4
Democracy, data and archaeological knowledge

J. Huggett
Department of Archaeology, University of Glasgow, Glasgow G12 8QQ, UK
Email: jhuggett@dish.gla.ac.uk

4.1 Introduction

In a recent article (Huggett 1993), I drew attention to what seemed to be an increasingly common phenomenon in papers discussing the application of computers in a variety of archaeological contexts: the concept of the democratisation of knowledge. This phrase has rapidly assumed the status of a buzzword over the last couple of years but it has rarely been questioned in terms of its implications for archaeology and archaeologists. In that sense then, I would suggest that the democratisation of archaeological knowledge using information technology is a classic instance of us, as archaeologists, hurtling down a road without any real idea of why we're doing it, or whether indeed it is a good idea in the first place. This paper addresses such issues as these, looking in particular at what is apparently meant by democratising knowledge and what it might mean for archaeology.

4.2 Computers and Knowledge

The impact of information technology on the production and distribution of knowledge is fundamental to any concept of democratisation. Computers can be integral to all stages: they may be the way by which we acquire knowledge, communicate that knowledge to others, and incorporate that knowledge into whatever is perceived as constituting the corpus of archaeological knowledge.

Much archaeological knowledge is communicated orally, and it is impossible to quantify how many 'oral archaeologies' exist now or have existed in the past without ever having been committed to paper. A substantial element in the creation of archaeological knowledge is the number of unpublished conference papers, seminars, and animated bar discussions that take place. Despite this, the primary academically and professionally recognised means of communicating ideas, concepts and knowledge is through publication – a more substantial and systematic method than intangible oral archaeology. However, disseminating published knowledge is slow, it lacks spontaneity, and is often the result of many stages of re-writing, editing, reconsideration, and reformulation.

Computers are increasingly being offered as a means of mediating between these two models of knowledge acquisition: they provide the opportunity for faster dissemination through electronic publication, without losing the freedom and spontaneity associated with oral communication. Computer-based communications have been identified as a major cultural revolution (for example, Harnad 1991) – representing not just a technological paradigm shift like the Gutenberg printing press, but also a symbolic shift towards the dematerialisation of culture. Tangible artefacts are replaced with ephemeral digital forms, words are substituted by pictures, and the solidity of the printed page is replaced by an insubstantial phosphor image on a computer monitor. It is the age of a post-textual archaeology (Sherratt 1993) that is driven by information technology. Nothing is permanent: the knowledge contained within these representations, as well as the representations themselves, can be recomposed by the viewer in a faster, more flexible manner than by using traditional methods. This is in the near-future: archaeological interest in hypermedia is already apparent (for example, Rahtz, Hall & Allen 1992, Wolle 1994), and increasing numbers of people are gaining access to the vast expanses of the Internet with a consequent increase in the amount of networked archaeological resources. In many respects, proponents of knowledge democratisation were initially captivated by the vast amounts of data that are freely available via the Internet, and this excitement led to the creation of an image of a brave new world of free access and use of information.

4.3 Democratising Archaeological Knowledge

What is the democratisation of knowledge? The concept covers a wide variety of different aspects, but it is essentially based upon the increased ease of transmission of information through the use of computers, and hence the ability to share data and to manipulate and reprocess that data in various ways and to different ends (for example, Reilly & Rahtz 1992, 18; Fukasawa 1992, 97).

At one level, this raises a host of new and exciting possibilities, including:

- the ability to access computerised databases on-line, archaeological databases that others have created and made available to us;
- the ability to reinterpret a site using different technologies to reveal new aspects;
- the ability to extract digital data for reprocessing within an entirely different context to that within which it was originally created;
- the ability to combine data from a wide variety of disparate sources.

The potential here is immense, but this vast reserve of data, information and knowledge also raises huge problems. In some respects, these problems are not

simply a result of the ability to access all this information, but are present in computerised data of any kind – as is often the way, new applications reveal the shortcomings of existing methodologies.

There are a number of issues that need to be considered.

- the nature of the information being made available. What is it that is being provided? Why? What do we actually want?
- the way that information is accessed. The whole concept of democratising information requires access to the material to be available, presumably by definition to everyone. Clearly there are major resourcing implications here.
- the ownership of the information provided. Is it public domain, or do people get charged for it? If it is in the public domain, is it supplied with conditions – can you freely use the resource or do you only have free access to some aspects? What about issues of copyright and intellectual property?

4.3.1 The content of the resource

What information would be made available in a democratised scenario? There is little doubt that what most people seem to desire is access to data, and in particular access to excavation data. A major problem in archaeology is the publication log-jam – the large numbers of excavations, in many cases undertaken years ago, still being worked on by their excavators, unpublished and sometimes inaccessible. However, this democratisation process has the potential to radically change our attitude to publishing archaeology. Instead of spending years assembling reports based on catalogues with in-depth multivariate analyses which by common consent turn out to be very cold, dull and boring to read, we could mount our datasets more or less immediately on a network and then concentrate on writing the kind of report we'd all like to read, concentrating on interpretation, rather than simply reproducing reams of tables and diagrams – the raw data would be available for anyone who wanted to see what the interpretations were based on.

Assuming the information is available, however, other difficulties arise. Foremost amongst these is the question of the nature and content of the information. The problem is simply that archaeologists do not record data in the same way. Archaeologists work with different theories and ask different questions. These issues will fundamentally affect the way we record our data and the type of data that we record. Archaeologists develop new methodologies, new twists to old techniques, in order to deal with situations that arise which make standardised methods inappropriate. Problems often require solutions that are specific to that situation. The alternative is to try and shoehorn our methodologies and observations into an inappropriate structure.

Of course, the same questions were raised in the past about standardisation, but the fact that there are as yet no formal recording standards, at least in British archaeology, should surely serve as a warning. The Archaeological Data Archive Project is designed to create a collection of site archives, assembled with the intention of making them available to other archaeologists via media such as the Internet (Eitlejorg, this volume). As those who subscribe to the Archaeological Institute of America List Digest circulated by Nick Eitlejorg on the Internet will be aware, there was a good deal of discussion about standards raised by the founding of this archive project, to the extent that a new section of the list was created. Somewhat ironically, perhaps, once the new list was set up, discussion dried up! The Archaeological Data Archive Project intends to handle such problems by ensuring that all data files are accompanied by full descriptions, detailing the contents, fields and relationships of the information so that they can be reconstructed by the user. Quite rightly, the discussion on the list concluded that without this type of documentation, the information provided could never be fully understood. A vitally important aspect of using someone else's data is the need to understand the rationale for their structuring and organisation of that data.

However, documentation of file structures is not enough. As has been pointed out many times (e.g. Reilly 1985) the records and observations that we as archaeologists make, our 'data', are theory-laden – the way we approach sites and deal with them is in a large part determined by the current state of archaeological inquiry. Anyone who has attempted to use old excavation reports will be well aware of what this means. Archaeological data are not the immutable truth that they are sometimes viewed as being. Data are recorded according to the theories and methodologies of the day, and data recorded for one purpose, or under one particular regime often cannot easily, if at all, be used for a different purpose. Data are not objective at all – observations are coloured by environment and perceptions together with levels of understanding and experience. In any database, different types of data will be missing, or recorded in different ways, all for quite valid reasons in most cases, but all of which limits the uses to which they can be put. There is still what seems to be a touching belief in the reusability of databases, but this is clearly problematic. Most databases are created for particular purposes; redefining that purpose after the event may require more than just a simple rearrangement of fields or restricted massaging of the content.

Realistically, both data and accompanying documentation (where present) will be of variable quality. It will be difficult to compare datasets recorded by different people separated in time and space. All information will be subject to what are generally a whole series of unstated assumptions about how we collect data and what type of data we choose to collect. On the other hand, the alternative would be no historical archive at all. However, we have to be very careful when approaching this type of resource, be fully aware of the context of the information and the problems that arise if data are removed from their original context.

4.3.2 Access to knowledge

Of course, such issues will be of only academic interest to those who do not have the means of access to this on-line information. After all, one of the key pre-requisites of democratising knowledge is to provide access to that information. There are two major means of disseminating information in a computer-mediated environment: on disk, or on-line. The disk-based solution is the computerised equivalent of microfiche, particularly given the increasing use of CD-ROM as a delivery medium. Its read-only nature offers a degree of security, but the continuing development of new standards makes it difficult to be sure that the disks can be read at all, unless they are mastered to the lowest common denominator.

It is no coincidence that developments in on-line access have entirely taken place in an academic environment, and that academics are the prime proponents of democratisation of knowledge. Academics are in a favoured position in that they have free access to the Internet, (although ultimately funded by their institution) but others are not so fortunate. The costs are not inconsiderable, although the appearance of companies offering gateways to the Internet is reducing the cost of access. But how democratic is a facility that is not freely available to all on the same terms, that depends on your affiliation to a particular type of institution or your ability to pay?

Perhaps the 'democracy' element of the democratising knowledge concept is being interpreted too literally here, but if the aim is to share and exchange information freely then the means to do this must be made available. There is the prospect of an unhappy distinction developing between the have's and the have-not's – an archaeological élite who have access to the international electronic highways, and those archaeologists who don't. The fact that this divide will to a large extent mirror the existing divide (imaginary or otherwise) between archaeologists working in higher education, and those working in field units makes this even more unfortunate. Although this appears to be an extreme image, in fact it already exists to some extent in areas like email access, and this clearly militates against true democratisation.

4.3.3 Ownership of knowledge

There is a more fundamental problem, which is connected to the ownership of data that are made available through published or on-line datasets. Who actually owns this knowledge? Archaeologists have long clung to the idea that the past belongs to everyone, that the data we derive are in the public domain once an (unspecified) time of study has elapsed to allow for analysis and publication. While some information may be withheld – to protect sites from looting, for instance – in general this knowledge is freely available. Yet paradoxically, access to that knowledge is often not without cost as a result of the publishing mechanisms that are traditionally used for dissemination. The problem is that once that same information is available via computer media, it can be freely copied without anyone knowing, modified, and changed to suit local purposes. In such an environment, the copyright and ownership of material becomes extremely problematic. As the Teaching and Learning Technology Programme (TLTP) in UK higher education is discovering, the issue of copyright when it comes to multimedia applications is extremely complex. How can you control access, restrict copying, limit manipulation, and so on in order to maintain your ownership and copyright intact within an environment that makes the breaching of that copyright so easy? Copyright statements may prohibit, but they cannot stop illegal copying. The chaotic, even anarchic, but jealously protected organisation of the Internet does little to help the situation.

At the moment, when approaches to existing copyright holders are made for permission to use pictures and other information, their response is in terms of traditional publishing methods and copyright in books and manuscripts. They want to know how many copies will be made, and look askance when told that once the information is released, no one will have any control over the number of copies. Furthermore, archaeological knowledge is not achieved without cost. Archaeological data are often perceived to be freely available, paid for out of public monies and yet rights to information are exerted by everyone from Government bodies, the Ordnance Survey, down to individual field units. Archaeological information often has commercial value, and access to such data costs money. Increasingly there are situations where service to public interest is outweighed or overridden by operational necessities. If a competitor in a difficult market place wants some archaeological information that a field unit collected, why should that unit give it away if it will help their rival compete more effectively against them?

Once information is published on disk or made available to the Internet community, control over its destination and use is lost. This raises issues of quality control which do not arise with traditional means of dissemination. As information is moved, copied, altered and becomes increasingly separated from its author in time and space, how is a viewer to know whether or not they are looking at the original, or some bastardised version? Is the database that has been downloaded the same as that originally mounted, or has it been massaged by some unseen person for unknown ends? And how can anyone tell? There is no way to mark originals as such that cannot be undone or that cannot be applied misleadingly to copies. The creation of centralised repositories such as that proposed by the Archaeological Data Archive Project would address the problem to some extent, but the very nature of the medium invites propagation, and the ephemeral nature of the digital form encourages a different attitude of mind to data.

4.4 Conclusions

The concept of democratising knowledge appears to be very exciting, and the processes involved certainly would represent a major cultural revolution in archaeology. The

exchange of information, the availability of data, the ability to call up remote databases, download excavation archives, re-analyse and compare datasets, can only be good for the study of archaeology, but we have to get it right. This new world of free access to and use of information will not be without its problems, and some of those problems could be so limiting and restrictive that this brave new world may remain nothing more than a few experimental islands of information. There is always a danger in new departures that the concept becomes so hyped and overblown that it becomes discredited, and it would be unfortunate if this was to happen with what could be such a radical development in a whole variety of ways for archaeology.

Archaeologists need to look closely at the issues raised by the concept of democratisation – they are issues that have to be addressed rather than swept under the table, otherwise we could find ourselves with a series of valuable but essentially useless resources on our hands, and also face severe problems over the ownership and use of that information. Some of these are problems that have been with us more or less since the beginning of time, in terms of archaeological computing – others are new issues that are raised by the potential of the technological advances that are becoming increasingly available to us. We need to make sure we capitalise on these possibilities, but in order to do so, we must have a clear idea of where we're going and the means of getting there before we start out on the journey.

References

FUKASAWA, Y. 1992. 'TRI: Three-dimensional imaging for recording and analysing stone-artefact distributions', *Antiquity* 66, 93–97.

HARNAD, S. 1991. 'Post-Gutenberg Galaxy: The fourth revolution in the means of production of knowledge', *Public Access Computer Systems Review*, 2(1), 39-53.

HUGGETT, J. 1993. 'Democratising archaeological knowledge with computers', *Archaeological Computing Newsletter*, 35, 7–10.

REILLY, P. 1985. 'Computers in field archaeology: agents of change?', in M. A. Cooper & J. D. Richards (eds.) *Current Issues in Archaeological Computing*, British Archaeological Reports, Oxford, 63–78.

REILLY, P. & Rahtz, S. 1992. 'Introduction: Archaeology and the Information Age', in P. Reilly & S. Rahtz (eds.) 1992 *Archaeology and the Information Age: A Global Perspective*, Routledge, London, 1–28.

RAHTZ, S., HALL, W. & ALLEN, T. 1992. 'The development of dynamic archaeological publications', in P. Reilly and S. Rahtz (eds.) 1992 *Archaeology and the Information Age: A Global Perspective*, Routledge, London, 360–383.

SHERRATT, A. 1993. 'Archaeology and post-textuality', *Antiquity*, 67, 195.

WOLLE, A. C. 1994. 'A tool for multimedia excavation reports', *http://avebury.arch.soton.ac.uk/Research/Anja.html*