DOCUMENT RESUME

ED 121 291 IR 003 268

AUTHOR Beale, Joel A.; Jacobson, Charles
TITLE The Design and Early Use of an Information Retrieval System for a Growing Diversified Agribusiness Technical Community.
PUB DATE Jun 75
NOTE 15p.; Paper presented at the Annual Meeting of the Special Libraries Association (Chicago, Illinois, June 8-12, 1975)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage
DESCRIPTORS Agribusiness; Cataloging; *Computer Programs; Coordinate Indexes; *Indexing; Information Retrieval; *Library Automation; Program Descriptions; Search Strategies; *Special Libraries
IDENTIFIERS Peavey Company

ABSTRACT

A library specialist and a computer specialist were asked to collaborate to solve the problem of a growing and diversified in-house, technical library in a diversified agribusiness organization. A computer program was written to: (1) generate alphabetical catalogs of collections; (2) index documents by uniterms; and (3) to retrieve documents by index terms. Books, periodicals, and documents can be rapidly indexed, updated, added, or deleted; and printouts of the latest holdings can be obtained instantaneously. (EMH)

THE DESIGN AND EARLY USE OF AN INFORMATION RETRIEVAL SYSTEM FOR A GROWING DIVERSIFIED AGRIBUSINESS TECHNICAL COMMUNITY

by

Joel A. Beale
Charles Jacobson

Peavey Technical Center
Chaska, Minnesota

Presented at a Contributed Paper Session
66th Annual Conference, June 8-12, 1975
Special Libraries Association
Chicago, Illinois

THE DESIGN AND EARLY USE OF AN INFORMATION RETRIEVAL SYSTEM FOR A GROWING DIVERSIFIED AGRIBUSINESS TECHNICAL COMMUNITY

ABSTRACT

Shows how a computerized library network grew from the collaboration of a library specialist and a computer specialist: how a single, interactive, minicomputer-based program is used to generate alphabetic catalogs of collections, to index documents by uniterms, and to browse through collections aided by threaded retrieval techniques: how at the entry level in a library of only 1,000 documents, the network not only serves users with very specialized interests, but also acts as a nucleus for a general, company-wide network to include all information found in a large, diversified, agribusiness community.

The Peavey Company is a Minneapolis based diversified agribusiness organization which began as a flour milling company 100 years ago. Although it is still very much in the flour and feed milling business, the company began to diversify and expand several years ago when it acquired Brownberry Ovens, Northwest Fabrics, and Fish Building Supply.

The Peavey Corporate Information Center Library (CICL), began as a department five years ago and was involved directly and indirectly in furnishing information and a variety of data in connection with the company's diversification and expansion mentioned above.

CICL came into existence as a brand new company operation charged with the responsibility of building a company industrial library of books, documents, periodicals and other pertinent information in the traditional sense. Intrinsic to the fulfillment of this charge has been the need for CICL to keep up with the new technical interests and information demands brought about by the company's diversification and expansion. Building a library collection to serve the information challenges of clients with very specialized interests functioning in a variety of business endeavors continues to be a challenge.

In a growing, diversified, agribusiness company, where geographical decentralization exists between corporate headquarters and other divisions, the problems of communicating library and technical information from a Corporate Information Center Library can be lessened by the use of a minicomputer-based program in library operations as described in this paper.

The principle of synergy or combined action states that it is "the effect of two or more elements working together which exceeds, by an unpredictable order of magnitude, the sum of their effects achieved separately."

This paper deals with the "working together" of a library specialist and a computer specialist in the early design and use of an automated library system for the purpose of providing an efficient information service to a technical business community.

Timing is significant in that a "go ahead" decision to automate was made by the librarian when the collection was very small and easy to handle and control. Far from the point when "10,000 documents are housed" as often advised by experts; a computer program was developed and experimentally put into operation near the beginning of CICL's formation. The librarian co-authored a report proposal to the company entitled An Automated Library Catalog and Retrieval System. The forward of this 1971 work stated that "the librarian, in concert with the electronic data processing professional, will be offering the foregoing innovative concepts as a mechanized data base library for Peavey Company."

Quoting from the forward of the 1971 report proposal we see how the groundwork for today's subject was formulated.

> The following pages are a proposal for an Automated Library Catalog and Retrieval System based upon the periodicals, books, documents, pamphlets, and other items of information within the Peavey technical library system, as well as other pertinent sources of information within the company (for example, various company files, letters, company catalogs, etc.) ... It would be possible to have information printed out in various forms.... all books could be printed out by author in book form for use at an individual's desk. This capability could bring the entire book holdings to an individual's attention. Such a system would be capable of retrieving information indexed by subject, or combination of subjects, such as "nutrition," "vitamins," "labeling," or any other subjects within the system.[1]

The above proposal was accepted immediately and work continued experimentally on the automated library catalog and retrieval system while at the same time the library was being built and the collection used in traditional form.

THE PERCEIVE SYSTEM

The system described here makes no claim for being "great," although that adjective, I am happy to point out, was used in unsolicited testimony by one of the early users of the system. (Slide 1 - Testimony) What we do claim here is that our Automated Library Catalog and Retrieval System is easy to use, that it does retrieve effectively, that it does have a necessity and that our system has proven value within the Corporate Information Center Library in the Peavey Company. We have named it PERCEIVE. P for Peavey, E for easy, R for retrieval, C for corporate, E for effective, I for information, V for value, and E for existence. It spells PERCEIVE, Peavey's Easy Retrieval of Corporate Effective Information of Value in Existence.

WHAT PERCEIVE DOES

BOOKS -- A book catalog of the holdings in the Corporate Information Center Library can be generated alphabetically by author/agency. Books are designated "T" for TOME in this mode and all books can be shown by the following: an accession number, title, and also by the holding classification number. PERCEIVE shows how many times the book has been indexed, a date, usually the copyright date or the publishing date; the book's location, price, a column headed "for," which tells who ordered it; as well as a column designated "acquired," to indicate whether the book was bought direct or purchased through a vendor. (Slide 2 - Catalog Books)

PERIODICALS -- The same data base columns as for books can be used, once the mode is shifted from books to periodicals. For example, the periodical item is numbered, as in the case of books, with an accession number. Under title is placed the name of the periodical; author/agency designation is optional. Holdings display the coverage in volumes and in years, that is, how far back the holdings go. "Type Refs," show the frequency (weekly, monthly, etc.) of the periodical and how many terms are used in indexing a particular journal. The date showing the expiration date of the subscription is also given. "Location," designates the place within the company where the particular periodical is housed, "Price," and "For" tells who ordered the particular periodical. "Acq," tells by what means this document was acquired, whether through an agency, direct, or through a vendor. (Slide 3 - Periodical Printout)

GOVERNMENT DOCUMENTS -- PERCEIVE systematically handles government documents similarly as it does in the case of books and periodicals using an accession number, title of document, the author or issuing agency, date, holdings, classification numbers and the type of document "G" which would be printed out.

JUNK -- Ephemeral or specialized information, such as maps, letters and items of one page or more can be handled by PERCEIVE as a separate specialized category called JUNK, J for justified, U for uniterm, N for Notable, and K for knowledge --or sometimes affectionately called "Joel's Uniterm Notable Knowledge." The

automated system PERCEIVE handles JUNK similarly as it does other holdings such as books and periodicals with slight variations where necessary.

PERCEIVE dictates by designated modes the name of the item being entered into the system or retrieved. Books, Periodicals, Government Documents and Junk maintained in the PERCEIVE system can be indexed, updated, mistakes corrected and items weeded, with the deletion being replaced by a substitute item. Printouts of PERCEIVE holdings can be obtained instantaneously.

THESAURUS

INDEXING BY UNITERMS

Indexing of documents is done with a sixteen character label called uniterm. The entire collection of uniterms is called the Library Thesaurus. Each uniterm is given a number when it is entered into the Thesaurus. This uniterm number serves as the computer random-access code. The Thesaurus is very easy to use. It is basically an agribusiness thesaurus, reflecting the Peavey Company's past and present interest and expressing current areas of diversification and ongoing activities. The strong points of this uniterm thesaurus are its simplicity, flexibility and universality.

THREADED RETRIEVAL

When a document is indexed, PERCEIVE creates a cross-reference entry which contains the number of the uniterm descriptor used by the indexer and the number of the next document in the uniterm "thread." The list of documents indexed under a particular uniterm is called a uniterm thread. To retrieve all documents indexed under a uniterm descriptor, PERCEIVE goes to the head of the uniterm thread to get the last document indexed by the descriptor. Proceeding step by step backwards along the thread, PERCEIVE stops the search when the first document indexed under the uniterm is located. (Slide 4)

BROWSING RETRIEVAL

The PERCEIVE user formulates a query by specifying the number and

relative weight for each uniterm descriptor in the thesaurus which interests him in a particular search. PERCEIVE follows each thread individually while building a weighted list of documents of interest. PERCEIVE then tells the user how many documents were found and asks the user what percent of documents to print out. This top percent of documents is printed out immediately upon demand to give the user a list of documents to browse through. This browsing retrieval program was chosen over Boolean Logic for simplicity, flexibility and broadness. (Slide 5)

EXTENSION TO RETRIEVAL

Eventually an automatic, ad hoc classification of the thesaurus will be implemented allowing the user to move from broad terms to related narrow terms automatically during a search.

HARDWARE AND SOFTWARE

PERCEIVE is a single interactive program written in FORTRAN IV and uses relative record random techniques to access the data base. PERCEIVE is implemented on a small timesharing system using a 16K, 16-bit word partition for its interactive input, retrieval and maintenance methods. However, the highly modular PERCEIVE program can easily be modified to run on virtually any minicomputer with a random access storage device.

THE NETWORK AVAILABILITY

The PERCEIVE system is updated and maintained on-line. Teletype terminals are located throughout the company, the system is

operable from any of these geographical points. In addition, other in-house network features are three facsimile transmitters, capable of transmitting or receiving an 8-1/2 x 11 page of information in 3 to 6 minutes between Chaska, Minneapolis, and Denver, Colorado. Installation of CRT hardware is planned for the future.

The Library bulletin routed to key management is an integral tool of the more traditional network of communication linking the Library to other departments within the company. Word Processing availability with phone dictating capabilities and an intra and inter-office mail system round out the communication network.

CONCLUSION

The knowledge or organized research and information generated by key individuals within a company constitutes an integral and vital component of the total corporate memory. In their book The Corporate Memory, authors Barbara N. Weavers and Wiley L. Bishop define the Corporate memory as:

> ... the total in-house information systems and services of an organization, which are established to collect, organize, and store efficiently all documentation generated into the company and which have the inherent feature of retrieving documents and/or information on a current and retrospective basis upon demand...[2]

The Peavey Company's interest in encouraging and stimulating personal growth and creativity of its employees reflects the organization's character. The computer specialist and the library specialist utilized this conducive climate to team their professional skills in the development of PERCEIVE. My colleague and I agree that unpredictable benefits were derived through our collective efforts, in designing PERCEIVE and the Peavey Company now has the advantage of an advanced corporate memory.

> ... Collectively, then all record centers and data processing systems within an organization constitute the corporate memory of that organization.[3]

BIBLIOGRAPHY

1. Holbrook, Iola Belle, "A Threaded-File Retrieval System," American Society of Information Science Journal, (Jan.-Feb., 1970) pp. 40-48.

2. Lefkovitz, David, "Automatic Classification and its Application to the Retrieval Process," File Structures for On-line Systems (New York, Spartan Books, 1969) pp. 186-201.

3. Segebarth, C., "Comments on Threaded-File Retrieval System" 'Letter to Editor,' American Society of Information Science, (Sept.-Oct., 1970) p. 372.

SOURCE FOOTNOTES

1. Beale, Joel A., and Charles Jacobson, An Automated Library Catalog and Threaded File Retrieval System, Peavey Corporate Research Report, 1971, p. 1.

2. Weaver, Barbara N., and Wiley L. Bishop, The Corporate Memory: A Profitable and Practical Approach to Information Management and Retention Systems. New York: John Wiley & Sons, 1974, p.1.

3. Ibid., p.2.