

BIBLIOGRAPHY

- (1) "The National Library of Medicine Index Mechanization Project," *Bull. Med. Libr. Assoc.*, **49**, No. 1, Part 2, (1961).
- (2) S. I. Taine, "The Future of the Published Index," Machine Indexing: Progress and Problems, Center for Technology and Administration, School of Government and Public Administration, The American University, 1961, pp. 144-169.
- (3) J. Markus, "State of the Art of Published Indexes," *Am. Doc.*, **13**, 15 (1962).
- (4) M. S. Bryant, "Bibliographies, Abstracts and Indexes," *The State of the Library Art*, **2**, Part 2, 1, (1960).
- (5) C. K. Schultz, and C. A. Shepherd, "The 1960 Federation Meeting: Scheduling a Meeting and Preparing an Index by Computer, *Federation Proc.*, **19**, 682 (1960).
- (6) C. K. Schultz, and P. A. Schwartz, "A Generalized Computer Method for Index Production," *Am. Doc.*, **13**, 420 (1962).
- (7) S. Herner, "Subject Slanting in Scientific Abstracting Publications," *Proc. Intern. Conf. Sci. Inf.*, 407 (1959).
- (8) J. T. MacMillan and I. D. Welt, "A Study of Indexing Procedure in a Limited Area of the Medical Sciences."
- (9) I. D. Welt, "A Combined Indexing-Abstracting System," *Proc. Intern. Conf. Sci. Inf.*, 449 (1959).
- (10) Report of the Commission on Enzymes of the International Union of Biochemistry, Pergamon Press (1961).
- (11) J. T. MacMillan and I. D. Welt, "Some Personnel Problems of a Small Indexing Project," *J. Chem. Doc.*, **2**, 79 (1962).
- (12) E. Garfield, "A Unified Index to Science," *Proc. Intern. Conf. Sci. Inf.*, 461 (1959).
- (13) C. Montgomery and D. R. Swanson, "Machinelike Indexing by People," *Am. Doc.*, **13**, 359 (1962).
- (14) I. D. Welt, "The Detailed Indexing of Biological Effects of Chemical Substances," *Bull. Med. Libr. Assoc.*, **46**, 367 (1958).
- (15) I. D. Welt, "Guide to the World Literature on Cardio-vascular Agents," *Medical Doc.*, **5**, 9 (1961).
- (16) G. M. Dyson, and M. F. Lynch, "Chemical-Biological Activities, A Computer-Produced Express Digest," Preprint of paper presented before the Division of Chemical Literature, 142nd National Meeting of the American Chemical Society, Atlantic City, N. J., September, 1962.
- (17) I. D. Welt, "Aspects of the CBCC Biology Code of Interest to Chemists," *J. Chem. Doc.*, **1**, 19 (1961).

Information Retrieval at the New Monsanto Information Center*

By JOHN R. VAN WAZER, PAUL LOGUE, and WILLIAM A. WILKINSON

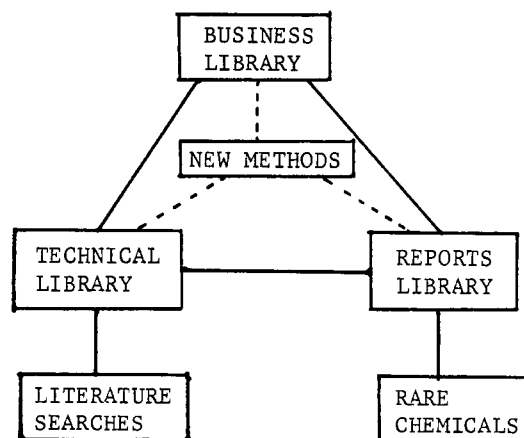
Monsanto Chemical Company, St. Louis 66, Mo.

Received February 20, 1963

In June, 1961, a large and widely scattered group of information services in Monsanto were centralized as a newly established Information Center. The Center consists of a Central Technical Library, a Central Business Library, a Central Reports and Chemicals Library, and a small group working on the introduction of New Methods of Information Handling to our operations. Because of the great strides presently being made in the area of information handling plus the ever-growing need for better and cheaper methods, we were prompted to include this small group doing research on new methods as part of the over-all Information Center. A brief diagram of the organization of the Information Center is shown in Fig. 1.

The establishment of the Center seemed to be an appropriate time for making changes to certain of the traditional but now outworn methods of information handling by more up-to-date ones. Outstanding among the changes were the introduction of a machine-prepared book catalog and the conversion of the indexing of Monsanto reports to a computer-based operation. The third area in which improvements were made was a Union List of Serials for all Monsanto libraries.

* Presented before the Division of Chemical Literature, 142nd National Meeting of the American Chemical Society, Atlantic City, N. J., September, 1962.



Section Heads:

Business Library	- Charlotte Perabo
Technical Library	- William A. Wilkinson
Reports Library	- Margaret E. Madden
New Methods	- Paul Logue

Fig. 1.—Organization of the Monsanto Information Center. The organization is led by an annually rotating chairman chosen from the section heads.

In the early consideration of potential improvements in our methods of information handling, we envisaged that much of the work of retrieving information would be relegated to computers in the reasonably near future. Therefore, the changes we have made have been carried out with this in mind. It is envisaged that the switchover from our present system to any type of entirely or partially mechanized computer operation would be a relatively simple process, not nearly as cumbersome as the job of transferring information from the outmoded library file-card systems to punched cards and tapes. Although some thought was given to going immediately to completely mechanized searching in which at least the abstracts would be stored in the computer to be printed out when needed, it was quite apparent that—at the present scale of our information-retrieval operations—such an approach would be uneconomical using the available computers and methods of mechanized information handling. Therefore, we turned to the next best approach—that of using machines for preparing indexes in such a way that the final information would be more readily available and that searches for the information could be done in less time to greater depth.

Machine-Prepared Book Catalog.—Our new book catalog typifies the difference between the old and new methods of information handling; therefore, the situation before and after its introduction will be reviewed in some detail. Before development of our new catalog, Library of Congress cards were employed in the usual manner. In the Monsanto research complex at St. Louis, four sets of these cards would have been necessary, with one at the Central Library, two others at reading rooms in other buildings, and the fourth at a sublibrary in downtown St. Louis, about 12 miles from our Central Library at Creve Coeur, Mo. Under the old system, incorporation of a new book into any of the four library locations would have necessitated ordering a set of Library of Congress cards for all of these locations so that the patrons of a given library site would be informed about the over-all book holdings of Monsanto in the St. Louis area. For an average technical book having, say, three author headings, five subject headings, a title heading, as well as a card for the shelf list (in which the books are arranged by Dewey decimal number), headings would have to be typed individually on 40 Library of Congress cards, and these cards would be handsorted into the file drawers at four different locations. Moreover, the accession of a book involves putting the Dewey decimal number on its spine, putting a labeled card pocket in the rear, and typing out two sign-out cards to be kept in the pocket. [Under our system, one sign-out card is filed according to the Dewey decimal number of the book and the other according to the name of the borrower.] In the old system the files of sign-out cards had to be checked quite often to see which books were overdue. Then an overdue notice was written by hand (or typed) and sent to the borrower.

With the present system, a single keying on an IBM key-punch handles all of the pertinent information concerning the book—authors(s), title, subject headings, Dewey decimal number. The resulting punched cards are then put into an IBM Document Writer for which simple programs were developed to print out the following things: (1) the Dewey decimal number on a sticker for the spine

of the book; (2) a similar sticker for the card pocket in the rear of the book; (3) two special IBM punch cards which are used as sign-out cards; (4) author(s), title, publisher, date, book number, and location in a listing of new books which is posted on the appropriate bulletin boards and distributed to various individuals by mail *via* Addressograph plates.

The newly punched cards are then stored with the sets of cards corresponding to all of the other books. Occasionally these are put through a card sorter and the IBM Document Writer to give a complete index of all the holdings. This index includes alphabetical listings by (1) author, (2) title, and (3) subject headings. The Document Writer prints the catalog on lithographic masters so that around 1000 copies of the catalog may be run off. About once a month, a cumulative supplement is issued; and, at the end of each year, the previous year's catalog and cumulative supplement will be thrown out and replaced by a new catalog very much as in the case of telephone books.

According to our new system, individual patrons of the library may have personal copies of the book catalog so that they do not have to go to a central set of file cards to get this information. Furthermore, the cumulative supplements keep the index about as up-to-date as a standard library card file would be, but the job of filing duplicate sets of cards has been abolished. A patron of the library now finds it convenient to look up a book in his personal catalog and then to call to see whether it is available to be sent to him in the Company mails (or, if it is signed out, to ask to be put on the waiting list). This is less frustrating than the former situation of having to go to the library, look up the book in those ubiquitous but most unhandy library card-catalog files, only to find in some case that someone else had signed out the book or that the book was not included in the library's collection. Such a trip to the library is now made only when a book is needed in a hurry and the patron has previously ascertained that it has not been signed out by someone else.

Index to Monsanto Reports.—The procedure formerly used for indexing Monsanto reports was a very lengthy and expensive one. It consisted of having the author in some Divisions and a librarian in other Divisions of Monsanto write an abstract and suggest title headings for the report. These then went to a central clearing house where an attempt was made to achieve uniformity in the indexing headings. The abstracts were typed on lithographic masters, without the subject headings. Although only 15 locations around the country had card indexes of the Monsanto reports, 17 sets of cards were made to allow for expansion into new sites. The sheets of cards, without subject headings, were again run through a lithographing machine to print the appropriate headings. The cards were then cut apart and assembled into sets to be sent to the various locations where these sets were alphabetically sorted into the card-index files by hand.

According to the new method,** a single typing on the key-punch is sufficient to put all of the information about a report into a form where it can be handled by the computer. We use coordinate-term indexing with links and roles. For our purposes, the 16 roles shown in Fig. 2 seem

** P. Logue, *J. Chem. Doc.*, 2, 215 (1962).

Role	Interpretation
08	Study of
10	Design of
12	Patent on
13	Review of
14	Government
15	Business information
21	Of, in, on
27	Product, output
33	Reactant
45	Environment
54	Special agent
56	By product
62	Independent variable
69	Dependent variable
71	Using, used for
90	(leave blank)

Fig. 2.—Role codes used in the bibliographic abstract.

to be sufficient. Instead of writing abstracts, we now, let these roles plus the appropriate terms form an abstract in themselves, as indicated by the example of Fig. 3. Although our abstracts are somewhat redundant, since a given role may be printed out a number of times in succession, they are easy to read. Moreover, these ab-

23838	23838
PHOSPHORUS PENTASULFIDE, ITS REACTION PRODUCTS AND RELATED COMPOUNDS	
MAIER L	
MOEDRITZER K	
VAN WAZER JR	
REPORT NO. 1520 TYPE F SOURCE RR 020	
JOB NO. R 9523 YR.62 MO. 3 DAY 7	
TOPIC A	
STUDY OF	EQUILIBRIA
STUDY OF	PREPARATION/CHEMICAL/
REACTANT	PHOSPHORIC ACID, TRIALKYL ESTERS
REACTANT	PHOSPHOROTETRATHIOIC ACID, TRIALKYL ESTERS
REACTANT	PHOSPHOROTETRATHIOIC ACID, TRIARYL ESTERS
REACTANT	PHOSPHORUS PENTOXIDE
ENVIRONMENT	PHOSPHORIC ACID, TRIALKYL ESTERS
USING, USED FOR	NUCLEAR MAGNETIC RESONANCE METASTABLE
TOPIC B	
STUDY OF	EQUILIBRIA
STUDY OF	PREPARATION/CHEMICAL/
PRODUCT, OUTPUT	PHOSPHOROTHIOIC ACID, POLY-, SODIUM SALT
REACTANT	PHOSPHORUS PENTOXIDE
REACTANT	SODIUM SULFIDE
USING, USED FOR	NUCLEAR MAGNETIC RESONANCE METASTABLE
TOPIC C	
STUDY OF	SOLUTION
STUDY OF	VISCOSITY
OF, IN, ON	PHOSPHORUS SULFIDES
ENVIRONMENT	BUTYL ALCOHOL
	AMORPHOUS
	CRYSTALLINE

Fig. 3.—Machine-made abstract of a Monsanto report.

stracts give considerable detail concerning the subject matter of the report without going into highly confidential data, such as product yields, whether a given process was adopted or not, comparison of the efficacies of alternate routes, etc. This is an advantage for an industrial concern and has proved to be of little or no disadvantage to the user of the new index since a person interested enough to look up a subject will probably find it worthwhile to page through a pertinent report whether or not he knows *a priori* that chemical synthesis A was preferable to chemical synthesis B or merely that synthesis A was compared with synthesis B without a statement being given as to which one was the better.

The Index to Monsanto Reports consists of three books. One of these is a permanent compilation of bibliographic abstracts which is kept up-to-date by insertion of new abstracts as they are issued. On each page of the bibliographic abstracts, there are approximately 11 abstracts of the type shown in Fig. 3. The second volume of the Index consists of alphabetical listings by author and by job number, as well as separate indexes for translations and for manuscripts to be published by Monsanto authors. Since approximately a year elapses between the time a scientific paper is submitted for publication and the time it appears in print, inclusion of manuscripts submitted for publication by Monsanto authors gives the company a lead time of about a year for possible exploitation of the subject matter of the manuscript in areas other than those with which the author was directly associated. Last but not least is the book containing the coordinate-concept index which consists of a dual dictionary from which the desired report is found from two (or more) terms or descriptors.

The distribution of the Index to Monsanto Reports was determined by the Research Directors of the various Divisions of Monsanto who, obviously, were in the best position to assign the indexes to the proper persons having need for this information. At the time of this writing, we are in the process of issuing the first index so we are unable to tell you about its practical usage. However, we expect that the majority of users will look up the accession number of the appropriate reports and request by telephone that these reports be mailed to them. Since the indexes are being distributed to the people needing them, we expect to eliminate the lost motion involved in having these people go to a library location in order to look up items in the internal literature. The contents of the volume which includes the author and other indexes as well as the contents of the dual dictionary of terms are disposable and will be replaced approximately annually, as with the Book Catalog. Between times, they will be kept up-to-date with cumulative supplements. As previously noted, the bibliographic abstracts are permanent, being continuously brought up-to-date by additions but never replaced.

Union List of Serials.—The IBM Document Writer has also proven very useful in preparing a list of serials subscribed to by the various Monsanto libraries and reading rooms. A typical excerpt is shown in Fig. 4, in which the symbols CC, DA, EV, IC, NI, SL, SP, and TC refer to Monsanto library sites in various part of the country. In addition, the single letter C stands for a current subscription and M indicates that part or all of

BER DEUT CHEM GES	CC	V1-77 1868-1944
	DA	V1 1868-1944
	EV	M V51-77 1918-1945 /MICROCARD/
	IC	M V1-77 1868-1945 INDEXES AS V1-70
	IC	M 1868-1937 /ON MICROCARDS/
	NI	V1-79 1868-1946
	SL	V1-77 1868-1944 CUM INDEX V1-29
	SL	1868-96, AUTHOR INDEX V30-70
	SP	V1-77 1868-1944 LACKS V 48, 53, 56,
	SP	69
BETTER CROPS WITH PLANT FOOD	CC C	V36 OCT, 1952-DATE
	CC C	
BIBLIOGRAPHY AGR	DA	V8-18 1946-54 LACKS V10#1 1947
	IC C	V7-12 1945-48, V15 #9 1951-DATE

Fig. 4.—Typical excerpt from the Union List of Monsanto Serials.

the holdings are on microfilm or microcards. As an appendix to this List of Serials, there is a listing of the

major reference works, particularly those that are brought up-to-date by regular reissues. By use of the Document Writer and an IBM 407 printer, this List of Serials is updated occasionally, for wide distribution within the Company.

Conclusion.—By use of modern methods of information handling, we have been able to offer more service to our patrons at considerably less cost. At the present time, we are employing computers and related information-handling machinery to prepare catalogs which are easy to use and have wide distribution. The act of information retrieval is still being done by hand, but, by improved make-up of the appropriate catalogs, it has been made easier and information may be sought in greater depth. As the operations of the Information Center grow, we will be in a position to use machine retrieval as soon as this becomes more economical than retrieval by hand.

Extensive Relations as the Necessary Condition for the Significance of "Thesauri" for Mechanized Indexing*

By MORTIMER TAUBE

Documentation Inc., Bethesda 14, Md.

Received February 20, 1963

Recently there have appeared a number of attempts to develop "thesauri" for use with mechanized information systems. H. P. Luhn, in discussing the possibility of auto-indexing, suggested that words could be arranged in natural families and noted the parallel between such families and the arrangement of Roget's *Thesaurus*.¹ One of the mechanical translation groups in England actually attempted to use Roget to supply contextual definition of words.² The Armed Services Technical Information Agency group, partly to distinguish the results of its work from the "ASTIA Subject-Headings" from which it started, called its product a "Thesaurus of ASTIA Descriptors." The name was also justified because a grouping of descriptors into 292 descriptor groups was provided, and this grouping suggests the grouping of words in Roget. The Integrated Engineering Control Group of the du Pont Engineering Department also constructed a thesaurus, which used as raw material a "word association matrix" based on certain suggested procedures developed by Documentation Incorporated. This thesaurus, as finally published by the American Institute of Chemical Engineers, is based only incidentally on the "word association matrix" and primarily on relations of terms determined by consulting dictionaries, handbooks, and usage. In the A.I.Ch.E. thesaurus no over-all grouping of terms is provided, but under each term there are displayed "related

terms" and terms which in some unspecified sense are "generic to" and "included in" the heading term. In a recent paper,³ Vickery has taken sharp exception to the use of the term "thesaurus" to describe authority lists, vocabularies, or subject-heading guides, and to the implication that a new name creates a new type of apparatus. However, the ASTIA organization and several other organizations and individuals have continued to insist that in "thesauri" there are provided structures of connections between words which are necessary to ensure satisfactory operation of mechanized systems of coordinate indexing. It is assumed that the connections of terms and the cross reference structures displayed in a thesaurus are necessary both as indexing and searching aids.

Let it be admitted that a structure of terms is a *sine qua non* of a complete indexing system. There remains to be determined whether such a structure is derivable from and reflective of the indexing operation or whether it can possess some sort of independent validity which is prescriptive, rather than descriptive. In other words, is a thesaurus or any authority list an independent semantic standard for an indexing system set up by a process of lexicography, or is it a description of a particular indexing system as developed from a concatenation of subject competence, the literature being indexed, and the requirements for efficient machine search?

In the history of the subject organization of scientific publications there have been many attempts to establish prescriptive systems—the Dewey Decimal System, the Universal Decimal System, the Library of Congress Classification System, the Bliss Classification System, the

* Presented before the Division of Chemical Literature, 142nd National Meeting of the American Chemical Society, Atlantic City, N. J., September, 1962.