

A Note on Information Retrieval experiment designed at Bangalore to establish an S.D.I. system with 5 subject profiles for using Belcom-1 system mini computer or a DEC 1090 system with DBS availability

T. K. S. IYENGAR⁺ AND K. K. NAMBIAR^{**}

A National Information Centre for Neuro Sciences will be set up at the National Institute of Mental Health which will be responsible for coordinating, organising and providing information at national level for large sectors of national endeavours. Its responsibilities include :

- (1) Developing and maintaining document collection, serving as the national base for information work.
- (2) Providing co-operative acquisition of foreign literature, so as to ensure that at least one accessible copy of everything worthwhile published in the world is made available within the country itself, for use by the scientists at all times to come
- (3) Compiling national union catalogues of collections in the subject.
- (4) Maintaining Data Banks, containing data relevant to the activities and the development of the sector.
- (5) Providing SDI Services in Neuro Sciences based on its own holdings and data and utilizing information services available, on magnetic tapes.
- (6) Bringing out Current Awareness Lists.
- (7) Providing reprographic facilities for copying documents and distributing information.
- (8) Providing translation service.
- (9) Providing documents on request from whichever source available.

* Librarian, I.I.Sc

** Chief, R & D Division, B.E.L. (Govt. of India)

- (10) Organising training facilities in information science, documentation, computer science, etc
- (11) Developing additional resources in the Neuro Sciences Area

Planning of selective dissemination of information

Neuro Sciences has become essential to specialist studies in various branches of applied technology, particularly in the context of the scientific and industrial developments that are taking place in the world. So much of literature on this comprehensive topic is published in the form of books, articles in periodicals, technical reports, etc., that any scientist or technologist cannot hope to keep himself abreast of the latest developments unless some service to disseminate information is offered to him on a selective basis. While this is true in the entire field of science and technology, a beginning was made offering SDI Service in Neuro Sciences.

Steps in planning

The planning of an SDI Service involves the following steps

- (a) Selection of important projects to be covered by the system
- (b) Preparation of User Interest/Project Profile
- (c) Preparation of Document Profile
- (d) Comparison by computer of the two files, of User-Profile and Subject-Profile for a close matching of terms and sending out a print-out to the user
- (e) Periodical evaluation of the system. An important element in the SDI is the provision for 'Feed Back'. The user is asked to assess the relevance to his requirements of each item of information he receives in the SDI System. His relevance assessment can then be used either to adjust his profile to meet any change in his interests or to alter the level of match when he wishes to receive more or less information.

The user group

In addition to all research workers in neuro sciences of the Institute, it is hoped to obtain the co-operation of local neuro science research workers to act as members of the user group for the subject. These will be drawn from Universities and colleges of advanced study and/or technology, industry, Government Research Establishments, etc.

Preparation of user-profile

If the SDI System is to be successful, it is essential that the profile of each user should exactly meet his requirements. The User-Profiles are to be constructed on the basis of

- (a) Description given in the project folder; and
- (b) Consultation with the project leaders or specialists concerned

By this process, it is possible to obtain from them, the Key-words or descriptors which adequately describe each project. If any ambiguity is found, then the word should follow a brief scope to make the concept more explicit in meaning.

On the basis of such a reconstructed vocabulary, a first approximation to his profile (User-Profile) will be compiled in the Index Language. This will be referred to him and after discussion with a member of the project staff, the profile will be used in the experimental phase of the project (See Fig. 1).

PATTERN OF SUBJECT SCATTER FOR SDI

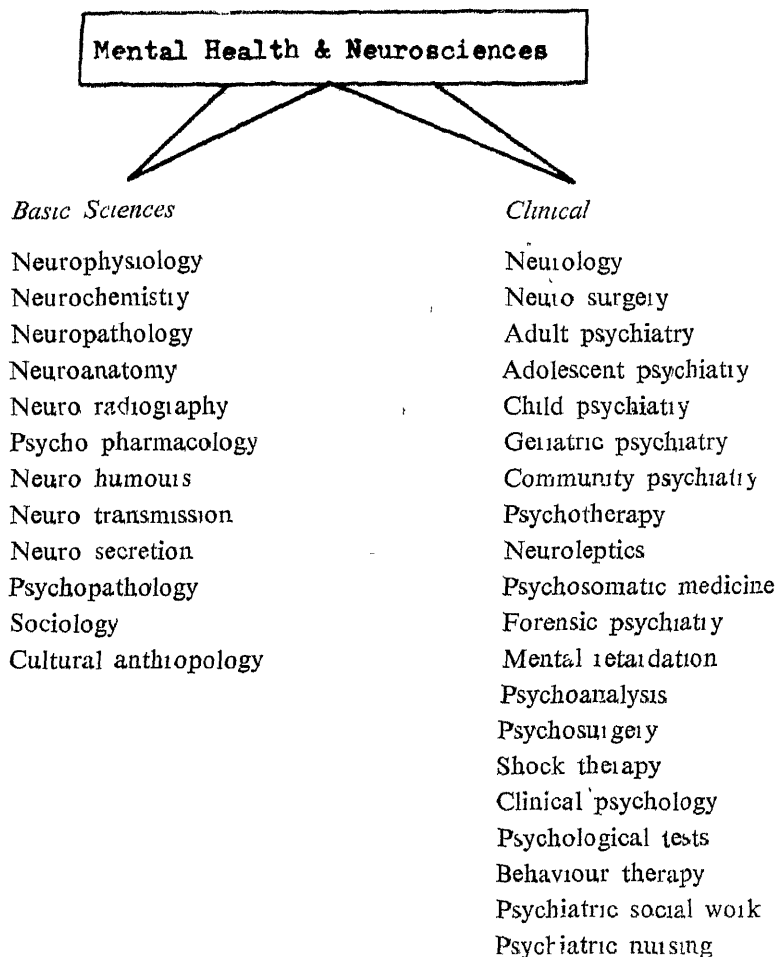


FIG 1

Each user profile consists of a set of cards, the number of cards being equal to the number of key-words used to describe the interest of the user. The items of information to be included in each card of the set is as follows :

1. Key-word, 2. Scope-note, 3 Project Number and 4. Name of the Project Leader

The cards so prepared for different User Profiles are merged together and a 'Master Profile' in one alphabetical sequence is made.

Matching of profiles

The Profiles are matched at regular intervals by the computer. The interval may be a week or a fortnight. In the comparison process, sufficient weightage is to be given for the brief scope note of the key-word in the user profile. When a match occurs, the concerned user (whose name can be traced on the User Profile Card) is informed by a print-out notification.

It may be necessary also to adopt different methods as well as different levels of matching for different users, since they vary greatly in their range of interests and their attitude to information. In the latter case, must be included the users' tolerance of 'noise', *i.e.*, the irrelevant material which they will receive in the SDI Service.

Project details and cost evaluation on this proposal

(i) Object

1. To do literature search, index, codify, store and retrieve information on the available literature in the form of computer paper print-out on all topics in the area of Neuro Sciences,
2. To develop an SDI Service with the availability of SDI Tapes procured through the N.L.M., Bethesda, Maryland, USA.

(ii) Methods of approach

1. (a) To collect the literature on each topic-books, articles in Journals, Reports, etc ;
(b) To classify topic-wise;
(c) To codify and store; and
(d) To retrieve and print-out on a computer paper printer.
2. Prepare User's Profile and Subject Profile in different disciplines in the field of Neuro Sciences.

(iii) Duration

Initially for a period of ONE YEAR

(iv) For creating data base system SDI is reflected in Fig. 2.

INFORMATION RETRIEVAL PROJECT No 1 CREATING A DATA BASE

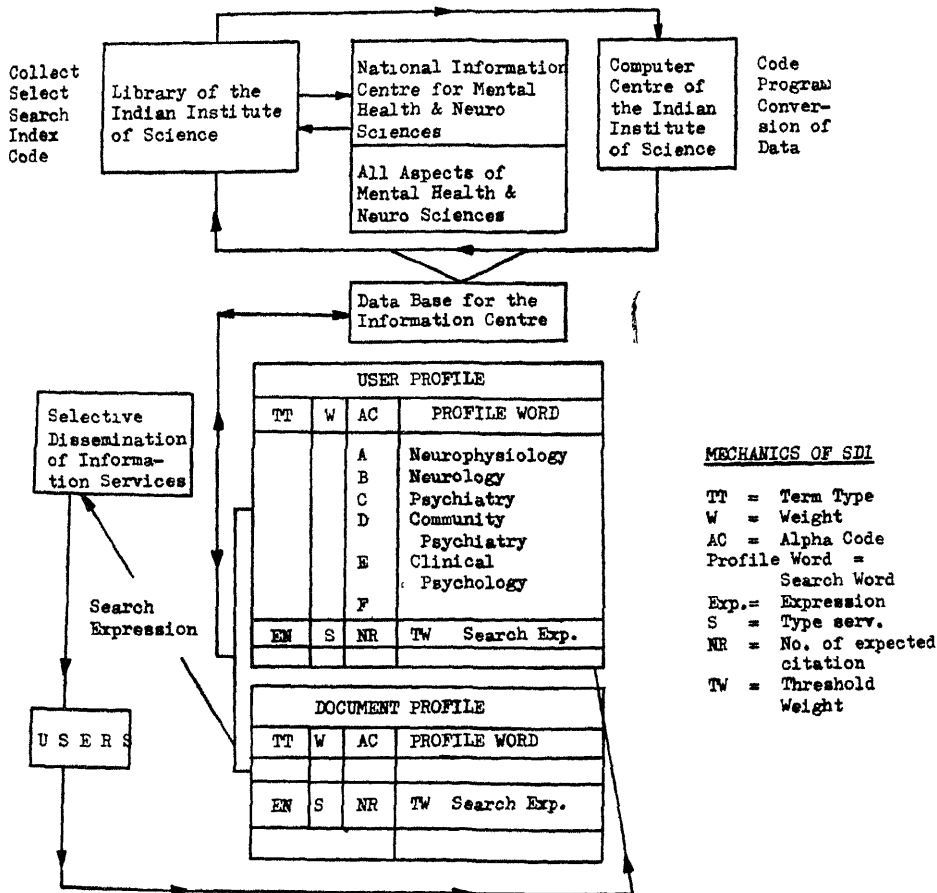


FIG. 2

(v) For creating DBS either the BELCOM-1 or for a large data base DEC 1090 which IISc has acquired could be used (Figs. 3 and 4)

I I Sc. DEC SYSTEM—1090 CONFIGURATION

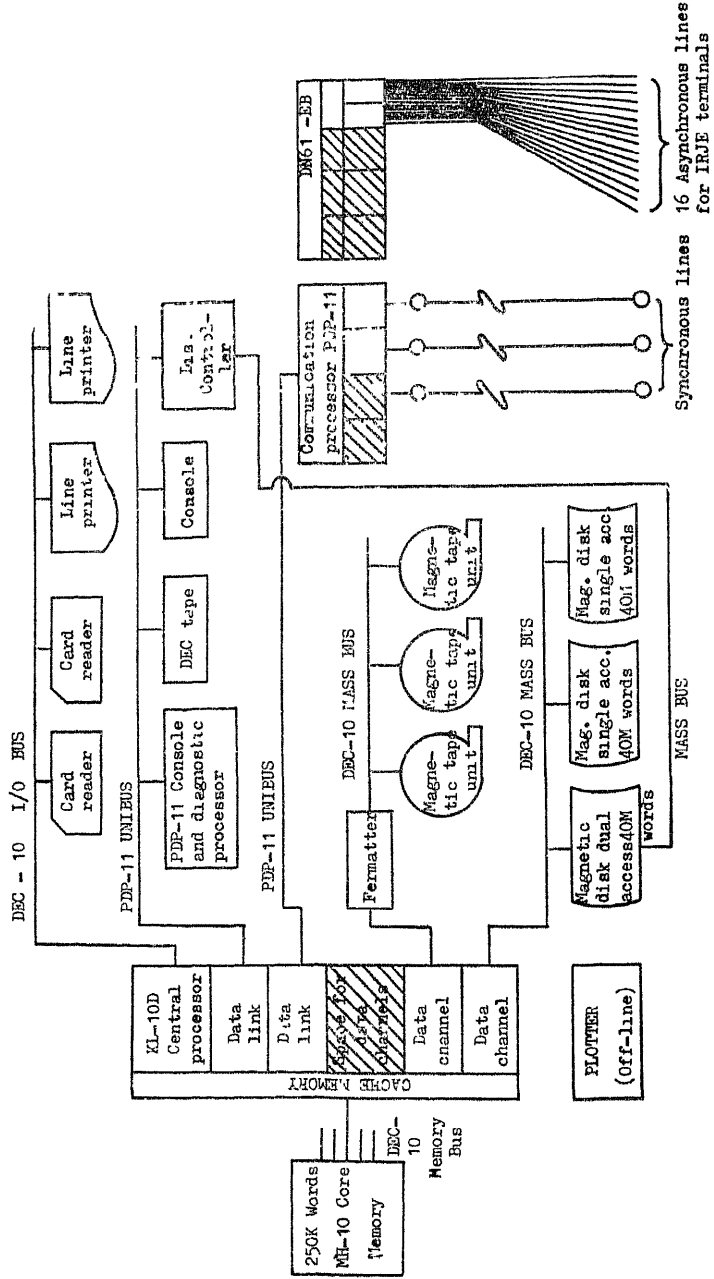


FIG 3

FIG 4

Features	BELCOM C810-S/C810-C
<i>Data Formats</i>	
Word length, bits	. 16
Fixed point operand length, bits	. 16
Instruction length, bits	16/32
<i>Main Storage</i>	
Storage type	.. Semiconductor/Core
Cycle time, microsecs/word	.. 0 35/0-65
Min. capacity, words	.. 8,192
Max capacity, words	.. 32,768
Max capacity, words in memory banking	. 2,62,144
Parity check	.. Optional
Storage protection	.. Optional
<i>Central Processor</i>	
No of accumulators	.. 2
No of Index Register	.. 1
Indirect addressing (only for word mode)	.. Multi-level
Add time (Register to Register), microsecs (full word)	.. 1 35/1 65
Add time (Memory to Accumulator) microsecs (full word)	.. 2-05/2 95
Hardware Multiply/Divide	.. Standard
Hardware byte manipulation	.. Standard
Immediate (literal) instructions	. Standard
Power failure protection	.. Standard
Real time clock	.. Standard
No. of instructions	.. 169
Hardware stack (16 registers)	.. Standard

Features	BELCOM C810-S/C810-C
<i>Input/Output Control</i>	
I/O word size, bits	.. 16
Direct Memory Address	.. Standard
Max. I/O rate, words/sec.	. 1,818,100/1,176,400
No. of external interrupt levels	. 12
<i>Peripheral Equipment</i>	
High Speed Paper Tape Reader	. 400 cps.
High Speed Paper Tape Punch	.. 110 cps.
Keyboard Send Receive Typewriter	. 30 cps.
Card Reader	. 400 cpm
Cassette Tape	. 4,800 bits/sec
Line Printer	.. 200 lpm
Visual Display	110 to 1,200 bauds
Magnetic Tape	.. 36,000 cps
Floppy Disc	.. Yes
Mini Disc	.. Yes
<i>Software</i>	
Assembler	.. 2-pass
Utilities and Loader	.. Yes
<i>Availability</i>	
Delivery period at present	.. 12 months after receipt of order

For a non-data bank retrieval system

MEDLAR Tapes are available which can be channelized through N.L.M., Bethesda, Maryland, USA (Fig 5)

INFORMATION RETRIEVAL PROJECT No. 2: PROGRAMMING ON AVAILABLE SDI TAPE

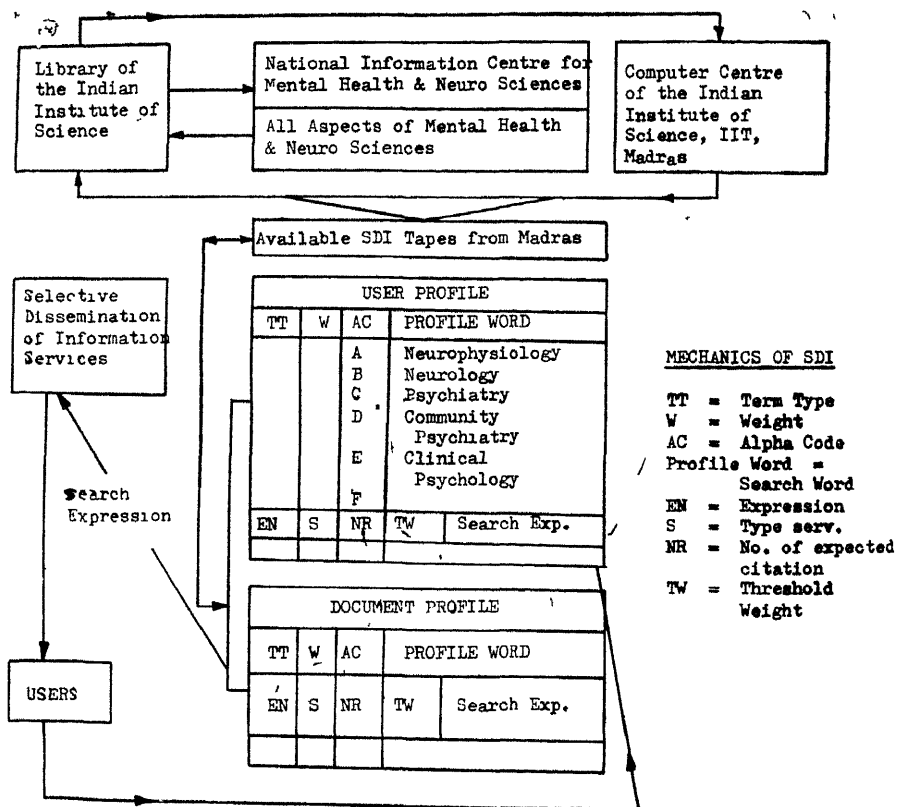


FIG. 5

Cost

The input to the computer-based system includes journal articles reports, conference proceedings, etc. The output includes individual notification, an Information Bulletin which may also be an Accession List, a quick reference key-word index for use of the library staff and may later, include catalogue entries, bibliographies and retrospective retrieval service. It is therefore neither practical nor helpful to compare the whole computer-based system with all the separate manual system because in many cases, there is no exact correspondence between the parameters being measured. However, the following analysis is made (Fig. 6).

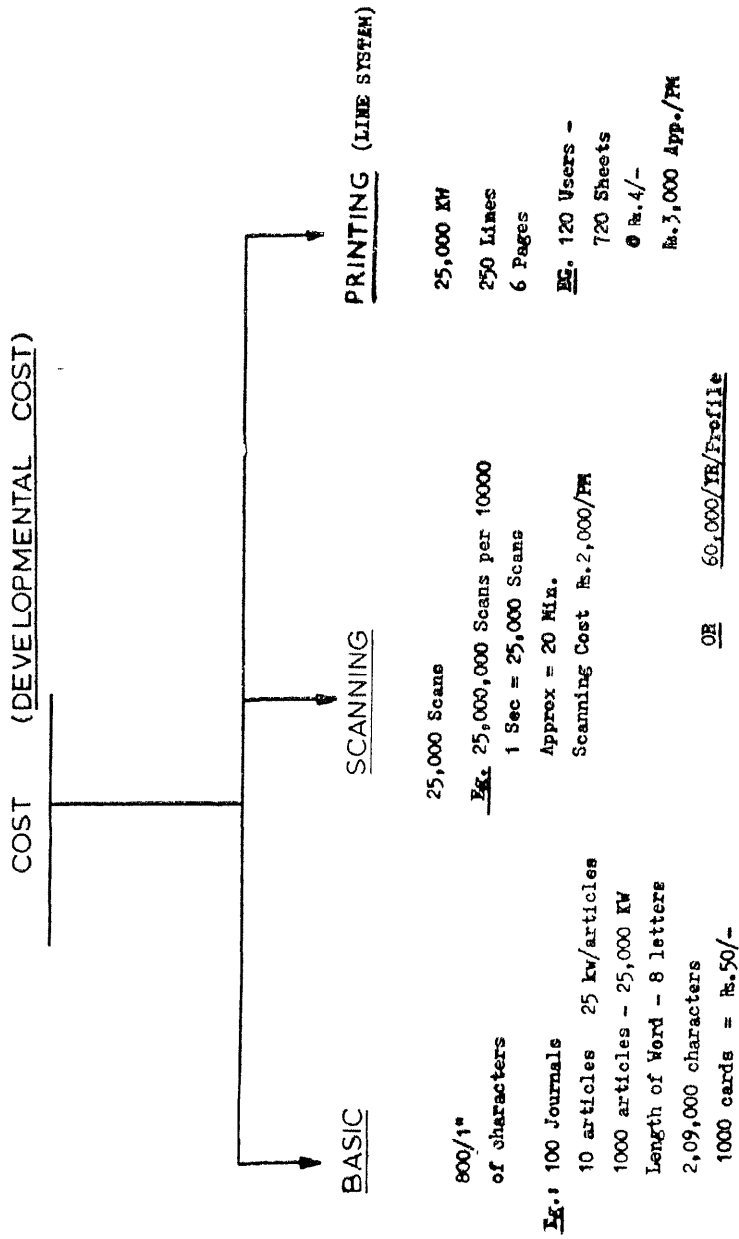


FIG. 6

ANALYSIS OF THE SDI AND THE MODE OF EVALUATION is as follows :

	RELEVANT	NOT RELEVANT
RETRIEVED	A	B
NOT RETRIEVED	C	D

$$\text{RECALL} = \frac{A}{A + C}$$

$$\text{RELEVANCE} = \frac{A}{A + B}$$

Fig. 7