A Digital Future for our Excavated Past

Tony Austin, Damian Robinson and Keith Westcott

Archaeology Data Service, Department of Archaeology, University of York

King's Manor, York, YO1 7EP, United Kingdom

e-mails: afa2@york.ac.uk, djr12@york.ac.uk, kaw4@york.ac.uk

1. Digital archives in archaeology

Computers have been used for archaeological research since early in the 1950s. Initial use was largely confined to specialists such as mathematicians and statisticians devoting some of their time and the processing time of a "mainframe" computer to archaeological problems (e.g. Wilcock 1999:35). In general archaeologists only started to use computers with the advent of the Personal Computer (PC) in the late 1970s. From this time on an irregular series of publications chart the rise of the PC as an increasingly necessary research tool for the archaeological community in Britain. In 1980 about 10 "micro-computers" were identified by the Museum Documentation Association (MDA) assessment of computer use in Archaeology (Stewart 1980). By 1986 the number of computers in use had risen to over 200 although some of these were being used as "intelligent terminals" to mainframe computers in university environments (Grant 1986:20). A further survey in 1989 suggests over 700 computers in use (Booth et al. 1989:4). In 2000 extrapolation from a recent survey suggests that between 2,200 and 3,000 professional archaeologists in Britain have computer access and that by the year 2004 computer access will be ubiquitous within the profession (Condron et al. 1999:24-7).

Using computers leads naturally to the creation of digital data and unsurprisingly as usage has grown the amount of data has grown and grown. Trying to quantify the amount of data produced by the archaeological community in Britain is something of a problem. Within the academic sector figures produced in 1998 suggested that as many as 1,800 of the 3,000 or so archaeological projects funded by the British Academy over a 30 year period had a digital content (Austin 1998). A broader survey by the Archaeology Data Service (ADS) in 1998 estimated the size of the digital resource produced by archaeologists in Britain as between 140 and 175 gigabytes (Condron et al. 1999:41). With the phenomenal growth in computer usage already noted the data mountain is undoubtedly much bigger today.

Modern excavations create huge amounts of digital information. Whether it is the on-site recording of the archaeology, specialist databases created during post-excavation or publication standard interpretative maps and plans, digital information has the potential to be created at every stage from assessment to publication. The existence of so much data raises problems in that it often represents primary data and sometimes is the only record of a necessarily destructive archaeological process. Synthetic accounts may also exist but here much of the functionality and complexity is lost in the move from dynamic data to a static account. An example might be the flexible query and retrieval of data from a database compared to the fixed table generated from it for inclusion in a report.

The same is true of many types of digital data; GIS, topographic and geophysical surveys, virtual reality models, CAD drawings and many others. In short, digital archives are important both in terms of re-use and of re-interpretation. The fragility of digital data should also be noted. While all storage media including paper have a finite life-span there are additional problems associated with the magnetic and optical media used to store digital data in the form of redundancy of both hardware and software technologies.

Within the discipline there has been an increasing awareness that this vulnerable data is as much a part of the primary site archive as the artefacts and paper records that have traditionally found their way into museum stores. Yet until comparatively recently there has been no secure home for such information and digital data simply languished in the virtual "desk drawers" of their creators. Clearly there was a need to create the structures; technological, managerial and administrative for the long-term preservation of and the provision of access to digital archives (e.g. Beagrie and Greenstein 1998).

The ADS was consequently founded in 1996 to provide just this kind of digital archive for the UK archaeological community. This paper looks at two specific but very different examples of digital archiving of archaeological projects undertaken by the ADS. The first involves the archives of several archaeological units based in museums and known by their penultimate resting place as the Newham Archive. When the Newham Museum Archaeological Service was closed down the digital archive eventually arrived at the ADS in the form of a large cardboard box containing dozens of floppy disks. The ADS was effectively thrown into a "rescue" situation in trying to save this data in battling against both deterioration and technological redundancy (Austin 1999). The second project was specifically designed to examine the problems associated with curating and providing access to digital archives. In 1999, the ADS, with English Heritage, the Oxford Archaeological Unit (OAU) and the Museum of London Archaeology Service (MoLAS) undertook a study into the needs and practicalities of the long-term preservation and dissemination of digital archives resultant from excavation. The Digital Archive Pilot Project for Excavation Records (DAPPER) set out to establish best practice for digital archiving in Archaeology.

1.2. The Newham Museum Archaeology Section digital archive

The Archaeology Section of Newham Museum, Greater London, was closed in late 1997 and although the closure was not entirely unexpected little thought was given to the fate of the material held on the Archaeology Service's computer system. It was not until there was a real threat of the archive disappearing that determined moves were made to save the archive, which was transferred to floppy disks and delivered to the Archaeology Data Service.

1.2.1. Accessioning the archive

The archive arrived on a total of 239 floppy discs from which a total of 6350 files were recovered. The floppy disks probably represented an image of the contents of the file server and individual hard disks on the network of the Archaeology Section of the museum. Thus it was possible to reconstruct the archive on the ADS server giving us a snapshot of the digital archive of a typical contracting archaeology unit. Initially the files were copied from the floppy discs onto the ADS server. A metadata record for each file was created as part of the accessioning process. This recorded the file name, path, format, size, last modification date and time, an abstract of the contents and comments such as problems encountered during transfer. This form of file management metadata is essential to help plan both the short term rescue of the data and also the long term maintenance of the resource. At this stage the fragility of the data was initially demonstrated when it was found that there were 25 corrupt files. The contents of around half of these files could be recovered after running scandisk on the discs although there was some loss of content. The remaining files could not be rescued and their contents were irrevocably lost.

1.2.2. The contents of the archive

The archive contains varying amounts of information for about 180 sites in northeast London that had been investigated over a period of over 15 years. Not all of the files include strictly archaeological data and there are also managerial files, project management files, correspondence and personal matter - in short all the sort of things one would expect to find on the computer system of a contracting archaeological unit. Table 1 breaks down the contents of the archive into functional categories.

The largest single category is survey data, which accounts for over a quarter of the files in the archive. This represents land surveys, borehole surveys and resistivity surveys. The second largest category is reports - Archive Reports, Level III reports, Evaluation Reports and Desktop Studies. Many of these files are draft copies created during the course of production of a printed report and consequently a substantial proportion of these files are unfinished or are duplicated in various stages of completion throughout the archive. Only the most current version of any report will be disseminated as part of the online archive, although the unfinished reports represent a rare opportunity to study the developing interpretations of a site throughout its post-excavation analysis.

Worryingly the third largest category is "unknown/no-value" which accounts for 14% of the archive. There are a number of files for which it proved impossible to determine what software was used to create them. Even worse, there are data files that cannot be associated with a site because they contain no information to specify what site they form part of the archive for and they were not in a directory associated with a recognisable site. These include over three hundred resistivity survey data files, a number of database files, and fragments of reports. Unfortunately without such contextual documentation about the site and the files associated with it the information contained on these files are to all intents and purposes lost to the scholarly community.

1.2.3. File formats

The Newham Museum Service began the process of computerising its post excavation practices relatively early on in the uptake of IT in the archaeological profession. Consequently a wide range of software has been used to create the files contained within its archive, which operate in both MSDOS and Windows environments. In order to archive and disseminate these files they have been migrated with no loss of information to formats that can be read by modern software. Table 2 shows the proportions of files created by the various package types.

Word processors were responsible for more than a third of the files in the archive. Five different word-processing package formats are present in the archive; the earliest files are in WordStar format progressing through Word 2 to Word 6 with occasional files in other formats. As modern word processors come with a variety of filters enabling a wide range of formats to be read and written migrating these files was a relatively straightforward task. The files created by the DOS version of WordStar, however, presented some problems as it only needed to support character printers for which 7 bit ASCII was sufficient. "Fancy" effects (emboldening, underlining) were coded into the non-printing ASCII characters. This left the eighth bit free which was used to flag the last letter in a word, presumably to help in wrapping text on the screen or page. Both Word and WordPerfect have filters to read WordStar files but they frequently cannot distinguish them from plain text files and use characters from the extended ASCII character set in place of the last character of every word. Unfortunately this makes reading these DOS WordStar files very difficult.

Table 2 appears to show that databases appear to be a significant proportion of the archive. This is misleading, however, as the majority of database files were created using dBase and consequently a proportion of the files are form definitions, report definitions and stored queries. The migration of old database files is also a straightforward task as modern database programs can also read a range of formats. Database files in the Newham archive were mainly in dBase format with a few Paradox and MS Access files, all of which could be read and thus migrated to ADS dissemination and preservation formats.

File Category	Percent
Survey Data	26.2
Reports and Report Drafts	22.3
Unknown format / no value	14.2
Administration / Project Management	11.0
Correspondence	10.8
Data and Databases	10.8
Project Designs and Tenders	1.7
Publication Articles	0.7
Manuals and recording templates	0.5

Table 1: Proportions of files within functional categories.

File type	Percent
Word Processor	35.1
Survey Data	26.3
Database files	17.2
Spreadsheets	10.2
Graphics	7.3
Project Management	2.9
Unknown	1.0

Table 2: Proportions of files created by software packages.

A much greater problem results from files, mainly databases, for which there is no site information. This left the data orphaned and therefore of very little value. In total about 250 files in the archive are orphaned.

Another problem arises from databases that contain coded information. In some cases the coding may be standard or easily recognisable by other specialists in the same field but there is one large database of inhumations from a monastic site in which the bones, all except for a lone patella, are coded as numbers with no explanation of the coding system. There are similar examples throughout the Newham archive, including pottery, building materials and environmental databases. This again clearly emphasised the need for effective documentation of the archive and all of its component parts before deposit. Unfortunately this could not be done with the Newham archive and in essence the data contained in these files has been lost to the archaeological community.

Spreadsheets in the archive mainly contain archaeological data tables, although they were also used in project management. DOS versions of Supercalc and QuatroPro were used before the Newham Museum Service adopted Windows and MS Excel. Unfortunately the two DOS programs are not widely supported by modern spreadsheets. A package from JASC software called Quick View Plus proved useful here. This program can view an impressive range of file formats including word-processor, spreadsheet, database, and graphics. The contents of the spreadsheets can be viewed and transferred to modern programs using copy and paste.

Many of the graphics files were created from survey data using Golden Software's Surfer and MapViewer programs. Unfortunately these packages only create files in their own proprietary formats which are not supported by other software packages. A modern version of Surfer is required to migrate these files to ADS archive standard DXF formats. This migration problem is even more acute with the CAD files created using TurboCAD 1.9. Again this is a format which is not supported by other packages nor, indeed, by modern versions of TurboCAD and thus presents major problems for migration. Effectively the functionality of the software itself will have to be preserved to enable the reuse of these files in the future. This is not a recommended archiving procedure for archaeological data (Richards and Robinson 2000) but is necessary in this case.

The Newham digital archive is probably a typical example of digital information resources of archaeological units. There are probably many archaeology units with archives of files in redundant formats, without explicit site relevance information, containing unexplained coding, and in unknown states of completion. The files may also be stored on unsuitable media in poor storage conditions. In short there may be large amounts of archived archaeological information, which can never be accessed again.

2. The Digital Archiving Pilot Project (DAPPER)

The Newham Museum Service digital archive is a depressing and salutary tale. The archive was developed as a working tool to help the Service write up and managed its archaeological projects, in this respect the archive was fit for the purpose for which it was developed. Equally the concept of digital project archiving was still in its infancy when the archive was developed and as there were no published strategies or methodologies in place to ensure the effective preservation of the data, poor condition of the archive is understandable. At a national level the ADS together with English Heritage, the Association for Local Government Archaeologists and the former Royal Commission for the Historic Monuments of England had begun a series of discussions out of which emerged the English Heritage funded Digital Archiving Pilot Project for Excavation Records (DAPPER). DAPPER aimed to demonstrate the concept of digital excavation archives in order to inform the development of best practice in this emerging field. The pilot project also aimed to encourage the re-use of these digital repositories and to explore the use of the Internet to enhance access to archaeological data. Consequently DAPPER also investigated ways to deliver the data to a potential re-user community via the Internet and in the process set up the world's first online digital excavation archive.

2.1. The excavations and their archives

In order to demonstrate the concept and potential of this new form of information provision it was necessary to choose the excavations to archive with care. After a lengthy consultation period two reasonably large and "high profile" sites were identified as being most suitable: the Royal Opera House excavation by the Museum of London Archaeology Service and Eynsham Abbey, excavated by the Oxford Archaeological Unit.

2.1.1. Eynsham Abbey

Between 1989 and 1992 the Oxford Archaeological Unit (OAU) undertook a full excavation of part of the precinct of the medieval abbey, revealing part of the cloister and south ranges, kitchens and domestic buildings. Evidence for earlier activity included the 11th century Anglo-Saxon abbey and earlier Minster, and a Bronze Age enclosure. The project was never designed as an exercise in digital data collection or management, but rapid changes in computer technology during post-excavation have meant that the project developed in this direction.

The downloadable archive contains text files, a database made available as comma delimited text, JPEG images, a 3D reconstruction of the medieval abbey available in DWG and DXF formats (i.e. AutoCAD) and numbers of digitised drawings also available as DWG, DXF and DWF files. All of the files in the Eynsham archive are also archived in ADS preservation file formats, for example, the preservation images are stored as high resolution uncompressed TIFF's.

The contents of the Eynsham digital archive can be summarised as follows:

Dataset documentation	Project background
	Excavation – aims, methodology
	Post-excavation – phasing, structures
	Specialist report summaries with database documentation and specialist bibliographies
Text files	Project documentation
Delimited text files	Architectural stone database
	Context database
	Context database keywords
	Glass lead cames database
	Vessel glass database
	Window glass database

	Lead objects database
	Post-Roman pottery database
	Environmental sample register
	Environmental sampling results register
	Small finds database
	Parallels for small finds database
	Structures database
	Tile database
Drawing files	Sections
	Phase plans
	Structure plans
	Site plans
	3-D reconstruction
Images	

2.1.2. The Royal Opera House

Following exploratory work in 1995 full-scale excavations were undertaken in 1996 by MoLAS at the site of the Royal Opera House in Covent Garden, Greater London. The excavation examined the largest area of the Saxon trading port of *Lundenwic* so far to become available and has provided a wealth of information on the form and economy of this settlement. All data from the Royal Opera House project was collected in computer based formats. This excavation was the first for which MoLAS used its state-of-the-art integrated database and GIS recording system. The online archive is currently incomplete but contains Geographical Information System (GIS) files, interpretive groupings and data files consisting of context, artefact and ecofact attribute sets. The rest of the archive will become available following publication of the site, which was projected for March 2000, and will be deposited with the ADS as part of an extension to the current project.

The contents of the Royal Opera House digital archive can be summarised as follows:

Text files	Group descriptions
Delimited text files	Field records by context
	Basic interpretation
	Roman pottery
	Post-Roman pottery
	Saxon pottery fabrics
	Registered finds
	Coins
	Loom weights
	Expansion codes for registered finds
	Assessment level animal bone
	Post assessment animal bone
	Animal Bone codes
	Animal species expansion codes
	Processed environmental samples
	Botanical remains – analysis
	Botanical expansion codes
ArcView themes	Context groups
	Trench edges
	Unreal edges

It is important to note that in neither case was the digital archive planned as part of the dissemination phase of the project and that both Eynsham Abbey and the Royal Opera House will be published as traditional monographs. The resultant digital archives are, however, remarkably similar in their content. This is due to their origin as the residues of large-scale, well-funded post excavation projects. Both archives are comprised of essentially the digital residues of the post-excavation process and represent a digital version of the "research archive" (cf. Richards and Robinson 2000).

2.2. The time taken to create the archives

The Newham Museum archive was hurriedly copied from the data server or individual hard drives before it was transferred to the ADS. Consequently the DAPPER project strove to avoid this problem through allowing the depositing units the time to adequately document their archive and to normalise their data.

The data selection and normalisation of the Eynsham Abbey excavation archive took almost twice as long as the Royal Opera House archive. This discrepancy in time is undoubtedly due to the fact that the Eynsham Abbey project was never designed as an exercise in digital data collection and management but had evolved into a digital format. Consequently much of the time allotted to the OAU was taken up in converting and normalising the numerous disparate databases into a single database structure and in converting the CAD files from the proprietary FastCAD to the DXF, CAD exchange file format. The Royal Opera House, however, was always envisaged as a "digital" project and standard file formats and software was used for both data capture and file management. Cost-effective archiving consequently can be encouraged through the use of non-proprietary software and exchangeable data formats. This is outlined in the ADS Guidelines for Depositors (ADS 1999), which is available online: http://ads.ahds.ac.uk/project/userinfo/deposit.html.

2.3. Documenting the archives

One of the aims of the DAPPER project was to inform the development of digital archiving practice. In this respect the project has been a success as it has enabled all of the participants to consider the types of documentation an excavation archive needed to facilitate its effective re-use. These recommendations will be published in the second edition of the Excavation Archiving Guide to Good Practice (Richards and Robinson 2000).

The experiences gained from working in partnership with MoLAS and the OAU have enabled the ADS to specifically identify:

- what documentation is necessary,
- the documentation already produced as part of the post-excavation process that should be included in the project level documentation, such as the site report, specialist reports and so on,
- the new forms of documentation that units need to write to accompany the archive; here the documentation of the digital archive is particularly important, if the codes used in the Newham databases, or if all the files for a particular site had been documented, a substantial proportion of the archive would today be available for reuse.

Much of this documentation can be adequately provided through the collation and editing of reports that would be written during the post-excavation or publication stages of a project. Although some additional writing is necessary and the preparation of project documentation would not represent a substantial additional burden for the units.

2.4. The cost of digital archiving

Digital archiving entails expenditure, both in the setting up and running of a digital archive and in preparing, depositing, accessioning and curating the information resources. If archiving of the digital component of fieldwork is to become a part of the everyday practice of archaeology, DAPPER also had to demonstrate the commercial case for such archives. The pilot has allowed us to quantify the effort and costs to both units and archives in the preservation and dissemination of computer-based data.

The experiences gained from DAPPER have proved invaluable in the development of the ADS charging policy (ADS 2000). More information on the charging policy, including a detailed breakdown of the costs and charging categories is available online: http://ads.ahds.ac.uk/project/userinfo/charging2.html

The central tenets of this policy are that:

- the use of ADS resources will be supported free of charge,
- archiving costs should be recovered from the body funding the archaeological investigation, or research, wherever possible,
- this would be a one-off payment collected at the point of deposit which should safeguard the long term future of the digital information.

Eynsham Abbey

Accessioning and migration

CAD	882 files	£4810
Images	24 files	£510
Database	13 data tables	£325
Spreadsheet	2 data tables	£50
Texts	2 documents	£50

Storage

CAD	477 Mb	
Images	998 Mb	
Database	2.8 Mb	
Spreadsheet	0.2 Mb	
Texts	0.1 Mb	
Total	1478.1 Mb	£739.05

Digital Archiving cost of Eynsham Abbey	£6484.05
Cost of Eynsham Abbey Project (total)	£540,000
Proportion of budget spent on digital archiving	1.2%

Royal Opera House

Accessioning and migration

GIS	3 files	£75
Database	3 data tables	£75
Spreadsheet	17 data tables	£455
Texts	2 documents	£50

Storage

CAD	4 Mb	
Database	0.1 Mb	
Spreadsheet	1.4 Mb	
Texts	0.8 Mb	
Total	6.3 Mb	£0

Digital Archiving cost of ROH Phase 1	£655
Cost of Royal Opera House Project (total)	£650,000
Proportion of budget spent on digital archiving	0.1%

The Eynsham Abbey and Royal Opera House digital archives are both essentially research archives. Yet they differ quite markedly in their cost. This is clearly due to the ways in which both sites were analysed during post-excavation. Although the stratigraphic drawings of both sites are available digitally, the use of GIS by MoLAS ensured a deposit of 3 files, as opposed to the 404 Eynsham Abbey site plan CAD files. The 80 structure plans, 15 phase plans, 2 sections, a trench plan, a composite plan and a 3-D reconstruction of the Abbey also swelled the number of Eynsham Abbey CAD files. Consequently the cost of archiving is closely related to the presentation of the drawing information. It should not be assumed, however, that GIS is preferable to CAD because it is more cost effective to archive. Both file formats have their own distinct advantages and disadvantages. The GIS structure of the Royal Opera House archive, for example, would require specialist software and training to facilitate effective understanding and reuse of the data. It is consequently difficult for a user with little knowledge of GIS to use the archive. On the other hand the Eynsham Abbey archive is structured like a traditional archaeological publication, with its separate structure and phase plans. With appropriate viewers, the CAD files can be accessed via the Internet. Consequently the CAD presentation of the data can open up access to the information for non-expert users. Yet for highly experienced users the MoLAS GIS offers a very powerful analytical tool, with interpretative functionality that is beyond CAD. It is recommended that units continue to undertake their post-excavation in the ways most appropriate to them and that the production of a comprehensive digital archive should come before matters of cost.

In its Charging Policy the ADS state that “it is usual for digital archiving costs to add an overhead of less than 5% to the total project budget (ADS 2000). DAPPER has illustrated that for projects of the scale of Eynsham Abbey and the Royal Opera House, the actual cost of digital archiving can be as little as 1% of the total project budget. It must be recognised, however, that this 1% figure relates only to large-scale excavation projects. For medium to smaller-scale projects with a variety of files, the actual cost of archiving will be represent a greater proportion of the total budget and be in the region of 3-5% of the total project budget.

2.5. Resource delivery

Both the DAPPER and Newham Museum Service archives are disseminated over the Internet via ArchSearch, the online catalogue of the ADS. ArchSearch now contains some 350,000 site index records based on the National Monuments Records of England, Scotland and Northern Ireland, regional sites and monuments records and localised gazetteers. This site-based index is increasingly being fleshed out through the deposition of resources such as the DAPPER and Newham Museum Service archives. An objects based model has been adopted to conceptualise and implement access to these resources. Active hyperlinks within the database holding the catalogue provide access to resources whose object types or classes are defined in terms of generic structure and functionality. Effectively, an objects based data structure has been bolted onto an underlying relational database (cf. Austin and Clarke 2000 for a fuller description).

One of the major data delivery issues concerned the user interface and revolved around the issue of whether a complex, visually pleasing, heavily packaged interface should be developed, or whether the raw data should be simply presented? The heavily packaged interface would be user friendly and consequently may open up access to the data. This approach, however, was rejected as it would have been prohibitively expensive to produce, would have set a precedent for all subsequent deposits and the heavily packaged nature of the data would probably present migration problems. More importantly perhaps, the packaging of data may restrict its potential for reinterpretation. Consequently it was decided to present the data in standard formats with sufficient on-line support documentation to enable the re-use of the data. This was regarded as the ideal, sustainable cost-effective solution for Research Level Archives and below.

The resource delivery issues brought the idea of who the data is aimed at into focus. The ADS receives its core funding from the Higher Education sector and consequently it is the scholarly community at which its data resources are aimed. Equally the ADS user community, as reflected in Strategies for Digital Data (Condron et al. 1999), clearly wanted raw, rather than packaged data. The simple, unpackaged DAPPER interface consequently reflects the professional archaeological community's desire for raw data.

2.6. DAPPER conclusions

The Digital Archiving Pilot Project was a necessary and timely project that enabled the Archaeology Data Service to develop and implement the world's first online digital project archive. The project also enabled its partners to think about exactly what forms of data should be incorporated into a digital project archive and how this should be documented and then delivered online. DAPPER provided the necessary funding and time for the Museum of London Archaeological Service and the Oxford Archaeological Unit to prepare their data for deposit and to reassess and develop their post-excavation procedures and documentation. Finally a critical review of DAPPER has allowed the ADS and English Heritage to consider how to ensure how other quality digital archives will be delivered to the archaeological community in the future.

DAPPER has demonstrated that an effective digital archive can be delivered for a fraction of the total project cost. For substantial fieldwork projects the DAPPER project suggests that the estimated archive deposition cost (approximately 5% of the total project budget) as reported in the ADS Charging Policy (ADS 2000) can be modified downwards, to less than 1% of the total project budget.

The results so far from DAPPER have been very encouraging. The Royal Opera House archive was released in August, followed by Eynsham Abbey in October 1999. Following the launch of the excavation archives the ADS website has received an increased number of visitors and has experienced successive months of record-breaking user figures. Indeed within the first four months of their release over 8,000 people have visited the project archives and either accessed or downloaded approximately 10,000 separate files within the archives.

When thinking about digital archiving, a more philosophical point came to mind; what is a digital archive for?

- a repository for the digital remains of project data?
- part of the project's dissemination strategy?

If a digital archive is simply a store of the digital products of archaeological projects, DAPPER has effectively delivered this form of archive at the research level and has also illustrated that it could also be implemented for lesser archives. The digital archive has the potential to develop into something more than this and an integrated archive could be part of a project's dissemination strategy. The integrated digital archive provides the opportunity to take unwieldy technical data out of the excavation report and locate it in a universally accessible environment. Modern archaeological datasets often do not translate well into the written reports, especially if much of the information is in the form of CAD drawings, GIS or relational databases, which loose their functionality when removed from their digital format. Consequently the digital archive offers the opportunity to place data in a facility that actively encourages its reuse and reinterpretation. If project data were located in a digital archive this would have profound implications for the ways in which archaeological research is published. If the correct linkages can be developed between the traditional paper or Internet publication and the digital archive this would enable the final project report to be slimmer and more synthetic and less bogged down in detailed description.

3. Conclusions

This paper has effectively illustrated through the use of DAPPER that a digital archive can be easily created from the digital residues of the excavation and post-excavation process where there is sufficient political will and funding. On the other hand Newham has demonstrated the opposite, that where the digital archive has never been considered, when the digital residues of the archaeological process are simply left to degrade, the reverse engineering of excavation archives can be an expensive and frustrating business, especially where documentation is lacking and file formats are obscure. It is becoming widely accepted amongst the archaeological community that the digital component of a project is a necessary and vital component of the primary record of a site and needs to be preserved into perpetuity. What DAPPER and Newham have illustrated is that the road to a successful deposit should ideally begin at the point of data creation and continue throughout the life cycle of the project until the point of its ultimate deposition with the digital archive. Only in this way can we truly ensure that through documentation, the use of appropriate file formats and by following adequate backing up and migration regimes that we can begin to create a digital future for our excavated past.

Acknowledgements

The authors wish to thank our colleagues at the ADS, Jo Clarke, William Kilbride, Maureen Poulton and Julian Richards for their efforts on behalf of this paper and on the DAPPER and Newham Museum projects.

Damian Robinson would like to thank Keith May and Tim Williams of English Heritage Archaeology Commissions, Ann Dodd and Alan Hardy of the Oxford Archaeology Unit and Gordon Malcolm of the Museum of London Archaeology Service for their contributions to the DAPPER project.

Keith Westcott would like to thank Sarah Richards and David Thorne, students on the MSc in Archaeological Information Sys-

tems at the University of York for their help with the Newham Museum archive.

References

ADS 1999. Guidelines for Depositors Version 1.1, http://ads.ahds.ac.uk/project/userinfo/deposit.html

ADS 2000. Charging Policy, http://ads.ahds.ac.uk/project/userinfo/charging2.html

AUSTIN, T., 1998. Tip of the iceberg in a digital sea: scale of the academic digital resource, *ADS Newsletter* 4: 4-6. Also on-line at http://ads.ahds.ac.uk/newsletter/issue4.html (18 June 2000).

AUSTIN, T., 1999. Rescue archaeology at the ADS, *ADS Newsletter* 6. Also on-line at http://ads.ahds.ac.uk/newsletter/issue6.html.

AUSTIN, T. and CLARKE, J., 2000. ArchSearch: under the bonnet, *Archaeological Computing Newsletter* 55: 15-20.

BEAGRIE, N. and GREENSTEIN, D., 1998. Managing Digital Collections: AHDS Policies, Standards and Practices: Version 1, on-line at http://ahds.ac.uk/public/srg.html.

BOOTH, B., GRANT, S. and RICHARDS, J., 1989. *Computer Usage in British Archaeology second edition 1989*, IFA Occasional Paper 3 (Birmingham).

CONDRON, F., RICHARDS, J., ROBINSON, D. and WISE, A., 1999 Strategies for Digital Data – Findings and Recommendations from Digital Data in Archaeology: A Survey of User Needs, The Archaeology Data Service. York.

DINGWALL, L., EXON, S., GAFFNEY, V., LAFLIN, S., van LEUSEN, M. (eds.), 1999. *Archaeology in the Age of the Internet*, Computer Applications and Quantitative Methods in Archaeology CAA97, BAR International Series 750 (Oxford).

GRANT, S., 1986. "Summary and recommendations". In Richards 1986: 13-32.

RICHARDS, J. and ROBINSON, D., 2000. *Digital Archives from Excavation and Fieldwork: A Guide to Good Practice Second Edition*, Oxbow Books. Oxford.

RICHARDS, J. (ed.), 1986. *Computer usage in British Archaeology*, IFA Occasional Paper 1, Birmingham.

STEWART, J., 1980. Integrated excavation and museum recording systems: methods, theories and problems, *Museum Archaeologist* 5: 11-27.

WILCOCK, J., 1999. Getting the best fit? 25 years of statistical techniques in Archaeology. In Dingwall et al. 1999: 35-51.