

Current Comments[®]

EUGENE GARFIELD

INSTITUTE FOR SCIENTIFIC INFORMATION[®]
3501 MARKET ST., PHILADELPHIA, PA 19104

**If You Think You're Ready to Search
Online, Here's How to Use *SciSearch***

Number 4

January 28, 1985

Just two years ago, I first described the *Sci-Mate*[®] Universal Online Searcher.¹ This software permits you to use just about any personal microcomputer as a terminal to search hundreds of databases on a variety of host computers. Not the least of these databases is ISI[®]'s own *SciSearch*[®] file. Possibly the largest scientific database available today, this file contains detailed information on more than five million articles we have indexed since 1974. Indeed, the file is so large that one host, DIALOG, had to segment the file into four separate chronological chunks. However, DIALOG has just introduced Version 2 of its search-support software, which will eventually make it possible to use all of *SciSearch* as one unit.

On the penultimate page of the *Sci-Mate* essay mentioned above, I discussed the relationship of *Current Contents*[®] (*CC*[®]) to *SciSearch*. We are regularly asked when *CC* will go online. I have, of course, discussed the idea of an electronic version of *CC*,² and I will have more to say about that later this year. It will become evident below how *CC* and *SciSearch* are related. It may surprise some of you, as it did me, to realize that we have never described *SciSearch* fully in *CC* during the 12 years since the file was mounted. By mounted I mean that a magnetic version of the *Science Citation Index*[®] (*SCI*[®]) is sent to a "host" computer where it is loaded onto large magnetic discs. Using special searching software, you are able to access the information from your local computer terminal or your personal microcomputer.

The fact is that all the journals in five *CC* editions covering *Life Sciences*; *Clinical Practice*; *Agriculture, Biology & Environmental Sciences*; *Physical, Chemical & Earth Sciences*; and *Engineering, Technology & Applied Sciences* are included in *SciSearch*. Also included are 96 additional journals covered in *CC/Social & Behavioral Sciences* and 8 journals covered by *CC/Arts & Humanities*. In fact, *SciSearch* covers even more journals than does the *SCI*. As of January 1985, the total coverage of *SciSearch* was over 4,100 journals. Every one of these journals is indexed by all search elements including cited references. In previous years, so-called *CC*-only journals were indexed by title, author, and address only. In addition, there are more than 200 journals from *CompuMath Citation Index*[®] and *GeoSciTech Citation Index*[®] also included in this file.

So the prefix "Sci" in *SciSearch* refers not only to the *SCI* but also to science in general. There is another important difference between *CC* and *SciSearch*. Timing! *CC* is produced weekly on a very tight schedule. The magnetic tapes for *SciSearch* are sent to the hosts (vendors) weekly. DIALOG chooses to update the file every other week. So there may be a lag of two to four weeks between coverage in *CC* and coverage in *SciSearch*.

There are other technical differences between the printed products you know and love and the *SciSearch* file mounted on DIALOG. But it is not essential to the discussion that follows to explain these differences in detail.

In 1972, it was still an experiment, but now *SciSearch* is an online version of the printed *SCI*. Formally introduced in 1976, *SciSearch* was one of the first online bibliographic information services. Since then it has changed and grown extensively. The intent of this essay is to help fill the gap in the documentation of this file. More importantly, we want to explain how you can use *SciSearch* to retrieve information. My purpose here is to summarize, for the occasional user, how to use *SciSearch* to help in the routine business of doing research and preparing manuscripts for publication.

There has now developed a group of professional online searchers. They have mastered the command languages necessary to use database systems like DIALOG, BRS, DIMDI, NLM, DATASTAR, QUESTEL, and so on. Many of them have a vested interest in acting as search intermediaries. They use online systems on behalf of individual "end-users" of the information being sought. The *Sci-Mate* software and other user-friendly packages of this type have changed that "monopoly." These so-called front-end interfaces provide end-users with menu-driven systems for searching that eliminate the need to master host command languages. But we all know that the fastest and most effective method of online searching is to master the native language for the most relevant databases. So I've tried to summarize below how you can search *SciSearch* with a minimum repertoire of commands.

Not coincidentally, just a few months before I introduced you to *Sci-Mate*, we provided a primer on how to use the *SCI*.³ That essay stressed that the primary purpose of the *SCI* is information retrieval. It is a tool for searching for literature. This purpose is sometimes overlooked. I have myself discussed non-retrieval uses so often that many readers need to be reminded that information retrieval is ISI's bread and butter.

Why would anyone spend from one to three dollars a minute to do searching

online when you could do it free in a library? Well, by now you should be aware that your time or that of your associates is valuable. While computerized searching has many other advantages, it is best described as a fast copying machine. It simply types or displays references or text much faster than you can. Yes, you can use a directory or a printed index like the *SCI* to determine the existence of the information that you want. However, manually copying all those references requires a lot of time and energy.

Another reason is local access. Very few individuals can afford the cost and space to store huge files of printed indexes. In addition to the access and copying time, there are certain types of searches that are just too cumbersome to do by conventional manual methods. These combinational searches are relatively easy by computer. So let's get down to brass tacks.

Suppose that your next-door neighbor says his friend Bloggs has recently published a paper in the *Astrophysical Journal*. Someone said nice things about the paper at lunch. Since you don't know exactly when it was published, you do a simpleminded search on the name Bloggs. Well you find that there are in fact dozens of Bloggs who have published papers over the past decade. What do you do? You can either call your friend to ask if he knows Bloggs's first name or initials, or you can ask the computer to limit your search to 1984.

In Figure 1, we've reproduced a portion of the *SCI* under the name Smith. This will have to suffice since I couldn't find a Bloggs in the *SCI*. As you can see, there are dozens of papers by many different Smiths, but there were only a few published in 1984 in the *Astrophysical Journal*, of which we have shown one. Figure 2 illustrates the results of a search for papers by Smith in the *Astrophysical Journal* in 1984 using *SciSearch* on DIALOG.

You could have approached the same search problem a bit differently if your

Figure 1: Portion of the printed *SCF[®] Source Index* showing several entries for authors named "Smith."

Smith H -----				
•Keats, John—Poet, Patient, Physician				
REV INFECTION	6(3):390-404	84	18R	
Royal Free Hosp, Coppets Wood Hosp, Dept Infec Dis, London N10 1JN, England				
see Chambers ST	Lancet	2	181	84
see Coates DM	Pediat Res	18	809	84
see Friedman Z	Pros Leuk M	14	279	84
see Husseini RH	Immunology	52	389	84
see Martin PMV	Br J Ven Di	60	151	84

Smith HA -----				
•Metal Abundances of Galactic Globular-Clusters				
ASTROPHYS J	281(1):148-157	84	64R	
Michigan State Univ, Dept Phys & Astron, E Lansing MI 48824, USA				
see Quinnan GV	J Am Med A	252	72	84

Smith HC -----				
see Cutten AEC	Endocrinol	115	600	84
see Fisher LD	Am J Card	53	C138	84
see Holmes DR	"	53	C 77	84

Smith HE -----				
see Bettman JW	Ophthalmol	91	889	84

Smith HH -----				
•Michel DJ—Effect of Environment on Fatigue Crack-Propagation Behavior of Alloy-718 at Elevated Temperatures— Meeting				
J METAL	36(7):22	84	NO R	
USN, Res Lab, Washington, DC 20375, USA				
see Michel DJ	J Nucl Mat	122	153	84

Figure 2: Results of searching *SciSearch[®]* file 34 (1984 to present) for papers by "Smith" in the *Astrophysical Journal*. First, the author name is searched with the "S AU=Smith?" command. The computer finds 5,074 papers and stores them in a temporary file called "Set 1." Then you search the journal field with "S JN= Astrophysical Journal." The computer finds 1,114 papers and puts them in "Set 2." Finally, you ask that Sets 1 and 2 be combined by using the "S S1 AND S2" command—Smith in the *Astrophysical Journal*. The computer creates Set 3 containing 19 papers. The question marks in front of the search commands are computer "prompts" telling you that the computer is ready to receive your instructions. The "?" following Smith, however, is the truncation command that causes all Smiths, regardless of first name or initials, to be included.

File34*.SciSearch - 84/wk46
See files 87, 94, & 186

	Set	Items	Description
? S	AU=Smith?		
	1	5074	AU=SMITH
? S	JN= Astrophysical Journal		
	2	1114	JN= Astrophysical Journal
? S	S1 AND S2		
	3	19	S1 AND S2

coffee-klatch friend had said that Smith's paper concerned globular clusters. In that case, you could have started by asking how many papers there are in *SciSearch* under those keywords. Having found that there are dozens, you could limit them to those written by Smith in 1984 or 1985. *Voilà!* The refer-

ence you wanted appears. Figure 3 illustrates the results of this search. Of course, you would now go to the library unless you are satisfied to write away for a reprint. When you use *SciSearch*, the author's address is usually included as it would be in *CC*. Indeed, the address could have been a search element. Since Smith is at Michigan State University, you could have used this affiliation (if you knew it) as a search entry point.

Aside from the simple purpose of copying out references, computerized searches usually provide two additional functions: searching by fields and searching on combined fields to limit or expand the search. Combining sets involves so-called Boolean "operators." They are named after the logician George Boole. The operators are simply the terms "AND," "OR," and "NOT." These terms indicate that you want a narrower or broader search. If you say "A AND B," then you get papers discussing both A and B. If you specify "A OR B," you will get papers on either topic.

Another useful way to describe access to *SciSearch* is to relate the online search commands to the four major

Figure 3: Using *SciSearch*® to find papers by "Smith" with the keywords "globular" and "clusters." The first command causes the computer to search for the word "globular." Then the word "clusters" is searched (Set 2). Then combine 1 and 2 into a set containing 54 papers. Next you search for papers by "Smith." Then you combine Sets 3 and 4. The resulting set (5) has two papers. You then ask for Set 5 to be displayed by issuing the command "T5/5/1-2." The second paper is the one you wanted.

	SS Globular AND Clusters		
	1	122	Globular
	2	608	Clusters
	3	54	1 AND 2
? S AU= Smith ?	4	5074	AU=Smith?
? SS S3 AND S4	5	2	3 AND 4
? T5/5/1-2			

5/5/1
0467625 Article OATS Order#: TJ913 60 REFS
On the Relationship Between Giant-Branch Color
and Metallicity for Globular-Clusters (English)
Michel A; Smith HA
Michigan State Univ, Dept Phys & Astron/
E Lansing//MI/48824
Publications of the Astronomical Society of the
Pacific, V96, N582, P588-592, 1984

5/5/2
030445 Article OATS Order#: SW432 64 REFS
Metal Abundances of Galactic Globular-Clusters
(English)
Smith HA
Michigan State Univ, Dept Phys & Astron/
E Lansing//MI/48824
Astrophysical Journal, V281, N1, P148-157, 1984

components of the printed *SCI*. In addition, there are *SciSearch* commands that correspond to features of *CC*.

There are four basic information-retrieval roles that are possible with the *SCI*. These correspond to *SCI*'s four component indexes. *SciSearch* provides access by these, but in some cases, there is much greater flexibility in searching than is possible with the printed volumes. Let's discuss literature access via *SciSearch* by comparing it with the *SCI*.

SCI's *Source Index*, the master author index, tells you what someone has published and where. It is an all-author index. Regardless of precedence of authorship, it covers all primary and secondary authors. After you specify the author's name and initials, you are provided the article title, journal name, volume and pagination, publication year, and address. In *SciSearch*, you can search any combination of the following key information elements or fields—au-

thor, title words, journal (source), publication year, and address.

SCI's *Permuterm*® *Subject Index* provides access by title words. Every pair of significant words in the titles of the indexed items is displayed in the printed version. This comprehensive display command is not available on any existing online system as yet. In *SciSearch*, you can search for any combination of title words you wish to specify. Using the Boolean operators *AND*, *OR*, and *NOT*, you can find those papers that contain any combination you want. However, experienced online searchers often consult the printed *Permuterm Subject Index* to determine all the different terms that have actually co-occurred before starting an online search.

SCI's *Citation Index* is, of course, unique. This is how you find out where a particular paper or author has been cited. Using *SciSearch*, you can find out how often each cited reference has been cited and where. You do this by specifying the first author of the article, the year, and then the volume and first page. Cited-reference searches are extremely fast and cost-effective. For a cited-author search, you can truncate after you key in the name. But remember, this covers only those papers for which this author was primary, and it may cover two or more authors by the same name. By using the "expand" command ("E") with the "cited-reference" ("CR") label, you can observe the number of citations for different articles and help reduce the homograph problem. When you key in the command "E CR= Smith J" you are asking the system to list the articles by an author, "J. Smith," in an abbreviated form and the number of times they have been cited. Figure 4 illustrates the results of executing this search. If you are familiar with the author's bibliography, you can compare it against the list of articles displayed after you key in the expand command. You can then identify the appropriate articles and determine how often each paper has been cited. You could also ask to see the titles of

Figure 4: Using the "expand" command ("E") with the "cited-reference" ("CR") label. You ask for papers that have cited articles written by J. Smith in 1983. The computer responds by showing you not only the hits for 1983 (highlighted with an asterisk) but also lists specific entries for 1982 and 1983.

? E CR = Smith J, 1983

REF	ITEMS	INDEX-TERM
E1	2	CR = Smith J, 1982, V35, P56
E2	1	CR = Smith J, 1982, V41, P797
E3	4	*CR = Smith J, 1983
E4	3	CR = Smith J, 1983, P198
E5	1	CR = Smith J, 1983, P5
E6	2	CR = Smith J, 1983, V15, P2422
E7	1	CR = Smith J, 1983, V16, P1595

each paper. Keep in mind that this would not include papers by J.A. Smith, J.B. Smith, and so on.

SCI's Corporate Index tells you what has been published at a particular institution. *SciSearch* has a field label (CS) that searches the address field so that you can list all papers from a particular lab. Figure 5 illustrates a corporate search.

SciSearch also permits you to limit a search by document type and by language, for example, a review article in Russian or an editorial in English. Figure 6 summarizes the online entry points to the *SciSearch* database. I will review a sample search later in this essay to show several approaches for accessing information via *SciSearch*.

Those of us who have used systems like *Sci-Mate* forget that large numbers of first-time users simply want to access databases the "old-fashioned" way. This means that you first have to dial up a local connection through Tymnet, Tele-net, or Uninet. For this, you need some type of modem. Of course, you must also obtain a password to gain access to computerized databases. If you are using *Sci-Mate*, you can contact ISI directly, or you can contact DIALOG, DIMDI, or other hosts to obtain a password for direct access.

Once your call has been cleared by the DIALOG host computer, you would specify file 34 to access *SciSearch* 1984 to the present. Similar procedures can be used to access *Social SciSearch*® (file 7). Once you are on the system, you

have to use your newly acquired command language. (There is a search guide available.)

For most purposes the simple command "S" gets things started. To do a search on "Salmonella" you would simply key in "S SALMONELLA." The computer responds by telling you there are 476 papers indexed under that term and that these have been placed in Set #1. Since that is a lot of Salmonella to swallow, you can decide to narrow your search to papers concerning an outbreak of Salmonella. To do so, you key in "S OUTBREAK." The response is Set #2 with 148 hits. Then you ask to combine these two sets by the command "S S1 AND S2." The response is Set #3 with two hits, and this seems to be manageable. If it weren't, you could make new sets by adding other terms (such as the date or address of publication) and then combine them to reduce the number of hits. Incidentally, you could have searched "OUTBREAK?" to cover both the singular and the plural. In that case, you would have found 41 additional papers in Set #2. If you then combined Set #2 with Set #1 (your hits on "Salmonella"), you would have a Set #3 with five hits, which is also manageable.

Your next problem is how to turn this preliminary data into usable references. So now we are ready to print or display the information. For this you use the command "T." I assume that your micro

Figure 5: Searching on the corporate source field in *SciSearch*® file 34 (1984 to present) for papers published by authors at Swarthmore.

```
? S CS = Swarthmore
  1 33 CS = Swarthmore
```

Figure 6: Codes for searching specific fields in the *SciSearch*® database on DIALOG.

<i>SciSearch</i> Field Name	DIALOG Field Label
Author	AU
Cited References	CR
Corporate Source	CS or /CS
Document Type	DT
Journal Name	JN or JO
Language	LA
Publication Year	PY
Update	UD
Title	/TI

or terminal has a local printer and a display tube. Since there are several dozen papers that can be listed, you may want to sample the output. You then extend the "T" command to indicate the number of hits you want to display.

There is a further command necessary to indicate the format of your display or printout. Each *SciSearch* record consists of many different fields. The standard or "default" format is in the following sequence—title, author, and so on. However, there are at least six alternative formats, short and long, that you can use. Some people just need to see the title and author of each hit. Others may want the full record, including addresses. Most people today use 1200-baud modems and printers, but many still have only 300-baud capability. It could seem like forever waiting for the full format to display if you are still running at 300 baud, which is approximately 30 characters per second. After you have transmitted data for a few minutes, you will learn which options you prefer.

Generally, whenever you use the command "S" alone in *SciSearch*, it is assumed that you want to search the title-word field. However, you stipulate the author field by using the command "S AU=Smith" as shown earlier. For each field, there is a two-letter code you must specify. If you want to search the address field, you specify "S CS," which means search "corporate source."

An interesting and important field is "JN," which stands for journal name. Remember when we wanted to find articles published in the *Astrophysical Journal*? As a matter of fact, you can search on this field label to display the contents of a particular journal from one issue to an entire year. Once you specify these fields, you create new sets, any of which can subsequently be combined into new sets for greater specificity.

Perhaps the most-used field label in *SciSearch* is "CR," which refers to the cited-reference field. I referred to this earlier. You can also use the "CR" field to do your own co-citation searching. Try a search in which you look for pa-

pers that have cited two of your favorite papers and/or authors. H.D. White, Drexel University, has reported extensively on the use of co-cited author searching using this simple method.⁴ The simplest form of the command for searching a co-cited pair is "S CR=Smith J,? AND CR=Jones R,?" You use the truncation symbol "?" to avoid specifying particular documents. Given enough co-cited pairs, you can create your own citation network or map. Incidentally, had you specified "CR=Jones R?," you would have found papers by R.A. Jones, R.B. Jones, and so on.

The crucial point to remember about "CR" searching is that you will often retrieve papers that use title words quite different than you expected. It is language-free searching. This is one of the great virtues of citation searching. Once you have found a few interesting papers this way, you can always expand your results. You can choose one of the words or phrases in the most relevant paper that you have found. Then extend the search by using one or more title words.

For those who have the time and energy, you can take a quick course on using *SciSearch* in a variety of ways. Just as we support our *Sci-Mate* software program, ISI also provides seminars on the use of *SciSearch*. You can attend a *SciSearch* demo at the Federation of American Societies for Experimental Biology meeting or any one of the dozen other conventions at which ISI will exhibit this year. You can obtain free, hands-on experience at these demos, so come prepared with your questions.

The *SciSearch* file is used by thousands of searchers each year. It is not out of the question for you to conduct a search for as little as a few dollars, but we would expect that the average search, properly planned, should take 5 or 10 minutes for which you will pay the applicable connect charges. There are also communication charges, which are now down to \$6.00 per hour in the US.

Many people want to use the online system to download records for their own personal files. It is much more effi-

cient to do this by pre-formulated searches so that you don't pay unnecessary connect fees. There are a variety of ways to do this, including the so-called SDI or save commands on DIALOG. In *Sci-Mate*, this is an important option because you can prepare your search offline. This applies whether you intend to use the native or menu-driven systems.

Perhaps the most important benefit in trying *SciSearch* is that you will later understand what the system can do for you when you use an intermediary. Your local librarians will be delighted to search *SciSearch* for you. Since they do this more often, there are a lot of tricks they learn, including the best and cheapest times of day to search. But when you are working at home and you want to finish that manuscript before the weekend is out, it will pay for you to learn that *SciSearch* is at your disposal to fill in that missing reference or idea.

Also available for your online information needs is *ISI ONLINE*, a quarterly newsletter about ISI's online products and services. The newsletter was launched last summer, and it is free to our online clients or anyone who wants to be added to the mailing list. *ISI ONLINE* provides readers with news about ISI's databases, tips on searching, and answers to questions about our products. In the future, we hope to discuss the evolution of *SciSearch* into a real-time CC online system.

All of the *SciSearch* data can also be accessed through DIMDI (Deutsches Institut für Medizinische Dokumentation und Information), a vendor of ISI data-

bases located in the Federal Republic of Germany. The data are divided into three separate files. File SO (*SciSearch*) is comprised of data from 1974 through 1978. File IB (*ISI/BIOMED*[®]) includes biomedical data from 1979 to the present, while file IM (*ISI/MultiSci*) includes all of the technical and scientific data from 1979 to the present that is not of a biomedical nature. Both *ISI/BIO-MED* and *ISI/MultiSci* are updated monthly.

Information on *SciSearch* can be obtained from ISI's Customer Services Department. Our Customer Services Representative will answer questions concerning ISI files and can provide you with a free *SciSearch* User's Guide and Source Journal List. The toll-free telephone number is: 1-800-523-1850, ext. 1371. Pennsylvania residents phone (215) 386-0100, ext. 1371. You can also write to Laura Weissenberg, Manager, Customer Services, ISI, 3501 Market Street, Philadelphia, Pennsylvania 19104.

For information on *Sci-Mate* software for personal text management and universal online searching, call the *Sci-Mate* Hotline weekdays from 8:00 a.m. to 6:00 p.m., EST. Phone toll-free: 1-800-523-4092. Pennsylvania residents phone (215) 386-0100, ext. 1418. Or write: Dan Giancaterino or Bob Rodgers, *Sci-Mate* Customer Services Representative, at the above address.

* * * * *

My thanks to Bernadette Freedman, Terri Freedman, John Harke, Barbara Schreiber-Coia, Theresa Rosen, and George Zajdel for their help in preparing this essay.

©1985 ISI

REFERENCES

1. Garfield E. Introducing *Sci-Mate*--a menu-driven microcomputer software package for online and offline information retrieval. Part 2. The *Sci-Mate Universal Online Searcher. Essays of an information scientist.* Philadelphia: ISI Press, 1984. Vol. 6. p. 96-106. (Reprinted from: *Current Contents* (14):5-15, 4 April 1983.)
2. What does *Current Contents Online* mean to you? *Current Contents* (10):3-6, 5 March 1984.
3. How to use *Science Citation Index (SCI). Essays of an information scientist.* Philadelphia: ISI Press, 1984. Vol. 6. p. 53-62. (Reprinted from: *Current Contents* (9):5-14, 28 February 1983.)
4. White H D. Cited author retrieval online: an experiment with the social indicators literature. *J. Amer. Soc. Inform. Sci.* 32:16-21, 1981.