BAR, computers, & publication: past, and present

Sebastian Rahtz
Production Department
Elsevier Science Ltd
The Boulevard
Langford Lane
Kidlington
Oxford OX5 7GB
email: s.rahtz@elsevier.co.uk

September 1994

Contents

1 Introduction

In considering the relationship between archaeological computing, and the history of *British Archaeological Reports* (henceforth BAR) , we may wish to ask three questions:

1. What did BAR do for archaeological computing?
2. What are the current problems in archaeological publication that can be solved with information technology?
3. What would BAR do if it was set up now?

The purpose of this brief paper is to try and answer these points, and in particular question whether BAR was important actually for its publications, or for the social phenomenon which it represented. I deliberately here use the past tense — it was the *quantity* of BAR volumes in its heyday which made it truly distinctive in a way which the (very fine!) new operation by Tempus Reparatum cannot match. I suggest that the current enthusiasm for the *World Wide Web* may be comparable to the origins and effect of BAR.

2 BAR and archaeological computing publication

BAR made an important contribution to archaeological computing in the UK — it legitimized the publication of the *Computer Applications in Archaeology* conferences. From 1973 to 1986 this conference had been held mainly in the University of Birmingham under the leadership of Sue Laflin (with forays to London and Bradford), and was an important (in some cases the only) outlet for workers in the field; the history of the conference is described in more detail in Reilly 1994 forthcoming. Unfortunately, the proceedings were published and distributed by the conference organizers in unedited A5 booklets prepared from authors' camera-ready copy (before the days when laser printers were commonplace), and never achieved widespread circulation. An important part of the 'democratization' of the conference in 1987 was production of the proceedings by a publisher with good access to world markets, and BAR were willing to take it on. At the same time, some degree of editing, peer review and systematic design was introduced, together with computerized page makeup from authors' files. This had two immediate benefits:

1. Citation of articles in a BAR volume became a much more reasonable element of a *Curriculum Vitae* than references to the old *ad hoc* A5 volumes;
2. The volumes had a more respectable typeset appearance (in advance of current practice in most BAR volumes), and the articles remained in electronic form for future republication.[1]

The results can be seen in the increasingly confident BAR International Series 393, 446, 548, 565 and 577.

That the first international CAA conference, held in Århus in 1992, was *not* published as a BAR (but as Andresen, Madsen & Scollar 1993) is an indication that the series was not entirely successful. This stemmed partly from problems with BAR itself, after the transition to Tempus Reparatum resulted in much decreased market presence, and partly from an increasingly sophisticated desire by authors and editors for better publication. The nature of the subject in some fields of archaeological computing demands high

[1] In fact, this electronic archive has not yet been utilized or indeed formalized; the files for 1987–1991 are held by the author of this paper in uniform format, and await archival storage or reissue.

Table 1: Articles in CAA proceedings

Year	*Articles*	*Location*
1974	13	
1975	13	
1976	11	
1977	12	
1978	10	
1979	6	
1980	6	
1981	12	
1982	23	
1983	22	
1984	22	
1985	18	
1986	17	
1987	31	Leicester
1988	38	Birmingham
1989	32	York
1990	30	Southampton
1991	30	Oxford

quality graphics, and colour pictures. The CAA volumes were printed from 300dpi laser printer output, which was not really good enough for graphics; as for colour, although BAR S446 had colour inserts, these were simply colour xeroxes kindly provided by IBM, and the problem has not yet been solved. There appears little chance of colour becoming cheap enough to justify its use in the small print runs appropriate for BAR (or indeed almost any technical archaeological publication).

It should also be noted that the undoubted world-wide spread of the BAR volumes was not accompanied by *depth*; while known in the USA, they were not standard items in many libraries, while aspiring computer archaeologists in Africa and India had no more chance of buying the BAR books than if they had been published in hard-back by a University Press at three times the price. The only solution to the latter problem would have been distributed local republication, and it is regrettable that this never seems to have been tried. It is not the purpose of this paper to analysis the whole range of BAR and decide whether it is too biased towards work in the First World (though one has the impression of almost no volumes from the developing world), but it is certainly clear in our computing field that the geographical range of authors is small. The success of the One World Archaeology series shows that the immediate, facile, conclusion that little interesting work is being *done* in the Third World cannot be upheld.

Five volumes of conference proceedings may seem trivial to other sub-disciplines of archaeology, but for the computing and quantitative methods professionals, they made a real contribution to British work being known around the world. Table 1 shows, at the least, the simple increase in *quantity* of material available in the BAR series.

Apart from the CAA conference volumes, and the earlier conference proceedings S271, it is perhaps more a reflection on the discipline of archaeological computing than BAR that few methodological volumes in this area appeared; few serious works have been produced by *any* publisher (Reilly & Rahtz 1992 and Richards & Ryan 1985 are perhaps the only main-stream volumes one can point at from recent years, apart from some very important books relating to quantitative methods). The theses of Paul Reilly (BAR 190) and Nick Ryan (BAR 183) are important works in the BAR canon, but this is not an adequate reflection of work over 20 years. This contradicts an obvious characterization of BAR as a forum for immature disciplines within archaeology (one may look at clay pipe studies, and maritime archaeology, as other areas which struggled to gain respectability during the first 20 years of BAR). The reader may prefer to take the perhaps legitimate view that 'computing' is not a discipline *per se* in archaeology, and that one had better look more closely at *all* the BAR volumes to see whether the impact of quantitative analysis and computerization can be traced in publications.

3 Archaeological data publication and computers

I would now like to address one particular area which BAR *appeared* to take up, the publication of 'real' excavation reports, with full backing data. In fact, Philip Rahtz points out that the proportion of excavation reports in the BAR catalog is actually rather small, and perhaps we are misled by the knowledge that the whole operation started with Hands and Walker's four volumes on Shakenoak. Nevertheless, Anthony Hands argues that data publication in general was an important feature of BAR, and it behooves us to consider the state of computerization.

In the earlier part of the century, excavations were, at least in Britain, on a relatively small scale, and not very numerous. Academics and professional archaeologists knew about many of them while they were taking place; and when they were eventually published, they were easily assimilated. This was made easier inasmuch as they were written largely in a reader-friendly narrative style, with little use of statistics or other numerical analysis, and with minimal scientific input.

Two factors have changed the situation radically since the 1960s. The first is the enormous expansion in field archaeology brought about by accelerating destruction of sites. The other is the equally dynamic improvements in technique, both in the recovery of vastly more information in the field, and in subsequent analysis; much of the latter has been in the realms of hard science, but has also been made possible by the development of computerization (though the uptake of quantitative methods appears to have been very patchy).

The result has been that there are now so many excavation reports that no-one can hope to give them detailed attention, or even to know of their existence. Moreover, they have become very difficult to read or comprehend. Text has become more 'objective', tedious and less narrative; and there is a noticeable loss in the ability of archaeologists to use graphics as a medium of expression. Part of the blame for problem can, of course, be laid at the door of BAR itself, which pioneered uncritical publication of 'data', and undertook no quality control of drawings (either on the author side, or the printing side).

One aspect of the problem that BAR *did* try to solve was the *speed* of publication, but their prompt and efficient transition from completed manuscript to printed volume did not help the underlying inability of the excavators to get their material from the record card to the analytical page. The major public funder in the UK, English Heritage, tried to tackle this in many different ways, with some success, but the problem remains rather intractable. The BAR approach was to just publish data, and the English Heritage approach was to publish extended abstracts and hide the data behind a cloak of 'archives' but neither really seems to help the future scholars.

Attempts have been made to deal with complexity by changing technology — micro-print, micro-fiche, archives etc, and an attempt to separate 'data' from 'interpretation'; but these have often led to increasing frustration, because they often made a report much more difficult to *comprehend*, albeit cheaper to produce. We need radical approaches to make the excavation data more accessible and more comprehensible. If we get it right, it means much more than saving money or getting reports out faster, it means new possibilities in the discipline itself. Neel Smith gave the right view in 1991:

> "As new technologies change our understanding of communication, dissemination of information and publication, they imply changes in our notion of the fundamental activity of our discipline. The theoretical discussions of archaeology as a production of texts have to date conceived of 'texts' solely in terms of material printed on paper, but we need to adapt our archaeological rhetoric to current information technologies. This is an enormous undertaking. Indeed, it represents nothing less than rethinking the place of archaeology in a new world view"

What exactly is going to make the difference? There are various information concepts and technologies available to disseminate archaeological information, of which the three most interesting[2] are formal reasoning about interpretation, hypertext, formal description of documents, and networking.

There has been a reasonable amount of success in formalizing archaeological knowledge and reasoning at a high level about processes of interpretation (see the work of Gardin (eg 1987), Barceló (1992), Dallas (1992) and, most important, Arthur Stutt's work at the Open University and Southampton University—see,

[2]Some readers may wish to add the techniques of three-dimensional visualization, but that moves away from the main bias of this paper.

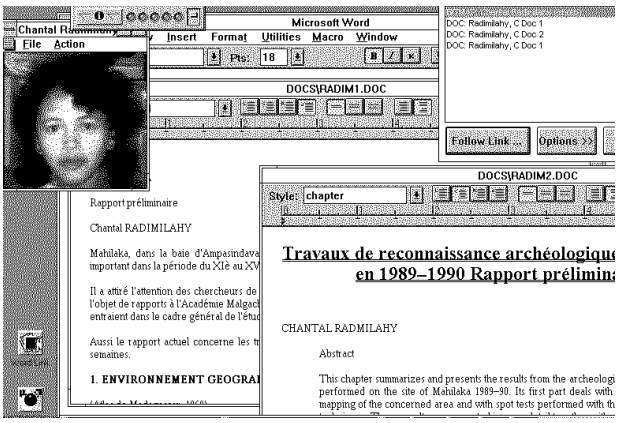

Figure 1: Example of archaeological hypertext

eg, Stutt & Shennan 1992). In general, however, this approach has not been the revolution it might have been; it depends less on technology than on a perception of archaeological evidence.

Hypertext in archaeology has been moderately well tried. Presenting information on a computer screen instead of in a printed book, it allows a form of non-sequential writing in which documents (which can be of any media type such as text, graphics, sound, video etc) are broken down into small self-contained units which are linked together in a variety of ways, and the reader is allowed to follow any path through the information. At its worst, it is little more than an electronic encyclopedia, at its best it has all the virtues of the traditional book, and adds valuable extra facilities. It is not difficult to build hypertext versions of archaeological reports or other textual databases; see the experiments of Rahtz *et al* at Southampton (eg Rahtz, Hall & Allen 1992); and the sophisticated success of the Perseus Project in the classical field (Smith 1992; Mylonas & Heath 1990). Figure 1 shows an example screen from a *docuvers* about East African archaeology using the Microcosm software (from Rahtz & Sinclair 1994). We have also seen very impressive museum-based multi-media systems (Makkuni 1992. Every archaeological unit or museum is aware of the possibilities of storing text and pictures on discs for access in some future environment.

The book, generally, forces a procedural, directed, approach to learning, while hypertext promotes a cognitive attack which puts the onus on the reader to make the associations between 'facts' and build a personal 'reality'. With such a change in delivery method, it is important that writers are aware of the different emphasis. It could be (fancifully) argued that the theory of BAR about just publishing data was an experiment of the same kind, trying to deemphasize just one interpretation by the original excavator, and giving the reader enough metaphorical rope to hang themselves on their own views.

Taking a different view of the problem, we have had plenty of success with highly structured databases of low-level data, which can now be linked to CAD or GIS systems. The SYASS project based at the University of Southampton showed with a number of experiments (including SyGraf and the current WinDig) that systems to teach excavation strategy could also function as form of publication, with an interactive site plan. Little, however, has been seen of the comparable formal models for the synthetic text of reports. Most books and papers lack a formal definition which will make them amenable to automatic processing and

```
<!element arch o o
        (article | report | book | archive>
<!element report - -
        (titlepag, header?, abstract?, toc?, lof?, lot?, p*,
         chapt*, (appendix, chapt+)?, biblio?) +(footnote)>
<!element titlepag o o (title, author*)>
<!element title - o (#pcdata, subtitle?) +(newline)>
<!element subtitle - o (#pcdata)>
<!entity %sect "heading, header?, p* " >
<!element chapt - o (%sect, sect*) +(footnote)>
<!element sect  - o (%sect, sect1*) +(footnote)>
.....
```

Figure 2: Document Type Description — elements

```
<!doctype arch system>
<title> Excavations at Roughground Farm Lechlade,
Gloucestershire: a prehistoric and Roman landscape
<author>Tim Allen
<author>Timothy Darvill
<author>Sarah Green
<author>Margaret Jones
<site>Roughground Farm</site>
<county>Gloucestershire</county>
<gridref>SP 216/009 <gridref>SP 221/005
<excavation start="1957" end="1990" digger="Jones">
<period name="Neolithic" start="-2000" end="-1000">
<period name="Iron Age" start="-500" end="0">
<period name="Roman" start="100" end="450">
```

Figure 3: Report header

effective archive searching. The often sterile work on standards in archaeology in areas like site databases, bibliographies, and standardized vocabulary has not yet resulted in agreement on textual material. Almost every modern archaeologist prepares her report electronically, but most are useless, because of the lack of structure and unambiguous markup. The classic paper of Coombs, Renear & DeRose 1987, and the huge effort of the Text Encoding Initiative, show how the humanist disciplines are now well ahead of archaeology in this respect. They use the comprehensive ISO standard for preparation of electronic book sources, using generic textual markup. This is the Standard Generic Markup Language (SGML; see Goldfarb 1990 for a full specification, and van Herwijnen 1990 for a more readable introduction), a language for describing the structure of documents. It does not dictate a specific set of codes to use, but provides a grammar in which a description of the codes must be written. This allows the recipient of a document to verify it against a Document Type Description, and to process it automatically. Extensions of SGML (HyTime) allow multi-media documents to be described, and this allows software-independent hypertext sources. The start of an example DTD is given in Figure 2; it is not necessary to understand this in detail, simply to appreciate that it lists unequivocally all the commands that will appear.

Given such a DTD, the start of a report would look like figure 3, containing the same information as a conventional title page, but specifying what the elements mean. This book could be trivially processed to find the name of the site, and the periods represented therein. The future of generic markup like SGML will probably be driven by software; when the big word-processors vendors add SGML-compatible export to their products, its use will increase dramatically. The advantages are:

1. It is not tied to any country, institution or system; it deals with all languages and publishing systems;

2. It is thorough, comprehensive, and linked to other standards;
3. It has a structure by which documents can be validated;
4. It encourages non-printing markup;
5. It is an ISO standard.

But its disadvantages for the average archaeologist are that:

1. It is not a system that does anything for the casual user; it needs formatting software to produce printed pages;
2. It is tedious to read and write at present;
3. It imposes a discipline upon writers which they are unwilling to accept.

Few archaeological groups will gain *immediate* benefit from adopting a policy of archiving material in SGML markup. There is no established academic practice of serious exchange, or searching, of archaeological documents, although there is plenty of uncritical archiving. But if archaeologists want to participate in the global database, there is very little choice but to start sorting out their textual archives.

Where does this leave networks, the fourth technology? The answer is that the new market leader in wide-area networking application, World Wide Web, neatly combines together not only the themes of this section, but also reflects on the history of BAR. Much has been written about the Internet and its philosophy already, and World Wide Web has often been described (see Krol 1992 for an excellent overview of the Net, and *World Wide Web* (henceforth WWW)); we may choose to remark on three aspects of WWW:

1. Its fundamental aim is distributed hypertext; dynamic multi-media documents built up out of links to smaller building blocks;
2. The underlying markup language, HTML, is an instance of an SGML-valid language, and meets (though not ideally) many of the archival needs;
3. It provides a single interface to any view of the data a provider can devise; it can therefore act as a vehicle for the formalistic approach as well as for the database approach.

Some of the parallels between WWW and BAR are instructive about both the past and the future:

- It is democratic, since it is effectively free for academics, and can be provided cheaply to many others; and because anyone can add material to WWW without reference to their academic status; this lets it function as an output for young scholars who are being frozen out of traditional journals;
- It is fast; as soon as something is written, it can be made available without any further ado; this can promote rapid, informed, development of a discussion;
- It is cheap, for the provider, since it costs little beyond disk space to make material available; the *reader* may not appreciate the volume of material if they pay for downloading time, of course!
- It provides a maverick cross-section of archaeology, but is the chosen outlet for the traditional scholar; what is now available on WWW is about as useful as (for instance) the first few blue BAR volumes;
- It promotes unedited, ephemeral, sloppy writing; it also allows unscrupulous authors to change old material if it becomes embarrassing;
- Its archival quality is poor at present, being up to the author / provider;
- It is unstable, since the only real copy of a document is held by its author, who may lose it or withdraw it at any time;
- Its physical quality is often poor (since people read on screen), though its potential quality is excellent (given, for instance, the ease of inclusion of colour pictures)

The obvious strengths of BAR volumes (timeliness, democracy, data publication), and their weaknesses (lack of editing, lack of archiving) are exactly the reasons why one is reluctant to wholeheartedly embrace WWW. Its obvious advantages and ease of use are, however, so great that it is to hoped that traditional publishers will soon take advantage and provide the stability and completeness so far lacking.

4 The future

Publication in a blue or red book does not *per se* necessarily mean that people can realistically look at it; many scholars around the world are disenfranchised by accident or economic design, and we should not be too confident that in 50 years time enough copies of, say, *New Forest Roman Pottery*, will survive to make it a genuine part of human knowledge. If the only copies we can be sure of are in the Bodleian Library and the Library of Congress, little more has been achieved than for a PhD issue. Librarians may well consider volumes of data as expendable.

The archaeological writer has to realize that there are two different powerful influences at work: on the one hand, we have the concept of *delayed evaluation*, whereby there is no longer a need to publish or promote one single interpretation, but leave it to the reader to follow her own path through a web of data; on the other hand, it is not acceptable or meaningful to just publish gobbets of raw data and think that ones job is over — one has a duty to justify and explain the work. In the well-written excavation report, this has always been to some extent true; the author will have clearly separated interpretation from data, but inevitably the linear nature of printed books will have emphasized one approach over another. The danger we face is that authors will feel completely absolved from the responsibility of any sort of interpretation.

In late 1994, it looks as if four pieces of technology (or their like) are going to dominate the near future:

1. World Wide Web — delivering structured hypermedia documents over the public networks;
2. CDROM — very considerable quantities of raw data in a reliable and simple format;
3. SGML — an agreed markup for archival reports;
4. Adobe Acrobat — *Portable Document Format*: an efficient subset of PostScript to provide effective on-screen typeset documents.

It is important that archaeological publishers develop their 'computer culture'. The technology *per se* does not do very much that was not possible before, but it does change the way in which we work, the possibility of examining more sources and data, and our willingness to revise and re-edit our own writing. As geological changes make a river take a different route to the sea, so the changes in the way information is presented electronically mean that the archaeologist will be pushed into different types of publication. Only by understanding the technology can we dictate what happens, and attempt proper understanding of material culture processes, rather than drowning in a sea of half-digested data. This is not to decry the revolutionary and stimulating experiment which BAR has represented for 20 years; rather it may be a warning to the practitioners and funders of field archaeology that their house is still far from being in order.

Acknowledgements

I am grateful to Peter Ucko and Philip Rahtz for discussions about archaeological publication at various times; at the meeting from which this book comes, I thank Richard Reece for stimulating ideas about BAR which made me want to write up this paper.

Part of this paper is derived from 'Archaeology, publication and computer culture', delivered to the first Hungarian conference on archaeological computing, held in Budapest in May 1993; the conference proceedings are awaiting publication in Hungarian.

Bibliography

Andresen, J., T. Madsen & I. Scollar (eds) 1993. *Computing the past. Computer applications and quantitative methods in archaeology. CAA92*. Århus University, Denmark: Århus University Press.

Barceló, J. 1992. Programming an intelligent database in archaeology. In *Computer Applications and Quantitative Methods in Archaeology 1991*, Lock, G. & J. Moffett (eds), 21–28, Oxford: British Archaeological Reports.

Coombs, J., A. Renear & S. DeRose 1987. Markup systems and the future of scholarly text. *Communications of the ACM*, **31 (7)**, 933–47.

Dallas, C. J. 1992. Syntax and semantics of figurative art: a formal approach. In *Archaeology and the Information Age*, Reilly, P. & S. Rahtz (eds), chapter 16, London: Routledge.

Gardin, J.-C., O. Guillaume, P.-Q. Herman, A. Hesnard, M.-S. Lagrange, M. Renaud, & E. Zadora-Rio 1987. *Systèmes experts et sciences humaines: Le cas de l'archéologie*. Paris: Eyrolles.

Goldfarb, C. F. 1990. *The SGML Handbook*. Oxford: Oxford University Press.

Krol, E. 1992. *The Whole Internet. User's Guide & Catalog*. Sebastopol, CA, USA: O'Reilly & Associates. 2nd edition 1994.

Lavell, C. 1986. Getting it back: some desiderata for information retrieval in archaeological computer archives. In *Computer usage in British Archaeology*, Richards, J. D. (ed.), IFA Occasional Paper No. 1, 75–9. Birmingham: Institute of Field Archaeologists.

Lock, G. & J. Moffett (eds) 1992. *Computer Applications and Quantitative Methods in Archaeology 1991*, International Series 577. Oxford: British Archaeological Reports.

Lockyear, K. & S. P. Q. Rahtz (eds) 1991. *Computer Applications and Quantitative Methods in Archaeology 1990*, International Series 565. Oxford: British Archaeological Reports.

Makkuni, R. 1992. The electronic capture and dissemination of the cultural practice of Tibetan Thangka painting. In *Archaeology and the Information Age*, Reilly, P. & S. Rahtz (eds), chapter 20, London: Routledge.

Mylonas, E. & S. Heath 1990. Hypertext from the data point of view: paths and links in the Perseus Project. In *Hypertext: concepts, systems and applications*, Rizk, A., Streitz, N. & André, J. (eds), 324–336, Cambridge: Cambridge University Press. Proceedings of the ECHT '90 conference.

Rahtz, S., W. Hall & T. Allen 1992. The development of dynamic archaeological publications. In *Archaeology and the Information Age*, Reilly, P. & S. Rahtz (eds), chapter 22, London: Routledge.

Rahtz, S. & J. Richards (eds) 1989. *Computer Applications and Quantitative Methods in Archaeology 1989*, International Series 548. Oxford: British Archaeological Reports.

Rahtz, S. & P. Sinclair 1994. Multimedia information systems for East African archaeology. *Archeologia e Calcolatori*, **5**, 219–236.

Rahtz, S. P. Q. 1986. AIE — archaeological information exchange. *Archaeological Computing Newsletter*, **8**, 10.

Rahtz, S. P. Q. (ed.) 1988. *Computer and Quantitative Methods in Archaeology 1988*, International Series 446. Oxford: British Archaeological Reports.

Reilly, P. 1994 (forthcoming). 21 years of CAA. In *Computer applications and quantitative methods in archaeology. CAA93*, Wilcock, J. & Lockyear, K. (eds), Tempus Reparatum.

Reilly, P. & S. Rahtz (eds) 1992. *Archaeology and the Information Age*. London: Routledge.

Richards, J. D. & N. S. Ryan 1985. *Data Processing in Archaeology*. Cambridge: Cambridge University Press.

Ruggles, C. L. N. & S. P. Q. Rahtz (eds) 1988. *Computer and Quantitative Methods in Archaeology 1987*, International Series 393. Oxford: British Archaeological Reports.

Smith, N. 1992. An experiment on electronic exchange and publication of archaeological field data. In *Computer Applications and Quantitative Methods in Archaeology 1991*, Lock, G. & J. Moffett (eds), 49–58, Oxford: British Archaeological Reports.

Stutt, A. & S. Shennan 1992. Designing a workbench for archaeological argument. In *Archaeology and the Information Age*, Reilly, P. & S. Rahtz (eds), chapter 18, London: Routledge.

van Herwijnen, E. 1990. *Practical SGML*. Kluwer. 2nd edition 1994.