digitalpreservationeurope

# DigitalPreservationEurope newsletter

ISSN 1752-9220 (Print)
ISSN 1752-9239 (Online)

Issue 2 A Newsletter on Digital Preservation Europe coordination action September 2007

## CONTENTS

**DigitalPreservationEurope newsletter**

ISSN 1752-9220 (Print)
ISSN 1752-9239 (Online)

September 2007 – Issue 2



**Director:**

HATII Seamus Ross s.ross@hatii.arts.gla.ac.uk

**DPE Newsletter content and managing editor:**

MIBAC Maurizio Messina messina@marciana.venezia.sbn.it

**Editor secretary:**

Giuliana Sgambati gi.sgambati@iccu.sbn.it

Susanna Spezia spezia@librari.beniculturali.it

Marialisa Cornacchia marialisacornacchia@yahoo.it

**Copyediting:**

Emily Witham e.witham@hatii.arts.gla.ac.uk

**Web developer:**

Brian Aitken b.aitken@hatii.arts.gla.ac.uk

# I. DIGITAL PRESERVATION CHALLENGE

Electronic resources are a central part of our cultural and intellectual heritage, but this material is at risk. Digital memory needs constant management, using new techniques and processes, to contain risks such as technological obsolescence. The DPE digital preservation challenge aims to raise awareness amongst researchers of the issue of digital preservation. Three competitions will be held over the lifetime of the DPE project, each including several tasks to solve. The first challenge invited participants to overcome the barriers hindering access to (sets of) digital objects. Each object was accompanied by a scenario based on a real-life situation. These scenarios were intended to make the challenge more accessible to participants from all backgrounds while not trivialising the serious nature of the digital preservation challenges facing society.

## *Who was Eligible to Enter?*

The first Digital Preservation Challenge open to individuals from any country or discipline has now closed. Post-graduate and undergraduate students in computing science and engineering tackled the Challenge and made submissions. This first Digital Preservation Challenge was not open to professional research teams or labs, however future competitions will be planned for these communities.

## *Evaluating submissions*

Submissions to the first Digital Preservation Challenge have been assessed by a panel of international digital preservation experts and practitioners. The incremental scoring method the panel uses places emphasis on the thoroughness and quality of the documentation of the processes used to render the objects rather than the overall outcome itself. In this respect, it may be possible for an individual to win the challenge even if he or she cannot ultimately render all of the objects.

© gábor suhajda

## *The six Scenarios*

### File Format Identification

### Scenario 1

You have met someone online and instant messaging has been your only form of contact. You agree to meet and your friend sends you a sound file using instant messaging to provide you with the details of where and when this meeting will occur. Unfortunately your computer crashes before the file transfer is complete and most of your hard drive is

destroyed. You find the sound file that was being transferred in a temporary folder but all of the contact information you had for your friend is lost. The only way you can contact your friend again is to recover this file and meet them in person as agreed - can you do it?

### Proprietary File Format
### Scenario 2

A friend contacts you asking for your help to access two files which she has been given by her PhD supervisor. She urgently needs to access these files so that she can complete her dissertation. She has been unable to open the files as they appear to have been created using a proprietary format that is not recognised by her operating system. She tried to contact her supervisor for additional information but unfortunately her supervisor is attending a conference and is not contactable for a week. Your friend needs these files before then to meet her deadline - can you help her?

### Client Server Database Application
### Scenario 3

You work for a small company that manufactures grand pianos. The company is currently facing financial difficulties due to low sales. A customer, named Mr. Miller, calls you asking for a quote for a large order. As a previous customer of your company, Mr. Miller assumes that you already have all his contact details on record, and therefore he hangs up without providing any. When you check the company's records however, you cannot find these details. He bought his piano during the month that your company trialled a new database application for managing customer relations. However, you still have a file from this trial and an accompanying note stating that the application has to be 'set up' and contains the 'current database'. You must recover Mr. Miller's contact details in order to make this important sale - can you obtain them?

### Legacy Emulator
### Scenario 4

You are a first year Computer Science student. For one assignment, you are set the task of recreating a computer game from your childhood for use on mobile platforms. You must incorporate as many of the game's original features as possible within your project including audio samples. Although you cannot remember the exact name of the particular game you have chosen to recreate, you have found a file relating to it on your hard drive. Can you find the application to run the game and retrieve the sound file?

### Legacy File Preservation Strategy
### Scenario 5

Your company is going through a certification process and requires access to business reports for the last ten years. You have discovered that many of these legacy files are partially incompatible with the company's modern applications and that there are many other files like these in the company's archive. You have been provided with representative samples of the files in question and have eventually been able to recover all of the information from the legacy files. However, it has taken a lot of time and effort to access these files and your boss is keen to avoid this kind of scenario in the future. She has asked you to design an appropriate preservation strategy that will facilitate access to such records. How will you do it?

### Digital Preservation of Multimedia Art
### Scenario 6

A large international art network has provided an interdisciplinary platform for developing digital art works since 1984. The rapid pace of change in software tools and frameworks used for multimedia authoring has meant that its archived artworks are in danger of becoming inaccessible and unusable. You have been asked to preserve two of these historical digital artworks for future generations and to develop appropriate digital preservation strategies. How will you achieve this?

*Awards*

First Prize 3000 Euros

Second Prize 1500 Euros

Third Prize 500 Euros

*Submission Deadlines*

The first DPE Preservation Challenge closed on the 15th of July 2007, evaluation of submissions is underway and the winners will be announced this autumn during this year's ECDL conference in Budapest on 16-21 September. The submission deadlines for the next challenges are as follows:

2nd Challenge: submission deadline March 15th 2008

3rd Challenge: submission deadline January 15th 2009

For further information and online submission, see: http://www.digitalpreservationeurope.eu/challenge

If you have any questions regarding the DPE Preservation Challenge please contact:

Stephan Strodl: strodl@ifs.tuwien.ac.at or Robert Neumayer: neumayer@ifs.tuwien.ac.at

# II. DIGITAL PRESERVATION EXCHANGE PROGRAM

Research and practice in digital preservation is patchy, fragmented, and disconnected. Communication between research groups is limited and does not always engage with the needs of practitioners. The exchange of professional practitioners and researchers provides a valuable way to understand and overcome these barriers. They can facilitate knowledge exchange, capacity building, and innovation. DigitalPreservationEurope recognises the value of exchange programmes as a mechanism to establish cross-institutional synergies, and offers exchanges to allow experts to visit a host institution of their choice for up to three months.

It is our hope that these DPE Exchanges will propagate knowledge, capacity and innovation as well as foster better cooperation among research institutions and industrial partners tackling the pressing challenges in digital preservation. DPEX aims to encourage innovative practice through research collaboration and to build bridges between practitioners and researchers.

### *What Funding Is Provided?*

The duration of the exchange can vary between two weeks and up to three months depending on the proposed activities. The total amount per requested grant should not exceed €3500, which can be used to partially meet the costs of accommodation, subsistence and travel. Unfortunately, DPEX support cannot be used to meet salary costs.

### *How to apply?*

Applications for the DPE Research and Practitioner Exchange Programme are made via the DPE website. All applications are assessed by an independent review panel on the basis of the applicant's potential, the scientific merit of the proposal and any added value that might arise for the digital preservation community.

### *Exchange Output*

In recognition of receiving funding for the exchange participants agree to write a report describing the activities and results of their exchange. A financial report must also be submitted.

### *Application Deadlines*

The application deadline for DPE's first Research Exchange closed on the 1st of August 2007. The evaluation of participant's submissions is now underway.

15th of November 2007
1st of January 2008
1st of June 2008


For further information and online application, see: http://www.digitalpreservationeurope.eu/exchange

# III. THE DRAMBORA TOOLKIT

## Digital information – valuable and vulnerable

Within every domain area and operational context digital information is capable of demonstrating tremendous value. No longer limited to highly specialist computer laboratories, digital information is of great scholarly significance in every scientific and academic research discipline, a commercial necessity within every business environment and an increasingly central part of our cultural heritage, both collectively and personally. Unfortunately though, the potential value of digital information is almost inversely proportional to the likelihood of its permanence. Numerous threats are posed to the availability, integrity, authenticity and usability of our digital assets from the moment of their conception. Software and hardware obsolescence are real dangers, given the quick development cycles that so often characterise the technology industry. Physical media fragility too is of concern, with the lifetimes of optical, magnetic and solid state devices often unknown, but certainly not infinite. Uncertainties related to organisational and legal landscapes pose further questions. How will the discovery of digital materials be facilitated in the future, and how will access take place? Less physical concerns relate to our ability to *understand* digitally encoded content in the years to come. It is likely that over time the parameters of our knowledge and understanding will change. Contextual factors that we know about now and take for granted and are vital requisites for interpreting digital information might slip from our collective consciousness.

## Repositories – a means for management

A great deal of recent work within the digital preservation community has concentrated on the development of centralised infrastructures aimed at combating these threats, capable of providing the appropriate stewardship to safeguard and make available our digital assets. Defined more broadly than in strict technological terms, these *repositories* encompass associated systems, policies, procedures and assets, including people, money and facilities. However, as Clifford Lynch, of the Coalition for Networked Information has warned, "Stewardship is easy and inexpensive to claim; it is expensive and difficult to honor, and perhaps it will prove to be all too easy to later abdicate". With this concern in mind, RLG and OCLC recommended in their seminal 2002 publication "*Trusted Digital Repositories: Attributes and Responsibilities*" that the priority was not only to develop and populate repositories, but in addition, to ensure that demonstrable trust relationships exist between these infrastructures and information depositors, users, funders, legislators, third party repositories or service providers and other stakeholders.

## Trust and Trustworthiness

RLG and OCLC argue that trustworthy repositories must display a number of key characteristics. Again, these take a broad perspective of the repository, and involve organisational, procedural, financial, legal and technological considerations, as well as issues associated more directly with digital object ingestion, management, preservation and dissemination. At a meeting in January 2007 between the US Centre for Research Libraries (http://www.crl.edu), the German *nestor* project (http://www.langzeitarchivierung.de),

the Digital Curation Centre (http://www.dcc.ac.uk) in the UK and DigitalPreservationEurope ten core criteria for digital preservation repositories were conceived (http://www.crl.edu/content.asp?l1=13&l2=58&l3=162&l4=92):

1. The repository commits to continuing maintenance of digital objects for identified community/communities.
2. Demonstrates organizational fitness (including financial, staffing structure, and processes) to fulfil its commitment.
3. Acquires and maintains requisite contractual and legal rights and fulfils responsibilities.
4. Has an effective and efficient policy framework.
5. Acquires and ingests digital objects based upon stated criteria that correspond to its commitments and capabilities.
6. Maintains/ensures the integrity, authenticity and usability of digital objects it holds over time.
7. Creates and maintains requisite metadata about actions taken on digital objects during preservation as well as about the relevant production, access support, and usage process contexts before preservation.
8. Fulfils requisite dissemination requirements.
9. Has a strategic program for preservation planning and action.
10. Has technical infrastructure adequate to continuing maintenance and security of its digital objects.

Given the availability of such criteria, the community's subsequent objective has been to determine the optimal means by which conformance to these criteria can be established, and to conceive mechanisms to effectively communicate the successes of individual repositories to a wider world. The proposal that has perhaps inspired the greatest enthusiasm is the development of formal mechanisms to support repository audit and certification.

## Repository Audit and Certification

To date, three primary outputs have contributed to the development of criteria, means and methodologies for performing assessments of digital repositories. The Trusted Repositories Audit & Certification (TRAC) Criteria and Checklist (http://www.crl.edu/PDF/trac.pdf) (maintained by the US Center for Research Libraries) and the Network of Expertise in Long-Term Storage of Digital Resources for Germany (nestor) Criteria Catalogue (http://edoc.hu-berlin.de/series/nestor-materialien/8/PDF/8.pdf) each present a set of organisational, technological and policy-based criteria capable of supporting evaluation, and the subsequent certification of a wide variety of repositories. Neither defines a clear process for assessment however, and the Digital Repository Audit Method Based on Risk Assessment (DRAMBORA) developed by the Digital Curation Centre and DigitalPreservationEurope adds this methodological value, defining a risk-based approach, where success corresponds closely to an organisation's ability to ameliorate the risks that it faces. Certification, while served by the availability of objective criteria and methods capable of supporting comparison remains some way from practical realisation. Numerous questions abound: Which (if any) organisations are sufficiently permanent and widely renowned to assume the role of accredited certifying agency? Who would be

responsible for certifying the certifiers? Where would liabilities lie in the event of a certified repository's failure? Nevertheless, irrespective of such questions, there do appear to be demands from the repository community for such formal expressions of success. It is therefore expected that this issue is set to be explored in much more detail in forthcoming months.


**DRAMBORA** (http://www.repositoryaudit.eu ) defines a process for repository assessment that is suitable for repositories of various maturities, from those in planning stages to those in development or full production. Within its six main sections it presents a methodology geared towards an organisational context-aware self-assessment, encouraging repositories to establish a comprehensive self-awareness of their objectives, activities and assets before identifying, assessing and managing the risks that they implicitly face. DRAMBORA characterises digital curation as a risk-management activity; the job of digital curator is to rationalise the uncertainties and threats that inhibit efforts to maintain digital object authenticity and understandability, transforming them into manageable risks. Initial stages require auditors to develop an organisational profile, describing and documenting the repository's mandate, objectives, activities and assets. Latterly, risks are derived from each of these, and assessed in terms of their likelihood and potential impact. Finally, auditors are encouraged to conceive of appropriate management responses to identified risks. The process facilitates effective resource allocation, enabling repository administrators to identify and categorise the areas where shortcomings are most evident or present the greatest potential for disruption. The process itself is an iterative one and therefore subsequent recursions will evaluate the effectiveness of previous risk management implementations. Rather than represent an onerous additional task for repository administrators, it is hoped that iterating through the DRAMBORA workflow will become increasingly indistinguishable from good repository management practice.

### Benefits of DRAMBORA

Following a DRAMBORA based self assessment, repositories can expect to have:

- Established a comprehensive and documented self-awareness of their mission, aims and objectives, and of intrinsic activities and assets;
- Constructed a detailed catalogue of pertinent risks, categorised according to type and inter-risk relationships, and fully described in terms of ownership, probability and potential impact of each risk;
- Developed an internal understanding of the successes and shortcomings of their organisation, enabling them to effectively allocate or redirect resources to meet the most pressing issues of concern;
- Prepared themselves for subsequent external audit whether that audit will be based upon the Trustworthy Repositories Audit & Certification (TRAC), nestor Catalogue of Criteria for Trusted Repositories, or forthcoming Consultative Committee for Space Data Systems (CCSDS) digital repository audit assessment criteria (http://www.digitalrepositoryauditandcertification.org).

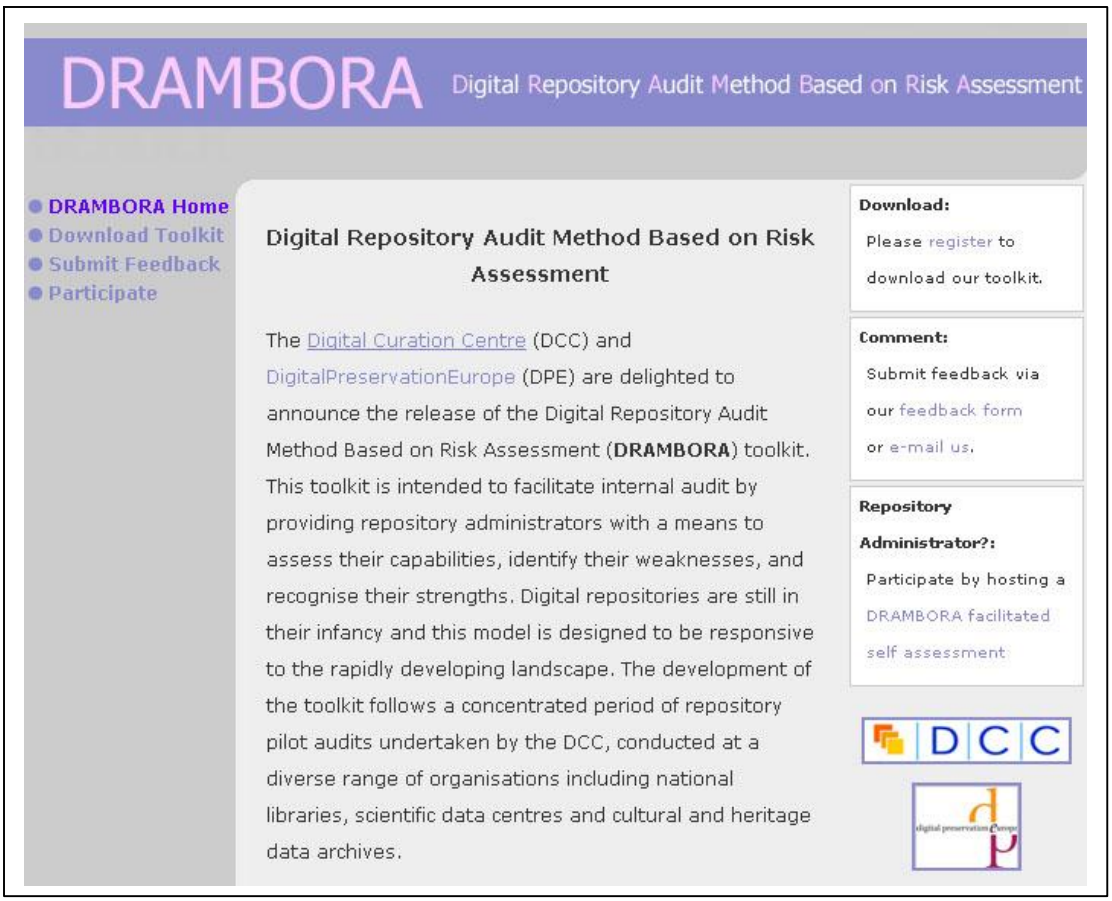
## Future Development

DRAMBORA's design reflects lessons learned during the Digital Curation Centre's international programme of pilot repository audits that took place in late 2006 and 2007. A second phase of facilitated self assessments is currently ongoing, led jointly by DPE and DCC in Europe and further afield to determine the limitations of the tool in its current form and areas in which it might be improved. In order to make the self-audit process as straightforward as possible the second release of the self-audit tool kit will be as an interactive web-based tool with semi-automated workflow mechanisms, options for pre-filled fields, and guidance materials in the form of examples, suggestions and comparisons.

Feedback from organisations that have used DRAMBORA is always very gratefully received. Furthermore, as part of the development process, we are actively seeking collaborations with repositories wishing to participate in a facilitated self assessment exercise. During these exercises, a member of the DRAMBORA team engages with the repository, providing on-site or remote assistance during its own self assessment. To provide feedback or to express interest in hosting a facilitated self assessment you can contact the DRAMBORA team via their website at http://www.repositoryaudit.eu/feedback or by email to feedback@repositoryaudit.eu.

## Getting DRAMBORA

DRAMBORA can be downloaded from the joint DCC/DPE Repository Audit website at http://www.repositoryaudit.eu/download.



# IV. BRIEFING PAPERS, POSITION PAPERS AND CASE STUDIES

The DPE project produces very short and synthetic documents providing key information elements, focussed on a specific topic and extremely communicative: Briefing Papers and Case Studies. Briefing Papers do not intend to be exhaustive on a certain topic or didactic materials but to address the reader's attention towards a problem or a set of possible solutions. In order to assure the final user point of view on practical requirements, some experts from different domains have been involved in Briefing Papers production. Case Studies illustrate the most relevant experiences carried out by organisations, companies or single experts in the domain of digital preservation. These documents are the description of "good practices", developed during a single experience or more durable projects. Their purpose is to be helpful for users and organisations by showing the faced problems, the open issues and the achieved assets.

Up to now DPE has produced the following Briefing Papers:

"*LOCKSS: Re-establishing Librarians as custodians of journal content*" by Adam Rusbridge.

This paper gives an overview of the LOCKSS (Lots of Copies Keep Stuff Safe) approach, a system that helps libraries regain custody of journal assets while maintaining the access and license restrictions stipulated by the publisher.

"*Automating semantic metadata extraction*" by Yunhyong Kim.

This paper faces the problems associated to the manual collection of metadata and addresses the challenge associated to the process of automating the extraction of semantic metadata from digital objects.

"*A data model for preservation metadata*" by Angela Di Iorio.

This paper illustrates the activities carried out by the PREMIS international working group that have defined an implementable "set of 'core' preservation metadata elements" for the digital preservation community.

"*Digital Preservation and Open Access Archives - Persistent access to open access digital assets*" by Valdo Pasqui.

This paper, after having evaluated potentialities and access of Open Access archives, analyses the strategies for long term preservation of these repositories.

DPE has produced the following Case Studies:

"**Ways to go about the clinical data management**" by Health Unit ULSS no. 8 in Asolo (Italy) Working Group. This document illustrates the process necessary to create a Digital Clinical Repository capable of managing, accessing and preserving more than 6.000.000 case files.

"**Digital preservation strategies in Public Administrations**" by Ilaria Pescini from Regione Toscana, Italy. This document describes the preservation system developed by Regione Toscana aiming at archiving and preserving the digital documents, and rationalizing the activities by improving the system of organisational workflow in Public Administrations.

### A Call to Ignite Debate

DigitalPreservationEurope is producing a series of thought provoking and controversial position papers on a range of issues surrounding digital preservation. It is our intention that these papers will promote vigorous debate within the digital preservation community and encourage people to think about digital preservation in new and innovative ways by exploring and challenging the received wisdom.

We would like to invite volunteers to submit position papers for us on topics such as:

- How to get rid of OAIS/OAIS is not the answer.
- Put your faith in Digital Archaeology.
- Digitisation is for access not for preservation.
- We should only keep what is most used.
- The future is paper.
- Future users don't matter.
- Digital Preservation is not a problem.

You can find the documents at: http://www.digitalpreservationeurope.eu/publications.

If you would like to share your experiences or focus the community's attention on a particular issue by writing a Briefing Paper, Case Study or Position Paper please submit your proposal to:

Maurizio Lunghi: m.lunghi@culturalheritage.it Chiara Cirinnà: cirinna@rinascimento-digitale.it

## V. USER COMMUNITY

DigitalPreservationEurope, building on the earlier successful work of ERPANET, works to improve coordination, cooperation and consistency in current digital preservation and curation activities to secure the longevity of digital assets and heritage. DPE promotes and supports the networking and coordination of research and innovation in digital preservation.

Individuals that would like to join the DPE community and take advantage of the benefits listed below can do so by completing the online registration form:

- Free access to all DPE publications.
- Discount of 30% for attendance at all DPE training events, workshops and conferences.
- Access to the DPE Exchange Program (DPEX).
- Access to the DPE Challenge.
- Access to the forthcoming online DRAMBORA toolkit.
- Access to the DPE user forum where DPE partners can share examples of best practice and exchange ideas and opinions with leading experts and practitioners at an international level.
- Access to a user community mailing list through which will provide advanced notification of DPE events, publications, tools and services.


Any enquiries concerning Partnership and the DPE User Community can be sent to:

Maurizio Lunghi m.lunghi@culturalheritage.it Chiara Cirinnà cirinna@rinascimento-digitale.it

# VI. COMPETENCE CENTRES STATE OF ART REVIEW

DigitalPreservationEurope (DPE) has carried out a State of the Art Review into competence centres in digital curation and preservation. This document reviews the current international landscape with regard to the availability and provision of digital curation and preservation expertise in the European Union and beyond. To enable the comparison of competence centres in different sectors and serving different communities DPE has developed a benchmarking model.

## Introducing the 7C's Benchmarking Model

Sources of digital curation and preservation competency exist in a wide range of institutions, projects and collaborative endeavours. Examples of the range and nature of these centres of competence include:

- University and scientific research community
- Ministerial structures
- National structures
- Industry
- Healthcare infrastructures
- EC-funded projects

Based on the salient features of the digital curation and preservation competence centres listed above, **DPE developed the '7C's' benchmarking model**. The criteria employed by this model are:

1. Capacity
2. Context
3. Credibility
4. Commitment
5. Certification
6. Competition
7. Communication

### Capacity

The centre must be able to clearly demonstrate its expertise in a specific aspect of digital curation and/or preservation. This expertise should be theoretical and/or practical and evident through participation in research activities, community building initiatives, international and standardisation initiatives, or the provision of services to a given user community or communities. Capacity may also be demonstrated through the take-up of services provided, community participation in the centre's awareness-raising events and training programmes, cutting-edge research outputs, or the success of the centre's efforts to influence change at the strategic decision-making level. The centre must also have an evident capacity for securing funding and attracting resources both at national and international level. These features determine whether the centre can provide a valuable and relevant service to a given user community.

### Context

Since a competence centre is part of a wider world populated by a community of users with internal mechanisms, constraints, rules and functions, its benefits and potential values must be explicitly defined in order to motivate that specific target community of practice. Context also includes governing bodies, legal regulations, inter-disciplinary connections and international research communities. In particular, the centre should demonstrate its ability to liaise with and influence sectoral governing boards, international initiatives and professional organisations. The first two elements of the 7C's model will determine whether the centre can provide **a contextualised service**.

### Credibility

Being considered trustworthy and reliable both by authorities and by the user community is essential for a competence centre. The credibility of tools, resources and methodologies produced by a competence centre is generally recognised through two distinct pathways – from the top-down or from the bottom-up. The top-down approach comes about when specific actions are mandated by a governing body that ensures take up is achieved via a system of rewards and penalties. The bottom-up approach comes about when the user community agrees to adopt best practices based on tangible benefits identified through research activity or practical experience. It is now clear that a convergence of these approaches is required to identify, promote, and ensure take-up and adherence to digitisation, curation and preservation standards and best practices. It is particularly important that the user community not only perceives the centre to be competent in relation to digital preservation and aware of the context of application and of legal-organisational issues, but also to be a member if of the community facing the same problems and speaking the same language. The first three elements of the 7C's model will determine whether the centre can provide **a trusted and contextualised service**.

### Commitment

A commitment or mandate is necessary both to hold a position of authority in relation to the user community and to secure financial sustainability. 'Authority' here is something distinct from the concept of 'credibility' discussed above. In this case there is an obligation on users to refer to the competence centre. The mandate can be 'limited in time and scope' and must be given by the country authorities, professional associations, international bodies, or the European Commission. The first four elements of the 7C's model will determine whether the centre can provide **a sustainable, trusted and contextualised service**.

### Certification

Compliance with international standards is essential nowadays for almost all the elements of our society. Standards dominate all areas of activity from professional behaviour to personal equipment. Certification can be obtained simply by respecting specific standards, or it can be granted by a 'super partes' agency, guaranteeing the quality of products and services offered by the centre. It is especially crucial that competence centres are able to demonstrate that they comply with the standards that they advocate to their user community. The first five elements of the 7C's model will determine whether the centre can provide **a certifiable, sustainable, trusted and contextualised service**.

### Competition

An environment that encourages competition can benefit all stakeholders. It can also help to ensure that competence centres do not become complacent but constantly strive to improve the quality of their services and resources. Adherence to standards can be a clear and transparent reference to test the performance and policies of competence centres, providing a 'super partes' reference to stimulate fair competitiveness. The first six elements of the 7C's model will determine whether the centre can provide **a competitive, certifiable, sustainable, trusted and contextualised service**.

### Communication

A competence centre must be able to communicate effectively with many different stakeholder communities. They must be able to communicate their user communities' needs and requirements to policy makers and funding bodies to influence change and be able to disseminate their expertise through outreach and training programmes aimed at a range of levels. Moreover communication with other competence centres ensures that duplication of effort is avoided and that limited resources are used to maximum benefit.

**By employing each of the seven criteria, a holistic view of the strengths and weaknesses of competence centres emerges.**

## Assessing the current competence centres landscape using the 7C's benchmarking model

The concept of competence is different in different environments and communities. It's different in terms of measurable parameters, indicators, credibility building, functions or services that might be offered and in terms of sustainability models. While the range and nature of competence centres vary, we have identified five generic models that may be used to help characterise the current landscape for the provision of digital curation and preservation expertise. While DPE recognises that other models may exist, for the purposes of this the state of the art review DPE examined the following five generic models:

- **Distributed centres of expertise**: There are many distributed centres of expertise and collaborative projects that bring together multiple institutions, often from a range of disciplines, to investigate specific research areas and to provide advice, guidance and tools for a specific user community. Examples of this type of centre are the DCC (http://www.dcc.ac.uk), KOPAL (http://kopal.langzeitarchivierung.de/index.php.en), DRIVER (http://www.driver-repository.eu), PLANETS (http://www.planets-project.eu), CASPAR (http://www.casparpreserves.eu), AHDS (http://ahds.ac.uk), DPE (http://www.digitalpreservationeurope.eu), CETIS (http://www.cetis.ac.uk) and nestor (http://www.langzeitarchivierung.de).

- **Single research-led institutions with preservation expertise**: Centres falling into this category are generally part of a larger organisation like a university or research centre. They tend to gain expertise through participation in short-term, externally funded research activity. Their range of expertise tends to be focused in a very specific area as funding for new projects is often awarded due to previous related experience. Examples of this type of centre are FUH (http://www.informatik.fernuni-hagen.de/ia), HATII (http://www.hatii.arts.gla.ac.uk), ULCC (http://www.ulcc.ac.uk), UKOLN (http://www.ukoln.ac.uk) and SUB (http://www.sub.uni-goettingen.de).

- **National libraries, archives or other organisations with preservation expertise**: These large, established institutions have proven their competence in the curation and preservation of analogue materials over decades and sometimes centuries. The majority of these types of centres are well aware of the risks facing our digital memory and have been actively working to improve their skill-sets to care for digital materials as well as they have cared for analogue materials. Examples of this type of centre are the NLA (http://www.nla.gov.au), BL (http://www.bl.uk), LoC (http://www.loc.gov/index), KB (http://www.kb.nl) and the LOUVRE (http://www.louvre.fr).

- **Commercial preservation centres/services**: These types of centres may be funded either through private membership fees or through consultancy fees. These centres often have good visibility and are generally regarded as highly competent by their user communities. As these centres serve clearly defined user communities, their work tends to be focused and user-driven. Examples of this type of centre are the DPC (http://www.dpconline.org), NDAD (http://www.ndad.nationalarchives.gov.uk) and ICCU (http://www.iccu.sbn.it).

- **International bodies or consortia, professional associations**: The formation of these consortium-based centres tends to occur voluntarily and in response to the specific needs of a particular community of practice. Participants in these types of consortia tend to represent international organisations aiming to influence the

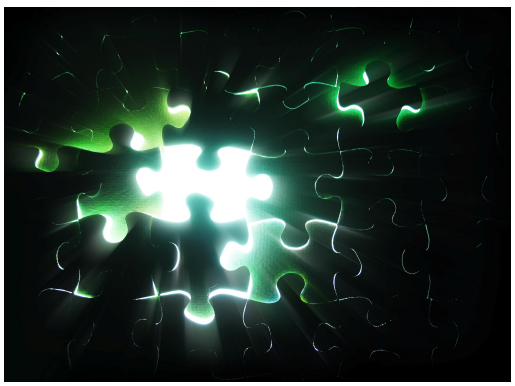
development of policy and standards and also to improve advocacy among their specific user community. Examples of this type of centre are the W3C (http://www.w3.org), DCMI (http://dublincore.org), ICA (http://www.ica.org), IFLA (http://www.ifla.org) and UNESCO (http://portal.unesco.org).

## Conclusion and recommendations

DPE is confident that the **7C's benchmarking model** will be of value for identifying possible new competence centre models that may benefit EU member states. Based on the results of its assessment using the '7C's' benchmarking model, DPE recommends that the European Commission consider measures that will result in a 'virtuous circle' approach to curation and preservation activity whereby user needs feed into research, development, service provision, dissemination and practice. The DPE 'Competence Centres: State of the Art Review' recommends:

- a federated approach to the provision of support and guidance,
- a life-cycle approach to the provision of support and guidance,
- better coordination of disparate competence centres' activities,
- improved integration of competence centres with industry,
- a more collaborative approach to the provision of training and outreach activities,
- increased research capacity within competence centres to help push forward the international research agenda,
- the introduction of competition between competence centres to drive performance forward, and
- the investigation of business models that will help to provide sustainable funding for competence centres' activity.

Please visit http://www.digitalpreservationeurope.eu/competence-centres to view the current list of competence centres. The list of competence centres is constantly updated and by registering your institution's details you will be helping to establish an international support network to promote the sharing of experiences and best practice in the long-term curation and preservation of digital assets.



Please visit http://www.digitalpreservationeurope.eu/publications for the full document 'Competence Centres: State of the Art Review'.

# VII. MARKET AND TECHNOLOGY TRENDS ANALYSIS

The National Library of the Czech Republic, a partner in the DPE project, carried out a representative survey focusing on recent developments and plans concerning the long term preservation of digital documents in European national libraries, archives and research institutions. Below is a summary of the findings of this survey.

### 1. Is digital long-term preservation (including migration, emulation, preservation metadata and planning etc.) one of the key strategic priorities of your institution?

82% of national libraries, 66% of national archives and 70% of research institutions indicated that long-term preservation is currently among their key strategic priorities. A further 12% of nation libraries, 27% of national archives and 11% of Research Institutions indicated that long-term digital preservation will be a strategic priority in the next two to five years.

### 2. Do you (or will you) have a trusted digital repository (according to the criteria listed in An Audit Checklist for the Certification of Trusted Digital Repositories)?

While only 30% of national libraries reported having a Trusted Digital Repository 61% indicated that they were working towards this goal. For national archives 32% reported having a Trusted Digital Repository and only 27% indicated that they were working towards this.

Despite highlighting a low number of Trusted Repositories currently available in Europe it is positive to see that 91% of national libraries and 59% of national archives recognise the importance of gaining Trustworthy status and see this as both a priority and a target.

### 3. Digital preservation is too big an issue for individual institutions to address independently. Who will your institution cooperate in this area with?

Respondents were given the choice of Memory Institutions, Research Institutions, Digital Document Producers and Software developers. For all three categories of respondents Memory Institutions were the first choice for cooperation in this areas. While the national libraries and archives rated the other three choices equally; research institutions indicated a preference for cooperation with software producers and vendors as a second choice.

### 4. The building and operation of a trusted digital repository is a big and expensive business. Will you create and operate the repository only for your library or share it with others?

While 85% of national libraries planned to cooperate with other institutions; only 53% of archives and 52% of research institutions planned to share in the creation and operation of their repository with other institutions. This difference could be explained by the existing experience national libraries have in the benefits of sharing their repositories of traditional documents and through coordinating activities in other areas such as Web archiving.

### 5. What system will your digital repository use?

The responses from national archives and research institutions were close with the majority, 38% and 53% respectively, favouring an open source system, and the rest split equally between commercial and in house systems.

53% of national libraries planned combined solutions the use of in-house, commercial and Open Source systems.

### 6. Which of the outputs listed in the model of DPE dissemination do you consider to be the most relevant for your institution?

Memory and research institutions alike indicated that the DPE website was the most relevant method of dissemination for them. The production of guidelines was also identified as relevant for both memory and research institutions. However conferences, seminars and workshops were considered more relevant by national libraries than by archives and research institutions.

**7. In the vision of FP7, national competence centres are seen as an integral way of ensuring effective development of expertise and services. Which institutions in your country do you consider to have the best background for becoming fully operational and trusted national competence centres?**

Memory institutions were identified by national libraries, national archives and research institutions as having the best backgrounds to become trusted national competence centres.

## Technical Solutions available for Digital Preservation

In the document, Market and Technology Trends analysis, DPE provides information on commercial and Open Source technological solutions that are already applied for digital preservation and that meet minimum standards. Bellow is a summary of the findings:

### Commercial:

**DIAS** (Digital Information Archiving System) – IBM

URL: http://www-5.ibm.com/nl/dias

The DIAS (Digital Information Archiving System) solution provides a flexible and scalable open deposit library solution for storing and retrieving massive amounts of electronic documents and multimedia files. It conforms with the ISO Reference OAIS standard and supports physical and logical digital preservation.

**DPS** (Digital Preservation System) – ExLibris

The DPS is a preservation solution for digital objects. The system conforms to the OAIS ISO standard and supports many of the standards in the library environment (METS, PREMIS, MARC, DC, OAI-PMH etc.). It is designed to support the acquisition, validation, ingest, storage, management, preservation and dissemination of different types of digital objects. It supports the E-legal deposit requirements.

### Open source based:

**CDS Invenio** – CERN, the European Organization for Nuclear Research

URL: http://cdsware.cern.ch/invenio/index.html

Developed by CERN, the European Organization for Nuclear Research, based in Geneva, CDS Invenio (CERN Document Server Software) is a suite of applications that provide the framework and tools for building and managing an autonomous digital library server. Its flexibility and performance make it a comprehensive solution for the management of document repositories of moderate to large size.

**Dspace** – Massachusetts Institute of Technology and Hewlett- Packard Company

URL: http://www.dspace.org

The DSpace digital repository system was designed to capture, store, index, preserve, and provide access to institutional digital research materials. It can accept all forms of digital materials ranging from text, images and datasets to websites, multimedia, video and

audio files. As DSpace's architecture mimics the structure of the organisation that uses DSpace, it is suitable for large and complex organisations that anticipate material submissions from many different departments. The architecture supports the implementation of workflows that can be customised for specific departments or other institutional entities.

**EPrints** – Developed by the School of Electronics and Computer Science of the University of Southampton/JISC

URL: http://www.eprints.org/software

Developed at the University of Southampton, the first version of the system was publicly released in late 2000. It was designed as repository software for e-prints, electronic versions of research articles, in pre-print or post-print versions or both. EPrints is a good candidate for many institutions as it is the least complex of the two above, and hence has the lowest skill level barrier to implement and maintain.

**Fedora** – Cornell University Information Science and the University of Virginia Library; supported by generous grants from the Andrew W. Mellon Foundation

URL: http://www.fedora.info

The goal of the Fedora Project is to provide Open Source repository software and related services to serve as the foundation for many types of information management systems. The Fedora digital object repository management system is based on the Flexible Extensible Digital Object and Repository Architecture (Fedora). The system is designed to be a foundation upon which full-featured institutional repositories and other interoperable web-based digital libraries can be built.

**Greenstone Digital Library Software** – New Zealand Digital Library Project at the University of Waikato

URL: http://www.greenstone.org/cgi-bin/library

Greenstone is a suite of software for building and distributing digital library collections. The aim of the Greenstone software is to empower users, particularly in universities, libraries and other public service institutions, to build their own digital libraries. Greenstone incorporates all the features of the existing system, and is backwards compatible: it can build and run existing collections without modification.

**LOCKSS** (Lots of Copies Keep Stuff Safe) – The Stanford University LOCKSS Program team – LOCKSS Alliance (Mellon Foundation grant)

URL: http://www.lockss.org/lockss/Home

LOCKSS ('Lots of Copies Keep Stuff Safe') is Open Source software that provides librarians with an easy and inexpensive way to collect, store, preserve, and provide access to their own, local copy of authorised content they purchase. Running on standard desktop hardware and requiring almost no technical administration, LOCKSS converts a personal computer into a digital preservation appliance, creating low-cost, persistent, accessible copies of e-journal content as it is published. Since pages in these appliances are never flushed, the local community's access to that content is safeguarded.

Please visit http://www.digitalpreservationeurope.eu/publications for the full document 'Market and Technology Trends Analysis'.

## VIII. VILNIUS CONFERENCE: DIGITALPRESERVATIONEUROPE (DPE) CONFERENCE

To coincide with the DigitalPreservationEurope project meeting in late June Vilnius University Faculty of Communication organised a conference on the preservation of digital content. The conference took place **in Vilnius on the 29th of June 2007** with 37 participants from the Lithuanian and Latvian library and computer science communities. Partners from DPE presented their latest findings and discussed the current situation of digital preservation in Baltic countries in detail with conference participants.


**Maurizio Lunghi**, the Scientific Director of the Fondazione Rinascimento Digitale, in his presentation "DigitalPreservationEurope: Community Building and Competences Centres", explained that a major objective of the DPE project is to foster collaboration and synergies among on-going projects and existing initiatives. DPE also works to raise awareness of Digital Preservation challenges among different user communities through community building and competences centres. Lunghi stressed the importance of cooperation in advocacy and outreach activities, training and continuing professional development. He also explained various models through which competence centres can act as focal points for member states offering advice and guidance on digital preservation issues.

**Prof. Seamus Ross**, Director of the Humanities Advanced Technology and Information Institute at the University of Glasgow and Technical Director of DPE, gave a presentation on "DPE and Managing Risk in Digital Repositories". He described the purpose and benefits of the Digital Repository Audit Method Based on Risk Assessment (DRAMBORA) toolkit developed in cooperation between DPE and the Digital Curation Centre. Prof. Ross also took participants through the main stages of a repository audit using DRAMBORA and encouraged the audience to audit their repositories and share their experiences.

DPE collaborates with the other major European digital preservation projects PLANETS and CASPAR and works to eliminate the duplication of effort and the sharing of best practices between them. In particular DPE works closely with PLANETS and CASPAR in the organisation of events,

training and other activities. **Birte Christensen-Dalsgaard**, Director of Development at the State and University Library of Denmark and a representative of the PLANETS project, gave a presentation on "The PLANETS Digital Preservation Project". She described the project, its partners and approach and also introduced delegates to the various tools and services they have developed such as the planning services, characterization services, preservation actions, the interoperability framework, and test-bed.

**Prof. Andreas Rauber**, of the Department of Software Technology and Interactive Systems (ifs) at the Vienna University of Technology (TU-Wien), in his presentation "Planets: Preservation Planning" described in more detail the preservation planning and workflow processes developed by PLANETS. He investigated the main planning questions such as how to decide what is most the most suitable solution for you, how to measure and evaluate the results of each preservation action and how to define a controlled and trusted environment and procedure for applying or testing preservation strategies.

**Holger Brocks**, of FernUniversität in Hagen, spoke on the subject of "Digital Preservation: the Future of Our Collective Memory". Brocks introduced the findings of the DPE Research Roadmap, a document based on an extensive crosswalk of existing preservation research agendas which aims to provide a concise overview on the core issues which have to be addressed in future digital preservation research. He argued that despite all discussions in recent years about what kinds of research are needed in the area of digital preservation, no concise and well-developed strategy that represents the views of a broad community has yet emerged.


Vilnius in 1576

After the presentations and productive discussions, DPE partners and participants agreed that it would be very useful to analyse the context of each Baltic country in the field of digital preservation. The key finding of the discussions was that Baltic countries are currently doing a great deal in the field of digitisation, but are not as yet concerned with the preservation of this content for future users.

# IX. TRAINING


The first joint DPE/Planets/nestor training event, 'Principles of Digital Preservation: a hands-on approach', will be held **in Vilnius, 1-5 October 2007**.

The different levels of awareness of Digital Preservation issues in each country and the variety of institutions and organisations involved in Digital Preservation activities has lead to an increased need for training in digital preservation that can meet local requirements while ensuring consistent and comparable knowledge and skills across the European Union.

By collaborating with other European funded projects DPE seek to implement a Pan-European framework of training and continuing professional development in the area of digital preservation. This will enhance individual modular training packages and ensure that project partners can provide comprehensive coverage of digital preservation topics and access to the widest possible network of expert trainers. It is hoped that this training framework will lead to a certified European qualification in digital preservation.

The courses offered will be used to train staff within organisations responsible for preserving society's scientific and cultural heritage such as Museums, Libraries, and Archives. DPE is planning to run five two-week long training events in Lithuania, Austria, Spain, Germany and Italy.

The first event to be held is a joint initiative between DPE, Planets and nestor and will take place in Vilnius (Lithuania) during 1-5 October 2007. Through attendance on this course participants will gain an understanding of the main principles of digital preservation and practical ways to tackle digital preservation problems. Students will be provided with access to a three day online course before and a two day online course after the training event with material and exercises covering the topics that will be discussed. In addition to this the training course will be linked to the international conference "Communication of memory in archives, museums and libraries: the interaction of science, policy and practice" providing participants with a unique opportunity to meet with other researchers, international experts, and practitioners across disciplinary and national boundaries.

Participants will leave with a coherent and practical understanding of activities surrounding digital preservation. In particular they will gain a developed understanding of key needs and challenges in digital preservation, the role and use of metadata, the file formats which are currently considered most useful for preservation, trust and trustworthiness in the context of digital preservation, and the benefits of a coherent preservation planning process to overall digital preservation strategies.

We hope the audience will consist of practitioners and researchers from the Archives, Libraries and Museums sector, as well as other institutions such as data archives, government departments and businesses with an interest in the topic. Internationally renowned lecturers will each lead one half-day session and some will be available through the week for further discussion.

## Vilnius course time table:

The two week intensive training will consist of three days pre-course online training, 7 half-day teaching sessions, a two day conference and two days of on-line post-course training.

| | |
|---|---|
| **Wednesday, 26 September 2007** | Access to three day long online pre-course training. |
| **Sunday, 30 September 2007** | 18.00: Overview of the joint training school and Introduction to the Lecturers |
| | 19.00: Opening Reception |
| **Monday, 1 October 2007** | 9:00 – 9:30: Registration |
| | 9:30 – 10:00: Welcome |
| | 10:00 – 13:15: Introduction in Digital Preservation |
| | 14:30 – 17:30: OAIS Reference Model |
| **Tuesday, 2 October 2007** | 9:00 – 12:45: Preservation metadata |
| | 14:00 – 17:30: File formats and significant properties |
| | 19:00 – 22:00: Social event (excursion in Vilnius) |
| **Wednesday, 3 October 2007** | 9:00 – 12:45: Trusted repositories |
| | 14:00 – 17:30: The Preservation Planning Process and Planets |
| **Thursday, 4 October 2007** | 8:30 – 12:25: Identifying essential characteristics of digital objects: the Planets approach |
| | 12:25 – 12.45: Feedback Dissemination of certificates |
| | 13:50 – 17:40: Conference |
| **Friday, 5 October 2007** | 8.30 – 18.40 Conference |
| **Monday-Tuesday, 8-9 October 2007** | Access to two day long online post-course training. |

Lecturers:

**Michael Day**, UKOLN, University of Bath.

**Manfred Thaller**, University of Cologne & Planets & nestor.

**Stefan Strathmann**, Göettingen State and University Library (SUB), nestor & DPE.

**Hans Hofman**, National Archives of the Netherlands. Planets & DPE.

**Christoph Becker**, Vienna University of Technology. Planets & DPE.

Trakai castle – Vilnius, Lithuania

Further details will be posted as they become available on http://www.wepreserve.eu/events/dpe-planets-vilnius-2007

# X. RESEARCH ROADMAP

After two decades of setting digital preservation research agendas there is little evidence of actual progress in developing solutions. Our heritage may now be at greater risk because many in our community believe that we are making progress towards solving the preservation challenges. A major cause of this has been that the community has consistently widened their research focus. As a community we have much more work to do.

The DPE research roadmap aims to contributing to the planning of our future research and development in Digital Preservation by means of analysing the state of the art in Digital Preservation research and existing research agendas on a global level. In addition to this the Roadmap will analyse the state of the art and future markets of computer science and information and communication technology, the needs and demands of the Digital Preservation user communities and their leading experts. It is hoped that this analysis will pave the way for the development of conceptual, technological and application oriented building blocks for advanced solutions to support Digital Preservation applications in a variety of possible scenarios.

The DPE team analysed research agendas and the research activities of the past twenty years, and offered experts the opportunity to contribute their ideas about what preservation research needs to be done. Our findings show that while much work has been done on the periphery, the central problems of digital preservation have yet to be addressed or resolved. We consider that although access, digital rights management, users and other aspects make digital preservation useful, this is outside the scope of core research.

DPE has released a draft Research Roadmap identifying ten core domains for preservation research:

**Restoration**, **Conservation**, **Management**, **Risk**, **Significant Properties of Digital Objects**, **Interoperability**, **Automation**, **Context**, **Storage**, and **Experimentation**

For further information, please contact: Holger Brocks Holger.Brocks@FernUni-Hagen.de

© Pzemyslaw Kucinski