MASTER

IS-1074

Mathematics and Computers (UC-32)
TID-4500 (43rd Ed.), July 1, 1965

SEP 20 1965

UNITED STATES ATOMIC ENERGY COMMISSION

Research and Development Report

AMES LABORATORY COMPUTER PRO-
DUCED PERMUTED TITLE INDEX AND
PUBLICATION UNIT RECORD GUIDE

by

C. H. Cameron and D. R. Fitzwater

June, 1965

RELEASED FOR ANNOUNCEMENT
IN NUCLEAR SCIENCE ABSTRACTS

Ames Laboratory
at

Iowa State University of Science and Technology
F. H. Spedding, Director
Contract W-7405 eng-82

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency Thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

IS-1074

This report is distributed according to the category Mathematics and Computers (UC-32) as listed in TID-4500 (43rd Ed.), July 1, 1965.

LEGAL NOTICE

This report was prepared as an account of Government sponsored work. Neither the United States, nor the Commission, nor any person acting on behalf of the Commission:

- A. Makes any warranty or representation, expressed or implied, with respect to the accuracy, completeness, or usefulness of the information contained in this report, or that the use of any information, apparatus, method, or process disclosed in this report may not infringe privately owned rights; or
- B. Assumes any liabilities with respect to the use of, or for damages resulting from the use of any information, apparatus, method, or process disclosed in this report.

As used in the above, "person acting on behalf of the Commission" includes any employee or contractor of the Commission, or employee of such contractor, to the extent that such employee or contractor of the Commission, or employee of such contractor prepares, disseminates, or provides access to, any information pursuant to his employment or contract with the Commission, or his employment with such contractor.

Printed in USA. Price \$ 2.00 . Available from the Clearinghouse for
Federal Scientific and Technical Information, National
Bureau of Standards, U. S. Department of
Commerce, Springfield,
Virginia

TABLE OF CONTENTS

	Page
ABSTRACT	1
I. INTRODUCTION	1
II. A KWIC INDEX	1
III. KWIC SYSTEM INPUT	3
A. Definition of a data set	3
B. General Information	3
C. Restrictions placed on a data set	6
IV. THE KWIC INDEX PROGRAM	7
A. Edit Program	7
B. Five IBM 1401 Programs	11
C. Sort 90.	14
V. PROCEDURE FOR SETTING UP A KWIC INDEX	15
VI. APPLICATION OF THE KWIC INDEX	15
VII. THE STOP WORD LIST	17
VIII. CONCLUSION	21
REFERENCES	23

AMES LABORATORY COMPUTER PRODUCED PERMUTED TITLE (KWIC) INDEX AND PUBLICATION UNIT RECORD GUIDE

C. H. Cameron and D. R. Fitzwater

ABSTRACT

A KWIC Index System adapted for the Ames Laboratory is described. This report is a guide for producing permuted indexes for personal files, annotated bibliographies and for the Ames Laboratory's annual progress reports. It also serves as a guide for maintenance of a multi-purpose bibliographical record of research and development reports and publications.

I. INTRODUCTION

The expanded application of permuted title indexes has created a need for a guide to aid in preparing a permuted title index. A knowledge of permuted title techniques may assist writers in forming more descriptive titles where the KWIC Index is used as a classification scheme in which the indexed items are differentiated into classes and subclass strata.

II. A KWIC INDEX

The Key Word-In-Context (KWIC) Index or permuted title index, designed by the late H. P. Luhn of the International Business Machine Corporation (IBM),¹ set up a standardized indexing scheme through

which important or key words in a title could be indexed. A permuted title index is an index of key words selected from a title and rotated in such a manner that each one appears in alphabetical sequence and serves to index the title. This is similar to reading a name on a cylinder; regardless of the position of the cylinder some portion of the name appears.

An example of the rotational effect of a permuted title is shown below: Item number 0900-P-35 has bronze as the indexed term. The words sodium, tungsten and conductivity may also be selected as key words and appear in alphabetical sequence in the list of indexed terms. The words that do not appear are the "stopped" terms, usually referred to as stop words. In addition, the sixteen most commonly used prepositions, conjunctions, and articles are treated as stop words. Stop words are discussed in detail on page 17.

VITY OF SODIUM TUNGSTEN BRONZE= ELECTRICAL CONDUCTI 900-P-35

The significant points of interest shown in the above example are:

1. The title is treated as an unbroken line.
2. The rotation of the words.
3. The equal (=) sign indicates the title termination.
4. Classification code associated with the title.
5. Indexed terms are listed in alphabetical sequence in the approximate center of the page.

There is, however, a drawback which isn't immediately apparent in the manner in which the indexed items are sorted and listed. Plural terms are not always immediately preceded by the singular term. Attempts are being made to overcome this difficulty.

The KWIC procedure offers more aid than conventional indexing procedures in locating information because descriptive words appear on either side of the key word.

The permuted title index (KWIC) as modified for use within the Ames Laboratory has been used to maintain an index of the status of research papers and research and development reports, and to produce an index for the Ames Laboratory annual progress report.²⁻⁴

III. KWIC SYSTEM INPUT

A. Definition of a data set

The type of input data employed in the Laboratory's permuted titles index system follows the guides and standards set by IBM. A data set is a set of cards which describe the items to be indexed.

The possible inputs for each item to be indexed are the local requirement card, author card, title card, source card, and local trailer card. An item to be described must consist of at least an author card and a title card. In most instances the other cards also are used.

B. General Information

The analysis of the general information fields leads to a breakdown of general information fields into "words" by use of the following rules:

1. A word is a string of characters separated from its neighbor by a blank or ends of fields.
2. All imbedding special symbols and characters are considered as part of the imbedding word.
3. All leading special symbols are treated as stop words and are separated from the following word.
4. Terminating special characters are treated as a part of the word.
5. Words may not be hyphenated and continued to a next card, because they will be treated as two words. These words are then used as potential index entries.

The general information fields are further delineated by class code (column 75) which determines the form of information on any particular card. The general information (columns 1-60) is coded or keypunched using the following formats: (NOTE: THE ASSOCIATED CLASS CODE (column 75) IS ALSO GIVEN)

CLASS CODE (0) - Local header class - Information is punched in columns 1-58 as follows:

Cols. 1-18	Group leader directing research
19-20	First and middle initial of group leader.
21-27	IS number.
30-35	Submitted data.
36-40	Published date.
41-45	Publication contribution number.

46 Rare earth code.

51-58 Activity or budget number.

The local header (requirement) information is usually punched on one card. Additional cards can be used if the research was jointly compiled by more than one group leader.

CLASS CODE (1) - Author class - Personal authors are punched in columns 1-18, 21-38, and 41-58. Additional personal authors may be punched in subsequent cards observing the above defined fields.

CLASS CODE (2) - Title class - Title information is punched in columns 1-60. Subsequent title information should be punched in columns 3-60.

CLASS CODE (3) - Source class - Source information is punched in columns 1-60.

CLASS CODE (5) - Abstract class - Abstracts are punched in columns 1-60.

CLASS CODE (9) - Local trailer class - Information (if any) is punched in columns 1-60.

A further breakdown of an input card by definition fields are:

COLUMNS 1-60 General information (see above)

COLUMN 61 Blank

COLUMNS 62-72 Classification and/or cross reference identification code.

COLUMNS 73-74 Continuation or sequence number within a class.

COLUMN 75 Class code.

COLUMNS 76-80 The last five columns are used for a unique sequence number which is assigned by the Edit Program within the KWIC System for its own housekeeping. The local requirement and local trailer cards are used only for an entire data set and are supplied by the operator of the system.

C. Restrictions placed on a data set

Generally speaking the KWIC Index System as designed by IBM has few restrictions and these have not been changed by our modifications. The primary limitation that accompanies the current computer-generated output is that general typesetting and typewriter conventions cannot be used. As a result, substitutes for punctuations are used widely. We are using the conventions found satisfactory for our own use.

Semicolons, colons, question marks, quotation marks, and exclamation points have been eliminated because they are impossible to reproduce on our computers and no substitutes are used. The substitute used for the apostrophe is the slash (/). Other special characters which are used in the Ames Laboratory system are the dash (-),

parentheses (), dollar sign (\$), comma(,), period(.), asterisk(*), plus (+), and "at" sign (@).

For simplicity and consistency, several other conventions have been established. All chemical symbols and abbreviations should be eliminated by spelling out the technical name or symbol. The only exception is the Greek letters α and β for which "A" and "B", respectively, are used as substitutes. The chemical name for an isotope is written followed by a hyphen (dash, -) and the isotope number; for example, U^{235} is written as URANIUM-235.

Examples of the coded and keypunched cards that make up an item to be indexed are given in Figs. 1 and 2.

IV. THE KWIC INDEX PROGRAM

Auxiliary information and options mentioned below are supplied by the operator and are included in order to provide a better understanding of the system. A generalized flow chart of the operation of the KWIC System is given in Fig. 3. The Ames Laboratory KWIC Index System is divided into three sections, the Edit Program, five IBM 1401 Programs, and the IBM Sort-90.

A. Edit Program

The Edit Program was written to scan all input prior to Run I. The program assigns a unique sequence number to each data set. Two blank cards (local trailer cards) are inserted after each item to be indexed, thereby providing a more readable bibliography listing of each set.

Fig. 1a. Research and Development Report

General Information for Codes 1 - 11																	CLASSIFICATION CODE			
1	5	10	15	18	21	25	30	35	38	41	45	50	55	58	60	61	CLASS PAGE OR PUB. NO.	CONT.	SEQ. NO.	CLASS CD
																	6263 65	69	737475	80
SAGE CR. ANDERSON, R.R. CAMERON, CH.																	I, S, 6, 4, - C, -	9, 4, 0	- 1 1	
AMES, LABORATORY, SDI, SYSTEMS, REFERENCE, MANUAL																	I, S, 6, 4, - C, -	9, 4, 0	- 1 2	

Fig. 1b. Publication

General Information for Codes 1 - 11																	CLASSIFICATION CODE			
1	5	10	15	18	21	25	30	35	38	41	45	50	55	58	60	61	CLASS PAGE OR PUB. NO.	CONT.	SEQ. NO.	CLASS CD
																	6263 65	69	737475	80
FLESCH, G.D. SVEC, H.J.																	PU 6 4, - C, -	1 0 3 8	- 1 1	
A MODEL OF A MASS SPECTROMETER TO SIMULTANEOUSLY COLLECT																	PU 6 4, - C, -	1 0 3 8	- 1 2	
, POSITIVE AND NEGATIVE IONS.																	PU 6 4, - C, -	1 0 3 8	- 2 2	
PROCEEDINGS OF THE 9TH ANNUAL MEETING OF ASTM COMMITTEE																	PU 6 4, - C, -	1 0 3 8	- 1 3	
REV. SCI. INSTR. 34, 897, (1963)																	PU 6 4, - C, -	1 0 3 8	- 2 3	

Fig. 1c. Annual Report Entry (Chemistry)

General Information for Codes 1 - 11																	CLASSIFICATION CODE			
1	5	10	15	18	21	25	30	35	38	41	45	50	55	58	60	61	CLASS PAGE OR PUB. NO.	CONT.	SEQ. NO.	CLASS CD
																	6263 65	69	737475	80
FITZWATER, DR.																	0, 9, 0, 0, - C, -	0, 7, 8 A	- 1 1	
CRYSTALLOGRAPHIC DATA ANALYSIS																	0, 9, 0, 0, - C, -	0, 7, 8 A	- 1 2	

Fig. 1d. Annual Report Entry (Engineering)

General Information for Codes 1 - 11																	CLASSIFICATION CODE			
1	5	10	15	18	21	25	30	35	38	41	45	50	55	58	60	61	CLASS PAGE OR PUB. NO.	CONT.	SEQ. NO.	CLASS CD
																	6263 65	69	737475	80
SMUTZ, M.																	0, 9, 0, 0, - E, -	0, 0, 3 A	- 1 1	
SEPARATION OF RARE-EARTH NITRATES IN A PULSE COLUMN																	0, 9, 0, 0, - E, -	0, 0, 3 A	- 1 2

Fig. 3. Generalized flow chart of Permutated Title Index System.

In addition, the Edit program provides a listing of all document input, allowing corrections to be made before the "edited card deck" is run.

The Edit Program also checks the class code that makes up an item. Continuation cards must be in the proper sequence. If any of the sequence tests are violated, the card in error and the remaining cards of that item are rejected. The program continues until the last card has been read. Each input card accepted by the Edit Program is then punched. The punched deck (edited input) is given the unique sequence number mentioned above. The Edit Program will internally assign a sequence number based on a constant supplied to the program by the operator.

The sequence number is usually incremented by three from item to item to allow subsequent item inserts if necessary, i. e., A0001, A0004, A0007, etc. This sequence increment also can be altered by the operator.

B. Five IBM 1401 Programs

1. KWIC program #1

Run #1 of the KWIC Index System prepares from a data set (card input) two tapes which are used throughout the remaining sections of the KWIC System. Each item is checked for an author record, a title record, a unique source record and a sequence number. The accepted data records are then written without a change in

form on tapes #1 and #2. Title records are written on tape #1 to form a Title Tape and all the remaining records are written on tape #2 to form a Bibliography Tape.

Run #1 has been modified so that expanded use of the reference code can be made. The modifications allow a special classification code to be generated if it has not been supplied. Generated codes are particularly useful for diversified input such as personal KWIC Indexes.⁵

2. KWIC program #2

Run #2 uses either a sorted or unsorted output tape #2 from Run #1 as an input tape. The unsorted tape #2 is used only if an unsorted author tape is to be produced. Sense switch settings allow for this option.⁶

Generally, the output tape #2 is sorted prior to use in Run #2.

Sense switch settings will permit the options of printing a bibliography listing and writing unsorted author records on an Author Tape, printing only the bibliography listing, or simply writing unsorted author records on an Author Tape. The only output tape from Run #2 is the unsorted Author Tape.

3. KWIC Program #3

Run #3 uses as input a sorted Author Tape. The author records are printed in a single column with the name of the author to the left of the first reference code. In the event there is a change to

a new page before all references to a particular author are complete, the author's name will be printed to the left of the reference code on the new page.

It has been our experience that a considerable savings in time can be obtained by generating the Author Tape before sorting the Bibliography Tape. The time required to generate the Author Tape from the unsorted Bibliography Tape is approximately two minutes. There is less tape handling and time overlap in the sorting of the two tapes, i. e., the unsorted Author Tape can be mounted and ready before the Sort-90 completes Phase 2 of the sorting procedure when sorting the Bibliography. Thus, the Bibliography Tape can be listed while the Author Tape is being sorted.

4. KWIC Program #4

Run #4 uses as input the output tape #1 (Title Tape) from Run #1. The primary function of Run #4 is to edit the titles in such a way as to allow indexing of the key words. Editing is by a rotational process wherein the entire title can be allowed to "wrap-around" or not "wrap-around". This option is determined by sense switch settings.

Wrap-around is a continuous title line in which some portion of the title precedes or follows the key word. It is analogous to labelling on a jar; if the label is read from any reference point, a portion of the label can be seen.

Each edited title (rotated) record is written on an output tape known as the KWIC Tape. The KWIC Tape at this juncture contains unsorted title records which must be sorted prior to Run #6.

5. KWIC Program #6

Run #6 uses as input the sorted KWIC Tape from Run #4. Run #6 will print, or write on an output tape, a permuted title listing of non-stopped key words. The stopped key words are determined by a dictionary of stop words. All "words" which appear in the text but do not appear in the stop word dictionary are used in the index. All stop words will appear in the title but not as sorted index terms (key words).

C. Sort 90

The IBM 7070/7074 Generalized Sorting Program⁶ Sort #7070-SM-077, Version 5 is used to sort the Author, Bibliography, and KWIC Tapes. Except for Run #1, each IBM 1401 program described above requires tape labels on each input tape. In addition, each program writes a label on the output tapes produced. Modifications have been made to the Sort-90 package to allow for reading (and checking) and writing tape labels. The pertinent information required for the Sort-90 control cards can be obtained from the operator of the System.

The Sort-90 program is run on the IBM 7074 utilizing the facilities of the Iowa State University Computation Center.

V. PROCEDURE FOR SETTING UP A KWIC INDEX

To set up a card file for a permuted title index one must first determine what type of identification code will be used for cross referencing. Consideration should be given to the scope of the index.

The primary purpose of the identification code is to provide any user of the KWIC Index with a key for looking up the reference in the bibliography. For this purpose a simple sequence number will suffice. If desired, the identification code may provide classification information in the KWIC Index itself. In such a case this classification identification can be used to sort the data into subfiles.

Special coding sheets provide a guide for coding information used in the index. Thus, coding the information to be punched is a relatively simple function. Examples of each card class coding are given in Fig. 1.

VI. APPLICATIONS OF THE KWIC INDEX

The permuted title index, illustrated in Fig. 4, is the primary reference section of the indexing system. Although this illustrates the application of the permuted title index within the Ames Laboratory, it is not limited to annual reports. Such indexes may also be used for many filing and retrieval problems both large and small⁷. For example, this system has been used with remarkable success in organizing a personal filing system of miscellaneous items for an individual scientist. This type of index allows the scientist to eliminate housekeeping and retrieve

Fig. 4. Four pages from a Permuted Title (KWIC) Index listing.

information in a far more flexible manner than in a catalog filing system. Additional applications have been made in the area of documentation studies, catalog reference, and retrieval. The relative significance of a reference can be ascertained by the context in which key words appear. The casual user will particularly enjoy fingertip access to any information that is a part of the sorted title listing. The index can be updated and expanded at any time.

The Bibliography Index, Fig. 5, is simply a bibliography listing in classification code sequence of all documents which are KWIC indexed. The classification code may be used to segregate bibliography entries in order to make browsing possible.

The Author Index, Fig. 6, is a listing of the authors whose names appear in the Bibliography Index and aids in quickly locating the contribution of each author.

VII. THE STOP WORD LIST

The Stop-Word List, Fig. 7., consists of words and integers. Stop words are punched into a deck of cards which are placed behind Run #6. Because some words in a title are not appropriate as sorted index terms, this list of stop words leads to a more concise but equally useful index.

Changes in the stop word list may be required for some indexes inasmuch as the relevance of a word is dependent on the use of the index. In general, a number in a title does not add any special

- IS64-C- 893 FITZWATER DR HIETBRINK EH LEDET MM
MCFARLAND DE
POGO PSAG MANUAL
- IS64-C- 913 FITZWATER DR MCFARLAND DE
TASK 64-A CONSEQUENT PROCEDURE REAL TIME COMPUTER LANGUAGE
- IS64-C- 940 SAGE CR ANDERSON RR CAMERON CH
AMES LABORATORY SDI SYSTEMS REFERENCE MANUAL
- IS64-E- 562 BROWN WF FISHER RW BLACK HM
DESIGN, CONSTRUCTION AND PERFORMANCE OF TWO COMPACT AIR
COOLED SODIUM HEAT EXCHANGERS
- IS64-E- 567 LONG BK SMUTZ M OLSON EH
GROWING OF BISMUTH SINGLE CRYSTALS
- IS64-P- 863 DOROTHY RG LYNCH DW
PRELIMINARY STUDY OF THE OPTICAL CONSTANTS OF
SILVER-CADMIUM ALLOYS NEAR 300 M MU
- IS64-P- 867 JEW TT LEGVOLD S
MAGNETIC ANISOTROPY IN DYSPROSIUM SINGLE CRYSTALS
- IS64-P- 870 SWENSON CA
THERMAL CONTRIBUTIONS TO THE EQUATION OF STATE OF SOLIDS
- IS64-P- 871 MURAO T SPEDDING FH GOOD RH
THEORY OF ZEEMAN EFFECT FOR RARE-EARTH IONS IN CRYSTAL
FIELD WITH C(SUB 3H) SYMMETRY
- PU64-C- 934 VOSSOS P FITZWATER DR RUNDLE RE
THE CRYSTAL STRUCTURE OF LICUCL3.2H2O
ACTA CRYST. 10, 1037 (1963)
- PU64-C- 974 CORBETT JD
THE SOLUTION OF METALS IN THEIR MOLTEN SALTS
FUSED SALTS (MCGRAW-HILL BOOK CO., INC. NEW YORK, 1964)
B. R. SUNDHEIM, ED., CHAP. 6, PP. 341-407.

Fig. 5. Bibliography Index listings

ANDERSON RR	1564-C-940	DANIELSON GC	900-P-33A	OSCHNEIDER KA	900-M-47B
BAGLIN JE	900-P-58A		900-P-32B		900-M-74A
	900-P-60	DEVRIES DB	900-P-32A		900-M-80B
	900-P-62A	DUKE FR	900-P-31B	GUNTHER RC	900-M-81
	900-P-62B		900-P-31A	HAMMER CL	1564-C-883
BANKS CV	900-P-63A	EDWARDS K	1564-C-86A		1564-P-676
	1564-C-737	ENGEBRETSON GR	1564-C-742		900-P-45A
	1564-C-777	ESPENSON JM	1564-C-665		900-P-50A
	1564-C-800		1564-C-667		900-P-51B
	1564-C-827		1564-C-651		900-P-52
	900-C-36		1564-C-62B		900-P-53A
	900-C-38		900-C-33A		900-P-53B
	900-C-40A	FAHLEN RW	900-C-32		900-P-53C
	900-C-41		900-C-31		900-P-54
	900-C-42		900-C-30B		900-P-55A
	900-C-44A		900-E-23A		900-P-57
	900-C-44B		900-E-23B		900-P-59B
	900-C-45		900-E-24	HANSEN RS	900-P-59C
	900-C-46A		900-E-25		1564-C-855
	900-C-46B		900-E-26A		900-C-25A
	900-C-47		900-E-26B		900-C-25B
	900-C-48A		1564-E-73A		900-C-26A
BARNES RG	900-P-17A	FASSEL VA	1564-E-73B	HATCH EN	900-C-26B
	900-P-16B		1564-E-816		900-C-27A
	900-P-16A		1564-E-817		1564-P-575
	900-P-14		1564-E-830		900-P-63B
	900-P-13B		900-C-54A		900-P-64A
	900-P-13A		900-C-54B		900-P-64B
	900-P-12		900-C-55A	HIETBRINK EH	1564-C-721
	900-P-9		900-C-55B		1564-C-769
	900-P-8B		900-C-56	HIOBY WJ	1564-C-693
	900-P-8A		900-C-57	HOPKINS EN	1564-C-74B
	900-P-36A	FERGUSON JM	900-C-58	HUDSON DE	1564-M-662
	900-P-44A	FINCH DR	900-C-59A		900-P-22
BARTELL LS	900-C-34C	FINNEMORE DK	1564-P-777		900-P-23A
	900-C-35A		1564-P-743		900-P-23B
	900-C-35B		900-P-17B	ILOPP EL	900-P-30A
	900-C-76B		900-P-19A	JOHNSTON TR	1564-P-743
	900-C-77		900-P-19B	KAYSER FX	1564-C-765
BECKWITH WF	1564-E-73A		900-P-20A		900-M-66A
BEKELMAN AM	1564-E-73B		900-P-20B	KELLER JM	900-M-82
BENSON JE	1564-C-67B	FISHER RW	900-P-20C	KERNAN WJ	900-P-34C
BLACK HM	1564-E-562		900-E-21A	KING LA	900-P-44B
BROWN BP	1564-E-810		900-E-32B	KINNEY PJ	1564-C-667
BROWN TM	1564-C-741		900-E-33	KLEINER KL	1564-E-72B
BROWN WF	1564-E-562		900-E-34		900-P-42B
BURNER BL	1564-C-739		900-E-36	KONOPIK AE	900-P-43
BURKHART	900-E-18B		900-E-37	LARSEN WL	1564-E-830
BURKHART LE	900-E-14		900-E-38A		1564-M-576
	900-E-16A		900-E-38B		1564-M-683
	900-E-19A		900-E-40		1564-M-639
	900-E-19B		900-E-41		900-M-30A
BURNET G	1564-E-849		900-E-42		900-M-30B
	1564-E-799		900-E-46		900-M-37
	1564-E-810		1564-E-562		900-M-45B
	900-E-5	FITZWATER DR	1564-E-605		900-M-46B
	900-E-8		1564-E-88B		900-M-105
	900-E-11A		1564-C-751	LASSILA KE	900-M-110
			1564-C-769		900-P-50C
			1564-C-893		900-P-58A
BURNET G	900-E-11B	FITZWATER DR	1564-C-913	LASSILA KE	900-P-58B
CANERON CH	900-E-13A		900-S-3	LEDET MM	1564-C-769
CARLSEN FS	1564-C-940		900-S-4		1564-C-876
CARLSON BC	1564-C-842		900-S-6	LEGVOLD S	1564-C-893
	900-P-55B		1564-C-67B		900-P-37B
	900-P-56A	FRAD WA	1564-C-721		900-P-37C
	900-P-56B	FRANDOLIG JE	900-C-78A		900-P-37D
CARLSON ON	900-P-56C	FRANZEN HF	1564-M-722		900-P-38A
	1564-M-745		1564-E-817		900-P-38B
	1564-M-812		900-C-20B		900-P-39A
	900-M-1A		900-C-33B		900-P-39B
	900-M-1B	FRITZ JS	900-C-34A	LONG BK	1564-E-567
	900-M-2A		900-C-34B	LORETAN PA	1564-E-757
	900-M-10B		900-C-48B	LUTMY JA	1564-C-742
	900-M-13A		900-C-49A	LYNCH DM	900-P-34B
	900-M-13B		900-C-49B		900-P-34A
	900-P-33B		900-C-50A		900-P-40C
	900-M-39A		900-C-50B		900-P-41B
	900-M-39B		900-C-51A		900-P-42A
	900-M-40A		900-C-51B		900-P-7B
	900-P-62A	FUCHS R	900-P-35	MACKINTOSH AR	900-P-7A
	900-M-62B		900-P-41A		900-P-36B
	900-M-62C	GERSTEIN BC	900-P-42B	MARCINKOWSKI MJ	900-P-37A
	900-M-64A		900-P-30B		900-M-52A
	900-M-64B		900-C-17A		900-M-67A
	900-M-65		900-C-17B		900-M-83A
	900-M-109B		900-C-18A	MARLOWE MO	1564-E-925
	900-M-109A		900-C-18B	MARTIN DS	1564-C-883
CASH RJ	1564-E-88B		900-C-18C		900-C-63
CHIOTTI P	1564-M-660		900-C-18D		900-C-64
	900-M-2A		900-C-20A		900-C-65A
	900-M-21		900-C-20B		900-C-65B
	900-M-32	GOLD AV	900-C-21A		900-M-68A
	900-M-44		900-P-1		900-C-68B
	900-M-85C		900-P-2	MCCARLEY RE	900-M-5
	900-M-85D		900-P-3		900-M-7
	900-M-89A	GOOD RH	900-P-5A		900-P-17B
	900-M-90B		900-P-40B		900-M-17A
CLYDE DD	1564-C-790		900-P-45A		900-M-40B
COLLINS SB	1564-E-849		900-P-45B		900-P-10B
COOK BC	900-P-60		900-P-50A		900-P-111A
	900-P-62A		900-P-49C		900-P-111B
	900-P-62B		900-P-49B		900-P-112
	900-P-62C		900-P-49A		900-P-114B
	900-P-63A		900-P-48C		1564-C-741
COPELAND JL	1564-C-665		900-P-48B		1564-P-1029
CORBETT JD	1564-C-739		900-P-48A	MCCORKLE WH	900-R-2A
	900-M-69A		900-P-50B		900-R-2B
	900-M-69B		900-P-51A		900-R-2C
	900-C-21C		900-P-65		900-R-8A
	900-C-22A	GRIFFIN JE	900-P-60		900-R-8B
	900-C-22B		900-P-62A		900-R-12
CORBETT JR	900-C-23A		900-P-62B		900-R-13
	900-C-23B		900-P-63A		900-R-14
	900-P-23C		1564-P-676		900-R-15
DANIELSON GC	900-P-24A	GROSSMAN ST	1564-C-855		900-R-17
	900-P-24B	OSCHNEIDER KA	900-R-20		900-M-19
	900-P-25A		900-P-19		900-R-24
	900-P-30B		900-M-49B		900-R-25
	900-P-33C		900-P-50	MC FARLAND DE	1564-C-721
	900-P-33B		900-M-51		1564-C-769

Fig. 6. Author Index listing.

[illegible]

Fig. 7. Stop Word list.

meaning to the title and so the small integers are also placed in the stop word list. However, an index of contract numbers, job descriptions, or section numbers of groups within a large installation would require removal of numbers from the stop word list.

The stop word list was developed to its present size after it was decided to include a KEY WORD INDEX SECTION in the Ames Laboratory annual progress report. The 0500 annual report served as test input and from it the majority of the stop words were derived. The size of the stop word list has not increased as rapidly as has the input. Some little-used words are presently not included in the list. It was determined after several word frequency tests that the little-used words, when they appear, do not increase the size of the index by more than 5%. Such an insignificant increase adds only a few additional sorted index terms and does not make the system bound by an inexhaustable list of words.

The removal of many little-used words is in keeping with the idea that the Permuted Title Index is used for more than the annual reports. Individual scientists have a very broad base for a personal file index. The use of many insignificant words in the stop list would restrict the individual file index with many delimitations.

VIII. CONCLUSION

The principal purpose in adapting a Permuted Title Index System for the Ames Laboratory was to provide a key word index to the annual

progress reports. This index section facilitates locating information and publications in a current fiscal year as well as providing base input for future cumulative indexes. A cumulative index is in the planning and preparation stage. However all input information is being stored for the cumulative index and for information retrieval purposes.

An additional application of the adaptive Permuted Title Index is in providing personal file systems for individual scientists. Such a personal file allows the scientist to merge new items with an existing file thus minimizing maintenance work. The scientist is thus provided with a very simple and inexpensive index which may be used for storage and retrieval. All file items are automatically cross-referenced under their key words, thus allowing simple and flexible access to the stored items.

REFERENCES

1. General Information Manual, Key-Word-In-Context, E20-8091.
2. USAEC Rept. IS-500, Ames Laboratory Annual Summary Report, Supplement (1962).
3. USAEC Rept. IS-700, Ames Laboratory Annual Summary Report (1963).
4. USAEC Rept. IS-900, Ames Laboratory Annual Summary Report (1964).
5. C. M. Lobron and D. H. Myers, IBM 1401 KWIC Index System.
6. Reference Manual, IBM 7070/7074 Generalized Sorting Program, Sort 90, C28-G111-1, IBM Corp. Data Processing Division, White Plains, N. Y.
7. D. R. Fitzwater et. al., USAEC Rept. IS-769, POGO Reference Manual (1963).