



*Memory is the human faculty
of retaining and reproducing present
and past thoughts, objects, habits, culture
for future generations independently
from circumstances that inspired them*

DPE consortium

ISSN 1752-9220 (print) ISSN 1752-9239 (online)

digitalpreservationEurope newsletter

Issue 3

October 2008



www.digitalpreservationeurope.eu



Coordination Action funded by the European Union
Information Society and Media Directorate General

digitalpreservationEurope



DigitalPreservationEurope newsletter

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

October 2008 – Issue 3

© [2008 - DPE] Annual periodicity

Free usable. DPE must be notified and credited when contents are used for any public work. No commercial use permitted. – Last updated September 2008

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



Contents

1. Community Building: Results and State of the Art	1
2. Briefing Papers, Case Study and Position Papers.....	6
3. Training activities summary and assessment.....	9
4. CeBIT 2008 participation	11
5. Digital Preservation Challenge	15
5.1. Miguel Ferreira My Personal Impressions on Digital Preservation..	16
6. DPE Unique Identifier Service	18
7. DRAMBORA: State of the Art.....	19
8. Planning the Trusted Digital Repository with PLATTER.....	22
9. DELOS appraisal Report review	24
10. News from the projects	26
10.1. PLANETS.....	26
10.2. CASPAR.....	29
10.3. nestor	32
10.4. SHAMAN.....	34
10.5. PROTAGE.....	37
10.6. LiWA.....	36
10.7. TextGrid.....	39
11. Social networking and Digital Preservation	42
12. Persistent identifier distributed system for Cultural Heritage digital objects .	46

1. Community Building: Results and State of the Art

DPE has been working hard over the last two years to build a community of experts, individuals and organisations with skills and interest in digital preservation. This community has grown with new Associate Members in many countries of the European Union. Some of which will be present at the annual We Preserve conference in Nice where a meeting with them will be organised.

DPE has been working with its community and Associate Partners in particular to bring down barriers to information and raise awareness on Digital Preservation issues by translating key documents produced by the project into many languages of the European Union. The series of briefing papers in particular will be of interest to many and we encourage you to visit our website to find out more.

The Associate Partners of DigitalPreservarionEurope are as follow:



AIB – Associazione Italiana Biblioteche - Italy
<http://www.aib.it>

The Associazione Italiana Biblioteche (AIB) is the professional association of Italian librarians. Founded in 1930, AIB is the general library association in Italy, the National Association Member of IFLA, and by far the oldest and largest association from this field in Italy.



Biblioteca Nacional de Portugal - Portugal
<http://www.bn.pt>

Founded in 1796, the National Library of Portugal holds the largest historical and cultural bibliographic collection in the country with over 3.5 million items from a wide range of types and genres, both in print and manuscript, spanning from 11th century documents to the personal papers of many Portuguese writers up to the 21st century. Promotion and co-ordination of digitization efforts, and all related management and technical issues, are amongst the NLP priorities for the short and medium term, aimed at supporting preservation of its heritage collections while enhancing global access to Portuguese cultural resources.



Charles University Computer Centre - Czech Republic

<http://uvt.cuni.cz/UVTENG-1.html>

Charles University, founded in 1348, is one of the oldest universities in the world and belongs to the most eminent educational and scientific establishments in the Czech Republic. Charles University aims to be recognized as a competitive research university on the world stage and recognizes the importance of international cooperation with prestigious educational and scientific establishments. As such Charles University has entered into 450 bilateral contracts and 170 international partnerships with foreign universities.



CIB – Centro Inter-Bibliotecario - Italy

<http://www.cib.unibo.it/cib>

Established in 1990 as a university centre dedicated to the development and the implementation of library automation projects, CIB (Centro Inter-Bibliotecario) hosts and maintains ACNP, the Italian Union Serials Catalogue created in cooperation with CNR (Consiglio Nazionale delle Ricerche). Late in 2001 CIB started Alma-DL (Alma Mater Studiorum Digital Library), a complex project with the aim of developing an organizational and a technological infrastructure to collect, organize, archive, integrate and provide access to the digital contents that the University acquires or produces and makes available primarily to institutional users and, when possible, worldwide.



CILEA Consorzio Interuniversitario - Italy

<http://www.cilea.it>

CILEA is a consortium of ten universities in Lombardy, established in 1974. CILEA provides Information and Communication Technology services to universities and related organizations and enterprises, promote the most efficient and innovative use of the most advanced computing and data transmission systems, to coordinate and develop research and to disseminate the culture of information technology and communication. In order to manage digital resources, CILEA has developed Codex[ml], an integrated system which performs import, creation, long-term storage, delivery and fruition of digital images and XML documents based on metadata standards like METS, MAG and PREMIS.



Cineca - Italy

<http://www.cineca.it>

CINECA is a not for profit inter-university Consortium of 32 Italian Universities, the CNR - National Research Council, the National Institute of Oceanography and Experimental Geophysics - OGS, and the MiUR - Ministry of University and Research. Established in 1969, CINECA is currently the most important computing centre in Italy and one of the major sites worldwide. Since the end of the 1980s, the Consortium has broadened the scope of its institutional mission by gradually embracing all IT sectors, developing management and administrative services for universities and designing telematic sys-

tems for the exchange of information between the Ministry of University and Research and the national academic system.



The Croatian State Archive - Croatia

<http://www.arhiv.hr>

The Croatian State Archives (CSA) is the central archival institution in Croatia and providing archival services to archival and current records created by state and public institutions, state and corporate enterprises, families and individuals whose activity covers the whole or a greater part of the Croatian territory, or is of State interest. CSA has begun the construction of a new archival information-evidence system ARHiNET. ARHiNET comprises of several modules including: Arranging of archival material, Register books, Archival documentation and evidences and Service for archival records outside archives and User service.



DANS – Data Archiving and Networked Services - The Netherlands

<http://www.dans.knaw.nl>

DANS - Data Archiving and Networked Services - is an institute of the Royal Netherlands Academy of Arts and Sciences (KNAW) and the Netherlands Organisation for Scientific Research (NWO), located in The Hague, The Netherlands. Since its establishment in 2005, DANS has been archiving and making accessible research data in the humanities and the social sciences. To this end DANS develops permanent archiving services and stimulates others to follow suit. One important result of the activities carried out by DANS is the establishment of EASY, a user friendly self archiving system. DANS also provides other services such as offering access to the files of large data collections, both nationally and internationally.



The Directorate-General of Portuguese Archives (DGARQ) - Portugal

<http://www.dgarq.gov.pt>

DGARQ is a Directorate General accountable for the planning and execution of national archival policy in Portugal. DGARQ has in one of its missions the support of all below administrations to set up systems of archive, doing reports about evaluation and selection of documents for permanent conservation, holds, preserves and allows the access to archival inheritance of national interest.

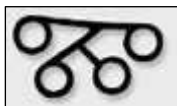


FEUP - Faculty of Engineering of the University of Porto - Portugal

<http://www.fe.up.pt>

The Faculty of Engineering of the University of Porto (FEUP) is one of the 14 faculties of the University of Porto, the largest university of Portugal. FEUP works closely with several institutions and agencies of note such as the European Space Agency, IBM, Microsoft and CERN, as well as research universities such as the Massachusetts Institute of Technology, the University of Texas at Austin and Carnegie Mellon. FEUP is a pioneer in the development of university information systems. SIGARRA, its integrated system on the Web, has been extended to the majority of the schools in the

university and integrates the management of courses, programmes, projects, staff and supporting services.



Faculty of Humanities and Social Sciences and Department of Information Sciences - Croatia

<http://www.ffzg.hr> and <http://www.ffzg.hr/infoz/web2>

The Faculty of Humanities and Social Sciences at the University of Zagreb is one of the largest institutions of higher education in Croatia. The Faculty traces its roots back to 1669, when the original Zagreb Academy was founded, and in its modern form this school of higher learning has existed since 1874. Because of this tradition, the Faculty is obliged to preserve all of the worthy accomplishments of its predecessors while meeting the challenges of the future and seeking out new solutions.

The Department of Information Sciences was established in 1981. The Department is active in the research in many areas of information and computer science including the Croatian identity in information institutions, knowledge management, records management, Croatian lexicography, machine translation, information processing, e-learning etc.



Health Unit ULSS 8 Asolo (Veneto) - Italy

<http://www.ulssasolo.ven.it>

In the last three years the Health Unit ULSS no. 8 in Asolo has laid solid groundwork for an intensive and widespread use of digital technologies in both in the clinical and in the administrative domains. Their aim is to make certain services, such as medical reports, e-learning, radiological territorial PACS (Picture Archiving and Communication System), on-line reservations, statements of account for suppliers and medicine flow, available electronically.



Kirtas

<http://www.kirtaseurope.com>

Since 2001 Kirtas Technologies has pioneered and perfected the technology used today in quality, high-speed, non-destructive mass book digitization. The company's revolutionary technology redefines digitization of all bound documents, delivering the gentlest handling and the highest image quality with quick handling and few errors. Kirtas' philosophy is to enable the rapid and economical mass digitization of books and bound documents at a level of quality that will stand the test of time.



NUK - National and University Library - Slovenia

<http://www.nuk.uni-lj.si/nukeng.asp>

The National and University Library (NUK) of Slovenia was founded in 1774 and serves as the central library in the national library system. The national library it undertakes the collection, cataloguing, conservation, access to and preservation of the national written cultural heritage, submitted under legal deposit and other 'Slovenica' publications, published abroad. In 2005 the Digital Library of Slovenia (<http://www.dlib.si>) was launched, a web portal providing access to digitised knowledge and cultural treasures, based on digitisation of materials, held by NUK.



Regione Toscana, Governance-network and information engineering for the regional system coordinating area - Italy
<http://www.regione.toscana.it>

The Regione Toscana is a regional public body, which was established in 1970 but started to operate only in 1972. Its organization consists of 8 General Departments in order to administrate the region more effectively. The Governance-network and information engineering for the regional system coordinating area, as part of the General Department for administration and information system, is in charge for the whole ICT sector by fixing a strong relation with its territory, especially with its local public authorities.



Roskilde Festival - Denmark
<http://www.roskilde-festival.dk>

Roskilde Festival is an annual music festival with an international audience. Originally inspired by the famous festivals at Newport, Isle of Wight and Woodstock, Roskilde Festival has been a playground for contemporary music since its foundation in 1971. All documents from Roskilde Festival's history have been carefully filed, thus forming a perfect time capsule for the history of international youth culture, as well as the history of the local community in Roskilde. In 2006, Roskilde Festival established an archive project in order to make its vast collection accessible to the public. The archive has initiated an electronic/digital registration of all files including audiovisual and paper documents.



VUL - Vilnius University Library -Lithuania
<http://www.mb.vu.lt>

Vilnius University Library (VUL), established in 1570, has a community of over 26000 of readers and holdings totalling over 5.38 million items dating from the 13th century to the present day. The Restoration Department was established in 1968 and is at present the largest library subdivision of this kind in Lithuania. The digitization policy of VUL has two major goals to preserve cultural heritage and promote collections of cultural importance.

See where DPE Associate Partners are on DPE website at:
<http://www.digitalpreservationeurope.eu/about/index.php#map>

2. Briefing Papers, Case Study and Position Papers

As part of its dissemination activities, DPE is getting started to publish a series of very short and concise documents providing information on specific topics. These documents include Briefing Papers and Case Studies: the first intend to direct the reader's attention towards a problem or a set of possible solutions; the second to provide examples of "good practices", the experiences of organisations, companies or single experts in the domain of digital preservation. In its first year DPE published four Briefing Papers covering a variety of topics such as LOCKSS as a system for library custody of e-journals, on metadata and its semantic extrac-

tion, on Open Access archives and on a data model for digital preservation (PREMIS). The two Case Studies were also delivered concerning Digital Protocols in Public Administration developed by Regione Toscana, and on an experience of clinical data management by Health Unit USSL 8 in Asolo, both in Italy. As announced in the previous Newsletter, DPE is producing some of provoking and controversial Position Papers on a range of issues surrounding digital preservation. The intention of these papers is to promote vigorous debate within the digital preservation community and encourage people to think about digital preservation in new and innovative ways.

In its second year, DPE produced the following Briefing Papers:

"Portico: a Collaborative Approach to Preservation" by Eileen Fenton

This paper gives an overview of the Portico model of archive and preservation infrastructure for the scholarly community. Through Portico, libraries and publishers can contribute toward a shared infrastructure which supports a mutually beneficial and long-term preservation of scholarly literature published in electronic form.

"Professional Development in Digital Preservation: a life-long requirement" by Ross Harvey

This paper takes stock of the e-learning context as mainstream delivery mode for education and

training. E-learning provides the blend of theory and practice necessary for effective digital curation and preservation. It is well suited to meeting the life-long learning requirements of personnel working in this field.

"Open Source in Digital Preservation" by Robert Neumayer

Having evaluated the Open Source concept and types of license, this paper examines the role of Open Source in the digital preservation context for open standards. It concludes that file format specifications and document formats can be also open source, and related to open standards, satisfying quite a number of preservation requirements, even if they cannot be

proclaimed as a one-size-fits-all solution for digital preservation.

“Preservation of Digital Audiovisual Content” by Richard Wright

This document deals with the problems of audiovisual records. These files have requirements (size; specific formats) that are not adequately addressed by current technology, and digitisation has been a solution, which has created a new problem: preservation for digital AV content.

“The Challenge of Appraising Science Records” by John Faundeen

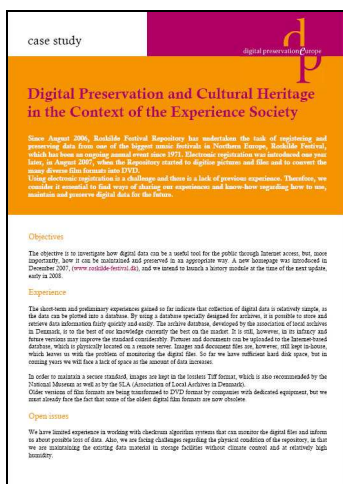
This paper examines how the records management function of appraisal can help determine the long-term value of science records. Various themes employed are briefly discussed to illustrate commonalities in differing approaches.

“Persistent Identifiers for Cultural Heritage” by Emanuele Bellini, Chiara Cirinnà and Maurizio Lunghi

In the Cultural Heritage domain it is essential not only to identify a resource but also to guarantee continuous access to it. This document evaluates questions about persistent digital object identification, including texts, music, video, still images, scientific documents, identifying some system requirements as a guideline.

“World Data Center for Climate: Preservation of Earth System Model Data” by Michael Lautenschlager

This paper analyses the strategies of the new archive developed by the ICSU (International Council for Science) World Data Centre for Climate (WDCC) for archiving and long-term preservation of Earth System model data.



In its second year, DPE published the following Case Study:

“Digital Preservation and Cultural Heritage in the Context of the Experience Society” by Henrik Christiansen

This document describes the preservation model developed by the Roskilde Festival Charity Society. The objective of Roskilde Festival Repository is to investigate how digital data can be a useful tool for the public through Internet access, but, more importantly, how it can be maintained and preserved in an appropriate way.

Finally, DPE produced the following Position Papers:

“Why Appraisal is not ‘Utterly’ Useless and why not the Way to Go either” by Robert Neumayer and Andreas Rauber

This provocative position paper argues that random selection should replace appraisal as the method of deciding what to keep and what to throw out in the archival world. The author argues that, in contrast to appraisal, random selection will provide 1 A fair unbiased view of contemporary life, 2 Simplicity and cost effectiveness, 3 Privacy protections and 4 A future-proof processes.

“So where’s the black hole in our collective memory?” by Ross Harvey

Harvey's position paper asks the important question, Have the digital preservation community cried wolf too often? Are our strident, alarmist proclamations about the loss of digital materials too extreme? He argues that our inability to bring evidence to bear in support of such claims leave us exposed and easily overlooked.

Marialisa Cornacchia
marialisacornacchia@yahoo.it



Join the debate on the series of DPE Position Papers at:
<http://www.digitalpreservationeurope.eu/mou/forum>

You can find all Briefing Papers, Position Papers and Case Studies at:
<http://www.digitalpreservationeurope.eu/publications>

3. Training activities summary and assessment

As part of its core objectives, Digital Preservation Europe (DPE) is implementing a suite of training events based on current best practices and identifying additional practitioner training and staff development needs. DPE training events aim to facilitate knowledge transfer between leading experts and established practitioners in digital preservation so that event participants could return to their own institutions with the practical skills needed to begin to tackle their specific digital preservation challenges. Since 2007 DPE has worked in close co-operation with other projects: Preservation and

Long-term Access through Networked Services (Planets) , Cultural, Artistic and Scientific knowledge for Preservation, Access and Retrieval (CASPAR) and Network of Expertise in Long-Term Storage of Digital Resources (nestor) to identify the knowledge requirements of our core user communities, and to distinguish areas where collaborative digital preservation courses could be delivered. This collaboration has made it possible for us to offer training events covering a number of inter-related topics that are out of the scope of any of the projects individually, but of relevance to our collective target audience groups.



In co-operation with Planets, CASPAR, Digital Curation Centre (DCC) , DELOS Network of Excellence on Digital Libraries and nestor, DPE has already delivered seven training events:

- DELOS Summer School 2007: International Summer School on Digital Preservation and Digital Libraries in Pisa (Italy), 3-9 June 2007.
- nestor/DPE Spring School 2007 'Introduction in Digital Preservation' in Gernrode (Germany), 11-15 March 2007.
- DPE/Planets/nestor joint training event 'Principles of Digital Preservation: a hands-on approach' in Vilnius (Lithuania), 1-5 October 2007.
- nestor/DPE Winter School 2007 - 'Practical Application Areas of Digital Preservation' in Gernrode (Germany), 11-15 November 2007.

- DELOS Summer School 2008 in Tirrenia (Italy), 8-13 June 2008.
- nestor/DPE Summer School 2008 'Introduction to digital preservation, its technologies and strategies' in Staufen/Breisgau (Germany), 16-20 June 2008.
- DPE/Planets/CASPAR/nestor 'Joint Training Event: 'Starting out: Preserving Digital Objects - Principles and Practice' in Prague (Czech Republic), 13-17 October 2008.

Nearly three hundred participants from twenty six countries (from Europe, Asia, South and North America) attended these training events, representing libraries, archives, universities and variety of other institutions from the cultural heritage and commercial sectors.

The titles of the training events themselves indicate that topics covered during these sessions include general as well as more specific issues of Digital Preservation. Lectures by international experts provided participants with an appreciation of the issues and challenges surrounding digital preservation; a coherent and practical understanding of digital preservation activities; experience with workflow modelling, metadata definition, and ingest process management; an appreciation of the different approaches to selection and appraisal; knowledge of the approaches to repository design and deployment; an understanding of the issues surrounding assessment, audit and certification

of digital repositories; knowledge of digital curation techniques and practices; a firm understanding of the issues of authenticity, integrity, and reliability in relation to digital libraries; a familiarity with file formats currently considered beneficial for preservation.

The training events have been a great success, with participants gaining new knowledge, having the opportunity to discuss their newly acquired skills and apply them to real-life examples. The majority of attendees were particularly impressed by the hands-on approach, which focused on exercises and practical work that drew upon their own experiences and work contexts. In addition to the sessions participants were often provided with access to online training materials before and after the course, including recommended reading, questions and exercises. This allowed for the development of knowledge and skills beyond the issues covered during the event.

Vilma Karvelyte
wilma.karvelyte@gmail.com

Jurate Kupriene
jurate.kupriene@gmail.com

4. CeBIT 2008 participation

CeBIT is regarded as the leading business event for digital technology and in 2008 hosted 5,845 exhibitors in total from 77 countries. The IT industry is highly specialised and fragmented and there are few means of communication that address the IT industry as a whole. CeBIT is one of the world's largest trade fairs showcasing the state of the art in information technology and addressing the whole IT industry. Held in Hannover, Germany, CeBIT 2008 ran from Tuesday the 3rd to Sunday the 9th of March 2008 and attracted 495,000 visitors up 3% on the previous year, with visitors from outside Europe totalling over 100,000.

Our primary objectives in attending CeBIT were

- to establish the current level of awareness of digital preservation in the IT and business communities,
- to raise awareness of digital preservation,
- to raise awareness on the European led initiatives in digital preservation and,
- to showcase progress and results of the four projects.

It is clear from the evidence presented here that we have achieved these aims.

DPE, in cooperation with CASPAR, Planets and nestor, lead the coordination of a shared WePreserve stand to showcase our cutting edge digital preservation research and technol-

ogy developments and to highlight their relevance to industry. A number of the projects' tools and research themes were demonstrated by DPE, PLANETS, CASPAR and nestor.

During the trade fare we welcomed a total of 990 visitors to our stand, an average of 165 people a day. We discussed the concept of digital preservation, the challenges faced when trying to preserve digital cultural heritage and scientific knowledge, and the specific research being carried out in the FP6 projects to address these digital preservation challenges. Particular care was taken to ensure that discussions with each visitor to our stand were aimed at an appropriate level to their existing knowledge and background with regards to digital preservation issues.

Visitors who had some previous knowledge of digital preservation, showed a particular interest, or were deemed a valuable future contact for our projects, were invited to complete a short questionnaire. This instrument gathered some basic information about the visitor and their level of awareness of digital preservation. Throughout the trade fare 181 questionnaires were completed representing 18.2% of all visitors. This represents an average of 45 in-depth conversations and contacts made on each day, which far exceeded our expected targets at the outset of the event.

In *Figure 1* we can see a break down of the Job titles of visitors to our stand. This shows that the majority of visitors to our stand, 42%, can be described as high level decision makers. This compares very favourably to that of CeBIT overall, *Figure 2*, where only 15% of visitors were identified as senior management or high level decision makers. However without further information from CeBIT regarding their methods of job title categorisation it is not possible to achieve a direct comparison.

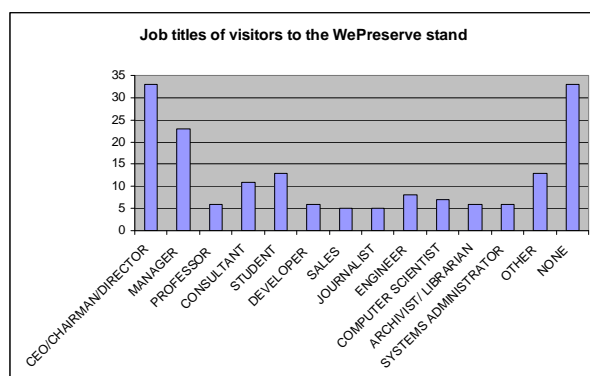


Figure 1 Job titles of visitors to the stand

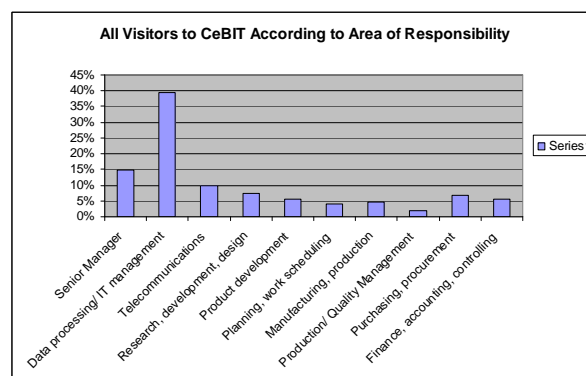
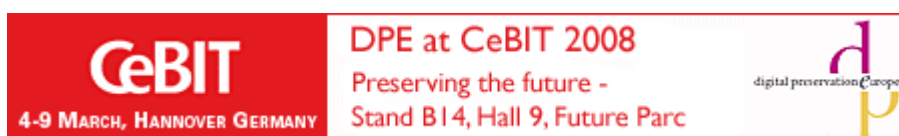


Figure 2 Areas of Responsibility of all visitors to CeBIT

Through our attendance at CeBIT we were able to speak face to face with senior decision makers in the ICT industry and influence them to take seriously the challenges we face in digital preservation. Reaching such a senior audience is the most cost effective and high impact method to meet our objectives to raise awareness of digital preservation issues and the work undertaken by European funded projects in this area. Individuals at this level are able to implement change within their businesses and cascade information to their employees. This is an extremely valuable opportunity and one we are not able to recreate outside CeBIT.



Of the 181 visitors we surveyed 52% were already aware of digital preservation and 48% were not aware. However these visitors were selected to survey because they had some previous knowledge of digital preservation, showed a particular interest, or were deemed a valuable future contact for our projects, so this figure cannot be taken to be representative of the total visitors to our stand. Figures 3 and 4 are very different with only 10% aware of digital preservation:

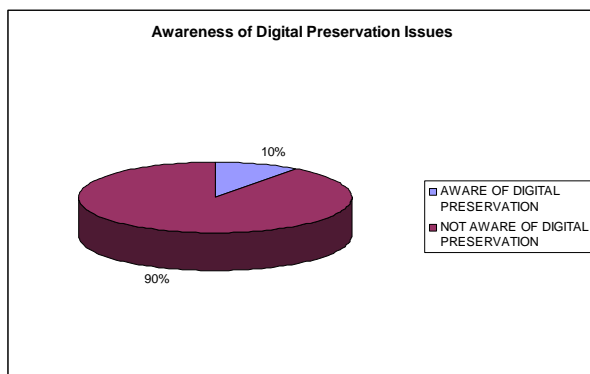


Figure 3 Awareness of Digital Preservation issues for all visitors to the stand

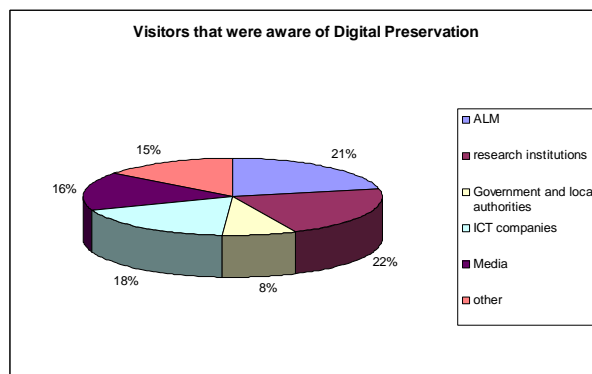


Figure 4 Sectors of the visitors that were already aware of digital preservation issues

When we examine the backgrounds of visitors that reported being aware of digital preservation issues we find that those from Archives, Libraries, Museums or Research Institutions are in the majority, with 43%. It is perhaps more surprising to note that of the groups identified those from a Government or Local Authority background make up the smallest portion of this group. Although it would be necessary to carry out further surveys in order to confirm this figure is not an anomaly of the demographic of CeBIT, this does suggest that there is work to be done in awareness raising of digital preservation issues in this group. This is especially important in light of the European Commission target for member states to have national strategies for long term digital preservation and deposit in place by mid 2008 (Annex: Priority Actions and Indicative Timetable, Council Conclusions on the Digitisation and Online Accessibility of Cultural Material, and Digital Preservation, Official Journal of the European Union, 7 December 2006, http://eur-lex.europa.eu/LexUriServ/site/en/oj/2006/c_297/c_29720061207en00010005.pdf).

It is positive to note that of the visitors to our stand that were already aware of digital preservation the majority, 62%, reported that this was a strategic priority for their businesses or institutions. For the total visitors to our stand this figure is 6.2%.

While the majority of people we spoke to felt that digital preservation is indeed a worthwhile objective, it is clear that there is considerable work to be done in articulating workflows, specific roles and responsibilities, and the range of practical tools that DPE and other projects' have to offer to assist them in realising this strategic priority. This objective would be best served through a joint presence at CeBIT 2009.

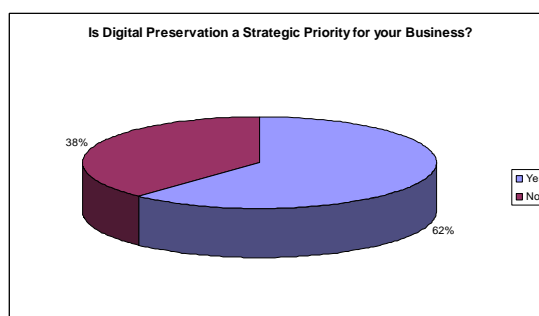


Figure 5 Digital Preservation as a strategic business priority of visitors to the stand

Our team of stand representatives from across the four projects must be highly praised for their enthusiasm, drive and commitment to our combined success throughout our attendance at the trade fair. Any passer-by that hovered for more than a moment was promptly approached by a member of the team and asked politely if they would like to learn a little more. This compares very favourably with the experiences of our team at other stands. It is certain that it

was this proactive strategy that led us to experience the success and high volume of interactions reported here and that the success of attendance at CeBIT 2009 will be determined by the individuals staffing the stand. Further we will use the experience and contacts gathered at CeBIT 2008 in order to make our attendance at CeBIT 2009 a success before we arrive.

Emily Nimmo

E.Nimmo@hatii.arts.gla.ac.uk



Representatives from DPE, Planets, CASPAR and nestor in the WePreserve Stand at CeBIT 2008

5. Digital Preservation Challenge

The Digital Preservation Challenge aims to promote innovation at all levels from students to researchers, and provide an insight into the range of digital preservation risks currently being faced by international research and practitioner communities. Digital-PreservationEurope's (DPE) 1st Digital Preservation Challenge ran from 25 May to 15 July 2007.

The first digital preservation challenge attracted numerous high-quality submissions and our first prize winner, **Miguel Ferreira** University of Minho, Portugal, successfully managed to solve all six sections of the challenge. A formal announcement of the winners and an overview of the submissions was presented at

ECDL in Budapest, Hungary on 18 September 2007.

The second DPE Preservation Challenge sponsored in part by XEROX ran from 17 January to 31 July 2008. The participants of the second DPE Challenge presented outstanding solutions and showed a deep understanding of digital preservation and the challenges associated with it.

DPE is proud to announce the winners of the second challenge:

1st Price Alex Mason, Durham University
2nd Price Juan-José Boté Vericad, Universitat de Barcelona, Universitat Oberta de Catalunya
3rd Price Mac Kobus Stuttgart Media University



More information about the Second Challenge at <http://www.digitalpreservationeurope.eu/challenge>

5.1. Miguel Ferreira

My Personal Impressions on Digital Preservation

Much has been achieved over the last ten to fifteen years in the field of digital preservation. From the point where there was not even a righteous term to describe the concepts inherent to “digital preservation” to the point where businesses and class studies are being built around it.

Early work on digital preservation was mostly focused on raising awareness among technicians and the general population about the terrible costs of losing information due to technological obsolescence as well the inherent difficulties in maintaining long-term access to contemporary cultural digital heritage. This problem was seen by many as a new kind of Y2K apocalypse. A lot of speculative articles, talks and news reports have been brought to light with a single purpose in mind - to incite panic amongst those who depend on digital information and to compel the ones ruling the money to create new funding programs to instigate research on digital preservation.

The fact of the matter is that those overblown, speculative articles, although not entirely accurate in terms of facts and predictions, were fundamental for the creation of a worldwide movement of researchers, scholars and funding agencies deeply concerned with digital preservation. The result of that was spectacular. Never before so many people had practical knowledge about digital preservation. Never before there were so many conferences, journals and publications dedicated to this cause. Never before there were grants for innovative research on digital preservation exceeding the 50M€ barrier.

Our understanding of digital preservation evolved to a point where it became evident that preserving digital information was more of a political and managerial issue than it was technical. It would eventually come down to a technical issue when the appropriate managerial tasks had been unable to predict users behavior or new trends in technology development. And that was exactly what the Digital Preservation Challenge was all about. The fundamental question that laid behind the challenge was - how may one preserve a group of disparate digital objects when no managerial work has been done over time to ensure its correct interpretation?

The problems proposed in the challenge made it clear that there are many diverse preservation scenarios and that it may be extremely difficult to find good information, tools and services for carrying out effective preservation interventions. The challenge also made clear that preservation may be extremely time-consuming and that general solutions, applicable to all sorts of preservation contexts, are extremely difficult to attain.

Even though research on digital preservation is pretty active these days, there is still a large demand for preservation tools such as freely available conversion applications with embedded quality assessment modules that report on any loss of properties or functionality. Format identification and characterization tools are fundamental for anyone doing deep object validation and long-term preservation. Such tools should be able to cope with rare niche formats, not just those ubiquitous formats already supported by most operating systems.

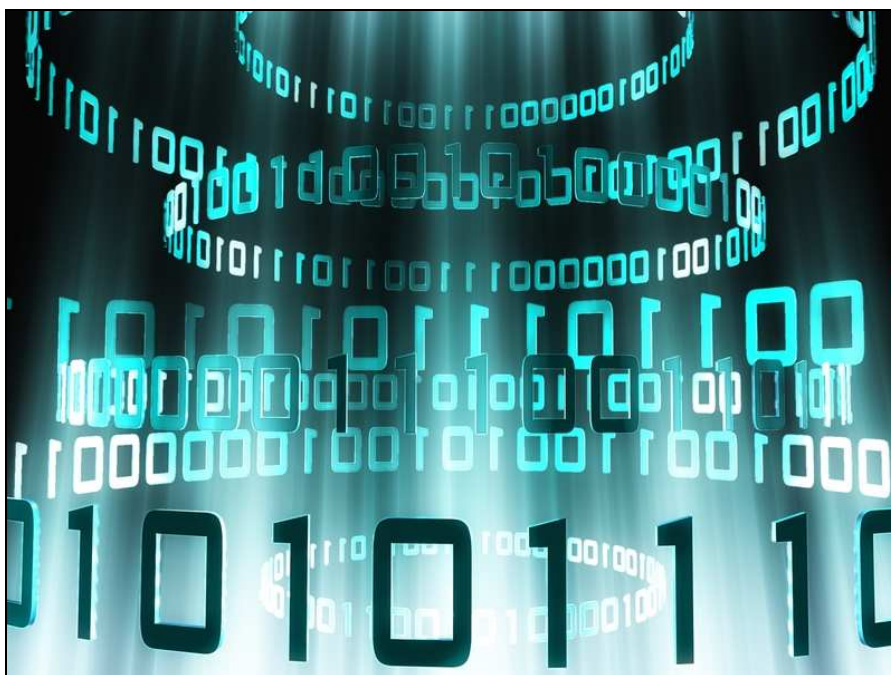
It is my personal belief that a large part of the future work on digital preservation should be centred on the education of the software industry on the concepts of digital preservation so that it may develop simple ways for its users to create archival versions of their objects without

having to pay additional fees or install any supplementary software modules. It should be an industry standpoint to fight this problem; this being a problem created by the industry itself in the first place.

Miguel Ferreira
Department of Information Systems
University of Minho
Campus de Azurém, 4800 Guimarães – Portugal
mferreira@dsi.uminho.pt

The complete document is available at:

http://www.digitalpreservationeurope.eu/publications/reports/DPC_revised.pdf



6. DPE Unique Identifier Service

The unique identification of digital objects has become a vital issue across a whole range of domains such as digital heritage and archival institutions or (digital) libraries. The ongoing transition to using digital documents has brought along a whole new level of complexity in document identification due to both the sheer number of publications available and the wide range of formats and forms of digital publishing. Online storage of documents has introduced problems of their

own such as broken links or countless copies of the same document.

DPE provides a persistent uniform resource locator (PURL), a naming and resolution service for general Internet resources, at <http://purl.wepreserve.eu>. It provides globally unique identifiers for digital objects. PURLS can be resolved using web browsers, also to access the actual resource. PURLs can be integrated in web pages, documents, or other resources. The service is open for everyone.

The service allows to:

- register users
- create PURLs, domains, groups
- modify PURLs, domains, groups
- search for PURLs, domains, groups
- validate PURLS

The screenshot shows the PURL Home Page. At the top right, it says 'digital preservation Europe' with a logo. Below that is a navigation menu: 'home & news | contact us | site map | rss | search | staff'. On the left, there is a vertical 'welcome' message. The main content area has a heading 'Persistent URL Home Page' and a paragraph explaining that a PURL is a Persistent Uniform Resource Locator. Below this, there are two columns of links. The left column is titled 'Further Information and Resources' and includes links for 'A brief introduction to PURLs', 'A longer introduction to PURLs', 'Frequently Asked Questions', 'Download the PURL software', 'PURL-L mailing list', and 'More info'. The right column is titled 'Interacting with this Resolver' and includes links for 'Create your first PURL', 'Register as a user', 'Create PURLs, domains, groups', 'Modify PURLs, domains, groups, users', 'Search this resolver', 'Validate PURLs', and 'Power user's page (all features)'. On the right side of the page, there is a sidebar with links for 'about DPE', 'news and events', 'exchange programme', 'DPE challenge', 'DPE user community', 'DPE registries', 'DPE publications', 'DRAMBORA', and 'Eprints'. At the bottom of the page, there is a footer with 'back to top | valid.shtml | Published: 17 March 2008 | Last Modified: 17 March 2008 | info@digitalpreservation.eu | © DPE 2008-2008'.

Stephan Strodl
strodl@ifs.tuwien.ac.at

More Information can be found at <http://purl.wepreserve.eu>
or contact info@digitalpreservation.eu

7. DRAMBORA: State of the Art

The Digital Repository Audit Method Based on Risk Assessment (DRAMBORA) was developed jointly by Digital Preservation Europe and the Digital Curation Centre (UK) as a self-assessment tool for digital preservation repositories. DRAMBORA approach requires repositories

to expose their organisation, policies and technical infrastructures to rigorous scrutiny. The assessment consists of a series of structured exercises, enabling the repository to build a comprehensive registry of its pertinent risks, arranged into a report that facilitates effective management.

Pilot audits with DRAMBORA

DRAMBORA was first released in March 2007 and has subsequently gone through a process of refinement and automation. The assessment methodology has been tested and commented on in nearly 20 pilot assessments undertaken in an internationally diverse selection of repositories: archives, digital libraries and data centres. Among these are institutions and services such as the UK National Digital Archive of Datasets, the National Archives of Scotland, National Library of the Czech Republic, National Archives of Netherlands, the Danish Internet Archive, Biblioteca Nazionale Centrale di Firenze, Charles University of Prague. Within the context of the DELOS Digital Preservation Cluster four audits of digital library environments were undertaken, using DRAMBORA, with a view to determining common characteristics of digital library repositories, in order to facilitate both knowledge transfer and comparison. The report, due to be published imminently, describes a range of common objectives, constraints, roles, responsibilities, activities and risks within the University of Michigan Library's MBooks, CERN's Document Server, Gallica at the Bibliothèque Nationale de France and the Swedish National Library's Digital Library.

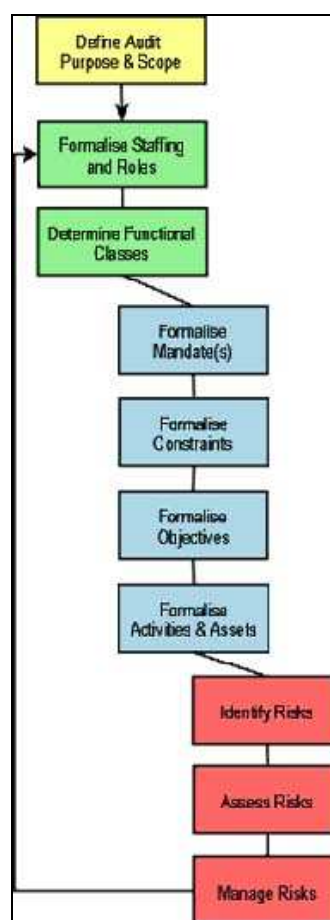


Figure 1. DRAMBORA workflow

The pilot assessments show that the DRAMBORA methodology itself is flexible, and responsive to the structural and contextual peculiarities of individual repositories, its metric for success directly linked with repositories own aims. The end result of the repository self-assessment is a register of risks the repository is facing expressed in terms of their likelihood and potential impact. Mapped to organisational aspirations and efforts this facili-

tates subsequent organisational development and resource allocation, and offers a quantifiable insight into the current severity of risks faced. The DRAMBORA process places considerable emphasis on demonstrable, and not just inferred, success. It can be used both independently and in association with the other repository assessment checklists.

DRAMBORA Interactive on-line tool

In early 2008 a second version of the self-assessment toolkit has been released as DRAMBORA Interactive, a freely available web based tool at <http://www.repositoryaudit.eu>. The on-line tool accumulates feedback and comments received from the pilot assessments and includes features that guide and support the auditor throughout the assessment process and documenting its findings. At the time of writing there are over 60 repositories registered and using DRAMBORA Interactive for assessment.

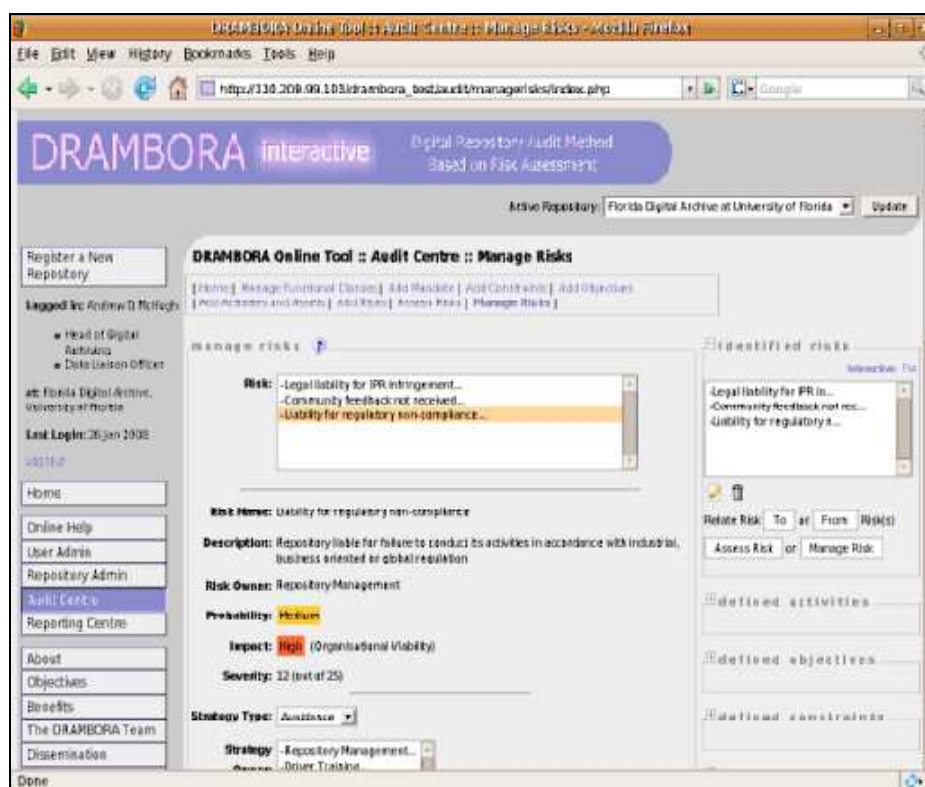


Figure 2. Screenshot from DRAMBORA Interactive.

You can find out more about DRAMBORA and the experiences of its users at <http://www.repositoryaudit.eu/dissemination>.

First DRAMBORA auditor's training course

To support the take-up of the DRAMBORA Interactive toolkit and the use of the DRAMBORA methods DPE in cooperation with the DCC are pleased to announce our first training course for auditor's:

DRAMBORA Auditors training
Prague, Czech Republic
13-17 October 2008

The course is designed for practitioners actively involved in funding, supporting, developing, implementing, and/or managing digital preservation repositories. The course will also be of value to auditors and risk managers from archives, libraries and museums sectors, and legal professionals and consultants. The course will enable attendees to:

- Comprehend the concept of risk with regards to digital preservation repositories
- Recognise the need for evidence-based evaluation for building trust in digital repositories
- Understand how the DRAMBORA toolkit can be used to help design and develop systems and workflows that can help build trusted digital repositories
- Obtain skills needed to undertake a thorough assessment of digital repositories using DRAMBORA toolkit
- Appreciate the range of staff and skill-sets required to implement and sustain a trusted digital repository
- Obtain a working knowledge of the issues surrounding assessment, audit and certification of digital preservation repositories.

For more information and registration please go to:
<http://www.repositoryaudit.eu/training/prague-2008>.

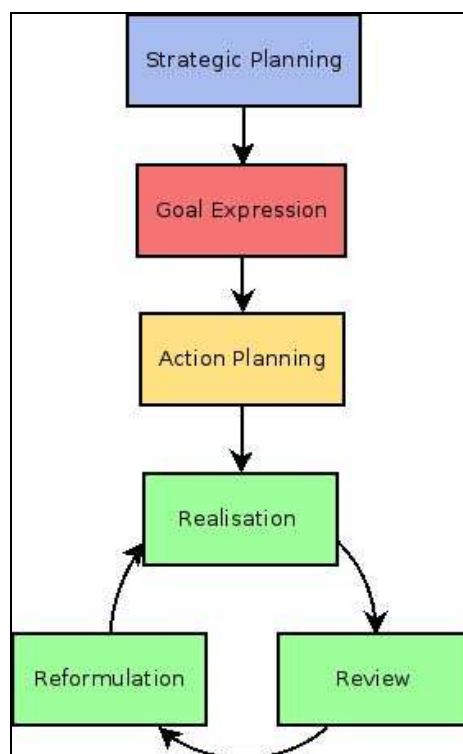
Raivo Ruusalepp
National Archives of Netherlands
raivo.ruusalepp@nationalearchief.nl

8. Planning the Trusted Digital Repository with PLATTER

DPE and other similar initiatives have greatly raised the general awareness of issues of long-term digital preservation, especially for memory institutions such as libraries, museums, and archives, and increasingly in the wider community of custodians of digital material, for example in the industrial, scientific, medical, and legal fields. Once enlightened as to issues surrounding digital preservation, many record holders naturally begin to look for advice on how to create a trusted digital repository with the minimum necessary investment in labour and skills. The problem they quickly discover is that trust cannot be purchased on a CD-ROM. Rather, a trusted digital repository is the result of implementing a strategic plan covering many different business processes and stretching far beyond the purely technical issues of storage, migration, emulation etc. Implementing such a solution can be a daunting prospect, which is why DPE has set itself a goal of providing repository planners with a toolkit to guide them through the process of planning for trust.

PLATTER (Planning Toolkit for Trusted Electronic Repositories) defines a planning cycle through which repository managers are able to set and monitor goals and objectives for their repository, based on the repository's mission and the widely-accepted "Ten Core Principles of Trusted Repository Design". The Ten Core Principles represent the crucial bridge between PLATTER, a planning toolkit, and DRAMBORA, the DPE-supported audit and assessment tool. Because the two toolkits are

based on the same underlying principles, their co-ordinated use will allow a new repository both to implement policies for trust and to demonstrate its trustworthiness to its stakeholders.



The PLATTER Planning Cycle

PLATTER steers a middle way between overly prescriptive recipes, which may not be applicable to every repository, and overly generalised frameworks which provide too little concrete guidance. This is achieved by specifying a set of nine SOPs (Business Plan, Acquisition Plan, Staffing Plan, Access Plan, Technical Plan, Data Plan, Succession Plan, Disaster Plan and Preservation Plan) which every trusted repository should ideally possess, and then defining within each plan a set of high-level goals which it is expected each repository will consider. It is then up to the individual repositories to decide which of these goals to adopt, based on criteria of relevancy and cost.

In PLATTER, goals are SMART (specific, measurable, assignable, realistic and time-related) and PLATTER not only guides repositories through the process of setting its goals but defines a planning cycle by which these goals may be refined and redefined throughout the repository lifecycle. By following this cycle, repositories will not only be able to establish trust but maintain it as their repository develops.

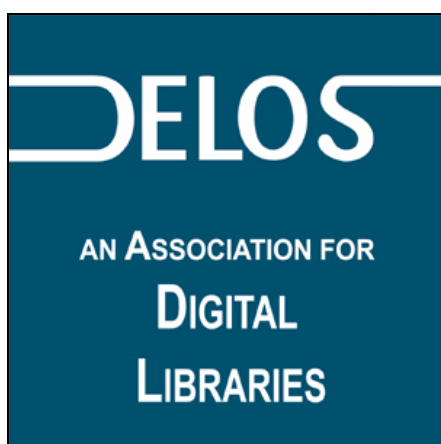
The road to trust is not only long and bumpy but hard to follow and with an uncertain end. With PLATTER as a guide (and DRAMBORA as a signpost) future travellers should find that road a little less forbidding.

PLATTER is available as a download from
http://digitalpreservationeuropa.eu/publications/reports/Repository_Planning_Checklist_and_Guidance.pdf

Colin Samuel Rosenthal
csr@statsbiblioteket.dk

9. DELOS appraisal Report review

The broad problem area considered by this report [Oliver, G., Ross, S., Guericio, M., Pala, C.: Report on Automated re-Appraisal: Managing Archives in Digital Libraries (Glasgow: DELOS NoE, January 2008)] is how the process of automatic re-appraisal of digital holdings (resulting in either the disposal or the retention of an object) may be effectively handled in the context of digital libraries. At different times after material has been ingested into a repository (which is what a digital library is) it may be necessary to re-assess whether the ingested material should be retained or disposed of. Given that this activity will often concern a substantial quantity of digital objects, it would be sensible to automate such a re-appraisal process, identifying what objects need to be removed based on pre-defined criteria.



<http://delos.info>

Given the proliferation of digital information of all types and the challenges of preservation, identifying what subset of that information is actually worth keeping is critical. This report investigates the relevance of processes to determine what is 'worth keeping' for digital libraries and suggests ways in which technology can be used to automate processes. The findings are particularly applicable to documents created in uncontrolled environments and to libraries.

Appraisal, the determination of the worth of preserving information, continues to be significant in the digital environment. Furthermore, the concept is applicable beyond the record-keeping domain in which it originated. A number of strategies have been identified to undertake appraisal, any one of which, or combination of, may be appropriate to a specific information community or domain. In considering the automation of the appraisal function in the context of a digital library or archives, it becomes clear that this will include the assessment of individual items. Automation enables a level of granularity that is rarely, if at all, possible in the case of manual appraisal methods, without loss of cognizance of the placing of items within the aggregation that they belong. The results of item-level assessment can inform the overall appraisal determination.

Research underway on metadata extraction, together with the structural view of genres, shows a great deal of promise for the digital library developer and user communities. In addition, the technological possibilities now present to facilitate input of other voices into the selection of information that has value for

communities open up a way forward to a new information age, one that need no longer be exclusively defined by dominant societal forces. As a result of our analysis of the approaches to appraisal we identified a series of appraisal criteria and structured these so that we can represent them as appraisal rules. Rules are susceptible to representation as active knowledge components. In considering the next steps, this representation suggests three models of automation:

Hybrid: A combination of manual and automated decision making. For instance, application of functional appraisal methodology supplemented by subsequent automated triage to determine the feasibility of preservation at the item level.

Appraisal engine: Where a document is submitted to an appraisal engine for analysis using a combination of text mining and rule-based reasoning.

Profiler: The development of a prototype to review a variety of information object types (image, document, dataset for example) and apply appraisal rules, probably again using rule-based reasoning methodologies.

It is though critical that when digital objects such as documents are selected for destruction or retention that why the disposal decisions were taken be recorded. One of the strengths of automation is that it provides this chain of evidence.

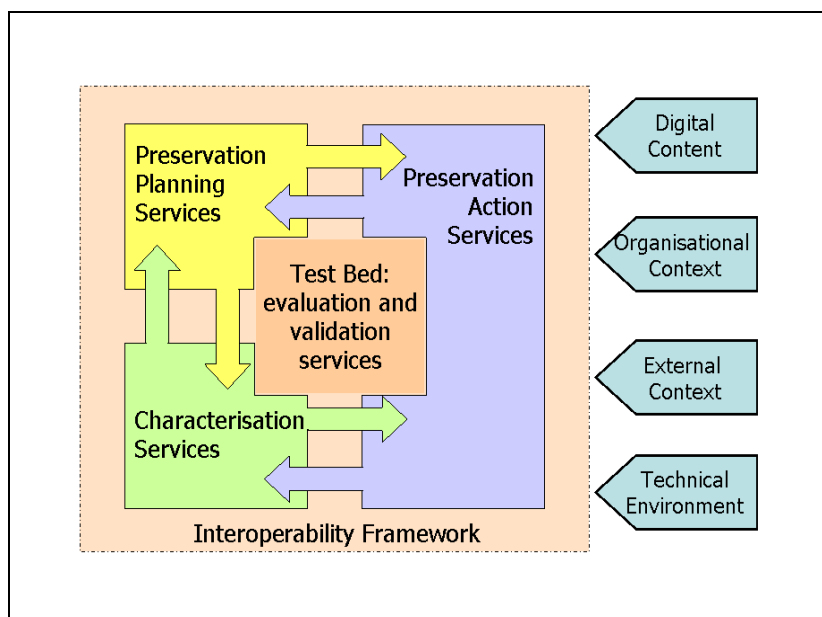
G. Oliver
Seamus Ross
Maria Guercio
C. Pala

10. News from the projects

10.1. PLANETS

Planets (project number IST-033789) is a four-year project, set up in June 2006 and co-funded by the European Commission under its Framework Programme 6, to address challenges in preserving digital assets. Coordinated by The British Library, the project brings together 16 national libraries, archives, research institutions and leading technology companies in Europe. These challenges include dealing with technology obsolescence and automating digital preservation to make it viable and cost-effective.

Planets provides a framework to underpin digital preservation activity, and is delivering practical tools and services to help libraries, archives and other institutions to preserve their digital content for the long term. Planets technology is focussed on the needs of content holding organisations with a core mission to preserve and provide access to digital material, such as national libraries, archives, universities, and government. Many of the core requirements, however, may be shared by large corporations or even long-lived smaller businesses. Plug-in capability will allow third-party tools and services to be incorporated easily.





Planets has now completed initial releases of each of the major software components in the Planets platform, including:

- Planets Preservation Planning Tool (Plato), which enables organisations to make informed selection of the most appropriate preservation plan against their own requirements. Plato allows organisations to identify preservation requirements for digital collections, define potential alternatives and select the most appropriate treatment plan. The selection of alternative plans is supported by the Planets Comparator, which provides the ability to measure and compare the characteristics of digital objects before and after treatment. Plato has been demonstrated at several events, including workshops where delegates have had hands-on experience of the tool, and a second release is expected later this year. Planets is also developing a model to assist organisations in creating preservation policies and to provide them with a framework to embed preservation planning into day-to-day activity.
- Planets Preservation Characterisation Registry, which automatically identifies the essential characteristics of digital materials. The registry contains technical information about the significant properties of digital objects. Planets has developed methods to express these characteristics, to extract them from files, and to compare them. This is enabled by two related languages:
 - XCDL – a characteristics description language
 - XCEL – a characteristics extraction languageTypical characteristics include colour depth of an image, number of images

within a textual document, metadata included within the file or the font type of text.

- Planets Preservation Action Services. Planets supports preservation of digital objects via migration (action which results in a transformation of the object itself) or via emulation (use of tool which replicates the technical environment required for rendering the object). To support this, Planets has created an inventory of the file formats most commonly found in digital archives. The project has also surveyed the tools that exist to characterise and preserve each file format and conducted a gap analysis to identify where tools do not exist or are insufficient. A blueprint has been defined for those that will need to be created, including emulation tools. Planets has integrated the emulators Dioscuri and Universal Virtual Computer (UVC).
- Planets Testbed, which is a controlled software environment that allows researchers and practitioners to conduct digital preservation experiments and gather data on a scientific basis. This enables us to learn which preservation strategies and tools work effectively for different types of content, and benchmark these against the experience of others and best practice. The Testbed makes it possible to experiment with preservation strategies and tools and analyse the outcome objectively. Groups of reference documents (the Testbed Corpora) act as benchmark content. The Testbed has been demonstrated at several events, and is expected to be available for institutions in Spring 2009.
- Planets Interoperability Framework, which provides shared functions and integrates the Planets tools and services into an easily managed preservation system. The Interoperability Framework underlies Planets applications such as Plato and the Testbed. It enables us to execute complex preservation workflows and also provides core software services such as authentication, logging and user management. Its extensibility allows plug-in of third-party tools and services. Planets is working on guide-

lines for wrapping preservation tools for integration into Planets. The term ‘wrapping’ refers to the provision of a web service interface so that the tool can be incorporated into preservation workflows.

The Planets project is now past the half-way point and is making good progress towards its end goals. The external review of the project in July 2008 was positive and encouraging.

Over the next few months, look out for:

- New version of Plato
- Collection Profile tool
- Gap Analysis – survey of Preservation Action tool provision
- Second versions of UVC and Dioscuri emulators
- State of the art report on preservation action technologies

- Preservation action toolset ready for demonstration
- Testbed available to external institutions
- Training and Outreach events.

Planets is active in sharing results and encouraging wide adoption of its tools and services. The project team presents its work at conferences and exhibitions around the world, and in scientific papers and articles and newsletters. Many of the papers and presentations from these are available via the Planets website.

Planets organises workshops, tutorials and training events around Europe. Those so far have been well-received, and a series of further events is being planned to take place during 2009. One of these will be aimed specifically at suppliers/vendors.

Lynne Chivers
Lynne.Chivers@bl.uk

For more information about Planets, publications, newsletters, training and outreach events, and to sign up to receive news about Planets, please visit
<http://www.planets-project.eu>

Planets encourages feedback and may be contacted at
info@planets-project.eu .

10.2. CASPAR

CASPAR (Cultural, Artistic and Scientific knowledge for Preservation, Access and Retrieval) Project, after the successful review by the European Commission, is now approaching the 3rd year of activities.

The core objectives of the CASPAR project are to design and implement a common digital preservation framework for heterogeneous data based on the OAIS Reference Model along with the research, development and integration of advanced components, for use in a wide range of preservation activities. These components are integrated in a common software platform that enables the building of services and applications for several digital preservation purposes. Starting from the common

CASPAR technology and organisational framework, the ambitious challenge for CASPAR is to derive a variety of innovative applications allowing OAIS-compliant digital preservation strategies to be undertaken in different domains, including Science, Cultural Heritage and Contemporary Arts.

During the 2nd year of work, the analyses of the theoretical foundations of CASPAR was pursued with a number of publications and conference communications reasserting the innovative nature of some of the digital preservation concepts and the conceptual model has been further detailed on the basis of the OAIS functional model.

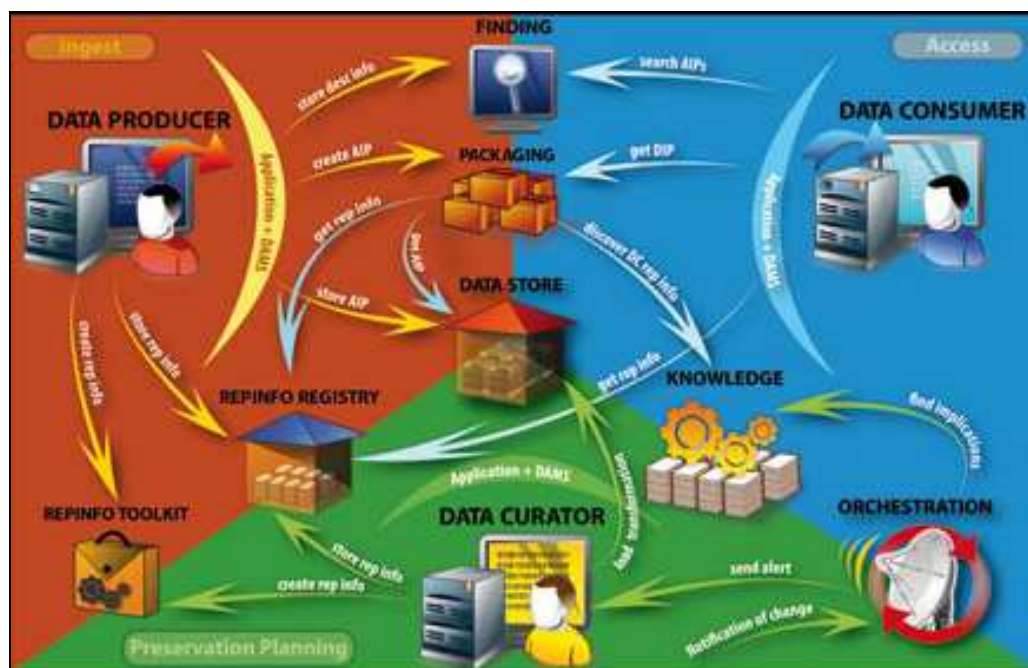


Figure 1: CASPAR Preservation Workflow

The development of the building blocks of CASPAR has also made significant progress. Based on the OAIS Functional Model, the CASPAR Architecture Team has identified 11 CASPAR Key Components as part of the 6 OAIS macro functional areas: ingestion, data management, storage, preservation planning, administration and access. These components are: registry (REG), knowledge (KM), orchestration (POM), representation information

(REPINF), preservation datastore (PDS), data access and security (DAMS), digital rights (DRM), finding aids (FIND), virtualisation (VIRT), packaging (PACK) and authenticity (AUTH).

The figure below shows how the CASPAR components are put in place into the OAIS Reference Model.

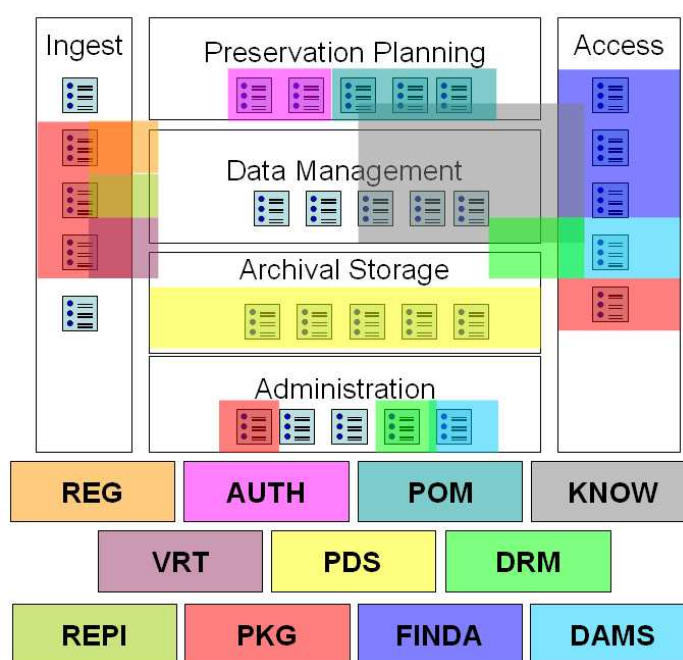


Figure 2: CASPAR Components vs OAIS Reference Model

The first release of the CASPAR software is now available and already covers the majority of the key components. A second one is in progress with a greater set of functionalities.

Each component is independent of each other and they offer web based services. That means that they can be used independently in different digital preservation systems. Moreover, a lot of effort has been spent in order to define an architecture model as independent from the technology as possible. In this way CASPAR can be an open system able to interoperate with as many different systems as possible and to be operated and re-implemented in the framework of existing preservation solutions.

CASPAR partners have also been looking beyond the CASPAR project, working with other European and global partners on a number of initiatives designed to provide a common entry point to digital preservation projects. Reference examples are:

- ICT 2008: 'I's to the Future: Lyon, France November 25-27th 2008. Europe's biggest event for research and communication technologies. CASPAR will be presented at a communications booth.
- SAMT 2008 3rd International Conference on Semantic and Digital Media Technologies. Koblenz, Germany De-

ember 3-5th 2008. CASPAR will participate in the event by hosting a poster session.

- DCC 2008: The 4th International Digital Curation Conference. Glasgow, Scotland December 1-3 2008. CASPAR partners have submitted 2 papers and will participate in a poster session.
- African Digital Scholarship & Curation 2009

In addition, CASPAR has been one of the leaders of an effort to integrate the EU funded projects, one of the latest developments being the WePreserve (www.wepreserve.eu) collabo-

orative platform highlighting activities of Digital Preservation Europe (DPE), Preservation and Long-term Access through NETWORKed Services (PLANETS) and CASPAR. This platform hosts many initiatives like conferences, forums and joint training modules in order to introduce the Principles of Digital Objects Preservation.

Join the CASPAR community to:

- Receive the latest news on CASPAR developments
- Interact with over 200 Digital Preservation Experts
- Participate at our events

For further information visit CASPAR website <http://www.casparpreserves.eu>



or the CASPAR developers community website

<http://developers.casparpreserves.eu:8080>

or contact us at

info@casparpreserves.eu

10.3. nestor

nestor - Network of Expertise in long-term **STO**rage and availability of digital **R**esources in Germany



If we wish to preserve the digital cultural heritage for subsequent generations, we need to tackle the issues of long-term accessibility of digital documents. The nestor project bundles the available long-term preservation know-how, energies and skills – forming the basis for a future alliance dedicated to preserving Germany's digital heritage.

nestor, the advisor of the Greeks in Troy, stands as symbol for the advisory and supporting function of the network of expertise in long-term archiving and long-term availability. The acronym nestor was taken from the English version of the official project name "Network of Expertise in long-term **STO**rage and long-term availability of digital **R**esources in Germany".

nestor is supported by the Federal German Ministry of Education and Research. A scientific advisory board accompanies and supports the project activities.

The nestor network was initiated

- to bring together all the parties in Germany concerned with the long-term preservation of digital documents
- to create an information and communication platform as a central point of contact for all matters regarding longterm preservation
- to provide an externally visible point of focus within Germany aimed at supporting

partnerships and serving as a first point of contact for international alliances.

The aim is to establish a cooperative infrastructure in which a wide range of specialist skills can interact to help solve the issues of digital curation.

nestor project partners include:

- Deutsche Nationalbibliothek (German National Library)
- Niedersächsische Staats- und Universitätsbibliothek Göttingen (Göttingen State and University Library)
- Computer and Media Service and University Library of the Humboldt University in Berlin
- Bayerische Staatsbibliothek (Bavarian State Library), Munich
- Fernuniversität Hagen (University of Hagen)
- Institut für Museumsforschung (Institute for Museum Research), Berlin
- Bundesarchiv (Federal German Archive), Koblenz

The group of the nestor project partners is surrounded by a broader group of institutions and individuals which are organised in working groups and working packages and which are forming the nestor network. At the moment more than 30 institutions are contributing effort to the network by cooperation with the nestor working groups. The groups include:

- WG Trusted Repositoris – Certification
- WG Media

- WG Cooperative long-term Preservation
- WG long-term Preservation Standards
- WG Grid/eScience/raw data and long-term Preservation
- WP Training & Education

nestor started its first project period in 2003 and the second project period will end in Summer 2009. At the moment the project partners are preparing a sustainable successor organisation, which will continue the work, after the end of the project founding.

Other workings of the last month include:

- The nestor WG on long-term Preservation Standards is working on a guide entitled "Into the Archive: Guide for ingesting information into the Digital Repository". This is aimed both at institutions wishing to transfer their digital objects to a repository and also at the repositories themselves. It is intended to support them in establishing structured transfer processes.
- The WG Trusted Repositories – Certification is preparing the revised second edition

of its Catalogue of Criteria for Trusted Digital Repositories (<http://edoc.hu-berlin.de/series/nestor-materialien/8/PDF/8.pdf>).

- WP Training & Education has – together with DPE and other partners - delivered several weeklong school events and is preparing more of this events for the next year.
- A revised version of 'nestor handbook – A small encyclopedia of digital preservation' (<http://nestor.sub.uni-goettingen.de/handbuch/index.php>) has been made available online to coincide with the 2008 Bibliothekartag (German Library Congress). In the future there will be further additions and changes, and a printed version of the handbook is planned on completion of the nestor project in summer 2009.
- The WG Cooperative long-term Preservation set up an "Digital Preservation and the Law Task Force" to support current digital preservation projects in Germany in legal issues.

More information about the project can be found on the nestor homepage:
<http://www.digitalpreservation.de>

Stefan Strathmann
strathmann@sub.uni-goettingen.de

10.4. SHAMAN

The first round of the FP7-ICT programme has funded three projects addressing digital preservation:

Name	Focus	Consortium
SHAMAN	Data Grid, Federated Digital Libraries, Persistent Data Archives and Multivalent Architecture. Test-beds: Documents in Memory Institutions and Governmental Collections, Objects in Industrial Design and Engineering, eScience	Universities and research organizations in EU and US, industry and government
LiWA	Web archiving: fidelity, coherence and interpretability, transforming pure snapshot into living web archive	Universities and research, new media archiving
PROTAGE	Explore software agent technologies to automate preservation processes (self-preserving objects)	Archives and universities, research, ICT

SHAMAN (Sustaining Access through Multivalent Heritage ArchiviNg) is a Large-scale Integrating Project (SHAMAN website will be available soon). More details on the project can be found at ftp://ftp.cordis.europa.eu/pub/ist/docs/digicult/shaman_en.pdf. The aim of SHAMAN is to develop the framework for the next generation of long term (more than one century) digital preservation systems and tools. It includes the definition of a SHAMAN theory of preservation integrating the analysis, ingestion, management, access to and reuse of information objects across distributed repositories. The data preservation capabilities offered will secure the authenticity and integrity of data objects through time. It has 18 academic and industrial partners from 9 European countries, and will run for four years. The project budget is €13.3m, of which about €8.4m is being contributed by the European Union.

The development work will be structured around four core components whose objectives can be described as follows:

- to establish an open distributed resource management infrastructure framework enabling grid-based resource integration (fig. 1), reflecting, refining and extending the OAIS model and taking advantage of the latest state of the art in virtualisation and distribution technologies from the fields of GRID computing, Federated Digital Libraries, and Persistent Archives;
- to develop and integrate technologies to support contextual and multivalent archival and preservation processes which are adapted and significantly extended from the fields of content and document Management and Information Systems;
- to develop and integrate technologies to support semantic constraint-based collection management to target one of the key challenges in automating one class of digital preservation core functions;

- to support the managing of future requirements by securing interoperability with future environments and maintaining essential properties of the preserved content.
- Three prototypical applications will support trialling and validation in the follow-

ing domains: 1) scientific publishing in libraries and documents in governmental (parliamentary) archives, 2) digital objects used (i.e. CAD) in industrial design and engineering and 3) data resources used in e-Science applications.

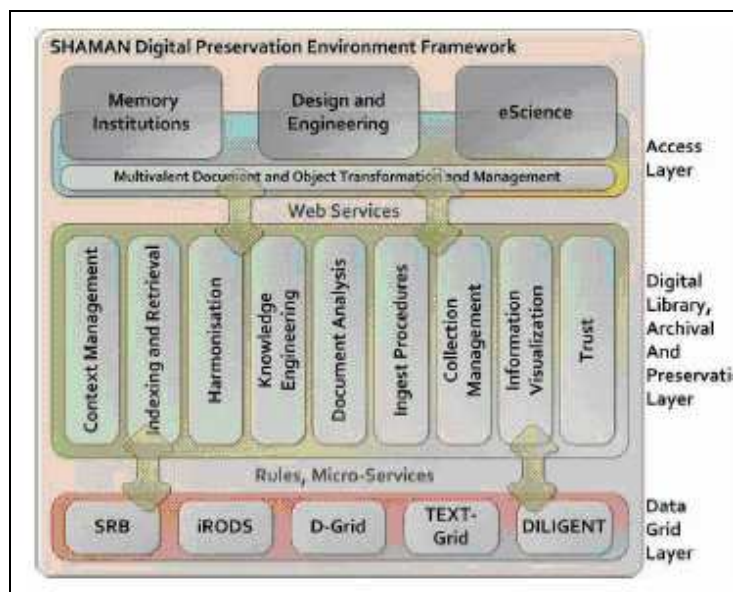


Figure 6. SHAMAN Conceptual Architecture & Grid Utilization

SHAMAN's dissemination and exploitation plans aiming at actively foster outreach and take-up of results will be tailored according to the specific needs of the scientific / academic world and of industry users. SHAMAN's work will be naturally coordinated with other digital preservation European projects (CASPAR, DPE, PLANETS <http://www.wepreserve.eu>) as well as initiatives at national (DGrid, Germany http://gks06.fzk.de/slides/D-Grid_Gentzsch.pdf) and international level (NDIIPP/NSF, US <http://www.loc.gov/today/pr/2008/08-004.html>).

Perla Innocenti
Researcher/SHAMAN Co-Principal Investigator
HATII at the University of Glasgow
p.innocenti@hatii.arts.gla.ac.uk

10.5. LiWA

The Web today plays a crucial role in our information society: it provides information and services for virtually all domains, reflects all types of events, opinions, and general development within society, science, politics, environment, business, etc. Due to the central role the World Wide Web plays in today's life, its continuous growth, and its change rate, an adequate way of Web archiving has become a cultural necessity for preserving knowledge. Consequently, a strong interest in Web archiving among library and archival organizations as well as emerging industrial services can be observed.

The intention of the LiWA project is to turn Web archives from mere Web page storages into “living Web archives”. Such living archives will be capable of: handling a variety of content types; improving the quality of their content by detecting capturing traps and filtering out irrelevant content such as Web spam; dealing with issues of temporal Web archive coherence as well as improving long-term content usability.



During the first six months the project focused on problem statement and detailed specification of requirements, as well as defining a framework which will allow seamless integration of LiWA components with the Heritrix and Hanzo crawlers. In addition, two application scenarios have been defined to demon-

strate the project results, namely “Streaming Archive” and “Social Web Archive”.

The Streaming Archive application, developed by Beeld en Geluid – Sound and Vision, will showcase the building of an audio-visual Web archive and how audio and video broadcast-related web information can be preserved in the Netherlands. It will demonstrate how a selection of Dutch broadcasting-related websites as well as of Dutch audio and video web content can be captured and archived in a web archive.

The goal of the Social Web application, developed by the National Library of the Czech Republic, the Moravian Library and Hanzo Archives, is to support archivists in the building of thematic social web archives, e.g. during natural disasters. Building thematic social web collections is a cumbersome, time-consuming and labour-intensive process due to their highly dynamic and user-centred nature. Therefore this application will demonstrate how web archives can capture the dynamics and the different types of user interaction of the social web.

The main expected outcome of the LiWA project is a set of innovative methods and services for capturing, preserving, analysing and enriching web content, which implement the vision of a Living Web Archive as described above. As a further outcome of the project, these services will lead to an integrated reference web archiving solution, based on LiWA results and open source developments and operated by the project partner European Archive as well as in an industrial-grade commercial solution developed by Hanzo Archives Ltd.

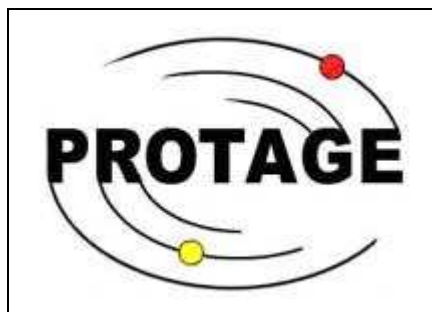
More information about the project can be found on the LiWA homepage: <http://www.liwa-project.eu>

Thomas Risse
rissi@13s.de

10.6. PROTAGE

GET BACK CONTROL AND LET THE AGENTS DO THE WORK

Many factors are involved in putting digital collections together, including copyright issues, text mark-up, metadata and retrieval standards, preservation, scanning and conversion. The vastly larger amounts of information created in a greater variety of formats than ever before makes it increasingly difficult for many organisations to identify what is of value and ensure its preservation needs over time. To meet the challenge and the growing feeling of losing control it is necessary to find new levels of automation and self-reliance in preservation solutions. You also need a permanent strategy for handling perpetual change.



The aim of the EU project PROTAGE project is to investigate and initiate complementary new approaches to digital preservation by linking digital objects to long-term digital preservation processes in a new kind of information system based on agent-based software technology which is robust, extensible and scalable. The synergy between the flexibility of agent ecosystems and the learning capabilities of smart adaptive systems will be exploited, in order to develop a hybrid smart multi-agent based architecture. The important capability of software agents is to be autonomous. Encapsulating information within agents reduces the

complexity and increases overview. Functionality is encapsulated in the tools.

The project will, by producing prototypes, build and validate flexible and extensible software agents for long-term digital preservation and access that can cooperate with and be integrated in existing and new preservation systems. The agents will not only perform tedious tasks, they can also support different kinds of activities such as monitoring preservation systems and transferring data smoothly between different repositories and institutions.

The PROTAGE consortium brings together cultural heritage institutions from the archival, library, and audio-visual sectors; academic institutions focusing on IT related research; industry partners with significant expertise in research and development within areas of intelligent agents, content management information technology, e-learning and e-publishing. A wide range of digital material will be tested within the project: web collections harvested from sites, sound and moving images from audiovisual production, databases, etc.

For individuals or institutions to be comfortable with the idea of delegating tasks to agents, they must first be able to trust them. Issues that need to be addressed include authenticity, privacy of communication, personal integrity, trust, auditing, accountability, and protection from malicious or incompetent agents. All used assessment tools and protocols within PROTAGE will, therefore, be verified by the project's Ethical Advisory Board.

The outcome of the PROTAGE project will enable memory institutions to increase their capacity and reduce costs and enable them to

meet also the challenges and intentions of the European Digital Library. One important area for action is preservation in its broader sense to

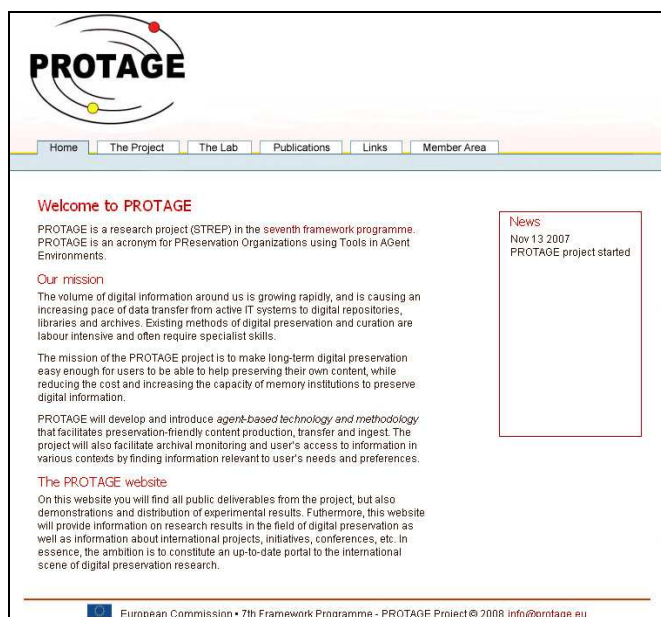
prevent valuable digital resources from being lost and to ensure that they are accessible and usable over time.

FACT BOX

PROTAGE is a three year project within the Seventh Framework Programme Theme Digital libraries and technology - enhanced learning. Grant agreement for: Small or medium-scale focused research project – STREP-CP-FP-INFISO. The projects full title is PReservation Organizations using Tools in AGent Environments

The beneficiaries are

- National Archives of Sweden – coordinator
- Luleå University of Technology (Sweden)
- National Archives of Estonia
- Fraunhofer Gesellschaft zur Förderung der angewandten Forschung e.V. (Germany)
- University of Bradford (UK)
- EASY Innova S.L.(Spain)
- Giunti Labs S.r.I (Italy).



Web site
<http://www.protage.eu>

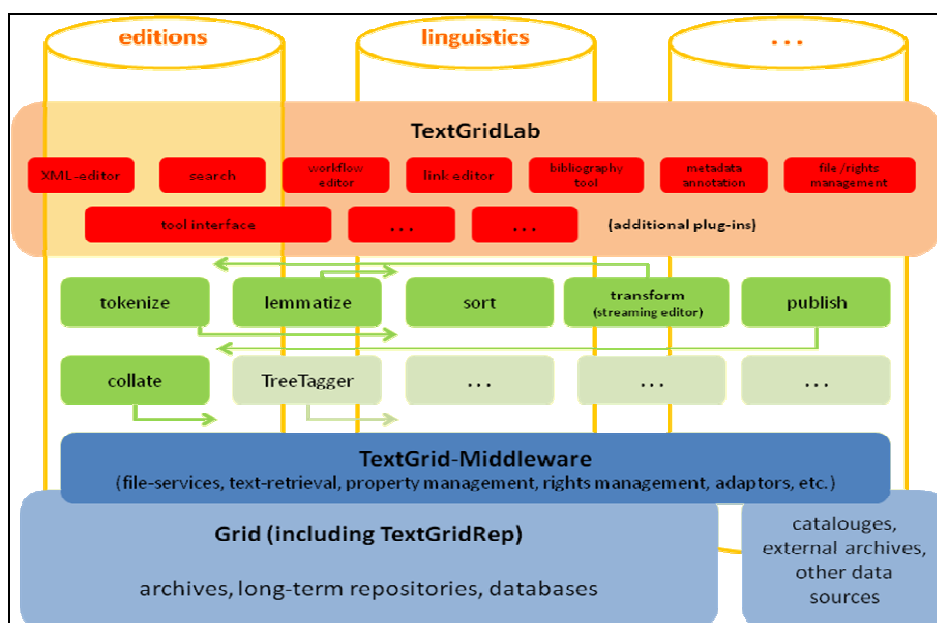
Contact
info@protage.eu

Ana Maria Durán
Ana.Duran@kb.se

10.7. TextGrid

TextGrid <http://www.textgrid.de> is the first project in the arts and humanities in Germany creating a community grid for collaborative editing, annotating, analysing, and publishing of specialist text resources. Providing a computational infrastructure, a collective network, and a comprehensive generic toolset for specialists in the arts and humanities, it is based on e-Science methods and forms a cornerstone in the emerging e-Humanities. Reaching out to the academic community, the project establishes a truly interdisciplinary platform, a virtual workbench for research and a trustworthy repository for research data. Open interfaces open the door for other projects to plug into TextGrid. Thus, any arts-and-humanities specialist can adopt TextGrid for their work. In its core functionality, however, TextGrid is, at this stage, focused on text as a data type since there is considerable demand in the community for processing

text data. In spite of modern information technology and a clear thrust towards collaboration, researchers in the arts and humanities can not currently make full use of the potentials of this development. Text scientists, for example, researching into the relations between language and discourse and into the complex processes in the genesis of literature, still mostly work in local systems and project-oriented applications. Current research initiatives also lack integration with already existing text corpora, and they remain unconnected to resources such as dictionaries, lexica, secondary literature and text processing tools. This integration and interconnection, though, bears a wealth of opportunities. With its architecture and integrated tools and services that satisfy, for the time being, the specific requirements of text sciences, TextGrid is able to provide such forms of integration.



The architecture of TextGrid is founded upon a Globus-based grid environment, enhancing it with a specific middleware layer and a service layer of specialized functionalities for textual processing including a metadata and ontology management. Additional tools can be integrated at any time. While the TextGrid middleware operates as an interface between the low-level grid and the high-level services, the service layer itself is conceived as an open web service environment that will easily elicit participation in active community processes. An Eclipse-based interactive client, TextGridLab, ties all available services and tools together into a workbench and grants intuitive access to TextGrid's data repository, TextGridRep, for content-providing users. The general public can access published contents using a web interface. The usage of standards including TEI and the XML family, RDF, SOAP, WSDL 2.0, GSI, WSRF, SAML, LDAP, and BPEL fosters openness and interoperability. A beta release of TextGridLab will be available in October 2008.

TextGrid offers infrastructural support for new, technology-informed methodologies in research at a time when these methodologies are beginning to emerge. Of course, the very utilization of grid technology with its wide potential of all kinds of collaboration is expected to transform the way scholars in the arts and humanities work, engendering new research topics and perspectives. Consequently, TextGrid can only be reasonably understood as one element within a national e-Humanities infrastructure. This infrastructure is characterized by integration, interoperability, and intercon-

nection. In this vision, it provides researchers with access to systems, services, networks and resources and enables the discovery of new resources. It guarantees the integrity, authenticity and quality of such resources and services as well as the long-term accessibility of data. An infrastructure of this kind enables scholars to collaborate, communicate and share research outputs in new dimensions. Exploiting the power of information technologies, it supports and encourages innovation and experimentation. Engagement with the industry, to date common only with natural and economic sciences, will be much more easily possible.

In order to succeed, a national e-Humanities infrastructure should be founded on a broad base of support and training structures so as to be acceptable and attractive to the community. The tracking of outputs, protection of individuals' work and privacy, management of rights as well as preservation and curation of research data are basic requirements that have to be met. Last but not least, creating an infrastructure for the arts and humanities is a political issue requiring well-planned strategies of community building. Only on the ground of reliable and operative infrastructures, a turn towards e-Humanities on a national and international level will eventually come about. In Germany, the foundations are being laid in form of an initiative exploring current e-Humanities developments, bringing together experts in the field, connecting with international initiatives like DARIAH, AHeSSC or GATE and conceptualizing a roadmap for the creation of a national e-Infrastructure for the arts and humanities.

Digital Preservation in TextGrid

One of TextGrid's key features is a digital preservation component. Long-term preservation of digital data is a fundamental requirement in all communities. Data preservation is, as it were, the memory of our culture, the foundation of sciences, arts, economy and politics. Bit-stream preservation of digital data is only part of protecting this memory. Equally important is securing permanent access to the

archived data in its original semantic context [UNESCO]. This makes digital preservation a most complex task that is, at this stage, far from being resolved. Therefore, combined efforts have to be made, even though the circumstances and specifications of data management may differ in different communities. Concepts, strategies and activities must be coordinated in order to succeed.

In spite of the fact that the immediate goals of digital preservation concern technology, all current data curation concepts are primarily organizational. While hardware and software environments are constantly changing, procedural and technological stability has to be guaranteed. Responsibilities in active data maintenance need to be clear. Standards and instructions carefully developed over the last few years by interdisciplinary research groups have to be observed.

Within the D-Grid initiative, the development of strategies for digital preservation of research data and content requires close communication among all grid partners and a transparent distribution of responsibilities. Service level agreements (SLAs) between the grid communities and resource providers need to be in accordance with data preservation requirements. Before SLAs can be defined, the communities have to plan their preservation strategies and communicate them transparently to their users. Of course, the communities should support each other and take advantage of the expertise accumulated by digital preservation expert groups (e.g. nestor, www.langzeitarchivierung.de).

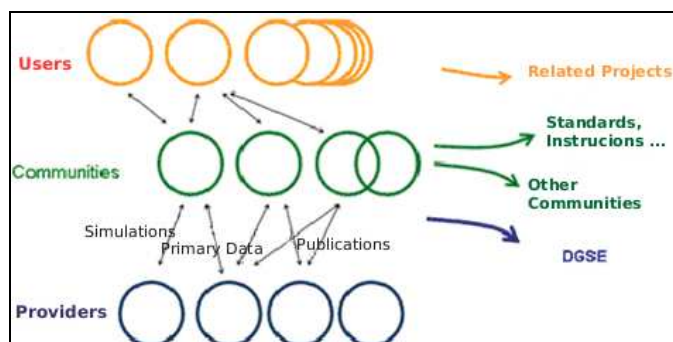
True long-term preservation of digital objects relies on systems of trusted digital repositories because only these are explicitly suitable for archiving purposes. TextGrid aims at developing a federated (i.e. distributed) system of trusted repositories that guarantees appropriate forms of data preservation in the grid. The approach being generic, any community grid can

adopt the solution for some or all of their preservation tasks.

Digital preservation systems powered by repository technologies facilitate the management of dynamic files in heterogeneous formats and a semantic referencing of objects, object parts, and object contexts. Thus, research data and content (publications, multimedia files etc.) can be better integrated in research environments and workflows.

Since TextGrid's digital preservation system will be based on a Fedora repository, the TextGrid utilities (especially TextGrid's XQuery search tool or its GridShib-based authentication and authorization module) have to be connected to Fedora. While data are exchanged via WSRF within the infrastructure, REST/SOAP-based interfaces allow open and user-friendly communication with external services.

Digital preservation services to be included in TextGrid's preservation system are format validation, conversion, metadata extraction etc. These services have to be embedded in organizational frameworks by the different communities. Workflows and responsibilities must be defined. In many communities, the very users share responsibilities in digital preservation processes because already at the creation of data both metadata and information about the context should be provided. Defining SLAs, it should be negotiated which of these services could be granted as supplementary services by the resource providers.



Agents, processes and responsibilities in digital preservation

Heike Neuroth
neuroth@mail.sub.uni-goettingen.de

11. Social networking and Digital Preservation

There is little doubt that the internet is used in a different way today than it was only a few years ago. More dynamic methods of communication, collaborative work, sharing and social-networking have influenced the behavior of all internet users. The term “web 2.0” was born as a metaphoric title for a brainstorming session by Tim O’Reilly in 2004 and is a synonym for the way people interact with the internet and with each other using the internet today. A survey by David White (*Results and analysis of the Web 2.0 services survey undertaken by the SPIRE project*, University of Oxford, 2007, <http://www.jisc.ac.uk/media/documents/programes/digitalrepositories/spiresurvey.pdf>) shows that the impact of web 2.0 applications is not limited to leisure activities or focused only on the youngest section of the total of 2 billion internet users. Businesses already recognized the web 2.0 potential for generating additional revenues, even when having to re-think customer behavior. Focusing on “long tail” consumers appears as an important part of the growing internet market.

Despite the symbols of web 2.0 (e.g. YouTube, Face Book, Wiki, or blogging) being perceived as entertainment, and therefore trivial, the role of YouTube in election campaigns of many countries documents the wider social impact of these technologies. The focus of this article is to identify how web 2.0 applications can support projects with limited, more specialized audience and how could projects like DPE benefit from it.

Clearly, we should split the question into two areas:

- how can web 2.0 applications be used for the communication in the DPE project and
- how do web 2.0 applications help DPE to address its target communities

On the project management level the impact of web 2.0 on DPE is the same as on other enterprises. Regular Skype conference calls, the use of instant messaging tools, RSS reader use, shared calendars and shared work on single documents on the internet (Google Docs, Microsoft Office Live, Zoho Office Suite etc.), the use of wiki systems for project management and information sharing, have all become part of the daily routine, and have helped reduce the financial implications of a geographically distributed team. However, this brings with it more dependency on externally provided services, on third party suppliers of often only seemingly free services. Perhaps we are too credulous in these behaviors; we submit our unfinished work, confidential information and private remarks to third party services, rarely knowing who runs them and under what motivation.

As one of the core goals of the DPE project is widening awareness about the digital preservation issues, use of social networking via web 2.0 applications comes to our mind. Here the traditional communication model, sender-medium-receiver, is no longer used. Until now DPE has influenced the “opinion leaders” in more traditional ways, addressing the target groups in conferences, journals and trainings and building specialist community networks. However, many web 2.0 applications are built upon the many-to-many communication-model. With this model, the advantages of web

2.0 applications for DPE are less obvious. DPE is only one of many submitters of presentations into slideshare.net (*image 1, image 2*) and disappears in the crowd of projects listed on the Wikipedia entries for digital curation or digital

preservation. DPE also plans to experiment with video sharing services like YouTube, exhibiting videos from events, trainings and staff presentations.

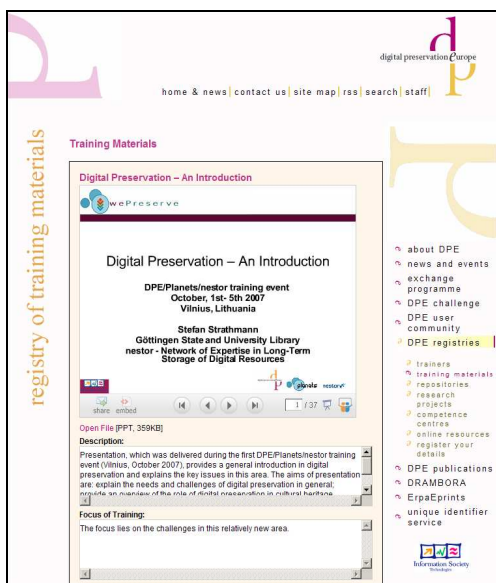


Image 1. Presentations uploaded to slideshare.net and embedded in the DPE website

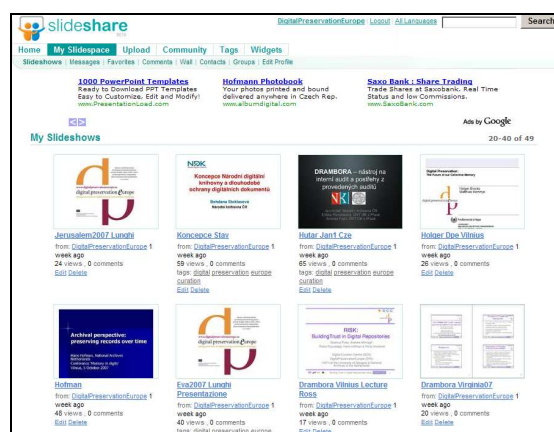


Image 2. DPE presentations on the Slideshare.net

Furthermore we could try streaming or recording of the DPE/CASPAR/Planets/Nestor training event in October and later posting the presentation on the YouTube. Unfortunately uploading content to e.g. YouTube alone attracts little or no attention. In addition to this, DPE partners and staff can communicate on Facebook and share their photos on Flickr, but will this affect the goal of the project: raising awareness on digital preservation issues? More

impact may be generated when people blog about DPE, link to the DPE website or share experiences from DPE trainings. DPE could also make more use of social bookmarking tools to promote information on its website. (<http://delicious.com> *image 3*). We have added the “bookmark and share” button to all DPE websites to allow users bookmark our websites to their bookmarks services very easily.

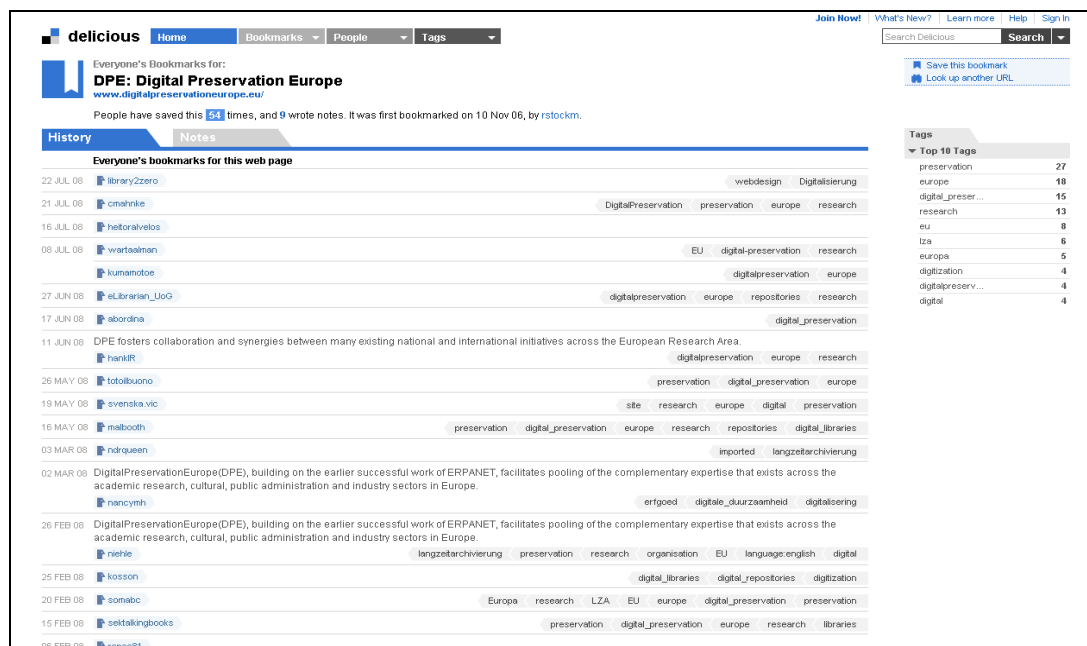


Image 3 DPE website information and reviews on delicious.com

In summary, web 2.0 applications influence communication and project management in DPE as much as in any other collaborative, international project nowadays. Experiments with web 2.0 are necessary to promote awareness of digital preservation issues to the widest possible audience. However, the success or failure of these activities depends not only on

DPE's intentions but also on the presence and activities of the digital preservation community and other target groups in network communities. The nature of the web 2.0 application communication model prevents one way distribution of information from a single authority, a many-to-many model must be adopted.

Recommendations for the future use of the web 2.0 social networking tools for the DPE project:

- Use the possibilities of slideshare.net and YouTube, embedding more presentations into the DPE web site
- Create DPE blogs in national languages and use them to promote preservation issues
- Address IT and developer communities, librarians, and the general public via web 2.0 networks with advertising campaigns raising the awareness about preservation issues
- Create web 2.0 tools aimed at digital preservation – collaborative web-based space for sharing experiences with preservation software, strategies, DOMS, formats

Types of existing web2.0 applications:

There is probably a couple of thousands of web 2.0 applications currently running in the web aimed at all types of users in many national languages (see <http://web2list.com>; http://edutechwiki.unige.ch/en/List_of_web_2.0_applications; <http://web2.ajaxprojects.com/?Title=Web+2.0> etc.). From the functionality point of view we can differentiate tools for:

- collaborative authoring (Wikipedia, writely, Everything2 etc.)
- sharing of images, videos or other content, (Youtube, Trooker, Google Video, Bolt, Singing fool, Meta cafe, tudou.com, youku.com, EbiTV, Eyevio etc.)
- blogging (Blogger, LiveJournal, WordPress, TextPattern, Blogadr, Edublogs, Classroom 2.0, Open Diary etc)
- social networking (Facebook, MySpace etc.)
- calendaring (Google Calendar, Remember the milk, yahoo calendar etc)
- communication and instant messaging (Skype, windows live messenger, Jajah, Google Talk, VoIPStunt, iChat etc)
- virtual social spaces and games (Second Life, Half Life, There etc.)
- social bookmarking (BookmarkSync, Delicious, Digg, Ma.gnolia, StumbleUpon etc.)

Jan Hutař
jan.hutar@nkp.cz

Carmen Heister
heister@ifs.tuwien.ac.at

Marek Melichar
marek.melichar@nkp.cz

Mark Guttenbrunner
guttenbrunner@ifs.tuwien.ac.at

12. Persistent identifier distributed system for Cultural Heritage digital objects

Persistent identification of Internet resources is an important issue within the life cycle approach to cultural and scientific digital library applications, preventing using the Internet as a platform for research and dissemination of scientific and cultural content, not only to identify a resource in a trustable way, but also to guarantee continuous access to it. The project presented will tackle this challenge.

It is well-known that Internet resources have a short average life; their identification and persistent location poses complex challenges affecting both technological and organizational issues, involving access and citation of cultural and scientific resources. The use of URLs can not be considered a reliable approach due to the structural instability of links (ex. domains no longer available) and related resources (relocation or updating). The current use of the URL approach increases the risk of losing cultural documents or under-using available cultural collections. In the Cultural Heritage (CH) domain it is essential not only to identify a resource but also to guarantee continuous access to it.

A trustworthy solution is to associate a Persistent Identifier (PI) to a digital resource that will remain the same regardless of where the resource is located. At present some technological solutions (e.g. DOI, ARK, Handle system, URN) have been developed but no general agreement has been reached so far among the different user communities: this scenario shows that it is viable to impose a unique PI technology. Moreover the granularity, the level of detail at which persistent identifiers need to

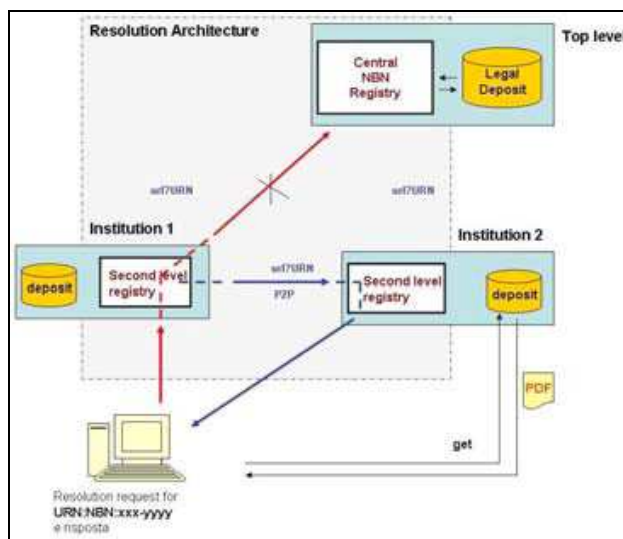
be assigned, is widely different in each user application sector.

To face these issues, a consortium led by Fondazione Rinascimento Digitale (FRD) has developed a prototype to lay the bases for a national identification/cataloguing service of digital cultural resources. The main objectives of the on-going project are the following:

- to allow easier and wider access to the digital resources produced by Italian cultural institutions, including not well-known or not yet published works;
- to encourage the adoption of long term preservation policies;
- to make costs and responsibilities for the service sustainable.

The National Bibliographic Number (NBN – RFC3188) is a namespace based on the Uniform Resource Name (URN – RFC1737) used by National Libraries. There are some important NBN implementations in Europe, for example at the National Library in Germany. Also the TEL project is in the process of implementing a unique system based on NBN namespace within the European Digital Library (EDL).

The developed prototype implements a new PI architecture: the approach is based on URN/NBN, with additional features and solutions recalling the DNS architecture. The prototype defines a hierarchical distributed system, in order to face the criticality of a centralised system and to reduce the high costs of management for a unique resolution service.



The project foresees a distributed authority and responsibility for the creation and resolution of names. The architecture identifies a central point, located at the National Library of Florence, and second-level institutions: each institution can self generate its NBN names for the resources; the central service harvests by OAI-PMH the new names, generated by the second-level registries, for validating and making them resolvable.

Each second-level node is identified by a sub-namespace expressed through the NBN name (for example NBN:IT:BNCR:xxx-xxxx for the National Library of Rome). The name resolution request can be submitted by the user to each resolution service of the second-level nodes: if the sub-namespace identifies the institution to which the request is submitted, the answer is given directly, otherwise the central registry will be invoked to redirect the resolution request to its appropriate second level node.

This architecture increases the robustness of the service and also foresees a peer-resolution between the second-level institutions, in order to maintain the resolution infrastructure opera-

tional, even if the central service is not available.

In order to avoid prohibitive management costs, the register does not include the descriptive metadata but only the administrative metadata for managing NBN name's lifecycle and an external pointer to authoritative metadata belonging to existing institutional repository. An additional feature of the system is its capacity to manage multiple URN-URLs associations, preventing the different copies of the same resources from having assigned different NBNs.

The central node allows interoperability functionalities with other namesystems, included other NBN systems, as well as the DOI system.

The expected impact is not only a great improvement in the quality of the web coverage of European cultural resources and the reduction of costs and efforts needed to maintain a stable reference of Internet resources, but also a general increase in the credibility and trustworthiness of digital libraries, by promoting the use of digital contents in different user sectors and applications.

Emanuele Bellini, Chiara Cirinnà, Maurizio Lunghi
(bellini, cirinna, lunghi)@rinascimento-digitale.it
Fondazione Rinascimento Digitale