

# Introducing and Implementing On-Line Bibliographic Retrieval Services in a Scientific Research and Development Organization†

MARY JANE RUHL\* and ELIZABETH J. YEATES\*\*

National Bureau of Standards Library, Washington, D.C. 20234

Received January 26, 1976

The National Bureau of Standards Library recently implemented on-line bibliographic retrieval services. Methods are given to orient and aid users in availing themselves of the services. Results are presented, based on appraisal of the services by users: value to users; most used data bases; problems requiring search revision; reasons for unsatisfactory results; purposes for request and use of search results; impact on subsequent library use; and future searching requirements. On-line capability impacts on library financing as a whole and on the library role in the community are described.

The NBS Library which serves the Bureau's scientific and administrative personnel assumed responsibility to contract for and monitor on-line retrieval services used within or outside the Library. The Library initiated purchase orders or contracts for all services and related equipment. The Library serves as the point of contact for service vendors, NBS users, and NBS administrative functions.

The intent was to provide search services for the infrequent user, the bench scientist, or the information specialist, and to allow each to have the Library's reference staff do his search or to do it himself.

Following preliminary testing, the program of on-line retrieval services was initiated in March 1974. The first service was obtained from Sciences Information Associates (SIA), who offered *Chemical Abstracts Condensates* and the National Technical Information Service (NTIS) bibliographic data file. SIA then left the on-line retrieval business and turned over most of their customers to System Development Corporation (SDC). Librarians trained to use SIA's system, Battelle's Basis-70, were trained to use SDC's system. In August we introduced Lockheed's service; in October, the New York Times; and in March 1975, the Transportation Research Information Service through Battelle.

Since most potential users of the service were generally unfamiliar with on-line bibliographic retrieval, the Library initiated a series of educational and promotional activities. Also, we offered one free introductory search per person. We believe that a successful search is a powerful selling device, and the opportunity for a free search helps overcome a user's initial indifference or reluctance to take a chance with his own project money. Later searches must be financed by the user. If the user anticipates spending as much as \$200 a year, Library personnel encourage him to request an I.D. number to access the services directly. This is done through the procurement office within existing contractual arrangements.

## EDUCATIONAL AND PROMOTIONAL ACTIVITIES FOR ON-LINE SERVICES

Since we realized that knowledge of data bases is as important as learning about retrieval systems, we initiated a series of seminars on the various data bases, and we sponsored workshops and training sessions on the on-line services. Figure 1 gives a chronology of events.

We promoted these activities by placing announcements in the *NBS Standard*, a biweekly internal publication; the *NBS Technical Calendar*, a weekly publication that announces

technical meetings and events of interest to NBS employees; and the Library's monthly publication that announces new library acquisitions and other Library-related news. We placed posters in the NBS Administration Building and Library lobbies and in the cafeteria. Before the paper shortage became severe, we had flyers prepared which we sent to technical and administrative personnel. We requested announcements to be made at Institute and Division meetings. We even served refreshments at the seminars. All of this promotion, which took advantage of most of the mass communication channels available within NBS, brought us from 50 to 100 attendees at the data base seminars and about 30-50 attendees at the workshops. We held training courses for both Lockheed's and SDC's services. There were eight attendees at each of those. These sessions will continue to be scheduled whenever six to eight persons sign up for training. Plans are now being made for regular updating sessions for users every two months. Participation is limited to six to eight persons because only that number can be trained in a day, with four or five terminals available. All workshop attendees are given hands-on-the-terminal experience.

## LIBRARY-PREPARED USER MATERIALS

We prepared a set of forms for collecting and recording information related to the searches. They include (1) a search request form, (2) a search processing form, and (3) a search appraisal form, of which there are two versions: one for searches performed in the Library, and one for searchers who have their own user I.D. numbers and perform their own searches. The search request form is usually filled out by the search requester. The processing form is used for recording librarian's or searcher's use, data such as bases and vendors' services used, terminal time, number of printouts, turnaround time for the search, searching strategies used, description of computer "downtime", and the like. The appropriate version of the search appraisal form is sent to all first-time search requesters and to others at regular intervals.

We prepared a descriptive booklet entitled "Usage Guide to Computerized Bibliographic Retrieval available through the NBS Library", which describes both administrative and technical aspects of the retrieval services. The guide is directed primarily toward the non-Library users. It contains a description of the administrative procedures to be followed in requesting I.D. numbers and describes optimal searching techniques for the various services. Appendixes contain information on each data base's availability and bibliographic coverage, each data base's contents and related services available from the data base supplier, and a comprehensive list of definitions for the various symbols and codes used by the vendors for the respective data bases. These aids are meant to assist the user in remembering which bibliographic elements are searchable or printable by Lockheed's or SDC's services,

† Presented before the Division of Chemical Literature, 169th National Meeting of the American Chemical Society, Philadelphia, Pa., April 9, 1975.

\* Address correspondence to Ruhl Associates, 905 Enderby Dr., Alexandria, Va. 22302.

\*\* National Oceanic and Atmospheric Administration, D822 Rockville, Md. 20852.

1974: March	SIA service in use
*May	Federal Information Suppliers' Seminar (DDC, NTIS, SSIE)
July	SDC training for library staff
*August	Chemical Abstracts Service Seminar
*September	Lockheed training for library staff
*October	Engineering Index Seminar
November	Lockheed Workshop, NBS-Wide
	New York Times training for Library Staff
	SDC Workshop, NBS-Wide
	New York Times Seminar
	New York Times Workshop, NBS-Wide
1975: February	SDC In-Depth Workshop
March	Lockheed In-Depth Workshop
April	American Institute for Physics Seminar

\*New service was introduced

Figure 1. Calendar of retrieval service activities.

Total No. Search Requests:	267
Average Data Bases Searched/Request:	2
Average Minutes/Request:	30
Average Cost/Request:	\$42

Figure 2. NBS Library search statistics (March 1974–February 1975).

and what are the various alpha codes for author, title, source, pagination, etc. This is especially true for infrequent users, which most of the NBS "do-it-yourselfers" seem to be.

### USER APPRAISAL OF THE ON-LINE SERVICES

We have tallied the results from the search appraisal forms that were mentioned earlier. Before these results are described, a few definitions are required. A "search request" represents the question posed by the user, which we serve by searching as many of the data bases as are relevant with as many iterations as are required. A "search" represents the perusal of one data base in satisfying a "search request". Therefore a "search request" may be filled by one or several "searches".

Figure 2 shows the search activity that we have experienced over a period of one year. A number of variables are embedded in these figures, e.g., the learning curve for both searchers and requesters, the inherent cost and content differences among the services and the data bases, and the wide variation of requirements among search requests.

As was mentioned earlier, we sent appraisal forms to all search recipients through September, and thereafter to first-time users only. We stopped sending forms to repeat users because they were not returning them. We have not ascertained whether they considered the one response as representative of their total attitude or whether they were just unwilling to fill out a two-page form more than once. We sent out a total of 239 appraisal forms from March 1974 through January 1975, of which 116, or 49% were returned. During that period there was a total of 256 search requests fulfilled. Also, we sent the specially designed forms to the "do-it-yourself" users who perform their own searches. At present there are 18 NBS divisions that hold Lockheed and/or SDC user identification numbers. Most numbers are used by more than one person. We sent out appraisal forms and received answers from 10 divisions representing about 60 non-Library users. Most of their responses paralleled those from the Library user group. Since there were so few non-Library user respondents, the following data are those from the Library user group. Also described are the areas in which the non-Library group seem to differ. Most of the users in both groups are research scientists.

### SEARCH USEFULNESS

Figure 3 shows the user satisfaction expressed on the returned appraisal forms. Please note that non-Library users were asked to respond about overall searching while Library users were being asked to respond about one specific search.

Of the on-line searches performed by reference librarians, *CA Condensates* was used most frequently with most relative

	Service Useful	Service Not Useful	No Response
Library-performed searches	73%	18%	9%
User-performed searching	100%	-	-

Figure 3. User satisfaction.

	% Yes	% No	% NR
Search revision required:	22	56	22
Search rerun required:	11	74	15
Second search or revision more responsive:	30	13	57

Figure 4. Search revisions.

	%
Needed redefinition/requester	37
Applicability of data base(s)	30
Additional ideas triggered	18
Needed redefinition/librarian	15

Figure 5. Why revisions were required.

success, according to user responses. It was also the "only data base used" most frequently. NTIS ranked second in use frequency and in "only data base used"; it ranked third in usefulness. INSPEC was used relatively infrequently, most likely because it was not available until NBS began using Lockheed's service in August, but was ranked second in usefulness. Compendex was rated "not useful" by most respondents. SSIE, DDC, and AEC/Recon searches, though performed remotely (not at NBS), were used to fulfill search requests and were included on the appraisal form; they ranked moderately useful to not useful, very likely because of their marginal coverage of NBS related topics. We must point out that "usefulness" here encompasses suitability to the search requests. A number of topics were searched, particularly in new technological areas, with poor results in all data bases used. We do not know whether this shows an inappropriately chosen data base, a lack of established subject terminology, or a genuine lack of coverage of a new area.

### SEARCH REVISIONS

Figure 4 shows the percentages of search revisions required and relative improvement of user satisfaction. Most reruns were required during June; none during September, October, December, and January.

Figure 5 lists the primary reasons respondents gave for requesting search revisions.

If at all possible, the librarian arranges for the requester to be present with him or her at the terminal. Of the respondents, 43% indicated that they would not prefer to have been involved with the searching to a greater extent than they were. Twenty percent indicated they would prefer to have been more involved with the searching; 37% did not respond. Eighty-six percent of the users indicated that the search did not require too much of their time; only 2% responded that the search required too much of their time; 14% did not respond. The highest percentage of those wishing to be more involved was reported for May searches. We might assume that, as users became more familiar with the service, they gained confidence in the librarians' ability to accurately interpret and perform the searches in the users absence.

### DATA BASE AND RETRIEVAL SYSTEM CONSIDERATIONS

Respondents attributed unsatisfactory search results to the following: 12% indicated that incorrect terminology was used; 9% stated that coverage by the data base(s) was too specific; 8% stated that data bases used turned out to be inappropriate to their requests; 6% stated that coverage by the data bases

On-going research	33
Research paper	25
State-of-the-art survey	22
New topic	22
Prospective research	17
Keeping up in area of interest	17
Review article	14
Program planning	13
Presentation based on research	8
Book chapter	3
Other	5

(more than one response permitted)

Figure 6. Use of search results.

was too broad for their requests; 3% indicated that the primary articles ordered as a result of the searches did not arrive in time; and 1% felt too many citations were retrieved to review. Both groups of respondents said that problem areas were primarily (1) lack of familiarity with indexing vocabularies, (2) searching in new areas of technology, and (3) requirements for material older than the files contain. The overwhelming response was related to the difficulty in second-guessing the correct index terms to use for retrieval.

Those who perform the searches, including reference librarians and "do-it-yourselfers", said that primary problem areas were as follows:

- System searched by multiple search terms, resulting in numerous search revisions.
- Delays in logging in because of vendors' and Tymshare's overloads.
- Brief time of data base availability resulting in inability to access appropriate data bases when needed.
- Delays in updating files.
- Noncompatibility of user languages.

#### PURPOSES FOR REQUESTS AND USE OF SEARCH RESULTS

Figure 6 shows the primary uses that were made of search results. Thirty-one percent of the respondents said their search results would be used by others; 12% said their searches would not be used by others; 57% did not respond to this question. Two respondents said that five to ten others would use the search results; and 21 said that fewer than five persons would use the search results. Seventy-five percent of the respondents said these were the first searches on their respective topics; 15% indicated they had searched these topics previously; and 10% did not respond to this question.

#### IMPACT ON SUBSEQUENT LIBRARY USE

As a result of their searches, 39% added their search results to a file; 37% of the respondents were prompted to use source documents, primarily from the NBS Library; 32% said this search service would substitute for other searching methods; 30% said that this service would replace routine literature searching. Some respondents indicated that their searches prompted them to contact a colleague; others said their searches helped them to solve a problem, that the search results instituted a change in their work, or that they wished to institute a recurring search on the same topic.

We have noticed some impact on the Library services; particularly an increase in photocopying of journal articles. Circulation and interlibrary loan have not changed significantly. Although we have not performed a controlled study, this might indicate that our library collection is adequate for users' need in following up their search results. Clearly, libraries will differ in which functions are most affected (interlibrary loan, circulation, photocopy, acquisitions) depending on existing capabilities and available funds.

#### FUTURE SEARCHING REQUIREMENTS

Figure 7 shows respondents' anticipated future search requirements. Figure 8 shows respondents' anticipated re-

<u>Anticipated Annual Search Requirements</u>	<u>% Rank</u>
Frequently (more than 12 searches)	2
12 searches	5
5-6 searches	5
2-4 searches	18
1 search	5
Possibly 1 more search	9
Do not know	7
No response	49

Figure 7. Future searching requirements.

<u>Anticipated Annual Search Requirements</u>	<u>% Rank</u>
12 searches	1
3-4 searches	6
2 searches	11
1 search	2
Fewer than 1 search	1
None	55
No response	24

Figure 8. Requirements for recurring searches.

quirements for recurring searches.

#### CONCLUSIONS AND IMPLICATIONS FOR LIBRARY FINANCING

The responses summarized herein show the following:

- The search activity is regarded as a valuable service.
- Chemical searching is required most frequently.
- Standardization of indexing vocabularies would be beneficial.
- Most searches performed by Library personnel are being performed effectively and satisfactorily; users are as involved in the activity as they wish to be.
- Most searches are being used in relation to on-going research, to plan new research, and in relation to reporting present or completed research.
- Most research topics require two or four searches per year.
- Most searches result in subsequent further use of library materials.
- Most users do not require regularly recurring search service.

This summarizes NBS's first year of experience with on-line retrieval services. Our experience also supports some significant implications for library financing as related to computerized information retrieval services.

Libraries (at least in nonindustrial settings) have traditionally received their budgets from taxation or overhead. Within a given budget the library performs its services without additional charge. Computerized services allow libraries to perform more work with the same number of people but also require more funds, from somewhere.

A case can be made for supporting computer use from general funds, without user charges, that is, the concept of information as an educational resource, which should be societally maintained. Nevertheless, most libraries do not obtain long-term funding for the added costs of computerized retrieval services. The alternatives for them are not to offer the services or to charge users for them. The first alternative appears self-defeating to library and users, since computerized information services offer improved means for the library to perform a basic service.

The introduction of fee-for-service in this case allows us to reevaluate the purposes which general funding can be expected to serve in a library in the current straitened economy.

For example, perhaps general funding can be expected to cover only expenses of an overhead nature in which specific charges to users would be impossible or difficult to assess and

in which benefits are truly generalizable to the entire user group: e.g., library staff in general; most books and journals; and perhaps certain basic services, particularly of a short duration, such as standard reference service. Possibly collection components which are specific to an identifiable segment of the user population (as well as significant services to a group) should be financed by that group. (These remarks are pertinent to multidiscipline libraries, rather than to small, specialized ones.)

In addition, the library management may have to redefine the uses which general funding can serve. Perhaps general funding is sufficient only for the "normal" or "average" user of a defined community. Possibly the user with a high volume of demand or from outside the community may have to pay usage fees.

Libraries should assume the responsibility to optimize service

through providing expertise on cost-effective selection of sources and on development of interview and query techniques, user aids, and control and evaluation methods. We have not had time or funding to carry out carefully controlled comparisons of cost or content effectiveness among the various services and data bases. However, we plan to study the services and make recommendations to our users for their optimal utilization and application at NBS. As we have said, the successful management of a computerized retrieval facility is both a significant service in itself and an expansion of the library's capability for information service of a sophisticated nature. This role of the library can be recognized by the user group. The library staff members may gradually find themselves becoming more involved as experts in development of user projects. The effect may spill over into all library functions, e.g., collection and other resource development.

## Use and Implications of On-Line Information Retrieval for Management†

SAMUEL A. WOLPERT

Predicasts, Inc., Cleveland, Ohio 44106

Received December 15, 1975

The scope, size, and variety of information available through Predicasts Terminal System (PTS) is discussed. Through sample searches in PTS File 16 (*Market Abstracts*) for chemical data, the advantages and disadvantages of search strategies involving natural language and hierarchical codes are compared. The development of on-line information systems is capsuled, and forecasts are made for the future, including comments about the use and implications of PTS and on-line retrieval in general for decision makers.

By the end of 1975, Predicasts Terminal System (PTS) files contained more than one million records and 400 million characters. PTS File 16, *Market Abstracts*, has 90 000 records giving extensive abstracts with descriptor statements. File 17, *PTS Weekly*, contains both market abstracts and F&S Indexes representing weekly updates of PTS Files 16 and 18. File 18, *F&S Indexes*, gives brief comments with descriptors, and totals 500 000 records on business and economic developments around the world. File 19, CIN, contains the extracts from *Chemical Industry Notes* of the American Chemical Society, approximately 150 000 records. File 19 is incremented weekly. File 20, *Domestic Statistics*, has about 150 000 records, including forecast abstracts, historical and forecast time series for the United States, and components of the U.S. File 21, *International Statistics*, presently contains forecast abstracts and forecast time series for all the countries of the world (except the U.S.), and by the middle of 1976 will include historical time series as well. *EIS Plants*, File 22, contains descriptive statistics and categorization of the 120 000 principal industrial facilities in the United States.

Of all the PTS files, File 16 (*Market Abstracts*) is the one with the best potential use for chemical industry management. File 16 contains digests of articles appearing in hundreds of worldwide journals on new products, acquisitions, capacities, end-uses, market data, technology, production, environment, foreign trade, countries, and regulations for the extractive, manufacturing, transportation, utilities, and other industries. General economics and financial industries are not covered. One of the industry groups reported most intensively is the chemical process area.

While File 16 covers all the countries of the world, about half of the records are from the United States. By type of information, 20% deals with organizations, 30% with technology, 20% with resources, 20% with market data, and 10% with other material. Each File 16 record contains a verbal abstract, a paragraph or more in length, and field descriptors. Included with each of these descriptors are hierarchical codes, which are a powerful tool for low-cost searching (as we will see).

In searching PTS one may use natural language, hierarchical codes, or a combination of both. If we are using natural language, we may first begin with the dictionary to see how many records the file has on a particular word or word combination. The dictionary alphabetically lists all nontrivial words from the abstracts, as well as word combinations and verbal names from the thesauri for the affiliated numeric codes, with the number of times each word or term appears. If we are searching, for example, for records about polystyrene, we discover that the dictionary lists about twice as many entries for polystyrene where codes were not mentioned, as polystyrene records with codes. In the latter case, a human indexer indicated that each of these records was, in fact, about polystyrene. The disparity is accounted for by two facts: first, some of the time when the word polystyrene was used, the human abstractor coded the entry as polystyrene film or foam and not simply polystyrene; secondly, the human indexer considered the polystyrene in some of the abstracts as being trivial, just mentioned in passing and therefore uncoded.

The PTS Thesauri include the numeric product, geographic and event codes with the affiliated verbal names for each code, again listing the number of times each appears. Thus, in the product code (PC) thesaurus we find polystyrene entries under PC 2821310 (polystyrene), 2821311 (regular polystyrene),

† Presented before the Division of Chemical Information, 170th National Meeting of the American Chemical Society, Chicago, Ill., Aug 27, 1975.