

System of Georgia, and the best promotion of the services has been word-of-mouth between colleagues. Overlap between data bases, which has caused considerable discussion in the Information Science group which operates the program, appears to be of little concern to the user. A large majority of users report that they overlook any overlap without reaction, either pro or con. Over 97% of the respondees indicated that the computer-based services had contributed either some or substantially to their professional activities, with the major types of contribution being in the amount of time it has saved in searching the literature (though not necessarily a savings in time spent using library materials) and that it has significantly broadened the subject areas which they can routinely monitor. In response to the question concerning the way in which library use habits have changed, two principal points emerge. For some users the service has been a method of by-passing the library reference works; they use the retrieval results as a request bibliography to obtain the original document by several means, including use of clerical or student help. On the other hand, there is clear indication from the responses that the computer-based services have brought many users back into the library and made them aware of a great many sources of information which were previously unknown to them. The subsequent ease of use of the library was stressed sufficiently often to postulate that it was the reference materials, including such resources as the indices and abstracting journals (and perhaps the card catalogs) which were the bot-

tleneck in efficient and easy use of the library materials. The staff is encouraged that the computer-based services have, indeed, proved a helpful supplement to the traditional institutional library resources, and plan to work much more closely with the libraries in the University System to improve this interface between the two organizations.

ACKNOWLEDGMENT

The authors would like to acknowledge the partial support of the National Science Foundation under grant GN-851 for this work. Al Garbin, Professor of Sociology and Anthropology at the University of Georgia, gave guidance and advice on the design of the survey instrument.

LITERATURE CITED

- (1) Park, M. K., Caughman, M. C., and Hamilton, H. J., "Education in the Use of Modern Information Techniques," *J. Chem. Doc.* 11 (2), 100-2 (1971).
- (2) El-Hadidy, B., and Amick, D. J., "Approaches to the Econo- mical Retrospective Machine-Searching of *CA Condensates*," Addendum to Final Report to NSF on Grant GN-738, University of Pittsburgh, Pittsburgh, Pa., p. 79, December 31, 1971.

Comparison of Service Centers and Document Data Bases—A User's View*

C. H. O'DONOHUE

Philip Morris Research Center, Box 26583, Richmond, Va. 23261

Received July 6, 1972

A study was undertaken to familiarize our users with selected data bases, and to determine the quality of services offered by various information centers. This paper presents an analysis of five commercial service centers and various data bases. A current awareness profile covering both chemical and botanical fields was sent to each service center. The responses were evaluated on the basis of promptness, interaction with the user, and physical format of the listing. Also evaluated were the various data bases on content, printout format, and ease of manually checking citations.

The Philip Morris Research Center has a Technical Information Facility (TIF) which serves over 200 professional staff scientists. The relatively modest size of the TIF staff and of our library holdings does not permit us to provide complete, in-house information services. Therefore, we employ various commercial services to complement our own efforts. These outside services have computerized data bases which we find useful, but too expensive to purchase for processing in our own facilities. The present paper is a comparative evaluation of the services we have employed.

About a year and a half ago we decided to develop a Selective Dissemination of Information (SDI) program for a laboratory group which had a definite need to be kept abreast of current developments in its area. The group had a project encompassing both chemical and botanical fields. We picked data bases which we knew, *a priori*, would be pertinent as well as others whose potential utility was less certain. Our major objective (besides providing assistance for the laboratory personnel) was to determine which data bases and service groups would best serve our present and future needs.

After discussions with various sales representatives, four services were selected (Table I): BioSciences Information Service (BIOSIS), Institute for Scientific Information (ISI), Knowledge Availability Systems Center at the Uni-

*Presented before the Division of Chemical Literature, 163rd Meeting, ACS, Boston, Mass., April 10, 1972.

versity of Pittsburgh (KASC), and the University of Georgia Computer Center (U. Ga.). A fifth service, the Computer Search Center at Illinois Institute of Technology Research Institute (IITRI) was added after the author attended a workshop held at the center. Each of these groups received a description of the areas in which information was desired. These areas included plant growth, isolation of plant constituents, radiochemistry, and aerosol chemistry.

EVALUATION

Table II shows the times required for initiation of the various services plus annual fees charged. Service response time refers to length of time from the point where the request was sent out until the service firm contacted the user by letter or telephone. The next column gives the time required until the first citation report was received, and this total initiation time ranged from two weeks to a month and a half. Annual costs varied from a low of \$100 (BIOSIS) to \$450 (IITRI) for the same tape service, BA. These costs, in some cases, have increased since the inception of the profile.

Most of the services processed more than one data base for us. The *Chemical Abstracts* (CA) and *Biological Abstracts* (BA) tapes were run at two different centers. Tapes of the National Agricultural Library (CAIN) did not become available until the Spring of 1971, but shortly thereafter University of Georgia ran our profile and sent us two runs for our evaluation. IITRI processed our profile for a trial period, prior to our formal subscription.

The types of primary literature covered by these data

Table I. Services and Data Bases Selected

Service and Data Base	
BioSciences Information Service (BIOSIS)	
Biological Abstracts	BA
BioResearch Index	
Institute for Scientific Information (ISI)	
Automatic Subject Citation Alert	ASCA-IV
Knowledge Availability Systems Center (KASC)	
Scientific and Technical Aerospace Reports	NASA
International Aerospace Abstracts	
Department of Defense Documentation	DDC
University of Georgia Computer Center (U. Ga.)	
Chemical Abstracts	CA
Nuclear Science Abstracts	NSA
National Agricultural Library	CAIN
Computer Search Center at Ill. Inst. Tech. Res. Inst. (IITRI)	
BA and CA	

Table II. Initiation Times and Annual Costs

Service and Data Base	Elapsed Time From Initial Inquiry, Weeks		Cost, \$/Year
	Service Response	First Citation	
BIOSIS			
BA	1.5	5.5	100
ISI			
ASCA	3.5	6.5	173
KASC			
NASA	1.0	3.0	170
DDC			
U. Ga.			
CA	<1.0	2.0	168
NSA	<1.0	2.0	120
CAIN	NA ^a	NA	120
IITRI			
BA	2.0	2.0	450
CA	2.0	2.0	225

^a Not Applicable

bases are given in Table III. ASCA and BA covered only journals and not the patent literature, while DDC and NASA covered only government documents plus the patent literature. Documents surveyed both by NASA and DDC represent about 9% of the NASA and 30% of the DDC files. There is a 23% overlap in journals monitored by BA and CA. The CA, BA, and Engineering Index people have a joint committee studying this problem of overlap in journal monitoring.

One of our major points of interest was the ability of the alerting service to refine our profile in directions that would make data screening more effective. As Table IV shows, BIOSIS and IITRI were the best in this category. These two firms allow the user to indicate the relevance of each citation to his needs. Both revise profiles automatically and corroborate the changes with the user by telephone. University of Georgia gives the profile at the beginning of each citation report for ready reference and checks the citation listing against the profile. When necessary, the profile is revised, and the changes are corroborated with the user. ISI's form is devised so that changes can be requested directly on the citation form. One copy is sent to ISI; a duplicate is retained by the user. KASC was deficient in this area. A form which was filed requesting a telephone conference about revising the profile (which the user does not see) received no response until we actually called the Center ourselves.

Table V is a summary of information about citation formats and procedures. All the services (except KASC) use computer printout. KASC sends hard copies of abstracts. An abstract gives a better handle to the material than just a title-keyword listing. BIOSIS and IITRI list citations on individual cards. This feature appeals to the laboratory personnel, since the cards are easier to handle than paper printouts and are more convenient for personal files.

IITRI has a limit of 25 keywords per profile for a set fee. A charge is made for additional words above this limit.

Table III. Source Materials for Data Bases

	Patents	Journals	Items/Year
ASCA	No	2,500	?
BA	No	8,000	220,000
CA	Yes	12,000	270,000
DDC	Yes	Government	22,000
NASA		Documents	72,000
NSA	Yes	3,500	54,000

Table IV. Alerting Service-Initiated Efforts to Improve Interest Profile

	Form for Citation Relevance	Profile Sent	Accuracy Check
BIOSIS	+	—	+
ISI	—	+	+
KASC	+	—	—
U. Ga.	—	+	+
IITRI	+	+	+

Table V. Characterization of SDI Output

	Format		Keyword Restriction	Citation	
	Cards	Paper		Title	Abstracts
BIOSIS	+	+	None	+	
ISI		+	Word Charge	+	
KASC		+	None		+
U. Ga.		+	None	+	
IITRI	+		25 Words-Set Fee	+	

COMPARISON OF SERVICE CENTERS AND DOCUMENT DATA BASES

Table VI. SDI Efficiency

Service and Data Base	Total Hits	Per Cent of Total Hits	
		Relevant	Ordered
BIOSIS			
BA	667	22	5
ISI			
ASCA	173	20	8
KASC			
NASA	210	9	1
DDC			
U. Ga.			
CA	462	13*	8
NSA	1106	4	2
CAIN	643	15	5
IITRI			
BA	553	32	7
CA	539	54	9

* Over-all per cent is low as U. Ga. in cooperation with us studied the effect of various keywords. Using a controlled keyword profile, the % relevant was 50+.

ISI charges a per-word fee, the value of which depends on the frequency of the word's appearance in ISI's files. The other companies set no restrictions on the number of keywords per profile.

Articles cited by KASC and ISI are available from these centers upon request. We found this to be an advantage, since we sometimes had difficulty obtaining documents cited by other alerting organizations. Now with the existing copyright situation, this service becomes even more valuable.

With the exception of KASC, all the services were prompt in sending SDI reports. These notices would arrive before the hard copy of the abstracts reached us. We were unable to correlate notices from ISI with *Current Contents*, another ISI service to which we subscribe. However, from experience we know that ISI frequently cites articles before they appear in print. KASC's reports always arrived a week or so after we had received the source materials. All services except ISI gave volume and issue number of the source in the citation report so that references could be checked. This was done by comprehensive scanning of source abstracts. References that were overlooked in computer scanning usually had inappropriate or insufficient keywords.

Table VI shows the total hits in a one-year period, the percentage of total hits which were relevant and the percentage for which hard copies were subsequently ordered. The per cent ordered is low (1 to 9%), as the laboratory personnel do not request articles from journals which they routinely see or which are held in the library. Most requests were for foreign-journal articles or symposium proceedings.

Running the keywords, plant and radioactivity, through *Nuclear Science Abstracts* precipitates references to nuclear plants as well as botanical plants. This appears to be a problem of the NSA data base rather than the service group. Services with tight controls on keyword profiles did not appear to have similar problems in other data bases and had improved relevance percentages.

All of the services facilitate renewal of contracts through such devices as postpaid cards, letters, or telephone contact. At such time, they all offer to review and revise the profile if necessary.

Except for IITRI, every organization involved in this study—as well as several others—have run retrospective searches for us. Table VII summarizes the impressions of the various searches. Service was considered prompt if the turn-around time from request to receipt of reports was less than two weeks. Occasionally, firms would furnish

Table VII. Retrospective Searches

Firm	Time Response (<2 Weeks)	Format		
		Paper	Cards	Abstracts
BIOSIS	Yes			+
ISI	Yes		+	
KASC	No			+
U. Ga.	Varied	+		
N.C. Res. Triangle	Varied			+
MEDLARS	No	+		
IFI	Yes	+		

Table VIII. Criteria for Judging Search Services

1. Turn-around time
2. Familiarity of personnel with data base
3. Ability to structure a search question
4. The personal element
5. Type of output

Table IX. Summary Evaluation

Service	Time	Familiarity-Data Bases	Phrasing Search	Personal Element	Type of Output
BIOSIS	+	+	+	+	+
ISI	+	+	+	+	-
KASC	-	-	?	-	+
U. Ga.	+	+	+	+	-
IITRI	+	+	+	+	+
NCRT	-	-	?	+	+
MEDLARS	-	+	+	+	-
IFI	+	+	+	+	-

partial results if a delay developed during the search. MEDLARS requires four to six weeks to conduct a search. U. Ga. has a set time schedule for retrospective searches, but will perform rush jobs—i.e., overnight, for a special fee. The NASA Center at the North Carolina Research Triangle responded in times which depended on whether it had the source tapes or had to ask another NASA center to conduct the search. KASC and N.C. Research Triangle personnel did not seem to be as familiar with the data bases as were personnel from the other firms.

SUMMARY

The most important points considered in evaluating search services (both SDI and retrospective) are summarized in Table VIII. Turn-around time refers to a response time of less than two weeks. Items 2 and 3 are interrelated in that the personnel have to be familiar with a data base in order to know which are the appropriate keywords to use for a particular data base and which data base contains material relevant to the search question. The personal element is the interaction between the service center and the user. Positive type of output, in our opinion, is either abstracts or individual cards.

Based on these criteria Table IX presents the over-all impressions of all services with which we have dealt. A question (?) means that there is insufficient data with which the firm may be evaluated with respect to a particular point.

We have found that for our particular needs, CA is the best data base for SDI and for retrospective searches. BA is very useful for biological fields, if it is searched by one knowledgeable in its format. ISI's Citation Index and IFI's Uniterm File are profitable sources for selected questions. In general, we conclude that great care is required in the selection of commercial information services. The spectrum of potential satisfaction is wide, and the user must analyze his needs and his suppliers' capabilities carefully to optimize results.