

THE PREPARATION OF PRINTED INDEXES BY AUTOMATIC PUNCHED-CARD TECHNIQUES*

EUGENE GARFIELD**

The author has previously differentiated between the three basic facets involved in the mechanization of documentation.¹ This paper discusses only one phase of the mechanization process — the utilization of machines to facilitate the compilation of conventional type indexes. It will be shown how automatic IBM punched-card equipment was experimentally employed to advantage in the preparation of a specific scientific index. This work was completed at the Johns Hopkins University Medical Indexing Project. The problem there was, given a specific indexing publication, the *CURRENT LIST OF MEDICAL LITERATURE*, could a punched-card machine system be devised that would facilitate its publication? The term facilitate is used to include the problems of economy in time as well as money.

The production of any large index is a complicated task. In the particular case of the *CURRENT LIST OF MEDICAL LITERATURE* there were specific operations that were especially troublesome and, as will be seen, were particularly susceptible to mechanization. In order to make the necessary parallels the methods and format of the *CURRENT LIST* will be briefly described.

The Present Method

The *CURRENT LIST OF MEDICAL LITERATURE* is a monthly publication of the Armed Forces Medical Library under the editorship of Mr. Seymour Taine. At the present time the *CURRENT LIST* indexes over 100,000

periodical articles and reports of medical interest per year. Each monthly issue is an index to 10,000 such items and current issues are only about two months behind receipt of journals, which justifies its title. Every six months there is a semi-annual cumulation of the author and subject indexes.

The *CURRENT LIST* is divided into three sections: Register of Articles, which contains the tables of contents of about 1400 journals each year, the Author index, and the Subject Index. Indirect index entries are used similar to those found in the indexes to *CHEMICAL ABSTRACTS* and *BIOLOGICAL ABSTRACTS*. In the indexes one is referred to a register number rather than a page number, as is shown in figure 1. At the *CURRENT LIST* offices in Washington periodicals are routed to indexers who examine the articles and prepare for each a 'data sheet' or more accurately a 'typist's information sheet.' (See figure 2.) The indexer selects the appropriate subject headings as well as modifying statements called 'modifications' that enable the index user to decide more easily which articles are pertinent to his needs. At this stage of indexing a Subject Heading Authority List is used and is more important for proper control of the operation. The subject heading list tells the indexer which headings may be used and provides hints as to what other headings may be useful to the reader. Beyond this actual indexing or subject analysis, preparing the index is, at present, a huge task of typing individual 3 x 5 slips, proofreading, numbering, alphabetizing, editing, etc.

*Based on a talk "Punched-cards in the Documentation of Technology" presented at the Annual Library Conference of the Case Institute of Technology, and the School of Library Science, Western Reserve University, Cleveland, Ohio, November 17, 1953.

**Documentation Consultant, 1530 Spring Garden St., Philadelphia 1, Pa.

¹"Mechanical Indexing, machine indexes, and the preparation of indexes by machine methods." Paper presented at the Symposium on Machine Techniques in Scientific Documentation, March 3, 1953, Johns Hopkins University, Baltimore, Md.

Ann. med. int. fenn. 40:1 1951

1-27

GASTROENTERITIS, blood in
calcium in inf.

CALCIUM, in blood
in gastroenteritis in inf.

28-36

NEUROSYPHILIS, complications
CSF contamination by blood in exam.
CEREBROSPINAL FLUID
contamination by blood in neurosyphilis exam.

Fig. 2. Typist's Information Sheets
Before Coding

The Punched-Card Equipment

It will be instructive to review briefly the punched-card equipment. Sorters (fig. 3) are

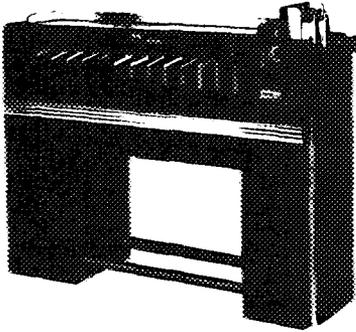


Fig. 3. Sorter

The IBM ELECTRIC PUNCHED CARD SORTING MACHINE automatically arranges punched cards in numerical or alphabetical sequence according to any classification punched in the cards. It provides a fast, automatic machine process of classification for the preparation of various reports and records—all originating from the same cards, but requiring a different sequence or grouping of information.

²For a more complete description of the basic principles of punched-card machine operations cf. Eugene Garfield, "The Preparation of Subject Heading Lists by Automatic Punched-card Techniques," *J. Documentation* 10:1-10 (1954).

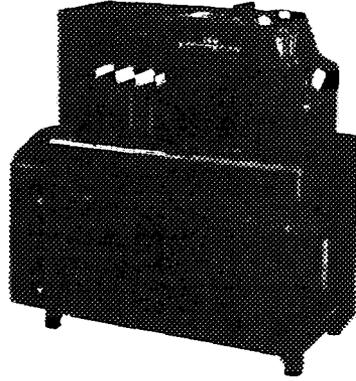


Fig. 4. Collator

The IBM Alphabetical Collator is designed primarily for manufacturers, distributors, and other organizations desiring to retain the various types of code numbers or other identifying information assigned to materials by their original suppliers.

This electronic unit will interfile, compare, select, and verify the sequence of IBM cards punched with alphabetical or numerical data, alone or in combination, and with interspersed special characters where needed. Operating at a speed of 240 cards a minute, this machine collates, in one operation, the IBM cards punched with controlling data of varying lengths up to 16 or 19 characters.

machines which distribute cards to one of several pockets according to which hole position is punched in the column. Sorting is controlled by a wire brush which makes contact with a metal roller. The card passes between the brush and roller—if a hole is present a circuit is completed, directing the card to the correct pocket. If the cards are passed through the sorter once for each column punched, it is possible to arrange them in consecutive order by number or letter, since letters are coded as pairs of numbers.² The collator (fig. 4) can read several

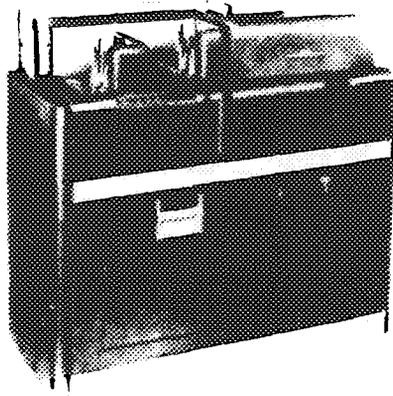


Fig. 5. Reproducer

The ACCUMULATING REPRODUCER, Type 528, combines in one unit functions of punched card accounting previously requiring a number of machines. Not only does it accumulate totals at the rate of 200 cards a minute and punch both totals and indicative information, such as account numbers, into cards but it verifies the accuracy of the total punching. This machine also can reproduce selected cards from a file, at the rate of 100 a minute, without first having to sort the cards from the file. This eliminates the need for maintaining duplicate files of the same data arranged in varying sequence.

cards simultaneously permitting two files to be merged or matched according to instructions provided in the control panel on the side. Reproducing machines (fig. 5) will duplicate a set of cards completely or in part. In tabulators (fig. 6) punched-cards activate fast moving type bars or wheels producing one or more lines of print for each card. Key-punches (fig. 7) are machines that prepare the cards originally and verifiers are machines which are used to edit them for errors, i.e. proofread. [The verifier, not shown, has the same physical appearance as the Key-punch.] Both the key-punch and the verifier are operated like a typewriter, whereas in the other machines an operator places in the

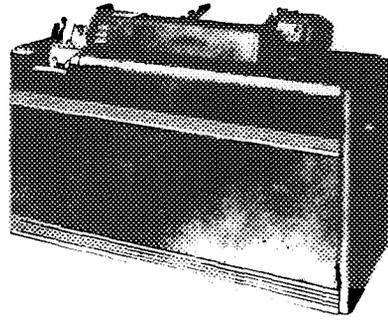


Fig. 6. Tabulator

This IBM ELECTRIC PUNCHED CARD ACCOUNTING MACHINE, Type 407, with wheel printing lists 150 lines of information a minute from IBM cards or accumulates information from IBM cards at 150 cards a minute.

A type wheel in each of the 120 printing positions can print 47 characters including all letters, numerals, and 11 special characters. Alphabetical as well as numerical information can be printed in any position on a report form; dollar signs, commas, and periods are printed where needed in amounts; asterisks are printed to the left of significant amounts for check protection. The 10 characters horizontally to the inch provided by the machine enable reduction in the overall size of forms and savings in forms costs.

Additional features of this accounting machine are a new card feed which holds a card stationary and allows it to be read repeatedly, provision for storing alphabetical and numerical information in the machine until needed, 112 net balance counters, increased selector capacity, ability to print and cross-add totals in any desired sequence, and inclusion of summary punch writing on the accounting machine control panel.

cards 500 to 1000 at a time. The interpreter (fig. 8) prints information on the face of the card, when it is desirable to read the information punched in each card.

The Machine Method

In the punched-card system for preparing the
CURRENT LIST OF MEDICAL LITERATURE

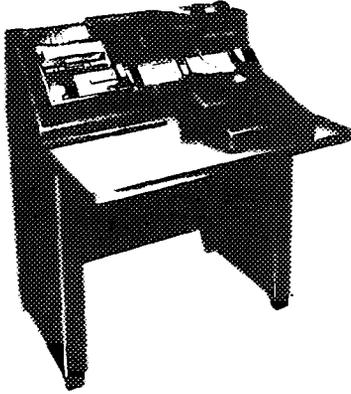


Fig. 7. Key-punch

The IBM ELECTRIC CARD PUNCH combines efficiency and speed with simplicity and ease of operation. Among its many features are: automatic card control of programming — an entirely new method of controlling skipping and duplicating which eliminates skip bars and tabular inserts; a simple, fast method of duplication in which information is duplicated into each card in a duplicating rack; a new design which permits more efficient and rapid operation, and a newly-designed, movable keyboard with a light touch and comfortably level keys.

This card punch is made in a non-printing model, and in a printing model for simultaneous interpreting of the information punched in the cards. It is available either for punching numerical data only, or for punching both numerical and alphabetical information.

the step following the preparation of 'data sheets' is coding. (see fig. 9). The code numbers are obtained from the Subject Heading Authority List. Coding could be combined with indexing, but it is inefficient to waste the time of skilled indexing personnel for coding. Data sheets are sent to a key-punch operator who prepares the necessary punched-cards. In fig. 10, there are shown sample register, author, and modification cards before numbering. Note that the register entry is *typed* on the face of

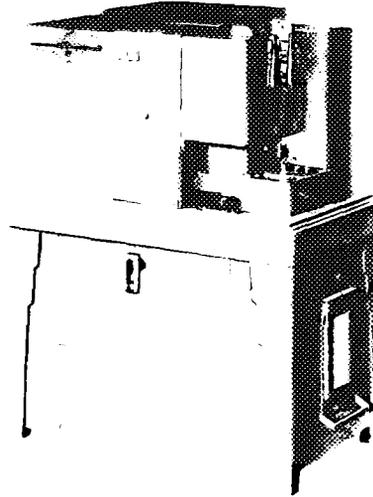


Fig. 8. Interpreter

IBM Electric Punched Card Interpreter translates the holes punched in IBM Cards and prints this data across the top edge of the card in any sequence. The flexibility of this machine permits the printing of account numbers and other information in the order that will best facilitate reference.

Ann. med. int. fenn. 40:1 1951

1-27

34860	31	
GASTROENTERITIS, blood in	(0336)	001
calcium in inf.	Journal	
13540	86	Code Number
CALCIUM, in blood		
in gastroenteritis in inf.		

28-36 (pagination)		002
56750	36	article
NEUROSYPHILIS, complications		number
CSF contamination by blood in exam.		
16690		
CEREBROSPINAL FLUID		
contamination by blood in		
neurosyphilis exam.		

Fig. 9. Information Sheets After Coding

the register cards, since the register card is used only once. There is little advantage to be gained by preparing punched-cards if they are

heading for each article. The seven digits are really a five digit serial number followed by a two digit sub-heading code number. In column

80 of all cards there is a punch which distinguishes the type of card. Thus, all author cards have a five punch in column 80 and modification cards a six punch in column 80.

After cards are prepared on the key-punch they are mechanically proofread on the verifier. The cards are then filed away in groups by journal and article code number. Since these journal numbers were assigned in alphabetical order this gives us an alphabetical arrangement of journals as shown in fig. 1. When a month's indexing is complete, register numbers are added to the cards mechanically, on the reproducing punch, by an interspersed gang-punching operation. In fig. 11 a group of cards

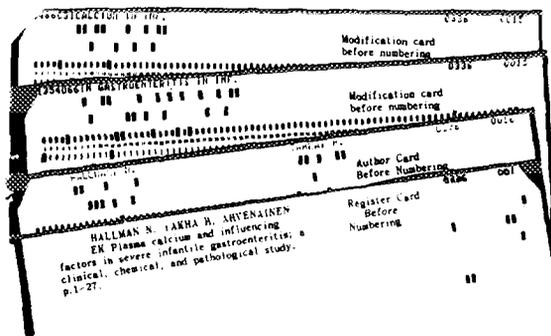


Fig. 10.

not to be used in repetitive operations. However, the register card does contain a journal code

number as well as article number, which was assigned during coding. These two numbers taken together constitute a unique identification tag for each article. Two authors' names have been punched on a single card to save key-punching time, but actual tests showed that it may be just as simple to prepare a separate card for each author during the key-punching. Note also the journal and article code numbers, which are automatically duplicated in all cards during punching. Also notice the serial numbers for subject headings followed by the modification. It is not necessary to punch the entire alphabetic subject

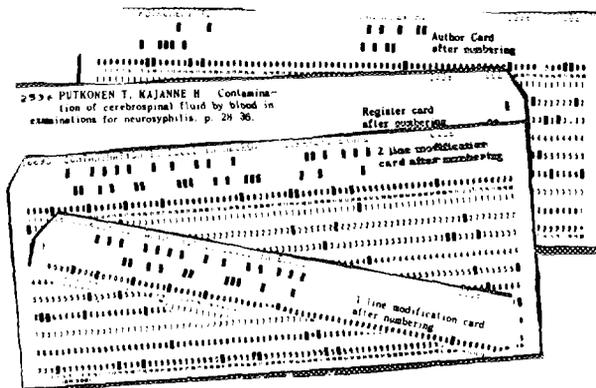


Fig. 11.

is shown after the numbers have been punched in columns 72-76.

CURRENT LIST of MEDICAL LITERATURE

24494 WISNER A, KAPPI R, GORDON E B
 Studies in RA sensitization. V. Imp-
 tance of the site of RA antibodies in the sensitive
 syngeneic RA-negative woman for prognosis
 p. 8-15

24495 LANDRIGAN R, BOOGLAS R G, EYDOR
 B B Rational changes in the toxicosis of
 pregnancy. II. Mild and severe hypertensive
 renal disease, and diabetic mellitus. p. 18-27

24496 MOORE J G, BARDALL J H Trends in
 therapeutic abortion; a review of 197
 cases. p. 28-40

24497 SCHWENK C

24498 UNDELSEDER J L Uterine ectositis
 associated with renal agnosia and
 malformation of the body cavity. p. 200-2

24499 REACMAN D W, REACMAN W D Dys-
 toxia due to fetal anguillulosis. p. 205-8

24500 CARPENTER R J Jr, JAMISON W J
 Uterine ectositis. p. 205-8

24501 MENDELSON C L Cesarean artery
 laceration in pregnancy. p. 201-31

24502 BALTSHERMAN C L

CURRENT LIST of MEDICAL LITERATURE

24503 ALING C A
 24504 BELLING H W
 24505 BELLING H W
 24506 BELLING H W
 24507 BELLING H W
 24508 BELLING H W
 24509 BELLING H W
 24510 BELLING H W
 24511 BELLING H W
 24512 BELLING H W
 24513 BELLING H W
 24514 BELLING H W
 24515 BELLING H W
 24516 BELLING H W
 24517 BELLING H W
 24518 BELLING H W
 24519 BELLING H W
 24520 BELLING H W

24521 BELLING H W
 24522 BELLING H W
 24523 BELLING H W
 24524 BELLING H W
 24525 BELLING H W
 24526 BELLING H W
 24527 BELLING H W
 24528 BELLING H W
 24529 BELLING H W
 24530 BELLING H W
 24531 BELLING H W
 24532 BELLING H W
 24533 BELLING H W
 24534 BELLING H W
 24535 BELLING H W
 24536 BELLING H W
 24537 BELLING H W
 24538 BELLING H W
 24539 BELLING H W
 24540 BELLING H W

24541 BELLING H W
 24542 BELLING H W
 24543 BELLING H W
 24544 BELLING H W
 24545 BELLING H W
 24546 BELLING H W
 24547 BELLING H W
 24548 BELLING H W
 24549 BELLING H W
 24550 BELLING H W
 24551 BELLING H W
 24552 BELLING H W
 24553 BELLING H W
 24554 BELLING H W
 24555 BELLING H W
 24556 BELLING H W
 24557 BELLING H W
 24558 BELLING H W
 24559 BELLING H W
 24560 BELLING H W

24561 COHEN W G The fate of the patient with
 advanced cancer of the cervix of the uterus
 p. 108-15

24562 BALTFREIDER D P Pelvic shape and its
 relation to multiple pregnancies. p. 128-31

24563 JAROD G S An analysis of four hundred
 consecutive cesarean sections. p. 128-8

24564 SCHAEFER G, EYDOR B Results
 following hysteropic abortion in
 tuberculosis. p. 128-32

24565 BOYBAUPE

24566 TURNER J J, ROUSE R E A new cir-
 cularization instrument. p. 220-1

24567 BASH G A Gastric ulcer in fetal twins
 complicating pregnancy. p. 228-8

24568 PETERSON W F A modified outlet
 forceps. p. 228-7

24569 LEVINE W, EURLAND J I Detection and
 treatment of unsuspected tuberculous endo-
 metritis. p. 425-8

24570 REDDALL S S, WEST R E Section of the
 placenta at cesarean section, a case of
 fetal asphyxia. p. 428-7

24571 MOHLER R W, MCCONNELL E L
 Chlorocephalotomy. p. 428-32

24572 BARRY

24573 ABDOU...
 24574 ABDOU...
 24575 ABDOU...
 24576 ABDOU...
 24577 ABDOU...
 24578 ABDOU...
 24579 ABDOU...
 24580 ABDOU...
 24581 ABDOU...
 24582 ABDOU...
 24583 ABDOU...
 24584 ABDOU...
 24585 ABDOU...
 24586 ABDOU...
 24587 ABDOU...
 24588 ABDOU...
 24589 ABDOU...
 24590 ABDOU...
 24591 ABDOU...
 24592 ABDOU...
 24593 ABDOU...
 24594 ABDOU...
 24595 ABDOU...
 24596 ABDOU...
 24597 ABDOU...
 24598 ABDOU...
 24599 ABDOU...
 24600 ABDOU...

24601 ABDOU...
 24602 ABDOU...
 24603 ABDOU...
 24604 ABDOU...
 24605 ABDOU...
 24606 ABDOU...
 24607 ABDOU...
 24608 ABDOU...
 24609 ABDOU...
 24610 ABDOU...
 24611 ABDOU...
 24612 ABDOU...
 24613 ABDOU...
 24614 ABDOU...
 24615 ABDOU...
 24616 ABDOU...
 24617 ABDOU...
 24618 ABDOU...
 24619 ABDOU...
 24620 ABDOU...
 24621 ABDOU...
 24622 ABDOU...
 24623 ABDOU...
 24624 ABDOU...
 24625 ABDOU...
 24626 ABDOU...
 24627 ABDOU...
 24628 ABDOU...
 24629 ABDOU...
 24630 ABDOU...

24631 ABDOU...
 24632 ABDOU...
 24633 ABDOU...
 24634 ABDOU...
 24635 ABDOU...
 24636 ABDOU...
 24637 ABDOU...
 24638 ABDOU...
 24639 ABDOU...
 24640 ABDOU...
 24641 ABDOU...
 24642 ABDOU...
 24643 ABDOU...
 24644 ABDOU...
 24645 ABDOU...
 24646 ABDOU...
 24647 ABDOU...
 24648 ABDOU...
 24649 ABDOU...
 24650 ABDOU...
 24651 ABDOU...
 24652 ABDOU...
 24653 ABDOU...
 24654 ABDOU...
 24655 ABDOU...
 24656 ABDOU...
 24657 ABDOU...
 24658 ABDOU...
 24659 ABDOU...
 24660 ABDOU...

Fig. 12.

After numbering, the cards are separated by type and the register numbers are printed on the face of the register cards as is shown in

page is shown in figure 12. The numbers are not as clear as they might be because the machine used for the experiment was not rented with this operation in mind.

Larger size type is available. The author cards are alphabetized mechanically and placed in a printing tabulator. For each author card there will be a line of print. Printed copy runs out of the machine in one continuous form. The printed copy is pasted up on boards and photographed. (fig. 12) Note the grouping of items under authors' names where more than one article has been indexed. (cf Bentolila & Eltorm). This is done mechanically. The register numbers under each author are also in ascending order because the cards were already in order by register number after numbering.

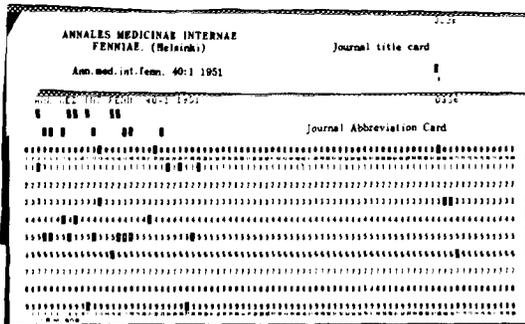


Fig. 13.

fig. 11. The 'two-line modification card' will produce two lines of print.

The register cards are now pasted up on large boards, photographed, and the resulting

remain and pose the most difficult though most interesting problem, i.e., the subject index. Before discussing the subject index there are some additional items of interest, important to that

discussion. In figure 13 there is shown a sample journal title card and journal abbreviation card. These cards can be added in a collator operation prior to mounting the register cards. In fig. 14 a subject heading card is illustrated followed by a see also card. Note that each of these has a subject heading serial number, similar to that shown previously on the modification cards (fig. 10 and 11). In fig. 14 there is also shown a main heading followed by a sub-heading card. These cards have a common serial number, but the sub-heading card also has the two digit code number. Finally there is a see reference card. Before going on, the reader should refer back to figure 1 where the present subject index page is illustrated. It will be seen that

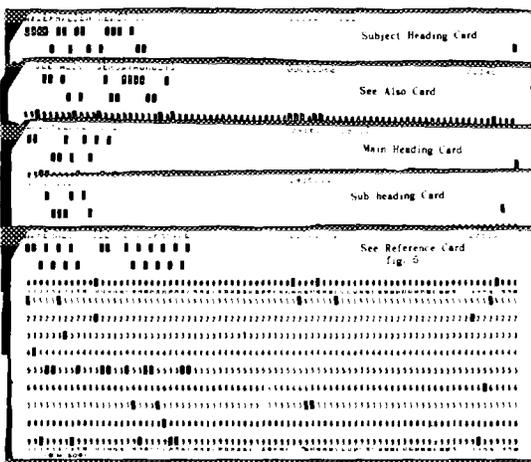


Fig. 14.

subject headings may be followed by cross-references, then under each heading there will be modifications that do not require sub-headings. Immediately after, one finds the indented sub-heading, again followed by additional modifications. Note that modifications as well as sub-headings are in alphabetical order.

The subject index cards are now in two groups. The modification cards, which have just been separated from the author and register cards on the sorter — and the subject heading cards of which little has been said so far. The subject heading cards are used to prepare the subject heading authority list. (For sample page from such a list see (2), page 6.) It is necessary to merge the two separate decks so that we may have the complete subject index. However, it is in this very procedure that we achieve mechanically a reduction in editorial work. If the modification cards are first alphabetized, and then arranged in serial number order, i.e. by subject heading, it is possible to do a matching operation with the complete subject heading deck and select only those main headings which have been used during the month's operations.

Finally, we can supply only the appropriate cross-references by another series of matching operations of subject heading cards against cross-reference cards. This prevents the index user from checking blind references. With the subject index cards finally merged it is possible to run the combined deck through the printing tabulator and obtain the results illustrated in fig. 12. This sample was prepared on an IBM 407 tabulator. Even with the limitation of upper case only, good legibility has been obtained.

In order to prepare the six-month cumulations similar operations would be performed. With the high speeds possible on the automatic equipment such large cumulations can be ready for the printer in a relatively short time.

The 101 Electronic Statistical Machine can

be used to some advantage in these operations, especially in alphabetizing and sequence checking. However, that machine's greatest utility may lie in its searching abilities.³

These detailed operations will seem at first excessively complicated. This is true of any large indexing operation. The analysis of the problem at hand is the first essential in machine methods and in certain instances machines force us to think out a problem to its ultimate. Machines will tolerate no deviations or sloppy thinking on the part of the human. It is also important to recognize that punched-cards, although quite versatile, have many shortcomings. These are not shortcomings inherent in machine methods. Punched-cards are a stepping stone to machine methods in development or yet to be developed. There is little doubt that computers like the UNIVAC or the IBM 702 could handle all of the manipulations described here with less rigid requirements demanded by an 80 column card.

Together with recent developments in printing techniques one cannot help but conclude that mechanical documentation is in sight. However, it is only by impressing the manufacturers with the need for more versatile equipment, designed with indexing problems in mind, that we can expect the best results. Perhaps such developments can help us fill the tremendous gaps that exist in our documentation of science and technology. Until that time, however, punched-card methods can be employed most economically and with favorable results.

Those readers who may be interested in applying these techniques should contact the author for a copy of "The Preparation of Printed Indexes by Automatic Punched-Card Equipment — A Manual of Procedures," Johns Hopkins University Medical Indexing Project, 16 p., 24 March 1953, which provides detailed instructions for punching cards, sorting, wiring, etc.

³Garfield, E. "Preliminary Report on the Mechanical Analysis of Information by use of the 101 Statistical Punched-Card Machine." American Documentation 5:7-12, January 1954.