AN EXPERIMENT IN THE USE OF A COMPUTER FOR ON-SITE RECORDING OF FINDS

by

Paul Buckland
York Archaeological Trust

Abstract

A site information retrieval scheme using a remote terminal at the DANVM *Roman fort site in Doncaster and an ICL 4-50 computer at North Staffordshire Polytechnic is reviewed. Field trials were carried out in 1972 using the* BASIC *language.*

The Doncaster project was designed at the outset to have a hierarchical data structure, allowing for unabbreviated plain-language keywords. Field trials revealed several points requiring revision in the final system, and these are detailed below.

1. A full time operative for the teletype is essential on a complex urban site since both director and pottery assistants are too involved with the excavation, washing and marking of finds to have time to deal with what can be a rather temperamental piece of equipment. Much of the trouble arises from the use of "noisy" telephone lines for transmission to the computer.

2. The hierarchical scheme, devised as a minimum means for the recording of groups of finds on urban sites in the East Midlands or North East, proved too rigid and complex for rapid use. In contrast, particularly amongst the Medieval and later material, grouping was found to be too broad to be of much use when information was required from the data bank; *e.g.* the entry 'BROWN GLAZE','SHERDS','1','RIMS', '2','*' could refer to any material of Late Medieval to modern age. The intention of the scheme was to facilitate comparison; the use of such broad categories defeats this object.

It is difficult to devise a system which would be sufficiently objective without becoming too dogmatic; modification must therefore progress with experience.

3. Despite the necessity for more objective detail, many entries need radical shortening in line length to reduce typing time. It is also evident that the teletype operative must be thoroughly conversant with the system before it is brought into operation, otherwise much time is wasted correcting typing errors; there are few typists familiar with archaeology and our attempts to recruit one failed.

4. Utilising an experienced typist, the combined errors of both archaeologist and assistant resulted in about 25 groups of roughly 20 items each being listed in two hours. When one is dealing with up to four times that number of groups in a day, often with a far greater number of sherds in each group, this is too slow for total recording of all finds from a large site or group of sites.

5. In order to improve the system so that it is easily applicable to any urban site, a series of "keywords" must be developed which are readily understandable to all archaeologists who may want to use the system in a given area. The present hierarchical system requires too great a degree of mental regimentation, something which is rarely met with in archaeologists.

6. The code system of find recording (*e.g.* 'DX' (site letters), 'AA' (group letters), 'S' (stratified), 'B23' (trench no.), '4' (layer number), 'C9/5' (correlation)) is suited to computerised recording, but it is necessary to obviate the repetition of this line before each individual piece description in each group. This was achieved by the use of an 'R' repeat character as a prefix to each line. To speed the process of tape punching an improvement would be to omit this facility and amend the program to check the first item punched and make the logical test:

'if not 'DX' (site letters) then repeat line 'DX' et seq.'

Line numbers are a further retarding feature. These unfortunately may be essential in some languages *e.g.* BASIC. The character 'P' (pottery) or 'S' (small find) can be omitted if the type of find is explicit in the group codes, as was the case with the codes used, 'AA' - 'ZZ' and '1' - 'n'. In order to achieve this the program modifications 'if AA to ZZ then pottery' and 'if 1 to n then small find' become necessary for use when recovering data. Animal bone was incorporated into the scheme as a separate type of finds in the later stages of the trials, and this can easily be extended.

7. The preceding broad grouping character (*e.g.* 'RB' or 'MED') can in many cases be omitted when it is obvious to the user that the piece can only be of one period. It is therefore necessary to write this into the program for use when lists of the stratigraphic distribution of certain types or forms are needed. Thus modifications such as 'if Samian then RB' or 'if S(outh) Pen(nine) Grit(ty) then Med' are required, complicating the original program but quickening the site recording considerably. A typical group entry could then be:

```
(1000DATA)'DX''FS','X','B15','12','C6/2','*'
(1001DATA)'DELFT','4','PIPE STEM','5','GLASS','3','*'
(1002DATA)'S.PEN.GRIT','SH','2','R','7','B','2',
          'B-THUMBED','1','*'
```

N.B. '*' is a terminator which indicates the end of a record.

For such grouping the necessary program modifications, omitting abbreviations, to obtain a printout which is useful to persons other than the devisor, are:

```
if 'DELFT' then 'POST MED'
if 'PIPE STEM' then 'POST MED
if 'S.PEN.GRIT' then 'MED'
if 'DALES' then 'RB'
if 'HUNTCLIFF' then 'RB'
```

Where RB1-8 is specified this becomes an unnecessary duplication as the program can be arranged to search for any RB character.

'CAL.GR' (calcite-gritted ware) is more of a problem since vessels in this fabric occurs from the Iron Age to the Medieval period (*e.g.* 'CAL.GR', 'R' '1', 'RB'.) and it is therefore necessary to add the period character; it is frequently not possible to ascribe a base ('B') or body sherd ('SH') to a specific period.

8. It is also essential to allow for entries in abbreviated form but printouts in fuller form in order to permit wider reference amongst users who may not be fully familiar with the terminology:

if 'N.G.127' then print 'NEAR GILLAM 127'.

This however, may create minor difficulties with line length.

9. Some of the more complex descriptions spill over onto two lines; whilst this is immaterial in printouts, in punching tape it would be a considerable advantage to restrict as many as possible to single lines or less, *e.g.*

```
    (1004DATA)'ROMAN','COARSE WARE','GREY BURNISHED
              WARE','BOWL','RIM',
    (1005DATA)'FLANGED','3','*'
```

requires reducing to:

```
    (1004DATA)'FLANGED','BOWL','GREY','3','RB5-8',
              '*'
or: (1006DATA)'ROMAN','COARSE WARE','JAR','RIM',
              'LID-SEATED','DERBYSHIRE WARE',
    (1007DATA)'1','RB4-6','*'
```

can be cut to:

```
    (1006DATA)'DERBYSHIRE WARE','R','1','RB4-6','*'
```

The necessary alterations being:

```
    if 'DERBYSHIRE WARE' then also 'JAR' and
    if 'R' then print 'RIM'
```

```
    (1008DATA)'ROMAN','COARSE WARE','WHITE OR BUFF
              FABRIC','JAR','RIM'
    (1009DATA)'CORNICE','1','*'
```

may be reduced to:

```
    (1008DATA)'RB','JAR','R','CORNICE','BUFF','1','*'
```

or if the period can be implicit, then:

```
    (1008DATA)'JAR','R','CORNICE','BUFF','1','RB2',
              '*'
```

(The RB being contained within a later character).

It is obvious from the above discussion that the whole program has to be written by either an archaeologist trained in programming or with close liason between the archaeologist and the computer expert.

10. To clarify these suggested modifications, a projected restructuring of the hierarchy in relation to Mortaria is here included with the necessary alterations to the program given in simple terms:

```
    'MORTARIUM' (if 'MORTARIUM' then 'RB')

(if stamped then 'S')

    'N.G.n' 'G.n' (see list below) 'SH' 'B'

    (where n is a number)              'PAINTED'
'OX''COL''VER''MID''LIN''NENE''SWAN''E.Y.''S.Y.''-'

                        No. of
                        'RB1-8'
```

The abbreviations for untyped material descriptions are listed below with the program instructions:

```
if 'B&R' then print 'BEAD & ROLL'
if 'R.B&R' then print 'REEDED BEAD AND ROLL'
if 'FL' then print 'FLANGED'
if 'R.FL' then print 'REEDED FLANGE'
```

(Care must be taken to see that abbreviations are not duplicated)

```
if 'H' then print 'HAMMERHEAD'
if 'R.H.' then print 'REEDED HAMMERHEAD'
if 'W' then print 'WALL-SIDED'
```

The suggested source contractions are:

```
if 'COL' then print 'COLCHESTER'
if 'VER' then print 'VERULAMIUM REGION'
if 'LIN' then print 'LINCOLN'
if 'MID' then print 'MIDLAND'
if 'NENE' then print 'NENE VALLEY'
if 'SWAN' then print 'SWANPOOL'
if 'E.Y.' then print 'E.YORKS'
if 'S.Y.' then print 'S.YORKS'
```

(Crambeck may be preferred)

To many archaeologists working in the area these abbreviations will be self evident, but reference may be necessary by students not thoroughly acquainted with the system and by this means the onus of remembering such abbreviations, some of which may be rather abstruse, is transferred to the data bank.

A typical abbreviated entry for a mortarium could then be:

```
    (1010DATA)'MORTARIUM','R.H.','MID','L','RB5-6',
              '*'
or  (1011DATA)'MORTARIUM','B&R','PAINTED','SWAN','2'
              'RB7-8','*'
```

Such a system could considerably speed the entry of data, provided the users were thoroughly conversant with it before it was put into practice.

11. One eventuality which must be allowed for, since archaeologists, apart from accepted abbreviations like BB1 for Black Burnished Ware type 1, do not talk in monosyllables, letters and numbers, is the occasional appearance of terms in full rather than in contraction; thus when a retrieval request is for 'SAMIAN','DR31','R','*', the computer must also be able to pick out these terms when they occur in full, *i.e.* 'DRAG.31' and 'RIMS'. Any differentiation between singulars and plurals must also be avoided.

All this gravitates towards an extremely complex initial program and makes considerable inroads into the hierarchical system but it is a far more viable proposition from an archaeological point of view.

12. For small finds a less rigid system may be employed. The listing system already employed on many sites fits into a logical pattern for data processing: *e.g.*

```
    'DX'(site letters), '97'(find no.),
    'S'(stratified), 'L5','7'
    'AE'(bronze), 'COIN','TETRICUS','*'
```

would be sufficient, utilising the same scheme as the pottery grouping but a numerical list of find numbers as detailed above. For some coins types or diameters rather than emperors must be specified depending on the degree of corrosion or wear. It may also be possible to include in the data bank a list of dates of the form 'if 'TETRICUS' then also print 270 - 273' but this may seem an unnecessary refinement, although particularly useful for medieval coin types.

For other small finds, provided the sequence - metal or other material, identification - is adhered

to, data retrieval is a simple matter, *e.g.*

```
(1012DATA)'FE'(iron),'SPEARHEAD','RB','*'
(1013DATA)'BONE','COMB','SAXON','*'
```

Further elaboration of this part of the recording system can be made if necessary, although the scheme is for a regimented list rather than a hierarchy.

Conclusion

This project was given a trial run during the Department of the Environment rescue excavations in Doncaster at Easter, 1972, and although the experiment was beset with both equipment and electrical problems, as well as a general shortage of staff, it was soon evident that it has a tremendous potential in accelerating the processes of recording on complex urban rescue sites, where the comparison of groups from layers on different parts of the site could be carried out rapidly from data printouts. The necessary marshalling of information on finds preparatory to writing the report can also be considerably speeded with this system. Of equal importance is the fact that the basic site information can be stored on magnetic tape for later reference by any party wishing to check information in the published report. Once such a scheme is implemented, it becomes easy to provide such facilities as histograms of coin frequencies, dot-density diagrams of the spatial distribution of specified finds over the site, etc.

The computer can be utilised as a recording as well as a mathematical tool; it is in the former archaeologically least developed capacity that it can prove of most use to archaeologists labouring under vast bodies of finds information too large and too complex for his poor brain alone to comprehend and organise.

Appendix

A rudimentary division into periods is suggested for Roman material:

RB1 - pre Flavian

RB2 - Flavian to Trajanic (*circa* 70-120)

RB3 - Hadrianic to mid Antonine (*circa* 120-160)

RB4 - Antonine to early Severan (*circa* 160-210)

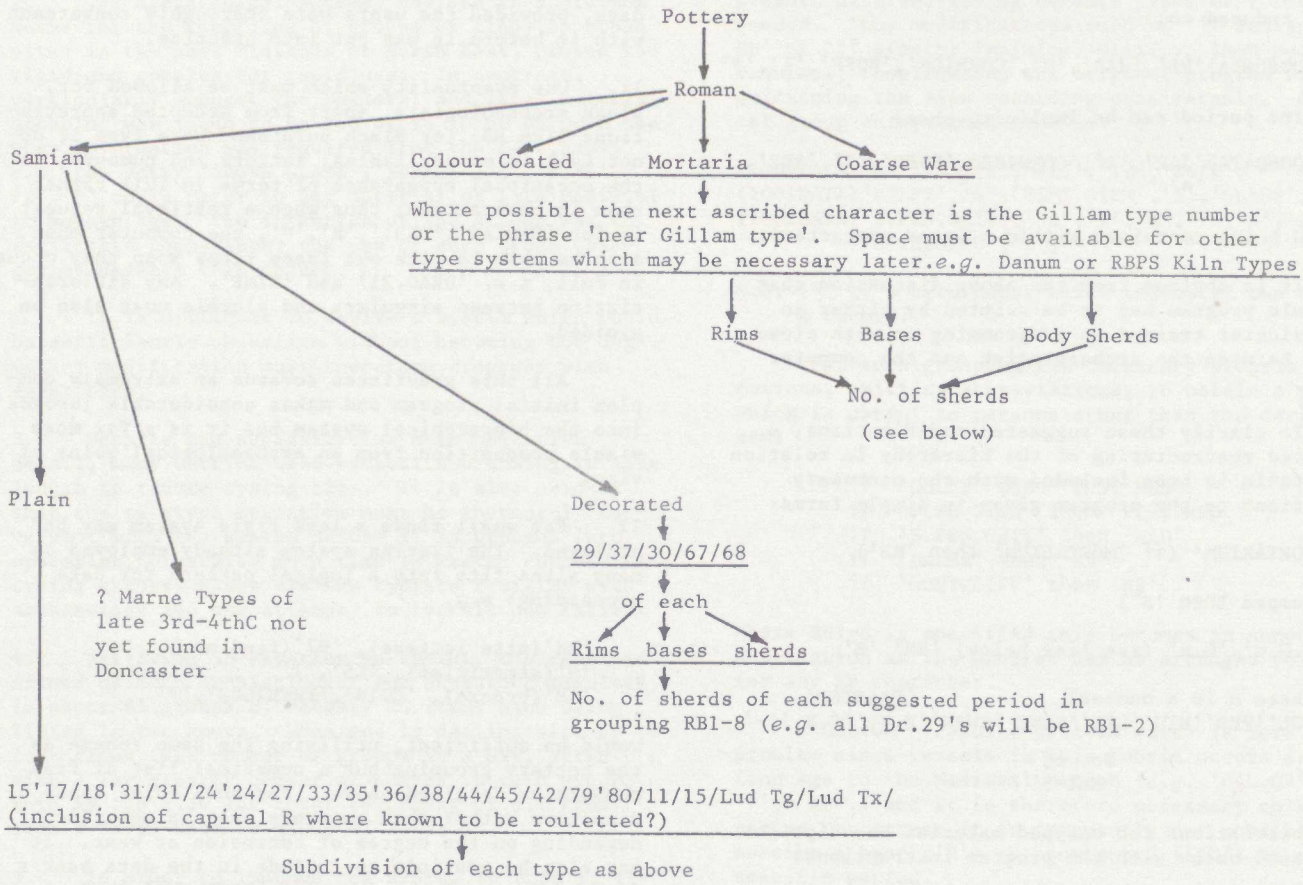
RB5 - Severan to late 3rd Cent. (*circa* 210-290)

RB6 - Late 3rd to early 4th Cent. (*circa* 290-330)

RB7 - Early to late 4th Cent. (*circa* 330-365)

RB8 - Late 4th to early 5th Cent. (*circa* 365-410+)

These are estimates for the general information of the user and are not objective rigid periods *i.e.* RB1 - 3 may occur, or even broader limits for some forms.



Paul Buckland,
York Archaeological Trust.

THE USE OF REMOTE TERMINALS FOR ARCHAEOLOGICAL SITE RECORDS

by
J.D.Wilcock
Dept of Computing, North Staffordshire Polytechnic

Abstract

The paper gives a general review of facilities used by archaeologists to transmit site data to a remote computer. The author collaborated with P.Buckland in the setting up of the Doncaster information retrieval project, which is also briefly described.

Communications networks are being used to an increasing extent nowadays for the transmission of data between computers, and between a user seated in his own laboratory and a computer which may be hundreds of miles away. The Post Office DATEL system is an example of a communications network which uses normal telephone lines. A 'remote terminal' consists of a teletypewriter, and a transmitter/receiver called a MODEM with a special telephone handset with keys to transfer between the voice and data transmission modes (or more conveniently an acoustic coupler into which a normal telephone handset can be placed). More specialised devices such as printers, plotters and card readers may also form part of the remote terminal. The user telephones the computer bureau in the usual way, waits for the transmission carrier wave to be connected, either by the operator or automatically, then presses his 'data' key or inserts the telephone handset in the acoustic coupler. The remote terminal then effectively becomes part of the distant computer. What this means for archaeologists is that excavation records can be entered into a computerised data bank from a teletype in the site office - the only services required are a power point and a telephone.

Archaeological data was transmitted by communications network at least as early as 1968 (Wilcock,1969). Newman (1969),with characteristic American flair, transmitted data from a site in Hawaii to a computer in California via communications satellite, while Gaines (1971) has described the successful day-to-day use of a computer in an archaeological field situation where the archaeologist concerned had not received any previous computer training. Thus the routine use of a computer for data recording, analysis and publication is a practical possibility right from the moment the artefacts emerge from the ground.

During the 1972 season of excavation at the DANVM site (Roman Doncaster), an experimental hierarchical keyword system was used for the recording of pottery and small finds. Records were checked automatically on input, thus guarding against the omission of data and the insertion of inadmissible data. It is an interesting point that the archaeologists, initially in favour of plain-language insertion of data with a minimum of stylised coding, quickly devised abbreviations for use in coded input. Using this system it is still possible for the computer to provide the equivalent plain-language printouts. During this exercise the teletype was located in a building near the site and the telephone was used to transmit data to a computer at North Staffordshire Polytechnic, Stafford, about 70 miles away from Doncaster. Suggested revisions to the system are detailed in the associated paper by P.Buckland.

References

Wilcock, J.D., 1969 — Computer analysis of proton magnetometer readings from South Cadbury 1968 - a long distance exercise *Prospezioni Archeologiche* 4, 85-93

Newman, S., 1969 — Correspondence in *Newsletter of Computer Archaeology* 5(2), 1

Gaines, S.W., 1971 — Computer application in an archaeological field situation *Newsletter of Computer Archaeology* 7(1), 2-4

J.D.Wilcock
Dept of Computing,
North Staffordshire Polytechnic,
Blackheath Lane,
Stafford.