

DST SEMINAR ON

Application of Computers to Bibliographical
Information Processing : Some Developments in India

(Bangalore;(10-13 July 1978)

EXPERIMENTS IN COMPUTER-BASED SDI SERVICE
AT BHEL AND PLANS FOR THE FUTURE

L J HARAVU, Deputy Manager, Bharat Heavy Electricals
Limited, Corporate Research & Development, Vikas Nagar,
Hyderabad 500 593

1 BACKGROUND

II BHEL and Its Engineering & Development Centres

Bharat Heavy Electricals Limited (BHEL) is among the top ten international manufacturers of power plant and related equipment. BHEL's nine manufacturing establishments produce a comprehensive range of products to meet the needs of power generation, transmission, distribution and utilization required by public utilities; large and medium industries and the transportation sectors of the economy. Consultancy and back-up services of BHEL together with manufacturing activity provides the total package of goods and services to customers. The Company's Corporate Research and Development Unit conducts basic and applied research in the fields of mechanical and electrical power

Haravu

equipment and systems, materials sciences applied mechanics, power electronics, insulation systems, electio-technology and related sciences: The Energy Systems and New Products (ESNP) division also undertakes basic and applied research in thermal plant engineering, heat transfer, fluidised combustion systems, etc.

The product engineering and design function at BHEL is carried on at a number of Engineering & Development Centres (EDCs). The EDCs were established in 1975 as a result of a major organizational change that itself came about out of the realisation that there was need for BHEL:.

a) to rationalise its design and production, which hitherto was based on the different design philosophies of more than one collaborator for a product;

b) to pool the technical know-how and information gathered from collaboration agreements; and

c) to develop where possible, an indigenous technology based on knowledge of different design philosophies (1).

The EDCs, together with the research divisions of BHEL, are envisaged as an important means of raising the engineering and innovative capabilities of the firm as opposed to mere manufacturing capabilities. The objectives of the EDCs are:

a) to ensure that engineering and design of the product line are according to the best available know-how; and

Computer-Based SDI Service at BHEL

b) to develop, competitive, new and improved designs for specified technical requirements,

Implied in the objectives of the EDCs is the need for support services that would enable the EDCs and R&D Division to meet their objectives. Among the support services envisaged for the EDCs. technical information services came to occupy a prominent position.

12 Genesis of the Experimental SDI Service

The objectives of the EDCs make it necessary for them to keep abreast of national and international technological developments. Among the different ways in which such a need can be fulfilled is through the provision of a variety of current awareness, information re-packaging and retrieval services to the EDCs. SDI service as a mode of current awareness provision is an attractive proposition since the SDI concept provides for a high degree of selectivity in the contents of the dissemination, personalisation and feedback ensuring that current user needs are always taken care of. The dynamic character of SDI makes it much more useful than traditional group oriented dissemination services,

A proposal to initiate experimental computer-based SDI service to the EDCs was therefore made in 1976 and was enthusiastically supported by the management of BHEL. The experiment was also endorsed by BHEL's Advisory Committee on Library & Information matters.

121 Options in Providing SDI

The two options available to an organisation considering the provision of SDI to its users are;

a) to buy profiles for its users with an established search centre in India or abroad;

b) to create its own data base by primary scanning, abstracting and indexing of new literature, and develop user profiles, software, etc to implement an in-house SDI.

The second option was chosen at BHEL. It was known at that time that INSDOC in collaboration with IIT, Madras, would be providing SDI based on IKSFEC tapes from the middle of 1977- It was, however, felt that it would be available from INSDOC. It was also felt that the in-house SDI would enable the better utilization of the available infrastructure of information resources and skills at BHEL, apart from providing the opportunity to develop indigenous capabilities to develop, organise and manage computer-based SDI services.

The provision of in-house SDI services requires:

a) a regular input of new documents ;

b) trained staff to index and describe the new documents and skills in interacting with and understanding users and their needs, to create and modify search profiles; and

c) accessibility to a computer and availability of computer programs for data base creation^ searching and printing of SDI notifications .

Computer-Based SDI Service at BHEL

The first requirement was met to a large extent at BHEL since it has already well established libraries in the older of its manufacturing establishments and at its Corporate. R&D Unit. New libraries and information facilities have also been set up in the recently established manufacturing and service units.

Trained library staff to index and describe new documents was also available at the different BHEL libraries. It was decided, however, to provide training to improve such skills.

Most of the manufacturing units have in-house computers. However, the software needed for SDI data base creation and SDI searching were not available.

It was decided to request the Documentation Research and Training Centre (DRTC; , Bangalore, to provide initial training in indexing, profiling and input preparation. The DRTC also was to provide the computer programs suitable for ICL 1900 series computers.

A one-week Workshop in May 1976 was organized by DRTC for the library and information personnel of the different BHEL Units. The workshop oriented participants to the principles, concepts and techniques of SDI , input preparation, indexing and search profiling. The experimental SDI service was started in July 1976. In what follows, the various features of the experimental service are described.

2 THE EXPERIMENTAL SDI PROJECT

21 Input Procedures

An important feature of BHEL's SDI was the fact that input to the SDI was contributed as a decentralised operation with each of the BHEL units doing the scanning and indexing of documents added to their libraries. Each library was allocated a distinct set of journals to be scanned. The journals were allocated to each library after taking into account the relative strengths of each library and the product profile of the Unit of which the library was a part. Scanning was done by each Unit to select items relevant not only to that particular unit but also to other units. A total of about 300 journals were scanned by the different units.

Standard worksheets, designed by DRTC, were provided to the different libraries to be used to record indexing and bibliographic information for different documents. The different kinds of documents that could be Input included books, technical reports, periodical articles, trade catalogues, patents and standards. In the experimental SDI, however, it was decided to input mainly periodical articles. A sample of the worksheet for periodical articles is provided as Annexure I.

211 Indexing

Indexing of input items was done using a standard thesaurus such as the Thesaurus of Engineering & ocie-intj-fie Terms (TEST). There was also provision to classify input items using the UDC.

Computer-Eased SDI Service- at BHEL

212 Central Processing

The input worksheets filled-in at the different units were sent to the Technical Information Centre at BHEL's Corporate R&D which acts as the Central Processing Facility. The information staff at the Corporate R&D edited the input; the editing work being confined mainly in ensuring that the conventions necessary for computer processing were in fact adhered to. Very little editing of the indexing work of the different units was done.

213 User Profiles

In accordance with well-established practice, users selected to receive SDI notifications were asked to express their information needs in as detailed a manner as they desired. They were also requested to cite one or two papers that they had found relevant to their needs. The statement of needs and cited references were used as the basis for creating the search expressions to be evaluated during the SDI runs.

A significant feature of the user profiles at BHEL was the fact that most of them were group-oriented profiles. One of the causes of the group orientation was the fact that each EDC was expected to contribute only one profile, with the result that the stated need attempted to reflect the needs of a large group of between 15 and 20 persons. This resulted in need statements that were very broad based in quite a few cases.

214 Computer Programs

Programs used In the SDI project were developed by

Haravu

DRTC and are described in detail in the manual prepared for the SDI Workshop (2). Only a brief description will be given here. Two programs are used in the SDI project. The first program (CORC) is used in order to create a bibliographic data base on magnetic tape from card input containing the bibliographic and indexing information for different documents. This program creates fixed length records of 512 characters. The format of the tape, however, is a directory oriented one as follows:

Character Position	Data	Remarks
0-3	Record Length	Binary
4-7	Sequence No.	Binary
8-29	Ordinal value of the Class Number	Used so as to achieve an ordering by Class Number in accordance with the order prescribed by the scheme of classification.
30*	Facet Structure	Binary. Used to represent the structure of a subject in a facet analytical scheme of classification
31-54	Tags	Binary. The data corresponding to a tag is indicated as a starting address relative to the first character of the record
55-310	Data Elements	
511	End of Record Symbol	
	Not used at BHEL	

Computer-Based SDI Service at BHEL

The tape format permits quick access to various data elements in the record. The program provides the option to output the bibliographical data base on the line printer in class number order.

A second program, SDBI, reads the user profile data on cards comprising the User's name and address, profile terms and the Boolean logical search expression. The search expression is first converted into Polish Notation. The availability of the logical expression in Polish Notation enables its evaluation in a single left-to-right scan. Also, in evaluating expressions in Polish Notation, it is not necessary to process all the operands and operators. The evaluation can terminate at a point at which there is no possibility of the search expression being true for a given document. The programming needed, however, to convert logical expressions to Polish Notation is fairly involved,

The SDBI program permits the use of the following term types in search expressions in addition to descriptors.

- | | |
|------------------|-----------------------------|
| a) Class Number | b) Author |
| c) Journal title | d) Place of publication |
| e) Publisher | f) Date of publication, and |
| g) Language. | |

The program evaluates one search expression at a time against the entire document data base on tape. The tape is rewound to start position after each search expression is matched against, the entire data base re-

Haravu

resulting in as many passes of the tape as the number of searches (profiles) in a SDI run. This limitation of the program is due to the limited core storage available in the ICL 1900 series computers.

SDBI outputs to the line printer the user's address, profile terms, the search expression and the bibliographical details of items that matched his profiles. A sample of the output is given as Annexure II. Both CORC and SDBI were written in COBOL.

3 PROBLEMS & SOLUTIONS

The problems faced in the experimental project pertained mainly to the following:

- a) Quality of input
- b) Quality of user profiles
- c) Quantum of input
- d) Computer programs.

3.1 Quality of Input

The input that was received from the different libraries was found wanting chiefly in the quality of indexing. It was found that items were indexed only superficially in a majority of the cases. Indexers in the different units had not used well established criteria for the identification of indexable concepts. The average depth of indexing per document was three. The capabilities of the computer for multidimensional searches was hardly utilized due to inexhaustive indexing.

Another aspect of the input indexing was the lack of vocabulary control. More than one thesaurus,

Computer-Based SDI Service at BHEL

depending upon what was available locally was used.

In a situation in which the file is to be used exclusively for current awareness provision and where the files are not going to be cumulated for retrospective searching, the lack of vocabulary control is not very serious. The users profile and search expression, however, need to be constructed taking into account this fact.

32 Quality of User Profiles

The profiles developed for the SDI were widely varying in quality. As already pointed out, the statement of needs given by most of the recipients were broad statements, e.g. 'Frequency response methods and measurement. Holography. Laser systems. Telemetry. Data acquisition. Electron microscopy'. At times, the need statement was vague, e.g. 'High voltage engineering - partial discharges in insulation'. In such situations, there is need for a meaningful interaction between users and the information analyst so that a profile that reflects the users requirements can be made.

Ideally, the development of a profile has to be seen as an iterative process : the user reacting to the outputs of the SDI over a period of time, and the information analyst making changes in the list of profile terms and search expression logic depending upon the user's reactions to outputs. The information analyst has also to be sensitive to the amount of noise that the user will tolerate.. It is useful If the user and the information

analyst can communicate **with each** other in an interpersonal manner. Interpersonal contact with users at BHEL is not always feasible. The result was that the profiles either retrieved too many references or missed references that were relevant-.

More recently, the problem resulting out of a broad need statement and lack of facilities for interpersonal contact has been tried to be resolved as follows.

The information analyst at BHEL's Corporate R&D does a manual search using an appropriate abstracts journal, e.g. Electrical and Electronics Abstracts, to receive a few references with a high probability of relevance to that user. The abstract or original is then studied to identify concepts that should be represented in the user profile. These are then searched in technical dictionaries, glossaries and thesauri to find out other terms that should be considered for inclusion in the users profile. In doing the search in an abstracts journals an attempt is made to identify concepts which are related to the concepts in the need statement in various ways. These relationships are sought as answers to questions such as, what test methods are reported in the literature? What, measurement techniques are reported? Is the work in the field mainly theoretical or is it experimental? The answers to these questions provide valuable clues that could be used to expand a need statement that was initially vague and ill-defined. A case study of a typical profile to illustrate the above method is given below:

Computer-Based SDI Service at BHEL

321 Case Study

Users need statement : High voltage engineering - partial discharges in insulation.

References provided by user: Nil

Method followed: As a first step, both Thesaurofacet and TEST were consulted to locate the term partial discharges. Both the thesauri did not contain the term. The term 'Electrical Discharges' was found. The facet classification of Thesaurofacet was consulted under the class number for Electrical Discharges. The terms 'Carona', 'Surface Discharges' and 'Flashover' were found to be other terms related to Electrical Discharges in a hierarchical manner. A search was done in two issues of Electrical and Electronics Abstracts to retrieve a few references on Electrical Discharges. Many references to articles on Electrical Discharges were found and two of them contained the term 'Partial Discharges'. The abstracts of these revealed that the term 'Partial discharges' was synonymous with 'Corona' and that it is a surface discharge. The articles also revealed that 'Partial discharge phenomenon' are studied in relation to insulation of high voltage transformers, transmission lines, switch-gear end rectifiers. It was also, found that the study of partial discharge was, by and large, theoretical in nature, though there was a reference to the use of ultrasonic methods of measurement of partial discharges. Most of the references revealed that the major concern was the testing and measurement of partial discharges in insulation.

Armed with the concepts found as a result of such search, it was possible to construct a profile that met the approval of the user when he was asked to get the terms in his profile along with his first SDI output. The user, in fact, was able to provide meaningful feedback. He indicated that he was not concerned much with the study of partial discharges in the insulation of rectifiers and that he wanted us to add the term 'impulse voltages' to his profile. Amendments have since been made to the users profile.

The success that this method has achieved leads the author to believe that similar spade work by information analysts would lead to more meaningful interaction even in cases where the interaction is interpersonal.

33 Quantum of Input

The volume of output that a user will receive is, to some extent, dependent upon the number of items that are input to the data base. At BHEL, the average monthly input was of the order of about 300 items consisting mainly of periodical articles. These came from periodicals mainly in the English language. This size of input together with the lack of exhaustivity in indexing and inadequate user profiles resulted in the output of an average of 3 to 7 items per user. In some cases, a user did not get any output at all. Justifiably, users felt that the size of coverage of the SDI was poor and that in order To make it more useful, the system should minitor a larger quantity of journals.

Computer-Based SDI Service at BHEL

The author found by simply counting the references in two issues of EEA, that about 6 out of 10 references of potential value to users at BHEL were written in languages other than English. Such a count, done as it was for a small sample, cannot obviously be used to make valid conclusions. It does point, however, to the possibility that unless items in non-English periodicals are also input, the coverage of the SDI would be poor. Since a majority of the periodicals received in the different libraries of BHEL are in English, the only way this can be achieved is to scan secondary sources to add to the input. This is now being done in addition to the scanning of primary journals.

Another dimension to the experimental SDI was added when INSDOC in collaboration with IIT, Madras, commenced the long awaited SDI service based on INSPEC tapes. It is obvious that the in-house SDI cannot achieve the size and coverage of INSPEC. Also depending entirely on periodicals as the source of input for the in-house SDI would only duplicate what would perhaps be available from the INSPEC tapes much earlier than what the in-house SDI can provide. One justification for duplicating in BHEL's SDI what would be available on the INSPEC tapes is the hope that the in-house SDI will one day serve a larger section of the technical personnel at BHEL at a lesser cost ~~compared~~ to using commercial services. Another justification is the need to build a data base, that after some years, could be used for retrospective searching. Such a data base would be

Haravu

selective, multidisciplinary and quite conceivably useful not only to BHEL, but to other users interested in power engineering.

For the present, a decision has been taken to continue the input items for primary periodicals as well as from secondary periodicals. Also, a decision to scan patent and report indexes so as to provide additional input to the SDI has been taken. These steps would increase the size of the monthly input to the SDI and at the same time cover document types not adequately covered by INSPECT

34 Computer Programs

Problems pertaining to the computer programs used in the SDI project stemmed mainly from the speed at which the programs operate. For an average monthly input of 300 items and for about 30 profiles, the processing time required is about 6 hours of computer time. On an average, it takes between 15 and 18 minutes to process one search expression. The reason for the relatively slow speed, as already pointed out, is the fact that it requires as many passes of the tape as there are search expressions. The speed of the search program is dependent upon both data base size as well as the complexity of the search expression.

The cost of computer time for a monthly SDI run is about Rs 4,800/- at Rs 800/- per hour on the ICL 1900.

Computer-Based SDI Service at BHEL

The speed of the programs seriously limits plans to extend the service to a larger clientele at BHEL. A new set of programs for data base creation, user profile validation and searching have been commissioned. These programs make use of disc storage and inverted file organization for the descriptor file. These programs are described in a later section of this paper.

4 FUTURE PLANS

4.1 SDI as a Network-Based Service

SDI cannot be viewed in isolation. It should be considered as one component in a service-mix comprising group oriented dissemination services, retrospective search services, technical enquiry services, reference, back-up and referral services. It would be necessary to integrate SDI with other services so that the intellectual effort, computer time, etc invested in such service is fully utilized for the provision of a package, of services. In addition, It would be necessary to consider how the network of BHEL Libraries might acquire a comprehensive range of documents in the field of power engineering and technology without unnecessary duplication, based on a well thought out and planned acquisitions programs. It need hardly be stressed that the quality, range and quantity of input material determines to a great extent the quality and value of various services.

The long range plan at BHEL is to produce a data base comprising of and describing a wide range of documentary material generated within BHEL and outside. The

Haravu

data base would be created out of input to be contributed by all the BHEL Libraries. An acquisitions policy for the company to be defined will ensure that the range and quality of input material is comprehensive enough to cover present and probable needs not only of BHEL but other industries, public utilities, educational and research organizations needing power engineering and related information.

The data base planned would be used to produce a number of outputs, such as group oriented dissemination services, personalised SDI, Union Catalogues, Directories of ongoing research. The data base will be cumulative in nature and capable of being used for retrospective searching. The services would need to be backed up by document copying, translation and referral services. Each library of the BHEL would need to develop a retrieval system to be able to respond to local user needs quickly. Such a retrieval system is seen as a sub-set of the total data base to be developed. At some future date, it is envisaged that subject specialists in the different EDCs and the R&D divisions of BHEL will be in a position to utilise the document base available in the company and the various services available locally and centrally, to be able to repackage information useful in design, production, standardization and problem solving in general.

The realization of such a plan at BHEL would require the augmentation of skills of library and infor-

Computer-Based SDI Service at BHEL

mation manpower at BHEL, development of common and mutually compatible procedures, formats and tools and the development of computer programs for data base creation and production of various products and services mentioned above. All these will be taken up as projects by the technical information centre at BHEL's Corporate R&D.

A second Workshop for the Library & Information personnel of BHEL was held in 1977 in collaboration with DRTC. Other intensive courses on specific areas such as indexing, bibliographic description, profile editing. etc are contemplated.

A project to compile a Union Catalogue of Periodicals available in BHEL libraries has been completed. Projects of a similar nature have been taken up to computer-produce Union Catalogue of standards and technical reports. These would be updated periodically since the files would be machine readable. Also, the data for these from all the BHEL libraries would ultimately become part of the data base planned.

42 New Computer Programs for the In-House SDI

A new set of computer programs have been developed for the in-house SDI. These are described briefly below:

1. A program that validates document and descriptor data on cards and creates two tape files. The tape

Haravu

files are later loaded on discs as indexed sequential files using the file organization software of ICL 1900.

2. A program to validate and create the user profile file on tape.

3. A search and output program. This makes use of the two indexed sequential files and user profile data on tape.

4. File maintenance programs to cumulate the monthly input files. The cumulated files will be stored on tapes.

The salient features of these programs are:

1. Improvement in search time requirements by a factor of 60 as compared to the programs that were previously used. In a data base size of 300 items and 30 search expressions, the time taken for search and output is about 6 minutes.

2. The possibility for cumulation of the files created for the monthly SDI runs to facilitate retrospective searches. Such searches can be made against files of upto 10,000 documents in a batch.

The details of the file structures used, file organization and evaluation of search expressions are being presented in a separate paper to be presented in this seminar,

Computer-Based SDI Service at BHEL

5 CONCLUSION

The experimental SDI project has provided valuable experience in the organization and management of a computer-based SDI as well as in the technical aspects of SDI. The project will remain an experimental one at least for another year, by which time the service will be extended to about 100 users in EFIEL, Data pertaining to costs of in-house SDI provision and reactions of users will be collected so that the viability or otherwise of in-house may be established.

6 ACKNOWLEDGMENTS

I am grateful to the General Manager, Bharat Heavy Electricals Limited, Corporate Research & Development, for permitting me to present this paper at this Seminar.

REFERENCES

- 1 . Engineering & Development Organization: Design and Concepts: New Delhi, BHEL, 1975- 32 p.
2. Workshop on BHEL Experimental SDI Project, Lay 3-7, 1976. Jointly organized by DRTC and BHEL, Corporate R&D.

Haravu

ANNEXURE - I

Form C: Article in
Periodical

BHEL SDI PROJECT

Work Sheet : Machine Readable Bibliographical Database

(Please fill in by typing)

Tag;	Data Elements	Description
1	Class Number	(See Note 2)
2	Descriptor	(including report code! number, if applicable) (See Note 3 and 4)
3	Name of the Author	
4	Title of the article	
5	Title of the periodical	(Abbreviated)
6	Volume Number, Issue Number	
7	Year of Publication	
8	Pagination	
9	Any other information,	e.g. ISSN for periodical
10.	Any other information	(such as Series)
11	Code for BHEL Unit	
12	Language of the Document	if other than English

Computer-Based SDI Service at BHEL

Note:-

1. IS:2381-1963 (Recommendations for Bibliographical Reference) or a cataloguing code (AACR, CCC, etc) may be used for the purpose of choice and rendering of data elements.
2. Give class Number, if used.
3. Descriptors may be obtained using a thesaurus, such as the EJC Thesaurus or TEST.
4. Group the descriptors according to the subject of the document. If a document discusses more than one subject, list together the descriptors pertaining to each subject. If a descriptor occurs in different subjects of the same document, it should be repeated. For example, a document may cover topics on the following two subjects:
 1. Evaluation of computer programs developed for the purpose of SDI.
 2. Evaluation of computer programs developed for the purpose of MIS in India.

The descriptors should be listed as follows:

For 1: Computer programs, SDI, Evaluation

For 2: Computer programs, MIS, Evaluation,
India.

ANNEXURE II

K. C. BANDOPADHYAY, SR. DEV. ENGINEER, HEAVY PLANT
ENGNN. BHOPAL

- B A TURBOGENERATORS
- B B ROTATING GENERATORS
- B C ALTERNATING CURRENT GENERATORS
- B D BRUSHLESS GENERATORS
- B E SYNCHRONOUS GENERATORS
- B F DIRECT CURRENT GENERATORS
- B G ASYNCHRONOUS GENERATORS
- B H EXCITATION
- B I EXCITERS
- B J ALTERNATING CURRENT EXCITERS
- B K CRYSTAL RECTIFIERS
- B L SEMICONDUCTOR RECTIFIERS
- B M THYRISTORS
- B N THYRISTOR RECTIFIERS
- B O WINDINGS
- B P ROTOR WINDINGS
- B Q FIELD WINDINGS
- B R TRANSIENT RESPONSE
- B S VOLTAGE CONTROL
- B T VOLTAGE CONTROLLERS
- B U COOLING
- B V COOLING SYSTEMS

Computer-base-SDI Service at BHEL

((A+B+C+D+E+F+G) + (H+I+J)) + (((A+B+C+D+E+F+G) + (H+I+J)) + (K+l+M+N+O+P+Q+R+S+T+U+V))

- 31 CEXCITATION,VOLTAGE CONTROLLERS,TURBOGE-
 NERATORS , EXCITERS , TRANSISTOR RECTIFIERS
 PRESTON (F.D)EXCITATION SYSTEMS FOR LARGE
 GENERATORS. G.E.C JOURNAL OF SCIENCE AND
 TECHNOLOGY 44,1;1977;31.
 BHEL-RD.
- 141 C621 313 32:621.3.013.8
 SYNCHRONOUS GENERATORS, ALTERNATING CURRENT
 GENERATORS, DIRECT CURRENT, EXCITERS,
 ELECTRICAL PROTECTION GEAR,
 LEDR (Z).NEW FIELD SUPPRESSION EQUIPMENT
 FOR SYNCHRONOUS MACHINES WITH A ROTATING
 DC EXCITER.
 CZEC':.HEAVY INDUSTRY.7;1977;31-35.
 BHEL-BHOPAL.
- 146 C621.314.634:621.313.2.
 THYRISTOR RECTIFIERS, CIRCUITS, TURBO-
 GENERATORS, STEAM TURBINES, EXCITATION,
 RECTIFIERS
 LAMF (L J) . SHUNT THYRISTOR RECTIFIERS FOR
 THE GENEREX EXCITATION SYSTEM.
 IEEE TR NS. ON PVR APPS.SYS, JULY,AUG.1977;
 1219-1225-
 BHEL-BHOPAL.