Milton J. Ruterbusch

Milton J. Ruterbusch heads up the Electronic Products Department of the Recordak Corporation, a Subsidiary of Eastman Kodak Company, in New York City. He studied Engineering at Central Michigan State College and received his background training in electronic systems at Point Loma, San Diego, where he studied Radar Counter-Measures while in the U. S. Navy.

Mr. Ruterbusch has had 11 years' experience in the microfilm industry, having spent three years in sales, four in sales management, and three years in systems and product planning. He became head of the Electronic Products Department of Recordak in April 1962.

He has conducted several seminars on Information Storage and Retrieval and recently (as Guest Speaker for the American Management Association) spoke on the subject of "Microfilm as a Management Tool."

Mr. Ruterbusch is the author of several patens for Equipment and Information Retrieval Concepts, and is very active in community and civic affairs in Ridgewood, New Jersey.

MIRACODE - A NEW BREAK-THROUGH IN AUTOMATED INFORMATION RETRIEVAL

M. M. Ruterbusch*

The new Recordak MIRACODE System permits an operator to locate an exact document image from a file of over 900,000 pages in less than 15 seconds. This is done by means of searching Binary Code, with associated document images, which has been recorded on the microfilm. The name MIRACODE was derived from the words Microfilm Information Retrieval Access CODE.

In recent years, "Information Retrieval" has taken on added significance somewhat similar to the term "Automation." These descriptive expressions were not generated overnight. They were both an outgrowth of long years of technical development to meet the increasing needs of industry. As a matter of fact, the first microfilm business system was installed May 1, 1928, by the Recordak Corporation. This first use of microfilm for Information Retrieval was a 16mm roll film system. By present standards, the first microfilmers and microfilm viewers were basic and slow. But they did accomplish the task for which they were intended.

One by one, new applications for microfilm systems were established. With each new application came the need for faster recording and even more rapid retrieval. It was this increasing demand for rapid access to Information and the need for a better method of location Information on microfilm that prompted the development of the Recordak MIRACODE System.

In some ways, the MIRACODE System is no different than any other system designed specifically for the storage and retrieval of documentary information. The fundamental requirement is that of indexing. Here, the user has a choice, and the involvement depends on the desired end result. If the indexing is carefully analyzed and programmed, the search is well guided and rapid. If the indexing is shallow and uncoordinated to speed up the input, the search can be difficult and time consuming. The new Recordak MIRACODE System was designed to give the desired flexibility for both input and search. A review of the system for Indexing through retrieval will

*Recordak Corporation of New York, New York

demonstrate the similarities and differences of this new Information Retrieval System to other IR systems.

Every Information Retrieval system starts with indexing. Even a standard correspondence file may be simply indexed by file drawer and then by folders within a drawer.

There are two standard methods of indexing documents for the Recordak MIRACODE System. The first method is simply to index by the number assigned to the report. This may be a project number, Case History number, policy number, etc., which will be referred to as the "fixed" document number.

The second method of indexing is by subject "descriptors," which, in other areas, may be called key words or subject identification terms. This is a convenient and highly desirable technique for indexing reports by their contents. In the MIRACODE System, the documents to be recorded may be either of these two methods or a combination of both.

When a subject index is desired, the user first determines which descriptors will best identify the various reports to be recorded. These descriptors are then organized and an indexing thesaurus is established. Each descriptor is assigned a three-digit number.

The indexer, after reviewing each report, determines which descriptors will be used and a report precoded index sheet is produced. The documents and associated precoded index sheets are forwarded to the microfilm operator for recording in the MIRACODE System.

Input to the Recordak MIRACODE System can be accomplished in several ways, depending upon the application requirements.

1. Manual code selection with operator placement of document pages.
2. Manual code selection with automatic document feeding.
3. Automatic code recording from 80-column EAM cards with operator placement of document pages.
4. Automatic code recording with automatic document feeding.

Manual encoding means that the operator has only to set the input keyboard selector switches to precode the numbers indicated on the index sheet and depress the code record button. For automatic encoding, the index sheets are sent to a key punch operator and an EAM card is prepared for each document. These cards are then forwarded to the Input Microfilmer for automatic card reading and code recording.

In either case, a document's fixed number or descriptor number is converted to Binary Code on the Microfilm. A code pattern (see Figure 1) of 14 bits forms a three-character code column. The lower 12 bits are used as data bits in three sets of four each to represent a numeric character. By adding the values of each opaque

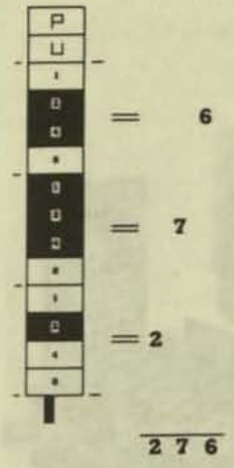


Figure 1

(or black) bit in each pattern of four Binary digits, the number 276 is generated.

NOTE: The Binary bits "U" and "P" represent the Utility condition and Parity Check respectively. These are used for equipment/search control and will be discussed later.

The MIRACODE System has the capacity for recording and searching up to 45 numeric characters in the fixed field (document number). Also, a choice of up to 1,000 descriptors can be used to identify any document. Needless to say, very few applications contain documents which require more than 6 or 7 descriptors, if such depth is necessary.

The multiple code columns in Figure 2 represent a typical code pattern. The first two

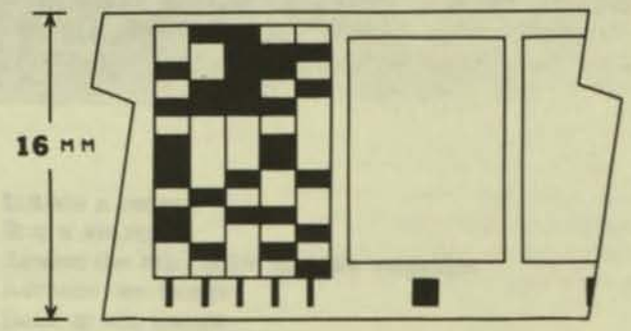


Figure 2

columns, recorded without a Utility bit, represent a 6-digit fixed document number. The remaining three code columns represent three descriptors and are considered to be in the "open field." Therefore, each column includes the recording of a Binary 1 in the Utility bit position. Each code column has its own "timing mark" at the bottom to signal the location of the code column. Also, each document image has a "document mark" exposed beneath it. This opaque square is used to automatically center the document image on the reader screen. Both the timing mark and document mark are recorded on the microfilm automatically during the code and document exposure cycles.

The new Recordak MIRACODE Microfilmer is a 16mm planetary unit capable of reduction ranges from 12:1 through 28:1. Documents up to 14 x 22 inches may be recorded in one exposure. The Microfilmer (Recordak Model MRK-1) is also equipped with a special Recordak Film Unit (Model MCK-1) which has an Integral Code Target assembly to permit constant and precise code exposure in a fixed size regardless of the document reduction ratio selected. The Input Control Keyboard (Model IDKM) is position on the Microfilmer Copy Table for code and document recording. Nine selector slide switches are provided to permit encoding nine numeric characters with each actuation of the code record button. The numbers encoded are also automatically displayed in the photographic area and recorded beneath the microfilm image for positive identification of the retrieved document images. This numeric representation of the Binary Code on microfilm is also contained on the paper printout for request verification. Recording progress indicators are included to assist the operators and assure controlled microfilming. Only half a second is required to record three numeric characters and it takes only one second to expose each document image.

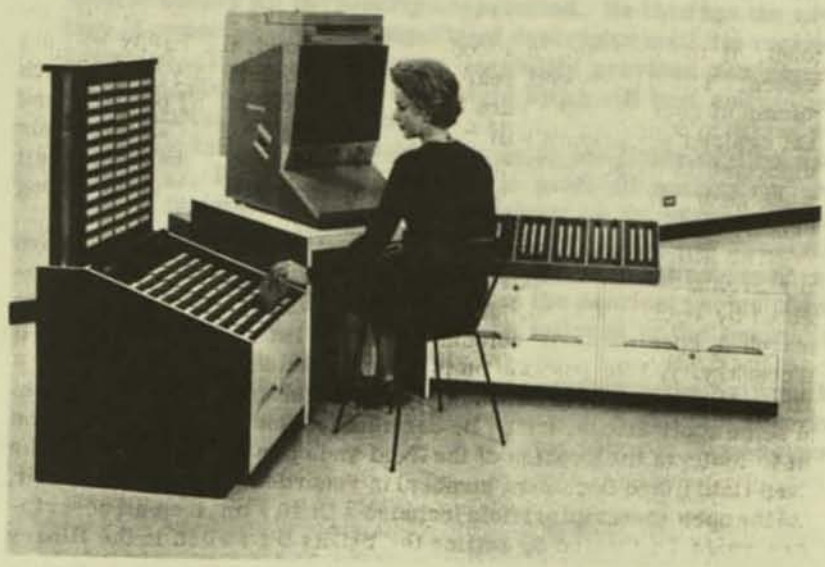
The Input Control Keyboard can be replaced by a Continuous Forms Feeder for automatic, high speed recording. By adding an Automatic EAM Card Reader - Control Converter, it is possible to record code and documents automatically.

The key element in this new break-through in automatic Information Retrieval is in the new Recordak MIRACODE Retrieval Station (see Figure 3). It features rapid document image access, paper printout, and extreme versatility of film file browsing.

Document storage density is high. If desired, a single retrieval station can hold as many as 490 film magazines. Assuming four code columns for each six-page document, 2000 document page images can be recorded on a single 100-foot roll of 16mm microfilm and mounted in a film magazine. This means that up to 980,000 document page images can be conveniently contained in only one retrieval station, and—within easy reach of a seated operator. Average access time for a desired report is only eight seconds from the instant the operator depresses the search button until the selected document image is positioned on the reader screen ready for review or paper printout.

A special Recordak Lodestar Reader-Printer (Model PEK) is the heart of the High-Speed Retrieval Station. This new Reader is equipped with a Keyboard Control Unit, MIRACODE Automated Retrieval Keyboards and a high-speed transistorized Logic Unit.

The MIRACODE Keyboard Control (Model KKC) has all the necessary controls to:



1. Initiate a search
2. Stop a search
3. Rewind the microfilm into the magazine
4. Advance one image
5. Back up one image
6. Position the image automatically for paper printout
7. Select a print mode
   - A. Manual
   - B. Print only the first page of the first "hit" automatically.
   - C. Print all pages of the first document located automatically.
   - D. Print only the first page of all the "hits" automatically.
   - E. Print all the pages of all the documents located automatically.
8. Indicate detection of a Parity error
   - A. Signal, and stop to determine cause of error.
   - B. Signal, but complete the desired search.
9. Search either negative or positive microfilm
10. Turn the system "on" or "off"

The system automatically generates and checks "odd" Parity in each code column. The 14th bit position is reserved for a "1" or "0" Parity bit. If the three characters in a code column and the Utility condition call for an even number of opaque code spots, the Parity bit will automatically be recorded to give an odd-number bit count. If the bit count in a code column is odd, the Parity will not record. The high-speed search reader automatically checks each column of code to make sure the system is operating properly and that desired documents will not be overlooked. This is extremely important in most applications as a missed report may very well cause an organization to duplicate a project or proceed in the wrong direction.

The MIRACODE Automated Retrieval Keyboards are designed to search a column of code. Three numeric characters may be selected by each keyboard. Also, each keyboard may be directed to search a specific code column from the 1st through the 15th, if necessary. A 16th position on the column selector switch directs a particular keyboard to search "any" code column for a code match. In some applications, it may be desirable to use the Utility condition bit to indicate the location of the fixed and open field. Assuming the fixed field (fixed document number) is recorded without a Utility bit, and the open (descriptor) field includes a Utility bit, then all descriptors would be located by setting the Utility bit switch in the Binary "1" position. In this manner, the keyboard can be directed to search only the fixed field or the open field, as desired. This particular use of the Utility bit provides the advantage of searching for descriptors in undetermined sequence. For example, a file category could be Defensive Missiles and would be so labled on one or more film magazines. One of the documents recorded might be indexed as follows:

Subject	Descriptor
250 Mile Range	108
Surface-to-Air	019
Solid Fuel	633
7J Booster	732
ACS Duration - 75%	957

The three-digit descriptors would be encoded on microfilm in the sequence shown, i.e., 108-019-633, etc. The searcher states the question in the order of importance to him. If the request were for (1) Solid Fuel, (2) Surface-to-Air, and (3) 7J Booster, the

document used in the example would be automatically located as it contains all three subjects.

A searcher might overdefine his request and not locate a document to answer all the descriptors selected. He then has the advantage of removing the least significant descriptor until his request is answered. This file browsing capability provides the searcher excellent selection of those documents which will best suit his needs in the shortest possible time.

A similar example could be made in the fields of Chemistry, Metallurgy, Physics, etc.—in fact, in most all application areas where it is desirable to locate a document by its subject content, a component by its characteristic, or an individual by his capabilities. A good Information Retrieval system supplies the best answers, to the right people, in usable form, and in the shortest period of time. Therefore, each application should be tailored to the desired end result.

I wish to take this opportunity to congratulate the National Microfilm Association upon an outstanding 1963 Convention Program and, to thank you for the privilege of being with you this afternoon.

Question and Answer Period:

Bob Anderson - 3M Company

Q. What happens if your thesaurus has more than 1000 words?
A. In this case it is probably advantageous if, in preplanning, it is possible to divide that category into two or more subcategories permitting the use of up to the 1000 descriptors for each.

David Varder - Bell & Howell

Q. How does the operator determine what magazine to select when a descriptor search is being made?
A. All of the documents, going into the system are placed in the film magazines by file category. Each magazine is then labeled with a small index tab which identifies the major file category it contains. A category may require the use of one or more magazines depending on the number of documents in the category.

J. E. Crow - Dupont Company

Q. Are you planning in the future to incorporate features to handle 6 digit descriptors along with links and roll indicators?

A. Thank you, Jim, I forgot to mention that the Inequality Decimal Keyboard can be used with another package which we call the Interpolation Logic Module. This permits you to couple more than three digits together for search on an extended word length by the equal to greater than mode. In fact, you can couple as many as 14 keyboards together and use a word length of all 42 digits. All of these add-on features are accessories to the basic system; nothing is replaced or caused to become obsolete.

Ralph Westgard - Magnavox

Q. In the automatic input preparation where you are reading from a keypunch, how do you correlate the punchcards with the code information with the proper documents that you're photographing, either when the operator is placing them on the copy table or when they are being fed automatically? How do you know you have the right punchcard data?

A. The documents and the index sheets generated by the indexer are sent to the punchcard operator. She prepares one card for each document in the order in which the documents are arranged. The tab cards and the documents are kept in sequence and sent to the microfilmer. The Card Reader/Control Converter is placed adjacent to the microfilm operation. During recording, the operator simply checks the document against the interpretation on the tab card to verify a proper match.

Charles Bourne - Stanford Research Institute

Q. First of all, I'll make the general comment that it's sure convenient when you've got a file that you can arbitrarily slice up into pieces that fit into the cartridges so that every time you do a search you only have to do it on one or two or three files, especially when you're using a subject index approach. If you've got magazines with one or two magazines for each different subject heading that you consider to be a major subject heading, how do you post your incoming file items to each of those magazines and keep them current?

N. A. Vogel, Session Chairman—That's a very good question.

A. Thank you, Chuck. There are three ways in which this can be accomplished. First, it is possible to keep the most current incoming information in its raw form for a number of days and to have access to it in its hard copy form. The second is to

create a daily addendum similar to what the telephone company does for changing phone numbers. You would then have a two-stage lookup, that of your major category and that of your addendum. With a Selective Roll-to-Roll Film Printer which is planned for this system, you could "file expand" your accumulated addendum and splice the images of the new documents to the beginning of each particular magazine category. It is possible that in the beginning of the system, you would have less than a full magazine of film. The new Press Tape Splicer makes this a very quick and convenient way of adding new information to the beginning of the roll and therefore you locate the most current information in the search first. And because each of the documents on the roll of film is individually identified with the binary code, it makes no difference as to what order they appear on the roll of film. Thank you for bringing it up.

N. A. Vogel, Session Chairman—Thank you very much. This was a very excellent presentation. I'd like to make an observation. Will we next year see automatic cartridge installation applied to these machines?

A. This is questionable. It becomes economically difficult to beat the young lady at the retrieval station in selecting a magazine and placing it in the reader. However, if there are enough applications that indicate this need, I think we might be led into making such a device.

PROCEEDINGS

of the

Twelfth Annual Meeting,

And Convention

Edited By

Vernon D. Tate

The National Microfilm Association

Annapolis, Maryland

1963