8

Pooling Resources: drawing on regional data from the computerized Sites and Monuments Records of the East Midlands

Roger Martlew*

Barny Creaser†

8.1 Introduction and background

County Sites and Monuments Records are the only detailed, constantly updated records of archaeological sites available at a local level. As well as their primary role in cultural resource management, SMRs also provide a major source of data for academic research and for educational use. However, a great deal of time and effort can be involved in compiling data from SMRs for research, especially when the records of several counties are involved. The computerisation of SMRs is improving internal access to information for planning enquiries and for local fieldwork groups, but the independent development of many Records restricts their usefulness to within their own county boundaries. Requests for information on a wider, regional scale are faced with almost as many difficulties as in manual systems. It is certainly easier for a list of sites to be printed out from a computerised Record, but with the large numbers of sites which can be involved at a regional level the end result can rapidly become unmanageable. The transfer of data in machine readable form is the obvious solution in this case, but differences of record type, hardware and software between counties create inevitable complications.

This situation has arisen because the need for information on a regional scale was not forseen when systems were being set up. Planning procedures operate within local authority boundaries, and existing archives reflected this. At that time it was a normal and accepted part of academic research for students to have to visit a number of different locations in person in order to collect information. The practical difficulties of this research activity ensured that it was not widely taken up. Following the Information Technology revolution, it is still necessary for individual locations to be contacted and in some cases visited in order to obtain printouts of information. The practical difficulties of fitting this into the timetable and workload demands of an undergraduate course ensure that these significant resources are still only used occasionally in dissertation work. It is hardly surprising that students provide a continuing source of irritation for SMR officers, with ill-formulated requests for 'anything you've got on the Bronze Age'.

The situation in universities contrasts with recent developments in education at secondary level. With the advent of GCSE, over 50% of British state schools are now using the Schools Council History Project, which was first taught in the 1970s. A major component of this project is the use of primary evidence to enable pupils to form their own opinions of an event in history. Primary evidence is used in

* Department of Archaeology,
University
Leicester LE1 7RH

† Department of Archaeology,
University
Leicester LE1 7RH

conjunction with secondary sources to explain how history, or indeed any report inevitably includes elements of bias. An important aim of this project is to encourage pupils to use their own judgement in dealing with several strands of evidence. The emphasis is on process as much as on content—pupils are not just memorizing strings of dates and events. With history being a designated foundation subject in the National Curriculum, an increasing number of entrants into universities will have come into contact with this approach to the use of primary evidence.

An educational pack has recently been produced which uses computerized SMR data as primary evidence in a simulation exercise for secondary level pupils. *Digging Deeper into History* (Martlew 1989) provides an insight into a county archaeologist's job of cultural resource management. This includes the extensive use of SMR data drawn from the North Yorkshire and Humberside SMRs, and made accessible to pupils by computer. The information is 'packaged' for a specific audience and for a specific purpose, but only a slight alteration of the 'packaging' would enable the simulation to be built into an undergraduate course. Similarly, the approach could be adapted to SMR data from other parts of the country, providing time and expertise are available to 're-package' the basic data.

8.2 Leicester University Regional SMR Database.

At Leicester University a computerized system has been set up which builds on the experience of compiling the datafiles for *Digging Deeper into History*. The system, on the university's VAX mainframe, consists of a database and a menu-driven suite of programs which allows data to be processed, analysed and plotted. In a professional context, the system provides a mechanism for the exchange, compilation and plotting of data from East Midlands counties. In an educational context, it also enables sets of regional data on specific themes to be built up for use by undergraduates. Working with such a resource can teach students the information handling skills which they so obviously lack at present. It also gives them experience in solving archaeological problems at levels which were not previously possible.

8.2.1 The database

Developments in the provision of computing facilities at universities seem to be moving towards increasingly distributed processing power. The main role of the central mainframe in the future is likely to be as a mass storage device, with a limited range of very large number-crunching applications. The information which is being compiled in the Leicester database is currently only that which is needed for teaching, but there is no technical reason why entire SMR databases could not be mounted given the scale of disc space which is potentially available. Such are the developments in mass storage at present that the same could equally be said of microcomputer-based storage devices such as CD-ROM. The main constraints are not imposed by the technology, but by politics and lack of funding. The project at Leicester University is in the nature of a pilot scheme to show what is possible when these constraints can be overcome.

The first stage in the establishment of the teaching database at Leicester was the compilation of machine readable data from SMRs, the counties involved being Leicestershire, Nottinghamshire, Derbyshire, Lincolnshire, Northamptonshire, Warwickshire, Humberside and Cambridgeshire. This grouping is based primarily on membership of the East Midlands SMR Working Party, with a few additions to fill in the irregular boundary of the region. The computerization of the Records in these counties is at different stages, and different approaches have been adopted in different counties. Leicestershire, for example, has a fully computerized Record, although

the information consists of a one-line index to the paper records. Lincolnshire is just starting on the computerization of its SMR, and information is only available for a couple of districts. Different hardware and software configurations are in use throughout the region: the Leicestershire record is on the County Council mainframe and is accessed by a program called MAPPA, while other counties use Superfile or dBase on microcomputers.

To obtain a test sample of data from each county, SMR officers were asked to supply information about Roman villas, either on a 5 1/4 inch IBM compatible floppy disk or on magnetic tape. There was a mixed response. Out of the eight counties from which data was requested, three sent data on floppy disks, three set data on printouts and two were unable to send any data since their SMRs were in the early stages of computerization. One of the counties which sent data on a printout is now able to put data on disks, so the situation is improving all the time.

The major problem in getting data from SMRs was the large amount of time which elapsed between requesting data and receiving it, and this is largely a consequence of the unusual nature of the request. Some SMR officers, especially when the SMR is kept on the county mainframe, have to go through other departments in order to put data on disks, and this process is time consuming. Some SMR officers had no experience of dumping data from a database onto a disk, so extra time was needed for them to sort this out. However, delays were also experienced in receiving printouts, and this is entirely due to the pressures on the time of overworked SMR officers.

The machine readable data was put onto the Leicester University mainframe using the Kermit package. This public domain software enables data to be transferred between computers in the form of ASCII standard text files. Once on the mainframe, the data was manipulated using Ingres, an SQL (Standard Query Language) based relational database system. Ingres, and similar packages such as Oracle, are also available on microcomputers, but the scale of the project required the use of Leicester University's central computing resources. Each county's data was read into a separate table within Ingres, preserving the record structure of the files. The key field, which appears on all the tables, contains the National Grid Reference. This approach makes *all* the original information from the SMR available to the system, and selections can subsequently be made in a flexible way. A different approach is being developed by HBMC, in which a standard grouping of fields is selected for transfer from each local database. This avoids the political and administrative problems of centralizing the information resource initially, as in the Leicester project.

Incompatibility of record structure between counties in the Leicester system is overcome relatively easily, but incompatibility in the cataloguing terms used remains a problem. Qualified retrievals require a fairly detailed knowledge of the keywords used by each SMR. Ingres tables can be searched either by using pre-determined forms, or by building up more flexible retrieval criteria using SQL. The program can create reports and produce graphs as output. In the Leicester system Ingres is also used to create data files for use with the menu-driven suite of processing and analysis programs.

8.2.2 Processing and analysis

The main aim of the Leicester project is to enable students to retrieve data easily for analytical work at a regional level. Students can appreciate the problems of collecting and using data of this type most effectively by having to use it themselves to support their own investigations and interpretations. A menu-driven suite of programs uses the grid references of sites drawn from the Ingres database to analyse site distributions and to produce simple distribution maps. Because different SMRs in the East Midlands record grid references in a variety of different ways, software had to be written to convert grid reference data into a common format. This data can

then be analysed using simple spatial analysis programs included in the package, or plotted onto a basemap of the East Midlands using GIMMS (Geographical Information Mapping and Modelling System). Standard base maps showing county boundaries, rivers, roads etc. are available for GIMMS. The main base map for the Leicester project was obtained from the University of Manchester via the Joint Academic Network (JANET), and consisted of administrative boundaries in England, Scotland and Wales down to ward level. Site locations can be plotted onto these maps, and discrete areas can be shaded according to data frequency. Maps can be plotted at different scales, corresponding to scales used by the Ordnance Survey, and GIMMS can also be used to calculate contours and search specific areas. The processing and analysis programs are user friendly, being designed for undergraduates with little or no computing experience.

8.3 The East Midlands SMR system: a test application

The system which has been developed at Leicester has been applied in a study of Roman settlement patterns in the East Midlands, using evidence contained in computerized SMRs. Non-computerized sorces inevitably had to be used to supplement the data for some areas, but as the computerization of county Records is an on-going process, it was felt that the work would usefully identify problems and potential for future applications. Using the system, a test sample of data on Roman villas and small towns was selected from the database and analysed. Modern county boundaries are assumed to bear little or no relationship to these distributions, and their investigation on a regional level has proved useful in the past. Patterns were detected by the analysis programs in the East Midlands data, but the main significance of the test application was the ease with which the process of hypothesis formulation and testing could be repeated. Having gained experience through such a system, students would bring a greater degree of research skill to their own projects and dissertations.

The use of the Leicester system in undergraduate teaching is dependant on a basic level of computing skill. Although the system is set up to be as user-friendly as is possible on a mainframe, the use of Ingres, GIMMS and the various statistical programs requires a knowledge of computerized information handling that most archaeological undergraduates do not possess. There is a lot of groundwork in basic computer literacy to be done (for both students and staff) before the system can be easily integrated into teaching.

8.4 Conclusions

The pilot project at Leicester has demonstrated the feasibility of creating a regional database using SMR data which allows the data to be manipulated, interpreted, and plotted and which also facilitates the adding of data in a machine readable form from SMRs. The availability of such a resource to academics would improve the efficiency of research, and would be beneficial to students by allowing them to have access to primary evidence. It would also encourage more constructive use of SMRs by both students and researchers, although there are serious manpower implications in adding to an already stretched service.

Hart (this volume) discusses the NAR's new ONLINE system. This, together with the data transfer standard proposed by English Heritage, reveals a growing interest in using new technology to improve the widespread flow of archaeological information. Although this has long been recognized as a worthwhile goal (Cleere 1984), it is something which is still a long way from being achieved. These latest developments will make the aquisition and transfer of machine readable data much easier, providing

a more efficient service for cultural resource management and research. However, there is still a need to improve information handling skills among both professional archaeologists and undergraduates, and merely providing access to large banks of data will do little to promote this. One example of the end-use of computerized SMR data (Martlew 1989) has shown the need for it to be carefully packaged for specific applications, particularly in education and training.

The ready availability of computerized resources such as SMRs to schools, colleges and universities in 'value added' forms can make an important contribution to the future of archaeology. The 'green' movement has already demonstrated that school and university education can be used with great effectiveness to get a message across, and to build up public and political support. In addition to improving professional access to information, the public image of archaeology could similarly benefit from this application of information technology to regional SMR data.

Bibliography

CLEERE, H. 1984. '"Only Connect"'. *in* Martlew, R., (ed.), *Information Systems in Archaeology*, pp. 9–20. Alan Sutton.

MARTLEW, R. 1989. *Digging Deeper into History*. English Heritage.