D 2.11 Research Agenda

Network acronym:	**EPOCH**
Network full title:	**European Research Network on Excellence in Processing Open Cultural Heritage**
Proposal/Contract no.:	**IST-2002- 507382**
Start date of project:	**15/03/2004**
Duration:	**4 Years**

Document Description

This document provides an update on the EPOCH WP2, Task 2.5 Research Agenda work as carried out since May 2006.

Authors:
David Arnold, University of Brighton
Guntram Geser, Salzburg Research

Preparation Date:	**17th January 2007**
Status:	**PUBLIC**
Revision:	**Final**

Contents

Executive Summary

This report presents an analysis of the current state and future directions of research in ICT specifically inspired by the needs of the cultural heritage sector. Starting from an overview of the place and challenge of Cultural Heritage in ICT research the report first highlights the need for close inter-disciplinary working between cultural heritage and ICT professionals if resources are to be focussed on the most appropriate ICT research. There follows a description of five hypothetical future scenarios for particular work situations involving cultural heritage professionals and the areas of ICTs that require additional research to support the scenarios envisaged are identified. The needs are then grouped over the five scenarios and the specific research issues elaborated.

The five scenarios are:

Scenario 1: Site excavation – Virtual Excavation Support Teams

Scenario 2: Community museums – Hybrid Eco-museum & Community Memory

Scenario 3: Educational experiences – Heritage classrooms without walls

Scenario 4: Heritage management

Scenario 5: Environmentally endangered sites – Large-scale industrial heritage site

The bulk of the report then discusses the research implications of the scenarios structured according to a number of business processes typically undertaken by cultural heritage professionals. These include:

- Data Capture over a full range of Cultural Heritage data types
- Documentation
- User created content
- Intelligent Tools
- Digitisation of Legacy Metadata
- Search and Research, including Semantic and Multi-Lingual Processing
- Visualisation and Presentation
- Specific Issues for Web Access and Dissemination
- Mobile, Distributed and Networked Systems
- Long term preservation and upwards compatibility

There follows a section discussing the mechanisms, pace and time horizons for the adoption of the results of ICT research in practice and the main report then concludes with a description of the next stages in achieving an agreed research agenda.

Appendices provide a list of the participants at a number of meetings in which aspects of the research agenda were discussed during the gestation phase of this report and an appendix commenting on the barriers to exploitation of ICT in Cultural Heritage.

1. Introduction and overview

1.1. Definition of relevant ICT research for Cultural Heritage

The Cultural Heritage sector is wide ranging and involves many different types of organisations. If these are to remain in existence it goes without saying that they must have viable "business models" – that is to say that their available sources of support match or exceed their resource needs. In fact the range of organisations spans:

- Memory Institutions (monuments, sites, museums, libraries, etc.)
- ICT commercial enterprises (often SMEs, but can be divisions of larger corporations) providing services to cultural heritage institutions
- Not for profit organisations (voluntary sector organisations such as preservation trusts, arts organisations, etc.)
- Local authorities and other public sector custodians of heritage
- Owners/custodians of private collections
- Organisations providing services involving cultural heritage venues (research organisations, tourism, education, etc.)
- Individuals and Organisations who exploit cultural content in products (publications, entertainment, souvenirs and other artefacts)

Many of the ICT uses of these organisations are generic to any enterprise – accounting functions, banking, payroll, invoicing, communications, etc. – and some will relate to generic business situations – point of sale, ticketing systems etc. This document is not targeted at these generic systems, but at the ICT applications which only relate to cultural heritage data and circumstances. In some cases there are overlaps – notably in the use of systems with cultural heritage data, where the systems may be suitable for processing data from other sectors.

As with many application areas the professionals working in those areas must have ownership of the responsible use of technologies. In this case Cultural Heritage (CH) stakeholders must find the right balance between different goals. For example, the conservation of fragile Cultural Heritage may conflict with the wish to provide better access for the citizen or, choosing to present diversity and richness of Cultural Heritage perspectives in contrast to choosing to present and interpret from limited perspectives in order to present clearer messages.

Information and communication technologies offer much potential and many opportunities to support CH institutions in such situations, but this requires a certain degree of "e-readiness" of the CH institutions (e.g. with respect to ICT-affinity, available budgets, technical skills, etc.).

Moreover, there is a need to take into account the full spectrum of CH "business processes" that could benefit from ICT applications. This spectrum includes processes in the following areas:

- Data collection /recording – the capture of the primary data about cultural heritage
- Organisation, Structuring, analysis, and interrogation – search and navigation of primary data to create cross-reference information, classification, indexes and knowledge.
- Cultural Heritage research (e.g. humanities, local/regional/national/European history, etc.) where ICTs offer potential intelligent tools,
- Interpretation and Communication – spanning from the interpretation of evidence to produce reconstructions of sites, through to producing explanations of the cultural significance of artefacts (e.g. contextualisation, different perspectives, etc.)
- Preservation and archive of records and secondary data
- CH site and resources management (e.g. monitoring and preservation),
- CH on-site and online visitors/users (e.g. requirements of researchers, professional, general public, etc.),
- CH exploitation/valorisation and regional development agendas.

It should also be noted that for many applications there still exists a considerable lack of available knowledge about best practices and benchmarks (e.g. regarding the impact of digital services delivered by local, regional and national CH institutions).

This document concentrates on research needed to realise the potential of ICT support to the processes that are specifically linked to the handling of cultural heritage data.

1.2. Potential benefits of a common Research Agenda

The EPOCH activity in defining a Common Research Agenda is an integral part of fostering the development of a European Research Area in research on the support of ICT for Cultural Heritage. The principal benefits of a Common Research Agenda on research and technological development in cultural heritage ICT are that it can

- provide cues for RTD investment decisions by funding agencies by identifying critical research strands, current limitations and gaps, and ways to leverage RTD investments by coordinating research activities.

- be a useful tool to mobilise stakeholders and form project consortia to target identified key RTD challenges.
- provide the members of the ERA community with a longer term sense of purpose and direction to research planning, superseding / overarching the short-term priorities of individual funding agencies and research programs.
- stimulate monitoring progress along the way, and help to identify required related activities, such as provisions and measures for fostering the uptake and broader use of research by technology companies and cultural heritage organisations.

In this way the existence of a Common Research Agenda is expected to also foster a better cohesion of the communities involved, yield more efficient spending of the available funding, and result in better and more sustainable ICT based solutions.

1.3. Socio-economic relevance of the Cultural Heritage sector in Europe

The EPOCH project includes an activity which seeks to evaluate the socio-economic impact of cultural heritage and the contribution to that impact that may be realised with appropriate use of technologies. This activity has resulted in a series of publications and events ([HI, 2005], [HI, 2006], [McLoughlin et al, 2006]).

It is clear that cultural heritage is considered to have intrinsic value for those to whom the heritage relates. For example, as recent study has been conducted of the value placed on the Brighton Pavilion by residents of, and visitors to, the city. This building, which is less than 200 years old, is nevertheless iconic and considerable economic value is placed on it, whether or not those surveyed actually visited the site.

Cultural Heritage also provides the backdrop to how we define the values of our communities and is an important motivator to the Tourism sector. The Tourism sector, however it is defined (and there are several variants) is worth many billions of Euros in annual turnover in Europe. Similarly Europe spends many billions annually on the education of its citizens, a component of which is in their education as citizens engaged in their history and heritage.

There are many direct and indirect values associated with cultural heritage assets. These vary from the physical environment, where property prices, inward investment potential and the general feeing of wellbeing of the citizenry have all been identified as positive impacts of valued heritage environments. The socio-economic impacts of these elements are, of course, difficult to measure, but are provably positive. However even though some of the positive contributions of cultural heritage may be difficult to quantify there is no doubt that the combined value of physical cultural heritage at the macro level across Europe runs into billions of Euros, even when only those aspects which can be quantified are considered.

Europe is the region most visited by international tourists. The World Tourism Organisation in its “Tourism Highlights. Edition 2006” for the year 2005 reports international tourist arrivals in Europe of 441.4 million, which is a share of 54.8% of the world market of 806 million. Of the international tourism receipts in 2005 of 547 billion, Europe earned about 280 billion (51.2%). (WTO 2006)

One major factor of the attractiveness of Europe is of course the cultural richness of its countries. For example, Europe has a larger share of UNESCO’s world heritage list than any other part of the world, with well over 300 entries of cultural and natural significance.

In a 2003 study [Ecosystems, 2003] “Historic interest” was cited as the 5th most common reason for the choice of tourist destination (by 32% of those surveyed), behind (1) “Scenery” (49%) (2) “Climate” (45%) (3) “Cost of Travel” (35%) and (4) “Cost of Accommodation” (33%). The citation of scenery here may also have a cultural heritage component. Looking at this list access and condition of cultural heritage is perhaps the influence most susceptible to active management by national and regional authorities, making cultural heritage an extremely important component of any strategic plan to benefit from potential tourism revenues.

Europe’s patrimony is an important asset both in cultural as well as economic terms. In fact, cultural tourism is good business, especially also in the new EU Member States. For example, a quarter of Cyprus’ gross domestic product (GDP) comes from tourism. Even in industrial countries like Germany and France, tourism accounts for 8% and 7% of GDP respectively. (cf. EC, DG Research 2004, 5)

What is harder to quantify is the impact that individual investment choices may make on the value. In part this is because the incremental effect on value may be difficult isolate and in part it is because the macro level economic advantage may not be directly realised by those organisations which make the investment.

Many organisations directly involved in cultural heritage are *social enterprises* where profit is not the primary motivation. This appears to extend to many of the SMEs involved in the sector where they are commonly led by individuals who are passionately committed to the cultural heritage sector and regard profit as necessary in order to survive but not “the reason for being involved”.

There remain a need for further work to understand the socio-economic importance and impact of cultural heritage and of ICT investment in it. There is a potential parallel between the attempts to evaluate return on investment in ICT in the Cultural Heritage sector and the attempts to quantify investment in other sectors – perhaps most notable the investment in Computer Aided Design in the 1970's.

It was noticeable that contemporary attempts to attribute direct economic advantage to the development of CAD systems in sectors such architectural design usually struggled to find direct payback on investment. The "easy" arguments of more efficient design completion rarely seemed to be backed by evidence yet it is clear that thirty years later the architectural profession has adopted ICTs in widespread areas of practice and it would be virtually inconceivable for a sustainable architectural practice to ignore the potential of ICT applications in support of design management, visualisation etc. The true benefits are significantly different to those that were originally envisaged and are phrased in terms of the ability to be responsive to clients and assistance in designing sound structures with more freedom of form than manual methods might have allowed.

Similar predictions can be made about the likely impact of ICTs on the professions associated with cultural heritage. It is likely that applications of ICTs will significantly alter the day to day working practices in these professions and enable whole ranges of previously unknown working practices which meet the strategic objectives of cultural heritage professionals in different ways. It is also likely that this evolution will take 20-30 years as the working practices and professional education of future generations adapt to the new potential. This should not be a surprise – it may well be a property of developing people's ability to adapt to fundamental paradigm shifts. The same processes can be seen as the mature computer games industry (that most techy of new sectors) evolved over perhaps 25-30 years from the initial interactive games to a mature industry sector where the consuming public were familiar with standardised paradigms for computer games usage.

In the cultural heritage sector these developments will continue to require nurture and effort should be expected for sometime to come in developing understanding of effective working practices and spreading the education and training required to change the profession's practice.

1.4. Towards a useful R&D matrix based on CH business processes

Recent discussions have seen an ongoing development of a framework or matrix for a Research Agenda that could inform current and future research and development of CH ICT. A major result of the discussion is that that the most useful approach to provide a framework where the ICT research considered is clearly of direct importance to cultural heritage activities is to concentrate on the notion of "business processes" within a number of areas of CH activity.

The frame work in Figure 1 was used to structure the debate at the EPOCH research agenda workshop and has proved a useful reference in considering questions such as the degree to which the research envisaged is actually part of generic ICT research in support of any applications or only requires as a specific consequence of enabling ICT applications for Cultural Heritage.

Level	Data Sources	Process	Information Flow			Process
Digital Frontier:	• Field recording • Artefact digitisation (2D, 3D, Dynamic)	Artefact Capture	↓	|	↑	Digital Artefact presentation
Asset processing:	• 3D Data (e.g. 3D artefacts (with documentation)) • 2D Data (e.g. Images (with documentation)) • Dynamic data (e.g. video and audio content (with documentation))	Data Transformation	↓	|	↑	Asset selection (e.g. index based search)
Semantic Processing:	• Collection documentation • Legacy metadata systems	Collection/linkage – metadata formation	↓	↔	↑	research/search and interrogation; experience authorship
Underpinning systems						
	Trans-lingual systems,					
	Data Management (Storage Archive and Retrieval, databases),					
	Web technologies					
	Graphics and Interactive Systems					

Figure 1.1. The discussion framework relating CH Business Processes to ICT layers.

Yet, in order to develop a systematic structuring and coverage of technologies several relevant axes of research and development need to be taken into account.

In addition, different time horizons for the research and development must be set that imply different levels of achievement in terms of new functionality.

Furthermore, the structuring should cover the processes from underpinning infrastructure to upper application layers. Common infrastructure is likely to form a layer of functionality which can be assumed to be available to support computation/interaction in any cell.

In this document emphasis is placed on the aspects of an agenda which address digital methods in cultural heritage relating to monuments, sites and museums, which is the focused remit of the EPOCH NoE.

We will also include analysis of the areas which overlap between digital collections and digital libraries, particularly in view of developments such as the European Digital Library initiative [European Commission, 2005] and the MICHAEL project [Rossella, 2006]. This is probably best addressed through consideration of the metadata area, although there are different implications in the business processes aspects.

Similarly we will collect views on the challenges of mass digitisation initiatives – in fact we may well need to predict the tools which will be needed in the longer term to cope with "recently generated legacy data". Many of the practical challenges here are in the effective deployment of effort.

1.5. Research needed for understanding operational situations

The above brief presentation addresses primarily the development or invention of technological solutions and processes, based on a perceived or anticipated need. This might be seen as corresponding to the "hard" view of system science – a view that tools can be devised which will satisfy the requirement and hence solve the problem.

However it is demonstrably true that the adoption of technological solutions rarely if ever follows the designed path and their significance needs constant re-interpreted as they become deployed within wider demands and contexts of CH business processes, and as they are confronted within wider (human) systems (business processes or user experience). The research agenda ought also to contain components of research addressed at other aspects of delivering the potential gains.

Hence, in order to deliver the Research Agenda we must also carry out research on particular operational situations. Such research includes:

- Usability research: This includes tool usability for CH professionals and others,
- Acceptability and cultural sensitivity guidance: This includes research on characteristics of successful communication/dissemination techniques, as well as frameworks for cultural sensitivity in interpretations
- Asset management including processes for IPR and copyright maintenance; licensing and secondary use policies.
- A related area are models for different software and systems licensing (open source, licensing etc) which, however, are less likely to have a sector specific component,
- Effective development of sectoral capacity, e.g. education and training (for example to overcome current misconceptions of 3D technologies and lack in use by archaeologists),
- Measuring and monitoring different types and levels of impact – socio-economic, cultural, sectoral contributions etc. at site, city, region, national and European scales. Interaction with policies and instruments/processes of policy making

1.6. Use-inspired basic research – “Pasteur’s Quadrant” and CH ICT research

Research in interdisciplinary fields such as ICTs for Cultural Heritage can be viewed from a variety of perspectives. In some areas basic, new computer science result must be developed specifically to address challenges which are unique to cultural heritage. In other areas the best of generic computer science research results must be applied to cultural heritage situations, potentially creating novel working methods for Cultural Heritage professionals. Such a range of situations can be understood neither as purely basic research in ICTs nor as a field of purely applied research. The research conducted is a mixture of basic ICT research targeted at enabling ICTs to solve specific problems for ICT applications.

EPOCH promotes a higher degree of inter-disciplinary use-inspired basic research. Such research increases the understanding of basic research issues in ICT and, at the same time, allows the development of improved technology for purposes that are specific to CH.

Use-inspired basic research has been promoted in many ways of recent years as a means of ensuring that publicly funded research is firmly based on providing solutions that have exploitation potential. Donald Stokes’s book “Pasteur's Quadrant: Basic Science and Technological Innovation” [Stokes, 1997] provides a framework to set the different types of research in context. Stokes analyses the relationships and his evaluation merits more detailed discussion from the viewpoint of cultural heritage ICT research.

In “Pasteur’s Quadrant” Stokes suggests abandoning the dichotomy between basic research and applied research and reconsidering the notion that “basic research” is the prime motor of scientific development and longer-term societal progress.

In the United States, this notion of basic research was stressed by the extremely influential federal report “Sciences, the Endless Frontier” which was released in July 1945 by Vannevar Bush in his role as President Franklin Roosevelt's director of the Office of Scientific Research and Development.

The emphasis on basic research was informed by the notion that this type of research is the starting point of a linear progress from basic to applied research and from applied research to market-orientated development of solutions. Moreover, the idea was that basic research should receive a larger part of public funding because it may not be able to ensure adequate levels of funding in the marketplace.

Research inspired by:		Considerations of Use?	
		No	Yes
Quest for fundamental understanding	Yes	Pure Basic Research (Bohr)	Use-inspired basic research (Pasteur)
	No		Pure Applied Research (Edison)

Figure 1.2 Stokes' Quadrant Model of Scientific Research [after Stokes, 1997, p73]

Stokes argued that scientific research should not be conceptualized as a linear progress and introduced four quadrants of research (Figure 2):

- In a first quadrant he placed pure basic research which is understood to be inspired by the quest for knowledge but not by potential use. A paradigmatic example for this type of research is the physicist Niels Bohr who worked on a model of the atom.
- In contrast, a second quadrant is reserved for pure applied research which is conducted to develop practical solutions and marketable products. Stokes example for this type of research is Thomas Edison and his work on electric lighting, sound recording and many other marketable, practical innovations.
- A third quadrant contains scientific work that is neither overtly theoretical nor directed at products. This work concentrates on the exploration of particular phenomena or the development of a taxonomy or other classificatory work. Rather than advance scientific knowledge or develop market-orientated solutions, the focus is more on already well understood research problems or formalising existing knowledge or academic practices (e.g. handbooks or guidelines).
- The final fourth quadrant is reserved for use-inspired basic science. This is understood to have potential practical utility, but researchers who conduct such research do not lose sight of the goal of advancing scientific understanding. The paradigmatic example here is the work of Louis Pasteur. Stokes suggested that "Pasteur's quadrant" should receive most of the interest in national research policies and public funding, providing a combination of advancing knowledge and potential exploitation and return on investment.

The notion of use-inspired research has significant implications for how scholars conceive of research which may face some tensions with current academic research cultures. In fact, if researchers concentrate on basic research they will usually do so within the confines of specific "pure" scientific disciplines that have their list of research priorities and established review and reward mechanisms. These priorities and rewards do not particularly foster considering practical, societal and policy-related considerations.

Typically research which can be located in Bohr's quadrant has received the highest prestige in peer assessment of the quality of research. Given that exercises in assessing quality often purport to be based on assessment of novelty, rigour and impact, this might be considered an odd result, but there is no doubt that such attitudes have a material impact on the research which is valued and hence on the behaviour and careers of professional researchers. In this context, attempts to become more "use-inspired" may be considered to be misguided, despite the obvious link between research in Pasteur's Quadrant and some degree of intended usefulness of the results.

For example, in EPOCH's "State of the Union" survey [EPOCH, SOTU, 2006] one researcher who addressed university research assessment criteria in the UK reported that "it would be hard to place 'Intelligent heritage' in the Research Assessment Exercise (RAE) and hard to persuade practitioners that they ought to be associated with it. In the 2000 RAE, heritage policy and research was specifically excluded from the RAE as being neither archaeology nor management, nor computing science nor politics. In simple terms, anyone caught doing 'intelligent heritage', heritage policy or applied computing is likely to be sidelined or dismissed in order to enhance an institutional response to RAE." (William Kilbride, Archaeology Data Service). Hence, there is a need to consider more deeply the values that prevent CH ICT becoming a more important interdisciplinary field of research.

In fact in this research agenda there are aspects which can be considered as lying in each of the four quadrants of Stokes' diagram. In ICT, basic research can be targeted at underlying theory or indeed, in Stokes's quadrants, research aimed at solutions which are independent of specific applications. Generic technologies, which are mentioned in various places here, would fit this later definition of basic research.

The majority of the basic ICT research topics would fit in Pasteur's Quadrant, including all of the research targeted at intelligent tools. The argument in favour of conducting research in this quadrant is that different basic research is undertaken with the limits on available resources. Whilst Kilbride's perceptions of the potential treatment of such research in the UK's Research Assessment exercise may have some truth in it, this would reflect some shortcomings of the exercise rather than the rigour, novelty or impact of the research. The academic argument favouring Bohr's quadrant is that the basic research is not sullied by potential distortion to known objectives, possibly not the personal objectives of the researcher, were they to be given free license to pursue any topic.

There are also serious motivations for research in Edison's Quadrant. Often this may be targeted at applying pure basic research results (Bohr's Quadrant) at applications in Cultural Heritage. Of particular significance is the need to prove that the generic results work effectively with the data arising in significant practical situations, which would typically exhibit special cases and data volumes which often may not have been tested in the original basic research.

Stokes states that the fourth quadrant should not be thought of as "empty" just because it is not labelled, but in his view it includes research that "systematically explores particular phenomena without having in view either general explanatory objectives or any applied use to which the result may be put." Stokes particularly mentions research into taxonomies (see section 4.8.1) and there are other areas where systematic classification and analysis is needed in this agenda. For example the systematic analysis of architectural design styles is required as background analysis to support the widespread adoption of the grammar-based procedural modelling considered in section 5.1. These do not in general add to the basic research results which prove the viability of the approach but they turn a prototype application (grounded on basic research - use-inspired or not) to a viable tool. In many ways research addressing the commonalities of different requirements in order to define best practice and standards also fit into the unlabelled quadrant as standards are intended to be useful in many different applications contexts.

2. Research Agenda scenarios

2.1. Scenario approach

In the following sections we present and examine five scenarios for practising cultural heritage professionals.

For developing each scenario,

1. A brief story presents the scenario in a specific area of the heritage sector;
2. A commentary on the scenario explores variations in the underlying issues, general trends regarding needs are described, and aspects of the scenario which will require technological research in order to realise them are identified

In the section which follows the technological research requirements of all the scenarios are drawn together and the implied technological developments that lie behind the situations described are elaborated.

Part of the overall technology need will be met by generic developments which can be expected to be the objective of developments in the wider ICT industry.

For the purposes of the EPOCH Research Agenda we concentrate on isolating those developments that require research and development which is specific to the Cultural Heritage Sector.

2.2. Foreword by Neil Silberman

The normal administrative methods of arriving at consensus in matters of official heritage policy are unlikely ever to address the most challenging and far-reaching directions that research could potentially take. Governmental heritage services and international conservation organisations already bear a significant burden in simply in keeping up with the day-to-day challenge of conserving and safeguarding the world's material heritage. But a longer-term vision of the "Future of Heritage" requires forms and modalities of recording, analysis, interpretation, and public dissemination that go far beyond those already available. The watchwords are place, network, memory, identity, and communication. Obviously technology can and will provide the context and tools for these new approaches to heritage.

From a strictly Cultural Heritage perspective, the big changes to be anticipated in the next ten years or so are unlikely to be about automation but rather about systemic changes in the way our heritage is categorized, protected, and interpreted. This will probably include the gradual dismantling of the rigid top-down structure of most heritage institutions and authorities and a much greater concentration on networking at the regional and even local level. The growing movement for "Heritage Ecology" – namely the recognition of the fragility and non-renewability of material remains from the past – is at least partially influenced by globalised information exchange about endangered sites and more general environmental and human threats. This Research Agenda seeks to balance the consumption of Cultural Heritage resources by accelerated digital capture, digital excavation, and elaborate presentation techniques with the conservation of that same heritage. This can be done through the development of an innovative, ICT-enabled stewardship of the world's Cultural Heritage resources through monitoring, enhanced and sustainable documentation, new communication networks, and more powerful knowledge discovery tools.

The scenarios below are intended to highlight major challenges facing cultural heritage professionals and others involved with cultural heritage as working methods and opportunities develop over the next 10 years or so. The challenges involve the practical requirements of several heritage subfields: Scientific Research, Museology, Education, Management, and Environmental Protection. In some ways the biggest challenges involve the embedding of technologies effectively and seamlessly into the working practices of professional disciplines that have been educated, deliberately and appropriately, to be conservative. At the same time the future directions that will become enabled by these developments have the potential to transform the both the working lives of professionals in the sector and the public's appreciation of, and engagement with, their own heritage.

2.3. Scenario 1: Site excavation – Virtual Excavation Support Teams

Early morning in a harbour town in Sweden:

Jan Anders, a junior archaeologist, did not get much sleep over the last few days. He is working at an excavation site that will be destroyed next week in order to proceed with the building of a new bridge in the harbour area.

When preparing the fundaments of the bridge, remains of an old tunnel were detected. This tunnel was used in the 16th and 17th century was used to transport goods on small boats to a market place in the centre of the town, but was abandoned after a relocation of the market place.

Jan was allowed, together with two volunteering students, to dig deeper into the tunnel which is filled with mud. This morning they found five coins which seemed to be not of Nordic origin. Furthermore, some fragments of pottery and fabric appeared of which the fabric also seemed unusual for this area. However, Jan also noticed that the ceiling of the tunnel could be rather unstable.

Therefore, he decides to ask for help from the International Virtual Excavation Agency (IVEA) to get clear about the finds, and whether he should invest the effort to stabilise the walls and try to rescue possible further interesting finds.

Using the hand-held 3D scanner from his portable excavation support set, Jan scans the pottery fragments and uploads them together with photographs of the coins onto the IVEA database.

In the meantime, the IVEA has issued a call for assistance which describes the local situation and required expertise. Within two hours a group of experts in numismatics, pottery and fabrics joins in a virtual environment equipped with a digital archaeology workbench and access to relevant databases from around the world.

For the coins a quick result becomes available through using automatic digital image comparison technology. Intriguingly, they prove to be bronze Spanish coins from the late 17th century.

In parallel pottery experts run the 3D objects of the shards through an application that suggests various likely shapes of the pottery. This demonstrates that one potential match would suggest that the pottery could be of Spanish origin and the virtual reconstruction takes this into account. But other results are inconclusive and do not verify the hypothesis. The most convincing results suggest that the pottery is a wine jug and after comparison with images of typical Nordic jugs of the 18th century the experts confirm this finding as the most probable.

Meanwhile, Marget an expert in the acquisition of chemical data assists Jan with the conservation and analysis of the fabric which is in danger of rapid deterioration. Jan is not acquainted with the infrared microspectroscopy tool available in the excavation support set. However, Margret guides him through the process as she can observe remotely Jan's handling of the tool. Margret's analysis of the data establishes that the fabric is damask and has traces of substances associated with crimson pigment.

The results are reported by Jan to the responsible Municipality department, and the decision is taken to shore up the site and explore it further; however, no further related remains are found over the following days. The 3D objects, photographs, chemical data, expert comments and annotations are assembled into a multimedia record of the excavation which is stored in the IVEA database of completed excavations. One month after the tunnel was destroyed in the construction of the fundaments for the new bridge another piece of information is added.

A city archivist who heard about the excavation sends Jan a scanned page from an old manuscript containing a few interesting lines. In the year 1712 a Spanish nobleman who visited the town with a group of merchants disappeared without a trace…

2.3.1. Commentary and Implications

Today, bringing together a multidisciplinary team of highly specialised experts is something most small, low-prestige excavations will not be able to afford. However, technologies such as the ones described in the above scenario can allow for forming ad hoc teams of experts as required in different phases of an excavation.

Virtual excavation support teams can not only bring required expertise to remote sites. This can also considerably speed up data acquisition, analysis and interpretation and thereby decrease the costs of excavations. Moreover, in the case of rescue excavations it can help in supporting local decision making and prevent a possible loss of valuable archaeological information.

Technological requirements for this on the one hand are tools for rapid on-site data acquisition and, on the other hand, a virtual environment that supports the remote experts in comparing and analysing data, exchanging opinions, and rapidly testing hypothesis.

In critical situations such as rescue excavations also the responsible heritage administration will need to be involved in the ongoing evaluation and interpretation of excavation results.

The need in the coming decade will be on technology to facilitate shifting constellations and collaborations of scholars—creating virtual multidisciplinary communities that can rise and fall according to specific research needs. Yet they will leave behind a growing body of data produced through this multidisciplinary synergy. In particular this means on the scientific level that formerly large excavations in most places are likely to become clusters of relatively small projects led by a variety of researchers and institutions. No more Great White Explorer watching the basket boys carry the dirt away, but shifting constellations of research and administrative interests that will deal with a variety of issues: scientific research, conservation, heritage administration, public interpretation.

So the key to the usefulness of a mobile device in the field will be to help the information flow from primary collection to relevant (and interlinked) repositories of analysis, conservation, management, and public interpretation. Anders, the field user in the scenario—a junior archaeologist on the staff of a busy municipal archaeological service—must and will be educated in a new way to be conscious of the various aspects of the discovery of unexpected evidence. That is he (or she) will be much more familiar with the post-excavation processes and will be educated to understand that traditional antiquarian studies are just a small part of the picture.

This scenario highlights the need to transform a view of the technology from the PDA as just an electronic field notebook of the traditional kind to become a vital communication link between the primary researcher and the growing range of information and expertise worlds involved in heritage.

Table 2.1 highlights the range of technologies already implied in the scenario presented. The broader category of working situations that this scenario represents will imply additional technologies and processes that could have been included in similar situations.

At the more generic level this scenario highlights the needs for:

(i) Mobile access to

(ii) Integrated, but distributed, resources and

(iii) Distributed expertise, informed by common views of the available information.

(iv) A variety of novel onsite data capture devices, capable of supplying different types of data and information to the collective and integrated enterprise.

Underpinning this scenario is a vision of a distributed research team with a variety of information needs, accessing data from distributed repositories from different perspectives. Their expertise is being applied to a combination of data direct from the site and that already held in the repositories. Finally there is an implication that their work will be used to influence the management of the ongoing excavation, and may change the directions of the onsite work, enabling more effective decisions to be made about site excavation strategy.

The scenario reflects a common current situation where substantial existing, but partial, documentation exists about a site and a great deal of research, currently would be done from documentary sources to aid planning before the onsite work begins. The scenario extends this process to allow use of data collected from the site to try and evaluate the relationship between the existing documentation and physical remains being uncovered and provide feedback from remotely located experts to adapt the on site management of the excavations. Such changes might link the existing descriptive text in the documentation to physical location where that was not described precisely in the documents (as is normally the case).

Other variants of the scenario requiring similar organisational and technological support might involve endangered sites and rescue archaeological investigations, whether for site development reasons or in the face of other climatic, environmental or political threat.

Table 2.1 Overview of specific technologies/applications and processes implied or mentioned in Scenario 1

Technologies / applications	Details about applications and processes
Portable excavation support set	
3D scanning technology:	Hand-held 3D scanner
Portable infrared microspectroscopy tool	Acquisition of chemical data of fabric
Virtual environment and digital archaeology workbench for a multi-disciplinary group of experts (numismatics, pottery, fabrics, etc.)	Evidence-based scientific research: verification / falsification of hypothesis
Automatic image recognition and comparison technologies	Comparison of photographs of coins
3D virtual reconstruction	Application suggests various likely shapes of pottery based on available 3D objects of pottery fragments, also taking into account information on pottery from different regions Comparison of 3D objects with images of pottery
Micro analysis of finds	Analysis of microspectroscopic chemical data (e.g. type of fabric, colour pigments, etc
Interpretation of finds	Expert comments and annotations of objects
(Semi-)automatic metadata creation	Metadata standards
Communication	
Call for assistance	Presentation of information about local situation and required expertise
Communication with in-field archaeologist	e.g. for carrying out micro analysis of finds
Communication among experts in/via the virtual environment	e.g. for interpretation of finds
Remote expert guidance in handling a tool	e.g. for acquisition of chemical data
Databases	
Database access	Access to distributed databases containing 3D objects of pottery, photographs of coins, information about chemical properties of fabric, etc.

Upload of content and metadata	Upload of digital information / objects to a central database - 3D objects of pottery fragments - Photographs of coins - Data about chemical properties of fabric
Aggregation and storage of complex multimedia object	Assembling, describing and storing: 3D objects, photographs, chemical data, expert comments and annotations Adding an archival record to a complex multimedia object
CH Management	
Decision making on an ongoing excavation	Reporting to a Municipality department in charge of local heritage

2.4. Scenario 2: Community museums – Hybrid Eco-museum & Community Memory

Maria Bauer, is a curator of an eco-museum in one of the alpine regions of Europe. Three years ago, the region had planned to build a museum, however, decided otherwise after Maria together with Max Frisch, an energetic owner of a small IT and digital media company, had presented a plan for developing a "museum without walls" for and with the people of the region.

Representatives of the local municipalities were very sceptic. However the museum concept and technical set-up proposed by Maria and Max proved to match much more closely the politicians' notion that a museum should promote the region to tourists.

The basic concept was that the whole region with all of its landscapes, local agriculture and traditional businesses, celebrations, objects of daily life, stories about places and events, cherished private objects, family albums and so forth is declared as Living Heritage of the region.

No objects are transferred to a museum, they remain where they are. No physical exhibitions of artefacts and visual representations are organised and presented. Instead, the technical set-up of the eco-museum allows for capturing, assembling, presenting and accessing digital representations and narrations of the region's heritage.

Maria's main task is to visit locals who want to "donate" objects, images and stories to the collective museum. Today, she has visited a retired blacksmith who donated some of his instruments, a family album and stories about what it meant for him, his father and grandfather, all of them blacksmiths, to earn a living in the region.

Over the last years the eco-museum has scanned thousands of photographs from family albums, scanned many unique and typical objects (3D) of the region as well as photographed landscapes, buildings, places and objects in villages and along streets and routes through the mountainous parts of the region. Hundreds of stories (written and recorded) about objects, places, and social life in the region have been collected, many of which have been collected through oral history projects at schools which directly upload the stories to the eco-museum's database. (cf. [Giaccari, 2006] on the key role of storytelling for the collective memory of a region).

A deep sense of ownership has emerged among the people of the region. This also extends beyond the local community, because, photographs, postcards, transcribed letters, and recorded stories are also donated from people around the world whose ancestors emigrated to other parts of the world.

The technical set-up of the eco-museum comprises:

- Digital objects (3D objects, photographs, stories about places and events, etc.) which are all geo-referenced.
- The digital entities are represented on a Web-based map of the region that can be panned and zoomed and for each object descriptive metadata and a direct link is provided.
- Speech recognition and text-to-speech conversion technology and semantic processing are employed and some of the metadata is automatically extracted from the stories using multilingual natural language processing technologies.
- Stories related to similar objects or events can be identified, browsed, selected, packaged and downloaded for more detailed study.
- Local historians, teachers, students, parents and grand-parents form virtual communities that engage in the study of historical developments in the region and organise virtual exhibitions that represent the past, present and likely future meaning of collective heritage from different perspectives.
- The museum also is part of a European Network of Eco-museums the members of which share digital objects and stories, explore common cultural and socio-economic topics, and create virtual exhibitions on common themes. Such exhibitions make use of enhanced machine translation and multi-lingual data processing tools
- Based on the geo-references, stories about, and historic photographs of, places, buildings and objects of the regions can also be accessed through mobile devices.
- For locals and tourists the digital resources provide an opportunity to gain a deeper understanding of the region, its tangible and intangible heritage and collective memory.

2.4.1. Commentary and Implications

Since the 1980s many regions have embraced and realised in different ways the concept of eco-museums. The term was coined by Hugues de Varine in 1971, and one of the realisations of the concepts has been the “Musée de l'Homme et de l' Industrie” in Creusot-Monceau-les-Mines which opened in 1974. (cf. [Varine 1993]; for a systematic description see [Davis, 1999])

According to the Laboratorio Ecomusei of the Regione Piemonte at present there exist about 230 eco-museums in Europe, most of which are located in France, Italy, Portugal, Spain and the Nordic countries. (cf. the maps and lists at Ecomusei.net; for a detailed presentation and discussion of the development of eco-museums in Europe see [Maggi/Falletti, 2000])

Museum doyen Kenneth Hudson about ten years ago suggested that Europe is “a giant network of potential eco-museums” [Hudson, 1996]. In fact, the concept has much future potential particularly through the use of novel technologies that allow for effectively representing the collective heritage and memory of a region.

Work such as that of the HICIRA Network (www.hicira.org) has shown clearly that local museums (and especially local site museums) are going to move away from the static displays of artefacts and concentrate on establishing the structures for the creation of long-term, sustainable local memory institutions, in which the input of the public is central. This view is supported by some policy work (e.g. the UK Department of Culture, Media and Sport's recently issued document on "Priorities for England's Museums" [DCMS, 2006]). The success of the local museum in the next decade is going to depend on how effectively it can function within a community context. It can no longer be just a "show" or a "tourist attraction," but needs to be an integral part of the community.

The local curator will, as always, be struggling for budget to sustain an environment in which the community will be recording and identifying with their own heritage. The budget to equip the museum for this role may be drawn from ticket sales and other revenue streams based on visitors (including School groups and tourists), or it may receive a revenue contribution from the community role. However there is no evidence to suggest that these revenues will be less stretched than budgets are at present and there will be an enhanced role to sustain. The key to the future is likely therefore to be as much about enhancing sharing of tools and activities with associations of museums in similar circumstances as it is about sharing artefacts and information.

In 10 years, it is likely that local museums will no longer be the poor and primitive shadows of the great national museums, but will be an active force in their own right in the heritage field – acting together in clusters to participate in original temporary exhibitions and sharing online educational programs (with the hierarchy of local, regional, and national museums becoming less rigid and less the source of all museum trends). This is the evolution that technology can and should facilitate.

In this scenario technology has to enable and maintain, in slightly different form to the first scenario, the creation of virtual communities as much as on producing virtual objects. In addition sharing of information across collections will assist in the definition and assembly collaborative exhibitions.

Alternative versions of the scenario involved trying to tie a locally discovered artefact with its distant origins and to establish the cultural historical connection. The reverse could also be true. The community might be want to locate artefacts and information concerning its own heritage that has migrated elsewhere and link this with memories in the community. If the community is one of oral traditions (e.g. in Aboriginal Australia or Africa) then the memories might be the results of generations of story-telling and require interpretation. The lost artefacts and associations might be distributed through a colonial power and via colonisation and trading routes. The scenario might actually relate to orphaned heritage (cf orphaned works in copyright terms) where both the colony and the colonial power have part of their heritage displaced and not very well appreciated. In this context the notion of digital repatriation and shared memories of the colonists and the colonised might form the basis of new and ongoing interaction which serves to help current generations appreciate their country's colonial past and identify the inheritance of that past in the country of today.

In a scenario, any specific hypothetical example which required linking of a specific significant artefact with a remote location and/or bygone time could be considered contrived. However the vision must be of a distributed environment involving integrated views of locally-held collections searchable on a common basis based on complex search criteria and characteristics (e.g. shape).

Table 2.2 highlights the range of technologies already implied in the scenario presented. The broader category of working situations that this scenario represents will imply additional technologies and processes that could have been included in similar situations.

This scenario therefore highlights the needs for:

(i) Systems for capturing, analysing and interrogating user-created content based on a variety of media including speech, story-telling, dance and music.

(ii) Integrated, but distributed, resources (both in terms of digital records and physical artefacts)

(iii) Data capture of physical artefacts

(iv) Recreation/Simulation of mechanisms from industrial heritage and the context of their use.

Extrapolating the potential demands would involve extremely complex and long term research questions, for example to define advanced search characteristics and mechanisms (e.g. mechanisms to search music or dance for particular structures in the plots of stories to detect common oral heritage which may have diverged over generations of retelling)

Table 2.2 Overview of specific technologies/applications and processes implied or mentioned in Scenario 2.

Technologies / applications	Details about applications and processes
Digitisation technology	3D acquisition, etc
Geo-referencing	All digital objects (3D objects, photographs, stories about places and events, etc.) are geo-referenced
Map-based access	The map can be zoomed and for each object descriptive metadata and a direct link is provided. Similar or related objects and stories are indicated and can be browsed, selected, packaged and downloaded for more detailed study.

Databases	3D objects Photographs Stories – text Stories – voice Music Dance Other
Metadata	Base descriptive metadata standards (Semi-)automatic extraction of metadata from written and recorded stories
Speech recognition and text to speech conversion technology	W.r.t. stories
Technologies for identification of similarity objects (not based on metadata)	- For images: Image recognition and comparison - For 3D-objects: Shape based identification
Technologies for identification of relatedness of stories (not based on metadata)	- For texts: Commonly used names (of places, objects, etc.) or common story structures - For voice: Commonly used names of places, objects, etc.
Machine translation and multi-lingual data processing	W.r.t. collaborations in a European network of eco-museum
Mobile devices and positioning technology	W.r.t. accessing geo-referenced stories and images
Virtual community and exhibition technologies	Study historical developments in the region and organise virtual exhibitions that represent the past, present and likely future meaning of collective heritage from different perspectives. Exhibitions of network of eco-museum in Europe and beyond
CH Management	Community-based

2.5. Scenario 3: Educational experiences – Heritage classrooms without walls

Julia, Paul, Philip and Veronica, 15 and 16 year old students from a secondary school, take part in a “museum without walls” project at the city museum. Together with staff from an interactive media company, a curator and the educational programme manager of the museum they work as co-designers of interaction concepts, story-boarding and applications which should allow for better mediating knowledge and learning about cultural developments and experiences in the city.

It’s the second time that the group works on what the company calls “experience prototyping”. This methodology concentrates on the potential users’ interactions with the novel applications and wants to ensure that the users’ experiences and learning processes are engaging and culturally enriching.

The students’ idea is to engage school classes in comparing the lives of young people in the city between the way it was centuries ago to the experiences of today. Students would also interview their grand-parents about their childhood and produce digital images of photographs from family albums and illustrate with objects cherished by the family and in some cases appearing on the old photographs.

Together with the staff of the media company and the museum curator the group of young co-designers develops 3D storyboards of how the future users of the “museums without walls” will interact with the information environment they conceive.

They want to make use of virtual narrators, avatars of a boy and a girl that can appear on the screen of any interaction device (e.g. a mobile, a kiosk system, a TV set; etc.) and tell stories about places, streets, buildings and objects from the young person’s perspective.

The educational programme manager suggests that the avatars should only tell their stories in exchange to a digital image and personal comments or a historic fact about what is shown on the image. (For example, this has been a major factor in the assemblage of a library of documented tourism images by the community of teaching staff in tourism studies – see www.tourismimages.org.uk). The images can be a geo-referenced and text sent in through a mobile or chosen from the museum database.

At some places in the city such as the central railway station the design team also wants to establish kiosks where visitors can meet the avatars and get information or ask them for stories by pointing to a place or street on a map, an image or certain parts of an image.

However, the students think that, ideally, all buildings, streets and objects of the city should be able to tell their history and stories about life in the city today as well as many centuries ago. There would not necessarily need to be avatars to interact with at special places such as kiosks. Instead people would wear special see-through glasses for perceiving changes of the environment through the centuries (augmented reality) and listen to the stories a building or place gently whispers into their ears.

2.5.1. Commentary and Implications

Systems for mediating cultural heritage knowledge on the one hand will need to be able to handle increasingly complex information environments and, on the other hand, make sure that the cultural experience and learning is stimulating and engaging for the users. The latter is of particular importance if the goal is not only to allow for enhanced access to digital cultural heritage resources, but also to invite users to provide their own content and stories to a learning environment such as the one described above.

With such environments the heritage sector becomes part of the so called "experience economy" in which customers of service, media and entertainment industries seek unique, meaningful and memorable experiences. (cf. [Pine and Gilmore, 1999]; [Schmitt, 2001]). Hence, cultural heritage organisations need to develop novel concepts of cultural experience and learning that inspire, engage and enrich the users of their resources, which can be achieved with greater predictability if the users participate in the creation of the experience and begin sharing own content and stories.

While such concepts are likely to make use of a new generation of "ambient intelligence" (i.e. distributed, embedded and context-aware computing and novel interfaces, in the development of effective concepts, the potential users will need to be involved in a more qualitative and effective way than carrying out some user testing before launching a new tool or service.

This has been emphasised by the IST Advisory Group with respect to future ambient intelligence systems and applications. They suggest that research and technological development will increasingly need to make use of "experience prototyping" which focuses on the quality of the users' interactions and experiences.

They write: "Requirements engineering for Ambient Intelligent systems design can no longer be seen as a task that can be accomplished through the development of scenarios and the translation of use cases into system requirements. System functionalities that generate true user experiences can only be determined in a reliable way from feasible prototypes providing proofs of concept. New approaches to prototyping are likely to be key to the successful development of AmI products and services." (cf. [ISTAG, 2003], 27-29; and [ISTAG, 2004])

Experience prototyping should enable design teams, users and clients to gain first-hand appreciation of existing or future conditions through active engagement with prototypes. This extends well beyond the kind of scenarios, use cases, requirements engineering for software design and usability studies that are in practical use today.

Historically a major defining purpose for museums has been the education of the public – a role which is continually developing and where further revolutionary change can be anticipated. DigiCULT has already explored this quite deeply (http://www.digicult.info). Cultural heritage institutions across the world are exploring a variety of educational approaches including e-learning, lifelong learning, neighbourhood, and cross-generational learning groups. In all of these, heritage education will move away from traditional heritage didactics to training kids (and their parents and grandparents!) to work with concepts and know how to contribute to the historiographical process. This scenario features a natural evolution away from the traditional school visits with groups of children equipped with clipboards and pencils making notes from the museum collection for inclusion in a home report, towards a more visionary "Heritage classroom without walls" of the future. In this future hypothesis the tools that the teachers/facilitators will need will be those that bring generations together to create and productively help the evolution of historical knowledge and collective memory. They will integrate with those of the previous scenario and with school based systems which put the work undertaken into context for curricula and learning objectives. The key will be not merely to help students memorize or mimic "expert" opinion, but to create their own perspectives—perhaps by participating in online initiatives in which local values can be stressed and local resources selected – and these must be carefully integrated into national curricula.

Following the specific exercise and follow up work the students' work might be completed by presentation of the results of their findings using a variety of digital presentation media, again linking to other areas of curricula.

The research challenges here are clearly linked to those in the previous scenario, but in this case with the added dimension that many of the tools must be usable by students of all ages, and their teachers, directly. Thus the mediation and support that the curator and museum staff might offer will be delivered on step removed by the educational establishment. In addition some of the offsite work might well be undertaken by students using a mixture of resources at, or from, home. This is both an interfacing challenge to operate with users of all ages, but a challenge in delivery which would need to be sufficiently ubiquitous and easy to allow equal participation by different socio-economic groups.

Table 2.3 highlights the range of technologies already implied in the scenario presented. The broader category of working situations that this scenario represents will imply additional technologies and processes that could have been included in similar situations.

This scenario therefore highlights the needs for:

(i) Systems similar to those required for the previous scenario for capturing, analysing and interrogating user-created content based on a variety of media but particularly speech and story-telling.

(ii) Similar integrated, but distributed, resources (both in terms of digital records and physical artefacts)

(iii) Interactive delivery systems for accessing resources from the home, based on technology convergance

(iv) Interfacing technologies enabling novice and non-specialist users to be creative with the resources, and assemble presentations drawing on and relating concepts from multiple sources

(v) Engaging presentation software and hardware technologies enhancing the learning experience

(vi) Integration between learning environments and museum resources

Table 2.3 Overview of specific technologies/applications and processes implied or mentioned in Scenario 3.

Technologies / applications	Details about applications and processes
3D storyboarding	Used to outline the interact of users in a hybrid information environment Allow for rapid "experience prototyping" focused on the quality of the users' interactions and experiences
Geo-referencing of images and texts	For location-aware provision of historic images and stories
Database of museum and user-generated content and metadata	In addition to the more widespread issues of identification of co-referencing through multiple heterogeneous, multilingual, and multi-cultural sources, user created content further complicates the automated systems by introducing terminology, vocabulary and concepts that may not exist in standardised cultural heritage thesauri, taxonomies and ontologies (see section 5.8.1). The user generated content will often be categorised and tagged by the users themselves. The creation and management of such "folksonomies" will be supported by novel, semantics-aware applications. The volume of data collected in this fashion also means that summarisation tools would be useful as a means of compressing the volume of data
Text to speech conversion technology	To convert user submitted stories in text format to natural language narration of avatars
Virtual narrators, avatars	The avatars can appear on the screen of any interaction device, e.g. a mobile, a kiosk system, a TV set; etc. Storytelling about places, streets, buildings and objects from a child's perspective
Kiosk systems	Where visitors can meet avatars and ask them for stories

Multi-modal interaction	e.g. pointing to a place or street on a map or an image or certain parts of the image to trigger narrations
Augmented reality	See-through glasses for perceiving changes of the environment through the centuries (and listen to the stories a building or place)
Ambient intelligence environment	Buildings, streets and objects able to tell their history and stories about historic life in the city

2.6. Scenario 4: Heritage management

It is 2016 and the Lord Mayor's office in London is planning how to handle the expected influx of visitors accompanying the hosting of the 2018 Football World Cup. The city had learnt much from hosting the 2012 Olympic Games and there is concern that the additional traffic at the city's principal heritage visitor centres. In 2012 experience was gained in the management of visitor demand through selective investment in visitor attractions associated with the less frequented heritage sites on the periphery of the city. Venues such as Hampton Court Palace, Kew Gardens and Greenwich had been busier than usual in 2012 but had failed to any great extent to alleviate the exceptional traffic on the Tower of London and other central venues. As a result of these experiences better computer models of visitor demand have been developed and an integrated strategy is being debated which is designed to encourage the visitors to go to the venues closest to their accommodation. The enormous numbers of expected visitors would mean that all accommodation over South-eastern England is expected to be fully booked and strategic marketing of the accommodation is being planned to minimise the likely travel between accommodation and the match venues. At the same time local investments in visitor experiences and marketing local visitor venues in packages with the adjacent accommodation is being planned as an integrated solution to spreading loads.

The planning is taking into account the mix of nationalities expected to be represented in the finals and the venues in which the teams will play new cultural experiences are being developed which address England's cultural interaction with relevant parts of the world. These experiences are being designed to emphasise positive elements of Britain's interactions with each region and downplay the many historic conflicts. Careful research has been commissioned to discover and retell appropriate stories and designers have been retained to turn these stories into engaging experiences. Emphasis is being placed on historic trade connections and on the human stories and cultural influences spread through trading. The stories are to be brought to life using large scale immersive displays and novel interactive technologies.

In parallel to the research into appropriate story lines local and regional museums are being identified and special exhibitions planned. Special collections are to be assembled to accompany the stories with the artefacts drawn from national collections. The integrated information sources developed in the last ten years are being used to identify the most appropriate artefacts and weave the information about them into the stories. The series of special exhibitions and visitor experiences is being planned to complement a full social program and other visitor opportunities include sport and recreational activities.

Today is an early design review in which the stories are being reviewed, venues identified and the associated special collections being proposed. The 10 person planning group comprises the Lord

Mayor (Ken), the Head of the Museums Service (Pauline), the consultant Historian (John), the professional writer (Andrew), the lead exhibition designer (Jasper), the Head of Tourism London (Adam), the chief technology advisor (David), the CEO of London Transport (Shirley), the Deputy Commissioner of the London Metropolitan Police (James) and a senior representative of the hotels association (Angelina).

The main purpose of this meeting is to decide outline budgets for investment and expected returns. The meeting is also to consider and imaginative proposal to exploit potential sharing of the experiences via networked systems with related complementary experiences in the competing countries. Part of the argument in favour of the investment in these systems is presented as the residual value of the technological investment after the event is over, but Ken is unconvinced by the early analysis and has asked for more socio-economic impact analysis on this aspect. The links so established may also be used for closely linked relaying of the World Cup matches…..

2.6.1. Commentary and Implications

In this scenario the region is supposed to be representative of many heritage regions in Europe where a few high profile venues attracts the majority of the visitors, but in this case a large influx of visitors to the region is predictable and inevitable. The expectation is that major wear and tear might be placed on very specific venues and the authorities are keen that the visitors have an enjoyable experience, but concerned at the potential damage to unique assets whilst other less valued venues are not used to capacity.

Much of the solution to this problem is a straight question of visitor management and transport and other capacity planning, but in this case the hypothesis is that the planning process may be assisted by tools which allow both prediction of the attraction each site is likely to present to the visitors and the degree to which this can be impacted. Thus the introduction of new exhibits across the museum system, the degree of sharing of artefacts and perhaps the ability to provide remote access to some content which is then integrated into more distributed presentations could all offer the local community leaders the opportunity to influence and spread the visitor loads and maximise the visitor satisfaction with the venue and events.

The use of new generation GIS and the interoperability with other evolving, dynamic repositories of information will help planning much as the ICTs have helped with the management of traffic patterns and solid waste. In 10 years time, there are likely to be many "heritage departments" in all levels of government, but they are not likely to be primarily staffed by art historians or archaeologists alone. The recognition that material heritage is an (endangered) part of the biosphere is going to spark a recognition that it must be managed and conserved, not just rebuilt and decorated for special events. For even if the crowds do come, what happens when the event is over?

On the community level, the shift is expected to be towards long range heritage management rather than specific event- or tourism- related promotion. This scenario is therefore exceptional in that an influx for a specific event is being planned, but the important aspect is that the situation can be accommodated and investment planning is enabled because of the continuing planning and impact assessment tools. The use of the major event (not at all envisaged as a heritage major event) was to demonstrate the inevitability of the situation where a step change in visitor numbers can give rise to political concern over whether the existing infrastructure would be able to cope. Hence the situation engenders the political will to actually engage in a more holistic view of planning the heritage assets of the whole area, rather than isolated and individual sites. Technology can be crucial in creating models of likely impact of changing economic conditions, population, zoning, etc. Heritage is not just an exploitable resource, it is non-renewable and structures and tools must be created for its effective management.

The intent of the scenario is to consider how to engineer a better distribution of the visitor flow to the more minor sites which often have much to offer, in order to relieve the pressure on the main sites. After the event the impact of properly orchestrated systems for visitor management would continue to help manage the heritage assets.

The scenario is obviously based on current issues for the City of London. In a July 2006 press announcement UK Culture Secretary Tessa Jowell cited the 2012 Olympics as a "unique opportunity" for tourism, and linked it directly with cultural heritage management issues as she launched the widest ever consultation of the industry:

"It [the consultation] suggests options for fully exploiting the benefits of the Olympics including:

> *"New links between tourism and the arts, media, and all other sectors which will contribute to making 2012 a success. This could include specially themed marketing in the run-up to the Games, highlighting individual aspects of what the UK has to offer – including culture, heritage, landscape and diversity."* [DCMS, 2006a]

It is clear that such events will generate increasing opportunity to link regional cultural heritage assets to a managed plan for handling the influx of visitors. In fact an event such as the World Cup offers potential better prospects than an Olympic Games with more separation between local events and the restricted capacity and appeal of individual events relative to the overall volume of visitors, meaning that a significant proportion will be at a loose end on a more regular basis during the tournament..

This scenario therefore highlights the needs for:

(i) Models for assessing socio-economic impact of sites, including the return on investment, likely visitor patterns etc where new developments are planned for existing venues.

(ii) Integrated information systems allowing assessment of visitor patterns and motivations.

(iii) Integrated views of the heritage resources of the region

(iv) Presentation technologies that create sufficiently entertaining and engaging experiences as to influence the visitors' choice of heritage venues to visit.

Table 2.4 Overview of specific technologies/applications and processes implied or mentioned in Scenario 4.

Technologies / applications	Details about applications and processes
Integrated CH management system	Modelling and analysis of regional visitor demand and visitor patterns Holistic integration of regional CH assets in event planning and management Provision of up-to-date information in visitor centres and through mobile information Traffic control, routing systems, access control systems, etc.
GIS technologies	Real-time, dynamic representation of visitor patterns Interoperability with other evolving, dynamic repositories of information
Mobile location-based information services	For visitor information and routing (e.g. market local attractions, direct visitors to alternative sites, etc.)
Multi-lingual data processing	W.r.t. mobile information services, real-time sharing of experiences in different countries
Cultural story telling technologies	Development of story lines Identification of most appropriate artefacts in special collections Distributed sharing of artefacts Virtual assembly of story lines and artefacts
Networked, real-time sharing of cultural experiences in different countries (e.g. Interactive TV)-	W.r.t. to sport & culture programme during World Cup
Large scale immersive displays	In large CH centres and football stadiums
Novel interaction technologies	Real-time staging of sport & culture events – e.g, interactive TV: different viewing angles, multilingual information, etc
Socio-economic impact analysis	Impact assessment tools, calculation of residual value of technological investments for CH experiences, etc,

2.7. Scenario 5: Environmentally endangered sites – Large-scale industrial heritage site

After several years of development Marek Wankiewicz is looking forward to the official opening to the public of an industrial heritage site in a country of the former Soviet block. Besides considerable investments from the region, European Union funding has been a great help in sustaining the efforts to realise a multi-functional site. The site now accommodates a community and visitor centre, exhibition spaces, a social history research institute and an environmental information agency.

The mission of the site is to narrate in a multi-faceted way the socio-economic, cultural and environmental history of the former chemical and textile factories, the origin of which date back to the beginning of the 19th century.

The environmental agency is a data centre that wants to raise awareness and understanding of the importance of sustainable development for the region's future. The long-term negative environmental impact of the former factories is used as one demonstrator for this need. Moreover, the agency agenda includes monitoring the effect of climatic changes and the planning of measures to reduce pollution on the regional level.

The agency has implemented a network of monitoring stations in several areas of the region amongst which is also a historic market town on the banks of the same river that passes the industrial site and located about 20 kilometres downstream. The traditional buildings of this town include a fine medieval town hall with particularly fine statuary. However they suffer from the combination of aerially borne deposits of pollutants and as well as acidic rain and are part of a national programme that wants to preserve such market towns in the region.

Environmental monitoring data are captured by the sensor networks and sent to the environmental agency where the data is processed, analysed and presented through GIS technology on Web-based maps as well as information displays in the visitor information centre. At the same time the condition of the stonework is closely monitored over time using highly accurate and detailed 3D scanning to detect small scale degradation of the stone surfaces, which are cross correlated with the environmental data.

The social history research institute investigates and interprets the history of the former industrial complex and contrasts it to life in the medieval town, its organisational and technical development, workers' social and cultural life, etc. As large parts of the factory archives have been lost, much of historical reconstruction builds on documents and recollections from workers' families.

Students of industrial archaeology also reconstruct machines and instruments using modelling and animation technology. The results are presented through virtual reality hologrammatic image and sound projection technologies in several halls of the factories interwoven with the stories recollected by the workers.

The researchers also track down the geographic spread and usage of the products that were produced by the factories and others that were typical of life in the medieval town. Missing artefacts, in particular, products of the factories, such as specimen articles made from textiles woven at the factory, are searched out and added to the collection, making use of databases worldwide in locating them. At the same time artefacts that typify the past of the medieval town are also tracked down.

The researchers weave the site's industrial record and stories about the preserved artefacts into the historic narrative of the region and its changing environment. One important semantic backbone of this narrative is an extension of the CIDOC Conceptual Reference Model for industrial heritage management. This extension is collaboratively developed and used by a network of industrial heritage sites.

2.7.1. Commentary and Implications

Industrial heritage is threatened all over the world. The risk of loss through destruction and abandonment is enormous. Among this rich heritage are plantations, mills, mines, forges, factories, workers' housing, warehouses, canals and bridges, harbour buildings and areas and whole industrial landscapes. Moreover, as Michael Nevell writes: "Today roughly 30% of all professional archaeology done in Britain examines archaeological deposits that include material from the industrial period (however that is defined)." [Nevell, 2006]

Louis Bergeron, Honorary President of the International Committee for the Conservation of the Industrial Heritage, writes: "Big industrial heritage sites are always at odds with their environment because of the consequences of the pollution. They are the kind of physical remains which are exposed to quick and radical decisions of demolition because of the kind of landscape they generated - which seems to be a symbol of a natural distress or of an historical failure." [Bergeron,1998].

Since the 1980s a number of industrial heritage sites have been included in the World Heritage List (at present the list includes 43 such sites). However, in general, the situation of most industrial heritage sites is problematic.

This is due to the reasons identified by Bergeron as well as the enormous financial resources that are required to preserve larger industrial heritage sites and carry out adequate programmes of re-use, historic research and cultural enhancement. (cf. [Nizhny Tagil Charter, 2003] for Industrial Heritage).

In Europe, particularly industrial heritage in the former Soviet block is endangered. For example, with respect to industrial sites in Riga (Latvia) Anita Antenišķe from the Faculty of Architecture and Urban Planning, Riga Technical University, writes: "One of the greatest challenges regarding conversion of industrial sites in Riga is the huge scale of those areas. The industries located in the city were seldom a result of local needs, they were a part of a larger economical system be it Russian Empire in the 19th century or Soviet Union in the 2nd half of the 20th century. Therefore it is not easy for the citizens to relate themselves to the industrial past." [Antenišķe 2006]

As illustrated in the above scenario, ICT can support the preservation and communication of industrial heritage in many ways. Of particular importance will be to integrate the site, its historic record and artefacts in a multi-faceted narrative. As emphasised in the scenario, this narrative should particularly also include the environmental and ecological dimension of an industrial heritage site.

This scenario is also a place holder for wider concerns in the area of preservation of any heritage in adverse circumstances, typified by the needs to engage in clean up work following periods of less than ideal industrial development. Whilst the scene is set in the former Soviet block, the situation could equally be represented by the extensive clean up operations required as part of London's preparation for the Millennium celebrations.

There are two big issues here that will only get more frightening in the next ten years: Global Climate Change and the clean-up of industrial waste, particularly in the former Soviet Union and the developing world. One of the ICOMOS Scientific Council's most interesting initiatives is the Global Climate Change Initiative. We know so little about the effect of climate change on various fabrics and materials – and there is little monitoring of any but the most exceptional sites like Venice. It is also likely that a serious effort must be made to document vanishing resources; the ICOMOS Polar Heritage committee (headed by Susan Barr of Norway) is working on this, as is the University of Ghent (http://remotesensing.lomitko.net/?id_page=61&lang=en). What should be done in that case, where unique frozen tombs are melting? Or what about a scenario where environmental clean-up workers have uncovered archaeological remains at a toxic waste dump? What can be done to preserve these resources and clean up the site? How can the linked technologies of environmental and heritage monitoring be effectively applied? There is a much larger context for this.

The hypothesis in the scenario is that a significant historic (and public) building has suffered from the effects of pollution from smoke, soot etc and been eroded by acid rain and perhaps has also been raided of artefacts, from a collection it had housed. The rural location may mean it’s some way from any industrial zone, but this is scant protection against environmental damage. The community-led re-assembly of the collection has much in common with aspects of Scenario 2. This scenario therefore concentrates on issues of ICT-enabled documentation, monitoring, restoration and preservation in a polluted landscape.

Table 2.5 highlights the range of technologies already implied in the scenario presented. The broader category of working situations that this scenario represents will imply additional technologies and processes that could have been included in similar situations.

This scenario therefore highlights the needs for:

(i) Accurate technologies for documenting the state of architectural heritage and analysing the changes over time.

(ii) Search facilities for distributed collections and other sources for locating missing artefacts. Potential linkage to police and insurance systems.

(iii) Restoration systems including potential modelling of environmental processes and digital restoration simulations as well as physical systems for material cleanup and preservation of polluted materials.

Table 2.5 Overview of specific technologies/applications and processes implied or mentioned in Scenario 5.

Technologies / applications	Details about applications and processes
Distributed wireless sensor networks	Monitoring climatic changes and the effect of pollutants and atmospheric factors on buildings of a historic village
Heritage monitoring data	e.g. humidity, erosion, material decay, etc.
Software for processing, analysing and presenting heritage monitoring data	e.g. novel visualisation methods for the impact of pollution on materials used for historic buildings in the region
Environmental impact modelling and simulation	Models for assessing the long-term environmental impact of an industrial site
GIS technology	Web-based maps (GIS), information displays in a visitor information centre
CAD and animation technology	e.g. for reconstructing how industrial machines and instruments were used
Virtual reality rendering and projection technologies	e.g. 3D projection technologies for the working of animated historic machines
Multi-faceted historic narrative	Applications should support different “lenses” on, and paths through, historic records and narratives

CIDOC Conceptual Reference Model – extension and adaptation for industrial heritage	Ontology development guidelines, preparation and implementation tools (note: there is a need for more effective tools for preparing the adaptation and implementation of the CIDOC-CRM as has become evident for example in the adaptation of the CIDOC-CRM for the English Heritage's Centre for Archaeology (cf. Cripps et al. 2004)
Industrial heritage resources management	Applications should support integrated approaches of CH resources management

3. Drawing together the threads from the scenarios

3.1. Technology needs identified in the scenarios

In the previous sections we examined a number of scenarios, each of which were currently unachievable, even if they appear tantalising close, based on the promise of demonstrations systems and the claims of their implementers.

Some of the advances needed are generic to any application sector, and will change the cultural heritage sector in the same generic ways as other sectors. However generic technologies can be made more effective in any individual sector by incorporating domain specific knowledge to develop domain specific tools and domain specific business processes. In this research agenda we will concentrate on those parts of the required developments which are specific to the cultural heritage sector and to the business processes required by memory institutions and Cultural Heritage professionals.

The five scenarios highlighted the needs for:

(i) Mobile access from geographically-remote, and probably environmentally-challenging, locations to remotely-located resources and expertise.

(ii) Integrated, but distributed, cultural heritage resources including catalogues, digital records, digitised collections of all types of cultural heritage data and management of physical artefacts.

(iii) Distributed expertise, informed by common views of the available information requiring common knowledge extraction, collaborative environments and the presentation tools (hardware and software) in which to explore them.

(iv) A variety of novel onsite data capture devices, capable of supplying different types of data and information to the collective and integrated enterprise.

(v) Systems for capturing, analysing and interrogating user-created content based on a variety of media including speech, story-telling, dance and music.

(vi) Interactive systems for accessing a network of integrated resources and expertise from the home, based on technology convergence.

(vii) Interfacing technologies enabling novice and non-specialist users to be creative with the resources, and assemble presentations drawing on and relating concepts from multiple sources. The range of functionality for these non-professional needs to be similar to that of the professionals' toolkits, but with intelligent interfaces requiring less manual intervention and assisting the novice and non-specialist users, using built-in domain knowledge to allow them to meet their challenges.

(viii) Presentation software and hardware technologies that create sufficiently entertaining and engaging experiences as to influence the visitors' choice of heritage venues to visit and/or that enhance the learning experience.

(ix) Integration between e-learning environments and museum resources.

(x) Models for assessing socio-economic impact of sites, including the return on investment, likely visitor patterns etc where new developments are planned for existing venues.

(xi) Integrated information systems allowing assessment of visitor patterns and motivations.

(xii) Integrated views of the heritage resources in a region.

(xiii) Accurate technologies for documenting the state of architectural heritage and analysing the changes over time.

(xiv) Search facilities for distributed collections and other sources for locating missing artefacts. Potential linkage to police and insurance systems.

(xv) Restoration systems including potential modelling of environmental processes and digital restoration simulations as well as physical systems for material cleanup and preservation of polluted materials.

3.2. Grouping of the needs

These needs can be grouped into some broad areas which broadly represent different types of processing digital cultural data and the specific needs arising from using generic technologies effectively for cultural heritage applications as follows:

- Data capture of many formats of data under a variety of conditions (Section 4)
    - Onsite data capture
    - Capture of Artefacts, Monuments and Architectural Heritage
    - 3D Data Capture
    - Documentation of 3D Digital Objects
    - Capture of user created content
    - Intelligent data capture tools using domain-specific Cultural Heritage knowledge including provenance and other metadata
    - Digitisation and enhancement of legacy metadata

- Search and research: Semantic and multi-lingual processing
    - ➢ Ontologies, taxonomies and thesauri
    - ➢ Multi-lingual and multi-cultural knowledge bases
    - ➢ Digital memories for cultural information integration
    - ➢ Reconstruction and simulation
    - ➢ Knowledge discovery (or "excavation in the digital domain")
- Visualisation and presentation
    - ➢ Asset management
    - ➢ Development and deployment systems for augmented reality interpretations used with replica and/or original artefacts
    - ➢ Authoring tools tailored to cultural heritage presentations, linked to digital cultural heritage assets embedded in a digital context
    - ➢ Authorship tools for cross-platform and multi-platform interactive systems (e.g. delivery via iTV, computer games machines and other domestic-level technologies, internet and location-based immersive VR/AR systems in memory institution visitor venues)
    - ➢ Tools and techniques to allow presentation of provenance, paradata (c.f. London Charter), interpretation and uncertainty
    - ➢ Adaptation and adoption of novel interaction techniques for domain specific applications (e.g. story-telling with multi-lingual, speech-enabled avatars accessing domain specific knowledge)
    - ➢ Frameworks for authorship of multi-cultural, multi-national, and multi-lingual presentations and multifaceted interpretations
    - ➢ Understanding and measures of engagement to inform authorship tools.
- Mobile, distributed and networked systems.
    - ➢ Many (some would argue mainly) generic technologies
    - ➢ Specific issues with design of system architectures suitable for integration in broader CH professionals' business processes
    - ➢ Cultural heritage component in design/implementation of CH-specific components for cross-platform systems
    - ➢ Integration and interoperability of data, coupled with the implementation of rich functionality which implements effectively CH requirements (e.g. recording of excavation data: contexts; artefact scans; images; textual descriptions; positional information; etc.) on less capable hardware (e.g. next generation PDAs).
    - ➢ Standards for cross-referencing and sharing cultural heritage data with remote sources
    - ➢ Maintaining and extending associated provenance etc. whilst extending information base

  - Interoperability with generic cross-platform applications (e.g. appropriate GIS systems)
- Long term availability (requiring long-term preservation and attention to issues of upwards compatibility)
  - Formats (Standards, Encodings, Metadata, Provenance, Paradata)
  - Business processes for long term preservation (media, regimes, security, resilience, redundancy)
  - Legal frameworks (IPR, copyright, licensing, royalties, grey literature/documentation, metadata rights, collected works, derivative works, orphaned works, etc.)
  - Business models for long-term preservation (Responsible authorities, Legislative requirements, Secure financial basis, etc)

In the sections that follow each of these lists is further elaborated with descriptions of the current state of the art and challenges anticipated in the short, medium and longer terms.

3.3. Current state of the art and research challenges

The areas requiring further work fall into a number of distinct categories:

a. Measures to achieve integration of the current partial technologies

b. Issues of deployment – achieving a critical mass of available data, expertise and adoption of the technologies.

c. Incremental improvements of current technologies

d. Fundamental research to develop new tools to achieve the more imaginative and intelligent functionality

These different perspectives are apparent in both the timescales for addressing the foreseeable issues and the nature of the work to be undertaken. The specific actions required to address integration issues include identification, adoption, deployment and further development of appropriate standards. Interoperability obviously requires that a range of compatible standards underpin all areas of developing technologies to ensure inter-operability of data and support systems. To avoid duplication of the discussion the standards discussion is treated separately.

The challenges of deployment also span all areas of effective development of future technologies for cultural heritage applications. Put simply – unless and until a critical mass of the raw materials and systems that support cultural heritage profession(s) are available in compatible systems that offer value-added to the delivery of their professional duties, new developments will remain prototype demonstrations. However it takes significant time to develop the more advanced tools envisaged and to produce the critical mass of data in interoperable formats. During this period it would be appropriate to be developing the professional environment so that in parallel to the creation of integrated data environments, the skill sets and business models suitable to exploit their potential is also developed.

Since the measures required will be equally applicable to the introduction of technologies in all aspects of cultural heritage activity, separate consideration is also given to the business support aspects of preparing and encouraging development of the sector for the inevitable changes that will accompany introduction of extensive, embedded use of technology.

One aspect of this period of ramping up the volume of available digital cultural heritage assets will be that the extensive period required to digitise legacy data will make it probable that the "ideal" format for encoding cultural heritage information will also develop. In these circumstances it is important to anticipate that even data that is being newly digitised at the moment will require further development in the future – almost as if it becomes legacy data before a widespread compatible infrastructure and associated critical mass of data is completed. It is consequently likely that significant effort will be needed in the future to produce tools to assist in adding additional material to early digitised data sets. The exact nature of these tools will depend on developments in the standards and systems over the next few years, and so the likely nature of these is difficult or impossible to predict at this stage. We therefore will not consider these aspects further here.

4. Data capture

The list included under this heading in the previous section was as follows:

- Onsite data capture including metadata
- Artefact/Collection digitisation including metadata
- Digitisation and enhancement of legacy metadata
- Capture of user created content including metadata
- Intelligent data capture tools using domain-specific Cultural Heritage knowledge including provenance and other metadata

Each of situations represented in the list of required research areas includes reference to the need to record metadata. The inclusion with each area of data capture is deliberate and intended to reinforce the message that digitisation of raw data (e.g. describing the shape and colour of an artefact) must be accompanied by recording information to qualify that data and the process used to capture them. Additional metadata will, of course, be added as more research is undertaken with subsequent analysis of whatever class of raw information is being recorded. This can vary from recording a context and later relating them to other contexts to adding subsequent results of analysis in laboratories or archives.

For the purposes of considering the types of data capture here the issues of the capture of the raw information and the linkage to metadata will be considered separately.

4.1. Activities and Processes in Onsite Data Capture

There is a requirement for better support of field data collection which comprises the planning of field surveys and excavations, the processing and structural analysis of finds data, resulting in archiving and publication of research results. The following processes can be identified:

Planning: This will typically involve accessing previously collected data and preparation of surveys and excavations based on known field data sources (e.g. literature, archival sources, databases and GIS data. Generally also issues of project strategy and project management need to be taken into account. There is thus a direct connection between the ability to link dispersed sources of prior information and planning of further work in the field.

Field Survey: This includes all data from the landscape such as topographic features, standing buildings, visible remains, field walking (intensive, extensive), documentation of geology & soils as well as geophysical aspects.

Excavation: This includes all data from the site such as textual documentation and numeric data, topographical and stratigraphical data, photography, and 3D data capture.

Finds processing: This includes processes of identification, classification, drawing (for archive and publication), weighing & counting, all typically carried out by specialists (note: there exists a particularly extensive literature on statistical analysis of artefact assemblages). Note that although field scanning of artefacts is potentially possible it is likely that, in terms of detailed recording of individual artefacts, the results will be improved by scanning under controlled conditions in the laboratory at a later date. In this case the onsite scanning will be useful for recording the find's location in context and relative to other finds.

Structural analysis: A particular research challenge here is to proceed from topographic and stratigraphic data to an understanding of land use through time. A basic approach is grouping of Stratigraphic Units, features, structures, building groups, and temporal groups.

Archive and publication: This represents the final stage of the data collection process, and the start of data and collection management; and communicating the results to other Cultural Heritage professionals.

The word "tools" is used in a variety of contexts when describing onsite data collection. These range from individual instruments used for recording specific data elements (e.g. levels, magnetic field or photographs) to integrated systems of recording using particular operational protocols, standardised forms etc.

The development of effective integrated recording systems for field data collection has been hampered by the tendency of many researchers to want to develop their own complete packages, which may include one or two innovative features, but to a large degree replicate facilities in existing packages.

There are many separate tools for different tasks involved in "conventional" recording. Hence, there is a need for a higher degree of integration of tools. This would be achieved most effectively by defining a common modular framework for integration - common interfaces to tools and common formats for the data they produce. The objectives will not be served by creating monolithic systems.

For example, a closer integration of surveying and geophysical instruments would be beneficial.

However the generation of a common modular framework is hindered because some producers of the better, extensible, tools have yet to be convinced of the advantages of publishing their API, encouraging others to continue to reinvent.

For deployment there are a number issues concerned with spreading more standardised working methods and resources – some relating to insufficient availability of tools based on agreed standards (e.g. CIDOC- Core Data Standard (CDS)) and some to standardised and agreed working practices. Regarding mobile tools the existing research questions are mainly methodological, not technical (though some developments are hampered by poor OTS technology).

Unfortunately, there exists a widespread ignorance of standards (e.g. CIDOC CDS) and for data standards, proprietary GIS and CAD de facto standards are in wider use.

With respect to guidelines on working practices some consolidation can be observed, e.g. in the UK there exists a widespread use of recording forms based on the Museum of London Archaeology Service (MoLAS) Manual [MoLAS, 1994].

There are a number of research issues relating to these tools and to the tasks listed above. For example:

- Linking field survey planning to existing data sources would be a potential research topic for improving search functionality as well as an infrastructure development need; issues here range from definition of compatible data sets, deployment to include a critical mass of historic sources and provision of search tools with a range of increasing sophistication. Research here ranges from incremental advance to potential fundamental advance.
- completing the considerable missing links in the driver architecture for survey instruments,
- further standardisation of data formats captured during excavations is required; e.g. there are considerable development issues with respect to giving structure to text and numeric data (i.e., the semantics of the information and numeric data organisation),
- in general, the CIDOC-CDS should be more widely incorporated via the use of recording tools which encode the data captured in the appropriate form,
- structural analysis could benefit from the integration of different scale data as well as easy availability of older survey data for comparison,
- 3D data capture technologies have yet to demonstrate a clear value to the excavation process on many site types, but the potential might only emerge as other technologies become used more widely and those using them become more technologically aware. The longer term challenges most commonly relate to intelligent instruments incorporating knowledge of their working environment – for example recognise object classes and adapt their modus operandi to optimally handle them.

3D techniques are most likely to be of advantage in documenting monuments and standing architectures, although one could conceive of intelligent recognition of partially uncovered finds, and it is these areas that we consider next.

4.2. Capture of Artefacts, Monuments and Architectural Heritage

The capture of artefacts covers a very wide range of situations. At a recent meeting on the nature of digitisation specialisms the following list of different circumstances was recognised:

- Digitisation of Collection Catalogues, including library catalogues. This is one of the earliest forms of digitisation for memory institutions. The original problem is to make legacy catalogues available in a digital format in conventional databases and most major memory institutions will have completed digitisation of at least aspects of their major collections' descriptions. However issues remain concerning the format and content of such catalogues which may be uneven with many national standard formats and other variations depending upon the language used and its character set. Current challenges concern standardisation of formats, perhaps requiring additional metadata, conversion tools and cross-lingual and multi-lingual search.
- Image based digitisation (digital version of historic processes to move documents onto microfilm/microfiche) concerns the digitisation of images of text works. Frequently these are later processed via OCR technologies to produce full text digitisations.
- Full text digitisations of written works allow all sorts of semantic and linguistic analysis, and producing them is commonly viewed as a solved problem. However evidence from a recent meeting of the centres of expertise in digitisation suggest that there are remaining issues concerning primarily earlier font designs with Gothic fonts before about 1840 specifically cited as presenting serious problems for a significant body of historic texts, particularly in German. Other problems concern texts in languages using characters beyond the standard English set and natural language processing in other languages. At the simplest level additional accents and diacritics complicate the character recognition, but more complex issues can arise concerning extracting the semantics in different languages.
- An additional level of complexity arises where the individual text object is itself regarded as an artefact of cultural importance. Examples would include illuminated manuscripts, early printed works and perhaps 1st editions of later printed works. Additional but similar complications might arise in digitisation of carved inscriptions, including hieroglyphics for example. In this case the artefact is clearly 3D but includes identifiable semantic information comparable to more conventional text objects. Figure 4.1 shows some examples of essentially text objects included in the context of 3D artefacts.

Figure 4.1 (a) Spreadsheet apparently carved in Stone and (b) Hieroglyphic carving at varying degrees of relief around curved pillars

In this category there are also significant issues of handling fragments of unique texts and of drawing inference from multiple partial copies of the duplicated texts (e.g. the USA Declaration of Independence or the Waitangi Treaty in New Zealand). Digitisation of documents in this category might well require more complex methods (e.g. multi-spectral scanning) and high levels of accuracy, including potential digitisation of details of surface indentation. Where handwritten manuscripts are involved there is the potential for the additional complication of recording alterations not just initially but over time. [Twycross, 2006]

- Predominantly 2D objects – images – this category concerns the digitisation of genuinely 2D artefacts, which there are predominantly photographs (since paintings and drawings have rather different characteristics described below. Collections of historic images are common place and a great deal of work has already been undertaken on defining appropriate formats. The issues here concern the need to digitise the vast numbers of images before they degrade; issues of restoration; the rapid improvements in available digitisation accuracy; and the evolving good practice in encoding associated metadata, provenance etc.

  Although an increasing proportion of images are being "born digital" different issues arise, including the provenance ones noted above and an increasing need for curatorial skills in determining the images that do not need to be archived.

- Paintings and drawings present a variation of digitisation of images with the additional features of methods of creating the images and of course the well-known use of techniques such as X-ray to show underlying structure and layers of paint. Drawings have similar properties but also include the potential to be representations of higher dimension objects (e.g. engineering drawings or architectural plans) Digitisation in these cases might include much more information about the objects represented and about the way in which the drawings were derived (sections, plans, perspectives, dimensions etc.)

- Moving image digitisation involves a separate set of issues including data volume (compression techniques) and mass digitisation needs to be undertaken in order to provide long term archive of the enormous quantities of material originally recorded on fragile media. Estimates indicate that around 100M hours of AV material including video, film and sound from the 20th century are awaiting digitisation. Given the rate of decay of the historic material this exercise is extremely urgent. The FP6 project PRESTOSPACE is targeted at this issue [PRESTOSPACE].

  Once digitised and placed in a long-term archival protocol to ensure preservation there remain significant research challenges in areas such as information extraction from the digitised content – for example analysing the material for particular content (places, people, events etc.) or automatic shot cataloguing.

- 3D artefacts are another class of object requiring specific technologies for massive digitisation. There are wide variations depending upon the scale of the objects being scanned from whole archaeological sites to individual items of micro carving. Challenges range from the opportunity (or not) to capture the objects under controlled lighting conditions, issues of scale etc. We shall examine digitisation issues for these classes of data below.

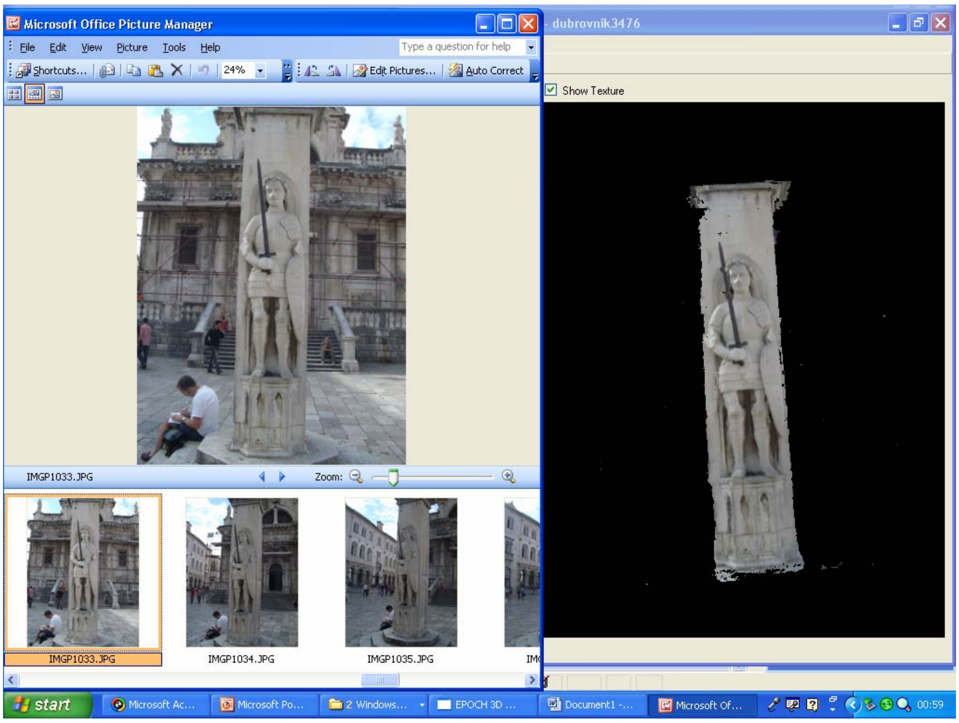


Figure 4.2 3D from a series of images using the EPOCH 3D web service

- 4D objects are not commonplace but might arise for example with the animation of 3D objects for cultural heritage.

- Structural representation (with linkages to component descriptions) also require capture though this would not typically be digitised directly, but assembled with manual intervention by the archivist or other Cultural Heritage professionals. There is some discussion of the issues here in the section concerning standards issues.

The different challenges represented by this wide range of data types means that each can be regarded as a distinct area of expertise, requiring distinct regimes of good practice and recording devices and it is true that devices have been purpose built for many of the situations. There are many different technologies and in general similar technologies may be used for different environments. The technologies can be grouped according to the dimensionality of the artefacts being captured.

Most of the outstanding issues relating to 1D and 2D capture – excluding interpretation and retrieval - relate to procedures for capturing and modifying metadata to produce integrated resources and to the sheer volume of material that is in need of digitisation. The techniques and technologies for acquiring the required digital representations have become sufficiently routine to regard further work as deployment and development rather than research. We will therefore concentrate in the next session on the issues, state of the art and future actions required for 3D artefacts.

4.3. 3D Data Capture

The last few years have seen much development in 3D applications for Cultural Heritage purposes (for example, compare the overviews of [Addison, 2000] and [Beraldin et al., 2005]). However, there remain considerable limitations to effective digitisation of 3D objects and to their use in applications in Cultural Heritage.

3D objects are an important part of digital Cultural Heritage, because, perceived reality IS in 3D with change over time frequently regarded as the fourth dimension, i.e. changes to heritage buildings, landscapes, etc. due to environmental effects and human activity, development, war, changes in land use driven by economic demand, etc.

In practical terms for 3D objects the following processes need more effective support:

3D acquisition:
The only way of substantially reducing the time, effort and costs of 3D acquisition in the long term is to create intelligent tools which will simplify the processes and reduced the level of ICT skills needed to undertake the tasks. This will be done by developing methods and tools which allow the operator to undertake the tasks based on working practices in the application domain, rather than becoming an ICT expert in order to be able to operate tools that intrinsically feel alien

to them. For example in 3D-Scanning and photogrammetry, a core problem in 3D data manipulation is at present merging multiple raw scans and smoothing; This is undoubtedly difficult operationally, but the development of tools that make this operation semi-automatic and user-friendly is a more sustainable approach than to alter the working practices and educational background of all Cultural Heritage professionals to accommodate these processes.

Documentation:
Importance of getting much clearer about the place and role of 3D objects in digital libraries; how can we achieve a deep integration of 3D in digital libraries? Of importance here is the conceptual integration into the world of documentation standards. These aspects also overlap with concerns for long term archival formats and are dealt with in the section on standards.

Visualisation and reconstruction:
The “simple” issue here is to remember the history of a reconstruction (c.f. version control). However this simple view is somewhat compounded by the process of reconstruction (interpretations, derived works and assemblages) where the basis of decision making and the evidence on which those choices are made may also need documenting and carrying forward with the digital artefacts.

Display/presentation:
There is a need to support a variety of display purposes such as scholarly research and public presentations (e.g. museum exhibitions) through 3D CH representations. These different purposes need to be sensitive to different requirements in terms of the balance between explanation and educational motivations and engagement and entertainment objectives. The balance needs to preserve the credibility of what is shown even where the objectives may be less educational and more entertainment orientated.

4.3.1. Requirements of Cultural Heritage 3D acquisition

There is a wealth of 3D acquisition techniques. Epoch Deliverable D3.1.1, section 6.1 [EPOCH, D3.1.1, 2005] gives an overview, discussing the general pro’s and con’s from a cultural heritage perspective, and listing the issues CH professionals may want to pay particular attention to when choosing a particular technology. Thus, rather than repeating such overview here, we simply reiterate our way of structuring the major technological 3D acquisition families.

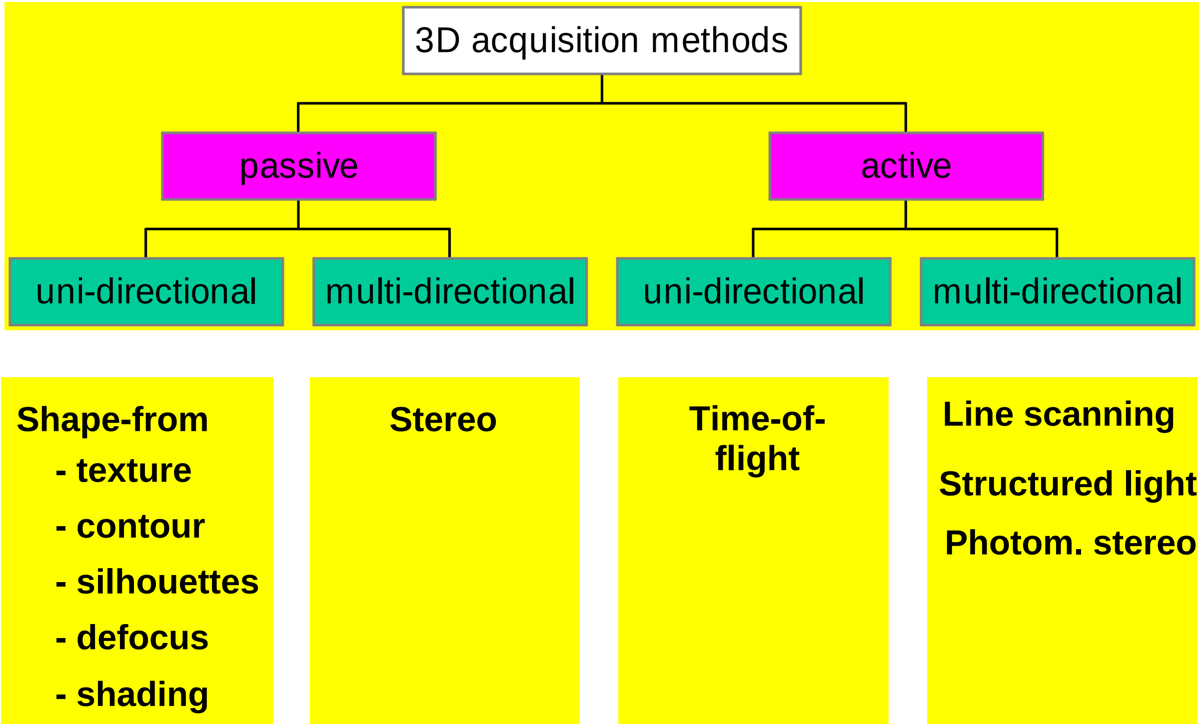


Figure 4.2 Taxonomy of approaches to 3D data capture

The first distinction that is made is between *active* and *passive* techniques. The former make use of special illumination, whereas the latter do not (i.e. they work with normal, ambient light).

Within each class a further distinction can be made between *unidirectional* and *multidirectional*, depending on the use of a single or multiple vantage points, respectively.

The bottom-line is that there currently are three technologies that dominate the market:

1. photogrammetric methods, based on multiple images – *passive, multi-directional*
2. structured light and laser scanners – *active, multi-directional*
3. time-of-flight methods – *active, unidirectional*

The 3D capture and modelling of cultural artefacts is a challenging task. As a matter of fact, the particular combination of demands from the field of cultural heritage turn it into one of the most critical testing grounds for 3D capturing technology. Claims that 3D scanning is a solved problem therefore are premature in this area. Simple adaptations of existing technologies will not suffice to offer adequate solutions. Here is an overview of some of those challenges.

1. *adverse working conditions*: for many applications, the equipment has to be brought on-site. This can mean bringing it to remote excavation sites, in a desert or a rain forest.

2. *hands-off conditions*: museum exhibits are often too fragile or valuable to be touched. The scanner should be moved around the object, without it being touched. Systems that are portable are to be preferred.

3. *intricate shapes*: many important pieces of art have intricate shapes. Scanning those means that great precision has to be combined with a great agility of the scanner. It has to capture narrow cavities and protrusions, deal with self-occlusions, fine carvings, etc.

4. *low price*: the area of cultural heritage may have a huge intrinsic value, much of which can be expressed in economic terms, but this usually appears elsewhere in the system (in hotels, restaurants, …) and not so much at the excavations or sites discovering or safeguarding it. In addition even when caring for priceless treasures this does not equate to an automatic revenue stream or available capital for investment. In practice, within memory institutions, the money that can be spent is usually very limited. Hence, solutions typically have to be cheaper than those affordable by industry, where the benefits may have commercial return on the investment.

5. *diversity of materials*: the types of objects and materials that are to be handled are very diverse. They range from metal coins over woven textiles to stone or wooden sculptures to gems and glass in jewellery. No single methods can deal with all these surface types at once.

6. *speed*: museum collections are often huge. Excavations tend to produce an enormous amount of finds. Even 3D modelling the most important and representative part means producing models for thousands of objects. Speed is of the essence to render such endeavour practical.

7. *size range*: things to scan range from tiny objects like a needle to entire landscape containing petroglyphs.

8. *non-technical users*: the users of the equipment usually did not get a technical training. This stands in sharp contrast to the use of similar equipment in industry.

9. *lack of predefined specs*: precision is a bit of a moving target in cultural heritage. There often is a type of analysis that would also become possible if even more precision can be obtained.

4.3.2. Technological research issues in 3D capture

The previous considerations lead to a number of desirable, technological developments.

1. *combined extraction of shape and surface reflectance.* 3D scanning technology is increasingly being aimed already at also extracting high-quality surface reflectance information. Yet, there still is still some way to go before high-precision geometry can be combined with detailed surface characteristics like full-fledged BRD (Bidirectional Reflectance Distribution) or BTF (Bidirectional Texture Function) information.

2. *in-hand scanning*. The first truly portable scanning systems are already around. But the choice is still restricted, especially when surface reflectance information is also required and when the method ought to work with all types of materials, incl. metals, and gem stones.

3. *on-line scanning*. The physical action of scanning and the actual processing of the data often still are two separate steps. This may create problems in that the completeness and quality of the data can only be inspected after the scanning session is over. This is the equivalent of creating scanned images of text and subsequently processing the images using OCR to create the full text version.
   By then it may be too late or cumbersome to take corrective actions, such as taking a few additional scans. For 3D digitisation, for example, it would be very desirable if the system would extract the 3D data on the fly, and would give immediate visual feedback. This should ideally include steps like the integration and remeshing of partial scans.

4. *opportunistic scanning*. There is no single 3D acquisition technique which is currently able to produce 3D models of even a large majority of exhibits in a typical museum. The techniques sometimes have complementary strengths and weaknesses. Untextured surfaces are a nightmare for passive techniques, but may be ideal for structured light approaches. Ideally, scanners would automatically adapt their strategy to the object at hand, based on characteristics like spectral reflectance, texture spatial frequency, surface smoothness, glossiness, etc. One strategy would be to build a single scanner that can switch strategy on-the-fly.

5. *multi-modal scanning*. Scanning should not only combine geometry and visual characteristics. Additional features like non-visible wavelengths (UV, (N)IR) have to be captured, as well as haptic impressions, all of which currently would need separate instrumentation and scanning processes. Haptics would then also allow for a full replay to the public, where audiences can hold even the most precious objects virtually in their hands, and explore them with all their senses.

6. *real-time, detailed 3D capture*. In the same vein as SLAM (Simultaneous Localisation and Mapping) activities in robotics, there will be an increasing need to quickly build detailed 3D maps of complex and dynamic environments. Part of the task will be to estimate the lighting conditions. Once this is possible, virtual objects can be included into real environments for portable augmented reality applications, from free vantage points, and where the real can occlude the virtual and v.v.

7. *semantic 3D*. Gradually computer vision is getting at a point where scene understanding becomes feasible. Out of 2D images, objects and scene types can be recognized. This will in turn have a drastic effect on the way in which low-level processes can be carried out. If high-level, semantic interpretations can be fed back into `low'-level processes like motion and depth extraction, these can benefit greatly. This strategy ties in with the opportunistic scanning idea. Recognising what it is that is to be reconstructed in 3D (e.g. a house), can help a system to decide how best to go about the task, resulting in increased speed, robustness and accuracy. More on this is also to be found under the research agenda for procedural modelling.

8. Obviously, once 3D data have been acquired, further processing steps are typically needed. These entail challenges of their own. Improvements in automatic remeshing and

decimation are definitely still possible. Also solving large 3D puzzles automatically, preferably exploiting broad shape characteristics in combination with surface detail and texture information, would be something in high demand from the cultural heritage domain. In addition techniques to cope with shape matching and reconstruction of sets of pieces where some pieces are missing and all are normally worn, would represent a more typical case. This would be assisted by an understanding of the processes of ageing and wearing (see also later challenges in modelling).

Against the general backdrop of 3D data capture, it is important to note that a series of recent techniques, generally referred to as image-based rendering, also hold good promise for cultural heritage. We consider these techniques to fall under the issue of combining (and balancing out against each other) the aspects of shape and surface reflectance and texture acquisition.

These technological challenges are further complicated by the need to develop solutions which are accepted in the Cultural Heritage professions. This implies ease of use with little technical training, provable accuracy and sensitivity to multiple potential interpretations and the need to demonstrate why particular solutions have been proposed. Furthermore there is a more pressing requirement that solutions for Cultural Heritage should be cheaper than perhaps is true in more commercial sectors.

During the EPOCH Research Agenda workshop at VAST 2006 discussion also focused on the following questions:

- How should targets on incremental improvements be set (for example, in the context of FP6 the objective was propagated to “halve cost of digitization in 4 years”)? This issue is complex because in many ways improvements in quality and achievability are as important as speed and tend to offset any improvements in speed. More typically the results may be better, but achieved in comparable time. As techniques are developed which allow adequate accuracy, speed-up may become a more independent target, but evidence from other sectors suggests that improved quality and functionality tend to attract more interest and attention than pure speed.
- Related to this the workshop perceived that precision is a moving target, not a given long-standing target, but one where the expectations change as the technical capability improves and applications learn how to exploit the accuracy (“squeezing the max out”).
- How rapidly “best practices” would change as equipment such as scanners improved? This issue addresses the question of deployment of technology and touches at least as much on moving the professions forward as on any explicit research into technologies. The underlying concern is that the technologies must demonstrate their worth before adoption in practice, but although this would be a necessary condition it would not be

sufficient for the technology to be adopted and that explicit sector development initiatives would need to be conducted in parallel.

- The final concern was that it would be essential, particularly given the rate of change in technologies and the massive digitisation exercises implied in many political agendas for developing the sector, that data would be recorded in many formats and by many techniques. Harmonisation would also imply format conversion being applied over time so it would be essential that the provenance of digital artefacts was recorded and available. This provenance information needs to be secure during many format conversions and other manipulations and in this context the role of the "London Charter" [London, 2006] on 3D in CH should be emphasised.

4.4. Documentation of 3D Digital Objects

In this section we examine some of the research issues which take us, literally, beyond the surface level data capture and representation of artefacts and into consideration of the semantics of the objects, sites etc – how we express, capture and encode our knowledge of their meaning. Since this deals with documentation it also inevitably leads us to discuss documentation standards and others that will enable the interchange of information and interoperability of services.

For the purpose of this discussion we consider 3D objects to be, in essence, documents which potentially encompass the full range of text and non-text based data in a structured, integrated and cross-referenced digital object.

4.4.1. Short-term, Urgent Open Research Problems

In this section we describe some of the components of documenting 3D digital artefacts that require systematic research at the interface between computer scientists and professionals in the various cultural heritage domains. These issues should be tractible in the short term unless otherwise noted in the text.

Define/classify *relevant* shape representations for CH

Surfaces: *Discrete:* point clouds, surfels, range maps, triangle meshes, b-rep meshes, subdivision surfaces, primitives (spheres, cylinders, etc.); *Parametric:* B-splines, NURBS; *Implicit:* metaballs, moving least squares

Volumetric: *Discrete:* voxel grids, tetrahedral grids, octrees, BSP trees; *Parametric:* trivariate splines, free-form deformations; *Implicit:* F-rep, radial basis functions

Structured: Articulated figures (bones), deformable models, procedural shapes, Boolean set operations (CSG), scene graphs, computer games, virtual worlds

Define a sustainable 3D file format.

This entails defining: fundamental requirements for all 3D representations; a basic set of exemplary shape representations; a *customizable encoding* that can accommodate all attribute variations (vertex/face normals, etc), and a well-defined mechanism and process to extend the basic set in ways which continue to enable interoperability.

Generic, stable, and detailed 3D markup

A method is required to reference a *portion* of a digital 3D artefact, irrespective of the particular *shape representation*, so that detailed surface and structural features can also be discriminated. The markup should survive simple editing operations (cutting, affine transformations) and be integrated with the paradata (as used in the context of the London Charter) which documents the digital artefact's provenance.

Define generic 3D query operations

People are well acquainted with the search for images, as for instance offered by Google or Ask. Yet, this search is currently still driven by the surrounding text and the image file names. The actual content of the images is not normally analysed, despite many attempts to design systems which undertake the analysis. For many CH applications, the direct access to the image content will be an absolute necessity, however. Queries will often be related to aspects that have to do with certain characteristics of the shapes of an object, of the ornamentation, etc. By their very nature these are not very well-suited for textual description in any case. Currently, much progress is being made in the recognition of particular objects and instances of more general object classes. The key is to shift away from the traditional, global and very low-level features (colour histograms, moments, etc.) towards configurations of well-selected local regions, that are characterized by certain levels of invariance to irrelevant geometric and photometric changes occurring under image projection. This said, it is clear that still a lot of further investigations will be needed and CH will again be a particularly challenging field. Queries often probe aspects also beyond the attention of the casual observer.

In the context of text based search there are established constructs which are domain based and define a known domain of discourse between the user and the knowledge base. The best known and understood is probably bibliographic search where operations such as searching for authors, publishers etc are well understood and supported directly in the metadata. More recently other more complex searches, based on content have become common place. For example searches on the basis of queries such as "Find all the citations of this work by others" are increasing regarded as normal.

In qualitative research it is common to search for concepts in free text and analyse sources based on juxtaposition of these concepts and other information in the text. Currently this is done primarily by manual techniques, perhaps assisted by coding assistance for the concepts defined in the context of individual investigations (e.g. Invivo). The next stages in the field of textual search will be to embed automated understanding of the language within the search tools, to enable higher level conceptual search. Making this effective is a longer term research target, with direct relevance to cultural heritage and requiring active engagement from professionals in each application domain.

In the context of 3D objects, research is substantially less well formulated and even the basic vocabularies for description and search operations require development. What sort of query methods should every shape representation be *required* to answer? What are the common concepts forming the base set of search operations? How would they be combined (i.e. are traditional Boolean logic operators as applied to typical text search sufficiently rich?)?

For example a basic query method for a *Bounding Box* interrogation of a shape representation – *Triangle Mesh* – might return a list of triangles inside the box. This operates in the shape domain which is already at a higher level than current search operators, but still a long way of allowing the Cultural Heritage professional to undertake research based on the cultural heritage domain of discourse. Hence defining these elemental operations is considered a shorter term research question.

An architect wishing to search for "all surviving examples of timber framed buildings which were renovated and adapted to include Georgian facades" would currently be reliant on keywords attached to data rather than searches of the underlying objects. Such searches must be considered a long term research challenge, and require a huge amount of prior research in defining the conceptual framework of codification and enquiry. The issues of enhancing metadata are considered further with the longer term issues below.

Provenance and processing history log

The provenance and processing of 3D digital objects needs to cover the initial data capture, the format and encoding of the individual elements and any refinements or abstractions that have been applied, documenting in each case the circumstances of the processing applied. If a processing history is complete then it is possible to envisage replaying the process of constructing the digital 3D object, perhaps varying some parameters in the process, or indeed substituting new or improved methods with later developments.

A significant issue in technology will be version control, not only of the digital artefacts or even of the tools used to process them, but particularly of the underlying technologies (computers, operating systems, compilers, and applications programs) which have been used to construct them. Issues here will be minimised if properly independent file formats are used, but manipulations may still have produced differing effects where different machine architectures have been used, with issues such as numeric accuracy potentially becoming important as the underlying accuracy of digitisation and processing arithmetic improves.

Two specific topics in the short term are therefore:

- A standard for describing the sources of digital 3D
- A standard way of recording how the source data have been processed, and how they were combined to obtain the result.

On a very basic level we also need to apply Unique Identifiers (e.g. DOI's)

Maintaining consistent relation between shape ↔ meaning.
Even at the level of shape representations derivation of a shape semantic is complicated by interaction with the specific inaccuracies of manufacture. For example:

- *Given that architectural feature will not be made with mathematic precision, is there such a thing as a circular arch in reality?*
- Putting it technically, "how can we relate a low-level primitive (a *triangle*) to the semantics of the overall shape?
- How can procedural 3D technologies help? (e.g. link triangle from scanned arch to procedurally generated arch).

Ideally, we would be able to achieve full Generative Surface Reconstruction, but some questions which arise are:

- *Where and how is the link between shape and semantic stored?*
- How can we *author*, *query*, *show* that knowledge?
- How can *knowledge* and *3D* information be kept **consistent**?
- How can we *navigate* through the *semantic 3D graph or semantic associations represented*?

At the technical level, in terms of current issues, questions such as "What is the relation between METS, CIDOC-CRM, and Collada?" arise. In broader terms the relationship between ontological standards being used and further developed from a Digital Libraries perspective and the EDL emphasis on a Digital Library spanning a much wider spectrum of artefacts than the traditional text based sources adds urgency to addressing these issues.

Although the immediate topics in this area may be considered and potentially solved in relatively short timescales, the challenge of achieving lasting standardisation agreements may well mean that the search for agreement lasts more into the medium term. It is also likely that the topics will need continuing attention and the standards require maintenance as new levels of semantic representation become tractable.

Other issues will arise during this process. For example if we assume the meaning of a number of shapes is known in particular contexts and we then change that context, can we derive new meaning based on the context or does that require explicit re-coding. Similarly, if the shape itself is edited, at what point do we decide that the semantics have changed? Can we derive new meaning of shapes?

We might assume that the meaning of a shape is known and unambiguous, but of course this is not the case as demonstrated by the current concerns of incidents with toy and replica firearms being wrongly interpreted and having life-threatening consequences. At a more relevant level in terms of 3D digital objects, the modelled version of a sci-fi weapon is designed to include many of the shape semantics of an actual firearm. In fact the designer typically relies on the shape's semantics to convey the threat and purpose of the object. It is therefore apparent that the semantic of shape exists within the context of other aspects of the 3D object and its environment, leading to questions of whether this context is an intrinsic part of the semantic or whether it is purely to resolve ambiguities.

4.4.2. Longer-term issues: Non-textual vocabularies and semantic 3D objects

For 3D digital objects adequate functionality for automation and integration is missing in a variety of areas including:

- Structural representation (some of this is reflected in the previous section but the full functionality would need to take into account effects of aging, damage, etc. as a integral process in the life of an artefact);
- Content categorization/classification & analysis;
- Non-textual metadata for 3D objects, automated mark-up, and semantic enrichment;
- Indexing / searching (at higher levels) based on non-textual metadata;
- Information extraction and abstraction;
- Search by similarity and type-based search.

These aspects are discussed below. It is important to address the research implied in these topics to get past the current limitations which current data capture and encoding techniques place on what digital libraries can support with respect to 3D digital objects libraries. For example by exploiting the semantics of artefacts on users of such libraries would benefit from novel search & retrieval methods or semantics-based navigation through distributed 3D documents.

In the context of searching in 3D documents there are a number of underpinning areas of computer science for which these issues have already received some focused R&D. Successful approaches so far include: Feature Vectors, Silhouette, Depth Maps, Lightfield Descriptors, Volumetric Abstractions: Sphere Functions + 3D Harmonics, Shape Distributions, Statistical Moments, Shape Spectrum, Skeletal Graphs, Self-organizing Maps (SOM) and others.

Among the more advanced European research initiatives in this field are the German Research Foundation's Strategic Initiative (V3D2) and the work carried out at the Visual Information Processing Lab of the University of Florence should be mentioned.

In the United States, the research group of the Princeton Computer Graphics Group led by Thomas Funkhouser merits mention (e.g. their 3D Model Search Engine and parts-based 3D object classification approach [Funkhouser, site]).

The current development of metadata including the semantic web approach is primarily text-based. However an adequate (non-textual) vocabulary to characterize the content and structure of 3D objects in the sense described above is not available. Systems which provide some information which begins to address this area have been driven by significant manual intervention. However manual categorization of artefacts is extremely time consuming and could not have a significant impact on the enormous backlog of digitisation, which includes:

- A backlog of undocumented artefacts (e.g. for London Museums 50% and more, according to Chris BATT, CEO UK Museums, Libraries and Archives Council, April 2004).
- A backlog of documented artefacts for which documentation is only available in paper based systems.
- A large proportion of digitised artefacts for which the digitisation has been considered a priority, but where the information so far recorded is partial and recorded using data technology.
- Even artefacts yet to be digitised where the potentially useful information is not yet recognised and routinely recorded ("future legacy data").

For typical library objects (e.g. text documents) nobody would use elementary pixel configurations to describe content and structure of pages. Instead, although scanning based on point data is the starting point, OCR is used and the content represented by characters and structural mark-up. There is however a continuum between objects whose primary interest is textual and those where the textual content is significant but other aspects become equally or more significant (e.g. illuminated manuscripts or inscriptions in stone).

If we scan a building (e.g. a cathedral) with a 3D laser scanner this typically will produce 10 million triangles and more, and the "representation" of the artefact remains within the domain of (hierarchical) triangle meshes. The content and structure of this artefact at present is mainly described manually by textual augmentation; no high-level elements are available to represent content and structural mark-up in a domain-specific way and indeed systematic and inclusive vocabularies, taxonomies and ontologies have yet to be defined to underpin such descriptions.

Development of proper "vocabularies" for a new generation of metadata capable of characterizing content and structure is key to 3D content categorization, indexing, searching, dissemination, and access. Humankind has traditionally adopted techniques which express our understanding of the higher dimensions in physical space by reference to lower dimensional representations to place order on higher dimensional structures. Thus 2D drawings (plans, elevations, sections and detailing) are used to organise and convey our understanding of 3D architectural constructions, plus an extensive set of conventional representations for drawings, features, labelling and indexing into structured data that can be systematically organised and searched in linear (1D) system.

Computer Science has for decades sought ever improved ways of sorting and searching based on multi-dimensional organisation and searching, exemplified by some of the higher dimensional methods of organisation of scenes in computer graphics (e.g. ray-classification schemes in ray-tracing) and it is to be expected that techniques will be developed to allow higher dimensional constructs of semantic information to be searched efficiently rather than the reduction of the search domain's dimensionality in ways which people might find an aide to simplifying the problem.

The important breakthroughs will need to be based on developing the understanding of the higher dimension semantic relationships that we seek to understand and represent, many of which are intrinsic and subconscious in human understanding but not established, documented, shared consistently and agreed. Only when this work is tackled will we be able to overcome the current treatment of 3D objects in Digital Libraries as BLOBs (Binary Large Objects) and generalized documents.

The goal is a *complete semantic 3D-model* instead of projections in lower dimensions (image, section, animation, text) or structure-less collections of polygons. What we need to achieve is a *deep* integration of 3D into Digital Libraries and Collections Management systems.

The benefits of such models can be assessed through a comparison with the traditional approach or representing 3D objects, which is based on (hierarchical) polygonal meshes. This approach implies the following problems: Loss of structure, content-based handling is (almost) impossible, inappropriate complexity measure (sphere: (centre, radius) is represented by a number of triangles which approaches infinity as the accuracy required increases), data compression is very hard, no object-specific LOD [level of detail] representation is possible.

3D Cultural Heritage data which is semantically enhanced by incorporating knowledge from the cultural heritage domain is a field of research which lies at the core of bringing technological work and Arts & Humanities research much closer together. A great deal of research and development work is required and, in fact, this is fundamental to realizing the potential of technological support for Arts & Humanities studies, analysis and presentation. Moreover, Cultural Heritage management could benefit much from having available semantically enhanced tools.

Before we can represent the semantic content of the 3D objects and assemblies of objects we must first define the vocabularies, taxonomies and ontologies that express the semantics and represent them. There are many studies of architectural, ceramics, decorative and other styles of man-made artefacts which document the variations in the ways styles are adapted to different circumstances. There are rather fewer studies attempting to model the systems by which styles are put together or to express them in an analytic set of relationships which could be used as a template against which artefacts can be matched. Some of the work currently being undertaken in procedurally based modelling is advancing the thinking in these directions. (See below). Some of the components of this exercise are beginning to emerge.

A semantic gap between a *shape* and its *meaning* has already been demonstrated above. Small shape or material differences can mean large semantic differences (for example in the classification of objects such as amphorae, jewellery, buildings, altars, etc.)

In practical terms we need to decide how a semantic 3D mark-up system can be created most effectively. Currently this would require segmentation and mark-up which is mostly manual and in which subject experts need to be directly and extensively involved, though some semi-automatic support may be provided.

We need highly structured data that enable detailed examination. This is not about descriptive metadata, but semantic interlinking of content and detailed search (e.g. for parts of heritage objects, e.g. parts of buildings or statues). CIDOC-CRM based semantic networks can play a core role in this.

For example, the Probado Project, [PROBADO], which is funded by the Deutsche Forschungsgemeinschaft, has a vision of a "Google for Architectural 3D Models". Probado strives to integrate 3D into a real library (TIB Hannover) and bases its work on standards such as METS, Fedora, Ajax, Axis 2 and server-side GML. The background to this project is that each year there are hundreds of new architects trained who could greatly benefit from this.

Interesting research questions in this work are, for example:

- What does "content based retrieval" mean for architectural models, and for shape in general?
- How can search queries be "formulated"?
- How do you create "*Shape abstracts*" (i.e. what semantic LODs (levels of detail) are feasible)?
- Current classification methods may work for shape classes but are too coarse; for example to create an abstract of variants in styles. How do we achieve that?

- To what extent is it reasonable to assume that text-based mark-up will be the preferred, or even available, encoding of the knowledge of the content of digital 3D object?

4.5. User Created Content

The concept of User Created Content as an integral part of our cultural heritage has a significant history of its own. In the European Framework Programs it received a significant boost under the "2001 Heritage for all" initiative, described in 2002 by Bernard Smith as follows:

"Here the objective was to foster sustainable online communities in creating and documenting the digital record of their societies, including safeguarding its accessibility for the future. Projects were expected to innovative and experiment in creating, manipulating or aggregating local resources and making them sustainable, visible and valid in the global context. One option was digital archiving applications integrating discovery technologies and tools, to provide easy access to the evolving digital record of the peoples of Europe at different levels of complexity and detail. Another option were tools and services that guarantee equality of opportunity and quality of discovery services and resources in support of social and cultural inclusiveness. The projects were expected to take account of ongoing national and regional heritage initiatives and digitisation programmes. They were also expected to promote cooperation between different types of memory and cultural organisations at local/regional level, as well as appropriate public/private sector partnerships. The technical focus was on resource discovery and data fusion, on authentication, integrity of services, on usability and ergonomics, on stable and reusable business models, and on the active participation of end users through new online communities.

"A new cluster is born: As a result of the above action four new research projects have been recently launched. It is hoped that they will form the basis for a cluster to be developed in future programmes. CHIMER, CIPHER and COINE all address the personal views and interests of ordinary people in order to build a living picture of regional heritage across Europe. The fourth project, MEMORIAL, focuses on digitising a wide variety of paper documents in libraries, museums, libraries and public records offices concerning the Holocaust and developing a methodology and tools for the creation of personal digital memories." See http://www.chimer.org/, http://cipherweb.org/, http://www.coine.org/ and http://www.memorialweb.net/" [Smith, 2002]

User created content takes on several forms:

1. Recording of personal experiences for a variety of reasons. Personal experience of the citizens may be recorded about a way of life, a place, an event, an organisation, a social movement or other collective experience. This would often be recordings of the experiences of the older generation in retrospect, but increasingly it concerns a cross-

section of people and may be relating to recording at the time of an event. Thus there were many attempts to record the significance of the millennium to different groups of people, countries etc. from the perspective of the full age range of those alive at the time.

2. Recording of other intangible heritage – for example, particular long-standing ceremonial occasions, traditional celebrations or orally transmitted cultural content (stories, songs, dances, etc.) which have both a personal component as well as a culturally-based and evolving set of cultural norms.
3. User created content may also be commentary on items in a museum's collections where "user" in this context may represent members of a community with specific knowledge. For example artefacts of social history are held in their millions in regional and local museums but the knowledge of specific artefacts purposes may be incomplete in many cases. Access to the knowledge and memories of the community may well be a useful source of documentation of such artefacts.
4. In the UK there is an increasing public involvement in documenting archaeological finds away from the known and protected sites with the volume of finds recorded by the public and officially registered rising by 45% in 2005/6 to 57,566 from 39,933 in 2004/5 [DCMS, 2007]

These sorts of data are expected to form an increasingly important part in the curatorial responsibilities of memory institutions, as highlighted in the description of Scenario 2 above:

"Museum doyen Kenneth Hudson about ten years ago suggested that Europe is "a giant network of potential eco-museums". [Hudson, 1996] In fact, the concept has much future potential particularly through the use of novel technologies that allow for effectively representing the collective heritage and memory of a region.

Work such as that of the HICIRA Network [HICIRA] has shown clearly that local museums (and especially local site museums) are going to move away from the static displays of artefacts and concentrate on establishing the structures for the creation of long-term, sustainable local memory institutions, in which the input of the public is central. This view is supported by some policy work (e.g. the UK Department of Culture, Media and Sport's recently issued document on "Priorities for England's Museums" [DCMS, 2006b]). The success of the local museum in the next decade is going to depend on how effectively it can function within a community context. It can no longer be just a "show" or a "tourist attraction," but needs to be an integral part of the community."

The implications for data recording are similar to those faced in collecting and analysing data for questionnaire-based surveys, combined with technologies aimed at creating a visual record of an event. At the surface level this involves recording oral interviews and visuals of events, but it is likely that the volume of such recordings would mean that without intelligent search and analysis tools the full richness of the data collected would be inaccessible. To make the data usable and useful a number of processing steps can be envisaged.

The oral data will require processing to produce transcripts and to associate metadata describing the context of the recording – what was being discussed, when was it being discussed (e.g. time separation from the memories), why was it being discussed (e.g. in the context of an anniversary, an event, an exhibition etc), who was enabling the discussion, etc.

As with questionnaire data there would then be post-processing depending upon the features that were being examined. This could include isolating concepts being expressed in the discussion, with subsequent analysis looking for particular associations of concepts. Interview analysis is frequently undertaken in social sciences research, and the distinctive features for the developments expected in cultural heritage institutions will be:

a. the volume of data becoming available

b. the concepts and questions being examined

c. the multi-cultural perspectives on particular events, ceremonies etc and the consequent likelihood of the set of transcripts being in multiple languages with additional difficulties of considering whether concepts from one cultural have any equivalent in another.

d. Relating multi-media recordings where the oral recording may be undertaken whilst viewing a visual experience for example.

As with the oral recordings, the visual recordings may require significant post-processing to make the resource useful.

There may well be aspects of recordings whose significance is not appreciated at the time of the original data capture. In the case of oral recordings, background noises, inflection, accent and emotion, these are all aspects that may not survive the transcription process and may become interesting in considering additional analysis later. For the visual recordings, background images are probably the most obvious additional information that may support additional analysis although not the primary subject being recorded. Other aspects such as the relationships between multiple recordings at a single event may also prove significant. This will also be shared with other current uses of multi-view data in reconstructing events in other disciplinary contexts:

(i) In crime and anti-terrorism where multiple perspectives from news cameras or from security cameras around an area can be used to try and reconstruct events in 3D

(ii) In sport where multi-view coverage of a sporting event (particularly in football) is increasingly used to provide a 3D reconstruction for use by commentators

Both of these are developments of technologies originally pioneered in environmentally controlled motion capture studios and it is easy to envisage these technologies being used to capture the essence of a folk dance, for example.

Other analyses may imply radically different views of the primary visual data. In the same way as a revisiting of a body of primary oral data might allow analysis of development of regional accents for example, visual data may provide opportunities for other research into societal change. An exhibit in the Millennium Dome recorded 250,000 individual, low resolution personal avatars of visitors. It is not difficult to envisage questions which might be asked of the social composition of visitors (apparent age distributions, for example) or perhaps examining other characteristics – clothing choices perhaps. As data sets are built up over time, analyses of the comparison and trends over the years might also become sustainable.

The degree to which datasets such as these can be used in alternative analyses will depend not only on the development of the technological tools, but also on the development of a societal willingness to permit alternative uses. This issue will be particularly focussed where visual information from surveillance cameras observing “normal” acceptable behaviour might be useful for other purposes, but nevertheless be seen as infringing the rights of individuals. However there is a thin line between recordings where the analysis of background images might be (and have been) undertaken to provide evidence in cases of antisocial or criminal behaviour (for example at a football match) and using the background of recording of a specific event (e.g. a may pole dance) to provide additional research material on cultural behaviour. In both cases the existence of a warning that the event is being recorded and that attendees have given their permission for alternative uses of the recordings might be a legal necessity, but there will be debate outside of a research agenda on the moral and ethical dimensions.

In this research agenda we consider the technological advances that would be needed to sustain the data collection and analysis. We leave the ethical dimension to others to decide.

4.6. Intelligent tools

Many aspects of intelligent tools are appropriately discussed in other sections of this report. This brief section is included in order that the reader may realise that they are included despite no specific listing and discussion in a centralised piece of text. Intelligent tools in the context of the research agenda include all those areas where the tools and techniques to be developed are based on semantic knowledge of the cultural heritage domain. This includes primarily:

- The use of domain specific knowledge to assist in digitisation or reconstruction. For example the digitisations of a row of very similar, but damaged, statues might be combined to deduce the shape of a perfect original prototypical statue, which could then be morphed to match the remains of the individual instances and define "perfect" versions of the complete set of unique statues. (Figure 4.3)

- Semantic encoding and ontologies to support advanced search and cultural heritage research;
- The cultural and style specific knowledge embedded in proposed tools for procedural modelling; and,
- In the presentation area, tools which exploit the internal structure of the environments presented and the data associated with them to produce improvements in the experience. This improvement may relate to value-added in the stories told or the underpinning real-time systems, which are enabled to operate more efficiently by exploiting the knowledge of the underlying semantics of the data.

Figure 4.3

It would be unsurprising for Cultural Heritage professionals to be sceptical about the building of intelligent tools into the processes they use. As the term professional implies, they operate by exercising judgement based on knowledge and experience built up over time. This is a true use of intelligence in processing cultural heritage information and no professional group should give up this responsibility.

The sensitivity to intelligent assistance for the professional in their work will depend upon the implication of the assistance being insufficiently intelligent. Recent years have seen a great deal of thought given to appropriate levels of professional responsibility in subjects allied to medicine. How much professional knowledge and experience is needed to allow a nurse to make decisions about patient care, or to allow a paramedic in an ambulance to provide care at the site of an accident before the patient can be transferred to more specialised care environments? How do we strike a balance between the care given in a doctor's surgery and that administered only in hospitals? The "implication of the assistance being insufficiently intelligent" in these cases is clearly life-threatening and the answers may well be different in emergencies. With Cultural Heritage professionals the implications may be less severe in terms of well-being, but may be as serious in terms of the subject of their domains.

The implication of the assistance being insufficiently intelligent is most damaging when there is no opportunity to reconsider and repeat a process. The most obvious cases in Cultural Heritage will concern:

- Archaeological or other data capture where the original circumstances are lost. Recording archaeological excavations is the most obvious case, but some other data capture may be difficult to repeat.
- Recording intangible heritage is an obvious case here, where it might be impossible to re-interview veterans of a particular conflict or re-create a particular event.
- Digitising fragile materials which is known to be deteriorating (e.g. ageing film or frescos open to the atmosphere) where there would be both a time element and an element of the resource being known to have limited life.
- Recording threatened heritage environments (e.g. buildings about to be demolished) where there is a planned "loss" due to a politically-determined balance of interest and an inevitable difference in opinions about the significance of the environment.

Of course in this last case, where there is a time limit on the access to the original site it may well be that rapid techniques using techniques that have produced good results in similar situations are preferable to attempting manual survey that cannot be completed accurately in time.

Of lesser concern should be those situations in which the experiment can be repeated. This is in principle true of all cases where the primary data is retained and is enabled by recording the provenance and paradata associated with the conclusions drawn by the (insufficiently) intelligent system. There are, however, other concerns that will naturally arise.

Systems which create believable results tend to be believed. The obvious cases of this come from the entertainment industry. Magicians rely on creating belief in the conclusion that a logically impossible set of apparent facts is nevertheless possible – the elephant really did "vanish into thin air" rather than "just" cease to be visible – no small feat in itself of course!

In the context of ICT in Cultural Heritage the most obvious concerns are where visual reconstructions are produced and become fact in the mind of the viewer. Even documenting uncertainty is not enough to relieve the concerns. Visual data is fundamentally different in terms of believability. There is no equivalent in visual paradigms of the footnote in text and a textual footnote might well not be read, or if read, accorded less impact on the messages taken away from the visual material. The message that "the events depicted in this film are entirely fictional" does not prevent the public from believing that there is an element of underlying truth in the film. The visual dimension represents a whole family of facts and the fictionality or otherwise of selected facts becomes buried in interpretation. Indeed the whole industry of product placement in films relies on this. The fact that a particular manufacturer makes cars we know to be true. In this science fictional account their cars are capable of amazing things and therefore we "know" that their cars are amazing.

There have to be concerns, even where unfounded, that the same can be true in cultural heritage messages. To take an example, the placement of avatars representing a particular cultural background in a virtual reconstruction of a different cultural background might be taken as implying a link in history that did not exist, or is highly speculative, or did not occur at the dates being reconstructed. Whilst this may be relatively easy to spot with avatars it is rather less easy with artefacts. The placement of particular styles of architecture in inappropriate regions or at inappropriate dates of reconstruction may be painfully obvious to a cultural heritage professional and pass unnoticed (but subliminally absorbed) to the lay member of the public. Whether such situations are important is of course debatable, but researchers in these fields need to be extremely conscious of the concerns and a sensitivity to the issues might be a requirement of any project as "inter-cultural ethics" in the same way as ethical considerations are raised for research with medical or social consequences.

4.7. Digitisation of Legacy Metadata

At several points the discussion elsewhere refers to specific issues to do with legacy data. Whilst these issues are in essence very similar to dealing with the creation and manipulation of digital objects from primary sources, there are different issues too. These issues concern:

- The quality and media of the secondary sources (including legacy digital data)
- The continuing availability of primary sources
- The relationship between sources

How do you link secondary sources which are themselves cultural artefacts of significant historic interest? In these circumstances the secondary source becomes primary in what it tells us of the times in which the investigation and recording were carried out and for what it tells us of the people and processes used. For the purposes of this section the primary source attributes have been covered above.

The aspects of the significance of secondary sources are addressed by the cultural heritage sector in non-digital systems. For example the existence of errors or inconsistencies between primary and secondary sources, or between multiple secondary sources recorded over time is a challenge which exists in any system (digital or non-digital). Do these variations represent accurate recording of changes over time or errors in one or more attempts to document the primary sources?

Recording and representing inconsistent information about a single set of circumstances or facts remains a challenge. By qualifying each of the "facts" with its attribution ("The archaeologist X believed that Y was true" rather than "Y is true") contradictory "facts" can be resolved even though the underpinning assertions remain in contradiction. The challenges are related to issues of representing provenance, uncertainty, interpretation and cultural values. The CIDOC-CRM [ISO21127] in its current version allows at least some of these aspects to be represented explicitly and recently-proposed extensions to the published standard include the capacity to represent the aspects of provenance relating to the development of digital objects.

The aspects which are more difficult to represent in ways which allow search and intelligent computer-assisted analysis concern the values and judgements embodied in the "facts". For example in past documentation of the evidence the cultural professionals of the day would have had to interpret with the information they had available. Clearly information discovered since their reports were completed cannot have been factored in.

In fact their recording would normally include some initial interpretations and be based on facts that we cannot now re-examine. For example a field archaeologist records contexts which are related to locally observed variations. If something is missed in the initial observation then the raw material will normally have already been disturbed and the experiment cannot be repeated. If the archaeologist interprets a particular pattern in the soil as a post-hole and records it as such, then revisiting the original evidence of soil-type boundaries to decide whether that interpretation is correct is normally not an option.

4.8. Search and Research: Semantic and Multi-lingual Processing

This section deals with the research background and needs for those seeking to enhance the following types of processing of Cultural Heritage data:

- Content analysis (e.g. extraction of base information and co-references from free text sources)
- Collection formation (classification and indexing)
- Cross-collection interrogation, (harmonisation and mapping between different collections formulation, classifications schemes etc)
- Knowledge discovery (or "excavation in the digital domain")

4.8.1. Ontologies, taxonomies and thesauri

A great deal of research has been undertaken into various levels, structures and mechanisms for providing frameworks suitable for supporting searches and organisation of complex bodies of information relating to specific domains of knowledge. The TEL-ME-MOR project recently published a report on subject access mechanisms which includes an analysis of some of the conceptual levels in use in National Libraries [TEL-ME-MOR, 2006, pp8-13]. The report's conclusions show the variety of systems in use across the National Libraries for access to the materials documented in their conventional catalogues and shows how far there is to go to achieve integrated cross-collection search in consistent bases.

The three related concepts in this section's title seem to engender a great deal of misunderstanding, debate and redefinition by different groups, both within cultural heritage and elsewhere, for different purposes. [Gruber] defines an ontology as follows:

> "The word "ontology" seems to generate a lot of controversy in discussions about AI [artificial intelligence]. It has a long history in philosophy, in which it refers to the subject of existence. It is also often confused with epistemology, which is about knowledge and knowing.
>
> "In the context of knowledge sharing, I use the term ontology to mean a *specification of a conceptualization*. That is, an ontology is a description (…) of the concepts and relationships that can exist for an agent or a community of agents. This definition is consistent with the usage of ontology as set-of-concept-definitions, but more general. And it is certainly a different sense of the word than its use in philosophy.
>
> "What is important is what an ontology is *for*. My colleagues and I have been designing ontologies for the purpose of enabling knowledge sharing and reuse."
> Tom Gruber

This definition of ontology appears to match closely the commonest usages in Cultural Heritage. The terms of ontology, taxonomy and thesaurus are used to refer to related but independent concepts in this report:

> "A formal definition of a thesaurus designed for indexing is:
>
> - a list of every important term (single-word or multi-word) in a given domain of knowledge; and
> - a set of related terms for each term in the list.
>
> Terms are the basic semantic units for conveying concepts. They are usually single-word nouns …" [Wikipedia]

The same entry describes how thesauri databases often include classification systems:

> "Thesaurus databases, created by international standards, are generally arranged hierarchically by themes and topics. Such a thesaurus places each term in context, allowing a user to distinguish between "bureau" the office and "bureau" the furniture. A thesaurus of this type is often used as the basis of an index for online material. The Art and Architecture Thesaurus, for example, is used to index the national databases of museums, Artefacts Canada, held by the Canadian Heritage Information Network (CHIN)." [Wikipedia]

In addition thesauri include synonyms providing a structure that relates similar concepts in the thesaurus. For many purposes a hierarchical structuring depending upon meaning would be construed as a taxonomy, whereas the inclusion of synonyms would be providing information on the meaning of the words. However, as the above reference to the word "bureau" highlights, many words have meanings which are context dependent.

A taxonomy as used in this report is a classification according to an ordered system that indicates natural relationships, with a resulting catalogue which can be used to provide a conceptual framework for discussion, analysis, or information retrieval.

An ontology provides a set of concepts and relationships in a domain of discourse. Thus glazing material and window frame are at one level independent terms that might be included in a thesaurus on architecture. Each represents a class of instances or sub-classes (types of glass, Sealed Unit Double Glazing, plastic, etc. or frames with different designs and materials, etc.) which themselves may be included in the thesaurus.

The taxonomy draws together those thesaurus entries into the set of relationships. In this case there is probably going to be a taxonomy entry which includes the concept of a window having both frame and almost always glazing material each of which have a number of potential values and sub-classes. However this arrangement of information classes would not encapsulate the concept of a consistent architectural style, in which a particular design of window detected whilst scanning an architectural ruin would provide considerable supporting evidence for interpreting other evidence. This might have implications when deciding the likely doors, room proportions, façade decorations, overall building scale, etc.

An ontology describes the structure of the concepts, bringing together the concepts in the thesauri with the classifications in the taxonomy. It incorporates knowledge of the structure of the information in the domain of discourse. At present the information systems that operate in support of cultural heritage organisations are based on classification systems that include the explicit encoding of a hierarchical taxonomy. Our ability to provide computer-based assistance in reasoning about the significance of the individual artefacts and the relationships between them will depend on being able to organise knowledge from different perspectives and in different domains. In addition quantifications of these domains should allow the development of more intelligent tools to detect and encode particular features.

Cultural heritage presents other unique challenges in that languages and concepts vary between different cultures and over time. This means that the thesauri and related taxonomies and ontologies have more contexts to take into account and yet the very nature of inter-cultural understanding requires that we are capable of moving meaning from one to another.

The definition of such classification systems for language used in cultural heritage is not only the province of the dry and precise vocabularies of the professional. Indeed with the rising importance of user created content it will become increasingly important that the evolution of language is taken into consideration at an ever faster pace.

Thus thesauri of popular cultural terms would need to accommodate “spiffing”, “hip”, “cool” and “book” as words which imply a similar position in terms of the cultural situation which is described in this way, but also imply different generations using the terms. We note in passing that the use of “book” in this way is still not in widespread use but arises as a product of a very specific technology as it represents the word “cool” as a favoured alternative in some predictive text systems used for SMS messages. It is also worth noting that the last three words all have very distinct meanings in other contexts and that in this case we are only dealing with English (although some popular terms are used in more than one language). The language of the professional in these terms is better documented and more stable.

4.8.2. Multi-lingual and multi-cultural knowledge bases

Even at the simplest level there are examples of issues which arise in multi-lingual operations. For example at a recent keynote presentation at VAST2006 Marc Küster highlighted the differences in operations as simple as organising a list of names in alphabetic order in different languages. Although this is only at a simple operational level the implications of this could be profound. For example this is of course more of an issue in searching through ordered digitised documents but might also become a factor in intelligent data capture of structure within documents. A further example would be the implications of particular character sets in the searching of full-text documents. A study by TEL-ME-MOR on the use of different character sets and the implications for classifying, sorting and searching sources has highlighted the potential difficulties.

“For instance, a search on ‘Böll’, ‘Boll’ and ‘Boell’ on current default collections shows that the British Library integrated catalogue ignores the Umlaut (‘Böll’ and ‘Boll’ give the same result set, ‘Boell’ does not), whereas SNL’s Helveticat takes it into account (‘Böll’ and ‘Boell’ give the same result set, ‘Boll’ does not).” [TEL-ME-MOR, 2005].

Languages have evolved in the localities in which they are used and reflect the needs of the local society to elaborate concepts relevant to them and the context in which they are used. The often repeated statement that the Inuit Indians have 26 words to describe snow and ice is clearly based on fact and on local need to distinguish between different forms. This may be of less concern in the Cultural Heritage field but the variations in popular culture terms in different societies all using variations of English would clearly not be insignificant. For most fields there are significant, though not necessarily extreme, variations, some of which can be handled by context or in extremis by special cases.

In archaeology however there are significant variations in the use of language at different times in different societies. These may also reflect needs for different levels of local distinction between similar concepts or indeed they may be a product of the gradual spread of particular cultural developments as people migrate. Hence a particular term might imply that an artefact was from a certain date in one region and from a different date elsewhere.

There have been some attempts at multi-lingual thesauri for cultural heritage. For example, the HEREIN multi-lingual thesaurus has a limited coverage of terms and the Canadian Heritage Information Network have added some 2600 French terms to the J. Paul Getty Arts and Architecture Thesaurus [Getty, AAT]. However these worthy efforts illustrate very clearly how much remains to be done in these fields if truly interoperable high level information systems are going to be developed.

An extensive and detailed survey led by the Minerva project [MINERVA, 2006] has shown that there is indeed a lack of multi-lingual thesauri. This is due to several factors that make the creation of a domain thesaurus quite a difficult task when several languages are involved:

(i) *there already exist implicit or explicit thesauri in the relevant languages, so the required task consists more in the creation of so-called inter-lingual thesaurus mapping than in the mere translation of terms into the various European languages.* (see for example the EU IST project SWAD, [SWAD, D8.3])

(ii) *there are concepts in some languages which have no, or multiple, correspondents in others. This is a typical problem of multi-lingual thesauri which is concentrated in those concerning culture, for the diversity of European history* (see for example [Doerr, 2001]),

(iii) *similar concepts have very different meaning in different places, starting from the very basic ones.*

For example let us consider time periods. As the EU Culture2000 project ARENA demonstrated, there is no uniformity throughout Europe on archaeological periods, as shown by the diagram on the ARENA web site [ARENA] which illustrates the extent of the usual archaeological periodization in Palaeolithic, Neolithic, Iron Age, Classical, Medieval, Post-Medieval and Modern, represented by different colours in the picture.

For example, the year 600 AD is Middle Ages in Poland, is still Iron Age in Norway, is Early Medieval in UK, and (very late) Roman period for Romania. It would be Byzantine in Greece and other parts of the Mediterranean region, not covered by ARENA, but not everywhere.

As suggested by Michael Buckland (ECAI) in his lecture at the DEN 2006 Conference in Rotterdam [Buckland, 2006], the most basic cultural concepts answer to the questions Who?

Where? When? – but these rely on each other. People may take different names in different countries, geographic names may change their meaning through time, and time periods may differ, as seen above, from place to place. So multi-cultural thesauri must face the challenge of being time- and place-dependent while defining time and place concepts, a sort of internal recursivity which has never been properly dealt with by information science.

The effort from knowledgeable professionals to recognise and document the specialist terms in use in each language and culture would in itself be massive. The effort to extend this understanding so that the different implications of variations in systems evolving over time and place so as to align the information sources in different languages is clearly not available in the medium term. It is to be expected that advances will be incremental and projects in should be required to think of the issues in general as well as the particular domain issues being addressed.

One of these general principles should be to try and understand the implications of including belief perspectives in systems attempting an interpretation of historic facts. Part of this complex set of interactions arises from the uncertainty of the "facts" themselves and part from the different cultural perspectives on the significance of those "facts" even where they are agreed. It is clearly true that the significance of an event is altered by the context in which it occurs – the most obvious examples being those where an identical criminal action may result if very different legal repercussions and hence the significance of the action must be viewed in context.

The recent French parliamentary debates on "making it a crime to deny that Armenians suffered "genocide" at the hands of the Turks in 1915" [BBC News, 2006] and the fact that perspectives on the debate have been reported and commented upon differently in different contexts is an illustration of the contemporary needs to be sensitive to different cultural perspectives in interpreting information. Clearly the varying perspectives promoted by fundamentalists and extremists in any belief system are another manifestation of the difficulties here. The degree to which civilised societies place limits on the actions taken as a result of deeply held views will also influence the potential fields of research. To present the reasons a fundamentalist interprets a set of circumstances in a particular light may be part of a cultural landscape which it helps to understand. To attribute a rational logic from whatever perspective to terrorism raised against innocent civilians is almost certainly an unacceptable interpretation of facts for the vast majority of society. Systems which seek to make logical deductions or interpretations based on an understanding of cultural perspectives would need to be aware of limits on the acceptable range of interpretations, at least from the point of view of presentation.

In the context of this research agenda such issues will not be addressed in any systematic way for the foreseeable future. What might be more tractable is to envisage systems in which potential belief is recorded as a perspective on facts or where the facts themselves are recorded within a range of uncertainty. Similar issues will arise in dealing with contradictory sources – for example secondary documentation or multiple perspectives on an event as primary sources from different witnesses.

In terms of future research priorities this analysis would suggest that the EPOCH research agenda needs to recognise that:

1. the fundamental problems underlying the semantics of cultural heritage are extremely complex and would require a great deal of basic research to address, if indeed they are addressable even in the longest timescales
2. in the shorter term issues should be examined of documenting the characteristics of different cultures' design styles in different forms and using this documentation to assist in examining the consequences of style based capture analysis and modelling tools. This underpins different cultural bases both for the terminology and for the reconstruction tools considered in the section on modelling and reconstruction. There are considerable paper based resources which have recorded the underlying elements of design styles, with some addressing them from mathematical perspectives or based on particular shape grammars (see the description of intelligent digitisation tools above and of generative modelling systems below).
3. The multi-lingual work to address challenges in terms of defining thesauri and inter-operability for documenting the cultural heritage of Europe is overdue and an essential underpinning to the semantic processing of CH data.
4. Creating and managing semantically recursive multi-lingual and/or inter-lingual thesauri as those described above is a task requiring new theoretical investigation and is probably a challenging one, definitely not trivial from an information theory perspective.

The next section considers the needs in terms of the typical processes in supporting the collection, long term availability and analysis of digital memories underpinning cultural heritage and the requirements in terms of creating such repositories.

4.8.3. Digital memories for cultural information integration

This section draws extensively from Doerr's presentation at the EPOCH research agenda workshop [Doerr, 2006].

Doerr defined his concept of *Digital Memories* (as opposed to Digital Libraries) as "Information systems preserving and providing access to source material, scientific and scholarly information, such as libraries of publications, experimental data collections, scholarly and scientific encyclopedic or thematic databases or knowledge bases." In fact this is quite close the definition of a Digital Library as conceived in the European Digital Library project in i2010. It is also in line with the new roles for local museums envisaged in Scenario 2 and in the future directions for museums in the UK [DCMS 2006b]. However, the implications, in terms of research which is required, relate not just to the range of sources, the search connectivity (i.e. integrated multiple sources), or availability from any location. All of these require some research and a great deal of development, but the truly long term research requirements lie in the need to structure the information sources in ways which enable novel semantically-based search functionalities. From this perspective there are a number of problems relating to current generation digital libraries.

There is still a widely held and traditional view of the task of libraries as institutions limited to the collection and preservation of documents and to providing assistance in finding specific items of literature or information. According to Doerr, this view of the library's role is completed "when *the* (one, best) document is handed out. 'All you want is in this document.'

This view has not helped much in raising the level of new functionality that semantic interoperability of resources would permit. In fact there are a host of problems in current generation digital libraries. In particular there is little or no support for the searches to produce new and informed responses from aggregated sources or to retrieve them by contexts (e.g. "Which excavation drawings show the finding of this object?"). There is little or no support to allow integration of complementary information in multiple sources into new insight (e.g., "What is known about the people who participated in this excavation"). Finally there is typically no support for cross-disciplinary search (e.g. to find relevant related information from the many disciplines that contribute to archaeological knowledge, such as ecology, ethnology, biodiversity, etc.). Such searches would be content based and span multiple areas of the typical library classification systems as well as requiring semantic analysis of complete sources (both text and other data types)

Doerr identified as a "Grand Challenge" (see [UKCRC] for a description of Grand Challenges and, [Arnold, 2006] for an example) the proposal that "Digital Memories should become integral parts of work environments as sources to find integrated knowledge and produce new knowledge, to create and defend hypotheses." Suitable knowledge management, which makes use of global networks of knowledge, would appear to be the key, distinguishing:

- Core ontological relationships for "schema semantics", such as: "part-of", "located at", "used for", "made from" which are localized atomic relationships, but and rich in potential structural information, relating to content.
- "Categorical data": taxonomies used for reference to and agreement on sets of things, rather than as means of reasoning, such as: "basket ball shoe", "whiskey tumbler", "Burmese cat", "terramycine". These terms define and order concepts rather than

providing structural information. They aggregate categories as opposed to integrating sources. The leaves of the taxonomic structure would be entries in a thesaurus.

- Factual background knowledge for reference and agreement as objects of discourse, such as particular persons, places, material and immaterial objects, events, periods, names. These would be elements of the taxonomic classes.

Doerr identified several preconceptions that hinder the evolution of digital libraries into digital memories with the required knowledge management support. (Table 4.1)

Table 4.1: Preconceptions hindering development of Digital Memories (After Doerr, 2006)

Preconceptions	Problems & solutions
"Libraries should not depend on domain specific needs. Domains are too many and too diverse. DLs need a generic approach."	This "seduces us" to only employ intuitive top-down approaches for generic metadata schemata. As a result, when the fantasy is exhausted, research stops. Deep knowledge engineering is required to generalize in a bottom-up manner from real, specific cases to find the true generic structures across multiple domains. Interdisciplinary work is needed on real research scenarios to identify the relevant semantics. This will involve studying professional information scientists and evaluating the support needs of the ways in which they actually work.
Preconception: "Ontologies are huge, messy, idiosyncratic and domain dependent. Mapping is the only generic thing we can do"	This statement is mainly true with "ontologies" used as "categorical data" (term lists), with which the sector seems to be transfixed. The different character of ontologies describing "schema semantics" tend to be overlooked. These may well pertain to generic classes of discourse and interdisciplinary work is needed to evaluate this potential.
Preconception: "Queries are mainly about classes. The main challenge of information integration is the integration of classes (terms)."	This preconception is not in fact sufficiently supported by empirical studies. It seems more likely that query parameters pertain to universals and particulars and relationships in real research studies. Original research questions need to be systematically analyzed to understand the way repositories are evaluated in real research situations.
Preconception: "Manual work is not scalable or affordable. Only fully automated methods have a chance"	This preconception allows us to overlook the quality of manual, intellectual decisions in favour of an "affordable automation". Yet billions of people produce content manually. Wikipedia demonstrates that the preconception is not true and that in some circumstances useful and interesting results can be generated by enabling large scale mobilization of affordable manual input. Designing interactive processes to involve users in massive Virtual Communities / Organisations in the operations of cataloguing, "data cleaning" and ontology, taxonomy and thesaurus development would allow huge quantities of data to be captured in appropriate formats. We need semiautomatic, highly distributed algorithms and genuinely interdisciplinary work to achieve this in compatible and consistent formats.

A recent TEL-ME-MOR report ([TEL-ME-MOR, 2006]) on subject access highlights the growing prevalence of user generated classification systems and quotes an extract from an online source by Elyssa Kroski:

"Today, the interest in subject access is underlined by the growing interest in categorising over the web (folksonomies).

> ”There is a revolution happening on the Internet that is alive and building momentum with each passing tag. With the advent of social software and Web 2.0, we usher in a new era of Internet order. One in which the user has the power to effect their own online experience, and contribute to others’. Today, users are adding metadata and using tags to organize their own digital collections, categorize the content of others and build bottom-up classification systems. The wisdom of crowds, the hive mind, and the collective intelligence are doing what heretofore only expert cataloguers, information architects and website authors have done. They are categorizing and organizing the Internet and determining the user experience, and its working. No longer do the experts have the monopoly on this domain; in this new age users have been empowered to determine their own cataloguing needs. Metadata is now in the realm of the Everyman.”[Kroski, 2005]

It is likely that the world of cultural heritage archives will also be affected by the combination of user-created content and folksonomies interpreting and adding to the information from both the user-created content and other more traditional sources. These user-created bodies of content and classification systems need to be related to each other and to more professionally generated content and taxonomies if they are to become part of the body of knowledge about our collective cultural heritage.

There is a growing awareness of the need for information systems which provide reasoning capability, but before any “reasoning” can be done over integrated knowledge resources, the data must be connected in a “global network of knowledge.” This requires:

- A sufficiently generic global model (core ontology with the relevant relationships).
- Methods of knowledge extraction / data transformation to populate the network.
- Massive, distributed, semiautomatic detection of co-reference relations (data cleaning) across contexts.
- Referential integrity of co-referencing needs to be curated in order to create, maintain and improve the consistency of global networks of knowledge as a continuous process (not making yet another independent database, but improving the quality of data through development of controlled new versions, whilst maintaining curation and provenance).

Only when a sufficiently rich information repository is available can we do advanced reasoning and intelligent query processing.

Research in cultural information repository generation should focus on “large-scale” information integration. It is feasible to create effective, sustainable, large-scale networks of knowledge, but the infrastructure needs have to be established to enable this.

For this he suggests four research directions:

1.) Interdisciplinary research on research processes, questions and discourse to analyse:

- relevant query and general reasoning mechanisms
- relevance of the ontological constructs we use
- necessary granularity of ontologies

2.) Research on widely applicable global models.

- Empirical evidence for global models pertaining to a generic discourse and validating generic principles in other domains.
- Formalization of intuitive concepts and understanding of their relationships.

The CIDOC-CRM is an example of such a model that requires study, development and application.

3.) Mapping and data transformation technology to extract and summarize information into a form suitable for integration.

- Extraction tools to analyse sources and extract structured information
- Architectures/techniques to integrate knowledge preserving links to the sources
- Mapping tools, data transformation generators.

4.) Identity negotiation and preservation of co-reference relations

Research in these areas is necessary to quantify the way in which ranges of interpretation contexts can be represented and analysed. This research touches on issues of national, ethnic and belief perspectives and needs to:

- research states of knowing and agreement on identity
- manage global processes of improving size and quality of co-reference clusters (generalize over authority files)
- build social organisation forums (communities) preserving the integration of knowledge.

In theory all four are needed together but in practice, without the first, research on other topics would continue to be “blind” what the real issues are.

5. Visualisation and presentation

The subject of this section is ‘visualisation and presentation’ which might be taken, by technologists, to refer in a strict sense only to image generation. However the word “Visualisation” is used in many contexts to describe the activity of reconstructing or recreating environments. Tools in this area are not capturing the shape and appearance of what exists – they are recreating an impression (hopefully an accurate impression) of the past from the present evidence.

Figure 5.1 M-PIRO Case Study. For further information - see [DigiCULT, 2005]

There are two fundamentally different paradigms to present computer generated imagery, namely films and interactive 3D. The first implies that images are rendered offline. In some ways this is less of a problem since almost all modelling tools incorporate a (nearly) photorealistic renderer, but the ease of authoring a script and assembling the assets has some way to go in order to allow effective production of filmed sequences by Cultural Heritage professionals. Hence in practice there are two other issues that are more difficult to accomplish:

(a) the creation of individual models that are to be incorporated in displays, and
(b) the assembly of compelling virtual worlds for interactive rendering.

The latter, in turn, has also implications for modeling, and this is why both related subjects are in focus in this report. The situation for visualisation and interactive rendering is characterised today by a fundamental dichotomy, the separation of the model creation from the interactive virtual inspection. The process is always basically as follows:

1. Individual model creation – either by shape acquisition or by manual modeling for synthetic shapes
2. Export and preprocessing – to make sure the created world is amenable to interactive inspection
3. Assembly and authorship of the complete 3D environment.
4. Interactive Viewing – ranging from simple 3D model inspection to rich responsive virtual worlds.

5.1. Procedural Modelling

3D modelling is still mostly an expensive affair and predominantly undertaken manually. As a result, such efforts have been focused mainly on major monuments and sites. Little attention has been given to the modelling of modest dwellings which constitute the majority of buildings such as houses, barns, small workshops, etc. Bringing the past to life again is a major goal often cited when producing 3D models. It goes without saying that such an endeavour is bound to fail if only the most spectacular aspects are visualised and the everyday part of people's living environments is omitted.

New developments in procedural modelling hold great promise for cultural heritage for modelling the vast numbers of relatively routine buildings which would be part of any historic, urban environment. Based on grammatical and parametrical descriptions of shapes (for instance of the architectural style of buildings for a targeted period) models consistent with these descriptions can be produced quickly and at low cost. As a result, large-scale modelling projects can be undertaken, as already exemplified through EPOCH's 3D model for the entire Pompeii site [Mueller et al, 2006].

Moreover, the ability to produce massive models efficiently opens up the opportunity to do so multiple times with parametrically controlled ranges of variation. This offers new ways of expressing degrees of uncertainty about the virtual reconstructions. The need to represent and convey uncertainty is fundamental for cultural heritage where images giving impressions of the past may contain large parts which are based on informed conjecture (i.e. made up, based on varying degrees of evidence).

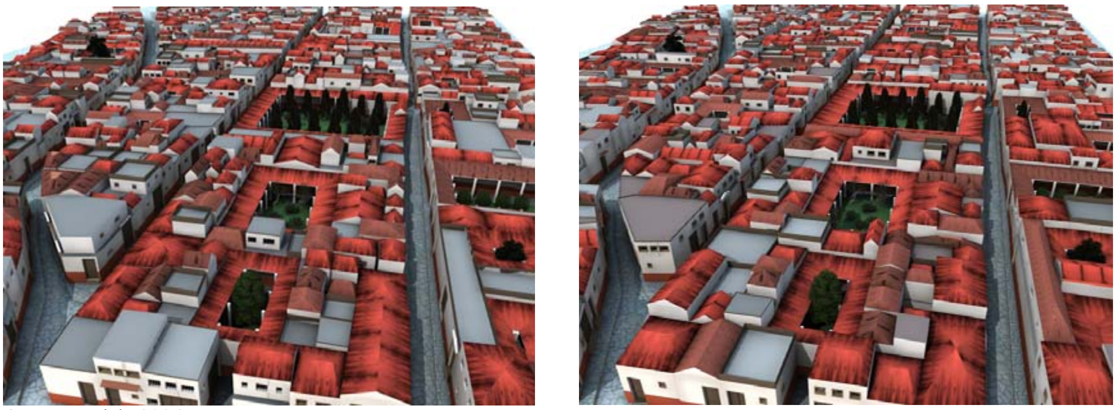


Figure 5.2 Two alternative procedurally generated models of Pompeii [Mueller et al, 2006]

Traditional approaches to expressing uncertainty have been based on showing that uncertainty within the display of a single model. However, when the opportunity is available to build reconstructions cheaply and easily, more than one can be produced (Figure 5.2) and the collection acts as `the model', rather than any of the individual reconstructions. The collection samples the space of possibilities, where elements with great certainty tend to reappear often in the form, whereas elements that are very uncertain will vary greatly from one reconstruction to the next.

This novel paradigm for expressing uncertainty would have been unsupportable without the developments in modelling techniques and should be evaluated for effectiveness alongside other, more conventional, approaches such as non-photorealistic rendering of areas of conjecture.

Technological research challenges in procedural modelling

The current ETHZ work for EPOCH on procedural modelling of urban environments [Mueller et al, 2006] has been recognised as some of the most advanced of its kind as evidenced by its inclusion in the ACM SIGGRAPH paper and tutorial sessions. Here is an overview of some of the research challenges, from an EPOCH specific point of view, as well as beyond.

1. Combining procedural modelling with multi-resolution and generative detailing. In EPOCH's terms this research area would combine the shape grammar functionality of procedural modelling to integrate parametric shape descriptions (such as those offered by the Generative Modelling Language (GML) in order to create shapes with adaptive, dynamic level of detail to be included at the level of leaf nodes in shape grammars) [Berndt et al, 2005].
2. Efficient level-of-detail (LoD) descriptions should involve some level of semantic understanding. Good results are difficult to achieve when only operating at the level of triangles and their textures. The shape grammar approach provides a natural, hierarchical description of buildings, and it should be investigated in how far this hierarchy can be automatically translated into better LoD representations. This is of paramount importance if one wants to efficiently render the large scenes that procedural modelling can help to generate.
3. *High-detailed 3D modelling for non-expert users*: Entry-level 3D modelling tools can have great impact, as for instance [SketchUp] which permits all users to easily create a rough sketch of their houses, and to insert them into GoogleEarth. However, creating highly detailed architectural models with such tools is too laborious. The CHARISMATIC modelling suite [Day et al, 2004] demonstrated the use of modelling operations based on constraint-based drawing-type interactions and provision of detailed style-based libraries for architectural details. The combination of easy sketching operations and interpretation, constraint-based modelling and style-based libraries of procedural detailing would provide a greatly improved tool for researchers and scholars to formulate their hypotheses in 3D in unprecedented precision. Improving the availability of such modelling for inspection by peers (including the decision making and selection of alternative interpretations) would also lead to more rigorous reconstructions.

4. *Integration of additional layers /GIS:* Current modellers (including procedurally-based and more conventional systems) already have some provision to cater for streets and vegetation. However there is insufficient integration of general purpose modellers and domain specific needs in these areas. These facilities need to be improved and additional layers should be included. Moreover, the reconstructed models need to be coupled directly to GIS systems, but the mechanism and formats for this will depend on ongoing discussions about appropriate geographical standards for cultural heritage (KML vs. GML, for instance).

5. *Standing structures as constraints*: Often, buildings are still intact or still have some substantial standing structures that remain. In that case, a modelling tool for architectural heritage should take images or scans of the standing structures as input, and use this information as constraints on the allowable reconstructions, potentially combining this with knowledge of the rules associated with the style represented by the remaining evidence of the standing structures. The task would be to build a 3D model from images or scans, but with the important, additional help of a grammar describing the architectural style of the building and the important additional task of extending the fragments of standing structure to achieve a complete building meeting both the constraints of the standing structure and other information about the full extent of the original structure. In case the standing structures are incomplete, part of the task will in most cases consist of filling in missing structures, based on the grammar and possibly additional iconography.

Figure 5.3 (a) The Site of the former Ename Abbey today (b) The virtual model of the abbey complex at Ename (Belgium) superimposed on the current setting

Figure 5.3 (c) Aerial View of the virtual model of the abbey complex at Ename (Belgium) superimposed on the current setting. For more information see [Vereenooghe, 2004]

6. *Coverage of styles*: In the same way as OCR techniques have developed to allow for increasing ranges of font styles (both current and historic) the vocabulary of design styles and composition rules needs populating and the range of compositional rules needs refining. This is a substantial piece of work, which may become informed by the research proposed under 7 and 8. However in the interim and in order to feed the research proposed there with base analysis and case study data, there will be a need to investigate a range of styles through more traditional analytic research. Some substantial work has already been under taken in these areas by the design community for some styles (most notably classical Greek and Roman architecture), but this needs to be elaborated and re-thought in terms of the parameterised algorithmic nature required by shape grammars and the reverse processes of style recognition.

7. *Learning grammars*: With procedurally-based modelling, based on rules which implement particular styles and classes of object, a lot of the detailed modelling work is built into the modeller and taken off the shoulders of the person (investigator) producing the final model. On the other hand, the underlying grammars and parametrical descriptions have to be generated. This is something the investigator might have to do, if a similar style has not been previously captured by someone else. If such style grammars could be created automatically, or semi-automatically, drawn from analysis of images of representative instances of a style this would further reduce the effort for future investigators and make the generation of reconstructions more efficient.

8. *Determination of style*: Once the grammars for a sufficient number of styles have been developed, and grammatical descriptions can be fitted against images as under item 4, one

could compare the success different styles have in explaining the visual data about an object like a building, and automatically determine its style or mixture of styles.

This section has concentrated on the specific research needs identified in procedural modelling as a technique for visualisation of architectural reconstructions and for compact representation of very large models of urban environments. The property of compact representation of highly detailed models is shared with other applications of procedural modelling including the sub-division surface techniques for architectural detail and the procedural modelling of materials. Procedural modelling appears to be the best practical alternative at present for detailed modelling of extensive environments. Architecture was taken as a good case in point, but procedural modelling should be seen as a paradigm with generic potential for any type of objects that are structured according to some kind of `rules', be it natural as certain fossil types in palaeontology or man-made like art nouveau vases in art history.

In the next section we will discuss briefly some other more general research needs particularly in the presentation area.

5.2. More general research topics in visualisation and presentation

Having addressed the issues of representing reconstructions of large scale historic environments efficiently the major outstanding issues relate to the needs to present these environments and their associated information to the target recipients. These presentation situations have different characteristics depending upon the purpose of the presentation, the major forms of which will be presentations to cultural heritage professionals for assistance in analysis and research and communication with a wider audience with educational and/or recreational objectives. Three major sub-areas are envisaged:

a. The linkage of models capturing the appearance of environments to the underpinning structured knowledge base about the history and significance of the environment's constituent components. This is an issue at both scene assembly and during the presentation of the environments:

   - Tools and techniques to allow presentation of provenance, paradata (c.f. London Charter [London, 2006]), interpretation and uncertainty
   - Asset management and version control (What is the significance of editorial actions in terms of derivative works? How does the investigator record, visualise and understand the provenance of the range of assets being assembled? What the fundamental operators for combining uncertainty and provenance of individual

components to produce the provenance and uncertainty measures of the total assembly?)

- Revealing the underpinning information (what are the appropriate paradigms for communicating the non-visual information associated with the components of the scene?). Given the full richness of future knowledge bases this needs to include appropriate filtering and packaging of the available information as well as navigation tools to allow the professional or interested amateur to explore beneath the surface. At the same time it must be possible for curators and other professional exhibition designers to create purposeful yet interesting and engaging presentations with sufficient accurate information.
- Representing uncertainty acceptably is still an important research topic with both the traditional means of varying the presentation mechanisms and the development of alternative content approaches to be investigated further for efficacy and impact.

b. Authorship tools to allow non-ICT specialists (and in particular cultural heritage professionals) to experiment, design and author using extensive digital assets in appropriate ways.

- Authoring tools tailored to cultural heritage presentations, linked to digital cultural heritage assets embedded in a digital context
- Authorship tools for cross-platform and multi-platform interactive systems (e.g. delivery via iTV, computer games machines and other domestic-level technologies, internet and location-based immersive VR/AR systems in memory institution visitor venues)
- Frameworks for authorship of multi-cultural, multi-national, and multi-lingual presentations and multifaceted interpretations. These frameworks will need to be based on mechanisms that allow the varying perspectives to be captured and represented.
- In the longer term, understanding the characteristics of an interactive experience that produce a sense engagement for the user and developing measures of engagement to inform authorship tools and assist authors in producing engaging experiences

c. Novel interface techniques for exploring the resulting environments. In general these areas relate to adopting and appropriately adapting generic techniques and technologies. These will become specific to the cultural heritage field where there are specific requirements. For example there will be specific issues in adapting generic haptic technologies and allowing these to address the needs of interfacing to measure of uncertainty (perhaps represented by deformable objects) or provenance.

- Development and deployment systems for augmented reality interpretations used with replica and/or original artefacts. This area involves the augmentation of the presentation of physical artefacts using virtual presentations involving digital assets. The technique has been used in a number of very promising early adopter schemes as a way of adding to the experience of the physical object in ways which preserve the security and integrity of the original pieces. In particular the ability to investigate the detailed internal structure and magnify the intricate aspects in a virtual surrogate allows the visitor to appreciate the detail of the original artefact at a level which could not be achieved with the original. This also helps to address issues of access for the visually impaired.

Figure 5.4 Augmented reality reconstruction of the Sagalassos Site

- Experimentation with ambient technologies in museum and visitor centre contexts. This topic covers the use of location based sensors, visitor identification systems and controlled, focused environments for sound. These technologies are under widespread research and development for many applications ranging from customer service environments, office operations and security systems. The museum and visitor centre sector provides a range of challenges which are different to those of other contexts. For example the need to offer multi-lingual capability with visitors of unknown origin might be shared with customer service environments in airports, but is more prominent here than in the controlled office environment. The major issues are likely to be those of integration into other museum and visitor centre systems – linking to the digital repositories and presentation systems for example. The issues to be addressed are nevertheless significant, particularly for wide-area, open-air sites and those with the densest throughput of visitors.

- Effective personalisation of experiences for repeat visits or linked venues. This area concerns the identification, profiling and linkage of activities of individual visitors. There are issues of privacy, which it is assumed will be overcome by making the enhanced services sufficiently attractive as to ensure the visitors will wish to use the services. Mechanisms for identification are a generic issue being widely trialled in commercial and security operations with technologies being continually improved for identifying individuals. More challenging is to understand effective ways of using the identification of the individual to profile their interests, tailor experiences and provide value-added services. Given the potential to offer multi-cultural perspectives on the information presented, the capability to tailor experiences must also take the research into the realms of ethical issues. For example should a tailoring system automatically challenge or reinforce particular individuals' perspectives by choosing confrontational or reassuringly familiar perspectives from which to present the stories of the past. The first might be considered more educational; the second might make better commercial sense.
- Appropriate use of haptic technologies. Of the five senses, most effort has been placed on sight and sound as the vehicles for presentations and interactive systems. Touch is probably the next most promising (i.e. more likely to be effective than smell or taste in Human-Computer communications). Haptic cover both technologies to provide tactile stimulation (i.e. sensations at the skin) and force feedback (i.e. exerting force on the muscular-skeletal system). Technologies targeted at exploiting this potential have existed for many years, but have made little penetration into cultural heritage applications. Where they have been deployed in the theme park sector, they have not been universally successful in operations despite a more favourable investment regime. There are a number of good reasons for this. In particular robust technologies have been very expensive and the cheaper devices lack the richness of interactive capabilities, quite apart from concerns about sustainable operations. Operational concerns include health and safety aspects (particularly hygiene) and the robustness of devices under repeated use. Nevertheless it can be anticipated that haptic devices may have a place in the range of interactive

techniques used in cultural heritage, particularly as cost: performance improves and devices become more robust.

Aspects of haptics that need more investigation include:

- What aspects of touch can be usefully exploited in cultural heritage contexts?
- If surface texture is an important communication opportunity how should it be represented in the environment?
- What are the architectural implications for applications using haptics, which require a much faster refresh rate than images, with the very large environments envisaged?
- What cultural heritage research investigations would be enabled by having haptics embedded, and what would be the consequent needs elsewhere in cultural heritage research?
- What additional presentation opportunities might be augmented by using haptics? Examples could include replaying the process of restoring an object, or the sensations of making an artefact (e.g. cutting gemstones to create jewellery, or knapping flints to make tools).
- For example, haptics have been proposed as assistance to an archaeologist in reconstructing large architectural features. The idea is that digital models of the heavy extant fragments of structures can be reassembled virtually using haptics to provide tactile feedback of the fit of the pieces. To enable such an application would of course require the accurate digitisation of all the potential fragments. Given that these may weigh several tons and, in situ where they lie, they would normally be partially buried, even where they are accessible. The application therefore is predicated on the assumption that acquiring digitisation of the stone fragments is practical – by no means a foregone conclusion.
- The INTUITION NoE on Virtual Reality has a working group developing a research agenda and road map for haptic technologies ([INTUITION, 2006]).

➢ Speech-enabled, multi-lingual systems accessing domain specific knowledge for story-telling and/or research. "Speech-enabling" involves a number of generic technologies. For conversational systems, speech recognition, speech analysis (commonly involving partial statements as opposed to the complete sentence structures assumed in most written text analysis), context awareness, negotiation of common understanding, response generation and voice production are all parts of a speech interaction cycle. All of these would be involved in any generic conversational system, although the breadth of untrained voices to be interpreted makes any visitor-orientated application including cultural heritage particularly challenging. Areas where cultural heritage presents significant additional challenges include the specialist thesauri involved as well as the range of languages (including historic languages).

- Avatars in interactive public environments. Avatars, including virtual humans are one presentational mechanism for interfaces including conversational interfaces. The generic issues which arise here are mainly to do with the acceptability of particular types of virtual communicator. There has been an ongoing search for techniques to define and present the most life-like of virtual humans, but there is still a degree of scepticism that a faithful reproduction of a human is attainable or desirable. The attainability concerns reflect the sensitivity of human recognition systems to slight imperfections, which may concern both the appearance, including the motion and deformation of soft tissue, and to the virtual human's reactions – which relates to a mixture of appropriate expression, responsiveness and apparent intelligence. An alternative to a faithfully reproduced virtual human may be a characterisation which makes no pretence at fidelity but provides an engaging communication style.

Figure 5.5 Virtual Tour guide in Virtual Wolfenbüttel

- Finally in the interfaces area it is important that interfaces and the technologies/techniques that support them are developed to allow cultural heritage professionals to work in their domain of expertise rather than become ICT specialists in order to use tools developed from the technologists perspective. To do this analysis is required of the ways in which real cultural heritage professionals undertake specific tasks and the interfaces to applications need to be built for the convenience of the CH professional in undertaking their normal working tasks. For example it would be desirable to produce interfaces such that a CH professional can use it to create and manipulate structures in a specific style/shape grammar, expressing operations in terms of the normal vocabulary of operations used in that domain. This sort of approach is essential to capture the skills of the professional in (in this case) evaluating alternative interpretations of the extant physical remains efficiently.

5.3. Additional ICT support in environmental assessment, restoration and reconstruction

In addition to modelling for visualisation there are also several potential applications for digital restoration techniques (e.g. simulations of stone deterioration).

Conservation specialists have an in-depth knowledge about erosion and weathering processes of stone, and how stone breaks when subject to strong forces (such as earthquakes). If we could model such processes in 3D and time, we could envisage using this knowledge to reverse the processes in digitised statues, monuments and archaeological remains. Having this kind of intelligent tool, combined with the style based, domain specific understanding of the range of likely shapes involved, means that it may become possible (probably semi-automatically) to reconstruct digitised weathered or damaged objects into the representations of the original and undamaged objects. (See for example the image below (Figure 5.6) where the damaged archaeological remains in Sagalassos are digitally put together through software that matches break surfaces. Turning this assembly into an unbroken and undamaged frieze remains a fully manual process currently even without specific objectives to understand and reverse the weathering processes).

Figure 5.6. Original and digitally restored frieze of Nymphaeum - Sagalassos

This kind of computer aided digital restoration would be very useful to show the splendour of the original object without altering the physical original (which would be, in most cases, a quite irreversible and inappropriate process). Similar reverse ageing processes could be envisaged applied to timber frame building construction, which distorts over time as the beams sag.

Figure 5.7 11th century ivory object with different types of erosion, damage and deformation

In the photographs of the object in figure 5.7, we see several forms of local erosion (due to the acid nature of the soil), damage (the ivory is broken in several places and parts are greenish due to corrosion of the bronze pins) and deformation (the parts do not fit together anymore) that could be corrected if good models and tools were available to “undo” these shortcomings on the digital model.

By experimenting with digital restoration, the restorer can learn a lot about the structure and creation of the object. By showing the restoration as a time-lapse animation, one can show the nature of the damage and create insight for the visitor into the work of the restorer.

This modelling of weathering, damage, colour bleaching, salt crystallisation, wear and erosion processes is a very multi-disciplinary approach, capturing the knowledge of conservation specialists. It is especially relevant in popular visitor centres where the visitors themselves are a major source of damage to the environments they visit. Figure 5.8 shows the hall tiling in the a Chateau in the Loire valley where generations of visitors have not only worn the surface of the tiles and removed all trace of imagery over most of the floor, but the underlying tiles themselves are worn to a depth of perhaps 1cm in areas of major footfall.

Figure 5.8 Worn tiles at main visitor entrance of a much-visited Loire chateau

Some research is going on in this field (MIT, Microsoft Asia) but it is still in its infancy.

5.4. Specific Issues for Web Access and Dissemination

This section draws on the contribution by David Bearman to the EPOCH Research Agenda workshop at VAST2006 in Nicosia. Bearman's contribution addressed various museum issues and open research questions in an emerging "Research Agenda for Museums on the Web". In this section the aspects most related to the specific needs of operating over the web are emphasised although the contribution also re-emphasised other aspects of the research agenda

One aspect of the potential for museums on the web is that it is important to articulate open CH ICT research questions in a way so that museum directors and staff see how the museums work could benefit from ongoing research in the field. It is important to recognise that (as Bearman stated) "Museums aren't well endowed or technically sophisticated so solutions need to be easily implemented", and "Evaluation, in actual contexts of use, is critical."

The role of web technologies in the "Museums on the web" context is to re-emphasise some attributes of application situations. In particular the limitations of communication speeds highlight the need to consider what constitutes sufficient accuracy. Museums often hold large numbers of digital objects, for use in a variety of situations, but efficient web delivery raises some additional issues:

- What constitutes the "best" representation for different kinds of objects and purposes?
- What will make content reusable and linkable?
- Why should some content be better retrievable than other content?
- What would constitute a full rendering of a(ny) museum object?

Information not embedded in objects helps us understand them. Such context is recorded

- in museum and other library publications,
- in registration records and archives,
- in reference sources on the web

Linking the digital objects to the additional data requires the metadata and the spatial information to be linked – which is precisely the issues raise in the discussion on 3D container formats in the standards section. All these questions touch on some of the issues raised elsewhere in this document, but are focused by the implications of distributed systems and internet technologies. For example the completeness of a rendering touches on the concept of "accuracy which is fit for purpose". Similarly the needs to link to other information will be influenced by whether the digital object is to be re-used or whether its display constitutes delivery of the museum's commitment to the user. An artefact which is to be incorporated in some derivative work might be expected to include additional information, not least the terms of its reuse.

Museums have historically created interpretive labels for each exhibition which are discarded after a single use. When the “exhibits” are available on line the information associated with them take on more significance and the museum knowledge contained in the “interpretive labels” will automatically have greater longevity. Integrating this data with web functions offers the opportunity for enhanced museum services, but raises the question of how to make the knowledge available most usefully.

Accessing resources over the web is a very personal experience for the user, but of course the opportunity to enhance this experience relies on identification of the user, and further enhancement to take into account the location of the user at the time of access would also potentially add value for the user. Internet visitors are at least as heterogenous in character (and disabilities) as those visiting in person and their specific requirements are not as easily identified when personalising services. All the previous comments on issues of multi-cultural perspectives are exaggerated when dealing with internet visitors and success in dealing with multi-cultural issues appropriately relies on all the same issue of representation and interpretation of varying cultural perspectives.

There is an important role of web technologies in the developing area of user-created content. The web provides an obvious vehicle for linking museums with their communities and for allowing users to add to digital content associated with museum resources in the form of social tagging and contributions to folksonomies. This engagement with the museum will also assist in promoting the accessibility of the physical museum experience – making it less alien for the typical member of the public.

As with all uses of technologies in museums additional research needs to be undertaken into understanding what works with visitors – what generates the sense of engagement and resulting understanding? What technologies work with real visitors? What are effective paradigms for making the message stick? How can monitoring of visitors and their use of interfaces (whether over the internet or in person) help in evaluating the effectiveness of the visitor experience. Over the internet there is the additional factor of the degree to which the museum’s curation of a collection is mirrored in the relatively unmediated access available over the internet. Specific control of routing through a web presence for an exhibition might be one solution to providing curated experiences, but these can be expensive to generate and more evaluation of the user acceptance of constrained navigation is needed. What other modalities of interaction are appropriate both for in situ and remote experiences?

The world wide access to museum collections will also highlight areas of controversy over the origins of museum collections and the ways in which they may have been collected historically. This may increase pressure for repatriation of some holdings to or restitution to indigenous communities. The alternative of virtual or digital repatriation may serve some purposes, but what technologies would be involved and how would assets be managed?

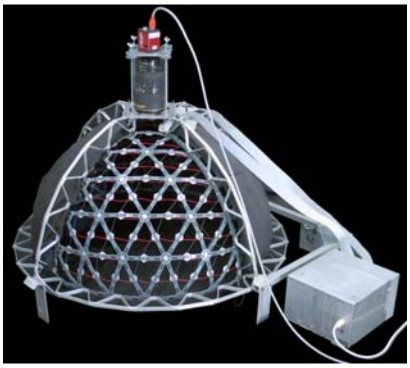
Finally there are the issues of the business models involved in providing web access. What are the most appropriate business models for museums offering web access, to enable them to stay in business? Open content may be desirable for the visitors, but how do museums stay in business? Can museums realistically augment their collective return-on-investment by adopting standard data and architecture today? Experience from online publications might suggest that online access to collections would encourage more physical visitors but this hypothesis would need more testing and evidence before it would be an appropriate planning basis for a sustainable and secure future.

6. Mobile, distributed and networked systems.

Many of the issues here concern generic technologies. However there are specific issues concerned with the design of system architectures which are suitable for use in supporting the business processes of the various cultural heritage professionals. The cultural heritage component consists of designing and implementing CH-specific components on the required variety of platforms and using communications protocols that follow agreed standards for sharing cultural heritage data, maintaining provenance etc. Thus the challenges are similar to those of integration and interoperability of data, coupled with the implementation of rich functionality which implements CH requirements (e.g. recording of excavation data: contexts; artefact scans; images; textual descriptions; positional information; etc.) on less capable hardware (e.g. PDAs). Some of these needs might be implemented on mobile devices interfacing to applications implemented over more generic GIS technologies.

- Many (some would argue mainly) generic technologies
- Specific issues with design of system architectures suitable for integration in broader CH professionals' business processes
- Cultural heritage component in design/implementation of CH-specific components for cross-platform systems
- Integration and interoperability of data, coupled with the implementation of rich functionality which implements effectively CH requirements (e.g. recording of excavation data: contexts; artefact scans; images; textual descriptions; positional information; etc.) on less capable hardware (e.g. next generation PDAs).
- Standards for cross-referencing and sharing cultural heritage data with remote sources
- Maintaining and extending associated provenance etc. whilst extending information base
- Interoperability with generic cross-platform applications (e.g. appropriate GIS systems)
- Design of mobile versions of CH-orientated recording hardware (e.g. mobile light studio)

Figure 6.1 Mobile Light Studio (a) Assembled and (b) Dismantled for transport



© ETH Zurich and KU Leuven, 2006

7. Long term preservation and upwards compatibility

There is increasing recognition within the ICT community of the different and extreme requirements placed on archival systems by the cultural heritage domain. Archival in ICT environments has rarely meant preservation over a period of more than perhaps 10 years. Recently the advent of several important initiatives for technologists (e.g. the opening and subsequent closing of the Boston Computer Museum and amalgamation into the Boston Science museum (http://www.mos.org/topics/tech_and_engineering?exhibits_shows/current_exhibits&d=214) and the 60th anniversary of the code cracking operations at Bletchley Park in the UK and the desire to develop replicas of the processing engines used there (http://www.theregister.co.uk/2004/06/01/colossus_remade/)) has led to an interest in the curatorship of computers and their associated information. However any of the timescales involved here pales into insignificance compared to even the shortest of timescales involved in cultural heritage and highlights the lack of long term preservation in the world of hi-tech. After all, an artefact is not technically considered an antique until it is 100 years old and the age of archaeological heritage is more commonly measured in 100's or 1000's of years.

In these circumstances it is not surprising that the archival regimes developed for the computing industry have a number of shortcomings when the archive of digital cultural artefacts is being considered. Obvious problems are concerned with both the physical media (both durability and the obsolescence of the equipment capable of reading it) and the logical formats (knowledge of the formats, version control and the software maintenance of systems for interpreting them). The following is a list of major sub-areas of concern:-

- Formats (Standards, Encodings, Metadata, Provenance, Paradata)
- Logistic processes for long term preservation (media, regimes, security, resilience, redundancy)
- Legal frameworks (IPR, copyright, licensing, royalties, grey literature/documentation, metadata rights, collected works, derivative works, orphaned works, etc.)
- Business models for long-term preservation (Responsible authorities, Legislative requirements, Secure financial basis, etc)

The Digital Preservation Europe project recently established under the 6th Framework is acting as a forum for research in this area and is working on a roadmap for digital preservation which it is expected would be adopted by EPOCH.

8. Standards and interoperability

Agreed standards are an essential part of reaching the vision of an integrated world of cultural heritage knowledge. The EPOCH project has worked with three distinct areas of standards:

- Metadata standards
- Technical infrastructure standards
- Charters and guidelines

In each of these areas there is considerable work to do to achieve genuine adoption of standards which maximize the opportunities for creating interoperable systems with rich functionality, capable of delivering on the potential benefits of ICT applications to cultural heritage.

There are several distinct but overlapping phases in standardisation

a. achieving agreed definitions

b. gaining adoption with defined interoperability with other standards

c. initial deployment in real applications as "proof of viability"

d. implementation with a critical mass of applications/data

A huge amount of effort is needed to achieve agreement between a representative international group of experts on the specification of any standard. This is in part because standards are defined in natural language (for technical standards, usually in English) which is inherently ambiguous and open to interpretation. The process of standardization consists not only of agreeing the functionality to be achieved but of ensuring that the functionality is described unambiguously.

The danger in seeking a fully considered standard is that the time taken to achieve agreement may be so long that the technologies it addresses, or the need it seeks to address, have been bypassed in the interim. Conversely a widely adopted standard with limitations may well continue to be used in both legacy and new systems for the degree of interoperability it offers, even when there are newer standards of much greater potential because the newer standards do not initially bring with them the same widespread adoption.

In some application areas, communities have taken general purpose standards and defined Application Specific Profiles or ASPs to agree on how a particular standard will be used in that application area. This profile might include specification of how optional parts of the standard will be used, what parameter ranges are included, and which encodings of the basic functionality are mandated. A metadata application resulting in the combination and re-use of different metadata element sets is also sometimes referred to as an application profile (see [Dekkers, 2001]).

Figure 8.1 summarises the EPOCH approach to contributing to work in standards, which involves engaging with the cultural heritage community and disseminating knowledge and training in the use of standards. In addition the EPOCH standards team examines the best specialist approaches to deploying appropriate standards. In return the community provides feedback on any specific shortcomings which are relayed to the standards community.

Although recent years have seen the development of standards targeted specifically at the definition and processing of cultural heritage data there is still a great deal to do before the vision of applications accessing widely distributed, interoperable, digital, cultural heritage resources.

The EPOCH Deliverable D4.2.1 [Niccolucci, 2006] gives an overview of current standardization activities and issues in EPOCH.

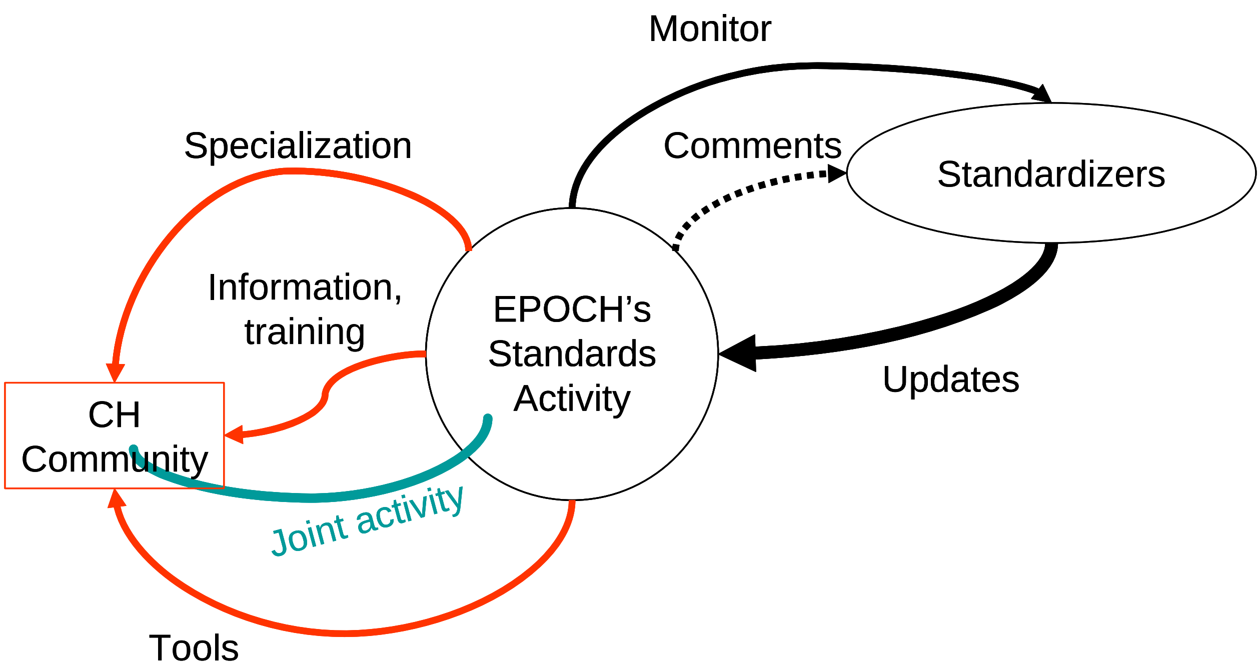


Figure 8.1.

8.1. Metadata standards

The most obvious and significant development has been in the finalization of the ISO version of CIDOC-CRM [ISO 21127]. Early work in investigating the mapping of CIDOC-CRM concepts to particular working practices has shown area where small extensions (which are permitted in the standard) would enhance the compatibility with national classification systems and hence the interoperability of applications using the standard with legacy systems based on national standards.

Other work on adoption of the CIDOC-CRM has suggested the desirability of an extension to accommodate the recording or the processing history and other provenance associated with a digital artifact.

CIDOC-CRM is in the relatively early stages of adoption with significant numbers of projects adopting the standard and working through the teething problems in operating with other systems. CIDOC-CRM is being introduced against a backdrop of considerable investment in metadata standards in the Digital Libraries community.

Much of the metadata standardisation in use in the digital libraries community is based on or mapped to the work of the Dublin Core Metadata Initiative (DCMI) and the Library Community within DCMI. The latter "is a forum for individuals and organisations in the library domain to exchange information and experience about the use of Dublin Core metadata in a library context and to encourage best practice. More specifically, it provides a forum for discussion of issues related to the DC Library Application Profile." [DCMI].

The Dublin Core also underpins work in a number of other communities and has been the target of mappings developed in other standardised environments. For example the Library of Congress has defined mappings from their systems to and from the Dublin Core Metadata Element Set [ISO 15836]. The Library of Congress also uses the METS (Metadata Exchange & Transmission Standard) as an interchange standard for metadata resources.

As the name implies the Dublin Core, however, only seeks to address a core set of elements and is intended to provide standardised definition of the elements likely to be included in the core metadata of library entries. Dublin Core is the announced basis for the MICHAEL project's exchange of museum catalogues to enable European museums to exchange information about their collections [Rosella, 2006]

The CIDOC-CRM approach provides a much richer set of relationship encodings and is based on the requirements of metadata suitable to encode information about cultural artefacts and their history. There is a simple mapping defined from Dublin Core Metadata Element Set to CIDOC-CRM, but this does not make appropriate use of the potential of the CIDOC-CRM to support functionally rich applications. To do this requires additional information relating to the primary sources to be encoded. The current major question is whether a sufficient body of effort will be put into generating a critical mass of digital resources with the additional information.

To enable the creation of a critical mass of data requires tools to semi-automatically assimilate and encode the required capture of new data and conversion of legacy metadata. These are rightly included in the sections above on digitisation of legacy data and metadata. There is also potential for intelligent tools to assist by using language processing technologies to extract metadata from free text and relating multiple sources to establish co-referencing.

The Canadian Heritage Information Network's website provides an excellent overview of standardization efforts at various levels [CHIN].

8.2. Technical Infrastructure Standards

Technical Infrastructure Standards is a term used within the EPOCH project to describe the standards used for the underpinning technologies such as graphics and GIS systems.

The major work to be addressed here is to define the relationship between such standards and the additional information held in cultural heritage collections, encoded in metadata and free text as covered by the standards in the previous section. There is a need to inter-relate the information held in each so that for example a real-time interactive display of a virtual reconstruction can draw information on the components of the virtual scene in response to interactions from the viewer.

The EPOCH project has adopted the OpenSG graphics standard for the scenegraph representations of 3D objects and environments. The OpenSG framework for 3D rendering is an Open Source, portable scenegraph system to create real-time graphics programs, e.g. for virtual reality applications, built on top of OpenGL. Additional standard mechanisms are required to incorporate the links to additional information within the representation of the scene and to provide the hierarchical nested structures implied by the scenegraph structure.

EPOCH has been investigating the use of COLLADA or METS as potential structures to incorporate the mix of 3D geometry and visual properties with links to metadata describing the provenance of the 3D objects. Since objects may be nested one in the other and instances repeated multiple times, the cross-referencing must support nesting and the aggregation of pieces of provenance for the individual components into the collective provenance of a derived work.

At present these considerations remain under review and there will be more research required to fully appreciate the consequences of selecting one standard over another. One concern must be that although the COLLADA standard is well supported by industrial backers from the games industry it remains a proprietary standard, whereas METS is based on internationally agreed formal standards.

8.3. Charters and Guidelines

EPOCH work in these areas has focused on three major topics: The London Charter for the use of 3D visualisations; Usability Guidelines for interactive experiences and the Ename Charter for heritage interpretation.

The initial version of the London Charter was debated at the **London Symposium and Expert Seminar** "Making 3D Visual Research Outcomes Transparent", which took place on 23-25 January 2006 (http://www.kvl.cch.kcl.ac.uk/Symposium/index.html).

Following a vibrant discussion in the symposium, a draft charter for the use of 3-dimensional visualisation in the research and communication of cultural heritage named "The London Charter" was eventually produced at the experts' seminar.

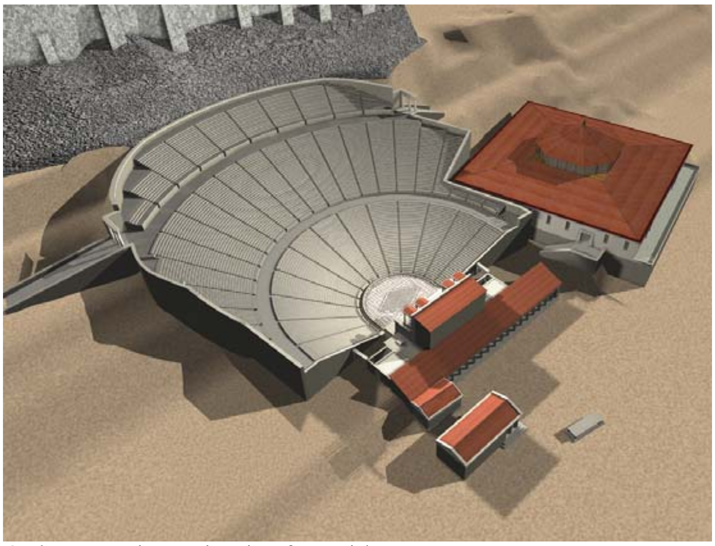


Figure 8.2 Reconstruction of the Greek Theatre of Dionysis [Baker and Beacham, 2003]

A separate topic has concerned usability of IT applications in which EPOCH has carried out a **survey on usability** the results of which have been published in a guideline whose contents have been presented and discussed at EVA 2006.

The **Ename Charter** has continued to be promoted and EPOCH is seeking to foster its adoption as an ICOMOS charter. A web site for its dissemination (http://www.enamecharter.org/) has been established and an international symposium on the Charter was successfully organised held in the USA in May 2005. More than 160 papers were submitted, most acknowledging the Ename Charter as a standard for heritage interpretation. At the conclusion of the Symposium, a declaration was formulated describing the main points of the conference discussions and offering continued support for the ICOMOS Ename Charter Initiative. The Ename Center has recently been appointed as Technical Secretariat of the ICOMOS International Scientific Committee on Interpretation and Presentation [ICIP].

It will continue to be appropriate to formulate guidelines on the best of current practice in areas relating to operations or practical and ethical behaviour and to liaise with representative bodies such as ICOMOS on the formulation of charters and dissemination to appropriate audiences.

8.4. Open data formats and Open Access to datasets

In the field of academic research and publication the Open Access movement has found its expression in several declarations such as the Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities [Berlin, 2003] and the OECD Declaration on Access to Research Data from Public Funding [OECD 2004].

However, as a matter of fact, in many disciplines projects aggregate large amounts of datasets, but only a smaller fraction of processed data finds its way into publications. In addition, such publications usually come in closed, inaccessible formats. According to researchers from ArchaoCommons, Alexandria Archive Institute, this may represent only about 5-10% of primary documentation, which is a tremendous wasted effort. (cf. [Kansa et al, 2005a], [Kansa et al, 2005b])

The reasons for this waste are manifold and include the fact that researchers tend to see the datasets as their property, there are limitations to effectively manage and make available datasets, and there is little professional reward for "raw data".

Open Access to collections of datasets would not only allow for valorising considerably the initial investment through further analysis and research publications. Such collections of datasets would also represent important material for teaching, coursework, exercises, and further studies of students.

8.5. Standards and future research

Table 8.1 summarises the future research requirements in different aspects of standardisation including entries relating to the tools required to assist in standardised encoding of Cultural Heritage data.

Table 8.1 Future Research Requirements in different aspects of standardisation

Activities	Monitor	Specialize	Analyse implications	Adopt EPOCH's	Propose
Topics					
Documentation standards/formats	yes	yes	yes	yes	Comments

Tools	yes	yes	yes	yes	New
Technological standards/formats	yes	yes	yes	yes	
Usability guidelines/interfaces	yes	yes	yes		
Quality guidelines & evaluation	yes	yes	yes	yes	New
Charters (guidelines)	yes	yes	yes	yes	New
Training and documentation	yes	yes	yes	yes	New

EPOCH has produced an overview of standards used in museums [EPOCH , D4.2.1]. This reveals that in the heritage domain there is very little standardization throughout Europe. There is a lack of accepted all-European documentation standards, which only exist in some areas, for example libraries, whilst no trans-national agreement exists in most other areas, such as archaeology or monuments. This adds to the difficulties of multilingualism and the lack of multilingual thesauri, which cannot be easily overcome by technology alone. The effect is to jeopardise digitization efforts, and it has been underestimated in establishing digitization policies, as far as they just rely on "core" metadata which actually guarantees very little information about the cultural content.

In this regard technology may offer substantial support to the unification of documentation by providing tools for mapping local, national or de-facto standards to each other or to an accepted international one. Although requiring a consistent effort, mapping appears in fact as the only way to overcome the idiosyncracy of heritage professionals to standardization and to deal with the huge amount of digitized legacy data. As shown by preliminary activity in this field, mapping is not a straightforward exercise and the technological aids must be able to cope with a number of complex and intriguing cases.

9. CH ICT maturity life cycles and different perspectives of stakeholders

A major component of the Common Research Agenda is to recognise that different communities have differing perspectives on the perceived maturity of individual technologies. These perceptions have a fundamental impact on the communities' perception of the usefulness of pursuing a topic and hence on the priorities embodied in the agenda.

For these reasons as part of the Common Research Agenda we are developing a model of the technology maturity life-cycle to serve as a tool to collect the views and assessments of the different communities and, in order to provide a consolidated shared framework, to summarise the different perspectives.

Some of the following sections already have been included in a previous version of the Research Agenda, however are presented again below with additions and further details from the ongoing research and experts consultations.

9.1. Technology maturity model

As a basis for developing a shared framework that acknowledges the different perspectives of technology researchers and developers, cultural heritage institutions, and companies and competency centres active in the field, we use the standard model of how technologies develop and achieve a broader level of use (cf. [Rogers, 1962], [Rogers, 1995]; [Moore, 1991]; [Hudson, 2002]).

9.1.1. Standard model of technology diffusion

This process model does not include the research and technological development which gives rise to new technological methods and prototypes, which we include in the diagram below.

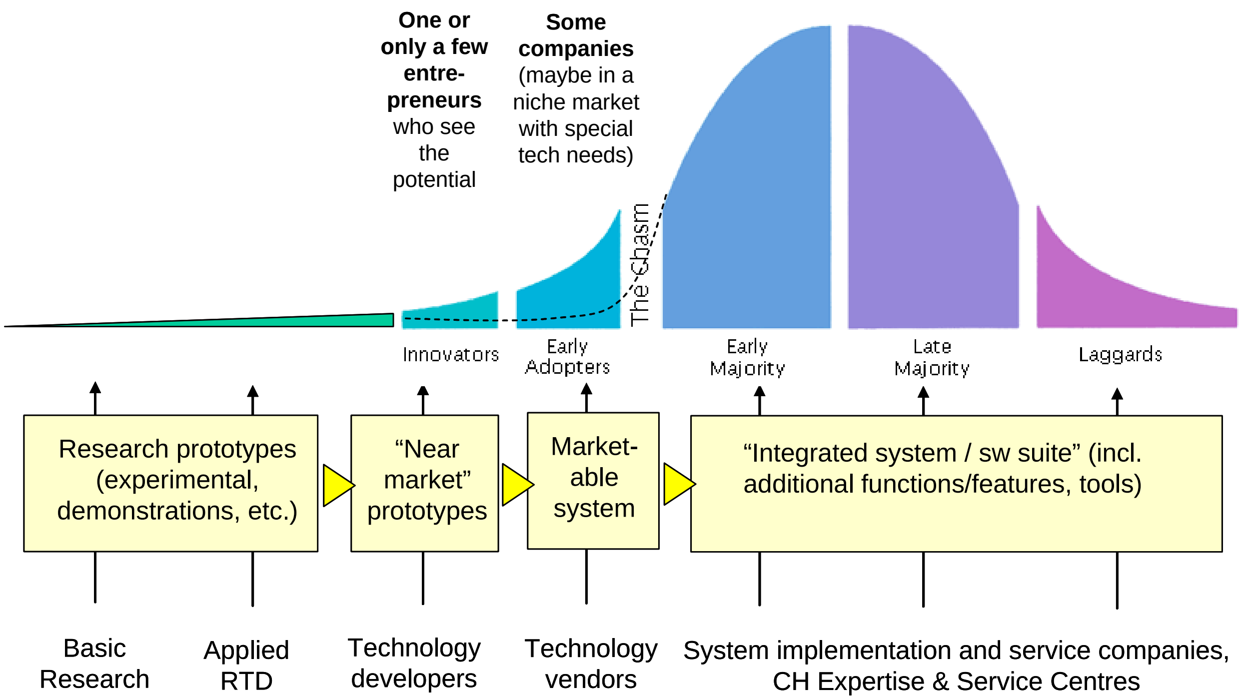


Diagram: Salzburg Research, 2006

The standard process model starts once technological research and development has reached a functioning and tested (prototype) solution, which is adopted by one or more innovative companies in search of a competitive edge. Then, industry solutions appear which usually target larger organisations, and find some early adopters, based on a more stable and scalable solution. Next, competing industry solutions appear which may also target smaller organisations, and are adopted by a much broader group of organisations, the so-called 'early majority'. Then, the mature and well-serviced technical solution will find a large, perhaps industry-wide 'late majority'. Finally, even the most confirmed sceptics will decide to use it.

The process model developed by Geoffrey Moore ([Moore, 1991]) specifically for "high-tech products", such as content management systems or new consumer devices, identifies a "chasm" between on the one hand the first users (innovators and early adopters) of such products, which may still to some degree be immature, and, on the other hand, later customers who will only adopt a mature product.

9.1.2. Closing gaps in the development of mature CH technologies

We understand that there is often a similar "chasm" between, on the one hand, plausible-in-principle solutions or prototypes of CH ICT developed in the framework of projects and, on the other hand, complete, turn-key software offerings – systems and tools that a user community needs and would like to use.

In most technology areas this gap is closed by certain types of technology companies who form an "interface" between RTD, market, and innovation-oriented customers (innovators and early adopters). However, such an "interface" has not so far evolved to a sufficient degree in the field of cultural heritage ICT (see: Non-RTD perspectives: Technology companies).

In the EPOCH Research Agenda activity discussions have started from the basis of asking "users" as represented by cultural heritage professionals via their user requirements, for perceptions of the priorities that they would place on particular developments. However the lack of technological awareness in such groups can mean that their conceptualisation of what might be considered a research issue represents real challenges, but challenges that a computing researcher would regard as operational. At the same time, the technologist's view of a real research topic is perceived by the cultural heritage practitioner as verging on science fiction. From their own perspectives both views may well be right.

Thus, some "blue-skies" research may be undertaken (for example the investigation of some interesting properties of a new material). After a potentially substantial period of research the issues may become more engineering orientated (for example "can the material be manufactured in sufficient quantities, economically?" or "what is its environmental impact?"). As these issues are resolved successfully, the material may become usable in the redesign of particular equipment, with commercial interests bringing the new material to market in innovative and attractive products. Only at this stage will the original blue-skies research be turning into applications with economic return for the original research.

This research cycle may have taken a significant time during which the will to pursue the original line of research needs to be maintained if the initial promise is to be realised. Of course, in many cases initial promise will not be realised because it may be found that the initial concept failed to take into account some important factor and the research demonstrates that this factor is so intractable as to negate the potential benefits of the line of enquiry. Maintaining a decision to invest in particular lines of research is an issue of judgement, based on perceived benefits relative to perceived or actual costs.

Such judgement is often exercised for a combination of political as well as economic reasons, and as with all political decisions the one to progress will normally be made based on widespread support for the potential benefits, that is widespread in terms of those who contribute to the decision to support the work.

9.1.3. Difficulties in areas of interdisciplinary research and development

A complication arises for agendas that are based on interdisciplinary collaboration in that the range of contributing perspectives inevitably reduces the concentration of support. In particular, where an “end-user” element is involved in the decision making processes the perspective of benefits and timescales to address research issues will be different. Thus, the decision to approve a research agenda in terms of developing a new drug or surgical procedure is normally taken by a group with similar professional understanding, weighing different potential developments against each other and prioritising between them.

Decision-making on research in the applications of ICT has only appeared to operate in this way where a perceived commercial return, in terms of product sales, is envisaged. Justification for investment is in terms of capturing a percentage of a potential market, etc. Where research is required to deliver in terms of a social agenda, the picture becomes more complex and the decisions have to be informed by support from the socio-political arena. Securing this support is an integral part of pursuing the research agenda.

The issue of timescales is vitally important here because the research with the highest impact is almost inevitably going to reach maturity when it has had a fundamental impact on the working practices of the very application constituencies whose support is required. Those supporting the research direction therefore have to also become knowledgeable about the implications of the work, and the potential impact on working practices as the research progresses. Indeed, concern over this impact and sensitivity to the implications is likely to be a serious component in determining continued support for the research directions and may well impact on the effective timescales involved. Too rapid a change in itself may lead to resistance to embarking on the direction of travel.

In order to share the definition of a common interdisciplinary research agenda, the perspectives of the contributing disciplines must all be discussed and a common, realistic understanding reached, which is likely to involve compromise. Issues which are likely to be at the top of an application domain’s priorities are also likely to be shorter term considerations than the potential future directions which professionals in the ICT domain might envisage.

The contrast is often classed as “technology push v application pull”, but in truth the gap is wider than that and there tends to be little overlap in the two perspectives. The overlap tends to be in the engineering required to make practical use of the results of research by implementing a set of operational pre-conditions (e.g. agree standards or evaluate/educate business practices) before a genuine take-up can be achieved. In many cases the operationalisation of research results requires extensive research in Stokes’ fourth quadrant. For example:

- systematic investigation, analysis and classification of cultural heritage situations (e.g. design styles)
- agreement on generic standards for representing classes of object in digital systems

Research in this quadrant very rarely tends not to achieve widespread acclaim and recognition, and hence tends rarely to attract the interest of career researchers (academic or otherwise). It is however a vital underpinning to the take-up of the highly visible, acclaimed prototypes of applications supported by basic research which demonstrates principles, but does not achieve integrated coverage of the domain. Populating the field with systematic results is a potential area in which it might be beneficial to consider mobilisation of community resources – in the same way as an open source community may contribute to the development and enhancement of a software system.

9.1.4. Difficulties in standards for interoperability

The issue is probably best highlighted in the area of standardisation for interoperability. In the discussions reported in the previous section, virtually every grouping of cultural heritage professionals recognised the importance of standardisation in the ways in which our knowledge of the past was archived. This will have a fundamental impact on the ability to design systems which can interoperate, for example, bringing resources together from a range of collections in order to respond to a need which an individual collection could not meet.

To many technologists the achievement of agreed technical standards is a tedious and time-consuming exercise which can only be undertaken after the research to demonstrate *potential* interoperability has been completed. Actually engineering a solution may be less interesting intellectually to the technological researcher, but is of fundamental importance to the business processes in the application domain, and to achieving market take-up of the research results.
To the application domain, the achievement of agreeing technical standards is a long-term goal and involves significant research on their part to understand the technical implications of the agreements being proposed. This process may well take several years and is normally an evolution as understanding is reached. The situation where, as is frequently quoted, "I like standards because there are so many to choose from" is a reflection of the continuing evolution of the proposed standards, as understanding of the implications of a particular set of agreements is realised in the application domain.

It is likely that the whole debate around the potential implications of take-up of "Dublin-core" (and its derivatives) and/or CIDOC-CRM as an approach to documenting knowledge about museum collections will become a manifestation of evolving understanding, complicated by existing investments and political willpower(s).

These debates also fuel the decision-making processes for research investment and may fundamentally influence the directions taken in the underpinning technological research and the evolving priorities in research there. For example, the assumption that multi-lingual applications will be based on a common standard for the ontology describing a collection might lead to research in one style of search based on embedded semantics. If the choice of common ontology is different, new constraints and search metrics may well need to be developed, and if the technology domain has to operate with multiple standards concurrently, then a profoundly different approach might be needed.

None of these individual scenarios is yet a solved research area and each would take a different research program to investigate. The priorities for the technologists must depend upon those of the cultural heritage domain and they in turn can only take the decisions based on the advice on implications from the technologists. The process of evolving the agenda must be truly interdisciplinary in order to be maximally effective.

In the first year of EPOCH the consortium tried to ease these communications difficulties by creating a number of showcases intended to enable a shared understanding of current potential applications and the work that would be required to realise them in a business-like context.

9.2. Different CH ICT perspectives

The Common Research Agenda will also need to augment the core perspective on RTD with a view on the requirements, likelihood and time horizon of heritage organisations adopting the future ICT systems and applications that may stem from the ongoing RTD efforts.

Such assessments and respective assessments will be of greater interest to stakeholders from technology companies and the heritage sector, as well as be useful for RTD planners and funding bodies.

Below we describe three perspectives which are important to consider with respect to the further development of market-near prototypes and potential uptake of new applications by cultural heritage institutions.

9.2.1. Technology companies

Technology companies develop, vend, implement and service technical systems and tools. With respect to the maturity life-cycle we distinguish between

(1) companies that are to a certain degree also engaged in technological research & development [RTD] activities, and

(2) companies that concentrate on marketing, implementing and servicing stable and proven technical solutions.

Both play an important role in the maturity life-cycle.

R&D driven companies

The first type of companies forms an “interface” between RTD, market, and innovation-oriented customers, i.e. the “innovators” and “early adopters” in the diffusion process of new technologies.

Such companies develop prototypic systems and tools into marketable solutions. In any technology field, they are rare examples, particularly if there are no large enterprises that would license or buy and market the solution.

This is the case in the field of cultural heritage ICT, where most of the companies are SMEs, and only little domain specific specialisation has taken place so far (e.g. in the area of collection management systems).

Most companies that engage in RTD activities and, among other target markets, deal with cultural heritage ICT, are spring-offs of university-based research centres. They build on results of some projects funded under various national and European programmes, and most often do not want to lose their foothold in the research community.

A typical example here may be EPOCH partner Imagination (Austria) which offers Virtual Reality services that include consulting, design, production of online interactive 3D applications and VR/AR installations for events, shows, exhibitions or permanent installations. The company is a spin-off of the Institute of Computer Graphics and Algorithms of the Technical University of Vienna. The institute participated in the long-term Austrian Joint Research Program on “Theory and Applications of Digital Image Processing and Pattern Recognition” (1994-2000), funded by the Fund for the Promotion of Scientific Research. After Imagination was formed, it participated in FP5-IST projects such as 3D-MURALE (11/2000-10/2003) and has been one of the industry partners in the K-plus Competency Centre “Virtual Reality and Visualisation (VRVis)”, funded by the Federal Ministry of Transport, Innovation and Technology.

Market driven companies

The second type of companies concentrates on customers who are not in a position or willing to take any risk. Their role includes representing “the face of technology” as mediated to such customers, who - as in any other domain - also in the cultural heritage sector form the large majority of organisations.

The unfavourable business situation of the companies that target customers from the cultural heritage sector is described in the previous Report on the Common Research Agenda ([EPOCH: D2.9, 2006]).

They have several major hurdles to take that include the different “business culture” of cultural heritage institutions and professions, small IT-budgets, and lack of technical stuff and background.

In practice, this can mean that CH customers sometimes use the companies for free consultancy, tenders may be ill defined, projects have long lead times and decision processes are not transparent. Expensive tendering exercises can lead to the cancellation of an initiative without appointment of a supplier, since it can often show an unrealistic perspective on the amount of investment required.

Consequently, most of the technology companies do not consider the CH domain as their core business. The degree of specialisation is rather low, which leads to the criticised situation that specific needs of the domain are often not met.

Results from a survey conducted in the framework of the EPOCH Sector Watch activity ([EPOCH: D.2.1], section 2.2.2) confirm and detail this overall unfavourable situation. In particular it should be noted that many technology developers concentrate on prototypes in the sense that they are often only applied in one cultural heritage site or museum as a test case. Hence, the work on these prototypes does not necessarily lead to a marketable product that could be easily integrated by different cultural heritage institutions.

9.2.2. Cultural heritage institutions

When assessing the feasibility of cultural heritage institutions making use of advanced information and communication technology (ICT) their capacity in terms of budget, staff, collections and users must be considered.

A study carried out by EPOCH partner Salzburg Research provides estimates of this capacity for small, medium-size and large institutions. [Geser, 2004] The study collected and analysed data from various surveys and other sources. The results are summarised in the following table:

	Small	Medium	Large
Annual operational budget in €	< 100,000	100,000 – 1 million	> 1 million
Staff in full-time equivalents (FTE); professional, support, volunteers not included	< 5 FTE	5-10 FTE	> 10 FTE
Number of collection objects	< 10,000	10,000-100,000	> 100,000
Number of annual visitors: for museums	< 7,000	7,000-30,000	> 30,000

Note: The focus of the study was to provide a better understanding of what distinguishes small from larger size institutions quantitatively. Therefore, the table does not include a category 'very large' or 'major' institutions, which may have an annual operation budget of over € 10 million.

The study points out that most of the smaller and even many of the medium-size institutions, which together make up more than 90 per cent of all organisations, will not find it easy to cover the total cost of ownership (TCO) for certain more advanced ICT applications beyond, for example, a simple web site or a collection management system.

The most pressing factor that hampers heritage institutions in their efforts to leverage their IT environment is the lack of staff. A typical small institution will have fewer than five full-time equivalents, with only a fraction of them being professionals concerned with the institution's core business (e.g. curators, librarians, archivists, pedagogues).

Furthermore, smaller institutions' efforts in following up new technology ventures are limited by lack of financial leeway. A typical small institution will work on an operational budget of no more that €100,000 while a medium-sized institution may have up to €1 million at its disposal.

Needless to say, these budgets leave scarcely any room to finance ICT projects out of the operational financial resources. For example, a survey conducted by Statistik Austria (2004) provides information about the ICT equipment of 389 museums and other institutions that exhibit cultural heritage objects. Of particular interest to EPOCH will be the information on available computers in, and websites of, historical and archaeological museums. Of the 49 museums in this category 11 did not have a computer. 37 museums had one or more computers, which were used for administrative purposes (26 museums), internet access (25), and collection management (19). 23 museums also had computers in place for visitor information. 44 of the 49 museums had a web presence through their own website (32) or/and on another website (13).

Institutions that are interested in developing and realising more advanced technology projects will need additional funding. Yet, a common problem for small institutions is that, while the limited number of professional staff available may be able to ensure that the institution provides its core services, there will be little time to track down the necessary funds that would allow them to finance any ICT venture. And if they identify a suitable funding opportunity, they will find it difficult to prepare an application due to a lack of expertise in drafting a possibly successful bid. (cf. the results of the IMRI studies on the effects of the "bidding culture" on local institutions in the UK, [IMRI, 2001] and subsequent reports).

Furthermore, experience from many initiatives shows that projects carry the risk of distracting institutions from core business and imposing activities that prove to be unsustainable after the funding period.

Critics also point out that the majority of such projects favour financing the technological infrastructure, that is, the hardware and software equipment, over the development of the 'wetware', i.e. the technical skills of the programmers, operators and system administrators. The cost of ownership for the technological infrastructure is usually underestimated or not even considered.

Given the institutions' 'trilemma' of lack of funds, lack of human resources, lack of technical skills, there is little likelihood of small to medium-size institutions being able to participate in research and technological development projects that develop new prototype applications and systems. Even the larger institutions may have difficulty engaging with projects to which they are required to bring their cultural and scientific heritage expertise and knowledge.

Some developments and issues in CH stakeholder needs

The following is a brief summary of a observation regarding developments and issues in stakeholder needs by Malika Hamza from the Ename Center for Public Archaeology and Heritage in the Research Agenda workshop at VAST 2006 conference:

- On the local/regional level there exist severe weaknesses with respect to strategic planning and funding of CH initiatives;
- Some progress has been made with respect to equipping local and regional museums with computers and Internet access;
- However, local and regional content digitisation initiatives are hampered by limited financial resources and a lack of relevant services;
- Much effort is invested in regional portals that provide information for cultural tourism purposes;
- More advanced virtual exhibitions are rare and institutional information is often only partial, not updated regularly, and most often not multilingual;
- However, priority must anyhow be given to the training needs of smaller and medium size institutions with respect to capacity and skills for basic technologies (e.g. collection management systems);
- Many institutions could benefit from services of Centres of Expertise for which, however, business models or/and funding mechanisms must be developed.

It should be noted that smaller and many medium size CH institutions are not be in a position to experiment with novel ICT-based approaches. They require tried and tested applications that show to deliver a clear cost/benefit ratio.

For those who may be involved in the further development of useful applications the following factors will be of particular importance for successful projects:

- Application development teams must show a good understanding of the practical demands of CH institutions, hence, practitioners must be involved in the set-up of the development project;
- The target must be applications that have a likelihood of being adopted by many institutions beyond the initial project partners (particularly, this requires that the applications are robust and affordable);
- Priority will need to be given to a “bottom up” approach, starting with data collection and processing, lowering the cost of digitisation, improving accessibility and usability of digital resources;
- This is closely related to the need of leveraging the inventory and collection management systems and external resource discovery, which can not be taken for granted for many CH institutions;
- Even for “bottom up” approaches considerable training of staff is required, which needs to involve experienced CH peers and be covered by extra project money (e.g. from regional funds).

Large vs. small museums: The case of Italian museums

According to the official statistics of the Italian Ministry of Cultural Properties and Activities ([Italy, 2005], section: Presences in State museums and archaeological sites) the following interesting facts can be noted:

- 9 museums (out of 394) make 50% of the presences,
- 20% of the museums make 90% of the presences,
- one third of all museums have less than 20 visitors per day.

The first two findings are illustrated by the following diagram

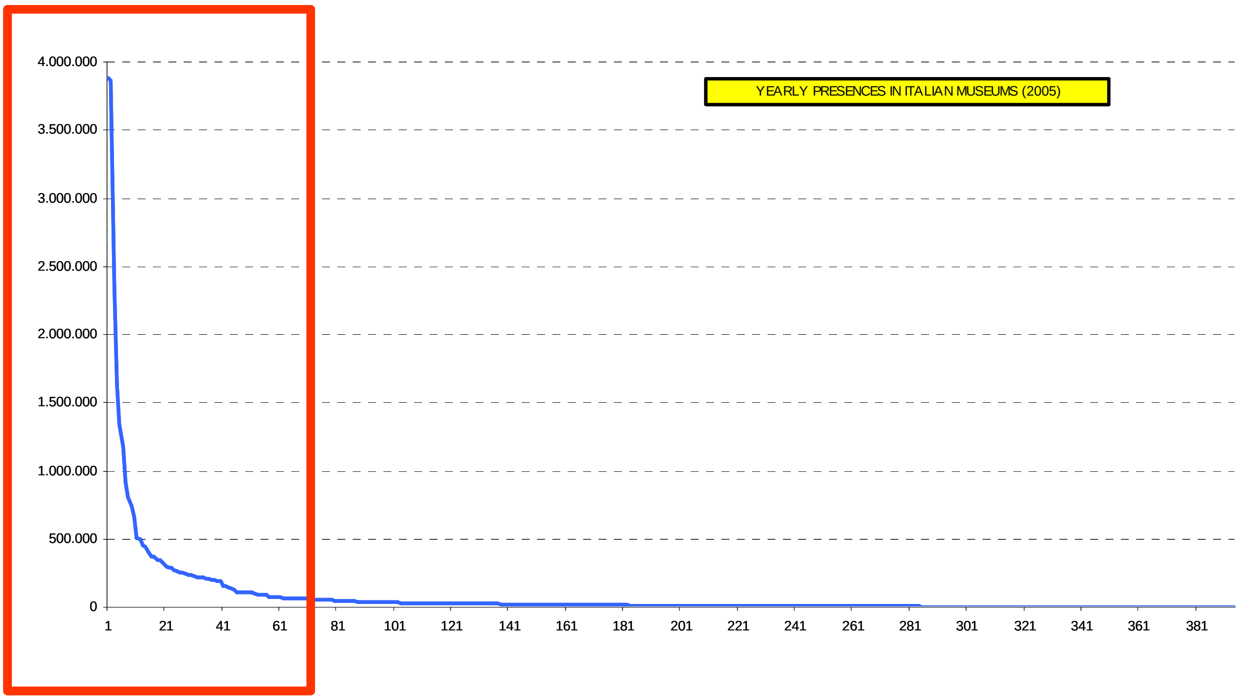


Diagram: Distribution of presences in Italian State museums, 2005 (diagram by F. Niccolucci, PIN, Italy) (data source [Italy, 2005])

These visitor numbers clearly indicates a considerable imbalance of the distribution of public and cultural tourism interest in cultural heritage. Renowned larger institutions that are located in tourism centres and intensively market holdings and exhibitions receive most of the interest while smaller museums (with more or less valuable collections) are neglected.

While differences in location, size, collections and exhibition activity must be noted, this also demonstrates a clear tendency of "cultural consumerism". The statistics clearly show a visit pattern that is different from the one that is still widely expected by many cultural heritage institutions and funding bodies.

At least the observable imbalance between "consumer behaviour" versus "cultural behaviour" (of educated visitors) should be noted, as this will considerable influence the presentation, communication and interaction with users of CH museums and sites.

9.2.3. Cultural heritage expertise & service centres

Given the severe barriers of most cultural heritage institutions, there is a need to create structures that prevent them becoming blind spots in the rapidly developing digital environment. There is enough evidence that this environment of next generation systems and tools evolve much faster than these organisations and smaller cultural networks can adopt and employ (cf. [EC, 2002]; [Geser, 2004]; [PULMAN, 2003]).

In fact, for smaller and also medium-size institutions the benefits of most current and future technologies will need to be realised within national and larger regional initiatives.

In such initiatives, a leading role will require to be played by new forms of cultural heritage expertise & service centres. There will, over the coming years, be an increasing demand for supportive digital services centres and ICT training programmes for technical and non-technical staff on how to handle new technologies.

Such funded mechanisms should enable smaller institutions to keep the costs and risks of digital heritage resources and services manageable, while not being excluded from new technological developments.

The establishment of CH expertise & service centres could also lay the groundwork for the required much stronger linkage between research & technological development and CH experts and practitioners, which should be based on true interdisciplinary efforts.

Particularly if such centres are established in conjunction with research centres that specialise in cultural heritage ICT, this may provide for a steady stream of knowledge between researchers and technologists and experts and practitioners from – and clients of – cultural and scientific heritage organisations.

In a much more effective way, curators, arts & humanities scholars, educational programme managers and experts from cultural hotspots, such as historic city centres or larger heritage sites, could be involved in the development of prototypes of new applications, and feedback from professional users and visitors of sites, monuments and museums be collected.

Model examples of cultural heritage expertise & service centres are the Dutch Digital Heritage Association (Vereniging DEN) which supports about 60 member institutions (cf. [Van Kasteren, 2003]), or the smaller EPOCH partner Interactive Institute.

9.3. Discussion of stakeholder interest in certain CH ICTs

In the further elaboration and application of a CH ICT maturity life-cycle model it will be important to consider the level and direction of interest the different actors take in certain technologies.

Because this interest will determine

- which areas of RTD will receive the most attention by the research community in cultural heritage ICT (e.g. when forming new research groups or preparing research proposals),
- which systems and applications technology companies and digital CH expertise centres will propose to the institutions, and
- which systems and applications the different institutions in the field of tangible heritage will consider adopting.

General dynamics of stakeholders' interest

The following diagram provides a hypothetical assessment of the current interest of the research communities, technology companies, digital CH expertise centres and institutions in the technological areas covered by EPOCH's Common Infrastructure activity.

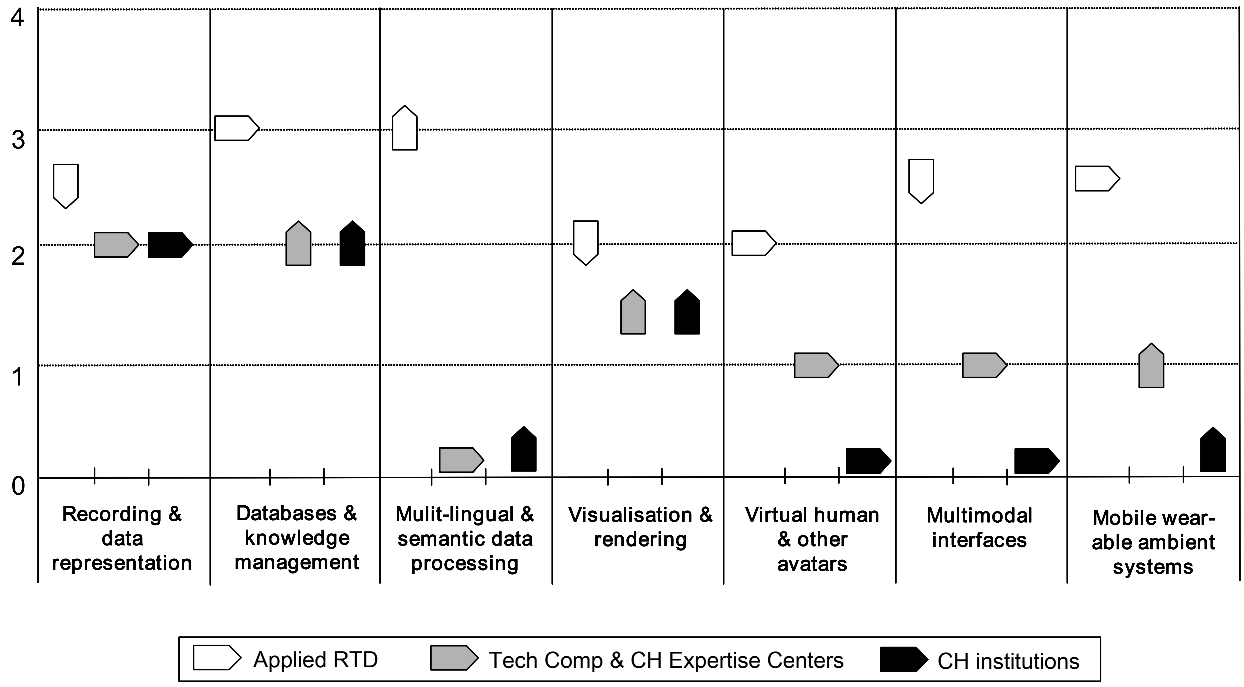


Diagram: Salzburg Research, 2006

The purpose of this diagram is to illustrate some general dynamics that the Research Agenda will need to take into account. Such general dynamics are that the different actors may show a stable, accelerating or declining interest in certain technologies.

For the research communities the interest may increase because of some new research priorities in major funding programmes that invite targeted research in the area. For example, this may be the case in the area of semantic data processing, where “intelligent heritage” has been specifically addressed in the European Union’s FP5 and FP6 as well as in some national funding programmes.

Alternatively, a decline in interest may be due to the observation of many researchers that a technology area is mature and provides only little further research potential. The area could even be already dominated and driven by commercial players.

One possible example for this dynamic is the area of “visualisation & rendering”. As shown in [EPOCH: D3.3.2, 2005], (pp. 32-38), in this area there exist high-end commercial tools for example for 3D CAD (Pro/Engineer, SolidWorks, UGS SolidEdge) and modelling (3D Studio Max and Maya).

There are also solid open source 3D tools such as OpenCascade and Blender, and the commercial players offer academic and educational institutions excellent conditions. Further, there is an abundance of special tools e.g. for terrain interpolation and landscape visualization.

With respect to the other actors we understand that there is on the one hand a rather close connection between which systems and applications technology companies and digital CH expertise centres will propose to the institutions, and what the institutions will actually consider to adopt.

On the other hand, there may be some technology areas where companies and expertise centres would like to see a stronger interest from the side of the institutions (e.g. with respect to avatars, multimodal interfaces, and mobile wearable ambient systems).

The reason behind this is that CH technology companies and centres who also engage in applied research & development would like to be able to point to reference implementations of applications. They need reference customers from the CH sector in order to win over further potentially interested institutions. Only thereby can they capitalise on the know-how acquired in the show case projects, in which they often invest effort that exceeds the financial gains.

However, there may also be cases where (larger) heritage institutions and networks start to take an interest in a technology area, but most technology SMEs and regional CH competency centres will not command the required expertise to develop and implement a state-of-the-art solution (e.g. in the area of semantic data processing).

Institutional differences: Monuments, archaeological sites, and museums

Furthermore, it is important to take into account possible differences between the institutions in the field of tangible heritage. In the diagram in the previous section these institutions are grouped together and an overall assessment of the level and direction of interest in the certain technological areas is given.

However, when considering the differences between the institutions the following diagram may present a more appropriate, yet again, hypothetical, picture:

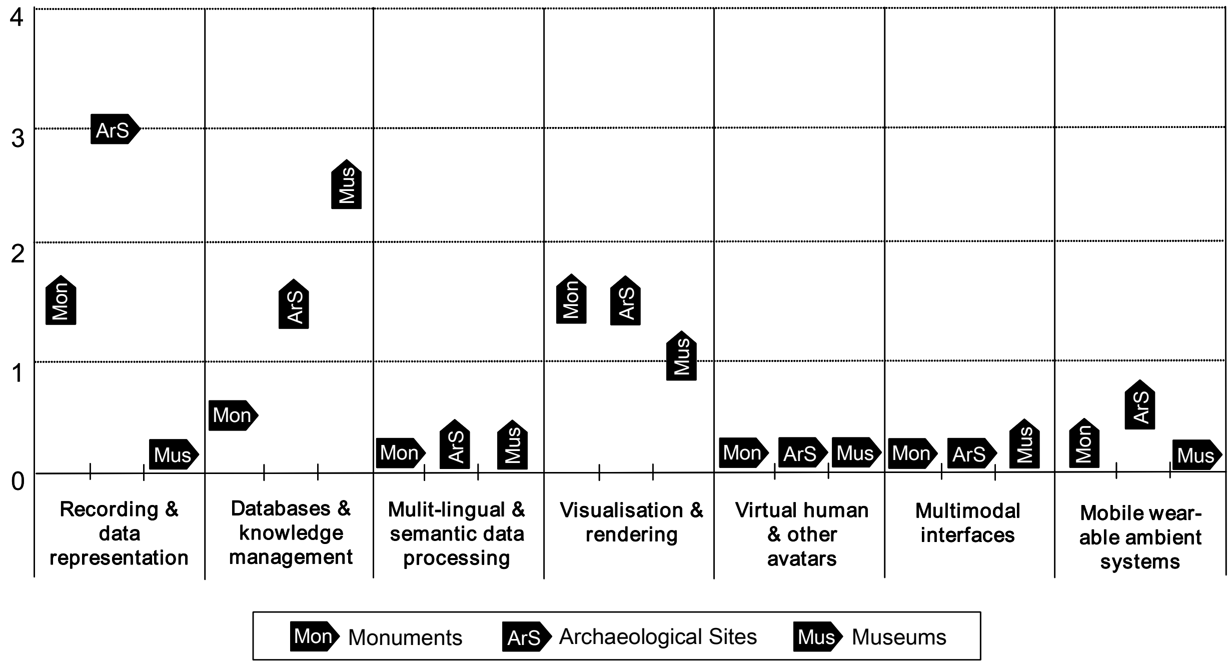


Diagram: Salzburg Research, 2006

Here of particular interest are the technology areas for which an overall assessment would level out the considerable differences in interest. Most notable are the first two areas:

In the area of "recording & data representation", besides the continuing high interest of the archaeology community in this area, we observe a considerable rise in interest of monument managers in digital monitoring technologies able to detect changes in physical conditions (e.g. humidity, corrosion, decay of material, etc.). (cf. the results of the EPOCH expert workshop at ICCROM Headquarters in Rome, 6-7 March 2006 [EPOCH, CHEDI, 2006])

On the other hand, the interest of museums in recording & data representation technology is low because most museums do not record, process and represent data sets such as archaeological field recordings, but digitise objects (e.g. 3D scans of museum objects such as statues) more as a means of presentation than as documentation of the sources and, hence, are much more interested in the technology area of "visualisation & rendering". The impetus given by initiatives such as the European Digital Library which emphasised a broader vision of collection digitisation may impact on the museum sectors interest in this area.

Finally, in the area of "databases & knowledge management" we would expect a rather low interest of monuments, at least in comparison to archaeological excavation areas and museums. The latter deal with large numbers of different finds or objects as well as today striving to provide and integrate access to the knowledge of different subject experts and results of ongoing scholarly research. A case in point here may be the considerable rise in interest in knowledge organisation systems such as thesauri and ontologies in these domains.

Clustering of the interests in technological applications

The discussion in the above sections of course simplifies many issues pertaining to the aim of developing a common Research Agenda.

Firstly, when considering the effects of the level and direction of interest the various stakeholders take in certain technologies, also the interrelations between the technologies must be taken into account.

Secondly, for particular (prototypic) research or market-near systems and applications the technologies and interrelations between them must be detailed.

Thirdly, for the (prototypic) systems and applications under discussion the perspectives and assessments of the different communities in cultural heritage ICT must be collected, analysed and clustered.

Hence, in the development of the Research Agenda the identification and prioritisation of RTD challenges of different ambition must start from a set of particular combinations of technologies as required to perform certain CH tasks.

The degree of ambition may range from a considerable enhancement of the performance up to the realisation of new technological platforms and attendant applications which allow for currently unachievable capabilities in cultural heritage research, management, and mediation (“grand challenges” [UKCRC], [Arnold,2006]).

Carrying out this work, and involving in it the stakeholders from EPOCH and other cultural heritage ICT communities, will be one of the core tasks as the agenda evolves. A major goal for this will be to have the results prepared and available for consultation by the stakeholders in the 7th Framework Programme as this becomes operational.

10. Adoption of ICT in the presentation and exploitation of Cultural Heritage

In the following sections we provide an overview of the current situation regarding the adoption of ICT applications in the Cultural Heritage sector. The presentation concentrates on perceived limitations that inhibit a wider uptake of ICT in the presentation and exploitation of CH.

The assessment of the different applications is based on a presentation of Vassilios Vlahakis from EPOCH partner INTRACOM S.A. Telecom Solutions, Greece, in the Research Agenda workshop at VAST 2006 conference.

INTRACOM is a commercial partner in EPOCH and has participated in several other RTD projects in the EU Framework Programmes and is involved in the development of practical solutions for the Cultural Heritage sector (in particular, guiding systems such as the ones developed for Olympia, Pompeii and various museum systems).

10.1.1. Widely used applications and devices

Applications & devices	**Short comments:**	**Current limitations**
CD-ROMs, DVDs:	Widely available either with educational, tourist, or scientific emphasis	None
Audio Guides:	Widely used in major (and smaller) museums and sites – analogue or digital – manual or automatic	None.
Interactive installations (open-air and museum-based):	Mainly based on touch screens or large projection screens. Their use is normally included in the museum ticket price	True **immersive installations** are not in wide use. Interactive installations in museums may sometimes feature **presence or motion detection.**
(Distributed) Digital Cultural Archives:	Proprietary formats or standardized, often not interoperable	Regarding the **database design** various schemas and standards are used to enable interoperability. However, as various “industry” standards have been developed and promoted by the major international museums and CH institutions the widespread adoption of a single standard or the interoperability of the existing ones has still not been achieved.
Web portals (info provision, event scheduling, virtual museums, etc.):	Individual institution efforts or combined efforts at national/international level	None. Some international standardisation of **portal design** could be beneficial.
Web forums, Weblogs:	More unstructured and uncontrolled than other media. Free of charge, requiring a simple user registration	None. **Social Software** applications are rather simple to implement and can allow for a more personal communication of curators with the museum audience (e.g. through a Weblog) or enhance the work of study teams (e.g. Wikis).
On-line shops (cultural artefacts, souvenirs, books, etc.):	Mainly owned by museums etc. Also by specialized companies	None.
Mobile telephony services for cultural information	Offered by mobile telephony operators as a basic guide and sometimes linked to on-line ticket booking.	Some examples include **location sensitivity**, though this is still in prototyping-evaluation phase

10.1.2. Applications showing considerable limitations

Applications & devices	Short comments:	Current limitations
Mobile MM Guides ("Commercial" on-site use of PDAs, e-guidebooks, etc.):	Featuring a wide variety of functionalities and presentation media. Include manual or automatic operation, location and/or context sensitive, content and/or UI personalization. Range from simple MM presentations to use of VR and AR.	The commercial use of mobile guiding systems is inhibited by the initial investment required (though in the medium to long-term operation they can provide significant revenues to a site / museum through the attraction of additional visitors, potential sale of souvenirs, etc.) The various installations (mainly experimental) provide a test bed for further development. Among the main technological challenges are **- location awareness** where a number of technologies are competing (infrared, WiFi, RFID tags, Bluetooth, ultrasound), **- context awareness** (including the understanding and exploitation of user behaviour and needs), - on-line access to **distributed archives** (problems arise from the non-unified standardization of these archives – CIDOC, Dublin Core, CDWA, are some of the standards used), **- Digital Rights Management solutions**, and - **Physical constraints of the devices, in particular, battery power** limiting the devices autonomy In combination the issues mentioned above also mean limitations regarding the **intuitive operation** and **personalization of the cultural experience**
Wireless MM Access and VoIP	Wireless MM access can be achieved by any of the existing wireless technologies. VoIP is used in culture as part of mobile guides in order to group several devices together and support group tours and educational scenarios.	Typical problems include the **tedious connection process** (common to mobile phone services), the **slow data transfer rates** as a result of high usage, and the still considerable costs involved. Free services exist though still at limited availability, mainly through WiFi and Bluetooth. VoIP is still limited to systems under evaluation while it has found a market outside

		culture for cheap telephony.	
Location-Based Systems:	GPS, infrared, WiFi, RFID tags, mobile telephony cells, Bluetooth, video tracking, and ultrasound are used in decreasing order of frequency. The most frequently used are GPS for outdoor spaces, and infrared for indoor spaces.	All current solutions suffer **from limitations of accuracy, ease of use, necessitate line of site (for infrared and video tracking), and cost.** The majority of mobile applications **still rely on the user identifying his location** of a digital plan or taping a code identifying an exhibit or location. More recently mobile phone applications enable their users to take photographs of their surroundings or exhibit of interest and send it to a central server for location detection by image matching to a visual archive.	
VR Installations (Caves, 3D theatres, interactive panoramic screen installations):	VR systems are used as part of on-site museum installations, or delivered through web sites, and more recently through mobile guides VR installations are mainly to be found in larger museums and commercial-entertainment establishments. Usually at extra cost to the normal entrance fee.	The use of VR systems is still rather limited due to the **high cost involved in creating the 3D models** of objects, building, and landscapes. A simpler and cheaper approach to VR is the use of panoramic view, possibly annotated. They do not provide all the functionalities of VR but they give the opportunity to explore museums etc. without having to visit them and help disseminate culture.	
Digital Interactive TV and Triple-Play Services	Digital Interactive TV and more recently Triple-Play Services are beginning to find their way to the consumer market. Their more common CH use is the transmission of documentaries mainly through specialized channels. However, the potential of these technologies is huge as consumers can browse and request cultural information from the comfort of their home using nothing more than their television.	Currently these technologies show a **limited market penetration and only a limited range of services** is offered, especially regarding cultural content. The high cost of setting the necessary infrastructure can be more easily covered since it will not be dedicated to cultural applications only.	Augmented Rea

10.1.3. Challenges in the further development and wider adoption of applications

While the overviews above mainly address technological limitations, in this section we want to address some general challenges in the further development and wider adoption of applications in the Cultural Heritage sector.

Better integration of efforts and synergies between all stakeholders:
A more decisive move from individual research to integrated research is required. In this integration care should be taken that available expertise in commercial companies and more advanced CH institutions (e.g. from best practice showcase implementations) is integrated.

Standardisation
As a core requirement of bringing the benefits of ICT to the Cultural Heritage sector more efforts must be dedicated to standardization. Besides a higher cost-effectiveness of single application development this is of particular importance for the overall infrastructure and application integration (e.g. for pervasive and ubiquitous computing in CH environments).

Process Automation
The success of a wider implementation of CH ICT hinges on a much stronger automation of many processes from data acquisition and metadata creation to support of content authoring (e.g. for multimedia or VR application) and database & application management.

Personalization and contextual sensitivity
The capability for providing enhanced personalization and taking into account user contexts (e.g. location, user behaviours, needs and interests, etc.) is currently rather limited. Various technical approaches have been explored over the last years each showing some shortcomings (e.g. with respect to employability, accuracy, ease of use). However, the most critical part to address in the future will be the modelling of the user experience which will determine whether applications will find a wider adoption in the sector.

Enabling storytelling in and beyond the museum
Michael Danks from Windfall Digital in the EPOCH Research Agenda workshop in Nicosia emphasised the importance of supporting CH institutions, particularly smaller local and regional ones, in telling stories about particular objects they hold. This would help in better contextualising their collections and make them more interesting for their on-site and online visitors.

Heritage story telling could also benefit much from leaving the boundaries of the museum and expand into urban places such as squares, railway and metro stations, etc. (see for example the projects of Local Projects, a museum and exhibit design firm, http://www.localprojects.net).

Better knowledge about requirements of institutions and users is required

It is important to gain much more knowledge about the current and future needs of CH visitors, CH management & staff, researchers and other professionals. This can be best achieved through ensuring that users participate in projects from the start (involving users only in some final testing of applications means that their needs and views have not been taken into account properly throughout a project).

In particular, the knowledge of institutions and their staff is important in overcoming many practical issues in the development of applications. Moreover, their commitment can become the driving force for further developments and ensure the financial viability of all endeavours.

It would also be highly beneficial to have the opportunity to work with Centres of Excellence in ICT-based Cultural Heritage and use them as test-beds for new technology development and evaluation.

Business models & financial issues

More knowledge is required with respect to various financial aspects of technology adoption and business models for CH ICT applications in general. This includes issues from initial development costs to the total cost of ownership (and sustainability) of certain applications in a museum or CH site.

Though benefits of technologies (visitor attraction, improved management of sites, etc.) may be easily communicated, the hard part is to get the cost/benefit equation right. Applications that are in “prototype stage” will often need a considerable commitment of heritage institutions and extra funding from governmental agencies that are interested in promoting advanced ICT in the CH sector.

Fragmentation in CH administration

In many European countries there exists a high fragmentation of responsibilities regarding CH administration and funding. Stakeholders in Cultural Heritage and CH ICT should wherever possible point out what this means regarding the development time required for new solutions that are required for responding to the change in demands on Cultural Heritage preservation, management and mediation.

11. Conclusions - the Role of the European Union

In this research agenda we have explored a range of topics related to the inter-disciplinary research required to realise the potential benefits of ICTs in support of the many professional roles in Cultural Heritage. In this final section we conclude with some observations on the justifications for European Union-led support of the progression of this agenda and relate this to other proposed developments.

In the introduction we included a section on the nature of different types of research and the roles that these types play in the creation of an ICT-enabled future Cultural Heritage sector. We also highlighted the prevailing attitude amongst professional researchers to the differing types of research and concluded that the current research environment is often philosophically at odds with the concept of use-inspired basic research and hence can under-value the careers of those who pursue it. There is no doubt that this position impacts at every level, from the attitudes of national funding agencies to the aspirations of individual researchers.

In some areas of interdisciplinary research these obstacles to all but Bohr's Quadrant in Stokes' classification are offset by the driving imperative of commercial advantage. An obvious example of this would be the Genome Project, where the initial basic science of understanding DNA has been complemented by huge efforts in developing the techniques of mapping the gene sequence and in carrying out the mapping by repeatedly and systematically applying known techniques. The social advantages in terms of disease understanding and control have all along been complemented by obvious potential for commercial advantages in medicine, insurance, food production, and other commercial activities with very large commercial interests.

As noted in the section on socio-economic impact of the cultural heritage sector, the advantages of success are perhaps not as obvious in the Cultural Heritage sector. There is no doubt at all that there are significant socio-economic benefits to be had from a healthy Cultural Heritage sector. What is perhaps less apparent is the direct connection between investment and the potential rewards along with inherent difficulties in quantifying the benefits of raising the "quality of life".

The connection between investment and return manifests itself in that although investment in Cultural Heritage may add disproportionately to its value, the beneficiaries are not necessarily the same as, or even focused on, the organisations that would invest. This is clearly a matter of concern, but equally if the social good can be defined and recognised, with careful choice of prudent investment, then this is an obvious area for government investment, with excellent prospects of payback to national economies and to perceived national quality of life. This would be a clear argument for national governmental investment, but perhaps not as clear for support at the level of the European Union. There are however clear reasons for the European Union also having an appropriate role. These are:

- Cultural Heritage within the European Union does not recognise national boundaries. The links between the heritage of different states of the Union span migrations of ethnic groups, changing political boundaries, trade links, developments and spread of technologies, cultural influences and the spread of design styles, along with the inter-cultural influences and freedom of movement within the EU of today. It is clear that national investment in techniques appropriate to particular elements of heritage will have significance well beyond national boundaries. The fact is that benefits of such investment are unlikely (for example) to include the development of a purely commercial new industry sector which an individual nation might exploit through exporting skills. It is therefore appropriate as a European investment.
- The Cultural Heritage Sector and the organisations that support and use it have a less commercial, but socially valuable mission. Many of the enterprises involved (including many of the SMEs) would be better classed as “social enterprises” rather than as “profit-motivated”, yet the sector as a whole is of very significant economic importance to Europe. If the sector is to grow healthily then investment is needed in the support infrastructure for these “social enterprises”. Organisational development, technological infrastructure, access to specialised expertise, shared best practices and market intelligence are just some the aspects that would benefit from a systematic, Europe-wide, support infrastructure.
- Individual National Research initiatives seem to accept the “purist” view that the most valued research should be unbridled by limitations imposed by consideration of usefulness. In part this appears to arise from the perceived need to have national research capability evaluated as at the forefront of international levels – measured using the same underlying philosophy of what counts. In fact this tends to devalue the interdisciplinarity that many claim to be rising in priority. The framework programs of the EU have in general been rather better at valuing use-inspired basic research, possibly because European Union programs are less about national advantage than about mutual benefit.
- One of the major barriers to integration of European digital cultural assets arises from the lack of widely used standards and relating to this the lack of sufficient appreciation of the implications of cultural diversity for ICT systems. These needs are reflected in the current lack of multi-lingual and multi-cultural thesauri, taxonomies and ontologies and in the multiplicity of national documentation standards in use. A European led initiative is much better placed to ensure the definition and adoption of inter-national standards (e.g. CIDOC-CRM)
- Individual member states tend to have national self-images which are less culturally diverse than the European Union taken as a whole. Promotion of a culturally diverse, but

multi-culturally aware society is a healthy objective for the Commission, but many of the more challenging research topics relate to the difficulty of achieving ICTs capable of personalised multi-cultural responses to queries. Research in this area must be suitable for European support, and is actually considerably longer term.

- A rather more obvious and immediate reason for the topics raised here to be supported at European level is that only continued investment in research in these fields can enable organisations to deliver on the vision of the European Digital Library, both in the area of mass digitisation of the full spectrum of cultural artefacts and their metadata and in providing appropriate access and exploitation of digital cultural assets.

In the next section we outline some relevant features of the European Digital Library initiative and describe the relationship between the research topics here and the successful pursuit of the objectives of the current initiative.

11.1. European Digital Library initiative

Most EU Member States have an Information Society strategy that covers at least to some degree the domain of Cultural Heritage. Often this has initially been stimulated by the eEurope initiative. However, the de facto level of funding often is felt to be too limited for achieving a wider implementation and use of available CH ICT.

Furthermore, cross-domain strategies – involving different Cultural Heritage domains as well as Cultural Heritage and other societal sectors (e.g. e-government, public sector information, creative industries, and others) – show a rather slow development (partly due to a high fragmentation of responsibilities among Ministries and regional agencies).

Some progress has been made with respect to (mass) digitisation of cultural heritage holdings on the national level, particularly regarding special collections of libraries and archives. On the European level the supportive role of the MINERVA and MINERVA-Plus projects in promoting commitment for digital Cultural Heritage and achieving a higher degree of coordination must be acknowledged. Though digitisation today is more widely seen as a Cultural Heritage priority, this is often limited to larger national institutions which receive most of the funding and can afford a digitisation department, or to the establishment of specific centres of expertise in the digitisation of classes of artefacts (most commonly textual objects).

Consequently, also the recent European Digital Library initiative will mainly build on available and newly digitised resources of those digitisation centres. Regional and local institutions may again be neglected. For example, the proposal of establishing a European Digital Library in April 2005 by the Heads of State and of Government of France, Germany, Italy, Hungary, Poland and Spain was followed by commitments of 19 national libraries of EU Member States. By 2008 two million books, films, photographs, manuscripts, and other cultural works are expected to be accessible through the EDL, and by 2010 the volume is planned to have grown to at least six million. (cf. Europa.eu 2006; Commission of the European Communities 2006)

In a recent speech, Horst Forster (Director Content, Directorate-General Information Society & Media of the European Commission) emphasised that the approach taken in the development of the European Digital Library will guarantee that the digital objects made accessible will "consist not only of objects from national libraries but from all kinds of cultural organisations at all geographic levels. After all Europe's heritage is not confined to a few privileged large organisations". (Forster 2006)

EPOCH observes that the European Digital Library initiative at present does not fully take into account representations of artefacts held by monuments, archaeological sites and related museums. The extensive section above on Data Capture highlights the many aspects of digitisation, and especially viable mass digitisation, that are currently unsolved research problems. Mass digitisation of a rich range of cultural heritage artefacts to produce sufficiently rich digital cultural asset repositories, will only be a viable proposition if digitisation tools.

The expertises of centres of excellence in mass digitisation in Europe are mainly drawn from the requirements of the National Libraries to engage in mass digitisation of the textual objects and in the museums world they are focused on reproducing legacy catalogues of museum holdings in digital form. The current gaps in tools for digitisation of a richer range of cultural artefacts and for the richer exploitation of the digital cultural assets are reflected in the extensive sections in this research agenda targeted at the broad interpretation of "data capture". This interpretation is an essential pre-requisite to achieving the vision of the European Digital Library.

In addition, as noted by Franco Niccolucci (PIN, Italy) at the EPOCH research agenda workshop, a different emphasis in the goals of digital libraries and museums should be taken into account. The main point was summarised in his formula: a Virtual Museum = Digital Library + Communication, which emphasises the mission of museums to interpret and pro-actively communicate cultural heritage rather than mainly provide access to digital documents, and indeed the widespread (but not uniform) mission of museums to "entertain".

More specifically, much of EPOCH's work concentrates on leveraging the use of 3D computer graphics in the presentation and interpretation of monuments, archaeological sites and museums. This interpretation requires a variety of tools from authorship support for professionals generating digital cultural content to database query engines, capable of interpreting their sources for different personalised responses.

Hence, the question needs to be posed of how 3D representations could be effectively managed and used for enhanced access to Cultural Heritage information held in a European Digital Library or other digitised repository.

As noted elsewhere, 3D objects require content classification of non-textual documents (similar to images or AV material but with some important further requirements); search & retrieval would benefit from using similarity algorithms rather than descriptive (textual) metadata; furthermore, linking in 3D documents and navigation through distributed 3D documents must be considered. Many of the issues also appear with images and depending upon the operations undertaken image data may be even more demanding than 3D. For example the search for similar objects in a collection may be simpler handled by searching for similar shapes in 3D than seeking to recognise similar 3D objects in images of them.

However, the European Digital Library, like any other major European initiative in digital Cultural Heritage, will first have to address the issue of cross-domain integration of information. For example, there is a legacy of different metadata standards and other factors that will make information integration a particularly difficult task.

While the library domain over about the last ten years has made much progress in joint cataloguing, authority files, OPACs, etc., other domains such as museums and archives are lagging behind. Compared to the library domain databases of museums and archives have remained by and large "information silos", i.e. the collection information is not accessible on the Internet due to shortcomings in technical infrastructures. Also considerations of (unclear) copyrights and effective DRM impact on the readiness of museums and archives to provide enhanced access to digital collections (where these are available). The MICHAEL project is seeking to address issues of interoperability of digital assets at the catalogue level but the interoperability at the level of primary source (artefacts) and the enhancement of underlying legacy metadata to support advanced search and exploration in the digital domain do not seem to be addressed.

11.2. In Conclusion

The considerable investment by the European Commission in ICT projects related to Cultural Heritage has had a substantial influence in building a research community targeted specifically on ICT research needed to support Cultural Heritage. As can be seen from this Research Agenda, the results to date have been considerable, but as with many exercises in knowledge discovery the "more you find out, the more you realise you don't know". There is some way to go to realise the potential of ICTs in support of the Cultural Heritage Sector. Two major areas have been identified in the EPOCH project:

- The continuing need for research into tools and techniques for mass digitisation of the many forms of Cultural Heritage data, including user created content.
- Effective tools for the use of the resulting data both for personalised search and research tools and for authoring of relevant, meaningful and engaging interpretations and interactive experiences of Cultural Heritage information.

To ensure that the knowledge and experience is taken up and used by the sector will require research in these areas to be backed by a number of support actions and accompanying networking infrastructure. In view of the nature of the sector, which is made up mainly of "social enterprises" it is certain that these measures will require public funding to make them effective.

12. Acknowledgements

This document has drawn on many sources and received helpful contributions from many people. A number of these contributors led the early assembly of the description of individual technology areas and they and many others have contributed by commenting on ideas and contributing to discussions at events such as the EPOCH Research Agenda Workshop at VAST2006 or directly on draft sections of text. Nick Ryan (University of Kent UK), (led on initial writing on "Onsite data capture and mobile technologies"); Paolo Cignoni, Dieter Fellner, Luc van Gool and Sven Havemann (CGV, Technical University, Graz, Austria) (led on "Ongoing work on 3D objects"; Martin Doerr (Center for Cultural Informatics, ICS-FORTH) ("Waking from a Dogmatic Slumber - A Different View on Research Challenges for Cultural Information Integration") at the EPOCH Research Agenda Workshop at VAST2006 in Nicosia, Nov 2nd 2006. The presentations were discussed at the time and have been discussed at other venues and with other groups since, before extensive editing and reformulation into this document. There have been additional contributions by David Bearman (Archives & Museum Informatics) on issues for Museums on the Web and Vassilios Vlahakis on Practical Issues in Cultural Heritage Exploitation.

The authors are very grateful to those who engaged in this process and apologise in advance if opinions have been unintentionally misrepresented in this document. It is also the authors' fault should errors have been introduced into this version of the EPOCH research agenda during the editing and restructuring process.

This work would not have been possible without the help and assistance of the European Commission through the EPOCH NoE (Project No IST-2002-507382)

13. Bibliography

[Addison, 2000] Addison, Alonzo C. (2000) Emerging Trends in Virtual Heritage, in IEEE Multimedia (Vol. 7, No. 2) April-June 2000 pp. 22-25

[Antenišķ, 2006] Antenišķ, Anita: Challenges and practices of industrial heritage conversion in Riga. XIII TICCIH International Congress, "Industrial heritage and urban transformation. Productive territories and industrial landscape", 14-23 September 2006, http://www.ticcihcongress2006.net/paper/Paper%20A/Anteniskeabstract.pdf

[ARENA] ARENA Project Home Page: http://ads.ahds.ac.uk/arena/project.html

[Arnold, 2006] David Arnold, (with Alan Chalmers, Andy Day, David Duce and Phil Willis) Bringing the Past to Life for the Citizen, Grand Challenge 9, http://www.bcs.org/server.php?show=ConWebDoc.4724

[Baker and Beacham, 2003] Baker, Drew and Beacham, Richard: "Bringing ancient Greek theatre to life with interactive virtual archaeology". In: DigiCULT.info 4, August 2003, p.12, http://www.digicult.info/downloads/digicult_newsletter_issue4_lowres.pdf

[BBC News, 2006] "French in Armenia 'genocide' row" – BBC News item on http://news.bbc.co.uk/2/hi/europe/6043730.stm

[Beraldin et al, 2005], J.-A. Beraldin, M. Picard, S. El-Hakim, G. Godin, L. Borgeat, F. Blais, E. Paquet, M. Rioux, V. Valzano, A. Bandiera. Virtual reconstruction of heritage sites: opportunities & challenges created by 3D technologies. In Recording, Modeling and Visualization of Cultural Heritage, International Workshop Proceedings, Ascona May 2005 (Manos Baltsavias, Armin Gruen, Luc Van Gool, Maria Pateraki (Eds.)), ETH Zurich Publication.

[Bergeron, 1998] Bergeron, Louis (1998), in: TICCIH bulletin, Nº 2, Autumn, 1998, http://www.mnactec.com/TICCIH/texts.htm

[Berlin, 2003] Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities. http://oa.mpg.de/openaccess-berlin/berlindeclaration.html

[Berndt et al 2005] Berndt, R. and Gerth, B. and Havemann, S. and Fellner, D. W. "3D Modeling for Non-Expert Users with the Castle Construction Kit", in Proceedings of ACM/EG VAST 2005 (pp49-57)

[Buckland, 2006] Michael Buckland, Access to Digital Heritage Resources Using What, Where, When and Who, DEN Conference 2006, Rotterdam http://www.den.nl/bestanden/Conferentie2006/07_buckland.pdf

[CHIN] Canadian Heritage Information Network. Standards Homepage. http://www.chin.gc.ca/English/Standards/index.html

[Christensen, 2006] Christensen-Dalsgaard, Birte: Research Challenges in the area of Digital Libraries. IST 2006 Conference, 22 November 2006, http://ec.europa.eu/information_society/istevent/2006/cf/document.cfm?doc_id=2492

[COLLADA]. COLLADA Public forum, https://collada.org/public_forum/welcome.php

[Congress] Library of Congress. http://www.loc.gov/standards/ Standards at the Library of Congress

[Davis, 1999] Davis, Peter Stuart: Ecomuseums: A Sense of Place (Leicester Museum Studies) Leicester University Press, November 1999.

[Day et al, 2004] Day, A. M., Arnold, D. B., Fellner, D., Havemann, S., "Combining Polygonal and Subdivision Surface Approaches to modelling and rendering of urban environments", Computers and Graphics, Volume 28, Number 4, Elsevier, August, 2004 (pp497-507)

[DCMI] DCMI Library Community Webpage - http://dublincore.org/groups/libraries/

[DCMS, 2006a] Dept of Culture Media and Sport, UK Government. Press Announcement 101/06 (July, 2006)

[DCMS, 2006b] Department of Culture, Media and Sport "Understanding the Future: Priorities for England's Museums, pp29, (October 2006)

[DCMS, 2007] Dept of Culture Media and Sport, UK Government. Press Announcement (January 2007)

[Dekkers, 2001] Makx Dekkers. "Application Profiles, or how to Mix and Match Metadata Schemas", January 2001. http://www.cultivate-int.org/issue3/schemas/

[DigiCULT, 2005] Ross, S., Donnelly, M., Dobreva, M., Abbott, D., McHugh, A., Rusbridge, A. (2005): Core Technologies for the Cultural and Scientific Heritage Sector. DigiCULT Technology Watch Report 3, January 2005, P-PIRO case study, pp. 92-96, http://www.digicult.info/pages/techwatch.php

[DigiCULT] DigiCULT Home Page. www.digicult.info

[Doerr, 2001] M. Doerr, Semantic Problems of Thesaurus Mapping, Journal of Digital Information, Volume 1 Issue 8, Journal of Digital Information, Volume 1 Issue 8 Article No. 52, 2001-03-26

[Doerr, 2006] Doerr, Martin (2006) "Waking from a Dogmatic Slumber - A Different View on Research Challenges for Cultural Information Integration" (Center for Cultural Informatics, ICS-FORTH) EPOCH Research Agenda workshop at VAST2006 in Nicosia, Cyprus

[DPE] Digital Preservation Europe Home Page. http://www.digitalpreservationeurope.eu/

[EC, 2002] European Commission, DG Information Society (2002): The DigiCULT Report. Technological Landscapes for tomorrow's cultural economy – Unlocking the value of cultural heritage. Authors: Guntram Geser and Andrea Mulrenin. Luxembourg. Available for download at: http://www.digicult.info/pages/report.php

[EC, 2004] European Commission: DG Research (2004): New technologies for the future of our past: EU Research for sustainable urban development and land use - Cultural heritage. Luxembourg: Office for Official Publications of the European Communities, http://ec.europa.eu/research/environment/pdf/her_tech_en.pdf

[EC, 2005] European Commission: i2010: digital libraries. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. COM(2005) 465 final, Brussels, 30.9.2005, http://ec.europa.eu/information_society/activities/digital_libraries/doc/communication/en_comm_digital_libraries.pdf

[EC, 2006b] European Commission: ICT – Information and Communication Technologies. A Theme for research and development under the specific programme "Cooperation" implementing the Seventh Framework Programme (2007-2013) of the European Community for research, technological development and demonstration activities. Work Programme 2007-2008. Agreed by Programme Committee. ftp://ftp.cordis.lu/pub/fp7/ict/docs/ict-wp-2007-08_en.pdf

[EC,2006a] European Commission: Commission Recommendation of 24 August 2006 on the digitisation and online accessibility of cultural material and digital preservation

(2006/585/EC). Official Journal of the European Union, L236/28, 31.8.2006, http://europa.eu.int/information_society/activities/digital_libraries/doc/recommendation/recommendation/en.pdf

[Ecosystems, 2003] Ecosystems Ltd., "Using natural and cultural heritage for the development of sustainable tourism in non-traditional tourist destinations" http://ec.europa.eu/enterprise/services/tourism/studies/ecosystems/study_sustainability.htm

[Ename Charter] Ename Charter home page. http://www.enamecharter.org/

[EPOCH, CHEDI, 2006] EPOCH: Information and Communication Technologies: The needs of museums, monuments and sites and their visitors. A Report by CHEDI, 14 April 2006 (incl. results of the EPOCH expert workshop at ICCROM Headquarters in Rome, 6-7 March 2006)

[EPOCH, SOTU, 2006] EPOCH, Digital Applications for Tangible Cultural Heritage: Report on the State of the Union: Policies, Practices and Developments in Europe" (2006) (Ed F. Niccolucci, G Geser, with Teresa Varricchio) ISBN 963 8046 68 6 pp168

[EPOCH: D2.1, 2006] EPOCH: D2.1 - Sector Watch Report. Ename Center, 31 March 2006

[EPOCH: D2.5.2, 2005] EPOCH: D2.5.2 - Report on Common Research Agenda. Ename Center, 29 April 2005

[EPOCH: D2.9, 2006] EPOCH: D2.9 - Research Agenda v2 (including background from Sector Watch). David Arnold and Guntram Geser, 12 May 2006

[EPOCH: D3.1.1, 2005] EPOCH: D3.1.1 – Overview of CH related IT research, related to stake-holder needs and the position of Europe therein (March 15th 2005).

[EPOCH: D3.3.2, 2005] EPOCH: D3.3.2 – 2nd 6-monthly EPOCH Pipeline Description. University of Leuven, 27 April 2005

[Funkhouser, site] Princeton Shape Retrieval and Analysis Group, http://www.cs.princeton.edu/gfx/proj/shape/index.html

[Geser et al, 2004] Geser, Guntram / Pereira, John (2004): The Future Digital Heritage Space. An Expedition Report. DigiCULT Thematic Issue 7, December 2004, is available for download at: http://www.digicult.info/pages/themiss.php

[Geser, 2004] Geser, Guntram (2004): Assessing the readiness of small heritage institutions for e-culture technologies. In: DigiCULT.Info Issue 9, November 2004, pp. 8-13. http://www.digicult.info/pages/newsletter.php

[Getty, AAT] J Paul Getty Trust, The Arts and Architecture Thesaurus Online, http://www.getty.edu/research/conducting_research/vocabularies/aat/index.html

[Giaccardi, 2004] Giaccardi, Elisa (2004): Memory and Territory: New Forms of Virtuality for the Museum, in: Proceedings of Museums and the Web 2004, Arlington, Virginia, March 31–April 3, 2004. Available at: www.archimuse.com/mw2004/papers/giaccardi/giaccardi.html

[Giaccardi, 2006] Giaccardi, Elisa (2006): Collective Storytelling and Social Creativity in the Virtual Museum: A Case Study, in: Design Issues, Vol. 22, number 3, Summer 2006, pp. 29-41 (which provides a detailed case study on the Virtual Museum of the Collective Memory of Lombardia)

[Gruber] Tom Gruber "What is an ontology" accessed at http://www-ksl.stanford.edu/kst/what-is-an-ontology.html

[Hart & Martinez, 2006] Hart, Jane K. / Martinez, Kirk (2006): Environmental Sensor Networks: A revolution in the earth system science? Earth-Science Reviews 78 (2006) 177–191, http://eprints.ecs.soton.ac.uk/13093/01/esn.pdf

[HI, 2005] J. McLoughlin, B. Sodegar and J. Kaminski (Editors), Heritage Impact 2005 Proceedings. EPOCH publication (ISBN 963 8046 66 X) (2006)

[HI, 2006] J. McLoughlin, B. Sodegar and J. Kaminski (Editors). Heritage Impact 2006 Proceedings., EPOCH publication (2007, in production)

[HICIRA] HICIRA Home Page, www.hicira.org

[Hudson, 1996] Hudson, Kenneth (1996): Ecomuseums become more realistic. In: Nordisk Museologi, No. 2, Vol. 4, 1996, http://www.nordiskmuseologi.com/www/nm/962/hudson962.html

[Hudson, 2002] Hudson, William (2002): Crossing the Wireless Chasm: A Standards Nightmare, http://www.syntagm.co.uk/design/articles/wirelesschasm.htm ((first published in SIGCHI Bulletin November/December 2002, http://bulletin2.sigchi.org/archive/2002.6/novdec02.pdf

[ICIP] ICIP - ICOMOS International Committee on Interpretation and Presentation. http://www.enamecenter.org/content/view/142/104/lang,en/

[IMRI, 2001] IMRI – Information Management Research Institute at Northumbria University (2001): The Bidding Culture and Local Government: Effects on the development of Public Libraries, Museums and Archives, http://online.northumbria.ac.uk/faculties/art/information_studies/imri/rarea/im/pubsec/bidcul/psbidcult.htm

[Industrial Heritage] European Route of Industrial Heritage project, http://en.erih.net/

[INTUITION, 2006] Research Roadmap of Cluster 2, Haptic Interaction Group INTUITION project document INTUITION_RDMP_WG2.10_V1 (2006)

[ISO15836, 2003] ISO 15836:2003 Information and documentation - The Dublin Core metadata element set

[ISO21127] ISO 21127 Information and documentation -- A reference ontology for the interchange of cultural heritage information (CIDOC-CRM).

[ISTAG, 2003] Ambient Intelligence: from vision to reality. September 2003, pp. 27-29. http://www.cordis.lu/ist/istag-reports.htm;

[ISTAG, 2004]: Experience and application research. Involving Users in the Development of Ambient Intelligence. Final Report - v1, 22 June 2004, ftp://ftp.cordis.lu/pub/ist/docs/istag-earc_wg_final_report_v1.pdf

[Italy, 2005] Italian Ministry of Cultural Properties and Activities / Ministero per i Beni e le Attività Culturali, Ufficio Statistica (2005): Visitatori e Introiti di Musei, Monumenti e Aree Archeologiche Statali (Presences in State museums and archaeological sites), http://www.statistica.beniculturali.it/Rilevazioni.htm

[Kansa et al, 2005] Eric C. Kansa and Michael Ashley Embedding Open Content in Instruction and Research, http://itsacast.blogspot.com/2005/10/kansa-and-ashley.html

[Kansa, Schult and Bissell]: Eric C. Kansa , Jason Schultz and Ahrash N. Bissell: Protecting Traditional Knowledge and Expanding Access to Scientific Data: Juxtaposing Intellectual Property Agendas via a "Some Rights Reserved" Modelhttp://journals.cambridge.org/action/displayAbstract;jsessionid=061A86BF4AC4DEEFC8F1E7757FF3243E.tomcat1?fromPage=online&aid=346256

[Kroski, 2005] Elyssa Kroski (2005) The hive mind: folksonomies and user-based tagging (http://infotangle.blogsome.com/2005/12/07/the-hive-mind-folksonomies-and-user-based-tagging/)

[Local] Local Projects, http://www.localprojects.net

[London, 2006] The London Charter (2006) 1st Draft (5th March 2006) http://www.londoncharter.org

[Maggi and Faletti, 2000] Maggi, Maurizio /Falletti, Vittorio (2000): Ecomuseums in Europe. What they are and what they can be. Instituto Ricerche Economico-Sociali del Piemonte. Working paper n. 137, June 2000, http://213.254.4.222/cataloghi/pdfires/535.pdf (102 pages) [see also: Maggi, Maurizio (2002): Ecomusei. Guida europea. Torino-Londra-Venezia: Umberto Allemandi & C.]

[McLoughlin et al, 2006] J. McLoughlin, B. Sodegar and J. Kaminski, "ICT Investment considerations and their influence on the socio-economic impact of heritage sites", in Proceedings of VAST2006 (Eds. M.Ioannides, D. Arnold, F.Niccolucci and K.Mania). (ISBN 3-905673-42-8) pp109-116 (October 2006)

[MDA, 1993] Social History and Industrial Classification: A Subject Classification for Museum Collections (SHIC) First Edition published 1983 by The Centre for English Cultural Tradition and Language, University of Sheffield, UK.(ISSN 0309 9229). Second edition published 1993 by The Museum Documentation Association, Cambridge, UK.(ISBN 0 905963 91 1) http://www.holm.demon.co.uk/shic/

[METS] METS Metadata Exchange & Transmission Standard: http://www.loc.gov/standards/mets/METSOverview.html

[MINERVA, 2006] "Multilingual Access to the European Cultural Heritage: Multilingual Websites and Thesauri" http://www.ifap.ru/library/book130.pdf

[MoLAS, 1994] Museum of London Archaeology Service (MoLAS) Manual, London 1994. ISBN 0-904818-40-3. 128pp

[Moore, 1991] Moore, Geoffrey (1991): Crossing the Chasm. Marketing and Selling High-Tech Products to Mainstream Customers. New York: Harper Business

[Mueller et al, 2006] P. Mueller, P. Wonka, S. Haegler, A. Ulmer and L. Van Gool. 2006. Procedural Modeling of Buildings. In Proceedings of ACM SIGGRAPH 2006 / ACM Transactions on Graphics (TOG), ACM Press, Vol. 25, No. 3, pages 614-623

[Nevell, 2006] Nevell, Michael: 50 years on: celebrating industrial archaeology. A major change in human evolution, in: British Archaeology, Nr. 86, January/February 2006, http://www.britarch.ac.uk/BA/ba86/feat3.shtml

[Niccolucci, 2006] Franco Niccolucci, EPOCH Deliverable D.4.2.1 Report on Standards and their roles in EPOCH, April 2006

[OECD, 2004] OECD Committee for Scientific & Technological Policy: Declaration on Access to Research Data from Public Funding. http://www.ub.uni-dortmund.de/listen/inetbib/msg23551.html

[Pine and Gilmore, 1999] Pine, Joseph / Gilmore, James H. (1999): The Experience Economy. Boston: Harvard Business School.

[PRESTOSPACE]. Prestospace Home Page. http://prestospace-sam.ssl.co.uk/

[PROBADO] PROBADO home page: http://www.probado.de

[PULMAN, 2003] PULMAN: Public Libraries, Museums and Archives: the eEurope Agenda for Local Services. Final Report of the PULMAN Network of Excellence. Edited by Rob Davies. Luxembourg: European Commission, Directorate-General Information (2003)

[Rogers, 1962] Rogers, Everett M. (1962): Diffusion of Innovations. New York: The Free Press

[Rogers, 1995] Rogers, Everett M. (1995): Diffusion of Innovations. Fourth Edition. New York: The Free Press

[Rosella, 2006] Caffo, Rossella (2006). MICHAEL Project : towards a trans-European portal of Culture, in Vito Cappellini and James Hemsley (eds) Proceedings of EVA 2006 Florence, Bologna:Pitagora Editrice, pp. 48-50.

[Schmitt, 2001] Schmitt, Bernd H. (2001): Customer Experience Management. New York: The Free Press.

[SketchUp] SketchUp Home Page. http://www.sketchup.com/

[Smith, 2002] Bernard Smith, (2002) "Chapter 19 - Digital Heritage and Cultural Content" in "A Guide to Good Practice in Collaborative Working Methods and New Media Tools Creation". AHDS,

[Statistik, Austria,2004]: Kulturstatistik 2003 – Museen und Ausstellungen, http://www.statistik.at/fachbereich_03/03_01_Museen.pdf

[Stokes, 1997] Stokes, Donald, E. (1997) Pasteur's Quadrant: Basic Science and Technological Innovation, The Brookings Institution, Washington (ISBN 0-8157-8177-6), pp180

[SWAD, D8.3] Semantic Web Advanced Development for Europe (SWAD-Europe), Deliverable 8.3 : Thesaurus Multilingual Report. http://www.w3c.rl.ac.uk/SWAD/deliverables/8.3.html

[Tagil, 2003] Nizhny Tagil Charter for the Industrial Heritage, July 2003, http://www.international.icomos.org/18thapril/2006/nizhny-tagil-charter-e.pdf

[TEL-ME-MOR, 2005] TEL-ME-MOR Deliverable D3.1: Report on TEL UNICODE requirements – June 2005

[TEL-ME-MOR, 2006] TEL-ME-MOR Deliverable: 3.3 – Subject access tools – January 2006

[Twycross, 2006] Twycross, Meg. Virtual Restoration and Manuscript Archaeology: A case study. AHRC ICT Methods Network Expert Seminar on History and Archaeology, Virtual History and Archaeology. Humanities Research Institute. University of Sheffield, 19 – 21 April 2006

[UKCRC] Grand Challenges in Computing Research, http://www.ukcrc.org.uk/grand_challenges/index.cfm

[UNESCO, 2006] Industrial Heritage Sites on the UNESCO World Heritage List, http://www.international.icomos.org/18thapril/2006/whsites.htm

[UNESCO-ICOMOS, 2006]. UNESCO-ICOMOS Documentation Centre: Industrial Heritage Bibliography, February 2006, http://www.international.icomos.org/centre_documentation/bib/2006industrialheritage.pdf

[Van Kasteren, 2003] Van Kasteren, Jost (2003): Development of the Semantic Web must begin at the grass roots level. An interview with Janneke van Kersen, Dutch Heritage Association. In: DigiCULT Thematic Issue 3, May 2003, pp. 12-13, available for download at: http://www.digicult.info/pages/themiss.php

[Varine, 1993] Hugues de (1993): Tomorrow's Community Museums, http://www.assembly.coe.int/Museum/ForumEuroMusee/Conferences/tomorrow.htm

[Vereenooghe, 2004] Vereenooghe, Tijl: Bridging the gap between technology and cultural heritage. In: DigiCULT.info 9, November 2004, pp. 14-16, http://www.digicult.info/downloads/digicult_info_9_xs.pdf

[Vlahakis, 2006] Vlahakis, Vassilios: Cultural Heritage Exploitation – Practical Issues. Presentation at EPOCH Research Agenda workshop at VAST 2006 conference, Nicosia, Cyprus, 2 November 2006

[Wikipedia] http://en.wikipedia.org/wiki/Thesaurus

[WTO, 2006] World Tourism Organization (2006): Tourism Highlights, Edition 2006. Madrid: WTO, http://www.unwto.org/facts/menu.html

Annex A:

1. Attendees and Contributors to the EPOCH Research Agenda Workshop

David Arnold, University of Brighton, UK

David Bearman, Canada

Bas Bogaerts, Playing The Past, Belgium

Vittore Casarosa, DELOS NoE

Paolo Cignoni, ISTI-CNR Pisa, Italy

John Clinton, University of Brighton, UK

Michael Danks, Windfall Digital, UK

Martin Doerr, ICS-FORTH, Crete, Greece

Dieter Fellner, Technical University of Darmstadt/FHG-IGD

Michel Genereux, University of Brighton, UK

Guntram Geser, Saltzberg Research, Austria

Halina Gottleib, The Interactive Institute, Stockholm, Sweden

Malika Hamza, Ename Centre, Belgium

Sven Havemann, Technical University of Graz, Austria

Jaime Kaminski, University of Brighton, UK

Christian Lahanier, C2RMF, France

Craig Moore, University of Brighton, UK

Franco Niccolucci, PIN, University of Florence, Italy

Denis Pitzalis, C2RMF, France

Daniel Pletinckx, Visual Dimension, Belgium

Nick Ryan, University of Kent, UK

Luc Van Eyeken, KU Leuven, Belgium

Luc Van Gool, KU Leuven, Belgium

Vassilios Vlahakis, Intracom SA Telecom Solutions, Athens, Greece

2. Presenters at the EPOCH Workshop on Standards (VAST2006). Oct 29th 2006

Andrea D'Andrea, University of Naples, Italy

Richard Beacham, KCL, London, UK

Tyler Bell, Oxford Archdigital

Bert Deknuydt, KU Leuven, Belgium

Hugh Denard, KCL, London, UK

Achille Felicetti, PIN University of Florence, UK

Sven Havemann, Technical University of Graz, Austria

Sorin Herman, PIN, University of Florence, Italy

Matteo Lorenzini

Franco Niccolucci, PIN University of Florence, UK

Sofia Pescarin

Steve Stead, Paveprime Ltd, UK

Go Sugimoto

3. Author/Presenters at the ICS-FORTH Symposium on "Exploring the limits of global models for integration and use of historical and scientific information. Oct 23-24 2006

David Arnold

N. Aussenac-Gilles

P. Constantopoulos

Martin Doerr

V. Dritsou

O. Eide

Michel Genereux

G. Goerz

N. Guarino

R. Kummer

K. May

C-E Ore

R. Smiraglia

R. Urban

4. EPOCH Museums Focus Group participants. November 10th 2006

Prof David Arnold, Coordinator, EPOCH NoE, University of Brighton

Dr Ciara Centazzo, Italy

Dr Carol Ebener, Park and Museum of Archaeology, Latenium, Switzerland

Dr Franz Hebestreit, Siemens Forum AG, Network of Company Museums, Germany

Prof M. Morgantini, Fondazione ADI (Italian Design Foundation), Milan, Italy

Prof Massimo Negri, Director, European Museum Forum

Prof Franco Niccolucci, PIN, University of Florence, Italy

Dr Andrea Rither, Museum of Recent History, Slovenia/Forum of Slavic Cultures

Dr Wim Wan der Weiden, Chairman, European Museum Forum

5. EU Workshop on Centres of Competence for Digitisation and Digital Preservation, Luxembourg. 14th November 2006

Emmanuel Amasio, Infotechnique, France
David Arnold, EPOCH/University of Brighton, UK
Mariann Backes, CVCE, Luxembourg
Markus Branti, Munchen Digitalisierungszentrum, Germany
Majlis Bremer-Laamanen, Centre for Microfilming and Conservation, The National Library of Finland
Theodor Bruggeman, Volltextsuche Online, Germany
Rafael Carrasco, Biblioteca di Storia Moderna, Italy
Stephan Cernohorsky, MEMORIA digitisation centre, National Library of the Czech Republic
Birte Chistensen-Dalsgaard, State and University Library, Denmark
Aly Conteh, The British Library, UK
Paulo Costa, Portugese Institute of Museums, Portugal
Mathias Hemmje, NESTOR, Germany
Adam Horvath, National Szechenyl Library, Hungary
Hans Jensen, National Library, Netherlands
Marek Jindrich, MEMORIA digitisation centre, National Library of the Czech Republic
Borje Justrell, National Archives of Sweden, Sweden
Minna Kaukonen, Centre for Microfilming and Conservation, The National Library of Finland
Monique Kieffer, National Library of Luxembourg
Stefanos Kolias, Image, Video and Multimedia Systms lab, Croatia
Catherine Lupovici, BNF, France
Teresa Malo de Molina Martin-Montavio, National Library of Spain
Istvan Monok, National Szechenyl Library, Hungary
Erlend Kolding Nielsen, Royal Library, Denmark
Patrick Peiffer, National Library of Luxembourg
Hans Petschar, Austrian National Library, Austria
Marc Pinter, Medea Services, Hungary
Frank Poireau, Infotechnique, France
Seamus Ross, DPE/HATII, UK
Lucien Scotti, BNF, France
Ute Schwens, Deutsche Bibliotek, Germany
Kataryzna Slazka, National Library of Poland, Poland
Stefan Strathmann, Gottinger Digitalisierungszentrum GDZ/NESTOR, Germany
Daniel Terrugi, Prestospace/INA, France
Guibert Vanhoof, Infotechnique, France
Edwin Van Huis, Netherlands Institute of Sound and Vision, Netherlands

6. Presentation/Discussion sessions on aspects of the research agenda

6.1. Where is High Tech going for Cultural Heritage: Europe at the dawn of the 21st Century. Panel Session 9th Mediterranean Conference on Archaeological Tourism, Paestum, Salerno, November 16-19th 2006

Panel Members

David Arnold, EPOCH Coordinator, University of Brighton

Andrea D'Andrea, University of Naples

Stefano De Caro, Regional Director for the "Beni Culturali e Paesaggistici della Campania"

Maria Pia Guermandi, Institute of the "Beni Culturali dell'Emilia Romagna"

Franco Niccolucci, PIN, University of Florence

Daniel Pletinckx, Visual Dimension, Belgium

6.2. IST Event 2006 Helsinki: EPOCH "Networking" Session, November 22nd: "Multidisciplinary v Interdisciplinary perspectives of ICTs for cultural heritage applications"

Session facilitators/speakers

David Arnold, EPOCH Coordinator, University of Brighton

Neil Silberman, Ename Center, Belgium

Session attended by about 40-50 conference delegates who recorded comments on research priorities on the issues raised

6.3. EPOCH Presence at NODEM, Dec 7-9th 2006

Session facilitators/speakers

Halina Gottlieb, The Interactive Institute, Stockholm

Franco Niccolucci, PIN, University of Florence

Daniel Pletinckx, Visual Dimension, Belgium

Nick Ryan, University of Kent, UK