Mechanization of Library Procedures in the Medium-sized Medical Library:

I. The Serial Record*

By Irwin H. Pizer

Associate Librarian and Research Assistant for Machine Methods
Washington University School of Medicine

Donald R. Franz

Senior Program Analyst
Washington University Computer Center

Estelle Brodman

Librarian and Associate Professor of Medical History
Washington University School of Medicine

St. Louis, Missouri

MECHANICAL means for expediting library work have been considered for many years, as an examination of the early pages of *Library Journal* bears out. The very first volume of the forerunner of the Bulletin, the *Medical Library and Historical Journal,* moreover, contained an article on the use of the "typewriting machine" in cataloging (1). Telephones have long been accepted as library machines, and, except for unusual items, hand bookbinding has given way to machine-bound books.

Nevertheless, when the term "machine methods" is used for libraries today, the prevailing idea is that expensive and complicated electronic equipment must be employed, and, since few libraries are able to afford such equipment for themselves, it is generally assumed that none but the largest libraries (the Library of Congress or the National Library of Medicine, for example) will be able to employ already existing mechanical methods or devise systems specifically applicable to their own situations.

Another stumbling block to the acceptance of machine methods in the majority of libraries is the investment already made in existing catalogs, serial records, order and billing systems, and the like. Even though the traditional systems do not do everything desired of them and even though

* This project was supported in part by funds from U. S. Public Health Service Grant No. SO1-FR-05089.

the manpower necessary to run them is difficult to obtain, the cost of scrapping them and installing a totally new system is so great it is unfeasible. A parallel has been pointed out between this situation and the one that existed when electric lights were introduced. In Edwardian Britain, where gas lines criss-crossed the country and made gas lamps cheap and easy to use anywhere, a substantial time lag occurred in the introduction of electric lines. In the United States, on the other hand, where vast stretches of the country had no access to piped gas, electricity was accepted more quickly and more easily than in Britain. Similarly, a new library today would have a freer hand in deciding to introduce mechanical methods than would a well-established, well-functioning one.

The enormous literature growth which has occurred and continues to occur in the sciences heightens the need for improved library methods. Control of this expanded literature is not an unfamiliar problem to those whose business is information. Much documentation of this problem exists (2), and much additional literature deals with solutions to parts of the problem (3). A solution to the entire problem appears remote; nonetheless, the staff of the Washington University School of Medicine (WUSM) Library felt that such steps should be undertaken as were presently possible.

General Objectives

The Library had two purposes in mind when it began seriously to work on machine methods. First, of course, was the motive of running its own shop efficiently, with all that this implies in the way of personnel and other costs. But it had a second objective.

Many medical libraries are too small and have too few staff to be able to purchase computers or other equipment for rapid storage and retrieval of masses of data or even to pay to rent time from commercial data processing firms. On the other hand, many such libraries in the United States are connected with hospitals or universities where this equipment is available in business offices, accounting departments, physics and engineering departments, and the like. If programs were worked out for the use of this equipment by medical libraries generally, such use would become feasible and many libraries would be able to make the contents of their collections available to users faster, with fewer library personnel, and at a greater depth of specificity than is now possible. In view of the enormous size of medical literature and the inadequate numbers of trained medical librarians, a goal such as this is in the general interest of scientific medicine.

A library as small as this one is easier to work with than one the size of the National Library of Medicine or the New York Academy of Medicine; moreover, it represents a large group of similar institutions. It was felt that methods worked out in an average-sized library could be more easily

adapted to other small libraries than methods from the giants. It was also hoped that the example of one small library experimenting in one segment of the problem would lead other, similar libraries to start experiments on other segments and thus aid all by multiple approaches.

Phases

Three phases for experimentation with machine methods were outlined at the WUSM Library, of which this is a progress report on one section of phase one. The three phases are:

1. The expediting of administrative work in the Library. This includes such matters as accurate determination of receipt of books and journals, correct billing and accounting, record keeping for binding, streamlining work flow, compilation of statistical reports, and preparation of book (as compared to card) catalogs, lists, and indexes for distribution widely throughout the affected community.

2. The meshing of published indexes with the Library's individual indexing schemes to provide a wider conspectus of the literature than is obtainable from either the home library or the literature keys. This includes studies on compatibilities of several systems, such as the National Library of Medicine's MEDLARS project, the BASIC (i.e., permuted title) index of *Biological Abstracts*, or the translations from the Russian in the *Federation Proceedings*.

3. The redividing of material in published indexes into subject groupings pertinent to the particular needs in a particular institution. Published keys to the literature can hope to break down their subject fields only to the level of the most widespread needs; they cannot be expected to do demand subject searches on very specific topics for large numbers of requestors in remote areas. This specificity is the field of local libraries and information centers. An attempt will be made in this phase to work out sample systems of specific coding for specific retrieval which will not re-do the enormous labor of handling material already indexed by other sources nor require the services of such highly trained people that finding and paying them would be difficult for society. It is realized that this is the most difficult phase of the work and may take the longest time to investigate.

Beginnings

In the fall of 1961 the WUSM Library underwent a reorganization. At that time a survey of the strengths and weaknesses of the organization was made; this showed first that, because of shortages and turnover of staff for a number of years previous to 1961, the records of library holdings were not to be relied upon. The second lack found was in fiscal account-

ing—it was difficult to determine how much money was due to which suppliers for what items, so that neither budget estimates nor logical planning for current expenditures was possible. These two items (records and fiscal accounts), together with the physical rehabilitation of the Library quarters, were chosen as areas for the first projects to be undertaken.

It was the consensus of the staff that present records had to be put in order first, because whatever system was adopted for the future would require accurate information. As a result, a physical inventory was taken of all the books and journals in the collection, and the serial-record Kardex was revised for completeness and accuracy. Only at that point was it possible to consider future methods in detail. Entirely new systems could be considered, as well as patching up of older systems, because it was estimated that the latter would be very time consuming and expensive.

It seemed obvious that punched-card equipment could help greatly in the fiscal processes of the Library. Like most such institutions, the Library did not itself have mechanical equipment, but accounting machines were available in the hospital's data processing section. The first approach was to that department, from which the Library was directed to the University's Computer Center.

In October 1961 preliminary arrangements were made with the Center, which was under the direction of Mr. Richard A. Dammkoehler, and a set of specific objectives was drawn up. Among these were publication of the Library's serial record and monograph holdings in book form, automatic check-in of journals and updating of records, compilation of a subject listing of currently received titles, and accounting tabulations for budget and acquisition purposes.

Some of these objectives had to be delayed because of the unavailability of suitable equipment at that time; however, in the succeeding year the accounting records, preliminary editions of the title list of journals currently received, and a subject list of those titles were produced at the Center using unit record equipment (keypunches, sorters, collators, and an accounting machine). The organization of the data, subject coding of the titles, and arduous proofreading and correcting chores were performed by Miss Isabelle T. Anderson, Assistant Librarian for Technical Services.

During the summer of 1962, new equipment (IBM 7072 and 1401 computer systems) was added to the Computer Center, and it was then possible to contemplate logical extensions of the work previously performed and to plan additional steps.

In 1961/62 considerable progress was being made in the field of library data processing in several institutions. At the end of the summer, for example, the University of California at San Diego issued a report on the mechanization of its serial records which took into account the problems

of check-in and updating (4). This program performed its functions in a manner analogous to that proposed at the WUSM Library; however, it used different computers and programming language.

The MEDLARS program of the National Library of Medicine was also proceeding according to schedule, and a report on the completion of Phase One was issued by General Electric (the contractor) in January 1962 (5). Since the WUSM Library had acted as a regional center for medical libraries for some time, it naturally wished to examine any changes in its role or its ability to carry out that role which MEDLARS might introduce. Specifically, it was interested in the possibility that it might participate in this program as one of a network of regional centers selected to receive bulk information stores for rearrangement and searching according to local needs.

At the same time, the report of the University of Illinois at Chicago appeared (6); this also seemed to indicate that other institutions were approaching the same problems from slightly different viewpoints and with different philosophies.

In September 1962, spurred by these progress reports and aware of the vastly increased capabilities of the University's Computer Center, the WUSM Library held further discussions with Mr. Dammkoehler on the feasibility and design of certain experimental procedures and with Dr. Edward W. Dempsey, Dean of the Medical School, who generously assigned to the Library some of the School's General Research Support Grant funds from the U.S. Public Health Service. The helpfulness of Dr. Dempsey and Mr. Dammkoehler cannot be overemphasized.

Mr. Irwin H. Pizer was assigned to the project on a full-time basis beginning in January 1963, and Mr. Donald R. Franz, Senior Program Analyst of the Computer Center, was designated as the Center's liaison, at which time work was begun on the compilation of the entire serial record in machine-readable form.

Output and Frequency

It was hoped that the research would produce several products, each of which would have different applications and functions. Some outputs, for example, would be used within the Library only, others outside it; some would give current awareness, others cumulative information.

In the first phase of the project the output designed for extralibrary use was to be of two types, *PHILSOM* and *PHILSOMS*.

"*PHILSOM*" (***P**eriodical **H**oldings **I**n the **L**ibrary of the **S**chool **O**f **M**edicine*)

The first and most immediately desired output was the creation of a publishable listing of the entire serial record of the Library, giving both

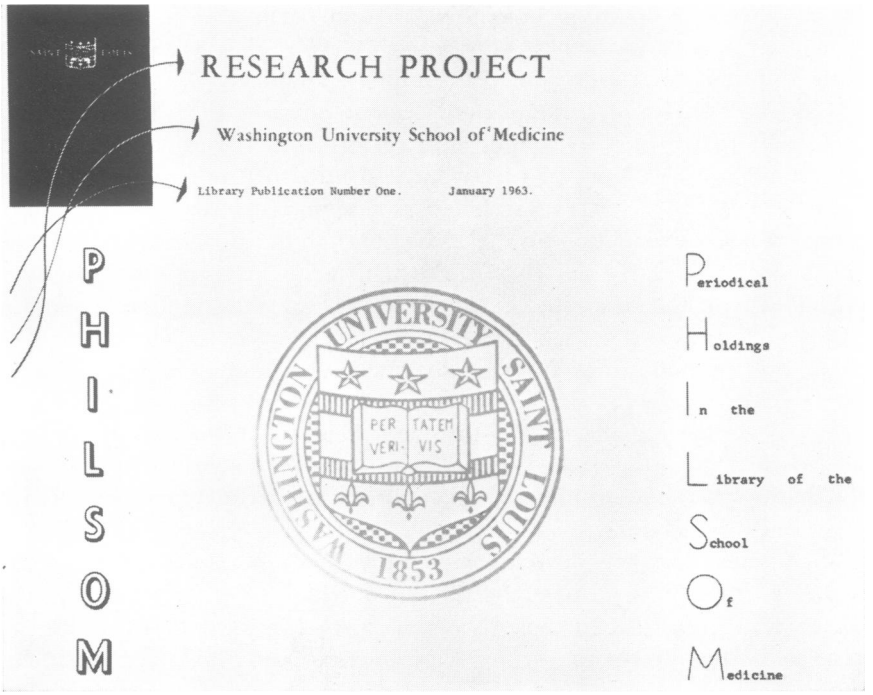

FIG. 1.—Sample of the covers designed for the extralibrary publications of the Research Project. The design was based on that used for other Library publications during 1962/63, and this cover was printed in red. For PHILSOMS, the letter "S" appeared on the left side below the "M," and the cover was printed in blue. The third line on the page was added by mimeographing. The column at the far right is not part of the cover.

current and noncurrent titles, the Library's holdings, and certain additional pertinent information (see figs. 1–2).

*"PHILSOMS" (**P**eriodical **H**oldings **I**n the **L**ibrary of the **S**chool **O**f **M**edicine by **S**ubject)*

In addition, a subject listing of the currently received titles, without their holdings, was also projected (see fig. 3). Since both *PHILSOM* and *PHILSOMS* were extensions of the work performed the year before, it was possible to build on that experience as well as to benefit by the amalgamation of data for those lists. Both of the projected listings could be obtained as by-products of the development of the total system, and their publication could be (and was) effected even before the entire system was operable; i.e., as soon as the basic data had been accumulated and keypunched.

It was finally decided to issue the lists semiannually, as the most practical periodicity from the economic viewpoint. Because of the sheer size of the

DIRECTORY OF CATHOLIC HOSPITALS
TO BEGIN IN 1963
FORMERLY HOSPITAL PROGRESS DIRECTORY ISSUE

DISCUSSIONS OF THE FARADAY SOCIETY SEE
FARADAY SOCIETY DISCUSSIONS

DISEASE A MONTH SEE
D M

• DISEASES OF THE CHEST
1/2N1-3,5-7,9-12/3-42/(1935-1962)

• DISEASES OF THE COLON AND THE RECTUM
2-4/5N1-5/(1959-1962)

• DISEASES OF THE NERVOUS SYSTEM
1-22/23N1-11/(1940-1962)

DISSERTATION ABSTRACTS
TO BEGIN IN 1963

DOKLADY AKADEMII NAUK SSSR
131-136/137N1-3/138N1-5/139N1-2,4-5/140-142/143N1-3/(19
60-1962)

DUBLIN HOSPITAL REPORTS AND COMMUNICATIONS IN MEDICINE AND
SURGERY
1-5/(1817-1830)

DUBLIN JOURNAL OF MEDICAL SCIENCE
1-95/104N308-309/114/1903N384/117/1905N402,406/1908/141
-150/(1832/1833-1921)

E

E E G JOURNAL SEE
ELECTROENCEPHALOGRAPHY AND CLINICAL NEUROPHYSIOLOGY

EDINBURGH HOSPITAL REPORTS
1/(1893)

EDINBURGH MEDICAL AND SURGICAL JOURNAL
1-54,56-57,59-82/(1805-1855)
UNITED WITH MONTHLY JOURNAL OF MEDICINE TO FORM
EDINBURGH MEDICAL J

FIG. 2.—Sample page from PHILSOM, showing the various types of information given in the final listing. An asterisk indicates that a title is indexed in *Index Medicus*. Slashes are used to separate volume information. "N" stands for "issue number," and commas indicate lacks within a volume, or missing volumes. Hyphens indicate inclusive volumes or issues. Also shown are cross-references and history information.

EXCERPTA MEDICA ENDOCRINOLOGY

* GENERAL AND COMPARATIVE ENDOCRINOLOGY

* JOURNAL OF CLINICAL ENDOCRINOLOGY AND METABOLISM

* JOURNAL OF ENDOCRINOLOGY

* RECENT PROGRESS IN HORMONE RESEARCH

* VITAMINS AND HORMONES

YEAR BOOK OF ENDOCRINOLOGY

ENZYMOLOGY SEE BIOCHEMISTRY

EPILEPSY SEE NEUROLOGY

EXPERIMENTAL MEDICINE

* ACTA BIOLOGICA ET MEDICA GERMANICA

* ACTA MEDICA SCANDINAVICA

* ACTA MEDICA SCANDINAVICA SUPPLEMENT

* ACTA SCHOLAE MEDICINALIS

* ANNALES MEDICINAE EXPERIMENTALIS ET BIOLOGIAE FENNIAE

* ANNALES MEDICINAE EXPERIMENTALIS ET BIOLOGIAE FENNIAE SUPPLEMENT

* ANNALS OF BIOCHEMISTRY AND EXPERIMENTAL MEDICINE

* ARCHIVES ROUMAINES DE PATHOLOGIE EXPERIMENTALE ET DE MICROBIOLOGIE

* AUSTRALIAN JOURNAL OF EXPERIMENTAL BIOLOGY AND MEDICAL SCIENCE

* BOLETIN DEL INSTITUTO DE ESTUDIOS MEDICOS Y BIOLOGICOS

* BOLLETTINO DELLA SOCIETA ITALIANA DI BIOLOGIA SPERIMENTALE

BULLETIN OF THE RESEARCH COUNCIL OF ISRAEL SECTION E EXPERIMENTAL MED

FIG. 3.—PHILSOMS, containing no holdings or cross-references from variant titles, is an alphabetical list of titles arranged by subject. Cross-references are given from unused subject headings. The headings are the same as those of the *Index Medicus*, adapted to fit local needs.

publications, it might be necessary to charge for them. Semiannual publication insured that if a fee had to be levied for their production, it could be kept at a nominal level. But up-to-dateness was as desirable a feature as low cost. In order to lessen the period between lists, it was decided to publish them alternately, so that the recipient of the lists had an up-to-date record of the titles received every three months. This schedule also has the advantage of spreading out the work involved in the production of the lists, so that excessively long periods of computer use are not required.

The lists for intralibrary use were more diverse in type and more complex in mode of production.

*"PHILSOM CURRENT" (**CU**mulative **R**ecurring **R**eceipts—**EN**try of **T**itles)*

This listing of all issues received during the current month was an obvious necessity, for it will be remembered that one of the purposes of the project was the development of new systems to replace those presently in use. Thus, eventually the serial record as presently known could cease to exist.

Frequency of issuance of *PHILSOM CURRENT* was important, and as a trial a print-out on an alternate day basis was selected. This meant that the time lag between lists would be two days; i.e., *PHILSOM CURRENT* could not give data covering the day a question was asked or the previous day. For this period a different record, *PHILSOM DAILY*, was devised. As an added insurance, it was noted that the journal arrangement in the Library provided for keeping daily and weekly receipts separate (because they do not circulate), and scanning would quickly reveal all titles received since the last issue of *PHILSOM CURRENT*. This periodicity of issuance is by no means immutable and can be lengthened or shortened as experience indicates. *PHILSOM CURRENT* will be self-cumulative up to one month.

"PHILSOM DAILY"

As was common in many scientific libraries (and still is in many), this Library used to produce a daily listing of the titles which it had received. These lists were posted in a prominent place and enabled the user who might not have been able to come to the Library on a certain day to check to see if any titles of particular interest to him had been received. Like many another library, however, this one was forced to discontinue the service in 1961 because of lack of personnel and the cost of manual production of the lists. Many readers expressed a strong desire to have such lists, however, and so they were designated as one of the desired outputs from any new system. Actually, the method of journal check-in from pre-

PHILSOM DAILY

THURSDAY MARCH 14 1963

	VOLUME	ISSUE	DATE
AMER NAT	97	1	1963
BULL GEISINGER MED CENT	15	1	1963
CANAD MED ASS J	88	10	1963
CLIN PHARMACOL THER	4	2	1963
INT J CLIN EXPER HYPN	11	1	1963
J MICROSCOPIE	1	6	1962
J NEUROSURG	20	2	1963
NEUROLOGY	13	3	1963
PATTERNS DIS	1963	3	1963
PRESSE MED	71	7	1963
THER NOTES	70	3	1963
TROP DIS BULL	60	2	1963
TRANS ROY SOC TROP MED HYG	57	1	1963
YALE J BIOL MED	35	4	1963

FIG. 4.—Sample page from the PHILSOM DAILY listing. The list gives abbreviated title, current volume number, current issue number, and year of the issue.

punched cards made this type of listing quite simple (see fig. 4). The cards are sent to the Computer Center daily, and, before they are interfiled with the previous cards for the month (see *PHILSOM CURRENT*), they are listed on the 1401 computer; the print-out is then returned to the Library. It should be noted that this type of listing can be accomplished equally well using equipment such as an IBM 407 accounting machine with a specially wired control panel to arrange the list and suppress undesired details.

Acquisition Records

In order to provide budget information, facilitate the checking of vendor or supplier lists, and provide statistical data, two lists were planned for the Technical Services Section.

The first of these was a list arranged by vendor or other source, with the titles arranged alphabetically within each group (see fig. 5). The list gives the full title, code number for source of supply, and the latest price. The price is totaled at the end of each section. The second list is strictly alphabetical by title, giving the same information as above, but with a cumulative price total for the entire list.

Automatic Check-in and Claim Cards

Each month the Library receives one card for each serial issue expected to be received during the month (see fig. 6). When a journal issue is re-

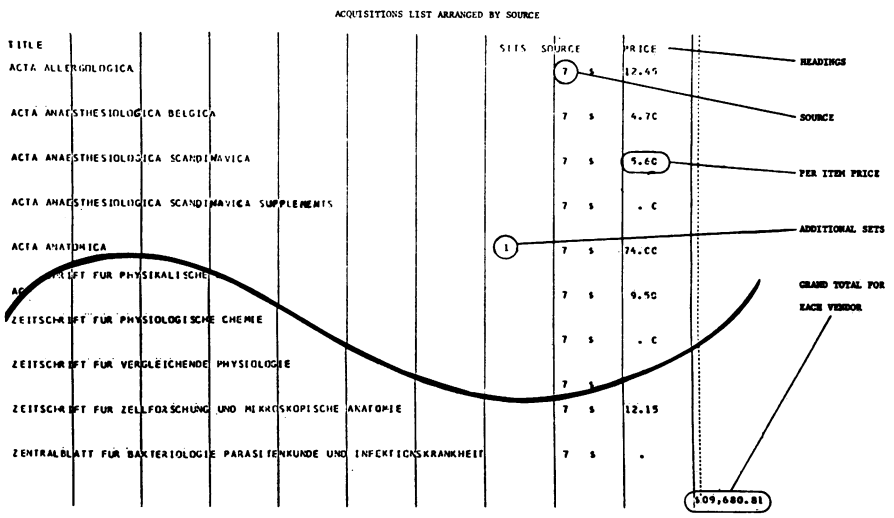

FIG. 5.—Two lists are prepared for the Acquisitions Section; this is a sample of the list arranged by source. The first and last pages of the list for one vendor are shown. The information given includes full title, price, number of additional sets (if any), and the total price for all titles purchased from a particular source.

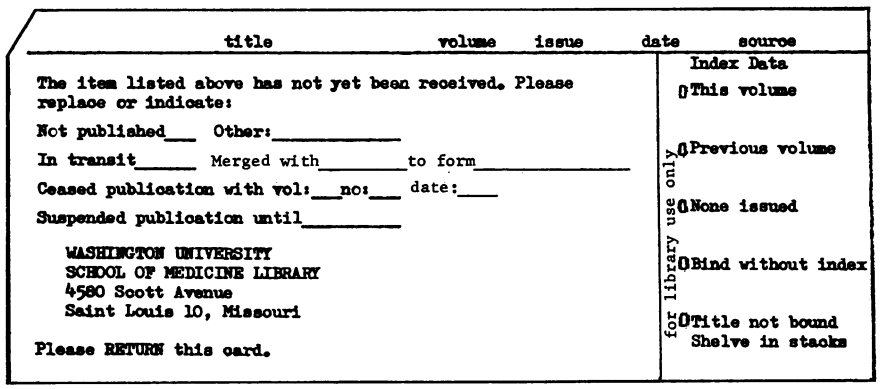

FIG. 6.—A test format check-in and claim card. All the cards which the Library receives each month are of this type. These are used to send to the vendor when an issue is not received. The right hand section titled "Index Data" is used by the Library to enter title page and index information.

ceived, the card for it is extracted from the deck, matched, and returned to the Computer Center for the updating of the master record and for listing in *PHILSOM DAILY*. The cards remaining at the end of the month then form the basis for claiming.

Claiming

Each month these remaining cards are interfiled with previous ones for issues which have not been received. Human judgment then determines when each will actually be sent to the supplier as a claim notice. At that point, a duplicate card is made, and one of them is sent to the source of supply as a request for the issue or for information about it. When the issue is received, the Library sends the other card back to the Computer Center for correction of the record.

Check-in cards are all designated with a special number (99) to distinguish them from the cards which are necessary for correction of other portions of the record; i.e., title changes, mergers, and new titles (see "Updating and Change Routines" below).

Bindery Notification

When the index to a volume is received, the serials assistant punches out the appropriate hole of the check-in card using a pencil or stylus. (The "Index Data" section of this card contains a row of squares which are actually prescored holes. See fig. 6.) When this punch is sensed by the computer, it scans the holdings record for the volume indicated. If the volume is complete, the title is punched into a card which becomes the bindery notification for the serials assistant. One of the advantages of putting this information into cards is that they can be easily interfiled for future use if the bindery pickups are less frequent than production of the cards.

At present no other outputs are planned.

Methods

The goals of this phase of the project having been decided, the next steps were determination of the information needed, the order in which it was to be recorded, the methods for control of the record, and finally the maximum size of the record. The record size was limited to 1250 characters.

A keypunch (IBM 026) was installed in the Library, and the full titles of all serials were recorded in a field of seventy columns. Most titles were fitted into this field unchanged; longer titles were either truncated in the latter portion of the title or abbreviated. At the same time all cross-references were keypunched, as was all history information, such as earlier or succeeding titles. The cards were proofread and alphabetized. To allow for machine alphabetization a seven-digit code number was assigned to each title automatically by the 1401 computer, based on a predetermined formula for assigning numbers proportionately throughout the alphabet. This served to tie together the multiple cards which were generated as

TABLE 1

FIELD DEFINITION FOR ORIGINAL RECORD CODING

CARD CODE	TYPE OF CARD	CONTENTS	COLUMNS
Blank	Full title	Alphabetical number	1–7*
		Card code	8–9*
		Indexed in *Index Medicus*	10*
		Full title	11–80
3	Abbreviation	Abbreviation	11–40
		Price	41–45
		Source	46–47
		Extra sets	48
		Current volume number	49–53
		Current issue number	54–57
		Frequency code	58–61
		Index data	62
		Language	63
		Blank (unassigned)	64–66
		Number of issues per volume	67–68
		Subject codes (maximum of 4 three-digit numbers)	69–80
4–7, 71–79	Holdings cards (nonseries)	Holdings	11–60
		Blank (unassigned)	61–80
8, 81–89	Series holdings	Holdings (1 series to a card)	11–60
		Blank (unassigned)	61–80
9, 91–93	History	History statement	11–80
1	Cross-reference	"From" reference (part one)	11–80
2	Cross-reference	"To" reference (part two)	11–80

* Columns 1–10 encode identical types of information on all cards.

parts of the record for each title. Each card belonging to a group was further coded to enable various portions of the record to be separated and to make possible the machine identification of special information (see Table 1). Columns 8 and 9 were designated as the card-code field and were punched as shown in figure 7. A new alphabetical number was assigned each time the computer sensed a blank or a "1" punch in column 8. At this time a print-out was also created to serve as the data collection sheet for the next steps and for proofreading.

Holdings were next coded onto the data sheets, corrections were made, and all new information was keypunched and listed for proofreading. This

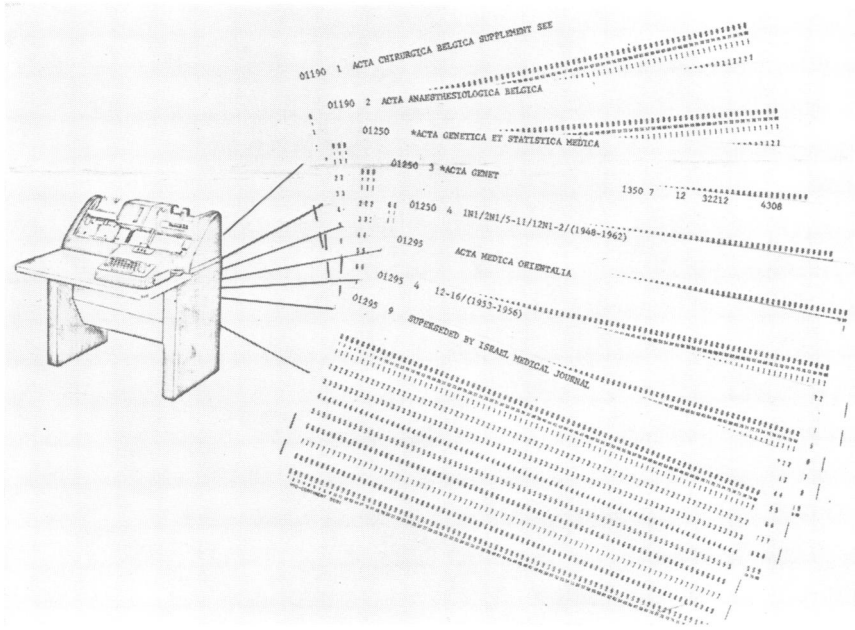

FIG. 7.—The information in the master record was assembled by punching the data into a series of IBM cards, of which the various types are shown here. Table 1 gives a detailed breakdown of the contents and arrangement of each card.

completed the assembly of the basic portion of the record and now permitted publication of *PHILSOM* from the cards.

Since this was one of the extralibrary output items, to be distributed in multiple copies, the most suitable mode of reproduction was studied (see "Equipment and Supplies" below). Mimeographing was selected as the best and cheapest method, and the entire computer output was generated on continuous-form mimeograph stencils by means of the IBM 1403 printer. The creation of the 179 pages required only twenty minutes of machine time.

Immediately after the publication of *PHILSOM* two further steps were taken. The data was collected to complete the abbreviation card (see Table 1), and work was begun on the format of the subject listing and the assignment of subject headings (7). Unassigned columns of this card are temporarily blank for additional information; e.g., the bindery schedule. A master list was prepared for entering corrections and changes which occur during the time between editions of *PHILSOM* (six months) and for obtaining alphabetical numbers for new entries (see fig. 8).

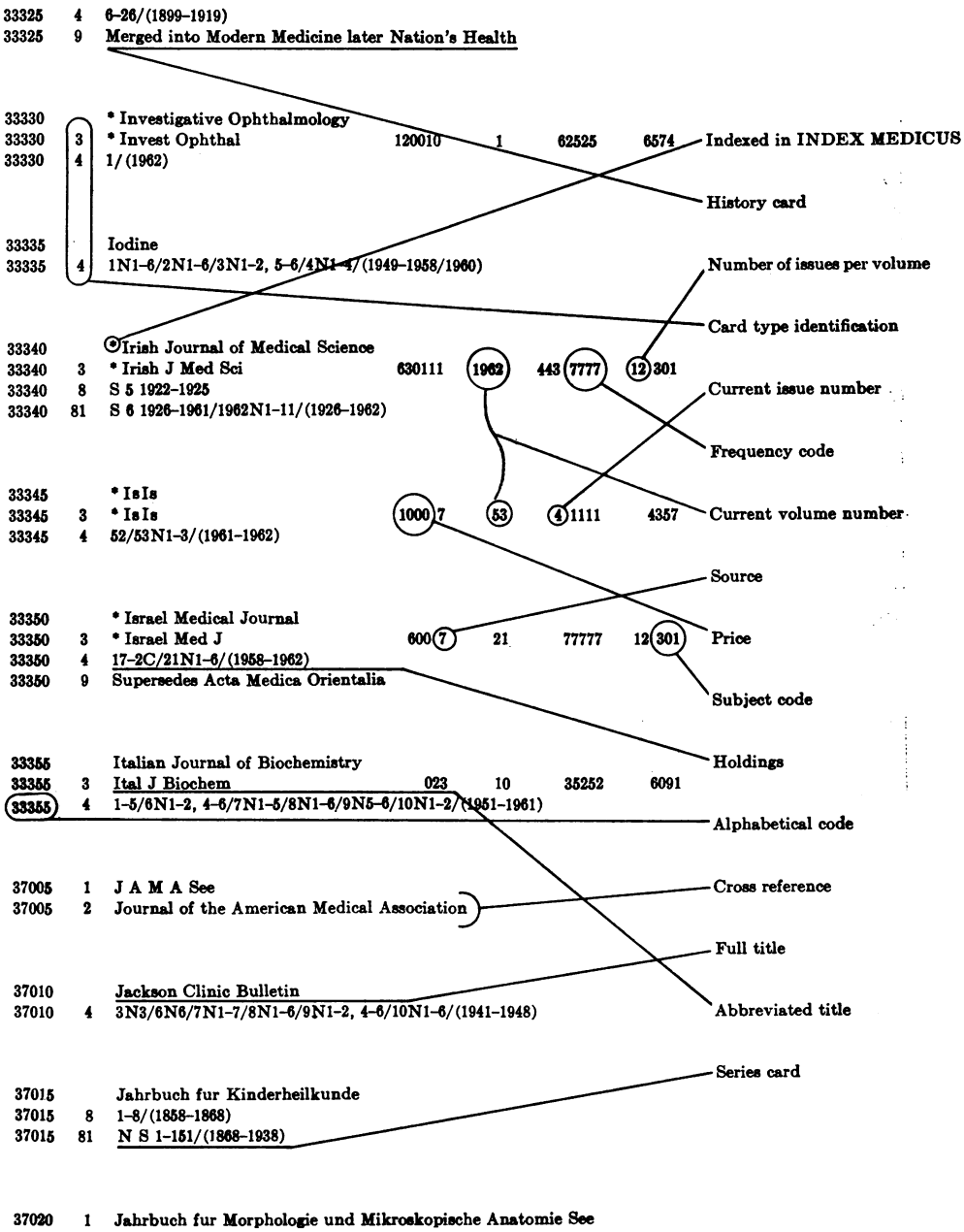

FIG. 8.—The master listing as produced for proofreading and for entering corrections and changes. It must also be consulted when new titles are added in order to obtain the alphabetical number, which is derived by interpolation between the two closest alphabetical items. (This is a typeset copy of the original computer print out, which is produced only in upper case.)

Frequency Coding

One of the major problems in any automated check-in system for journals is the varying frequency of the individual titles; in some libraries, especially very large ones, this problem may be a major stumbling block to the success of the program. We believe that the system evolved here would be applicable in all medical school libraries, and the frequency code may be adaptable for other cases with which we have not had to reckon.

The problem was approached from the standpoint that, of the 1200 titles which the Library currently receives, the majority are published on a regular, announced schedule. While this assumption helped to clarify the situation and enabled frequency variations to be regarded as exceptions to the rule, it did not in itself suggest a method of coding the information for efficient machine handling.

This problem was compounded by the limitations of space available for frequency information. It was finally solved by formulating a code number system which reduced practically all possibilities to a four-digit number which could be composed entirely of digits or combinations of digits and blanks (see fig. 9).

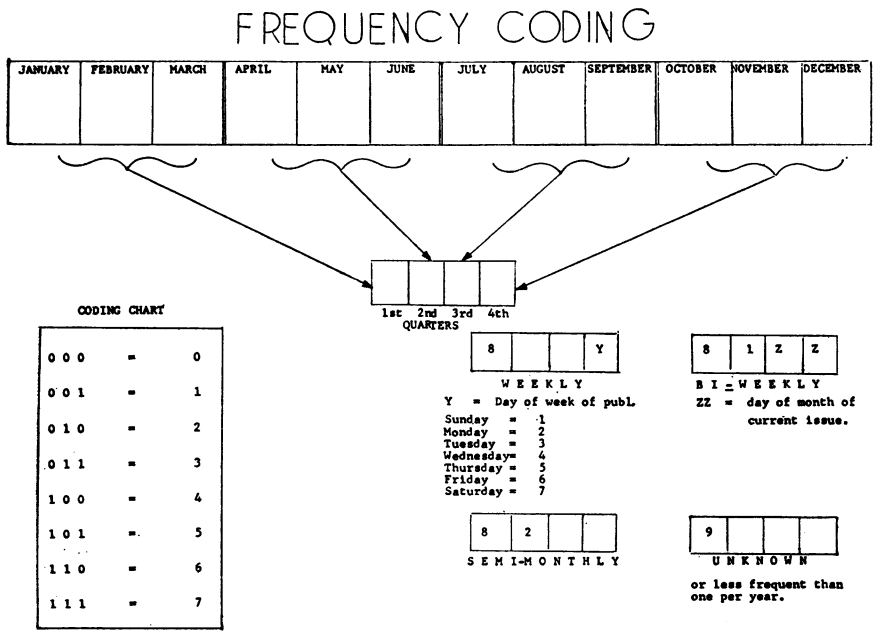

FIG. 9.—Frequency coding. The chart is used to convert the three-digit number to a single digit, saving space and enabling all frequencies to be given a unique number no matter how the publication frequency varies.

In order to clarify this concept, let us examine the case of two journals of different frequency. In the following examples and in actual practice, the digit "1" indicates that an issue appears at the time indicated, and a "0" (zero) means that no issue is expected during that period.

Developmental Biology is published seven times a year: in February, March, April, June, August, October, and December. If we plot these months on a chart such as that shown in figure 9, we obtain the following:

JAN	FEB	MAR	APR	MAY	JUNE	JULY	AUG	SEPT	OCT	NOV	DEC
0	1	1	1	0	1	0	1	0	1	0	1

which gives four three-digit numbers, one for each quarter. If we look these numbers up in the frequency code table in figure 9, we find that the "pattern" of "1" 's and "0" 's which we have established equals the four-digit number shown below:

1st quarter = 011 = 3
2nd quarter = 101 = 5
3rd quarter = 010 = 2
4th quarter = 101 = 5

→

3	5	2	5

Final four-digit code

On the other hand, if we examine a monthly publication such as the *American Review of Respiratory Diseases,* we find that the "pattern" consists of four groups of "1" 's; i.e., "111 111 111 111." On looking up this "pattern" in the code table, we find it equal to four "7" 's, and so, in this case, the frequency code number is "7777."

This type of coding will take care of all journals with any frequency from one month to one year. For serials which are more frequent than monthly, a variation on the system is used, as shown in figure 9. For titles which appear less frequently than once a year, the code number "9bbb" is given, in which the "b" 's represent blanks (no punch at all). It should be noted here that the purpose of the code is to enable the computer to determine when to generate a check-in (receipt) card. In this system, a new card is generated for irregular titles and titles which appear less frequently than once a year whenever the previous one is fed back into the updating routine.

When all the data had been collected, the development of the necessary computer programs was possible.

A flow chart for the entire system had been prepared as a general guide (see fig. 10), and now block diagrams were drawn up for the various steps in the system and the computer programs were written for the 7072 and 1401 computers to produce *PHILSOMS.* These were more complicated than those needed for *PHILSOM,* since the new list was to be created

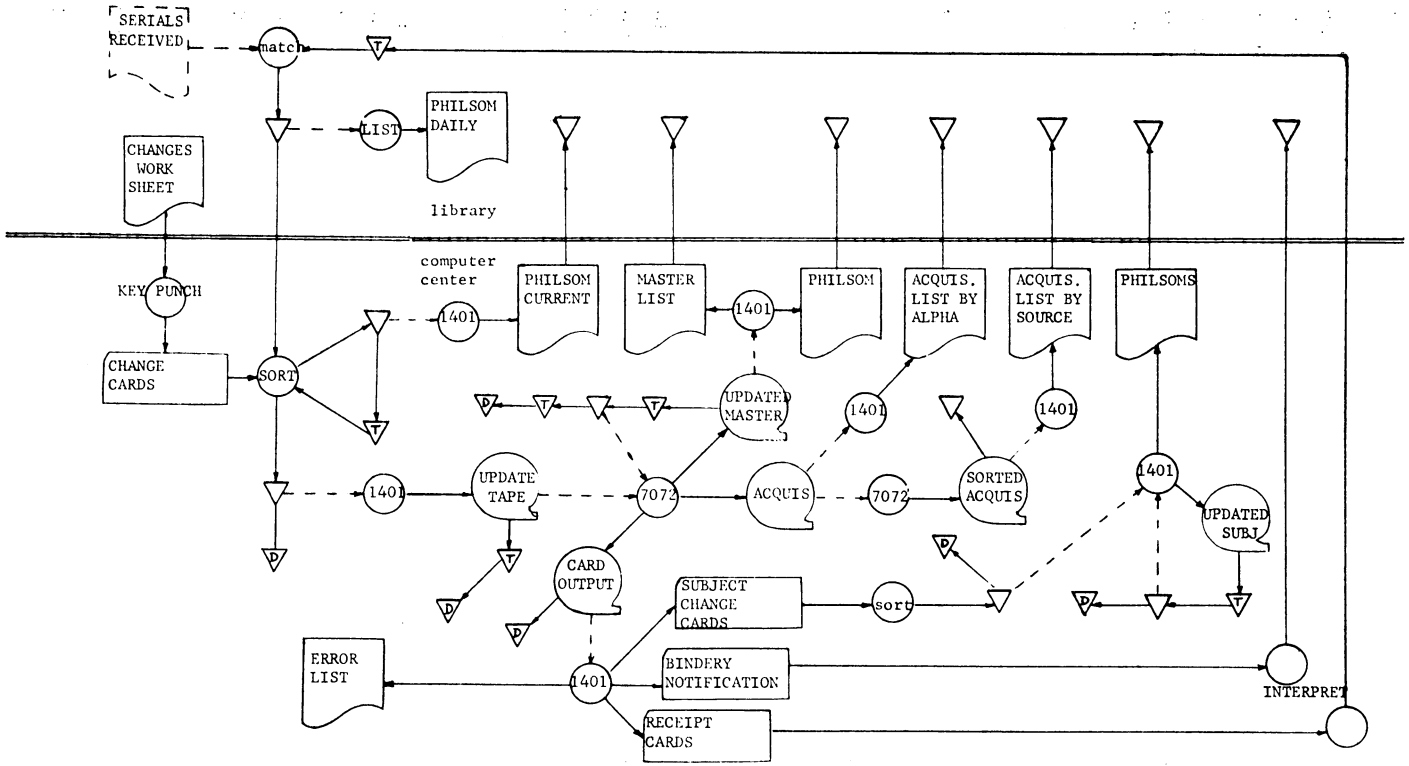

FIG. 10.—Overall flow chart for the entire system. The steps which occur above the double line are performed in the Library; those below the line are performed at the Computer Center.

from magnetic tape records. Moreover, it required a machine step to generate the various entries for each subject heading and a machine sort of the magnetic tape thus created in order to insert the subject headings and to arrange the titles alphabetically within each group. The programs already written were still useful, however, since they became parts of the complicated subroutines in the overall system.

Updating and Change Routines

The data needed for changing and updating the master tape records must be identified in some way other than that used for the original record. Columns 8 and 9 were again used for card-code identification (see fig. 11), but with a different series of numbers to indicate what type of change was to be made in the record. In effect, the digit "2" was added to all the previous card codes; e.g., cross-reference changes were added by designating the card code as "21" and "22" instead of the previously used "1" and "2." To identify long portions of the record, such as holdings which may require more than one card for the data, all cards except the last must have a "+" (plus) punched into column 10 to indicate that there is more to come. With this method, various portions of the record can be altered selectively.

Other methodological details, such as the organization of the record on tape by the use of index numbers, are not described here, but will be published at the completion of this phase of the project.

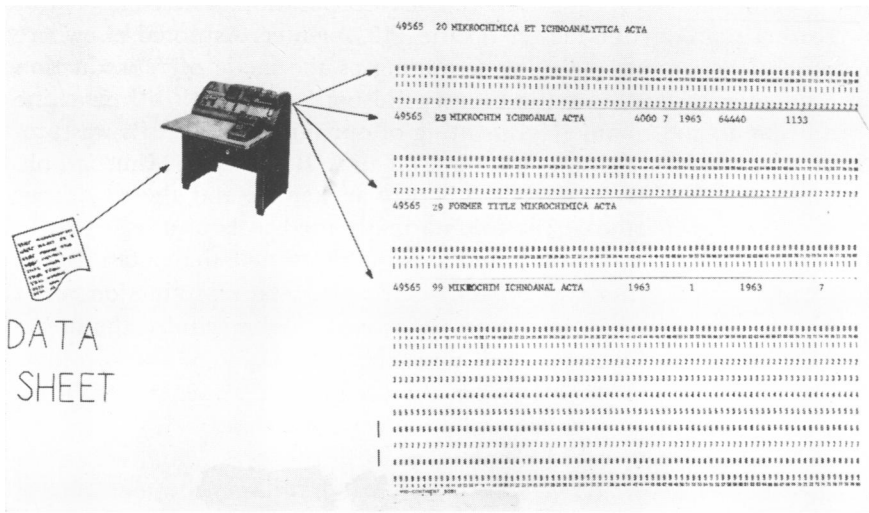

FIG. 11.—Change cards, in which the card code is altered so that machine identification of the type of information and the action to be taken is combined.

Equipment and Supplies

The Library's programs are designed for the IBM 1401 and 7072 computer systems. The 1401 system is made up of the 1401 processing unit, 1402 card punch-read unit, 1403 printer, and two 7330 tape-drive units. The printer operates at a speed of approximately 400 lines per minute when printing the Library's output, although it is capable of printing some types of data faster. The 1401 processing unit has 4 K (4000) memory positions (the equivalent of 400 ten-letter words). The 1401 system serves both input and output functions of the program; i.e., generation of updating tapes, format creation of output lists, and printing. The 7072 system uses six 7330 tape drives and has a memory capacity of 10,000 ten-digit words. This is used for sorting, merging, updating, and correcting the master holdings record and creating submaster tape records for output on the 1401.

All input to the system is done in card form, the cards being prepared on an IBM 026 printing keypunch, which allows for visual proofreading. Machine verification of keypunching is not done. The cards may then be sorted on any of the machines made for the purpose. Many lists can be created from cards alone, as mentioned earlier, by using unit-record equipment.

Two kinds of special supplies were required for production of the output in the form we desired. One was continuous-form mimeograph stencils, which were obtained from the A. B. Dick Corporation. These enabled us to create a reproducible master on the 1403 printer. As noted elsewhere in this article, we selected mimeographing as the mode of reproduction for economic reasons, since the Library did not have a multilith machine at its disposal, and commercial printing of our multilith stencils was prohibitively expensive. Printing was done on an A. B. Dick 434 Mimeograph. This machine is able to accept both regular stencils and those without the familiar paper band at the top, normally used to hook the stencil to the ink drum. A paste ink was found to be more suitable than a fast drying liquid ink for this latter type of stencil. Spirit-process reproduction could have been used; however, such a process would have precluded the use of two sides of a sheet of paper, and doubling the paper costs would have eliminated the saving obtained from the use of the cheaper spirit-process stencil. Esthetically, also, mimeographing was preferred.

The second kind of special supplies was cards used for the daily receipts, which contain prescored holes in designated positions to allow for the easy recording of index data. These cards are also preprinted for use as claim notices.

PROGRAMMING

The Library's programs are written in two machine languages, Autocoder for the 7072 system and SPS (Symbolic Programming System) for the 1401. Fortran, which is usable on the equipment available, was felt to be unsuitable for our purposes for several reasons. Fortran programs normally produce a greater number of machine instructions than do programs written in the languages chosen here, thus taking up a large percentage of the computer memory area and leaving a small amount of space for the actual data. Fortran is also an unwieldy language for data processing. It was designed for mathematical computations primarily and is unable to handle very large records efficiently.

In a recent article, Don S. Culbertson of the University of Illinois makes a plea for standardization of program language for library data processing, specifying the selection of COBOL (**CO**mmon **B**usiness **O**riented **L**anguage) as the most desirable (8). Standardization is, of course, a laudable goal; however, it is not without numerous problems. Since one of the most expensive factors in data processing is equipment, one must consider whether any language chosen as the standard is compatible with a large number of machines. In our case, we would be unable to use COBOL, since it requires a memory size of 8 K, or double the size of the memory on the 1401 system now used. It also requires a greater number of tape-drive units than we have presently available on the IBM 7072.

In addition, COBOL is a more general language than either Autocoder or SPS. It is relatively easy to translate a specific program into general terms (as a recipe for apple pie into a general theory of cooking), but not, however, to convert a general language to a specific system (as a general theory of cooking into a specific recipe for apple pie). Conversion of COBOL to other languages for other systems would thus be rather difficult. Culbertson further suggests that alternate solutions be made mandatory for parts of the program which are written in variations of COBOL; if this were done, one would be spending time writing programs which were of no use to the library which wrote them.

Standardization of combinations of equipment would seem to be the prime requisite for any system of standardized programming. Yet this is likely to be the area in which the library has the least control; i.e., except in enormous libraries such as the Library of Congress or the National Library of Medicine, most libraries will have to use equipment already available in the parent institution. Thus, until hospitals, medical schools, and universities standardize their equipment, the medical library's ability

to solve the problem by using COBOL is small. Standardizing the language is certainly not the immediate answer.

The major objection, nonetheless, remains the necessity for obtaining a larger computer memory to handle the language. It seems more desirable in setting up a standard procedure to start with the premise that one uses a minimum of equipment necessary to establish the goals set, rather than an intermediate or maximum array of machines. (This is the modern corollary of the medieval "Occam's razor"; namely, "entia non sunt multiplicanda praeter necessitatem"—entities should not be multiplied except by necessity.)

Statistics

One of the benefits derived from the entire program in machine methods has been the production of previously unknown statistical data concerning the WUSM Library. Many of these were unknown prior to mid-1961 or known only in part; for a year and a half the staff has been assembling data regarding such administrative concerns as budgets, estimation of work loads, use of the collection, its real size, and planning for a new building. In addition, new nonadministrative data was uncovered with wide implications for general bibliographic and bibliothecal knowledge. For example, when the subject listing of journals received in the Library was prepared, the program called for a sort procedure. As the initial tape record was read into the machine memory, the number of pieces of data were automatically counted and recorded on the typed print-out. This information was primarily for the benefit of the machine operator, to let him know where the machine was in its program and also to let him check to see that no items of data were lost by the time the procedure was completed. At the end of the program the machine typed out the number of items on the final tape. From this figure we could easily derive the number of subject entries assigned to each title. Not all machine generated statistics come out so easily—many must be programmed—but this particular programming is basically a simple matter, and the total number of journals in the library, the number currently received, total amounts paid to vendors, total amounts paid for all journals, etc., were all obtainable automatically. Once the program is written to elicit or accumulate this information, it need not be redone.

One of the most intriguing statistics derived when the count of the total number of titles in the Library was obtained may help us in evaluating the present state of medical literature publication. It was found that half of all the titles in the Library were still being currently received. This percentage forms an interesting datum which might have some correlation with statistics on the number of present-day scientists. As Gerald

deSolla Price and others have noted, approximately 80 to 90 percent of all scientists are still alive today (9). What hypothesis about the future of medical literature and medical libraries this fact raises is still unclear, but that there is some significance to it seems very likely.

Costs

If a new system is to obtain widespread acceptance, it must not be more costly than the one it supersedes (if it provides additional benefits not possible under the old, of course, one is not comparing equivalent things). The question frequently asked, "How much did it cost to produce *PHILSOM* or *PHILSOMS*?" is therefore both a logical and a pertinent one.

Because, however, these publications are only two of the products which will eventually be produced by the new system, it is hard to determine the costs of these particular segments. Several operations (such as keypunching, placing the record on magnetic tape, working out data sheets, designing special forms, or training personnel to handle the system) are fundamental to all portions of the system, and it is difficult to figure out how to prorate the total costs for any one segment. In struggling with this problem, our staff has come up with an analogy to a commercial bakery—probably because the delicious odor of new-baked cinnamon rolls is wafted over to the Library every morning from the bakery across the street.

In a bakery making rolls, bread, cake, buns, etc., the initial cost of preparing the dough can be divided into the quantity percentage of the total output of each item. If 500 1-lb. loaves of bread, 1000 2-ounce rolls, 75 pound cakes weighing two pounds each, and 3600 3-ounce cinnamon buns are produced from a batch of dough which costs $750 to produce, it is possible to work out the cost of dough per loaf of bread, per roll, per cake, or per bun. Unfortunately, initial costs for the machine project cannot be divided up among all the products of the system by any such method; this is so because, unlike the dough which is made in larger or smaller amounts depending upon the size of output desired, the total preliminary work of the machine project has to be accomplished even if only one end product is to be produced. In a sense it might be logical to charge the entire preliminary costs to *PHILSOM* or *PHILSOMS*, since without the keypunching and other early work these products could not have been produced. Yet to do so would be to give a false impression, because the basic coding will be used for several products with no common factor (like weight of baked goods) on which prorating can be based.

Another reason that costs of our program are not clear-cut, of course, is that it is an experiment, with the usual false starts, blind alleys, and inoperable way stations adding to the development costs. The true appraisal of what it would cost a library to use the system developed here

TABLE 2

COSTS OF EXTRALIBRARY PUBLICATIONS

	Price
PHILSOM	
Continuous-form mimeograph stencils (179)	$ 22.00
Paper, drilled for three-holed binder	55.00
Printing and design of cover	5.00
Machine time	12.00
Clerical time for mimeographing, collating, distributing	33.38
Mimeograph 434 (amortized over 10-year period)*	14.00
Mimeograph supplies: Ink	8.25
Stencil wrappers	6.74
Drum cover	0.04
Stencils (preface etc.)	1.02
Total cost for 100 copies	$157.43
Cost per copy	1.57
PHILSOMS	
Continuous-form mimeograph stencils (71)	$ 7.81
Paper, drilled for three-holed binder	11.75
Printing and design of cover	5.00
Machine time	45.00
Clerical time for mimeographing, collating, distributing	9.00
Mimeograph 434 (amortized over 10-year period)*	14.00
Mimeograph supplies: Ink	2.75
Stencil wrappers	2.96
Drum cover	0.04
Stencils (preface etc.)	0.34
Total cost for 100 copies	$ 98.65
Cost per copy	.99
Paper fasteners	$ 1.00
Postage	2.50
Total cost for PHILSOM and PHILSOMS	$259.58

* Yearly amortization cost quartered.

can come only when all parts of the system are operational and the entire program functioning.

This is not to say that we have no idea of our costs, or that all the costs of the project are to be charged off to research and development. We have enumerated in Table 2 the actual costs for the production of the published lists: the equipment and supplies necessary for their production alone and the cost of the machine time needed to produce them. What we have not included, then, are professional salaries, the cost of writing pro-

grams, and the machine time which was spent in "debugging" and manipulating the programs to make them more effective.

"PHILSOM" and "PHILSOMS"

Since the first production of *PHILSOM* was from cards, we are not yet able to estimate the machine-time costs which will eventually be required for its production from magnetic tape (this estimate is further complicated by the fact that as the tape for *PHILSOM* is corrected, so also are the tapes for *PHILSOMS*, the acquisitions lists, and the master tape record). Included in this production procedure is updating of the master record, addition to or correction of the record, and production of the following month's check-in cards.

At the time of its initial production, the card file was some 9,000 cards in length (for approximately 2,500 titles). These cards required twenty minutes to be read through the 1401, while suppressing undesired information for this listing, arranging format, and printing out the stencils.

PHILSOMS was more expensive to produce, even though it is shorter, because it required more machine manipulation. (It should be noted here that cost for the publication and distribution of both *PHILSOM* and *PHILSOMS* were borne by the Library rather than the General Research Support Grant. The Grant, of course, paid for the experimental work.)

Current Status

At the end of March 1963, the developments and current status of the project can be summarized as follows:

1. *PHILSOM* was published in January 1963 from cards.
2. *PHILSOMS* was published in March 1963 from tape.
3. Acquisitions lists, both by source and alphabetical, were printed in March 1963 from tape.
4. Work was proceeding on the production of the various programs for portions of the system, both on the 1401 and the 7072. It was expected that work would be completed on the programs by the end of April and that May and June would be spent in correcting and changing various details.
5. Production of the second issue of *PHILSOM* was scheduled for the latter part of June and the second issue of *PHILSOMS* for the end of September 1963.
6. Automatic check-in of journals and correction of the master record for the production of *PHILSOM* was to be undertaken in June.
7. A symposium on the current state of the project was held on April 26, 1963, for twenty-five librarians from all parts of the country. In

addition to talks on the development of the system and outlines of programs, etc., a demonstration was planned showing the steps in the compilation of the subject listing (*PHILSOMS*).

8. A three-day symposium will be held on September 5–7, 1963. Limited to fifty registrants, the meeting will concern the development and theory of operation of unit-record machines and computers, in addition to reports on current developments in this project and in others of related nature.

Immediate Next Steps

Once the check-in system has become operational (June 1963), it will require close attention to determine what types of problems arise and how they may best be solved.

This stage will bring to a close the outlined program of experimentation with the serial record. The Summer months will be spent developing and putting into operation a semiautomated circulation system for library materials. In the Fall, it is expected that work will be begun on the next phase of the program, which aims at varying degrees of coordination of acquisitions and cataloging and at the eventual production of a distributable book catalog.

REFERENCES

1. Fisher, Charles Perry. The typewriter in cataloging and shelf-listing. Med. Libr. Hist. J. 1: 201–204, 1903.
2. Brodman, Estelle. The Development of Medical Bibliography. Medical Library Association, 1954, p. 102–104.
   Price, Gerald deSolla. Science Since Babylon. New Haven, Yale University Press, 1961, p. 92–124.
3. International Business Machines Corporation. Literature on Information Retrieval and Machine Translation. C. F. Baly and R. H. Stanwood, eds. New York, 1962.
4. University of California, San Diego. Report on Serials Computer Project, University Library and UCSD Computer Center, 1961/62. La Jolla, California, 1962.
   Vdovin, George; Voigt, M. J.; Newman, D.; and Perry, C. Computer processing of serial records. Libr. Res. Techn. Serv. 7: 71–80, 1963.
5. General Electric Company. MEDLARS Phase One, Project Report. Bethesda, Maryland, 1962.
6. Schultheiss, L. A.; Culbertson, D. S.; and Heiliger, E. M. Advanced Data Processing in the University Library. New York, Scarecrow Press, 1962.
7. Brandon, Alfred N. Subject list of journals indexed in *Index Medicus*. Bulletin 50: 353–406, 1962.
   (This list was taken as the basis for subject coding, in order to achieve uniformity and to relate the Library's journal collection to current *Index Medicus* subject classification. Adaptation was made to fit specific local needs.)
8. Culbertson, Don S. Another tower of Babel? Libr. J. 88: 940–943, 1963.
9. Price. *Op. cit.*, p. 107.