6

The Greater London Sites and Monuments Record—a case study

Hugh Jones*

6.1 Introduction

CAA conferences are concerned with the application of computer technology to archaeology. However, the computing problems and opportunities which are presented by the Greater London Sites and Monuments Record can only be properly understood when placed in the context in which the record has developed and is used. For this reason this paper falls into three main parts. The first (sections 6.2–6.5) describes the organisation of archaeology and conservation in Greater London, and the way in which this has affected the development of the record. The second part (sections 6.6–6.9) goes into some detail about the hardware and software currently used, and the advantages and disadvantages which they bring. Finally (sections 6.10–6.11), the paper explores directions which might be taken in the future.

6.2 History of the SMR

Greater London was one of the last English counties to start developing a county-wide sites and monuments record. It was also one of the first to consider keeping information on archaeology and historic buildings on the same record. This is because the initiative for setting it up came from the Greater London Council, more particularly the GLC's Historic Buildings Division, where archaeology was very much subsidiary to a variety of work relating to standing historic buildings. This initiative came in the early 1980s, and happened when and where it did for a number of reasons: the GLC wished to have more sophisticated ways of handling the information it both used and generated in its work on historic buildings; there was an enormous amount of archaeological work, both past and present, whose results needed to be systematically available to a wide range of users; and it was recognised that almost every English county had or was developing a computerised sites and monuments record while London did not possess one. In addition, it was felt that the very real fear of abolition which was then looming meant that there may not long be an opportunity to establish such a record London-wide.

In 1982 the GLC funded a survey by the Museum of London to establish the range of source material for the archaeological part of the record. At the same time the GLC looked at how it wanted to manage its own information relating to historic buildings, and assessed the computing facilities available within its own Central Computer Services. In 1983, the GLC employed an SMR Officer to carry out the detailed design of the record, to set up and co-ordinate the team to work on it, and to oversee the creation of computer systems for it. Compilation started in 1984, and was, and still is, carried out by a team whose members are employed variously by the Museum of London, the Passmore Edwards Museum, and the Borough of Kingston

* English Heritage
HBMC
Chesham House
30 Warwick Street
London WIR 6AB

on Thames. Most of the funding for these posts came from the GLC, and to date the same arrangements have been maintained by English Heritage. On the abolition of the GLC in 1986, English Heritage took over the majority of the functions of the Historic Buildings Division of the GLC, which, with minor changes, became the London Division of English Heritage. The effects of this change are important to the future of the computing strategy, as will be shown later.

6.3 Conservation in London

One of the major determinants of the structure of the record, and the computer system which supports it, has been the range of interests which the SMR has to cater for. In London, the GLC had, and English Heritage now has, powers to direct local planning authorities in the granting or refusal of listed building consent. Elsewhere in the country, these powers are exercised solely by the Department of the Environment. There are 33 local planning authorities in Greater London (31 boroughs plus the Cities of London and Westminster), and these should get appropriate advice on the archaeological implications of planning applications from the Museum of London and the Passmore Edwards Museum, which use the SMR for this purpose. There is also the London Docklands Development Area in which planning rules are somewhat different from those elsewhere. London has, in the DUA and the DGLA, what are popularly believed to be the two largest archaeological field units in Europe, as well as the Passmore Edwards Museum active in North-East London, Kent Archaeological Rescue Unit in South-East London, and numerous museums and local societies. There is also the need in London, as elsewhere, to consider the interests of the major national bodies in the field, the RCHM and English Heritage, and the interests of academic researchers, and other public and commercial users. One way of making the SMR responsive to the needs of this wide constituency has been to base the members of the SMR team in a number of the organisations both providing and using data. Another has been the existence, from the beginning of the project, of an Advisory Group, on which representatives of many organisations can monitor progress and provide an input into future policy making.

6.4 The structure of the record

The information needs of the different interests identified above are many and various, but they had to be reconciled in a single working system. This required decisions on a number of points, and decisions according to a variety of criteria: some philosophical, some practical, some technical, and some no doubt entirely arbitrary. One of the first decisions to be taken was that it was appropriate to put both historic buildings and archaeological items on the same record. This was decided partly because it is often an accident of history whether something has survived to the present day as a standing structure, and partly because there can often be very real uncertainty as to whether something should be considered as an archaeological site or a historic building, The Tower of London for example, or Hampton Court Palace, where the debris resulting from the recent fire was removed archaeologically; there will also be times when the management interests of archaeology and historic buildings will overlap. Another decision to be taken was the level of detail which it was appropriate to allocate to a single record. The solution adopted was to take the idea developed in the North Yorkshire system, whereby each record has an equal status in the database, but groups of individual records may be related to one another as primary records, component records of a primary, or sub-components. This means that there is great flexibility, so that a single record can refer to a complex archaeological site at one extreme, or a single listed bollard at the other.

It had to be decided whether free text retrieval was desirable, and if so, if it was achievable. It also had to be decided what classes of information to record, and whether there would be a substantial storage overhead from having a large number of fields defined, many of which might be empty in any particular record. The choice was to record archaeological sites in a fairly standard SMR-type way, but include further fields particularly relating to historic buildings, and yet others reflecting the particular management responsibilities then enjoyed by the GLC. A decision was needed on what terminology to use, whether to use abbreviations or whole words, what to validate, how many indexes to define, and so on and so forth. However, the main point is that computer-based record systems end up looking and performing how they do for a variety of reasons. Some of those are indeed to do with the hardware and software used, but technical considerations are only part of the picture. The opportunities or limitations provided by any particular computing environment can often have less of an impact on the effectiveness of an application than other, non-technical factors, such as resources, objectives—and that includes not just the objectives set at the beginning of an exercise, but the ability to review those objectives and amend them accordingly—the corporate culture of the organisation running a record, the adequacy of communication with other relevant organisations, and the closeness of the match between the skills possessed by those working on the record and the skills needed to carry out the work effectively. For these reasons, while it is important to obtain the best possible computer system for any particular application, it is at least as important to take a holistic view of the problem to be solved, and ensure that whatever computing solution is chosen will form a coherent whole with the other factors which must come together to make an application succeed.

6.5 The contents of the record

In the circumstances in which the Greater London SMR was set up, it was decided that it was necessary to develop a large record. It is believed that the current phase of compilation, which began late in 1984 and is expected to finish in 1991, will produce something in the order of 65,000 records, of which rather under half will relate to archaeology. This 'basic record', as it is called, consists of all archaeological sites and isolated finds about which information has been available from reasonably accessible sources, and all statutorily listed buildings. The need to add to this basic record in the medium term has been reviewed, and the resources available to do this have been assessed, and it is believed that a good case can be made for doing further research into the less accessible or productive archaeological sources, and adding information on historic parks and gardens. Some people feel that the SMR should become a much broader-based record dealing with all aspects of the past of London, but the resource implications of this are enormous for a record which is already many times the size of an average county SMR. The potential size of each individual record on the system also had to be quite large. In order that the database could truly hold both archaeological and buildings data together, there is a single record structure to cater for all data. There is thus a significant amount of redundant space in every record. As there was a wide range of user interests identified from the start, there is a large number of fields defined.

6.6 The current computer service

These two things, the number of records, and the potential size of each record, meant that the computer system had to be quite large and powerful. In addition, the need for direct access to the record from a variety of locations around London meant that

there had to be good communications facilities. In these circumstances, a PC-based record was not a viable option. The GLC therefore chose, not surprisingly, to use their own in-house mainframe computer service for the record. This had the advantage that hardware and software were already in place, communications across a wide area were well established, and the cost of developing and running the system was internal to the one organisation.

There have been minor changes to the SMR computer system since development started, but these have been evolutionary. The computer system is run by the Central Computer Services, or CCS, of what used to be the GLC and then became the London Residuary Body. In 1988 Hoskyns plc acquired CCS as part of the process of disposal of all the former GLC assets. Since the abolition of the GLC, CCS' biggest customer has been the Inner London Education Authority, which has a number of applications including payroll, supplies, building maintenance, property register and management information systems. Other major users include the London Fire and Civil Defence Authority, the London Research Centre, and a number of the London boroughs. Since abolition, English Heritage has paid for the computer service on a 'cost only' basis—that is to say that the Residuary Body only covered its costs and did not make a profit. Hoskyns have agreed to increase their prices only by the rate of inflation for three years, and will introduce profitability by cutting operating costs. CCS are responsible for provision and maintenance of hardware and software for the SMR, and have day-to-day responsibility for all technical aspects of the record such as ensuring the availability of the computer service, security, data prep., and programming (apart from a certain amount of updating and report generation carried out by SMR staff). English Heritage retains, through SMR staff, all responsibility for the data in the database except where it has been corrupted by faulty CCS software, and for controlling access either to edit or to read the database.

6.7 Hardware and software

The computer itself, which is currently at County Hall on the South Bank of the Thames opposite the Houses of Parliament, is an IBM 3090 running the MVS/XA operating system, which controls the internal processes of the computer. Between the operating system and the application is a facility called TSO (which stands for Time Sharing Option)—this splits up the available computer time between all the users. The database used for the SMR is called ADABAS, which is produced by a company called Software AG. This was chosen as the more appropriate of the two database packages available on the County Hall mainframe. ADABAS is sometimes called a 'semi-relational' database—each database can be split into up to 255 files, but the information in each of those files need not be related to the other 254 in any way. The mainframe at County Hall has only two databases on it, and the SMR uses only two files out of the 255 available on one of those databases, one for the main descriptive material and the other for bibliographic and archive reference data, so in fact it is essentially stored in two intensively indexed flat files. Each ADABAS file can hold up to 16.7 million records, and each record is up to 500 fields long. ADABAS has an associated programming language called NATURAL, claimed by the manufacturers to be a fourth-generation language. NATURAL is used by SMR staff to produce customised reports from the record.

ADABAS has a number of features worth commenting on. As with many systems, screens can be set up to suit the needs of the user, and information can be retrieved using either a menu-driven or a command-driven system. The menu-driven system is easier to use but slower and less flexible, while the command-driven system is more suitable for experienced users. A maximum field size is defined for ADABAS fields, but the database only stores what data is put in, and not the blank space in

vacant or half-empty fields. Currently, only about 20% of the total available space in the SMR database is actually used, so this facility makes for much more economical storage than would otherwise be possible. It also allows records to contain free text, although that free text is not indexed or otherwise retrievable in any way.

Indexed fields are called descriptors, and it is possible to set up descriptors at any time, after data has been entered as well as before. It is also possible to set up indexes on a specific combination of fields or on a specific part of a single field: these are called super-descriptors and sub-descriptors respectively. It is possible to search on fields which are not indexed, but this requires a sequential search through the whole database, and is very expensive, not just in terms of machine time and disk accesses, but also in cash terms. Certain categories of data may recur an unpredictable number of times in any particular record. ADABAS caters for this through multiple fields or periodic groups. These are single fields, or combinations of them, which may be so defined that they will be present in any record the number of times that there is data to go in them. They must however be used with caution, as it is not possible to sort records according to data categories defined as multiple fields or periodic groups.

6.8 Computing blind alleys

Before stating the advantages and disadvantages of the current computing arrangements, and by that is meant not just the technical but also the organisational factors, two computing blind alleys should be mentioned which have been explored in the past. The first is STAIRS, a free text retrieval system. This was the second of the databases available from CCS whan the SMR was set up. There was a pilot study to assess the value of STAIRS using some of the text from the statutory lists of historic buildings. The particular problem which ruled it out was that it was designed for very static data, and therefore did not possess any in-built editing facility. In order to amend any typing errors, or make any of the changes or additions which happen to the lists from time to time, it was necessary to download the whole file into the somewhat primitive text editor available on the mainframe, make the necessary changes, and reload the file. This was not a workable system.

The second piece of software which has been tried and rejected is called FOCUS. This is the report generator element of another database system, which is available without the main part of the FOCUS package and can be interfaced to other databases. We have experimented with it because it has 2 modes of operation, menu-driven and programmed, whereas NATURAL, which is currently used, has no menu-driven report generation facility. The FOCUS menu-driven system was indeed easier to use than the programmed version, but was also very much less powerful, so much so that it could not perform the required tasks. Rather than re-train the relevant staff to program in FOCUS, we have elected to continue using NATURAL.

6.9 Advantages and disadvantages of the current computer system

The following are seen as the advantages of the current system. It is designed for London-wide use, and for large records. It is run by a large computing department which can afford to have specialists in different aspects of hardware and software. ADABAS is a mature product and is not subject to frequent bugs, and it can mix structured data and free text and provide a range of facilities such as data validation and sophisticated indexing, including the ability to create new indexes on an existing database. New fields can also be defined and added at any point. System security is also well established—as well as regular back-up procedures, access to databases is properly controlled according to need.

The disadvantages are, firstly, that the system is expensive. Secondly, its future is uncertain under new management. Thirdly, the level of service cannot be guaranteed. Users are obliged to accept the advice of CCS on technical matters, and work has to be paid for even when it is abortive—the arrangement offered is that payment is for time expended, not results achieved. The SMR database is perhaps more complicated than, and certainly different in kind from, the majority of databases managed by CCS, and programs do not always seem to be adequately tested before they are run. Where these programs make significant changes to the database, more harm than good can result, causing extra work for SMR staff who have to make good the damage. Every user, including the SMR, is subject to problems affecting the whole mainframe system—there are many ways in which such a system can be brought down, and from time to time the system is unavailable for use, and we have no control over when it comes back up. On a more specific level, there are genuine problems caused by the inability of ADABAS to sort on multiple fields, and there is not sufficient flexibility in the formatting of printed reports to allow them to look how we would wish.

6.10 The future direction of Greater London SMR computing

The disadvantages of the current computer system are now believed to outweigh the advantages. There is also a changed standpoint since the SMR has been run by English Heritage. Under the GLC, there was a need to set up a system, and get it running, quickly, before abolition. Cost was perhaps not as important a consideration as it is now, as the computer service was free at the point of use. An awareness of the true costs of mainframe computing might make many archaeological users of mainframe computers think again about the appropriateness of that technology. In English Heritage, unlike the GLC, conservation is the primary function, and the organisation is already involved in a considerable way in archaeological computing, in such areas as the Central Excavation Unit and the Ancient Monuments Laboratory, as well as the initiatives in SMR computing, including grant-aid, as well as the support of the so-called 'Version I' software and Superfile.

In this context, English Heritage is considering the computing strategies of three organisations: English Heritage, the RCHM and the Museum of London. In particular, we are looking to the agreement on computing standards reached by English Heritage and the RCHM to guide the future development of the SMR. The objective is therefore to develop a replacement computer system in-house for the SMR, based on the relational database ORACLE. It is possible, with many SMRs now outgrowing their current computing environment, PCs getting more powerful and cheaper, and ORACLE now available on them, that this may become a new standard in SMRs.

It is not easy in the public sector these days to specify computer hardware for more substantial systems, due to the Common Market GATT regulations which require such purchases to go out to tender. English Heritage will look to follow the RCHM's lead in getting DEC hardware, with either the VMS operating system or a version of Unix, but failing that, there are many other hardware platforms for running Unix. The SMR will wish to retain as many of the advantages as possible of the current system, while removing as many of the problems as possible, and in particular we consider it essential to allow direct access from different locations around London. Another important development will see the implementation of easy-to-use enquiry systems to cater for the most common types of query.

6.11 Mapping

In conclusion, some brief ideas on potential developments in mapping. The items recorded in the SMR are currently marked on film overlays attached to 1:1250 or

1:2500 scale O.S. maps. This is adequate for the moment, but no more, and in the medium term there will be a need to develop a new system, which it is hoped would be computerised. Two systems are immediate possibilities. One is available from CCS at County Hall and is based on the Ordnance Survey's own computerised map base—London will be one of the first areas of the country for which the entire map base will be available. At the moment, we fear that the cost and level of service which, based on our experience with the SMR, we might expect, will probably rule out this option, but we are watching its development with interest. The other possibility is the system based on raster-scanned maps used in the English Heritage Records Office in Fortress House to record the location of scheduled ancient monuments. At present the map coverage is only available at 1:10000, but such a system with large-scale maps could form the basis of a solution for the SMR. In the long term, there is no doubt that SMRs will make use of fully-fledged Geographic Information Systems, preferably in conjunction with many other categories of data, but it may be some time before such systems are fully functioning in a form which is affordable to most archaeological users.