

Stansted inquiry: indexing documents using a microcomputer

Martin Woodrow

County Information Librarian, Hertfordshire Library Service

The Stansted Inquiry needed indexes to 1800 documents in 8 locations, 3 of them not local; some of the documents were generated in the early months of the Inquiry itself. An Apple microcomputer was used, with the D.B. Master software package including a Report Generator.

There were 8 items of input for each document, viz title, personal author, corporate author (e.g. a Council), subject, type of document, its date, location, and number, and the machine produced 5 index volumes, one for each of the first four of these and one for the last.

Details of material preparation and input routines are described, with precautions that had to be taken.

The Hertfordshire Library Service originally purchased an Apple microcomputer to replace an ageing teletype terminal which was used to access the bibliographic data bases available from Europe and the USA. I proposed purchase of a microcomputer as it would allow us not only to access these data bases but also to build up various information and housekeeping files for internal use.

Just after obtaining the micro the Library Service was asked to assist in the Inquiry on Stansted Airport that the County Council were involved in, by indexing the documents that were being gathered ready for the Inquiry and those that would be generated by it, especially during the first few months. As the person in overall charge of the departmental libraries in County Hall, I was asked to take control of the project. Naturally, having just acquired the Apple microcomputer, my thoughts quite quickly turned towards it and its possible use for this project. It was not simply the desire to 'play' with the micro that influenced me, but a number of important factors.

Of these the most important were shortage of time and lack of staff. Eventually we were able to fund a person for six months, but the delay in making these arrangements reduced still further the time available to us. What we were actually expected to do was to index approximately 1800 documents in 8 different locations (3 not local) by subject, corporate author, personal author, title and document number—no mean task!

I had already obtained a number of software packages or programs that might be applicable, and one in

particular seemed ideal. This was D.B. MASTER, and was, I felt, most appropriate, as it not only would store and sort the records we should be creating, but it also incorporated a Report Generator that would lay the records out on the page exactly as needed. This feature enabled us to produce an informative index rather than merely an indicative one, and certainly made it much easier for the people who had to use it, who were, in the main, not librarians or indexers.

D.B. MASTER is quite simple to use, both in the setting up of the file and entering the data. To set up the file one has to create a number of fields into which relevant parts of the document details can be entered, so that one will have a field dedicated to the title, one for the subject, one for the author, and so on. Once the fields have been set up and one has started to enter documents it is difficult to change them (although certainly not impossible), and so care must be taken in planning the file. One has to decide on the length required for each part of the entry. There are several factors to be considered here since, once decided, the length of the field can be altered only with difficulty. Related to this is the fact that the number of entries on an individual floppy disk, the storage medium, is dependent on the overall size of the record, so that restraint needs to be exercised and thought given to the maximum number of characters one will in practice need for each field. This all may seem to make it a rather difficult thing to use, but in reality it is quite simple, provided that one does not rush headlong into it.

The documents themselves presented some problems. One was that many of the official documents published by the County and District Councils and the British Airports Authority had, and were usually referred to by, an official document number, so naturally this had to be entered. Many others, however, did not have these numbers, but they might have one assigned to them later on, and so still needed to be listed by document number. We took the most simple option and assigned them simulated numbers, using the style of the official numbers, and thereby still linking together like documents. Further problems were the spread of the documents over a number of locations and the need to allow for locational changes, but both were solved by the editing feature of the D.B. MASTER, which allowed changes to be made quickly and easily to parts of indi-

vidual entries. One serious limitation quickly made itself apparent: for each entry we could assign only one subject heading, and so if it was felt necessary to have more than one for a particular document we had to create another entry, identical except for the subject heading field. Eventually we decided not to try to create detailed subject headings but to use standard, fairly general ones, and then we found that for the vast majority only one entry was needed. I am sure that the specialist and legal staffs using the indexes preferred these general headings.

Finally we had the preparation of the print-out using the Report Generator of D.B. MASTER. Here, particularly for our Stansted Indexes, is the most attractive feature of the package; the ability, from the entry of

records once only, to produce print-outs in varying orders and layouts, as not only were we able to sort in alphabetical or numerical order, but we could also arrange that the field we were using as a base for sorting was brought to a position on the printed page which gave it obvious prominence. This can best be illustrated by the example below, which shows a single entry from the file (i.e. how it looks on the screen of the visual display unit), and then how it was laid out in the final printed indexes. Note that the numbers given against each field are not normally present, but have been inserted to make clear the movement of the fields from entry to entry.

The other two volumes, corporate author and document number, were similarly laid out.

SCREEN LAYOUT

- 1 TITLE: Letter to BAA—Details of Buildings
- 2 PERSONAL AUTHOR: Knight, c j
- 3 CORPORATE AUTHOR: Uttlesford District C.
- 4 SUBJECT HEADING: Buildings of Spec. Arch. Int.
- 5 DOCUMENT TYPE: A4 letter
- 6 DATE OF PUBLICATION: 6 July 1981
- 7 LOCATION: Counsel
- 8 DOCUMENT NUMBER: PC12

TITLE VOLUME PRINT LAYOUT

Letter to BAA—Details of Buildings(1)	
Uttlesfield District C(3)	A4 letter(5)
Knight, c j(2)	Counsel(7)
Buildings of Spec. Arch. Int.(4)	PC12(8)
6 July 1981(6)	

SUBJECT VOLUME PRINT LAYOUT

Buildings of Spec. Arch. Int.(4)	
Letter to BAA—Details of Buildings(1)	PC12(8)
Uttlesfield District C(3)	Knight, c j(2)
A4 letter(5)	Counsel(7)
6 July 1981(6)	

PERSONAL AUTHOR VOLUME PRINT LAYOUT

Knight, c j(2)	Letter to BAA—Details of Buildings(1)	
	Buildings of Spec. Arch. Int.(4)	6 July 1981(6)
	Uttlesford District C(3)	Counsel(7)
	A4 letter(5)	PC12(8)

The D.B. MASTER software we used allows only word-by-word alphabetization which, to a computer, is more 'logical', and easier for the programmer who originally writes the software. Spaces, punctuation and hyphens are sorted in that order by the computer and have priority over letter characters. This requires, of course, that every entry be just right. Spaces and punctuation not entered correctly, or even not to the set standard, will confuse the order—something that was brought home to us when we found that some titles had

been entered enclosed by inverted commas and some without. This was quite easily corrected, but took time we could ill afford.

With the size of record we were creating we were able to get about 750 records on each floppy disk. When all the records needed will fit on one disk the sorting (i.e. alphabetization) is quite straight forward; one sets it going and then leaves it for half an hour until it has finished. When, however, the records take up three disks, as ours eventually did for the Stansted indexes,

things are not so easy, and for two or three hours someone must sit by the computer swapping over disks as required.

There are disadvantages in using microcomputers for indexing, but with D.B. MASTER these stem from the fact that it has to compromise in certain areas to be able to offer special features such as the Report Generator, which was essential for our particular need. There are many other software packages that are of use to

indexers, particularly if all that is required is a typescript to be passed to a printer.

Although I had overall charge of the project the day-to-day management was in the hands of Linda Anton and a fair slice of the actual data entry was done by Paul Curtis.

*

I should be happy to give more detailed descriptions to anyone who cared to enquire.

Bias in indexing

Recently I was asked to revise the original index for the second edition of a sociological book—which I will refrain from naming here. The text was liberal in outlook, critical of government policy; but the index seemed less fair-minded, laying undue emphasis on all references derogatory to the establishment, to the extent that indexing was unbalanced. Whereas most names were listed without qualification, such as,

Harris, Lord, 7, 17
Summerskill, Dr Shirley, 110, 111

and strings of undifferentiated page numbers were allowed to follow neutral topics, e.g.,

complaints, 73, 74-5, 78, 79-80, 91
sentencing policy, 24, 26, 38-40, 97-8

—some single entries, even if referring to less than half a page of text, were qualified with relish:

Golder, Sydney, court victory against Home Office, 137
Ince, George, maltreatment and forcible administration of drugs, 9-10, 11
Nightingale, John, and difficulties of study, 61-2
Walton (Liverpool) prison, National Front membership of officers, 125
welfare work, prison officers' failure to gain responsibility for, 121-3, 125.

Two long entries in particular seemed to show sub-headings selected with quite unprofessional glee:

Home Office: defence of use of drugs . . . judgment against in European Court of Human Rights . . . 'lack of leadership' . . . 'out of touch' . . . and prison department . . . refusal to transfer Irish prisoners to Northern Ireland . . . secrecy and misinformation . . .
Strangeways prison: military training of officers, 121; National Front membership of officers, 124;

overcrowding, 16; study facilities lacking, 62

I condensed, rephrased or regrouped these entries to produce:

Home Office, in prison system . . .; and Board of Visitors . . . dismisses visitor, 78; and drugs in prisons . . . and Human Rights Court . . . and information . . . and Irish prisoners . . . and medical service . . . and Open University . . .

Strangeways prison: officers, 133, 136-7; overcrowding, 17; study facilities, 66-7

—Less dramatic reading, but, I trust, neater and more objective indexing.

H.K.B.

Index simplicissimus . . .

Perhaps we should all spare ourselves our agonizing over the exact phrasing of headings, division into sub-headings, and other mysteries of our craft. *The Irish Kama Sutra* by Peter O'Regan and Sean Dunbar (Futura Publications, 1981) does away with all such tedious niceties, thus:

Index
FIND THE PAGE YOU WANT

For quick and easy access to the page you want we have included this comprehensive index.

PAGE NUMBER:	NUMBER OF PAGE:
Page One	1
Page Two	2
Page Three	3
Page Four	4
Page Five	5
Page Six	6
[. . . and so on, up to]	
Page One Hundred and Twelve	112