities of digital preservation and emphasised the need for long-term strategies. Existing methods have many drawbacks. Mr Wheatley described advanced techniques of data migration which can be used to support preservation more accurately and cost effectively.

To ensure that preserve works can be rendered on computer systems over time "traditional migration" has been used to convert data into current formats. As the existing format becomes obsolete, another conversion is performed, etc. Traditional migration has many inherent problems as errors during transformation propagate through future transformations. Mr Wheatley described how the Camileon project had developed new approaches to extending software longevity ("C--") which had been applied in experiments and demonstrated improvements over traditional migration. This new approach is named "Migration on Request".

Migration on requesting shifts the burden of preservation onto a single tool, which is maintained overtime. Always returning to the original format enables potential errors to be significantly reduced. Mr Wheatley also described how preservation quality emulators were being produced and strategies of migration on request and/or emulation were being applied.

The need for public-private partnerships in the field of digital preservation is crucial, said **David Bowen**

of Audata Ltd. He went on to outline what industry was currently doing in this field - e-mail, document and database migration, as well as promoting standards, while software suppliers are also improving backward compatibility (Word, Wordperfect, RTF), and increasingly adopting and promoting standards themselves too. Mr Bowen called for R & D partnerships, like the Testbed Digitale Bewaring in the Netherlands, which is leading to the sharing of results and advice, and sound record creation and metadata practices. Particularly important, he concluded, was the need for software suppliers to be brought into the growing public-private partnerships that are

developing.

The final session of the day was a discussion session. Key themes that emerged included:

- the importance of archiving software and technical documentation. It was felt by participants from all sectors that this is a major gap and there was an urgent need to develop appropriate repositories;
- the need to develop case studies and tools for modelling costs. It was felt this is a major area that should be covered in future DPC forum;
- the necessity of developing national funding for the preservation of intangible heritage assets. It was noted there is no "Superfund" or legislation which allows digital heritage to be gifted in lieu of tax to or purchased by, the nation;
- further work by the Digital Preservation Coalition to establish contacts with industry and to build on the dialogue commenced at the forum.

It was felt that it was important, as David Bowen had said, to include software and hardware suppliers in future developments as their actions could be crucial, in particular in providing the tools and products for end-to-end solutions which were needed in this area. Once again the importance of using both migration and emulation strategies was emphasised, as was the question of the criteria for choosing what should be preserved; we are not in a position to judge easily what will be in demand in the future. Therefore sampling could be of crucial importance as one strand in our overall strategy.

Some delegates from industry felt that there were gaps of responsibility between the organisations, and that it was therefore important for the DPC to coordinate and facilitate activities in this area.

The forum ended on a note of optimism that the first steps in the dialogue with industry had been taken, and with a number of concrete action points which Lynn Brindley, Chair of the DPC, promised would be followed up in the coming months.

*End of Meeting Report*

### Programme and Presentations

To open PDFs you will need <a href="#">Adobe Reader</a>		10.30 - 11.00	Registration and coffee
11.00 - 11.10	Welcome and Introduction		
11.10 - 11.35	<a href="#">Keynote Address - "Trends and Future Opportunities"</a> (PDF 92KB) Neil Be		
11.35 - 12.00	<a href="#">Preserving digital records in Industry</a> (PDF 258KB) Philip Lord ex GlaxoSm		
12.00 - 12.30	<a href="#">Preserving digital records in Government</a> (PDF 71KB) David Ryan Public F		
12.30 - 1.30	Lunch		
1.30 - 1.55	<a href="#">Preserving TV and Broadcast Archives</a> (PDF 449KB) Adrian Williams BBC		
1.55 - 2.15	<a href="#">Preserving Digital and Historic Images</a> Julian Jackson Internet Consultant a		
<a href="#">Digital File Longevity</a>			
2.15 - 2.35	<a href="#">The Camileon and Cedars Research projects</a> (PDF 500KB) Paul Wheatley		
2.35 - 3.05	Coffee		
3.05 - 3.30	<a href="#">Practical Experiences of Preservation: R&amp;D partnerships in</a>		
<a href="#">the private and public sector</a>	(PDF 397KB) David Bowen Audata Ltd		
3.30 - 4.10	Discussion		
4.10 - 4.30	Concluding Address - Lynne Brindley Chief Executive British Library		