This is a summary of selected recent activity in the field of digital preservation compiled from the Preserving Access to Digital Information (PADI) Gateway and the digital-preservation and padiforum-l mailing lists. Additional or related items of interest may also be included.

Contents:

1. News from organisations and initiatives
   1.1 U.K. Digital Preservation Coalition (DPC)
   1.2 Digital Curation Centre
   1.3 Museums, Libraries and Archives Council (UK)
   1.4 Library of Congress (US)
   1.5 National Archives and Records Administration (US)
   1.6 Long-lived Data Collections Task Force (US)
   1.7 Center for International Earth Science Information Network (US)
   1.8 National Consultation on Access to Scientific Research Data (Canada)
   1.9 The Daniel Langlois Foundation for Art, Science and Technology (Canada)

2. Specific themes
   2.1 Digital Preservation Research and Directions
   2.2 Digital Preservation Readiness
   2.3 Digital Repositories
1. News from Organisations and Initiatives

1.1 U.K. Digital Preservation Coalition (DPC)

The Digital Preservation Coalition (DPC) has released the 2005 Digital Preservation Award Shortlist and has added the finalists' presentations to their web site. The five finalists are:

- Vienna University of Technology - Choosing the optimal digital preservation strategy
- National Archives of the Netherlands - Digital Preservation Test-bed
- PREMIS Working Group - PREMIS (Preservation Metadata: Implementation Strategies)
- British Broadcasting Corporation - Reverse Standards Conversion
- UK Web Archiving Consortium - The Consortium

The winner will be announced at the 2005 Conservation Awards event held at the British Museum on 22nd November.

The DPC has also launched the UK Digital Preservation Needs Assessment survey, its biggest project to date. This 'state of the nation' survey is designed to reveal the extent of the risk of loss or degradation to digital material held in the UK's public and private sectors.

It will analyse and synthesise existing sources of data on digital preservation activity in the UK and gather additional information to present a detailed analysis of the national status quo regarding digital...

Press Release: *State of the nation survey to reveal threat to the UK's digital heritage.*

---

### 1.2 Digital Curation Centre

The Digital Curation Centre (DCC) has recently published the *Digital Curation Manual.* The Manual will comprise various instalments on topics related to digital curation. Each instalment is written and peer reviewed by experts in that particular field and is regularly revised to encompass new developments.

The first instalment, *Open Source for Digital Curation,* (see also 2.9) was published in July and the abstracts of a further 45 identified instalments are available, covering topics such as appraisal and selection, business models, user requirements, rights management, preservation strategies, workflows and metadata.


Instalments and abstracts available from: http://www.dcc.ac.uk/resource/curation-manual/chapters/

---

### 1.3 Museums, Libraries and Archives Council (UK)

The Museums, Libraries and Archives Council (MLA) in the UK, in conjunction with the Digital Preservation Coalition, has published a report on a survey and assessment carried out from December 2004 to March 2005 regarding digital preservation preparedness and needs in regional organisations. The results supported general assumptions that long term management of digital material needs to be more firmly and clearly embedded in corporate planning processes, strategic thinking and planning, and that there is a general need for awareness raising on the whole issue of digital preservation. The organisations surveyed expressed a need for help on a range of issues to do with digital preservation such as training, procurement, support networks, and resources. A list of actions for all the stakeholders from a local to
national level was developed based on the survey results. The results of this survey will also be incorporated into the UK Needs Assessment (see 1.1).


### 1.4 Library of Congress (US)

In August, the Library of Congress launched the Section 108 Working Group. The purpose of this group is to recommend changes to copyright law that recognise the need for exceptions to the law for libraries and archives in the digital age. This is especially important for the collection, preservation and accessibility of born digital material. The website has links to the relevant section of the Act as well as background papers and will post any findings and recommendations on how to revise the copyright law in order to ensure an appropriate balance among the interests of creators and other copyright holders, libraries and archives.


### 1.5 National Archives and Records Administration (US)

In September, the National Archives and Records Administration (NARA) announced that Lockheed Martin has been awarded a $US305 million, six-year contract to build the Electronic Records Archives (ERA) system. The purpose of ERA will be to preserve the digital output of the United States federal government. It is intended that ERA will preserve all types of electronic records, free from dependence on any specific hardware or software.


### 1.6 Long-lived Data Collections Task Force (US)
The National Science Board of the US National Science Foundation has formed the Long-lived Data Collections Task Force and undertaken an analysis of the policy issues relevant to long-lived digital data collections in research and education for science and engineering. Digital data collections are now recognised as unique and high-value resources for the Foundation that need to be addressed collectively. A pre-publication of the resulting report provides the findings and recommendations arising from that analysis. The recommendations presented after the survey analysis call for clarifying and harmonizing NSF strategy, policies, processes, and budget for long-lived digital data collections. The report is considered to be the beginning of much broader consultation on these issues.


1.7 Center for International Earth Science Information Network (US)

In June, the Center for International Earth Science Information Network (CIESIN) published the Guide to Managing Geospatial Electronic Records. This guide is part of the Managing and Preserving Geospatial Electronic Records project conducted by the Center for International Earth Science Information Network (CIESIN) at Columbia University. The guide offers recommendations for managing and preserving geospatial data and related electronic records. Areas covered include infrastructure issues such as policies and planning; data accession issues such as selection, authenticity and intellectual property rights; and custodianship issues such as metadata, media migration and preservation masters.


1.8 National Consultation on Access to Scientific Research Data (Canada)

In June, The National Research Council Canada (NRC) released the Final Report of the National Consultation on Access to Scientific Research Data (NCASRD). The Report includes many recommendations including the creation of a task force, dubbed Data Force, to prepare a full national implementation strategy, and the instigation of a pilot project to show the value and impact of access to research data. The Report also proposes the establishment of a dedicated national infrastructure to
assume overall leadership in the development and execution of a strategic plan which includes action to stop degradation and loss of the country's research heritage. Discussion of the issues, recommendations and proposed timetable for action are given.


1.9 The Daniel Langlois Foundation for Art, Science and Technology (Canada)

The Daniel Langlois Foundation for Art, Science, and Technology is commencing a five year interdisciplinary research alliance project that will focus on important issues regarding the preservation and documentation of media art. Entitled Documentation and Conservation of the Media Arts Heritage: Research and Case Studies, the research will focus on three areas of expertise - technological history, art documentation and art conservation. The project will also focus on three main research fields - historical studies and development of descriptive and analytic tools; documentation, dissemination and access via research networks; and preservation of works featuring technological components. Tools, guides and methodologies such as case studies, best practice guides and specifications will be produced as part of this project.


2. Specific themes

2.1 Digital preservation research and directions

The staff of RLG DigiNews have reviewed digital preservation and digitisation research and development projects around the world and have compiled a list of 10 initiatives in progress which have the potential to contribute to the broader cultural heritage community. The ten projects were

- The National Digital Information Infrastructure and Preservation Program (NDIIPP) - ECHO
2.2 Digital preservation readiness

Anne Kenney and Elle Buckley have reported the results of a survey of participants in Cornell University Library's digital preservation management training program on the readiness of their institutions to develop and maintain an ongoing digital preservation program. The survey was conducted from 2003-2005. The results are also compared to a previous survey conducted in 1998 in order to identify shifts in emphasis, concerns and issues in digital preservation readiness.


2.3 Digital repositories

There have been a number of developments in the area of digital repositories which are grouped below
into those addressing the creation of digital repositories, those discussing the issues around trusted digital repositories and those on digital sustainability including repositories.

### 2.3.1 Creation of digital repositories

DSpace has published a document *Creating an Institutional Repository: LEADIRS Workbook*. Intended for managers and repository developers, this extensive document covers all aspects of building institutional repositories such as planning, choosing repository software platforms, legal and regulatory environmental and policy development and cost modelling. Each section has associated worksheets and key questions. There are many references to case studies which are used to highlight approaches to the various issues in developing an institutional repository.


reUSE is a cooperative project of libraries and universities from Austria, Estonia, Germany and Slovenia funded through the European Commission eContent Programme. The reUSE partners will focus on the publications of public sector institutions such as government, universities and research institutions. The purpose of the reUSE project is to ensure that digital master files are preserved and managed as well as the corresponding hardcopy versions. Part of the project includes four demonstrator services which are designed to operate permanently after the project itself is completed.

The project has also published a whitepaper, which covers the project's partners and stakeholders, interoperability and digital preservation, reUSE demonstrators, and voluntary deposit agreements.

Universität Innsbruck. *reUSE*.
Retrieved 27 October 2005 from: http://www2.uibk.ac.at/reuse/


Formerly known as e-Archive, the Portico initiative was launched in 2005. Portico is a not-for-profit archive of born digital scholarly E-journals maintained by Ithaka Harbors Inc. One of the main goals of Portico is to ensure that electronic resources remain accessible in the future and it aims to do this by developing an economic model as the basis of long term sustainability; developing relationships with publishers who will deposit their publications in the archive; and developing a production and technological infrastructure that will support preservation and accessibility.

Ithaka Harbors Inc. *Portico*. 
2.3.2 Compliance and certification of trusted digital repositories

A draft paper for public comment, *Audit checklist for the certification of digital repositories*, was compiled by the RLG-NARA Task Force on Digital Repository Certification. The paper identifies the audit and certification criteria repositories must meet in order to reliably store, migrate and provide access to digital collections. The criteria include organisational aspects, such as sustainability; repository functions and processes such as ingest, storage and management; usability, such as descriptive metadata; and technologies and technical infrastructure. The audit instrument is also included in the document. The period for comment on the draft closes on January 15, 2006.


The National Archives (UK) and UK Data Archive have released a report comparing their preservation practices to the OAIS reference model and contemplating use of the METS standard. This work also considers the recommendations of the RLG/OCLC work on attributes of trusted digital repositories but predates the checklist above.

The project aimed to determine whether each organisation is compliant with the OAIS model and to present the results in order to provide a model for other organisations to test the compliance of their own systems. To this end Appendix 5 contains a useful list of questions to assist other archives assessing their compliance to the OAIS model. The results of their own comparison showed that they both not only comply to the OAIS model but exceed the stated requirements.


Ronald Jantz and Michael J. Giarlo from Rutgers University Libraries have written an article discussing what it means to be a trusted digital repository and how Rutgers believes it is meeting the challenge. They discuss a range of technologies available today that can be utilised to support trust in a digital preservation process and how these technologies can be integrated into an operational digital preservation framework.

2.3.3 Sustainability of digital repositories

Published online in June 2005, the Australian Project for Sustainable Repositories (APSR) released its Sustainability Issues Discussion Paper. Written by Kevin Bradley, this comprehensive paper examines various aspects of digital sustainability. These include economic aspects, such as costing, funding and business models; technological aspects, such as repositories, formats, bit stream and persistence; as well as risk management. Also included is a summary of specific issues and their associated tasks.


2.4 E-prints

In June the Research Councils UK (RCUK) released a position statement on the dissemination of and access to research outputs in the form of journal articles or conference proceedings. The statement describes their support for a cooperative effort to preserve and make available digital publications in the UK and encourages researchers to deposit in e-print repositories. One of the fundamental principles recognises the importance and value of digital research collections for the entire community: "The outputs from current and future research must be preserved and remain accessible not only for the next few years but for future generations."

RCUK Position Statement on Access to Research Outputs.

Thoughtful and enlightening discussion on the above position statement and open archive repositories can be found in the Open Access Newsletter archive.

SPARC Open Access Newsletter, issue 87 July 2, 2005.
Retrieved 27 October 2005 from: http://www.earlham.edu/~peters/fos/newsletter/07-02-05.htm

SPARC Open Access Newsletter, issue 84 April 2, 2005.
Retrieved 27 October 2005 from: http://www.earlham.edu/~peters/fos/newsletter/04-02-05.htm#oara
2.5 Web archiving

The International Internet Preservation Consortium (IIPC) has released a set of tools developed for Web archiving which are available for download. The tools include acquisition software Heritrix and DeepArc and access and finding aids NutchWAX (Nutch + Web Archive eXtensions), WERA (Web ARchive Access) and Xinq (XML INQuire).


Two papers, both by staff of the National Library of Australia, deal with acquisition costs and current issues in PANDORA Web archiving activities. Paul Koerbin discusses the curatorial, technical and legal issues currently problematic for Web archiving activities by PANDORA at the National Library of Australia. Margaret Phillips discusses the costs of maintaining PANDORA. She finds that acquiring Web publications is considerably more expensive than printed publications and describes ways in which the costs could be reduced.


2.6 Preservation Metadata

In September, the DPC published a new technology watch report on preservation metadata. The report, authored by Brian Lavoie and Richard Gartner, provides a comprehensive view of developments in preservation metadata, including METS, OAIS and PREMIS. This is recommended reading for both the beginner to preservation metadata and more experienced practitioners.

2.7 Standards

The International Organization for Standardization has recently released a new standard for PDF documents. The standard, ISO 19005, *Document management - Electronic document file format for long-term preservation - Part 1, Use of PDF 1.4 (PDF/A-1)*, will ensure that the PDF/A is an internationally recognized digital preservation standard.

Press release: *New ISO standard will ensure long life for PDF documents.* (7 October 2005)
Retrieved 27 October 2005 from: updated 24 February 2006:

2.8 Open Source

The first issue of the DCC Digital Curation Manual, published in July, is called *Open Source for Digital Curation*. It provides a practical and in-depth analysis of all aspects of open source within the context of digital curation. The article covers the background and development of open source software, the application of open source to digital preservation, including costs, legal issues, functionality and security. It also discusses the various applications in governments, science and humanities institutions.


2.9 Personal Archiving

Now that millions of people are creating their own personal collections of digital materials such as email messages, digital photographs and music collections there is growing interest in how to keep these materials alive.

An interesting article on hard-disk life and personal archiving by John Naughton appeared in the UK newspaper, *The Observer*. It covers differences between industry measures of performance (Mean Time Between Failure) and service life, and implications of changing patterns of personal digital storage and archiving.

Retrieved 27 October 2005 from:
D-Lib Magazine has also published an article by Neil Beagrie discussing current research and emerging services in the area of personal digital libraries and collections and discusses potential implications (including preservation) for individuals, libraries and their institutions.

Retrieved 27 October 2005 from: http://www.dlib.org/dlib/june05/beagrie/06beagrie.html

2.10 Storage Media

In a FAQ for RLG DigiNews, Dr. Vivek Navale from NARA presents a helpful overview of three studies that have looked at the life expectancy of high density magnetic tape, CD-ROMs and recordable CDs and DVDs. The overview includes a summary of work undertaken by NARA on digital tapes, by the Library of Congress on CD-ROMs, and a study by NIST on CD-Rs and DVD-Rs

Retrieved 27 October 2005 from:

2.11 Digital Preservation Training

The Cornell Digital Preservation Management Workshops are continuing in 2006. The program consists of an online tutorial and a series of one-week workshops. The workshops were scheduled for revisions in 2005.


Cornell are also partnering with a UK consortium, led by ULCC, and including the DPC, who are developing an intensive training programme for the UK, funded by JISC as part of its Digital Preservation and asset Management programme (JISC 4/04). Like the Cornell workshop, the programme will last for one week and will limit delegates to 24 to ensure maximum participation. The programme will pilot in October 2005 and will then be offered during 2006. Further details available from:
3. Other publications

All you ever wanted to know about the development and use of checksums is well presented by Richard Entlich from Cornell University in this FAQ for the June 15 issue *RLG DigiNews*.


Blogs continue to be a popular medium for information exchange in the world of digital preservation and archiving. Two new blogs featuring information on aspects of digital preservation have begun recently.

Media Matters Inc. has launched the *DAVA : Digital Audiovisual Archiving blog*. This Weblog is updated weekly and will feature news and announcements on topics related to the digital transformation and preservation of audiovisual materials. The blog will also list upcoming conferences and workshops pertinent to archives and digital libraries, as well as links to metadata resource sites and other digital archiving resources.

*DAVA, Digital Audiovisual Archiving blog.*

*Digitalkoans* is another blog that has recently been launched. This Weblog, created and maintained by Charles W. Bailey, provides commentary on the broader issues of scholarly electronic publishing and digital culture.

*Digitalkoans.*

4. Events

4.1 Recent events
Joint Conference on Digital Libraries 2005 : Digital Libraries Cyberinfrastructure for Research and Education

The conference was held 7-11 June 2005, in Denver, Colorado, USA. A report of the conference was published in D-Lib Magazine, July/August 2005, Vol. 11 no. 7/8.

Retrieved 27 October 2005 from: http://www.dlib.org/dlib/july05/sumner/07sumner.html

DCC workshop on persistent identifiers

The DCC workshop on persistent identifiers was held 30 June to 1 July 2005 at the University of Glasgow, Glasgow (UK). Presentations and MP3 recordings are available from the DCC website.

Retrieved 27 October 2005, from: http://www.dcc.ac.uk/training/pi-2005/index

DCC/DPC Workshop on Cost Models for preserving digital assets

The report for the DCC/DPC Workshop on Cost Models for preserving digital assets held at the British Library Conference Centre on Tuesday 26th July 2005 is available from the DPC website.


The International Conference on the Preservation of Digital Objects (iPRES)

This conference was held 15 and 16th September in Goettingen, Germany. Focus areas of the conference were preservation policies, technical workflows and web archiving as well as recenty developments. Links to abstracts, slides and video streams of the presentations are now available from the website.

Retrieved 27 October 2005, from: http://rdd.sub.uni-goettingen.de/conferences/ipres/home

ECDL 2005

The 9th European Conference on Research and Advanced Technology for Digital Libraries was held September 18-23, 2005 in Vienna, Austria. The proceedings will be published by Springer in the series of Lecture Notes in Computer Science (LNCS), in hard copy and electronic form.


5th International Web Archiving Workshop (IWAW05)
The workshop was held over 2 days this year, September 22 and 23 2005, in Vienna, Austria in conjunction with the 2005 ECDL conference. The workshop provided a cross domain overview on active research and practice in all domains concerned with the acquisition, maintenance and preservation of digital objects for long-term access, with a particular focus on Web archiving and studies on effective usage of this type of archives. Papers and presentations are now available from the official web site.

Retrieved 27 October 2005, from: http://www.iwaw.net/05/

Refresh! the 1st International Conference on the Histories of Media Art, Science and Technology

This conference was held 29 September - 1 October 2005 in Banff, Canada. The sessions on 30 September were of particular interest, covering active projects and methods for the long-term preservation of new media art. A prominent theme was documentation of new media art to ensure its appropriate and accurate preservation. Various documentation concepts were discussed including the use of: thesauri, keywords, video, object-oriented vs. relational databases, ontologies (including hierarchical), scope notes, 'free-tagging' (i.e. user annotations), documentation of works in their context, source code, documentation layers including subject (i.e.technical, user, image description) and the implementation of the Variable Media Art Questionnaire. Project and initiatives of note include:


4.2 Current and forthcoming events

2005

October

Joint Digital Curation Centre (DCC) and ERPANET Workshop on Long-term Curation of Medical Databases, 13-14 October 2005, Gulbenkian Institute, Lisbon, Portugal. Retrieved 27 October 2005, from: http://www.dcc.ac.uk/mdbworkshop.html


November


Updated 10 January 2006 Link disabled. Document no longer available.


December


Updated 10 January 2006 Link disabled. Further information at http://www.nottingham.ac.uk/ecriuk/


2006

February


March


A comprehensive and frequently updated list of forthcoming events is available from the PADI Web site:
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.