The Department of Health and Social Security (DHSS) computerized library system has been operational for more than a year. (For a brief description of the system please see below). One function of the system is the production of current awareness bulletins. At the moment two of these bulletins, Hospital abstracts and Social Service abstracts (SSA), have keyword indexes. The production of the SSA index is described below.

SSA Index

SSA is a monthly bulletin containing about 150 brief summaries of periodical articles, reports, books and pamphlets, etc., on the personal social services, with particular emphasis on British literature. About 250 periodicals are scanned for information.

The first stage of indexing is a subject analysis of the document to determine the relevant keywords to be assigned to the computerized record. Using a visual display unit (VDU), linked to the library's central Prime minicomputer, the indexer may search an online thesaurus of some 23,000 terms for the preferred terms. Once the terms have been selected the indexer uses a data entry format on the VDU, (this is really an electronic worksheet), to add the document to our main data base—DHSS-DATA. Within this worksheet there is a section for the entry of keywords.

Once data entry, i.e. of bibliographical details, keywords, abstract, etc., is complete, the record is entered directly into the data base. However, before it is accepted there is an automatic check against the thesaurus to find whether or not terms entered in the keyword section are in the thesaurus or are non-preferred terms (synonyms). For example, if the indexer entered NHS in the keyword section the following response would appear on the VDU screen:

1 NHS
2 USE National Health Service
3 to Re-enter or Omit the Term

Enter the Number of the Option Required

2

In this case National Health Service would replace NHS in the record. Option three would delete the entry from the record.

If Guys Hospital had been entered the indexer would see:

Guys Hospital
Has not been found
Do you Wish to Accept it?

Y

In this case the term has been accepted, since the thesaurus includes only a few proper names. If it were a new term it would not be added to the thesaurus until it had been checked by the thesaurus editors; however, the term would not be removed from the record—as it would if a negative answer had been given.

Indexers may enter as many keywords as they wish. The first four keywords are automatically included in any printed index, unless the term is preceded by an apostrophe, e.g., 'Social policy. More than four terms can be included in a keyword string by the use of the @ character before the term, e.g., @Care.

At the end of the month a search is carried out on the data base to create a file of all new SSA entries. This usually produces about 150 records. Using only one command the search is carried out, the records are downloaded to a file, and the file is processed and then produces two files:

1. The main text of the bulletin is formatted as follows: A4, double-column, title bold, periodical title in italics, pages numbered, each record sequentially numbered for index location, monographs in AACR2 format, subdivided into broad subject areas and the subheadings inserted.

2. The keyword index (it is possible also to produce an author index at the same time.) Initially a rotated keyword index is produced as a single column and appears as follows:

- Children in care
- 1332
- Children in care
- 1347
- Children in care
- Access to children
- 1334
- Children in care
- Access to children
- Parents
- Rights
- 1336
- Children in care
- Care
- Research
- 1337

Abridged version of the address given to the Society of Indexers, London, 1 November 1984.
Using a word-processing package the indexer removes duplicate entries and makes any other alterations that may be required to produce the following file, using the example above:

Children in care 1332, 1347
Access to children 1334
Parents Rights 1336
Care Research 1337
Family relations 1333
Grief work 1294

This process mainly entails the use of the delete-line facility. However, we believe that we can remove this stage also by fine tuning of the existing program.

Once the indexer is satisfied with the contents of the index, it is processed by a final program that formats it into double-column, A4, inserts the page heading 'Subject index' and inserts print codes so that the lead term is printed in bold—in the example above this would be 'Children in care'. There will then be a small amount of word processing to include any carry-overs if a string has been split by the double-column or by a new page. Also, page numbers have to be added.

The bulletin is then ready to print on the library's double daisy-wheel Qume printer, and the master copy is sent to HMSO (Her Majesty's Stationery Office) for reproduction. The whole bulletin production process usually takes one working day.

The format of an index need not necessarily be the same as the SSA index; for example, Hospital abstracts index appears as:

Accounting, Cash flow 955
Accuracy, medical records, obstetrics, assessment 950
Administration, international cooperation, in West Germany and Saudi Arabia 938
Anaesthesiology, quality of patient care, medical audit 1005

Although the rotated keyword index is obviously a relatively crude form of indexing, the basic index produced by the computer, as shown above, has proved to be a useful aid to the indexer and has cut down the time and effort required for index production. Indexers always have the opportunity to amend the index as they think fit. For instance, in the example above the 'in' and 'and' have been added to 'in West Germany and Saudia Arabia'—replacing commas.

The Library will be using the same program for the production of other bulletins, e.g., Nursing research abstracts, and it can also be used when producing bibliographies. In fact, any listing from the data base can have author and subject indexes produced if required.

The DHSS Library system

Hardware: Prime 750 minicomputer, two 300MB disc drives, tape deck, 30 VDUs, 14 small printers, Printronix high speed line printer, Qume double daisy-wheel quality printer, 10 modems (for connection to 5 branch libraries), and all the library rooms at Alexander Fleming House, DHSS HQ, are wired directly to the minicomputer.

Software: The main package is STATUS—a free text information storage and retrieval package developed by Atomic Energy Research Establishment at Harwell. Other software, including the indexing package, has been written by BNF Metals Technology Centre, Wantage. BNF are the main contractors for the DHSS system and they are the franchise holders for the use of STATUS on Prime computers. STATUS allows the creation of any number of databases, and at present the DHSS Library has seven databases dealing with all aspects of library administration. The main database, DHSS:DATA, which acts as the library's online catalogue of monographs, periodical articles, circulars, etc., covers the whole field of health and social services. This database is publicly available on two database host systems: DATA-STAR and SCICON.

Paul Brewin is Systems Librarian, DHSS Library.

Eurolux triumphant

Godfrey Smith writes in The Sunday Times (7 Oct. 1984) of Eurolux, 'a computer system in which 300 million words of law reports are stored. Lord Denning, asked to advise, said it could never have the subtlety of human search. Eurolux riposted by choosing one of Denning's favourite phrases ("lo and behold") and asking the computer for any case where he'd said it. In less time than it takes to write this, it had spun out eight.

'A sceptical barrister said they'd never find a rape case he'd been involved in where the victim had alleged the rapist appeared at her window wearing only his socks. They fed in the three words, rape, window, and socks and, lo and behold, up came the whole case.'

© Times Newspapers Ltd.