

Computer-Based SDI: Cobol Program.

(Non-conventional methods in document retrieval 23).

IK Ravichandra Rao* and Hemalatha Iyer, *Documentation Research and Training Centre, Indian Statistical Institute, Bangalore 560003.*

[Experimental project on a computer-based SDI service, using COBOL software].

1 Scope of the Paper

This paper presents a report on an experimental and pilot project on a computer-based SDI service using COBOL software developed at the Documentation Research and Training Centre, ISI, Bangalore. The studies were carried out in the Indian Institute of Science, Biochemistry Department (IISC) and the Bharat Heavy Electricals complex (BHEL). In BHEL, the SDI service is no w an on-going project.

2 User Interest Profile

A specimen of the worksheet used for preparing user interest profiles (UIP) is given in Appendix A. Each of the six scientists in the IISC included in the project were interviewed by one of us (HI) in filling the UIP worksheet. At the outset, the scientist was apprised of the project, the objective of SDI service, its advantages etc. The dialogue that followed helped to specify the interest of the scientist, the types of documents he desired to have etc.

In the case of the BHEL units (about 10 of the units located in different parts of India participated in the project) the information scientist in each unit met with the users for filling in the UIP proforma. In the experimental stage, the filled-in profile worksheets were received by DRTC for further processing.

3 Document Description (Document Profile)

For the IISC pilot study, current issues of some 15 periodicals covering biochemistry and other related areas were scanned and about 60 papers selected. The BHEL units had earlier agreed on a list of about 150 periodicals and each selected 10-15 titles for scanning and selection of papers. Other documents, such as technical report, thesis etc were also selected for inclusion in the data base. To facilitate the adoption of a uniform cataloguing procedure for the selected documents, document description worksheets were designed and used by the units. A specimen of this worksheet is given in Appendix C. To begin with a separate worksheet was used for each type of document. However, it was later pointed out that all the data

elements for the different types of documents could be enumerated in one worksheet as is being done in the INIS or AGRIS worksheet. It was also agreed that the choice and rendering of the different data elements in an entry should conform to accepted standards such as the *Classified Catalogue Code* (= CCC) and the *Indian Standard for Bibliographical References* (3). The following additional guidelines were given:

1 Mention class number, if used.

2 Subject vocabulary control (selection of appropriate descriptions etc) may be obtained using thesauri such as TEST, MeSH etc., as applicable or an appropriate faceted scheme for classification.

3 Group the descriptors to represent each of the relevant subjects discussed in the document.

In the experimental stage, the filled in document worksheets from the IISC and BHEL were received by DRTC for further processing and creation of the data base.

After the initial experiment and a workshop to explain the whole operation and future possibilities, the SDI operation is now being done at the BHEL, R&D Unit, Secunderabad, which receives the inputs from the different BHEL units, in collaboration with the Ramachandrapuram unit in Hyderabad which does the computer processing of the inputs to obtain the SDI output listings.

4 Vocabulary Control and Search Strategy

40 NEED FOR CONTROLLED VOCABULARY

The Feature Heading, that is a systematized string of descriptors, is obtained from a system for vocabulary control such as a thesaurus or a scheme for classification. The Feature Heading allows the user of the computer-based system — by specifying a precise set of subject terms — to obtain only those records which are relevant to his/her query. Needless to say, the terms selected by the indexers must adequately represent the subject treated in the document. In addition, the person who receives the query from a user must know how to translate the users query into a meaningful formulation of the query acceptable to the system using the system for vocabulary control.

* Present address: School of Library and Information Science, University of Western Ontario, London, Ontario, Canada.

41 ANALYSIS OF UIP

Some reference sources in the area of biochemistry were read for the purpose of understanding the scope of the subject.

The keywords from the statement in UIP and also from the articles cited in UIP were picked up. The keywords furnished by the users were also studied. TEST and MeSH were used for the purpose of vocabulary control. In the case of synonyms, one of the terms was chosen and consistently used. For example, Vitamin E was used instead of Tocopherol. The descriptors were then listed and assigned an Alpha Code and Term Type. The design of the profile coding sheet and a specimen copy of the analysis of subject with search expression is illustrated in the Appendix 92.

42 FORMULATION OF SEARCH EXPRESSION

From the narrative statement of the user's requirement and dialogue with the user whenever possible, the exact requirement of each user was identified. Search expressions were then formulated for each profile, using Boolean operators.

For example, a user may be interested in the *Chemical methods, especially the spectrophotometric method, of estimation of Vitamin E, but not the biological method.* Information analyst now picks up the necessary descriptors from the statement of the subject, and cited articles. Then using the thesaurus, gets the standard terms and lists out as follows:

- A Vitamin E
- B Chemical
- C Method
- D Spectrophotometric
- E Estimation
- F Biological

The required search expression is:

$$(E + ((B + C) * (D + C))) - (F + C)$$

Details regarding the card design for search expression are discussed in the earlier reports (2, 7). An example is given in Sec 92.

42.1 Steps in Formulating Search Expression

The following steps are involved in formulating search expression:

- 1 Collect user profile. An efficient way of collecting user profile is through personal interaction between information analyst and user.
- 2 Analysis of user profile — pickout relevant descriptors
- 3 Assign term type and alpha code (2,7)
- 4 Formulate search expression using the Boolean combination of logical connectors—AND, OR and NOT.

43 SEARCH METHODOLOGY

The search expression is Sec 42 is similar to any arithmetic expression of High Level Programming Language, such as FORTRAN, COBOL, etc. Therefore Boolean operators instead of arithmetic

operators can be used and Polish notation is helpful to processing the search expression.

Polish notation is used to represent arithmetic or logical expressions in a manner which specifies in simple and exactly the order in which operators are to be evaluated. In Polish strings, no parenthesis is needed. In this notation, the operators come immediately after the operands. Therefore, sometimes it is called suffix or post-fix notation. For example, $A * B + C$ as $AB * C +$ and $A * (B + C / D)$ as $ABCD / + *$ etc.

Search expressions in Polish notation can be evaluated by a single left-to-right scan with the aid of a stack. The stack will hold all the operands which have been scanned or produced as the result of some operation, but not used. The system starts with the left most symbol, process it and go to the symbol on its right, process it, and so on.

All the operands and operators in a Search Expression need not necessarily be processed. It depends upon the value of the operand and operator. For example, let us consider the Search Expression $A + (B * C)$. The required Polish notation is $ABC * +$. In this expression, suppose the value of B is 1, it is not necessary to find out the value of operand C. Also, if the value of $B * C$ is Zero, it is not necessary to search A, as we require $A \text{ and } (B * C)$. Therefore, a well defined system for skipping off sub-trees in searching a tree structure information results in saving computer time and money. It may be noted that such provision is built into the systems developed at DRTC. Polish notations are processed using the technique of Recursive Programming.

While searching it is preferable to take into consideration of Facet Structure of the Subject Heading String.

43.1 Facet Structure

The structure of a subject formulated on the basis of facet analysis can be used conveniently for achieving economy in search time. Based on Ranganathan's General Theory of Library Classification it is possible to identify different facet structures. Therefore, search can be restricted to one of these 64 different facet structures. This is derived on the basis of occurrence in combinations to six different facets.

Assuming that the data base adopts a faceted scheme for arrangement of files, the 64 different facet structures may be conveniently used for retrieving, provided the user or the computer software system is able to specify the specific facet structure. However, facet structure need not necessarily be incorporated.

44 DATA BASE AND OUTPUT

The printed form of data base and output have been published in earlier reports (2, 7).

5 Program Description**5A PROGRAM 1**

Name of the Program: Computer-Readable Catalogue
Acronym: CORC
Author: I K Ravichandra Rao
Date of Current Version: 24 June 1975

5A1 Purpose

The purpose of this program is to create a Machine Readable Bibliographical Data base. The program can also be used for retrieving the references by specifying a precise data element, such as Call Number, Subject term, Author's name, Title, Edition and Volume statement, Place of publication, Publisher, Date of Publication, and Language of document.

5A2 Input

1 The input entries for CORC are prepared by punching each of the entries on a standard 80-column card. An entry can have a maximum of six punch cards. However, with slight modification in File Description of Input-Output Section in Source Program, an entry can use a maximum of 24 punch cards. The number of characters used in an entry varies with each document and therefore, it is preferable to punch continuously the characters in each entry using continuation card, wherever necessary. The Sections in each entry were separated from each other into fields by a Hash mark (#) or virgule (/). Hash mark is used when Section is printed in separate lines. For example, Call Number and Feature Heading are to be printed in the independent lines and therefore hash mark should be used as field separator between them. But title statement is to be printed as continuation to author statement in such a case, a virgule may be used as a field separator between the two fields.

2 The characters in an entry are punched continuously in columns 1 to 76. Columns 77 and 78 are used for punching alphanumeric Document Number. Columns 79 and 80 are used for punching a Serial Number in the continuation cards for an entry. The first character in a Leading Section must be either space or virgule or any other symbol developed to denote bibliographical level. A specimen copy of the input is given in Sec 94.

3 The Last Card of a set of cards is identified by placing immediately after it a card in which an asterisk is punched in each of the first four columns.

5A3 Output

1 The output file from CORC is a magnetic tape file called TESTDATATAPE. The access mode and file organisation are sequential. It contains exactly four bibliographical records per logical record or block. Each logical record is 2048 characters in

length. The output may be obtained on on-line printer, whenever necessary.

2 The tape file called TESTDATATAPE may be organised into ascending or descending sequence according to the ordinal value of Call Numbers, that is, in the classified sequence by using manufacturers software for sorting the tape file.

3 Using the program CORC Sequence Numbers may be assigned to the records in TESTDATATAPE. This requires an additional scratch tape. The output file from CORC is again a tape file called CORC1975 FILE.

Data format of the tape file is given below.

Table 1. Data Format of the Output File

SN	No of character	Character position	Description	Remarks
(a)	(b)	(c)	(d)	(e)
1	4	0-3	Record length	The characters are in binary form
2	4	4-7	Sequence number	—do—
3	22	8-29	Ordinal value of Call Number	The characters are in numeric form
4	1	30	Facet structure	The character is in binary form
5	24	31-54	Tags	—do.—
6	456	55-510	Main entry	The characters are in alphanumeric form
7	1	511	Last character	The character is in alphanumeric form; its value is "59"

5A4 Procedure

The procedure is simple and uses many times the principle of Push-Down Stack. To begin with, control cards are read. There are four control cards including the Last Card. The flowchart for creating TESTDATATAPE is given in Sec 95 (10).

5A5 Retrieval Part of CORC

1 The input to the program CORC: Retrieval Part, consists of the Machine Readable Bibliographic Database called 'CORC1975FILE' and the Query on Punch cards.

2 Each Query is punched on a standard 80-column card. Table 2 illustrates the design of the punch card.

TABLE 2. Card Design for Punching Query on Punch Card

SN	N of characters	Character position	Description	Remarks
(a)	(b)	(c)	(d)	(e)
1	1	1	Bibliographic level	Punch either A or B or C etc, depending on the bibliographical level.
2	1	3	Data field identifier (Tag)	A = Leading Section; B = Subject Term; 3 = Author C = Title; D = Edition; Vol statement E = Place of publication; F = Publisher; G = Date of publication H = Language Punch 0, if Facet structure is punched in 7th column
3	1	5	Facet structure	
4	1	7	Data elements	Maximum 40 characters

3 In the Retrieval Part of the program, output file is on-line printer. If there is no document for a given query, the message "NO DOCUMENTS ARE AVAILABLE" appears at on-line printer.

5A6 Flow Charts

The Block Diagram of CORC is illustrated in Sec 95.

5A7 Sample Input

1 The sample input for CORC to create a bibliographical data base, including the control cards, is illustrated in Sec 94.

2 The sample input of CORC, Part 2 — Retrieval program — is illustrated in Sec 941.

5B PROGRAM 2

Name of the Program: Selective Dissemination of Bibliographical Information

Acronym: SDBI

Author: I K Ravichandra Rao

Date of Current Version: 2 July 1975

5B1 Purpose

The purpose of this program is to retrieve documents from Machine Readable Bibliographical Database for a given Search Expression.

5B2 Input

1 The input to the program consists of Machine Readable Bibliographical Data base, called CORC 1975FILE and query on punch card. For punching Query on punch card, four card designs were developed as given below;

5B21 Card Design Type 1

This is designed to punch user's name on punch card. The user's name can be punched from Col 1 to 80. The user's name card should be preceded by a Control Card.

5B22 Card Design Type 2

There may be a maximum of 26 cards having the card design of type 2. The cards should be followed by a Last Card. The Card Design Type 2 is given below.

TABLE 3. Card Design Type 2

SN	N of characters	Character position	Description	Remarks
(a)	(b)	(c)	(d)	(e)
1	1	3	Term Type	See col d, Table 2, SN 2
2	1	5	Alpha Code	
3	40	7	Data Element	Maximum 40 characters

5B23 Card Design Type 3

There is only one card having the card design type 3 per search expression. This card provides information regarding bibliographic level, and Facet Structure, if any, see Table 3.

TABLE 4. Card Design Type 3

SN	N of characters	Character position	Description	Remarks
(a)	(b)	(c)	(d)	(e)
1	1	1	Bibliographic level	Punch A or B etc depending on bibliographic level.
2	1	3		Punch 0 if Col 3 is blank
3	1	5	Facet structure	

5B24 Card Design Type 4

There is only one card having the card design type 4 per Search Expression. This card contains search expression as discussed in Sec 42. The Last Character in the Search Expression should be the hash mark. Search Expression should be punched from Col 1 onwards. This card should be followed by a Last Card if there is no Search Expression to be processed; otherwise, cards having card design types 3 and 4 are to be repeated.

5B25 Control Card

As mentioned earlier, control card should precede the card having card design type 1. However, if there are more than one user's query for processing, it is sufficient to provide only one Control Card and it should be preceded by the first card having card design Type 1.

Control Cards provide some control information and also few literals such as "\$", "#", "£", etc. Control information is used to identify the last character of descriptors or Kernel Terms in Feature Heading Section. These characters may be comma, semicolon, colon, single inverted comma, double inverted comma, dot, hash mark, virgule, etc. Generally, hash mark is used as the last character in the Feature Heading. Card design of the Control Card is given below.

TABLE 5. Card Design of Control Card

SN	Number of characters	Character position	Description	Remarks
(a)	(b)	(c)	(d)	(e)
1	2	11-12	Zero's	
2	2	20-21	One's	
3	1	24	Zero	
4	3	29-31	Zero's	£
5	1	32	One's	
6	8	65-73	Literals	Punch " = \$ / - b * + - ..

Note.—To punch control information, select the field separators used in the Feature Heading Section, as discussed above. Find out the ordinal value of the Field Separator in ICL 1900 series. Add 1 to the ordinal value. Then punch zero in the corresponding column (punch 1 if the field separator is the last character in Feature Heading Section). For example, consider the field separator: comma. The ordinal value of comma is 30 in ICL 1900 series. So punch zero in the 31st col.

Thus, a Query of single user requires several cards having card designs as indicated above. The sequence of these cards is as follows:

- Control card;
- User's name card: Card design Type 1;
- Profile words on cards: Card design Type 2: (There may be a maximum of 26 profile words on 26 different punch cards);
- Last card (having 4 asterisks in col 1 to 4);
- Search Expression: Card design Type 4; (Cards as discussed in Steps 5 and 6 may be repeated if there is any more query of the same user);

- Last card (having 4 asterisks in col 1 to 4); and (Cards as discussed in steps 2 to 7 may be repeated if there is any other user's query);
- Last card (having 4 asterisks in col 1 to 4).

5B3 Output

The output file from SDBI is taken on-line printer. Search Expression is processed and if the value of the search expression is true, the corresponding Bibliographic Record is printed on on-line printer. If the value of the search expression is false for all the Bibliographical Records in CORC1975FILE then a message "NO DOCUMENT" appears on the printer.

5B4 Procedure

1 The program reads a Control Card and user's name card. Then it reads cards having Card Design Type 2. After reading the Last Card, that is, immediately after reading all profile cards, it reads cards, having card Design Type 3 and 4. Provision is made in the program to get the printed list of complete set of Query Cards except Control Cards.

2 Information on cards having Card Design Type 2 is stored in Working Storage Area called Q-TERMS. Q-TERMS consist of 27 fields. In the first field, information on card having the card design type 3 is stored.

3 Search Expression is stored in the Working Storage Area called STC.

4 Search Expression is converted into Polish Notation and stored in the Working Storage area called POL.

5 Polish Notation is then processed. To pick up the relevant profile word from Q-TERMS, table look-up procedure is adopted. For example, if the operand is E, the binary value of 31 will be subtracted from 37 which is the binary value of E in ICL 1900 series. Now, the 6th field of Q-TERMS is referred. In the 6th field, the 3rd subfield, that is, from 7th character, gives the profile word. Then, using the 1st subfield of the 6th field in Q-TERMS, Term Type may be obtained to pick up the relevant data element from the bibliographic record. After picking up the data elements to be matched, the computer matches the two data elements. If the match is perfect, the value of operand is true; otherwise, it is false.

5B5 Flow Chart

The Block diagram of SDBI is illustrated in Sec 95.

5B6 Sample Input

The sample input (only query on punch cards including control card) is illustrated in Sec 942.

6 Operating Procedure**6A PROGRAM CORC**

1 The program is compiled, loaded and executed. After the compilation, the object program

may be dumped on punch cards. Then, next time the program is run, the compilation step may be omitted and the binary file, which is on cards, may be loaded directly and executed.

CORC, when executing, reads the data base from punch cards. It outputs on to the magnetic tape.

To begin with, the following message will appear on the console typewriter:

DISPLAY—ON 1, IN CASE OF RETRIEVAL
HALTED 01

Note.—If the program is used to create a data base, type GO and press the ACCEPT key on the console typewriter. The following message will then appear on the console typewriter:

DISPLAY—ON 15, IF CORC ON PRINTER
DISPLAY—OFF 15, IF CORC TO BE CREATED
DISPLAY—ON 23, IF SN TO BE ASSIGNED
HALTED 02

Note.—(a) If program is used to create a data base, type GO and press the ACCEPT key on console typewriter. It requires a scratch tape. The output file name is TESTDATATAPE.

(b) If program, i.e., used to assign the sequence number to bibliographical record, switch 15 and switch 23 should be made ON through console typewriter. Load CORC1975FILE and a scratch tape.

(c) At any stage, if the program is used to print on on-line printer, switch 15 should be made ON through console typewriter.

(d) After creating a data base on tape, called TESTDATATAPE, a message appears as follows:

DISPLAY—SORT THE TESTDATATAPE
DISPLAY—LOAD THE PROGRAM AGAIN
HALTED 03

Then, using the control cards, sort the tape file and load CORC for assigning sequence numbers to bibliographic record. It requires a scratch tape. The output file name is CORC1975FILE.

(e) To begin with, if the switch 1 is on, then program requires a data base called CORC1975 FILE and Query cards as is discussed in Sec 5A91.

(f) The message "PROGRAM IS WRONG" may appear while running the program. This implies data field tags are not numeric and therefore, program may have to be suspended.

6B MAJOR STEPS OF OPERATING CORC

6B1 Creation of Data Base

- 1 Reading control cards—four control cards are read. Fourth card should contain four asterisks in the first four columns.

6B2 Reading Catalogue on Cards (A scratch tape is required)

- 2 Computer reads the first card of the catalogue entry.
- 3 Converts the class number into its ordinal value. It converts only the first 9 digits and ignores other digits. However, all the digits of the class number are stored in record

and may be used for the purpose of retrieval. Ordinal value is stored in OVALUE. This information is used for the purpose of loading the tape file.

- 4 Reads other cards of the same catalogue entry. Herein, it assembles other cards and thus, formulates a record. Meanwhile, it also constructs a facet digit and stores it in record. Facet digit is stored in the area called FACEFFP (For detailed flow chart for formulating record refer article in *Library Science*, 5; 1968; Paper Q).
- 5 Write record on tape. The above steps 2-4 are repeated for all catalogue entries on cards.
- 6 Print out of the Date Base may be obtained by operating the switch 15 to ON status.
- 7 Sort the tape.
- 8 Read and execute the program (CORE). Operate the switches 15 and 23 to ON status. Herein each record is assigned a sequence number. Output is available on magnetic tape. File name is CORC 1975 FILE.
- 9 Printout of the sorted file may be obtained operating the switch 15 to ON status.

6B3 For Retrieval using CORC

- 10 Load and execute the program CORC. Operate switch 1 to ON status.
- 11 Reads user name card. (Punch user name from first column; only one card is permissible).
- 12 Read query card. (Card design is given in the manual) (Table 3). If the 1 is space, entire data base is searched, otherwise only particular type of records are searched, and in such cases, code for particular record should be punched in col 1.
- 13 Tag, which is available in col 3 of the Query Card is used for picking up the corresponding data element from the record, and then it is matched with Data Element from the record, and then it is matched with Data Elements given in Query Card. If perfect match exists between the two, the bibliographical record is printed on printer. A user may provide several query cards. At the end of the Query card of each user, Four asterisks card should be loaded, and at the end of the process two Four asterisks card should be loaded, consecutively.

6C PROGRAM SDBI

The program is compiled, loaded and executed. After the compilation, object program may be dumped on punch card. Then, next time the program is run, the compilation step may be omitted and the binary file which is on cards, may be directly loaded and executed.

SDBI, when executing reads the data base from tape file called CORC1975FILE. It outputs on to

on-line printer. Also it requires control card and query cards as discussed in Sec 5B6.

The message "PROGRAM IS WRONG" may appear on console typewriter which implies either search expression or control cards are wrong. Therefore, program may have to be suspended. However, on continuation of executing the program, by giving the command GO through console typewriter, program will write "NO DOCUMENT" on on-line printer and start processing the next query, if any.

6C1 Steps involved in SDBI

(This supplements the information given in background material).

- 1 CORC1975FILE is the input for SDBI.
- 2 Load and execute the program SDBI. It reads the control cards. This card gives information regarding the indicator digits and literals (only one card).
- 3 Reads user name card (only one card). After reading this card information is printed on the left side of the top of the page.
- 4 Reads profile cards. Maximum of 26 cards.
- 5 A single search expression is read. It is converted into Polish Notation.
- 6 Polish Notation is processed (using recursive programming technique) for each record. If the value of the Polish Notation for a given record is True, that is 1, record is printed on printer.
- 7 At the end, another search expression is read, if any, and procedure as given in Step 5 and 6 are repeated.
- 8 At the end of reading all the search expressions for a given user, another user name card is read and procedure is repeated as given in Step 2, 3, 4, 5, 6 and 7.
- 9 At the end of the process load the *Four Asterisks* card.

6D IMPLEMENTATION

The programs are written in COBOL language and were implemented on ICL 1901-A computer, with memory size 16K words, 4 bytes per word, 6 bits per byte or one character per byte. The following peripherals were used:

- 1 ICL 2105 Card Reader
- 2 ICL 1920 Card Punct
- 3 ICL 2405/2 Line Printer
- 4 ICL 1971 Magnetic tape; at least 2 tape units are necessary
- 5 ICL 2821 Magnetic disc (twin exchangeable disc store; two units)

7 Steps Involved in Design and Implementation of SDI Services

71 Job A: USER PROFILE

- 1 Identifying user for profiling. A dialogue between user and information analyst is helpful.

- 2 Assign descriptors for user profile. This also involves assigning term type, alpha code, and formulation of search expression.
- 3 Checking and verification of steps 1 and 2.
- 4 Punching and verification of user profile.

72 Job B: CREATION OF DATABASE

- 5 Identifying subjects of interest of users.
- 6 Scanning of publications such as articles in periodicals, technical reports, standards, etc.
- 7 Filling up of worksheet, writing bibliographical description. This involves classification/indexing of documents assigning descriptors to each document.
- 8 Checking and verification of steps 6 and 7.
- 9 Punching and verification of bibliographical description.
- 10 Transfer the input on to magnetic tape. File name is TESTDATATAPE.
- 11 Sort the tape and thus file Name of the sorted file is TESTDATATAPE.
- 12 Assign the sequence number to each bibliographical record. Output file name is CORC 1975FILE.

73 Job C: MATCHING

- 13 Matching data base with user profile.

74 Job D: NOTIFICATION

- 14 Despatch the result to user.

75 Job E: FEEDBACK

- 15 Feedback from the user. Some the above steps are to be repeated, if user is not satisfied. This may involve updating the user profile as well as database.

76 Job F: MAINTENANCE

- 16 Updating both user profile and Data base
- 17 due course, develop a manual, which should cover time analysis, job analysis and cost analysis.

8 Conclusion

The computer programs used in the experimental study envisaged a batch processing information system. In batch processing, the user cannot have direct access to the data base. The query is delegated to an information analyst. However, recent developments have shown that on-line systems can be used in a non-delegated search mode—that is, the user of the system is able to undertake productively the literature search without the interference of the information analyst, thereby avoiding the problems of misinterpretation and miscommunication. In the near future, when on-line facility becomes operative in the country, it is essential that the planner and manager of information systems take cognizance of it.

Therefore, with suitable amendments for the programs developed at DRTC, an on-line real time

Information Storage Retrieval and Dissemination System can be implemented in the library which has access to on-line facilities.

Programs are to be suitably amended so that truncated words such as *NAL*, LIPID*, *CHEM* etc. may be used for retrieving purposes. Such provision increase the recall of the systems. Also, it would takes care of the variation in spellings. Programs are to be suitably amended to accommodate more than 26 profile words for each profile.

In the pilot study, an attempt was made to present a budget for computer-based SDI services for the scientists in the Department of Biochemistry at the Indian Institute of Science. For the purpose of estimating the budget, it was assumed that the system may process about 20 UJP each having about 5 search expressions. Also, it was assumed that the data-base will have approximately 6,500 bibliographical entries which cover about 50 periodicals in the area of Biochemistry. It was then estimated that the cost of computer-based SDI service would be approximately Rs. 80,000 per annum (1, 2). But in the case of BHEL it may be noted that it costs approximately Rs. 57,000 per annum and this would be shared by about 10 units in the complex.

It is fairly clear that even if one takes a more pessimistic view and the assumptions made in estimating the budget might have erred on the conservative side, there seems to be substantial cost benefit in the computerbased SDI services.

Part of the cost of services can be collected by circulating the *data base* to other libraries at a reasonable price.

The use of standards (4, 5, 6, 9) such as MARC, INIS, CAN/SDI, etc, wherever possible, will certainly increase the efficiency of the systems and therefore, cost benefit. The programs are available at DRTC.

91 Bibliographical References

- | | | | |
|-------|--------|-----|---|
| | 2 Sec | 42 | HEMALATA IYER. Proposal for SDI service for the Centre of Advanced Study in Biochemistry, IISc (Project submitted in partial fulfilment of the requirements for the degree of Master of Library Science, Bangalore University, 1974-75). |
| | | 44 | |
| | 3 Sec | 3 | IS: 2381-1963. Recommendation for bibliographical reference. Indian Standards Institution. 1963. |
| | 4 Sec | 7 | ISO: 2709-1913. Documentation—Format for bibliographic information interchange on magnetic tape. International Organisation for Standardisation. |
| | 5 Sec | 8 | MARC MANUAL used by the Library of Congress. Part 1. Tape format. 1969. American Library Association. Chicago. 25-70. |
| | 6 Sec | 7 | RAVICHANDRA RAO (I K). Data format and data identifiers in bibliographical data base. (Paper presented at the 16th Indian Standards Convention (Bhopal) (1975). |
| | 7 Sec | 3 | — Generalised, approach to computer-based information storage, retrieval and dissemination. (Project submitted in part fulfilment of the requirement for the Associateship in Documentation of the Indian Statistical Institute, 1975). |
| | | 42 | |
| | | 44 | |
| | 8 Sec | 3 | RANGANATHAN (S R). Classified catalogue code with additional rule for dictionary catalogue code. Ed. 5. Assist by A Neelameghan. 1964. Bombay. Asia Publishing House. |
| | 9 Sec | 8 | VENKATARAMAN (S) and NEELAMEGHAN (A). Catalogue-on-tape. (Lib. Sc. 5; 1968; Paper Q). |
| | 10 Sec | 5A8 | WAINWRIGHT (Jane). Standard used by bibliographical tape services. A comparison. 1974. COSTI Project. SI/11423). |
| 1 Sec | 8 | | COURSE MATERIAL on the use of computer in information storage retrieval and dissemination. 1975. (Course material of special course organised jointly by DRTC and INSDOC, Bangalore. March-May 1975. Sponsored by the Department of Science and Technology, Government of India). |

COMPUTER-BASED SDI : COBOL PROGRAM

91 APPENDIX A: SPECIMEN COPY OF USER INTEREST PROFILE
 - FOR IISc STUDY
 (M-PAPER)

USER INTEREST PROFILE
 (Please fill in by typing)

Page No. 1 of Profile No.1

- 1 Name Prof H C Gama
- 2 Department Biochemistry
- 3 Field of Study Fat-soluble vitamins
- 4 Please state the specific subject on which you require information
- 1 Chemical method of estimation of Vitamin E
 - 2 Distribution of Vitamin E in body fluids
 - 3 Its role in membranes
 - 4 Anti oxidant properties of Vitamin E in normal and diseased conditions
- 5 Please furnish the keywords of the topics of your interest
- Vitamin E, Tocopherol, Biological membranes, Vitamin E deficiency and Thrombosis, Antioxidant, Spectrophotometric methods
- 6 Please state any particular aspect(s) of subject interest on which you do not require information
- Biological methods of estimation of Vitamin E
- 7 Please mention the name of any particular author whose works you are interested in
- J A Lucy, Anne Morday, Nofstad I, L K Bjornson
- 8 Please state the titles of journals you generally read for getting the information
- 1 Journal of Lipid Research
 - 2 Lipids
 - 3 J Chemical Investigation
 - 4 Circulation
- 9 Please give below titles of articles/reports which are relevant to your subject. (Not more than two)
- 1 KAYDEN (H J), CHOW (E K) and BJORNSON (L K).
 Spectrophotometric method for determination of tocopherol in red blood cells. (J Lipid Res. 14;1973;533).
 - 2 Comparison of exchange of tocopherol and free cholesterol between rat plams, lipo proteins and erythrocytes by C.K Bjornson, C_Grieskowskie and H J Kayden. (J Lipid Res. 16;1975;39).

USER INTEREST PROFILE

(Please fill in by typing)

Page No. 2 of Profile No.1

State your search request in the narrative form. Add two references of papers published by you or of a colleague working in your field.

I am interested in the chemical methods of estimation of Vitamin E. The content of Vitamin E in biological membranes, such as red blood cells, platelets and its role as antioxidant in imparting stability to biomembranes is also of interest.

I would also like to know the effect of Vitamin E deficiency and consequent effects on membranes particularly the platelets and thrombosis.

1 J A Lucy : Functional and structural aspect of biological membranes: a suggested role for Vitamin E in the control of membrane permeability and stability. Ann N Y Academic Sciences 203, 1972; 4.

2 R E Olson. Vitamin E and its relation to heart disease. Circulation. 48, 179.

3 H G Lovelady. Separation of and tocopherols in the presence of and tocopherolis and Vitamin A acetate. J Chromatography, 78, 449.

COMPUTER-BASED SDI : COBOL PROGRAM

92 APPENDIX B: SPECIMEN COPY OF USER INTEREST PROFILE
FOR BHEL STUDY

Form P: User Profile Sheet

(Please fill in by typing)

Page 1 of Profile No. 1

(To be filled by the user)

Profile No.

Name and address
of user

K Ragavaiah
Manager
Control Planning
BHEL; Hyderabad - 500 032

STATE YOUR SEARCH REQUEST IN NARRATIVE FORM AND TWO REFERENCE OF PAPERS PUBLISHED BY YOU OR A COLLEAGUE WORKING IN YOUR FIELD

The main field of interest is in Long Range Planning comprising of manpower planning, investment planning, technology planning for heavy electrical industry. This involves knowing the technological trends in the industries using our products, e.g. petrochemical industries, sugar, paper, steel, etc. Studies includes techno-economic evaluation of projects - methodology as well as case studies; methods of technology transfer and the variables in it.

Reference

- 1 KATRALAS (Harvey) and LEININGER WAYNE (2). Manpower planning model for organizational effectiveness. Long Range Planning, 8, 4. 1975.
- 2 SURREY (A J). World market for electric power equipment. Sussex, science policy. Research unit University.

RAVICHANDRA RAO and HEMALATHA IYER

Form Q: Profile Coding Sheet Part 1

(Please fill in by typing)

Page 1 of Profile No. 1

(To be filled by information analyst),

TT	AC	Profile words
(a)	(b)	(c)
B	A	Planning
B	B	Long range
B	C	Manpower
B	D	Investment
B	E	Technology
B	F	Heavy
B	G	Electrical industry
B	H	Trend
B	I	Products
B	J	Petrochemical
B	K	Sugar
B	L	Paper
B	M	Steel
B	N	Evaluation
B	Ø	Techno economic
B	P	Methodology
B	Q	Case studies
B	R	Transfer
B	S	Organizational
B	T	Effectiveness
C	U	Katralas (Harvey)
C	V	Leininger Wayne (E)
C	W	Surrey (A J)

COMPUTER-BASED SDI : COBOL PROGRAM

Form Q: Profile Coding Sheet: Part 2

(Please fill by typing)

Page 1 of Profile No.1

(To be filled by information analyst)

EN	Search expression
1	a 0
	b $(A*B*C*D*E)+(F+G) \#$
2	a 0
	b $(E-H) + (I+(J*K*L*M)) \#$
3	a 0
	b $(\Theta+N) + ((P*Q) * (E+R)) \#$
4	a 0
	b $(A+S+T+(B*C*D*E)) \#$
5	a 0
	b $U*V*W \#$

Note: In each row, a = card having design type 3 and b = card having card design type 4.

93 APPENDIX C: SPECIMEN COPY OF WORK SHEET FOR BIBLIOGRAPHICAL DESCRIPTION

Form C: Article in Periodical

WORK SHEET
Machine Readable Bibliographical Data Base

Tag	Data Elements	Designation
1	Class Number	621.314:538.23
2	Descriptor (including report code number and product code number, if applicable)	Hysteresis, Magnetic flux, Power transformers, Transients, Transformers, Ferroresonance, effects
3	Name of the author	Teape (J W), Slater (R D), Simpson (R R S), Wood (W S)./
4	Title of the article	Hysteresis effects in transformers including ferroresonance
5	Title of the periodical (Abbreviated)	Proc IEE./
6	Volume No., Issue No.	123,2;/
7	Year of Publication	1976;/
8	Pagination	155.
9	Any other information, such as series	/
10	Any other information, such as ISSN for periodical	
11	Code for BHEL unit	BHEL - Hardwar.
12	Language of the document, if other than English	£

COMPUTER-BASED SDI: COBOL PROGRAM

94 Appendix D: Input

To create a data base

```

1 484950515253545556571112 15 01 03 08098407021613141085 58596061626364
2 65666768697071727374757677787980818283 17042252425262728293031323334353637
3 38394041424344454647
4 ***
1 X2,Y#HUMAN RESOURCE, WOMAN WORKER, SOCIAL STATUS, GERMANY#HARTMANN (N) /TXAX01
2 E ENTERPRISING WOMAN; A GERMAN MODEL#COLOMBIA J OF WORLD BUSINESS. /S1 Z1 / AX02
3 1970: /61-6#(ENGLISH)E AX03
4 X2,Y#HUMAN RESOURCE, WOMAN WORKER, GERMANY#US SOCIAL SECURITY ADMINISTRATION#X01
5 #, OFFICE OF RESEARCH AND STATISTICS, /LIFE TIME ALLOCATION OF WORK AND LEIS#X02
6 UPE, #WASHINGTON#1968, /44P.// (ENGLISH)E BX03
7 X2,Y#HUMAN RESOURCE, WOMAN WORKER, FRANCE#MICHEL (A), /NEEDS AND ASPIRATION#X01
8 S OF MARRIED WOMEN WORKERS IN FRANCE, #INT LAB R, /94: 1; /1966: /39-53#(ENGL)CX02
9 ISH)E CX03
10 08,2#PLANNING MACHINE, THE DOUBLE SPIRAL DRIVE, WARD LEONARD SPEED CONTROL,AA01
11 PENDANT CONTROL, HIGH SPEED MILLING HEAD#COOPER ENGINEERING LTD #POOMA #PLANAA02
12 MING MACHINE, #DH-4, #COOPER, #VOLTA LIMITED, POST BOX 900, BOMBAY-1.#/ENGLISAA03
13 CJE AA04
14 /D7, B #633#CHAPIN(N), /AN INTRODUCTION TO AUTOMATIC COMPUTERS, #/NEW YORK, #D VAC01
15 AN NOSTRAND COMPANY INC, /1963#(ENGLISH)E AC02
16 /E #67#CEMISTRY#GESS(FRED C), /CEMISTRY MADE SIMPLE, #B2152, #LONDON, /W @ ALLAA01
17 EN, /1067#(ENGLISH)E AA02
18 /D7, B #6KNUTH(DONALD E), /THE ART OF COMPUTER PROGRAMMING, #/CALIFORNIA, /ADDISAD01
19 ON, #ESLEY PUBLISHING COMPANY, /1969#(ENGLISH)E AD02
20 /D7, B #6#NITTPANN, /DICTIONARY OF DATA PROCESSING, #/AMSTERDAM, /ELSEVIER SCINTAE01
21 FIC PUBLISHING COMPANY, /1973#(ENGLISH)E AE02
22 /D7, B #6DESIGNOF(V H), /COMPUTERS AND THEIR USES, #/ENGLEWOOD CLIFFS N J, /PRENAF01
23 TICF-HALL I.P.C, /1964#(ENGLISH)E AF02
24 X#ECONOMICS#//////////E AB01
25 X2#RESOURCE#//////////E AF01
26 X2,Y#HUMAN RESOURCE#//////////E BB01
27 X2,Y#HUMAN RESOURCE, DEVELOPMENT THROUGH JOB ORIENTED EDUCATION, INDIA#CONCT101
28 ENTREATED EFFORTS TO IMPROVE HUMAN CAPITAL THROUGH JOB ORIENTED EDUCATION, #PNT102
29 LIT LEON RER /1,29; /1970: /P4, #/ENGLISH)E TI13
30 /#5#DISCUSSION METHODOLOGY#FREELY(A J), /ARGUMENTATION AND DEBATE, #R2765, #6EJ101
31 LMOINT, #WADSWORTH PUBLISHING CO INC, /1971#(ENGLISH)E JJ02
32 /7#SYSTEMOLOGY#NEGREENE(K B), /SYSTEMS PSYCHOLOGY, #B2418#NEW YURK, /MCGRAW HILWH01
33 L BOOK COMPANY, /1970#(ENGLISH)E
34 X2, Y #HUMAN RESOURCE MOTIVATION INFLUENCED BY INITIAL HOURS OF WORK #ARGINA#01
35 L UTILITY, #SANENA(C P), /WORKERS CLASSIFICATION BY ACTIVITY STATUS, /ARTHA VJ#RR02
36 ANA, /11; /1969/40Q-10#(ENGLISH)E RR03
37 X2,Y#HUMAN RESOURCE CANADIAN NATIONAL EDUCATED IN USA PERMANENTLY STAYING TM01
38 L US: #ECONOMIC ASPECTS#COMAY(P Y), /BENEFITS AND COSTS OF STUDY ABROAD AND MIM#02
39 GRATION, #CAVAD J ECON, /3; /1970: /3V0-8#(ENGLISH)E MM03
40 /E #7C#STARK (J G), /MODERN CHEMISTRY, #B2424, #USA, /PENGUIN BOOKS, /1970#(ENGL001
41 ISH)E 0002
42 /E5 #64#ORGANIC SUBSTANCE CHEMISTRY #FIESER(LOUIS F), /ADVANCED ORGANIC CHEM#X01
43 ISTRY, #R922, #BU#BAY, /ASIA, /1964#(ENGLISH)E XX02
44 # N70#MANAGEMENT SCIENCE, #COVENTRY(WILLIAM F), /MANAGEMENT, #B3265, #LO-00M, /YY01
45 L N ALLAN, /1973#(ENGLISH)E YY02
46 #P3#COMMUNICATION, #CHERRY(COLIN), /ON HUMAN COMMUNICATION, #B2301, #LONDON, /MIZZ01
47 T PRESS, /1966#(ENGLISH)E ZZ02
48 08,2#TURNING AND BORING MILL, DOUBLE MILING VERTICAL, HIGH CUTTING SPEEDS LW01
49 URICATION AUTOMATIC QUICK AND RELIABLE CONTROL#PRECISION ENGINEERING WORKS LW02
50 GITTALDOV NATIONAL CORPORATION MULIN WORKS, #PRAHE C26CHUSLOVAKIA#DOUBLE MULI#W03
51 SING VERTICAL TURNING AND BORING MILL#SK 12, #STROJIMPORT, #/ (ENGLISH)E WW04
52 08,2#GRINDING MACHINE TWIST DRILL AUTOMATIC IN FEED DRILL DIAM#STAVELEY MACHWV01
53 THE TOOLS LTD ASQUITH DIVISION, #HALIFAX, #TWIST DRILL GRINDING MACHINE, #/ASQUITHV02
54 N, #/ (ENGLISH)E VV03
55 08,2#LATHE, SAL4MEN AND SCHLESNGER LIMITS, HIGH PRECISION SPINDLE BEARINGS, #S01
56 #AR (UX#RFLISATIONS D APPAREILS ET DE MACHINES OUTILS, #PARIS, #LATHE, #A42, #RS02
57 #I#U#P/ (ENGLISH)E SS03

```

RAVICHANDRA RAO and HEMALATHA IYER

941 To retrieve document

I K RAVICHANDRA RAO

A D8,2

A O D D8,2

942 Query Cards

1 Control card as discussed in Sec
 2 I K RAVICHANDRA RAO
 3 2 A PLANNING MACHINE
 4 2 B THE DOUBLE SPIRAL DRIVE
 5 3 C COOPER ENGINEERING LTD
 6 1 D X2, Y
 7 2 F HUMAN RESOURCE
 8 2 F COST
 9 2 G HOBBING MACHINE
 10 2 H MACHINE TOOL METAL CUTTING
 11 3 I HEMMERMESH (D 5)
 12 2 J USA
 13 1 K D8,2
 14 ****
 15 0
 16 E-F
 17 0
 18 A* B-C
 19 0
 20 C* (A+B)
 21 0
 22 A+B
 23 0
 24 (D* E)+(I*J)
 25 0
 26 -H
 27 D
 28 K
 29 4
 30 D
 31 ****
 32 ****