

## DPC/PADI What's New in Digital Preservation

A joint service of the Digital Preservation Coalition and the [PADI](#) (Preserving Access to Digital Information) gateway

**Issue no. 7 (October 2003 - January 2004)**

Compiled by Michael Day ([UKOLN](#), University of Bath) and Gerard Clifton ([National Library of Australia](#))

26th February 2004

---

This is a summary of selected recent activity in the field of digital preservation compiled from the Preserving Access to Digital Information ([PADI](#)) Gateway (<http://www.nla.gov.au/padi/>) and the [digital-preservation](#) (<http://www.jiscmail.ac.uk/lists/digital-preservation.html>) and [padiforum-l](#) (<http://www.nla.gov.au/padi/forum/>) mailing lists. Additional or related items of interest may also be included.

### Contents:

<b>1. NEWS FROM ORGANISATIONS AND INITIATIVES.....</b>	<b>2</b>
1.1 THE UK DIGITAL CURATION CENTRE .....	2
1.2 NSF-DELOS WORKING GROUP ON DIGITAL ARCHIVING AND PRESERVATION.....	2
1.3 AUSTRALIAN RESEARCH INFORMATION INFRASTRUCTURE COMMITTEE .....	3
1.4 NESTOR NETWORK OF EXPERTISE IN LONG-TERM STORAGE OF DIGITAL RESOURCES .....	4
1.5 PROJECTS FUNDED UNDER THE EUROPEAN UNION'S 6TH FRAMEWORK PROGRAMME .....	4
1.6 WORLD SUMMIT ON THE INFORMATION SOCIETY .....	5
1.7 US NATIONAL DIGITAL INFORMATION INFRASTRUCTURE AND PRESERVATION PROGRAM .....	6
1.8 UK DIGITAL PRESERVATION COALITION .....	6
1.9 ERPANET PROJECT.....	6
1.10 JISC SUPPORTING INSTITUTIONAL RECORDS MANAGEMENT PROGRAMME .....	8
<b>2. SPECIFIC AREAS OF ACTIVITY.....</b>	<b>9</b>
2.1 WEB ARCHIVING.....	9
2.2 FORMATS AND TOOLS.....	11
2.3 METADATA .....	13
2.4 LEGAL ISSUES .....	16
2.5 INSTITUTIONAL REPOSITORIES, E-PRINTS AND E-JOURNALS .....	17
2.6 DIGITAL LIBRARY FUTURES.....	19
<b>3. OTHER RECENT PUBLICATIONS.....</b>	<b>20</b>
<b>4. EVENTS.....</b>	<b>22</b>
4.1 RECENT EVENTS .....	22
4.2 FORTHCOMING EVENTS .....	23

# 1. News from organisations and initiatives

## 1.1 The UK Digital Curation Centre

On 5 February 2004, the Joint Information Systems Committee and the eScience Core Programme announced the award of GBP1.3 million a year for a UK Digital Curation Centre based in the UK. The centre will be run by a consortium made up of four partner institutions: the universities of Edinburgh (the National eScience Centre, the School of Informatics, and other units) and Glasgow (Humanities Advanced Technology & Information Institute, and Information Services), the Council for the Central Laboratory of the Research Councils (CCLRC), and the University of Bath (UKOLN). The Centre includes a research programme funded by the Engineering and Physical Sciences Research Council (EPSRC), and service, development and outreach activities funded by the JISC. Peter Burnhill, currently the Centre's acting director, notes that it is not intended to act as a digital repository itself, instead, "based on insight from a vibrant research programme that addresses wider issues of data curation, it will develop and offer programmes of outreach and practical services to assist those who must curate data."

Joint Information Systems Committee. (2004). "New UK centre to help secure the future of digital data" [press-release], 5 February 2004. Retrieved February 24, 2004, from: [http://www.jisc.ac.uk/index.cfm?name=dcc\\_news\\_040204](http://www.jisc.ac.uk/index.cfm?name=dcc_news_040204)

Burnhill, P. (2004). "Edinburgh at forefront of data curation." *University of Edinburgh, Bulletin of IT Services*, February 2004. Retrieved February 24, 2004, from: [http://www.ucs.ed.ac.uk/bits/2004/february\\_2004/](http://www.ucs.ed.ac.uk/bits/2004/february_2004/)

## 1.2 NSF-DELOS Working Group on Digital Archiving and Preservation

The final report and recommendations of the NSF-DELOS Working Group on Digital Archiving and Preservation was published in late 2003. The report identifies significant challenges in digital archiving and long-term preservation, and seeks to define the research agenda that would be necessary to overcome these. Recommendations for the agenda cover three main areas:

- Emerging research domains - the report concludes that further research is needed on: repository models and repositories for software, formats and peripheral devices; archival media, storage abstractions and data rescue; digital object documentation, functionality, the preservation of knowledge content and age-testing.
- Re-engineering preservation processes - in order to reduce costs, the report calls for more research into preservation process modelling and automation, detection of trustworthiness, information quality and collection completeness, development of smaller-scale affordable solutions and mechanisms for distributed storage.
- Preservation of systems and technology - the report calls for more research into management and preservation of complex or dynamic objects and formats, the automated creation and long-term management of metadata, methods to assess the impact of potential information loss, and the efficient re-use or repurposing of digital information.

The three areas that the Working group agreed would have greatest impact on effectiveness and efficiency of digital preservation were self-contextualising objects, metadata and ontology development, and mechanisms for preservation of complex or dynamic objects.

Hedstrom, M., Ross, S., et al., (2003). *Invest to save: report and recommendations of the NSF-DELOS Working Group on Digital Archiving and Preservation*. Retrieved February 24, 2004, from: <http://delos-noe.iei.pi.cnr.it/activities/internationalforum/Joint-WGs/digitalarchiving/Digitalarchiving.pdf>

### 1.3 Australian Research Information Infrastructure Committee

On 22 October 2003, the Australian Commonwealth Minister for Science announced funding of AUD 12 million through the Department of Education, Science and Training for four new projects to improve Australia's research information infrastructure and the management of university information.

The four demonstrator projects implement an action plan derived from the findings of the Higher Education Information Infrastructure Committee's 2002 report, which recommended improvements in information creation and management, access to information resources, and the discovery and dissemination of new information to researchers and institutions. The strategy includes the storage and dissemination of scholarly information and research outputs, including books, articles, theses and research data. The four projects involve fifteen Australian universities, Australian and international libraries, as well as industry representatives and international organisations. The projects will be co-ordinated by the newly established Australian Research Information Infrastructure Committee (ARIIC).

The four projects are:

- **Meta Access Management System (MAMS)** - This will develop the conceptual architecture and essential "middleware" capable of supporting multiple independent models for access management systems, allowing for integration of multiple solutions to manage authentication and authorisation, digital rights management, metadata and search systems.
- **Towards an Australian Partnership for Sustainable Repositories (APSR)** - A three-part project which will focus on critical issues for continued accessibility and sustainability of digital collections; will build a base of demonstrator repositories within partner institutions; and will co-ordinate and encourage development of skills and expertise across the sector nationally, providing national services and international linkages.
- **The Australian Research Repositories Online to the World (ARROW)** - This will identify and test open-source software solutions to support institutional repositories of e-prints, digital theses and electronic publishing, and will include development and testing of national resource discovery services.
- **Australian Digital Theses Program Expansion and Redevelopment (ADT)** - This will increase the coverage and utility of the existing central metadata repository of the Australian Digital Theses Program, expanding it to include metadata about all Australian higher degree theses, regardless of format. Some retrospective digitisation of non-digital theses will also be included.

Further information on projects and participating institutions is available in the following documents:

Department of Education, Science and Training. (2002). *Final report of the Higher Education Information Infrastructure Advisory Committee, Systemic Infrastructure Initiative - Executive Summary*. Canberra: Department of Education, Science and Training, November 2002. Retrieved February 24, 2004, from:  
[http://www.dest.gov.au/highered/otherpub/heiac/exec\\_summary.htm](http://www.dest.gov.au/highered/otherpub/heiac/exec_summary.htm)

Department of Education, Science and Training. (2003). "\$12 million for managing university information" [press-release]. Canberra: Department of Education, Science and Training, October 2003. Retrieved February 24, 2004, from:  
<http://www.dest.gov.au/Ministers/Media/McGauran/2003/10/mcg002221003.asp>

National Library of Australia. (2003). "National Library to participate in 'Backing Australia's ability: an innovation action plan for the future'." *Gateways*, 66, December 2003. Retrieved February 24, 2004, from:  
<http://www.nla.gov.au/ntwkpubs/gw/66/html/p18a01.html>

## 1.4 NESTOR Network of Expertise in long-term Storage of digital Resources

The German Federal Ministry for Education and Research (*Bundesministerium für Bildung und Forschung*) is funding an initiative called the NESTOR Network of Expertise in long-term STORage of digital Resources (*Kompetenznetzwerk Langzeitarchivierung digitaler Ressourcen*). The network will provide a forum on digital preservation topics and will produce guidance on standards and best practice. Topics that will be considered by the network include: criteria for trusted digital repositories, certification procedures, selection for preservation, and the principles of long-term preservation. Partners in the network are *Die Deutsche Bibliothek*, the Bavarian State Library (*Bayerische Staatsbibliothek*), the Göttingen State and University Library (*Niedersächsische Staats- und Universitätsbibliothek Göttingen*), the Humboldt University of Berlin, the Bavarian State Archives (*Generaldirektion der Staatlichen Archive Bayerns*), and the Institute for Museum Studies (*Institut für Museumskunde*) in Berlin.

*NESTOR Kompetenznetzwerk Langzeitarchivierung digitaler Ressourcen*. Retrieved February 24, 2004, from: <http://www.langzeitarchivierung.de/>

## 1.5 Projects funded under the European Union's 6th Framework Programme

The European Commission (EC) unit on the 'Preservation and enhancement of cultural heritage' has announced eight new projects funded under the European Union's 6th Framework Programme (FP6). The structure and funding of projects under FP6 is different to that of previous framework programmes, resulting in a smaller number of much larger projects. Bernard Smith, the head of unit, notes that these projects bring together more than 250 participants, and that the projects have been allocated a 'funding envelope' of around EUR 37 million. The projects with the most relevance to digital preservation are:

- **DELOS Network of Excellence on Digital Libraries** - This is a 'network of excellence' that intends to integrate and co-ordinate the ongoing activities of European research into digital libraries. The project is defining a joint programme of activities, which is organised into seven research clusters. One of these clusters (Work Package 6) concerns preservation, and is led by the University of Glasgow, UK.
- **PRESTOSPACE**: Preservation towards storage and access. Standardised Practices for Audio-visual contents in Europe - This 'integrated project' builds on an older EU project (PRESTO), and its objective is to provide technical solutions and integrated systems for the preservation of all types of digital audio-visual collections. The focus is on developing a semi-automated 'preservation factory' approach to preservation of audio-visual collections. The project is co-ordinated by the French *Insitut national de l'audiovisuel* (Ina), and partners include major broadcasting and film archives - including the archives departments of RAI (*Radiotelevisione Italiana*) and the BBC (British Broadcasting Corporation) - together with applied research institutes, universities and various industrial partners. Web site retrieved February 24, 2004, from: <http://prestospace.ina.fr/>
  - Insitut national de l'audiovisuel. (2004). *PrestoSpace: le projet européen de recherche sur la sauvegarde et la numérisation des archives audiovisuelles* [press dossier]. Bry-sur-Marne, France: Ina, 5 February 2004. Retrieved February 24, 2004, from: [www.ina.fr/presse/dossiers/prestospace.pdf](http://www.ina.fr/presse/dossiers/prestospace.pdf)
- **BRICKS**: Building Resources for Integrated Cultural Knowledge Services - This 'integrated project' aims to establish the foundations of a digital library for integrating collections of multimedia digital documents in the cultural heritage domain. The BRICKS project is co-ordinated by *Engineering Ingegneria Informatica SpA*, Italy.
- **EPOCH**: Excellence in Processing Open Cultural Heritage - This 'network of excellence' is focused on improving the quality and effectiveness of the use of information and

communication technologies for cultural heritage applications. The network is co-ordinated by the University of Brighton, UK.

- **MinervaPLUS:** Ministerial Network for Valorising Activities in digitisation PLUS - This is a 'coordination action' - building on a similar activity funded under the 5th Framework Programme - that will harmonise activities relating to the digitisation of cultural and scientific content. The scope of the project specifically includes promoting recommendations and guidelines about digitisation, metadata, long-term accessibility, and preservation. MinervaPLUS is co-ordinated by the *Ministero per i Beni e le Attività Culturali*, Italy: Web site retrieved February 24, 2004, from: <http://www.minervaeurope.org/>

More information on all eight FP6 projects controlled by the EC 'Preservation and enhancement of cultural heritage' unit can be found in:

European Commission, DG Information Society. *DigiCULT Newsletter*, 5(1), February 2004: Retrieved February 24, 2004, from: <ftp://ftp.cordis.lu/pub/ist/docs/digicult/february2004.pdf>

Some projects funded under the 'Semantic-based knowledge systems' action line in FP6 may also have some relevance to digital preservation activities in the longer term, but these mostly relate to ontologies and Semantic Web developments rather than to preservation itself:

- **DIP:** Data Information and Process Integration with Semantic Web Services - This is an 'integrated project' co-ordinated by the National University of Ireland Galway with other research and industrial partners. Its objectives include the semantic enrichment of Web Services. Web site retrieved February 24, 2004, from: <http://dip.semanticweb.org/>
- **Knowledge Web:** This is a 'network of excellence' focused on fostering the use of ontology technology in industrial contexts. The network is co-ordinated by the University of Innsbruck, Austria. Web site retrieved February 24, 2004, from: <http://knowledgeweb.semanticweb.org/>

More information on FP6 can be found on the Cordis Web site, retrieved February 24, 2004, from: <http://fp6.cordis.lu/fp6/home.cfm>

## 1.6 World Summit on the Information Society

The World Summit on the Information Society (WSIS) takes place under the patronage of the UN (United Nations) Secretary-General and is an activity led by the International Telecommunication Union (ITU) in co-operation with other UN agencies. It aims to bring together member states, UN agencies, non-governmental organisations, private sector organisations, etc. to address the challenges of the fast-evolving global information society. The first phase of WSIS was held in Geneva, Switzerland on the 10-12 December 2003. It adopted a *Declaration of Principles* and *Plan of Action*. The principles uphold the role of existing cultural heritage institutions in the long-term preservation of digital information and records:

*Public institutions such as libraries and archives, museums, cultural collections and other community-based access points should be strengthened so as to promote the preservation of documentary records and free and equitable access to information* (Article 26).

In the Plan of Action, governments are encouraged to establish legislation on the preservation of public data (paragraph 10). Elsewhere, the plan promotes the "long-term systematic and efficient collection, dissemination and preservation of essential scientific digital data" (paragraph 22) and supports efforts for "developing systems for ensuring continued access to archived digital information and multimedia content" (paragraph 23). The second phase of WSIS will take place in Tunis, Tunisia on the 16-18 November 2005.

World Summit on the Information Society. *Declaration of principles*. WSIS-03/GENEVA/DOC/4-E, 12 December 2003. [http://www.itu.int/dms\\_pub/itu-s/md/03/wsis/doc/S03-WSIS-DOC-0004!!PDF-E.pdf](http://www.itu.int/dms_pub/itu-s/md/03/wsis/doc/S03-WSIS-DOC-0004!!PDF-E.pdf)

World Summit on the Information Society. *Plan of action*. WSIS-03/GENEVA/DOC/5-E, 12 December 2003. [http://www.itu.int/dms\\_pub/itu-s/md/03/wsis/doc/S03-WSIS-DOC-0005!!PDF-E.pdf](http://www.itu.int/dms_pub/itu-s/md/03/wsis/doc/S03-WSIS-DOC-0005!!PDF-E.pdf)

## 1.7 US National Digital Information Infrastructure and Preservation Program

The National Digital Information Infrastructure and Preservation Program (NDIIPP) has published a report compiled by Margaret Hedstrom (University of Michigan) entitled *It's about time: research challenges in digital archiving and long-term preservation*. The report is a product of a workshop on 'Research Challenges in Digital Archiving and Long-Term Preservation' held in Washington, D.C. in April 2002, jointly sponsored by the National Science Foundation (NSF) and the Library of Congress. Hedstrom's report summarises the discussions and recommendations of the workshop in order to help develop a research agenda for digital archiving and long-term preservation. The report first outlines what is at stake if the challenges of digital preservation are not addressed. After a survey of the current situation, it goes on to propose a research agenda based on the following categories:

- Technical architectures for repositories
- Attributes of collections
- Tools and technologies
- Organisational, economic and policy issues

There follows some notes on implementing the research agenda, concluding that "it's about time to launch a new research initiative" (p. 26).

Hedstrom, M. (2003). *It's about time: research challenges in digital archiving and long-term preservation*. Retrieved February 24, 2004, from: [www.digitalpreservation.gov/repor/NSF\\_LC\\_Final\\_Report.pdf](http://www.digitalpreservation.gov/repor/NSF_LC_Final_Report.pdf)

## 1.8 UK Digital Preservation Coalition

The British Broadcasting Corporation (BBC) and the Arts and Humanities Data Service (AHDS) joined the Digital Preservation Coalition (DPC) in October 2003 as associate members. Maggie Jones, the DPC's recently appointed (May 2003) co-ordinator, was interviewed in the February 2004 issue of RLG DigiNews. She talks about past achievements and future priorities, the DPC's membership and stakeholders, etc. The first DPC Technology Watch Report, by Brian Lavoie (OCLC Research) on the Reference Model for an Open Archival Information System (OAIS), was published in January 2004.

Lavoie, B. (2004). *The Open Archival Information System Reference Model: introductory guide*. DPC Technology Watch Series Report 04-01. OCLC Online Computer Library Center, Inc. and Digital Preservation Coalition, January 2004. [http://www.dpconline.org/docs/lavoie\\_OAIS.pdf](http://www.dpconline.org/docs/lavoie_OAIS.pdf)

RLG DigiNews. (2004). "Editor's interview: Maggie Jones, Digital Preservation Coalition (DPC)." *RLG DigiNews*, 8(1), February 2004. Retrieved February 24, 2004, from: <http://www.rlg.org/preserv/diginews/diginews8-1.html#feature2>  
<http://www.rlg.org/preserv/diginews/diginews8-1.html>

## 1.9 ERPANET project

ERPANET continues to provide new products and services to support digital preservation.



## ERPAePRINTS

At the IFLA Conference and General Congress, held in Berlin in August 2003, ERPANET launched 'ERPAePRINTS,' an Open Archive repository set up in conjunction with the DAEDALUS project and the Swiss National Archives to provide an ePrints preservation and access facility for the cultural and scientific heritage community. The repository seeks to acquire through deposit and preserve the full text of the research output of experts, scholars and researchers in the field of digital preservation and is intended complement existing publication routes and services.

The ERPAePRINTS service runs on the GNU EPrints Archive Software, a freely distributable archive system developed at the University of Southampton (retrieved February 24, 2004, from: <http://software.eprints.org/>). Deposits to the repository are encouraged and other institutions are invited to set up their own open archives for author self-archiving, using the EPrints.org software.

ERPAePRINTS Web pages, retrieved February 24, 2004, from: <http://eprints.erpanet.org/>

## ERPANET tools

ERPANET have developed a number of tools (ERPAtools) to promote proactive approaches to digital preservation challenges. Four tools are available in the first suite (retrieved February 24, 2004, from: [www.erpanet.org/www/products/04tools.html](http://www.erpanet.org/www/products/04tools.html)):

- Cost Orientation Tool - This tool provides an overview of cost factors that an agency might consider in determining the costs of digital preservation, presenting a table of factors and issues, cost impacts and considerations, along with a proposed approach for use of the tool. PDF document retrieved February 24, 2004, from: [www.erpanet.org/www/products/tools/ERPANETCostingTool.pdf](http://www.erpanet.org/www/products/tools/ERPANETCostingTool.pdf)
- Selecting Technologies Tool - This tool outlines the role of technology in digital preservation, the need to develop a digital preservation policy, strategy and set of requirements, and includes a table of factors for consideration when evaluating technologies to support them. PDF document retrieved February 24, 2004, from: [www.erpanet.org/www/products/tools/ERPANETSelect\\_Techno.pdf](http://www.erpanet.org/www/products/tools/ERPANETSelect_Techno.pdf)
- Digital Preservation Policy Tool - This tool describes the reasons for developing a digital preservation policy and provides guidance on areas for inclusion in such a policy, including scope and objectives, requirements, roles and responsibilities, context, coverage, costs, monitoring and review, and implementation. PDF document retrieved February 24, 2004, from: [www.erpanet.org/www/products/tools/ERPANETPolicyTool.pdf](http://www.erpanet.org/www/products/tools/ERPANETPolicyTool.pdf)
- Risk Communication Tool - This tool is intended to assist organisations to determine which digital resources are at risk in an organisation, to identify, categorise and prioritise the risks, to enable communication about risk areas and to stimulate risk management strategy development. PDF document retrieved February 24, 2004, from: [www.erpanet.org/www/products/tools/ERPANETRiskTool.pdf](http://www.erpanet.org/www/products/tools/ERPANETRiskTool.pdf)

## ERPANET events

The report from the ERPANET workshop on 'The long-term preservation of databases,' held at the Swiss National Archives in Bern, 9-11 April 2003, is now available. The Workshop report covers the practical experiences and development projects of seven initiatives from across Europe and the United States, as well as presenting steps in the database preservation process, including extraction, appraisal, description and access. Migration was favoured as an approach for database archiving in all the cases presented.

- ERPANET. *The long-term preservation of databases: ERPANET workshop report, Bern, April 9-11, 2003*. December 2003. Retrieved February 24, 2004, from: [www.erpanet.org/www/products/bern/Bern\\_Report\\_final.pdf](http://www.erpanet.org/www/products/bern/Bern_Report_final.pdf)

- ERPANET. *Metadata in digital preservation: ERPANET Training Seminar, Marburg, September 3-5, 2003*. Retrieved February 24, 2004, from: [www.erpanet.org/www/products/marburg/marburg.htm](http://www.erpanet.org/www/products/marburg/marburg.htm)

Presentations from these and other recent ERPANET events are also available on the project's Web site (retrieved February 24, 2004, from: <http://www.erpanet.org/>). These include the workshop jointly organised with the Italian Accademia Nazionale dei Lincei on 'Trusted Repositories for Preserving Cultural Heritage,' held in Rome, 17-19 November 2003, and the joint ERPANET and CODATA workshop on the 'Selection, Appraisal and Retention of Digital Scientific Data,' held in Lisbon, 15-17 December 2003. More details of these events are included in the Recent Events section (below).

### **ERPANET Topic of the Month**

In January 2004, ERPANET launched its *Topic of the Month* pages, which aim to bring together the knowledge and expertise ERPANET has accumulated to date on various themes and topics in the field of digital preservation. The focus topic for January 2004 was "metadata" and the pages provide an editorial on this topic plus links to a range of associated resources, commentaries, studies and reports, including a transcript and summary of the 'erpaCHAT' expert online discussion of digital preservation metadata, which took place on 6 November 2003 following the ERPANET training seminar on metadata, held in Marburg, Germany on the 3-5 September 2003.

ERPANET *Topic of the Month* pages. Retrieved February 24, 2004, from: [www.erpanet.org/www/products/topic/topic.htm](http://www.erpanet.org/www/products/topic/topic.htm)

## **1.10 JISC Supporting Institutional Records Management programme**

Twelve projects were funded under Theme 1 of the UK Joint Information Systems Committee's Supporting Institutional Records Management programme (JISC 09/02) to explore the practical and theoretical problems associated with applying the JISC-funded *Study of the Records Lifecycle* in higher education institutions. In December 2003, a series of studies from these projects were made available. These focus on different aspects of implementing the lifecycle study, e.g. student and employee records, e-mail, Web sites, primary research data, etc.

JISC Supporting Institutional Records Management programme. Retrieved February 24, 2004, from: [http://www.jisc.ac.uk/index.cfm?name=programme\\_supporting\\_irm](http://www.jisc.ac.uk/index.cfm?name=programme_supporting_irm)



## 2. Specific Areas of Activity

### 2.1 Web Archiving

The changeability of the web and the ephemeral nature of web-published documents provide the basis for several articles appearing recently.

#### The longevity of Web references

A study by Dellavalle, *et al.* (2003) examined Internet citations from papers published in three major scientific and medical journals (*New England Journal of Medicine*, *JAMA*, *Science*) after three, 15 and 27 months, and found the prevalence of inactive links increasing from 3.8 to 13 percent over the period. Possible solutions to the problem are discussed, including the use of persistent identifiers. Dellavalle's team have also suggested (in a letter published in *The Lancet*) that authors could "submit all referenced internet information to the Internet Archive" (Johnson, *et al.*, 2004). Dellavalle's findings and other studies on the problem were also discussed in an article appearing in the *Washington Post* (Weiss, 2003).

Wallace Koehler (2004) provides an update on the extension of a study, to May 2003, on the longevity of a single set of URLs (360 web pages and 343 web sites), monitored weekly since December 1996. While only 33.8% of the sample pages persisted at the original URLs, some inferences are drawn about the relative stability of navigation pages and content pages and distribution of page stability across top-level domains with age.

From a slightly different perspective, Burke, Germain and Van Ullen (2003) raise the issue of the inclusion of Web sites in library catalogues. The volatility of Web resources and the fact that they are not under direct control of the library means that their inclusion in library catalogues risks its accuracy and reliability. The authors' survey of ARL (Association of Research Libraries) institutions revealed a large variation in the number of linking errors in library catalogues. As with Koehler's paper, the authors conclude with some observations on the stability of certain domains and types of Web page and suggest that these could inform future collection development policies.

Burke, G., Germain, C. A., & Van Ullen, M. K. (2003). "URLs in the OPAC: integrating or disintegrating research libraries' catalogs." *Journal of Academic Librarianship*, 29(5), 290-297.

Dellavalle, R. P., Hester, E. J., Heilig, L. F., Drake, A. L., Kuntzman, J. W., Graber, M., & Schilling, L. M. (2003). "Going, going, gone: lost Internet references." *Science*, 302 (5646), 787-788, 31 October 2003.

Johnson, K. R., Hester, E. J., Schilling, L. M., & Dellavalle, R. P. (2004). "Addressing Internet reference loss" [letter]. *The Lancet*, 363, 660-661, 21 February 2004.

Koehler, W. (2004) "A longitudinal study of Web pages continued: a consideration of document persistence". *Information Research*, 9(2), January 2004. Available at Retrieved February 24, 2004, from: <http://informationr.net/ir/9-2/paper174.html>

Spinellis, D. (2003). "The decay and failure of Web references." *Communications of the ACM*, 46(1), 71-77.

Weiss, R. (2003). "On the Web, research work proves ephemeral." *Washington Post*, 24 November 2003, p. A08. Retrieved February 24, 2004, from: <http://www.washingtonpost.com/ac2/wp-dyn/A8730-2003Nov23?language=printer>

#### Web-archiving initiatives

Cruse *et al.* (2003) reported on a project addressing the roles of libraries and other memory institutions in preserving web-based government information in the United States. The report outlined the project team's methodology, including a demographic review of the US dot-gov

domain, and also noted the challenges in the capture, curation, and persistent management of web-based materials, with discussion of cost elements and a route map for service implementation.

Juha Hakala (2003) provided an update to a report on European experiences in archiving the Web, extending the discussion of the NEDLIB project and the Nordic Web Archive (NWA) to include current figures and a description of the work being undertaken by the International Internet Preservation Consortium to develop new web archiving tools.

Margaret Phillips has described Australia's Web archive PANDORA and its digital archiving system PANDAS in a recent article published in *DigiCULT.Info* (Phillips, 2003). As well as discussing the roles and differing approaches of national libraries in relation to web harvesting, the article provides an overview of the selection criteria used for PANDORA, the processes of digital archiving, and the main functions of the PANDAS software. Future directions for the PANDAS archiving system and digital archiving initiatives such as the 'Deep Web' were also described.

A newly formed UK Web Archiving Consortium is in the process of tendering for a common Web archive infrastructure for the UK. The members of the consortium are the British Library, the Joint Information Systems Committee, the National Archives, the National Library of Scotland, the National Library of Wales, and the Wellcome Trust. The consortium is looking for a contractor that will provide the infrastructure for the capture of selected sites, to provide public access and technical support. The initial contract period is for a two-year pilot, in which it is expected approximately 6,000 sites will be collected.

Cruse, P., Eckman, C., Kunze, J., Christenson, H., Colvin, J., Hutton, C., & Greenstein, D. (2003). *Web-based government information: evaluating solutions for capture, curation and preservation*. Oakland, Calif.: California Digital Library, November 2003. Retrieved February 24, 2004, from: [http://www.cdlib.org/programs/Web-based\\_archiving\\_mellon\\_Final.pdf](http://www.cdlib.org/programs/Web-based_archiving_mellon_Final.pdf)

Hakala, J. (2003). "Archiving the Web: European experiences." *Teitolinja*, 2/2003, December 2003. [URN:NBN:fi-fe20031951]. Retrieved February 24, 2004, from: <http://www.lib.helsinki.fi/tietolinja/0203/webarchive.html>

Phillips, M. E. (2003). "PANDORA, Australia's Web archive, and the digital archiving system that supports it." *DigiCULT.info*, 6, 24-28. December 2003. Retrieved February 24, 2004, from: [http://www.digicult.info/downloads/dc\\_info\\_issue6\\_december\\_20031.pdf](http://www.digicult.info/downloads/dc_info_issue6_december_20031.pdf)

### **Technologies and tools**

On the technology front, the NWA (Nordic Web Archive) Toolset has been released as open source software under the GNU General Public License. The Toolset allows search and navigation functions within web archives via standard web browsers, and comprises a Document Retriever, which delivers documents to other modules, an Exporter, which prepares documents for search engine indexing, and an Access module which interfaces with a search engine and the Document Retriever.

NWA Toolset. Retrieved February 24, 2004, from:  
<http://nwa.nb.no/index.php?doc=aboutNwaToolset>

The NWA Toolset and associated documentation are available for downloading from the National Library of Norway Web site. Retrieved February 24, 2004, from:  
<http://nwa.nb.no/index.php?doc=download>

### **Archiving dynamically-generated content**

A recent paper by Fitch (2003) discussed an approach to archiving dynamically generated responses from a web site as they are delivered. The results of changes to content and content generation systems, as manifested in the changed pages a site generates in response to delivery

requests, are inspected via the outgoing server, and materially different responses are archived as a record of the site as delivered at particular points in time.

Fitch, K. (2003). "Web site archiving: an approach to recording every materially different response produced by a Website." AusWeb 2003: the Ninth Australian World Wide Web Conference, Hyatt Sanctuary Cove, Gold Coast, Australia, 5-9 July 2003. Retrieved February 24, 2004, from: <http://ausweb.scu.edu.au/aw03/papers/fitch/>

### **Full-text searching of the Internet Archive**

The Internet Archive has now added a feature called "Recall" for full-text searching of the archive. This is a search engine (currently in beta) that indexes the text of over 11 billion pages, dating from 1996. Users can search for text strings and limit result by date.

Internet Archive Recall. Retrieved February 24, 2004, from: <http://recall.archive.org/>

Background information at: Retrieved February 24, 2004, from:  
[ia00406.archive.org/about.html](http://ia00406.archive.org/about.html)

## **2.2 Formats and Tools**

### **JHOVE (JSTOR/Harvard Object Validation Environment)**

JSTOR and the Harvard University Library (2003) have collaborated to produce JHOVE (pronounced "Jove"), an open source, extensible, Java-based application designed to automate the identification, validation and characterisation of file formats. The process is based on analysis of files themselves against defined profile schemata rather than filename extensions. Several standard modules are currently available (ASCII, UTF-8, GIF, JPEG, TIFF, PDF, XML and "bitstream") and the ability to define further modules is a feature of the software. The software, modules and documentation are available for download under a GNU General Public License.

Harvard University Library & JSTOR. (2003). *JHOVE: JSTOR/Harvard Object Validation Environment: Digital Format Specific Validation*. Cambridge, Mass.: Harvard University Library, December 2003. Retrieved February 24, 2004, from:  
<http://hul.harvard.edu/jhove/jhove.html>

### **Release of PRONOM 3**

The Web-enabled version 3 of the UK National Archives PRONOM database is now available online. The database can be searched by file extension, product or vendor name, and release or support dates, and can be used to identify formats, compatible products and support coverage. The initial data load is intended to include the most commonly used office products for PC operating systems, with the aim to load information on around 450 products shortly, followed by ongoing work to verify information on a further 3,000 formats. Contributions to the data collection are encouraged and may be submitted via an online form. Access the web-enabled PRONOM database and format submission form from the National Archives Public Record Office Web pages, retrieved February 24, 2004, from: <http://www.records.pro.gov.uk/pronom/>

An overview of the purpose, content development strategies and future development plans for PRONOM was published in the October 2003 issue of *RLG DigiNews* (Darlington, 2003).

Darlington, J. (2003). "PRONOM: a practical online compendium of file formats," *RLG DigiNews*, 7(5), October 2003. Retrieved February 24, 2004, from:  
<http://www.rlg.org/preserv/diginews/diginews7-5.html#feature2>

### **MPEG-21 Digital Item Description Language**

An article in the November 2003 issue of *D-Lib Magazine* (Bekaert, Hochstenbach & Van de Sompel, 2003) looks at the major characteristics of the MPEG-21 Digital Item Description Language (DIDL), an XML packaging format for multimedia, which provides broad flexibility and extensibility for complex object representation. Studies at the Research Library of the Los Alamos

National Laboratory (LANL) into use of DIDL for the representation of complex objects have led to a decision to make DIDL-conformant documents the unit of storage in the LANL repository, and suggest that DIDL may be useful for other digital library applications. A further article in *D-Lib Magazine* (Bekaert, *et al.*, 2004) looks at the application of MPEG-21 Digital Item Processing (DIP) and NISO OpenURL for the dynamic dissemination of complex digital objects.

Bekaert, J., Hochstenbach, P., & Van de Sompel, H. (2003). "Using MPEG-21 DIDL to represent complex digital objects in the Los Alamos National Laboratory Digital Library" *D-Lib Magazine*, 9(11), November 2003. Retrieved February 24, 2004, from: <http://www.dlib.org/dlib/november03/bekaert/11bekaert.html>

Bekaert, J., Balakireva, L., Hochstenbach, P., & Van de Sompel, H. (2004). "Using MPEG-21 DIP and NISO OpenURL for the dynamic dissemination of complex digital objects in the Los Alamos National Laboratory Digital Library." *D-Lib Magazine*, 10(2), February 2004. Retrieved February 24, 2004, from: <http://www.dlib.org/dlib/february04/bekaert/02bekaert.html>

## **PDF/A**

Progress in the development of PDF/A, a restricted set of characteristics for Adobe's PDF format amenable to long-term management of text-based documents, is described in a recent RLG DigiNews article (LeFurgy, 2003). Characteristics of the proposed standard, now under consideration for further development by a Joint Working Group of the International Organization for Standardization (ISO), include legally-compliant embedding of fonts, device-independent specification of colour spaces, and strict exclusion of audio and video, JavaScript, executable file launches and encryption.

LeFurgy, W.G. (2003). "PDF/A: developing a file format for long-term preservation." *RLG DigiNews*, 7(6), December 2003. Retrieved February 24, 2004, from: [http://www.rlg.org/preserv/diginews/v7\\_n6\\_feature1.html](http://www.rlg.org/preserv/diginews/v7_n6_feature1.html)

## **University of Leeds, Rendering and Representation Project**

Further outputs from the Rendering and Representation Project at the University of Leeds are available. Results on further development of the Migration on Request concept are presented, including a case study using bitmap colour spaces to investigate development of canonical intermediate structures for interchange, discussion of use of a minimal subset of Java for long-life programming, and the design of a tool for the Migration on Request of text from HTML to other formats.

Further work on the Representation Networks concept has produced a case study on the creation and maintenance of a representation network for a particular file type, and the initial design framework for a database to support the rendering of preserved digital objects.

Representation and Rendering Project Web site. Retrieved February 24, 2004, from: <http://www.leeds.ac.uk/reprend/>

## **XML as a preservation format**

Andreas Aschenbrenner (ERPANET) has published in the February 2004 issue of *RLG DigiNews*, a study of data formats, focusing in particular on XML. XML formats are often proposed as candidates for standard formats that can be used for preservation. In response, Aschenbrenner notes the lack of suitability of XML for many data types (e.g. images, multimedia, large-scale scientific data), finds that its human-readability is dependent on an understanding of tag labels and semantics, as is its reusability and interoperability.

Aschenbrenner, A. (2004). "The bits and bites of data formats: stainless design for digital endurance." *RLG DigiNews*, 8(1), February 2004. Retrieved February 24, 2004, from: <http://www.rlg.org/preserv/diginews/diginews8-1.html#feature3>

## 2.3 Metadata

### Preservation Metadata

The PREMIS (PREservation Metadata Implementation Strategies) Group is composed of representatives from library, academic, museum, government and commercial communities, and seeks to carry forward the efforts of the OCLC/RLG Preservation Metadata Working Group, with a focus on the practical aspects of implementing preservation metadata in digital preservation systems.

Also sponsored by OCLC and RLG, the PREMIS group's objectives are to develop an implementable set of broadly applicable "core" preservation metadata elements and a data dictionary, to examine strategies for encoding, storage, management and exchange of preservation metadata, to conduct pilot programs to test recommendations and best practice in a variety of settings and to explore options for creation and sharing of preservation metadata.

The PREMIS group is divided into two subgroups:

- The **Core Elements Subgroup** - charged with determining a core set of elements, with consideration of what is implementable, and development of a data dictionary.
- The **Implementation Strategies Subgroup** - charged with examining encoding, storage, management and exchange strategies and testing implementation of recommendations through pilot programs.

Activities of the Core Elements Subgroup have been to define terminology and a methodology for comparing element sets, to map elements from various preservation metadata implementations against the OCLC/RLG framework, and work through the elements. Recent discussions have focused on identifiers, types of relationships between objects and how they apply to objects at different levels, the treatment of events within various implementations, and development of a typology of entities for digital objects, to clarify at which levels given elements would apply. Work on a data dictionary for core elements is progressing.

The Implementation Strategies Subgroup have developed, tested and revised an Implementation Survey on metadata practices in digital preservation repositories, which was subsequently distributed to a range of preservation initiatives and made available on the OCLC web site for response. The Subgroup is currently compiling a summary of results for analysis.

For more information about PREMIS activities and Subgroups, visit:

- PREMIS: Working Group. Retrieved February 24, 2004, from: <http://www.oclc.org/research/projects/pmwg/>
- Core Elements Subgroup: Retrieved February 24, 2004, from: [http://www.oclc.org/research/projects/pmwg/core\\_elements.htm](http://www.oclc.org/research/projects/pmwg/core_elements.htm)
- Implementation Strategies Subgroup: Retrieved February 24, 2004, from: <http://www.oclc.org/research/projects/pmwg/implementation.htm>

### DCMI Preservation Working Group

At the DCMI-2003 conference in Seattle, a DCMI (Dublin Core Metadata Initiative) working group on preservation was proposed to act as a discussion forum for individuals and organisations with an interest in preservation metadata. The DCMI Preservation Working Group exists to collect information on and review existing preservation metadata schemas, to investigate the need for domain-specific schemas, and to liaise with other preservation metadata initiatives like PREMIS. Joint chairs are Andrew Wilson of the National Archives of Australia and Heike Neuroth of the Göttingen State and University Library. Membership of the group, as with all other DCMI Working Groups, is open to interested members of the community with the time and expertise relevant to the domain of the group.



DCMI Preservation Working Group: Retrieved February 24, 2004, from:  
<http://dublincore.org/groups/preservation/>

### **Open Archives Initiative Protocol for Metadata Harvesting**

The use and development of the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) have been the subjects of two recent papers. Boston (2003) describes the use of the OAI-PMH at the National Library of Australia, in a paper given at the 'International Seminar on the Use of Standards in the Development of Online Access Systems for Archives,' held in Canberra in October 2003. The paper looked at the advantages of using the protocol on Library services such as PictureAustralia, the Australian National Bibliographic Database and the library's digital collections, and also provided a brief description of OAI software.

The aim to incorporate structured rights expressions about metadata and resources into the OAI-PMH is the basis of a recent white paper (Lagoze, *et al.*, 2003), which describes the scope and issues to be investigated, and which will result in a set of OAI-PMH guidelines, scheduled for release in the second quarter of 2004.

Boston, T. (2003). *National Library of Australia initiatives using the Open Archives Initiative Protocol For Metadata Harvesting*. Canberra: National Library of Australia, October 2003. Retrieved February 24, 2004, from:  
<http://www.nla.gov.au/nla/staffpaper/2003/boston1.html>

Lagoze, C., Van de Sompel, H., Nelson, M., & Warner, S. (2003). *OAI Rights White Paper*. Open Archives Initiative, 26 September 2003. Retrieved February 24, 2004, from:  
<http://www.openarchives.org/documents/OAIRightsWhitePaper.html>

### **DFG Long-term preservation of digital publications project**

Researchers at the University of the Federal Armed Forces in Munich published a paper on extending and implementing the OAIS (Open Archival Information System) reference model (Rödig, *et al.*, 2003). This reported on research carried out as part of a German Research Foundation (DFG) funded project called "*Langzeitarchivierung digitaler Publikationen*" (long term preservation of digital publications), led by the Bavarian State Library. The paper looks at three main topics: modelling the 'detachment' of digital objects from their media as part of the ingest function, investigating the role of database management systems in supporting the functions of an OAIS, and looking at the prospects for the semi-automatic generation of metadata. Other project outcomes have been published (in German) in Rödig, Pfeiffer and Borghoff (2002) and in a book (Borghoff, *et al.*, 2003), described in section 3 (below).

*Langzeitarchivierung digitaler Publikationen* project Web site: Retrieved February 24, 2004, from: [ist.unibw-muenchen.de/Inst2/Research/LZA/index.html](http://ist.unibw-muenchen.de/Inst2/Research/LZA/index.html)

Rödig, P., Borghoff, U. M., Scheffczyk, J., & Schmitz, L. (2003). "Preservation of digital publications: an OAIS extension and implementation." *Proceedings of the 2003 ACM Symposium on Document Engineering, Grenoble, France*. Full-text available to subscribers only through the ACM Digital Library. Abstract retrieved February 24, 2004, from: <http://doi.acm.org/10.1145/958220.958245>

Rödig, P., Pfeiffer, E., & Borghoff, U. (2002). *Langzeitarchivierung digitaler Medien*. Bericht Nr. 2002-02. Munich: Universität der Bundeswehr München, Fakultät für Informatik, June 2002. Retrieved February 24, 2004, from: [www2-data.informatik.unibw-muenchen.de/Research/Projects/LZA/lza\\_techrep.pdf](http://www2-data.informatik.unibw-muenchen.de/Research/Projects/LZA/lza_techrep.pdf)

### **Library of Congress metadata overview**

In a short paper based on a workshop presentation, Sally McCallum reports on three main aspects of the Library of Congress's metadata activities. Firstly, she notes the continued primacy of stable standards like the Anglo-American Cataloguing Rules (AACR), MARC21, and the Encoded Archival Description (EAD) for bibliographic description and control. Secondly, she introduces the use of an XML-based MARC21 format (MARCXML) for promoting interoperability



between MARC21 records in ISO 2709 format and XML-based usage scenarios, e.g. for harvesting by the OAI protocol, for managing transformations to and from formats like the Metadata Object Description Schema (MODS), ONIX, Dublin Core, etc. Thirdly, McCallum notes broader metadata needs, e.g. for technical and rights metadata, noting Library of Congress involvement in developing and deploying the Metadata Encoding and Transmission Standard (METS).

McCallum, S.H. (2003). "Library of Congress metadata landscape." *Zeitschrift für Bibliothekswesen und Bibliographie*, 50(4), 182-187, July-August 2003. Retrieved February 24, 2004, from: [zfbib.thulb.uni-jena.de/servlets/DerivateServlet/Derivate-85/j03-h4-auf-2.pdf](http://zfbib.thulb.uni-jena.de/servlets/DerivateServlet/Derivate-85/j03-h4-auf-2.pdf)

## Two perspectives on Dublin Core

In the first of two short articles, Wendy Duff (University of Toronto) argues that while the goals of the Dublin Core Metadata Initiative (DCMI) are admirable, the Dublin Core metadata element set has many limitations from an archivist's perspective. She notes its relatively low deployment on the Web and her doubts about the element set's 'domain neutrality,' arguing that all taxonomies are "grounded upon, shaped by, and reflect the world view of their creators" (p. 32). In reply, Andrew Wilson (National Archives of Australia) defends the Dublin Core, noting that it was only ever intended to be "a set of signposts for digital surfers." On the positive side, Wilson argued that the DCMI had successfully demonstrated that a diverse group of individuals could work together to form a global community. He also noted the importance of DCMI for supporting work on tools and technologies that support all types of metadata, e.g. with regard to metadata registries and the Semantic Web.

Duff, W. (2003). "The Dublin Core and its limitations." *DigiCULT.Info*, 6, 31-32, December 2003. Retrieved February 24, 2004, from: [http://www.digicult.info/downloads/dc\\_info\\_issue6\\_december\\_20031.pdf](http://www.digicult.info/downloads/dc_info_issue6_december_20031.pdf)

Wilson, A. (2003). "Why the Dublin Core Metadata Initiative (DCMI) is important." *DigiCULT.Info*, 6, 32-34, December 2003. Retrieved February 24, 2004, from: [http://www.digicult.info/downloads/dc\\_info\\_issue6\\_december\\_20031.pdf](http://www.digicult.info/downloads/dc_info_issue6_december_20031.pdf)

## Open Language Archives Community (OLAC)

There are an increasing number of language resources becoming available in digital form, including both data (lexicons, grammars, field notes, recordings, etc.) and the software tools used to create and view them (Simons & Bird, 2003). Linguistic scholars and others have long noted problems with, for example, the discovery of language resources, the non-standard ways in which they are created, and their long-term accessibility. In order to respond to these problems, the Open Language Archives Community (OLAC) was formed in 2000 to help develop consensus on best practice for the digital archiving of language resources, also to act as a network of interoperating repositories and services. Steven Bird (the universities of Melbourne and Pennsylvania) and Gary Simons (SIL International) have recently produced a series of publications on various aspects of OLAC. These include papers on its basic infrastructure (Simons & Bird, 2003) and an application profile of Dublin Core metadata that can be harvested using the OAI protocol (Bird & Simons, 2003a), and a study of the 'portability' of languages over time and between application domains and scholarly communities (Bird & Simons, 2003b).

OLAC: Open Language Archives Community. Retrieved February 24, 2004, from: [www.language-archives.org/](http://www.language-archives.org/)

Bird, S., & Simons, G. (2003a). "Extending Dublin Core metadata to support the description and discovery of language resources." *Computers and the Humanities*, 37(4), 375-388. Retrieved February 24, 2004, from: <http://arxiv.org/abs/cs.CL/0308022>

Bird, S., & Simons, G., (2003b). "Seven dimensions of portability for language documentation and description." *Language*, 79(3), 557-582. Preprint retrieved February 24, 2004, from: [www.language-archives.org/documents/portability.pdf](http://www.language-archives.org/documents/portability.pdf)

Simons, G., & Bird, S. (2003). "The Open Language Archives Community: an infrastructure for distributed archiving of language." *Literary and Linguistic Computing*, 18(2), 117-128. Preprint retrieved February 24, 2004, from: <http://arxiv.org/abs/cs.CL/0306040>

### Other recent publications

Caplan, P. (2003). *Metadata fundamentals for all librarians*. Chicago, Ill.: American Library Association. ISBN 0-83890-847-0.

An introduction to metadata for librarians and others working in library environments by Priscilla Caplan (Florida Center for Library Automation). The book introduces metadata schemes and profiles, syntax and content rules, and provides overviews of many standards - organised by domain. The book was reviewed by Stuart Sutton (University of Washington) in the December 2003 issue of *D-Lib Magazine*. Retrieved February 24, 2004, from: <http://www.dlib.org/dlib/december03/12bookreview.html>

Gorman, G. E., & Dorner, D. G., eds., (2004). *Metadata applications and management*. (International Yearbook of Library and Information Management, 2003-2004). London: Facet Publishing. ISBN 1-85604-474-2.

The *International Yearbook of Library and Information Management* (IYLIM) is an annual edited volume covering a specific subject area. The topic of the latest volume is metadata and the book contains fifteen chapters on various metadata-related topics, including a chapter by Michael Day (UKOLN, University of Bath) on "Preservation metadata" (pp. 253-273). In the United States, the volume is distributed by Scarecrow Press (ISBN 0-8108-4980-1)

## 2.4 Legal issues

### Copyright and licensing

A seminar on Copyright and Licensing for Digital Preservation was held on the 19th November 2003 in London. Attendees included publishers, librarians, legal and digital preservation experts. The seminar featured a presentation on the UK Arts and Humanities Research Board (AHRB) funded Copyright and Licensing for Digital Preservation (CLDP) project as well as discussion focused on preservation copying requirements, copyright issues and relationships with publishers. A seminar report describes constraints to digital preservation and proposed recommendations on how these may be overcome:

Copyright and Licensing for Digital Preservation Project. (2003). *Copyright and Licensing for Digital Preservation Project Seminar*. Loughborough University, Department of Information Science. Retrieved February 24, 2004, from: [http://www.lboro.ac.uk/departments/ls/disresearch/CLDP/Project\\_reports.htm](http://www.lboro.ac.uk/departments/ls/disresearch/CLDP/Project_reports.htm)

Further papers are available from the CLDP project, which is investigating how copyright legislation and licensed access to digital content may limit the ability of libraries to provide long-term access to that content and suggesting ways in which these problems can be overcome. The aims of the CLDP Project, and copyright issues involved in copying for digital preservation, extension of legal deposit legislation, licensing and perpetual access are described in papers by Adrienne Muir. A report for CLDP by Michael Norris examines the copyright laws of twenty-five countries to identify clauses allowing the copying of copyright material for preservation, as well as other clauses which may in other ways help or hinder preservation, including those dealing with adaptation of computer programs or circumvention of copy protection.

Muir, A. (2003). "Copyright and licensing for digital preservation." *Library & Information Update*, 2(6) June 2003, 34-36. Retrieved February 24, 2004, from: <http://www.cilip.org.uk/update/issues/jun03/article2june.html>

Muir, A. (2003). "Copyright and licensing issues for digital preservation and possible solutions." In: *Proceedings of the ICC/IFIP Seventh International Conference on Electronic Publishing*, Universidade do Minho, Guimarães, Portugal, 25-28 June 2003.

Norris, M. (2004). "An international survey of archival, preservation and related clauses in copyright law." Loughborough University, 19 February 2004. Retrieved February 24, 2004, from: [http://www.lboro.ac.uk/departments/is/disresearch/CLDP/Project\\_reports.htm](http://www.lboro.ac.uk/departments/is/disresearch/CLDP/Project_reports.htm)

Peter B. Hirtle discusses US copyright law and possible options for libraries and archives for copying of digital materials for preservation purposes or under "fair use" provisions.

Hirtle, P. B. (2003). "Digital Preservation and Copyright." In: *Copyright and Fair Use*, Stanford University Libraries, November 2003. Retrieved February 24, 2004, from: [http://fairuse.stanford.edu/commentary\\_and\\_analysis/2003\\_11\\_hirtle.html](http://fairuse.stanford.edu/commentary_and_analysis/2003_11_hirtle.html)

### **Anti-circumvention laws**

One of the legal issues that are often raised in digital preservation contexts is the right for libraries and other repository institutions to reverse engineer computer programs for preservation. While his review does not focus on the preservation aspects, Carlos Fernandez-Molina has provided a good review of anti-circumvention laws in the US Digital Millennium Copyright Act (DCMA), European Union directives on conditional access (1998) and copyright (2001), and the Australian Copyright Amendment (Digital Agenda) Act 2000. He concludes that rights holders currently benefit from several cumulative layers of protection, "copyright, technological protection, legal protection of the technological measures, and contract law" (p. 63). Claire Elizabeth Craig has produced a more focused study on the impact of e-books on the US library system. She uses the arrest of the Russian programmer Dmitry Sklyarov under the DCMA for developing software that circumvented digital-rights management (DRM) restraints built into proprietary e-book reader software as the basis of a study of the use of e-books in US libraries. She argues for a better balance between the needs of rights holders and readers, concluding that, unless this can be achieved, the e-book market has "no real prospect of ever expanding beyond an esoteric niche market" (p. 1113).

Craig C.E. (2003). "'Lending' institutions: The impact of the e-book on the American library system." *University of Illinois Law Review*, 2003(4): 1087-1113. Retrieved February 24, 2004, from: [http://home.law.uiuc.edu/rev/publications/2000s/2003/2003\\_4/craig.pdf](http://home.law.uiuc.edu/rev/publications/2000s/2003/2003_4/craig.pdf)

Fernandez-Molina, J.C. (2003). "Laws against the circumvention of copyright technological protection." *Journal of Documentation*, 59(1), 41-68.

### **Changes in UK legal deposit law**

The UK Legal Deposit Libraries Act 2003 received Royal Assent on 31 October 2003. This legislation, resulting from a Private Members Bill introduced by Chris Mole MP, enshrines the general principle that electronic publications and other non-print materials should be subject to legal deposit. The Act allows for secondary legislation to be approved by the UK Parliament that will ensure that such publications can be included within the legal deposit system.

United Kingdom. *Legal Deposit Libraries Act 2003*. 2003 Ch. 28. London: The Stationery Office. ISBN 0-10-542803-5. Retrieved February 24, 2004, from: <http://www.legislation.hmso.gov.uk/acts/acts2003/20030028.htm>

British Library. (2003). "Historic change in Legal Deposit Law saves electronic publications for future generations - Bill to extend legal deposit to UK non-print materials receives Royal Assent" [press-release]. London: British Library Press & Public Relations, 31 October 2003. Retrieved February 24, 2004, from: [www.bl.uk/cgi-bin/press.cgi?story=1382](http://www.bl.uk/cgi-bin/press.cgi?story=1382)

## **2.5 Institutional repositories, e-prints and e-journals**

### **Final versions of JISC reports**

The Joint Information Systems Committee (JISC) has now published the final versions of reports into the preservation of e-prints and e-journals:

James, H., Ruusalepp, R., Anderson, S., & Pinfield, S. (2003). *Feasibility and requirements study on preservation of e-prints*. London: Joint Information Systems Committee, 29 October 2003. Retrieved February 24, 2004, from: [http://www.jisc.ac.uk/uploaded\\_documents/e-prints\\_report\\_final.pdf](http://www.jisc.ac.uk/uploaded_documents/e-prints_report_final.pdf)

Jones, M. (2003). *Archiving e-journals consultancy: final report*. London: Joint Information Systems Committee, October 2003. Retrieved February 24, 2004, from: [http://www.jisc.ac.uk/uploaded\\_documents/ejournalsfinal.pdf](http://www.jisc.ac.uk/uploaded_documents/ejournalsfinal.pdf)

### Open-access journals

The publication in October 2003 of the first issue of the Public Library of Science's first 'open access' journal *PLoS Biology* (retrieved February 24, 2004, from: <http://www.plosbiology.org/>) has generated much comment on the open access publishing model for scholarly and scientific journals. The basis of the open access model is that journal content is made freely available to end-users while production and review costs are covered in other ways, e.g. by charges levied on authors. The ultimate success of the open access business model will depend upon the willingness of those who fund research to pay for publication 'up-front.' Some funding organisations are now beginning to take note of this. In early October 2003, the Wellcome Trust - a UK charity that supports medical research - issued a position statement that supported "open and unrestricted access to the published outputs of research" and said that it was prepared to meet the cost of publication charges. Later the same month, an international group of research funding organisations - including the *Deutsche Forschungsgemeinschaft* (DFG), the *Max-Planck-Gesellschaft*, and the French *Centre national de la recherche scientifique* (CNRS) - signed the 'Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities,' encouraging grant recipients to publish their work according to the principles of the open access paradigm. In December 2003, the World Summit on the Information Society (WSIS) endorsed the promotion of "universal access with equal opportunities for all to scientific knowledge and the creation and dissemination of scientific and technical information, including open access initiatives for scientific publishing" (WSIS Declaration of principles, paragraph 28). See section 1.6 (above) for more information on WSIS.

The implications of open access publishing for preservation and long-term access have yet to be fully understood. That said, BioMed Central - one of the commercial pioneers of the open access approach - has made some efforts to address preservation concerns by making an agreement with the National Library of the Netherlands (KB) for it to act as an "official archival agent" for the publisher. The library is committed to updating the archive as technology changes, thus facilitating what the press-release calls "open access in perpetuity." BioMed Central conceives of the KB's archive as being part of an interim service system in the event of a disaster, or as a means of maintaining open access if the publisher (or its successors) no longer make the journals available.

Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities. Retrieved February 24, 2004, from: <http://www.zim.mpg.de/openaccess-berlin/berlindeclaration.html>

Wellcome Trust position statement in support of open access publishing. Retrieved February 24, 2004, from: [www.wellcome.ac.uk/en/1/awtvispolpub.html](http://www.wellcome.ac.uk/en/1/awtvispolpub.html)

BioMed Central. (2003). "National Library of the Netherlands and BioMed Central agree to open access archive" [press release]. BioMed Central, 17 September 2003. Retrieved February 24, 2004, from: <http://www.biomedcentral.com/info/about/pr-releases?pr=20030917>

### Other publications

Chang, S. -H. (2003). "Institutional repositories: the library's new role" [editorial]. *OCLC Systems & Services*, 19(3), 77-79.

Falk, H. (2003). "Digital archive developments." *The Electronic Library*, 21(4), 375-379.

Gadd, E., Oppenheim, C., & Proberts (2003). "RoMEO studies 2: how academics want to protect their open-access research papers." *Journal of Information Science*, 29(5), 333-356.

Shearer, M. K. (2003). "Institutional repositories: towards the identification of critical success factors." *Canadian Journal of Information and Library Science = Revue Canadienne des Sciences de l'Information et de Bibliothéconomie*, 27(3), 89-108.

## 2.6 Digital library futures

The future of digital libraries has been a topic for discussion at several recent events, including ECDL 2003 (Trondheim, Norway, August 2003) and the NSF Workshop on Research Directions for Digital Libraries (Chatham, Mass., USA, June 2003). In late 2003, the *Journal of Academic Librarianship* published three papers outlining visions of future digital libraries from the Digital Libraries Symposium held at the ALA Midwinter meeting in Philadelphia, Penn., on the 25th January 2003. The papers, by Deanna Marcum (Library of Congress), Ann Okerson (Yale University) and Clifford Lynch (Coalition for Networked Information) cover a wide range of issues, e.g. the building and long-term maintenance of digital content, copyright, changes in scholarly communication, etc.

A draft report from the NSF workshop on research directions for digital libraries is also now available. The workshop was convened to look at requirements following the two-phase Digital Libraries Initiative (DLI) and other research programmes. The draft report recommends continued funding of research (including some into the 'long-term availability' of resources) as well as for significant investment into key digital resources and services.

Larsen, R., & Wactlar, H. (2004). *Knowledge lost in information: report of the NSF Workshop on Research Directions for Digital Libraries, June 15-17, 2003, Chatham, MA* [working draft]. Retrieved February 24, 2004, from:  
<http://www.sis.pitt.edu/~dlwshop/report.pdf>

Lynch, C. (2003). "Digital library opportunities." *Journal of Academic Librarianship*, 29(5), 286-289.

Marcum, D. (2003). "Requirements for the future digital library." *Journal of Academic Librarianship*, 29(5), 276-279.

Okerson, A. (2003). "Asteroids, Moore's Law, and the Star Alliance." *Journal of Academic Librarianship*, 29(5), 280-285.

In 2003, Seamus Ross (University of Glasgow and EPRANET) conducted a review of the work of the Digital Library Transition Team at the National Library of New Zealand. His report contains recommendations on how the national library can take forward the development of a digital library service.

Ross, S. (2003). *National Library of New Zealand: digital library development review: final report*. Wellington, N.Z.: National Library of New Zealand (Te Puna Matauranga o Aotearoa), July 2003. 88 pp. ISBN 0-477-02797-0. Retrieved February 24, 2004, from:  
[http://www.natlib.govt.nz/files/ross\\_report.pdf](http://www.natlib.govt.nz/files/ross_report.pdf)



### 3. Other recent publications

Abrams, S. L., & Rosenblum, B. (2003). "XML for e-journal archiving." *OCLC Systems & Services*, 19(4), 155-161.

In a special journal issue on 'XML and e-journals,' Abrams and Rosenblum report on a study undertaken by Harvard University Library as part of the Andrew W. Mellon Foundation's E-journal Archiving Project, including the development of an XML DTD for the archival interchange of articles.

Bishoff, L., & Allen, N. (2004). *Business planning for cultural heritage organisations*. Washington, D.C.: Council on Library and Information Resources. Retrieved February 24, 2004, from: <http://www.clir.org/pubs/abstract/pub124abst.html>

Borghoff, U. M., Rödiger, P., Scheffczyk, J., & Schmitz, L. (2003). *Langzeitarchivierung: Methoden zur Erhaltung digitaler Dokumente*. Heidelberg: dpunkt.verlag. xv, 283 pp. ISBN 3-89864-245-3. More details from dpunkt.verlag Web site, retrieved February 24, 2004, from: <http://www.dpunkt.de/buch/3-89864-245-3.html>

The first part of the book consists of a methodological overview that introduces archiving models (e.g., the OAIS Reference Model, the NEDLIB project's DSEP model), emulation, migration, documentation, and some standard formats. The second part looks in more detail at selected initiatives and projects, concentrating on markup, metadata, and preservation strategies (migration and emulation). The authors are all based at the Institute for Software Technology of the University of the Federal Armed Forces, Munich.

Byers, F. R. (2003). *Care and handling of CDs and DVDs: a guide for librarians and archivists*. Washington, D.C.: Council on Library and Information Resources; National Institute of Standards and Technology. Retrieved February 24, 2004, from: <http://www.clir.org/pubs/abstract/pub121abst.html>

Cain, M. (2003). "Being a library of record in a digital age." *Journal of Academic Librarianship*, 29(6), 405-410.

Cornell University Library. (2003). *Digital preservation management: implementing short-term strategies for long-term problems* [online tutorial]. Ithaca, N.Y.: Cornell University Library, 2003. Retrieved February 24, 2004, from: <http://www.library.cornell.edu/iris/tutorial/dpm/index.html>

Depocas, A., Ippolito, J., & Jones, C., eds. (2003). *The variable media approach: permanence through change*. New York: Solomon R. Guggenheim Foundation; Montreal: Daniel Langlois Foundation for Art, Science, and Technology, 2003. 137 pp. ISBN 0-9684693-2-9. Retrieved February 24, 2004, from: <http://www.variablemedia.net/pdf/Permanence.pdf>

DFG-Arbeitsgruppe 'Informationsmanagement der Archive.' (2004). Die deutschen Archive in der Informationsgesellschaft - Standortbestimmung und Perspektiven." *Zeitschrift für Bibliothekswesen und Bibliographie*, 51(1), 17-27. Retrieved February 24, 2004, from: <http://zfbf.thulb.uni-jena.de/servlets/DerivateServlet/Derivate-131/j04-h1-ber-1.pdf>

A paper (in German) on preservation and the information society in Germany, compiled by a German Research Foundation (DFG) working group on 'Information Management of Archives,'

Digital Preservation Testbed. (2003). *Emulation: context and current status*. Den Haag: Digital Preservation Testbed White Paper, October 2003. Retrieved February 24, 2004, from: [http://www.digitaleduurzaamheid.nl/bibliotheek/docs/white\\_paper\\_emulatie\\_EN.pdf](http://www.digitaleduurzaamheid.nl/bibliotheek/docs/white_paper_emulatie_EN.pdf)

Jantz, R. (2003). "Public opinion polls and digital preservation: an application of the Fedora Digital Object Repository System." *D-Lib Magazine*, 9(11), November 2003. Retrieved February 24, 2004, from: <http://www.dlib.org/dlib/november03/jantz/11jantz.html>

Lazinger, S., Negin, B., & Berman, Y. (2002). "Preservation of electronic records in Israeli government offices." *Journal of Government Information*, 29(5), 319-331.



Mathieson, S.A. (2003). "Libraries embrace digital age." *The Guardian*, 28 January 2004. Retrieved February 24, 2004, from: <http://society.guardian.co.uk/e-public/story/0,13927,1132475,00.html>

Müller, E., Klosa, U., Andersson, S., & Hansson, P. (2003). "The DiVA project - development of an electronic publishing system." *D-Lib Magazine*, 9(11), November 2003. Retrieved February 24, 2004, from: <http://www.dlib.org/dlib/november03/muller/11muller.html>

Murray-Rust, P., & Rzepa, H. S. (2003). "XML for scientific publishing." *OCLC Systems & Services*, 19(4), 162-169.

An article that looks at the potential uses of XML in scientific publishing while noting that take-up so far has been relatively limited.

Pennavaria, K. (2003). "Nonprint media preservation: a guide to resources on the Web." *C&RL News*, 64(8), September 2003. Retrieved February 24, 2004, from: [www.ala.org/ala/acrl/acrlpubs/crlnews/backissues2003/september8/nonprintmedia.htm](http://www.ala.org/ala/acrl/acrlpubs/crlnews/backissues2003/september8/nonprintmedia.htm)

Renear, A., & Dubin, D. (2003). "Towards identity conditions for digital documents," DC-2003: 2003 Dublin Core Conference: Supporting Communities of Discourse and Practice - Metadata Research and Applications, Seattle, Wa., USA, 28 September - 2 October 2003. Retrieved February 24, 2004, from: [www.siderean.com/dc2003/503\\_Paper71.pdf](http://www.siderean.com/dc2003/503_Paper71.pdf)

In a paper presented at the 2003 Dublin Core Conference, Allen Renear and David Dubin describe recent work in determining identity conditions for digital documents represented in markup. They illustrate some of the problems encountered in comparing documents at too low a level of abstraction, e.g. in using comparisons of bitstream, character stream or canonical serializations of data structure as methods to determine identity between documents. They propose a solution through use of XML semantics to represent and compare documents for identity. The approach has implications for the comparison of documents preserved through migration.

Royal Netherlands Academy of Arts and Sciences, Social Sciences Council. (2003). *Networked data services: towards a future data infrastructure for the social sciences in the Netherlands*. Amsterdam: Royal Netherlands Academy of Arts and Sciences, September 2003. ISBN 90-6984-399-4. Retrieved February 24, 2004, from: <http://www.knaw.nl/publicaties/pdf/20031077.pdf>

This report by the Social Sciences Council of the Royal Netherlands Academy of Arts and Sciences (KNAW) calls for better access to social science data in the Netherlands. It proposes a new infrastructure based on a network comprising a national centre for data services (mandated by the Netherlands Organisation for Scientific Research (NWO) and KNAW) with separate 'topical expertise centres' based at research institutes.

Stiegler, B., Fingerhut, M., & Donin, N. (2003). "The IRCAM Digital Sound Archive in context." *DigiCULT.Info*, 6, 19-22, December 2003. Retrieved February 24, 2004, from: [http://www.digicult.info/downloads/dc\\_info\\_issue6\\_december\\_20031.pdf](http://www.digicult.info/downloads/dc_info_issue6_december_20031.pdf)

US National Coordination Office for Information Technology Research and Development. (2003). *Supplement to the President's Budget for FY2004*. Retrieved February 24, 2004, from: <http://www.itrd.gov/pubs/blue04/>

This 'Blue Book,' a supplement to the US President's FY2004 Budget produced by the National Coordination Office for Information Technology Research and Development (NITRD), gives an overview of future research priorities for the US federal government and its research agencies (NSF, DARPA, NASA, etc). Preservation and access are given prominence in the program component areas, "which focus on fundamental, long-term research in computing and networking technologies" (p.3).

Vitiello, G. (2004). "Identifiers and identification systems: an informational look at policies and roles from a library perspective." *D-Lib Magazine*, 10(1), January 2004. Retrieved February 24, 2004, from: <http://www.dlib.org/dlib/january04/vitiello/01vitiello.html>

## 4. Events

### 4.1 Recent events

#### Digital Objects Repository Management Forum, Sydney

Presentations and other material from the Digital Objects Repository Management Forum 'Information Infrastructure: Backing Australia's Ability' held in Sydney, Australia, 19-30 May 2003 are now available from the University of Sydney Web site. Retrieved February 24, 2004, from: <http://www.library.usyd.edu.au/dest/forum.html>

#### European Conference on Digital Libraries (ECDL 2003)

Some new conference reports of ECDL 2003, held in Trondheim, Norway on 17-22 August 2003, are now available:

Huxley, L. (2003). "ECDL2003: conference notes." *Ariadne*, 37, October 2003. Retrieved February 24, 2004, from: <http://www.ariadne.ac.uk/issue37/ecdl2003-rpt/>

Day, M. (2003). "Report on ECDL 2003." UKOLN, University of Bath, November 2003. Retrieved February 24, 2004, from: <http://www.ukoln.ac.uk/metadata/publications/events/ecdl2003/>

Day, M. (2003). "3rd ECDL Workshop on Web Archiving." *Ariadne*, 37, October 2003. Retrieved February 24, 2004, from: <http://www.ariadne.ac.uk/issue37/ecdl-web-archiving-rpt/>

#### The Future of Digital Memory and Cultural Heritage, Florence

Some of the presentations from the The Future of Digital Memory and Cultural Heritage held in Florence, Italy on the 16-17 October 2003 are now available on the conference Web site, retrieved February 24, 2004, from: [http://galileo.imss.firenze.it/memorie\\_digitali/](http://galileo.imss.firenze.it/memorie_digitali/)

At the conference, Alessandro Ruggiero presented six case studies of the preservation or rescue of digital data that had been at risk. These included the 1960 US Census, the Combat Air Activities File (both NARA), the Database of the *Consorzio Neapolis*, BBC Domesday, the "Kaderdatenspeicher" of the German Democratic Republic, and the Web site of the city of Antwerp. The case studies reveal the high cost of rescue strategies and some of the risks involved. The recovery of the Neapolis database (created in 1987-89 to combine map data with archaeological monuments, discoveries, etc. from Pompeii), for example, took two years to recover at a cost of around 200,000 euros. Also, it was dependent on the continued availability of certain hardware (a mainframe with tape drive) and expertise. The continued availability of hardware was also a significant factor in some of the other cases, e.g. for recovering the 1960 US Census (UNIVAC II-A tape units) and for emulating the BBC Domesday Project. The Firenze Agenda was also published in issue 6 of *DigiCULT.Info*.

Ruggiero, A., ed. (2003). *Conservazione delle memorie digitali: rischi ed emergenze: sei casi di studio = Preservation of digital memory: risks and emergencies: six case studies*. 'The Future of Digital Memory and Cultural Heritage,' Palazzo Vecchio, Florence, Italy, 16-17 October 2003 (48 pp.).

"The Firenze Agenda (17 October 2003)." Retrieved February 24, 2004, from: <http://www.minervaeurope.org/structure/nrg/documents/firenzeagenda031017draft.htm>

*Agenda di Firenze*. Retrieved February 24, 2004, from: [www.iccu.sbn.it/PDF/Firenze-agenda-17-Oct\\_ITAL.pdf](http://www.iccu.sbn.it/PDF/Firenze-agenda-17-Oct_ITAL.pdf)

## 6th International Conference on Asian Digital Libraries (ICADL 2003)

A brief report on the 6th International Conference on Asian Digital Libraries (ICADL 2003), held in Kuala Lumpur, Malaysia, on 8-11 December 2003, appeared in the January 2004 issue of D-Lib Magazine.

Cunningham, S. J. (2004). "Report on the 6th International Conference on Asian Digital Libraries (ICADL 2003), 8 - 11 December 2003, Kuala Lumpur, Malaysia", *D-Lib Magazine*, 10(1), January 2004. Retrieved February 24, 2004 from:  
<http://www.dlib.org/dlib/january04/cunningham/01cunningham.html>

## ERPANET events

Briefing papers and presentations are available for some recent ERPANET events:

- ERPANET workshop on the Long-term Preservation of Databases, Berne, Switzerland, 9-11 April 2003. The briefing paper (PDF), presentations, and final report (PDF) are available from the ERPANET Web site. Retrieved February 24, 2004, from: [www.erpanet.org/www/products/bern/bern.htm](http://www.erpanet.org/www/products/bern/bern.htm)
- ERPANET training seminar Metadata in Digital Preservation, Marburg, Germany, 3-5 September 2003. The briefing paper (PDF), presentations, final report, and other material are available from the ERPANET Web site. Retrieved February 24, 2004, from: [www.erpanet.org/www/products/marburg/marburg.htm](http://www.erpanet.org/www/products/marburg/marburg.htm)
- Accademia Nazionale dei Lincei and ERPANET workshop on Trusted Repositories for Preserving Cultural Heritage, Rome, Italy, 17-19 November 2003. The briefing paper (PDF) and selected presentations are available from the ERPANET Web site. Retrieved February 24, 2004, from: [www.erpanet.org/www/products/rome/rome.htm](http://www.erpanet.org/www/products/rome/rome.htm)
- ERPANET and CODATA workshop on the Selection, Appraisal, and Retention of Digital Scientific Data, Lisbon, Portugal, 15-17 December 2003. The briefing paper (PDF) and selected presentations are available from the ERPANET Web site. Retrieved February 24, 2004, from: [www.erpanet.org/www/products/lisbon/lisbon.htm](http://www.erpanet.org/www/products/lisbon/lisbon.htm)

## RLG Members' Forum on Metadata and Institutional Repositories

Presentations from the RLG Members' Forum 'To Have and to Hold: Metadata and Institutional Repositories' (held at Washington, D.C. and Chicago, Ill. in December 2003) are available from the RLG Web site. Retrieved February 24, 2004, from:  
<http://www.rlg.org/events/haveandhold2003/index.html>

## 4.2 Forthcoming Events

*March:*

Preservation and Access for Electronic College and University Records (ECURE 2004), Tempe, Arizona, USA, 1-3 March 2004. Retrieved February 24, 2004, from: [www.asu.edu/it/events/ecure/](http://www.asu.edu/it/events/ecure/)

10th Global Grid Forum, Berlin, Germany, 9-13 March 2004. Retrieved February 24, 2004, from: <http://www.gridforum.org/>

Preserving the Knowledge Base of Science, ALPSP (Association of Learned and Professional Society Publishers) Seminar, London, 17 March 2004. Retrieved February 24, 2004, from: [www.alpsp.org/events/s170304.htm](http://www.alpsp.org/events/s170304.htm)

2004 International Conference on Digital Archive Technologies (ICDAT 2004), Taipei, Republic of China (Taiwan), 18-19 March 2004. Retrieved February 24, 2004, from: [www.iis.sinica.edu.tw/ICDAT04/](http://www.iis.sinica.edu.tw/ICDAT04/)

Museums and the Web 2004, Arlington, Virginia, USA, 31 March - 3 April 2004. Retrieved February 24, 2004, from: <http://www.archimuse.com/mw2004/>

*April:*

Long-Term Stewardship of Globally Distributed Storage: 12th NASA Goddard, 21st IEEE Conference on Mass Storage Systems and Technologies (MSST04), College Park, Maryland, USA, 13-16 April 2004. Retrieved February 24, 2004, from: <http://storageconference.org/2004/>

Audit and Certification in Digital Preservation, ERPANET Workshop, Antwerp, Belgium, 14-16 April 2004. Retrieved February 24, 2004, from: <http://www.erpanet.org/>

Archival Perspectives in Digital Preservation, Society of American Archivists professional education offering (Instructor: Paul Conway, Duke University Libraries), New York, USA, 15-16 April 2004. Retrieved February 24, 2004, from: [www.archivists.org/prof-education/workshop-detail.asp?id=932](http://www.archivists.org/prof-education/workshop-detail.asp?id=932)

IS&T Archiving Conference, San Antonio, Texas, USA, 20-23 April 2004. Retrieved February 24, 2004, from: <http://www.imaging.org/conferences/archiving2004/>

*May:*

File Formats for Preservation, ERPANET Training Seminar, Vienna, Austria, 10-11 May 2004. Retrieved February 24, 2004, from: <http://www.erpanet.org/>

13th International World Wide Web Conference, New York, USA, 17-22 May 2004. Retrieved February 24, 2004, from: [www.2004.org/](http://www.2004.org/)

Data Futures: Building on 30 Years of Advocacy, International Association for Social Science Information Service and Technology Annual Conference 2004 (IASSIST 2004), Madison, Wisconsin, USA, 25-28 May 2004. Retrieved February 24, 2004, from: <http://dpls.dacc.wisc.edu/iassist2004/>

Accountability and Ethics in the Archival Sphere, Association of Canadian Archivists 2004 Annual Conference, Montréal, Québec, 17-29 May 2004. Retrieved February 24, 2004, from: <http://archivists.ca/conference/>

*June:*

Distributing Knowledge Worldwide Through Better Scholarly Communication, 7th International Symposium on Electronic Theses and Dissertations (ETD 2004), Lexington, Kentucky, USA, 3-5 June 2004. Retrieved February 24, 2004, from: <http://www.uky.edu/ETD/ETD2004/>

13th IEEE International Symposium on High-Performance Distributed Computing (HPDC-13), Honolulu, Hawaii, USA, 4-6 June 2004. Retrieved February 24, 2004, from: <http://hpdc13.cs.ucsb.edu/>

Global Reach and Diverse Impact, Joint Conference on Digital Libraries (JDCL 2004), Tuscon, Arizona, USA, 7-11 June 2004. Retrieved February 24, 2004, from: <http://www.jcdl2004.org/>

Computing and Multilingual, Multicultural Heritage, Joint International Conference of the Association for Literary and Linguistic Computing and the Association for Computers and the Humanities (ALLC/ACH 2004), Gothenburg, Sweden, 11-16 June 2004. Retrieved February 24, 2004, from: <http://www.hum.gu.se/allcach2004/>

Building Digital Bridges: Linking Cultures, Commerce and Science, 8th Conference on Electronic Publishing (EIPub 2004), Brasilia, Brazil, 23-26 June 2004. Retrieved February 24, 2004, from: <http://www.elpub.net/>

Preserving the AudioVisual Heritage - Transition and Access, Joint Technical Symposium (JTS 2004), Toronto, Canada, 24-26 June 2004. Retrieved February 24, 2004, from: <http://www.jts2004.org/>

*July:*

AusWeb04: The Tenth Australian World Wide Web Conference, Gold Coast, Queensland, Australia, 3-7 July 2004. Retrieved February 24, 2004, from: <http://ausweb.scu.edu.au/aw04/conf/>

Joint Information Systems Committee/Coalition for Networked Information Meeting 2004, Brighton, UK, 8-9 July 2004. Retrieved February 24, 2004, from: <http://www.ukoln.ac.uk/events/jisc-cni-2004/>

IADIS e-Society 2004 Conference, Avila, Spain, 16-19 July 2004. Retrieved February 24, 2004, from: <http://www.iadis.org/es2004/>

*August:*

Society of American Archivists Annual Meeting, Boston, Massachusetts, USA, 2-8 August 2004. Retrieved February 24, 2004, from: <http://www.archivists.org/conference/index.asp>

Libraries: Tools for Education and Development, World Library and Information Congress: 70th IFLA General Conference and Council, Buenos Aires, Argentina, 22-27 August 2004. Retrieved February 24, 2004, from: <http://www.ifla.org/IV/ifla70/>

Archives, Memory and Knowledge, International Congress on Archives 2004, Vienna, Austria, 23-29 August 2004. Retrieved February 24, 2004, from: <http://www.wien2004.ica.org/fo/index.php>

30th International Conference on Very Large Data Bases (VLDB 2004), Toronto, Canada, 30 August - 3 September 2004. Retrieved February 24, 2004, from: [www.vldb04.org/](http://www.vldb04.org/)

*September:*

Digital Resources for the Humanities (DRH-2004), Newcastle-upon-Tyne, UK, 5-8 September 2004. Retrieved February 24, 2004, from: [www.ncl.ac.uk/niassh/drh2004/](http://www.ncl.ac.uk/niassh/drh2004/)

European Conference on Digital Libraries (ECDL 2004), Bath, UK, 12-17 September 2004. Retrieved February 24, 2004, from: <http://www.ecdl2004.org/>

*October:*

Ensuring the Long-Term Preservation and Adding Value to the Scientific and Technical Data, Frascati, Italy, 5-7 October 2004. Retrieved February 24, 2004, from: [www.congrex.nl/04a08/](http://www.congrex.nl/04a08/)

9th International Symposium on Information Science = 9. Internationales Symposium für Informationswissenschaft (ISI 2004), Chur, Switzerland, 6-8 October 2004. Retrieved February 24, 2004, from: <http://www.isi2004.ch/>

Metadata Across Languages and Cultures, International Conference on Dublin Core and Metadata Applications 2004 (DC-2004), Shanghai, People's Republic of China, 11-14 October 2004. Retrieved February 24, 2004, from: <http://dc2004.library.sh.cn/english/>

*November:*

The Information Society: New Horizons for Science, 19th CODATA International Conference, Berlin, Germany, 7-10 November 2004. Retrieved February 24, 2004, from: <http://www.codata.org/04conf/>

3rd International Semantic Web Conference (ISWC2004), Hiroshima, Japan, 7-11 November 2004. Retrieved February 24, 2004, from: <http://iswc2004.semanticweb.org/>

Archiving Web Resources: Issues for Cultural Heritage Institutions, National Library of Australia, Canberra, Australia, 9-11 November 2004. Retrieved February 24, 2004, from: <http://www.nla.gov.au/webarchiving/>

Great Technology for Collections, Confluence, and Community, 2004 Museum Computer Network Conference, Minneapolis, Minn., 10-13 November 2004. Retrieved February 24, 2004, from: <http://www.mcn.edu/>

Managing and Enhancing Information: Cultures and Conflicts, American Society for Information Science and Technology (ASIS&T) Annual Meeting, Providence, Rhode Island, USA, 13-18 November 2004. Retrieved February 24, 2004, from: <http://www.asis.org/Conferences/AM04/>

*December:*

Digital Library? International Collaboration and Cross-Fertilization, 7th International Conference of Asian Digital Libraries (ICADL 2004), Shanghai, People's Republic of China, 13-17 December 2004. Retrieved February 24, 2004, from: <http://icadl2004.sjtu.edu.cn/>

A comprehensive and frequently updated list of forthcoming events is available from the PADI Web site: <http://www.nla.gov.au/padi/format/event.html>

---

Problem links last disabled or updated: 17 September 2009

Warning! Web site links tend to have very short lifetimes, as documents are frequently updated or deleted, Web sites are restructured, domains are renamed or moved, etc. The compilers of this bulletin, therefore, cannot guarantee that all of the URLs in this document will successfully resolve to the resources described here. However, in these cases, try searching for the same resource on the PADI gateway (<http://www.nla.gov.au/padi/>), which will provide updated URLs wherever possible.