DigiCULT.Info

Issue 4 A Newsletter on Digital Culture ISSN 1609-3941 August 2003

Information Society Technologies

INTRODUCTION

In this the fourth issue of the DigiCULT.info Newsletter not only do we have the opportunity to bring you some exciting articles but we have begun our move towards integrating more media types within the Newsletter. In this instance we have incorporated audio into an article. This issue includes information about forthcoming events, reports on emerging standards and tools, and examines some recent work in digital preservation, the digitisation of audio archives, and 3D imaging. We also note some developments in RFID applications since the last Technology Watch Report was published in March 2003.

Continued next page

DigiCULT - Engaging the Sector

LINK Thematic Issue 3 now available
LINK Fourth Forum held in The Hague: Learning Objects for the Cultural and Scientific Heritage Sector
LINK DigiCULT Technology Watch Briefings online for community comment

Reports on Culture & Technology Events

LINK European Museums' Information Institute (EMII) public meeting

News in the Spotlight

LINK Business Models of News Websites
LINK Calls for Papers
LINK Archives Hub
LINK US National Plan for preserving digital information
LINK Report on preserving e-mail
LINK Recent study shows growth in educational links for US museums
LINK Free sheet music for library users
LINK Economic and Social Data Service (ESDS)
LINK "What's New in Digital Preservation"

Challenges / Strategic Issues / New Initiatives

LINK Overcoming the dangers of technological obsolescence: Rescuing the BBC Domesday project
LINK Techniques and Challenges of emulating obsolete software. Interview with Paul Wheatley
LINK Bringing Ancient Greek theatre to life with interactive archaeology
LINK Archaeology museums & 3D in the 21st Century
LINK LESTER (Learning Science and Technology Repository)
LINK The Digital Object Identifier system explained
LINK The YLE Digital Sound Archive

For subscription to the DigiCULT.Info please go to:
http://www.digicult.info

In our first Technology Watch Report (2003) we examined the potential of virtual reality in the cultural heritage sector. Since its publication the ORION project, to develop a Research Roadmap and Network on 3D for Archaeology Museums, has reported. This newsletter includes an overview of the results of their work, which are more fully available at the ORION website (http:// www.orion-net.org). The study of user practices and needs uncovered higher than expected take-up of 3D applications as well as a very strong interest in future uses. There is widespread recognition that 3D visualisations enable learning and open up learning to wider social groups. Many european museums with archaeological collections already use 3D and they have found a place in cultural scholarship as well. This theme is also taken up in a short piece about the work that Richard Beacham and his colleagues at the University of Warwick have done to demonstrate the power of virtual reality technologies to enable investigations of ancient theatres. Their work has shown that 3D reconstructions make it feasible to uncover new insights into the uses and cultural meanings of the buildings.

Among the other shorter pieces is a discussion of the Digital Object Identifier System (DOIs), which explains how DOIs can help to manage intellectual property in the digital environment. They are persistent and extensible and can be used to facilitate access to individual digital objects and collections (e.g., e-books).

In Europe the Network of European Museum Organisations (http://www.ne-mo.org) (NEMO) collects and makes accessible 'information to museums on relevant EU initiatives, key legislative policy and funding concerning the cultural heritage'. They press the message that museums make a major contribution to not just the cultural life of European citizens, but also have a tremendous impact on the economy which stretches well beyond the some 250,000 jobs in the European Museum sector. There is a great disparity in the numbers of museums per 100,000 inhabitants across European countries according to the Guide to European Museum Statistics (Institut für Museumkunde, Berlin 2003). For instance in Norway there are 11.4, Italy 6.3, Netherlands 5.8, and the Slovak Republic and Spain each have 1.9. Several of the editors of this newsletter have commented on new museum developments in Spain such as the Museum of Science in the Cuidad de las Artes y les Ciencies (http://www.cac.es) in Valencia and the Museo Arqueológico Provincial de Alicante (MARQ) (http://www.marqalicante.com/) which are at the forefront of museum design. The MARQ is most notable for how it has developed a dynamic space which uses a diversity of media, both physical and virtual, to introduce visitors to the archaeology of the region and to the processes of archaeology itself. In both instances there is a focus on museums as learning environments and technology plas a role in enabling this. We have noted several reports that have been published during the past few months among them the Institute of Museum and Library Services (IMLS) study True Needs, True Partners 2002, which found that the links with and commitment to education by US museums has greatly increased since 1994 and that museums were activity promoting more engaging ways for children to learn.

Lisa Spiro, Director of the e-Text Centre at Rice University, describes a project she is leading to create a Learning Science and Technology Repository (LESTER). It provides a gateway to events, discussions and information about the development of learning technologies. The major aim of LESTER is to enable collaboration in the LST communities and to promote awareness of high

quality research. Since the last newsletter the DigiCULT team has been joined by Daisy Abbott (picture left) who completed both her undergraduate studies in Theatre Film and Television and a postgraduate degree in information technology at the University of Glasgow. As well as contributing to the overall work of DigiCULT Daisy will be managing the production of newsletter and occasionally contributing articles of her own to it. Among the first of these is her piece on the rescuing the BBC Domesday Project. BBC Domesday aimed to create a modern-day equivalent of the record that William I created when he sent out recorders in 1086 to every holding in England. While the 900 year old volume remains accessible today, the data in its 20th century equivalent had been rendered inaccessible as a result of rapid technology change. Two different approaches were taken to rescue the BBC Domesday Project, CAMiLEON developed an emulator, and LongLife Data Ltd rescued and represented the data on a newer system. The work of both groups was made more difficult by the lack of access to the original system itself and to information about how it worked. Some of the BBC Domesday information can now be viewed by the public on the LongLife Data System at the United Kingdom's National Archives (Kew).

This article is accompanied by an interview with Paul Wheatley from the CAMiLEON project at the University of Leeds Library who spoke to DigiCULT.info about digital preservation and in particular the techniques and challenges of emulating obsolete software.

Experts from Finland's national public service broadcasting company, YLE, have contributed an article

on the steps that it is taking to improve access to and preservation of its sound archives through digitisation. The archive recognises that while it is essential to capture the analogue material before the tapes deteriorate YLE realises they are creating a new problem because the digital media themselves are not that stable. In addition the collection is so extensive that YLE has to prioritise material. The article draws attention to some best practices such as the importance of transferring all the information on the analogue media to the digital environment and only carrying out restoration on copies of the digital material (a conclusion that work in the area of digital images has also demonstrated. This article also represents a significant departure for DigiCULT.info because it is the first time that we have been able to include audio in this Newsletter.

The DigiCULT Technology Watch Report released in March 2003 examined the role of RFID (Radio Frequency Identification) tags in the cultural heritage sector. Two difficulties with RFID tags remain the costs of the tags and distances at which they can be read. Recent reports continue to show their increasing prevalence in the commercial sector and a growing diversity in the types of applications for which they are used.

Among the TWR case studies on the use of RFID tags was included the Saïd Business School (University of Oxford) Library. Since then Intellident, the company that did the work at SBS has completed the tagging of 280,000 items at the Colchester Public Library and more than 450,000 items at the Nottingham Trent University Library. One of the problems remains the high cost. In large sales such as the purchase by Gillette of more than 500 million tags it appears to have been possible to move the cost of the tags down towards about 10 US cents a tag, but in general they tend to

remain between 50 cents and $1 each. Although in reviewing proposals by Benetton to include tags in the clothing it manufactured, the RFID Journal reported that 'most analysts say it makes sense to track items that cost more than $15 with RFID tags that cost 50 cents' (23 June 2003, News Section of the RFID Journal, http://www.rfidjournal.com/article/articleview/471/1/1/). Computerworld reported that Delta Airlines in its pilot project to use tags to manage passenger baggage believes that the costs of tags will fall significantly over the coming year (18 June 2003, http://www.ti.com/tiris/docs/news/in_the_news/2003/6-18-03.shtml). Indeed the European Central Bank is examining ways that by 2005 it can in clude RFID chips in euro bank notes (http://www.eetimes.com/story/OEG20011219S0016), currently this could only be cost effectively done for high denomination notes.

As the costs of the tags continue to fall and the technologies for collection information from them become more sophisticated it will be feasible to use them for increasing imaginative applications. For example, museums might use visitor RFID cards in combination with geographical information systems (GIS) and CCTV recordings to distil moving image footage of individual visits—'automatically generated museum visitor video ethnographies'. New technological applications would be needed to automating the linking the RFID time codes giving the position of the target at any point in them with the time codes on CCTV recordings to extract automatically the moving image sequences related to each visitor. Such usage might raise privacy issues, as Benetton discovered when news leaked earlier this year (2003) that they were considering putting RFID tags into their Sisley clothing line.

We continue to make progress on the next technology watch report. Readers who wish to comment on the segments of the report as they are drafted are reminded (and indeed encouraged) to do so on the DigiCULT website. The 2004 Technology Watch Report will focus on the XML family of technologies, Application Service Models, collaboration and virtual communities, mobile access to cultural institutional information resources, and cultural agents and avatars. If readers wish to recommend projects, institutions, or other activities that would make good case studies in any of these areas please contact us.

We are looking at new ways to improve the visibility and take up of the DigiCULT products. We are pleased that Salzburg Research's Head of Corporate Communications, Birgit Retsch, who holds a degree in Communication Science from the Paris Lodron University of Salzburg, has agreed to lend her expertise to DigiCULT's promotion of its activities and publications.

Seamus Ross & John Pereira
Editors, DigiCULT.info

DigiCULT Forum 4 in the Hague: "Learning Objects for the Cultural and Scientific Heritage Sector"

Nestling among the documented collections of Dutch history, language and culture of the National Library of the Netherlands, DigiCULT found a natural home for the fourth in its series of seven Fora, to concentrate the minds of the gathered experts on the pressing needs of tackling "Learning Objects" within the Cultural and Scientific Heritage Sector.

The opening presentation set out a conceptual framework for understanding learning objects in the wider context of enabling teaching and learning with technology. It was, however, quickly pointed out that technology should be seen as an enabler and not as the driver in the take-up of learning strategies in the cultural heritage sector. There was general agreement that the cultural heritage sector needed to encourage convergence and improve consensus building to overcome the issue of scale in the sector, whereas on the technology side the day was dominated by key words such as 'reusability' and 'durability'.

We invite you to return in September 2003 for the complete coverage of this Forum with the release of DigiCULT Thematic Issue 4.

Thematic Issue 3: "Towards a Semantic Web for Heritage Resources"

DigiCULT Thematic Issue 3 is now available for downloading from the DigiCULT Website: http://www.digicult.info/pages/publications.php. It addresses the questions: What is the Semantic Web? What will it do for heritage institutions? And what is the role of certain languages, in particular XML and RDF? In an effort to raise the veil of mystery surrounding the Semantic Web, this issue includes an example from the sector on the implementation of semantic interoperability of metadata, combined with a primer that explains core building blocks such as XML, RDF and ontologies.

DigiCULT Technology Watch Briefings Online

The first two in our current series of Technology Watch Briefings are now online and available for public comment: "The XML Family of Technologies" and "Technologies and New Socio-economic Business Models", which covers areas such as Digital Rights Management (DRM), online selling, payment methods and issues of sustainability. http://www.digicult.info/pages/publications.php

Business models of news Websites

The online journal First Monday (http://www.firstmonday.dk) has published a new report: "Business models of news Web sites: A survey of empirical trends and expert opinion" by Frederick Schiff. The report outlines eight business models in order to assess the commercial trends of news Websites in a competitive marketplace and to comparatively assess different approaches to online news provision. You can view the article at: http://www.firstmonday.dk/issues/issue8_6/schiff/index.html

Connect to the DigiCULT Events

DATABASE FOR A COMPREHENSIVE LIST OF UPCOMING EVENTS AFFECTING THE CULTURAL HERITAGE SECTOR.

HTTP://WWW.DIGICULT.INFO/PAGES/EVENTS.PHP

European Museums' Information Institute (EMII) public meeting

On Friday 20 June 2004, around 80 representatives of Europe's leading cultural organisations met at the Tate Modern on London's South Bank to participate in an EMII-DCF (http://www.emii-dcf.org/) joint discussion entitled "A Joint Solution to Managing Cultural Digital Assets". The afternoon's lynchpin was Louise Smith, Director of the meeting organiser mda Europe (http://www.mda.org.uk/), who outlined the afternoon's work and provided a quick overview of the European Museums' Information Institute's Distributed Content Framework (DCF) project. The project is aimed at guiding cultural heritage organisations in formulating best practice strategies for dealing with the dissemination of their valuable content.

The afternoon's first invited speaker was Sandy Nairne, Director of the National Portrait Gallery, London, who provided a broad overview of the issues involved in the management of digital assets, and in particular the challenges faced by content provider organisations. The conflict between the urge to present content and the need to protect it from misuse was a key theme, and Nairne proposed that the widespread attitude that Web content is somehow 'commonly owned' be questioned and realigned.

Next to speak was John Wyver, Chairman of Illuminations Television Ltd, who provided a different perspective on the subject: that of the content user. According to Wyver, many cultural heritage professionals are unprepared to deal with increasingly complicated rights issues, and fail to understand fully the needs of consumers.

A counterblast to the two preceding talks was offered by Michael Spearman of the National Museums of Scotland's Multimedia Team, who put forward a passionate and provocative argument against recourse to copyright, and what he sees as "our increasing obsession with very small amounts of money". Spearman was followed by Sharon Page of Goldsmiths College, who gave a detailed legal analysis of current copyright and other relevant legislation, as well as stressing the implications of increasingly complex multimedia objects.

After a short break for coffee and conversation, the real meat and bones of the afternoon was the group discussion session, where the delegates were grouped into threes and fours and asked to brainstorm what the EMII Distributed Content Framework would mean for them. Delegates from organisations as diverse as museums, libraries, governments and academia, and from countries including Ireland, Portugal, Greece and the Netherlands, explained what digital rights management meant for them personally, and what they thought of it. The use of open standards was a common theme, as was the importance of user-focused strategies. The mda team will certainly have had much to ponder in the weeks following the event, as different organisational viewpoints produced a wide range of responses to the increasing use of rights management software.

The event concluded sophisticatedly at 5 o'clock with a champagne reception in the glamorous surroundings of the Tate's topmost floor. Delegates continued conversations begun in the participatory session and struck up new ones, while at the same time enjoying panoramic views of the famous London skyline.

(A case study on the EMII-DCF is currently in preparation, and will be included in the next DigiCULT Technology Watch Report, due to be published in February 2004).

Call for Papers

NG2I 2003, International Workshop on Next Generation Geospatial Information
Cambridge (Boston), Massachusetts, USA
19-21 October 2003
http://dipa.spatial.maine.edu/NG2I03/
The scope of this intensive workshop is to address research advancements in Digital Image Processing & Analysis, Geographic Information Systems, and SpatioTemporal Databases that are relevant to geospatial information. Its main objective is to bring together specialists from these overlapping, but not necessarily interacting, scientific communities and provide a high-quality forum for the presentation and discussion of related research activities.

High-quality, original contributions are invited. Topics of interest for potential paper submissions include (but are not limited to) the following:
- Automated methods of geospatial information extraction and change detection using digital imagery;
- Image sequence analysis and video processing for dynamic events;
- Spatiotemporal knowledge management;
- Distributed computing for geospatial applications;
- Semantic and geometric integration of heterogeneous spatial information and sources;
- Spatial and spatiotemporal data mining, queries, and content-based information retrieval methods;
- Digital libraries and Web-based geospatial environments;
- Mobile computing and location-based services;
- Handling data quality and uncertainty in geospatial information, including modelling, propagation, visualisation and communication.

If you are interested in submitting a paper, please e-mail: questionsNG2I@spatial.maine.edu for information on deadlines and formats.

2004 ACM SIGMOD International Conference on Management of Data
Maison de la Chimie, Paris, France 13-18 June 2004
http://www.sciences.univ-nantes.fr/irin/SIGMODPODS04/
The annual ACM SIGMOD conference is a leading forum for database researchers, practitioners, developers and users to explore cutting-edge ideas and to exchange techniques, tools and experiences.

Submissions are invited for original research contributions and industrial papers, as well as proposals for panels, tutorials and demonstrations. Work on topics of emerging interest in the research and development communities is particularly encouraged.

Important areas of interest include (but are not limited to) the following:
- organising and querying information on the Web;
- integration of text, data, code and streams;
- large-scale information fusion; peer-to-peer data management;
- knowledge discovery;
- personalised information systems;
- data privacy and trustworthy systems;
- new user interfaces;
- embedded and self-organising databases.

For more information, visit:
http://www.sciences.univnantes.fr/irin/SIGMODPODS04/SIGMODcfp.html

Contact the People Behind DigiCULT Forum

Salzburg Research
Project co-ordinator
John Pereira
john.pereira@salzburgresearch.at
Guntram Geser
guntram.geser@salzburgresearch.at

HATII - Humanities Advanced Technology and Information Institute, University of Glasgow
Seamus Ross
s.ross@hatii.arts.gla.ac.uk
Martin Donnelly
m.donnelly@hatii.arts.gla.ac.uk
Daisy Abbott
d.abbott@hatii.arts.gla.ac.uk

Overcoming the dangers of technological obsolescence: Rescuing the BBC Domesday project

By Daisy Abbott

The ill-fated BBC Domesday project of 1984–86 is often cited by archivists as an example of the dangers of technological impermanence. The format on which BBC Domesday was stored, LV-ROM videodisc, was quickly superseded and manufacturers soon stopped supporting the technology, rendering this £2.5 million project practically worthless. For the project directors this was bitterly ironic, given the fact that their intended aim was to create a modern-day equivalent of the Domesday Book, a record that has survived for more than 900 years. However, access to this previously 'lost' information is now possible once more due to systems that allow the data to be understood by today's computers. A working version of BBC Domesday is now available for public use at the British National Archives at Kew.

THE DOMESDAY BOOK

In December 1085, William the Conqueror commissioned a huge survey of the lands and people under his control. It is thought that a primary reason for William ordering that this record be produced was to ascertain how much tax he was collecting from the country, money which was often used to buy off Scandinavian armies, the greatest threat to his power.

The survey is a detailed statement of the land, its regions, people and resources, the first draft containing records for nearly 13,500 separate English settlements. The scale of the research is apparent from a quote of the time:

> *"there was no single hide nor a yard of land, nor indeed one ox nor one cow nor one pig which was there left out: and all these records were brought to [the King] afterwards."*

This vast survey, completed in less than two years, became known as the Domesday Book.

ABOUT BBC DOMESDAY

The BBC Domesday project was carried out between 1984 and 1986 to celebrate the 900th anniversary of the Domesday Book, and was organised on a massive scale reflecting that of the original survey. Around 60 BBC staff, 14,000 schools, 2000 other groups and 1 million children and other volunteers were involved in collecting, analysing, recording and disseminating geographical and social data about small areas of the UK. Over 24,000 maps and 200,000 photographs were processed, along with moving images, articles and over 8000 data sets (traffic congestion, radiation levels etc.) to provide a 'snapshot' of life in Britain in the mid-1980s. It has been estimated that, at a rate of 40 hours per week, it would take over 7 years to view all of the information of the BBC Domesday project. [1]

The information was recorded onto two 12" videodiscs that could be played using a BBC Master computer (fitted with a second processor and an SCSI card) connected to a special new type of videodisc player called an LV-ROM player. Much of the technology was developed specifically for this project by a collaboration of BBC, Acorn, Philips and Logica. In 1986, a full BBC Domesday system was presented to the Keeper of Records at the Public Record Office, to be kept alongside the original Domesday Book.

LV-ROMs have now been superseded by CD-ROM and DVD technology and almost all existing LV-ROM players are no longer functional. Less than 20 years later, the information contained in this innovative £2.5 million survey had become virtually inaccessible.

RESCUING THE DATA

Finding ways to understand the obsolete 'digital language' of the data and BBC Micro software stored on the disc can be likened to having to learn Latin to read the original Domesday Book. There are different approaches to rendering data in an understandable way, each with advantages and disadvantages. 'Emulation' preserves not only the data but also the software and the original interface, and simply makes it readable by modern computers. Presenting the information in its original context helps to preserve the authenticity of the original data and furthermore allows study of the original interface itself – the ways in which information is presented may also have historical significance. Another approach is to 'interpret' the original data, that is, to represent the underlying information within a new interface. This has the advantage of using the benefits of the most up-to-date user interfaces, which can provide easier and more efficient access to information.

For the BBC Domesday project, Creative Archiving at Michigan and Leeds Emulating the Old on the New (CAMiLEON, http://www.si.umich.edu/CAMILEON/) developed an emulation of the original Domesday system hardware, which included the co-processor, SCSI

[1] http://www.iconbar.com/news/features/camileon.html

communication and the many functions of the laserdisc player. In other words, knowledge of how the original system worked is encapsulated in the emulation software. Together with the abstracted data of the BBC Domesday project, this software provides a record of the original system and gives an accurate reproduction of almost all the original functionality.

In a separate project, Adrian Pearce from LongLife Data Ltd designed a new application that allows new views of the BBC Domesday data. For example, the interface can show a map of an area and its accompanying text in the same window, which was not possible on the lower resolution screens of the original BBC computers. Adrian worked to understand and partially rebuild the original hardware and collaborated with Jeffrey Darlington from The National Archives (previously the Public Records Office) and Andy Finney of ATSF (http://www.atsf.co.uk) – one of the original BBC Domesday team – to incorporate digitised versions of the original analogue film tapes held by the BBC.

The CAMiLEON team managed to obtain access to a semi-working Domesday system (donated by the School of Geography at the University of Leeds) and transferred data from the laserdiscs as a byte-stream readable by modern hardware. This information totalled around 70 Gigabytes per disc side. Both of these systems employ the original byte-streams and formats (e.g. data structures) and fulfil different needs; some users require new ways to access the data, while others will want to see it in its original context.

In making data accessible it was, of course, imperative not to fall into the same trap as in 1986 and to understand that a crucial part of digital preservation is not simply maintaining a copy of the data but maintaining a means of accessing that data. Current media such as CD-ROMs are likely to be just as short-lived as the rapidly changing media of the past: paper tape, punched cards, magnetic tape (e.g. QIC 24), 8" floppies, to name but a few examples. Therefore, to avoid obsolescence of both data and emulation software in turn, the CAMiLEON team emphasised media-neutral, platform-independent formats crucial for accurate and efficient preservation over time.

A Windows PC version of the Domesday project, provisionally known as '1986 Domesday Community', presents the contents of one of the Domesday laserdiscs and is now available for visitors to use in the public reading rooms at the National Archives at Kew, Surrey (http://www.nationalarchives.gov.uk/).

CAMiLEON are currently seeking funding to develop their demonstrator into a system to serve Domesday to the public (subject to IPR permissions).

The BBC Domesday project encapsulates many difficult issues relating to the preservation of an interactive system: a huge amount of multimedia data, technological complexities, the IPR issues and the need for data longevity so that it can provide a useful historical and sociological resource for the future. William the Conqueror's Domesday Book is an invaluable document for understanding the past and it must be ensured that digital archives in both linear and interactive formats can ideally also be accessed 900 years from now.

The Domesday Book Online can be found at: http://www.domesday-book.co.uk/ Further information about BBC Domesday: http://www.pro.gov.uk/about/ preservation/digital/domesday/default.htm Andy Finney's (ATSF) report on Domesday: http://www.atsf.co.uk/dottext/domesday.html

PAUL WHEATLEY FROM THE CAMILEON PROJECT AT THE UNIVERSITY OF LEEDS LIBRARY SPOKE TO DIGICULT.INFO ABOUT THE TECHNIQUES AND CHALLENGES OF EMULATING OBSOLETE SOFTWARE.

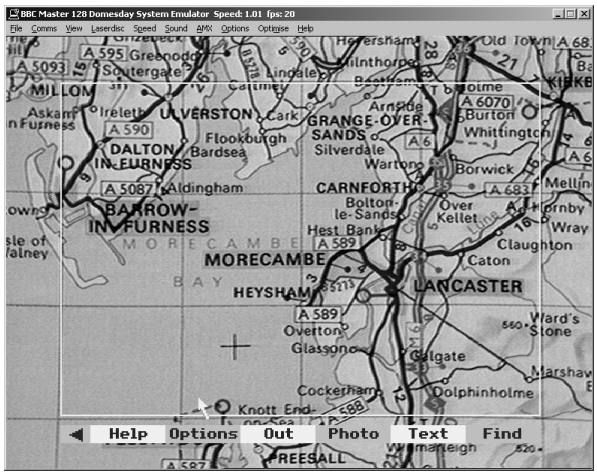
What led to the CAMiLEON project being set up?

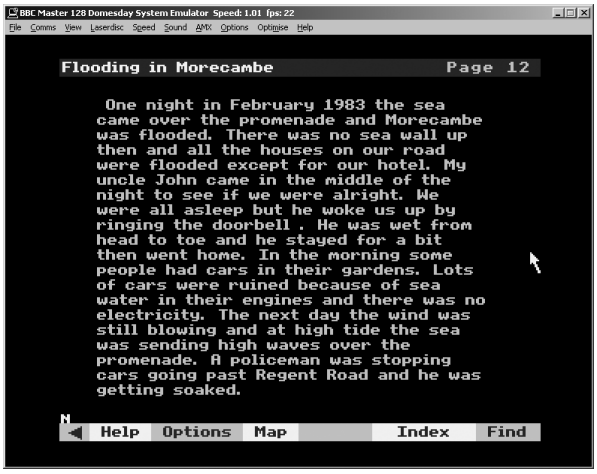
At the time that CAMiLEON was formed there was an unresolved debate about the role of emulation in digital preservation. Some thought that emulation would simply not work in practice and that migration (changing data files into new formats as necessary so that they can be used by new software) would be the answer. A minority argued that migration would not be sufficiently accurate in the long term, and that the only way of being sure that your preserved objects have really survived over time is to see them emulated in their original environment. Overall, emulation was viewed with a high degree of scepticism. So CAMiLEON was proposed as a way of getting to the bottom of the emulation question and testing if it could work in practice.

Of course, emulation has been used in many aspects of computing for decades. What was needed was a strategy to give longevity to emulation implementations and hence make it practical in a preservation context. We found that emulation and migration, rather than being opposing sides to the preservation coin, actually work very well in partnership. One solution does not satisfy every eventuality.

Why was it important to tackle the BBC Domesday project?

BBC Domesday is a fascinating resource and a wonderful story of vision and innovation but it has acted as a prime example of the digital preservation





Screenshots from CAMiLEON's emulation.

© CAMiLEON, 2003, http://www.si.umich.edu/CAMILEON

problem for some time. Domesday contains a wealth of data in a myriad of file formats. But it is the user interface and the software behind it that really add so much value to the resource. You can do really quite advanced things with it; overlaying local statistics on maps for example – in effect pulling together and combining information to provide more useful and informative views. The role of integration that the software performs demands full preservation of the whole. From a historical perspective it would be virtually criminal not to find a way to preserve the user interface. BBC Domesday was a massively innovative and influential landmark in the development of multimedia and came way ahead of its time in almost every aspect of its delivery. Apart from the fact that its complexity made it a challenging test case for our work, by successfully tackling Domesday the CAMiLEON team felt we could go a long way to convincing the emulation sceptics that there was potential in this technique.

Were there any particular stumbling blocks when it came to capturing the original data from the discs, or in producing an emulation of the system?

Producing the emulation was not without its difficulties. There were many stumbling blocks, not least of which was the lack of knowledge and documentation we had about the system. Until we got our hands on a BBC Master and opened it up, we didn't know what special hardware we would have to emulate! We started with just one original Domesday set-up and had a difficult time getting it working to see what the resource actually looked like. Other than seeing it demonstrated on Micro Live (BBC1) in the 1980s I'd never actually seen it running before.

A considerable amount of time was spent tracking down documentation, licence agreements and staff who worked on Domesday in 1985–86 in order to inform our preservation work. A key problem with the development was the emulation of the co-processor. With little

documentation and virtually no software to actually test the emulation, debugging the system was particularly time-consuming. We had to use a great deal of experimentation to get things to work. In extracting the data from the discs we exploited the very early SCSI 1 support used in the laserdisc player. But modern operating systems wouldn't recognise the player as an SCSI device and we had to go through a lot of trial and error to get the communication working. We were fortunate that members of the public saw the media coverage of our work in early 2002 and donated bits of Domesday hardware found in their lofts. We went through a number of the specialist LV-ROM videodisc players and without the donations we simply wouldn't have got the work done.

Fortunately, some examples of the original system were still available to you, albeit incomplete. Was the data rescued just in time?

Most of the really difficult problems we faced were due to the long time gap between the creation of Domesday and its preservation. If we had conducted the rescue 10 years earlier it would have been far easier. The timeliness of preservation work is a crucial issue that Domesday really underlines. Would we be able to rescue Domesday if we left it another 10 years? I'm sure we could, but it would be at far greater expense.

Storage technology becoming obsolete led to the need for CAMiLEON's emulation in the first place. How did you ensure the longevity of your work?

The Cedars Project (http://www.leeds.ac.uk/cedars/) pointed out quite rightly that the medium upon which digital resources are recorded is not part of the message that the resource is conveying. Yet many resources are inextricably tied to the medium on which they are published despite this being a very short-sighted policy. By selecting an appropriate abstraction for the Domesday data we were able to capture the elements of the resource significant for its use and encapsulate these in byte-streams that are easy to preserve and can move from medium to medium over time.

One rather alarming issue was highlighted by the range of offers we had from commercial vendors to 'assist' with our work. One manufacturer offered us a special polymer that they guaranteed would preserve a CD-ROM for 100 years. They were unable to answer how we would preserve a CD-ROM player for that length of time and I could guarantee them there won't be CD-ROM players in the computers of 2103AD! The myth that long-lived media equals long-lived preservation is still worryingly popular.

What are some of the other issues that surround CAMiLEON's Domesday emulation?

The data in Domesday comes from many different sources, and copyright was cleared for use specifically in the Domesday publication, meaning that conforming with intellectual property rights standards is not a simple matter. Does an emulated version of Domesday that runs on a modern computer without the original discs or hardware that it ran on when published constitute a new version of the resource or a copy? There are little or no precedents in law for these kinds of issues. The precedents of the paper world do not map simply to those of the digital domain and this makes the IPR situation very complex. We would argue that the whole point of the preservation work is to create a 'copy' of BBC Domesday which as far as the user is concerned looks exactly like the original, while ensuring that this copy can be preserved over time. It is interesting to consider that in some cases emulation may involve a simpler copyright question than migration, simply because less is actually changing when preservation action is taken.

Based on your experience with CAMiLEON, what do you think are the most important issues for a project to consider when undertaking long-term data preservation?

It is too easy to consider preservation action as 'long-term digital preservation' without actually thinking about the real long-term practicalities of what is being attempted. Both Cedars and CAMiLEON attempted to use examples from the past to inform what our challenges will be in the future. Most assumptions that current standards will survive are flawed and, although technologies like XML will certainly make digital preservation easier, I do not believe they are the all-encompassing solution that many are suggesting. A common sense approach to what is and what isn't likely to survive into the future, based on the lessons of the past, is a good starting point. I think the CAMiLEON approach to software longevity and the identification of elements of a programming language that will be present in future languages is a useful way of looking at preservation as a whole. Whatever kind of digital technology you are looking at, some aspects of it will be inherited by future developments and some will fall by the wayside due to poor design. If you can minimise your use of what will probably be lost in the digital evolution of the future then preservation of what you create will become much simpler.

A discussion of the legal and IPR issues relevant to making CAMiLEON's emulation of BBC Domesday accessible is available at http://www.si.umich.edu/CAMILEON/ reports/IPRreport.doc

Archives Hub

Hub (n.) – A centre of activity or interest; a focal point.

The Archives Hub service allows users of various levels of expertise to access over 10,000 online descriptions of archival collections held by higher and further education institutions in the UK. It provides a single point of access to ever-increasing numbers of descriptions of archival collections that are available for research in universities and colleges.

The co-operation of small to medium-sized organisations in offering a single point of access to distributed collections is growing in popularity. An experimental project, Pouce (http://pouce.valoris.com:8080/pouce/), has performed such a task in a number of museums in France. A full case study on Pouce will feature in the next DigiCULT Technology Watch Report, due in Feb. 04

The Archives Hub forms one part of the UK's National Archives Network and is guided by a steering committee which includes representatives of the Public Record Office and the Historical Manuscripts Commission. Anyone may use the Hub by logging on to: http://www.archiveshub.ac.uk

An international resource, the United Nations Educational, Scientific and Cultural Organisation (UNESCO) Archives Portal, can be found at: http://www.unesco.org/webworld/portal_archives/)

US national plan for preserving digital information

In February this year, the plan for the National Digital Information Infrastructure and Preservation Program (NDIIPP) was approved by the US Congress. This approval releases $20 million of a total $100 million in funding which will allow the programme, led by the Library of Congress, to begin the initial phase of creating a national infrastructure for the collection and long-term preservation of digital content.

The Plan for the NDIIPP outlines both the reasoning and the strategy of the programme. It opens: "Digital technology is radically transforming the ways that we create and disseminate information. This new technology has created a surfeit of information that is extremely fragile, inherently impermanent, and difficult to assess for long-term value… Never has access to information that is authentic, reliable, and complete been more important, and never has the capacity of libraries and other heritage institutions to guarantee that access been in greater jeopardy." In other words, there is a plethora of culturally and historically significant material that is here today, gone tomorrow and therefore steps must be taken to preserve it.

The main goals of the NDIIPP are summarised by the Plan:
"…to seek national solutions for
- the continuing collection, selection and organisation of the most historically significant cultural materials and of important information resources, regardless of evolving formats;
- the long-term storage, preservation and authenticity of those collections, and
- persistent, rights-protected access for the public to the digital heritage of the American people."

It is generally considered to be easier to obtain funding for digitisation for access than for preservation. The long-term benefits and requirements of preservation seem often to be overshadowed by the immediate benefits of current access initiatives, therefore the core funding of the NDIIPP is an important step towards a long-term commitment to preserving 'fragile' information.

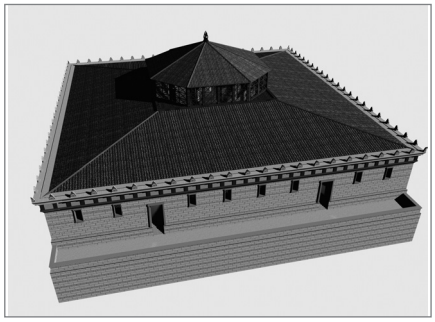
The programme is co-operative and the Library of Congress will work jointly with both federal and non-federal institutions – including the National Library of Medicine, the National Agricultural Library, the Research Libraries Group (RLG), the Online Computer Library Center (OCLC) and the Council on Library and Information Resources (CLIR) – to develop a network of collaborative partners. "The Library of Congress has gained an enormous amount of knowledge from its partners in this initiative. We look forward to a continued successful collaboration as we work together to preserve digital materials before they are forever lost", said Associate Librarian for Strategic Initiatives Laura Campbell. It is hoped that all participating institutions will be able to contribute expertise to the project and will benefit from the technological infrastructure set up.

Although the Library can now move forward with developing the details of the plan and set up the policies, protocols and strategies for the long-term preservation of materials, a further $75 million of private 'matching' funds must be raised in order to receive the remaining $75 million from Congress.

The Plan for the National Digital Information Infrastructure and Preservation Program can be downloaded at: http://www.digitalpreservation.gov/ndiipp/repor/repor_plan.html

Bringing Ancient Greek theatre to life with interactive virtual archaeology

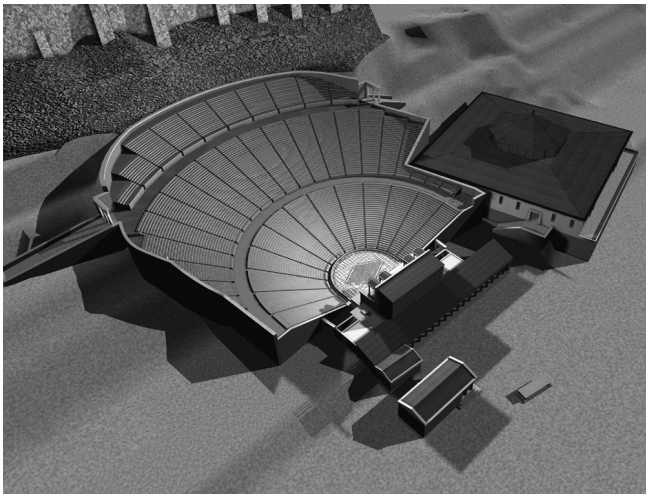
The Odeon of Pericles was the first indoor theatre, built 2500 years ago in Ancient Greece. The theatre today is no more than a pile of ruins on the south slope of the Acropolis but, thanks to the Theatron project at the University of Warwick (http://www.theatron.org), you can now enter a virtual reconstruction of the Odeon, find the best seat, or 'stand' on the stage.



VR reconstruction of the Odeon of Pericles

Archaeological research coupled with state-of-the-art 3D technology allowed the team to reconstruct the theatre, complete with its polygonal roof and around 80 supporting columns. Mr Drew Baker, a Multimedia Designer with e-lab (http://www.warwick.ac.uk/3d/) at the University of Warwick, explains the benefits of 3D imaging: "Virtual reality models of historically important but long-lost ancient sites inject new life into study and enable students and researchers to walk around theatres, many of which have long since disappeared. The creations enable people to look at intricate details and produce 3D representations to help experience time, space and lighting in a way far more engaging than a lecture or set of slides."

In fact, the virtual Odeon indicates that the rows of supporting columns would have obscured parts of the stage for around 40% of the audience and the best view was actually from the beams of the roof. This finding indicates that emphasising the grandeur and spectacle of the 3000-seat auditorium itself was probably more important to the Ancient Greeks than providing clear sight lines to the stage. It is likely that the great plays of Periclean Athens would have been held in the nearby open-air theatre of Dionysus (also digitally reconstructed by the Theatron team) and the Odeon would have been used more for events such as recitals or auditions.



© Theatron project, University of Warwick, 2003, www.theatron.org

Reconstruction of the Theatre of Dionysus

Project leader Professor Richard Beacham, from the School of Theatre Studies at the University of Warwick, states that recreating the Odeon with interactive virtual archaeology allows us to better understand the theatre both in terms of its drama and its historical context: "Most people who walk past it are not really aware that there was quite an important building there in terms of culture, in terms of its architecture, in terms of the influence it had on later buildings." The team have already completed reconstructions of almost 30 theatre sites in Europe, ranging from the Theatre of Dionysus to the Globe Theatre in London, in order to gain a better understanding of ancient plays and the places where they were first staged. Professor Beacham says, "By learning about the setting of a play, you can understand aspects of the drama much better. You can really feel what it was like to watch a performance thousands of years ago." These sorts of applications of visualisation technologies create opportunities for research that have never been possible before. As we begin to develop the methods and theory for such studies new understandings will emerge, as will whole new lines of inquiry.

For more information please contact: Drew Baker, e-lab 3D Visualisation Group, University of Warwick, d.baker@warwick.ac.uk, or Professor Richard Beacham, School of Theatre Studies, University of Warwick, R.Beacham@warwick.ac.uk.

A case study on Theatron can be found in the recent DigiCULT Technology Watch Report (pp. 104-105), which explores the benefits of VRML in Cultural Heritage applications. This Technology Watch Report can be downloaded from: http://www.digicult.info .

Archaeology museums & 3D in the 21st century

David Clarke and James Hemsley, Department of Archaeology,
National Museums of Scotland, Edinburgh, Scotland [d.clarke@nms.ac.uk, jamesrhemsley@cix.co.uk]
Neil Shaw and Michael Spearman, The Multimedia Team, Edinburgh, Scotland
[n.shaw@nms.ac.uk, mike@theteam.demon.co.uk]
Monica Kaayk, Researcher [acinom@cix.co.uk]

ABSTRACT

The ORION project to develop a Research Roadmap and Network on 3D for Archaeology Museums began with the User Practices and Needs Study reported in this article as well as a technology state-of-the-art survey. The user survey revealed higher than expected take-up to date and a very strong interest in future 3D applications. Although certain industrial areas such as electronic games and television are naturally ahead, archaeology museums are in a leading position in the cultural heritage area as regards 3D and are well placed to stimulate as well as apply technology advances. The growing ORION network is intended to help these efforts. ORION can be found online at: http://www.orion-net.org.

INTRODUCTION

3D literacy has always been a part of the museum environment and general learning. Curators and other museum professionals read 3D objects to derive meaning. No other discipline does this to a greater extent than archaeology, where objects are absolutely core to gaining a basic understanding of the societies that produced them. The archaeologist's relationship with the 3D object is unique and it is for this reason that the ORION project chose, as its main focus, European archaeology museums. Nowadays '3D' has come to mean something quite specific in the computer field, i.e. creating the effect of 3D on a 2D computer screen in what is often termed '2.5 D'. It is this sense of 3D, which we are using in this paper, and not the 'true' 3D usually achieved with glasses or virtual reality, although the boundary lines with the latter can be fuzzy. This kind of 3D is already an established part of European society. We are familiar with it from Playstation-type games and television where it often appears in a variety of contexts from news and sports programmes to archaeology and history documentaries. In addition to entertainment and information, its educational potential and its capacity to show, engage and inspire is much greater than 2D images can provide.

This article reports on the results of a survey of museum practices, needs and future aspirations carried out in six European countries in late 2002. The work was conducted as part of the ORION (Object Rich Information Network – IST-23001-39122) project to develop a 'Research Roadmap'. This report may be regarded as 'work-in-progress 'and we hope that it will lead to feedback for ORION.

Begun in mid-2002 and due for completion in mid-2003, the EC-supported ORION Object Rich Information Network has two main aims. First, it aims to create an informed, authoritative picture of current and emerging 3D technology, literacy and usage for object and site interpretation in Europe and beyond. This is to help in determining a 'Research Roadmap' for the European Commission. The second main aim is to build up a balanced, structured network of leading cultural heritage, business and technical partners with a direct interest in the development of 3D technology, literacy and usage that will be ready to support subsequent R&D work The project consortium includes archaeology museums from six European countries (Provincial Archaeological Museum of Alicante, Archaeological Laboratories of Dion, Mont Beuvray (archaeological site and museum), National Museum of Ireland, National Museums of Scotland, Neanderthal Museum) and also technical partners from these countries.

METHODOLOGY

Four main approaches were utilised in this user survey:

1. Workshops in each partner country and international workshops such as at EVA 2002 London, bringing together 16 to 40 participants from archaeology museums, universities, companies and other interested organisations. This was the most important method used to obtain information.

2. Questionnaire survey with responses from 67 archaeology museums in the partner countries, with this indicative sample ranging from very small museums (3000 visitors per year) to large ones with over a million visitors.
3. Interviews/meetings with curators, museum management, government representatives and technical experts from the ORION countries along with other European countries, Canada, Japan, Russia and the USA.
4. Literature and Web research of museum and other relevant sites, taking advantage of current and previous work (see references [2] [3] [4] [5] [6]).

We believe this four-pronged approach provided the team with a reasonable overall picture including 'leaders and laggards' in technology take-up. However, we do not claim to have followed a rigorous statistical methodology since only six countries were covered and we opted to have equal sample sizes from each country, thus under-weighting France, Germany and Spain.

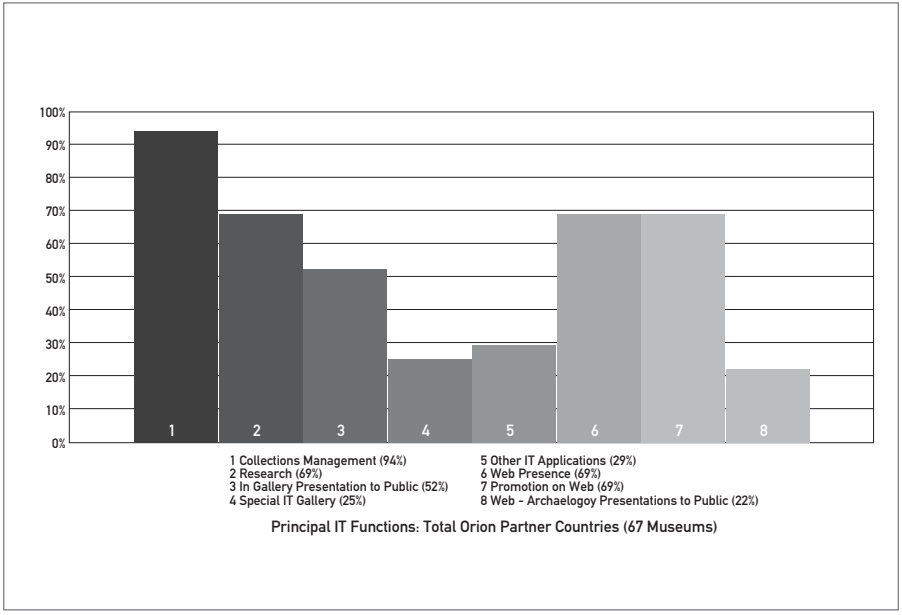
In the ORION study we take 'Archaeology Museums' to mean all museums with archaeological collections, thus embracing a range of institutions from small specialised archaeological museums such as the Museum of Cycladic Art in Athens to the Louvre and the British Museum.

Standing at the end of the 'archaeological chain' from excavation through to display, museums have a crucial role in preserving and restoring items, and in educational and research access for the public. In the European Union there are around 1 billion objects in 15,000 museums receiving over 600 million visits per year, as reported by the Museums and Galleries and New Technologies (MAGNETS) project [7]. Of this total we estimate that there are over 5000 archaeology museums in Europe. In the UK, for example, there are 739 museums with archaeological collections – about a third of all museums (www.24hourmuseum.co.uk). Greece is a special case; the majority of museums have archaeological collections.

GENERAL USE OF IT AND ENGAGEMENT WITH 3D BY EUROPEAN ARCHAEOLOGY MUSEUMS

First, there was a universal acceptance across all six countries analysed that IT has an absolutely core role to play in managing cultural assets and presenting cultural heritage to the public. This was recognised by those representing small museums with little IT capacity as much as by large institutions with sophisticated IT installations. There was, however, a great variety of experience regarding the extent to which IT is currently being used and a concomitant range of future expectations and priorities. At the lower end of the spectrum, a number of smaller museums had a very basic level of IT support. Accordingly, the immediate priorities for these institutions tended to be fundamental IT activities such as computerising their collections management system, training staff in using computers, establishing a Web presence. Larger institutions, by contrast, had some very different priorities, for example the incorporation of IT and increasingly 3D in major in-gallery exhibitions and the development of sophisticated online exhibitions. However, a striking aspect was the high level of IT use for research (e.g. the Web and Internet) by both small and large museums.

Consultation with museum professionals and other key experts in the cultural sector highlighted great interest in 3D and considerable awareness of its potential. Around 65% of archaeological museums said 3D had either an 'important' or a 'very important' role to play in presenting archaeology to the public and in the study of material culture. This varied from around 50% in Spain to around 85% in Greece. One surprising finding from our survey of European archaeological museums was just how many had already used 3D in presenting or promoting their collections, sites or museums. Around 35% of museums we surveyed had already used 3D in some form. Often this was for only one initial presentation or project but further



inputs from the questionnaire survey, workshops and interviews showed that museums are beginning to realise the potential that 3D offers and take action. Moreover, around 10% of museums consulted held some 3D images of their objects and around 15% held 3D movies of sites. The fact that a significant number of the museums consulted already have the beginnings of a 3D library illustrates a serious awareness of the growing relevance of 3D to the cultural sector. The Scottish Cultural Resource Access Network (SCRAN) has the largest collection to our knowledge in the European cultural sector with several hundred 3D scenes, although a much lower number of 3D images of objects.

The use of 3D provides many specific functional advantages especially as regards social inclusion, as indicated by this quote with regard to Public Access and 3D Applications:

> *'The visualisation offered by 3D serves as a better stimulation of the human system of cognitive perception: less abstraction and more real understanding of archaeology. The transfer and the marketing of archaeological or cultural content by 3D opens this field to a greater public and does not focus only on few people of a well-educated elite.'*
> (Dr Joachim Paul, Medienzentrum Rheinland des LVR, ORION interview, 2 October 2002)

Based on the study work, some of the main forces affecting the take-up of 3D by Euro- pean archaeology museums are as follows:

Negative forces:

- Perceived costs (65%)
- Lack of appropriate knowledge of 3D
- Conservation
- Lack of awareness of commercial benefits
- Confusion between accessibility and 'dumbing down'
- Fear of obsolescence

Positive forces:

- Falling costs as technology develops and experience grows
- A number of role models
- Increasing public familiarity with 3D, e.g. games and TV
- Influence from other areas, e.g. medical and tourism, industry
- University, research institute and company interest in museums

Most museums that provided inputs to the study were agreed as to the benefits of using 3D as a vital means of enhancing public access to both objects and sites. They feel 3D offers potentially unprecedented opportunities for:

- Getting nearer towards using real objects than can be achieved by 2D presentations
- Increasing users' sense of connection to and ownership of cultural resources
- Increasing effective learning by showing more clearly than can be achieved with 2D
- Promoting interest in and understanding of not only the objects and sites butthe societies and cultures that produced them.

The opportunities come in many forms. While using 3D representations users can:

- Virtually examine a treasure which only specialists and curators could normally handle, see all of the object – including reverse, base and, where applicable, the interior – and even view a magnification of the object to several times its size to analyse details.
- See how objects, such as tools, were actually used.
- See where objects, such as statues, would originally have appeared within a site.
- Virtually explore objects and sites that no longer exist in their original form.
- Virtually explore alternative views of what original objects and sites might have looked like.
- Virtually compare different objects or different sites.
- See how an object or a site developed over time.
- Virtually explore existing sites.
- See how sites, such as mills, operated and what processes were involved.
- Virtually explore sites that are not easily accessible, such as underwater wrecks.

In addition to these uses for the public, which it was generally agreed should be the main priority, there are also many other advantages and uses both for museum professionals and also for marketing applications such as physical reproduction manufacture.

ROLE MODELS AND FUTURE ASPIRATIONS

One of the ORION questionnaire topics was on examples of good practice. This led to a surprisingly high number of responses that cited more 'foreign' cases than national cases. This is indicated by the replies from the Greek museums surveyed who cited as examples: Byzantine Museum of Athens, Louvre, Sanctuary of Isis at Dion, Archaeoguide at Olympia, University of Newcastle, British Museum, Boden Museum Berlin, Metropolitan Museum NY, and the Virtual Reconstruction Cultural Map of Rome CD-ROM.

It may be noted that in general there was considerable willingness to look outside one's own country for best practice and indeed outside Europe, reflecting both the openness and the nature of the museum archaeologists. However, for some countries the references were mainly national and a number stated that they knew of no example of excellence. There is thus no reason for complacency.

A telling finding was that around 45% of museums said that 3D was part of their future IT development strategy. In some cases, this involved specific, identified projects; in others, it represented more of a declaration of intent. There was a large variation across countries on this question, from around 35% in Scotland and France to around 80% in Germany. Regardless of whether or not 3D was part of their institutions' present plans, there was almost universal consensus among museum professionals that certain uses of 3D in an archaeological museum context constituted a high priority despite the costs. There was a strong consensus across all countries that the main purpose of using 3D was, in the first instance, to create better public access to museums' collections. Presentation of 3D resources was rated a significantly higher priority than other uses such as facilitating academic research, documentation and security, restoration and conservation, or marketing. However, as institutions increasingly use 3D for public access purposes, they are more likely to use 3D resources in other areas of their day-to-day work.

In researching the current and potential future role of 3D in archaeology museums in six European countries, it is important also to consider the broader context, first in the context of the archaeological museums in the rest of Europe which we believe have similar experiences and future aspirations to those studied in the six ORION countries. For example, in Italy there are many examples of leading 3D practice but there are also many small archaeological museums that have not yet been able to 'enter the 3D world'. Secondly, there are archaeology museums worldwide and consideration of published papers, conference presentations and, perhaps most importantly, visits (both physical and using the Web) and personal contact with leading international figures indicates that the European trends are being shared worldwide with pioneers such as the Smithsonian and the Takayama Museum in Japan. Other types of museums are also important to consider, one example being Science museums (who are also in the forefront of 3D use). Returning to archaeology, ORION has also focused on cutting-edge research and development in archaeological excavations (e.g. the 3D MURALE project, http://www.brunel.ac.uk/project/murale/) and cultural tourism (e.g. Archeoguide project, http://archeoguide.intranet.gr/). The games, manufacturing, medical and television industries are areas of intense activity in 3D, as is the 3D systems supplier sector. ORION technology work includes consideration of these industries and corresponding work in universities and research institutes.

The Research Roadmap sets out the key lines for research and development meriting possible EU support in the Sixth Framework Programme. In the study of user needs therefore, particular attention was paid to the views of users and technical specialists attending the user-oriented workshops. The workshops resulted in the following initial (partial) list of priority topics:

- Developing culturally aware Centres of Excellence to disseminate to the museum community information and advice about 3D – including on best practice in IPR and business standards – and provision of 3D technical and design expertise.
- Research into delivering high-quality, high-resolution 3D images of objects at affordable costs. N.B. The European Commission has, as a priority, reduction of digitisation costs by 50% and this is a goal that may well be achieved in the timeframe up to 2006 supported by well-focused R&D efforts.
- Research into technology to provide high-quality presentations of different restoration alternatives for objects and reconstruction alternatives for sites.
- 3D and disabled access, including for example haptics ('touching' technologies).
- 3D watermarking technology and 3D content management systems.
- Automated 3D measurement systems for registration purposes including objects from recent excavations.
- Showcasing such new results in a major international touring archaeological exhibition.
- Exhibition to major world centres supported by an advanced 3D Website.
- Investigation into social and psychological behaviour with regard to 3D.
- Research into overcoming 'soft' organisational barriers to 3D take-up by museums and use of benchmarking management approaches.

Using the parallel-technology state-of-the-art study in ORION these issues are being transformed into more detailed R&D topic descriptions. Moreover the entire 3D and Archaeology issue can be regarded as just one 'cell' in the wide-ranging Research Roadmap matrix from the eCultureNet project (see http://137.120.135.183/FP5/ for the latest information).

CONCLUSIONS AND NEXT STEPS

'You see great 3D used in archaeology and history programmes on television in Britain. TV has really grasped the opportunities that 3D offers. Why don't museums?' Mark Wright, Edinburgh Virtual Environment Centre, Edinburgh University, ORION Workshop, Edinburgh, 3 December 2002

The potential for 3D use in museums is clear; however, it should be recognised that archaeology museums, especially in Germany, are more advanced in this regard than other types of museums and far ahead of other memory institutions such as archives and libraries which naturally concentrate on 2D images, despite calls to get engaged in 3D (Lesk, 2002).

Inevitably, there are, as with 2D, practical issues inhibiting full development, one example being watermarking: 'Without effective 3D watermarking museums will naturally be unwilling to distribute 3D images of their objects.'
John Taylor, National Research Council of Canada, ORION Workshop, London, 24 July 2002

This leads to specific conclusions for the Research Roadmap, as already indicated above and similarly for other business and management issues.

Next steps include publishing online the full ORION report on User Needs and Practices as reported in this paper, as well as the report on Technology State of the Art. However, the key document is the ORION Research Roadmap, which has recently been completed and will be made available online in the near future.

For more information on the ORION Object Rich Information Network, visit: http://www.orion-net.org/index.asp

REFERENCES

[1] Lesk, M.: 2D Good, 3D Better Higher: Dimensions in Digital Libraries. EVA Harvard Symposium 2002, VASARI UK, Fleet, Hampshire, UK

[2] Lowe, A., Franquelo, M. & Clark, J.: The Production of Conservation Facsimiles: the Tomb of Seti 1. EVA 2002 London Conference Proceedings, VASARI UK, Fleet, Hampshire, UK

[3] Chen, L.: Excursion to an Ideal Landscape: West Lake Beyond Time and Space. EVA 2001 Florence Proceedings, Pitagora Editice, Bologna 2001

[4] Brewster, S.: The Impact of HAPTIC Touching on Cultural Applications. EVA 2001 Scotland Proceedings, VASARI UK, Fleet, Hampshire, UK

[5] McLaughlin, M. and Sukhatme, G.: Touch in Immersive Environments. EVA 2000 Scotland, VASARI UK, Fleet, Hampshire, UK

[6] Mulrenin, A.M., DigiCULT Primary Results: Unlocking the Value of the Cultural Heritage Sector. EVA 2001 Berlin Proceedings GfaI, Berlin, 2001

[7] MAGNETS report on Museums & Galleries and New Technologies, 1997 http://www.vasari.co.uk/magnets/VMAG1.html

Images of The Hague

Report on preserving e-mail

Digital information is often ephemeral and difficult to capture for a permanent record. Additionally, to maintain their value as accurate historical artefacts, it must be ensured that digital objects are both authentic and inviolate. Testbed Digitale Bewaring (Digital Preservation Testbed http://www.digitaleduurzaamheid.nl/home.cfm) was set up in October 2000 by the Dutch government in order to carry out research into the long-term preservation of four kinds of digital object: e-mail messages, text documents, spreadsheets and databases.

Recommendations for the preservation of e-mail can now be found in a new publication, 'From digital volatility to digital permanence: Preserving emails', in the Digital Preservation Testbed Virtual Library at: http://www.digitaleduurzaamheid.nl/index.cfm?paginakeuze=185&categorie=2

The report covers topics related to the long-term preservation of e-mails: technical, cultural and legal issues (including the potential of XML); the roles of managers, record keepers, ICT professionals and end-users; and the practical aspects of digitisation.

Digital Preservation Testbed is planning to produce recommendations for preserving text documents, spreadsheets and databases in three more separate publications.

RECENT STUDY SHOWS GROWTH IN EDUCATIONAL LINKS FOR US MUSEUMS

US museums from art galleries to zoos commit more than a billion dollars and over 18 million instructional hours annually to providing learning programmes for schoolchildren under 12.

A recent study by the Institute of Museum and Library Services (the primary source of federal grants for museums and libraries) found that median educational programming costs increased four-fold, from 3% of the total budget in 1995 to £22,500 (12% of budget) in 2001.

The Institute of Museum and Library Services (IMLS) study shows that museums present educational programmes in core school subjects that are reaching an increasing number of teachers and pupils. By surveying more than 15,000 museums of all sizes, types and locations, the IMLS study found that, in 2000-01, nearly 70% of responding museums said the number of schools, students and teachers they serve had grown in the past five years.

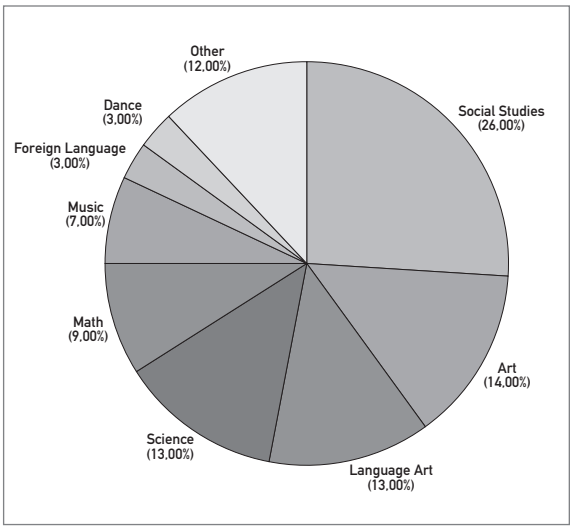
The survey also found that 71% of museums work with curriculum specialists to tailor educational programming to support school curriculum standards and that museums are increasingly using new technologies to improve access to their resources: 72% use Websites for educational programming, 58% e-mail teachers and 24% e-mail students directly.

Educational activities offered by museums include field trips for school groups (77%), museum staff visits to schools (54%), teacher training (32%), resource kits at school sites (34%) and travelling exhibits (17%).

IMLS Director Dr Robert Martin said, "Museums are no longer depositories of the past, but are centres of learning. The resources that they bring to school partnerships provide unique object-based and visual learning opportunities that help schools meet their learning objectives. IMLS wants every child to have access to museums and libraries, to encourage a love of discovery and learning that they can return to again and again in their lives."

The survey highlight publication, True Needs, True Partners 2002, can be viewed at: http://www.imls.gov/pubs/pdf/m-ssurvey.pdf

For more information about the Institute of Museum and Library Services, visit: http://www.imls.gov



LESTER
(Learning Science and Technology Repository) By Lisa Spiro

In a fast-growing sector it is often useful for projects, researchers, organisations and communities to be able to communicate with one another. Unfortunately, the speeds at which innovative technologies change sometimes make it difficult to quickly get an accurate overview of the field in question. Project Director Lisa Spiro describes the Learning Science and Technology Repository (http://lester.rice.edu), a gateway to events, discussions and information about the development of learning technologies.

Tracking and facilitating innovation in learning technologies - The lack of a forum focused on ongoing research in learning science and technology hampers efforts to (1) track innovations, (2) identify areas requiring further research, and (3) implement new learning technologies. LESTER fills this gap.

Developed by Rice University's ETRAC (Educational Technology Research and Assessment Cooperative, http://antioch.rice.edu/etrac/about.html), and sponsored by Microsoft Research, LESTER serves as an information gateway to support the creation and implementation of innovative learning technologies as well as the dissemination of new research about learning. LESTER features a robust, Web-accessible database inventorying leading learning science and technology (LST) projects, researchers, research organisations, and funding agencies around the world. Through this continuously updated resource, users can discover valuable information about significant LST initiatives, such as their research priorities, timelines, funding sources, personnel and sponsoring organisations.

Just as online educational resource communities such as the Multimedia Educational Resource for Learning and Online Teaching (MERLOT, http://www.merlot.org) provide extensive, carefully catalogued records about educational content in a range of disciplines, so LESTER serves as a dynamic portal to innovative educational technologies and research projects, as well as the researchers and organisations developing them. LESTER can be searched, analysed and updated by researchers, and is based on a robust, extensible data model that captures the relationships among LST entities and activities. LESTER provides the basis for a composite portrait of ongoing research in learning science and technology, highlighting new technologies and their application to teaching and learning.

Goals

The specific goals of LESTER are:

- To improve educational practice and research by making available detailed, current information about significant learning science and technology (LST) organisations, projects and researchers.
- To facilitate collaboration among LST researchers.
- To make visible research paths and reveal gaps in current research priorities.
- To serve as a community space for learning science and technology.
- To address the gap between innovation and implementation.

Approach

To achieve these goals, LESTER's approach is to:

- Build an active, engaged user community.
- Design LESTER so that it can be scaled, extended and federated with similar databases.
- Actively collect information.
- Develop an LST Taxonomy and Thesaurus to be used in creating records.

KEY FEATURES OF LESTER

Offers high-quality, useful information about new technologies for education

LESTER makes visible otherwise hard-to-find information about the state of the art in learning science and technology by aggregating data culled from journal articles, conference presentations, Websites and other sources, as well as data contributed by the researchers themselves. When LESTER was released on 2 December 2002, it contained information on more than 265 projects, 290 researchers, 180 research organisations and 110 funding organisations throughout the United States and Europe.

Constructs a collaborative and inclusive community

Through LESTER, researchers can discover potential partners, and instructors and instructional technologists can identify promising technologies and methodologies. Not only does LESTER provide a window into the LST community, it also enables the members of this community to add new records and edit existing ones about their projects and organisations.

Provides language for describing emerging technologies

To facilitate accurate searches and provide a precise descriptive language, LESTER employs a set of keywords that is based upon the ERIC Thesaurus, a standard vocabulary in the field of education, but

that includes new terms and keyword hierarchies drawn from the computer science and high-tech communities.

Complements other information sources on learning science and technology LESTER offers snapshot views of ongoing research, enabling users to dig more deeply by visiting related Websites for further information. By providing information about researchers and organisations as well as projects, LESTER highlights centres of activity in particular areas of educational technology.

FUTURE DEVELOPMENTS

As LESTER grows, it will provide key elements of the basic research infrastructure for LST and support communication and collaboration. LESTER will include a wider variety of data, including information about conferences and current funding opportunities in educational technology as well as white papers and citations submitted by researchers. In addition, we would like to enable greater information exchange and interactivity in the LST community by providing online forums, bulletin boards, and other community-building tools, as well as by enabling users to contribute peer annotations about projects. It is hoped that LESTER will contribute to research by providing information about recent innovations, promoting interdisciplinary collaborations, and offering both an overarching taxonomy and data to support the development of research roadmaps. LESTER also aims to address the gap between research and implementation in the classroom by providing a lens through which instructors and technologists can examine recent research. Success will be measured by number of registrands, number of records included in LESTER, Web traffic, and usefulness and quality of LESTER as determined by user studies.

Images of The Hague

For more information, please visit: http://lester.rice.edu or contact Lisa Spiro, Project Director, at: lester@rice.edu

THE DIGITAL OBJECT IDENTIFIER SYSTEM EXPLAINED

WHAT IS A DOI?

The Digital Object Identifier (DOI®) system is for persistent identification and interoperable exchange of intellectual property on digital networks. It provides an extensible framework for managing intellectual content in any form, at any level of granularity and in any digital environment. A DOI differs from commonly used Internet pointers to material (e.g. a URL) because it identifies an object as a first-class entity, not simply the place where the object is located.

WHAT IS THE IDF?

The International DOI Foundation (IDF) was created in 1998 and supports the needs of the intellectual property community in the digital environment by the development and promotion of the Digital Object Identifier system as a common infrastructure for content management. The Foundation is a registered not-for-profit organisation, controlled by an Executive Board elected by the members of the Foundation. The activities of the Foundation are controlled by its members. Membership is open to all organisations with an interest in electronic publishing and related enabling technologies.

In April this year, an informal consortium of three national libraries (The British Library, Die Deutsche Bibliothek and the Koninklijke Bibliotheek) joined the IDF.

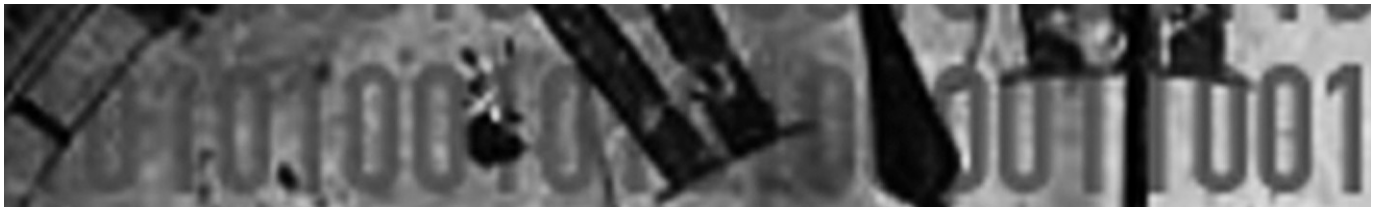
More libraries, particularly national libraries, are being encouraged to join the IDF by the International Federation of Library Associations (IFLA).

A 2002 steering group meeting of the IFLA and International Publishers Association (IPA) recognised that national libraries are extremely well placed to lead the adoption of DOI within library communities, in particular given their role in associating bibliographic information to national collections.

Using DOIs can make managing intellectual property in a networked environment much easier and more convenient. For example, a library or a publisher can construct a central directory of digital objects, all uniquely identified with DOIs, which can be accessed over a network. Each item within the system and its associated metadata have persistence if moved or bookmarked and are interoperable with data from other sources. The DOI system is also extensible and can provide functionality to deal with groups of objects, management of data for multiple output formats and dynamic updating of data and services.

The initial focus of DOI applications was on created resources rather than natural objects (e.g. specimens in a natural history museum, people, places, events, etc.). However these other types of resource are also necessarily involved in intellectual property transactions, and so may be identified by DOIs where appropriate. As a DOI can be used to identify any resource involved in an intellectual property transaction, it facilitates the construction of automated services and transactions for e-commerce. Practical applications of DOI in use in an e-book can be viewed at: http://doi.contentdirections.com/phase2.html

A detailed description of the DOI system, including the system handbook, DOI tools and latest news can be found at: http://www.doi.org



FREE SHEET MUSIC FOR LIBRARY USERS

Thousands of pieces of music can now be accessed, free of charge, from seven of Denmark's libraries. Users can download and print out high-quality original editions of classical music and jazz from world-famous artists, composers and publishers. The owners of the music receive a royalty from the library for each copy that is distributed. The seven libraries (Frederiksberg, Gentofte, Copenhagen, Odense, Herning, Aarhus and the State Library) will pilot the system before it is expanded to a possible 250 libraries in Denmark.

This new access scheme is a result of co-operation between Denmark's National Library Authority (Biblioteksstyrelsen, http://www.kb.dk/index-en.htm) and Sheet Music Now A/S (http://www.sheetmusicnow.com/). The project provides enhanced access services for members of the public and also ensures that the artists' rights are protected.

"WHAT'S NEW IN DIGITAL PRESERVATION"

Issue 5 of the DPC/PADI quarterly bulletin "What's New in Digital Preservation" is now available. It covers a selection of the recent digital preservation activity of various organisations and initiatives, as well as specific areas of activity. This issue is collated from the Preserving Access to Digital Information (PADI) Gateway and covers February-May 2003.

You can view or download the report from the Digital Preservation Coalition (DPC) Website at: http://www.dpconline.org/graphics/whatsnew/ or from the National Library of Australia (NLA) at: http://www.nla.gov.au/padi/qdigest/jun2003.html

Images of The Hague

Economic and Social Data Service (ESDS)

The Economic and Social Data Service (ESDS, http://www.esds.ac.uk/) is a new national data archiving and dissemination service which came into operation in January 2003. It is a jointly funded initiative sponsored by the Economic and Social Research Council (ESRC) and the Joint Information Systems Committee (JISC). The ESDS has been established as a distributed service, bringing together the following centres of expertise in data creation, preservation and use:

- UK Data Archive (UKDA), University of Essex (http://www.data-archive.ac.uk/)
- Institute for Social and Economic Research (ISER), University of Essex (http://www.iser.essex.ac.uk/)
- Manchester Information and Associated Services (MIMAS), University of Manchester (http://www.mimas.ac.uk/)
- Cathie Marsh Centre for Census and Survey Research (CCSR), University of Manchester (http://www.ccsr.ac.uk/)

ESDS provides preservation, dissemination, user support and training for an extensive range of key economic and social data, both quantitative and qualitative, spanning many disciplines and themes. ESDS comprises a number of 'specialist' data services that offer enhanced support for the secondary use of data across the research, learning and teaching communities:

ESDS MANAGEMENT

The overall direction and management for the ESDS is the responsibility of the UKDA, aiming to provide consistency and standards across the service. It has a broad strategic role, relating to a variety of stakeholders concerned with the supply, funding and use of social science data, and creates a coherent publicity, promotion and out-reach strategy for the whole service. The key deliverables include a universal portal to help discover and locate social and economic data resources; a central registration service operating across the ESDS; a 'first-stop' help desk service; and a coherent and flexible collections development policy.

ESDS ACCESS AND PRESERVATION

This service represents the core of the ESDS and focuses on the central activities of data acquisition, processing, preservation and dissemination. The key services to be provided are a collections development strategy working to identify and acquire a range of data from across the social sciences; ingest activities including adding value to data and documentation, creating metadata, negotiating appropriate rights management frameworks and preparing data for preservation and dissemination purposes; a flexible data dissemination service; and long-term preservation and collections management for all datasets acquired across ESDS.

ESDS GOVERNMENT

Large-scale government surveys, such as the General Household Survey and the Labour Force Survey, are key data resources for social science researchers. ESDS Government, led by the CCSR, aims to promote and facilitate increased and more effective use of these surveys in research, learning and teaching. This will include methodological guides, short courses and support for key statistical packages. The service will work closely with ONS, other data producers and users to maximise the synergy between these groups. Data will be made available via the UKDA Web-based download system and the Nesstar system for online browsing and visualisation of the data.

ESDS INTERNATIONAL

ESDS International, led by MIMAS, provides access to, and support for, a range of international datasets – both macro and micro sources. Key services include Web-based access to regularly updated macro-economic time series datasets produced by organisations such as OECD (Organisation for Economic Co-operation and Development) and IMF (International Monetary Fund); helping users to locate and acquire international micro-level datasets, such as Eurobarometer or ISSP (International Social Survey Programme); and the provision of end-user training, a range of Web-based support and a dedicated help desk.

ESDS QUALIDATA

ESDS Qualidata provides access and support for a range of qualitative datasets. The new service builds on Qualidata's expertise and international reputation in this area, developed over the past eight years. It aims to create, and provide user support and training for, enhanced data and documentation resources for use in research, learning and teaching. Emphasis is being placed on creating more flexible access to digital qualitative data, via the UKDA Web-based download service and online browsing of data marked up using XML standards. The Edwardians Online project

(http://www.qualidata.essex.ac.uk/edwardians/), which has developed an online multimedia demonstrator, represents the first phase of this work.

ESDS LONGITUDINAL

The work of ESDS Longitudinal will be undertaken jointly by the UKDA and ISER. ISER includes the ESRC-funded UK Longitudinal Studies Centre (ULSC). The service aims to provide value-added data enhancements, user support and training for key longitudinal data collections including the British Cohort Study (BCS70), the British Household Panel Survey (BHPS), the Millennium Cohort Study (MCS) and the National Child Development Survey (NCDS). The service will create additional derived variables, enhanced data and documentation stan dards across all studies, and improve information on weighting, missing data, and statistical adjustments. Online access to data, data samplers and documentation is also planned to facilitate the exploration, matching and subsetting of data.

THE YLE DIGITAL SOUND ARCHIVE

Yleisradio (YLE) is Finland's national public service broadcasting company. It is among the first broadcasters to face the challenge of creating a digital archive from radio broadcast material. YLE has now adopted a full digital radio archive solution where the digitised material can be located, listened to and transferred to the user by browser-based technology. Jouni Frilander, Pekka Gronow, Petri Home, Markku Petäjä, Pekka Salosaari and Lasse Vihonen describe some of the issues and challenges that faced them when designing and creating this digital sound archive.

THE PROBLEM

For more than half a decade, most broadcasting companies have recognised the need to archive some of their productions. The most obvious motive for this was the need to use old broadcasts or parts of them again at a later date, but there has also been recognition of the intrinsic value of broadcasts and the need to preserve them as part of cultural history.

For those of us who work in broadcasting, all this is well known, but it is useful to describe the established practice for those who come from other fields of archiving. Radio and television programmes were usually produced on audio or videotape before transmission. After transmission, it was easy to keep the tape reels in storage. In the same manner,broadcasting companies also acquired large collections of recorded music.

Most broadcasters kept card catalogues of their archival tapes, stored with the tape number and some additional information. In the simplest case, the cards were filed alphabetically according to the programme title or the name of the producer. In more elaborate systems, there were multiple cards, which permitted searches by programme title, name, tape, and some additional criteria. In the 1980s, the YLE radio archive had a catalogue consisting of more than a hundred thousand such file cards.

At this stage, fortunately, computers came along to help us. Instead of file cards, the information could now be typed into a database. It was now easier to search for information, and if the computer was put on a network, anybody working for the company could search the catalogue. The computerised catalogue was initially used only for new productions, but gradually we were even able to convert our old card catalogue into database form.

With the computerised catalogue, we could have managed for a long time, if there had not been the problem inherent in our archival medium: magnetic tape. Professional-quality audiotape has a maximum life expectancy of about 50 years, in practice often less. By the 1990s we had an increasing number of old tapes approaching the end of their lifetimes. These tapes had to be copied onto a new medium if they were to be saved.

From an archival point of view, the easiest solution would have been to copy them onto a new analogue tape. Although analogue copying always results in a loss of sound quality, with professional equipment the loss is so small that acceptable standards could have been maintained for a few hundred years more with analogue tape, just by copying our archives every 50 years.

However, digital recording technology was introduced at this time. It was so much cheaper, and in some respects better, than the old technology that analogue tape more or less disappeared within a period of 5 to 10 years. Suddenly all radio production was done on digital media such as DAT tape, recordable compact discs and minidisks. At YLE, we do not even have analogue tape recorders in our studios any more, therefore we have to copy all archival tapes onto new media when they are used in broadcasts.

In many ways, the 1990s were a nightmare for the radio archivist. It was soon discovered that the new digital carriers were not reliable archival media. Although they could, under ideal conditions, preserve high-quality sound for a long time, their preservation demanded such a high degree of quality control that we found it impossible to maintain. And although a digital copy is in theory identical to the original, this does not help much if your DAT or CD player is unable to reproduce the original. But there was little that could be done. At the YLE radio archives we have today more than 50,000 digital sound carriers of various types (not including commercial CDs) from the 1990s, and they cause us more problems than our old analogue tapes.

By the mid-1990s, these digital carriers were gradually replaced by 'tapeless recording', recording onto computer hard disc. This was introduced with the development of computer-assisted radio (CAR), where programme planning, production and transmission were integrated. However, CAR systems had very limited archiving capacity. To preserve the programmes permanently, we had to copy the sound files on CDs or DATs and retype the content information to our database. This was obviously not economical. Something had to be done.

THE PROBLEM IN BRIEF

- 250,000 deteriorating analogue tapes
- 50,000 unstable digital carriers (DAT, CDR, MD)
- 'Tapeless recording', born-digital sound files

The solution

The main goals of our planned archive system were:

- the possibility of archiving 'born-digital' material directly from production systems (CAR systems), and using metadata created in the production process as the basis of the catalogue database;
- the digitisation of existing sound carriers;
- to make the archives available online to all producers in the company and improve the service and effectiveness of the archives;
- permanent storage ('forever').

When we started studying new archival solutions in the mid-1990s, we considered various options. One option was to use giant CD juke boxes, which had the advantage that they could be used for normal music CDs as well as CDRs. Soon, however, we concluded that in the long run the only feasible alternative would be a digital mass storage device – in practice, a tape robot. The original 'production quality' sound would be stored on the tape robot and the system would automatically create a bit-reduced 'listening copy' of all sound files for quick browsing.

We also decided that the system should be able to handle all of the audio formats used in our company. All programmes are archived in the format in which they are created, but the system will create conversions required by various users.

At this stage we had to decide what our priorities were: what we wanted our system to do. We wanted to integrate the digital archive into our existing catalogue database, which had been recently revised and we were fairly satisfied with it. This was a critical decision, because we found out that in practice this excluded certain potential suppliers who wanted to sell us their own database model as part of their archive solution. Using the existing database would guarantee that we could keep information about the old analogue collection in the same system as new digital material and effectively control the digitisation process.

THE SOLUTION IN BRIEF

- Tape robot (digital mass storage) with 100 TB capacity
- Integrated with existing catalogue databases (RARK, FONO)
- All sound files duplicated (archival copy and browsing copy)

It was equally important to decide what the archive system should not do. The system was not planned for the capturing

of broadcasts off-air. Neither was it designed to support on-air broadcasting. Before archival materials could go on air, they would first have to be moved to a CAR system. It was also decided that the archive would not be used for temporary storage. All materials that go into the archive will, in principle, stay there forever. The link to the existing catalogue meant that the system does not support the digitisation of uncatalogued material. It cannot – at present at least – be accessed outside our Wide Area Network.

THE USER'S VIEWPOINT

From the user's viewpoint, the digital archive is not much different from the old catalogue database, which already has more than a thousand users within our company. The users approach the archive through BRS-SEARCH, a Web-browser based user interface. They can search all of the items in the database, or digitised material only.

We also had to consider files with complex relationships. How, for instance, should we handle multi-part musical works? In classical music, symphonies and sonatas typically consist of three or four parts. Should we handle sonatas as one sound file, or as three or four sound files? We chose a solution that allows us to use both approaches: the user can retrieve the whole work or one part only. The same principle is applied to news broadcasts which are treated hierarchically as 'multi-part works', allowing the user to retrieve the whole programme or just one story.

In principle, it would also be possible to index other radio programmes in great detail, so that the user could go directly to a specific part of the programme. However, this would demand a detailed documentation of the contents. We do not have the capacity for this. Instead we have built a feature which allows users to listen to the complete programme and mark those parts that they need. Instead of ordering a large file consisting of, say, half an hour of linear sound, the user gets only the three-minute part that he wants.

SYSTEM FUNCTIONS

- Access through search and retrieval system.
- User can search digital and/or analogue materials.
- Browsing-quality sound streamed to all users – production-quality files take longer.
- Format conversion on demand.
- Special features: pre-selection of excerpts, support for multi-part works.

SYSTEM COMPONENTS

The metadata management system contains descriptive information on the material (title, authors, performers, duration etc.) and also technical metadata. Each item is assigned a unique ID that is used when creating names for individual audio files, then, during the last stage of the process, the audio files are imported to the digital audio archive system and linked to the related metadata.

All audio files that are stored in the archive are controlled by an essence management system which processes the files and also optimises access time by caching the most recently stored or retrieved files within the online storage. All pre-listening copies are kept online for fast access time.

The search and retrieval system has a database structure, which can grant access rights for different user groups: text data search (open to all users), access to pre-listening quality material (open to a restricted group of users) and access to production quality material (open to a restricted group of users).

ARCHIVING DIGITAL PRODUCTION

Most YLE radio programmes today are created in various CAR (computer-assisted radio) systems that integrate planning, text and sound editing, transmission, and reporting to various collecting societies. If a programme is going to be preserved in the archive after transmission, ideally it should be possible to move the sound files and the accompanying metadata online to the archival system.

In theory, this should be simple, as everything is already in a digital format. In practice, this proved to be the most difficult part of the archival system due to differences between and changes in the CAR systems.

However, parts of the system already function and the general principle is simple. The decision to archive a programme is made by its producer. After the programme has been transmitted, the producer commands the CAR system to send it to the archive. Metadata such as title, content description, duration, and names of participating individuals are converted into XML files and sent to the catalogue database. At the same time, the sound file is sent to the digital mass storage. The archive system links the sound file to the metadata, creates a bit-reduced pre-listening version of the sound file and tells the BRS SEARCH user interface that the programme can now be found in the archive. When the system is complete, the archive will monitor all incoming programmes on a daily basis, and the accompanying metadata will be checked. Typing errors will be corrected, personal names will be subject to authority control, and subject headings added. It is also expected that archivists will give feedback to producers on the quality of metadata received.

Today, YLE transmits about 200,000 hours of radio programmes every

year. Originally, we planned to preserve about 5% of them permanently, although changes in programming schedules and the addition of another radio station mean that we may have to revise our plans.

DIGITISING ANALOGUE MATERIALS (AND CDS)

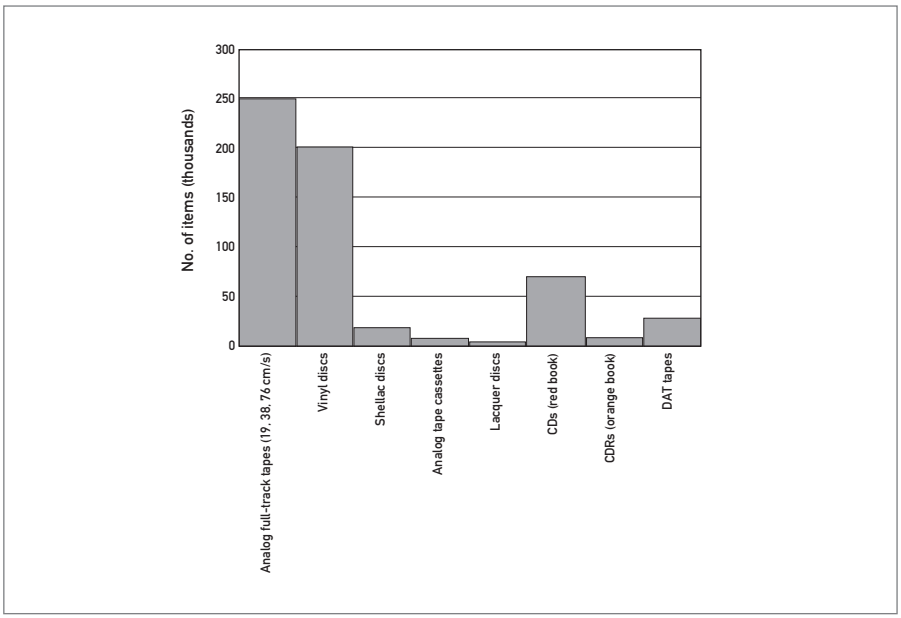
As we have mentioned, there are about 600,000 radio programmes and commercial sound recordings in the YLE archives. Moving these collections into a digital archive involves both philosophical and technical problems. From the practical viewpoint, each format presents a different set of problems. From the philosophical viewpoint, radio programmes and commercial discs are quite different.

The YLE record library has about 300,000 different records (many in duplicate) in various formats. Altogether these records contain at least 1.5 million 'tracks'. These items are not unique, that is other copies exist.

Last year YLE broadcast more than 800,000 pieces of music. If we study the frequency of radio plays, we will find that a relatively small number of 'evergreens' and 'hits' are used daily or weekly. The most popular titles in our library were broadcast more than 500 times last year. There is a much larger number of records which were played several times, and more than 100,000 titles which were played once only. On the other hand, 1.4 million tracks in our library were not used at all last year, although they may be used in the future.

For the convenience of our producers, we will naturally attempt to store the most popular titles in the digital archive so that they are always available. If a title will be played only once, and it is available on a CD, it will be much more economical to use the original CD. Therefore, for seldom-used material, digitisation is only necessary if the title is available only on vinyl or shellac.

At the moment we are concentrating on the music we estimate will be most frequently used in broadcasts. Most of this music, even if it is older, is already available on CD. Hence we have concentrated our efforts on converting the most important part of our CD collection to the digital archive, and we are happy to report that we have developed a method that is, in our opinion, well integrated into our cataloguing process. The digitisation of vinyl and shellac records will have to wait.



RECORDED MUSIC IN THE DIGITAL ARCHIVE

- Only the most frequently used records need to be digitised.
- Emphasis on old and new material already on CD (and music produced and owned by YLE).
- Vinyl and shellac digitised on demand, restoration if required.

'Back catalogue'
- Today there are 40,000 tracks in the digital archive.
- We intend to 'digitise' 40,000 additional old tracks.

New acquisitions
- Each year we buy 6000 new CDs (= 50,000 tracks).
- We plan to 'digitise' 10,000 new tracks (20%).

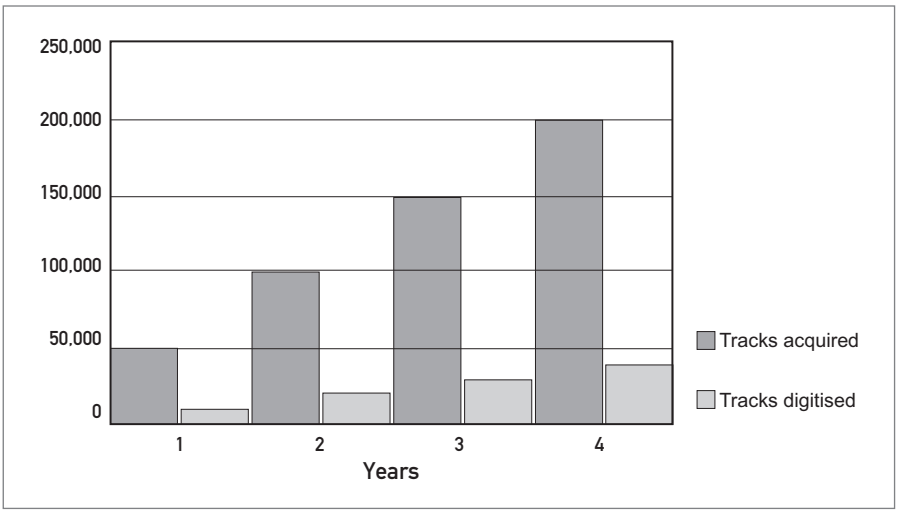
Capturing music from CDs and moving it into the digital archive is quite simple; unfortunately, it does not work with copy-protected CDs. Unlike recorded music, practically all the radio programmes in the archive are unique in the sense that there are no copies of them in other collections. If we do not take care of them, they may be lost forever. Yet it is clear that it is possible and necessary to make priorities within the 300,000 tapes in our collection.

THE RADIO ARCHIVE IN BRIEF

- Around 300,000 radio programmes.
- Unique material: if we do not save them, they will be lost.
- Huge quantity makes selection and prioritisation necessary.

1. The archival collection
- 15,343 selected programmes from 1935-1970, detailed documentation, copied on new stock .

2. The music collection



- 10,000 musical performances owned by YLE (Radio Symphony Orchestra etc., 1950-2000).
- Detailed documentation.

3. The working collection

Unselected material with minimal documentation

- 1950-1954: 218
- 1955-1959: 2666
- 1960-1964: 3802
- 1965-1970: 15,329
- 1971-2002: c. 200,000

4. Digital carriers: DAT, CDR etc.

Since 1961 the radio archive has selected valuable programmes and inserts for a special archival collection. The tapes have been copied onto new archival-quality stock and documented in detail. This collection today consists of about 12,000 analogue and a number of digital tapes, with a total of 40,000 programmes or inserts, dating from 1935 to 2000. When we installed our first capturing station, Quadriga (http://www.quadriga.com/home.htm), a year and a half ago, we decided to start with this archival collection, and to date we have digitised 4700 tapes in numerical order. We expect to finish the task by 2005.

The archive also has a large collection of music productions, many of them featuring the Finnish Radio Symphony Orchestra. Many of them have already been transferred onto CDR, but there are still about 10,000 analogue tapes, each containing one piece of music. We shall soon start digitising this collection with Quadriga as well.

The radio archive also has a huge collection of so-called working tapes, the earliest dating from 1950. This collection consists of about 250,000 analogue tapes and about 50,000 digital carriers. We recently acquired a NOA capturing station (NOA Audio Solutions, http://www.noa.at/), which will be used to digitise this collection.

RADIO ARCHIVE DIGITISATION PLAN IN BRIEF

The first stage:

- Digitise 15,343 tapes in archival collection with Quadriga or NOA (4700 already digitised).
- Digitise 10,000 tapes in music collection with Quadriga.
- Digitise 6468 collection tapes from 1950-64 with NOA.
- Other materials digitised on demand.

The oldest tapes are clearly approaching the end of their life-cycle, and if they are to be preserved, they have to be digitised soon. We still need to study the condition of the tapes more closely, but we feel that all tapes recorded before around 1965 belong to the endangered group, and we have decided to digitise all of them as soon as possible. More recent tapes would seem to have at least ten years of lifetime left. Fortunately, our engineering department always insisted on using expensive, high-quality BASF and AGFA tapes, so we do not have the problems with decaying tapes from the 1970s that are so common in American archives. The number of tapes dating physically from 1965 or earlier is only about 7000, so the task is limited, but on the other hand many tapes suffer from a number of problems: bad splices, unusual speeds such as 76 cm etc., in addition to the generally declining quality of the tapes. This will inevitably slow the process down.

This still leaves us with more than 200,000 tapes to digitise, dating from 1966 to 2000. Of course we would like to start with the tapes most likely to be used in broadcasts. Most of our 'greatest hits', such as war documentaries, have already been moved into the archival collection and will be digitised in that group. However, it is clear that we will also need other important programmes such as interviews with important statesmen and great moments in sport. As an experiment, we have already digitised all news commentaries from 1993. If there seems to be a demand for material from 'ten years ago', we will continue with the years 1994, 1995 and so on. But how do we continue from there?

This is a question which we intend to put to all our producers as soon as the digital archive has been up and running long enough for them to begin to have some idea of the system's possibilities. Here again we should be prepared for changes in demand. If digital radio starts growing as predicted, we are likely to see the creation of new, specialised radio channels such as a sports channel which could feature repeat broadcasts of past Olympic games, or a religious channel. We already have enough material for such channels.

From a technical viewpoint, the material most in need of digitisation consists of open-reel analogue tapes of various sizes. So far our efforts have concentrated on finding a method for digitising such recordings as effectively as possible. In 2001, we started using a Quadriga workstation for this purpose. As the archival system was at that time not yet ready, tapes were temporarily copied on AIT (Advanced Intelligent Tape) tapes, and are now being transferred to their final destination.

After analogue tapes, we will have to consider DATs and CDRs. Although they are considerably younger, our digital carriers already show all kinds of problems: tapes that will not start, CDs that have BLER rates (BLock Error Rates) far beyond acceptable standards and so on. For CDRs, we can use the same digitisation process that we are using for commercial CDs. For DATs, we have used an outside supplier to transfer 700 DAT tapes to sound files on AIT tapes, which we have then moved to the archival system ourselves. These tapes were already catalogued in our database, and each tape had a bar code with the tape's UID. Technically, the experiment was successful, but we still have to evaluate the economic aspect and determine which procedure is more effective – outsourcing or doing it ourselves.

RESTORATION

As a general principle, a sound archive aims to preserve the recordings in its collection in a condition which is as close as possible to the original. However, our main focus in preservation is on the signal rather than the carrier. All magnetic tapes will crumble with the passing of time, but we can hope to preserve the sound that has been recorded on the tapes.

When professional-quality tapes from the past three decades are digitised, the process is fairly automatic. There is no reason to alter or 'remix' the recordings, and if a future producer wishes to change, say, a concert recording made in the1980s, we see this as post-production which may also be archived one day but will not replace the original.

Images of The Hague

However, when it comes to vinyl or shellac records or magnetic tapes from the 1950s or 1960s, there are frequently problems which limit their usefulness in current production: clicks, dropouts, hiss, mains hum and so on. There are many efficient methods of handling such problems in the digital domain; for instance, we have used Sonic Solutions NoNoise with good results for several years.

There is always a subjective element inherent in restoration. One has to wage a constant battle between removing as much noise as possible and changing the original signal as little as possible. For this reason, we treat the restored sound file as an alternative to the original sound rather than as its replacement. In practical terms this means that, when we decide to digitise and restore an old recording, we store both the untreated sound with the noise inherent in it, and the restored files. Our customers can hear both and decide which one they will use.

We also know that in ten years' time restoration methods will probably improve considerably. We know from experience that it is generally useless to attempt further restoration work on a recording that has already been treated. In ten years' time we might also have better equipment to play our old discs and tapes, so we intend to keep them as long as possible after they have been digitised.

THE FUTURE

We have learned a lot from our experiences digitising this archive and now need to use this experience when we look to future digitisation efforts. It is of utmost importance that the YLE Digital Sound Archive continues to provide users with a means of quickly finding and accessing high-quality material if they are to fully exploit this rich and varied resource.

Listen to two examples of the digital sound archive: a German language interview with the composer Sibelius and a rendition of The Magic Flute:

LINK Sibelius

LINK The Magic Flute

Sounds © YLE Digital Sound Archive, 2003, http://www.yle.fi