

Current Comments®

Introducing *Sci-Mate*— A Menu-Driven Microcomputer Software Package for Online and Offline Information Retrieval. Part 1. The *Sci-Mate* Personal Data Manager

Number 12

March 21, 1983

In 1982, about three million microcomputers were delivered in the US alone.¹ The average number of users easily exceeds double that number. Among them are thousands of scientists and scholars who are *Current Contents*® (*CC*®) readers. By the end of this decade, I estimate that over 90 percent of our readers will rely on microcomputers or intelligent terminals to store and retrieve information.

The first time I wrote about ISI®'s plan to produce software for microcomputers² we received several hundred letters—more than on any other subject I've ever discussed in *CC*. The microcomputer system for managing reprint files I described was designed by A.E. Cawkell, ISI's former director of research. Cawkell continues as a consultant to ISI. He also publishes a newsletter, *Information Technology and People* (CITECH Ltd., P.O. Box 5, Ickenham, Uxbridge, Middlesex UB10 8AF, England).

Let's recapitulate the reprint filing problem and the solution first proposed by *PRIMATE*, that is, Personal Retrieval of Information by Microcomputer And Terminal Ensemble. Scientists accumulate reprints through various means—writing to colleagues, using reprint cards such as *Request-A-Print*®, photocopying journals in the library, or by ordering tear sheets from ISI. When the reprints arrive, you peruse most of them quickly, while a selected group is read carefully or discussed in your lab.

In either case, the next problem is how to file them.

While there are many well-organized reprint collections, most scientists lose their initial enthusiasm for maintaining them. They start out with the best intentions, but soon abandon any attempt to be systematic. *PRIMATE* was designed to help scientists keep up with their reprint files by eliminating the indexing and filing problems in the ordinary sense of those terms.

The enthusiastic response to the idea of a personal reprint management system encouraged us to continue the development of *PRIMATE*. *PRIMATE I* and *PRIMATE II* represented the early stages of this system.³ We have since decided to name the system for the main audience it will serve—the scientific community. So we changed the name to *Sci-Mate*™ to eliminate the implication that it serves the exclusive needs of those in primate research! It will also serve to highlight the fact that the new *Sci-Mate* system is significantly advanced over its predecessors and can be used in conjunction with ISI's scientific and scholarly data bases.

The original system was designed almost exclusively to deal with the problem of filing and retrieving personal reprint collections in an *offline* situation. This distinction from *online* means that it does not have to be operated through a centralized time-shared computer system. Each microcomputer has its own central processing unit (CPU). We are

also developing a version of *Sci-Mate* which can be used in minicomputer systems that serve many terminals in a single institution. In that situation each terminal is used in an online mode in relation to its CPU.

In its four-year evolution to *Sci-Mate*, the ISI system has become much more than a reprint filing system. First, it is a personal offline *data base management system* with unique features and capabilities not available in a single package elsewhere. Second, it is a system for accessing and using numerous commercial *online* time-shared data bases that does not require knowledge of their special command languages. The *offline* and *online* components are available separately or in combination. In a future essay, I'll discuss *Sci-Mate's* online capabilities in detail. In this portion of the essay, I'll focus on its *offline* functions. Figure 1 summarizes these features and their benefits.

A microcomputer data base management system is a software package, that is, a series of integrated computer programs which enable you to create and manipulate many types of records. Most of the commercially available software packages were designed for use in business settings. They are commonly used to store lists of customers, subscription records, inventories, etc. They are designed mainly to handle numeric information, and are limited in their flexibility for handling textual information.

The *Sci-Mate Personal Data Manager* is a data base management system designed specifically to accommodate *textual* material—especially bibliographic material. You can, of course, use it to store full bibliographic citations. However, it can also handle notes and abstracts, or comments that you wish to add to help you retrieve them or to evaluate or sort them once retrieved.

Sci-Mate's forerunners were designed mainly with reprint collection management in mind. *Sci-Mate* itself has retained that characteristic. In fact, the ability to manage reprints may be the function that will be initially the most useful to most *CC* readers. However, its usefulness in other record keeping situations will become apparent here.

Let me make clear that in the following discussion the machine readable bibliographic records are only *representations* of the original reprints, letters, or other documents. The actual reprints are filed separately. The same would be true for correspondence, lab notes, and medical case reports, unless the full text is so short that it constitutes an entire document.

To use *Sci-Mate* as a reprint file manager, you have several options. You could assign each new reprint a serial number. For each reprint, you'd enter this accession number and a complete bibliographic description or summary. You could do this one reprint at a time or you could have an assistant add the in-

Features	Benefits
Random access storage.	Fast access to any stored record.
Menu-driven command system.	No need to learn a special language.
Tutorial subsystem.	<i>Sci-Mate</i> lists options for you.
Variable length records.	You can store textual data of any length.
No pre-indexing.	You can search any phrase, word, or character alone or in combination.
Flexible record layout.	You can design your own record format.
Report generator.	Columnar reports to your specifications.
Immediate record/file updates.	Records searchable as soon as entered. Eliminates need to create or update lists.

Figure 1: Features and benefits of the *Sci-Mate™ Personal Data Manager*.

WELCOME TO THE ISI SCI-MATE SYSTEM

Do you want to:

1. Enter the UNIVERSAL ONLINE SEARCHER
2. Enter the PERSONAL DATA MANAGER
3. Enter or change date
4. Delete the WORK file
5. Receive instructions on the use of SCI-MATE
6. Leave the SCI-MATE system

Select a number or enter ? for help:

Figure 2: This is what you see upon entering *Sci-Mate*™.

formation for a series or batch of papers, perhaps after reading *CC* or as reprints arrive each week.

However, with *Sci-Mate* you don't have to assign a number "manually." You can use the number that the program *automatically* assigns to each entry. In Part 2 of this essay I'll describe how much of this keyboard work can be eliminated if you choose to input records by using the online option to "capture" records.

As the information for each paper is keyboarded, *Sci-Mate* automatically adds the new records to the computer's memory. The reprints themselves can be filed in numerical order in filing cabinets or loose-leaf binders. If you prefer, instead of filing *numerically*, you can file the papers *alphabetically*—either by author or by subject categories. However, the serial number system maintains your file essentially in chronological order. This is handy when you want to browse through the reprints you acquired over a particular time period. However, if you prefer to file the actual reprints by lecture topics, you can include the name or code for the file folder as part of the document description. Incidentally, as you are keyboarding all this information, or capturing it from another file, it is displayed on the screen for visual verification.

Having stored several hundred document descriptions, how do you retrieve them? *Sci-Mate* is designed to be "user-friendly," which is computerese for "easy to use." The validity of such a claim depends upon the amount of artificial intelligence built into the system. We believe *Sci-Mate* is justifiably called user-friendly because it is a "menu-driven" or multiple-choice system. At each step the computer asks you a question and presents a short list of options. For each choice you press one key to execute the option or command chosen.

Figure 2 shows the first *Sci-Mate* question after you insert the magnetic floppy disk (diskette) on which the program is stored. You are presented with six choices. Figure 3 shows what happens when you press "2", the request to search your personal data base. Again you have six options, one of which simply takes you back to the beginning. The first option is the one you can expect to use a lot for searching. The second option is used when you want to enter new reprints into the collection. Figure 4 shows what you see when you press "1"—"Search the USER file."

Figure 5 shows what you see when you press "2". You enter a search term, which is used to search the entire file. This means you do not have to specify whether you want to search titles, author

OFFLINE

You may:

1. Search the USER file
2. Enter a new record
3. Display/copy WORK file records
4. Create and update templates
5. Generate document request status report
6. Return to the SCI-MATE menu

Select a number:

Figure 3: This is what you see after selecting option 2 in Figure 2.

names, journals, etc. Everything in the records you have entered is automatically searched. If you had added your own annotations or comments to your files, those comments would also be searched. This "string searching" capability is a simplifying element in *Sci-Mate* that is presently available only on large text searching computers.

One feature I should stress is that every entry or change you make is *immediately* searchable. No matter how many times a day you add or modify records, you can search them as soon as you create or change them.

Sci-Mate also allows you to refine or extend your search through use of Boolean logic. This means you can combine terms. For example, you could search for papers on fission. But in a less restricted search both fusion *and* fission could be stipulated. This in turn could be narrowed to a particular journal or laboratory.

Perhaps the most common bibliographic search involves the location of a paper by a known author on a particular topic. All you remember is one key word of the title. So you simply use the *and* command (or "operator") to find the Jones paper on mitochondria. Or you remember Smith wrote a paper in *Scientific American* a few years ago, but you simply can't recall the title. Just key in the author's name and the journal name.

Sci-Mate also includes a way to broaden your search. It gives you the option of "truncating" your term. Suppose you wanted to review the papers in your collection which contained the term *bronchitis*. You can broaden the search by inserting a "wild card," represented in *Sci-Mate* by this symbol: #. So if you entered "bronch#", you would get hits using the terms *bronchial*, *bronchial-secretions*, *bronchoconstriction*, *bronchosopic*, and so on. You can also put the "wild card" at the front of a term to find prefixes to the key word. You can even insert it in the middle of a term.

Almost all data base management systems claim to "search" files. Even a word processor can do limited searching. But it is absurd to discuss the search capability without specifying the time required to do the average search. We believe that *Sci-Mate* is extremely fast for large, randomly organized files. But this statement is best demonstrated on your own or comparable files. In any case, the speed is affected by many factors.

There is a significant difference depending on whether you are storing data on a *hard* rather than a *floppy* disk. The *average* search rate for the much larger hard disk is about 100 records a second, given an average of 300 characters per record. For *floppies*, the speed is about one-third of this.

If you are used to working with a mainframe or minicomputer system, it is easy to make invidious comparisons be-

SEARCH

Would you like to:

1. Search by accession number
2. Search text: find total hits before display
3. Search text: display hits as retrieved
4. Return to the OFFLINE menu

Select a number:

Figure 4: This is what you see after selecting option 1 in Figure 3.

TEXT SEARCH

Enter a search statement

computer

Figure 5: What you see when you ask *Sci-Mate*™ to search by text. The user has entered "computer." The search term will be used to "string search" the entire record.

tween *Sci-Mate* and software designed to operate on such relatively large time-shared computers. Most large-scale systems use an inverted file system. *Sci-Mate* has eliminated the need to create inverted files and the updating problems associated with them. But this means it will be used optimally for files that are typical of a personal reprint collection. Do not expect to store a year of *CC* or *Science Citation Index*® on *Sci-Mate*. In test files of 5,000 records containing an average of 300 characters per record, a typical search question involved less than one minute of search time. Incidentally, the search time goes down as your request becomes more specific.

Sci-Mate is geared up for a large number of bibliographic or other texts such as correspondence or even patient records or lab notes. The number of records you can store depends on how lengthy those records are, and what your disk capacity is. The maximum length of a separate record is 1,900 characters. Even at that size you can store 2,200 records on a 5-megabyte hard disk device. There are two such devices on a 10-megabyte disk.

However, most people use records of less than 1,000 characters. In this case, you can store up to 6,000 records if they are only 600 characters long. The typical journal article title usually consists of fewer than 100 characters. To exceed the storage limit of the hard disks you can use a series of floppy disks or an additional hard disk.

In principle there is nothing that prevents you from storing the complete text

of a short manuscript, e.g., a letter to the editor of *Nature* or *Science*, as a single record. Furthermore, you could in principle store sections of a large manuscript as a series of "linked" separate records. It should be obvious, however, that when you use *Sci-Mate* for such lengthy records you sacrifice your ability to store many smaller records. However, you could also use a hybrid system combining storage on hard disks for the basic reprint file and use floppy disks for one or more complete article texts.

The *Sci-Mate Personal Data Manager* not only allows you to store the bibliographic data for your reprint requests, but also review their status periodically. Suppose, for example, you want a paper you saw in *CC* or identified while using the *Sci-Mate Universal Online Searcher*. You order a reprint before you add the reference to your permanent data base. Then you enter this information in your data base, which contains all your recent requests and other items. You turn on a "flag" or code to be reminded that you've sent out a reprint request card or wrote the author. When the reprint arrives, the flag can be turned off. An example of a flag screen is shown in Figure 6. You can use it to be reminded about up to six different things: *OATS*® (Original Article Text Service), interlibrary loan, Lockheed's *DIALORDER*, *Request-A-Print*, telephone calls, etc.

Subsequently you can review your records to obtain a printout of all

FLAG	STATUS
1. OATS	ON
2. ILL	ON
3. REP. REQ.	OFF
4. DIALORD.	OFF
5. SCHED.	OFF
6. RESPOND	ON

Would you like to change a flag status (Y/N) ?

Figure 6: An example of a *Sci-Mate*™ "flag," to keep you up to date on the status of document delivery.

Field No.	Name	Field No.	Name
1	Acc. No.	11	Compound
2	Title	12	LD50
3	Language	13	Lab
4	Doc. Tp.	14	
5	Author	15	
6	Corp. So.	16	
7	Journal	17	
8	Year	18	
9	OATS #	19	
10	Org Test	20	

Figure 7: An example of how you might arrange a *Sci-Mate*™ "template" to help you sort and search.

"flagged" items. This status report will remind you that you've not received, for example, the reprints or interlibrary loan items you requested. This type of follow-up system also can be used as your "tickler" file to check on correspondence that is overdue. Flag 6 of Figure 6 shows how you can indicate that a letter needs to be answered. I'm using *Sci-Mate* for this purpose and, more importantly, as a way of checking on the status of small and large writing or research projects.

There is a feature of *Sci-Mate* that causes quite a stir wherever it is demonstrated. It is the menu-driven system for formatting, or "templating," your files. A template is simply the sequence of informational elements for each record. In computerese, these elements are called "fields," and should not be confused with fields of science or research. Using templates will help you maintain consistency in indexing and storing records. They also label the records you display, and are used to sort the results of searches.

A sample template is shown in Figure 7. Like all *Sci-Mate* templates, it consists of up to 20 fields. In this example, fields one through nine have been taken up with the accession number, title, original language, document type, author, corporate (or organization) source, journal, year, and OATS order number. Fields ten to 13 have been taken up by organism tested, test compound, LD50 (lethal dose, 50 percent) test results, and laboratory. (The LD50 test measures the dose of a substance that kills half the test animals.) Although the field names are limited to eight characters, there is nothing to prevent you from putting the full titles, authors' names, and so on, into the appropriate fields. Fields 14 to 20, in this example, have been left blank. They can be filled with other items at a later time. You might, for instance, want to add a field for "notes," and enter your opinion of the data in the record. You can arrange fields in any order you wish, and you can use as many as 50 different template formats.

These templates enable you to sort the records obtained from a search by selected fields. If you choose to sort the "author" field (field 5 in Figure 7), the result is a display of records, shown in alphabetical order by author. You might, however, want to sort by compound. So you would order the system to sort by field 11. The results would then be a display of records in alphabetical order by compound. (The sorting program automatically alphabetizes.) You can also sort by language, document type, journal, organism tested, laboratory, notes, or any other field.

ACC. NO.	ORGANISM	COMPOUND	LD50
105	Rats	Caffeine	192 mg/kg
283	Rabbits	Fenoflurazole	28 mg/kg
1205	Rabbits	Phenylbutazone	146 mg/kg
1299	Mice	Carbofuran	2 mg/kg

Figure 8: An example of a report generated by *Sci-Mate*™.

With *Sci-Mate* you can view hits on the screen, or use a feature called the "report generator" to print lists of specified information. This feature enables you to print lists of fields of a record in a horizontal rather than a vertical format. Suppose you wanted to report on the organisms you've tested, plus the compounds you used and the LD50 test results. *Sci-Mate* can give you simple step-by-step instructions to produce a report like the one in Figure 8. The first column shows the record accession number. The second lists the organism; the third, the test compound; the fourth, the result. This figure is arranged by accession number order. But you could generate this report in alphabetical order by any field, using the sort facility. You can also generate lists using other fields, and get reports on, say, the authors in your file, the status of projects, journals, and so on.

The microcomputer software market is an enormous and growing industry. One indicator of this is the number of magazines catering to the microcomputer buff, such as *Byte*, *Personal Computing*, and *InfoWorld*. One of the most important functions performed by these publications is evaluation of software packages. One of the prevalent criticisms is that software is not adequately tested on different microcomputers. Since software packages are sold at bargain prices, suppliers will not back up their software with "service." But all of *Sci-Mate's* functions have been tested by several outside groups. And ISI will stand behind its software claims.

However, it is not enough to say that our software will be backed up by ISI's reputation as a user-sensitive company. While *Sci-Mate* has been tested in numerous commercial and academic institutions, there is no way that one can anticipate all the myriad problems that can arise when a system is used by thousands of individuals. We have taken concrete actions to ensure that the software works to your satisfaction.

Nothing can replace adequate documentation and service. Each *Sci-Mate* customer receives a detailed manual. ISI has also set up a special customer hot line so that unanticipated problems can be dealt with quickly. Don't laugh—some scientists forget to plug the cord into the electrical outlet. Beyond the obvious, those problems that cannot be dealt with over the phone will be referred immediately to our research and development department where they will be dealt with by the team that has helped develop *Sci-Mate*. We cannot be responsible for any equipment failures but we can attempt to tell you if you are having a hardware rather than a software problem. And quite frequently it may be a matter of improper execution of your own specifications for format, etc. Using a new system to organize your files is a significant learning experience.

The ISI staff is excited about the potential of *Sci-Mate*. We think it's a logical step in our mission to improve access to scientific information. When we started this project four years ago, we planned to build our own special purpose microcomputer. Now they are so cheap and diversified that it is neither necessary nor practical for us to manufacture or assemble hardware. But *Sci-Mate* may well be the first reason you may acquire a microcomputer. Whether you acquire a new microcomputer or use the one you have, do not expect *Sci-Mate* to make it the ultimate intelligent system. As with any decent software development program, we are making provisions for continuing improvements and updates to the software. This will be particularly relevant for the online component to be described in the second part of this essay. Updates in our software will include such features as automatic re-formatting of citations depending upon the journal in which you publish.

Sci-Mate software is currently available for the IBM Personal Computer, Vector 3 or 4, the Apple II, the TRS-80

Model II, and any CP/M-80 (control program for microprocessors) system with a standard eight-inch disk. This includes dozens of brands like Altos, Industrial Microsystems 8000, and Hewlett-Packard 125 (8-In.). For the IBM, a CP/M-86 operating system is required. For the Apple II, you need cards (circuit boards) for a Z80 microprocessor, 16K RAM (random access memory), and an 80-column screen. Of course, a printer is also needed. In addition, to use the *Sci-Mate Universal Online Searcher*, you need a modem, a device you connect to your telephone so you can dial up large online data bases. We also recommend 500 kilobytes of disk storage for an effective system. *Sci-Mate* requires at least two disk devices. It can also operate with two floppy disks as a minimum. However, it's preferable that one of your disks be a hard one, to store your records. Hard disks store much more information than floppies.

It is inevitable that readers who do not own a microcomputer will ask us to recommend the best. That is not a simple question because each microcomputer will perform better under certain conditions. However, within the next few months we will be evaluating the results of *Sci-Mate* use in the field and will be glad to pass along to you our experiences and recommendations.

It will be obvious to experienced microcomputer users that the *Sci-Mate*

system is not a toy. It requires 64K RAM. The machines required are in the price range of \$4,000 to \$7,500 in the US. These prices include all the necessary peripherals.

The offline *Sci-Mate Personal Data Manager* costs \$540. The online *Sci-Mate Universal Online Searcher* option is \$440. Together, both cost \$880—a savings of \$100. For more information on *Sci-Mate*, contact Joel Sartorius, ISI, University City Science Center, 3501 Market Street, Philadelphia, Pennsylvania 19104, USA. Telephone: (800)523-4092. Telex: 84-5305. All others should write to ISI's US headquarters.

In the near future, I'll describe the *Sci-Mate Universal Online Searcher*, which can be used for online searching, and how the two *Sci-Mate* systems work in tandem. The online system not only dials up the host computer automatically but also eliminates the need to learn command languages. As they say in the US, "Look ma—no hands."

* * * * *

My thanks to Thomas Marcinko for his help in the preparation of this essay.

© 1983 ISI

REFERENCES

1. Friedrich O. The computer moves in. *Time* 121(1):14-24, 1983.
2. Garfield E. Introducing PRIMATE—Personal Retrieval of Information by Microcomputer And Terminal Ensemble. *Essays of an information scientist*. Philadelphia: ISI Press, 1980. Vol. 3. p. 551-5. (Reprinted from: *Current Contents* (29):5-9, 17 July 1978.)
3. On the 25th anniversary of *Current Contents/Life Sciences* we look forward to the electronic online microcomputing era. *Current Contents* (4):5-11, 25 January 1982.