Grand challenges: Grand opportunities? archaeology, the historic environment sector and the E-science programme

William Kilbride, Glasgow Museums
Monday, 24 July 2006

0. About this document

This document provides a report on the findings of an expert seminar group that met at the British Academy on a very hot day in June 2006. It includes a series of recommendations agreed by the group and it provides some context for how the group came to these recommendations. Although based on a transcript of proceedings, the text has been ordered differently and much of the detail has been edited out. The document starts with an introduction to how the group defined 'E-science' then presents in summary four salient themes that emerged in discussion before moving onto three sets of recommendations.

This document is very firmly a collective work and thanks are due to the following who contributed to its construction: Sheila Anderson, Tobias Blanke, Luke Blaxill, Stuart Dunn, Paul Cripps, Vince Gaffney, Isabel Holroyd, Jeremy Huggett, Lorna Hughes, Stewart Jeffrey, Ed Lee, Keith May, Stewart Waller and Katrin Weideman. Contradictions or inconsistencies that may exist come from the dynamic inherent in bringing so many people together. Errors, biases and misrepresentation, however, are purely the fault of the author.

1. Introduction: E-science and E-archaeology

One of the characteristics of E-science is a legitimate uncertainty on what can be ruled in and what can be ruled out. This is in partly because the E-science community is replete with its own jargon and partly because the discursive attraction of association with E-science gives it the feel of a bandwagon. Consequently, numerous projects and teams are actively re-describing their processes and tools as E-science in the hope of sharing in the investment that the E-science programmes offer. The problem is exacerbated by the lack of a precise and comprehensive definition of E-science.

Rather than try to define a term whose meaning is unclear, a number of existing E-science projects were presented in preliminary documentation, a selection of which were presented in more detail at the workshop. These provided basic criteria to characterise existing E-science projects.

- Collaborative but distributed interaction between scholars based on necessarily bespoke IT applications
- Collaboration that meets 'grand challenges' in a discipline
- Massive and distributed computing power
- Massive and distributed storage and transport of data
- Virtual access to shared scientific facilities

- Data mining and analysis
- Hypothetico-deductive reasoning

Crucially, none of these criteria are antithetical to archaeology. Our language may be different: hypothetico-deductive reasoning may be part of a broader interpretative approach; shared scientific facilities such as mass spectrometry and electron microscopes are often shared with colleagues in other disciplines; we may call our grand challenges 'research frameworks'. But these are minor quibbles. The prospects for archaeology and the historic environment engaging in E-science are strong. Uniquely in the arts and humanities, archaeology would be able to present research projects which matched *all* of these criteria.

As well as using the set of projects to infer the boundaries of E-science, the panel also took the opportunity to review whether our experience of E-science projects met their own stated aspirations. There was consensus that the E-science had enjoyed a high profile, especially within those institutions which have received investment in E-science infrastructure. It was proposed that the benefits of the investment had not been distributed evenly and that results of specific targeted research had been more modest than either the investment or profile would suggest. Gartner's 'hype cycle' was introduced to the discussion. While being generally acknowledged as a recognisable model for the adoption of technology, there was disagreement as to where we should place ourselves on it. It was agreed however that, by joining the E-science programme relatively late in the day, archaeology should be able to develop more quickly.

2. Archaeology as Virtual Research Organisation

Characteristically, E-science is about collaboration. Superficially, therefore, archaeology is well placed to contribute and benefit from the programme since almost all archaeological research projects are shared endeavours. Archaeology is smaller in terms of number of active researchers than many of the scientific disciplines. This makes collective decision-making and agenda-setting easier. But there is also a tendency to localisation of research interests which leads to a degree of fragmentation between research groups.

Collaborative tools, loosely specified as Virtual Research Environments, have been used in a number of E-science projects to facilitate project management and interaction. Archaeology has a long history of document sharing and personalisation tools and in cases such as the Silchester VRE project, a widely used site database has been explicitly re-described as a Virtual Research Environment to underline the collaborative nature of archaeological excavation to the E-science community.

Work on portals-based tools are relevant here also, especially the development of portlets for personalisation – such as implementation of WSRP and JSR286 – which are designed to allow remote content to be re-cut according to the needs of distributed but personalised working environments (Awre et al 2004, Awre et al 2005). Such work extends several generations of research and development tools to enhance archaeological field recording and post excavation analysis. There is no question that archaeology can benefit from adopting some of these tools and, by adopting the language of E-science, make existing archaeological applications available to other researchers.

Over and above collaboration on specific research projects, it is arguable that archaeology is closer than almost any other discipline to transforming itself into a complete virtual research organisation in which the different agents within the disciplinary body act as persistent consumers and producers of information through a continuous collaborative cycle of analysis and interpretation. The roots for this transformation exist in development control archaeology and the private sector rather than the academic sector. The OASIS project has migrated this existing process to a digital information flow and has opened this process to the academic community. Although development is painstaking, the use of the OASIS data pipeline means that contracting fieldworkers, local government curators, national heritage agencies and academic researchers are linked into a single virtual organisation that supplies and supports research outputs.

Though incomplete this transformation worthy of analysis by other disciplines: although it has only come to the table late, archaeology may well be the first complete disciplinary grouping to transform itself into a comprehensive virtual research organisation.

The expert seminar noted and welcomed this development. It also became clear that better exchanges of metadata and access to grey literature, however useful, were not ends in themselves, nor were they all that archaeology could achieve as a Virtual Research Organisation. For example, the extensive textual archive that OASIS generates could usefully be open to full-text indexing and other text mining applications might be appropriate. Moreover the ability to add more complicated digital objects to the reports – raw geophysical survey data or site GIS for example –would open interesting possibilities for collaborative virtual research. In addition there is no formal mechanism to capture the re-use of OASIS products and the identification of research in progress is hard. There was recognition of the mixed quality and utility of the research produced in development control archaeology and of the very great backlog of material not currently available. Each of these headings point to work that might be done to support the emergence of our Virtual Research Organisation.

3. Infrastructure and Networks

One of the familiar metaphors in E-science is the idea that an entire network (or grid) of computing power acts as a single but infinitely extensible computer. Following this metaphor, if different parts of the grid behave as components of a single system, then the duplication of services becomes wasteful in more than simply financial terms. As well as reducing revenue costs de-duplication frees up processing power and storage and allows researchers to move to substantive questions more quickly. The geospatial data (GD) sector was presented as a useful analogy of this process: in part because it is relatively more advanced in its preparedness for E-research, and in part because it is a sector with which archaeologists routinely engage.

The problems and opportunities of grid computing were mentioned throughout the workshop, noting that the grid is itself poorly defined as a term and subject to continuing change as an infrastructure. Consequently it was agreed that the physical infrastructure and networks upon which E-science depends is a moving target: it's hard to know what the infrastructure can actually support. It was proposed that experience was of a retreat from widely distributed grid computing to more narrowly distributed cluster computing. This latter approach had the crucial advantage of having simple lines of institutional management and credit and was not constrained by wrangling over inter-institutional service delivery.

It was noted that infrastructure reached from local through institutional through regional to national and international domains. The lines between these infrastructures are invisible to the researchers but are very real organisationally. A single academic will have little real control over national investment but should have more involvement in local matters. The question arises of how to link the local infrastructure to the regional and national ones, and whether the national investment is sufficiently cognisant of local pressures. Three risks exist. Firstly there is a risk of supposedly national investment being sidelined into a few personal projects, for which strong external management is the only practical solution. Secondly there is a risk that research projects are turned into services without sufficient investment or understanding that the researchers may not be best placed to offer services. This risk emerges when E-science falls into the interstices between a university computing services and computing science department, which was noted as a familiar experience. Finally there is the risk of institutional politics preventing genuine collaboration.

These three risks mean that the research councils and national funding agencies need to pay much closer attention to the local infrastructure of individual scholars in the development of E-science, especially true in disciplines. The stark fact is that not all institutions and individuals are equal. Either investment should be made more even to ensure consistency or by failing to investment or share resources, institutions and consortia effectively rule themselves out of e-science.

4. Visualisation, Simulation and Modelling

Many E-science projects have used or developed visualisation and simulation tools as a basis to solve problems, especially in disciplines where visualisation and simulation are familiar methodologies. Examples include using E-science resources to visualise, analyse and share the outputs of CT scans, microscopes and other sensor arrays. In addition, grid computing power means that visualisation need no longer be a solitary experience: collaborative virtual environments allow groups of scholars to interact in real time with a shared visualisation.

Irrespective of the infrastructure, the seminar agreed with one member who observed that 'the majority of applications are still going to break down into what we do now but on a bigger scale' listing imaging, surface modelling, volumetric modelling, remote sensing, numeric simulation modelling, geo-spatial and geo-temporal processing and text mining as the principle application areas. Consequently applications which previously failed because of lack of computing power would be become possible. However these developments would have to be cognisant of related theoretical changes in archaeological thinking. Behavioural simulation for example, was not widely used because the computing power was insufficient but also because it fell out of favour intellectually. A firm prediction that this would become more popular again was qualified by the clear sense that the theoretical underpinnings would have to be included as part of the processing.

A four-fold model of the research processes associated with all modelling and textual analysis was proposed.

1. Resource creation
2. Documentation and preparation
3. Analysis
4. Output and dissemination

Resource creation remains the subject of research and development but for analytical purposes is not an end in itself. Once created, this data goes through a process of documentation and preparation, such as data mining and sampling, to enable the core research processes of analysis. The analysis phase– characterised by textual analysis, remote sensing, imaging, temporal modelling, simulation modelling, surface modelling or volumetric modelling creates an output which is evaluated and either published or discarded.

This fourfold process of resource creation, development, analysis and dissemination provides useful reference model for all E-science development. It is sufficiently abstract to allow a whole range of activities to be encompassed within it. Crucially a balance needs to be struck in investment. Hitherto a great deal of energy had been expended on phases documentation and dissemination, with relatively little investment in innovative or shared resource creation or innovative and shared resource processing.

5. Resource discovery, delivery and processing

E-science relies on the ability to locate, deliver and process information so a number of core applications and tools have emerged that support or extend these functions for the E-science community. In effect E-science is neutral about the approaches used to discover and deliver resources. Nonetheless, the dependence on the ability to do so means that changes in technologies which support these functions are important.

It was noted early on in the workshop that one of the grand challenges faced by archaeology was to ‘unlock’ the very great potential that already exists in the very many data sets and tools that have already been accrued. Unlocking this potential, however, meant reaching far beyond resource discovery and into processing. There was general agreement that if all we managed to achieve were more sophisticated mechanisms for searching and retrieving files then we would have missed the opportunity to do a great many things more interesting.

How could archaeology develop shared processing of resources? Three inter-connected lines of argument emerged. Firstly it was noted that despite the very great concentration on resource discovery, these technologies had not yet developed fully: so action was still required on topics like de-duplication and indexing. Secondly, action on resource discovery characteristically progresses as standards development. There was consensus that while standards were useful they were not – nor should ever be allowed to become – ends in themselves. Moreover there was agreement with one member of the panel who insisted that we must not privilege standards creation and adherence over blue skies research and that standards should never be allowed to inhibit innovation. Best practice should not be used to prevent good practice.

Returning to the question of processing existing data, the group were presented with the sorts of distributed modelling undertaken by SETI and by the BBC Climate Change predictions. These provide clear examples of where emphasis is on processing rather than discovery. However one member of the panel objected – that computational power wasn't an issue for archaeology. Our data is not in a fit state for large scale computational processing, so a mass of hard work is needed to make the data ready for computation. To achieve this we need to re-think data gathering and processing, which takes us well beyond the possibilities of E-science and into issues of organisational and professional cohesion: considering the virtual research organisation is the only way for archaeology to proceed beyond resource discovery since that is where we begin to question and assess the quality of data supplied. If the virtual research organisation were to act as a venue for shared data creation, dissemination and archiving, then there is potential for shared processing.

The problem of data preparation and enabling then introduced a three-fold division of data sets which also epitomises the functionality and purpose of processes applied to the data. Assuming that data has three forms: existing enabled, existing created and newly created, then there are three parallel sets of processing required. An existing enabled data set may be subject to processes (including augmentation or production of the other two types of data) which lead towards blue skies research. An existing created data set may go through processes of enablement (such as preparation for resource discovery) as well as being augmented by the other two types of data. A newly created data set may be derived from an existing enabled data set or be the product of new data creation from fieldwork, and thus lead to its publication for subsequent processes of resource discovery.

The question of data processing also raised the problem of effective data linking – and the discussion turned for a moment to consider the issue of persistent identifiers. There was consensus that data processing would need some sort of deep and dependable linking between data sets and tools – beyond file level to items within files. Although such tools seem to be fundamental it was surprising that they seem to be missing from the current E-science prospectus. Their development is essential.

Assuming that we can share data in such a way as it will be useful to others, the panel asked whether the tools and processes could also be shared. Tools sharing is an obvious step from data sharing but contains a different series of challenges. Like data, tools are developed for a specific purpose, so subsequent re-use of tools implies publication of the original purpose and a willingness to adapt tools to new purpose. In data this can be done through the publication of standards, but the language of tools capabilities has not yet properly emerged. Nonetheless, if tools could be abstracted through different cycles, then possibilities for tools development are clearer. Following this line of reasoning, there may be layers of tool fragments from generic tools that are widely applicable to specific tools that are refined and unlikely ever to be re-used. Efforts would therefore focus on fitting and refining generic tools to produce specific ones. Self-assembly tools, based on interoperating layers of abstract and specific tool fragments can therefore emerge. At the very least this process would lower the cost of tools development.

Like data, the development of tools and processes has historically attracted little credit or reward. Consequently, if a generic and distributed tool kit is to emerge, then processes need to be put in place to encourage researchers to take risks in tools development. Tools development is some way behind data in this regard. Data publication is beginning to bring credit and mechanisms for peer review and quality assurance are being posited. Tools will need similar treatment.

6. Conclusions from the seminar

Archaeology is not only ready for E-science: it is able to introduce resources that in turn will assist the E-science community.

Perhaps uniquely in the humanities, archaeological research involves collaboration, access to shared instruments, production, access and analysis of large scale data sets and a degree of hypothetico-deductive reasoning. Archaeology has already provided a generation or more of experimentation in visualisation and simulation and it already offers at least one rudimentary but extensive virtual research organisation. Although IT skills remain in short supply the computer-based research undertaken in archaeology is as varied as any discipline so the skills available are equally diverse. Archaeology has been receptive to electronic publication and has a proven its commitment to long term viability of digital objects. The expert seminar resolved quickly that archaeology could benefit from E-science and moved onto a detailed dissection of how the two might move forward together.

The key task identified at the workshop was how to unlock the potential of the very large amounts of data already available, and how best to exploit the expertise, enthusiasm and collaboration already at our disposal.

The key benefit that archaeology can bring to E-science is its a nascent, though incomplete transformation into a Virtual Research Organisation in which many partners act as persistent consumers and producers of information. This is worthy of analysis by other disciplines.

In order to move the debate forward, a number of basic principles were suggested against which any proposals should be evaluated and a series of recommendations were made regarding current research that could be extended using E-science. Finally a set of tools were proposed that will help archaeology integrate more fully with the E-science community.

7. Archaeology, the Historic Environment and E-science: Some principles

Participants at the workshop agreed a series of generic principles on which archaeological E-science projects should be judged.

- *E-science activities in archaeology should focus on the needs of archaeology*

The group noted the dangers inherent in adopting a technology for its own sake, and projects that took a 'technology first' approach are to be avoided. The benefits for good research projects in E-science ought on balance to favour archaeological research and be appraised by archaeologists.

- *E-science activities in archaeology should be considered as fundamental research*

There was a strong feeling that computer-based and science-based applications in archaeology tended to suffer from the expectations of the research councils as they often deliberately set themselves the task of applying lessons learned from cognate disciplines. This has had the result that first-rate research has been sidelined as insufficiently innovative by the research councils or deemed insufficiently mainstream for research assessment. Academics engaging in the E-science programme need to be convinced that their work will be considered fundamental research or the programme will be ignored.

- *E-science activities in archaeology should be assessed on the basis of archaeological values*

Following this, the group agreed that the measure of success or failure of an ICT project can only reasonably be assessed by relation to its importance to the discipline in question – not by the standards of computing science or the disciplinary group which may have initiated or invented a specific application.

- *E-science activities in archaeology should push the boundaries in the use of ICT in research*

The group recognised the value in extending the use of ICT in research as an end in itself, partly so that it can build capacity and partly to maintain a momentum behind the use of ICT in research in archaeology.

- *E-science activities in archaeology should involve the whole sector not just ac.uk and be embedded within the community*

The group recognised that the majority of research in archaeology in the UK is currently commissioned and executed outside the academic sector. If E-science were to be successful in archaeology, then it needs to be open to non-academic partners and agencies. There would be little merit in the academic community pursuing an E-science agenda in isolation from the many other agencies and projects that contribute to archaeological research.

- *E-science activities in archaeology should involve existing standards and professional groupings.*

There are already a number of well-established groupings in archaeology which consider both practical and strategic development in ICT. These include the Forum for Information Standards in Heritage (FISH) and the Historic Environment Information Resources Network (HEIRNET). E-science developments should complement these groupings and support them where appropriate.

- *E-science activities in archaeology should be international as archaeology in UK universities is international*

It was recognised that a proportion archaeological research and teaching in the UK is based overseas, with particularly strong groupings across Europe, the Mediterranean and the Middle East. These networks of researchers collaborated with very many external and foreign agencies. E-science is collaborative and international in scope and there would be merit in exploring the collaborative and international possibilities of E-science with reference to international archaeological research.

- *E-science activities in archaeology should go beyond resource discovery*

The group agreed that attention to resource discovery issues has been very productive for archaeology but that resource discovery was not the only area for possible further work. Applications might involve remote access to data processing tools such as text mining, modelling and visualisation tools, access to large-scale computing infrastructures, access to scientific instruments, access to collaborative research organisations – or all of these together. The purpose of E-science activities should not be restricted to digitisation, standards, improvements in how we find documents, or extensions to existing metadata or ontological tools.

- *E-science activities in archaeology should include work on training and skills*

There was consensus that ICT skills in archaeology were generally higher than other humanities disciplines, but that even so there was a great need for detailed practical training in advanced ICT as well as broader education on the appropriate deployment of ICT in research.

- *E-science activities in archaeology should distinguish between analysis and interpretation*

Archaeological study is characteristically interpretative and hermeneutic rather than strictly hypothetico-deductive reasoning. Scientific analyses are appropriately and widely applied but are seldom sufficient on their own. They are embedded in a continuous cycle of interpretation and debate. Consequently, where we use E-science to process large quantities of data, we should not seek new answers so much as new questions. The distinction between the results that computing processes can produce and the interpretation that we chose to make of those results should not be lost.

- *E-science activities in archaeology should scalable, sustainable, implementable*

Any activity should look to the long term and the potential for re-use by others. We should not wait for a single big project, but if small projects are funded these will have a greater impact if they can be shown to lead towards a shared set of tools and conclusions.

- *The success of E-science activities in archaeology should not simply be measured by successful outcomes.*

Projects should not be funded to fail, but experimentation is essential and some degree of failure within the projects we develop is not necessarily unhealthy, provided lessons are learned.

8. Archaeology, the Historic Environment and E-science: recommended actions

Having reviewed the principles on which E-science applications should be evaluated, the seminar turned its attention to the sorts of practical activities that might be considered. A long list of recommended actions emerged which has been subdivided into three smaller groups. We start with a generalised set of headings for archaeological research and move onto a set of specific 'wayfinding' studies that will clarify the relationships between existing tools, their application to archaeological problems and their relationship to the E-science agenda. A further set of recommendations emerged that might be classified as research-support initiatives that are of wider relevance than just archaeology.

The seminar was clear that archaeology and the historic environment would principally be interested in pursuing and applying eight modes of E-science research. These have the advantage that they have already been the subject of limited development in previous E-science projects but have not been fully exploited outside of their native domains. Archaeology would not only benefit from collaboration with those who have developed such tools, but would be able to contribute novel insights and refinements that would be of wider benefit:

- Textual analysis
- Temporal analysis and modelling
- Surface modelling
- Volumetric modelling
- Remote sensing
- Imaging
- Geo-temporal computing
- Simulation modelling and intelligent agents

For example, it was noted that archaeology often required large scale computing power – such as is now available through the computational grid – for geo-spatial computation. In fact these computing processes are designed to help account for dynamic, 3 dimensional landscapes not the static 2-dimensional maps currently available for processing. If, through grid computing, we can remove some of the restrictions of processing power then archaeology and other disciplines would be able to develop more sophisticated geo-temporal applications that allowed more depth of time and introduced a 3rd dimension to landscape. This would be of interest across a range of disciplines concerned with change through time in three dimensional landscapes such as environment, oceanography and earth sciences.

In a similar way, visualisation, imaging, remote sensing volumetric and 3d modelling were noted as examples of where E-science tools could be applied and developed to extend current archaeological research. The group made a strong case that archaeology would benefit from conversation with E-scientists already engaged with such tools and that given the fragmentary and temporal nature of archaeology, there would be potential not simply to apply existing tools but to develop and extend them. In this way archaeology would both benefit from and extend the E-science programme.

It was noted that previous attempts at simulation modelling and intelligent agents in archaeology had failed in part because they were constrained in terms of computing power but also because they were 'black box' applications. A grid architecture would not only eliminate the constraints on computing power but also make such tools more transparent. Even so, the group noted that such research was further compromised by association with outmoded functional and processual approaches to anthropology so would need to address these broadly-based philosophical objections if they were to progress. These objections cannot be resolved by reference to E-science, thus reinforcing the recommendation that E-science activities should focus on the needs of archaeology and be assessed on the basis of their contribution to the discipline,

These generalised areas of activity could be usefully supported by a number of more specific activities. The expert group proposed six specific but interconnected way-finding initiatives to help archaeology along the road of E-science development.

- Studies of large scale data gathering and delivery
- Data mining to extend data sharing beyond metadata
- Ontology developing and testing as a support to data mining and integration
- Development of simple data exposure and integration tools
- A review of web services and their relation to E-science
- Reconnaissance of E-science tools for archaeological sciences
- From Research Frameworks to Grand Challenges

For example it was noted that OASIS is a Virtual Research Organisation, but that the current architecture is based on metadata sharing and it currently only makes available relatively simple digital objects (textual reports). This could be enhanced in two ways. Firstly it would be possible to use more complicated data types and allow novel forms of data integration. So for example, users with an interest in geophysical survey data might be able to manipulate and visualise existing data sets, if only the appropriate access mechanisms were added to the existing tools. Secondly, the several hundred reports currently available through OASIS are only accessible through a user-initiated metadata search. Appropriate text mining tools could allow full text analysis, especially if used in conjunction with appropriate pre-coordinated or post-coordinated indexing tools. This raises the question of indexing terminology and markup of text and means that facetted classifications or ontology connectors could also be built on such text mining tools. In theory at least this would allow a deeper engagement with source material by reducing our reliance on intermediate metadata brokerage services. It would make data integration more possible and facilitate much more than resource discovery.

It was noted that, even though there is a very well developed process for releasing data in archaeology through OASIS, it is relatively slow and relatively demanding. Simplified data exposure tools could be developed – based perhaps on Web Sevices – and which in turn could extend the potential of archaeology to act as an virtual research organisation.

It was also noted that archaeology produces some very large data sets which need a co-ordinated approach to preservation and delivery. In particular, the increasing use of precision 3d digital scanning, and continuing improvements of technologies like sonar, interferometry, ground probing radar and other geophysical techniques point to an imminent crisis in data supply from field to lab to archive to publication. The E-science programme has developed mass transport and storage tools, such as UKLight, from which these emergent but increasingly popular archaeological techniques would benefit immediately.

The expert seminar spent some time considering the possibility of identifying ‘grand challenges’ for archaeology, which have provided a useful framework for existing E-science research. Although inherently collaborative, archaeological research tends to be particularist, in part because research tends to be geographically and temporally isolated. There have been efforts to draw such diverse research together, such as the recent programme of ‘Research Framework’ documents (e.g. Glazebrook 1997, Nixon et al 2002). However these have tended to be more influential in the private sector research and have had only limited impact on the academic sector. This mismatch underlined a repeated them of archaeological debate in the last decade: the vast majority of research in the UK is undertaken under the provisions of development control, from which the academic sector is often remote. This has two implications for the E-science programme. The desire for tools and infrastructure to be access to all parties not just the academic sector has already been mentioned. An additional approach might be to encourage the academic sector and research councils in particular to evaluate and where appropriate adopt the same regional research frameworks. Consequently a project to match existing academic research to existing regional research frameworks could make some helpful recommendations on whether the regional framework documents could evolve into ‘Grand challenges’.

Finally it was noted that, although there are very many projects in E-science currently developing tools, there was not an authoritative list of tools that specified how these tools could be applied to archaeological research – especially archaeological science. Consequently, a basic catalogue of tools annotated for an archaeological audience would be of immediate benefit, especially if that could then be turned into a way of matching research groups.

A series of fundamental tasks were noted as being of much wider relevance than archaeology alone. These are presented as recommendations for action though it is recognised that these may in fact be adopted by others or delivered as sections of wider initiatives.

- Infrastructure investment and development
- Capacity building
- Dissemination of tools, processes and research outputs
- Collaborative project management

It was noted that there has already been some investment in infrastructure but that this has been uneven. In particular it was argued that headline-catching national initiatives had very limited local impact for archaeology. The successes that could be identified for archaeology from these national initiatives in fact seemed to depend on a small number of enthusiasts and experts. Consequently local infrastructure is relatively weak and needs further investment.

This problem is exacerbated by a shortage of skills: the number of graduate students or researchers in archaeology that are able to use grid tools is small and the very small number of technicians and supervisors available to support their work means that the situation is unlikely to improve. Training and scholarships are necessary.

Both of these problems make it hard to disseminate tools, processes and research outputs. Even so it was recognise that simply fixing local infrastructure and providing training would not be sufficient. Systems of credit and professional advancement need to reward those who make available their digital tools and research outputs, otherwise senior researchers and research managers who are naturally risk-averse would be unwilling to commit large amounts of institutional time and effort to E-science.

It was further argued that the development of scientific techniques in archaeology – including mass data gathering and publication – had historically risked being considered insufficiently innovative as science and insufficient mainstream as archaeology. A gap was noted between the AHRC and NERC which had hindered legitimate research. The development of dissemination mechanisms therefore needs to be more than technical access. If dissemination means credit then it must include some form evaluation, such as peer-review. Moreover, the gap between research councils on the experimental uses of information technology needs to be filled and that the RAE which exists to encourage research should not have the effect of penalising those who engage with this area of research. The former requires policy development and is likely to be wider than just archaeology. It was recognised that the latter was to some extent within the gift of archaeology to develop on its own: peer review of tools, processes and data sets are essential and mechanisms to undertake such peer review need to be established.

Archaeology has had some success in the development of collaborative research management through initiatives like the Virtual Silchester project. This and other projects have extended the development of collaborative project management tools which go some way to changing how archaeological research is carried forwards. Real time interaction between distributed experts who have different levels of confidence and expectations of the technology presents a new area for research in the anthropology of research. These projects expose strengths and weaknesses in existing working practices and will contribute to new forms of research behaviour. The changes implicit in collaborative project management and virtual research environments are not trivial. Consequently, continuing experimentation with virtual research environment will provide a significant long-term understanding about the evolution of the academic community.

9. E-science tools in archaeology

Basing discussion on our high level principles and areas for activity, the workshop ended by identifying specific tools that could be of particular interest to archaeology. The list of possible tools is as follows. This list is not exhaustive nor are proposals self contained: it is only intended as a series of possible tools to which other proposals may be added in time.

- Tools to extend OASIS beyond metadata into data mining and analysis
- Ontology connectors to allow the archaeological information community to share knowledge more effectively and to access cognate data in other disciplines
- Tools for declaring and sharing capabilities of data sets, such as geophysics, volumetric and surface modelling, geospatial and geo-temporal data sets and aerial photography
- Deep and persistent identifiers to encourage deeper linking and to facilitate access for machine-to-machine processing
- Tools for trust mechanisms and digital rights management to extend collaborative project management while ensuring that rights are properly managed and to ensure that archaeology can exploit 3rd party data sets or tools such as commercially sensitive mapping data or image visualisation applications
- Tools for concordance and de-duplication of datasets to help link versions of the same document and thus ensure that users can access the most up-to-date version of a document and reduce the impedance caused by the proliferation of sources
- Tools for mapping, declaring and managing standards which open the process of standards creation and refinement to a much wider audience
- Data enrichment tools which allow users to bring diverse data sets together into a single research environment and then re-publish the enhanced data set back to the original source with minimal delay and without compromising the integrity of the originating data sets.

10. Bibliography

Awre C, Dovey MJ, Hunter J, Kilbride W and Dolphin I 2004 "Developing Portal Services and Evaluating How Users Want to Use Them: The CREE Project" in *Ariadne Issue 41*, Originating URL: http://www.ariadne.ac.uk/issue41/awre-cree/intro.html, last visited 24th July 2006

Awre C, Waller S, Allen J, Dovey MJ, Hunter J and Dolphin I 2005 "Putting the Library into the Institution: Using JSR 168 and WSRP to Enable Search within Portal

Frameworks" in *Ariadne Issue 45*, Originating URL: http://www.ariadne.ac.uk/issue45/awre/intro.html, last visited 24th July 2006

Glazebrook, J (1997) *Research and Archaeology: A Framework for the Eastern Counties*, East Anglian Archaeology Occasional Paper 3

Nixon T, McAdam E, Tomber R, Swain H, (2002) *A Research Framework for London Archaeology*, Museum of London, London