The International Journal of Digital Curation
Issue 2, Volume 3 | 2008

3rd Annual WePreserve Conference 2008: A New Generation of Preservation Tools and Services

Emily Nimmo
Humanities Advanced Technology and Information Institute,
The University of Glasgow

Silvia Boi
University of Urbino
Italy

Summary

This is a report from the third annual WePreserve conference held in Nice, France on October 28-30, 2008. The WePreserve consortium is currently made up of three Digital Preservation projects funded by the European Commission, DigitalPreservationEurope (DPE), Preservation and Long-term Access through Networked Services (Planets) and Cultural, Artistic and Scientific knowledge for Preservation, Access and Retrieval (CASPAR), but is in the process of expansion to include other relevant projects. The theme for 2008 was 'a new generation of tools and services' and was designed to showcase the tools and services available now for use in tackling the digital preservation challenge.

The *International Journal of Digital Curation* is an international journal committed to scholarly excellence and dedicated to the advancement of digital curation across a wide range of sectors. ISSN: 1746-8256 The IJDC is published by UKOLN at the University of Bath and is a publication of the Digital Curation Centre.

Introduction

The WePreserve conference series began in 2003 and was designed to showcase the activities of the three projects DPE[1], Planets[2] and CASPAR[3]. The third conference[4] in this series was held in Nice, France, on October 28-30, 2008 with the theme "a new generation of tools and services". Emphasis was placed on practical demonstrations of preservation solutions that can be adopted and implemented today. This Conference attracted digital content owners and managers, bringing them together with the leading experts in digital preservation, to make them aware of the challenges and to present the solutions developed and identified by DPE, Planets and CASPAR, the main preservation projects co-funded by the European Commission and also Shaman[5], Driver[6], the DCC[7], and the Alliance for Permanent Access[8].

Sessions

Session One: Opening Keynote Presentations

Manuela Speiser, of the European Commission Directorate-General Information Society and Media[9], opened the conference by framing the case for digital preservation and providing a perspective on European action and research in this area. Actions at an EU level highlighted by Speiser included the 'i2010 Digital Libraries' initiative, begun in 2005, which sets out to make all Europe's cultural resources and scientific records – books journals, film, maps, photographs, music etc – accessible to all and to preserve them for future generations.[10] Also highlighted was the Commission recommendation of digitisation and online accessibility of cultural materials and digital preservation (European Commission, 2006) stipulating the creation of national strategies and quantitative and qualitative targets, including financial planning and improved conditions for digital preservation.

In addition to highlighting the work of the current research projects funded by the European Union under FP6: DPE, Planets and CASPAR, and under FP7: LiWA[11], PROTAGE[12] and SHAMAN[13], Speiser also outlined the future research challenges identified by the European Commission such as scalable systems and services, including end-to-end workflows for different types of digital resources. She also

[1] DigitalPreservationEurope (DPE) http://www.digitalpreservationeurope.eu/
[2] Planets http://www.planets-project.eu/
[3] CASPAR Digital Preservation User Community http://www.casparpreserves.eu/
[4] DPE, Planets and CASPAR Third Annual Conference: Costs, Benefits and Motivations for Digital Preservation http://www.wepreserve.eu/events/nice-2008/
[5] Sustaining Heritage Access through Multivalent ArchiviNg | SHAMAN http://shaman-ip.eu/
[6] DRIVER http://www.driver-repository.eu/
[7] Digital Curation Centre http://www.dcc.ac.uk/
[8] Alliance for Permanent Access http://www.alliancepermanentaccess.eu/
[9] CORDIS : ICT : Programme : Digital libraries and content http://cordis.europa.eu/fp7/ict/telearn-digicult/home_en.html
[10] Europe - Information Society : i2010: Digital Libraries Initiative http://ec.europa.eu/information_society/activities/digital_libraries/index_en.htm
[11] LiWA - Living Web Archives http://www.liwa-project.eu/
[12] CORDIS: The Seventh Framework Programme (FP7) http://cordis.europa.eu/fetch?CALLER=FP7_PROJ_EN&ACTION=D&DOC=1&CAT=PROJ&QUERY=011dd31cb2c9:1620:3d2d498a&RCN=85354
[13] SHAMAN http://shaman-ip.eu/

mentioned large-scale test-beds where organisations can demonstrate the feasibility of systems and services and assess their use and advanced preservation, focusing on challenging problems which cannot be handled by current models.

Following on from this, **Seamus Ross,** Director of the Humanities Advanced Technology and Information Institute (HATII) at the University of Glasgow[14], investigated how we can engage multi-disciplinary communities, encourage investment in preservation research and solution delivery, and posed difficult questions to the audience about the way we have communicated these challenges in the past. Sharing experience gained through decades of successful leadership of preservation research projects, such as ERPANET[15] and DigitalPreservationEurope (DPE), Ross attempted to answer these difficult questions. He concluded that we, the preservation community, need to approach our awareness-raising activities differently and described the innovative activities DPE is currently undertaking under his leadership to achieve this.

DPE is working to capitalise on web 2.0 technologies, exploiting viral marketing techniques, producing more online materials that are dynamic, visual and interactive and ensuring DPE has a presence in the social networking services in which our target communities are active. You can now find DPE active in Facebook, Delicious, slideshare and Wikipedia to name but a few. In addition to this, DPE is developing training videos for their successful tools DRAMBORA and PLATTER, which will ensure access to these resources irrespective of time, location or the ability to travel to training events. DPE is also developing an animation which will be disseminated via YouTube and other online video sharing sites that will help to raise awareness among the general public of the digital preservation challenges we are facing.

Session Two

Session two looked at who we are keeping data for and what they need. We heard from four individual communities of interest providing us with a fuller understanding of the range of communities we serve and kinds of data they need to preserve. **Ian Upshall** of the UK Nuclear Decommissioning Authority[16] presented the challenge facing managers of radioactive waste, **Mario Hernandez** presented the challenges faced by UNESCO[17] in preserving the cultural heritage data relating to world heritage sites, **Luigi Fusco** presented the digital preservation challenges faced by the European Space Agency[18] and **Salvatore Mele**, CERN[19], described the needs of the high-energy physics community.

Despite many differences in the individual contexts described, there were many common issues and themes evident. It was especially clear that the data generated by all these communities are increasingly varied and complex and that they are being generated at an ever-increasing and, from the point of view of those attempting to preserve the data, alarming rate. Moreover, a major theme to arise from these case studies was that there is an urgent need to incentivise and motivate researchers, either

[14] HATII, University of Glasgow http://www.hatii.arts.gla.ac.uk/
[15] ERPANET - Electronic Resource Preservation and Access Network http://www.erpanet.org/
[16] The Nuclear Decommissioning Authority http://www.nda.gov.uk/
[17] UNESCO http://portal.unesco.org/
[18] The European Space Agency (ESA) http://www.esa.int/
[19] CERN - European Organization for Nuclear Research http://public.web.cern.ch/public/

through funding or mandate, to carry out the necessary work to ensure that their data remains accessible in the long term. If data creators are uninterested or disengaged with the digital preservation issues of their data, ensuring their preservation will continue to be a struggle.

Session Three

Session three looked at what we need to keep and how can we ensure that it is kept, and we were introduced to concepts of authenticity and appraisal and selection. **Sergio Albani,** European Space Agency[20] (EO Applications Strategy Office), provided us with an insight into the requirements for accessing the ESA historical archives which have expanded considerably in the last ten years, a trend which is likely to increase in the future. Albani described the ESA's objective of ensuring the perpetual preservation of their data in coordination with institutions of its member states, its activities in developing appropriate techniques and strategies by promoting and participating with projects related to long-term data preservation and its participation in the CASPAR project, playing a role of both user and infrastructure provider.

Following on from Albani, **Jerome Barthelemy,** IRCAM[21], discussed authenticity from the point of view of musical creation using audio digital processing. Authenticity, he argued, is dependent on the point of view of the community. Developers, musical assistants and curators, he explained, each have different criteria to assess authenticity. Barthelemy also argued that the OAIS model provides the concepts needed by non-archival organizations to be effective participants in the preservation process as it provides a framework for understanding archival concepts needed for long-term digital information preservation and access.

Hans Hofman, National Archive of the Netherlands[22], brougt the session to a close by providing the conference with an archival perspective of appraisal in a preservation context. He stressed that appraisal is always biased and time-bound, it is not a once-only but an ongoing activity and is related to all processes of managing information. Hofman illustrated that archival appraisal supports decision-making about digital preservation. This included identifying criteria for preservation within a specific context, defining workflows for evaluating preservation plans, and developing methodologies for assessing the risks of applying different preservation strategies for different types of digital objects.

Session Four

In session four we looked at developing preservation infrastructures. To illustrate the need for representation information to preserve the semantics of a digital object in this context, **Esther Conway,** STFC[23], used as an example a chemical "recipe" from 1706. Conway posed a number of questions regarding the types of additional information needed to preserve in parallel, to ensure that the intended user community would be able to interpret the information successfully in the future. This was a topic we revisited by Steve Rankin in session six.

[20] The European Space Agency (ESA) http://www.esa.int/
[21] Institut de Recherche et Coordination Acoustique/Musique (IRCAM) http://www.ircam.fr/
[22] Het Nationaal Archief http://www.nationaalarchief.nl/
[23] Science and Technology Facilities Council (STFC) http://www.scitech.ac.uk/

Dirk von Suchodoletz, University of Freiburg[24], then closed the session by providing the audience with an overview of the benefits of using an emulation strategy especially in the context of dynamic digital objects. This presentation also introduced a further community of interest, one that perhaps we do not frequently consider, that of the computer gamers. He explored the importance of maintaining the look and feel or experience of the user, echoing points made by Jerome Barthelemy, and reiterated the need to preserve additional information that will help render the digital object understandable.

Session Five

Session five brought the first day of the conference to a close with a panel discussion on Sustainability and Interoperability in digital preservation featuring David Giaretta, Seamus Ross and Clive Billeness. This discussion really animated the conference delegates who provided excellent contributions reflecting on their experience, at an organisational and political level, of advocating the importance of digital preservation.

The second day of the conference provided a greater focus on the tools and services which are now available for use in digital preservation processes, bringing to the fore one of the key emphases of this conference, practical demonstrations.

Session Six

Session six aimed to provide an overview of tools and resources facilitating digital curation and preservation at an organisational level. **Steve Rankin,** STFC[25], opened the second day of the conference with a presentation on OAIS representation information and representation information networks. He focused in particular on the CASPAR and DCC tools and services which support this, such as the representation infromation registry/repository, tools for curating representation information and representation information tools training material. An important point made by Rankin was that the representation information itself needs representation information, so that it too can be interpreted and understood. However, he also explained that this potentially recursive network of representation information can be limited by the designated community to which that network applies, and the level of existing knowledge that can be reasonably expected to allow that designated community to interpret that digital object.

Following Rankin, **Colin Rosenthal** of the State and University Library (Statsbiblioteket)[26], Aarhus, provided an overview of the Planning Tool for Trusted Electronic Repositories (PLATTER) developed by DPE and complementing DRAMBORA. The PLATTER tool is designed to address how the core principles for trust in digital repositories can be incorporated into the design and planning of a repository so that it is "trust-ready" from the start. This contrasts with DRAMBORA and other tools assessing the trustworthiness of a digital repository, as they address more established organisations. The PLATTER process is centred on a group of Strategic Objective Plans (SOPs) through which a repository specifies its current objectives, targets, or key performance indicators in those areas which have been identified as central to the process of establishing trust. The intention is that the SOPs

[24] Universität Freiburg http://www.uni-freiburg.de/
[25] Science and Technology Facilities Council (STFC) http://www.scitech.ac.uk/
[26] Statsbiblioteket http://www.statsbiblioteket.dk/

should be living documents which evolve with the repository, and PLATTER therefore defines a planning cycle through which the SOPs can develop symbiotically with the repository organisation.

Hans Hofman then returned to introduce the audience to the recently released Digital Repository Audit Method Based on Risk Assessment (DRAMBORA) interactive toolkit. This is a freely available online version of the DRAMBROA toolkit, making the self-audit process easier and more efficient for repository managers. Hofman explained the DRAMBORA approach which characterises digital curation as a risk management activity which enables an auditor to establish the organisational context and goals of a repository and assess how it is achieving these goals.

Manfred Thaller, of the University at Cologne[27], then outlined the work the Planets project has undertaken to understand the properties of digital objects which are significant to their long-term accessibility. He also spoke of work to develop methods for expressing and measuring these properties to inform preservation planning, to validate the results of preservation actions, to support the needs of user communities, and to promote the development of object types with 'preservation-friendly' characteristics. **Luigi Briguglio,** Engineering Ingegneria Informatica SpA[28], rounded up the session by providing a hands-on demonstration of the CASPAR web desktop[29].

Session Seven

Session seven provided an overview of the current landscape with regard to shared/distributed infrastructures and **David Giaretta,** STFC[30], opened the session by providing an overview of the PARSE.insight project and the PARSE.insight Infrastructure Roadmap. PARSE.insight have used this roadmap to structure a survey to gain insight into science in Europe to establish what researchers think about access and storage of research data.[31] Following on from this, **Dale Peters,** Göttingen State and University Library[32], looked at how the DRIVER Infrastructure could support the long-term preservation of enhanced publications, which she defined as a compound digital object that may consist of various heterogeneous, but related web resources.

Luigi Briguglio then returned to show how CASPAR is addressing preservation with an OAIS-based infrastructure. He highlighted that what differs about the CASPAR Infrastructure is that each "CASPAR Key Component" provides functionality independently from the environment conditions, allowing it to meet the conditions of the "preservable equation" as he described it, which are: self-contained, well described, adaptable, replaceable and preservable. **Ross King,** Austrian Research Centers GmbH[33], brought session seven to a close by giving us an insight into an alternative and complementary approach taken in the Planets Interoperability Framework. Planets is a preservation architecture based on Web Services which supports interoperability and a distributed environment and the Planets service oriented

[27] Universität zu Köln http://www.uni-koeln.de/
[28] Engineering | Ingegneria Informatica - it http://www.eng.it/
[29] The CASPAR Developer Community http://developers.casparpreserves.eu:8080
[30] Science and Technology Facilities Council (STFC) http://www.scitech.ac.uk/
[31] PARSE.Insight Permanent Access to the Records of Science in Europe http://www.parse-insight.eu/
[32] Niedersächsische Staats- und Universitätsbibliothek Göttingen http://www.sub.uni-goettingen.de/
[33] Austrian Research Centers http://www.arcs.ac.at/

architecture has been extended with Grid Services, cloud and virtual infrastructures providing a powerful solution for obtaining on-demand access to computational resources.

Session Eight

The Final Session of the conference looked at putting evidence at the heart of curation and preservation and focused on demonstrations of the testbeds developed in Planets, CASPAR and SHAMAN. The three demonstrations given by **Steve Rankin,** CASPAR, **Brian Aitken**, Planets, and **Matthias Hemmje,** SHAMAN, not only showed how one can test the efficacy of a preservation action and its level of appropriateness within an organisational context, but also illustrated the way in which each of the projects showcased within this conference are working together, providing alternative, imaginative and complementary solutions to what is an extraordinarily complex problem.

Conclusion

This conference highlighted the progress that has been made by the projects DPE, Planets and CASPAR, under the leadership of the European Commission Directorate-General Information Society and Media in the years since the beginning of this conference series. It is clear that we have moved on a great deal from building an "agitating buzz", as Seamus Ross put it in his opening address, to delivering real solutions. It was also positive to see that the key recommendations for further research outlined in the DPE Research Roadmap (DigitalPreservationEurope [DPE], 2007b) have been taken on board by the community and that scientific experimentation has been put squarely at the heart of digital preservation research. The key note address given by Manuela Speiser highlighting the previous successes and future research priorities of the European Commission showed that the European Commission continues to be crucial in creating a rich research and development environment which has fostered the developments showcased at this event and will go on to generate the additional solutions needed. It is also clear from their contributions to this conference that newly funded projects such as SHAMAN will be instrumental in taking these achievements even further.

Feedback from the participants of this conference was very positive, and when asked what they would be able to take from this event to their own organisation, participants answered that the event was "thought-provoking" and "a nice round-up of various approaches that was inspiring" and that they will "be better able to inform colleagues about preservation work, what they can do better and how important it is to preserve digital data". Indeed the high level of participation from audience members and interaction between delegates and speakers throughout the event can be said to be key to the success of the event.

One salient point from the floor was that competence centres, as recommended in the DPE report on the Design, Value and Impact of Competence Centres (DPE, 2007a), are likely to emerge as increasingly important in this field as they can provide and communicate a competitive, certified, sustainable, trusted and contextualised service though the use of tools such as DRAMBORA[34] and PLATTER[35].

[34] DRAMBORA Interactive http://www.repositoryaudit.eu/
[35] DPE Publications: Planning Tool for Trusted Electronic Repositories (PLATTER) http://www.digitalpreservationeurope.eu/platter/

You can find out more about this conference, view all the presentations discussed here and find out about upcoming events by visiting the WePreserve website at http://www.wepreserve.eu/events/nice-2008/

References

DigitalPreservationEurope. (2007a). *Competence centres: State of the art review*. Deliverable 5.1: Report on the Design, Value and Impact of Competence Centres. March 30, 2007. Retrieved November 18, 2008, from http://www.digitalpreservationeurope.eu/publications/reports/competence_centre_SoAR_1.pdf

DigitalPreservationEurope. (2007b). *Research roadmap.* October 31, 2007. Retrieved November 18, 2008, from http://www.digitalpreservationeurope.eu/publications/reports/dpe_research_roadmap_D72.pdf

European Commission. (2006). *Commission recommendation on the digitisation and online accessibility of cultural material and digital preservation*. (2006/585/EC) August 24, 2006. Retrieved November 18, 2008, from http://eur-lex.europa.eu/LexUriServ/site/en/oj/2006/l_236/l_23620060831en00280030.pdf