Centre for Archaeology Report 78/2004

Revelation : Phase 1 Assessment

Sarah May, Brian Attewell, Paul Cripps, Vicky Crosby, Thomas Cromwell, Karla Graham, Jen Heathcote, Claire Jones, Eddie Lyons, Keith May, Andy Payne, Sarah Reilly, David Robinson, Kirsty Stonell Walker, Jörn Schuster and Mary Walkden

ISSN 1473-9224

The Centre for Archaeology Report Series incorporates the former Ancient Monuments Laboratory Report Series. Copies of Ancient Monuments Laboratory Reports will continue to be available from the Centre for Archaeology (see back cover for details).

Centre for Archaeology Report 78/2004

Revelation : Phase 1 Assessment

Sarah May, Brian Attewell, Paul Cripps, Vicky Crosby,
Thomas Cromwell, Karla Graham, Jen Heathcote, Claire Jones,
Eddie Lyons, Keith May, Andy Payne, Sarah Reilly, David Robinson,
Kirsty Stonell Walker, Jörn Schuster and Mary Walkden

Summary

"Revelation" is an English Heritage project to provide a coherent digital information system that will make the capture, analysis and dissemination of Historic Environment research faster and more effective. Stage 1, described here, was a comprehensive review of information systems and work practice at the Centre for Archaeology in the context of the broader profession.

This report constitutes an IS-related business analysis of the CfA to assess the scope of the requirement. The aim of this stage has been to define how we use data throughout the life of an archaeological project so that, in the next phase of Revelation, we can design and build an archaeological information system that will be used by all the CfA.

The Methods section covers the main strands of research and the project working practice more generally. The Results section relates to the main reviews of the project. A discussion of issues arising from the research and not directly covered in the reports follows. The Conclusions highlight key findings and directions for future work. The Recommendations section is directed towards five audiences: EH Senior Management, CfA Management Group, CfA Staff, Systems Development, and the Sector.

Keywords
Information systems
Work practise
Data management
Digital data

Author's address
English Heritage Centre for Archaeology, Fort Cumberland, Fort Cumberland Road, Eastney, Portsmouth PO4 9LD

Revelation: Phase 1 Assessment

Version 3.1

16 February 2004

The Revelation Project Team

"The Holy Grail of unit computing is the Integrated Information System where information flows seamlessly from excavation, through post-excavation to publication and archive, offering an efficient process that would give a competitive edge to any organisation managing to achieve it. In theory this is possible, as any consultant will tell you, and indeed different levels of success have been claimed (Rains 1995, Beck 2000 for example) although the real picture is more likely to be one of ad hoc development within an environment of under resourcing, a lack of expertise and intense time constraints".

(Lock 2003 p 265)

Executive Summary

"Revelation" is a English Heritage Centre for Archaeology (CfA) project to provide a coherent digital information system that will make the capture, analysis and dissemination of CfA research faster and more effective. Stage 1, described here, was a comprehensive review of information systems and work practice at the CfA in the context of the broader profession. The purpose of the work is to support and improve current research practice at the CfA.

The project is part of the CfA's commitment to the Modernising English Heritage agenda, as it will increase our capacity to deliver advice, support and expertise on the development and preservation of the Historic Environment. The project builds on a number of other Modernisation initiatives already active in the organisation.

Corporately, HEIST , Tech Refresh and the English Heritage Research Strategy have established frameworks for realising the benefits of research data as a corporate asset. Within the Archaeology Department the project responds to the developing strategy on digital dissemination to facilitate access for a wider range of audiences. The CfA Digital Archiving Strategy, (Brown 2000), laid the foundation for considering how the way we create and hold our data affects the way that we use it, and how it can be used by others.

This report constitutes an IS-related business analysis of the CfA to assess the scope of the requirement. The aim of this stage has been to define how we use data throughout the life of an archaeological project so that, in the next phase of Revelation, we can design and build an archaeological information system that will be used by all the CfA.

Recommendations for English Heritage Senior Management focus on strategic management changes and the support required for future stages of the project succeed. Those for CfA Management Group include procedural and cultural reviews as well as the prioritisation of the project within the Team plan, including training and development. All CfA Staff are recommended to play an active role in future stages of the project and in the cultural changes that will go with it. Recommendations for the Sector pass on the lessons learned in our review of Sectoral Practice.

One of the recommendations for CfA Management Group is to establish a Revelation Programme board. Systems Development recommendations are addressed to this board with the aim of building a system based on a spatially-indexed database with appropriate interfaces for spatial, graphic, numeric and textual entry, manipulation, and dissemination.

English Heritage, as sectoral champion, has a responsibility to build capacity in areas critical for the enhancement and protection of the Historic Environment. The importance of developing high quality digital resources for archaeological research is recognised in the strategic initiatives described above. The scale of work involved in the Revelation project is not usually possible in commercial archaeological units. This project is a real opportunity to provide leadership, advice and assistance to the archaeological sector in a fundamental area, which requires long term commitment to research and development.

1 Introduction

The Centre for Archaeology (CfA) provides in-house expertise for English Heritage (EH) on all aspects of field archaeology and archaeological science. We also collaborate with colleagues in other sections of the Archaeology Department and EH Regions to provide broad-ranging research and strategic support on the Historic Environment. This produces a vast amount of archaeological data in a mixture of digital and traditional formats (more detail on the business context of research in the CfA is found on page 32)

To meet our obligations to efficiency and accessibility as part of Modernisation a coherent digital information system (IS) is required to collect, manage, analyse, disseminate and archive these data. The current systems do not fully meet our needs locally, corporately or within the sector, partly because they have been developed individually over time for particular circumstances.

The Revelation project was established to investigate the use of information systems in the research activities of the CfA and recommend improvements. The most time consuming and challenging part of this work has been defining the processes that IS support, and the work practices and procedures that form the human contributions to the current systems. CfA Management have provided substantial resources for the assessment and have shown the courage to allow a group of staff to review work processes in such detail.

1.1 Purpose

The principal purpose of the project is to improve our internal processes for archaeological research. This will speed analysis and dissemination, allowing us to meet our strategic goals more efficiently. Our aim has been to understand how we use data throughout the life of an archaeological project so that we can design an information system that will be used by all of the CfA. In the process of this assessment we have had the chance to reconsider how our working practice feeds our understanding, and how it can be supported and improved by better designed systems.

We hope to develop an information system that:

- Clarifies and supports the relationships between the different functional and project teams which form the CfA, and integrates their work in the corporate context of EH.
- Speeds up extended work programmes where individual member's input varies and key people's time is limited.
- Facilitates integrated working between dispersed teams of specialists, all of whom have responsibilities to their sub-disciplines as well as to the team.
- Builds a data library for the archaeological research of the CfA, particularly enhancing the understanding of those sites we have statutory or commercial responsibility for.
- Facilitates digital dissemination of archaeological information to a range of audiences.

- Allows for broad scale comparative work to examine detailed research questions on the condition of the archaeological resource in relation to issues such as agricultural damage, erosion and climate change.

1.2 Scope

The project will run in two stages. Our aim in this stage has been to understand how we use data throughout the life of an archaeological project so that, in the next phase of Revelation, we can design the desired information system. Given the interdependencies between work processes and system design, this assessment has focussed on what is often referred to as 'business analysis'.

We comment on a wide range of issues regarding work processes but we do not expect the Revelation project to solve all the problems it identifies. In fact, we identify many problems solely because they must be dealt with before a successful information system can be implemented.

The design of any system should be based on process and we need to support the entire process of archaeological projects from fieldwork to dissemination and archive. Therefore, we have considered all aspects of our work in the field and in the lab, throughout the life of a project. We have conducted reviews of existing systems, sectoral practice, field practice, including an investigation of digital drawing in the field and, finally, a two-stage survey of user needs.

We have examined practice and procedure as an integral part of systems. All information systems are composed of data stores and processes that control the creation, manipulation and movement of those data. In a manual information system, like a reference collection, both the data and its storage space are physical objects. Procedure and practice define all the processes in a manual system. In a digital information system, the data stores are digital and the processes can be augmented and automated with software and hardware. So there is a trinity: procedures/practice, software and hardware. In this initial stage of work we have focussed more heavily on procedures/practice in order to determine how software and hardware can support work process.

Although our aim has been to support and improve current practice, the system should be designed with completion and dissemination in mind. Our outputs are diversifying, with a greater emphasis on digital and popular dissemination and this imposes requirements on the rest of our project work. Revelation will enhance archiving of the data and facilitate both paper and electronic dissemination strategies.

Just as we use a common recording strategy, the information system should be applicable to all of our work, and provide a model that can be used by other archaeological practitioners. While individual customisation may be necessary for specific projects, we aim to integrate the CfA systems for standard practice.

Although the principal purpose of the project is to improve our internal processes, it also has the potential to contribute to the development of broader archaeological practice. A section for the sector has been included in our recommendations (p. 50 ff.).

1.3 Context

English Heritage is currently not at the cutting edge of this field. The use of databases for post-excavation is a fairly standard part of current professional practice in archaeology. For some units this is a fairly *ad hoc* process, others have more coherent systems. Almost all of these systems replicate paper records, which are still the main data collection format. The Landscape Research Centre has been pioneering the use of computers in the field for the last decade, and EH has been involved in this through the DigIT project. A few units have integrated GIS into their systems, but most systems take 'the context sheet' to be their spine. Many enter data during the excavation in site offices, and a handful of units are beginning to enter data directly in the field.

The CfA and its predecessors, have explored, developed and relied on digital recording and analysis for a range of projects over the last 25 years. For example, the use of survey equipment for planning and 3-D recording at Battle Abbey, and the development of the Delilah database for contextual information, were both groundbreaking pieces of research. We have also developed a range of cutting-edge scientific information systems. Some of these are analytical, such as the systems used for XRF analysis and the analysis of geophysical data. Others are for data management, such as Labfile and ABCD. Projects such as the Raunds Iron Age and Romano-British project and the Windsor Castle excavations have used digital methods extensively in the phasing, assessment and analysis of stratigraphy and building fabric. The Revelation project builds on this experience, especially in Stage 1, and an outcome of Stage 2 will be that EH is a leader in this field once more.

We have already provided that leadership through the Digital Archiving Strategy (Brown 2000), which is accepted as a model of good practice (Richards and Robinson 2000) and contains detailed procedures and standards. This is an evolving document, and will certainly need to change as our practices change.

The system must work in harmony with established data frameworks and standards. English Heritage is actively engaged in promoting digital dissemination of archives as part of the Archaeology Commissions programme. We take as an axiom that field practice is the first stage of our publication and dissemination strategy. If we encourage others to seek new methods of dissemination we must also change our practices to suit these new end products.

1.4 Team

It was important to ensure that this assessment of IS was user-led. We were particularly concerned that the team not be dominated by staff already involved in IS development. The project team consisted of people from twelve specialist areas of the CfA and one member of the Archaeology Department

Strategy Group. All are co-authors of this report and team working was very strong within the project. Smaller groups and sometimes individuals were given responsibility for specific aspects of research and reporting and are credited in the relevant appendices. Regular team meetings and workshops ensured that the results reflect experience across the CfA. Further, all CfA staff have contributed substantially to this project through meetings, interviews and questionnaires.

1.5 Structure of this report

Like all research outputs, this report summarises a huge body of data. The Methods section covers the main strands of research and the project working practice more generally. The Research Results section relates to the main reviews of the project. A Discussion of issues arising from the research and not directly covered in the reports follows. This is mainly concerned with setting the research in a context within the CfA. The Conclusions highlight key findings and directions for future work. The Recommendations section is directed towards five audiences: EH Senior Management, CfA Management Group, CfA Staff, Systems Development, and the Sector. There is a Glossary of all technical terms and acronyms and a Bibliography.

We have managed the records of this project digitally, and the holdings of this digital archive are described in Appendix 3. This includes databases, internal reports, GIS and AutoCAD files, conference papers and presentations, as well as minutes of seminars, workshops and meetings. All of the arguments in the report are based on these data. We would like to be able to disseminate this archive digitally, but our current systems will not support it. We are, however, happy to make the archive available to interested parties, subject to the constraints of the Data Protection Act, and we will look at methods for disseminating more detail as part of planning for later stages of the project.

2 Methods

This assessment report is derived from a number of other reports, workshops, team meetings, discussion groups, interviews, fact-finding visits, field trials, literature researches, conference papers, surveys and seminars. The following section sets out the most substantive methods used during the assessment and in particular those that resulted in the production of specific reports (see list of List of Appendices).

2.1 Review of Existing Systems

The Review of Existing Systems (Appendix 1.A) was undertaken between November 2002 and February 2003. It involved the entire sixty-plus staff working for the CfA: in the field, in Savile Row and at Fort Cumberland and aimed to ascertain:

- what systems the CfA currently has and uses, for what purposes and by whom;
- how information for these is gathered, structured, processed and disseminated;
- if and how this information interfaces with other CfA, EH and external systems;
- the expected lifetime of the requirement for the systems and of the systems themselves.

The review consisted of:

The CfA elements of two existing surveys carried out by EH: the Year 2000 (Y2K) exercise, checking on EH software compliance, and the Redundant Software Survey (2001), intended to rationalise and modernise the EH software 'estate' around its transfer to Microsoft Office 97 from Lotus and Word Perfect.

Adrian Brown's (CfA Information Manager 2000-2002) Digital Archiving Project (DAP) internal consultation report, produced in 2000. The Digital Archiving Project's main aim was *"to develop a comprehensive, integrated, and achievable digital archiving strategy for the Centre for Archaeology (CfA), within the context of the evolving corporate policy."* The first project stage asked for much information relevant to the review : *"…undertake an internal consultation exercise with a representative sample of CfA staff, chosen to include all constituencies of data creators and users"* (Brown 2000).

A full scan of Fort Cumberland networked storage looked for files that were likely to represent existing systems. This scan aimed to find a number of systems on the server that were missed by the other surveys; either genuinely 'orphaned' systems where the owner no longer worked for the CfA and our procedures failed to document them, or forgotten systems ready to be reintroduced to the owner!

The results from a questionnaire, presented to all CfA staff by members of the Revelation Project Team. After introducing the project and the role of the survey a questionnaire was completed for each system identified by the interviewee(s). 146 questionnaires were completed and loaded into a simple

Microsoft Access database. The responses to each of the questions were looked at together and an analytical version of the responses produced for each of the 109 different perceived systems.

The production of formal Data-Flow Diagrams based on the questionnaire responses and subsequent systems analysis. Diagrams were produced for each of the ten CfA teams plus one for the CfA Raunds Project Team and one for the production of CfA reports.

2.2 Field Practice

Initially only a small-scale, targeted piece of fieldwork was planned to assess available handheld digital drawing systems (Appendix 1B). Because the original project timetable was extended, an opportunity arose to broaden the assessment to verify the CfA data-flow model during the summer fieldwork season, resulting in the Report on Current Field Practice (Appendix 1C).

2.2.1 Handheld Drawing Systems (Appendix 1.B)

This small-scale fieldwork trial was focused upon testing software capabilities rather than the suitability of specific hardware platforms for handheld field drawing. The primary purpose of the Assessment of Handheld Drawing was to test the success and limitations of digital drawing methods using a number of software packages on available handheld hardware platforms.

The main methods employed in the assessment were:

A number of short 2-3 day field trials were carried out based around an ongoing CfA fieldwork project at Whitby. Practical experience was gained in recording examples of excavated archaeology using digital drawing on site. The recording comprised digital planning; digital sketch plans and sketch sections; and, photo rectification to enable annotating photographic records either on or off site. Staff on site also provided feedback as to the feasibility of developing handheld digital drawing methods for the CfA excavation process.

A number of existing platforms were tested that have been used elsewhere and which are currently available to the CfA "*off-the-shelf*." Separate reports were produced on the outcomes.

An Assessment report was produced including an overview of recent developments and current practice in digital drawing. The Whitby fieldwork results were incorporated along with experiences from a number of other recent EH fieldwork activities (including the 2001 DigIT project and 2002 fieldwork on Dartmoor) within the context of the wider Revelation project.

2.2.2 Current Field Practice (Appendix 1.C)

The Report on Current Field Practice is based upon participant observation at the CfA excavations at Barrow Clump, Figheldean, carried out as part of the Badger Damaged Barrows Project in September and October 2003.

The aims were:

- to compare the model of the CfA field data flow (derived from work on the Review of Existing Systems) with actual data flow in the field.

- to describe how current practice aids on-site interpretation.
- to comment on the trial use of digital indexing, especially relating to finds recording.

These aims were achieved through the following:

Field observation and participation in the excavation as a site assistant. Information was also derived from asking questions of, and discussing issues with, the other team members. Longer more structured discussions were carried out with several members of the excavation team in which specific aspects of the data-flow model were considered. These were used to produce a table of notes on specific data flows and more general comments on the recording process.

The pilot digital indexing system comprised the indexes for context, small finds, samples, skeletons, drawings and photos in an Access 97 database. Preparation and creation of some of the basic tables took place before the fieldwork started with advice from project Finds Officer and Environmental Officer on how they required their information to be presented, and what information would be useful during the excavation and in the immediate post-excavation phase. Much of the database development took place during the excavation - in particular the forms, reports and standard queries - through discussions with the on-site Finds and Environmental Assistant and the Project Manager. This development work was part of the excavation project and not part of Revelation. The value of the indexes was assessed by monitoring their use throughout the fieldwork.

2.3 Sectoral Practice

The purpose of the assessment of sectoral practice was to find out whether and how other archaeological units and organisations are already using IT to enable an integrated approach towards data handling – from retrieval of data at excavation through to dissemination and publication. The resulting report (Sectoral Practice Report, Appendix 1.D) demonstrates the areas where we can rely on other people's experience and those where we will have to undertake substantial development ourselves. It also points out some of the difficulties that might occur in the course of such a process and which have kindly been indicated to us by our colleagues.

In order to achieve the aims laid out above, a tripartite approach was taken to gather relevant information. This included:

- meetings and fact-finding visits, organised with a range of archaeological organisations or individuals in England who are widely acknowledged to have experience with well-tried and tested IT systems, or are using cutting-edge IT;
- a list of relevant software recording the publicly and commercially available software from the visits, a literature search of 1500 books and articles and comprehensive web searches;
- a literature review bringing together our thoughts on 100 different papers under 11 headings derived from keywords in the database: Work Practice and Knowledge Generation, Digital Dissemination, Archiving Systems,

Data Flow, Database Design, CAD and Digital Drawing, Digital Recording, GIS in Archaeological Fieldwork, Post-ex Use of Contextual Data, Analytical Systems and Integrating Different Data Sets.

The Sectoral Practice Report summarises these elements into a brief, readable assessment of the use of IS in the archaeological fieldwork process (excavation-analysis-dissemination), and provides a context for the other aspects of the Revelation project.

2.4 User Needs

The assessment of user needs was carried out in order to obtain detailed information about the current working practices of CfA staff (users) with respect to data acquisition, handling and dissemination. This information provided the basis for a preliminary statement of the needs of users based on the said aspects of their work, including the identification of areas where improvements could be made.

To achieve these objectives, the assessment was undertaken by means of:

- a series of meetings (focus groups), each discussing the case history of one *CfA* project;
- a questionnaire administered to all *CfA* staff.

2.4.1 User Needs Focus Groups

The focus group meetings allowed detailed information about working practice to be obtained from a range of *CfA* projects. Staff were asked to attend the relevant meeting for all projects in which they were involved, even if this meant attending multiple workshops. This was done so that differences between projects, and the role of individuals or teams within them, could be highlighted if necessary. The projects were selected to represent the variety of those undertaken by *CfA* staff. By including projects at different stages of completion – from recent fieldwork to completed publication – it was envisaged that issues inherent to specific phases in a project's life would become apparent.

Each focus group meeting lasted for approximately 2 hours and during this time staff were asked to consider the following questions:

Sources - how and where are data acquired/captured?

Routes - how and where do data move within the project team?

Bottlenecks - are there any obvious obstacles to data flow?

Duplication - where does duplication/double-handling occur?

Access - who has access to the data, both formally and practically?

Responsibility - who decides levels of access to the data?

Outputs - who receives data from the project and in which forms?

Together with issues raised during visits to other organisations that formed part of the preceding Sectoral Practice Assessment, the results of the focus group meetings helped to devise the questions put to *CfA* staff in the User Needs Questionnaire.

2.4.2 User Needs Questionnaire

The questionnaire was designed to:

- test our understanding of current and desired working practice developed through focus-group meetings and other aspects of the project.
- provide a statement of evidence to support the recommendations to be made in this *CfA* report.
- The majority of questions were multiple choice with one free-text question for respondents to record their top three wishes for improved working practice. The questionnaire was structured to allow breakdown of responses by *CfA* sections (Archaeology, Archaeological Science, Archaeological Resources) and length of service, but not designed for formal statistical analysis. It was also not possible to differentiate between fixed contract and permanent staff. The responses are individual’s perceptions of their working practice, and are unlikely to be derived from any formal examination of records of work practice, eg timesheets.

Thirty four questions were asked (see User Needs Survey Report, Appendix 1.E for details), arranged under headings to obtain information regarding:

- work culture
- training
- work practices
- procedures
- tools available/tools required.

2.5 Pulling ourselves up by our bootstraps

Throughout the course of this assessment stage there have been occasions where the whole team, or individuals working on the project, in particular the project manager, have felt a strong sense of “*pulling oneself up by the boot-straps*”. Because the Revelation project is itself a CfA project - albeit with the goal of reviewing and updating existing processes and systems for archaeological recording - the project has had to work within the existing project management systems and other procedural structures currently used by the CfA and EH.

As a result there have been times when the Revelation project has encountered problems with existing systems and procedures. We have learned as much by the process of carrying out project work in a reflexive manner as by any specific analysis of other projects, or examination of previous work practice.

Some examples of this reflexivity encompass large-scale issues such as attempting a project about changing culture with no organisational change management procedure. At a smaller scale are more mundane ironies (appreciated most by the archaeologists on the team) of creating a sequential numbering system to model the different recording processes in the CfA system, only to find that the group had collectively failed to allocate one of the numbers.

Quite often, especially in the user needs survey, the need for better systems for team working was raised as a key requirement for improving existing CfA systems. Identifying it as a User Need however did nothing to help in relieving the problems that Revelation team members still encountered with matrix managing conflicts of interests over deadlines for other internal projects. More of this is discussed below in the section on Work Practices.

The broader issues of change management and changing systems were also an area where the project often found people showed a natural ambivalence. Many of the respondents in the CfA User Needs Survey declared that they would be generally happy about the prospect of change. But when specific issues emerged during project work that recommended a change of practice, there was often a much more cautious response from Revelation team members. This highlighted that there were no clear internal mechanisms available - other than this report - whereby the project could recommend, let alone simply establish, new procedures. Many times during the project there was a recurring feeling of *"can we really say that"*. This needs to be recognised as a reflection of a broader corporate cultural issue.

3 Research Results

3.1 Review of Existing Systems (Appendix 1.A)

"*The CfA creates an enormous variety of data which has potential long-term value, and in very significant quantities. It is also clear that much of this data is managed in an unsatisfactory manner…The exercise... [the Digital Archiving Project]... has provided important information on key issues such as documentation, organisation, version control, storage, access and reuse, and identified a number of other important factors which are significant for the management of resources.*" (Brown 2000).

CfA's current systems are designed to collect, store, manage, analyse and disseminate the results of archaeological fieldwork and scientific analysis. They should also enable CfA staff to undertake research and development into archaeological theory and practice to enable them to give authoritative advice, set standards and develop new techniques.

There is a fairly equal split between systems used for the collection, storage and management of data and those employed in analysis and dissemination. The systems comprise 70% digital, 30% manual and 10% a combination of both. Most are databases (with a few spreadsheets), and the records contained range from 100s to 100 000s and the storage they occupy from Kilobytes to Gigabytes. There is a trend towards increased use of graphics-based systems.

Well over half the systems are seen as being fit for their purpose. They are widely used and are perceived as being quick, easy and flexible to use and producing quality results. There were, however, also many weaknesses identified such as incomplete data sets, difficulty of use, problems of keeping up to date, isolation from other systems, not widely used, lack of access for other users, limited performance, functionality and flexibility and no security copy/back-up. The greatest problem is seen as lack of integration between systems leading to data-duplication, reduced functionality and inefficient working.

Most requirements (ie projects) are said to be ongoing, as are the systems. There is some concern about "life of project" systems. There is a need to focus on systems which are not project-based but which have a wide use and applicability. There are also issues of support, upgrades and replacements.

Almost all CfA's digital systems run under Microsoft products, but much of the software is very specialised and, in some cases, there are not enough software licences for users. There are many and diverse problems with the existing hardware. These will not be dealt with in detail here as they presumably will be addressed satisfactorily via Tech Refresh.

Most systems exist, fortunately, as a single version and can be categorised as being single user, multiple copy, sequential multi-user and concurrent multi-user. Many are self-designed, showing great initiative, but more formal recognition of the requirement for continued management, data standards, interoperability, documentation and support is needed. Systems are almost exclusively user/team supported and in some cases there is considerable uncertainty and confusion over ownership.

Documentation is very varied and often inadequate or completely lacking, with varying perceptions among staff as to what is available for specific systems. Lack of formal metadata for these systems is also a real concern.

The majority of systems require data (digital or otherwise) to be input from other systems as shown by the data-flow diagrams (see Appendix 4). The main exceptions are where the system itself creates the primary record (eg CfA Recording system, Geophysics Field Data Capture). Any quality controls are primarily user-dependant although a small number of systems do have built-in checks. Data are also output to other systems - to in-house and external specialists and for management, research and publication purposes within EH and externally. These data are available as copies of systems, parts of systems, via intermediate formats or hard copy.

In conclusion, it is clear that many of CfA's systems are efficient and easy to use, some are simple in design and operation. The majority do the job asked of them by users and a few are seen as exemplary. There are, however, some recurring weaknesses. The system "owner" may be non-existent or uncertain. Data integrity can be poor, sometimes awful, and duplication of data is common. Some old but needed data exists solely in formats now unreadable within CfA. There are also many problems associated with entering and extracting data from systems. Hardware and software is often inadequate and, in some cases, there are not enough software licences for users. Some systems are difficult to use, with a lack of documentation, user guide and/or user training. Finally, some systems are not supported and some are not backed up.

A number of opportunities for future systems can be identified. In most cases a development towards concurrent multi-user systems is seen as desirable. Many of the manual systems - card indices, hanging files, filing cabinets and proformas - should be digitised for security, ease of use and interoperability. Automatic checks and validation should be built in and switched on!

A formal and responsible owner should be identified for each system and this knowledge made freely available. There should also be strict enforcement of CfA procedures so that systems are not "orphaned" when staff leave. Lack of documentation is a threat to business continuity; it should be standardised and widely known. Standardised system names, although not vital, would also be beneficial.

Usability shortfalls should be addressed by a mixture of design and training, whereas keeping up to date and standardisation are issues of management and professionalism.

Integration of requirements, information and outputs will give users what they want quickly and be cost-effective. Integrated graphics- rather than text-based systems and spatially-aware systems are the way ahead. This trend needs to be recognised relative to future IT infrastructure management.

Systems should be made available to "everyone" not just specialists but for all stakeholders throughout the archaeological process. In the future there should be direct output by controlled Internet access.

There are clear threats to this vision in the form of systems not designed and operated within a proper IS/IT strategy/framework. These will lack the necessary integration and continuity to be efficient. Without user understanding, acceptance and support a system will fail. Inadequate IS/IT resources can prevent initiation or completion of a project or compromise design, function and performance. Systems must be designed “properly” with due attention to documentation, training and support structures. Time scales and deadlines must be realistic or they will handicap the project, stress the project team and may precipitate failure. Lack of senior management, understanding, endorsement and commitment can allow all of the above and more to occur and destroy staff morale.

And a final point: systems that are easy to use will be used.

3.2 Field Practice (Appendices 1.B and 1.C)

3.2.1 Handheld Drawing Systems

Key aims of the Revelation project are to reduce double handling and to make data more broadly available in the field. Direct digital recording of all data in an integrated fashion would advance both of these. Since this is not part of current sectoral practice (see section 3.3, below) initial feasibility work was included in this stage of the project.

Since the drawn record had been identified as a particularly difficult area, a number of preliminary trials were conducted to explore the feasibility of on-site hand-held digital drawing to inform proposals for future work. The scale of this initial work was designed to indicate whether such approaches were worth pursuing.

The results did prove successful enough to recommend that further R&D work should be carried out including the use of more suitable hardware platforms such as ruggedized Tablet PCs. While there are a number of obstacles that would need to be overcome, the promise of an integrated record available during excavation with minimal disruption to work flow is a strong incentive to pursue this matter further.

It may be that direct digital drawing systems can be developed that are used in appropriate circumstances, such as for annotation of geo-referenced photographs or for recording standing buildings. For the full recommendations for further R&D work on digital drawing see Appendix 1.B.

The principal criteria used in initial trials to assess the digital drawing methods were the following:

- Can it do the task at least as well as current practice?
- Does it provide new additional functionality?
- How compatible is it with other software?
- What were the hardware advantages or limitations?
- How easy is it to use on site?
- Is it compatible with processes for post-excavation & dissemination?

3.2.1.1 Can it do the task at least as well as current practice?

Compared to pencil drawing on permatrace and subsequent off-site digitization, the results of the hand-held trials indicate that direct digital drawing can greatly enhance the speed of delivery, degree of accuracy and ability to re-use the majority of drawing information while still on site. However, some aspects relating to particular tasks, such as section drawing, need further research before digital drawing can be fully implemented.

3.2.1.2 Does it provide new additional functionality?

There are a number of areas where there should be added functionality:

- Excavators are able to check and update data graphically on site.
- Improved speed and accuracy of recording both survey and archaeological data on site and in relation to the OS national grid.
- The earlier availability of digital drawings to inform excavation strategies on site.
- Ability to deliver drawn data more widely, quickly and efficiently to all members of project teams including external specialists.
- Far greater ability to analyse, group and interpret site records in a GIS system both on site and during post-excavation analysis.
- Ability to attach a range of intelligent information to the digital drawings.
- Drawn information in reports can be disseminated far more readily via web-based technologies.

3.2.1.3 How compatible is it with other software?

The GIS products upload directly to ArcView with no loss of spatial or attribute data. The data structure of the AutoCAD drawings needs to be improved to allow ease of use within a GIS data structure.

The final archiving format of the drawing data should also be addressed. SGML, or more likely its derivative SVG,XML should be considered as the archiving format for digital drawings and the requirements of the Government's eGMF metadata standards should be incorporated into any long-term archiving needs.

3.2.1.4 What were the hardware advantages or limitations?

For initial trials, three readily available CfA hardware platforms were used for the digital drawing trials; the *Hammerhead*, *Pencentra* and *IPAQ*. Of these, the Hammerhead was shown to be inappropriate due to its bulkiness and lack of portability, but this hardware is due to be replaced in the near future by a Tablet PC under the EH Tech refresh programme. The two other available platforms were more portable, but each had its own constraints. The screen of the Pencentra was difficult to view under certain weather conditions and the IPAQ hand-held was good for text entry but did not have a screen size suitable for large area drawings. Despite their various limitations these hardware platforms were used because of their ready availability for testing the software, not because they were expected to provide the final hardware platforms for an on-site drawing system.

An additional preliminary trial was also carried out at Fort Cumberland using a hired Tablet PC and the initial results, were favourable enough recommend further R&D work on using a ruggedized, all-weather version of a Tablet PC in a future Revelation R&D project to develop digital drawing systems.

3.2.1.5 How easy is it to use on site?

None of the devices in the preliminary tests had purpose-designed outdoor screens, therefore in normal daylight there were some difficulties seeing coloured lines, particularly when using pale colours. In general the black and white screens were more visible. For most current field-drawing purposes black and white drawing appears to be adequate. Further research is needed into the suitability of different screen options under varying light conditions and particularly into the use of water-proofed and ruggedized versions of any hardware platforms.

During the trial, communications between devices caused considerable practical difficulties, although much of this could be attributed to the experimental nature of the testing. The appropriate compatibility for communications between devices will need to be addressed by the systems design. The trials suggest that digital drawings will need to be saved as early as possible to a server in the site office and that further R&D should explore how this might be done using wireless networking, to enable two-way access to the drawings from field and office.

3.2.1.6 Is it compatible with processes for post-excavation and dissemination

The quality of digital drawings must at some point in the process already achieve a standard that is acceptable for “what we do now” in regard to publication and dissemination as all archaeological publications are now printed using computer technology.

3.2.1.7 Work practice and platform choice.

The outcome of the initial field trials suggests that, to support current working practice of individual excavators completing the drawn record for the contexts they excavate, a digital drawing system would probably require both a hardware drawing platform and a context database recording device, but only if the necessary data was available to both hardware devices. This is a reflection of the current work practice of using paper-based context sheets that are separate from the main drawing records made on permatrace. However, it is also possible that a single hardware platform could be developed that would enable both drawing and database entry to be carried out on a single piece of equipment. However such a solution would probably be more expensive if each excavator were to have such a device, and it may prove unduly cumbersome for excavators to carry tablets around site when only needing to make text based records.

The indications from fieldwork experience are that individual archaeologists will need a text and numeric data entry device, such as a small and portable handheld PDA, during the process of excavating the site; the digital equivalent of the excavator's clipboard and context sheets. There may be factors of cost effectiveness that make it more practical to give all excavators a handheld PDA but only have a smaller number of drawing tablets to be used on a "need-to-draw" basis. It is however absolutely crucial that the resulting drawings can be viewed by excavators in combination with any non-spatial digital records that they may be cross-referencing.

3.2.1.8 Conclusions

The initial recommendations suggest that we should aim to create digital drawings in the field, which can be used for each of the required purposes of the primary record, analytical interpretative models, dissemination of the interpreted results and the site archive. It is already the case that we have examples of high quality digital drawings created at each of the above stages. Full-scale feasibility work will be required to pursue the issues described above. Further work should focus on the design of a digital drawing process that takes all uses into account when the drawings are being created. This will also require the design of tools that enable that process to be carried out more quickly and effectively than the existing systems. For detailed recommendations for further R&D please refer to Appendix 1.B.

3.2.2 Current Field Practice

Data-flow and entity-relationship models for the CfA excavation recording system were produced following a workshop involving most of the Revelation project team. These were tested by a 'participant observer' working as a site assistant on the excavation of a round barrow in Wiltshire, part of the Badger Damaged Barrows Project. This enabled the actual data flows to be compared with the model, and the processes by which data were transferred to be described. Detailed notes on this are presented in an appendix to the Field Practice Report.

The initial data-flow model reflected the composition of the group which did the groundwork for the model, and some aspects of the recording system were not well represented. Some key relationships were missed, for example between survey data and the small finds record. The absence of site supervisor-level staff from the meeting (which reflected the structure of permanent CfA staff), also meant that a level of expertise and understanding was missed, and probably led to a bias towards the post-excavation use of completed records rather than the way in which the record itself is formed.

This project may not be fully representative of recent CfA fieldwork. For example, some of the recording forms were not used because of the nature of the site (eg the Built Structure record). It was not possible to examine the data flows involved in on-site environmental sample processing, as this did not take place because of equipment problems. Additionally, the project manager and site supervisor had not previously worked on CfA excavations. While methods followed the CfA Recording Manual, they may have differed in some ways from established 'custom and practice'.

A key observation is that while the data-flow model envisages flows between the formal elements of the recording system (especially the forms), on site much of the data moves in a less formal, less structured way. The present recording system does not appear to support this well, and the work of the Revelation project offers the opportunity to consider how this could be improved.

The on-site data flows relied heavily on memory and/or asking others for information. Labels, tags and personal notebooks (or odd bits of paper) were also important, and the context and small find indexes (whose main function had been seen as controlling number allocation) were found to have a more significant role than envisaged.

Data flows were often more complex than modelled. A flow shown as one-way on the model was frequently two-way (for example, the Cut/Deposit record was shown as supplying information to the Working Matrix, but in practice relationships might be recorded on the matrix and the information later transferred to the Cut/Deposit record). The drawings were an important source of data for the Cut/Deposit record and the Working Matrix, but neither relationship was in the model.

Some specific suggestions for improving data flow were made. These fall into three groups:

- Improving the way the recording system operates, by completing the Recording Manual, and modifying some forms to improve ease and accuracy of data collection. Current photograph records require the same descriptive information to be written on three separate forms. Some gaps in the system were identified – the lack of a Skeleton Index form or a section on skeleton recording in the manual, for example.
- Improving access to information, particularly via the drawings. Ensuring the drawings were easily available and organised by grid square, with each sheet showing a plan matrix, would make them more useful on site.
- Recognising the use of the 'transitional' information in the indexes and improving its quality to increase its value. Information in the context index, for example, tends to be regarded as transitional in the longer term, but improving its quality would increase its value in the recording process.

But the key point is that the distinct nature of the site data flows should be recognised and supported by our systems. Two data-flow networks were in operation. One data-flow relates to the records to be used in post-excavation work - the formal 'end product' of the fieldwork stage and the way in which data is passed on. A different (if overlapping) and less formal data-flow process supplied the information excavators needed to do their job of producing that 'end product'.

The trial implementation of digital indexes seemed to work well, and saved the finds supervisor a great deal of time by avoiding the need to copy data already held digitally (context descriptions, co-ordinates) onto paper recording forms. Regular printouts of information were provided for filing with the paper finds records, to build confidence in this approach. They also made tasks such as writing co-ordinates on small finds bags easier. In retrospect, the indexes could also have been used to assist in information retrieval by the other site staff. Although there can be problems with 'hybrid' recording systems (ie partly manual and partly digital), this digital-indexing approach does seem to offer considerable potential for increasing efficiency and availability of data in the short term.

It had been hoped to assess the usefulness of the digital indexes in initial post-excavation work (the stage before the full records have been entered and made available digitally). Owing to the relative project timetables, this has not yet been possible.

Based on the observations, a revised data-flow model has been produced, and some changes will be made to the entity-relationship model.

3.3 Sectoral Practice Report (Appendix 1.D)

This report was based on a literature review and fact-finding visits with other members of the sector who are using information systems research

3.3.1 Fact-finding Visits

For practical purposes only certain individuals could be assembled as representatives of organisations and therefore it was agreed that for the purposes of reporting no individuals would be named directly *and that specific information from the organisations involved would not be directly attributed*.

Most archaeologists primarily want to be doing archaeology not worrying about how their computer might record information, but we found that there was wide spread interest and concern amongst all groups of archaeologists about how the archaeological process might be improved and re-designed and recognition that the use of IS could contribute.

It was widely felt that there is unlikely to be a “one system that suits all circumstances” solution for English archaeology. Different recording strategies can be adopted by different organisations for different types of fieldwork, analysis, and publication under very varying conditions. Most commercial organisations contacted, felt their systems “did the job required”, but many felt that better integrated systems could do more, if time and resources allowed. In particular a systematic organisational approach to data integration ought to provide the following benefits:

- clearer understanding by staff in all areas of what to record and why
- easier data transfer both within and between organisations with less data loss or corruption
- less wasted effort from staff “reinventing the wheel”
- higher quality and consistency of end products

None of the commercial units interviewed are directly recording primary context information digitally in the field, but a number of research-led organisations have developed systems for doing this on specific research projects and one organisation has been using handheld recording devices in the field for over fifteen years. Some of the latest R&D work was beginning to look at the applications of wearables (electronic devices in clothing) and wireless networking for fieldwork. This however tended to be seen as exciting cutting edge work, but with not enough practical application feeding through to “mainstream fieldwork”. This was partly attributed to a lot of R&D systems being funded on a research project basis and therefore the systems developed may not continue beyond the life of each academic project.

A concern was raised that “blue skies” research could be seen as irrelevant to a large part of the development control sector unless the systems and technology could be applied to day-to-day fieldwork. It was felt that improved efficiency (costs) of digital systems was considered a crucial factor in their adoption by commercial operations.

Commercial fieldwork applications of IT are almost exclusively dedicated to the recording of spatial data using equipment such as TST or GPS rather than providing excavation staff with IT as a tool in their fieldwork practice. Some organisations do use GIS in the field on laptops but usually only on large sites, because of the capital expenditure, and training, involved and the likelihood of the need to write-off some of the equipment. The degree of use varies from project to project and can depend a lot upon the IT literacy of the project director or manager.

Most commercial units (and all those interviewed) record information in the field on pro-forma paper context sheets. Usually the data from the paper record is then transferred on to a computer but often not until after the excavation is complete and the subsequent uses of the digital data vary. In many organisations there is considerable inefficiency from double-handling and errors when transferring information between paper and digital media, particularly where digital systems have been implemented in a very piecemeal fashion.

When assessing how data is used in off-site analysis, there were few systematic mechanisms adopted for controlling what sorts of digital information specialists returned. Some felt the definition of being “a specialist” is the particular knowledge, which separates you from the rest and therefore requires its own way of working. It was often found difficult to integrate results from specialists because of difficulties in understanding the research questions that specialists were addressing.

There is growing recognition of the value of archaeological data for academic research and re-use and therefore acknowledgement of the importance of disseminating information in a variety of media and including digital archiving. But dissemination and archiving still take second place, particularly in the commercial sector, to “inevitable” practical project constraints such as lack of budget. The approaches to dissemination therefore tend to be project specific, whereby larger (well-funded) projects are able to build in more provision for digital dissemination & archiving.

Most data creators and users recognised the considerable peer group prestige and academic kudos derived from “paper based” publications and some felt it was unlikely that people would totally abandon printed publications, although this should not exclude using digital media as well.

3.3.1.1 Problems & Opportunities

The primary intention of our visits was to gather information, ideas, and examples of both pitfalls and opportunities that would be of most relevance to CfA’s future needs for the design of a new IT system.

Many people identified that there is a major training issue with implementing new systems or in getting new and innovative technologies adopted. It was considered very important to get “user buy-in” to the design of any system, in particular the user interfaces, and important to find ways that encouraged *all* staff to use the system. If the system is too complex then the users can be stranded when it falls over and will be too reliant on technically competent people. Some archaeologists raised concerns that computer based records might tend to restrict thought and it will be important to enable inclusion of additional interpretation beyond the minimum asked for from pick-lists or check boxes. It is important to recognize that change itself, and particularly introducing new IT, can be threatening to people and needs careful planning and management.

Often in archaeology the use of information systems tends to be on a project-by-project basis. A purely project centric approach to data collection can lead to silos of information with very little integration of data at the analysis or dissemination stages. CfA need a system that plans for integration at the point that projects are conceived. Such planning should incorporate projects that move through various different stages and which may be geographically dispersed.

Other opportunities presented by the development of integrated systems include much better feedback of information to the excavator on site to increase the engagement of the staff with the research objectives and interpretive processes that drive the project. The availability of the databases and web-based delivery of information while on site increases the opportunity to share data amongst all the project team members, to enable people to use the data in a good way for checking quality, and encourage comments and feedback to improve team working. Planning early for digital dissemination and having archives in a well structured digital format should better enable their dissemination and further use online.

In the course of visits a number of more general problems were identified that seemed to extend beyond the immediate scope of Revelation. One such was the perception that many of the most able people don’t remain in fieldwork because of poor skills development and a lack of CPD. To improve the level of intelligent data collection, we need to bring specialists and training onto sites and generate more skills development and feedback from specialists and the systems to those in the field. This will help reduce problems such as poor site sampling and let staff engage with the research questions, by establishing a better dialogue between people and process. The appropriate application of ICT on site could help alleviate some of these problems.

3.3.2 Literature Review

The Literature Review considered nearly 100 articles and books out of the 1500 identified in the Literature Search. Given the vast number of references in the Literature Search Database the publication gaps for our purposes were substantial. Generally speaking, there was a lot more material relating to technical issues than theoretical discussions or methodology.

People publishing in archaeological computing are more interested in development than implementation. Some of this sense of being on the cusp of great things is connected to the discourse surrounding IT in general, which is closely tied to 'progressivist' narratives (Huggett 2000).

There is very little published that gives a 'from Dig to Dissemination' overview of Information Systems in archaeology. Much is written on how to record (Roskams 2001), some on interpretation in the field (Hodder 1999, Lucas 2001), some on phasing and the assessment (Roskams 2000), almost nothing on analysis. The most general overviews were in the Digital Dissemination and Archiving Systems topics perhaps because they focused most on what might be needed from the 'end product'. Nonetheless, it is still not clear what a digital publication report will actually entail. Many papers are advocating XML for use in digital dissemination of data to facilitate flexibility in end use.

Additionally, the emphasis is on the experience and work of individuals. The role, nature and support of teams is barely discussed at all. The material on 'dynamics' that exists is heavily focussed on theory (Hamilton 2000, Bender et al 1997), which makes it difficult to identify patterns in relationships. As a result, there is little discussion of data flow as we conceive it, since getting data from or to another team member is not discussed.

In all areas of the review, papers repeatedly stressed the need for planning, documentation and integration in the design of systems. There is a desire to integrate IS into the archaeological mainstream, but little sense that this is happening. Interest in the impact of IS on work practice and knowledge generation is building (Denning 1997, Shirky 2003). But there are also papers considering the social position of IS from a more critical perspective (Huggett 2000).

3.3.3 Conclusions

When first considering the Revelation project, we had a sense from many people, particularly those using IT, that the main issues relating to systems development for archaeology had been solved. Indeed there is a fairly large community of people developing information systems for archaeological projects. Most archaeological units now use IT as a part of everyday work practice. Nonetheless few, if any, organisations have achieved systems that they are happy to consider as fully integrated.

The discipline’s current use of computers can be characterised as widespread use of relational databases for holding core archaeological records, structured according to a variety of differing data models based on a range of different recording methodologies. As new technologies have become widely adopted for drawing, photography, analysis, publication and other aspects of the archaeological process, further software has been adapted with an increasing trend towards more fully integrated systems that enable the core data to be used, and added to, throughout the archaeological process.

No-one we spoke to remotely suggested that they might have an off-the-shelf system that could deliver a plug-and-play solution to handle efficiently the capture, analysis, dissemination and archiving of all excavation data. Instead, we have found systems that suit their users needs **quite** well. Some of these systems will provide inspiration and may be adaptable so that they form a component of our system. They may be particularly helpful in the development of technically complex issues.

The main barrier to the development of good systems is not technology, but time and expertise. Contemporary hardware and software is powerful and flexible. Lock summarises, "*Despite the variation in recording systems in use, it does appear that relational databases, based on the concept of single-context recording, do offer a stable core system for the written record. With the increasing flexibility of modern software, such databases can now be routinely linked with the drawn and photographic records moving the whole recording process into an integrated digital environment*". (Lock 2003, 98).

The least well supported aspects of the archaeological process are not data management, or even analysis, but communication, project management, contemplation and the development of ideas. Many professions share our concerns with regard to data management and analysis as well as communication.

Archaeologists suffer with poor systems largely because they have rushed and scrimped, particularly in the systems analysis and design phases of development. Allied to this is a desire to use existent systems which often stems from expediency, but also from a reasonable desire for inter-operability.

Training is essential and often overlooked. Some organisations employ dedicated trainers teaching and encouraging people to use the new systems. There is a substantial need for Continuing Professional Development in Information Communication Technology.

We must plan to review and develop any system. Most systems that work either have a constant programme of development, or bring out new versions every 3 - 5 years. If the system is clearly designed and documented, update will be easier. Integration is key - most successful systems are based on a well-structured relational database with defined links and interfaces to a GIS and other programs for graphic and photographic display.

Legacy data are not going to go away, but we should design our system to deal with what we are creating now. The clear message from a number of organisations was that we should adopt a 'year-zero' policy, where everything and everyone after a certain date would use the new system.

The assessment of current practice should help inform CfA's development of a core system that can manage existing data requirements. But IT continues to change and develop, and with the design of a new system there are opportunities to also consider other innovative areas that the sector is currently experimenting with that might become integrated into archaeological systems in the short to mid-term future. If we can design a system that is compatible with such developments then we should be able to carry out R&D using new recording technologies that allow the data from any such projects to be integrated into the new system.

Whatever solutions Revelation provides, they must go beyond the simple "we can do everything we do now, only digitally". The best systems developed by others to implement new technologies have been based on wide-reaching assessments of user requirements and, where necessary, re-appraisal of working practices.

3.4 User Needs (Appendix 1.E)

This report builds on the outcome of the Focus Groups and responses to the User Needs Assessment Questionnaire

3.4.1 Focus Groups

The focus groups examined the flow of data in a project team and management of these data. Five CfA projects at various stages were examined, with an additional group looking at external projects which involve CfA resources.

All of the projects examined cited paper records created during fieldwork as the main source of data. However, topographic survey and 3-D data are stored digitally as ASCII text or CAD drawing files. Geophysical data also are almost exclusively digital.

In principle, standardised paper records are input to computer databases after being photocopied for security. All drawings are scanned for security after they come off site and some are digitised as CAD drawings. In practice a considerable length of time may elapse from the end of fieldwork before these tasks are done.

Bottlenecks that restrict the flow of data appear for a variety of reasons, such as an increase in the extent of fieldwork without a concomitant increase in post-excavation resources. Problems with data consistency, arising from proformas being completed incorrectly on site, can result in excessive requirements for data checking. If sufficient resources are not available to conduct data-checking on-site the problem is exacerbated in post-excavation, where such resources usually are not available in the short term. The current lack of a mechanism to track bulk finds, and inflexible recording systems which cannot be digitised, also are identified as bottlenecks in the flow of data.

It is not only bulk finds that are problematic. Duplication in object-related tasks often arises because of problems in tracking them. This applies in particular to specialist analyses and illustration, when objects need to be transported to and from controlled storage. There also is frequent duplication in the data provided by excavators to various specialists working on a single project.

Varied approaches by managers and specialists in logging and tracking tasks, and a lack of records coherency, are seen to be at the heart of many current problems. Insufficient communication between the different CfA sections and teams adds to the overall problem.

Changes in management priorities can lead to delays (of up to several years) between succeeding project stages, leading to much duplication of effort or the need to transfer data to newer systems. Such transfers as one consequence may necessitate data cleaning, a laborious and time-consuming task. Site data inherited from older or disparate sources often need to be collated and standardised before further use.

A common barrier to access to data is that much data exists solely on an individual's personal computer, or in their dedicated computer network folders. However, there also can be problems with centrally-stored data, where read and write access is available to all. Concerns include the vulnerability of data to accidental or deliberate corruption, version control, and who has responsibility for keeping data up to date. One project still had all of its data on paper only, to which the project manager had sole access. This example shows a lack of awareness of the CfA's digital archive strategy and current data management procedures.

The administration team expressed a need to have access to summary descriptions of individual projects, as essential background information for their work – particularly in recruiting site staff.

The Geophysics team runs a database accessible outside the CfA via the Internet. Maintenance is an integral part of the team's duties, which stretches already tight timetables. The Scientific dating team is currently developing another such database for Radiocarbon dates. There are concerns regarding the sensitivity of some content in these databases, but general access can be limited by blocking out some fields for non-specialist users.

Most staff thought that project managers should be responsible for determining and providing levels of access for others to project data. Exceptions to this were finds and environmental data, where it was perceived to be the responsibility of the appropriate project team member(s).

Outputs from a project primarily are in the form of site archive deposition, production of an assessment report and updated project design (for analysis), and a published analytical report. External projects have wider circulation to include the client, land or property owners, and others with a current interest in a site.

The origins of many of the problems identified by the focus groups can be traced back to what can be regarded as ill-considered or even unrealistic management decisions, a lack of management enforcement of procedures, insufficient provision of staff training, or a lack of quality control. Inffective records management lies at the heart of most of the problems revealed during these discussions.

In addition to those mentioned above, there are other bottlenecks that affect data flow, but raised only briefly in one of the focus groups. These concern the availability of staff to conduct project work. Specifically, a small number of key staff appear on all project teams. These can include external specialists as well as those from within CfA, from the Archaeological Conservation and Environmental Studies teams, or finds specialists from the Archaeologists' team. It also concerns the provision of other specialist services. For example, on the five project teams interviewed, one person each provided or oversaw the following roles: topographic survey (3 projects), computing (3 projects), AutoCAD support (4 projects), and archiving (5 projects). Only one other individual attended more than one of the project team-based focus group meetings. Another limitation on staff availability (not covered in the focus groups) arises from other demands on their time, such as monitoring Archaeology Commissions projects.

3.4.2 Questionnaire

The responses to the questionnaire produced a wealth of data concerning trends in the culture of the CfA, training issues, work practices, procedures, and the availability of or requirement for suitable computer tools.

The survey shows an even spread of new, mid-range, and long-term staff members. There is much cultural commonality between the three sections of the CfA, but some differences are apparent, eg Archaeological Scientists are more often involved in external projects than are the Archaeologists, and therefore have less involvement with existing CfA procedures. There is broad agreement that the CfA should have a shared philosophy on the purpose of archaeological fieldwork, and a broadly similar agreement that such is currently lacking. Variation on this question is more pronounced according to length of service with CfA than between the three sections.

Most staff spend a greater proportion of their working time on computers, and most staff feel competent in their use. Very few consider themselves to be technophobes, with some correlation between those who classify themselves as such and their longer length of service. There is little reluctance evident to the need for CfA staff to adopt new software solutions – they demonstrate an ability and willingness to adapt quickly and get to grips with new software to improve their work practice.

A large proportion of staff often do not place their work files on the Fort Cumberland computer network, either because they have incomplete work in progress, or frequently work away from the office on laptop computers. There also are problems importing files from the network, sometimes because of limitations of an application, or because of network underperformance. It also is the case that staff knowledge of the architecture of the network is not all that it should be. Suitable training to address this shortcoming, and the development and application of clear, relevant, and simple procedures, should encourage greater use of the network, making data more accessible to all. However, this would require a rigorous use of metadata to describe files and help maintain version control.

Most staff see the task of acquiring information as a mid-range time investment that they are unlikely to want to alter much in the future. However, approximately one quarter of staff would like to spend less time doing this. Obstacles to accessing or obtaining data are recurrent themes.

Communication is an important issue, emphasised by the answers to the questionnaire and discussed in the focus groups. There is a wish to reduce the amount of time spent communicating, while there was a 50:50 split between those happy or unhappy with the level of communication between colleagues, with some variation between the three CfA sections and by length of service. Even though there was no strong correlation between those who were dissatisfied with the level of communication and those who wanted to spend less time communicating, the interpretation placed on the findings is that, while there is plenty of communication happening, it appears not to be of the quality needed.

Most staff said they would also like to spend less time having to enter and/or prepare data. This could be achieved by adopting a more automated approach to recording and capturing data.

Analysing project data currently has a low time-investment. Most staff want to increase this, and to incorporate other data sets to improve the overall level of interpretation. Almost half consider writing and drawing as their primary tasks and saw little need to increase the amount of time devoted to these outputs. Nevertheless, there are cultural differences apparent here, with a strong desire in the Archaeological Resources section to increase the time available for producing outputs, and a similar but less pronounced desire amongst the Archaeologists' section. The Archaeological Science section appeared content with their current level of output.

A lack of awareness of existing CfA procedures is a serious source of problems. Formulation of more useable procedures and the provision of training in their application is needed to improve this situation. Over one third of staff are unaware of any of the procedures either for project, collections, or data management. Those who are aware of them consider those for collections management and data management to make sense. In contrast less than half of staff think the same about management procedures. Many staff also suggest that procedures are helpful only if enforced, and if adequate resources are provided for them to be implemented. Currently this is not the case. It is interpreted that there is a degree of passivity at management level which leads to the lack of enforcement of these procedures, and the inevitable problems that ensue.

Quality control of digital data most commonly is either by inspection during use or record by record inspection. Validation controls at input or intermittent inspection are employed less often. The least popular method is double entry of data.

Strangely, in only about 50% of cases do CfA staff provide a specification for the format in which data should be supplied to them by colleagues. Even more oddly, just over 40% of internal and external colleagues specify the format in which they need to receive our data. This suggests either ignorance of common data standards, a blasé approach to the matter, or an ability to adapt and juggle with data in diverse formats. This inevitably is a less than efficient use of data. The formulation of appropriate procedures and provision of training should improve matters.

Over three-quarters of CfA staff have lacked the correct IT tools for the job. This prompted them either to develop an inventive solution using existing resources, ask a colleague for help, or ignore the problem and hope it would go away. It was pointed out that the use of external resources to compensate for this lack of tools relies positively on the development and maintenance of networks of personal contacts, but negatively advertises EH inefficiencies and deficiencies to the outside world.

Standard office computer applications have the highest levels of use amongst staff, followed by Internet browsers, databases, and spreadsheets. The remainder comprises more specialised software, eg graphics packages, GIS, or specialist data processing. Just over one third of staff currently analyse their data spatially. This would increase to three-quarters if the appropriate tools and training were provided. One note of caution was sounded: data should be analysed spatially only when it is useful to do so.

In general staff have received a greater amount of training for more complex software (eg graphics, statistics, GIS, databases, and spreadsheets). Few staff want training in simpler, more general applications, such as email.

Three quarters of CfA staff across all three sections regularly use databases or spreadsheets. Sources of irritation are difficulty of use and unreliability, indicating a degree of poor implementation or limitations in functionality. Specialist statistical and plotting packages (not office-standard Excel) also are required for some functions, eg Archaeomagnetic Dating, to provide the specialised forms of data handling and graphical presentation these subjects require.

A large proportion of staff use physical reference and storage systems. A large number would prefer these to be made digital. The remainder wish to stay with what they have, because they prefer to stick with what is familiar and they are comfortable with, or it is not worth investing in digitising these records, or the physical system does not lend itself to a digital format. Restriction of access to these physical systems to single users is far from ideal – wider availability of these data is desirable. The provision of wider access is generally thought to be a “good thing”, with little evidence to suggest that staff are restrictive or overprotective of their data.

3.4.3 Conclusions

The CfA currently operates with a degree of fragmented cultures. These are not aligned entirely along either the three sections (Archaeology, Archaeological Science, Archaeological Resources), or by the length of time staff have worked for the organisation. One of the fundamental issues to be highlighted is the number of CfA staff who are unclear about what constitutes a CfA project as opposed to an external project. Similarly, the variability in knowledge of existing procedures for project, collections and data management is worrying and appears to be at the root of a number of the problems experienced currently. Training needs to be strengthened to increase awareness of basic internal procedures, especially those related to what should happen during and after the return from site. Revision and streamlining of these procedures would be beneficial.

Many of the problems highlighted during the focus group meetings stem from poor communication. However, in the questionnaire staff indicated that they want to reduce the amount of time they spend communicating. Given that a lot of time currently is invested in this, it is vital that a means for more effective communication is found.

Staff are concerned about the lack of access to data both within CfA and across the organisation. The strongest constraint is felt to be the lack of access to data in other parts of EH. When asked where they tended to store data and why, responses showed that CfA staff rarely work exclusively on networked drives, either for pragmatic reasons (eg working away from the office) and/or personal preference (eg work is incomplete). The network's Share drive gradually is seeing an increase in usage which will improve access to data, but there needs to be a comparable increase in the use of metadata to describe files and aid version control.

Staff want this increased access to data and many would like them also to be available in a relational database format, both for storage and interrogation. A key element is that there must be flexibility in how data are visualised (eg forms, graphs, tables, lists) according to both the task being undertaken and the eventual audience.

Overall, staff want to use systems that a) reduce duplication of effort and double-handling of data across the CfA and EH, and b) produce reliably accurate data from the point of origin. With respect to the latter, it would be desirable to increase the use of automated validation controls as the principal method of quality control for records. Currently, more labour intensive methods (eg record-by-record inspection) typically are used.

Two overwhelming user needs have been identified during the course of the User Needs Assessment:

- The development of a genuine CfA system that can cope with the full range of tasks and projects currently undertaken by CfA staff.
- The streamlining, standardisation and enforcement of procedures for data collection, recording and management.

These two needs are linked inextricably – both are essential if either is to succeed.

4 Discussion

4.1 Business Context of Research at CfA

"*The Centre for Archaeology integrates archaeology and archaeological science in research projects, in providing advice, by monitoring work done for EH, by training and by dissemination and outreach*." *– EH Functional Directory*

As the strap-line from the *Functional Directory* shows, the CfA has a complex role within EH, even though it fails to mention the other role of developing archaeological strategy. All of these functions are specifically tuned to address EH's aims, corporate priorities and goals, especially where those goals are those of our partners and clients such as DEFRA. Revelation is concerned with the research aspects of this business, seeking ways to make the work more efficient and effective. It should be noted that the Revelation team are not accountants, and that much of the outcome from CfA research is difficult to quantify in a strict "pounds per product" system because of the nature of intellectual processes.

The backbone of CfA is the ability to provide expert advice to the rest of the organisation. It is used in our Archaeology

Commissions monitoring as well as our involvement with the Regions and other parts of EH. Our expertise is kept current through our project work, and the projects themselves are chosen because they address the research priorities of our clients.

CfA takes on intensive rather than extensive projects. We look in detail at individual sites through a battery of techniques ranging from remote sensing to excavation to analysis of artefacts and environmental evidence. In this way we complement the work of the Metric Survey team who provide detailed mapping of sites, and that of the various regionally based Archaeological Investigation teams who give an extensive picture of the archaeological resource through thematic studies.

4.1.1 Clients

CfA has clients both within EH and external to it. Since our work is part of a wider picture of understanding the past and planning for the future the links between CfA and our clients are extremely important. We rely on data and services from our clients, and we provide them with data and services in return. We must therefore maintain our skills and systems so we can continue to fill a useful role. After all, our clients approach CfA because we can provide added value in our projects.

Internally, our list of clients include the EH Executive Board, EH sections such as Properties and Presentation and the whole of the Archaeology and Survey Department. Service to these clients takes a number of forms including direct advice and participation in specific projects. Within the Department we provide information, advice and services to the Chief Archaeologist and his advisory group, particularly in support of the development of strategy. We also count the Metric Survey and Archaeological Investigation teams as clients with whom our work often dovetails. In the Regions, we work with the Inspectorate and the Regional Teams in support of casework and maintenance of our own properties stock by providing advice, developing briefs for archaeology, and undertaking some direct work where appropriate. In short, CfA supports the whole of EH in the attainment of its goals.

Our external relationships are just as important, and include addressing direct requests, involvement in Archaeology Commissions casework, and provision of a research “data library” for the archaeology of England. Further, we support the development of strategy with partners such as DEFRA through direct advice and through project work.

4.1.2 Resource profile

There are currently 58 complemented staff at CfA, and at any one time there can be up to a dozen temporary staff at Fort Cumberland recruited for specific tasks or projects. Where fieldwork calls for excavation the number of temporary staff (mainly based on site) can increase considerably.

Projects are developed by a staff member who will act as Project Manager responsible for the creation of a project outline in which project aims and rough resource estimates are set out for CfA Management Group. Then follows a project design (PD) which lists and justifies detailed tasks for the required resources. Since there are a number of other projects at various stages of completion that also require resources the needs of each new project must be carefully woven into the existing CfA work programme.

Projects may draw upon full-time CfA resources such as our various departments – Graphics, Archives, Administration, Technology, Archaeometry, Conservation, Environmental Sciences – as well as our archaeologists. There are physical resources also, such as vehicles, tools, IT equipment and facilities. Where these are not sufficient for the needs of a project additional resources can be found in the form of fixed-term staff, external contractors and hired equipment.

Management of a CfA project involves the management of colleagues of various grades as and when they have tasks to contribute, creating a network of matrix management that is separate from the hierarchical system of line management applied to all EH employees. Conflicting pressures from different projects and non-project work require careful planning by line managers and the whole system is affected by changes in priorities.

4.1.3 Project spend profile

Project timescales vary considerably depending on complexity as well as changes to priorities that affect resource availability. Fieldwork projects that involve only one short episode of field data collection might be completed within a calendar year depending on post-excavation resource availability. Many of our projects, however, run to multiple seasons of fieldwork that extend the overall programme by years. In this case many of the post-site archive completion tasks are fitted between field seasons. Where extra seasons are planned from the outset the overall project life can be predicted, but when projects grow organically this becomes problematic. Projects can also fall foul of changing priorities that affect resource allocations, leading to substantial gaps between phases of work. Career progression of key players also affects completion timetables, especially when they leave the organisation. Under the current system it can be difficult to pick up the threads of a project after a long gap, especially if the original team members have been replaced in the interim.

Time management within CfA is supported by a system of time recording based on a mix of project tasks predetermined through the development of project designs and other items that are not. Time recording data can be used along with existing project plans to map the availability of resources at any point in time, but the system is open to disturbance by un-planned tasks that cannot be predicted through the project-design process.

One model of archaeological project planning is that post-excavation work generally costs about as much as the excavation fieldwork that created the archive. That means that half of a project's budget will be spent in the short period near the start when the team is in the field, while the rest is spread out over the following months and years. The field component is spent at a steady rate, but the post-excavation process is prone to episodes of work interspersed with periods of little or no work, due in part to the way that information flows or is held up at various stages in the process.

The result is that post-excavation programmes are extremely vulnerable to changes in priority or resource availability that can drag out a project if not tightly controlled. In theory, building investigation projects are less vulnerable since their post-site work is generally only half of the field costs due to the nature of the process of examination, interpretation and recording of built structures that completes more of the archive and analysis while still in the field.

Further analysis is necessary to determine whether our spend profile corresponds to these models of project planning. Since post-excavation programmes can be so drawn out, there may be more costs than in organisations where full time project dedicated staff can complete projects quickly. Many key personnel in CfA projects find it difficult to spend two days uninterrupted on a project after leaving the field. This constant moving between tasks has effects on efficiency.

While Revelation cannot address the pressures that stall projects, it can help to front-load the archive completion, assessment and analysis tasks in an effort to minimise the amount of post-site work that is exposed to outside influences. The fragmented nature of time spent on project work should also be borne in mind during systems design. We need to facilitate both clear ends to working sessions, and methods for getting up to speed with project progress when returning to the project after time doing other work.

4.1.4 Portfolio

There is a very large and constantly changing set of active projects within CfA, ranging from large-scale fieldwork to post-excavation analysis, scientific analysis of specific materials and non-fieldwork projects looking at subjects like digital archives. In addition there are projects that are dormant, projects being developed and even occasional “finished” projects that are re-opened when new opportunities arise to use their data to address new corporate priorities. They can address methodological development, academic research or analytical work, but in practice our projects are a mix of all three.

4.1.5 Outputs

CfA work produces a number of outputs including formal advice, archaeological records and publications of various types.

Advice takes the form of comments, briefs for investigation that the client can implement, and guidance notes.

Published reports include monographs such as the recent Catterick volume, journal articles and reports published in other works and popular books such as *Windsor Revealed*. So-called “grey literature” takes the form of the CfA Report series (incorporating the AML Report series as well as CAS Reports), CfA News, and CfA Update. The nature of the proposed Revelation system (see Recommendations section below) will aid the creation of these reports through well-structured data, and will enable CfA to present more output as on-line reports.

At present our archives remain as stored collections of data with all the inherent difficulties of wider public access. Where Revelation will make these easily transferable in digital format, the opportunity exists to provide full dissemination of the data rather than just the results of analysis as is common today. The dissemination of research archives will be an example to the profession.

4.2 Technological context

Technology is central to the work of the CfA, both in its day-to-day tasks and its role of methodological and technical development. In the field it manifests itself in the survey equipment used on excavations, geophysics equipment and computers within the site hut. All of these generate data that are used in post-excavation. Back at the CfA there is a wide array of instruments used by the scientists, desktop computers used by virtually all staff members for a variety of data-handling tasks and dedicated computers for the production of publication graphics. Most - but not all - of these machines share data via the local area network (see Brown 2000 for details).

CfA and its predecessors have been heavily dependent on computer technology to manage the volumes of data created by fieldwork and analysis. There are several important databases that hold our primary records as well as interpretations, reports, and management information, and a number of computer systems that support specific pieces of equipment.

The CEU developed a database in the late 1980s called Delilah aimed at on-site computer input and automatic checking of records at the beginning of the archaeological “computer age”. When CAD software made it possible to “draw” with survey data in the early 1990s the CAS promptly explored this process, and also developed methods to enhance photogrammetric output digitally without having to redraw it on paper. 3-D modelling of excavated features via survey instrument was employed at Battle Abbey (1993-4), and planning directly to CAD via survey instrument using TheoLT software was trialled at Birdoswald. Digital enhancement of Photogrammetry began at Windsor Fire Damage (1992-5) and Battle Abbey Courthouse (1993-4).

The primary aims of these increasingly technology-dependent developments were to improve the range and quality of information that could be captured on site within the constraints of time and staff numbers and to provide data to our partners in more useful formats, although some efficiency benefits could be seen too. The majority of on-site recording was however still paper-based, with digital entry happening either off-site in a site hut or back at Fort Cumberland – a situation that continues to this day.

Computer use in the AML also developed as the technology became available, and the Labfile database was created around 1975. As geophysical equipment went digital, masses of data could be stored and manipulated by computer, leading the geophysics team to develop their own software for data processing and storage that is still at the forefront of the discipline. The SEM, XRF, XRD and photographic workstation in the lab all generate data, but currently these are all on separate non-networked equipment-specific machines despite the presence of networked PCs at everyone's desk.

The concept of data sharing lies behind the project folders on the CfA LAN, but these tend to hold mostly field-generated data from excavations. All CfA staff have access to word-processing, spreadsheet and database software over the network, but the resulting files are often kept in individual staff folders. The piecemeal development of technology across the CfA has left a legacy of unconnected systems with "pools" of data that are not easily shared, and the situation is exacerbated by the rise of mobile computing for tasks that traditionally required access to a desk-bound PC.

All of this technological development is set against a backdrop of greater contact with the wider profession. CfA is not alone in the use of computing, since the rest of the sector also adopted computers as they became affordable by the end of the 1970s. Word processing, spreadsheets and CAD have become standard tools of the profession and the development of networks is widespread. Drawings, photographs, and reports that used to be posted to colleagues as hard copy at great expense and delay can now be sent digitally via email. Even humble letters often travel without being printed. Where this data is project-critical, CfA needs to retain the capability to work with it, and the ability for all concerned to access it. Examples include survey data from other EH teams as well as databases and reports from external sources.

There are other units with archaeological databases, and some such as Dominic Powlesland's G-sys are marketed commercially. Data still flows as paper, or on media such as CD-ROM and floppy disk, but the rise in power of Internet-capable computers means that more information now travels as digital files via email and websites. In the wider profession there are already developments in the use of web-based systems to share data. The Interactive Archaeological Data Base (IADB) uses Internet access to share project data with specialists and other team members who may need remote access, and the system holds the working digital archives of all the Trust's projects. The Internet is also used by the ADS as a means of access for its clients, and is also used for dissemination of a number of products by organisations around the country.

CfA can retake the lead on technology through the Revelation project, since none of the organisations examined in the sectoral practice survey have managed to create a system that works from start to finish of a project. Our lead could see all of the sector eventually adopting digital recording, integrated data holding, and web-based wider access to reports and primary and archival data to the benefit of all in the profession.

The Tech Refresh programme currently under way will certainly improve the CfA's present ability to share and manipulate data through the replacement of old equipment and commercial software that can no longer cope with present-day demands. However, it is up to the CfA to knit together all its disparate systems into something more cohesive if it wishes to tackle the inefficiency and data-flow bottlenecks identified in this report. Technology will only be one element of this programme.

4.3 Research context

4.3.1 Understanding our Processes

In the course of the project we have investigated data flow and work practice within CfA projects, and have looked at the process of knowledge generation as described in the literature. We have also discussed these issues with our colleagues in the rest of the sector. Some aspects of the processes that we are hoping to support have been clarified. But some, notably the processes carried out in the 'Analysis' stage remain less well defined.

Although the conceptual and practical issues surrounding work practice and knowledge generation have been discussed a good deal in the literature, there are two sources of bias in the way it is discussed. Firstly, there is a drop-off in the amount of literature written as one gets further through the life of archaeological projects (see page 21).

Secondly, people working with Archaeological Information Systems (Stewart 1997) often describe the archaeological process in a linear fashion. The following stages are commonly identified from a data-driven perspective:

- data collection
- data management
- data manipulation and analysis and finally
- dissemination (of both data and interpretation)

But this sequence rather quickly glosses over the core of the process and the most important element in the design of a system - data manipulation and analysis. Contemporary archaeological theory is less keen on linear processes than earlier incarnations and most archaeologists accept that interpretation and analysis are also part of data collection (Hodder 1999).

An alternative model might propose five analytical tasks that can recur at all stages of research from trowel to monograph and beyond. Each of these needs to be supported by an integrated IS.

• **Atomisation** - the construction of the units of analyses is a choice, a process that often takes place before data collection begins, although repeated through analysis. Key elements in atomisation are the data structure itself, controlled vocabulary and the spatial definition of boundaries.

• **Quantification** - can be as simple as measuring the diameter of a posthole, or more complex as in estimating the mean level of a particular element in an assemblage of slag.

• **Ordering** – this may include classification, categorisation and sequencing.

• **Integration/aggregation** - regrouping, also bringing together different classes of data (and interpretation).

• **Explanation -** where the other processes are incorporated with broader understandings and belief.

The key difference in the second model is that the processes are iterative rather than linear. Archaeological data is not atomised, or explained, or subjected to any other process only once and therefore the IS needs to be able to track or audit these processes and their more complex relationship(s) to each other.

There is inconsistency in the use of the formal project management phraseology of MAP2 and the day-to-day language of archaeologists. The concept of "Post-ex" is woolly and should be abandoned, or at least clearly defined if it is to be used meaningfully. This ambiguity of terminology is not helped because the nature of the processes carried out under the MAP2 Analysis Stage are also poorly described, and need better definition if we are to design a system which supports our work from fieldwork to dissemination.

4.3.2 Supporting Research Agendas

Though CfA field projects are often small in comparison to much developer-funded work they carry considerable weight in supporting English Heritage's Research Strategies. While each project is linked to appropriate corporate, regional and national research agendas in the project design, the relationship between these high level objectives and the dynamic research questions that underpin the interpretation of data is not explicit.

Many people advocating the 'reflexive' methods (Andrews et al 2000, Chadwick 1997, Hodder 1997, Lucas 2001) talk about the importance of the questions that we are asking and how they develop through the course of a project. It is regularly pointed out that the questions being considered when data are collected or analysed deeply affect the data themselves.

There is, however, no formal place for these research questions in the recording methods or data structures currently used in English archaeology. We need to record and hold our data in relation to the research questions being asked at the time of recording. An event-based data model, in an integrated system would allow us to do this dynamically (see Appendix 4). This would create a permanent link between research objectives and questions and the data used to address them.

Holding questions as part of data structure would facilitate tracking data holdings against corporate priorities. The current method of tracking records the research objectives of projects prior to excavation. Using this system, a user may, due to the way the project turns out, find that there is little actual data relevant to some of the original research issues, while other projects with unexpectedly much more useful data, cannot be identified. Recording questions, in relation to the data used in answering them, would also allow selective use of larger archives where only parts of the data may relate to any given research theme. It would also allow us to track when the archive of one project is used as part of a new project with new questions.

This is quite important if there is any suggestion that 'assessment can be completed in the field'. The MAP2 process of assessment determines what research questions can actually be answered by the data. This method of tracking the questions asked of data would make the process of assessment transparent - since the amount of data linked to a question would be clear. Assessment is not just phasing and quantifying, which we can support easily, but drawing a link between the data and the questions that they can answer

Recording questions in the data structure would both require and generate a much more self aware working practice. The research objectives and questions would be alive through the project rather than something set at the beginning and forgotten about.

It would also facilitate team working in analysis (see below 4.3.4). A project member, begining work with a dataset, could see what questions other team members have asked (eg if geoarchaeologists have been asking about waterlogging, then that will be apparent to conservators and to the person doing the phasing).

Finally, it would allow a 'reader' (a person using the disseminated archive) a number of research-based routes into the data, which reflect our thinking, and the way we approach the material ourselves. This would provide an intellectual index, which could be greatly enhanced through web-based dissemination using hyperlinked text, and the latest developments in semantic web technologies for tagging and searching of data. Other projects that are disseminating their archives are intending to do this through providing the queries used in analysis. But such queries cannot be easily archived, whereas if the questions are part of the data structure then they will be safely held in the archive.

4.3.3 Supporting Complex Thought

Speeding up the research process is a double-edged sword. Front loading analysis into recording makes sense, but only if you give it proper resources. Otherwise it is simply a greater pressure. We need to identify which slow tasks are irritations and sources of error (eg - multiple transcriptions), and which are contemplative exercises (eg manipulating a Matrix).

The site matrix (as a graphical representation of stratigraphic relationships) can show more than just the relationships between context numbers. The matrix needs to be a more dynamic tool for modelling the data, with the functionality to show different views of stratigraphically-related data (eg phase colouring, types of finds in contexts, or which samples come from which contexts, etc). A user should be able to select a group of contexts and view their drawn plans, click on a context and bring up its 'deposit/cut details', or a list of its plant macrofossils. Much of this functionality is available in Jnet (Ryan 2001) which is still under development.

Archaeologists in general do not usually expect IT to give them support in intellectual or contemplative work. That is why people working with matrices for the purpose of understanding site formation do not expect a tool like Jnet to exist. Computers re often seen as management tools, communication tools and potentially useful for statistics. We need to develop the perception that computers can help them think about their data.

4.3.4 Supporting Analysis at all Stages

The most time consuming and important process in analysis is integration/ aggregation. An integrated data structure enables a single process of aggregation rather than separate aggregation for each data type. For example Roskams, in his review of excavation methodology, recommends attaching a hierarchy of captions to photographs at the archive completion stage to facilitate the use of the photographic archive by different users (2001, 242). But we cannot know the needs of all users at the archive completion stage.

A more useful mechanism might be to link the photographs to primary reference material (eg Context numbers) with potentially a single description (eg looking north, showing the collapse of A over B). When the context numbers are aggregated into higher level interpretations such as phases, then the user searching on these higher level interpretations gains access to the data (including photos) linked to these interpretations. This can be made more useful by recording which photographs have been used in what circumstances. Graphics Office quality checking can be used to tag the photos as 'well composed', 'clear but boring', 'unreadable' etc. This will speed archive completion and add useful metadata to the photos.

There is a very strong emphasis on 'post-ex' in current IS in archaeology. "*Thus the requirements of post-excavation work - the 'analytical destiny' of the data (Carver 1991) - are the fundamental consideration here. The problem of over computerisation, and therefore any solution, lies with the archaeologist, not the system analyst, still less the designer of software*". (Roskams 2001, 172). The second sentence underlies the approach we are taking with Revelation. The first sentence, however, stems from a sense that analysis does not take place in the field. We need to identify elements of data that may be re-used in the field, either for the development of on-site analysis or for communication of interpretative ideas, but particularly with an awareness of how they might be carried through later in the process.

Just as analysis is not confined to the laboratory, data collection is not confined to the field. Many systems manage field data well but pre-suppose that laboratory data will be managed separately, attached to the site only through reporting. What is often missing from systems is the ability to hold the changes in interpretation and understanding which occur throughout the analysis of a site. The same principles for on-site recording need to apply to laboratory recording, with the same considerations of change management.

Most organisations only attempt to integrate summaries of specialist data into any system that they design (IADB, Frameworks, Ryan 1995). This is generally argued for on the basis that specialists don't want to provide their primary data for fear of misinterpretation. Some would also argue that 'nobody wants' this degree of primary data. We really need to confront this as a cultural issue, because it leaves important and expensive data detached from its context and often unsecured. In fact, many of our external specialists say that they don't retain these data. All of this is counter to scientific practice in other disciplines and derives from mistrust and disregard, on both sides, and is compounded by isolation through work practice. Another argument is that primary data can be analysed more easily in relation to other aspects of work when they are in the same system, whereas working with *ad hoc* copies of primary data leads to difficulties with version control and communication.

4.3.5 Feedback to and from the Recording Process

The review of sectoral practice highlighted the advantages that can be derived from creating a system that provides feedback to staff as early as possible in the recording process (Andrews et al 2000, Beck 2000). One important outcome of on-site digital recording is an improvement in the speed and quality of data verification in the field. But this is not simply a matter of improving the quality of data by enabling staff to cross-check their own and others data in the system while on site. Even more importantly it is a key mechanism for engaging staff much more directly in the how, what and why questions about the archaeology they are recording.

Recording data digitally during excavation enables sharing the information available on computer while the fieldwork is still in progress. This means that data could be served to the fieldworkers who may be able to amend update, and enhance their records and interpretations by easier cross-checking of associated records. It would also mean that information (such as digital photos of enigmatic discoveries) could be available for hard-pressed specialists to view and send guidance and advice without recourse to costly or unnecessary journeys.

Increased feedback from an on-site digital recording system, when combined with the explicit recording of research objectives, becomes a powerful means of increasing the whole project team’s overview, and understanding, of the archaeology on a site or project.

4.3.6 Supporting 'Background Reading' through increased Communication

A real problem for the CfA in functioning as a centre of excellence is the extremely limited time that most people have for reading. The literature review undertaken as part of the Sectoral practice report provided a remarkable opportunity to assess important developments in the field of IS and archaeological theory. This concentrated effort is usually associated with projects. General 'keeping up with the literature' is harder.

CfA staff are encouraged to spend time reading, and this activity has a code for time accounting, but it is seen as a private enterprise. Choosing what to read can take a bit of time in itself. In the end it tends to be project driven or completely *ad hoc*. Some people 'keep up' with specific journals. While this reinforces the expertise in their specialism, it can mean that broader conceptual articles about what we are, or should be, doing can be missed. This is critical given our strategic role. If we write strategy that is not informed by broad reading as well as focussed reading, then it will read as a collection of buzzwords.

Some of this is just a matter of looking for the right balance, but it would also be eased if we saw this as a team endeavour. Most of us take notes of some sort when reading, if we shared these, or some aspect of them, using bibliographic software such as Endnote, we could save each other a lot of time, and create useful dialogue. We could also see who else had referenced an article, and in what papers, and therefore whether it was useful (our very own citation index).

As with all aspects discussed here, software won't solve the problem. CfA Management needs to define what it hopes to gain from 'background reading'. This could range from project specific goals, to CPD, to engaging in strategic development. The key is to implement a strategy to ensure that these benefits are realised.

4.4 Work Practices

While we considered many aspects of work practice in the User Needs Survey Report, we also found our experience of conducting the project within current procedures illuminating in this respect (see discussion in Methods, p. 12 f.). The project has taken six months longer than initially expected. It was re-planned (varied) to increase its scope and to reflect other commitments of project members. These are common problems affecting CfA projects. This section uses our own experience of work practices and considers improvements.

4.4.1 Project Management

Project management is one of the most important sets of procedures for ensuring efficiency. The original Project Design for Revelation was initiated under existing CfA project management systems and procedures and was therefore structured according to archaeological project management requirements. However, it was a non-fieldwork-initiated project examining the very procedures it was running under (see p. 12). As a result we adapted as we encountered problems or gaps in existing CfA systems and procedures.

Many staff were uncertain about procedures for specifying project roles and outcomes, and as a result project programming was variable. Only three days were allocated to project management when the true figure was over 30. The role of the project manager is poorly defined in the current system. The project manager is usually the field director/principal investigator. In many ways the current role is closer to client than manager. This blurs lines of responsibility, since team members see themselves as providing a service to the manager, rather than the manager supporting and managing the work of the team.

"Matrix management" for projects is nominally in place within CfA, but it is not well supported as it relies upon slightly different skills to line management. There is more need to build and maintain personal networks, influencing members of project teams as a project manager without authority, dealing with multiple managers and coping with other teams' multiple (and often conflicting) priorities.

Project management documentation is also weak within the current structure, with little being required beyond project initiation. It was decided early on, due to a greater degree of emphasis on the IS requirements for the project, that it was appropriate to use an IS project management approach over and above the more archaeologically based CfA project management approach (eg MAP2 (English Heritage 1991). Thus the Prince2 project toolkit, that is already used for IS projects within EH, was incorporated into project planning. It was found particularly useful to set out the product descriptions for the project tasks already identified, and these were soon seen to be a valuable aid in identifying more precisely what the nature, scale, format and responsibilities were for the main products of this assessment stage - including this report.

As with many large or complex projects there were some tasks that were not identified, or fully specified, until after the project began. A number of the single tasks identified in the original PD actually required several other related tasks which were not included in time estimates eg task estimates for fact finding visits did not include time for writing up the notes from each of the visits, and compiling a report. Similarly the review of existing systems did not allow time for analysis of the results, or for creating the data-flow diagrams (a task that eventually took 12 days). When project re-planning proved necessary, the mechanisms for doing this were not well defined and caused problems with priority clashes in other areas of work.

With fieldwork there is usually a clear demarcation that the person running the site has authority over the people carrying out the fieldwork. However for projects (and project stages) that are not based in the field the responsibilities and authority of the field director are taken on by the person (not necessarily the original field director) in the role of project manager. In CfA, there is no well-established procedure for the project manager to enforce project management responsibilities, because all the staff on non-fieldwork projects are primarily responsible to line managers, before project managers. This is not to suggest that staff on Revelation were not carrying out their responsibilities, but rather to emphasise that other line-managed tasks inevitably tended to get priority over project managed tasks, because "*that's the job you were taken on to do*".

Some of these difficulties can be dealt with by a management review of project management. They also need to be taken account of in systems design. The INTRASIS team told us that they wished they had put more project management documentation into their first version of their fieldwork software, and will be revising this in further releases.

4.4.2 Programme Management

This assessment project is only an initial stage in a longer-term project. It has been recommended already at senior management level that the best way to take Revelation forward within EH is as a Programme of several projects, each with clearly specified aims, tasks, resources and outputs. Of course, senior management could only agree this after initial scoping work was carried out.

It is always difficult to pre-define the scale of a project. At what point does an archaeological project begin, and how is the information from additional pieces of fieldwork integrated into a project archive? With many CfA projects the first season of excavation is seen as the beginning of the project archive. Often there will have been previous investigations, perhaps geophysics or geochemistry work (not necessarily by CfA) on the site, which need to be integrated as part of the archive. However, the earlier work may initially have been carried out as the equivalent of 'scoping', work and therefore not necessarily anticipating any further archaeological investigations.

An integrated CfA system should allow for internal data management that recognises that each separate project has a folder which is 'owned' by someone. Such a system should allow original investigators to hold on to the identity of their data as a separate project, while making the data available to others planning future fieldwork, analysis or any other wider project stages.

The CfA has a history of building up projects over a long period of time with many projects developing from different research angles (Whitby is the classic example but the Hadrian's Wall work, Richborough and, of course, Stonehenge also fall into this pattern). CfA needs a system that can manage and integrate data from large-scale excavations, but also a series of short investigations with clearly-defined goals that can be understood as a single site over time. We need to be clear where projects are nested within programmes.

4.4.3 Team Working and Communications

Archaeology is essentially a team enterprise. There are very few archaeological projects that can be undertaken by a single researcher, and most of those rely on other people's work in museums and archives. Joining a team in the field is one of the things that draws many people to archaeology. Obviously the intense experience of fieldwork can create difficulties but the sense of 'team' is very strong. When fieldwork finishes this positive identification dissipates rapidly. Specialists who are not involved in fieldwork often feel isolated within projects for this reason. The team responsible for Assessment and Analysis has much less contact, support and communication.

One of the clearest messages that came from the User Needs Survey was that internal communication within CfA isn't working effectively. Ironically, even the message communicated by the report was double-edged. Many people felt they spent too much time “communicating” (eg answering emails) while at the same time a lot of the problems highlighted in the project focus groups derived from poor communication. There was at least one heartfelt request that people would phone, visit and write letters more, which rightly underlines the importance of personal contact.

Within the Revelation project we relied very heavily on email for communication between project meetings. It was good for delivering information to the dispersed team. This was particularly important as many members of the team regularly worked away from Fort Cumberland, and couldn't access the network where project files were stored. It was less useful for discussing issues and maintaining a sense of 'team'.

One possible way that new systems could help with this is to increase the availability and use of ‘social software’, software that supports social interaction (Shirky 2003) over and above email, although this would have to be as a supplement to direct communication, not instead of it. The use of social software may particularly help maintain and develop team communications throughout the extended life of a project.

Arguments for and against the potential value of social software for discussion groups using new web-based mechanisms such as Wikis (a browser-based bulletin board for on-line discussion groups) and Bloggs (a kind of online diary/log book for tracking of ideas and events) emerged during the course of the Revelation project. It became clear that particularly the discussion of new ideas and quick sharing of project developments was exactly the sort of area where bulletin-board style communication would be advantageous. Unfortunately the lack of the software made the communication (and demonstration) of the use of such things more difficult.

It will also be important to recognise where some formal methods of communication need to be agreed and maintained. Project team communication is a different issue to general communication of ideas, and should be under-pinned by guidelines for holding regular Project Team Meetings and establishing protocols about when team members must communicate with the project manager, and when the project manager must communicate with the team. This is largely a matter of good project management and team building, but again requires a culture that enables best practice. It is therefore recommended that CfAMAN define types and levels of communication that are minimum requirements for different programme stages.

Overall the message seems to be that better methods are needed for both informal and formal means of communication. Clearly any IT based solutions can only be successful as part of a much broader cultural change management process. That includes staff agreeing to and adopting changes in practice (*Yes! Cultural change means YOU*).

4.4.4 Individual Working

Although one of the key drivers for an integrated system is the sharing of information, it must also be recognised that the quality of an individual’s work can at times depend upon being able to think without being interrupted. An integrated system should help remove some of the interruptions of having to go looking for information by providing appropriate data when needed.

There are individual thought processes that need to be shared - bouncing ideas off others. But there are also thought processes that are contemplative and single - such as working up a matrix diagram. Any new system (along with procedures, management and culture) will need to ensure there are places that allow users to hold their own “working drafts”. Providing such 'private spaces' may save people finding ways of circumventing or not using the systems. The development of document management protocols and best practice should be established building upon the Tech Refresh roll out and training.

5 Conclusions

5.1 Key results

The current use of IS is both a product of, and a contributor to, the fragmented work practices that undermine productivity at the CfA. A well-designed IS, including procedures, software and hardware would help us meet our obligations under the modernising agenda. It would speed project programmes, allowing us to conduct a more active programme of research covering more issues identified in departmental, corporate and national strategies. It would also make the data produced through this research more accessible and more useful for other researchers, strategists and planners.

Many of the causes of extended project programmes within the CfA relate to management and culture rather than systems development. A weak sense of team working is connected to fragmented management and working practice. These issues are being addressed as part of Modernisation. A comprehensive review of procedures is badly needed. Poor risk management is a broader issue for the organisation. When key people leave, or problems with data arise, or one aspect of the programme slips, we are slow to respond. Projects are put 'on the back burner' and are more difficult to pick up.

No appropriate “Off the shelf” solution exists already within the sector, which can fulfil CfA’s needs and requirements as revealed and identified by Revelation’s work so far. Revelation will similarly not result in an “Off the shelf” solution which can be adopted as it stands by a wide range of organisations across the sector. We will provide a “core” conceptual framework, which can in full or in part be adopted and more-or-less readily adapted by a wide range of archaeological organisations to suit their specific needs.

We have not found a single organisation that manages the whole excavation process digitally. In fact many of the people we spoke to had well developed integrated systems for assessment and analysis of excavation data, but maintained that the data must be recorded on paper. Some of these people employed data entry staff on site, so that a usable product was available during the excavation (Beck 2000). This is a workable approach on large sites with complete single context recording. On smaller sites, or sites with multi-context plans, it puts another strain on an already stretched documentation system.

Planning needs to take the entire project into account. Trying to move more quickly through one stage may make things slower later. Producing a coherent high quality digital archive before leaving the field will make fieldwork more expensive but should bring considerable savings over the life of the project.

Digital recording has real benefits for archaeology. It can produce a more useful product more quickly than paper equivalents. What's more, it allows us to think about archaeological recording in new and interesting ways. But a key benefit of digital recording to archaeology lies in the wide dissemination of our high quality archival material. This will facilitate synthetic work, making use of the huge number of small investigations that characterise contemporary archaeology. This will not be possible until digital recording is standard practice, or at least common in archaeology.

Despite the potential of digital recording, sloppy implementation undermines all benefits. We need to ensure reliability and introduce change only when the system has all necessary supports in place. This requires testing, validation, and demonstration.

Organisational change is always challenging, inertia is a strong force (Moss Kanter 2001). Our users are nervous about changes to their work practice, particularly changes that centralise data management. Any new system will have to address these concerns and offer benefits that outweigh the disruption of change. We must produce a system that *supports and improves* current practice. When new users become familiar with digital procedures, they may well come to want more complex functionality and will be willing to make bigger changes. In our assessment of User Needs (Appendix 1.E) we have found that there is plenty that we can offer people to improve their current practice.

5.2 Future work

The introduction of Revelation will not be a one-off exercise. The implementation of the results of this assessment will require a programme of work stretching over a number of years. Further, it will require an ongoing commitment to development to ensure that the system does not become fossilised.

The system must be used in all archaeological research conducted by CfA. It should make substantial and measurable savings in staff time over the life of a project. It should make our research archive easily accessible to a range of audiences through digital dissemination. Finally it should increase the security of this archive in both digital and traditional formats.

We are currently developing project initiation documents for Stage 2 of Revelation. This stage will cover formal requirements documentation, development of procedures, identification of data standards, design, and production of the system. The overall programme should be arranged to ensure that implementation co-ordinates training, management change and systems change.

A key element in the next stage will be more comprehensive data modelling resulting in an ontology for the CfA based on the CIDOC CRM. The aim is to reduce our understanding of our data to its key concepts, processes and relationships and then using this model as a basis for developing procedures and systems.

Some aspects will require further research and development. Probably the most pressing is the field-recording interface. The review of Field Practice was only a scoping exercise. Considerably more work will be required before these ideas can be used as part of daily practice.

The systems development aspects of Revelation are the simplest in many respects. Development of the theoretical/academic framework is much more demanding. Beyond data modelling, this includes understanding the work processes that make up the Analysis stage of a project. It also includes the conceptual changes that are attendant on changing dissemination method. Many of the recommendations listed below deal with issues that cannot be part of a systems development programme, but are vital to the successful implementation of a system and development of good practice at the CfA.

6 Recommendations

Each section begins with a single page in numbered bullet point style. Further explanations follow where necessary with reference to the numbers on the first page.

6.1 Recommendations for English Heritage Senior Management:

1. **To *endorse* the findings of this report. In particular:**
   a) **Research and Standards** to Support within EH committees and management structures, the recommendations to CfA Management group.
   b) **Archaeology Department Management** to engage with and maintain currency with the project to enable the effective management of internal and external expectations of the project's primary deliverables.
   c) **Archaeology Department and Archaeological Strategy** to endorse new models for archaeological research to the wider profession.

2. **To *provide resources* for the outcomes of this report. In particular:**
   a) **Research and Standards and ISIS** to establish links and secure the assistance of other projects, programmes and initiatives within EH that can provide help with achieving Revelation project aims.
   b) **ISIS** to leverage plans for a corporate web-based GIS Server for use in web-based delivery of Revelation data.
   c) **Executive Board and Archaeology Department** to Provide the necessary resources to complete the programme.
   d) **Archaeology Commissions** to ensure the availability of CfA staff resources for Revelation by keeping Archaeology Commissions casework loadings under review.

3. **To *take responsibility* for some of the report findings. In particular:**
   a) **Executive Board** to recommend a representative to attend the Revelation programme board.
   b) **ISIS** to implement the latest EH Information Policy along with the necessary IS infrastructure to enable sensible data and document management.
   c) **Archaeology Department** to support the need to prioritise Revelation tasks over and above other projects and therefore consider this before taking on new tasks for CfA.
   d) **Archaeology Commissions** to clarify any further dependencies and products required from Revelation for other EH projects or initiatives.

4. **To *address* the training requirements highlighted in this report. In particular:**
   a) **Executive Board to** accept that the introduction of new ICT through Revelation will require complementary investment in training.
   b) **Research and Standards** to address the need for an ongoing programme of CPD within EH to deliver the best results from Revelation.
   c) **Archaeology Department** to build a minimum requirement for reading into FJPs to enable policy, strategy and guidance to be based on current understanding.

5. **EH senior management are asked to *advise* upon the requirements for provision of *change management* in this report. In particular:**
   a) **Executive board** to acknowledge the need for an agreed change management process and procedures - not just for IS projects.

6. 1.1 Further Explanation

1a This report delivers a range of recommendations which are not just about creating a new IS system. For any changes to be most effective they will need support from Archaeology Department managers and other senior mangers across a number of EH committees and management groups, to give the fullest backing to CfA managers and staff to carry out the work proposed.

1b Clear senior management endorsement of the project is important. A lack of senior management engagement with the project could seriously affect how further developments are received and implemented throughout CfA and the rest of EH, and will have a negative impact upon the staff trying to deliver further work on Revelation. Senior managers, particularly in Archaeology, will need to maintain currency with the progress of the Revelation project, and attend external meetings where Revelation is presented in order to deliver accurate and up to date views, both within EH and to the wider discipline.

In the course of this assessment various views have been expressed, even within EH, of what different archaeologists expect Revelation to deliver. Some have talked about “re-engineering” English archaeology and others are considering a need to update or revise the MAP2 guidance for archaeological projects, depending upon Revelation outcomes. A clear message needs to be delivered that MAP2 is about archaeological project management, but that Revelation deals with the whole archaeological process and therefore involves other and more complex models.

From the broader archaeological discipline there have been concerns from two almost opposite viewpoints. Some fear that Revelation will attempt to impose a single, rigid and inflexible “National System for English Archaeology”. An alternative view has also been expressed that if Revelation is carried out using public money then its products should necessarily be made available free of charge to all (tax-paying?) archaeologists.

Regardless of whether developing a “National” system is feasible or desirable, it remains the primary goal of Revelation to produce an IS system that works for the CfA. If any resulting system is successful then it may well be that others will want to adopt it. But it is worth noting that one of the findings of the sectoral practice review was that there is hardly any evidence in England of archaeological organisations implementing systems built for others in recent years.

1c It is also important to emphasise that the Revelation project does not simply require an IT-based solution. The products from Revelation will not just be objects which can be given or sold. Some of the recommendations for systems development may require changes to archaeological practice, which will need to be endorsed and promoted to the rest of the archaeological discipline.

2a EH is already in the process of modernising many aspects of its business, and some of these changes, particularly those with an IT element may be of significant help if co-ordinated with Revelations aims.

The planned provision of new IT equipment under the project known as Tech Refresh, will provide an important platform upon which Revelation can plan for much more up to date systems development. Advice on how to maximise the benefits from Tech Refresh and other such corporate IS initiatives, will be of benefit to future work on Revelation. The availability of new software following Tech Refresh should provide opportunities to increase general levels of IT skills amongst CfA staff in order to both maximise the benefits of the Tech Refresh programme, and introduce changes required by Revelation.

2b It is recommended that the publication route of CfA reports and dissemination of results should follow a web-delivered dissemination strategy. One opportunity for this would be to make space available on a corporate web-GIS server. By doing so information from CfA projects could be made available internally to the rest of Archaeology Department; to other Departments and Regions; but most importantly, where appropriate, published in the public domain.

2c This assessment stage will result in a project design setting out a further programme of work to produce a new system. This programme will have associated costs that will need to be agreed. Also, as part of this Revelation assessment, a number of more innovative R&D elements have been identified (eg the use of handhelds and PC Tablets for field recording). The development of these R&D elements might be accelerated by incorporating them with other corporate initiatives for developing mobile computing solutions.

2d At a staff level, the time required to design, develop and implement the new systems and procedures will be considerable. Any additional resources that can be found by other teams to alleviate the numbers of CfA staff working on monitoring the Archaeology Commissions Programme will be of benefit to the Revelation project.

3b Implementation of better corporate systems for document management would enable staff to follow the new policy rather than feeling at odds with it under current provisions.

3c The current assessment stage has been completed while continuing to carry out other existing commitments, but this has considerably impacted upon its timetable. It is important that the next stage of the project does not get de-railed because other work is suddenly given greater priority.

3d It is important for those who have genuine requirements from the implementation of the new system specify these clearly in the project planning and at the programme board so that they can be incorporated in the design, and genuine outputs delivered that match expectations. In addition it would be helpful if those with requirements can also specify where they have elements or resources they can put back into the project.

4a A recent OECD report notes that "*The impacts of ICT depend on complementary investments, eg in appropriate skills, and on organisational changes such as new strategies, new business processes and new organisational structures*" (OECD 2003)

According to current time recording data the Archaeology Department spends less than 2% of its time on training. Ideally the department would like to see this nearer 10% and it is suggested that a rapid increase to at least 5%+ for CfA will be needed if Revelation is to be implemented effectively.

4b One possible method for addressing this would be for EH management to consider enabling sections of the organisation, such as CfA, to apply for the Investors In People (IIP) accreditation.

4c Some of this may be done pro-actively by setting out specific time allocations for research reading in FJPs, along with better follow-up mechanisms for genuine knowledge management and information sharing, which could be included within Revelation (eg using networked bibliography software such as Endnote).

5a This is required in the first instance by CfA for the next stage of Revelation, but is an area that Archaeology Department and EH as a whole may need to consider. Some of the issues in the report about change relate to broader corporate practice beyond the remit of just the CfA. There are cultural questions about the delivery of management information. Communication is a two (or three) way process - not just a matter of pushing out information. The key is to capture the ideas that are flowing in communication, rather than making paper trails.

6.2 Recommendations for the CfA Management Group

1. **Understand** and **Endorse** the findings of this report and the consequent recommendations.

    a) **Recognise** that the systems development aspects of Revelation are simple in comparison to development of the theoretical/academic framework.

2. **Champion** the project and proposed changes to both CfA staff and middle/senior management in particular.

    a) **Establish** agreed procedures for "change management" including "day zero" implementation.

3. **Recognise** that the introduction of Revelation will not be a one-off exercise.

    a) **Establish** a Revelation Programme Board.

    b) **Plan** for continued development on a long-term basis.

    c) **Explore** methods to replace the systems development post and expedite recruitment in Archives section.

4. **Provide** resources from the CfA budget to support projects proposed as part of the Revelation programme.

    a) **Assign** a clear and high priority for Revelation projects within the CfA team plan.

    b) **Offer** advice and assistance in the production of business cases for Revelation project bids to EB.

5. **Conduct** a review of functions and procedures at CfA, leading to the **introduction** of agreed and functioning procedures.

    a) **Complete** the current CfA manual and review the recording system.

6. **Establish** a proactive training and development strategy aimed at acquiring necessary skills and linked both to team plans and individual programmes.

7. **Strengthen** CfA's research/professional identity.

    a) **Reintroduce and support** the CfA seminar series.

8. **Complete** the review of project planning.

    a) **Clarify** responsibility for decision making, lines of communication and authority within projects and **promote** a culture of completion.

    b) **Review** the role of principal investigator/ field director and separate this from project management of the same project.

    c) **Clarify** the need for active and strong line management and conflict resolution.

6.2.1 Further Explanation

1. It is crucial that the report findings are understood and accepted as well as endorsed (ie not just "rubber stamped").

1a The theoretical/academic framework includes understanding the work process inherent in the analysis stage of a project and the conceptual changes attendant on changing methods of dissemination. Most of the recommendations made by this report deal with issues outside a systems development programme but are vital to the successful implementation of a system and the development of good practice at the CfA.

2 This "championing" should include a readiness to provide answers with regard to what the Revelation programme is and is not and the expected return on investment (time and money). It will also be necessary to maintain the commitment and high priority given to Revelation in the face of both demands made by middle/senior management with regard to new projects, and also possible resistance to change from staff.

2a Change management must be taken seriously and not left to chance. The ultimate responsibility for managing the implementation of Revelation should be clearly defined and assigned, and appropriate resources allocated to achieve this. A day zero approach to the introduction of the main core of the programme is fundamental to success. Other aspects may require staged implementation.

3a The Revelation Programme board will oversee the implementation of the recommendations **-** its composition should be determined according to standard EH procedures.

3b Although a day zero approach is strongly recommended for most aspects of the Revelation programme, many aspects will need appropriate support during subsequent staged introduction and adjustment.

3c This post became vacant in December 2000, was subsequently frozen and in August 2002 it was lost as a consequence of savings made as part of EH's Modernisation process.

4 While there will be other funding bids made, the CfA must commit resources (time and money), for development and training, equipment and accommodation. The time commitment must be clear in Forward Job Plans.

5 There is a need to map and assess functions, update some procedures and create others. The function map should be reduced to its simplest form (the tube map approach).

5a The Recording Manual should be completed as soon as possible. It should then be subjected to rigorous practical testing in the light of Revelation's findings. This will ensure its usefulness on a day-to-day basis. There is also a need to review the recording system within a similar frame of reference. These are key early stages in implementation of the Revelation programme.

5b It is recommended that CfAMAN define/adopt procedures, including types and levels of communication, that are minimum requirements for different project stages.

6 This will of course primarily be 'Revelation' training, but should also include considerable organisational awareness, and other CPD, especially in management. It is important that the necessary skills are acquired by staff and that management provide the opportunities and impetus for appropriate training.

7 This should be done in whatever way possible – both with regard to the outside world and, even more importantly, with regard to CfA itself.

7a This can best be achieved by giving high priority to extended programmes of talks by internal and external speakers on “serious” topics (archaeological and non-archaeological), and a culture of attendance – an important part of everyone’s contribution to CfA

8a In order to address CfA’s “culture of non-completion” it is necessary to give the satisfactory completion of projects a high priority. This requires the application of effective management with clear decision-making, efficient lines of communication and appropriate credit for, and acknowledgement of, completion.

8b The project findings strongly suggest that separation of the roles of project manager and principal investigator/field director in the same project is very advantageous. This would give the project manager the necessary “distance” from the project to enable them to manage efficiently.

8c Conflicts are inevitable - regardless of how well thought out and implemented the Revelation programme is. It is easy to imagine conflicts arising between the processes of line and project management, or from the reluctance of some groups or individuals to adopt some or all aspects of the Revelation Programme. It is crucial that CfA Management takes on the clear responsibility for resolution of these conflicts so that they are not permitted to “de-rail” the process.

6.3 Recommendations for CfA Staff

1. **To acknowledge that there are problems due to the lack of a coherent approach to data collection and management.**
    a) **Recognise** that ‘non-completion culture’ is undermining CfA's professional standing.
    b) **Acknowledge** that the CfA operates as a series of fragmented cultures: there is a lack of cohesion in work perceptions and practices.
    c) **Acknowledge** that all EH data is used by others. Appreciate that this can be a beneficial symbiotic relationship.
2. **To take appropriate responsibility for necessary changes in culture and work practice. In particular staff need to:**
    a) **Understand** that cultural change needs to include every individual within CfA and it will require acceptance and engagement by all.
    b) **Challenge** mistrust between CfA staff with respect to sharing data and adherence to management protocols.
    c) **Increase** of awareness of and interest in the work of other CfA staff especially where they may impact on each other.
    d) **Expect** clear priorities and support from management.
3. **To establish clear ownership and responsibilities within projects. To improve teamwork and communication of data:**
    a) **Adopt** team-based approach to collecting data and data handling.
    b) **Clarify** team communication. It needs to be focussed and more effective, not taking more time.
    c) **Allow** more time to produce an overview of a project to help develop interest in other aspects of the project and a sense of their place and value within a project.
    d) **Build** more time in to projects to make data accessible and usable. This includes defining terminology and holding metadata.
    e) **Tailor** outputs to audiences. Non-specialist users need to understand how data relate to conclusions.
4. **To improve their knowledge of and use of existing internal procedures for project, collections and data management.**
    a) **Get involved** in streamlining, standardising and reinforcing procedures and terminology.
    b) **Prioritise** transparent and consistent collections and data management.
    c) **Understand** the context and purpose of their data. **Define** why and for whom they are collecting data.
5. **To include Revelation in their long-term planning**

a) **Recognise** that Revelation will be a long-term commitment requiring adjustments and re-directions.
b) **Adopt a** 'Day Zero' approach to including data on the new system.
c) **Reassess** methods of task estimation to ensure that time scales and deadlines for projects and tasks are realistic.
d) **Expect** the Revelation system to reduce the overall length of projects.
e) **Identify** their Revelation related training needs and build these into their forward job plans including secondary training

6.3.1 Further explanations

CfA staff are central to the success of Revelation. The aim of these recommendations is to bring about benefits to CfA staff through increasing the speed and efficiency of analysis and dissemination of the fieldwork results. The recommendations are based very substantially on the results of the User Needs report.

1 The Sectoral Practice report shows these problems are not unique to the CfA. We are, however, in a position to demonstrate leadership by addressing them.

1a Completion of projects is not celebrated and this undermines staff morale. We can sometimes think 'why should I bother?' Further, the pressures of other responsibilities put research on the back burner, projects have a history of slipping. Deadlines are seen as flexible. We need to focus on completion of what's in hand before taking on new tasks.

1b The fragmented nature of the CfA is largely due to historical circumstance. Fragmentation needs to be challenged both because it undermines our collaborative work practice and because it can present conflicting views to our clients, both corporate and external.

1c It needs to be recognised that the data each individual gathers needs to have a life outside the individual (creator), and the project and be comparable with other data. Accessible documentation, feedback on work and access to data produces effective, good quality data.

2a The user needs survey indicated that many people felt that they followed procedures but that others did not. Everyone within the CfA has both good work practices and those that need improvement. We all need to work actively to bring about change.

2b Many staff members are reluctant to share data with others because they worry that others will either damage it or misuse it. We need to build trust within teams to enable them to work effectively. Similarly, procedures are sometimes viewed as unreasonable constraints on professional judgement. We need to use them as statements of agreed process, endorsed and followed by all staff.

2c The historical circumstances leading to fragmented work practices have also led to a lack of knowledge of our colleagues and their fields of expertise. We are often unaware of support and advice that may be available, and as a result waste time replicating tasks or doing them less efficiently. Also we may underestimate our colleagues interest and knowledge of our own research and miss chances for useful collaboration.

2d As in many workplaces, rapid and repeated structural change has produced a sense that management concerns are divorced from the realities of our work programmes. As professionals we often feel able to prioritise our own work. Nonetheless, it is the role of management to provide strategic priorities, and we should expect them to help us in dealing with conflicts arising from the pressures of our broad range of responsibilities.

3 All archaeological research is collaborative by nature. We often feel isolated within teams, especially when we aren't in the field. But we rely on data, and on other support from project teams at all times. The sense of teamwork needs substantially more support than it currently has.

3a Although we work within teams of expertise, we also work in project teams. All data for a project should be collected and managed with the team in mind. This will prevent bottlenecks in the transfer of data, duplicating of tasks and double handling of data. Data handling should not be a constraint on the project.

3b Many of the problems highlighted during the focus group meetings stem from poor communication. Many staff already invest a lot of their time in communication and indicated that they would like to reduce this. It is therefore vital that means for more effective communication of data are found.

3c Staff often focus quite closely on their own role in a project, partly in order to accommodate busy schedules. This undermines teamwork, causing delay in the long run.

3d Data management is rarely budgeted for as a separate task in project designs. People end up doing this kind of work as an add-on. Sometimes data are left inaccessible. Sometimes the project over-runs. All projects will need time for this, though a well-developed system should establish a baseline for most projects.

3e The use of specialist language and terminology needs to be balanced against the need for the non-specialist to understand and use the data (without having to pursue the specialist)

4a The next stage of Revelation will involve a range of projects to define procedures, standards and terminology. These need to be staff-led to ensure that these arrangements work and are followed.

4b Many of the problems outlined in the user needs report were not related to poor systems, but to systems not being implemented. We all need to take responsibility for the day-to-day management of collections and data.

4c The next stage of Revelation will involve projects to describe our processes in more detail (to ensure that the systems support them). This is an opportunity to be explicit about things which we currently understand implicitly (with the inevitable confusion).

5b Legacy data are not going to go away and are not going to be ignored by Revelation. The clear message from the Sectoral Practice report was to design our system with what we are creating now i.e. to create a “Day Zero” where everything after a certain date would use the new system. Backlog projects will be dealt with on a migration basis, according to the CfA digital archiving strategy

5c Task estimation is a difficult process relying on experience and prediction. Many projects in all organisations slip because the original estimates were unreasonably conservative. We need to make more use of our past experience to produce realistic estimates that don't slip. This is essential for programming planning since slippage in one project affects resources for other projects.

5d Revelation will be a flexible system. This is not one that allows data to be of variable quality, rather it will take what has already been entered and generate new and useful outputs, saving time in the long run.

5e For example, learning to use standard computer packages to feel more confident and ensure that everyone is speaking the same language.

6.4 Recommendations for Systems Development

The recommendations for CfA Management Group call for the establishment of a Revelation Programme Board (hereafter RPB) to manage the development, implementation and maintenance of Revelation systems. The RPB are recommended:

1. **To adopt a general design philosophy to:**
   a) ***Design*** the system around the needs of the users.
   b) ***Base*** the system on modular design with basic high-level design completed before any modules are built.
   c) ***Support*** digital data capture in the field.
   d) ***Support*** group working through system design.
   e) ***Ensure*** the system design supports early dissemination, both digital and traditional.
   f) ***Provide*** a mechanism for effective change management.
2. **To endorse the following concepts and:**
   a) ***Base*** the system on a spatially-indexed database with appropriate interfaces for spatial, graphic, numeric and textual entry, manipulation, and dissemination. It should be written to industry standard and professionally documented to allow upgrading.
   b) ***Identify*** elements that make up the practical minimum or “core” and develop them together for initial roll-out.
   c) ***Allow*** project directors some discretion to customise interfaces, and protocols at the beginning of projects to better link data collection to research questions, within parameters and procedures set by the RPB.
   d) ***Investigate and implement*** systems of version and quality control.
   e) ***Implement*** the CfA digital archive strategy through system design, including object tracking and collections management, and work with the Data Services Unit to define appropriate standards for form and content.
   f) ***Extend*** the data modelling exercise to include all of CfA and comply with CIDOC CRM standards.
   g) ***Conduct*** a process modelling exercise to explore poorly understood or variable processes in the Analysis stage of projects.
3. **To design the system around adequate resources and support and:**
   a) ***arrange*** appropriate professional systems development resources.
   b) ***Consider*** it as an EH corporate system to be maintained by the EH IT Facilities Management company (currently Schlumberger-SEMA).
   c) ***Ensure*** that proper training is developed alongside the system.
   d) ***Develop*** a programme of R&D to facilitate cutting-edge developments such as digital recording.

6.4.1 Further explanation

1a It must be intuitive to use. It must support dynamic querying by users, rather than a set of rigidly defined queries that would stifle creativity.

1b The whole system must be designed as a complete entity to avoid the pitfalls of incompatibility with new features. This design will leave room for growth as new tasks are identified, but must set the framework in which new functions can be developed. Individual modules can then be developed, written and implemented in a rolling programme.

1c Digital capture in the field is the first step towards streamlined dataflow. The Revelation programme will need to embrace R&D on digital drawing, handheld computer technology, wireless networking, web-served data linked to site systems, and web publication to make real gains from on-site recording. Development of the system must support this research.

1d Design must support group working, with all data seen as belonging to the project team rather than individuals. Data must be available to the whole project team (including remote access) as early as possible, allowing feedback to inform data collection. Protocols must ensure access to all, while modification rights are carefully defined throughout the life of a project based on project-defined goals to avoid corruption of data. The system should provide updates on recent changes to database when users log on.

1c, 1d, 1e, 2d To ensure the high standard of data and allow traditional “archive completion” tasks to happen while fieldwork is still in progress, the system must support automatic quality-checking of entries based on existing data within the database. This includes dynamic matrix-building and cross-referencing of records, which should highlight inconsistencies instantly while they can be easily investigated.

1e Dissemination strategy must include on-line reports and data.

1f Design must include a system of effective change management that allows (and even promotes) development of new features as needs are identified, and ensures that components are kept up to date with proprietary software developments. The system must also support research and development, but remove the need for numerous not-so-temporary *ad hoc* systems.

2b The system needs to be of benefit when launched, so must have sufficient modules to replace key parts of our current data-gathering system. These should focus on known process bottlenecks to gain instant visible benefits.

2d There must be robust security controls. The system should track edits/authors by tagging all data with user IDs and dates, subject to terms of Data Protection Act. It must also provide audit trails and allow roll-back.

2e The system should provide automatic creation of metadata.

3a & 3b Development and support cannot be left to CfA alone, or to *ad hoc* attempts to update the system. Making it an EH corporate system will ensure professional systems development resources are available for the continual support and evolution of the system, as well as providing an overview that will allow greater integration with other EH corporate systems.

3c Training can be divided into “driver” and “mechanic” levels of use as appropriate to the responsibilities of the project team members. All users will need basic training to collect and interrogate the data (the “drivers”) as part of their jobs, while a select few would need detailed knowledge of the inner workings of the system (the “mechanics”) to keep it running.

6.5 Recommendations for the Sector

EH has a role in establishing standards for archaeology in England – through the publication of guidelines, the provision of advice and the commissioning of archaeological research. Considering this role, we feel that – beyond the scope of establishing requirements for an archaeological recording system for the CfA – many of the results of this assessment will be of use to other people designing such systems. The recommendations and suggestions listed below are results of the assessment which we feel deserve particular consideration by a wider audience. Our thanks go to the organisations which we visited as part of this assessment and we explicitly acknowledge that our recommendations draw partly on good practice seen elsewhere.

1. **Training and *Continuing Professional Development (CPD)***
   We suggest that Practitioners and Organisations:
   a) **Engage** better with computer information systems and recognise the potential benefits offered by these systems to the practice of archaeology and disciplines associated with it.
   b) **Provide** more INSET training for field staff.
   c) **Identify** specific training needs of senior staff to keep up-to-date with archaeological practice in fieldwork.
   d) **Consider** a Change Management Process throughout the sector.
2. **Digital dissemination. Considering the feedback from the various surveys carried out in this assessment, we are certain that the Sector would benefit from:**
   a) **Recognising** the benefits of digital dissemination and publication of archaeological research and plan for it as early as possible in the process.
   b) **Giving** “Good Practice” specifications when asking for submission of digital data from archaeologists/specialists.
   c) **Enabling** the use of primary data across the Sector.
   d) **Following** “Good Practice” concerning intellectual property and copyright for synthesis of others primary and archival data derived from web-based (or other non-traditional) sources.
3. **Assessment of existing work practice. Other practitioners may want to use the Revelation results (systems) if appropriate. However, when starting a similar assessment, we think it might be worth remembering the following:**
   a) Revelation cannot provide an “off the shelf” system for everybody involved in English Archaeology.
   b) Revelation might be useful as a starting point for discussion in units/organisations. To make it work for you it will, however, need your own feedback, subsequent revision and adaption of your working practice.
   c) The data models you come up with may have a wider application and therefore deserve publication.
   d) Software development will achieve best results when done professionally, rather than on an *ad hoc* basis by those who do not have sufficient expertise as software developers.

6.5.1 6.5.1 Further explanation

1 There is a substantial need for CPD programmes in archaeology, particularly for IT, but also other areas of the sector. As well as improving systems, obtaining better data depends on how people are working. TORC is addressing these problems and we need to keep them abreast of issues surrounding IS.

1a The perception of IT as a threat to some people, sometimes even in senior management positions, needs to be overcome. Computer literacy will no doubt improve in the years to come, due to the more positive attitude of the younger generation towards IT. However, there is a need to have “multi-taskers” as field archaeologists with proficiency in survey, databases, excavation skills, GIS and good communication skills.

1b Improvement of the quality of data collected on site depends on enabling site staff to record the best quality data they can. Apart from designing an adequate system, this issue needs to be addressed by providing relevant training. To meet this need for training some people during the review of sectoral practice suggested INSET training days on site. By ‘INSET training’ we mean training in specific activities/tasks prior to the actual start of fieldwork, provided by the relevant specialists to the whole team at the job site during working hours, or at relevant points during prolonged fieldwork. This kind of training should become an integral part of work programmes including PPG projects.

1c Some interviewed as part of the Review of Sectoral Practice pointed out that many project managers and senior staff cannot maintain currency with archaeological practice in fieldwork. This needs to be addressed by providing more CPD for project managers and similar staff to keep them involved in fieldwork and survey, and continually develop site supervisor skills and general post-ex skills.

1d The introduction of a widely adopted culture of digital recording, from excavation through to publication/dissemination, will need to be reflected, or perhaps even driven by, a change management process in the sector as a whole. Currently there is a feeling of patronage in how people are given access to technology. We need to overcome the element of prestige attached to the use of technology reported in the review of sectoral practice (eg Survey is seen as another “specialist” discipline; The sense of “I wouldn’t trust my diggers with that”). Equally, changes need to be considered to the way archaeology is taught in universities and elsewhere.

2a The review of sectoral practice indicated that peer group prestige and academic kudos is mostly attached to “paper-based” publications. To facilitate a quicker and more cost-effective way of dissemination / publication of excavation results, a change of attitude towards digital dissemination/publication is needed throughout the profession.

2b Digitally well-structured archives are an essential pre-requisite for dissemination and use online, both by the unit that provides them and national online resources, such as OASIS or the ADS catalogue. Models for ‘good practice' specifications for profession-wide dissemination already exist in the format of the CfA Guidelines (eg dendrochronology, geophysics, environmental archaeology, waterlogged wood, leather, textiles and human bones).

2c The User Needs Report showed that there is a wish for better access to data sets generated by other members of the same project team. We feel that it would make sense to facilitate a profession-wide access to this kind of primary data. For it to be of proper use to people outside the initial project team it will, however, need necessary interpretation, metadata and documentation attached to it.

2d Should the previous recommendation find widespread acceptance throughout the sector, it will become increasingly important that the intellectual property/copyright of authors or creators of primary data available from web-based sources is acknowledged as such, and cited accordingly, if their work is incorporated in synthesis or analysis elsewhere. We therefore suggest that the profession adopts the model provided by the ADS in making digital archives available:

(see http://ads.ahds.ac.uk/project/goodguides/excavation/sect62.html).

3 Considering the scale of an assessment like Revelation, we are fully aware of the fact that other organisations could probably not afford to carry out an assessment on that scale (see above in Sectoral Practice).

3c Since an explicit discussion on team work in archaeological process (especially post-ex) is lacking in current literature, the feedback, constructive criticism and your own experiences you will come up with when conducting a "Revelation"-type revision of your work practice, will no doubt be of value to other members of the archaeological community.

3d Considering the scope of different applications that will have to be integrated, only IT professionals will be able to provide the required expertise to make a system like Revelation work. *Ad hoc* "solutions" might alleviate problems in one area but could lead to even greater ones in other parts, or even all, of the system.

Glossary of all technical terms and acronyms

6.5.2 Acronyms

ABCD	ArchaeoBotanical Computerised Database
ADS	Archaeology Data Service. The ADS promotes good practice in the use of digital data in archaeology, it provides technical advice to the research community and supports the deployment of digital technologies.
ALSF	Aggregates Levy Sustainability Fund. The Department for Environment, Food and Rural Affairs (DEFRA) have made available resources, through the Fund, for an initial two-year pilot scheme. It is managed by EH (in conjunction with English Nature and the Countryside Commission) and provides funds to help tackle a wide range of problems in areas affected by the extraction of aggregates.
AML	Ancient Monuments Laboratory, English Heritage. Co-located In 1999 with the CAS to form the Centre for Archaeology.
ArcView	ESRI's desktop GIS software.
ASCII	American Standard Code for Information Interchange. The basis of character sets (numbers, letters and other keyboard characters) used in almost all present-day computers.
AutoCAD	Computer-aided drafting software package developed by AutoDesk
Blog	A blog, or weblog, is a personal Web site updated frequently with links and commentary.
BMP	Bitmap digital image format.
BUFAU	Birmingham University Field Archaeology Unit
CAA2003	Computer Applications and Quantitative Methods in Archaeology Meeting, Vienna 2003.
CAD	Computer Aided Drafting.
CAS	Central Archaeology Service, English Heritage. Co-located in 1999 with the AML to form the Centre for Archaeology.
CASPAR	Central Archaeology Service Project Archives Record.
CD-ROM	Compact Disk – Read Only Memory.
CEU	Central Excavation Unit, English Heritage. Renamed CAS in 1992? then co-located with the AML to form the Centre for Archaeology.
CfA	Centre for Archaeology, English Heritage.
CfAMAN	Centre for Archaeology Management Group.
CIDOC CRM	The International Committee for Documentation of the International Council of Museums (CIDOC) Conceptual

	Reference Model. Now the agreed model for systems design in EH.
CPD	Continuing Professional Development.
CPU	Central Processing Unit.
CSV	Comma-separated values file format.
3-D	3-dimensional.
DAP	Digital Archiving Project.
DEFRA	Department for Environment, Food and Rural Affairs.
DigIT	Digital recording and Information Technology project.
DSU	Data Services Unit. An EH team based in Swindon responsible for establishing data standards and interoperability within EH.
DTP	Desk Top Publishing. Software which produces files which can be sent directly to printers.
DXF	Drawing Exchange Format file format for AutoCAD.
EAA2003	European Association of Archaeologists 9th Annual Meeting, St. Petersburg, Russia 2003.
EB	Executive Board – the principal English Heritage managerial decision-making body.
EDM	Electronic Distance Measure, part of a TST.
eGMF	e-Government Metadata Framework.
EH	English Heritage.
ESB	Environmental Sciences Branch, CfA.
FDA	Field Digital Data Capture.
FIS	Financial Information System. EH's software for managing purchasing.
FISH	Forum for Information Standards in Heritage. National body, largely conducts business through a mailing list, with occasional conferences. Advises on standards of content and form.
FJP	Forward Job Plan. The individual planning element of the EH appraisal system.
G-sys	Excavation GIS software developed by Dominic Powlesland at the Landscape Research Centre.
GIS	Geographical Information System.
GPS	Global Positioning System, a satellite navigation system.
HEIST	Historic Environment Information System Team.
HSIS	Heritage Spatial Information System. The bespoke GIS designed for EH by IBM as part of a private finance initiative. It has 30 seats in London and 30 in Swindon and holds statutory data (designations) as well as the spatial data of the NMR.

IADB	Integrated Archaeological Database. Originally designed by SUAT, now based at York Archaeological Trust, but available for use by other units.
ICT	Information & Communications Technology.
ID	Identification.
IMC	Information, Management and Collections, CfA (renamed Archaeological Archives).
INSET	In Service Training.
INTRASIS	Intra-site Information System. Recording and Analysis software for archaeological fieldwork in Sweden. Designed by the state archaeology service and widely used in the profession.
IPAQ	Compaq/Hewlett-Packard's brand name for its version of the handheld computer (PDA).
IS	Information System.
ISIS	IS Integrated Strategy - EH internal committee.
IT	Information Technology.
Jnet	Nick Ryan's stratigraphic modelling software for visualising, and editing the stratigraphic sequence (matrix). Previously called Gnet, but re-written in the Java programming language.
JPG	Probably the most common and current image compression format.
LAN	Local Area Network.
LRC	Landscape Research Centre. An Archaeological Unit with a focus on the Vale of Pickering.
MakeDXF	Bespoke survey software written at CfA. Creates DXF and CSV files of survey data.
MAP2	Management of Archaeological Projects 2.
MIS	Management Information Systems.
MoLAS	Museum of London Archaeology Service.
MPP	Monuments Protection Programme (English Heritage).
NAPEX	New Approaches to Post Excavation. A project funded by Archaeology Commissions which aims to reassess post-ex practice in English Archaeology.
NMP	National Mapping Programme. This EH programme aims to identify and record all archaeological sites and landscapes visible on aerial photographs.
NMR	National Monuments Record. The national version of an SMR.
NMRC	National Monuments Record Centre. The offices of the NMR, in Swindon.

OAD	Oxford Arch Digital. An independent software company focussing on development for archaeology. A spin-off from the Department of Archaeology in the University. Not connected to the Oxford Archaeology Unit.
OASIS	Online Access to the Index of Archaeological Investigations.
OAU	Oxford Archaeology Unit (now Oxford Archaeology).
OECD	Organisation for Economic Co-operation and Development.
OS	Ordnance Survey.
PalmOS	Palm Operating System used on some PDAs rather than Microsoft software.
PC	Personal Computer.
PD	Project Design.
PDA	Personal Digital Assistant. A handheld computer with functionality focussed on communication and workflow.
PNG	Portable Network Graphics. An emerging digital imaging format that gives "lossless" compression.
Post-ex	Post-excavation. All of the archaeological work that occurs outside of the field.
PPG	Planning Policy Guidance (Governmental Guidance notes).
PPG15	PPG15 is the Government's guidance on Planning and the Historic Environment.
PPG16	PPG16 is the Government's guidance on Archaeology and Planning.
Q&A	Questions and Answers.
RAM	Random Access Memory.
RCHME	Royal Commission Historic Monuments and Environment.
R&D	Research and Development.
SEM	Scanning Electron Microscope.
SF	Small Find.
SGML	Standard Generalised Mark-up Language. A mark-up language like HTML but more powerful due to not having a fixed set of tags.
SMR	Sites and Monuments Record.
SSD	Site Subdivision.
SUAT	Scottish Urban Archaeological Trust.
TORC	Training Online Resource Centre for Archaeology.
TST	Total Station Theodolite.
TWA	Trust for Wessex Archaeology now Wessex Archaeology.

Unix	The operating system used to run the CfA servers (Madras, Robin and Wren) and the applications on them.
UPD	Updated Project Design.
VR	Virtual Reality.
WAC5	5th World Archaeological Congress, Washington DC 2003.
WAN	Wide Area Network.
XML	Extensible Markup Language.
XRD	X-Ray Diffraction.
XRF	X-Ray Fluorescence.
YAT	York Archaeological Trust.
Y2K	Year 2000.

6.5.3 Technical Terms

Archaeology Commissions	The English Heritage Archaeology Commissions Programme is a central strategic grant budget that enables EH to fund important archaeological activities and strategic initiatives, to fulfil EH's remit under the 1983 National Heritage Act. The programme is explicitly linked to the EH research agenda Exploring our Past 1998 Implementation Plan, with an emphasis on strategic national research.
Attributes	Data that describe properties of a point, line or polygon in GIS. In CAD attributes are non-spatial data appended to a symbol (eg of a Block entity in AutoCAD).
Data Protection Act	Data Protection Act 1998 on the protection of individuals with regard to the processing of personal data and on the free movement of such data.
Delilah	DOS flat file database for storing and manipulating text data from CEU / CAS / CfA projects. Long overdue for replacement.
Endnote	Software for managing bibliographic references in word-processed documents.
Framework Archaeology	Framework Archaeology is a joint venture of Wessex Archaeology and Oxford Archaeology set up to work at BAA airports. It is carrying out excavations at Heathrow Terminal 5 and Stansted.
Geomatics	The profession of metric survey.
Georeference	To position a drawing in geographic space on a computer system using known coordinate points.
Hammerhead	Ruggedised handheld pen computer with A5-size screen, runs on standard software.
Labfile	An Oracle database with a wide range of functions including managing the movement of objects, recording conservation work carried out, images created (photos, x-rays) and analyses carried out. For all projects.
Lotus	Software suite similar to Microsoft Office.
Matrix	A graphical representation, most commonly of the type developed by Ed Harris, using numbered boxes and lines to represent the logical relationships between archaeological units of stratigraphy. Can also be used to model relationships between other archaeological entities such as plan drawings.
Munsell chart	Reference standard for soil colour description.

TabletPC	Portable computer, similar in size to a laptop, designed for portability and usually with a rotating screen. It has a touch sensitive screen that is its main data entry mode, and has well developed handwriting recognition software as standard.
TheoLT	Survey recording software developed by Bill Blake of EH. Connects TST directly to CAD software for direct creation of drawings.
Pencentra	Handheld pen-computer by Fugitive, runs on Windows CE.
Penmap	Pen-based computer drawing software developed by Richard Trainer. Usually used with at TST.
Permatrace	Drawing film used as a more robust alternative to paper for recording archaeological drawings.
Photogrammetry	Stereo photography used to record elevations and plans on computer.
Polygon	The area defining a single enclosed entity in a computer drawing.
Polyline	A single entity consisting of one or more line segments between sequential vertexes, defining an open or closed area. It is distinct from a Line entity, which is a single line segment only between two vertexes.
Prince2	PRojects IN Controlled Environments. PRINCE was first developed by the Central Computer and Telecommunications Agency (CCTA) now part of the Office of Government Commerce (OGC) in 1989 as a UK Government standard for IT project management.
Raster	Raster files hold attributes in the form of a grid. This is useful for continuous data (such as elevations or geophysical survey data) particularly prior to further interpretation. Images can also be raster files.
SEMA	SchlumbergerSema are EH's IT Facilities Management contractors.
Tech Refresh	English Heritage Technology Refresh programme
Vector	Vector files hold attribute data in the form of linked points creating bounded entities in a drawing - eg polylines or polygons of context edges. This is most useful for spatial data that has been interpreted or categorised to include context record information.
Wiki	Wiki is a piece of server software that allows users freely to create and edit Web page content using any Web browser. The word 'wiki' is Hawaiian for 'quick'.
Windows CE	Windows CE operating system used on a range of device types but particularly mobile handhelds.
Word Perfect	Text processing software.

Acknowledgements

We would like to thank all CfA staff for their input, experiences and expertise. Thanks to everyone within the profession who generously shared their experience of IS in archaeological fieldwork. In particular thanks to Tony Austin, Catherine Hardman, Simon Buteaux, Glyn Barrett, Vince Gaffney, Dominic Perring, Geoff Avern, Dominic Powlesland, Anthony Beck, Duncan Lees, Phil Treveil, Natalie Cohen, Angela Wardle, Jackie Keilly, Anne Davis, John Giorgi, Nick Shepherd, Paul Backhouse, Laura Highmarsh, Paul Miles, Niall Donald, John Lewis, Emma Noyce, Catriona Gibson, Fraser Brown, Nicholas Cooke, Chris Brayne, Tom Goskar, Jens Neuberger, Mike Rains, Häken Thorén, Karin Lund, Henrik Skousen, Craig Cessford, Tyler Bell and Gary Lock. Jo Cooper cheerfully entered the massive set of references into our literature search database. The Senior Responsible Officer for the project was Brian Kerr (Head of CfA Archaeology) and the Project Board consisted of Brian Kerr, Andrew David (Head of CfA Archaeological Science) and Mary Walkden (for the time when she was Acting Head of Archaeological Resources). We received support and advice from many colleagues within EH, particularly Adrian Olivier, Nick Seal, Chris Mould, Michael Crich, Chris Scull, Bernard Thomason, Simon Probert, Paul Everson, Jacqui Huntley, Jane Sidell, Ian Panter, Dominique DeMoulins, Jonathan Coad, David Gander, Edmund Lee, Mike Evans, Nigel Clubb and Karen Kennedy. Opinions are those of the project team, as are errors and omissions.

Bibliography

Andrews, G., Barrett, J. C., Lewis, J. 2000 Interpretation not record: the practice of archaeology *Antiquity* **74:** 525-30.

Beck, A. 2000 Intellectual excavation and dynamic information management systems, pp. 73-88 in G. Lock and K. Brown (eds) *On the Theory and Practice of Archaeological Computing* Oxford University Committee for Archaeology, Monograph No. 51.

Bender, B., Hamilton, S. and Tilley C et al 1997 Leskernick: stone worlds; alternative landscapes; nested landscapes, *Proceedings of the Prehistoric Society* **63**: 147-78.

Brown, Adrian 2000 *Centre for Archaeology Digital Archiving Project; internal consultation summary report*.

Carver, M. 1991 Digging for data: archaeological approaches to data definition, acquisition and analysis, pp 45-120 in R. Francovich and D. Manacorda (eds) *Lo scavo archeologico: dalla diagnosi all' edizione* (Florence, All'Insegna del Gioglio).

Chadwick, A. 1997 'Archaeology at the Edge of Chaos: Further towards reflexive excavation methodologies' *Assemblage* **3.** http://www.shef.ac.uk/assem/3/3chad.htm

Denning, K. 1997 I Link? or Ink? Therefore I Am: More Ruminations about Electronic Publishing and Archaeology *Assemblage* **3.** http://www.shef.ac.uk/assem/3/3kdrumin.htm

English Heritage 1991 *Management of Archaeological Projects* (MAP2), HBMC.

English Heritage 1995 *Geophysical survey in archaeological field evaluation* Centre for Archaeology Guidelines.

English Heritage 1995b. *Guidelines for the care of waterlogged archaeological leather* Centre for Archaeology Guidelines.

English Heritage 1996 *Guidelines for the conservation of [historic] textiles* Centre for Archaeology Guidelines.

English Heritage nd. *Waterlogged wood: Guidelines on the recording, sampling, conservation, and curation of waterlogged wood* Centre for Archaeology Guidelines.

English Heritage nd. *Dendrochronology: guidelines on producing and interpreting dendrochronological dates* Centre for Archaeology Guidelines.

English Heritage 2001 *Redundant Software survey* internal report.

English Heritage 2001b *Archaeometallurgy* Centre for Archaeology Guidelines.

English Heritage 2002 *Environmental archaeology: a guide to the theory and practice of methods, from sampling and recovery to post-excavation* Centre for Archaeology Guidelines.

English Heritage 2002b *Human bones from archaeological sites: guidelines for producing assessment documents and analytical reports* Centre for Archaeology Guidelines.

Hamilton, C. 2000 Faultlines, the construction of archaeological knowledge at Catalhöyuk, p in I. Hodder (ed.) *Towards Reflexive Method in Archaeology: The Example at Çatalhöyük* Cambridge, McDonald Institute for Archaeological Research and British Institute of Archaeology at Ankara.

Hodder, I. 1997 Always momentary, fluid and flexible: towards a self reflexive excavation methodology *Antiquity* **71**: 691-700.

Hodder, I. 1999 *The Archaeological Process: An Introduction*, Blackwell.

Huggett, J. 2000 Computers and archaeological Culture Change, pp. 5-22 in G. Lock and K. Brown (eds) *On the Theory and Practice of Archaeological Computing* Oxford University Committee for Archaeology, Monograph No. 51.

Lock, G. 2003 *Using Computers in Archaeology* Routledge.

Lucas, G. 2001 *Critical Approaches to Fieldwork* Routledge.

Moss Kanter, R. 2001 *Evolve!: Succeeding in the Digital Culture of Tomorrow* Harvard Business School Press.

OECD 2003 *ICT and Economic growth: evidence from OECD countries, industries and firms*. OECD.

Rains, M. 1995 Towards a computerised desktop: The integrated Archaeological database System, pp. 207-210 in J. Huggett & N. Ryan (eds*) Computer Applications and Quantitative Methods in Archaeology 1994,* BAR Int. Ser. **600**, Tempus Reparatum.

Richards. J., & Robinson. D. 2000 *Digital Archives from Excavation and Fieldwork: A Guide to Good Practice* (2nd edition) AHDS.

Roskams, S. (ed) 2000 *Interpreting Stratigraphy: site evaluation, recording procedures and stratigraphic analysis: Papers presented to the Interpreting Stratigraphy Conferences 1993-1997* BAR Int. Ser. 910, Archaeopress.

Roskams, S. 2001 *Excavation* Cambridge University Press.

Ryan, N. 1995 The excavation archive as hyperdocument? Pp. 211-220 in J. Huggett & N. Ryan (eds*) Computer Applications and Quantitative Methods in Archaeology 1994,* BAR Int. Ser. **600**, Tempus Reparatum.

Ryan, N. 2001 Jnet: a successor to Gnet, paper presented to In W. Borner, (ed), *Archäologie und Computer, Workshop 6*. Forschungsgesellschaft Wiener Stadtarchäologie, November 2001. http://www.cs.kent.ac.uk/pubs/2001/1547/

Shirky, C. 2003 Social Software and the Politics of Groups, "Networks, Economics and Culture" Mailing list. http://www.shirky.com/writings/group_politics.html

Stewart, J. 1997 Has 25 Years of Computing Provided Greater Physical and Intellectual Access to Archaeology?, In *Computer applications and quantitative methods in archaeology.*

List of Appendices

Appendix 1a

English Heritage Centre for Archaeology

Revelation Project

Report on the Review of Existing Systems

7 May 2003

Brian Attewell

Contents

Introduction

The fundamental aims of the Revelation Project are to greatly improve the capture, analysis and dissemination of Centre for Archaeology (CfA) research by creating a coherent digital information framework.

The CfA's current Information Systems (IS) are based on decades of research and development, practical experience, skills and procedures; in the field, office and laboratory. Whilst a 'clean start' in creating new systems would be ideal, pragmatically we cannot afford to abandon all of our IS heritage; also we may find ourselves discarding good systems and practices along with the bad.

Before we leap forward we need to see where we are standing, by ascertaining:

- what systems we currently have and use, for what purposes and by whom;
- how information for these is gathered, structured, processed and disseminated;
- if and how this information interfaces with other CfA, EH and external systems;
- the expected lifetime of the requirement for the systems, and of the systems themselves.

We assessed user's needs for their systems, whether they were 'fit for purpose', and finally whether they came within the scope of the Revelation Project.

The review was undertaken between November 2002 – February 2003. It involved the entire sixty-plus staff working for the CfA: in the field, in Savile Row, and at Fort Cumberland.

It consisted of:

- The CfA elements of two existing surveys carried out by English Heritage: the Year 2000 (Y2K) exercise, checking on EH software compliance, and the Redundant Software Survey, intended to rationalise and modernise the EH software 'estate'.

- Adrian Brown's (CfA Information Manager 2000-2002) Digital Archiving Project (DAP) internal consultation report, produced in 2000.

- A full scan of CfA networked storage looking for files that were likely to represent existing systems.

- The results from a questionnaire, presented to all CfA staff by members of the Revelation Project Team.

- The production of formal Data Flow Diagrams based on the questionnaire responses and subsequent systems analysis.

Data Gathering

Y2K Survey

In 1999 English Heritage Information Systems Branch (ISB) started preparing for the year 2000 IT systems switch. This was to identify where computer hardware and software designed to use the last two digits for designating the year, eg 98 for 1998, would possibly fail or misbehave when the year 'clocked over' to 00 (2000). Within the survey the CfA IS Team identified a number of systems used at Fort Cumberland so that contingency plans could be made.

Redundant Software Survey

This was a survey undertaken by the CfA IS Team in September 2001 as part of the English Heritage-wide 'Removal of Legacy Software' project. The driver for this was the impending replacement Facilities Management contract; EH sensibly wanting to lose as much old and obsolete software, and rationalise around its transfer to Microsoft Office 97 from Lotus and WordPerfect.

A scan of network files was undertaken with the aim of highlighting Access 2 databases and Lotus 1-2-3 spreadsheets. Manuela Lopez investigated the Access 2 databases found, talked to owners as to whether they were still required and converted them to Access 97 if requested. Lotus spreadsheets could be relatively easily converted to Excel and this was left to users.

All of the databases highlighted in this survey that still exist were picked up by the 2002 network server scan below.

Digital Archiving Project (DAP)

The Digital Archiving Project's main aim was *"to develop a comprehensive, integrated, and achievable digital archiving strategy for the Centre for Archaeology (CfA), within the context of the evolving corporate policy."*

Adrian Brown's first project stage was to *"…undertake an internal consultation exercise with a representative sample of CfA staff, chosen to include all constituencies of data creators and users"*. For this he asked for much information relevant to our current review.

Adrian Brown's interviews with various system owners and subsequent analysis, although now out of date in some areas, led to some very useful, but gloomy conclusions, primarily that *"…data is managed in a very unsatisfactory manner…"* For this reason it is reproduced in its entirety in Appendix 1 and its results must be taken into consideration in stage 2 of the Revelation Project.

Scan of CfA network server

We knew that there would be a number of systems on the server that were missed by the other surveys; either genuinely 'orphaned' systems where the owner no longer worked for the CfA and our procedures failed to document them, or forgotten systems ready to be reintroduced to the owner! A complete scan of all the files stored in the data areas of the CfA NetWare fileserver was undertaken and imported into an MS Access database.

Existing Systems Questionnaire

This was based around a pre-designed form of nineteen questions (see Appendix 2). Ten members of the Revelation Project Team were assigned to interview CfA staff, sometimes on a one-to-one basis, sometimes to functional teams. After introducing the project and the role of the survey a questionnaire was completed for each system identified by the interviewee(s). Where arranging an interview proved impossible, staff completed their own questionnaires. The interviewers were also briefed by Manuela Lopez on how to best elicit information suitable for the production of the data flow diagrams.

Lessons learnt from undertaking the questionnaire

The work involved in extracting useful information from the questionnaires was greatly underestimated. Much of this can be blamed on their design and implementation, specifically:

Better briefing of interviewers - There was a lack of consensus amongst the interviewers as to the meaning of some of the questions and the nature of the responses we were looking for. The discrepancies resulting devalued some answers and needed extra questioning of users, delaying the analysis. A more formal interviewer briefing with a chance to discuss likely problems would have been of benefit.

Better phrasing of questions - There were ambiguities in some of the questions *like "What software and hardware does it use?"* which led to a spectrum of answers from the short and simple 'PC' to lengthy descriptions of IT. The resulting data needed to be reinterpreted to be of any use. Perhaps a 'tick box' approach should have been used.

Quality control of completed questionnaires - The returned forms were extremely variable in terms of completeness, spelling and grammar, and comprehension. More emphasis should have been given to interviewers checking their questionnaires before analysis.

Data Flow Diagrams

Data flow diagrams were produced based on the answers given in the Existing Systems Questionnaire and also by follow up meetings. Diagrams were produced for each of the nine CfA teams plus one for the CfA Raunds Project Team:

Administration
Archaeological Conservation
Archaeologists
Archaeometry
Archives (previously Information Management and Collections)
Environmental Science
Graphics Studio
Information Systems
Technology

The diagrams in this early form showed both data and work flows. They highlighted the complex web of data dependencies, both internally to the CfA and externally to other EH and external users. They also illustrate some ‘interesting’ perceptions, and sometimes realities, of who and what data interact with whom; with perhaps less adherence to standards and procedures than there should be.

It is intended to produce higher level, more generic diagrams to inform subsequent stages of the Revelation project.

Analysis

Definition of a system

The definition of a 'system' obviously had a substantial part to play in what was identified, described and captured by the survey and used in subsequent analyses. The definition chosen was amalgamated from a number of 1980/1990s texts on Information Systems theory:

"… a process that assembles, stores, manipulates and delivers information."

It had to be Information Technology (IT) independent, as we had a considerable number of manual processes that were readily identifiable as suitable, if not critical, for the digital framework Revelation is aspiring to. Business Processes and Work Flow were not separated out but included where necessary as they would have a large part to play in the design of future systems.

Scope

Systems whose function was unrelated to the collecting, processing and dissemination of archaeological information, eg equipment inventories, personnel data, some Management Information Systems (MIS) etc, were included in this analysis, but probably will not form part of the Revelation Project.

Systems not owned by the CfA, but used by us, such as HSIS, SMR, and FIS were not included in the review, but will obviously be included in subsequent data flows and requirements if necessary.

Analysis by survey

Y2K Survey and Redundant Software Survey

These surveys were based on application software rather than the systems themselves. All the significant systems 'picked up' in them have also been recorded in the Existing Systems Questionnaire and are analysed below.

Digital Archiving Project

The project's main purpose was to look at the storage and management (or lack of) of those data destined to be archived by the CfA. For this reason it was not particularly concerned with the creation of data, its use, and the data flows involved. The DAP Consultation Summary Report concluded that:

"the CfA creates an enormous variety of data which has potential long-term value, and in very significant quantities. It is also clear that much of this data is managed in a very unsatisfactory manner… The exercise has provided important

information on key issues such as documentation, organisation, version control, storage, access, and reuse, and identified a number of other important factors which are significant for the management of our resources"

Adrian Brown recorded around forty 'digital resources' early in 2000, thirty-four of which could be mapped across to systems identified by the Questionnaire, and analysed below; only a couple were not traceable.

Scan of CfA network server

For this snapshot, taken in December 2002, it was decided that potential systems would be identified as utilising Microsoft Access (.MDB), Microsoft Excel (.XLS) or dBase (.DBF) files. There are obviously other systems using other types of files, but these were either too difficult to disentangle from the myriads of non-system files (eg Microsoft Word .DOC), or too obscure as they used non-standard file naming conventions. The option of quizzing staff over files located proved to be too difficult as many parts of the network directory structure were inherited from generations of predecessors.

From the eight thousand identified files, duplicates, samples and tutorials where obvious were removed before analysis. This left 491 MDB, 1573 DBF and 4900 XLS; the task of mapping these files to known systems has not yet been completed and may be abandoned as too resource intensive.

Existing Systems Questionnaire

There were 146 questionnaires completed and these were loaded into a simple MS Access database (Appendix 3). The original responses were corrected for obvious factual errors and had their terminology somewhat standardised; where necessary further information was sought from the interviewee or system owner.

The responses to each of the questions were then looked at together and an analytical version of the responses produced for each of the 109 different perceived systems.

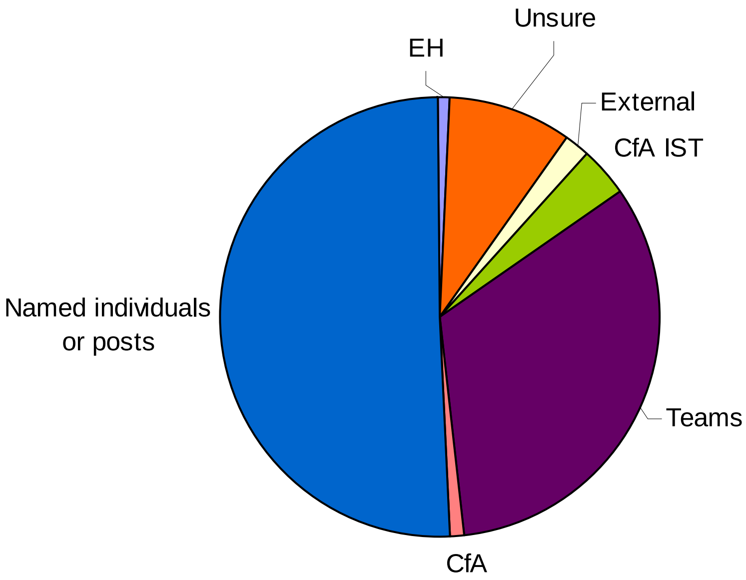
Q1) System name?

The 'naming' of systems did not seem to have much importance for interviewees. There were a number of answers that were given for the same system, with different staff supplying different names. Where the system names chosen were used as the file system name, these were often not descriptive or intuitive in any way. These are not really critical issues but do fall within the general area of file, document and information management; sensible and to an extent standardised names would make life easier.

Q2) System owner?

There was significant uncertainty over who owned or was responsible for systems. In certain cases this could be seen as the major reason for why systems failed. To the interviewees the difference between user and owner was often unclear for multi-user systems. The owner was often quoted as being the 'team' having the most users, yet there was no formal acceptance of this role.

- Unsure: 10 systems (from 2 to 4 owners)
- EH: 1
- CfA: 1
- External: 2
- CfA IS Team: 4
- Teams: 36
- Named individuals or posts: 56



A formal and responsible owner should be found for each system, and this knowledge made available to all. The role and responsibilities of a system owner should also be made clear. Current CfA procedures for handling of staff leavers and their EH files, documents and systems need to be more strictly enforced.

Q3) What is the system designed to do?

To try and understand the nature of these systems and how they were used they were categorised as follows:

Collection - an identifiable, classified, physical set of items or records (in specific locations). For example the pottery from an excavation stored at FC; the photographs taken and stored by the Technology Team; a number of specialised reference libraries.

Index – an organised list referring to the existence of items, data or information held elsewhere. For example the Raunds drawing catalogues and database.

Inventory – an organised list referring to the location of items, data or information held elsewhere. Normally for management and tracking purposes like the Taxon/Taxref plant microfossil reference collection system.

Transaction – an organised list recording the processes undertaken on items, data or information. For example Labfile.

Information Store – an organised collection of records, manual or digital, designed for interrogation. For example the Environmental Team's Species Tables.

Analytical – data and information held to be used to generate further information. The Technology Team's site spreadsheets are a good example.

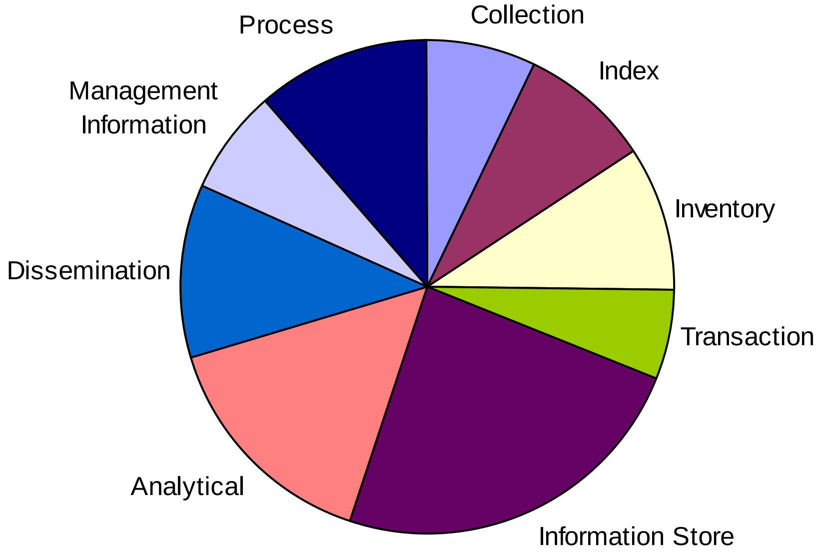
Dissemination – data and information capable (and designed) to be output for access by others. The World Heritage Site GIS systems are designed to output information to be widely disseminated.

Management Information System – data and information used to inform management decisions. CASPAR is used here to monitor the progress of CfA project completion.

Process – the input, manipulation and output of data or information, but not including storage. Includes systems like the Graphics Team reconstructions and the MakeDXF utility.

These categories were not exclusive; most systems fell into a number of types:

- Collection: 16
- Index: 19
- Inventory: 21
- Transaction: 13
- Information Store: 53
- Analytical: 34
- Dissemination: 25
- Management Information: 16
- Process: 25

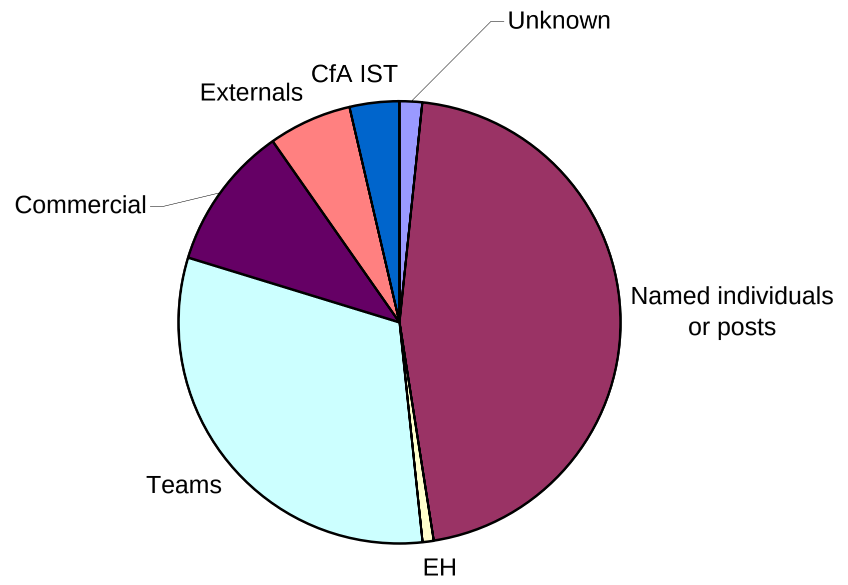


In summary, and unsurprisingly, the systems are designed to collect, store, manage and manipulate, analyse and then disseminate the results of archaeological fieldwork and scientific analysis. The systems are also designed for the CfA to undertake research and development into theory and practice so that we may authoritatively provide advice, set standards, and develop new techniques. There is a fairly equal split in our systems between the collection and management of data, and it's analysis and dissemination.

Q4) By whom, when and in what context was it designed?

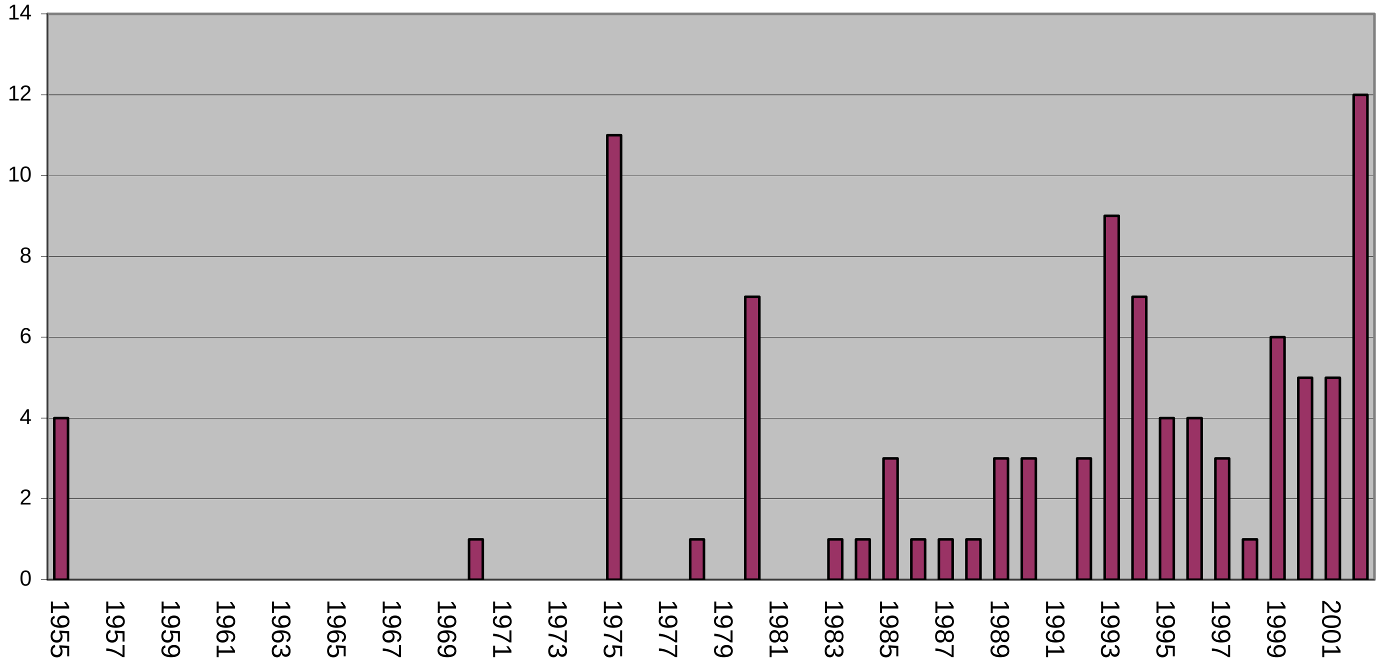
Whom:

- Unknown: 2
- Named individuals or posts: 52
- EH: 1
- Teams: 36
- Commercial: 12
- Externals: 7
- CfA IS Team: 4



The large proportion of systems designed by individuals or teams can be looked at positively: staff are largely self-reliant and capable. Where the systems are manual or simple digital ones, this is a practice to be encouraged. However, the responsibilities of the system designer(s) for continued management, data standards, interoperability, documentation and support need to be formally recognised. Where the systems are complex, or are used outside of the original scope, then more formal design, development and support practices must be considered.

When: (Number of Systems Designed by Year**)**



The earliest systems were manual ones developed by the Ancient Monuments Laboratory; the Central Excavation Unit was set up in 1975 and a number of manual and digital systems had to be created. The 'burst' of systems development starting in the late eighties reflects staff taking advantage of the wider availability of desktop applications.

Context:

- Specific projects or tasks: 18
- Ongoing projects or tasks: 80
- As a replacement for a duff system: 3

Most systems seemed to have been created to deal with the regular and relentless tide of data, both archaeological and management, that are the stuff of fieldwork, scientific analysis, and operations. The systems were developed to organise and automate a large number of business processes and work flows. Increasingly for much of our work digital data and information is the primary and only record.

The systems set up for specific projects or tasks are of some concern perhaps where, with some foresight, economies of creating and using shared systems would have been better.

Q5) What kind of information does it hold and in what format?

There are a myriad record types used in the CfA, reflecting the variety of physical things we deal with, analyses we undertake, and reports we produce. The vast majority of these records are classifications and descriptions of archaeological sites or the materials removed from them.

A simple breakdown shows that most of our systems are digital, as should be expected in modern archaeology:

- Digital: 72
- Manual: 27
- Both:11

There are a significant number of manual systems that should be digital, and this fact was recognised by the interviewees. The reasons given for not digitising these resources were often quoted as lack of time or expertise, this highlights the requirement for the CfA to have appropriate access to systems developers or, even better, a CfA based resource.

Not surprisingly most of the digital systems utilise databases or spreadsheets:

- Database: 30
- Spreadsheet 8
- GIS: 4

A large number consist of or utilise:
Graphics (vector and raster), Images (of which Photos: 16), Plans, Drawings: 23

The use of digital graphics is fundamental to our archaeological work and is increasing rapidly. It may be that, in a few more years, the majority of our systems will be graphically rather than textually based; this is a trend that needs to be recognised for IT infrastructure management purposes.

In general the lack of availability of any formal meta-data for these systems is a real concern. There is a high risk that data and information, often gathered with considerable effort and at considerable expense, can become lost in a poorly described system, especially where the original designer, owner, or user has left the organisation.

Q6) How much information does it hold?

- In terms of individual records, from 100s up to 100,000s
- In terms of storage used for each system, from Kilobytes up to 20Gigabytes (although we are including a number of systems holding graphical records)

For a digital system, the volume of information it holds should only be an issue if the system is required to be accessed remotely over a low bandwidth (ie slow) connection. The performance of a system should not be affected by the size of information stored. The intended EH Technical Refresh Project should remove any hardware infrastructure constraints on size. Needless to say this does not apply to manual systems!

Q7) How many people use it?

There was a lot of uncertainty in the responses to this: partly over the definition of ‘use’ and also because no time-frame for the use was specified. The ‘intended’ definitions of the terms were:

Single user – only one person or post uses the system (a replacement post-holder would count as single user).
Multiple copies – more than one person uses the system, but each has their own copy.
Sequential multi-user – more than one user has access to the system, but only one at a time.
Concurrent multi-user – true shareable systems with simultaneous users.

- Single user: 31
- Multiple copies: 14 (from 3 users to all the CfA)
- Sequential multi-user: 34 (2 to12 users)
- Concurrent multi-user: 28 (2 users to ‘available on the Internet’)

Should single user be multiple? A number of comments were made that other people would be interested in using the system if they knew it existed or had access. There were also issues over the actual ownership of systems, perceived to be personal by some staff, and the property of EH by others.

Multiple copies can be very dangerous and interviewees normally knew this. There is a requirement for a formal file, document and system management procedure.

Should sequential use be concurrent where possible? (a large percentage of the sequential multi-user systems are manual, eg recording proformas and card indices, or based on a single piece of equipment). Some of the database systems could be modified to allow multi-users.

Q8) What strengths and weaknesses do the users perceive in the system?

Well over half the systems were seen as being fit for purpose; the other identified strengths were:

- Quick to use and get results: 8
- Easy to use: 25
- Quality of product: 5
- Flexible to use: 7
- Standard (in terms of being widely used): 5

There were some interesting differences in how users defined ‘ease of use’; it being menu driven suited some and command line driven others! It is worth noting that, after “fit for purpose”, ease of use was the most frequent response. Quality of product was only stated as a criteria by the Graphics Team, reflecting its obvious visibility in graphics, rather than its absence in other systems.

Weaknesses mentioned included:

- Incomplete data sets: 3
- Keeping system data up to date: 10
- Isolated from other systems: 20
- Not standardised: 8
- Lack of access for other users: 3
- Lack of performance and or functionality: 4
- Lack of flexibility: 5
- No security copy: 4
- Not easy to use and or needed to be experienced to use: 9

The most frequent problem by far was the lack of integration between systems, resulting in data duplication, reduced functionality and inefficient working. It is not coincidental that remedying this is one of the fundamental objectives of the Revelation project. It will also need to address the usability shortfalls of many existing systems, hopefully by the correct mixture of design and training. Keeping data up to date and standardised are both problems that Revelation cannot magically put right, they are issues of management and professionalism.

Q9) How many versions of the system exist and how many are currently in use?

There was some confusion between ‘versions’, ie modified replacements of an earlier system, and ‘copies’, duplications of a current system; neither should really be in use at the same time.

- 1 version: 82
- 2 versions: 6
- 3 versions: 2
- Multiple: 4

It was good to see that most systems only existed in a single version. Where more than one version existed, this tended to be with the software supporting the system (eg AutoCAD r14 or AutoCAD 2002) rather than the system itself. Tech Refresh should remove a lot of these anomalies.

Q10) What quality controls are in place, how reliable is the data in the system?

The majority of answers stated that the user was responsible for the quality control of the system. The manual methods stated included double-entry and peer review; for digital systems: glossary control, referential integrity, and type validation. There were a significant number of responses stating that system data were unreliable, and this was quoted as a major reason for not using the systems.

- Depends on user: 61
- Built in checks: 18
- Unknown: 3
- None: 10

Wherever possible, automatic checks and validation should be built into systems, <u>and switched on</u>; if not standards should be available and adhered to. Systems data should not be allowed to ‘degrade’ to unacceptable levels.

Q11) What software and hardware does it use?

Manual:

We still use a large number of card indices, hanging files, filing cabinets, and *proformas*. Wherever possible these systems should be digitised for security, ease of use and interoperability. For distinct datasets and information systems this is a major goal of the Revelation Project, for others future CfA and English Heritage record and document management strategies will be of great interest.

Digital:

The vast majority of our systems run under Microsoft Windows and use Microsoft Office Professional applications; however due to the specialist nature of some of our IT equipment and applications, and the long lifetimes of many systems; we are using a few other operating systems:

- 'WWW browser': 3
- MS Windows: 106
- Unix: 3
- DOS: 5
- Other: 2

Unix is a robust operating system that, as long as we undertake the necessary maintenance and upgrades, will pose little threat. There are advantages to porting some Unix applications across to Windows, and we have started doing this, but the case for a larger conversion has to be looked at in detail. The DOS based applications were either for legacy systems that we have yet to find a Windows alternative for, or for where the required raw processing speed needed a less complex environment than Windows.

Software:

The list of software applications and utilities required for CfA systems is extensive.

Name	Description
Adobe Acrobat	Electronic document exchange tool
Adobe Illustrator	Presentation Graphics Package
Adobe PageMaker	DTP package
Adobe Photoshop	Graphics editor
Autodesk AutoCAD	Computer aided design package
Autodesk AutoCAD LT	Computer aided design package
Autodesk AutoCAD MAP	AutoCAD light GIS extension
Autodesk AutoSketch	Computer aided design package
Autodesk Land Desktop	AutoCAD terrain extension
Autodesk Raster Design	AutoCAD raster editor extension
Autodesk View	Graphics file viewer
Autodesk Volo View	AutoCAD file viewer
CA SuperProject	Project planning
Corel Paradox	Database package
CorelDraw	Presentation graphics package
Corel PhotoPaint	Graphics editor
dBase	Database package
Delilah	Database package
ESRI ARC/INFO	GIS
ESRI ARCPAD	Lightweight GIS for field use
ESRI ArcView	GIS
ESRI ArcView 3D ANALYST	ArcView surface modeller extension
ESRI ArcView IMAGE ANALYSIS	ArcView image processing extension
ESRI ArcView SPATIAL ANALYSIS	ArcView spatial analysis extension
Geosoft	Geophysics Image processing
ISI ResearchSoft EndNote	Bibliography organiser
KeyTERRA-FIRMA	Digital Terrain Modelling package
Kodak Photo CD Access	Photo CD display software
LEICA SURVEY OFFICE	Data logging package for survey equipment
Mathsoft Axum 6	Graphics package
MetaCreations Painter	Graphics package
Microsoft Access 2 & 97	Database package
Microsoft PowerPoint	Presentation Package

Microsoft Project 98	Project Planning
Microsoft Visual Basic Professional	Graphical Programming Language
Oracle Client Software	Database client software
Oracle Developer/2000	Database development tools
OxCal	Radiocarbon dating
PocketGIS	GIS system for on-site use
PowerBASIC	Programming application
Sensors & Software EKKO_PRO	Ground Penetrating Radar software
SKIPRO	GPS logging software
SPSS	Statistics package
TheoLt	Data logging and surveying tool
Tilia	Pollen data analysis
Trimble Geomatics Office	Data logging package for survey equipment

This is not a complete or definitive list but gives a good idea of the specialist nature of much of our software.

Hardware:

This is harder to categorise as systems can run on a number of platforms or combinations of hardware. The figures quoted are for a system's main platform ie nearly all software could be installed on a standalone, or on a mobile laptop, but this is not normal practice. Also around a third of our 'networked PCs' are actually docked laptops.

- Application server: 6
- File server: 78
- Networked PC: 78
- Standalone PC: 4
- Mobile laptop: 4
- Proprietary: 4

Issues raised by interviewees were many:

Wide Area Network: as we have a very low bandwidth connection to the rest of EH this played havoc with remote applications.

LAN performance issues: the response times of the network for file intensive tasks (CAD, GIS, DTP etc) were very poor. Many 'power' users found it necessary to work on files locally, even with all the problems of backups not being carried out and file reconciliation and sharing difficulties.

Fileserver: the performance of the main NetWare fileserver was seen to be a problem as with the LAN. The disk capacity of the server was a real problem as it has always been at around 90% full, which with our ever increasing use of digital graphics has required a lot of on-line off-line data juggling.
Application servers: Two of the three Unix servers are seriously beyond their expected life and are a threat to systems running on them. There is a proposal to move the functionality of these across to the newer Unix server, if it is upgraded.

The 'PC' hardware and software available as part of the EH IT infrastructure was, in the majority of cases, not chosen for system optimisation. 'Standard' PCs or Laptops are grossly under-powered; the average installed RAM being 64MB and

CPU speed 200-400MHz. The lack of CD-ROM drives also made data import a much more troublesome process.

The systems also required a number of specialised hardware and software platforms, either for field data-logging purposes or where the IT was required to drive scientific equipment.

Finally a number of standard and specialised IT peripherals were required for inputting (digitising tablets, digital cameras, scanners of all sizes) and outputting (printers and plotters) to and from systems.

Q12) What is the expected lifetime of the requirement?

The majority of systems were seen as being ongoing, either for operational purposes as in transaction, inventory, and analytical systems, or for systems supporting the CfA's role to carry out research & development and provide advice to EH.

- Ongoing: 84
- Life of project: 21
- Until replaced: 3
- Finished: 1

The relatively large number of 'life of project' systems may be a concern if the systems designated were not seen as being of use in any other project. It would be hoped that wherever possible systems are designed to be used widely, both for economies of development (re-inventing the wheel) and to provide standardised tools.

Q13) What is the expected lifetime of the system?

The system, hardware or software, may not last as long as the need for it, especially where its use or support is withdrawn by the manufacturer, or by corporate policy.

- Ongoing: 57
- Life of project: 10
- Until replacement: 16
- Finished: 1

Obviously, where the lifetime of the system is less than the lifetime of the requirement, then there is the issue of upgrades or replacements. The EH Technical Refresh project will hopefully assist in maintaining the good systems by providing a stable, supported platform that will be upgraded as necessary.

Q14) What documentation is available?

This varied greatly between the systems, but more worryingly, in the fourteen systems with multiple interviewees, four had significantly different ideas of what was available:

- None: 46
- On-line: 9
- Crib/desk instructions: 21
- Manufacturers: 19
- Standards: 4
- Users: 17
- Tutorials: 3
- Training: 3

The availability of documentation should be made known to all users (and potential users) as a matter of course. The large number of systems without proper documentation, even single-user ones, is a genuine threat to Business Continuity.

Q15) Who supports the system?

It is taken for granted that Sema supports all underlying hardware, software and those applications in the EH Contract. The nature of support that could be expected by users was not obvious, especially with the more specialist applications (CAD, GIS, and statistics). There was also considerable confusion as to who supported many of the in-house produced systems.

- User: 57
- Team: 32
- CfA IST: 10
- External: 4
- Manufacturer: 8

There was an impression that ongoing support was not considered when systems were developed. Systems that are not supported, or where the support provider is not documented, are a significant risk. The level of support available as part of the Sema contract should be clearly communicated to all staff.

Q16) What data, digital & otherwise, is input from other systems?

This information is best presented in the Data Flow Diagrams. The vast majority of systems do require data to be input from other systems, the main exceptions being where the system actually creates the primary record itself (eg CfA Recording System, Geophysics Field Data Capture).

These data are best input by direct links between systems eg Access databases or spreadsheets; next best is via an intermediate format eg CSV or DXF; more inefficient is cut and paste between systems; and finally there is re-entry.

Q17) What data, digital & otherwise, is output to other systems?

This information is again best presented in the Data Flow Diagrams. CfA systems providing project data to in-house and external specialists is a fundamental workflow process. Our systems are also a valuable source of data for management, research and publication in EH and externally.

Currently these data are either provided as copies of systems, or parts of systems; exported via intermediate formats; or produced as hard copy. It is to be hoped that such necessary data can be directly 'output' in future eg by controlled Internet access.

Q18) Are there any data or functions that you would like in your system that you do not have, what do you do about these 'missing' at the moment?

About 30 systems needed no improvements, as judged by the interviewees; for the rest requirements could be briefly summarised as follows:

Digital systems needed to replace manual ones.
Digitisation of old records required for inclusion in datasets.
Automation of systems -
- basic statistics and charting included;
- data input mechanised;
- labels generated from systems.

Improved hardware -
- local A3 scanner;
- more powerful laptops;
- networking of all FC based hardware.

Improved systems design -
- linking to other systems;
- converting systems to multi-user;
- improving database design;
- in-built checking.

All of these requirements seemed reasonable with sound business justifications. The fact that many interviewees seemed frustrated that nothing was to be done about them is a serious concern likely (not counting the Technical Refresh project).

Mitigating missing functions –
- re-entering and copying of data;
- relying on paper records;
- frequently editing or configuring systems;
- manual calculations performed.

It is encouraging that interviewees found methods for getting around system limitations so that they could do their work; however we should never had got into this predicament in the first place. Many of the work-arounds currently used are resource intensive or do not produce the best results possible.

The questionnaire system wish lists will obviously be used in the next stage of the Revelation Project to ensure we design the systems that are really wanted by the users.

Summary

The analyses undertaken on results from the various surveys are by no means exhaustive; there are many patterns of use (and abuse) that can still be extracted. This brief summary is solely based on the lessons learnt so far. It should support the User Needs Survey to produce a Revelation system design fit for purpose.

Strengths of Existing Systems

- Many of our systems were efficient.
- Many were easy to use.
- Some were simple in design and operation.
- The majority did the job users asked of them.
- A few were seen to be exemplary.

Weaknesses of Existing Systems

- The System 'Owner' was non-existent or uncertain.
- Data integrity was poor, sometimes awful!
- Duplication of data was common.
- Old but needed data was in formats now unreadable within the CfA.
- There were many problems with getting data and information into systems.
- There were problems with getting data and information out of systems.
- Hardware and or software was often inadequate.
- There were not enough software licenses for users.
- Systems were difficult to use.
- Systems were not supported.
- There was a frequent lack of user training.
- There was a lack of documentation or user guides.
- A number of systems were not backed up.

Opportunities for Future Systems

- Clarity of ownership and responsibilities will make everyone involved happy and lead to a better utilised system.
- Integration of requirements, information and outputs will give users what they want quickly and deliver it cost-effectively.

- Integrated multi-media (a picture is worth a thousand words).
- Spatially aware systems, everything in its place as it should be in archaeology.
- Systems designed to be easy to use, will be used.
- Have systems accessible to others, not only specialists but for all stake-holders; not just as a final product, but also during the archaeological process.

Identified threats to a successful Revelation

- Systems not designed and operated within a proper IS/IT strategy and framework will lack the necessary integration and continuity to be efficient.
- If users do not give their support, and they do not understand, like or use it, then the system will not work.
- A lack of IS/IT resources can prevent a project starting or finishing, or result in major compromises to design, function and performance.
- Systems must be designed and developed properly, with all the necessary documentation, training and support structures in place.
- Unrealistic time scales and deadlines, for obvious reasons will handicap the outcome of a project, create stress amongst the project team, and if severe enough result in the project failing.
- Lack of senior management ‘buy-in’ can allow all of the above to occur. It can also disenfranchise staff and destroy their morale.

Conclusions

The Centre for Archaeology has and will always need a large number of systems to carry out its role in English Heritage. These systems are nearly all vital for efficiently handling the capture, analysis and dissemination of our work. There are inescapable pressures on English Heritage from government, public and profession to make our information stores and the results of our analyses available in a coherent digital fashion. To do this there is a considerable amount of work required to understand, consolidate, re-engineer and produce a new systems framework. This should build on the best of our existing systems, utilise proper information systems theory and practice, and most importantly have the input and ongoing support of our users.

Acknowledgements

Thanks must go to all of the staff of the CfA who were grilled over their systems, provided comments on processes and products, undertook interviews, and assisted in data entry. Thanks also to Manuela Lopez and Paul Cripps for producing the Data Flow Diagrams.

Appendices

Appendix 1: Digital Archiving Project
{currently s:\projects\pr3591\review of existing systems\appendix 1.doc}

Appendix 2: Existing Systems Questionnaire Form
{currently s:\projects\pr3591\review of existing systems\appendix 2.doc}

Appendix 3: Existing Systems Questionnaires, collated and analysed responses
{currently a MS Access 97 database in s:\projects\pr3591\review of existing systems\appendix 3.mdb}

Appendix 4: Data Flow Diagrams
{nine draft diagrams are in s:\projects\pr3591\DFDs\DFDgraphics_080403, currently in Adobe Illustrator .ai format, which can be easily read into Adobe Acrobat}

Appendix 1b

An Assessment of Handheld Drawing Systems

English Heritage internal assessment report
for the Revelation project
Version 1.2

Keith May

December 2002

1. Introduction

1.1. Purpose of this assessment report as part of the Revelation project

This report is one of a number of reports to be produced as part of a project called Revelation. The purpose of the Revelation project is "**to provide a coherent digital information system that will make the capture, analysis and dissemination of CfA research faster and more effective**".

The Revelation project is seen as a two-stage project. The first stage, which this report is a part of, is an initial review of existing CfA systems and user requirements and an assessment of the needs for integrated systems in the future. The assessment will produce a CfA report that sets out the needs for future systems and a project design for the next stage of the project with a plan for how the required systems can be delivered. The second stage of the Revelation project will be based upon the project design and the outcomes of the assessment stage.

The primary purpose of this report, as part of the first assessment stage of Revelation, is to assess the successes and limitations of digital drawing methods using a number of handheld hardware platforms and software interfaces. The report will make recommendations to aid in design of future stages of the project and will raise any resulting issues along with recommendations for future action. It will highlight where a decision making process is needed before proceeding beyond the assessment stage. The report should assist the project team in planning the next stage of the project.

1.2. Current circumstances and CfA fieldwork trials carried out to date

A number of pieces of fieldwork involving digital recording have been carried out by the CfA over recent years. It is not intended here to attempt to report on the outcomes of all these different and varied pieces of work, although a short overview of some of this work is presented in section two, along with some other background information deemed relevant to the issues addressed by this report.

This assessment is based primarily upon fieldwork trials that took place over a short period of 3 days at Whitby cliff in October of 2002. A number of other relatively recent fieldwork activities have also provided further information and experience to aid in this assessment. These include the author's experience from the DigIT project fieldwork carried out in August 2001 and 3 days Fieldwork on Dartmoor in July 2002,

It is not intended that this fieldwork is seen as an exhaustive fieldwork trial, or that it somehow supersedes the experience that has been gained by other practitioners operating in a variety of organisations across the archaeological community. It is the intention of the Revelation project to draw upon this valuable experience from a review of published literature and by a series of fact-finding visits to those in the sector with experience of developing such systems. In addition a seminar will be held at the British Academy to develop a wider perspective on how the use of these systems will enable better integration and interoperability of CfA information with colleagues both within EH and in the wider archaeological community.

1.3. Organisation of this report

This report has been divided into a number of sections. After this introduction to the assessment report and how it fits within the wider Revelation project, the report sets out some of the background and context for why the digital drawing component of the whole project has received particular attention at this stage (section 2). There is a short summary of the aims of the assessment (section 3) followed by a detailed account of the actual work undertaken in the field to assess some current examples of handheld drawing tools (section 4). The conclusions from the fieldwork are given (section 5) along with a listing of recommendations and suggestions for future work (section 6)

2. Some background to developments in digital drawing

2.1. CfA brief history of digital lines

The CfA have been developing the use of digital drawing methods over at least the last 10 years. Probably the earliest use was on the Windsor Castle Fire Project, around 1992, where attempts were made using an early version of Penmap to "digitise corrections directly into the CAD files" but this was largely seen as an off-site, post-excavation process as the Penmap data could not easily be incorporated into existing CfA fieldwork systems. AutoCAD was also applied to the Fire project, being deployed on site quite early on. Because the basic building recording was done using photogrammetry, a lot of detail was needed to fill in gaps in the data. This information was recorded by hand on overlays to printouts of the elevations, then digitised into AutoCAD via a digitising tablet.

Between 1992 and 1995 a number of seasons of fieldwork at Battle saw the development of attempts at 3D modelling while recording data on site using software such as Digital Ground Modelling v2.5. Unfortunately the nature of the equipment at that time meant that the data could not be seen until it was taken from site and processed. This inevitably led to difficulties in matching the series of "point-to-point" drawings in the record to what was on site and a number of features needed more than one attempt to record a reasonable representation. In addition at Battle Abbey here was digitising of hand drawn overlays to fill in the gaps in the photogrammetric record of the Courthouse. In this respect, it was similar to Windsor Fire. This approach also was applied to the Bolsover Castle Terrace Range in 1992

At Birdoswald in 1998, the work related to the visitor centre developments in the NW quarter of the Fort saw the introduction on a CfA project of EH Survey branch's TheoLT. This enabled the TST to be connected to the AutoCAD drawing platform while in the field. This was a major step in practice in that it allowed on site viewing of the information as it was entered. The recording strategy adopted at Birdoswald was to record only a skeletal outline of the context using a process involving TST/TheoLT/AutoCAD. The strategy was adopted in part because of the perceived difficulty in establishing a grid peg-based site grid on the sloping site setting. TheoLT was seen as providing a suitable method of georeferencing the plan record without setting up a physical site grid. The result was plotted out from AutoCAD and used as an underlay to guide the archaeologist in manually planning the context on drafting film. The hand-drawn plans were digitised into AutoCAD to complete the data set.

The TST/TheoLT combination used at Birdoswald was not intended by itself to produce final plans of contexts. The data was still entered in the form of a series of "point-to-point" vertices recorded by taking a series of readings using a survey staff.

There is however a substantial difference between the process of taking points with a survey staff by physically placing the staff around the edges of a feature and the archaeologist who excavated the feature being able to record points with a 'pen' on the drawing surface.

As demonstrated by the development of TheoLT many of the recent advances in this area have been developed for use in the recording of standing buildings. The development of reflectorless TSTs for use on site around 1998 meant that real-time true-3D drawing of elevations could be delivered in CAD with results that were previously only possible using photogrammetry.

Even though acknowledging these advances there are however some particular practical problems associated with the recording of buildings from the fixed position of the tripod which is required by even a reflector-less TST. Maintaining line-of-site can sometimes be particularly difficult on buildings with scaffolding when ironically the scaffolding and the recording often go hand in hand. A very similar problem occurs with equal regularity on archaeological excavations with deep or narrow trenches that do not allow easy viewing by an operator whose tripod is fixed on one or other side of the trench.

Experience of more recent attempts to disseminate and archive digitally produced drawings as part of the Archaeology Commissions digital archiving and dissemination programme, have suggested that no particularly consistent system has yet been devised by archaeological practitioners for producing a digital record from site which is also capable of dissemination without some fairly complex and often quite lengthy re-processing. This often involves either digitising hardcopy plans or “re-engineering” digitally derived survey data to produce drawings of suitable quality for either interim or final publication.

This also raises the point that every archaeological drawing will become digital if it is to be published, since publishers now print from digital copy - even if it is just a scan of the final inked drawing. What's more, most archaeologists now require a digital version of their plans for assessment and analysis - again the degree of functionality required varies, but the reason that so many approaches have been used is that there is such a broad need.

It seems that currently there are a range of commonly used approaches to producing digital drawings with differing degrees of detail or data attached to them depending largely upon the final purpose for which they are required. The most commonly used current approaches to creating archaeological digital drawings within CFA and elsewhere are summarized below.

2.2. Current practice in production of digital drawings

2.2.1. Digitising from hardcopy drawings using a digitizer

Site plans can be transcribed from the original pencil or pen drawn hardcopy versions of drawings using a digitising tablet to georeference the drawing and some form of tracing tool, usually called a puck. The digitising process involves accurately tracing over the lines that have been drawn on the paper or Permatrace. The resulting digitised drawing is in vector format and can therefore have attribute data attached which is most commonly done to enable its use in a GIS.

There are at least two techniques that can be used in this digitisation process. The more common is probably a 'connect-the-points' technique. An alternative is continuous digitisation which involves guiding the cross-hairs of the puck along the whole of each line being digitised. Depending on the software the connect-the-point digitising can be done with straight lines or splines (curved lines).

2.2.2. Scanning hardcopy drawings for on screen digitising

It has become relatively common practice to produce digital copy of drawings made on drafting film by scanning these, usually on a flat-bed scanner, or for larger drawings using a drum scanner.

The scanned copies of excavation drawings can be used as the basis for creating a vector drawing from the fieldwork hardcopies by a process of “on screen digitising”, using a desk-based computer and an on-screen pointing device. This process is carried out by tracing round the lines of the scanned image on screen with a mouse-pointer, or on a tablet using a pen.

On-screen digitising is usually considered a superior approach to digitising directly from the hardcopy. The georeferencing of the drawing is more reliable, and is not lost at the end of a work session, such as when the digitising tablet is switched off at the end of the day. On screen digitising is a more intuitive method for the person doing the digitising, particularly if they are inexperienced in the process and enables the person digtizing to be able to zoom in and out on the image in areas where the drawing is less definite.

Further advantages of on-screen, 'heads-up', digitising come from the fact that the person doing the digitising is looking at the place that the drawing is appearing. If the person is looking directly at the drawing on screen as it is being created there is more control over the accuracy of the result. In addition if the person digitising is looking at the drawing directly there is less strain on their neck and back, and they have a better working posture - which are significant health and safety concerns.

Some tools do exist for semi-automating this tracing process (e.g. ArcScan fro ESRI) but inevitably require some manual intervention to resolve problems when areas of a drawing need additional definition. The overall success of any such approach depends very much on the quality of the original drawings, the quality of the scanning and the quality, and therefore usually cost, of the vectorizing tool.

2.2.3. Scanning hardcopy drawings

Another use for scanned drawings may be to insert the images in a publication report or for similar dissemination purposes. In this case the scanned images remain in a raster format and often do not have any data associated with the drawing other than a file name identifier and if lucky some metadata to document when, where, why and by whom the scanned version of the drawing was created.

In C*f*A the primary application of scanning drawings has been for use as a security copy for archiving purposes.

2.2.4. Digital drawings created by survey instruments

As suggested by the summary of CfA work, most of the developments in digital drawing to date have been made using fixed position survey equipment. Such systems are usually based on some form of measurement recording device, such as a total station or EDM fixed on a tripod, and using a survey staff with a reflector to mark the position of the point on site that is being recorded. More recently pen computers, and reflectorless TSTs, which can directly locate the point being surveyed using a laser range-finder, have enabled this to become a single person task without the need for anyone to hold a survey staff. Both the staff and the reflectorless versions of this method require some data acquisition software to process the points being gathered and generate the lines of the digital drawing on a computer. This type of fixed position equipment was originally developed for larger-scale field survey and initially began to be used on site for surveying the location of excavation trench positions in relation to the OS national grid. As time has passed the equipment has become more sophisticated and is now used on a regular basis for recording all forms of spatial data, on both urban and rural excavations by archaeological organisations such as the Wessex Archaeology and MoLAS (Ziebart et al. 2002). This form of data acquisition includes digital planning of contexts on site, 3D location of finds and samples as well as the more standard survey of location information about trench and section positions. The generation of the digital plans on site enables the acquisition of non-spatial data attributes such as context number to be added to the digital drawings at source, leaving far less room for error.

With the advent of differential GPS it is now feasible to record digital plans in a similar way using a GPS staff to mark drawing points around a feature in the ground and create digital drawings derived solely from GPS acquired data. The quality of such drawings depends ultimately on the accuracy in terms of the number of points taken when plotting a feature using the survey staff and again, like most technical processes, the cost of the GPS equipment used. In addition differential GPS currently also requires a considerable amount of post-processing work to achieve fully accurate positional accuracy of the recorded data in relation to the real world. This limits the use of even of differential GPS for real-time on site recording as it would not be possible for excavators to see the drawings of the archaeology as they are recording it.

Some experimental work has been carried out using handheld GPS devices to fix trench positions but this has only succeeded in locating trenches to an accuracy of the nearest 5m. As yet the level of accuracy for the handheld machines is insufficient to enable digital planning of individual contexts, and differential GPS.

2.2.5. Digital drawing by hand on the computer screen

The ability to create digital drawings directly on a computer is becoming more common as software for digital drawing devices becomes more sophisticated. Digital drawing is carried out using a stylus or pen and a handheld computer and the lines of the drawing are generated directly on to the computer screen at the point where the stylus touches the screen. This method of digital drawing can be used to record drawings in both a raster format such as a bitmapped image file or in vector format where attribute data can be attached as the drawing is created.

It is this latter method of creating digital drawings which was particularly under assessment during the Whitby field trials and is considered in more detail in section 4 of this report.

2.3. Thinking the unthinkable - Can one drawing serve all? What are we drawing it for? Objective record, interpretive record, analytical tool, publication drawing and archive?

This paper is primarily focused on the question of digital drawing but that is only part of the Revelation project. Revelation itself is a CfA project sitting within the wider activities of English Heritage Research & Standards Group, which in turn conducts archaeological work within the much wider framework of archaeological research carried out by the whole archaeological community. In recent years there have been major advances in the way that much of this research is carried out using computer tools at all stages in the archaeological process from archaeological investigation and recording, through the assessment and analysis stages, to the dissemination of results and the archive.

One of the aims of the Revelation project is to review the actual purposes for which we use different processes and conventional recording methods in the archaeology carried out by the CfA, in order to review whether those processes and methods are still the best way of achieving the required results. Primarily this is a review being carried out within the CfA, but in order for it to have the correct perspective it must be in accord with the modernizing requirements of Research & Standards Group and take into consideration the broader needs of the wider archaeological community.

One of the principle aims in archaeological research must be to produce the best possible record of the site to enable the appropriate degree of analysis to be carried out in order to produce a report on the results of the investigation which is available for dissemination as soon as possible.

It has long been a problem recognised across the whole discipline that the time taken from leaving site to achieving the publication of results is disproportionately long. One area where it should be possible to improve the speed, accuracy and usability of the archaeological record is by reviewing the processes involved in creating drawings on site and how these drawings get used for different purposes during the course of the various post-excavation processes that result in an archaeological report.

Some of these fundamental questions underlie the approaches we might adopt for making any drawn record either on or off an archaeological site. The requirement of creating the drawn record has become such an unquestioned pre-requisite of any archaeological recording process that perhaps established processes have been carried through almost unquestioningly to the current era of digital recording. There is no clear recognition that the same drawn record can actually be “re-used” with a series of transformations at a number of different stages in the whole process of going from site record to publication.

Archaeologists use drawings to record and present our interpretation of what was found during an archaeological investigation. This assessment primarily focuses on the on site excavation processes that generate drawings, I will therefore leave aside the need for specialist drawings of finds or interpretive site reconstruction drawings, but the methods and processes for integrating these with the rest of the record should be examined as part of the review of existing systems. The main elements of the excavation processes where archaeologists make site-based drawings currently are:

- We draw plans.
- We draw sections.
- We use drawings to understand and (re)interpret the site
- We make publication drawings of plans and sections for dissemination
- We archive the drawn records

Each of these different types of drawing can be created in a digital form, and it is common practice that drawings from each of these different stages are required in digital format at some point in the archaeological process. The fundamental question is whether the information contained within the original drawings made on site can be created in a way that, with an appropriate degree of intervention, adaptation, and change control, it can fulfil more than one, and possibly all, of the purposes required during the whole archaeological process?

There are immediate questions about the differences between the use of plans and section drawings that suggest that these two types of drawing serve different purposes. Plan drawings are used to record the shape and extent of deposits prior to excavation, or the shape and extent of cut features after excavation. In either case the information recorded is usually about a single context (although often several contexts may be included on the same piece of physical drawing film – often for cost saving purposes, but also to simplify the quantity of post-excavation drawings to be handled). Section drawings are used to record and represent a visualisation of the relationships between a number of different contexts. Because of this difference it is likely that archaeologists will continue to record both plans and sections as separate types of drawing.

The most obvious way of approaching this would be to create drawings in digital format at the point of first recording the information on site. The small-scale Whitby fieldwork trial was carried out to investigate a number of available ways of making digital plan drawings on site in order to better inform the assessment stage of the Revelation project and to provide a sounding board with which to begin the work of assessing the wider sectorial experience of this form of digital recording. As part of the further investigation it is recommended that more research be done on the incorporation of digital section drawings into the record.

3. Aims of the assessment of handheld drawing

The main aims of the assessment of handheld drawing systems were as follows:

- Gain practical experience of recording examples of excavated archaeology using digital drawing methods during a typical CfA fieldwork project.

- Gather feedback from staff on site as to the feasibility of developing handheld digital drawing methods for the CfA excavation process.

- Test a number of existing platforms that have been used elsewhere and which are currently available to CfA “off-the shelf”.

- Produce a report of the outcomes of the field tests within the context of the wider Revelation project and to inform the full CfA report and the Project Design for any revised CfA systems.

4. *Methods tested and results of fieldwork to date*

The software tested in the course of the Whitby fieldwork was AutoCAD R14, PocketGIS v 1.6 and ESRI's ArcPad v 5.0.1.23. The hardware used to run these various pieces of software was a Hammerhead P233 computer running Windows95, a Fujitsu Pencentra 200 running Windows CE and a Compaq IPAQ PDA running Windows PocketPC. It should be noted that the trial was focused more upon testing software capabilities rather than suitability of specific hardware platforms for handheld field drawing. Therefore the results of the testing of each piece of drawing software are given in detail below while the issues arising from the hardware platforms used are dealt with more generically under section 5.

4.1. Digital Planning - software tested

4.1.1. Digital planning using AutoCAD

A field trial was carried out at Whitby Cliff 3 site in October 2002 by Eddie Lyons of CfA, an experienced AutoCAD user (see Eddie Lyons report Appendix A). The purpose of the trial was to determine whether a version of the AutoCAD drawing software could be used on a portable computer platform (in this case a Hammerhead computer) to produce drawings of a quality to match existing CfA records drawn in AutoCAD (Fig 1).

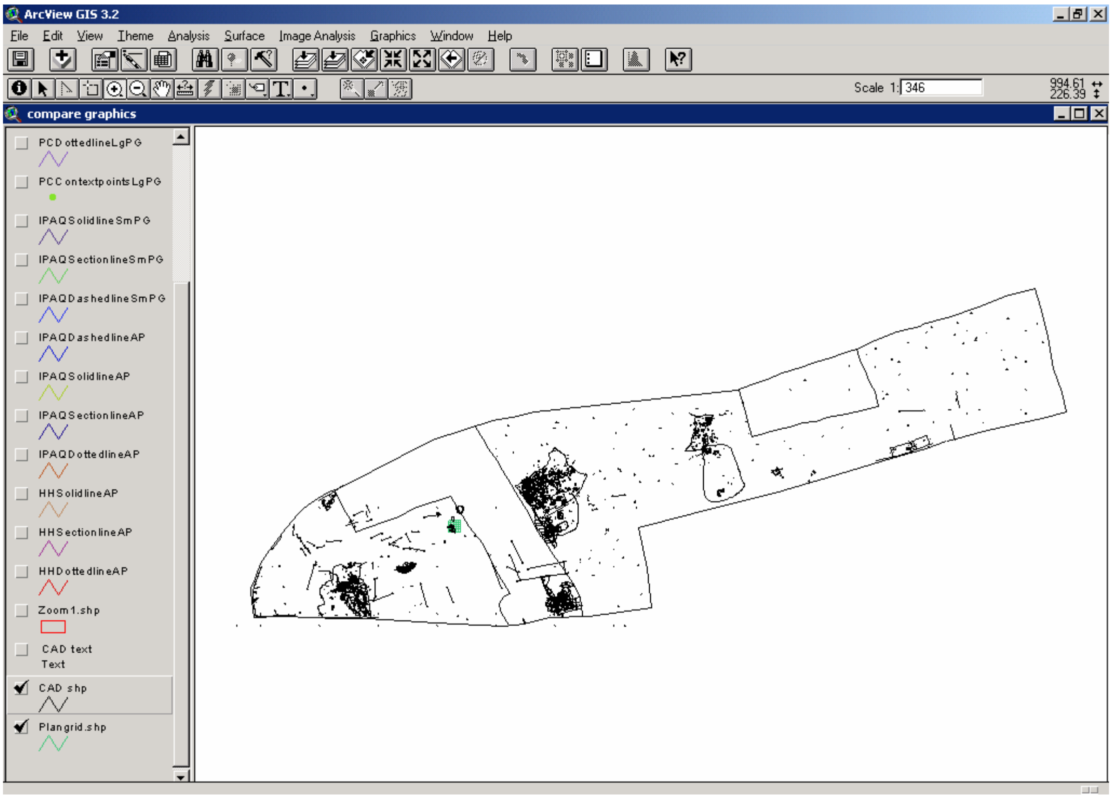


Figure 1. Overall site plan of Whitby Cliff 3 created in CAD and imported as a Shape file into ArcView 3.2 (the green digital planning grid marks the area used for the digital drawing trials).

For this to be achieved required a certain amount of R&D work to customize the AutoCAD software to enable "full direct planning capability". These customisations included a number of functions to enable the closing or entry of drawn lines while on the screen. Particularly of interest was the need to create a "digital planning frame" to be visible on screen to the drawer while using a real planning frame on site (Fig 2). In digital terms this was the equivalent of having squared paper under plain permatrace or grided permatrace (as used by organisations such as MoLAS).

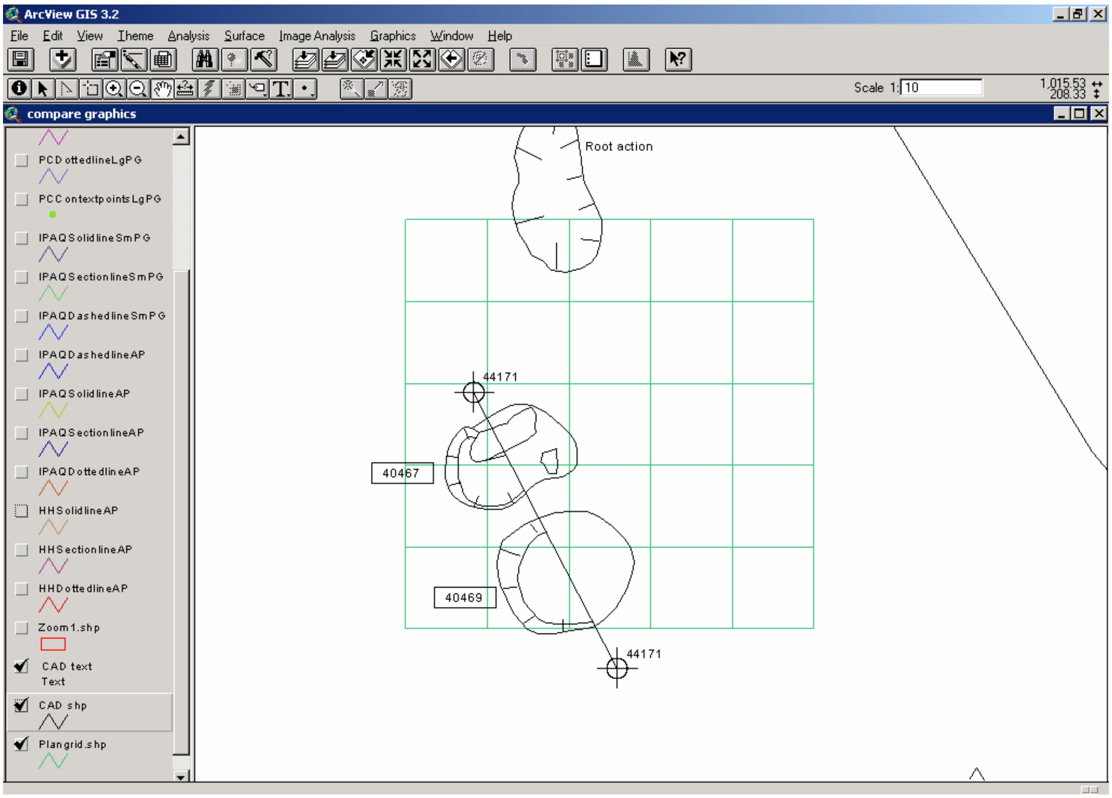


Figure 2. CAD plan drawn on site then imported into ArcView 3.2 with text labels of postholes and "digital planning frame".

Whilst the creation of the specific "Enter" & "Close" user tools for the front-end is an issue specific to the use of AutoCAD, it is likely that whatever form of software is used in the field, it will need a graphical representation of either the planning frame or some functionality which allows the viewing of grid lines of various resolutions on the screen. This is essential to enable the accurate location of drawn lines.

The conclusions of the AutoCAD testing were that the onsite planning was equally as effective in producing digital plans as planning offsite by digitising from hardcopy drawings. What the testing also revealed was a number of practical issues concerning the user interface and the hardware platform which would need to be successfully addressed in order to create a handheld recording platform which could adequately do the job as well as conventional permatrace and pencil.

The biggest issue with creating site drawings in AutoCAD is the question of how at the same time to record the data that needs to be associated with the digital drawing. In CAD the approach is to record each drawn context on a single layer and give that layer the context number. Attributes can be attached to the lines by turning them into "blocks" in CAD, but then these blocks cannot be further edited easily and the attributes are not available for sorting and searching in the way that a GIS database can do across all the drawn lines in the GIS. The CAD drawings with their blocks can be imported into a GIS but currently this then requires a considerable degree of re-processing and editing of the CAD drawing information to create fully intelligible spatial data within the GIS (Fig 3).

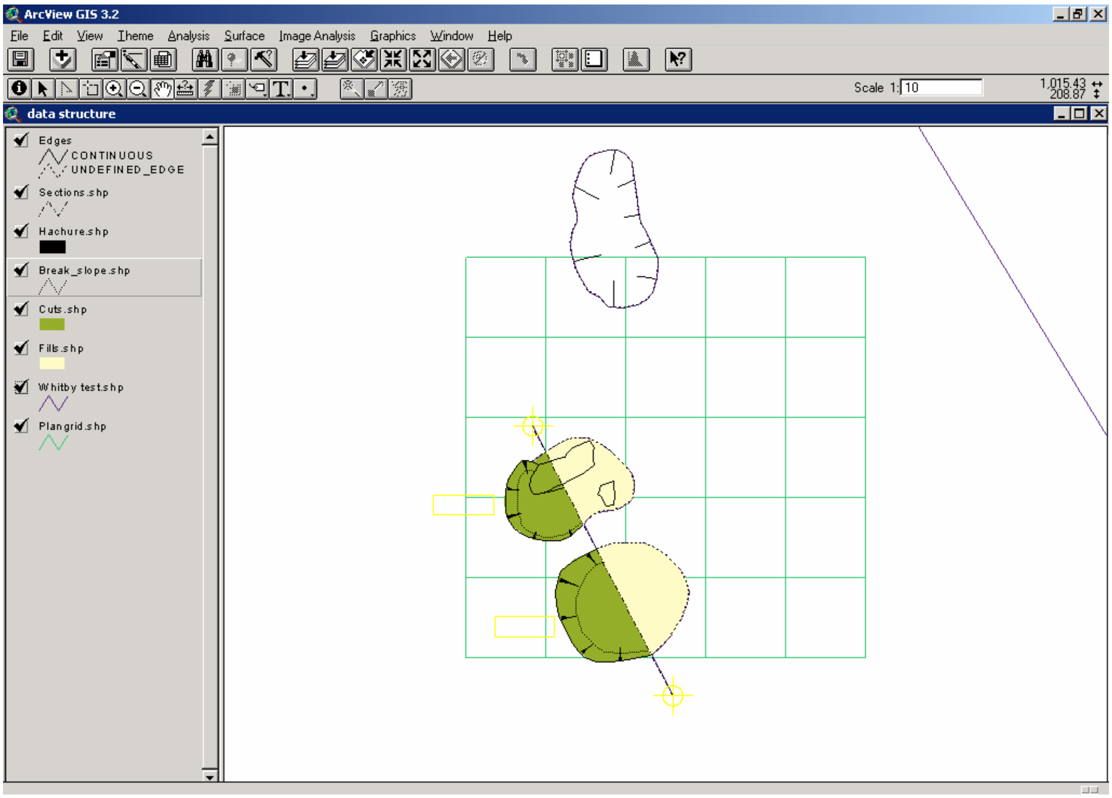


Figure 3. GIS data structure created in ArcView 3.2. The CAD plan of the postholes has been imported, but includes unwanted attributes e.g. text boxes, etc

Some GIS functionality may be provided by the Autodesk Map version of CAD although the fundamental point remains that CAD is at its best as a drawing package and is not expected to provide the full functionality of a GIS.

4.1.2. PocketGIS

PocketGIS was tested at Whitby Cliff using both the Pencentra and the IPAQ handheld devices. It was immediately noted that setting up the PocketGIS on the handheld devices required relatively good IT skills and experience. To get a copy of a database for entry of records onto either hardware device required carrying out all the necessary design processes to create a pre-formatted database on a laptop first. This was because once the database was on the handheld it was not possible to make alterations to the database design without going back to the original laptop set-up and downloading a revised version of the database.

PocketGIS is similar to AutoCAD in the way that the user has to tap the pen on the screen to create a series or group of points that make up the overall shape of the feature being drawn. This "connect-the-points" approach means that rather than using a freehand drawing method, the user may have to enter a fair number of separate points (referred to as vertex points) in order to produce a drawing of a curving edge – probably the most common drawing requirement of any drawn record in archaeology.

When using PocketGIS it was found to be difficult to draw on the screen at 1:20 scale as it proved very difficult to accurately judge where the precise point of the pen was on screen. (Fig 4 showing "small-scale" drawing).

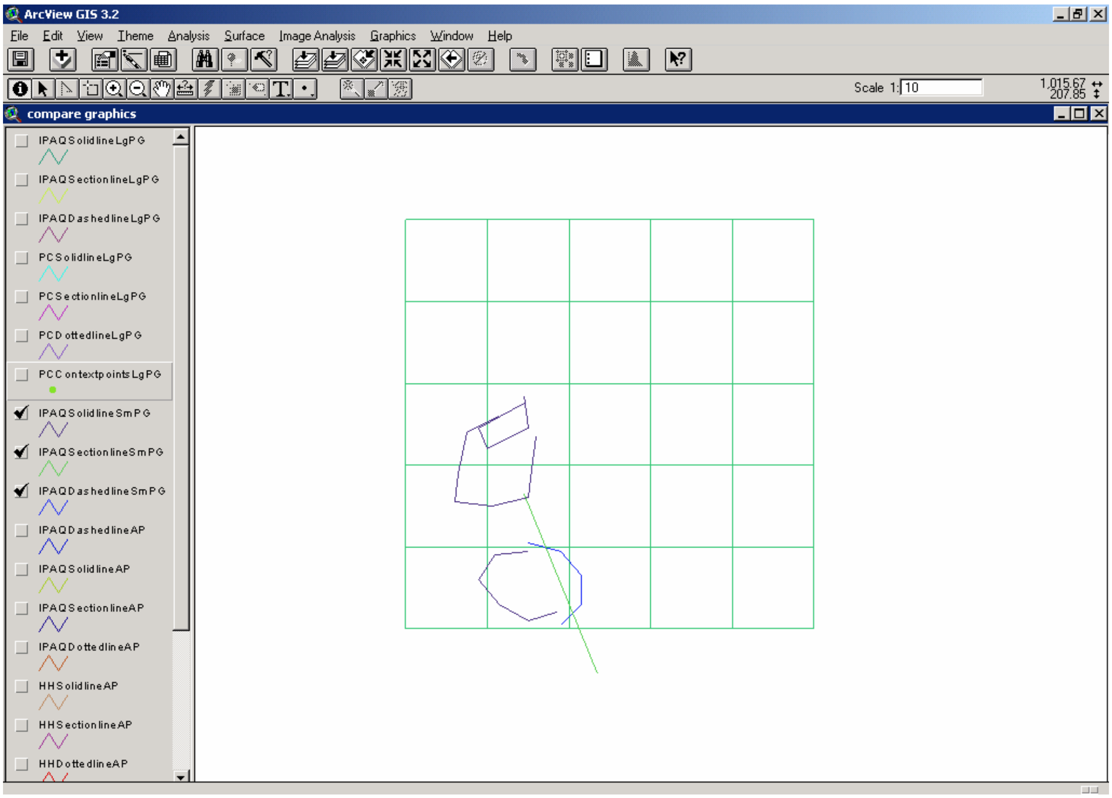


Figure 4. Posthole drawings made on the IPAQ at 1:20 scale (SM) using PocketGIS

Some improvement was possible by adjusting the scale of the drawing to a larger scale to improve the drawing resolution and then re-adjusting back to 1:20 (or 1:10) afterwards (see Fig 5 drawn at "large-scale" but displayed at 1:10).

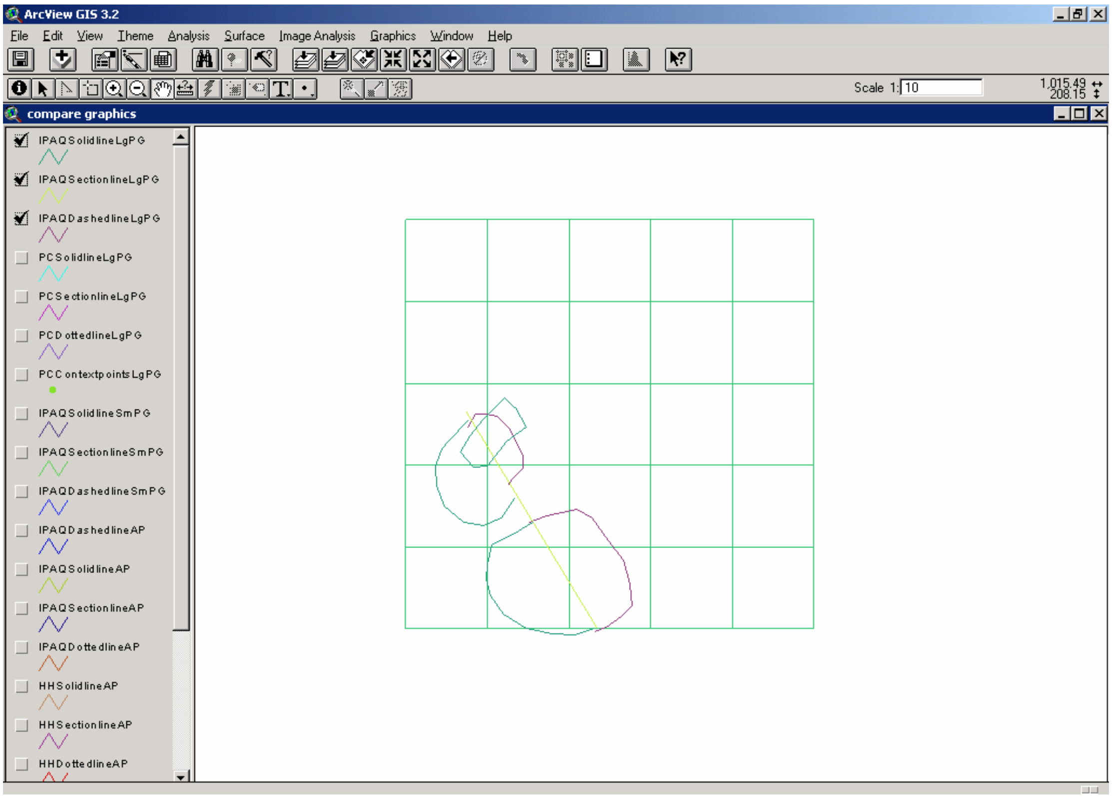


Figure 5. Posthole drawings made on the IPAQ in large scale (LG) using PocketGIS

This problem may have been exacerbated by the screen size on the available hardware, but it will remain an issue, as it requires prior consideration of the scale at which the digital drawings will be made. It should be an advantage of digital drawing that the drawer is able to adjust the scale to draw more detailed areas if appropriate. However unless the drawing is to be displayed at an appropriate scale the detail recorded may not be visible in the final presentation of the drawing.

PocketGIS has only two line types, either solid or dashed, unlike CAD, which allows adjustment of properties such as line thickness, spacing of dashes etc. This is rather limiting for use of most current archaeological drawings which use various types of dashed line conventions to represent different meanings such as edge of excavation, edge of layer, change of slope etc. The number of different line types used in archaeology is not however large and the fact that the GIS software allows the attaching of attributes which can hold information describing an almost limitless level of details about the line should enable a range of different line types to be displayed in future systems.

4.1.3. ArcPad GIS

ArcPad has the additional functionality of allowing freehand drawing on screen that most closely replicates the same process as drawing with a pencil on permatrace. To a novice user of ArcPad the interface is considerably easier and intuitive than PocketGIS and the combination of pull-down menus and keyboard entry allows much more scope for the user to interact and learn from feedback while using the tools.

However there are problems associated with the freehand tool. From the results of the drawings made at Whitby it appears that the accuracy of the freehand drawn lines can vary quite considerably depending upon the confidence and speed with which the pen-holder draws the line. Lines drawn slowly tend to appear jagged in the resulting drawing because the software attempts to interpret all the data it receives and the effect of moving the stylus more slowly means that more data is being transmitted to the screen and thereby captured as a jagged line in the resulting vector drawing. (Fig 9)

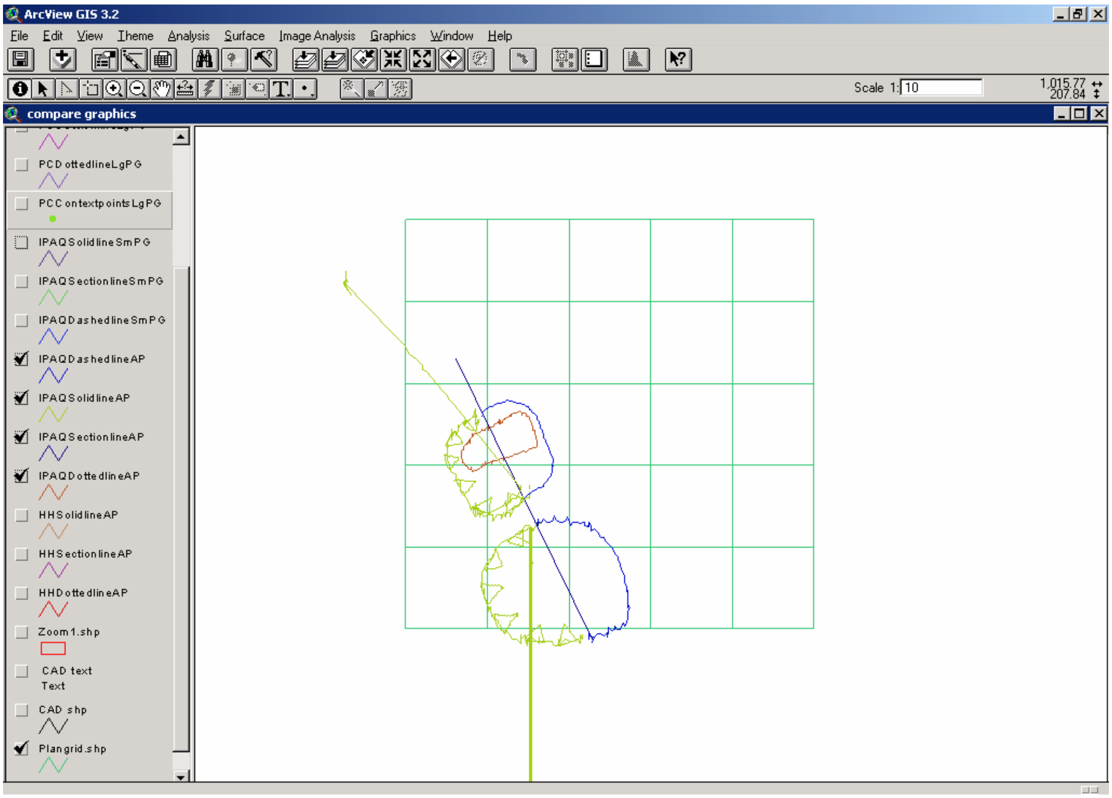


Figure 6. Posthole drawings made on the IPAQ using ArcPad's freehand tool

There is also a question of what scale to view the drawings at, as the lines appear relatively smooth when viewed at the scale used to record them but then become more jagged (pixcelated on screen) if the drawing is zoomed in upon (Fig 10).

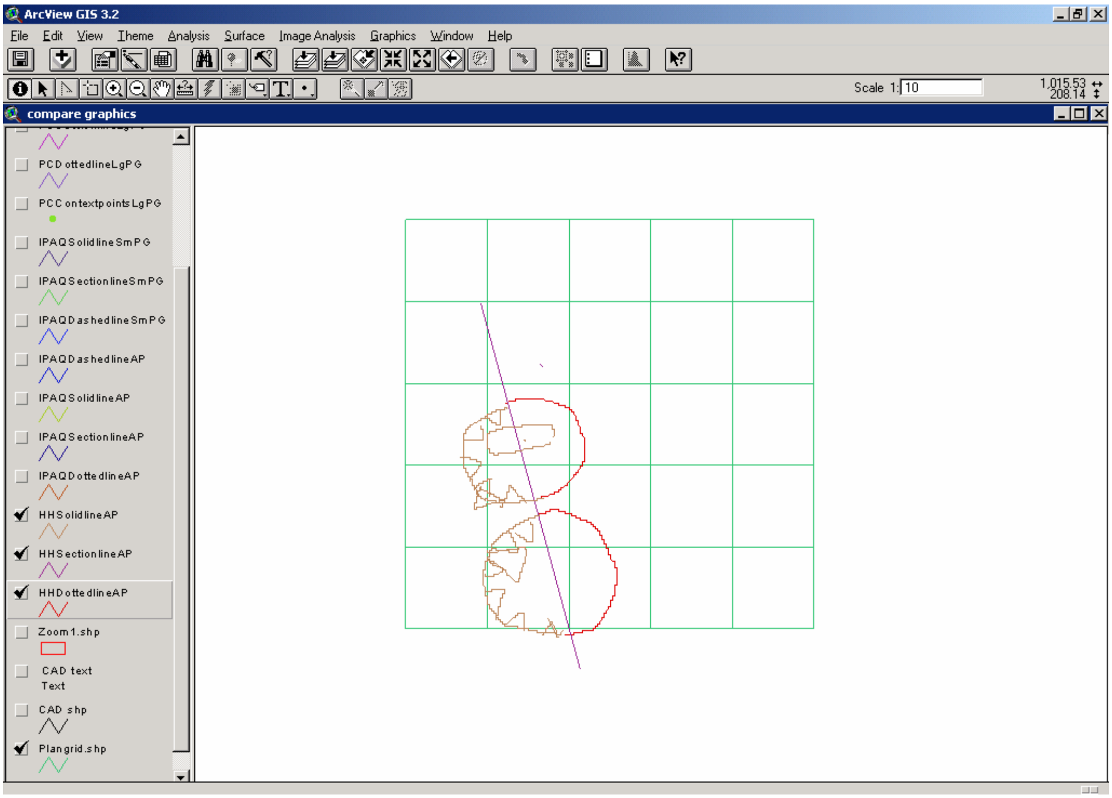


Figure 7. Posthole drawings made on the Hammerhead using ArcPad, beginning to show "jagged" lines when displayed at 1:10 scale

This is not just a problem of digital recording in the field, as of course the same will be true of drawings made from hardcopy plans and would also occur if a conventional pencil drawing were printed at an over-magnified scale in a book. The key requirement here is to define what use, or uses, the drawing is going to be needed for at the time it is created.

One problem encountered particularly in ArcPad during the trial was that once a drawing item was begun it was not possible to zoom in and out of the feature or pan around the screen until after the drawn line had been completed. This meant it was essential to make a careful choice of scale and initial position of the drawing on the screen before beginning to draw a single item to make sure it could all be fitted on the visible screen area. (Fig 9)

Like PocketGIS, ArcPad only has two available line types although it is likely that some increased functionality in this area could be developed and would be required as part of the development of any future digital drawing package.

4.2. Digital sketch-plans & sketch-sections - hardware & software tested

AutoCAD does have a "sketch" mode, and some custom tools have been developed in CAD that allow the user to drag the cursor around the screen and then (at the push of a button) turn the sketch into a polyline entity. Pocket GIS also has a sketch function built in, but given the limited time available during the Whitby fieldwork neither of these particular functions were tested at Whitby. It should be noted that the ArcPad free-hand tool which was used for the measured drawing tests at Whitby could serve as a sketching tool as well.

The results of handheld sketching described below are derived from fieldwork carried out during the DigIT project in August 2001. The software used at that time was TealPaint running on a PalmOS PDA platform and a piece of hardware called a Smartpad which included sketching software which was then further updated to an application called InkNote running on an InkLink pen. The InkLink pen was also used during the Whitby testing.

4.2.1. Sketch drawing on a handheld using Tealpaint PalmOS software

The main issue with sketching on the handheld screen was, not surprisingly, problems associated with scale and being able to move around the drawing and viewing all the necessary information on a small screen. It was possible to produce reasonably coherent sketches of records like sketch-sections and sketch-plans of cut features (see example in Fig 1).

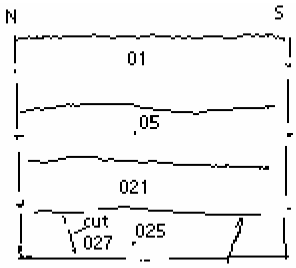


Figure 8. Example of sketch-section produced on handheld PDA using TealPaint

It did not prove so easy to create worthwhile sketches of larger areas such as trenches or long linear features such as plough-marks. These suffer particularly from lack of scale information although in principle that is an element that could be built in to a handheld drawing package such as ArchPad.

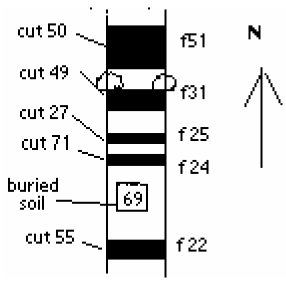


Figure 9. Example of sketch-plan of southern end of DigIT Trench AC produced on handheld PDA using TealPaint

4.2.2. InkNote & Smartpad – Seiko digital drawing devices

The Smartpad uses a conventional ballpoint point pen for drawing on a conventional piece of paper in ink. The infra-red detector in the pen then transmits a digitised version of the drawing to the PDA using a spatially referenced sensor on the sketchpad. This device was used to make a quite successful detailed sketch-plan of an area of chalk rubble, but unfortunately because of the proprietary format of the resulting drawing it was not possible at the time of the DigIT project to export the resulting sketch to any of the windows based packages for use in the final record. Such a proprietary format would not be considered appropriate for creating drawn records that need to be shared across a number of different software platforms.

Further to using the Smartpad on DigIT, about 8 months after DigIT finished Seiko released a version of the InkLink pen which has a sensor that can be clipped on to a clipboard. The principle largely remains the same but now the drawings can be exported in BMP, JPG & PNG formats. Although this may be a step in the right direction for recording some sketch information there are still major limitations on its actual use for archaeological drawing. Firstly the pen is designed to draw using an ink ballpoint pen, as it was principally designed to record hand written notes using handwriting recognition software. This currently means that without some modification it is not possible to draw in pencil so that it is not feasible to draw on Permatrace, unless one is prepared to ink over, and thereby digitise, existing pencil drawings. Even more limiting was the lack of an erase function while drawing so that any sketches produced either have to be without errors or would need re-editing once exported which is time-consuming and likely to incur further errors.

The InkLink pen does demonstrate that technologically the possibility of producing freehand drawn digital drawings directly in the field is tantalizingly close to being achievable. The principle limitation of the InkLink pen is that it produces only raster images and not vector drawings. Although the resulting images can be exported in various formats it is as yet not possible to attach other information, such as the context number of the drawn feature, to the individual elements within the raster images.

The sketch pen does however at least fulfil the same sort of function as that provided by the level of a sketch on the back of a context sheet where the excavator is seeking to illustrate interpretive detail rather than record an exact measured drawing to be used as the final record of an archaeological feature.

4.3. Digital Section drawings

4.3.1. Do we want or need to record sections digitally?

The Whitby trials did not attempt to draw any sections using the available digital recording methods. Recording sections using the digital drawing tools outlined above would be possible, but there is a more fundamental question about the nature of sections which make them different to plans in both function and use and this needs to be recognised when deciding on the best approach to creating any digital drawings of sections. Unlike plans, which are usually recorded to show separate individual contexts and often using a separate drawing for each context, sections (and to a lesser degree profiles) are drawn to show the physical vertical and stratigraphic relationships between a number of different contexts, layers, cuts and fills. A typical section drawing will usually contain considerably more than one context number.

GIS deals primarily with planometric data (i.e. data viewed in plan) and is not designed to handle vertical data of the type represented in section drawings. Because of this any sections that were drawn digitally using CAD would need to be handled differently to plans in a GIS. This may affect the choice of approach to how best to record digital section drawings.

One possibility is to treat any digital record of sections rather like the photographic records made on site. In this way the sections might be spatially referenced according to the current practice of surveying-in the positions of the ends of the section lines but the record of the section could be held as a image, either as an annotated photographic record or as a scanned image of conventional pen-drawing. The sections could then be held as any other digital image data in the GIS database and could be viewed in a pop-up window. Questions remain about the best way to integrate non-spatial section data (such as section numbers and label information) with other non-spatial data in the context record.

4.4. Other digital drawing – further considerations

4.4.1. Photo rectification to enable annotating photographic records either on or off site?

An interesting developmental aspect of the Whitby trial fieldwork was a short piece of work on rectification of digital photographs taken on site, to form a raster-based image as a back-drop for the digital drawings. This was achieved by "rubber-sheeting" the digital photo into position using known coordinates for georeferencing (Fig 10).

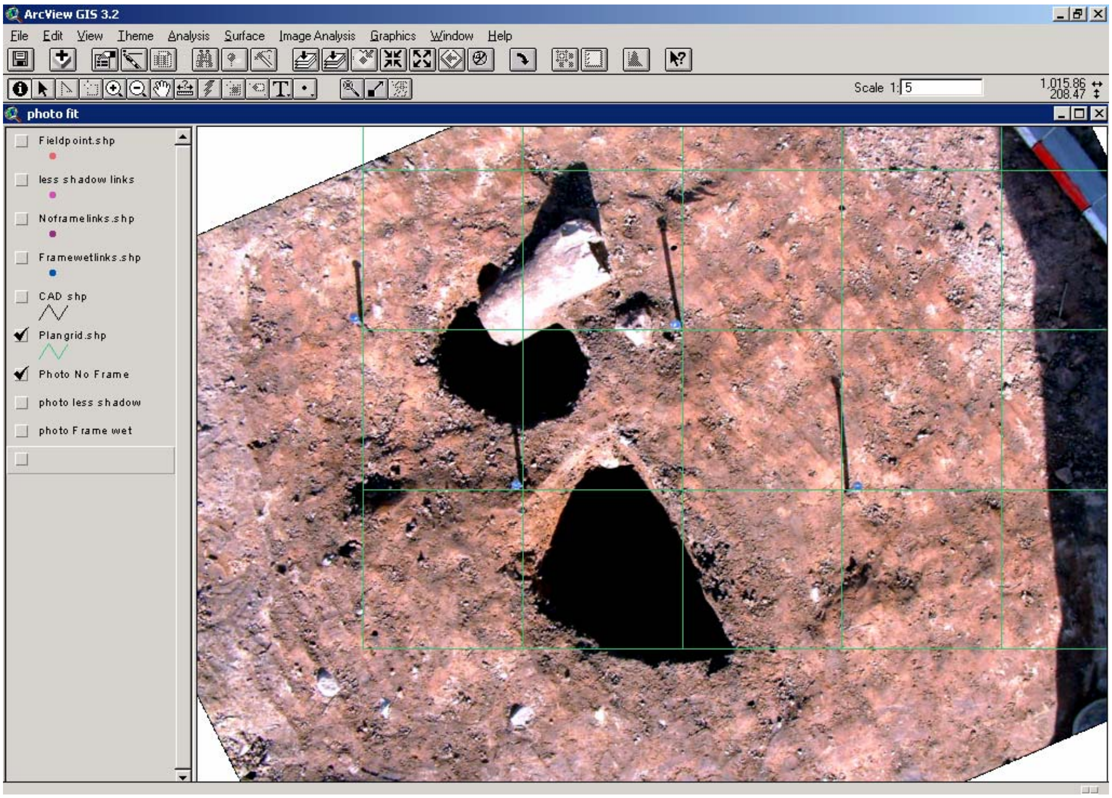


Figure 10. Photo of postholes taken in strong sunlight with no physical planning frame but showing digital planning grid and georeferencing points

The results quite dramatically enhance the level of background visual information, particularly when used as a background to the drawn digital record (Fig 11).

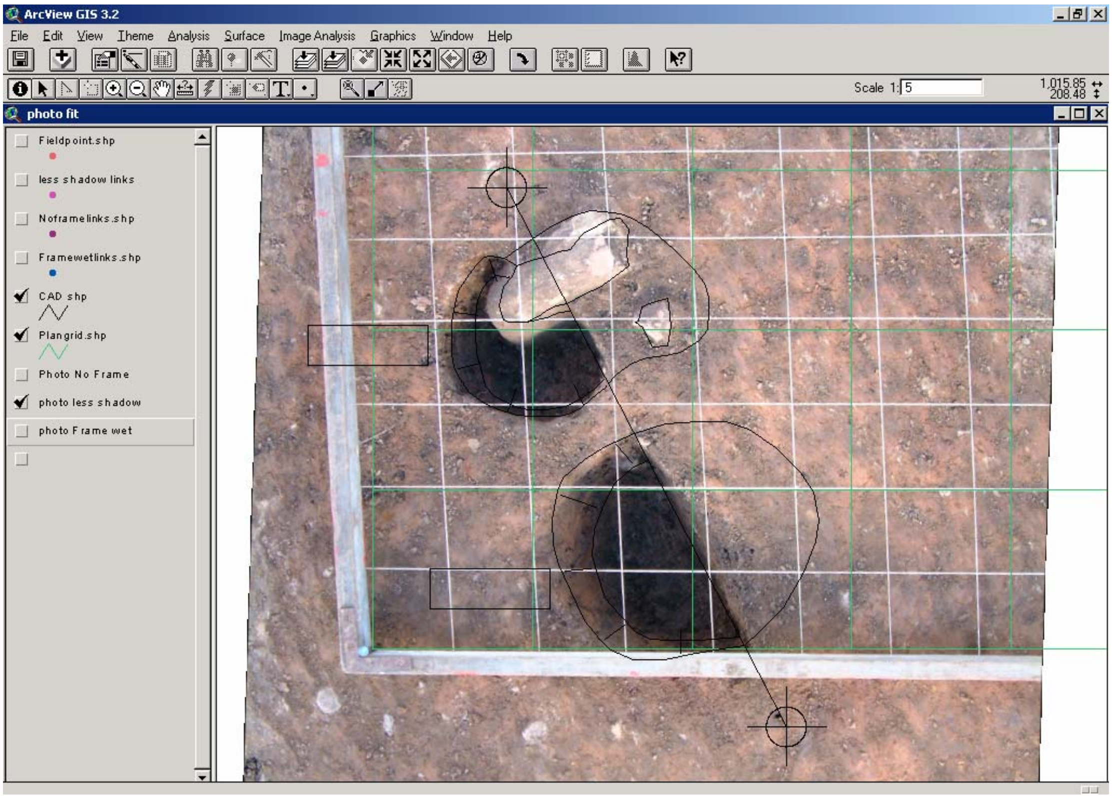


Figure 11. Rectified photo in less sunlight with CAD overlay at 1:5 and physical frame & digital planning grid (note discrepancies)

Questions remained about the degree of accuracy of the rectified photograph, when compared to the site grid and the likelihood of greater inaccuracies on sites with steeply sloping ground. In particular the process also raised some questions in the minds of the recorders about whether future excavators might be tempted to “cut corners” by tracing from the rectified photo off-site rather than actually recording from the archaeology present in the field.

The results obtained from the limited amount of photo rectification carried out suggested that it would definitely be worth carrying out further research into the most appropriate applications and methods for this technique in the future, bearing in mind some of the variables that could result from varying degrees of light and weather conditions (compare Fig 12 with Figs 10 & 11).

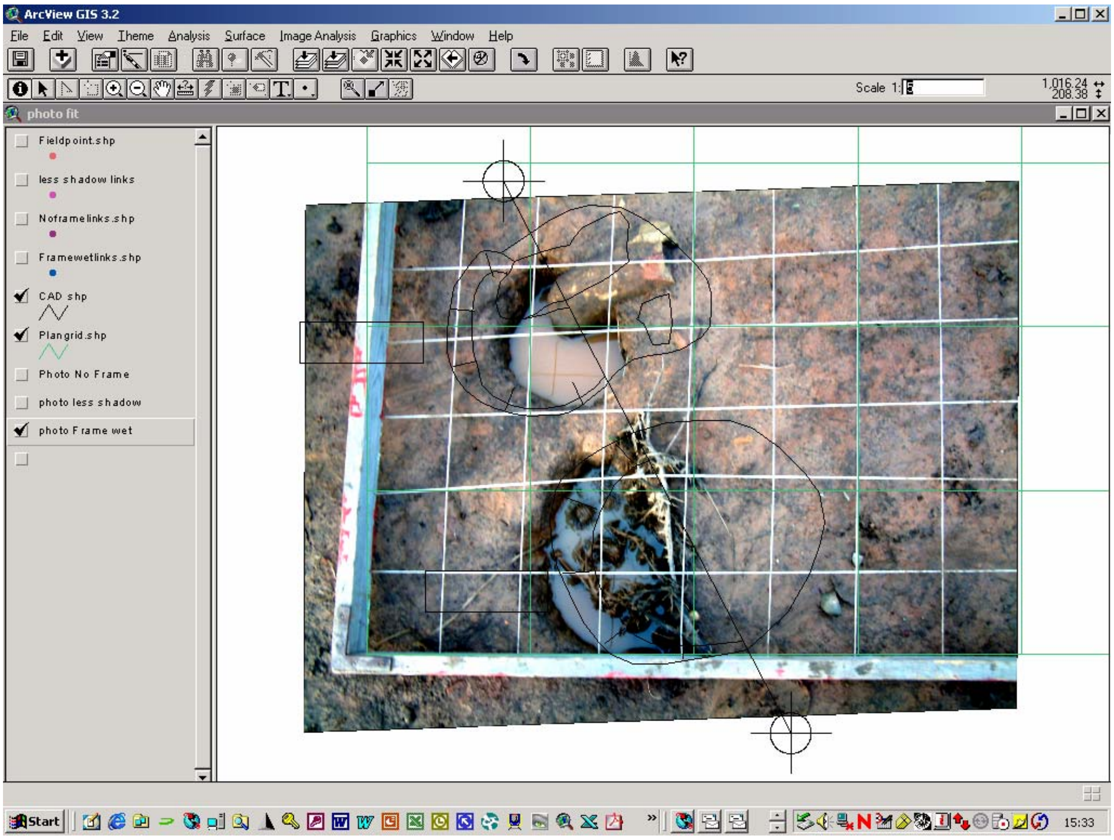


Figure 12. Georeferenced photo taken following heavy rain with planning frame and digital grid

4.4.2. Finds drawings

Drawing of finds was not investigated as part of the fieldwork trials as it is normally carried out as a post-excavation process. It is raised here in this assessment because if we are creating drawings on site, which incorporate spatial data about finds, then it would be beneficial to consider including a process in the eventual publication drawing of those finds that enables integration of the finds drawings with the information that was gathered about them on site.

4.5. What worked and what didn't – Criteria for assessment

4.5.1. Criteria for Assessment

The principle criteria used to assess the digital drawing methods were the following:

- Can it at least do the task as well as what we do now?
- Does it provide new additional functionality?
- How compatible with other software?
- What were the hardware advantages or limitations?
- Ease of use on site
- Compatibility with processes for post-excavation & dissemination

4.5.2. Can it at least do the task as well as what we do now?

General Issues

Given the current level of technology available (in 2002) it is unlikely that every conceivable archaeological site will be suitable for using on-site digital drawing. However it does seem from the experience of the CfA and other organisations that current digital drawing systems are now able to produce results which are as good as the conventional recording methods under appropriate and favourable circumstances.

Most importantly it was felt from the results of the trials and other CfA work practice that the introduction of digital methods and systems of recording can greatly enhance the speed of delivery, degrees of accuracy and ability to re-use drawing information. A system that incorporates digital drawing should be able to improve the processes that we currently carry out and the question must also be asked "do we need to continue to do things exactly the way we do now"?

A comparison could be made between the adoption of computer drawn plans over pencil drawings and the adoption of computer written letters over hand-written letters. If as the end result you want a beautiful hand-written letter you may take up calligraphy, but if the actual requirement is have the information held within that letter then it is more effective to have the letter as a word-processed document and probably as an email that can be sent most efficiently to the person needing the information. The key here is defining the purpose for which the letter - or drawing - is intended.

The simple fact is that a computer drawing is not a pencil drawing and no amount of argument is going to change that. Any drawing package that uses "connect-the-points" line technology intuitively seems to be reducing the quality (i.e. level of detail) of the eventual result. However when planning using a pencil we actually draw by measuring a series of "fixed points" with the tape or in relation to a planning frame and then "eyeing in", with varying degrees of skill and accuracy, the placement of the line in between. This process does not seem dramatically different from marking the fixed points on the screen and then using the software to "eye-in" the line in between the points. Some further research should be carried out to see how good the computer drawing tools are at producing closely spaced curving lines, particularly in complicated circumstances such as a plan of the undercutting in steep sided pits dug into soft sand.

The question of how best to record sections needs to be resolved. This should be addressed from the viewpoint of what uses do we need sections for, how can we best record the data contained in sections to fulfil those needs, and how can the chosen method be best integrated into the rest of the recording system.

AutoCAD Issues

To date the CfA approach to recording digital drawings using CAD has primarily been based on the need to display visual representations of information on the screen or in printed versions of the drawing (e.g. see AutoCad.

For CfA this has mainly been for purposes of analysis, report production and archive. As such the existing data structures for CfA use of AutoCAD do not fully utilise the ability to attach data to the drawing objects for incorporation of spatial data with non-spatial data. Rather AutoCAD uses a number of different layering conventions and drawing symbols to represent different attribute data.

If AutoCAD is to be used as an on-site drawing tool it must fulfil the requirement to record and export all the information contained in the drawing in a data structure that can be effectively managed by GIS software such as ArcView. For example this would mean that data such as levels should be recorded as point data with an attribute field in a database that records the level height as intelligent and searchable data, rather than simply using a drawing convention symbol that is not then available as intelligent data if the drawing is used in other software.

If AutoCAD is to be used on site then a range of additional tools may need to be developed as additional software features to make AutoCAD fully suited to direct planning. Further research should be carried out to see if such tools have been designed for AutoCAD or other compatible graphics software.

4.5.3. Does it provide new additional functionality?

There are a number of areas where digital drawing can provide new functionality. The first is the ability to be able to attach a whole range of intelligent information to a digital drawing in a way that simply is not possible with a paper-based drawing.

This should be of benefit to archaeologists at all stages in the process. There are advantages for excavators on site who should be able to review and update site data and relationships between different drawn records using an intra-site geographic information system. There are advantages to staff during the post-excavation processes of analysing, grouping and interpreting site records using all the available spatial information from the site. There are advantages for external specialists in being able to access the excavation records and drawings that pertain to their specialism and being more informed when adding their data into the relevant areas of the analysis process.

The advent of electronic publishing using new media such as the internet and web browser technology means that by creating drawings digitally and including links to the associated information there is a greatly increased ability for using the information contained within the drawing. The information can be disseminated far more quickly, efficiently and widely than if it remains on a piece of paper or permatrace.

A useful summary of some important advantages of direct digital drawing and data capture gained from commercial organisations such as the MoLAS are give below. (Ziebart et al. 2002):

- Speed and accuracy of recording both survey and archaeological data on site
- Ability to relate sites accurately to the OS national grid
- On-site graphic capacity to check data
- Use of site digital data to develop excavation strategies on site
- Ability to deliver digital data rapidly to clients
- Ease of transfer to a GIS system for post-excavation analysis

4.5.4. How compatible with other software?

Both the GIS products uploaded directly to ArcView (and therefore presumably also into other vector GIS software) with no loss of spatial or attribute data (see Figs 5 & 7). The AutoCAD has considerable problems importing into ArcView 3.2 but far fewer problems in ArcView 8. The main issue for GIS software compatibility, as stated already, is to improve the data structure of the AutoCAD drawings to allow ease of use within a GIS data structure.

The final archiving format of the drawing data should also be considered. At present there are issues about how completely all the attributes from CAD drawings can be exported and archived (even using the standard DXF export format). SGML or its derivative XML should be considered as the archiving format for digital drawings and the requirements of the Governments eGMF metadata standards should be incorporated into any long-term archiving needs.

4.5.5. What were the Hardware advantages or limitations?

Hammerhead

The hardware used for the AutoCAD testing was found to be quite problematic. In previous fieldwork the CfA have mainly used the Hammerhead computer while attached to a surveying tripod. As a result although it has a reasonable size screen for drawing which made it useful for test purposes, it lacks genuine portability and carries a number of additional elements which make it quite an awkward shape to hold or move around with. In particular it's overall weight proved a noticeable burden for the user after only one days use. These factors together make it unsuitable for any long-term use as a handheld drawing platform for sustained fieldwork in the future.

This is not seen as a viable device to use for day-to-day on site digital drawing

Pencentra

The screen size of the Pencentra was similar to the Hammerhead and this made it a useful platform for testing the different software. The weight and general portability of the Pencentra made it a much more appropriate platform for digital drawing on site. However there were problems with the Pencentras colour screen and it's suitability for sustained outdoor recording work.

One particular issue was the Windows CE operating system. This resulted in issues of not being able to alter the design of the database on the Pencentra without re-establishing communication links with the parent laptop and copying a revised version of the database structure on to the Pencentra

IPAQ

The IPAQs size makes it easily the most portable of the devices tested. However this portability is countered from a drawing point of view by the necessarily small screen size when drawing. This raises considerable doubts about it's suitability as a handheld drawing platform for recording digital drawings.

Where the IPAQ is more likely to have a role is in the collection of digital records, such as the standard context record information. In addition the display functionality of the IPAQ is probably good enough for allowing a browse capability of plans that have already been created as part of the site record.

One could therefore envisage an archaeologist using the IPAQ to record and update the context records on site with the ability to cross-check the plans and drawn records that pertained to the context being recorded.

4.5.6. Ease of use on site

The Whitby trial was not expected to identify a hardware platform that would be instantly suitable for all the requirements of digital drawing on site. As anticipated none of the hardware devices tested at Whitby proved ideal for the task of drawing on site, although possibly the Pencentra came closest to providing a viable hardware drawing platform.

What the trial did help focus attention on was the range of requirements that would be needed from a digital drawing platform and in particular the need for adequate screen size. It was felt that this might be provided by a version of the new Tablet PCs that are now emerging on the commercial market and which are specifically designed with mobile and fieldwork computing in mind.

Screen visibility

The field testing was carried out using both colour and black and white screens. When used outdoors in normal daylight there were some difficulties with seeing colour lines on screen, particularly when using light colours. In general the black and white screens were more visible. For most current field drawing purposes black and white drawing would seem to be sufficient as archaeological field drawings are currently only in black and white. Some thought should be given to how best to create field drawings that might need the addition of colour during the post excavation process.

Indoor/outdoor screens

It was noted during the Whitby tests that none of the devices tested had purpose-designed outdoor screens and that some further research was needed into the difference between different screen options when working indoors or outdoors in varying conditions from bright sunlight to heavy rain.

It was clear that further research and experimentation needs to be carried out using other screen devices and different colour drawing options in varying degrees of light, both indoors and outdoors to find the best available hardware solution.

Weather conditions

The AutoCad testing using the Hammerhead was carried out while the weather was dry, (although cold and windy). The Hammerhead has however often been used on other fieldwork projects in wet weather conditions and has proved waterproof. The remaining testing at Whitby was carried out at a time when heavy rain was affecting access to the site. No attempt was made to record during the rain, as no specific testing using waterproofing covers had been planned. However the site access was restricted to all excavation staff during the rain regardless of the activities they were carrying out and it was pointed out that the rain also brought a halt to paper-based recording.

Need for electrical power

This was not a major issue for the trial as the Whitby excavation was using a generator and the equipment used had adequate battery power, although a number of additional plug adaptors and cables were needed. It is however a logistical factor to be born in mind and some research will be needed into the best and most efficient means of powering digital drawing devices if they are to be used in numbers and on a daily basis, particularly on more remote archaeological sites.

Communications between devices

During the trial work at Whitby this was one area that caused considerable practical difficulties, although some of this could be attributed to the experimental nature of the testing. The problems of transferring information between machines were generally overcome as the final results show but this will be an area that needs considerable planning as part of a viable fieldwork system. Many of the detailed hardware issues were specific to the devices used for the test (see report on trials Appendix B).

In general communications between computers – and particularly mobile communications for out of office working – is probably the fastest moving industry on the planet at this time. This suggests that the best solution for communications will need to be sought at the time a systems design is put in place. One area that is highlighted by the digital drawing trial is that there will need to be good procedures and mechanisms in place for saving any digital drawings as early as possible on to a server in the site office and ideally that if the full functionality of digital drawings are to be utilized then the use of wireless networking should be investigated both to save drawn records to the back-up server and also to enable the drawings to be served back to excavators and those making records in the field.

4.5.7. Compatibility with processes for post-excavation & dissemination

The quality of digital drawings must at some point in the process already achieve a standard that is acceptable for “what we do now” in regard to publication and dissemination as all archaeological publications are now printed using computer technology.

What is needed, as part of the overall review of systems and processes, is to identify the appropriate levels of Q&A required and build the requirements for post-excavation and dissemination into any new system so that the drawing information generated on site can be used with the minimum need for additional transformation.

5. Outcomes & conclusions

It does seem feasible that by reviewing the existing systems and processes used by CfA that a system could be designed that incorporates the use of on-site digital drawing, especially given that the plan records from the latest work at Whitby cliff have already been digitised during the excavation of the site.

At the broadest level the feasibility of digital drawing directly on site was demonstrated under test conditions in the field but with many practical issues still to resolve to enable a genuine working system to be developed. However the advantages that would accrue from being able to develop such a system, definitely make it worthwhile pursuing.

The outcomes of the Whitby trial work on the currently available software and hardware suggest that a possible solution may be offered by a combination of using drawing software such as AutoCAD that can provide the best available digital drafting tools but with the requirement that the process of producing any drawings on site must automatically include the creation of data in a database format which will enable the functionality of GIS in managing both the spatial data (e.g. drawings) and the non-spatial data that goes with them (e.g. context records in a database). A representation of this at a conceptual level is given below.

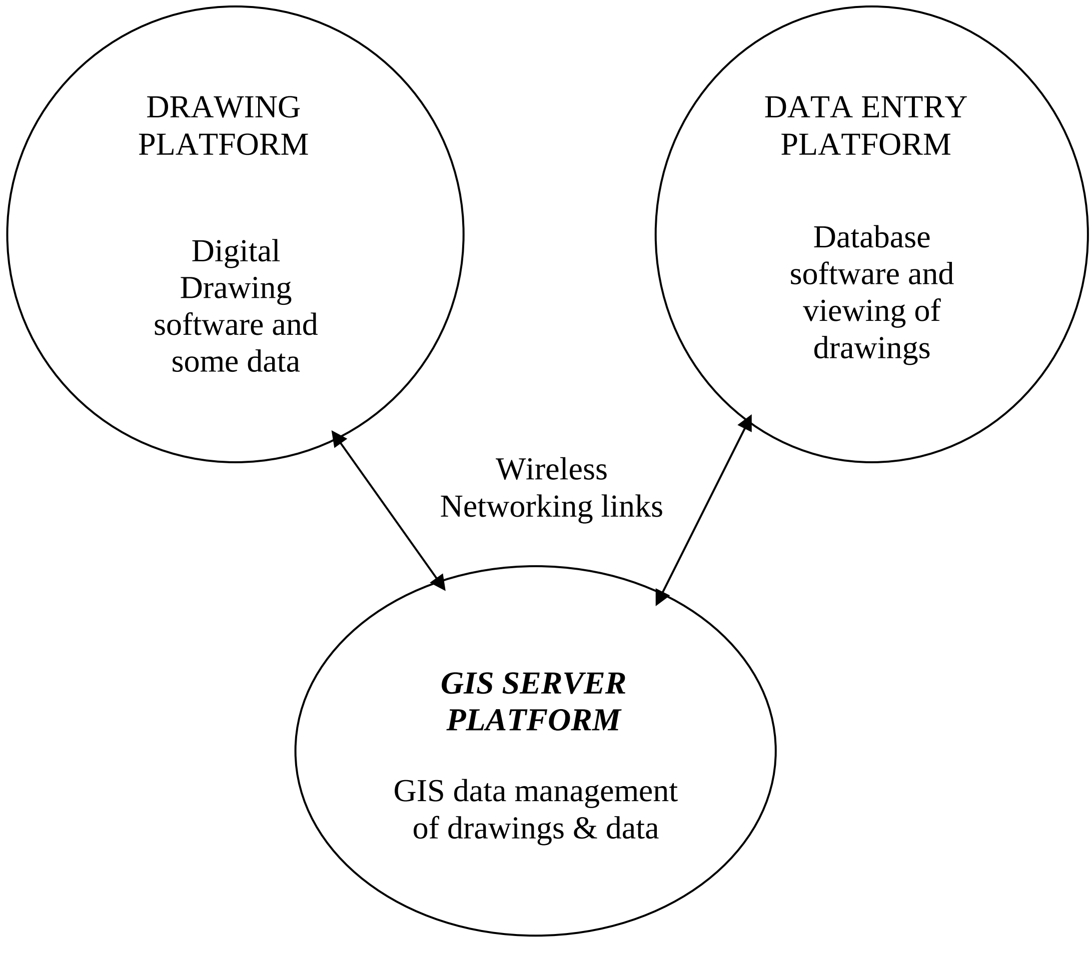


Fig 3. Model of a fieldwork recording system

It is possible that a single hardware platform might be developed that would enable both drawing and database entry to be carried out on a single piece of equipment. The model above suggests that under current working conditions the system would probably be based upon using separate hardware drawing platforms and context database recording devices, but crucially with the necessary data available to either hardware device. This is a reflection of the current practice of using paper based context sheets that are separate from the main drawings that are done on permatrace. Current field practice suggests that archaeological excavators spend considerably more time recording context record data as text, rather than drawing plans or sections.

The experience of current fieldwork practice suggests that there is likely to continue to be a far greater requirement during site work for individual archaeologists to need a text and numeric data entry device, such as a small and portable handheld PDA, during the process of excavating the site. This would be the digital equivalent of the excavator's clipboard and context sheets. It is rarely, if ever, the case that an excavation requires a drawing board for every archaeological excavator on site. Common practice is to have a number of drawing boards that are shared by excavators on site and there may be factors of cost effectiveness that make it more practical to give all excavators a handheld PDA but only have a smaller number of drawing tablets to be used on a "need-to-draw basis". It is however seen as absolutely crucial to the best use of creating digital drawings on site that the resulting drawings can be viewed by excavators in combination with any non-spatial digital records that they may be cross-referencing.

The feelings of those involved in the field trials was that the best approach to producing a viable system was to consider a "toolkit" approach whereby a combination of different digital tools would be used in digital recording just as is used in current recording work. The main difference is that by using a digital media for capturing and storing the data then the same piece of information is available for re-use for a range of different purposes throughout the whole archaeological process. Creating a single digital drawing as early as possible in the process must be more appropriate than having to take a drawn record on a piece of permatrace, have it inked in as a permanent record, have the drawing scanned as an image to allow someone then to trace it using a digitising tablet to create a drawing which is then "cleaned" and re-processed to produce an image that is considered appropriate for conventional publication in a book or journal. The digital toolkit should be designed to enable a range of tools that best create the drawings that are fit for all the required purposes.

So can digital drawings be made in the field that can be used for each of the following purposes?

- Primary record
- Analytical interpretive models
- Dissemination of the interpreted results
- Archive of the site

The answer is surely yes as it is already the case that we have examples of digital drawings created at each of the above stages. The aim should therefore be to design a digital drawing process that takes all these uses into account when the drawings are being created and to design the tools that enable that process to be carried out more quickly and effectively than the existing systems.

6. Potential & proposals for future work

Inevitably the fieldwork trials have highlighted some issues that may require further investigation and potentially further research and design work.

6.1. Recommendations

6.1.1. **It is recommended** that further work be carried out to examine how best the drawing capabilities available in drawing software such as AutoCAD can be combined with the appropriate tools and data structures needed to create an integrated digital site record that can be utilised throughout the whole archaeological process. This would require that the drawn plans of individual contexts are intelligible and have the required data structure to associate all the necessary individual context data to the drawing at the point that the drawing is created digitally. Testing of drawing platforms should include experienced planners working with the systems for at least a week and commenting.

6.1.2. **It is recommended** that further research be carried out to more accurately identify where a range of methods of digital data capture may be most appropriate and cost-effective dependent upon the site conditions and location, the complexity of the archaeology and the amount of time available to make the record.

6.1.3. **It is recommended** that further consideration be given to how best to record sections using digital drawing methods and integrate these with other computer records in the system.

6.1.4. **It is recommended** that further consideration be given to the choice of appropriate scale for recording and publication/dissemination of drawings. It may be appropriate that digital drawing can take place at various scales depending upon the complexity of the archaeology being recorded. However some form of convention will still be needed for the minimum scale that is used on site for any plans that are to be used in a printed publication format. The minimum recording scale should be twice that required for publication. The scale used should be appropriate to the requirements for record, analysis, publication and archive.

6.1.5. **It is recommended** that further research and experimentation be carried out into the accuracy and suitability of photo rectification as a means to gain additional pre-excavation and post-excavation data for use as a supplement to the drawn record, but also potentially for additional analysis work.

6.1.6. **It is recommended** that further research is carried out using a Tablet PC device running AutoCAD and/or Autodesk Map, ArcView 8, ArcPad and/or PocketGIS software.

6.1.7. **It is recommended** that the requirements of the Governments eGMF metadata standards should be incorporated into any long-term archiving needs for digital drawings.

6.1.8. **It is recommended** that further research and experimentation is carried out using alternative screens and different colour drawing options in varying degrees of light, both indoors and outdoors to find the best available hardware solution.

6.1.9. **It is recommended** that further research and experimentation is carried out as to the feasibility of using digital drawing devices during moderately wet conditions, such as those where current drawing methods are still possible (e.g. relatively light rain showers). This should include some testing of wet weather covers such as those used in underwater diving and as used for PDAs on the DigIT project.

6.1.10. **It is recommended** that further background research be carried out in to the use of wireless networks to investigate methods for the updating of drawing data recorded on a portable device in the field to be transmitted and saved to a server in the site office.

6.1.10 **It is recommended** that considerations of spatial data structure form a key component in overall consideration in the design of data structure for any further Revelation stages.

6.1.11 **It is recommended** that special attention is paid to issues of data flow and data structure when dealing with drawing issues in the literature review.

7. Bibliography & Glossary

7.1. Bibliography

Biswell, S., Cropper, L., Evans, J., Gaffney, V. and Leach, P.1995. GIS and excavation: a cautionary tale from Shepton Mallet, Somerset, England. In G. Lock and Z. Stancic (eds) *Archaeology and Geographical Information Systems: a European Perspective*: 269-285. London: Taylor & Francis.

Ziebart, M., Holder N., & Dare, P. 2002. Field Digital data acquisition (FDA) using total station and pencomputer: A working methodology. In D. Wheatley, G. Earl and S. Poppy (eds) *Contemporary themes in archaeological computing:* 58-64. Oxbow. Oxford.

7.2. Glossary

CfA – Centre for Archaeology.
DigIT –Digital recording and Information Technology project.
TST – Total Station Theodolite.
CAD – Computer Aided Design.
TheoLT – Survey recording software developed by Bill Blake of EH.
Photogrammetry – Stereo photography used to record elevations and plans on computer.
Attributes – Data that describe properties of a point, line, or polygon in GIS. In CAD attributes are non-spatial data appended to a symbol (e.g. of a Block entity in AutoCAD).
Georeference – To position a drawing in geographic space on a computer system using known coordinate points.
Permatrace – A proprietary name for drawing film used as a more robust alternative to paper for recording archaeological drawings.
GPS – Global Positioning Satellite
Polyline – A single entity consisting of one or more line segments between sequential vertexes, defining an open or closed area. It is distinct from a Line entity, which is a single line segment only between two vertexes.
Polygon – The area defining a single enclosed entity in a computer drawing.
Raster – Raster files hold attributes in the form of a grid. This is useful for continuous data (such as elevations or geophysical survey data) particularly prior to further interpretation. Images also can be raster files.
Vector – Vector files hold attribute data in the form of linked points creating bounded entities in a drawing - e.g. polylines or polygons of context edges. This is most useful for spatial data that has been interpreted or categorised to include context record information.
Penmap – Pen based computer drawing software.
PDA – Personal Digital Assistant.
PalmOS – Palm Operating System used on some PDAs rather than Microsoft software.
BMP – Bitmap digital image format.
JPG – Probably the most common and current image compression format.
PNG – Portable Network Graphics. An emerging digital imaging format that gives “lossless” compression

8. *Appendices*

8.1. Appendix A – *Whitby Cliff 3: AutoCAD R&D.* Report by Eddie Lyons.

8.2. Appendix B – *Revelation field testing at Whitby Cliff 3:* Report by Tom Cromwell, Sarah Cross, Paul Cripps.

8.3 Appendix C - Addendum to Hand-Held Drawing Assessment. Report by Paul Cripps and Tom Cromwell, April 2003

9. *Figures & Illustrations*

Revelation Project 3591

Appendix 1c

Report on Current Field Practice

version 2.3

Vicky Crosby
November 2003

Contents

Appendices

1. Introduction

1.1 Investigation of Current Field Practice

This report presents the results of the participant observations of the excavations at Barrow Clump, Figheldean, carried out as part of the Badger Damaged Barrows Project (CfA Project 984) in September and October 2003. The two main aspects were observing the flow of data during site recording and implementing a pilot digital indexing system.

Data-flow and entity-relationship models for the CfA recording system were prepared by Paul Cripps, based on work on the paper-based records and the discussions of these held at the second Revelation away day workshop. The data flow model represents the movement of data through the CfA recording system during the on-site stage of a project. The fieldwork compared this model with actual data flow during the excavation. The resulting revision of the model is also included (Appendix 3).

A simple digital indexing system was created and its usefulness evaluated, particularly in the area of finds processing.

Additional information used includes an informal account of site recording and process at the Whitby Cliff 3 excavations (Autumn 2002). This is presented as Appendix 4, written by the senior site supervisor, Liz Muldowney.

The report is presented as a discussion of the observations.

1.2 Aims

The aims are:

Aim 1: to observe the recording system in use and compare the observed procedures with the CfA Recording Manual

Aim 2: to compare the data flow model with actual data flow in the field

Aim 3: to identify any areas where data flow could be improved

Aim 4: to describe how current practice aids on-site interpretation

Aim 5: to pilot a digital indexing system, and comment on its use, especially relating to finds recording

Aim 6: to present a revised data flow model

1.3 Badger-Damaged Barrows

The Badger-Damaged Round Barrows Project is addressing a range of management and research issues relating to severely badger-damaged round barrows in important prehistoric landscapes in Wiltshire. In particular, it will assess the nature and extent of badger damage to barrows, with widespread application to the management of archaeological sites. It will recover archaeological and palaeoenvironmental data as a mitigation exercise, in light of the continuing damage to the earthwork by badgers. Additionally, it will re-assess and/or enhance understanding of previous (antiquarian) investigations of Wiltshire barrows. The Project Manager is Jonathan Last

The barrow examined in 2003 was Figheldean 25 (Barrow Clump). The intention was that the southern half of the barrow mound would be fully excavated and any features cut into the natural chalk beneath investigated. However, due to the complexity of the archaeology and some logistical issues, only two areas of the mound were excavated.

Figheldean 25 (Barrow Clump) is a spread round barrow of bowl barrow form located on the Salisbury Plain Training Area, 6 km north-east of Stonehenge (SU 1655 4690). It is the only surviving earthwork among a group of 20 barrows recorded on aerial photographs and mapped by the National Mapping Programme (NMP) (*Fig. 1*). A third trench was excavated across the ditch of one of these levelled sites.

Burials were found in the barrow during the1890's, when excavations by Hawley uncovered a primary inhumation, accompanied by an Early Bronze Age pottery vessel, and four secondary inhumations. It was not recorded whether these burials were removed or left *in situ*, and at the start of the excavations it was not known whether there were further undisturbed burials to account for the human bone recently found in the spoil around the badger sett entrances.

The work was complicated by disturbance caused by past military activities, but useful information about badger activity and the sequence and construction of the mound was recovered. The find in 1935 of a Saxon spearhead in a rabbit scrape had implied that there might be later intrusive burials, and during the excavations three early Anglo-Saxon graves, one containing an adult and child, were found. These contained grave goods, all having a number of objects of mid sixth century date. A fourth excavated grave contained no objects and may be different in date. Two additional graves were identified but not excavated.

The fieldwork took place over six weeks, with an excavation team of up to twelve.

There is no such thing as a 'typical' CfA project, and in some ways the project may not be fully representative of recent CfA fieldwork. In particular, the project manager and site supervisor had not previously worked on CfA excavations and while methods followed the CfA Recording Manual, they may have differed in some ways from established 'custom and practice'. For example, it was intended that all finds would be three-dimensionally recorded for analysis of the effects of badger activity on their distribution. (In the event, practical considerations during excavation led to the modification of this strategy, and objects were recovered as bulk finds from many contexts.)

Some of the recording forms used on Romano-British and later excavations were not required on this site (for example, the Built Structure and the Timber Structure Record Forms). Because of problems with the equipment, on-site wet processing of soil samples did not take place.

2. Methods

2.1 Observation

The main method of observation was participating in the excavation as a site assistant (it is relevant to note that this was also the first time I had worked on a CfA site). Additionally, I informally discussed issues with and asked questions of other team members, and had longer more structured discussions with several members of the team in which specific aspects of the data-flow model were considered.

A detailed table describing data flows (Appendix 2) was prepared from this information.

2.2 Digital Indexing

A pilot digital indexing system was tried out on site. The indexes for Contexts, Small Finds, Samples and Skeletons, and the records for Drawings and Photographs formed the basis for this. The indexes are held in an Microsoft Access 97 database, named Badger. During the fieldwork, this was held on a laptop, and regular security copies were made to recordable compact discs (CD-Rs).

Preparation and creation of some of the basic tables took place before the fieldwork started. The Project Finds, Environmental and Surveying Officers provided advice. However, much of the database development took place during the excavation, in particular development of the forms, reports and standard queries.

The data from the Indexes and Context sheets were entered in the evenings. Initially, some additional information from completed Small Finds sheets (see 2.3) was added at the same time, but during the later part of the fieldwork, it was added directly by the Finds/Environmental Supervisor Tracey Clark (see below).

The usefulness of the digital indexing, especially relating to finds recording, was assessed by discussions with the Finds/Environmental supervisor and the Project Manager.

2.3 Description of the Digital Indexes

The database has 12 tables covering the Context, Small Finds, Samples, Drawings, Skeletons and Photography records, and the Small Finds and Samples co-ordinates. There are additional tables for the daily Total Station

(TST) readings, and nine glossaries. There is also a table of record numbers used and record type.

The database opens to a Menu from which forms and selected frequently used queries and reports can be accessed. There are data display forms for the Context, Small Finds, Samples, Drawings and Photography records, and separate data entry forms for Contexts and Small Finds. It is intended to ensure the database is easily used by those with little knowledge of Access, and a User Guide will be produced once development of the application is completed.

Although the original intention was to digitise only the indexes, following advice from the Project Finds Officer, Sarah Jennings, a few additional fields from the Small Finds records were added. This had two roles: increasing the reliability of the data, as the data entered had by then been checked by the Finds and Environmental Supervisor, and making a fuller level of information available to the Project Finds Officer during the immediate post-excavation period for initial work and completion of the archive summary. The Context Type, Simple Name and Fill Of fields from the Context Records were also included in the Context table, for similar reasons (the nature and use of the Description field on the Context Index is discussed in 3.3 below).

The Project Environmental Officer, Gill Campbell, supplied a list of the fields to be digitised from the Sample records, the glossary terms needed, and answered questions about the form of the data.

Project Surveying Officer Tom Cromwell ensured the TST data were easily converted to a suitable format for import in CSV files. These were imported into the database, and currently remain held as separate tables. Selected fields were then appended to a table holding the co-ordinates for Small Finds and Samples, allowing the information to be linked to the relevant records.

Documentation for the Digital Indexes is in Appendix 5.

2.4 Other Digital Data

As part of the general digital data management required, the digital photographs were also downloaded to the laptop, and additionally backed up onto CD-Rs. This enabled a CD of photographs from the Saxon burial with the grave goods *in situ* to be sent to Fort Cumberland for the Project Finds Officer. It was also possible to create a slideshow of photographs to show visitors to the site.

The data from the Total Station (TST) used for surveying were downloaded to the laptop, usually daily, and imported into the database. Regular security copies were made.

3. Observations

3.1 Data Flow in the field: general comments

Barrow Clump is situated in the Salisbury Plain Training Area. This area has, in addition to the normal network of roads and lanes, a separate system of tank roads and track ways which at times runs parallel to or intersects the road network. Traffic can therefore move in unexpected ways and reappear in unexpected places. This struck me as a good model for the data flows on site. While the data flow model envisages flows between the formal elements of the recording system (especially the forms), on site much of the data moves in a less formal, less structured way. The present recording system does not appear to support this well, and the work of the Revelation project offers the opportunity to consider how this could be improved.

On site, two data flow networks were seen in operation. One data flow relates to the records to be used in post-excavation work - the formal 'end product' of the fieldwork stage and the way in which data is passed on. A different (if overlapping) and less formal data flow process supplied the information excavators needed to do their job of producing that 'end product'.

In fact, three distinct if overlapping areas of data acquisition could be defined, relating to:

- information needed by the field staff as they excavate and record the site (a sub-set of this is the data needed by finds/environmental staff processing on-site)
- information needed for on-site interpretation as the work progresses
- the site records required for post-excavation work – the primary 'product' of the fieldwork stage

These are three processes, with three separate sets of information needs. Recognition of these separate networks and their needs could make things work more efficiently, not least if the indexes used on site were improved so that information could be recovered more readily.

Meeting the needs of on-site data collection and interpretation may lead to duplication of data - one thing which surprised me was that some staff stressed the value of making context sheet sketches which were simply copies of the site drawings. Their usefulness in the field was felt to justify the time spent. This perhaps reflects difficulty in accessing or using the drawings. (From a post-excavation point of view, I would regard the value of a context sheet sketch as explaining or recording things which were not shown clearly in the drawings, such as interpretation, location of finds, or relationship to other contexts.)

The Assessment of Handheld Drawing Systems Report raised the possibility of separate hardware platforms for digital drawing and database entry on site. The Barrow Clump observations support the conclusion that it is crucial that completed digital drawings can be viewed in combination with the database, and not solely through the drawing creation platform (*An Assessment of*

Handheld Drawing Systems, Section 5, which forms Appendix 1B of the CfA Report *Revelation: Phase 1 Assessment*.).

"How do you find information you need?" (for example, which layer was above the feature you are excavating) was a key question used to find out about data flows. The responses demonstrated the marginalisation of the main record forms in the excavation process itself.

The on-site data flows relied heavily on memory and/or asking others for information, and several people commented that errors or 'fluffiness' could result from this. Labels, tags and personal notebooks (or odd bits of paper) were also important. The context and small find indexes (whose main function had been seen as controlling number allocation and acting as a check on what full records should be present) were found to have a more significant role than envisaged.

However, formal records were often not as useful as they could be. The quality of information was often poor, especially on the Context Index Form. Some site staff with finds/environmental or post-excavation experience commented that they did try to fill this out as fully as possible for this reason. Until the finished Deposit/Cut forms are placed in the completed records file, the Context Index Form is the only generally available source of information, and restructuring this form to capture better and more specific data should be considered.

Suggested ways of finding information (or how people accessed information elsewhere) included:

- looking through plans, where they were accessible and filed in a structured way so those for an area were readily located (a plan matrix on each drawing was said to aid this greatly)
- better quality data on indexes
- accessible working matrices
- building up sketches of areas and features
- personal notebooks

Improving access to information was a clear issue we should consider further. The question of finds staff access to context information is considered below, but the converse was also true. Excavators wanting to check the Small Finds Index for cross-referencing were hampered by the fact that finished sheets were removed to the finds office (some did visit the office to check records in their lunch breaks). With hindsight, it would have been easy for me to printout the digitised indexes each day, perhaps sorted by Context number, but this problem was not mentioned to me until almost my last day on site. Better organisation of the plans would have been possible.

This relates in part to the logistics of the Barrow Clump site, where the offices were about a mile from the site, and equipment transported daily in a very full Land Rover. But the situation could have been improved by clearer recognition that records are also for continuing use on site.

One data concern related to the use of the level (see 12 to 23 data transfer in the Appendix 2). There were two reasons for this:

- the lack of a level book, and consequent potential for loss of data.
- the lack of integration (on-site) between data collected using the Total Station (TST) and data recorded using the plans and level.

In the event, no levels data were lost, and the levels will be converted to be compatible with the TST data during post-excavation.

I have since discussed this with Tom Cromwell, who has told me that each site should have a level book, which ultimately forms part of the paper archive. However, this was not the case here, and nor was it the case on several other CfA sites I have since asked about. Tom's comments illustrate an interesting point about our current practice:

> "Usually there's an unspoken division of responsibilities, where I'll be asked to sort out the TST because it's "technical", while the Project Directors retain control of the dumpy level since it's a "traditional" field tool that they are comfortable with."

This kind of "unspoken division" is perhaps precisely the sort of thing that Revelation must challenge.

It is interesting that even with a TST on site, levels were almost always taken with the level. At one point, the fact that the TST was set up over the station normally used as the backsight for levelling was seen to be preventing levels from being taken. This presumably reflects the fact that most site staff are not familiar with using the TST (although the Site Supervisor went through the set up procedure and use of the TST with all site staff in turn). Taking the levels with the TST would have been as efficient in terms of transferring levels to drawings and provided an additional digital record of the information. There was a TST book, which kept a thorough record of the surveying done, in line with CfA procedures.

Co-ordinates for Contexts were recorded to the 5m grid square. With the TST on site, it might in future be worth recording a central point on each Context. This could be linked with the digital index (in the same way as for the Small Finds), and providing this information as printouts might be one way of enhancing data availability on site.

3.2 The Recording Manual and related issues

Aim 1 was to observe the recording system in use and compare the observed procedures with the CfA Recording Manual

All staff received copies of the Project Design and Recording Manual, but as far as I am aware there was not a Site Master File on site, listing the site-specific detailed recording decisions which were not in the PD (e.g. the decision that context co-ordinates should be recorded to the 5m grid square level). I doubt this is unique to this project - reactions to asking colleagues about Site Master File produced reactions varying from “The what?” to “It’s all in the PD”.

As all site staff now receive the Recording Manual, a good option might be to insert several site-specific pages including these recording decisions and the Record Number Allocation Form at the start or end of the Manual. This would fit happily into its modular format.

There were several areas where the Recording Manual (including the paper recording forms) did not document or support procedures adequately, and some specific examples follow.

3.2.1 Cross-referencing

The Context Form requires cross-referencing of other records, such as Small Finds, Photographs, Samples, and Drawings. The data flow issues are detailed in Appendix 2. The value of this process was recognised especially by those site assistants who had post-excavation or finds/environmental experience, and this may be reflected in their records.

In some cases, noting the information does provide an independent check on the data held in the other records (particularly for samples). But in other cases, the information was in fact copied from those records or indexes. If this is acknowledged procedure, then the digital indexes could be used to aid checking and completion of paper context forms (although if the fully cross-referenced paper records are needed primarily for initial post-excavation work, the digital indexes might make this unnecessary). Similarly, if CfA had a database which allowed rapid data entry during or immediately after fieldwork, the case for cross-referencing on paper might be removed. On the other hand, if the aim is an independent check on these relationships, then ways of ensuring this happens should be considered. (The question is of course superseded if CfA moves to full on site digital data capture, but in the meantime it should be considered.)

3.2.2 Soil description

Several people were distinctly proud of the fact they never record Munsell numbers. If and/or when this should be done seems to be an issue which should be decided for each site, and noted in the Site Master File. If Munsells are useful, recording them should be enforced, and if they’re not, time shouldn’t be wasted on them. Since one use for them is comparing contexts, consistency is needed.

Gill Campbell has commented that it would be useful to review the use of soil descriptions, and organise training in this aspect of field recording.

3.2.3 Drawing and Photograph records

Both Photograph and Drawing Records raise some problems. With the Photograph Records, the main problem is that the same descriptive information needs to be written on three separate forms (colour, B&W and digital), which both wastes time and causes entirely predictable errors (I've corrected the records before entering them) where the number sequences fail to match (usually due to additional shots being taken on one camera only, but sometimes because someone fails to write on all three forms). I have since seen a Digital Photograph Record Form, but there were none on site and my copy of the Recording Manual lacked it (the general Photograph Record Form was used instead). One form (showing the Photograph Number and the descriptive information) which gave the film numbers for Colour and B&W and the date and image number for digital would be much better.

The use of drawing sheet numbers was misunderstood on site for the first few days leading to errors. Where drawing is normally done by 5m grid square, a sheet number in the form grid square/sheet number would be easier to understand, and would make identification of relevant drawings quicker. This could assist on site use of information held in the drawings considerably. A Drawing Sheet Index (currently there is none, again resulting in errors) could then be in grid form, with 5m squares down the page and sheet numbers across.

3.2.4 Recording burials

The burials showed up some marked inadequacies in the Recording Manual. There is currently no section on Skeleton recording, and no crib sheet. [Two people asked me where they could find these…] Site staff felt the Skeleton form itself was clear and easy to use, though a quick survey of the five completed forms suggests that there were inconsistencies in how they were filled in, for example, how articulated and disarticulated bones were indicated on the diagram. Staff were unsure whether skeleton numbers form a separate block in the number sequence (they do, and this was defined in the Record Number Allocation Form, which should also have been in the Site Master File). The problem with the Skeleton Form (which has a Context field but no Skeleton Number field) caused confusion as to whether the context field on the sheet should hold the skeleton number or the context it was within (and therefore where the other piece of information should go). There is not a Skeleton Index; an *ad hoc* version was created on site to control number allocation, but did not for example include information on the cut or fill containing the burial.

Subsequent discussion with Gill Campbell revealed that the section of the Recording Manual dealing with skeletons and present in earlier drafts is missing in the current draft. This may be because it was in need of updating.

3.3 Comparison with the data flow models

Aim 2: to compare the data flow model with actual data flow in the field.

Data flow and entity-relationship models for the CfA excavation recording system were produced following a workshop involving most of the Revelation project team. The data flow and entity-relationship models created in the workshop are presented as Appendix 1 of this report. Appendix 4 of the CfA Report (*Revelation: Phase 1 Assessment*, December 2003) describes the background to the modelling and the methods used.

This section aims to identify and comment on some key points about the data flow diagram and about on site data collection. The points raised are believed to have some general or wider significance in terms of CfA site recording procedures. Appendix 2 presents the detailed observations made on the data flows at Barrow Clump. By describing specific data flows, it adds processes to the data stores identified in the data flow model.

The initial data flow model reflected the composition of the group which did the groundwork for the model, and some aspects of the recording system were not well represented, with some key relationships being missed, for example between survey data and the small finds record. The absence of site supervisor-level staff from the meeting (which reflected the structure of permanent CfA staff) also meant that a level of expertise and understanding was missed, and probably led to a bias towards the post-excavation use of completed records rather than the way in which the record itself is formed. As a result, the model was perhaps a rather idealised version of how data 'ought' to move between forms.

The model misses one major type of data store – labels (including small finds bags, finds bar code labels, sample bucket labels, marked grid points, context labels pinned onto sections, temporary markers for small finds, and metal detecting tags). These form a key part of the on site data flow network.

Data flows were often more complex than modelled. A flow shown as one-way on the model was frequently two-way (for example, the Cut/Deposit record was shown as supplying information to the Working Matrix, but in practice relationships might be recorded on the matrix and the information later transferred to the Cut/Deposit record). Some staff used the Working Matrix form in this way, and then discarded it (on other sites, where required to put the form into the records, a neat copy is often made). Working Matrices drawn up for areas of the site by the supervisor were, of course, retained. The drawings were an important source of data for the Cut/Deposit record and the Working Matrix, but neither relationship was in the model.

Perhaps more importantly, data was not always used in the way envisaged. While the main role of the Context Index is regulating the issue of context numbers, it is surprisingly important in supplying data to the finds process. Current practice is that the Description on the Context Index becomes the Siting Description on the Small Finds form (a relationship not shown in the model). This reflects the availability of data (access to the Context Index is easier for finds staff than access to the full context records). However, I think this is of concern because the Description on the Context Index is often poor quality information – an initial brief comment before the context is investigated. Where specific siting information was noted on the Abbreviated Small Finds Form (as it was for the Saxon burials), it did not go into the Siting Description on the Small Find Form, but was noted under Comments.

The information transferred may be all that is readily available and prove useful during the field work, for example for noting where finds are alleged to come from a cut, or identifying finds from different fills of the same feature. However, if Siting Description is only to hold Context information, during post-excavation work information directly taken from the Deposit/Cut form would be more reliable. On site, improving the Context Index would enhance the quality of the data passed on to the finds recording process. Additionally, a specific Siting Information field on the Abbreviated Small Finds form might encourage this information to be noted where it was useful.

3.4 Areas for improvements of data flow

The distinct nature of the site data flows should be recognised and supported by our systems. There are different data flow structures and different needs at different project stages.

Some specific suggestions for improving data flow can be made. These fall into three groups:

- Improving the way the recording system operates
- Improving access to information
- Improving the quality of the ‘transitional’ information in the indexes

Improving the way the recording system operates requires the completion of the current Recording Manual, including the modification of some forms to improve ease and accuracy of data collection. Some of these issues were discussed in 3.2 above.

Improving access to information on site requires both physically facilitating access, for example by storage drawings in an ordered form, and considering how the use of records can be encouraged.

One way of achieving that could be by recognising the use of the information in the indexes and improving its quality. Information in the context index, for example, tends to be regarded as transitional in the longer term, but improving its quality would increase its value during the recording process.

This applies to both digital and paper indexes. The use of the digital indexes did greatly improve data flows and save time in the finds recording – but I did not use their potential to assist with the context recording. For instance, I should also have put printouts on site for cross-checking in the recording process.

In discussions of on-site digital recording systems, it is sometimes said that indexes will be redundant in a fully digitised system. At one level this is true, but if we see indexes as interim statements before the full records are completed, the picture is different. Their present hidden or unofficial role as a major source of on-site data transfer needs to be recognised, and they could be improved and tailored to meeting these needs and increasing data accessibility. Similar arguments apply whether paper or digital records are involved.

As a general point, if we are making changes to the Recording Manual and procedures to improve site data flows, we should ensure that the people who are actively extracting data from site staff are involved, and include site supervisors and site finds/environmental supervisors in the process.

3.5 Supporting Field Interpretation

One aim of the report was to describe ‘the current use of recording and data flow in the process of interpretation in the field’. This is not something which is easily directly observed. Points already raised relating to the accessibility of information and the value the indexes have on site are again relevant here. Improving the quality of the indexes would enhance their role in interpretation.

he Project Manager felt he had not really used the records for field interpretation. This reflected the small size of the site and team – interpretation was based on discussions, memory, and observation – and also pressure of time. He felt he would have used records more if the project were larger, especially if it was necessary to modify/develop the excavation strategy as work progressed. The printouts of small finds by context (with context information as available) were useful early on, but during the later stages of the fieldwork the pressure of work on site prevented him making use of this.

The Site Supervisor compiled working matrices in discussion with site assistants working in each area. He regarded the site drawings, when filed in order and given plan matrices, as another important source of information. He was strongly in favour of copying the drawings onto the Context sheets to increase their usefulness in this respect. Otherwise he relied largely on memory and discussion with staff for interpretation. He used the Context Index to monitor progress (to check that Deposit/Cut forms were being filled out for all Contexts listed).

One site assistant mentioned looking at other people’s context records to help understand his area/features. This was, however, unusual, and referred at least in part to his experience on other sites. It may be significant that this was someone with supervisory/initial post-ex experience.

3.6 The Digital Indexes and Finds recording

The trial implementation of digital indexes seemed to work well, and saved the Finds Supervisor a great deal of time by avoiding the need to copy data already held digitally (context descriptions, co-ordinates) onto paper recording forms. Regular printouts of this information were provided for filing with the paper finds records, to build confidence in this approach.

Printed and ordered listings also made tasks such as writing co-ordinates on small finds bags easier (although this was discontinued during the course of the project, once it was clear the co-ordinate information was readily available and therefore unnecessary on the bags).

A report based on the Context Finds Register was created. This avoided the need to copy the Description from the Context Index into the Simple Name field on the Context Finds Register, and also gave the total number of Small Finds for each context. Bulk finds information was not digitised at all, as the original Project Design specified that all finds would be treated as Small Finds. The report included a Bulk Finds column, which was filled out in the same way as normal – it would be easy to add this information to the digital indexes if required. [This is a second case where the Context Index Description is used in the finds records in a way which the entity-relationship model does not represent.]

The Finds Supervisor quickly learnt to use the database confidently and add information to the digital Small Finds records, despite little or no previous Microsoft Access 97 experience. This took very little additional time while she was filling out the full Small Find Form. A separate data entry form was created to make this process as smooth as possible. The only problem encountered was that Access cannot support continuous forms with a sub-form (needed as objects can have more than one Material). The Finds Supervisor felt that single forms were less easy to use for data entry than continuous forms (this may be individual preference rather than a general issue).

One issue raised was a difference in approach to data structure of the relationships between Small Finds, Samples and Contexts. The Project Finds Officer requires the key relationship shown to be that between Small Find and Context, while the logic of the data structure is that, for Small Finds from Samples, the key relationship is Small Find to Sample. The Small Find is therefore recorded as 'within' a record which can be either a Sample or a Context, avoiding holding the relationship between Sample and Context in multiple locations in the database. This approach was taken in the digital indexes, but the Small Finds data can readily be presented in reports in the format required, masking the actual data structure. At present this requires two-stage querying, but it should be possible to automate this using a macro.

One practical problem encountered was that the printer would not work satisfactorily in a portacabin which suffered badly from condensation. Actually producing the printouts was a time-consuming and frustrating process, as the printer refused to print more than two or three pages in a row. The quality of the printing was poor, and a more portable printer which can easily be removed from site overnight should be regarded as a priority. The paper also needs removing overnight, or storing in an airtight container. The printed sheets were also very unsatisfactory if they got wet, when the ink smudged and ran badly.

One issue remaining to be resolved is how these digital indexes are to be integrated into CfA data entry procedures.

In retrospect, the indexes could also have been used to assist in information retrieval by the other site staff, particularly for cross-referencing Small Find numbers and other records to the Deposit/Cut forms.

Although concern has been expressed within the Revelation team about the problems of ‘hybrid’ recording systems (ie partly manual and partly digital), this digital indexing approach does seem to offer considerable potential for increasing efficiency and availability of data in the short term.

It had been hoped to assess the usefulness of the digital indexes in initial post-excavation work (the stage before the full records have been entered and made available digitally). However, owing to the relative project timetables, this has not yet been possible.

3.7 Revised model

The revised data flow model forms Appendix 3 of this report.

4. Conclusions

The participant observer work at Barrow Clump has enabled the data flow model of CfA recording systems to be compared with practice on site. Both general and very specific observations and comments are presented above and in Appendix 2, and Paul Cripps has revised the data flow model to reflect the results. A full account of the data modelling is presented in his report, *Data Modelling in an Archaeological Context*, which forms Appendix 4 of the CfA Report *Revelation: Phase 1 Assessment*.

The report’s conclusions are a combination of straightforward observation and interpretation by the participant observer. The comments of CfA colleagues are awaited with considerable interest.

Some unexpected and expected observations can be picked out.

Unexpected observations:

- The initial data flow model represented a post-excavation perspective, and hence an ‘idealised’ model of data flows.
- Informal data flows outside the paper recording system are very important in on-site information retrieval and use. (I am not entirely sure why I found this unexpected, because with hindsight, it seems obvious. This may reflect the structure of the Badger-Damaged Barrows team – perhaps it is more usual to work in established teams where the necessary informal information networks are taken for granted.)
- Indexes have a more important role in the on-site use of data than their formal purpose of controlling number allocation suggests.

- Recognising this and enhancing the indexes, especially the context index, could improve access to data on site and assist in completing the context records.
- There is some deliberate duplication of information on site, to aid on-site interpretation and use of data.
- At present, interim context information from indexes is transferred onto the small find records in a way not shown in the data flow models: improving the context indexes would also improve the information available to finds staff.

Expected observations:

- It is important that site staff understand the later stages in the archaeological process. (This can be as simple as knowing that the pencil type used is important because drawings will be scanned.)
- The need to produce a fully updated version of the Recording Manual has been recognised for some time.
- The digital indexes proved to be useful and timesaving in finds processing.
- Importing TST data into the digital indexes database and thus linking co-ordinates to the small finds data was straightforward and saved considerable time.

In this report, the suggestions for improvements in data flow are still made within the framework of an initial paper recording system. The approach taken to digital indexing in the pilot study does appear to offer considerable potential for increasing efficiency and the availability of data in the field and during initial post-excavation work. The indexes were successful in their aim of assisting the finds recording process and providing information to the Finds Officer.

However, during the excavations use of the indexes was largely restricted to the finds work. More explicit recognition of the role of indexes on site and their potential to improve the quality of the data collected should be considered. Enhancing the context index forms to improve the data collected is one option. Once digitised, the indexes would assist in subsequent stages of the site recording process (for example by allowing site staff to identify related contexts and to cross-reference contexts with other records). This should seen by site staff to be offering them an advantage, and encourage fuller data collection and amendment.

Clearly a fully digital site recording system offers the potential to improve both access to data during the excavation period and data quality (by building in checks, glossaries, targeted help, etc). But it should not be assumed that direct digital recording in itself will miraculously solve all the issues. A thorough understanding of CfA archaeological data structure, data flows and real life recording processes is needed to inform the work, and it is hoped this report has made a start towards achieving this.

5. Acknowledgements

Thanks to:

- Colleagues in the Revelation and Badger-Damaged Barrows project teams, especially Brian Attewell, Gill Campbell, Tom Cromwell, Sarah Cross, Sarah Jennings, Claire Jones and of course Jonathan Last for allowing his excavation to be dissected in this way
- Paul Cripps for the Data Flow Models
- Liz Muldowney for much useful discussion on CfA field practice and her informal report on Whitby Cliff 3
- The Badger Damaged Barrows site staff: Site Supervisor Duncan Stirk, Finds/environmental Supervisor Tracey Clark, Site Assistants (Stella Bickelman, Liz Chambers, Richard Gregson, Ben Moore, Gareth Watkins, Kim Watkins) and Trainees (Adrian Smith, Nicola Fraser)
- All those CfA colleagues who have answered questions regarding practice on other sites

Appendix 1d

Sectoral Practice Report

English Heritage internal report for the Revelation project

Version 1.0

Sarah Cross & Keith May

October 2003

Contents

1. Introduction

1.1 Purpose

When first considering the project we had a sense from many people, particularly those using IT, that the main issues relating to systems development for archaeology had been solved. Most archaeological units now use IT. Most organisations of any size have systems that they have been running for some years.

So perhaps it was just a matter of defining our needs and then locating the most appropriate, or the latest system that fits our requirements? There was a nagging sense that this would prove less likely in practice, but we certainly were keen to not have to re-invent systems or cover territory which had already been well explored and mapped effectively.

This review has been an opportunity to discover what other archaeologists are already doing that might be of relevance to developing new CfA systems. It demonstrates the areas where we can rely on other people's experience and those where we will have to undertake substantial development ourselves. It also maps some of the elephant traps that our colleagues have been kind enough to point out for us.

1.2 Methods

Attempting to summarise the development and state of the use of IS in archaeological fieldwork practice in 20 pages or less is a foolhardy proposition. This report rests on its three appendices, which in turn rest on a larger set of 'raw' data. The summaries of fact finding visits (Appendix A) to different organisations condense the notes of all the people involved into a single report for each visit. The literature review (Appendix B) brings together our thoughts on 90 different papers under 11 headings. The list of relevant software records the publicly and commercially available software from the visits, the literature review and comprehensive web searches (Appendix C). This report in turn summarises these elements into a brief, readable, assessment of the use of IS in the archaeological fieldwork process (excavation-analysis-dissemination), and provides a context for the other aspects of the Revelation project.

2. Highlights from the Summaries of the Fact Finding Visits

This section will present some of the key findings that have emerged from meetings and fact-finding visits organised with a range of archaeological organisations or individuals in England who are widely acknowledged to have experience with well tried and tested IT systems or are using cutting-edge IT. The primary intention was to gather information, ideas, and examples of both pitfalls and opportunities that would be of most relevance to CfA's future needs for the design of a new IT system. For practical purposes only certain individuals could be assembled as representatives of organisations and therefore it was agreed that for the purposes of reporting no individuals would be named directly *and that the organisations involved would not be directly identified*.

Those specifically interviewed could be roughly characterised into three groups:

i) commercial field units - MoLAS, YAT, TWA, OAU, Framework Archaeology,
ii) research organisations/individuals - BUFAU, LRC, G. Avern
iii) end users of archaeological information - ADS, OAD, NAPEX

This said, most of the discussions centred upon issues that spanned across all three of these groups, which is why their responses have not been presented separately. However there were some emphases made by the groups which were different according to their differing funding sources and end needs. It may also be possible to detect some differences in emphasis over the relative importance of fieldwork – analysis/research – dissemination/archive

2.1 Trends across different organisations

2.1.1 Systems and intellectual processes

It was widely felt that there is unlikely to be a "one system that suits all circumstances" solution. Different recording strategies can be adopted by different organisations for different types of fieldwork, analysis, and publication under very varying conditions.

However although it seems likely that different organisations will continue to want to develop their own internal systems, there was a sense that units with more unified data systems produced better results in a more efficient manner than units that didn't have a unified approach. This can be attributed to several factors, including:

- clearer understanding by staff in all areas of what to record and why
- easier data transfer both within and between organisations with less data loss or corruption
- less wasted effort from staff "reinventing the wheel"
- higher quality and consistency of end products

Most commercial organisations felt their systems "did the job required", but many felt that better integrated systems could do more, if time and resources allowed. Most archaeologists primarily want to be doing archaeology not worrying about how their computer might record information, but there was wide spread interest and concern amongst all groups of archaeologists about how the archaeological process might be improved/re-designed.

Many practitioners see the context sheet as a prompt for contemplation and interpretation rather than a data collection device, but often systems are still based on an attempt to provide preservation by record. These can be viewed as two ends of an ideological spectrum but in practice most archaeologists seem to lie somewhere between the two.

2.1.2 User training, CPD & change management

There is a major training issue with implementing new systems or in getting new and innovative technologies adopted. If the system is too complex then the users can be stranded when it falls over and are reliant on technically competent people.

Digital recording can be threatening to people and there is resistance to change, often even more so from senior people because they have to acknowledge their own inexperience. In general it was felt the situation would improve over the next ten years with increasing computer literacy but there would always remain a group of users who were wary of technology.

It was considered very important to get "user buy-in" to the design of any system and in particular the user interface and important to find ways that encouraged *all* staff to use the system.

Getting better data will depend on getting people on site to record the best quality data they can. This is a training issue not simply a matter of designing an adequate system. There was a perceived problem that the best and most able people "move on" from fieldwork because of a lack of CPD.

To improve the level of intelligent data collection, we need to bring specialists and training onto sites and generate more skills in the field. This will help reduce problems such as poor site sampling and let staff engage with the research questions, by establishing a dialogue between people and process.

2.1.3 Trends in fieldwork practice

None of the commercial units interviewed are directly recording primary context information digitally in the field, although a number of research led organisations have developed systems for doing this on specific research projects and one organisation has been using handheld recording devices in the field for over fifteen years.

Fieldwork applications of IT are almost exclusively dedicated to recording of spatial data using TST or GPS equipment. Units tend to establish "specialist" teams (Geomatics) who carry out a range of digital recording services. The output from these geomatics teams are often used in post-excavation analysis or further re-processed as graphics for publication requirements, but the information is less often actually available for use while excavation is still in progress.

Most commercial units (and all those interviewed) record information in the field on pro-forma paper context sheets. In most cases the paper record is then transferred on to a computer. Transferring the paper records to the computer usually uses one of the approaches listed below although sometimes a combination of the following occurs:

- entry of records into databases while on site
- entry of records into databases during post-excavation processing
- scanning of paper context sheets as security copies
- scanning of paper copies for use as part of a digital record indexed by a database

In many cases there is considerable double-handling of information between paper and digital media, particularly where digital systems have been implemented in a very piecemeal fashion.

Some organisations do use GIS in the field on laptops but usually only on large sites, because of the capital expenditure involved and the likelihood of the need to write-off some of the equipment. The degree of use varies from project to project and can depend a lot upon the IT literacy of the project director or manager.

More work had been done by research organisations on developing methods for capturing digital data directly in field. Some of the latest R&D work was beginning to look at the applications of wearables (electronic devices in clothing) and wireless networking for fieldwork. This however tended to be seen as exciting cutting edge work, but with not enough practical application feeding through to “mainstream fieldwork”. This was partly attributed to a lot of R&D systems being funded on a research project basis and therefore the systems developed may not continue beyond the life of each academic project.

A concern was raised that the “Blue skies” research could be seen as irrelevant to a large part of the development control sector unless the systems and technology could be applied to day-to-day fieldwork. It was felt that improved efficiency (costs) of digital systems was considered an important factor in their adoption by commercial operations.

One organisation has been recording excavation data directly on handheld devices for well over a decade. Nevertheless even in this case the organisation does not record all digital data directly in the field, as site drawings are hand drawn first and then digitised and conventional photography is still required despite R&D with annotated rectified digital photos.

2.1.4 Data use & analysis

There were few mechanisms adopted for controlling what sorts of digital information specialists returned. Some felt the definition of being “a specialist” is the particular knowledge, which separates you from the rest and therefore requires it’s own ways of working.

One organisation sent material to specialists that included digital records of all object data and context data. Initially specialists felt this was too much data but they make more use of the more of other peoples data to inform their analysis in the end.

It was often found difficult to integrate results from specialists because of difficulties in understanding the research questions that specialists were addressing.

2.1.5 Dissemination and archiving

There is growing recognition of the value of archaeological data for academic research and re-use and therefore acknowledgement of the importance of disseminating information in a variety of media and including digital archiving. But dissemination and archiving still take second place to "inevitable" practical project constraints such as lack of budget. Again there tends to be a project specific approach, whereby larger (well-funded) projects are able to have more provision for digital dissemination & archiving built in.

The point was made that data alone is not very useful to people. Rather than just presenting raw data, there is also a requirement to provide interpretation, and to make the level of interpretation more explicit (particularly with reconstructions and Virtual Reality).

There was greater emphasis by end users of the need to look first at archive deposition before designing recording systems, in order to ensure that any digital outputs and archive is of a suitable quality for long-term conservation (e.g. stable file formats, useful metadata, etc).

It was also argued by at least one end user that certain data have natural origin and expiration points within the whole excavation - analysis - publication cycle, and may not need to be carried from cradle to grave at all costs.

Most data creators and users recognised that there was considerable peer group prestige and academic kudos that people derive from "paper based" publications and some felt it was unlikely that people would totally abandon printed publications.

2.2 Problems and Pitfalls

Project specific approach to data

Problems occur when there is a lack of a unified system and a lack of overall IS strategy. But in Archaeology the use of digital systems tends to be on a project-by-project basis, with smaller projects usually not meriting, or being able to afford, much computer provision. Development cycles can be limited due to funding issues and innovation is often undermined by piecemeal implementation. Computer programme licenses and training are often a problem where the cost has to be covered by any one single project.

This is a recurring issue, as the return time for the benefits and consequences of a more integrated digital system can often be longer than the life of any one single project. This can make it difficult to 'sell' the advantages of an integrated system to management.

A purely project centric approach to data collection can lead to silos of information with very little integration of data at the analysis or dissemination stages. What is needed are systems that plan for integration at the point that projects are conceived.

Larger projects, with more funding, tend to be able to plan for a more ambitious dissemination strategy and this usually requires more planning of data capture and data management to achieve dissemination requirements.

Divisions by project stages
The divisions between fieldwork – analysis – dissemination are useful for project management purposes (e.g. MAP II) but do not necessarily reflect the needs for information flow and continuity within a large long-term project and an integrated systems design.

Geographic scope
Identifying the geographic scope of the organisational needs for data can greatly aid in producing coherent systems, rather than simply incorporating a series of projects as they come along.

Lack of training and CPD for archaeologists in relevant IT skills
Often the adoption of digital systems is very dependent on the enthusiasm of the individual running a site, and therefore relies on informal, as opposed to formal procedures.

There is limited training for introduction of new skills across archaeology. For IT in particular this perpetuates the idea that only "the anoraks" need be concerned with development of IT skills. There is a need to have "multi-taskers" as field archaeologists with skills in survey, databases, excavation skills, GIS, and basic interpersonal skills.

Change management
IT can be threatening to people and there is almost inevitably resistance to change, sometimes even from senior people because they have to be acknowledge their own inexperience. In general it was felt that computer literacy would improve this situation over the next ten years but there would always remain a group of users who were wary of technology.

Poor integration of Specialist data
Often difficulties were encountered when transferring data to and from outside clients and partners due to different software formats. The issues will be as much about data-flow between users as the design of an IT system to hold the information.

Audit Trails

Many using and changing digital data had not really considered this as an issue at all. One organisation with a lot of digital data said that they did have an audit trail but that it had got too big to be useful.

Digital archiving
There was a lack of clear specifications for digital archiving within units. Projects were often completed to necessity and not to standard, creating problems with archive deposition.

IT based solutions that were not User focused
A very common complaint was that there was too much focus on systems not users. Many felt recording systems should be driven by archaeologists not programmers, but recognised that archaeologists often don’t understand enough about the IT to achieve the best solutions.

Some felt that computer based records tend to restrict thought and are thus less likely to carry added-value interpretation beyond the minimum asked for from pick-lists or check boxes. Any system should also provide the ability to include free-text descriptions along with database driven pro-forma style data entry.

Can you trust the diggers with that?
There may be a culture in some field archaeology units that people don't look after equipment. Again this is as much an issue for training and CPD for individuals as it is about the choice of ruggedized computers that you can write off.   It does highlight the need to accept a degree of right-off investment in the technology which perhaps only “major units” were really serious at costing into projects.

2.3 Opportunities

Field system feedback
A well designed system that delivers information back to the excavator can increase the engagement of the staff much more in the investigate

Building Research into the process
Recording is inextricably linked to research objectives that drive the project. Working with this as a positive factor gives focus to all stages of the project.

Integrated work practice
Server side databases and web-based delivery of information gives the chance to share data amongst all the project team members.

Better data management
There is the chance to develop a system that “It pulls together the whole archive in one digital “place”, allowing for sensible version control and archiving”

Better planning of end products

Data should be captured with the end requirements built in to the system. This should include analysis requirements for both dissemination purposes and deposition (both digital and physical).

Re-use of information at all levels
Having the information more accessible should enable people to use the data is a good way of checking quality and will encourage comments and feedback that should improve team working.

Online dissemination of integrated reports and archives
Having archives in a well structure digital format will enable their dissemination and use online, both by the unit and through national online resources such as the OASIS online form and the ADS catalogue.

Training & Continuing Professional Development
Training and development issues were the crux of gathering data efficiently and if people could recognise the difference between good and bad data, sites would get finished more quickly.

Change Management
This is a chance to agree and implement a much needed change management process.

3. Highlights of the Literature review

The Literature Review considers approximate 90 articles and books out of the 1500 identified in the Literature Search and entered into the database. The review is structured under 11 headings derived from keywords in the database: Work Practice and Knowledge generation, digital dissemination, archiving systems, data flow, database design, CAD and digital drawing, digital recording, GIS in archaeological fieldwork, post-ex use of contextual data, analytical systems, integrating different data sets. There was considerable overlap between these headings, which were established at the beginning of the literature search. Generally speaking, there was a lot more material relating to technical issues than theoretical discussions or methodology (see Figure X). Different topics flourish and fade through the 30 years of records we examined. A few articles form the cutting edge of a topic under discussion, followed by a flood of people using the technology as it becomes available, followed by a trickle of people actually developing the field.

3.1 Publication gaps

Given the vast number of references in the Literature Search database the publication gaps for our purposes were substantial. There is very little published that gives a “from Dig to Dissemination” overview of Information Systems in archaeology. Although Lock's recent volume (Lock 2003) summarises a huge amount of work, and covers all the stages we are interested in, the research base that he cites shifts widely through the book and each element in the research process is treated as a standalone entity.

The impression is that archaeological computing is in a continuing state of ongoing development, “Archaeological Computing is in a liminal time” (Lock 2003 p 263). In fact, looking at the scope of the literature over the last 30 years, this has always been the case. People publishing in archaeological computing are more interested in development than implementation. Some of this sense of being on the cusp of great things is connected to the discourse surrounding IT in general, which is closely tied to 'progressivist' narratives. (Huggett 2000).

Although the conceptual and practical issues surrounding work practice and knowledge generation have been discussed a good deal there are two sources of bias in the way it is discussed. Firstly, there is a chronological drop off through the life of a project. Much is written on how to record (Roskams 2001), some on interpretation in the field (Hodder 1999, Lucas 2001), some on phasing and the assessment (Roskams 2000), almost nothing on analysis. This means that there is little on post-ex use of contextual data, and integrating different data sets. Similarly, the heading 'analytical systems' really refers to analytical software, since the main characteristic of this material is how isolated it is from the rest of the process.

Secondly, the emphasis is on the experience and work of individuals. The role, nature and support of teams is barely discussed at all. The material on 'dynamics' that exists is heavily focussed on theory, which makes it difficult to identify patterns in relationships (Hamilton 1996, Bender et al 1997) As a result, there is little discussion of data flow as we conceive it, since getting data from or to another team member is not discussed.

Generally speaking, much of the material focuses on technical issues and solutions to specific IT or computer hardware and software issues (e.g. CAD literature and archiving issues and field digital data capture)
Although many people have published papers on their databases, and even on the process of development, only Schloen has published the data model that he uses for holding excavation data (Schloen 2001). Given that within the UK the recording systems tend to vary more in implementation than structure, such a model would be immensely helpful in advancing the quality of our systems.

There is some bias in the literature on dissemination that presupposes academic use of the Internet (exceptions which discuss this are Hodder and Huggett). Discussion of dissemination on the internet is still very much about the type of material traditionally published in a monograph, despite the fact that the Internet has a much broader audience. While there is a sense that VR will bridge this gap, there is little published on how our excavation reports, can reach these audiences.

3.2 Trends in well understood areas

In all areas of the review, papers repeatedly stressed the need for planning, documentation, and integration in the design of systems. Developments in the broader field of IS have greatly facilitated integration, but changing fashions have also created 'lop-sided' development. Recently there has been much less interest in databases, and more interest in modelling and representation.

There is a sense that all archaeologists are becoming "informational workers" (Lock 2003). The notion of 'computer archaeology' is being questioned (Tschan and Daly 2000). Everyone uses computers but only some people understand them, which leaves a 'high priest' culture and a lot of bad practice (Huggett 2000). There is a desire to integrate IS into the archaeological mainstream, but little sense that this is happening.

Interest in the impact of IS on work practice and knowledge generation is building (Denning 1996). Mostly this focuses on the increased range of communication and cross comparison (Hodder 2000). But there are also papers considering the social position of IS from a more critical perspective (Huggett 2000). Given the problems described above, the real assessment of the interplay of technological change, theoretical movements, and methodological developments remains to be written.

The most general overviews were in the Digital Dissemination and Archiving Systems topics perhaps because they focused most on what might be needed from the "end product". Nonetheless, it's still not clear what a digital publication report will actually entail. Many papers are advocating XML for use in digital dissemination of data to facilitate flexibility in end use.

Both relational databases and GIS have become fairly standard parts of the archaeological toolkit. There is a clear pattern of using databases to replace books of context sheets and the GIS to replace the plan chest, and there are a large number of projects using slightly different implementations of this basic idea. There is a push toward 3D recording and display, but little on 3D assessment or analysis.

4. Comparison of the two sources

4.1 UK vs. International

One of the main purposes of the literature review was to pick up international patterns in IS use. A further attempt at widening the perspective was attempted by presenting papers and through discussions at conferences such as CAA2003, WAC5 and EAA2003.

There is little difference between UK and international sources. The same issues of project specific development, divisions between IT and the rest of the profession and lack of clarity in requirements bedevil archaeologists internationally. The use of IS in archaeological fieldwork is well established in most countries with reasonably well-established archaeological infrastructure. It is also used fairly extensively in countries that host foreign researchers, though it's not clear in this circumstance where the balance lies in the establishment of requirements.

Most of the distinctions between systems from different countries lies in the recording systems in use, though many countries have used information systems as a 'wedge' to update recording systems, and some have changed (or are considering changing) theirs to make use of a good system. (Some people in France are considering the excavator records model in order to use INTRASIS).

The structure of archaeology in the UK also makes a difference to development of systems. Most of the people that we spoke to in the UK operate in private companies or charitable trusts. Most of the people publishing and whom we spoke to elsewhere worked for government departments, museums, or universities (though there were exceptions). There are obviously different constraints and opportunities in these circumstances. Given our position as a public body, CfA may be more like the international scene. Certainly there is no reason for us to develop our systems in a project specific fashion.

Another international distinction is in data standards and archives. The UK is generally recognised to be substantially ahead of other countries in digital archiving (which is a bit scary really). This is largely related to the work of the ADS, which is also active in data standards. The support for data standards (especially standards of content) varies considerably, with Italy and Denmark being more focussed on interoperability than standards, an issue which will be interesting to see develop.

4.2 Publication vs. personal contact

Not surprisingly, people were more willing to discuss problems in person. That being said *Archaeological Computing Newsletter* contains quite a lot of honest comment and criticism. Also people are often happy to discuss 'problems they have overcome' in a spirit of helping others from falling into the same traps. In fact it was amazing how many papers described circumstances which must have been awful to live through, though they did so with a brave face.

There is more emphasis on training in personal contact than in publication. In fact the only real discussion of training in the literature relates to using computers for training, rather than training people in using computers. This is linked to the 'upbeat' focus of most publications. Describing a training programme and its success or otherwise falls between the stools of theory and technical practice described in the literature review.

Another difference between publication and personal contact is that the focus on technology is much stronger in the published record. This is probably mostly because we were able to direct discussion, but also because most of the people we spoke to were aware that the processes and concepts were the important things and the details of implementation (what machines, what software, specific code) will change. The focus on technical details in publication may be because people see themselves as presenting solutions to problems which are often discussed elsewhere. Sharing the detailed methodologies of working with computers has also been an important feature of this field of publication from the early days, as a primary method of building technical capacity.

4.3 Currency of information

Although technical changes in IT make some literature outdated very quickly, the information we got from personal contact is not substantially more advanced than the information in the literature. This is due to three factors.

Firstly, many of the outlets, notably the CAA proceedings, are aware of the rapid change and are published considerably more quickly than many other parts of the archaeological literature. Most of the CAA proceedings are out by the time of the next annual conference. Secondly, people tend to write about material still in development. This is one of the reasons that there is less 'reflection' in the literature than there could be, but it also means that material has a longer 'sell by date'.

Finally, while the details of technology change quickly, the archaeological and methodological issues have changed far less. People are still looking for better integration, clearer standards, improved communication, and greater feedback. Since people publish along the way it’s unlikely that there is a perfect system out there which we haven't seen. IS in archaeology doesn't leap forward, it wanders. This isn't a bad thing, since it means that there's much more discussion along the way, much more room for feedback.

5. Overall assessment of archaeological computing practice

It is almost impossible to give a comprehensive summary of the current state of computing in archaeology because the picture across the discipline is so varied and changing. This section is therefore not meant to be a summary of all aspects of the uses of computing in archaeology. For something approaching this see Lock 2003. It aims to inform the scope and requirements of the Revelation project by examining the broad context of current archaeological computing practice.

5.1 Tackling some Myths

IT systems for archaeology – it’s already sorted

It is true that most archaeological units now use IT as a part of everyday work practice, indeed most organisations of any size have developed systems over a number of years. But what has emerged is that few if any, organisations have achieved systems that they are happy to consider fully integrated, and no-one we spoke to remotely suggested that they might have an off-the-shelf system that could deliver a plug-and-play solution to efficiently handle the capture, analysis, dissemination and archiving of all excavation data.

The picture seems rather that most Archaeologists now use IT for some aspects of their work and Archaeology, like many other sectors, is becoming much more of an "information industry" (Lock 2003 p265). But the degree to which IT has become an integral part of archaeology still varies immensely across different sections of the discipline. There are enthusiasts who are keen to try any new piece of software or kit that comes along. But the majority see the technology merely as something that will enable them to do what they've always done, only hopefully faster and more efficiently. In certain well-defined branches, which are often treated as specialisms, IT is used as an integral part of work, but because such work is usually regarded as specialist, any broader understanding of the IT used is slow to permeate to others.

As discussed elsewhere this is partly about training and development, but it is also about a willingness to accept change which is often harder to instil in a discipline that tends to focus far more on the past than the future.

It must also be recognised that the information technology will continue to change. Revelation is unlikely to provide solutions to all aspects of digital recording in a single project, but what it should produce is a plan for how to deal with the required changes to the work practice of CfA.

The data will not be safe unless it's on paper

It is important that archaeologists recognize that we are already creating a large amount of data digitally and that we need to create systems that follow best practice in the use of standard file formats and creation of documentation to preserve these records.

The CfA is actually in a good position with regard to the preservation of digital data. The CfA already has a Digital Archiving Strategy which details the requirements for implementation of a preservation strategy based upon data migration. But this should not be a reason for complacency and the CfA will need to ensure that new systems take the additional requirements of digital archiving into account.

There is no ultimate guarantee of the safety of any information, be it on paper or computer. Digital data inevitably seem more ephemeral than paper records because paper media has a much longer track record (although there are plenty of examples of people losing draft texts of books and reports on trains along with fires, floods and earthquakes destroying whole libraries).

The idea that paper publication is somehow a means of presenting the final record of an excavation is also something addressed by the Publication User Needs Survey report. "The survey exposes a muddle, and a paradox. Historically, one of the principal reasons for full print publication was the tenet of preservation by record. It was subscribers to that doctrine who nevertheless took the first explicit steps to conceding the inevitability that publication must in practice be something less than full (see Section 2). The concept then shifted to the idea of the archival record, and a full archival report. However, it is now widely understood, if not yet universally accepted, that excavation does not ‘preserve’ by substituting written observations for deposits that have been destroyed. For one thing, the complexity and potential of a given deposit will nearly always exceed scientific capacity for its exhaustive interrogation. For another, the idea that archaeological recording can be value-free is a delusion. And for yet another, serendipity aside, it is difficult to collect data in response to questions which have not yet been asked” (PUNS 6.2.1).

Computers will only provide numbers, not knowledge
This seems a bit like saying that books will only provide words not understanding. Many people still conceive of computers as number crunching devices (which they are) but not a means to provide information or knowledge. But with current web-based technologies it is possible to design databases to present information along with explanatory texts and prompts that explain why information has been collected and what it means.

Taking computers into the field will provide the solution
None of the organisations interviewed felt that they had been able to implement a “fully digital recording system” in the field and only in the case of one R&D project had anything approaching such been attempted. The majority of “digital systems” are either hybrid systems that combine paper-based records with digital records for end use, or systems which require double-handling to transform paper records into digital records for re-use.

However, many of those spoken to considered that the aim of capturing data digitally as early in the process as possible, did provide considerable operational and interpretational advantages.

Some of the major issues about the design of a system are not just about the use of the technology but are also are about how and what archaeologists need to record. Without a genuine re-appraisal of what information is required and how it is to be used the adoption of “digital methodologies” is likely to be at best superficial and potentially counter-productive.

Whatever solutions Revelation provides they have to go beyond a simple “we can do everything we do now, only digitally”. The best systems that others have developed to implement new technologies have been based on wide-reaching assessments of user requirements and re-appraisal of working practice where necessary.

5.2 Use of IT for field data capture and management

Again the aim here is to give a sectoral context for the Revelation project and in particular an assessment of how other organisations have attempted digital recording of information as early as possible in the archaeological process.

The Strategies for Digital Data Survey (Condron 1999) found that between 70% and 80% of field archaeologists used computers to access information. This survey shows that, even five years ago, a majority of field practitioners were already using IT in some aspect of their jobs. However this may actually reflect the numbers that use IT during post-excavation or report writing and seems in contrast with the numbers who actually use computers directly as part of fieldwork.

Only a few organizations are currently directly recording context data on computers during excavations. The organization that has been doing so the longest is the Landscape Research Centre (LRC), which pioneered the use of hand-held devices for on-site recording on the Heslerton Parish Project in the mid eighties. The LRC has continued the development of G-Sys; it's computerized geographic data management system, to hold the spatial and textual information that they record in the field.

Framework Archaeology have developed a "dynamic information management system" (Beck 2000) which includes the use of differential GPS to plan features digitally and enables excavation strategies to be informed by, and therefore adjust to, new data that is made available to the excavation team by use of the system.

MoLAS has a specialist Geomatics team who have been developing Field digital data capture (FDA) (Ziebart et al 2002). This methodology uses a total station (TST) to measure archaeologically significant points and a pen computer to view the digital drawings as they are recorded and enable text based data to be attributed to the surveyed spatial records. The work in London suggests that not all sites, and particularly those with dense complex stratigraphy, are suitable for FDA, and that a combination of manual planning with subsequent digitizing and FDA where practical, provides the best returns in a competitive commercial environment.

An increasing number of the larger (better resourced) field units (including CfA) have developed similar approaches to field data capture and some organizations have also developed systems for integrated management of archaeological data. One of the most advanced examples, the Integrated Archaeological Database (IADB) (Rains 1995) was originally developed for use by the Scottish Urban Archaeological Trust (SUAT) and is now used by YAT. Reading University has also recently adopted it for use on excavations at Silchester.

The picture in Europe is fairly similar, with the major users of digital recording tending to use systems which are enable a combination of paper-based and digital recording depending upon circumstances in the field. One system which has been widely adopted by at least eleven different organizations in several northern European countries is INTRASIS. This system originated in Sweden and has been used widely on over 700 interventions varying from watching briefs to large-scale excavations.

5.2.1 Summary

The discipline's current use of computers can be characterised as widespread use of relational databases for holding core archaeological records structured according to a variety of differing data models based on a range of different recording methodologies. As new technologies have become widely adopted for drawing, photography, analysis, publication and other aspects of the archaeological process, further software has been adapted with an increasing trend towards more fully integrated systems that enable the core data to be used and added to throughout the archaeological process. As Lock puts it, "Despite the variation in recording systems in use, it does appear that relational databases, based on the concept of single-context recording, do offer a stable core system for the written record. With the increasing flexibility of modern software, such databases can now be routinely linked with the drawn and photographic records moving the whole recording process into an integrated digital environment". (Lock 2003, 98).

6. Issues Beyond the scope of the Revelation project

The CfA is fortunate in being able to take an overview of sectoral practice. The research involved is beyond the project focussed budgets of most units. Since CfA requirements are more research led than many commercial organisations, the lessons from this review are not simply transferable to the project. The solutions for Revelation may still be internally driven. Further, it is clear that other organisations will still have their own problems and circumstances that they will wish to address independently. None the less, we hope that other organisations will be able to derive benefit from this sectoral overview.

There are a number of issues that have been picked up during the course of the sectoral practice review, which are considered beyond the scope of the Revelation project. Some of these include:

6.1 Bridges between "academic", R&D and contract archaeology

Problems were perceived in bridging the gaps between "blue skies" academic R&D and the day-to-day fieldwork realities of the commercial archaeology. This is an innately difficult problem, as the nature of "blue skies" is precisely to take research beyond the more mundane applications. As such any use of results may only ever be possible to achieve to a certain degree. But there has to be some measurable value in blue skies research. This gap is not just a problem for archaeology but it does seem to be one we are less good at bridging than say disciplines like Medicine, and the key may be for more developer funded academic research (perhaps something that Archaeology Commission and ALSF is attempting).

6.2 Continuing Professional Development (CPD)

There is a substantial need for CPD programmes in archaeology, particularly for IT, but also other areas of the profession. TORC is addressing these problems and we need to keep them abreast of issues surrounding IS.

As well as improving systems, getting better data depends on how people are working. A general point was made by many during the review of sectoral practice that there was a need for the profession to consider, better CPD provision. In particular aspects of IT such as survey, databases, digital publishing were cited. Some suggested INSET-style training days on site, building them in to work programmes including PPG projects.

Some interviewees pointed out that many project managers and senior staff can't maintain currency with archaeological practice in fieldwork and there may be a need for more CPD for project managers and similar staff to keep them involved in fieldwork & survey.

6.3 Change management in the profession

The design of most fieldwork systems in England is based upon the principle that individual excavators are responsible for recording all the site based information relating to the particular contexts they excavate. As a result, developing a good working system will be as much about enabling the users to learn the appropriate skills to use the system as in the design of what the system does.

It may need a sea change in the profession to bring about a wide-spread adoption of digital recording "from the trowels edge" right through to publication because the majority of people have learned what archaeology is on paper. Any such changes will need to be reflected, or perhaps even driven by, changes to the way archaeology is taught in universities and elsewhere.

7. Lessons for Revelation

7.1 Context of systems development

There is a fairly large community of people developing information systems for archaeological projects. Many of them would be willing to work with us both formally as consultants and less formally as peers. While we have found no system that we could simply adopt, we have found systems that suit *their* users needs quite well.

Some of these systems, and indeed some modules, will at least provide inspiration to Revelation and may be adaptable so that they form a component of our system. This may be particularly helpful in the development of technically complex issues such as the management of stratigraphy, metadata, XML schemas, data modelling, free text searching, "social" software (see below). Other aspects, particularly dealing with interface such as web served systems, direct entry in the field, and integration of archive into digital syntheses are still in relatively early stages of development.

The main barrier to the development of good systems is not technology. Contemporary hardware and software is powerful and flexible and many professions share our concerns with data management and analysis as well as communication - indeed communication is the central concern of most software development at the moment. It is true that many people involved in IS and archaeology are pushing at technological barriers - for example the development of topologically aware 3D spatial software - but core issues of integration, interoperability, reliability, and user focussed interfaces should not present insurmountable technological problems.

Where archaeologists suffer with poor systems it’s largely because they have rushed and scrimped, particularly in the systems analysis and design phases of development. Many systems are developed 'on the fly' throughout the life of the project, so that as the requirements of each stage come into view, developers are faced with panicking users. The classic example of this is organisations only addressing the archive specifications (both digital and hard copy) when they have finished analysis. Allied to this is a desire to use existent systems (or parts of systems) which often stems from expediency, but also from a reasonable desire for interoperability.

7.2 Context of implementation

Training is essential and often overlooked. Some organisations employ dedicated trainers teaching and encouraging people to use the new systems. We do not run enough projects for that to be useful or feasible, but we need a coherent training plan.

There is a substantial need for Continuing Professional Development in Information Communication Technology. Although we will be recommending training specific to the changes that we are proposing there is also a broader need for staff to keep abreast of changes and developments in a technology which is increasingly core to their work (even when they aren't aware of it).

We must plan to review and develop the system. Things change quickly, both in IT and in archaeology. Most systems that work either have a constant programme of development, or bring out new versions every 3 - 5 years. The continuous development route can only be done when there are full time in house developers who understand the system. Of course if the system is clearly designed and documented, update will be easier. By defining the end point, the products, that we're aiming for we can control the impacts of changes external changes.

7.3 Characteristics of successful systems

Integration is key. It is seen as the 'Holy Grail' only because people don't design with dissemination in mind. It has been growing since the 70's and what would have seemed integrated in ten years ago seems disparate now. Most successful systems are based on a well structured relational database with defined links and interfaces to a GIS and other programs for graphic and photographic display.

People are afraid of big systems; 'one size fits no-one'. What we need to recognise is that, in some senses we already have a big system, just the links are mostly 'procedural'. Our work is interdependent. Far from keeping things simple, the 'island' approach adds to the complexity and problems.

Legacy data is not going to go away, but we shouldn't be afraid to design our system to deal with what we are creating now. The clear message from a number of organisations was that we should adopt a 'year-zero' policy, where everything after a certain date would use the new system, but earlier projects would not necessarily be brought forward. Of course this is complicated by our backlog projects, but this needs to be dealt with as a migration issue, rather than designing the system around the existent data.

The least well supported aspects of the archaeological process are not data management, or even analysis, but communication, project management, contemplation and the development of ideas. Some of the systems out there do cover these aspects, but many of them fall over because they didn't put the same effort into these aspects as others. Others focus very tightly on data management and analysis but they rely on other structures to support the other aspects.

7.4 Further directions for research & development

The above assessment of current practice should help inform the CfA development of a core system that can manage existing data requirements. But as already stated, IT continues to change and develop and with the design of a new system there are opportunities to also consider other innovative areas that the sector is currently experimenting with that might become integrated into archaeological systems in the short to mid-term future. Some of these are considered below as possible areas that CfA might wish to carry out further R&D on. If we can design a system that is compatible with such developments then we should be able to carry out R&D using new recording technologies that allow the data from any such projects to be integrated into the new system.

Tracking Research Agendas

The increasing recognition that the research objectives determine the nature of the archive suggests that these should be integrated into the data structure, rather than being linked only through discursive text.

Mobile computing

There is growing recognition by a number of organisations of mobile computing as an opportunity to move to on-site digital recording at source. The use of hand-held recording devices has been well field tested by a few organizations. In particular the use of hand-helds for large area survey and field-walking and there may be opportunities for CfA to develop R&D projects to further develop such methods (Ryan 2002).

Wireless networking
Most uses of this technology are indoors and are intended to remove the need for cabling within office environments. There is however potential to set up an external system of network beacons on an archaeological site to enable data recorded on portable devices in the field to be transmitted and saved to a server in the site office, and served back to other users on site.

Tablet PCs
Some very limited initial work using a tablet PC, both in this project and by some of the units we spoke to, suggests that they may be particularly versatile for digital drawing directly in the field. Further R&D should investigate software flexibility and durability of the hardware. Tablet PCs may be particularly appropriate for buildings recording, and providing a mobile interface for TST or GPS recording.

Rectified photography
It is recommended that further research and experimentation is carried out into the accuracy and suitability of recording using photo rectification with annotation. This could be either in place of, or as a supplement to, the drawn record and as a means to gain additional pre-excavation and post-excavation data for use in analysis. This could be particularly useful for sections (where practical) and may be more appropriate for planning in some cases that require less on-site interpretation (e.g. skeletons or brick surfaces).

XML for dissemination and digital archiving
It is recommended that the requirements of the Governments eGMF metadata standards should be incorporated into any long-term dissemination and archiving requirements. The development of archaeological schemas for archaeology should be investigated in consultation with the appropriate standards bodies (e.g. NMR DSU FISH, ADS and others).

Social Software
Communication is central to teamwork. Most organisations have added email to their communication strategy, where it hold a place somewhere between the memo and the letter. Other businesses are beginning to look at dynamic web environments (Blogs, wiki's etc) to provide more team based communication. Archaeologists have yet to look formally at team communication.

Supporting/Defining the Archaeological Process
There is still little agreement on what is required in the process of ‘analysis’. There are clear guidelines for assessment so our system can be designed to support that stage. We need to get a clear idea of what kinds of activities take place during analysis. It has been said that you can’t write procedures or set standards for this stage because it is so variable. But we should be able to

identify a suite of activities that the system will be used for at this time, which lead to identifiable dissemination products. Initially we could go through Commissions Updated Project Designs (UPD's) to get a sense of the range, but this probably needs to be taken on by CfAMAN, Commissions, Strategy and the discipline itself.

7.5 Our Contribution

It would be useful for Revelation to publish certain elements of this Assessment stage in more widely disseminated fora than the CfA report. Given the split between 'technical' and 'theoretical' aspects it would be useful to discuss the issues of each in journals and conferences more usually focussed on the other.

The data models should be published in a peer reviewed journal, as well as being linked to CIDOC CRM in any useful ways. The 'field practice' work at Barrow Clump could well form the core of a useful article on 'the archaeological team' discussing the way the people use each other's data.

In addition to publishing articles describing any systems we build, we should publish on the process of implementation, with an emphasis on change management and training, since these are areas of substantial publication gap.

Appendix A - Fact Finding Summaries
Appendix B - Literature Review Report
Appendix C – List of software identified

APPENDIX 1E

USER NEEDS SURVEY REPORT

Report for the Revelation Project

Version 2.2

Jen Heathcote, Eddie Lyons and Andy Payne

October 2003

Contents

1 Introduction

The Revelation Project aims to improve the capture, analysis and dissemination of Centre for Archaeology (C*f*A) research by creating a coherent digital framework. Before making recommendations to achieve this goal, it was first necessary to:

- examine the systems currently used by staff at the C*f*A to carry out their work
- determine how effective these are
- examine how they might be improved in the future.

This process has formed the User Needs Assessment that looked at work processes, data structure and data flow to understand the needs of users in all of their different roles; as individuals, as members of project teams and within the C*f*A section in which they work.

This report summarises the results of the Assessment and presents an overview of the working environment at C*f*A, particularly with respect to data collection, handling, storage and management systems. An overview of the whole of the C*f*A is described and variations between categories of users (e.g. according to length of service or section) are also identified.

2 Objectives (including a definition of 'users')

The objectives of the User Needs Assessment were to:

- present detailed information about all C*f*A staff (users) concerning their current working practices with respect to data acquisition, handling and dissemination

- provide a preliminary statement of the needs of the full range of users based on working practices with respect to data acquisition, handling and dissemination, including the identification of areas where improvements could be made.

Users are defined as English Heritage staff working within the C*f*A, regardless of the nature of contract (i.e. permanent, longer-term fixed, or shorter-term fixed) and regardless of geographical location (i.e. Fort Cumberland or Savile Row).

3 Methods and sources

The assessment was undertaken by means of:

- a series of meetings (focus groups), each discussing the case history of a C*f*A project: these varied in size and stage of completion.
- a questionnaire administered to all C*f*A staff

3.1 User Needs Focus Groups

The focus group meetings (see Appendix 1 for details) allowed detailed information about working practice to be obtained from a range of C*f*A projects. Staff were asked to attend the relevant meeting for all projects in which they were involved, even if this meant attending multiple workshops. This was done so that differences between projects, and the role of individuals or teams within them, could be highlighted if necessary. The following projects were selected as representative of the variety of those undertaken by C*f*A staff:

- **Chesterton** – recent fieldwork
- **Longstone Edge** – completed assessment
- **Raunds Roman** – mid-way through completion
- **Silkstone** – in analysis
- **Catterick** – completed to publication

  plus
- **external projects** – non- C*f*A projects where EH resources are invested.

Each focus group meeting lasted for approximately 2 hours and during this time staff were asked to consider the following questions:

- **Sources** - how and where are data acquired/captured?
- **Routes** - how and where do data move within the project team?
- **Bottlenecks** - are there any obvious obstacles in data flow?
- **Duplication** - where does duplication/double-handling occur?
- **Access** - who has access to the data, both formally and practically?
- **Responsibility** - who decides levels of access to the data?
- **Outputs** - who receives data from the project and in which forms?

Highlights of these meetings are presented in Section 4 which identifies where projects have similar needs and where different or even conflicting needs became apparent. Results are structured according to the seven questions (Sections 4.1-4.7) rather than by project.

3.2 User Needs Questionnaire

The questionnaire was designed to:

- test our understanding of current and desired working practice developed through focus group meetings and other aspects of the project.
- provide a statement of evidence to support the recommendations to be made in the final C*f*A report.

Thirty four questions were asked (see Appendix 2 for details), arranged under headings to obtain information regarding:

- culture
- training
- work practices
- procedures
- tools available/tools required.

The questions were derived from issues arising in both the focus group meetings and during visits to other organisations that formed the preceding Review of Sectoral Experience. The majority of questions were multiple choice with one free-text question for respondents to record their top three wishes for improved working practice. The questionnaire was largely administered by online database, although hardcopies were supplied to those staff working away from Fort Cumberland or who felt uncomfortable filling it in online: all completed questionnaires were anonymous. The questionnaire was structured to allow breakdown of responses by section and length of service, but not designed for formal statistical analysis. It was also not possible to differentiate between fixed contract and permanent staff. The responses are individuals perceptions of their working practice, and are unlikely to be derived from any examination of timesheets, etc.

Highlights of the questionnaire analysis are presented in Section 5 which identifies variations in user needs and perceptions. The distribution of these variations has been analysed according to factors such as length of service with the organisation or section (Archaeology, Archaeological Science, Archaeological Resources).

4 Highlights of User Needs Focus Groups

See Appendix 1 for details of each of the focus group meetings.

4.1 Sources – how and where are data acquired/captured?

For all projects, the major sources of data were paper records created during excavation: context, sample, object and drawing records, plans, sections and photographs. These usually comprise in-house records on standardised proformas. Occasionally, alternative records (e.g. site notebooks) inherited from previous work, sometimes taking place several decades previously, may need to be incorporated into the C*f*A system as might data generated by external groups (e.g. local societies).

Topographic and geophysical survey data are acquired routinely. For recent projects these are in digital formats, but older projects may require analogue topographic and geophysical data to be processed. Topographic survey data and 3-D data are variably stored as text or CAD files. Very occasionally, only hard copy versions of digital files exist (e.g. an AutoCAD contour plan) due to lost digital data. Geophysical survey data may be generated by either the in-house team or external teams. Externally collected geophysical data can be of variable quality and format depending on the proficiency and working methods of the organisation undertaking its collection (both amateur and professional). Such data often requires re-formatting and re-processing to meet internal C*f*A standards of data presentation. This requires a site plan showing the location and relationships of data collection units (individual grid squares) that should be archived to allow any subsequent incorporation of the data into larger mosaics.

4.2 Routes - how and where do data move within the project team?

Data currently moves through projects on the basis of personal contacts. There are no established patterns of data transfer, with most transfers occurring in response to particular requests. The project manager is seen to be the focus point for this transfer, but not everyone is satisfied with this. For the majority of projects, digital resources are used in the transfer of data, though hardcopy may be made for particular circumstances. There are more difficulties with data transfer between external project team members than within the C*f*A.

Excavation data on standardised proformas are transferred to a digital database by Archaeological Resources after hardcopies have been made for security as part of the coming-off-site process. Also, all drawings should be scanned to provide security copies. However, it was noted that considerable time may lapse between coming off site and security copies being made resulting in the following problems:

- lost data due to lack of security copies and/or inventories of excavation records
- delay in transfer of basic excavation data within project team
- difficulty controlling the number of versions of data in circulation (also a quality control issue)

Lost data: poor documentation of the movement of data leads to difficulty in knowing where it is; ie. where a report, small find or sample is currently stored. Is it within C*f*A, has it been sent out to a specialist, or has a specialist report been received?. Finds may be "lost" as no detailed lists of what has been sent to whom (i.e. which specialists) exist: this information relies on the knowledge of individuals which raises problems when staff leave.

Data received by C*f*A from external specialists can be equally problematic to locate. Commonly, the receipt (by C*f*A staff) of specialist reports is not confirmed and there can be ambiguity about whether they have reached the right person. It is also unclear whose responsibility it is to create a security copy; is one made automatically on arrival of the report at the C*f*A or is it assumed that one is kept by the specialist author? An example of this was identified for a current project, where records show a geophysical survey was conducted several years ago by an external team. However, the report can't be found and details of the group commissioned to carry out the survey are ambiguous. Accessioning data as soon as it is received would help eradicate this problem.

Version and quality control: one project limited the number of versions in circulation by providing hardcopy print-outs that were worked on manually after the records had been digitised. In addition, all tasks were recorded in a single version paper-based logbook to allow control of what had been done and by whom. Varied logging of tasks by managers and specialists and the lack of record management is seen as at the heart of many problems currently experienced.

Checking the quality of data (stored in either analogue or digital forms) to be input to systems currently has to be done manually and is extremely labour intensive. For one project, this task alone required around 8 months of data cleaning, *i.e*. standardising / finding missing data, prior to the data set being transferred. In large part, this was due to data being stored on multiple databases that had to be checked for consistency (also a version control issue) and accuracy.

There is no standardised data route across C*f*A, or even within a particular section. However, for a number of disciplines (e.g. human bone, archaeobotany), data created is transferred to the relevant external professional database.

For those working on external projects, the Project Manager is often their only point of contact with the rest of the project team and all data is routed through them. Reports are often sent to the relevant English Heritage Regional Science Advisor and may be sent directly to other specialists involved in the project.

4.3 Bottlenecks - are there any obvious obstacles in data flow?

Major bottlenecks are thought to occur for two main reasons. Firstly, when the size of an excavation increases without a matched increase in post-excavation resources. This results in delayed data processing, subsequent data transfer and ultimately, project completion. Secondly, changing management priorities can result in time delays, sometimes of several years duration, between project stages (e.g. assessment and analysis). This results in repetition of tasks as assessment reports are rewritten to account for changes in, for example, the understanding of the region/period or for deterioration of materials.

In addition, insufficient communication between sections at the project planning stage can later result in delays due to unrealistic programming of tasks and bottlenecks also occur as a consequence of:

- not adhering to current working practice procedures and
- using current data collection/management systems whose design is not fit for purpose.

Working practice procedures: problems for the post-excavation process are created when, during excavation, standard proforma sheets are incorrectly or incompletely filled in. This could be avoided by increasing staff awareness and use of the C*f*A recording manual. Data checking of completed proformas also needs to be undertaken on-site and not relegated to the post-excavation process where it is more difficult to achieve. Procedures for both of these tasks exist but clearly require management enforcement to be effective. Similarly, post-excavation procedures need to be enforced and explained. For example, one projects' experience of scanning plans at too low a quality to allow their subsequent use serves to highlight why quality control is important. Being unaware of, or not adhering to, specified standards inevitably results in duplication of effort; in this case many drawings had to be rescanned.

Systems unfit for purpose: an effective system for tracking the location and status of bulk finds does not exist. The current system (Labfile) works for items (e.g. small finds and samples) that have unique identifiers, but it is insufficiently flexible to be used for pottery and bulk finds. This means there is no way to log either when or why items are temporarily removed (e.g. for specialist study) from a storage box. A proxy card system used to exist but has lapsed as it was time-consuming and inefficient to implement. At present, it is almost impossible to monitor what happens to these items either within Fort Cumberland, or once they leave the site.

The nature of this system (Labfile) also means that certain data cannot be captured digitally. Whilst working on a recent project, the Conservation branch found that the stabilisation treatments required for a major category of finds could not be recorded digitally, necessitating additional hand-written records to be made and stored. Similar problems exist for transferring data from environmental sample sheets to the database as currently only 50% of the information held on the paper copy can be entered digitally.

4.4 Duplication - where does duplication/double-handling occur?

Duplication of data currently occurs in many aspects of working practice and is particularly associated with the transcription of data from form to form and from hardcopy to digital formats. Although inefficient, it is not always detrimental to the archaeological process as occasionally, it has allowed lost records (e.g. context records) to be partially reconstructed from data recorded elsewhere (e.g. sample and finds records). Double-handling of data also occurs when old data is stored in a number of disparate formats that requires collating before analysis can occur, although this is unlikely to be an issue for recent projects. Some teams feel it impractical to avoid data transcription during working practice. For example, Conservation reported that information was routinely written down at the bench whilst working on an object and later transferred to computer. However, it was noted that it is impractical to have a bench-top computer and workers wouldn't necessarily want one anyway. Occasionally, double-handling of data was felt to be an inevitable part of the archaeological process as, for example, when phasing revision necessitates specialist analytical reports to be revised.

Duplication of individual tasks appears to be particularly problematical for the Graphics and Conservation teams. Object-related tasks were occasionally repeated in Conservation as objects cannot be tracked efficiently in Labfile. Similarly, illustration work has been replicated for a number of projects when data has been lost or not logged effectively. Occasionally, environmental specialists have been asked to assess the same sample twice. All of these duplicated tasks could be avoided by more effective logging of tasks completed and easier access to such records.

When dealing with external specialists, C*f*A archaeologists double-handle data in the course of providing each specialist with data in the format they require. Similarly, when analytical data is returned to C*f*A, it often comes back in very different formats which may need to be reconciled before it can be incorporated into the C*f*A system. The specialists themselves may unknowingly duplicate effort when they record the same information as other specialists working on the material (e.g. reporting the same information on the same find). However, individuals often do not know who other specialists on the team are, or what other relevant material exists if it hasn't been sent for specialist analysis.

4.5 Access - who has access to the data, both formally and practically?

Lack of access to data is currently experienced within project teams, between C*f*A sections and between C*f*A and the rest of English Heritage. For example, there is currently only limited access to unpublished data produced by other sections of English Heritage (*e.g*. MPP data) which has to be circulated as hard copy. Staff also feel that the drive to make information more publicly accessible sometimes takes precedence over addressing in-house needs.

In-house problems of data access are commonly caused when data is stored in personal directories rather than on the network share drive. Some staff were concerned that when work is stored centrally, it may be difficult to know whether you were using the most up-to-date version of a document, how many versions of the same document were in circulation and who was responsible for update control. However, these concerns are just as applicable, if not more so, when work is stored in personal directories.

Data seems to be stored in personal directories for reasons of personal preference, but also when staff are unaware of data management protocols. In one instance, this lack of awareness has meant the task of transferring excavation records onto the digital database was not programmed into Sheila Keyte's schedule (Sheila is currently responsible for digital input of site records) and access to data by the rest of the project team is delayed until time can be found in her schedule. A further problem relating to this is that inventories of the list of records to be input are not routinely provided. Consequently, the digital version of the excavation record cannot be checked for completeness.

Specific requests arising in the discussion of data accessibility were the need for:
- greater accessibility of primary excavation data across C*f*A
- a clear inventory of project resources (by individual project and by storage location)
- increased commitment to the use of central databases for data storage, including resources being made available for their maintenance.

In particular, specialists often find it difficult to access information other than that concerning their own area of study. For example, Environmental Studies would often like to know what pottery and small finds were found in the context from which a sample was taken as this may help to inform the high-level interpretation of the deposit.

Currently, access to this information relies on individual, explicit requests to other members of the project team.

In addition, the administration staff in Archaeological Resources often have limited (or no) access to information about the nature and aims of a project. Condensed site information is always needed by them, especially for project recruitment: the lack of access to this data often results in the team feeling pressurised. An executive summary stored in the share drive project folder at project approval stage would alleviate this problem.

A number of issues surrounding the use and maintenance of databases for archiving and interrogating data arose during the meetings, with a general consensus that more use of centrally stored databases should be made. It also became clear that the nature, amount and accessibility of information created by staff is highly variable across the C*f*A. There is no clear pattern of who may gain access to this information which ranges from databases that allow internet access by external users through to those which must be interrogated on a specific computer terminal within C*f*A. However, if wider access to information is to be created, it was noted that provision must be made to restrict access to some sensitive data according to the type of user.

Considerable concern was expressed at the lack of commitment to providing resources for the maintenance and quality control of databases, both within the C*f*A and in external organisations. With respect to the latter, the fact that Archaeology Commissions will fund the setting up of databases but will not fund their upkeep, even in partnerships, was highlighted.

4.6 Responsibility - who decides levels of access to the data?

Overall, most projects felt that it is the Project Manager's responsibility to determine who should have access to data although not everyone is happy with this and some categories of data fall outside this general rule. These exceptions are finds and environmental data, which are perceived to be the responsibility of other members of the project team. Very occasionally, no-one has taken responsibility for a site data set, creating major problems during the post-excavation process.

During one meeting, a question was raised concerning whose responsibility it was to ensure that the formats used for creating (external) specialist reports/databases were readable. Although initially the responses were a little ambiguous, overall it was felt that the responsibility should lie with C*f*A.

4.7 Outputs - who receives data from the project and in which forms?

Outputs were rarely explicitly discussed and where they were mentioned, it was in relation to the nature of specific products:

- site archive deposition - external deposition as paper record (within relevant museum) and C*f*A deposition as both paper and electronic data.
- assessment and UPD reports (paper and digital records within C*f*A)
- published analytical report (EH or external imprint).

Some specialist reports are supplied to non-archaeologists (i.e. public stakeholders). These include land- and property-owners or interested parties such as architects who may supplied with a paper copy of geophysics reports and property owners who will be given a copy of dendrochronology reports resulting from analysis of their property.

Staff were concerned that they often have limited or no control over the way their data is presented in external reports. Also, specialist finds work is usually published in such a way (i.e. without small finds numbers) that it is impossible to reconcile data contained within C*f*A Reports with that in the published reports.

It was noted that data generated during technological analysis is not formally archived in its raw state as it is not qualified and thought to be easily misinterpreted. It would require considerable time to produce it in a form suitable for archive deposition. However, the data is in a sense ‘archived’ within Technology Lab. (i.e. as files on the analytical machines). Also, samples analysed are kept and therefore can be reanalysed if necessary.

5 Highlights of results from questionnaire

(See Appendix 2 for an analysis of the User Needs Questionnaire.)

The questionnaire was designed to allow analysis that would highlight broad trends across the C*f*A, and identify potential differences of culture between the three C*f*A sections, Archaeology, Archaeological Science, and Archaeological Resources. It was not structured so that cultural differences *within* each section, where different teams exist, could be identified.

5.1 Culture

One cultural divide between users stands out (Q.5 and Q.6); the Archaeological Science section spends comparatively more time working on external projects and less time on internal C*f*A projects. The archaeological science section may therefore be less engaged with internal C*f*A procedures (where these exist) than staff in the archaeology and archaeological resources sections, who devote more of their time to C*f*A projects.

There is broad agreement that the C*f*A should have a shared philosophy (Q.4), although this is not defined clearly and there could be conflicts with external partnerships. Opinion is more varied on whether the C*f*A currently has a shared philosophy, especially when examined in terms of the length of time staff have been employed by the organisation (Q.3).

There is an even spread within the C*f*A overall of staff employed in the four service length categories (Q.1). Those in post for less than 3 years includes all staff employed on shorter fixed-term contracts. Newer staff members are dominated by archaeological science posts, especially in the 3 to 5 years category, coinciding with the relocation of the AML to Fort Cumberland in 1999 when the C*f*A was founded.

Poor access to data (Q.15 and Q.16) sadly is a part of EH culture resulting from the fragmented nature of the organisation. The strongest constraint is lack of access to data within other parts of EH (60% of responses) followed by data within C*f*A and then externally held data (both between 40-50%). Making our own data available to these groups is seen to cause us less of a problem (about 30% of each section).

5.2 Training

Awareness of internal procedures (Qs.18-20) is one area where training perhaps needs to be strengthened. Lack of awareness of members of project teams of procedures can often cause problems (see above). There is also a need for induction training/refresher training for C*f*A staff in procedures, especially those related to coming off site.

Procedures also perhaps need to be revamped, to make them simple, straightforward, sensible, and non-bureaucratic.

The majority of the C*f*A staff responding to the survey (68%) felt competent or better in use of computers with a relatively small percentage (13%) in the "awkward" category (Q.10).

Q11 shows that C*f*A staff generally possess the ability and willingness quickly to adapt and get to grips with new software for carrying out their work. Staff also are good at devising their own solutions to problems often improvising with the software and hardware tools they already have at hand (Q.27). The majority of staff (72%) began to feel comfortable using new software within one month of being introduced to it, although this must depend to a great extent on how complex the software being learned is. There will be wide differences between AutoCAD for example and a simple data download program. This was pointed out in a few of the responses to the survey. The one month or less figure was generally the case in all sections of the C*f*A, but those staff who took longer than a month were more likely to work within either the archaeological science or archaeological resource sections. This may reflect the more complex specialised non-standard office software employed by these groups such as graphics applications and specialist hardware related applications.

A greater proportion of those staff who had worked at C*f*A for over 10 years took more than 3 months to become familiar with new software. There is a clear correlation that some staff who have served longest with the organisation class themselves as a technophobe (Q.8), regard themselves as awkward with computers (Q.10), and take longest to feel comfortable with new software. It is possible that these longer serving staff also are dealing with more complex software.

There is evidence that C*f*A staff are generally very self reliant and can usually get by with self-learning and minimal training, but training as well as immediate access to the software involved is obviously desirable for getting up to speed quickly and efficiently. Training is often wasted if subsequent access to the software is limited by lack of availability (for example where there are not enough licences) to become quickly familiar with it.

Few staff are trained with email and internet usage and only just over half with Word processing (Q.12). Only about half of the users of databases have formally trained and a smaller proportion for spreadsheets. People who use graphics packages, specialist data processing software and statistical packages are largely self-taught. AutoCAD and GIS stand out as having a high level of related training. Few want more training with email, Word and the Internet but the wish for more training jumps up considerably for graphics packages, statistical packages and GIS. Database and spreadsheet users want further enhanced training.

A more detailed analysis of computer training (Q.12), though, did show up some worrying trends. There is a high percentage of software users who have had no formal training provided – they therefore must have self-trained. This is the case even for the categories of more complex software. A significant percentage of these non-trained users do not want formal training. In contrast, significant numbers of those who have received training in the past want additional training to be provided. There also appears to be a significant demand for potential new users, especially of GIS and statistical packages.

5.3 Work practices

Only a minority of staff spend little time working on computers (Q.7) while the majority of staff (62%) spend a high proportion of their time (over 60%) doing so. More archaeological scientists are in the two highest categories of computer usage (60-100%) than either the archaeologists or archaeological resources staff, indicating that computers are the main tools for manipulating and analysing scientific data. In the 40-60% bracket the scientific staff are outnumbered 2:1 by archaeologists. These figures might reflect a fairly even divide among the archaeologists between management type tasks (specifically project monitoring) that involve less time on computers and more project related time. No scientists spend below 20% of their time on computers.

In cases where PCs are seldom used this probably relates to jobs where traditional methods are still the best (for example hand illustration, filing with paper-based systems) or an emphasis on spoken communication (answering the telephone and attending meetings etc).

Length of service seems to be immaterial to the amount of time spent working on computers. The vast majority of staff do not consider themselves technophobes and there seems little evidence of reluctance among staff to adopt new software solutions. The small minority (13%) who admit to being technophobic range across the different sections of the C*f*A, with those who have worked the longest for the organisation being the largest proportion of the group. This perhaps is not a surprise, since those who began their careers in a purely paper-based office might find it hardest to adjust to an increasingly digital culture.

For the majority of staff, getting information is seen as a mid-range time investment (Q.9) which is unlikely to alter much in the future. Ten percent of respondents said that it consumes most of their time, suggesting there are obstacles in accessing information. This impression is supported by a significant group who would like to spend less time rather than more time on getting information. Obstacles in accessing data and getting access to information also are a recurrent theme of the focus group meetings and the feedback expressed in the wish-list comments section of the questionnaire.

Currently communication is seen as an activity requiring a high time investment and ideally staff would like this to be more mid-range in terms of time. Communication is important however as a lot of the problems highlighted in the focus groups stem from poor communication (see Section 4.2 Routes, Section 4.4 Duplication).

Organising things – the overall trend across the C*f*A is a desire to spend less time on organisation tasks.

Entering/preparing data – the majority of staff would like to spend less time entering and/or preparing data. This could be seen as an indication of the desire to streamline the process by adopting more automatic recording and data capture.

Analysing data – the questionnaire indicates that this is currently a low ranking task in terms of time investment, but the staff are clearly not satisfied with this situation and ideally would like to increase the amount of time they can spend analysing data. This probably reflects a desire to analyse data at a higher level incorporating other data-sets that may be relevant to improving the overall level of interpretation as expressed frequently in the wish-list section of the questionnaire.

Writing and drawing (producing output) – approximately 40% of staff both currently and ideally place this as their primary task on which they spend the majority of their time. Given the nature of the task and the expectation on staff to produce identifiable outputs such as documents, reports and publications this is no surprise. Performance is generally judged on written output and other paper-based deliverables.

To summarise in the future the majority would like to spend:-

Less time -	Communicating, Organising things, Entering/Preparing data
More time -	Analysing data, Writing and/or drawing
Little or no change –	Getting Information

Diaries/day-books (Q.14) – the ratio of keepers to non-keepers is slightly less than 2:1 for all sections showing that it is a common practice but not universal. This is probably simply down to personal preference, since there is no corporate requirement to keep them.

A high percentage of C*f*A staff do not always work on networked drives (Q.17). The main reasons given for this is that files or data relate to work in progress and is incomplete, and because people are working away from the office. Network problems including slowness or incompatibility of file transfer and software with the network, and problems importing files from the network, were also cited as reasons for storing files on non-networked local drives. The fact that significant amounts of work is stored on individual computers' hard drives, and as a result is not available for regular network-based backing up, has serious implications in the event of hardware failure, damage, loss, or theft – unless individual users arrange for regular backups to be made of their files.

Training to improve awareness of the architecture of the C*f*A computer network, as well as the development of clearer procedures so that people are more willing and able to put their work on main servers, also would aid accessibility to other peoples data and easy exchange of data. The network's "Share" drive increasingly is being adopted as a means of doing this but probably requires rigorous use of metadata to describe files, and for reasons of version control.

A specialist shared system works well for the geophysics team in the archaeological science section. They have a shared local network which provides them with access to each others projects and data as necessary. While this works well for this team, it is isolated from the rest of the C*f*A.

An index of project codes in relation to site or project names within the main projects directory of the network "Share" drive would be a good way of sign-posting data more effectively to help occasional users not familiar with particular project codes find it more easily.

5.4 Procedures

Levels of communication with work colleagues

Overall C*f*A staff are evenly split between those who are happy with the level of communication between colleagues during project work and those who are unhappy (Q.13). The lowest level of dissatisfaction with the level of communication is in the archaeological science section. The highest level of dissatisfaction with the situation is in the archaeological resources group. The archaeologists are evenly split on the issue.

Procedures for project management, collections management and data management
38% of respondents currently are unaware of any procedures for either project management, collections management, or data management (Qs.18-20) and only 17% were aware of all three management procedures.

Those with an awareness of procedures were clearly outweighed in all categories by members of staff who are currently unaware. This was especially evident for data management procedures (only about 25% of respondents).

For those aware of procedures a reasonable majority thought they made sense in the cases of collections management and data management but less than half thought the project management procedures made sense. Two thirds to three-quarters followed them for project management and collections management but only just over half for data management and there was a general perception that the figures were much lower for other people. In all cases the majority thought that the procedures were helpful even though in the case of project management a majority thought they didn't make sense.

Respondents remarked that procedures are helpful only if enforced and adequate resources provided to allow them to be followed. One respondent noted that whilst procedures made sense and were helpful, neither they nor others followed them since time, resources and enforcement were in short supply.

Passivity of upper management in relation to enforcing working practice protocols leads to problems further down the line. This could in large part be achieved by releasing resources for the completion, updating and enforcement of the Procedures Manual. The Procedures Manual also needs to be a useable and user-friendly document. Too large and unwieldy a document is a disincentive to staff to become familiar with and use procedures.

Collections management relies on honesty, diligence and consistency. Lack of records management lies at the heart of most of the problems revealed during the focus group discussions.

Quality control of data
Of the 92% of C*f*A respondents with responsibility for data (Q.21) the most common methods of quality control are inspection during use (63%) and record by record inspection (59%). Under 40% employ validation controls on entry, less than 30% use intermittent inspection and the least popular quality control method at below 10% is double entry of data.

Protocols for exchanging data
In only about 50% of cases do C*f*A staff specify the format data supplied by colleagues should arrive in. This is a bit peculiar really, because it really does matter what format the data is in for it to be usable. Even more oddly, only just over 40% of internal and external colleagues specify the format our data should arrive in. This suggests a general lack of awareness of common standards or a blasé approach based on the assumption that it will take care of itself. Alternatively it could mean that people are

fairly adaptable and have no problem manipulating a variety of formats or converting the data into a common usable format. (Qs.22-24.)

Those who have worked for the organisation for over 10 years are the most likely to specify formats they require for data exchange (perhaps indicating that they are less flexible, or that they are more aware of incompatibilities due to experience over the years).

5.5 Tools

Office applications have the highest levels of use across the whole of the C*f*A (Q.12), with word processing and email being used by 100% of the staff who completed the questionnaire. The use of internet browsers, databases and spreadsheets also was high. The remaining software is more specialist with around half using graphics packages and specialist data processing software. Smaller groups (1/3 of staff or less) use AutoCAD, GIS and statistical packages. A lot more staff would like to add GIS and statistical software to their skills base. CAD usage currently seems to be close to its natural limit – there are only a few additional potential users, but more than half of trained current users want more training.

<u>Spatial interrogation of data</u>

39% of respondents currently analyse their data spatially (Q.25). This would nearly double to 75% of staff if the tools and training were made available to all who would like to. All sections of the C*f*A are represented in both groups. How we define analysing data spatially is not clear in the survey – it could mean use of GIS at the most sophisticated level involving multiple layered data-sets, or simply plotting distributions of data at a more basic level.

A similar trend is apparent for analysis of data in relation to data gathered by colleagues (Q.26) with a clear increase from those who currently undertake this (49%) to those who would if the tools were available (79%).

One respondent commented that just because we can analyse our data spatially and in relation to other peoples data doesn't mean it is always worth doing so, or even applicable. The capability should be there if needed, but is not necessary a routine requirement.

<u>The right tools for the job</u>

77% of respondents indicated that they had been in the position of lacking the correct IT tools to do their job efficiently (Q.27).

The most common solution to this problem (73%) was to devise a way of doing the task with existing tools or devise a different way of dealing with the problem (39%). Although these solutions are far from ideal, they do indicate that most C*f*A staff are very capable and creative at improvising and making do with what they have. While this is an undoubted strength, it should not be taken as an excuse not to invest in better and more efficient tools – it underlines a significant weakness in systems provision.

The opposite side of the coin is that 66% of staff did the job less efficiently. Providing the right tools for the job would speed up tasks and the process of project completion.

The other most common solutions were to ask a colleague in the C*f*A to help with the task (56%) and put the task off and allow it to slide without completion (44%). The first solution is problematic because it has a knock on effect on the work programmes of the other people brought in to help. Sorting out the problem at the point of origin would therefore free up staff time and enable more time for doing other things. Putting the task off preventing completion is really no solution at all but is all to common when such obstacles are encountered.

Around 30% of the respondents said they had asked colleagues in EH and external contacts for help with the task or arranged for an appropriate tool for the job to be purchased or developed at the C*f*A. The use of external contacts relies on the ability of individuals to create and maintain networks outside the organisation and will also highlight areas of EH inefficiency to external organisations.

Less than 10% had commissioned an appropriate tool to be developed externally, probably indicating an absence of an appropriate budget.

Database usage

A high proportion of *CfA* staff (79% of the survey respondents) regularly use databases and spreadsheets (Q.28). Difficulty of use and unreliability stand out only marginally as greater causes of irritation to users. The least problematic aspect appears to be the software interface.

By far the most popular form of data visualisation used by the database and spreadsheet users (Q.30) is a list or table (just under 90%) with just under 50% of users also opting for forms and graphs. Other categories included a distribution, a report, “on paper” and “spatially”. Two individuals pointed out that the nature of the data presentation depends on the type of data and the audience for the end product. It was also noted that specialist scientific spreadsheet programmes (other than the office standard Excel) are required for some applications such as Archaeomagnetic Dating that require specialised forms of graphical presentation.

Physical collections (such as index cards, reference collections)

A high proportion of respondents (72%) currently use physical systems of some kind (Q.31). Of these 59% would prefer to be using digital systems (Q.32) whilst the remainder prefer physical. This may be because either:

a) users would prefer to stick with what they know and are comfortable with,
b) it is not thought to be worth investing resources in converting the physical records to digital records, or
c) the physical system does not lend itself to being digitised.

Only 28% of the users of physical collections indicated that there was some form of disaster protection control in place for their systems (Q.33).

Access to physical resources (currently and ideally)

There is general recognition in the questionnaire responses that access to physical systems restricted to single users is far from ideal (Q.34) and wider availability of these data is desirable. Availability of often fairly specialist data to the whole of EH is not thought to be particularly useful or important (which would make sense) but availability to more professionals within a particular specialism (internal and external) is an aspiration making the most gains in the “ideally should have access” category. The availability to “my team” figure comes down in the ideally category in favour of wider availability.

5 out of the 7 respondents in the single access ("only me") category want to provide wider access, indicating that the motive for single access is not a selfish desire to be restrictive and over protective of data. There is also a desire for greater public availability (an increase in response from 0 to 3).

Any data made available to the public also would be available to EH staff. A safeguard should be put in place to ensure that C*f*A and EH staff can have immediate access to such data. This appears to have been overlooked by the NMRC, where EH staff from other offices have to stand in line with the general public before gaining access – often with a long waiting time. Greater availability of collections to the public, if demand is high, generally requires a dedicated post to supervise and coordinate the required access.

6 Comparison of Focus Groups and Questionnaire

The C*f*A currently operates as a series of fragmented cultures that are neither aligned entirely along the sections (Archaeology, Archaeological Science, Archaeological Resources) nor by the length of time that staff have worked for the organisation. One of the fundamental issues to be highlighted in the focus groups and the questionnaire, was the number of C*f*A staff who are unclear about what constitutes a C*f*A project (as opposed to an external project). Staff were not only unsure of the definition of a C*f*A project but also were often unaware whether a project was Archaeology Commissions funded (and therefore internal to EH) or truly external e.g. when a university project uses EH staff (usually 'in kind'). The reasons for this lack of understanding are ambiguous, but it is a clear illustration of the lack of cohesion in work perceptions and practices that is found across C*f*A. In particular, the variability in knowledge of existing procedures for project, collections, and data management is worrying and appears to be at the root of a number of the problems currently experienced. Training needs to be strengthened to increase awareness of basic, internal procedures, especially those related to coming off site. In addition, revision and streamlining of those procedures would be beneficial.

Many of the problems highlighted during the focus group meetings stem from poor communication. For example, most of the bottlenecks experienced in transferring data and many of the reasons why tasks are duplicated or data is double-handled result from either ineffective or a lack of communication between members of project teams within C*f*A. It is interesting to note that in the questionnaire, staff indicated that they would like to reduce the amount of time they spent doing communication tasks. Given that many staff currently invest a lot of their time in communicating, this suggests it is vital that a means for more effective communication of data is found.

Staff are concerned about the lack of access to data both within C*f*A and across the organisation. In-house problems of data access are commonly caused when data is stored in personal directories and a common request during the focus groups was that more people stored more data on the network share drive. However, the strongest constraint is felt to be the lack of access to data within other parts of EH. The questionnaire asked staff where they tended to store data and why they chose that location. Responses showed that it is rare for C*f*A staff to always work on networked drives. This is for both pragmatic reasons (ie. working away from the office) and personal preferences (ie. the work is incomplete). Although the network's Share drive is gradually seeing increased usage which will help with accessing data, there needs to be a comparable increase in the recording of metadata to describe files and aid version control. Concern was expressed both during the focus groups and in the comments section of the questionnaire regarding whether additional resources would be made available for the creation of metadata: this is clearly seen as a separate (and optional) task at present.

Staff want increased access to data and many would like it to be available within a relational database format, allowing both storage and interrogation of data. A key element is that there must be flexibility in how data is visualised (forms, graphs, tables, lists) as data presentation needs vary according to both the task being undertaken and the eventual audience of the work.

Finally, six themes were identified from analysis of the wish-list and comments recorded as free-text in the questionnaire (Appendix 3) and in the 'Improvements' discussion of the focus groups. The six themes, in no particular order are:

- improved access to, and integration of, data
- simplicity, flexibility and ease of use of systems
- improved databases and data quality
- (enforcement of) procedures for data collection, management and standards
- provision of the right tools for the right job
- digital recording

Access to, and integration of, data: easy and effective information retrieval is a high priority for staff as indicated by requests for an "*Openly accessible shared data resource in compatible formats*"; "*a single commonly agreed integrated digital system for all aspects of CfA work*"; and "*a single database with access for everyone*". Easier access to data generally refers to the wish of project team members to have better access to project data-sets that are often currently stored in disparate locations. Standardised data formats were also seen as being important for enabling better integration and shared access. A team based approach to collecting data was advocated, as was the integration of data at the immediate post collection stage. Better cross departmental communication was seen as important for facilitating this. Improved access to data would also include the ability to access data remotely from locations outside Fort Cumberland including excavation site offices via a modem. Greater accessibility to data also includes better access to electronic resources outside EH including library records and on-line journals.

Simplicity, flexibility and ease of use of systems: there was a common desire for systems that are simple to use. This is not necessarily seen as conflicting with wishes expressed above. In an ideal world, data handling should not be a constraint on a project; systems that are simple and easy to use also have the benefit of saving time. A system that streamlines a process and saves time while achieving its intended function is likely to be universally popular and is unlikely to require high levels of management enforcement. In this respect, the ideal is a reliable, stable system that possibly includes an element of automated data capture.

Flexibility is equally as important as simplicity and ease of use. For this reason it will be important to avoid any temptation to develop a monolithic IT infrastructure with one size fits nobody data-standards that requires additional staff time to comply with standards that have no visible return for the extra work involved. There is also a continuing requirement to retain the flexibility to use specialist software, freeware software solutions and develop our own tools where necessary.

Database issues and reliable data from the point of origin: the most frequently recurring theme was the request for a reliable relational database for the organisation, storage and querying of project records (contexts, drawings, objects, photos, samples, etc). Part of the required function would be the automation of cross-referencing between records *eg.* drawing numbers, photo records, context numbers and sample numbers in order to increase the ease with which information can be accessed and decrease the time spent locating it. Creating a system with a single point for data entry was also highlighted, not only in order to reduce the time currently spent copying data, but also to reduce the potential for errors to be introduced during data transcription. There should also be the ability to alter records after initial entry (records not necessarily 'frozen' on entry), however this would also require that data changed in one location was automatically updated in the other areas in which it appeared.

The creation of a specific, additional database was requested to either replace Labfile or to act as an add-on. This would perform the same functions as Labfile but would provide greater flexibility for recording treatments of objects and in addition would allow the processing and location of bulk finds to be tracked.

Procedures for data collection, management and standards: there was a request for clear procedures for the collection and management of data that are enforced so that all users understand what they are being asked to do (and also why they are doing it). These procedures should be simple, straight-forward and easy to use to encourage staff to adhere to them. They should also be commonly agreed, documented and the details readily accessible to staff. The absence of enforcement of extant procedures (by managers) was felt to lie at the heart of many of the problems experienced during the post-excavation process.

The right tools for the right job: many of the staff stated that, at some time, they had found themselves without the IS tools required to carry out their work efficiently. However, improved technology (eg. faster PCs) rarely appeared on anyone's wish-list. Users who work away from the office want reliable hardware that works away from the office as well as faster, more reliable remote access to networked data. The key request by C*f*A staff is that everyone is equipped with effective IT tools for doing their job, and that regular reviews of needs are undertaken. More up-to-date versions of certain software packages were required to support exchange of data with internal and, in particular, external colleagues.

Digital recording: very few individuals mentioned this desire on their wish-list and it is clearly not felt to be as important as other issues such as provision of the "*right tools*" and "*better access to and integration of data*". However, many staff wish to spend less time transferring data from paper based records to digital systems. Removing the need for data transcription would reduce delays in accessing data for the project team and may help to prevent loss of cohesion/momentum of a project. This would also help to address the request for reliable data from the point of origin issue (see above), removing the need for some aspects of data double-handling and potentially speed up the archaeological process enabling faster dissemination of result: these benefits are explicitly stated by some individuals.

7 Conclusions

C*f*A staff (users) need reliable, comprehensible, and traceable digital data sources for all aspects of project work. Variable and limited access to this data, and poorly defined responsibility for data is undermining project progress across the C*f*A. Staff need greater structural support for communication and managerial backing for clear procedures, laying out standard working practices and responsibilities. They need a wide range of IS tools, many of which they already have but there is variable access to some tools and a number of staff lack access to the full IS toolkit that would enable them to fulfil their role efficiently.

There are few things, which users see themselves as completely unable to do under the current circumstance, but they do feel that they could do them much more efficiently. They need to spend less time developing and maintaining systems, collections and *ad hoc* data management tools and more time analysing and producing the outputs they see as central to their jobs.

Staff interviewed for the User Needs Survey are generally happy about the prospect of change, but are (like most people) unwilling to 'lose any more time' to it. Although people need and expect better IS tools to help with the problems they experience, they have become adept at creating their own, individual solutions. They also recognise that many of the solutions are cultural and procedural and can only be resolved by increased training in, and enforcement of, extant procedures. The very existence of procedures needs to be given a higher profile within the C*f*A and the worrying culture of mistrust between staff with respect to adherence to management protocols needs to be addressed. Team work (as opposed to a collection of individuals) within projects is variable and inter-section communication needs to be improved as does basic organisational awareness.

Overall, staff want to use systems that a) reduce duplication of effort and double-handling of data across the C*f*A and EH, and b) produce reliably accurate data from the point of origin. With respect to the latter point, it would be desirable to increase the proportion of staff employing automated validation controls as their principle method of controlling the quality of records. Currently, more labour intensive methods (e.g. record-by record inspection) are typically used.

Two overwhelming user needs have been identified during the course of the User Needs Survey:

- the development of a genuine C*f*A system that can cope with the full range of tasks and projects currently undertaken by C*f*A staff
- the streamlining, standardising and enforcement of procedures for data collection, recording and management.

These two needs are inextricably linked: both are required if either is to succeed.

Appendix 2. Conference papers presented.

1. Revelation: practice, technology, dissemination and the design of a field recording system.

Presented by Keith May & Sarah Cross at CAA, Vienna April 2003 (proceedings forthcoming)

On site computer recording of archaeological information has been a reality for 30 years and yet it is still seen as an experimental or at least innovative approach. Why? While most archaeologists use computers in some aspects of life, field recording is still done largely on paper. What is the challenge in field recording that makes it resistant to change? The problems may be as much cultural - to do with the way in which we see 'the field' in relation to the rest of our work, our colleagues and our audiences - as they are technological.

The Revelation project has undertaken a comprehensive review of information systems at the English Heritage Centre for Archaeology in the context of the broader profession. Our aim has been to understand how we use data throughout the life of an archaeological project so that we can design a field recording system that is used by the majority of our field teams. In the process of this assessment we have had the chance to reconsider how our working practice feeds our understanding and how it can be supported and improved by better designed systems.

This paper presents the results of the assessment and plans for implementation.

2. Revelation: data quality as fitness for purpose.

Presented by Vicky Crosby at CAA UK, Southampton, May 2003

Archaeological field data is notoriously flaky. Inconsistent, incomplete, undocumented and unchecked datasets are common, even in well-funded research contexts. Information systems hold the promise of increasing data quality through supporting standardised approaches, as well as automating validation and verification. Unfortunately, poorly designed systems can make data quality worse because they cut across the quality control practices of standard archaeological practice.

Since quality is fitness for purpose, we can only increase data quality through defining what it will be used for. Given the range of audiences for archaeological data, this is no mean feat. Once requirements are defined, standards can be applied and systems can be designed to meet them.

The Revelation project has undertaken a comprehensive review of information systems at the English Heritage Centre for Archaeology in the context of the broader profession. Our aim has been to understand how we use data throughout the life of an archaeological project so that we can design a field recording system that is used by the majority of our field teams. In the process of this assessment we have had the chance to reconsider how our working practice feeds our understanding and how it can be supported and improved by better designed systems.

This paper presents the results of the assessment and plans for implementation.

3. Revelation: practice, technology, dissemination and the design of a field recording system.
Presented by Sarah Cross at WAC5, Washington, June 2003

On site computer recording of archaeological information has been a reality for 30 years and yet it is still seen as an experimental or at least innovative approach. Why? While most archaeologists use computers in some aspects of life, field recording is still done largely on paper. What is the challenge in field recording that makes it resistant to change? The problems may be as much cultural - to do with the way in which we see 'the field' in relation to the rest of our work, our colleagues and our audiences - as they are technological.

The Revelation project has undertaken a comprehensive review of information systems at the English Heritage Centre for Archaeology in the context of the broader profession. Our aim has been to understand how we use data throughout the life of an archaeological project so that we can design a field recording system that is used by the majority of our field teams. In the process of this assessment we have had the chance to reconsider how our working practice feeds our understanding and how it can be supported and improved by better designed systems.

This paper presents the results of the assessment and plans for implementation.

4. Revelation: Researching the structure of the digital archive
Presented by Claire Jones at the Digital Research in the Humanities conference, August 2003

At the English Heritage Centre for Archaeology we create and hold a large quantity of digital data, created through our archaeological and scientific research.

Once we capture the raw data and start the assessment and analysis phases the amount of data we hold expands considerably. Most of this data is stored on our systems in relation to individual projects. However, although research is a collaborative exercise, which requires shared access to large and complex datasets, it is conducted by a disparate group of people, who create, stored and managed their data separately. This, and a proliferation of our IT systems, has led to data being increasingly difficult to integrate, complicated to archive and slow to disseminate.

In 2002 the CfA started the Revelation project, the aim of which is to provide a coherent digital information system that will make the capture, analysis dissemination and archiving of CfA research faster and more effective. We wish to improve the articulation of these systems to save duplicated effort and support integrated research and professional standards. The project will investigate the use of new technologies and the World Wide Web to deliver a truly interoperable workflow-based system.

This paper will discuss the active role of the archives team in the development of the information system for archaeological research. We have the opportunity to reverse our relationship with the archive; through inputting into the creation and structure of the archive from its conception, rather then when the project is complete. This will enable us to re-use the data and support future research, teaching and learning.

5. Digital Recording Systems and Archaeological Practice: if everyone else is doing it, why shouldn't we?
Presented by Keith May at EAA2003, St. Petersburg September 2003

The Revelation project has undertaken a comprehensive review of information systems at the English Heritage Centre for Archaeology in the context of the broader profession. Our aim has been to understand how we use data throughout the life of an archaeological project so that we can design a field recording system that is used by the majority of our field teams. In the process of this assessment we have had the chance to reconsider how our working practice feeds our understanding and how it can be supported and improved by better designed systems.

This paper presents results from the assessment

Appendix 3 Statement of Archive

The project has generated a substantial amount of data, which has been managed digitally throughout the project. Metadata conforming to Dublin core and extended for version control has been recording as data is created in a Access database designed as part of the project. The archive will be deposited according to the CfA Digital Archiving Strategy (Brown 2000) at the end of this stage of the project. This appendix describes the material as used by the project. Some of the software formats will change when the archive is deposited to conform to archive standards. Some issues, such as how to manage correspondence created as email, require further clarification of our existing strategies and procedures.

In addition to correspondence, minutes and other administrative data the archive will contain all of the internal reports produced for the project and their appendices. Many of these appendices are large datasets held in Access databases. There are also some GIS, CAD and other graphics files particularly relating to the digital drawing research strand. PowerPoint presentations and written texts used both in conferences and internal presentations are also stored

The archive is structured according to research strand and administration. There are 7 major directories: Existing Systems, Field Practice, User Needs, Sectoral Practice, Project Management, Presentations, and Data Modelling. Each of these has several sub directories reflecting different aspects of the research. In addition the correspondence of the project (both administrative and relating to intellectual issues is currently held as email files. The appropriate structure for this archive is still being determined.

Existing Systems contains a database detailing responses from CfA staff about their current use of information systems (of all types) as well as information from the Fort Cumberland servers. This data was originally captured on paper proformas and these paper records should be considered the primary archive for the information. This directory also contains the report on the review of Existing Systems.

Field Practice contains subdirectories supporting the digital drawing report as well as participant observation at Barrow Clump excavations. The digital drawing subdirectory contains field notes on Trials at Dartmoor, Whitby, and at Fort Cumberland. It contains photos, CAD drawings and GIS files in both ArcView and PocketGIS. It also contains the report on Digital Drawing. The field practice subdirectory contains field notes, the field practice report and its appendices. It does not contain any of the archaeological archive of the Barrow Clump excavation.

User Needs contains two major data sources: the focus group reports and the User Needs Survey. The focus group reports are a series of text files reporting on the data flow and user needs of individual projects within the CfA and a group representing staff who work on external projects. The User

Needs survey is a pair of Access Databases one holding the responses of CfA staff to 32 questions regarding their work practice and IS, and the other containing the forms used to administer the questionnaire. The questionnaire is anonymous but there is data on which staff responded to the questionnaire and their comments on it. The database with the data also contains queries used in analysis. Most of the further analysis of the data is held in Excel spreadsheets. The directory also holds the User Needs report and its appendices.

Sectoral Practice contains two separate strands, the Literature Review and the Fact Finding Visits. The first of these is open access, the availability of the second has yet to be negotiated. The Sectoral Practice report mentions no units specifically given that there could be commercially sensitive information that was discussed during our visits. Our archive reports, however, record each visit separately, access to this section of the archive may therefore be restricted. The literature review consists of a literature search held in a purpose designed Access database, and a literature review, in the form of a report which synthesises the material reviewed under thematic headings. There are 1500 items in the literature search and 97 items reviewed. The directory also contains the Sectoral Practice report and its appendices.

The Project Management directory contains all information relating to project management, with the exception of correspondence (see above). This includes the original project initiation and planning documentation from the summer 2002. Also its revision from the variation to the project plan in spring 2003. It contains agendas and minutes of meetings (but not workshops), all of the product descriptions for the project as well as other documentation including the risk log and issues log.

The Presentations directory contains information on presentations, both internal and external, seminars and workshops. This material contains many of the sources for our broader discussions of issues. The project organised one seminar, to which key players in EH and in the profession were invited. There were four presentations in the morning and facilitated discussion in the afternoon. Presentations, preparation documents and results are held in the files. Five conference papers were presented, presentations with speakers notes and formal text are held for these. Four internal presentations are similarly held. There were four workshops held, one for commissions staff and three for the project team. Preparation documents and reports on outcomes are held in the directory.

The Data Modelling directory contains all of the data models described in Appendix 4, as well as supporting documentation.

Appendix 4

Data Modelling in an Archaeological Context

Report for the Revelation project

Version 1.2.2

November 2003

Paul Cripps

Data Modelling in an Archaeological Context

Prepared by Paul Cripps, Centre for Archaeology, November 2003.

Introduction

Data modelling is the name given to a range of associated processes that are used to help us understand the nature of the data we use and how we use it. This modelling process then informs the design or improvement of data handling systems.

The modelling undertaken for the revelation project has three components. The first, undertaken primarily by Manny Lopez, looked at the processes which involve data and how data moves through a system in terms of its use and storage. This produced a number of models illustrating the complexity of data flow within the CfA.

The second component looks at the specific nature of some of the archaeological data being processed: This component is called Entity-Relationship modelling and is based on the work of Codd (1975) who first proposed a mathematical model for data structure. The basic premise is that data can be modelled in terms of entities or objects and the relationships between them. In terms of the Revelation project, the Entity-Relationship modelling process is being applied to the data gathering phase of the work of CfA, specifically archaeological fieldwork as defined by the (draft) fieldwork recording system currently employed within CfA. In archaeological fieldwork as undertaken by the CfA, data is gathered using pro-forma context sheets. These context sheets hold attribute data regarding archaeological contexts but also identifiers which are used to relate forms to one another. By examining the complete set of forms associated with the recording system, it will be possible to build a model of the recording system which stores data at the most appropriate point.

A third component applied a data flow type approach to the fieldwork recording system. The entire Revelation team attended a workshop where the set of forms used by the fieldwork recording system was examined in order to identify where data is transferred. The data flow diagram produced was then taken into the field at the Barrow Clump excavation and compared to actual field practice by Vicky Crosby.

Data flow modelling of the CfA

Background

This document describes the Data Flow Modelling undertaken Autumn 2002 – Spring 2003 which examined the processes undertaken by CfA staff as well as the data involved in these processes and means of storage.

The aim of this exercise was to support the Review of Existing Systems by providing a structured assessment of systems and pathways for data transfer and representing this in a graphical form.

Process

The bulk of the work was undertaken by Manny Lopez using the Data Flow elements from a system called SSADM (Structured Systems Analysis and Design Methodology). Associated with this system is a software application which allows the creation of complex Data Flow Diagrams according to the systems development rules defined in SSADM. The source data used for the production of these diagrams came from a series of interviews with CfA staff. DFDs were created to represent the following:

- Archaeologists team
- Administration team
- Archives team
- Conservation team
- Environmental team
- Information Systems team
- Geophysics team
- Dating team
- Raunds project team

The DFDs are all Current Physical DFDs, as per the defined methodology.

In order to make the diagrams accessible, each was exported from SSADM and converted to SVG format, making it possible to browse even the largest of models, zooming in and out and searching for keywords using a web browser. Selected parts of the models also were used to create graphics to illustrate specific aspects of the systems.

Discussion

By representing CfA processes using data flow modelling techniques, it was possible to gain an overview of how data actually moves within the CfA. An example of a DFD is given in Figure 1, an extract from the IS team DFD, which shows data (site records) moving from an external entity (the Project/Site director, not a member of the IS team), being processed by a member of the IS team resulting in the data being stored in a digital store (Delilah).

It was possible to note that duplication of effort existed with double-handling of data and that this is often associated with the use of additional systems which provide specific functionality. Figure 2 shows how site data is entered twice, once into Delilah and once into a site specific database.

In other parts of the models, this transfer was shown as being an sequential data entry followed by export and import rather than two separate data entry processes, but this still results in data residing in two separate systems with potential for divergence. One particular process which was identified is just such an export/import process whereby a member of the IS team transfers data between systems for the conservation team (Figure 3). In circumstances such as this, it is much more difficult to maintain the integrity of data.

The models also revealed the complexity of our interactions with other parts of English Heritage and the rest of the sector. Not only does a significant amount of our data come from sources external to CfA, but we also supply data to a number of external organisations for different purposes (see Figure 4).

Conclusions

It became apparent during the process or working with the DFDs that a good deal of variation is evident in the processes described due mainly to the way in which data was gathered: By asking individual staff members about the processes they undertake, each part of the model was biased by the responses of the staff member. Thus some parts of the model look ‘neater’ than others. There was also a varying degree of detail present, with some processes being split into every component part while others contained (and hid) more processes.

It was certainly possible to identify more than one instance where data flow is not supporting the work of the CfA, with processes existing solely to prop-up the flow of data and taking resources from essential processes. It was also possible to note how the use of ad-hoc systems supports the more formal systems of the CfA.

Data-flow modelling of the CfA Recording System

Background

This document describes the results of the data flow modelling process undertaken Summer 2003 which examined the flow of the data used within the CfA Fieldwork Recording System. This work was based on work undertaken on the paper-based recording system and the output from the second Revelation Away Day at Portsmouth Museum.

The aim of this exercise was to identify how data moves through the recording system whilst on-site in order that a Data Flow Diagram could be constructed in a similar fashion to those created for the individual teams and projects. This Data Flow Diagram was then taken out on site and compared with actual field practice.

Process

The individual forms that comprise the recording forms collection, as described in the (draft) recording manual, were used as a starting point, the forms being the physical manifestation of the recording system. The forms were distributed amongst the group and the following questions were asked of each form:

1. Is this form used on every site?
2. Who is responsible?
3. Where can you find it?
4. How long is it active?
5. What data on the form comes from another form? Which form?
6. What data on the form goes to another form? Which form?
7. Who adds data?
8. Who takes data?
9. Still required in the future?

The final stage of the process was to actually build the model at the workshop based on the answers to questions 5 and 6 and discussions regarding the answers. This model was recorded by Tom Cromwell and subsequently digitised. The answers to the remaining questions were stored digitally and used as reference.

This is different from the process used for the other DFDs: The other process involved interviewing each person from each team individually then joining the individual elements together. This resulted in a highly detailed, very complicated model, incorporating every nuance and oddity in the data flow within the CfA, often representing personal approaches rather than agreed CfA systems. The workshop approach focussed on an agreed system and encouraged debate and the resolving of issues within the team *before* the creation of the model, hence the model is much less detailed and far more generic. The resultant DFD is somewhere between a Current Physical and Logical DFD, using SSADM terminology.

The model also includes some real-world objects, observations of which are recorded on forms. This include Finds, Samples, and Photographs. Also, Drawings were included. Drawings represent both a physical object and a recording form, with data being recorded from observation of the world and copied from forms.

This model was taken out to the Barrow Clump excavation where the relationships identified in the model could be compared to actual field practice. The model was then augmented with detail from the actual fieldwork practice.

Discussion

The discussion is divided into two sections.

INITIAL OBSERVATIONS ON THE FIELDWORK DFD

For the most part, the Data Flow Diagram is a fairly intuitive output. It represents graphically the way in which data moves around the fieldwork system as understood by the workshop team.

It appears that the Working Matrix currently acts as an output, drawing data from a number of sources but not being used as a data source by any other forms. The Record Number Allocation form can be seen to be the source of all indices, supplying a number of other forms with data but not receiving any from other forms.

Two of the physical entities created at the workshop did not feature in any of the data flows as described. These were Photographs and Digital Photographs. While theoretical discussions at the workshop led to their inclusion, the fact that they were not then identified as participating in any of the data flows is indicative of the nature of photographs as used in the context of archaeological fieldwork. Where traditional film-based cameras are used on site, it has been observed elsewhere in the Revelation project that the site staff rarely see the photographs they take, let alone derive any information from them. This is due to the way in which unprocessed films are sent 'back-to-base' for processing and copies of the processed images are not sent back to site. While an interesting observation, the fact that digital photographs also did not feature in any data flows, despite being usually instantly accessible on the camera itself, suggests that information is not drawn from the photographs on-site. This confirms their nature as record photographs, captured for use once the excavation itself is complete.

Indeed, the inclusion of physical entities in the model requires further discussion. The model produced was based on the set of forms defined in the Recording Manual. It was noted, however, that at various points in the fieldwork process, the data was relayed between the various forms by means of physical objects, notably finds and samples. Drawings are an obvious case of a physical object which carries data, as are the forms themselves, but in this abstraction, the workshop group determined it was necessary to include Finds and Samples within the model.

OBSERVATIONS ON THE FIELDWORK DFD, POST-FIELDWORK

The trials at Barrow Clump were able to check identified flows and provide more information about how data actually moves around in the field. This process highlighted the way in which the identified data flows are often bi-directional: This ensures reciprocity between records, with key data copied across from one form to the other and vice-versa. It was apparent however, that with the current paper based system, access to the most appropriate data is an issue resulting in detailed interpretations of excavated features not being used in favour of the initial statement for finds recording and subsequent analysis. This issue did not arise at the workshop, perhaps due to the lack of a finds specialist in the workshop group, but it is a good demonstration of the importance of a shared understanding of the data. Data needs to be carefully defined within any proposed system and these definitions need to be agreed upon and conformed to in order that a piece of information is gathered with a clear understanding of its source, limitations, its lifetime within the archaeological process and the uses to which it will be put during this lifetime.

A major point to arise from the field-work experience was how much additional data transfer occurs on site. While it would appear from the model that data flows directly from one form to another, it is often the case that intermediary steps are necessitated by external factors and facilitated, as described previously, by physical objects and also interactions with other site staff (see the more detailed discussion of this issues in the report on Field Practice). In this way, data is transferred from the indices to record forms via, for example, labels on site, the site grid pegs, personal communication between site staff and human memory. Some of these additional data flows can be incorporated into a digital field recording system where appropriate (e.g. printing labels for finds bags) but the real benefit of a digital field recording system would be the improvement in shared access to appropriate and correct data thus reducing the need for less formal means of transferring data from inaccessible repositories (e.g. a context index kept in a site hut a mile from site).

Looking at the position of the working matrix in the data flow model, it appears that the matrix is an end product, with no data flowing from it. The experience in the field was very different. The process of producing the matrix informed the stratigraphic interpretation recorded on the context record forms. This highlights the need for a working matrix tool in any digital field recording system. This also highlights how a working matrix is different from the kind of final matrices drawn up during post-ex.

Conclusions

The main conclusion to be drawn from this process is that despite there being an agreed system, there is also a collection of supporting, undocumented procedures used by site staff in the process of excavation as highlighted by the field experience. These procedures result in data flows existing outside of the system as defined which are not made explicit and can vary. This also results in data being collected for one purpose and used for an entirely inappropriate purpose.

A secondary conclusion is that the sort of idealised data flows produced by examining the physical manifestation of the recording system, i.e. the recording forms, do not take into account how much the physical locations of data stores and access to them affects the flow of data on an archaeological site. It may be that a particular attribute should be used for subsequent work in favour of another, but if the former is inaccessible while the latter is accessible, in order that work can proceed, the less favourable data may well be propagated. It should also be noted that there is no formal mechanism for updates to be cascaded around the system.

The work on this DFD and the report on the fieldwork experience have shown that the data flow exhibited in our current fieldwork practices are the result of a number of contributory factors which have also been noted in other Revelation reports and do not represent an ideal situation. These factors include lack of access to data and uncertainty regarding quality of data. It is highly likely that this is related to the work intensive ways in which we currently manage our predominantly paper-based fieldwork data and the lack of integration with digital data (e.g. coordinates generated using a total station, transcribed, copied onto finds bags and context records and finally re-entered onto Delilah).

Entity-Relationship modelling of the CfA Recording System

Background

This document describes the results of the entity relationship modelling process undertaken Summer 2003 which examined the nature of the data used within the CfA Fieldwork Recording System. The modelling process made use of the third draft of the recording manual (June 2001). The aim of this exercise was to look in detail at attributes, their data types and domains in order to support the Data Flow modelling exercise, given that ER modelling ignores data flow and DF modelling ignores the nature of the data.

The principle behind ER modelling is that by using this approach, a data structure can be gleaned from the data itself without the need to impose an arbitrary structure: An understanding of the inherent structure of the data can be gleaned from a dataset by using a form of mathematical set theory and a more recent object-oriented approach. This structure is based around entities to which data can be attached, representing real-world objects, and relationships which describe how these objects relate to each other. In Figure 5, Cut and Deposit entities are shown with examples of two relationships between them and their cardinality. This diagram summarises the following observations:

- A Cut may contain one or many Deposits.
- A Deposit may be contained within a Cut.
- A Deposit can be cut by none, one or many Cuts.
- A Cut must cut at least one Deposit.

In the diagram, a dashed line represents an optional relationship while a solid line represents a compulsory relationship i.e. a Cut must be involved in at least one cuts/cut by relationship.

Process

The methodology used involved starting from the (agreed) paper-based system and each attribute held on each form. The next step was to define sensible entities to which this data could be most sensibly attached. For the most part, these represent the paper forms although it was necessary to create new entities to represent some relationships and also to rationalise some aspects of the recording system.

Each form was inspected in turn and attributes recorded on each form noted. The data type for each attribute was also recorded. Within the paper-based system, a numeric index was applied to attributes held in Delilah. This index is based on the concept of records and tables and as such, attribute number 1 is used to store Context Number, Drawing Record Number, Record Number, Sample Number and Small Find Number at different points in the system, i.e. the primary record number. This represents the way in which all record numbers are assigned from a single sequence and represent UIDs within the system (i.e. no two records of any type will ever share a primary record number). Additionally, many attributes on the forms do not have a corresponding attribute number. This represents attributes added to the system since the system was set-up or attributes which are recorded on the paper forms only and have no subsequent use (ie they do not have corresponding attributes in Delilah). As such, the attribute numbers were recorded but will serve no future part in the fieldwork recording system.

The following questions were also asked of each attribute:

- **Mandatory?** Must the field be completed or can it be left blank. UIDs will always be mandatory.
- **Glossary?** Does the attribute require a list of controlled terms or can anything be entered. Essential for categorical data which will subsequently be used to group or search records.
- **Post-ex use only?** Is this data captured in the field as part of the field recording process or is it added to the form at a later date.

Having defined attribute lists for discernible entities, the model could then be constructed. Where relationships between objects exist, these were recorded and attributes were moved where appropriate to be stored at the most logical position within the model. Where attributes served other purposes in the paper-based system (such as duplication across forms as a cross-check) these were removed.

Discussion

Of the 22 forms looked at, 18 fed directly into the ER model, as listed in Table1. 154 unique attributes were recorded with 328 instances of these attributes across the 18 forms. Observations regarding these forms and their data fields is held in an Access database (*recsys.mdb*) to facilitate building the model diagrams and analysis.

Table 1; Forms used in the modelling process

Long Name	***Short Name***	***Purpose***
Abbreviated small finds form	ABVSF	Minimal recording for individual objects; not a summary of small finds form (SF)
Record number allocation form	ALLOC	Allocation of UIDs
Built structures record form	BLTST	Recording built structure type contexts
Bulk finds form	BULK	Recording bulk finds from contexts
Box index form	BXIND	Index of boxes and their contents created at end of fieldwork as finds are packaged
Context finds register	CFR	Checklist of finds by context
Context index form	CTXIND	Assigning numbers to observed contexts ready to record the context
Deposit and cut form	DEPCUT	Recording deposits and cuts
Drawing index form	DRAWIND	Index of drawings created on site
Film index form	FILMIND	Provides an index to the films and cameras on site.
Photographic record	PHOTO2	Index of photos taken on site
Pottery record form	POT	Detailed recording of pottery by context completed by pot specialist.
Sample index form	SAMIND	Index of samples
Sample record form	SAMREC	Detailed record of each sample

Sample evaluation form	SAMEV	Records detail regarding processing of samples; method, quantities used, contents,
Small find form	SF	Recording individual objects
Skeleton form	SKEL	Recording skeleton type contexts
Timber structure record form	TIMST	Records timber structures, generally by a specialist

The following forms were not included in the data model for the reasons given:

Table 2; Forms not used in the modelling process

Long Name	***Short Name***	***Reason***
Small find index form	SFIND	Form not present. No information available.
Small find wood form	SFWOOD	Form not present. No information available.
Record numbers used form	USED	This form does not hold data as such, simply collates a subset of all context numbers available (i.e. those used). As such, the functionality provided by this form will be replaced in a digital or hybrid system by a query of the record allocation table.
Working matrix form	WKMTRX	It is anticipated that the matrix will not be used as a static output. Rather the matrix will act as a dynamic view of any data stored in the rest of the database. Indeed, it was reported that field staff generally use plain paper for matrix generation on site as the process is an iterative, creative one as the matrix is used to understand the site formation processes.
Skeleton Processing Checklist	SPF	This form simply records on a single form information that could be recorded elsewhere on appropriate forms i.e. small finds, samples and bulk finds data relating to skeletons. It is also not described in any way.

For the most part, each of the forms represented the recording of a discreet archaeological concept (e.g. a deposit). Also present were index forms used to manage the allocation of record numbers. Entities were created to represent the archaeological conceptual objects and additional entities were created to represent the indices: The retention of the index entities related to the main recording entities by a 1:1 relationship allows the creation of records to be split into a two stage process, the first simply consisting of registration while the second pass completing the detailed records. This then enables the indices to be used to control data entry of other data (e.g. environmental data) without necessarily having to enter complete context records from the outset. Of course, this is only applicable in a system which involves some paper-based recording: As soon as the system is fully digital, the need for immediate data entry to facilitate subsequent tasks becomes irrelevant as all data will be entered onto a live database accessible to all. Indeed, it would be possible to model the system without any index entities, but this would then require any data entered to comply with *all* enforced relationships and validation rules from the outset, removing the scope for an initial registration phase (unless relationships are not enforced and the validation procedures are turned off, which is entirely counter-productive).

The complex nature of archaeological data is also apparent as is the way the way in which a relational system can be used to successfully model such complexity. The ‘skeleton’ entity, for example, on one level is treated as a deposit type context: Skeletons are treated as a fill of a grave on one hand and yet, once excavated, they are numbered and treated like finds. As such, the skeleton is the only deposit on site to ever have it's current location recorded (the current location for most deposits being the spoil heap post-excavation).

Some aspects of the recording system require clarification.:

- **Skeletons and contexts relationship –** The context in which the skeleton was found cannot be recorded on the skeleton form as there is no field for such data. As such, a relationship has been created based on the skeleton being a fill of a cut; the same *filled by/fill of* relationship exhibited between cuts and other deposits.
- **Deposits and deposits relationship –** take the situation where a lens of material is wholly contained by another context. This is an unusual situation but not impossible. In this case, one deposit is within another. As such, an additional *within* relationship has been created relating deposits to deposits. In the context of the first clarification, this relationship can also be seen to exist between deposits and skeletons (assuming a skeleton can be found within a deposit as well as between deposits).
- **Composite type contexts –** These appear to be a form of rudimentary grouping to allow contexts to be grouped in the field for ease of manipulation. This appears to be purely for simplicity and such groupings are not recorded explicitly to be retained as interpretations for later parts of the archaeological process. The specific use of composite contexts is not adequately described and requires clarification.
- **Sample Evaluation attributes –** As this form is only ever completed by a specialist there is no detail in the manual regarding the attributes in use.

The first point to notice about the model is that parts of it appear to represent more traditional hierarchical relationships between entities while other parts appear to represent more complex object-oriented relationships. This is entirely to be expected: While it is clear to see how projects, site-subdivisions and contexts form a hierarchical tree with projects at the top and contexts at the bottom, the relationships between different types of context or object are more complex (e.g. a single drawing can be of many contexts; a single context can appear in many drawings).

The second point to note is that there are a number of new entities which do not appear on the paper forms. As stated previously, new entities are often needed to model specific relationships, notably many-to-many type relationships, but also where a single attribute is used to hold a range of data (e.g. the sent to/date attribute is replaced by current/previous location entities). While it is not clear how the sent to/date attribute functioned as a locational trail for objects in the paper-based system, it is abundantly clear that such a simplification of the real-world situation will not work in a digital system. *(Further clarification of this issue suggests that the sent to/date attribute is only used for the first move from field to base; subsequent moves are managed using the LabFile system at Fort Cumberland).*

The final thing to note is that the ER model is exactly that: A model. It can be used as the basis for subsequent database design and also as a spur to encourage discussion of theoretical issues and through such activity, be refined. By defining our data structure in terms of objects and relationships between objects, every object and its relationships must be explicitly defined: Unlike a paper-based system, any implementation of the ER model will only allow the storage of data that complies with the defined structure. In other words, if the model states (as this one does) that a deposit can only fill one cut, it will be impossible to record a deposit as the fill of two cuts. The theoretical argument would argue that such a deposit is not a fill at all, but a deposit stratigraphically above both cuts and physically above the fills of both cuts.

One area that was not resolved was that of composite contexts. The way in which this concept appears to be used is as a means of associating contexts to form a group. Given that there is no other means of grouping contexts and this topic is not described in the manual, this concept has been left out of the model, the nature of the process meaning that this concept can be included once clarified.

Indeed, the model as it stands currently represents the data capture phase of archaeology, being based on the fieldwork recording system. While some aspects of post-ex work are represented, this is generally where that part of the post-ex process sometimes happens on site and hence a form has been incorporated into the fieldwork recording system. Examples of traditionally post-ex activities which can be undertaken on site as part of the excavation process include environmental sampling and finds processing. Grouping and phasing are examples of activities which currently do not occur on site and hence there are no entities to represent these concepts in this model.

Conclusions

The modelling process was an informative and useful process. Having now modelled two such recording systems (for CfA and Wessex Archaeology), I would argue that the nature of archaeological data as used within such systems is pretty consistent, with very similar theoretical underpinnings.

A case in point would be the skeleton object as used within the CfA recording system. On one level, a skeleton is treated as a deposit (a fill of a grave cut) and on another, it is treated as a find whose location and movements are recorded. The reasoning for this is clear: The skeleton *is* a deposit type context but unlike other deposits, has a presence as physical remains; the bones themselves. The ER model accounts for these relationships and this duality of concepts can successfully be implemented.

There are, however, limitations to this approach. By focussing on the nature of the data to the detriment of the processes that generate it, we end up with a ‘perfect’ scenario, one where the processes that surround the system result in only appropriate data being stored. Even with draconian validation routines that ensure data entered is appropriate for a given attribute, there is still a requirement for interpretation: Many attributes are not black and white and depend on expertise to determine appropriate values (e.g. pottery typologies). For example, if context type is recorded as an attribute of the context entity, there is no scope for differences of opinion or subsequent reinterpretation *within the system.* Supporting information regarding why the change occurred has no place in the data structure. The old record is replaced by the new and even with transaction log files which can show which user performed which system action when, it is not apparent to system users that there has been a modification to the interpretation unless one rolls back the database to a previous state.

This problem has been resolved in a number of situations by the use of what is referred to as Event-based modelling: This is an object-oriented approach, using an Entity-Relationship modelling process, in which the core entities represent events, with additional entities representing real-world entities to which base data can be attached. Thus we have archaeological entities such as contexts with core attributes (e.g. inclusions, descriptions, measurements, stratigraphic relationships, etc) and event entities representing the processes in which these archaeological entities are involved in (e.g. context type *interpretation,* context *spot-dating*, feature *phasing*). In this way, both the data itself and data regarding any data transformations are recorded. The ExeGesIS SMR system has an event-based model at its core, as does the Stonehenge Condition Survey database and an event-based model is also behind the CIDOC Conceptual Reference Model and its implementations. This approach therefore, allows concepts such as interpretation and the discourse surrounding the interpretive process, the heart of the archaeological process, to be stored within the system data structure rather than being completely external to the system and thus the archive, this vital information only existing in the memories and notes of the participants.

General Discussion

The data modelling exercises conducted as part of the Revelation project have highlighted a number of key points including some from other parts of the Revelation project. The User Needs survey highlighted both access to data and quality of data as being important to users while the Review of Existing Systems highlighted duplication of data as being an issue. The modelling exercises found evidence to support this.

While undertaking the modelling exercises, innovative solutions to particular problems were also observed. Where a particular problem exists, CfA staff consistently produce solutions using existing resources (e.g. the CAD-based phasing system for the Raunds Roman project or the extensions to the RRAD itself). While this may not be an ideal way to plan for structured systems development, the ability to use existing resources in innovative ways to produce results is something that should be helped not hindered by any proposed Revelation system.

Duplication of effort

In both the Data Flow modelling processes and the Entity-Relationship modelling process, it was clear to see that there is duplication of data and data processing within the current CfA systems. This is associated with the number of ad-hoc systems in place for various purposes and the lack of dynamic linkages between systems. It is also related to the use of paper as the primary recording medium and the necessary manual cross-transcribing, checking and data entry, often into multiple systems. As a result, CfA staff often spend time processing data to support the system rather than the system being capable of supporting the work of CfA staff.

Duplication of data

There are two issues arising from the modelling process. Firstly, the Data Flow modelling process demonstrated that as a matter of course, data is transferred between systems. Secondly, the Entity-Relationship modelling process demonstrated that there is redundancy built into the field-recording system.

The transfer of data between systems is an operation that will undoubtedly need to be performed from time to time with any set of systems. It should not, however, be undertaken without careful consideration of version control and other data integrity issues. Where it is deemed necessary to transfer data, there must be a means of reconciling datasets afterwards and this should not be a manual process due to the amount of time it takes to manually reconcile diverged datasets.

The redundancy built into the field-recording forms is a product of the development of the paper-based recording system and fulfils a number of roles, mainly to provide access to information from multiple points and to act as a cross-check or validation routine. With a paper-based data collection process, these functions are necessary, but it must be recognised that such duplication should not be carried over to digital systems, where it is the data structure and associated automated validation routines that ensure the validity of data entered. It may seem counter-intuitive to those used to working with paper-based systems, but redundancy within an information system can undermine data quality.

Quality of data

It is clear from other reports for the Revelation project that users rate quality of data very highly. This is evident in the number of processes identified within the CfA that involve ‘checking’ various forms of data post data-entry. This is a very expensive means of maintaining data quality in terms of CfA resources. Attempts have been made in the past to support this labour intensive process with automated validation and error-checking routines: Delilah (the current CfA digital context recording system) has a number of automated routines to ensure the validity of data (e.g. stratigraphic cross-checks) as does the RRAD (Raunds Roman Analytical Database, a relational database) but the majority of the ad-hoc systems in place have little or nothing to ensure data integrity.

Contemporary thinking about quality emphasises quality systems over quality checks. Quality checks are a last step to ensure the rest of the system is working. We will need to include quality systems in our modelling in the form of feedback and audit trails plus accessible context-sensitive help to ensure users are aware of their environment.

Types of data

Being based on the existing systems and processes within the CfA, the modelling processes undertaken only took into account the types of data currently in use. The Entity-Relationship modelling process dealt solely with textual/numeric attributes and photographic images used by the field-recording system while the Data Flow modelling process did not look explicitly at the nature of the data being processed. Other types of data used within areas of the CfA not studied in this exercise and other types of data that may well be useful need to be assessed for incorporation into any future system.

Innovative solutions vs. Ad-hoc systems

During the background research to the modelling exercises a number of innovative solutions to particular problems were observed. These often represent necessary development where existing systems have been found not to support the task in hand and include specialist databases and custom-built software applications. The use of a digital indexing component developed by Vicky Crosby as part of the recording system at Barrow Clump, for example, had significant positive benefits for the site staff in terms of easing access to data and reducing unnecessary transcription of data. The construction of databases or entire applications to satisfy particular requirements of projects has been key to systems development at the CfA and the experience gained through such developments, both the successes and failures, must not be lost.

Conversely, a number of systems have been less successful with some projects suffering from less than ideal ad-hoc systems development which have had a negative impact on the project. These system problems are not evident in the data models themselves, simply that the problematic systems exist as part of CfA processes. While some of the ad-hoc systems are undoubtedly useful and well-developed, it is the production of such systems that has led to the complexity of systems present within the CfA today as exemplified by the data flow diagrams. The specific functionality of these systems rather than the systems themselves should be retained. For example, a system which can produce phase diagrams. This can be achieved by means of a straightforward thematic mapping module (almost certainly a GIS interface to the data) rather than the current assemblage of custom written scripts.

Revelation needs to provide a framework in which development takes place within. With a defined conceptual model, it will be possible for qualified users to extend the data model where appropriate. In this way, the system is not rigid and inflexible, forcing users to develop ‘bolt-ons’ which compensate for system deficiencies yet remain external to the main system (eg the custom project databases), but instead provides a means of developing *within* the system. The proposed modular approach means that the system is flexible and rather than ad-hoc development being necessary to accomplish particular tasks, the resources normally devoted to such development can be used to develop reusable system components which comply with the overall schema.

A conceptual approach

Much work has been done in the world of IS and KOS on ways of describing knowledge domains using a common structured language. By describing a domain in this way it is possible to create a formal conceptual model or ontology.

A formal conceptual model of CfA data and processes is the next step on the path to Revelation. This could be accomplished by continuing the SSADM approach and developing the DFDs and ERDs. However, we intend to prepare a formal ontology of the CfA domain using the CIDOC CRM. As well as being the newly adopted standard for systems development within EH, a CRM based approach would afford greater interoperability with external systems also developed using the CIDOC CRM due to there being a shared conceptual base. This will result in an event-based, object-oriented system capable of supporting the archaeological process and its evolution.

Conclusions

To conclude then, the data modelling exercises have provided a more explicit view of the issues surrounding both the nature of archaeological data and the processes used by the CfA to manipulate this data. Observations on the data models support those made in other Revelation reports, particularly the Review of Existing Systems and the User Needs Report.

Further to this, it is clear that the system should be made to support more closely the working practices of the CfA. This includes such things as designing and implementing schemas for data collection with associated validation routines to ensure data is fit for purpose. Access to data is vital and the current practice of multiple versions of the same data existing for different purposes needs to be eliminated in favour of systems capable of support a data-warehousing type approach, with all CfA data stored centrally[1] and available to all.

Any proposed transfer to more digital systems must take into account the extensive work already undertaken on the CfA recording system, Delilah, and other existing systems but must use structured systems development and a strong conceptual basis to avoid carrying over inconsistencies inherent in, and masked, by paper-based systems.

The idea of removing inconsistencies, in this context, does not mean imposing a rigid, monolithic, one size fits no-one solution; rather, any proposed solution must be flexible and extensible. Any system developed must be able to handle diversity of thought and multiplicity of interpretation explicitly.

Acknowledgements

Thanks are due to all CfA staff who participated in this exercise. Particular thanks are due to Manny Lopez for producing the DFDs representing CfA work processes and to Vicky Crosby for her work with the field-recording DFD and information on the RRAD and Delilah.

Glossary

Attribute – a piece of information relating to an object or concept; a 'field' in a database table; a column in a spreadsheet.

Cardinality – the degree to which objects can participate in relationships i.e. one to one, one to many, many to one or many to many.

Domain – a subject area, used in IS/KOS to refer to an area of application of data.

DFD – Data Flow Diagram

Entity – an object/concept to which data relates; in database terms, entities are represented as tables in which data is stored.

ERD – Entity-Relationship Diagram

IS – Information Systems

KOS – Knowledge Organisation Systems

Ontology – a formal description of knowledge relating to a particular domain.

Relationship – a formal description of how two entities relate to each other, including cardinality.

SSADM – Structure Systems Analysis and Design Method

SVG – Scalable Vector Graphics, an XML based graphics format for storing two-dimensional vector and mixed vector/raster graphics

UID – Universal Identifier

[1] This does not imply a server-client model or any other network model, simply that data is stored 'somewhere' and that it is available from any specified access point (eg from within a GIS application or a statistical package).

Figures

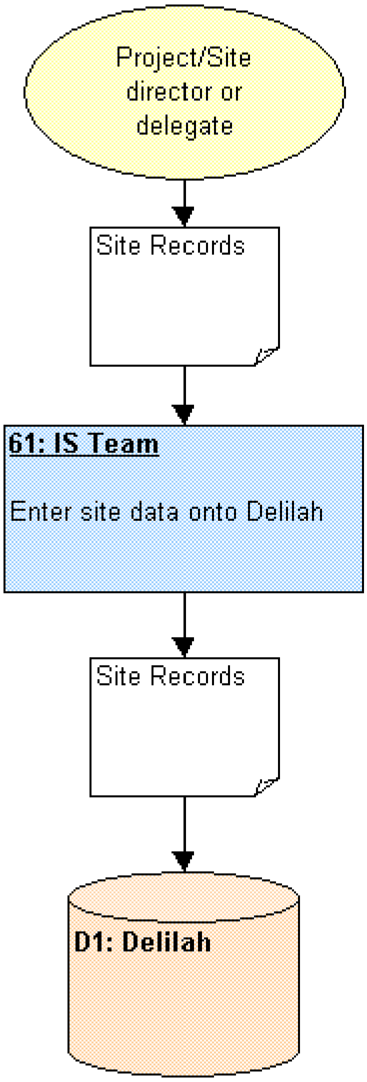


Figure 1; A CfA process represented as a Data Flow Diagram.

The numbers and letters assigned to each element in the diagram are those assigned by the SSADM system to identify processes and data stores.

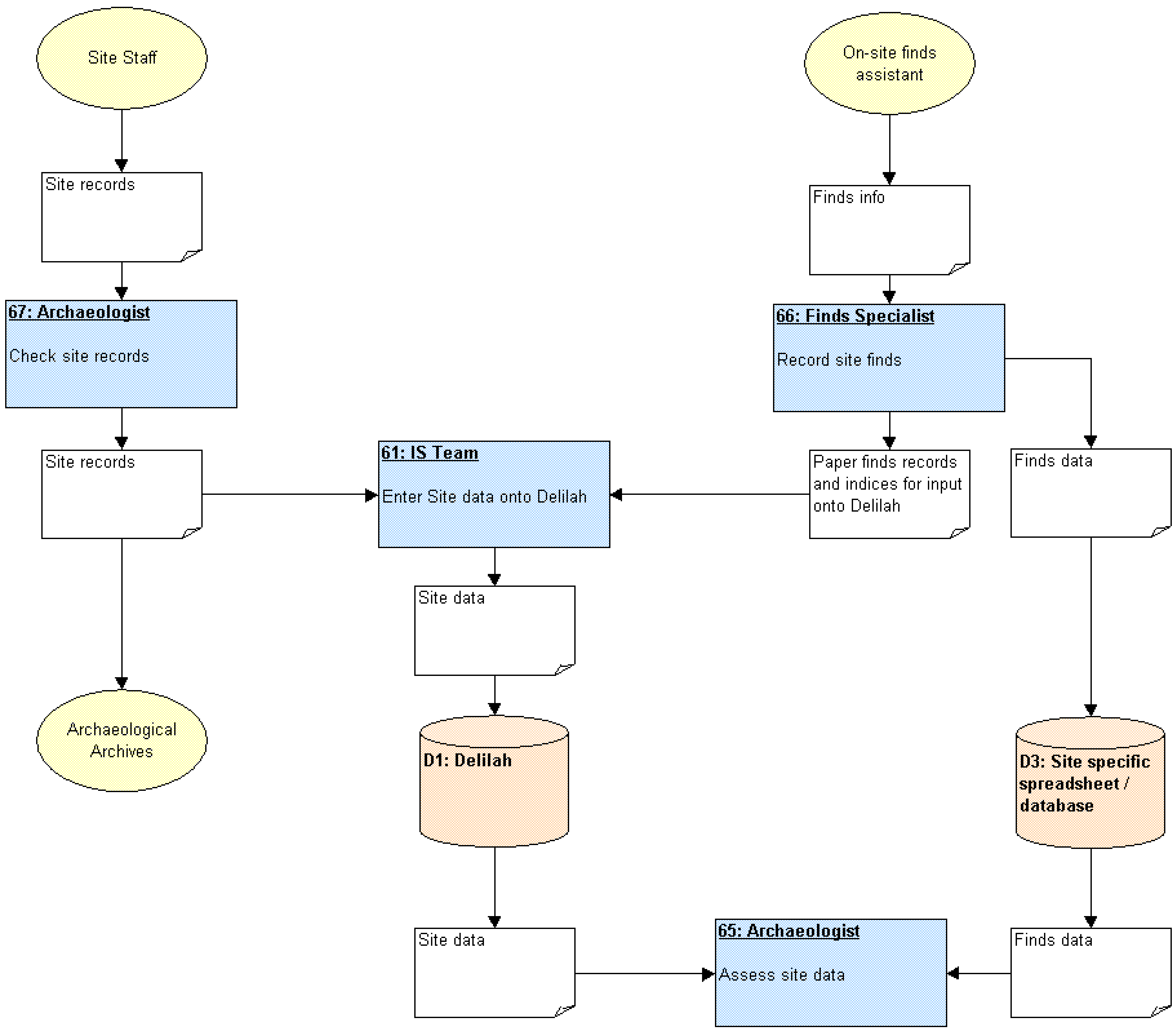


Figure 2; Example of duplication of effort identified.

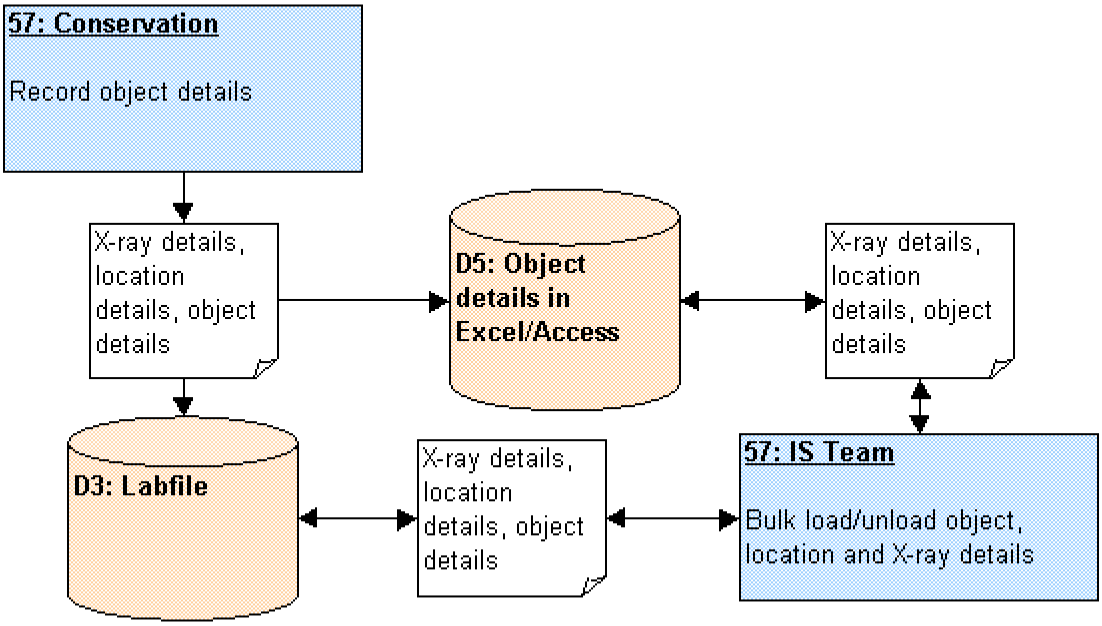


Figure 3; an example of data transfer between systems.

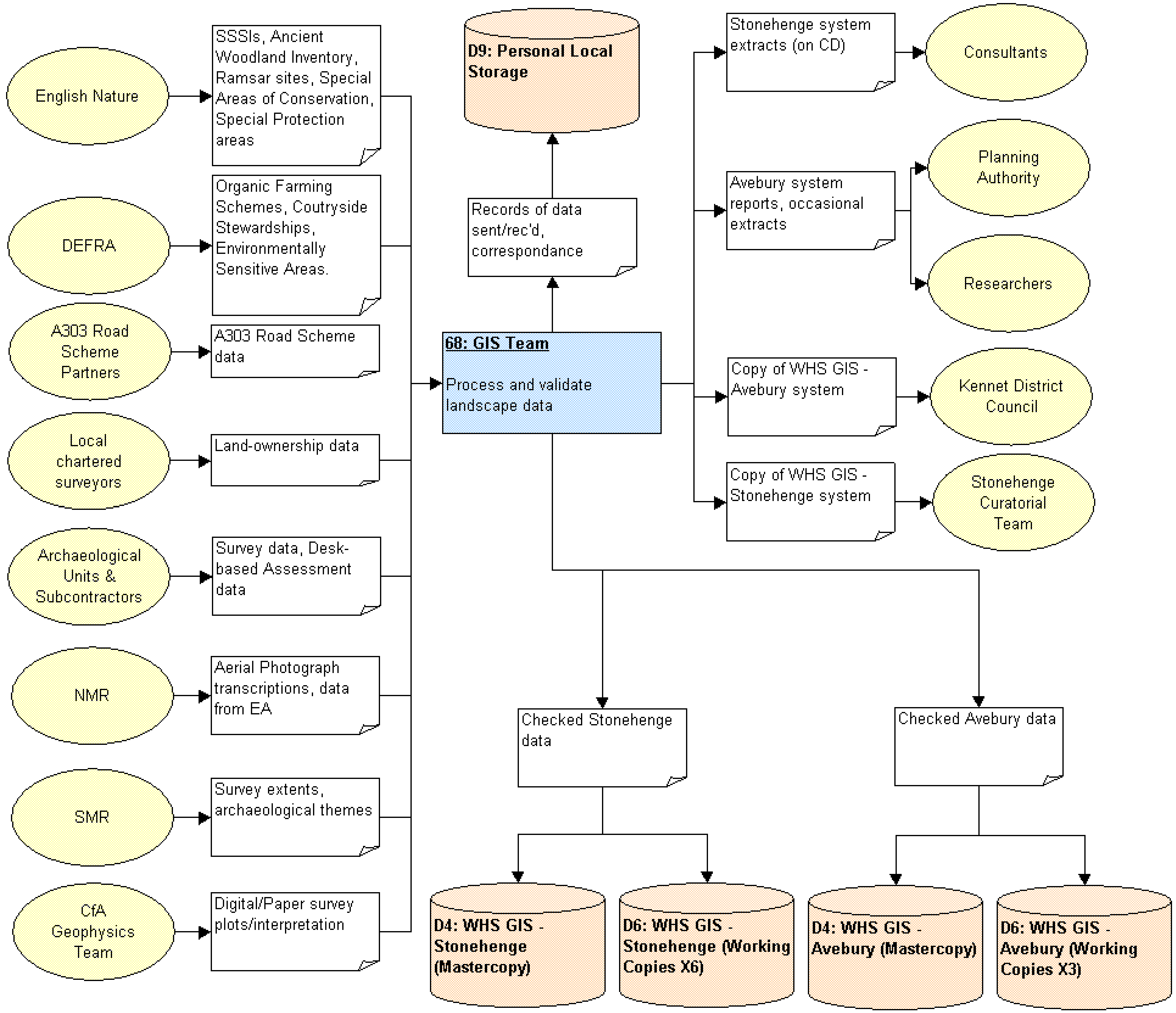


Figure 4; an example of extensive links to external organisations

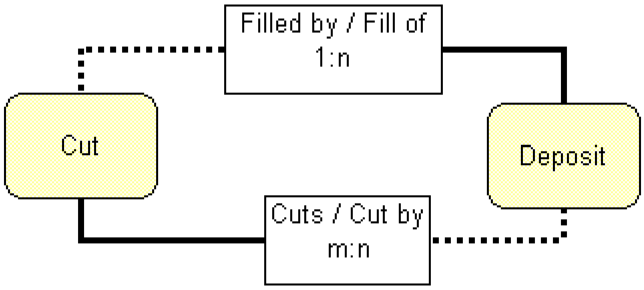


Figure 5; Example of a simple Entity-Relationship diagram

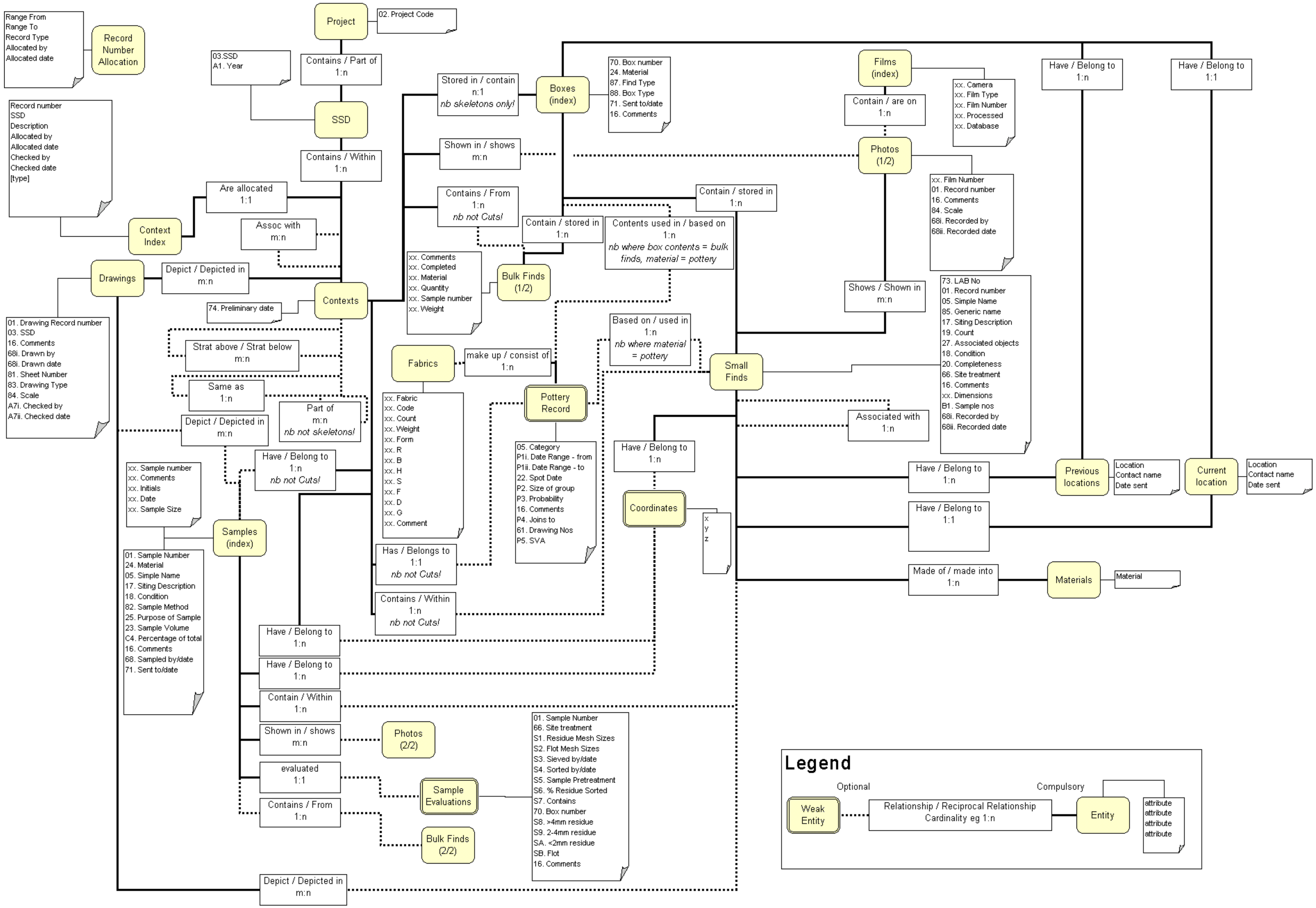


Figure 6; Entity-Relationship diagram for the recording system – main view

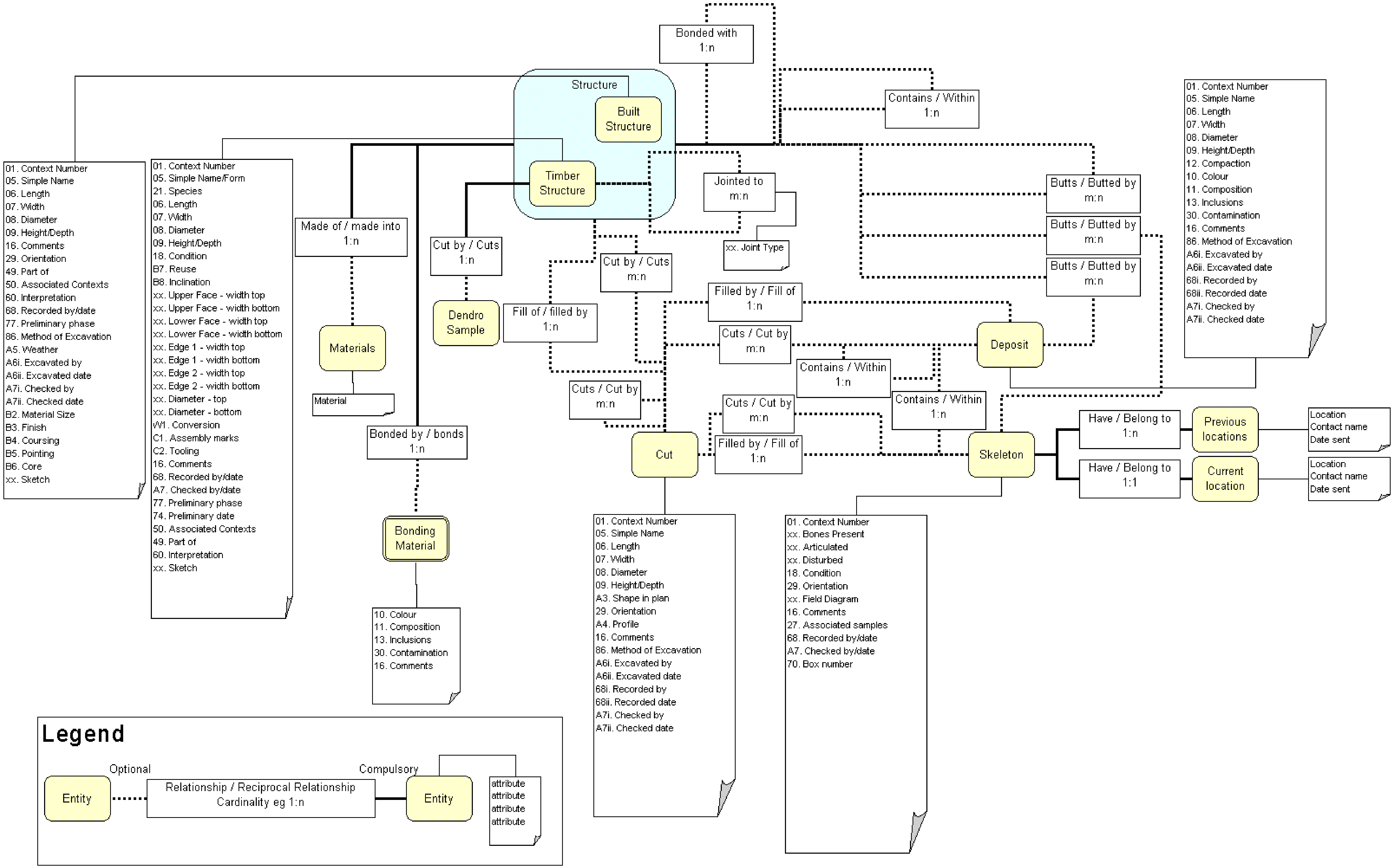


Figure 7; Entity-Relationship diagram for the recording system – context detail

Appendix 1 – List of attributes identified during ER modelling

The following list contains all attributes identified during the modelling process. Where attributes did not have an index number, they have been allocated an xx. prefix to maintain consistency. Such attributes generally exist on forms that have been developed and included on an ad-hoc basis.

A number of attributes occur more than once (e.g. Context Number as 01, 37 and xx) representing the different uses of attributes in the paper-based system (i.e. as record UID or as a reference to another context). Attribute 01 represents the record UID in every case.

Attribute	***Data Type***
01. Context Number	String
01. Drawing Record number	String
01. Record number	String
01. Sample Number	String
01. SF Number	String
02. Project Code	String
03. SSD	String
04. Coordinates	Numeric
05. Category	String
05. Simple Name	String
05. Simple Name/Form	String
06. Length	Numeric
07. Width	Numeric
08. Diameter	Numeric
09. Height/Depth	Numeric
10. Colour	String
11. Composition	String
12. Compaction	String
13. Inclusions	String
14. Bonding material	String
16. Comments	String
17. Siting Description	String
18. Condition	String
19. Count	Numeric
20. Completeness	String
21. Species	String
22. Spot Date	String
23. Sample Volume	Numeric
24. Material	String
25. Purpose of Sample	String
27. Associated objects	String
27. Associated samples	String
29. Orientation	Numeric
30. Contamination	String
33. Stratigraphically Above	String
34. Filled by	String
35. Cut by	String
36. Butted by	String
37. Context Number	String
37. Within	String
38. Contains	String

39. Bonded with	String
40. Same as	String
41. Stratigraphically Below	String
42. Fill of	String
43. Cuts	String
44. Butts	String
48. Cut by sample	String
49. Part of	String
50. Associated Contexts	String
60. Interpretation	String
61. Drawing Nos	String
62. Context numbers	String
63. Photo Nos	String
66. Site treatment	String
68. Drawn by/date	String
68. Recorded by/date	String
68. Sampled by/date	String
70. Box number	String
71. Sent to/date	String
73. LAB No	String
74. Preliminary date	String
77. Preliminary phase	String
78. Shows Contexts/Objects	String
81. Sheet Number	
82. Sample Method	String
83. Drawing Type	String
84. Scale	Numeric
85. Generic name	String
86. Method of Excavation	String
87. Find Type	String
88. Box Type	String
A1. Year	Date/Time
A2. Context Type	String
A3. Shape in plan	String
A4. Profile	String
A5. Weather	String
A6. Excavated by/date	String
A7. Checked by/date	String
A9. Small find nos	String
B1. Sample nos	String
B2. Material Size	Numeric
B3. Finish	String
B4. Coursing	String
B5. Pointing	String
B6. Core	String
B7. Reuse	String
B8. Inclination	String
C1. Assembly marks	String
C2. Tooling	String
C4. Percentage of total	Numeric
P1. Date Range	Date/Time

P2. Size of group	Numeric
P3. Probability	Numeric
P4. Joins to	String
P5. SVA	String
W1. Conversion	String
xx. Allocated by/date	String
xx. Articulated	String
xx. Bones Present	Graphical
xx. Box Number	String
xx. Bulk	String
xx. Camera	String
xx. Checked by/date	String
xx. Comments	String
xx. Completed	String
xx. Contains	String
xx. Context Number	String
xx. Database	String
xx. Date	Date/Time
xx. Description	String
xx. Diameter - bottom	Numeric
xx. Diameter - top	Numeric
xx. Dimensions	Numeric
xx. Disturbed	Binary
xx. Edge 1 - width bottom	Numeric
xx. Edge 1 - width top	Numeric
xx. Edge 2 - width bottom	Numeric
xx. Edge 2 - width top	Numeric
xx. Field Diagram	Graphical
xx. Film Number	String
xx. Film Type	String
xx. Initials	String
xx. Joint Type	String
xx. Jointed to	String
xx. Lower Face - width bottom	Numeric
xx. Lower Face - width top	Numeric
xx. Material	String
xx. Processed	String
xx. Quantity	Numeric
xx. Record Numbers	String
xx. Record type	String
xx. Sample number	String
xx. Sample Size	Numeric
xx. SF number	String
xx. SFs	String
xx. Simple Name	String
xx. Site Name	String
xx. Sketch	Graphical
xx. Type	String
xx. Upper Face - width bottom	Numeric
xx. Upper Face - width top	Numeric
xx. Weight	Numeric

Appendix 2 – instances of attributes

Attribute	Form Name	InstanceNote	Mandat	Gloss	PostE
01. Context Number	BLTST	UID for the context record taken from block	True	False	No
01. Context Number	SKEL	UID for the context record taken from block allocation	True	False	No
01. Context Number	DEPCUT	UID for the context record taken from block allocation	True	False	No
01. Context Number	BULK	UID for the context record taken from block allocation	True	False	No
01. Context Number	TIMST	UID for the context record taken from block allocation	True	False	No
01. Drawing Record number	DRAWIND	UID for the drawing record taken from block allocation	False	False	No
01. Record number	SF	UID in this case refers to record number for object not context; context is	True	False	No
01. Record number	PHOTO2	UID for photo record taken from block allocation	False	False	No
01. Sample Number	SAMEV	UID for the record	True	False	Yes
01. Sample Number	SAMREC	UID for sample record taken from block allocation	True	False	No
01. SF Number	ABVSF	Delilah number conflicts with context number: both = 01.	False	False	No
02. Project Code	ABVSF	As issued by CfA and common to all contexts on a particular site.	True	False	No
02. Project Code	ALLOC	As issued by CfA and common to all contexts on a particular site.	True	False	No
02. Project Code	SF	As issued by CfA and common to all contexts on a particular site.	True	False	No
02. Project Code	BLTST	As issued by CfA and common to all contexts on a particular site.	True	False	No
02. Project Code	CFR	As issued by CfA and common to all contexts on a particular site.	True	False	No
02. Project Code	CTXIND	As issued by CfA and common to all contexts on a particular site.	True	False	No

02. Project Code	DRAWIND	As issued by CfA and common to all contexts on a particular site.	True	False	No
02. Project Code	PHOTO2	As issued by CfA and common to all contexts on a particular site.	True	False	No
02. Project Code	SKEL	As issued by CfA and common to all contexts on a particular site.	True	False	No
02. Project Code	DEPCUT	As issued by CfA and common to all contexts on a particular site.	True	False	No
02. Project Code	BULK	As issued by CfA and common to all contexts on a particular site.	True	False	No
02. Project Code	FILMIND	As issued by CfA and common to all contexts on a particular site.	True	False	No
02. Project Code	BXIND	As issued by CfA and common to all contexts on a particular site.	True	False	No
02. Project Code	POT	As issued by CfA and common to all contexts on a particular site.	True	False	No
02. Project Code	SAMEV	As issued by CfA and common to all contexts on a particular site.	True	False	Yes
02. Project Code	SAMIND	As issued by CfA and common to all contexts on a particular site.	True	False	No
02. Project Code	SAMREC	As issued by CfA and common to all contexts on a particular site.	True	False	No
02. Project Code	TIMST	As issued by CfA and common to all contexts on a particular site.	True	False	No
03. SSD	ABVSF	Trench or area number	False	False	No
03. SSD	ALLOC	Trench or area number	False	False	No
03. SSD	BLTST	Trench or area number	True	False	No
03. SSD	CTXIND	Trench or area number	False	False	No
03. SSD	DRAWIND	Trench or area number	False	False	No
03. SSD	PHOTO2	Trench or area number	False	False	No

03. SSD	SF	Trench or area number	True	False	No
03. SSD	SKEL	Trench or area number	True	False	No
03. SSD	DEPCUT	Trench or area number	True	False	No
03. SSD	BULK	Trench or area number	True	False	No
03. SSD	POT	Trench or area number	True	False	No
03. SSD	SAMREC	Trench or area number	True	False	No
03. SSD	TIMST	Trench or area number	True	False	No
04. Coordinates	ABVSF	non-atomic. 2d location of object.	False	False	No
04. Coordinates	SF	non-atomic: a small find may have multiple sets of eastings and northings	False	False	No
04. Coordinates	BLTST	non-atomic.	False	False	No
04. Coordinates	SKEL	non-atomic. up to 2 sets of eastings and northings	False	False	No
04. Coordinates	DEPCUT	Non-atomic coords	False	False	No
04. Coordinates	POT	up to two pairs of eastings/northings	False	False	No
04. Coordinates	SAMREC	non-atomic. Up to 2 sets of 2d/3d locations.	False	False	No
04. Coordinates	TIMST	non-atomic: up to two pairs of 2d coords.	False	False	No
05. Category	POT	? No idea - not mentioned in manual at all.	True	False	No
05. Simple Name	ABVSF	Stores simple name of object this time rather than context.	False	False	No
05. Simple Name	SF	Simple name for object	True	True	No

05. Simple Name	DEPCUT	Simple term for the context based on site specific glossary	True	True	No
05. Simple Name	BULK	Referred to as context simple name in this instance	True	False	No
05. Simple Name	SAMREC	of the sample, e.g. bulk processing, technology. No glossary in place but	True	True	No
05. Simple Name/Form	BLTST	Same attribute as for other context types (05) but different glossary.	True	True	No
05. Simple Name/Form	TIMST	Same attribute as for other context types (05) but different glossary.	True	True	No
06. Length	BLTST	Numeric dimension	False	False	No
06. Length	DEPCUT	Numeric dimension	False	False	No
06. Length	TIMST	Numeric dimension	False	False	No
07. Width	BLTST	Numeric dimension	False	False	No
07. Width	DEPCUT	Numeric dimension	False	False	No
07. Width	TIMST	Numeric dimension	False	False	No
08. Diameter	BLTST	Numeric dimension	False	False	No
08. Diameter	DEPCUT	Numeric dimension	False	False	No
08. Diameter	TIMST	Numeric dimension	False	False	No
09. Height/Depth	BLTST	Numeric dimension	False	False	No
09. Height/Depth	DEPCUT	Numeric dimension	False	False	No
09. Height/Depth	TIMST	Numeric dimension	False	False	No
10. Colour	DEPCUT	Applies to Deposit only - not cuts. Non-atomic - stores description and	False	False	No

11. Composition	DEPCUT	Applies to Deposit only - not cuts	False	True	No
12. Compaction	DEPCUT	Applies to Deposit only - not cuts	False	True	No
13. Inclusions	DEPCUT	Applies to Deposit only - not cuts. Non-atomic: breaks down into	False	True	No
14. Bonding material	BLTST	Seriously non-atomic: requires full deposit record plus information	False	False	No
16. Comments	ABVSF	free text comments	False	False	No
16. Comments	SF	free text comments	False	False	No
16. Comments	BLTST	free text comments	False	False	No
16. Comments	DRAWIND	free text comments	False	False	No
16. Comments	PHOTO2	free text comments	False	False	No
16. Comments	SKEL	free text comments	False	False	No
16. Comments	DEPCUT	free text comments	False	False	No
16. Comments	BXIND	free text comments	False	False	No
16. Comments	SAMEV	Free text	False	False	No
16. Comments	POT	free text comments	False	False	No
16. Comments	SAMREC	free text comments	False	False	No
16. Comments	TIMST	free text comments	False	False	No
17. Siting Description	SF	Context type for the parent context i.e. pit 1024, post-hole 2234,	False	False	No
17. Siting Description	SAMREC	detail of where the sample was taken from - free-text. Also a sketch,	False	False	No

18. Condition	SF		False	True	No
18. Condition	SKEL	Small finds also have condition attribute - different glossary required	False	True	No
18. Condition	SAMREC	The condition of the sample using standard terms: Dessicated ==>	False	True	No
18. Condition	TIMST	non-atomic. Glossary of descriptive terms used, use as many as apply.	False	True	No
19. Count	SF	Default = 1 unless (on rare occasion) multiple objects recorded as one	False	False	No
20. Completeness	SF		False	True	No
21. Species	TIMST	non-atomic. Holds species plus initials of person making the	True	False	No
22. Spot Date	POT	Different from 74. Preliminary date which is based on bulk finds - spot	True	False	No
23. Sample Volume	SAMREC	Duplicate of sample size on SAMIND	False	False	No
24. Material	ABVSF	non-atomic. Stores all component materials in decreasing order of	False	True	No
24. Material	SF	non-atomic. May be a composite object made of different materials.	True	True	No
24. Material	BLTST	non-atomic. Uses same delilah number as object materials. Specifically	True	True	No
24. Material	BXIND	non-atomic. Should be possible to pull this through from the objects	False	True	No
24. Material	SAMREC	Material(s) observed in the sample. Non-atomic.	True	True	No
25. Purpose of Sample	SAMREC	why the sample is being taken, not the process ie not 'for radio-carbon	False	False	No
27. Associated objects	SF	non-atomic. Relationship with other small finds.	False	False	No
27. Associated objects	SKEL	non-atomic. Relationship with small finds.	False	False	No
27. Associated samples	SKEL	non-atomic. Relationship with samples.	False	False	No

29. Orientation	BLTST	Measured according to agreed strategy	False	False	No
29. Orientation	SKEL	Measured according to agreed strategy. Cardinal point glossary no good as	False	False	No
29. Orientation	DEPCUT	Applies to Cuts only - not deposits. Not strictly necessary in GIS based	False	True	No
30. Contamination	DEPCUT	Applies to Deposit only - not cuts	False	True	No
33. Statigraphically Above	BLTST	non-atomic. Has space on form for initial and revised, but only 1 delilah	False	False	No
33. Statigraphically Above	SKEL	non-atomic. Has space on form for initial and revised, but only 1 delilah	False	False	No
33. Statigraphically Above	DEPCUT	non-atomic. Has space on form for initial and revised, but only 1 delilah	True	False	No
34. Filled by	DEPCUT	non-atomic. Reciprocal of Fill of. Applies to Cuts only. Relationship.	False	False	No
35. Cut by	BLTST	non-atomic. Reciprocal of Cuts. Applies to Cuts only. Relationship with	False	False	No
35. Cut by	DEPCUT	non-atomic. Reciprocal of Cuts. Applies to Cuts only. Relationship.	False	False	No
35. Cut by	TIMST	non-atomic. Relationship with cuts.	False	False	No
36. Butted by	BLTST	Relationship with other contexts.	False	False	No
36. Butted by	TIMST	non-atomic. Relationship with other contexts.	False	False	No
37. Context Number	ABVSF		False	False	No
37. Context Number	SF	Relationship with context	True	False	No
37. Context Number	POT	Relationship with context	True	False	No
37. Context Number	SAMEV	context from which sample was taken	True	False	Yes
37. Context Number	SAMREC	Relationship with context	True	False	No

37. Within	BLTST	Uses same delilah number used for context ID elsewhere.	False	False	No
38. Contains	BLTST	Confusion in manual between this and 37. Relationship with other	False	False	No
38. Contains	SAMREC	non-atomic. Relationship with small finds.	False	False	No
39. Bonded with	BLTST	Relationship with other contexts.	False	False	No
39. Bonded with	TIMST	non-atomic. Relationship with other contexts.	False	False	No
40. Same as	DEPCUT	non-atomic. Relationship.	False	False	No
41. Stratigraphically Below	BLTST	non-atomic. Has space on form for initial and revised, but only 1 delilah	False	False	No
41. Stratigraphically Below	SKEL	non-atomic. Has space on form for initial and revised, but only 1 delilah	False	False	No
41. Stratigraphically Below	DEPCUT	non-atomic. Has space on form for initial and revised, but only 1 delilah	True	False	No
42. Fill of	BLTST	non-atomic. Reciprocal of Filled by. Applies to Deposits only.	False	False	No
42. Fill of	DEPCUT	non-atomic. Reciprocal of Filled by. Applies to Deposits only. Definition	False	False	No
43. Cuts	DEPCUT	non-atomic. Reciprocal of Cut by. Applies to Cuts only. Relationship.	False	False	No
44. Butts	BLTST	Relationship with other contexts.	False	False	No
44. Butts	TIMST	non-atomic. Relationship with other contexts.	False	False	No
48. Cut by sample	TIMST	non-atomic. Relationship with other samples, esp. dendro. Samples where a	False	False	No
49. Part of	DEPCUT	non-atomic. Composite contexts only ie groups. Relationship with other	False	False	No
49. Part of	BLTST	non-atomic. Composite contexts only ie groups. Relationship with other	False	False	No
49. Part of	TIMST	non-atomic. Composite contexts only ie groups. Relationship with other	False	False	No

50. Associated Contexts	DEPCUT	non-atomic. Composite contexts only ie groups. Relationship with other	False	False	No
50. Associated Contexts	BLTST	non-atomic. Composite contexts only ie groups. Relationship with other	False	False	No
50. Associated Contexts	TIMST	non-atomic. Composite contexts only ie groups. Relationship with other	False	False	No
60. Interpretation	DEPCUT	free text interpretive statement	False	False	No
60. Interpretation	BLTST	free text interpretive statement	False	False	No
60. Interpretation	TIMST	free text interpretive statement	False	False	No
61. Drawing Nos	DEPCUT	non-atomic. Relationship with drawings	False	False	No
61. Drawing Nos	SF	non-atomic. Relationship with drawings.	False	False	No
61. Drawing Nos	SKEL	non-atomic. Relationship with drawings.	False	False	No
61. Drawing Nos	BLTST	non-atomic. Relationship with drawings.	False	False	No
61. Drawing Nos	POT	non-atomic. Relationship with drawings.	False	False	No
61. Drawing Nos	SAMREC	non-atomic. Relationship with drawings.	False	False	No
61. Drawing Nos	TIMST	non-atomic. Relationship with drawings.	False	False	No
62. Context numbers	DRAWIND	non-atomic. Relationship with contexts.	False	False	No
63. Photo Nos	DEPCUT	non-atomic. Relationship with photos.	False	False	No
63. Photo Nos	SF	non-atomic. Relationship with photos.	False	False	No
63. Photo Nos	SKEL	non-atomic. Relationship with photos.	False	False	No
63. Photo Nos	BLTST	non-atomic. Relationship with photos.	False	False	No

63. Photo Nos	SAMREC	non-atomic. Relationship with photos.	False	False	No
63. Photo Nos	TIMST	non-atomic. Relationship with photos.	False	False	No
66. Site treatment	SF	description of on-site treatments - possible to categorise?	False	False	No
66. Site treatment	SAMEV	non-atomic, contains three sample volumes	False	False	Yes
68. Drawn by/date	DRAWIND	non-atomic.	False	False	No
68. Recorded by/date	SF	non-atomic.	True	False	No
68. Recorded by/date	PHOTO2	non-atomic.	False	False	No
68. Recorded by/date	SKEL	non-atomic.	True	False	No
68. Recorded by/date	DEPCUT	non-atomic.	True	False	No
68. Recorded by/date	BLTST	non-atomic	True	False	No
68. Recorded by/date	TIMST	non-atomic.	False	False	No
68. Sampled by/date	SAMREC	non-atomic.	True	False	No
70. Box number	SF	box in which object resides.	False	False	No
70. Box number	SKEL	Box in which skeleton resides.	False	False	Yes
70. Box number	BXIND	UID for box records	False	False	No
70. Box number	POT	non-atomic. If there is pottery in both small finds and bulk finds then there	True	False	No
70. Box number	SAMEV	Container identifier.	False	False	No
71. Sent to/date	ABVSF	For when the object is sent to the Fort or an external specialist. Non-	False	False	No

71. Sent to/date	SF	non-atomic. Needs a full audit trail rather than the current situation.	False	False	No
71. Sent to/date	SKEL	For when the object is sent to the Fort or an external specialist. Non-	False	False	No
71. Sent to/date	BXIND	non-atomic. Needs a full audit trail rather than the current situation which	False	False	No
71. Sent to/date	SAMREC	non-atomic. Probobaly no need for audit trail in this instance: Sample moves	False	False	No
73. LAB No	ABVSF	UID assigned by scientists upon entry into LabFile system.	False	False	No
73. LAB No	SF	Can only be entered once record created on LabFile	True	False	Yes
73. LAB No	POT	non-atomic. Relationship with small finds via their lab number.	False	False	No
74. Preliminary date	DEPCUT	Spot date generated by specialist and added to form later by supervisor	False	False	Yes
74. Preliminary date	BLTST	Spot date generated by specialist and added to form later by supervisor	False	False	Yes
74. Preliminary date	BULK	"This column is intended as a guide and entries can be as specific or	True	False	No
74. Preliminary date	TIMST	No details as to where this data comes from, potentially from specialist	False	False	No
77. Preliminary phase	DEPCUT	Preliminary phase once phasing commences and added to form by	False	False	Yes
77. Preliminary phase	BLTST	Preliminary phase once phasing commences and added to form by	False	False	Yes
77. Preliminary phase	TIMST	Preliminary phase once phasing commences and added to form by	False	False	Yes
78. Shows Contexts/Objects	PHOTO2	non-atomic. Relationship between photos and contexts. Relationship	False	False	No
81. Sheet Number	DRAWIND	The sheet number on which the drawing resides, for sheets with	False	False	No
82. Sample Method	SAMREC	How the sample was taken. Not sure if this would benefit from a	False	False	No
83. Drawing Type	DRAWIND	Glossary: Plan/Section/Elevation/Timber/Matrix/Profile	False	True	No

84. Scale	DRAWIND	Scale or scales used: May be possible to use multiple scales for an	False	False	No
84. Scale	PHOTO2	non-atomic if more than one scale bar used in a shot	False	False	No
85. Generic name	SF	post-ex (analysis phase) only!	False	True	Yes
86. Method of Excavation	DEPCUT	Statement regarding excavtion method	False	False	No
86. Method of Excavation	BLTST	Statement regarding excavtion method	False	False	No
87. Find Type	BXIND	Whether the contents are small finds, bulk finds or a skeleton.	False	True	No
88. Box Type	BXIND	Type of box, based on approved box list	False	True	No
A1. Year	ABVSF	Year of excavation	True	False	No
A1. Year	SF	Year of excavation	True	False	No
A1. Year	ALLOC	Year of excavation	True	False	No
A1. Year	BLTST	Year of excavation	True	False	No
A1. Year	CFR	Year of excavation	True	False	No
A1. Year	CTXIND	Year of excavation	True	False	No
A1. Year	DRAWIND	Year of excavation	True	False	No
A1. Year	PHOTO2	Year of excavation	True	False	No
A1. Year	SKEL	Year of excavation	True	False	No
A1. Year	DEPCUT	Year of excavation	True	False	No
A1. Year	BULK	Year of excavation	True	False	No

A1. Year	FILMIND	Year of excavation	True	False	No
A1. Year	BXIND	Year of excavation	True	False	No
A1. Year	POT	Year of excavation	True	False	No
A1. Year	SAMEV	Year of excavation	True	False	Yes
A1. Year	SAMIND	Year of excavation	True	False	No
A1. Year	SAMREC	Year of excavation	True	False	No
A1. Year	TIMST	Year of excavation	True	False	No
A2. Context Type	DEPCUT	Deposit/Cut. Determines which attributes are applicable.	True	True	No
A3. Shape in plan	DEPCUT	Applies to Cuts only - not deposits	False	True	No
A4. Profile	DEPCUT	Applies to Cuts only - not deposits. Non-atomic: breaks down into Break of	False	True	No
A5. Weather	DEPCUT	Statement regarding weather at time of excavation	False	False	No
A5. Weather	BLTST	Statement regarding weather at time of excavation	False	False	No
A6. Excavated by/date	DEPCUT	non-atomic	False	False	No
A6. Excavated by/date	BLTST	non-atomic	False	False	No
A7. Checked by/date	DRAWIND	non-atomic.	False	False	No
A7. Checked by/date	SKEL	non-atomic.	False	False	No
A7. Checked by/date	DEPCUT	non-atomic	False	False	No
A7. Checked by/date	BLTST	non-atomic	False	False	No

A7. Checked by/date	TIMST	non-atomic.	False	False	No
A9. Small find nos	DEPCUT	non-atomic. Relationship with small finds.	False	False	No
B1. Sample nos	DEPCUT	non-atomic. Relationship with samples.	False	False	No
B1. Sample nos	SF	non-atomic. Relationship with samples.	False	False	No
B2. Material Size	BLTST	non-atomic. Dimensions of ALL materials used, including variations within	False	False	No
B3. Finish	BLTST	non-atomic. Description of finish on all surfaces.	False	True	No
B4. Coursing	BLTST	non-atomic. May be a number of courses within a context to describe.	False	True	No
B5. Pointing	BLTST	Overlaps partly with 14. Problematic as pointing may apply only to a	False	True	No
B6. Core	BLTST	free text description of any rubble core present.	False	False	No
B7. Reuse	TIMST	binary - has timber been reused. Supported by comments (16)	False	False	No
B8. Inclination	TIMST	Angle off horizontal, ie 90 for vertical, 0 for horizontal.	False	False	No
C1. Assembly marks	TIMST	Description of assembly marks, with sketch(es) overleaf.	False	False	No
C2. Tooling	TIMST	Description of evidence for tool marks.	False	False	No
C4. Percentage of total	SAMREC		False	False	No
P1. Date Range	POT	non-atomic. Will consist of a from-to date pair or a to-from period pair or	True	False	No
P2. Size of group	POT	? Unknown ? Need clarification	False	False	No
P3. Probability	POT	Confidence in spot date ascription ranked 1-5	False	True	No
P4. Joins to	POT	? Unknown ? Need clarification	False	False	No

P5. SVA	POT	? Unknown ? Need clarification	False	False	No
S1. Residue Mesh Sizes	SAMEV	non-atomic, multiple choice	False	True	Yes
S2. Flot Mesh Sizes	SAMEV	non-atomic, multiple choice	False	True	Yes
S3. Sieved by/date	SAMEV	non-atomic.	True	False	No
S4. Sorted by/date	SAMEV	non-atomic.	True	False	No
S5. Sample Pretreatment	SAMEV	Non-atomic, although not clear what is being recorded: length of	False	False	No
S6. % Residue Sorted	SAMEV	The relative quantity of residues?	False	False	No
S7. Contains	SAMEV	Actually an array of data not a single attribute. Some materials marked	False	False	No
S8. >4mm residue	SAMEV	need clarification	False	False	No
S9. 2-4mm residue	SAMEV	need clarification	False	False	No
SA. <2mm residue	SAMEV	need clarification	False	False	No
SB. Flot	SAMEV	need clarification	False	False	No
W1. Conversion	TIMST	How timber was cut from original lump of wood.	False	True	No
xx. Allocated by/date	ALLOC	non-atomic.	False	False	No
xx. Allocated by/date	CTXIND	non-atomic	False	False	No
xx. Articulated	SKEL	binary attribute (tick box on form)	False	False	No
xx. Bones Present	SKEL	Currently stored as a schematic representation of a skeleton with bones	False	False	No
xx. Box Number	ABVSF	UID for the box in which the object is stored. Cannot be entered on site	False	False	Yes

xx. Box Number	BULK	UID for the box in which the object is stored. Cannot be entered on site	False	False	Yes
xx. Bulk	CFR	binary	False	False	No
xx. Camera	FILMIND	Replaces the camera1/camera2 form split.	False	False	No
xx. Checked by/date	CTXIND	non-atomic	False	False	No
xx. Comments	BULK	Additional free text comments	False	False	No
xx. Comments	SAMIND	Additional free text comments	False	False	No
xx. Completed	BULK	Whether context has been fully excavated and completed.	True	False	No
xx. Contains	DEPCUT	Bulk finds attribute does not have an attribute number. Non-atomic.	False	True	No
xx. Context Number	CFR	copied from context sheets	False	False	No
xx. Context Number	SAMIND	relationship with contexts	False	False	No
xx. Database	FILMIND	Binary tick-box to indicate that database entry is complete; no indication of	False	False	No
xx. Date	SAMIND	equivalent to recorded by/date, but atomic!	False	False	No
xx. Description	CTXIND	Free text description	False	False	No
xx. Diameter - bottom	TIMST	Diameter across bottom of post	False	False	No
xx. Diameter - top	TIMST	Diameter across top of post	False	False	No
xx. Dimensions	SF	Currently stored in free-text comments field: dimension and	False	False	No
xx. Disturbed	SKEL	binary attribute (tick box on form)	False	False	No
xx. Edge 1 - width bottom	TIMST	Width across the bottom of one of the side faces of post, opposite edge 2	False	False	No

xx. Edge 1 - width top	TIMST	Width across the top of one of the side faces of post, opposite edge 2	False	False	No
xx. Edge 2 - width bottom	TIMST	Width across the bottom of one of the side faces of post, opposite edge 1	False	False	No
xx. Edge 2 - width top	TIMST	Width across the top of one of the side faces of post, opposite edge 1	False	False	No
xx. Field Diagram	SKEL	Sketch showing salient details.	False	False	No
xx. Film Number	PHOTO2	Sequential number for film, sequences run for each film type	False	False	No
xx. Film Number	FILMIND	UID for film, in sequential order by film type.	False	False	No
xx. Film Type	PHOTO2	colour slide, colour print, B+W, Digital	False	True	No
xx. Film Type	FILMIND	Duplicate of attribute on Photographic record. Best stored as attribute of film	False	True	No
xx. Initials	SAMIND	equivalent to recorded by/date, but atomic!	False	False	No
xx. Joint Type	TIMST	non-atomic, space for up to four joints on form. Associated with Jointed	False	False	No
xx. Jointed to	TIMST	non-atomic, space for up to four joints on form. Context numbers for other	False	False	No
xx. Lower Face - width bottom	TIMST	Width across the top of the upper face of post.	False	False	No
xx. Lower Face - width top	TIMST	Width across the top of the lower face of post.	False	False	No
xx. Material	BULK	Material present	False	True	No
xx. Processed	FILMIND	Binary tick-box to indicate that film has been processed.	False	False	Yes
xx. Quantity	BULK	Quantity of material	False	False	No
xx. Record Numbers	ALLOC	non-atomic. Stores record number range as allocated. No system id	False	False	No
xx. Record Numbers	CTXIND	potentially non-atomic. Stores record number taken to be used to record	True	False	No

xx. Record type	ALLOC	type of records number allocated to	False	True	No
xx. Sample number	BULK	relationship with samples	False	False	No
xx. Sample number	SAMIND	UID for sample record	False	False	No
xx. Sample Size	SAMIND	sample size in litres.	False	False	No
xx. SF number	BULK	for small finds from the context. Duplicates small finds recording.	False	False	No
xx. SFs	CFR	binary	False	False	No
xx. Simple Name	CFR	of context; copied from context sheets.	False	False	No
xx. Site Name	ABVSF	The name of the site.	True	False	No
xx. Site Name	ALLOC	The name of the site.	True	False	No
xx. Site Name	CFR	The name of the site.	True	False	No
xx. Site Name	CTXIND	The name of the site.	True	False	No
xx. Site Name	DRAWIND	The name of the site.	True	False	No
xx. Site Name	PHOTO2	The name of the site.	True	False	No
xx. Site Name	FILMIND	The name of the site.	True	False	No
xx. Site Name	BXIND	The name of the site.	True	False	No
xx. Site Name	SAMIND	The name of the site.	True	False	No
xx. Sketch	BLTST	a sketch which never leaves the context sheet.	False	False	No
xx. Sketch	DEPCUT	a sketch which never leaves the context sheet.	False	False	No

xx. Sketch	TIMST	annotated sketch of aspects of the recording, plus annotation.	False	False	No
xx. Type	BULK	for small finds from the context. Duplicates small finds recording.	False	False	No
xx. Upper Face - width bottom	TIMST	Width across the bottom of the upper face of post.	False	False	No
xx. Upper Face - width top	TIMST	Width across the top of the upper face of post.	False	False	No
xx. Weight	BULK	Mass of material	False	False	No