An Index to Ordnance Reports*

By ALFRED M. ANZALONE
Plastics Technical Evaluation Center, Picatinny Arsenal, Dover, New Jersey
and
GUNTHER COHN
The Franklin Institute, Philadelphia 3, Pennsylvania
Received September 27, 1962

Ordnance researchers require that published information on initiators and initiating compositions be collected and stored in a single file available to all Government agencies, their contractors, and suppliers. As spokesman for these Ordnance workers, the Explosive Components Subcommittee of the Joint Army, Navy, and Air Force Fuze Committee requested the Technical Information Section of Picatinny Arsenal to set up a file of technical data. It was intended that this file would be a preliminary step to establishing an Initiator Information Center. This paper discusses the data file which resulted.

Equally important to gathering literature into one collection was the development of a suitable index which would satisfy the following criteria[1]:

1. Accessiblity
2. Capability of including all information which is to be retrievable in the future
3. A language common to both indexer and searcher
4. Document classes which are small enough to be acceptable by the searcher and
5. The presence of all document classes required by the searcher, either specifically or as obvious members of more generic classes which are themselves small enough to be acceptable for search.

The Technical Information Section had previous experience in compiling and using coordinate uniterm indexes. The uniterm index of Picatinny Arsenal Technical Reports is a compilation of some 2,500 documents with a total of about 5,000 terms compiled in two separate but identical sets, both bound side by side under one cover. This arrangement permits coordination of terms such as silver and azide.

Some noticeable deficiencies of this report index were corrected when we developed the initiator index. First, uniterms were precoordinated where prior experience indicated that they were used together. A high frequency of postings for any given term would negate the intended ease of use.

It was necessary, then, to devise an economical and expeditious way of editing these terms. Manually sorting through the single aspect cards, counting the postings, and making changes, seemed to be laborious and costly. At that point, The Franklin Institute, having experience with information systems such as a Fuze Catalog,[2] reviewed the problem. To determine whether the desirable features of uniterms could be salvaged from those which would limit its use in this project, a survey was conducted by a team of Institute specialists from the Laboratories for Research and Development and from the Computing Center.[3]

A computer was used advantageously to arrange the terms and to print out a rough draft to permit editing of those terms having large numbers of postings, synonyms, mistakes, etc. A further advantage was that all noted corrections could be made by machine techniques and the final version printed in a form suitable for photo-offset reproduction.

A sample of the final data file is shown in Fig. 1. The file consists of two parts, abstract and index. The abstracts are copies of citation and abstract cards while the index is a machine compilation of coordinate uniterms.

Citations and abstracts were prepared for each of 990 documents. Citations are in a standard form and abstracts are informative but short, having an average of 120 words each. When author abstracts were written, we used them as guides but usually expanded them. In case of contractor's reports, the monthlies were lumped with the final to make one document. Final citations and abstracts were typed on the front of preprinted 5 × 8 cards (Fig. 2). Run-on cards were occasionally needed for abstracts of greater length. Cards were alphabetized by corporate author and serially numbered. If we inadvertently omitted an entry, we assigned an A number (example: 865, 865A, 866). Ultimately, the numbers are posted with their uniterms.

The coordinate uniterms associated with each document were assigned by the abstracter and typed on the back of the first abstract card pertaining to each document. No master vocabulary list was used but we attempted to be fairly consistent. While this task was simplified by the fact that the bulk of the terms were assigned by one man, it turned out later that we wished we had been more consistent. We assigned an average of 15 terms per document. Citation terms (such as author) accounted for about four of these while the rest were precoordinated subject uniterms. Term length was not controlled, and varied from short words, like *cap*, *stab*, and *wax*, to the longest one in the set, *lead styphnate/barium nitrate/antimony sulphide/tetracene mixture*.

These coordinate terms were alphabetized by Univac. Of the several ways of preparing the text, in a form the machine could digest, namely to unitype magnetic tape, to flexowrite paper tape, or to keypunch cards, we selected the third. Even though punched cards required the extra

* Presented before the Division of Chemical Literature, 142nd National Meeting of the American Chemical Society, Atlantic City, N. J., Sept. 12, 1962.

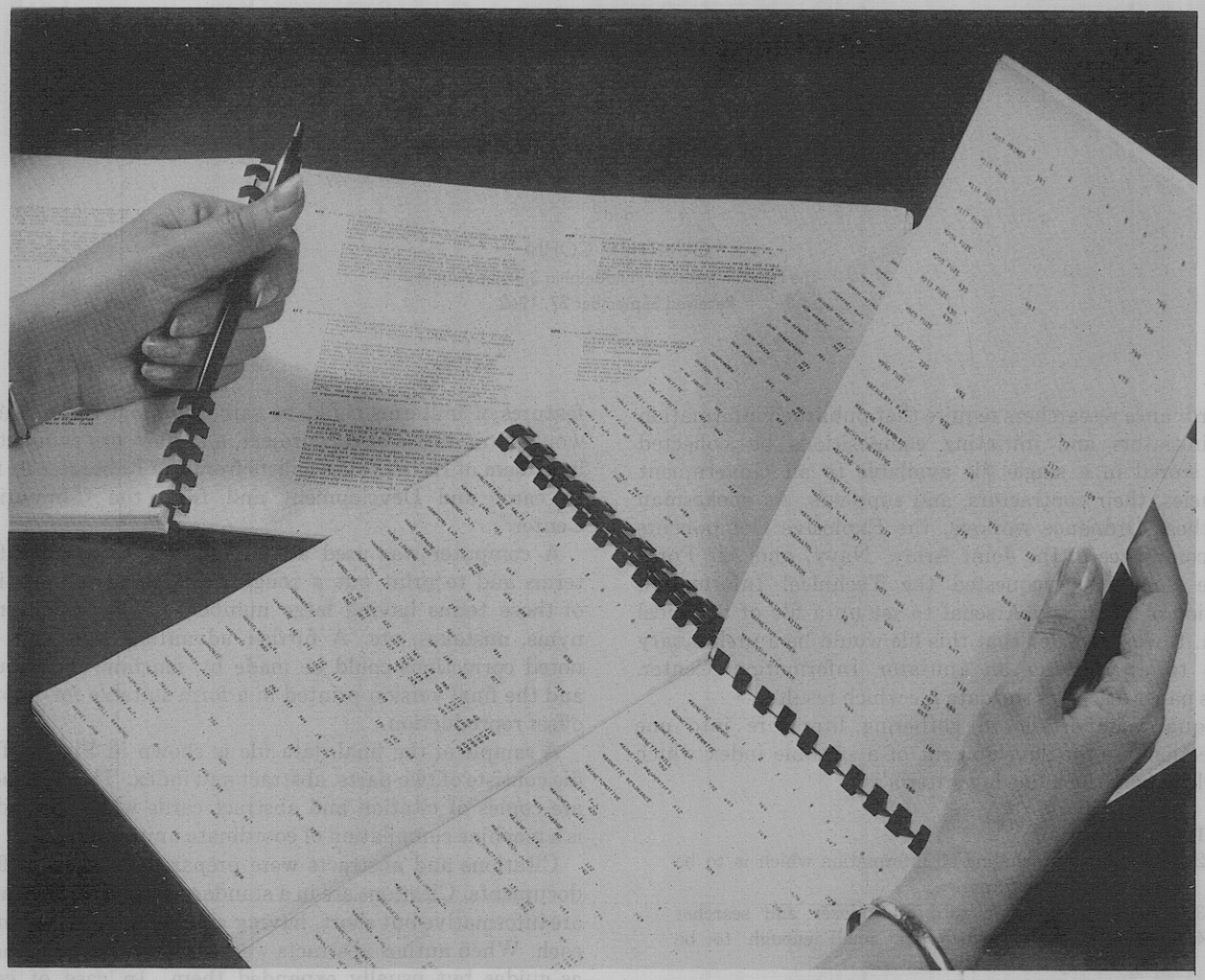

Fig. 1.—The index itself is prepared in duplicate to permit coordination of terms. Other volumes contain citations and abstracts.

step of card to tape conversion, they were the most expedient. Each coordinate term was keypunched on one IBM card together with its serial number, and the card set was converted to magnetic tape to form the Univac input.

The central computer, using an existing routine, alphabetized the terms and arranged them in numerical order within each term. We called this output file *A*. With another existing routine, Univac also dropped the repeated listings and added the number of times each term was used. We called this frequency file *B*. We then printed out files *A* and *B*.

420 Picatinny Arsenal.

SURVEILLANCE STUDY OF NICKEL-ZIRCONIUM TYPE DELAY POWDER FOR M205 HAND GRENADE FUZE, by M. T. Hedges and T. J. Mahler. 17 August 1953. PA technical report 1952. Unclassified report.

Surveillance tests on M205 hand grenade fuzes were made to determine the effect of the age of Nickel powder on the burning time of nickel-zirconium delay compositions. The fuzes were loaded with nickel which had been aged for 0, 5 and 11 months. No correlation between age of nickel and increase in burning time of the composition after storage was observed. (vis)

Fig. 2.—Citations and abstracts are typed on 5 × 8 manila stock. Preprinted guide lines, shown here, are not reproduced in final copy.

File *B*, with a total of 5,605 terms, was found to have an average frequency of only 2.9 postings per term. Although five or six terms had a frequency of over 100 at one extreme, the bulk of the entries contained a single posting. Many items with minor differences were listed separately by Univac. To the machine, *20 mm* and *20mm* are different. Another example is illustrated in Fig. 3. By contrast, typographical and spelling errors were few in number.

The next step was to select a style and change all entries which differed from it. Our style rules are listed in Table I. Most of these rules conform to standard indexing conventions while others were selected to assure the uniformity demanded by the machine.

The preliminary output file could be corrected by hand or by machine. A hand change would require marking up file *B* and then punching a new IBM card for each change that would list item location and new version. A routine exists which instructs Univac to find the old item and replace it with the new one. A machine change, on the other hand, would involve preparing detailed style instructions that the computer would follow. For example, Univac could be told to look for 1-, 2-, or 3-digit numbers

```
TEMPERATURE & HUMIDITY TEST                1
TEMPERATURE AND HUMIDITY TEST              9
TEMPERATURE AND HUMIDITY TESTS             1
TEMPERATURE AND HUMIDITY CYCLING TEST      1
TEMPERATURE CYCLING                        2
TEMPERATURE CYCLING TEST                   1
TEMPERATURE CYCLING TESTS                  1
TEMPERATURE SENSITIVITY TEST               1
TEMPERATURE-HUMIDITY TEST                  7
TEMPERATURE-HUMIDITY TESTS                 1
```

a. Original Frequency File *B*

```
TEMPERATURE & HUMIDITY TEST               25
```

b. Final Frequency File

```
102   TEMPERATURE & HUMIDITY TEST
107   TEMPERATURE & HUMIDITY TEST
258   TEMPERATURE & HUMIDITY TEST
 .
 .
 .
955   TEMPERATURE & HUMIDITY TEST
```

c. Final Output File

```
TEMPERATURE & HUMIDITY TEST
          720   491   102   305   414   465   476   107   258   849
          130   621   622   733   884   475   546         628
                731   952   483         865A  646
                            943         955
```

d. Final File Arranged for Print-Out

Fig. 3.—Index terms are edited to eliminate duplicate or synonymous terms. Postings are then arranged in columns by their terminal digits.

followed by the unit *mm* and to insert a space between number and unit if one is not already there. Making up precise style corrections is a formidable task.

In this instance, hand corrections were indicated. Not only was the number of style rules large, but we also wanted to make changes which were not covered by rules, such as correcting spelling and adding *see* and *see also* references. In addition, we wanted to examine all entries which had a frequency greater than 25. References in such groups might be defined more precisely to achieve a smaller, more useful group. File *B* was therefore carefully reviewed. In cases of frequencies above 25, we referred to file *A* and marked changes there.

Table I
General Style Rules

1. Terms are singular except for collective nouns.
2. Terms are in the common, non-inverted form (examples: Hand grenade, Firing time), except that classes of certain items are inverted to keep the group together (examples: detonator electric, detonator stab).
3. Combined terms use &, while chemicals use / (examples: Temperature & humidity test, Nickel/chromium wire).
4. First symbol is always a letter, never a number (examples: Ammunition 37 mm, Primer 20 mm).
5. In military nomenclature, designation comes before name (examples: B36 Bomber, T 905 Fuze).
6. In designations, numbers are spaced to fall into their correct position (example: T 7 so it will line up with T2055).
7. Words are not abbreviated except initials, designations, and units.
8. Abbreviated designations and units are not punctuated (examples: FA Mixture 70, delay 1 sec).
9. For proper names, last name is listed first, then initials after comma.
10. No space is used when - serves as hyphen but a space comes first when used as a dash (examples, Navy Project N-21-123, Clip - see Fastener).

IBM cards were then punched for the changes. About 4,000 cards were needed or one-fourth the total number in the set. Changes to file *B* require only one IBM card per new term regardless of frequency, but changes in output file *A* demand a card for each entry to be changed. After corrections were inserted, the file was re-alphabetized. The file now contained 3,816 terms with an average frequency of 3.9.

Our most challenging problem was that of incorrectly arranged contract numbers. This curious problem arose because Univac, primarily a computer, has certain columns (the first in each 12-digit word) reserved for algebraic signs. A typical Army Ordnance contract number has a series of letter-and-number groups separated by hyphens. It so happened that one of the hyphens fell occasionally into a sign column. The machine recognized the hyphen as a minus sign and proceeded to arrange all following numbers by the convention of negative numbers, largest one first. Rearranging the numbers in absolute order required a special routine; the offending numbers were fed into the machine, displaced one digit to the right, arranged in this position, and then shifted back.

With uniformity problems behind us, we prepared for final layout. The alphabetized coordinate terms were to be listed with their serial numbers arranged by terminal digits. Univac instructions for page layout are shown in Table II. We wrote a program to have the computer follow these instructions and to control layout by means of non-printing symbols. The great amount of detail required is characteristic of machine operations where all steps must be defined carefully.

Print-out was by means of the high-speed printer. It prints 26 letters, 10 numbers, and these 15 symbols: - & (, . ; #) $ * : + / %. The symbols are shown in the order in which they would be printed if Univac were to list them in sequence. Actually, symbols start with a space. After the hyphen, they are interrupted by numbers and letters arranged in a pattern that turns out to be convenient for digital notation. This is the arrangement we get when we ask for alphabetization. The machine follows the "nothing before something" rule of alphabetization be-

Table II
Detailed Rules for Index Layout

1. A page has a maximum of 74 lines.
2. Serial numbers under a coordinate term are not carried over to a new page. If rest of page cannot hold the entire group, it is left blank and term starts on next page.
3. A coordinate term is on a line by itself. Serial numbers, single spaced, start on the next line. A line is skipped before the next coordinate term.
4. A coordinate term has a maximum of 73 spaces. If a term should be longer, it is cut off at that point.
5. The first line on each page consists of numbers 0 to 9 in the unit places of columns.
6. The first column of serial numbers is indented 10 spaces. Thereafter a triple space is left between columns.
7. Columns are three spaces wide. Serial numbers are printed flush right in the column except that capital letters are printed in the space following the unit digit.
8. The last digit of the serial number assigns the column.
9. Vertically, numbers in each column are arranged in increasing order. Horizontally, numbers are "closed up" so that the first number in each column is in the first line, the next in the second line, etc.

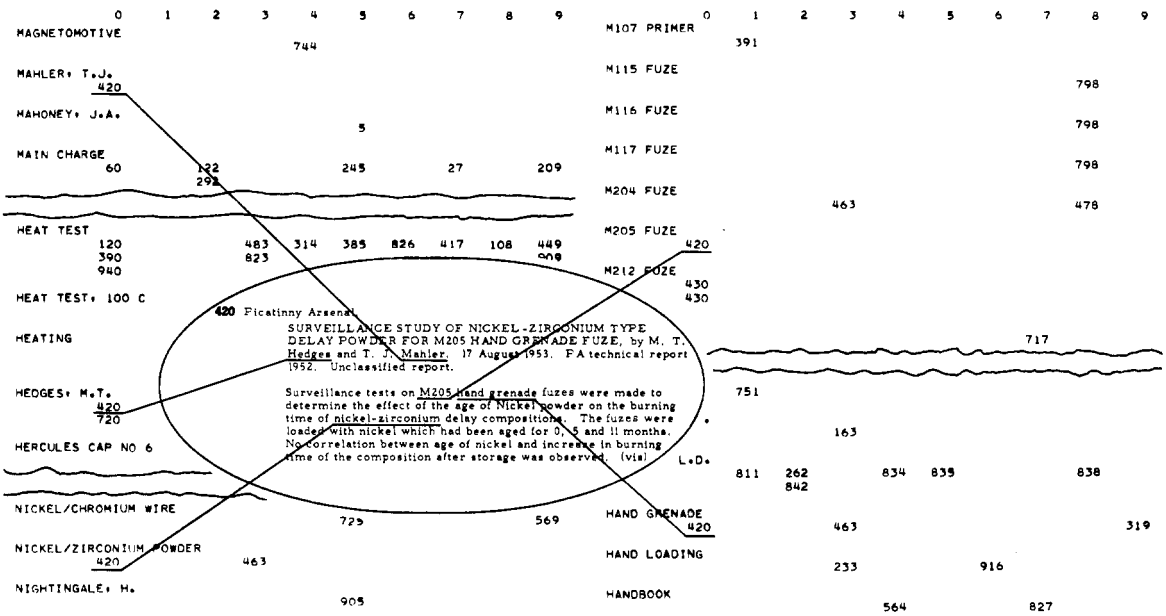

Fig. 4.—Any number of terms, two at a time, can be coordinated by comparing postings from both index sets. This step coupled with a glance at the abstract permits evaluation and selective choosing of indicated references.

Table III
Breakdown of Machine Costs (Dollars)

	Programmer and method analyst labor	Key-punch and unitype labor	Card to tape converter	Univac computer	High-speed printer	Total
			Hourly rate			
	7	3	10	120	15	
Initial List	375	456	10	300	12	1153
Changes	1790	222	4	1080	24	3120
Final Index	450	...	..	578	49	1077
Total	2615	678	14	1958	85	5350

cause the space heads the alphabet. This rule, accepted as standard by libraries and directories, differs from the "letter by letter" rule used in some dictionaries. With the "nothing before something" rule, *Acoustic Mine* comes before *Acoustical*, the space after *Acoustic* coming before the letter *a*. The order would be reversed with the "letter by letter" rule because *a* precedes *M*.

A sample of the final print-out is shown in Fig. 3. Each sheet was reduced photographically to half-page size and reproduced by the photo-offset process. Resulting letter size was 6-point. Two identical sets were mounted side by side in cardboard covers and bound at the top with flexible binding so that the book would lie flat when opened. The coordinate index made up one volume of the three-volume set. The companion volumes contained abstracts arranged six on a page. Abstracts, photocopied from the cards, were divided into two volumes to permit segregating the ones which carry a security classification. Use of the index is illustrated in Fig. 4.

Detailed costs of our machine indexing are shown in Table III. Total costs were as follows:

Abstracting and assigning coordinate terms	$12,680
Machine indexing	5,350
Reproduction (150 copies, 435 pages)	2,300
Total	$20,330

The total cost works out to $20.50 per document which is low in comparison with previous searches made by the Information Section of Picatinny Arsenal.

This paper is based on the report entitled "Initiators and Initiating Compositions: A Literature Search," published as Technical Report 14 of Feltman Research Laboratories of Picatinny Arsenal and dated September, 1960. Volume I contains unclassified citations and abstracts, Volume II lists classified citations and abstracts, while Volume III contains the coordinate index. Qualified Government agencies and their contractors may obtain the three volumes from ASTIA.[4]

REFERENCES

(1) W. A. Wildhack and J. Stern, "The Peek-A-Boo System-Optical Coincidence Subject Cards in Information Searching," in "Punched Cards-Their Applications to Science and Industry," 2nd Ed., R. S. Casey, J. W. Perry, M. M. Berry, and A. Kent, Ed., Reinhold Publishing Corp., New York, N. Y., 1958.
(2) G. Cohn, "Technical Data on Border-Punched Cards," *Am. Doc.*, **X**, 2, 116 (1959).
(3) Contract DA-36-034-ORD-3162.
(4) Armed Services Technical Information Agency, Arlington Hall Station, Arlington 12, Va.: Vol. I (U) AD-244 064; Vol. II (S) AD-319 506; Vol. III (U) AD-244 065.