

Solving Information Storage and Retrieval Problems with Miracode*

EARL S. DANIEL

Tennessee Valley Authority, Muscle Shoals, Alabama

Received May 13, 1966

The Miracode (Microfilm Information Retrieval Access Code) system, in which the material is indexed in depth at the time of input, is described and its advantages are discussed. The system not only met our imposed requirements for speed, thoroughness, and minimum storage space but also permitted recording of entire articles rather than just bibliographic references.

Several factors brought about our move toward automation. We were continually faced with the age-old problem of available floor space. This problem may not be most important, but it is most noticeable. During the past year or two this had become critical, especially filing cabinet and back-file storage space. We could eliminate the problem in two ways: Move into larger quarters or install some kind of automated retrieval system that would confine large amounts of printed material to a small area.

Another common problem is that a considerable amount of time is required for processing, filing, and retrieving information. Even though automation does not eliminate the processing problem, it does considerably decrease the number of man-hours required to file and retrieve material.

Although these seemed to be the most common problems encountered, and quite real to us at TVA, they were not the problems causing us the most concern as library administrators, and causing us ultimately to seek retrieval through automation.

We were unable to find material because the requester could not fully or properly identify the item needed. It was not practical to retrieve manually every possible reference in the area of his request in order that he might identify the one he wanted. We needed a system that would show the patron, in a matter of seconds, all the possibilities from which he might make his selection. For example, we repeatedly have the following type of request: Find a journal article having a flow sheet on nitric acid plant operation. Patron does not know the date, title, or author, but thinks the article appeared in *Chemical Engineering* or *Chemical Engineering Progress* sometime within the past 10 years. In order to answer this query it was necessary to search indexes for articles on nitric acid plant operation, then check the article itself for a flow sheet. To solve this problem we needed equipment that would show the patron in a matter of seconds or minutes all articles with flow sheets on nitric acid plant operation which had appeared in both *Chemical Engineering* and *Chemical Engineering Progress* during the past 10 years.

Another problem was caused by the fact that our literature collection is catalogued and filed in several different

places, according, in some instances, to the source of the materials and, in others, to the type of material. These divisions are necessary for acquisition and processing control; however, we found that occasionally one or more places would remain unsearched in the quest for information. For example, we might neglect to search our U. S. Bureau of Mines file for information on the ammonia industry in North Dakota, either because the searcher might forget this collection or because it seemed to him unlikely that there would be information on the subject in that file.

We found occasionally that pertinent material that did not fit conveniently into any of the divisions of our literature collection became lost or forgotten. No matter where and how it is stored, a fact on file is useful only if it can be found.

In our search for suitable automated retrieval equipment we covered the field as completely as possible. We found that at least 32 companies make automated retrieval equipment; but fortunately, we were able to determine the applicability of 90% of this equipment to our situation from the literature available from the manufacturers and the trade journals.

We first considered, though not very seriously, the simple punched card and "ice pick" system already in use by some TVA offices. This was not suitable for several reasons. Maintaining this type of system seemed from observation even more difficult than our old methods and the amount of material for input would have been greatly limited. Cards would have to be manually refiled after each use, and they soon become mutilated and difficult to handle. As compared to electronic devices, the search rate was too slow.

Some thought was given to using a combination of microfilm and microfiche. Although this would have saved considerable space, there was the added disadvantage of having to make printouts of all the file material when needed. Neither would this method provide greater speed in retrieval.

More serious consideration was given the aperture card system. This system would have solved most of our problems such as thoroughness of search, speed in search, identifying difficult requests, saving space, and preventing duplication of effort; but it had two serious handicaps for our particular use. The amount of material which

*Presented before the Division of Chemical Literature, Symposium on Problems of Small Information Groups, 151st National Meeting of the American Chemical Society, Pittsburgh, Pa., March 25, 1966.

can be put into the aperture card "window" is limited (we did not consider multiple cards for one item satisfactory), and the cost of the equipment necessary to use this system was prohibitive. Equipment of this type does exist in our computer section, but it is in continuous use and sufficient time could not be scheduled for our needs.

We were interested in the more sophisticated equipment using magnetic tape. Several installations using this system were visited, including the MEDLARS complex at the National Library of Medicine. However, we found that even a small installation of the magnetic tape equipment was prohibitively costly. Neither is it practical to put entire articles and booklets on magnetic tape.

We heard about Miracode at a seminar in Florida. After their representatives demonstrated the equipment, using material from our literature collection, we concluded that Miracode could solve our problems. The cost was within our budget. The system offered rapid retrieval; it would save a great amount of storage space; it had printout capability; and it offered an opportunity to scan the collection without printing it out. Material ranging from bibliographic references and abstracts to entire books can very easily be put into the system for retrieval.

The data must be selected and coded. Herein lies the success or failure of any automated retrieval system. For this portion of the operation we use professional personnel trained in the areas of interest, primarily chemical and agronomic research as it relates to the fertilizer industry.

Miracode operates in two parts, input and retrieval. Although the input portion is necessary, its purchase is optional, as this work can be handled on a contract basis with the manufacturer, Recordak Company, a subsidiary

of Eastman Kodak. The input consists of a specialized 16-mm. microfilm camera and electronic coding device, which film and code a document in one operation easily handled by clerical personnel.

The most interesting part of the Miracode equipment is the retrieval station or keyboard console. This retrieval station allows one to search a maximum of four million different subject areas or combinations of areas. The number of possible combinations depends on the number of keyboards in the console. This means that a searcher may retrieve all entries in a given subject matter area, no matter what the source, so long as they are in the Miracode system.

The Miracode equipment has even greater possibilities than we at first realized. One of our latest innovations has been the coding, filing, and retrieval of black and white and color photographs, color slides, and other visual material. The simplicity of input has made it possible for us to keep on file material of undetermined value that might otherwise have been discarded. Miracode searches at the rate of about 8000 pages per minute, or about 3000 code characters per second. Simplicity of operation eliminates the need for highly trained personnel. In fact, we hope that in the near future local patrons will carry out their own searches with minimum library assistance. Miracode also allows us to search our entire collection without regard to the divisions or categories in which it is kept. Miracode has also made it possible to acquire collections from other organizations and use them as our own. This would have been much more difficult to accomplish with hard copy or even with the more sophisticated hardware.

A Decade's Experience With A Primitive Machine Retrieval System*

FRANK S. WAGNER, Jr.
Celanese Chemical Company, Corpus Christi, Texas

Received May 31, 1966

A machine retrieval system is described which uses a superimposed coding adapted to mechanical sorting. The system has been in operation for the past 10 years. Several problems concerned with formula-functionality indexing, number of terms, etc. are examined. Most of the shortcomings are well-recognized deficiencies of all coordinate indexes.

Mechanized retrieval systems usually imply large numbers of documents or relatively complicated coding systems. Our system has neither. It is primitive in the sense that the minimal necessities are supplied by as simple means as possible. Machine methods are used to speed

up and cut the cost of input; machine methods are used to recover the references needed from the file. In our decade of experience, this comparatively unsophisticated system has encountered surprisingly few difficulties. This may be in part attributed to its limited subject field, the small size of the file, or the primitive adaptability of the system; but in any case, it is perhaps instructive to examine some of the troubles we discovered and how they were managed or overcome.

* Presented before the Division of Chemical Literature, Symposium on Problems of Chemical Information Groups, 151st National Meeting of the American Chemical Society, Pittsburgh, Pa., March 20, 1966.