Mr. Adams

THE MEDLARS STORY at the NATIONAL LIBRARY OF MEDICINE

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE
PUBLIC HEALTH SERVICE

THE MEDLARS STORY

at the

NATIONAL LIBRARY OF MEDICINE

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE
PUBLIC HEALTH SERVICE
1963

FOREWORD

The problem of providing improved bibliographic access to the voluminous medical literature is critical. The number of publications is growing rapidly. New specialties and subspecialties appear. The increasing support of research is resulting in a greatly enlarged output of findings. These and other factors threaten to overwhelm traditional means of helping the scientific community keep abreast of new knowledge.

The National Library of Medicine has felt acutely the need to improve upon its own bibliographic techniques. It has needed a system which would satisfy increasing retrieval requirements and at the same time provide for the publication of the Index Medicus and its byproducts. MEDLARS (Medical Literature Analysis and Retrieval System) is the system it is adopting.

The study, design, and implementation of the MEDLARS project was performed under contract by the Information Systems Operation of the General Electric Company. We have welcomed the General Electric Company's offer to prepare this description of the system, and we are grateful for the Company's permission to reproduce it.

Although this document is primarily the work of General Electric, many of the ideas and concepts were provided by National Library of Medicine publications and staff. Of particular note are the National Library of Medicine Index Mechanization Project Report (Bulletin of the Medical Library Association, vol. 49, part 2, January 1961) and the Library's recent Annual Reports.

Library and General Electric personnel have worked together closely on this project, and continue to do so. Overall direction of the Library's participation has come from Mr. Seymour I. Taine, the Project Officer and Chief of the Bibliographic Services Division; contributions concerning aspects of the preliminary design, from Mrs. Claire K. Schultz, Consultant to the Library; detailed program assistance, from Mr. James P. Riley, who was initially responsible for MEDLARS liaison, and from Mr. Charles J. Austin, Head of the Data Processing Section; and assistance regarding the subject heading authority lists from Winifred Sewell, Deputy Chief of the Bibliographic Services Division.

MEDLARS will become operational early in 1964.

All of us look forward eagerly to that event.

Frank B. Rogers, M.D.
Director
National Library of Medicine

TABLE OF CONTENTS

Chapter		Page
	FOREWORD	iii
	CONTENTS	v
	LIST OF FIGURES	vi
	LIST OF TABLES	vi
	LIST OF EXHIBITS	vii
I.	INTRODUCTION	1
II.	WHAT MEDLARS WILL DO	3
	Introduction	3
	MEDLARS Objectives	3
	An Overview	4
	The MEDLARS Unit Record	6
	Input and Conversion Subsystem	11
	Manipulation Subsystem	15
	Output Subsystem	32
III.	MEDLARS AT WORK	36
	Introduction	36
	The Input and Conversion Subsystem	36
	The Manipulation Subsystem	39
	The Output Subsystem	45
	Conclusion	46
IV.	DESIGN CONSIDERATIONS	47
	System Criteria	47
	Computer Selection	49
	The Output Subsystem	53
	Selection of Data Conversion Equipment	55
V.	MANAGEMENT CONSIDERATIONS	56
	Introduction	56
	Time Schedule	56
	Equipment	58
	Operating Personnel	58
	System Operating Times	58
	Miscellaneous Operating Details	60
	Secondary Objectives	61
GLOSSARY		69
DEFINITION OF SYMBOLS		74

LIST OF FIGURES

Figure No.		Page
1.	Subsystems of MEDLARS	7
2.	Hypothetical Journal Article Record	8
3.	Hypothetical Non-Journal Article Record	12
4.	Input and Conversion Subsystem	14
5.	The Compressed Citation File: Arrangement of the Elements of One Citation	17
6.	Major Inputs and Outputs of Each Program Module	20 & 21
7.	Block Diagram for Input Processing Module	22 & 23
8.	Block Diagram for Demand Search Module	25
9.	Block Diagram for Output Processing Module	27 & 28
10.	Block Diagram for Report Generator Module	30
11.	Block Diagram for Statistical Module	33
12.	Block Diagram for Output Subsystem	34
13.	Journal Processing Flow Chart	37
14.	Computer Area	40

LIST OF TABLES

Table No.		
1.	Comparison of Capacity of Exiating System and MEDLARS	4
2.	File Usage for Each Program Module	19
3.	Computer Time Requirement by Type of Usual Daily Task, 1964	60

LIST OF EXHIBITS

		Page
Exhibit A	Journal Data Sheet (Preliminary)	64
Exhibit B	Unit Record on Typed Copy Unit Record Information on Paper Tape	65
Exhibit C	Request for MEDLARS Literature Search (Preliminary)	66
Exhibit D	Demand Search Data Sheet (Preliminary)	67
Exhibit E	GRACE I M Page (Premiminary)	68

I

INTRODUCTION

The National Library of Medicine dates back to 1836, when its predecessor, the Library of the Surgeon General's Office, was established. The first large-scale indexing of current medical-journal literature using institutional team-approach methods had its start in 1879 with the publication, by this Library, of the first volume of the first work to use the title Index Medicus. This was followed in 1880 by the first volume of the first series of the Index-Catalogue of the Library of the Surgeon General's Office.

In 1916 the American Medical Association started the publication of a Quarterly Cumulative Index to Current Medical Literature. In 1927 this material was merged with that of Index Medicus to form the Quarterly Cumulative Index Medicus. In succeeding years, circumstances brought about other changes, with the result that two indexes are currently produced: the monthly Index Medicus, published by the National Library of Medicine, and the annual Cumulated Index Medicus, accumulated by the National Library of Medicine and published by the American Medical Association.

The preparation of the basic publication, the monthly Index Medicus, is a monumental task. In 1961 the publication averaged 450 pages in length and contained references to more than 10,000 articles per average issue. The annual total of items indexed has increased from 120,000 in the first annual volume (1960) to 140,000 in 1961, and it will reach an estimated 250,000 in 1969. A total of 14,000 journal issues were indexed in 1961, and this figure is expected to reach 25,000 by 1969.

In addition to Index Medicus, the Library regularly publishes one special recurring bibliography, the Bibliography of Medical Reviews, and services assorted nonrecurring requests for bibliographies of various complexities. It is expected to expand the list of regularly-published bibliographies to a total of 50 by 1964. The nonrecurring "demand" bibliographies will reach an estimated total of 2,500 in 1964 and 22,500 in 1969.

With the installation of MEDLARS, the Library expects to add monographs to its list of source material to be indexed, and a total of some 5,000 English-language monographs is foreseen for 1964. These will be combined eventually with foreign monographs, and total of both should reach 13,000 by 1969. The grand total of articles and monographs indexed is expected to reach 160,000 by 1964 and 250,000 by 1969.

The predecessor publication of the monthly Index Medicus was the Current List of Medical Literature. Its preparation was a manual operation which increased in difficulty as the publication grew in physical size and scope. In 1950, a shingling technique, a method of manually arranging the publication's entries in sequence for photographing, was adopted as the first step in the printing process. Shingling gradually became a bottleneck, preventing future expansion. In 1960 a new mechanized system was adopted for the publication of the Index Medicus.

The mechanized system - which is the existing system - involves the use of tabulating cards, upon which the citations are imprinted. The cards are also punched with machine-readable sorting information. After machine sorting and matching, the cards, suitably arranged, are automatically photographed with a Listomatic camera upon rolls of film which are then cut and assembled into page arrangement for printing.

Although the existing mechanized system offers distinct advantages over the older technique, it is limited solely to a publication system. It cannot satisfactorily meet growing demands for rapid retrieval of complex requests specified according to multiple subject axes. Some way had to be found to accomplish more, to do it better, to take less time, to operate more efficiently, and generally to provide a greater all-round versatility.

This was the situation that led to the initiation of the MEDLARS Project. That system is described in the following pages; in order that the reader be allowed a comprehensive knowledge of MEDLARS he is introduced to the system as a whole, then by observing it at work, and then is given the technical particulars of its individual components.

II

WHAT MEDLARS WILL DO

INTRODUCTION

MEDLARS is designed for use by the National Library of Medicine to perform various functions of literature analysis and retrieval. It will involve the use of a digital computer and special composing equipment capable of providing excellent typographic quality. Knowledge of the specific objectives of the system is basic to an appreciation of what MEDLARS is designed to do and how it will be done. Accordingly, this chapter begins by first outlining the objectives of the system. This is followed by a description of how these objectives are to be met. In the description, a brief overview of the entire system precedes a somewhat detailed examination of the system's component parts.

MEDLARS OBJECTIVES

The objectives of MEDLARS may be summarized as follows:

- Improve the quality of and enlarge Index Medicus and at the same time reduce the time required to prepare the monthly edition for printing from 22 to 5 working days.

- Make possible the production of other compilations similar to Index Medicus in form and content.

- Make possible for Index Medicus and other compilations, the inclusion of citations derived from other sources, as well as from journal articles.

- Make possible the prompt (a maximum of two days) and efficient servicing of requests for special bibliographies, on both a demand and a recurring basis, regularly searching up to five years of stored computer files.

- Increase the average depth of indexing per article by a factor of five, i.e., 10 headings versus 2.

- Nearly double the number of articles that may be handled annually - from 140,000 now to 250,000 in 1969.

- Reduce the need for duplicative total literature screening operations.

- Keep statistics and perform analyses of its own operations, to provide the information needed to monitor and improve system effectiveness.

- Permit future expansion to incorporate new and as yet not completely defined - and hence secondary - objectives.

An idea of the anticipated capacity and capability of MEDLARS over the present system may be gained from Table 1.

Table 1

Comparison of Capacity of Existing System and MEDLARS

Item	Existing System	MEDLARS	
	Current	1964	1969
Journal issues indexed	14,000	16,000	25,000
American monographs indexed	-	5,000	-
American and foreign monographs indexed	-	-	13,000
Articles and monographs indexed	140,000	160,000	250,000
Recurring bibliographical requests filled*	-	50	50
Demand (non-recurring) bibliographical requests filled	-	2,500	22,500

* The bibliographies produced will be prepared at an average rate of three or four issues per day; they will recur with periodicities ranging from one week to six months.

OVERVIEW

Under MEDLARS, the regularly published products will continue to be, for the foreseeable future, the monthly issues of Index Medicus, the annual Cumulated Index Medicus, annual printings of Medical Subject Headings and List of Journals Indexed, and such recurring bibliographies as Bibliography of Medical Reviews. The typography for these will be of a quality comparable to the currently published issues of Index Medicus, through use of the specially-designed Graphic Arts Composing Equipment (GRACE). GRACE will compose finished copy on film suitable for offset printing. In addition, demand (non-recurring) bibliographies will be produced to meet

specific requests as they are received by the Library. Most demand bibliographies and all statistical information prepared by the system will be printed by a standard computer printer.

The various operations required to produce these publications may be logically separated into three subsystems (see also Figure 1):

1. Input and Conversion Subsystem

   In this subsystem, journals, monographs, and other documents are received and indexed or cataloged; a "unit record", consisting of a citation and its associated headings, is prepared for each article, book, and serial title; search requests are received and prepared; and unit records and search requests are transformed into machine-readable form with punched paper tape as the primary medium.

2. Manipulation Subsystem

   The heart of this subsystem is a high-speed, digital computer. It accepts the unit records on paper tape, checks them for the presence and correctness of those elements for which such checks are possible, does some preprocessing to facilitate and speed the subsequent processing, and stores the unit records on magnetic tape. In response to search requests, it searches its cumulation of unit records for those that qualify for retrieval and edits and composes them for output on magnetic tape.

3. Output Subsystem

   This subsystem transforms the magnetic-tape output of the Manipulation Subsystem into exposed film from which publications may be printed. When high-quality typography is not required, the standard output of the computer's mechanical printer is utilized.

The way in which the three subsystems, or groups of operations, will relate to each other may be visualized in a capsule description of the operations of the system as a whole: Journals received by the Library will be distributed to a staff of indexers for the selection of articles to be indexed, the translation of foreign article titles, and the indexing of article citations with appropriate subject headings and other tags. Documents such as monographs will be distributed to a staff of catalogers who will perform somewhat similar functions. A unit record will be prepared for each article indexed and for each document cataloged. The unit records will be punched on paper tape in a form suitable for computer input. The paper tape will be read into the computer, which will accept the information, process it, compress it, and store it on magnetic tape. The unit records will be ordered in chronological sequence and will be stored for at least a five-year period.

Once a month, for the publication of Index Medicus, the unit records placed into storage over the past month will be expanded, sorted, edited, and rewritten in final form on magnetic tape. Once a year the material will be merged for the Cumulated Index Medicus.

The magnetic tape output of the computer will be used by the Graphic Arts Composing Equipment, which will compose each citation, properly arranged, on photographic film. This film will be developed, inspected, and transmitted to the printer. It will be used to make the offset plates from which the pages of Index Medicus will be printed. Film will also be prepared for printing the annual Cumulated Index Medicus and recurring bibliographies.

Recurring bibliographies may cover one or more clinical or research areas which are responsive to the bibliographic needs of specialized consumer groups. Periodically, the subject-heading parameters and output format requirements will be prepared by search specialists and entered into the computer on punched parameter cards. An increased depth of indexing, as well as the combining of the subject headings and other search tags (such as relate to journals, authors, and year of publication), will be possible. The magnetic tape output from the computer will be handled in the same way as the tape produced for Index Medicus. The printed format of recurring bibliographies will be flexible and can be similar to that of Index Medicus.

Requests for demand (non-recurring) bibliographies will be processed by a staff of bibliographic searchers and be converted to machine-readable form on punched paper tape. The tape will be read into the computer, which will then search its files for the citations that meet the request requirements. These retrieved citations will be properly arranged by the computer and then printed by means of the computer printer. In some instances, the Graphic Arts Composing Equipment will be used for printing demand bibliographies.

All of these operations, plus storing information and preparing statistics necessary to monitor and improve system effectiveness, will be performed speedily. The monthly issues of Index Medicus will require five working days for completion; both recurring and demand bibliographies will be processed in one or two days, with priority demand requests being given same-day service.

Broadly outlined, these are MEDLARS' operations. Let us proceed now to examine them in more detail.

THE MEDLARS UNIT RECORD

An entry into the MEDLARS system will consist of a citation plus its associated tags and is called a unit record. The unit record will be the fundamental entity prepared by the system. From the time that it is initiated until it leaves, it will both activate and be acted upon by the system. However, the medium on which it is stored, the language in which it is expressed, its syntax, its length, and the frequency of its occurrence will all change. There are two types of unit records: (1) the journal article record and (2) the non-journal article record.

Journal Article Record

Journal article records will be the type of unit record associated with the vast majority of citations appearing in Index Medicus. Journal article records will consist of the following elements (see Figure 2).

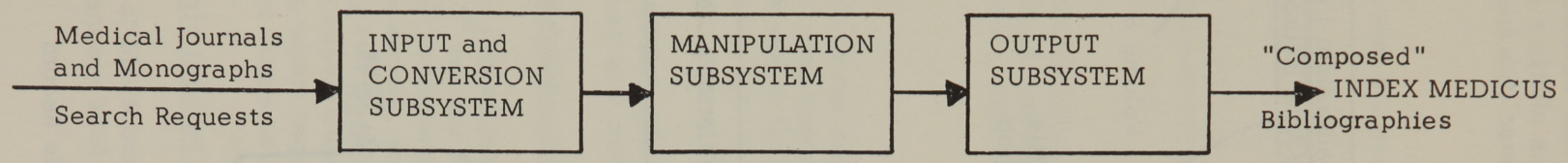


Figure 1. Subsystems of MEDLARS

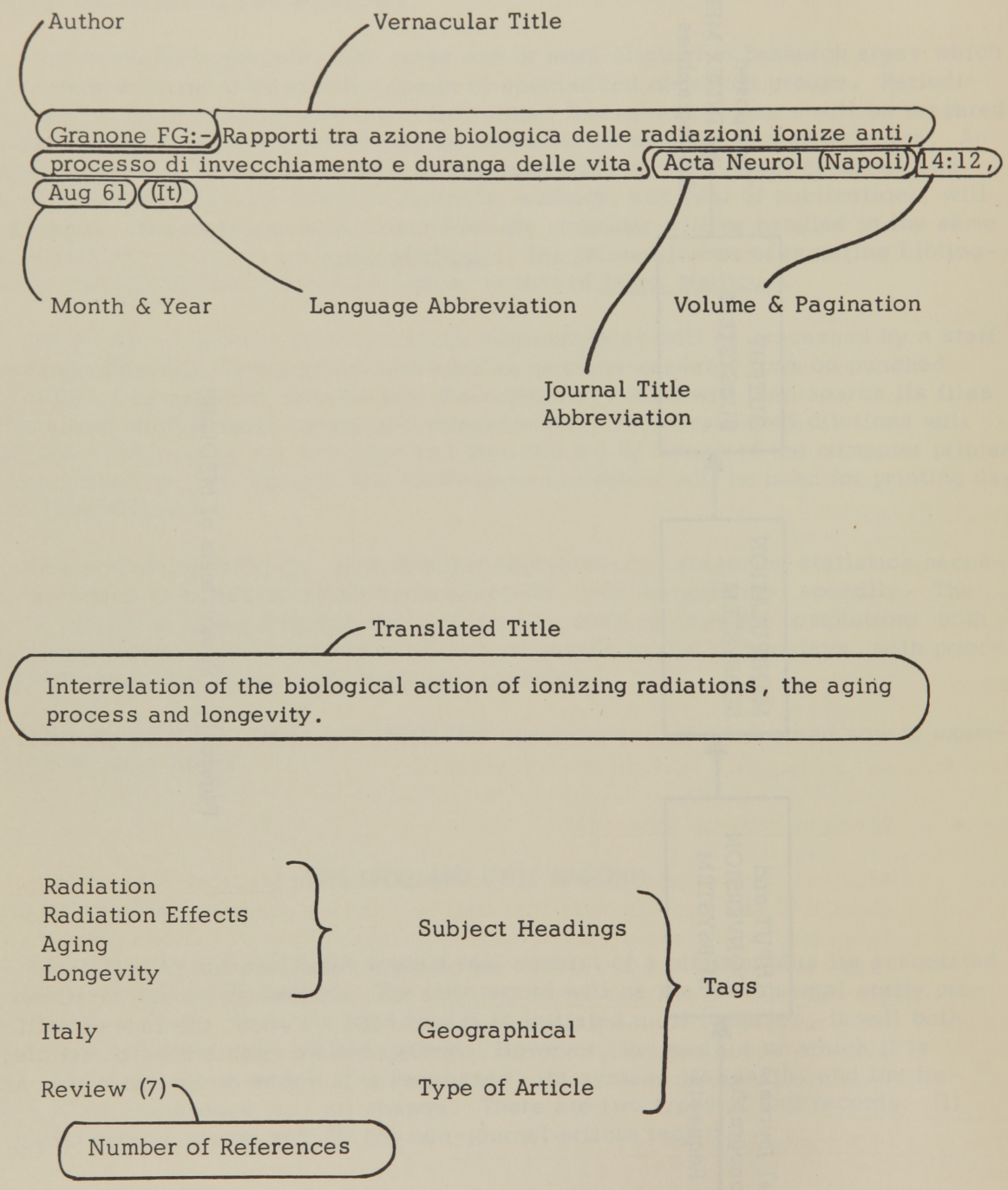


Figure 2. Hypothetical Journal Article Record

Author's Name

Experience indicates that 1.5 percent of the unit records will have no authors, 59 percent will have a single author, 23 percent will have two, 10 percent will have three, and 6.5 percent will have four or more authors.

Title in English

It is expected that 35 percent of the unit records will refer to articles written in English.

Foreign Language Title

This will be a vernacular version, always using the Roman alphabet. Some 65 percent of the unit records will cover foreign language citations.

Journal Title

This will be abbreviated in accordance with a standard list.* Currently, some 2,200 journals are being indexed. This number is expected to increase to 3,000 in 1964 and 6,000 in 1969, with the increase expected to come primarily from foreign language journals.

Language

An abbreviation of the language in which the article is published will be taken from a standard list. The current number of languages listed is 32; it is expected to rise to 40 in 1964 and to 55 in 1969.

Tags

Tags which will describe the contents and form of the article will be provided by the indexers. They will include the following.

- Medical Subject Headings

  Medical subject headings which are published in Medical Subject Headings (MESH)** are the only tags that may be used as headings in Index Medicus. However, with an anticipated increase in the depth of indexing from two tags per article, the physical size of the subject section of both the Monthly and Cumulated Index Medicus would increase if each citation were to appear once under each assigned tag. Hence, the indexer will indicate which tags are most descriptive of the contents of each article. In effect, there is one level of indexing for Index Medicus and two levels

---

* The List of Journals Indexed in Index Medicus (LJI) is a regular publication of the Library.

** National Library of Medicine, Medical Subject Headings. 2nd Edition, Part 2 of Volume 4, No. 1, January 1963, INDEX MEDICUS

of indexing for both demand and recurring bibliographies. Thus, a specific subject heading might be checked under one citation as appearing in Index Medicus and for a second citation as not appearing in Index Medicus. It is anticipated that each citation will have an average of 10 tags assigned, but only an average of three of these tags will be checked by the indexer as those tags under which the citation will appear in Index Medicus.

In addition to the tags in MESH, there will be other tags which never appear in Index Medicus but are used for both demand and recurring bibliographies. Some of these tags are described below.

- Provisional Headings

  "Provisional headings" will be new descriptors which may be required to adequately describe an article. Examples might be tentative or investigational drugs. Provisional headings will never appear in Index Medicus but may be used for coordinate searching or as headings in recurring bibliographies. If at a later date it is decided that a provisional heading has general use, then it will be added to MESH, and it will appear in Index Medicus.

- Age Group Tags

  These tags will be used whenever they are applicable.

- Public Health Service Support Group Tags

  These tags will be used to identify medical literature resulting from grants authorized by the Public Health Service.

- Geographical Tags

  Such tags will be used, when applicable, to designate the country about which an article is written (not the country in which the article was written). There will be about 200 geographical tags.

- Volume and Page Data

  These will consist of information in standard bibliographic form.

- Date and Year

  This information may appear in various forms or arrangements and include the year only (about 24 percent); the month and year (about 53 percent); the day, month, and year (about 19 percent); and other designations (about 4 percent).

- Entry Date

  This will be the date that the unit record is received by the computer. The computer will supply it automatically.

Non-Journal Article Record

Non-journal article records (NJAR) (Figure 3), the second type of unit record, will be derived primarily from published monographs. They will represent the body of material that the Library catalogs. When MEDLARS becomes operational, the Library will discontinue the section of Index Medicus called "Recent U.S. Publications" and instead cite cateloged material by subject and author in both the monthly and annual volumes. The printed form of the citations will be similar, though not identical, to the form of the citations entered for journal articles.

The non-journal article records will contain not only the elements required for their citation in Index Medicus but also such additional elements as are required to prepare standard catalog cards. Such elements will include the following:

- Title (substituted for journal title).
- Edition note - information covering editions, printings, revisions, etc.
- Descriptive holdings statement - information describing the issues of serials available in the Library stacks.
- Place of publication.
- Publisher's name.
- Series note - the title or description of the series of which the entry is a part.
- Miscellaneous drop notes - such information as caption title, titles of specific editions, thesis information, and form of reproduction.
- The call number of the National Library of Medicine.
- Any additional entries and cross-references.

INPUT AND CONVERSION SUBSYSTEM

Introduction

Reference material must be selected and prepared for entry into the computer. It must be done timely, to assure that the reference information in storage is always kept up to date. Requests for bibliographies must also be prepared for entry into the computer. The requests must be precisely defined to enable the computer to retrieve the relevant citations from storage and to format them properly for printing. These are the

Transplantation of homografts of sclera on the eye of the dog. including volume determinations: an experimental study. — Title

Author — JOHNSON WA. [Minneapolis] 1961. 450 illus. — leaves

Thesis (M.S. in Ophthalmology) - Univ. of Minnesota. — Citation

Misc. Note #1 — Thesis

Misc. Note #2 — Typewritten copy.

Call No. — W4A J72t

Transplantations Sclera	Index Medicus Headings	Tags
Dogs	Non-Index Medicus Heading	
Thesis	Type of Article	

Figure 3. Hypothetical Non-Journal Article Record

major tasks of the Input and Conversion Subsystem. The activities required to perform may be divided into two categories, those associated with updating the computer files and those associated with file searches.

Updating The Files

Journals received by the Library will be distributed to a staff of indexers (see Figure 4). For each article, an indexer will prepare a "data sheet" on which he specifies all of the descriptive tags and certain other unit record elements. The information on the data sheets and other information from the journals (such as journal title) will be utilized by a staff of input typists to prepare the unit records. The journal information will be typed only once, followed by all of the unit records for the journal, which will equal the number of articles having data sheets.

The input typists will use typewriters that simultaneously produce typed copy and punched paper tape that contain the same information, though in different form. Each machine will have a program control feature that will simplify and reduce the typing task by automatically providing control characters on the typed copy and the paper tape and producing a fixed format on the typed copy. Typically, when the typist depresses the control key prior to typing a unit record element, the typewriter carriage will return, control characters will be typed, and the typewriter will tab once and stop. Each unit record element will have a unique control character. The computer will be able to identify the different characters when the paper tape is read in, thereby identifying the different elements.

Periodically, reels of punched paper tape, the typed copy, journals and data sheets will be moved together to the proofing area. The proofreading staff will check the typed copy and indicate needed corrections. A separate paper tape containing all the corrections and necessary control information will be prepared. A complete unit record, any element of a unit record, or individual words will be automatically added, deleted, or changed by the computer on the basis of the instructions on the "correction tape". When the original reels of paper tape and their associated correction tapes are read into the computer, the computer will not only make the specified corrections but also look for other uncorrected errors. Any errors that it finds will be printed out on the computer printer. This error information will be reviewed primarily by the proofreaders to determine what corrective action is indicated.

File Searches

In addition to preparing unit records for entry into the computer, the Input and Conversion Subsystem will also process requests for the retrieval of stored information that is already in the computer files. These requests will be of two broad types: (1) those calling for the preparation of periodically printed publications, such as Index Medicus, and (2) those calling for the retrieval of bibliographic reference in-information to satisfy the individual needs of Library customers, namely, requests calling for the preparation of recurring and demand bibliographies.

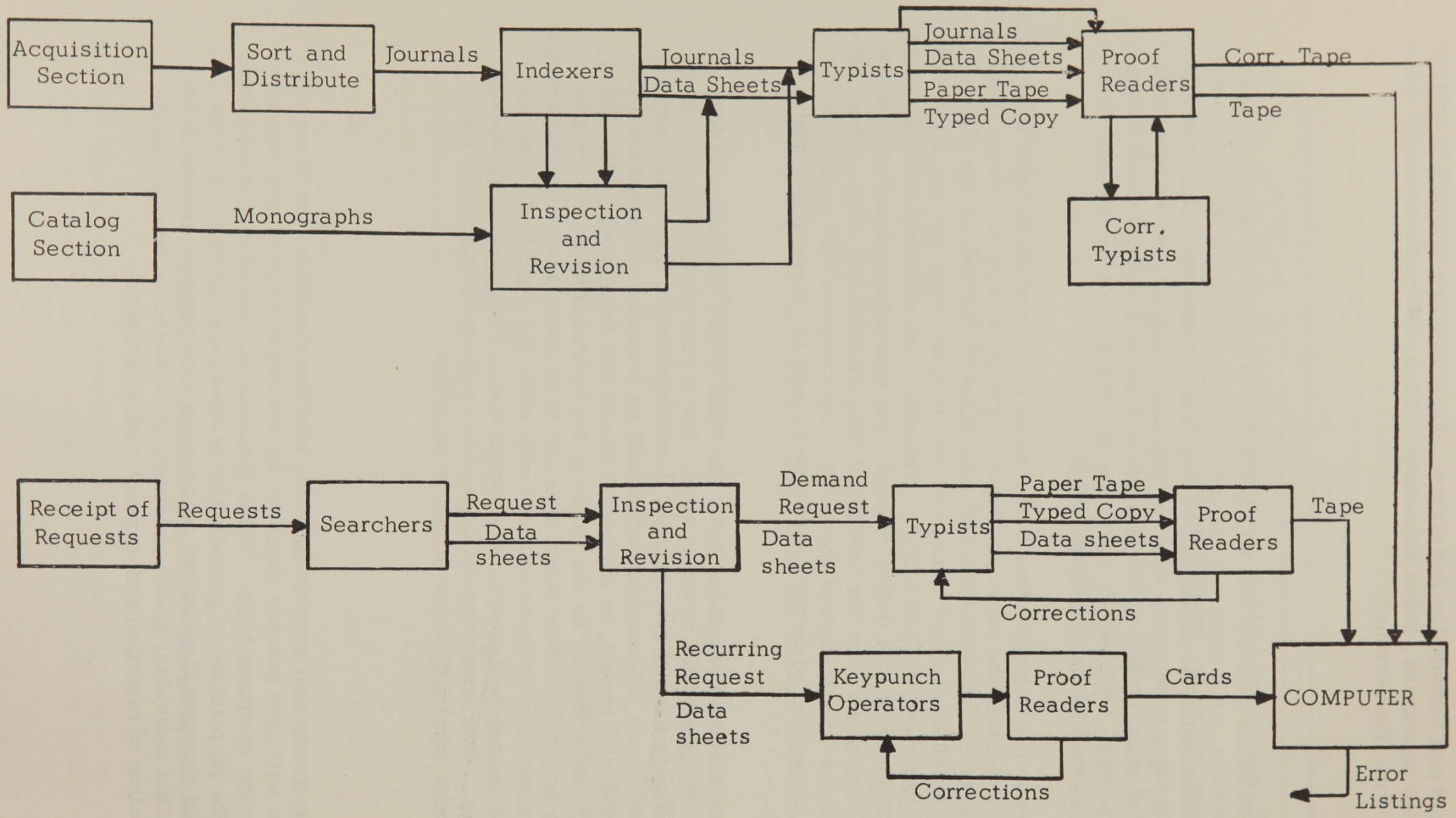


Figure 4. Input and Conversion Subsystem

As demand requests (for bibliographic information) are received by the Library, they will be sorted and distributed to a staff of bibliographic specialists called searchers. These searchers will translate the requests into the language, syntax, and format required by the system. In the same manner that indexers select those subject headings that adequately classify an entry, the bibliographic searchers will select the headings that identify the entry for retrieval. The retrieval criteria elements will be listed and organized into the form of logical expressions following the detailed demand request rules.

The demand bibliographic requests will be typed/punched by the input typists in the same manner as the citations from medical literature. The typed copy will be proofread, corrections made, and the paper tape read into the computer. The information retrieved by the computer to satisfy these search requests will be screened by the searchers to make sure that it satisfies the original requests before being transmitted to the original requester.

Transition Period

During the first year of MEDLARS operation, the existing mechanized system will continue to prepare Index Medicus and associated publications. The unit records generated will, however, be fed into the computer's files for use in meeting future demand bibliographic requests. Early in 1964, such requests for bibliographies will be met by the computer. When MEDLARS finally and completely supersedes the existing system, it will have a backlog of stored information to call upon in filling idividual demand bibliographic requests.

MANIPULATION SUBSYSTEM

Introduction

The high-speed, digital computer, which is the heart of the Manipulation Subsystem, will accept the unit records on paper tape as supplied by the Input and Conversion Subsystem, check the tape records for the presence and correctness of those elements for which such checks are possible, perform some preprocessing and compressing operations to speed up the subsequent processing, and store the unit records on magnetic tape. In response to search requests, it will search its cumulation of unit records for those that qualify for retrieval, edit and compose them for output, and deliver magnetic tape to the Output Subsystem.

In the performance of these tasks, the Manipulation Subsystem will utilize the Honeywell 800 computer and associated tabulating equipments, and will be staffed by Data Processing Section supervisory personnel, programmers, equipment operators, and maintenance personnel.

The unit records and other data will be stored on reels of magnetic tape. These tape files will be prepared, updated, and used by various computer programs which

manipulate the files to provide the desired outputs. These programs will consist of a number of component modules, each of which will be used on a demand basis, independently of the other modules. Operations will be scheduled in the interests of efficiency. In the material which follows, first the magnetic-tape files and then the program modules are described.

Magnetic-Tape Files

The most important magnetic-tape files are described briefly. Working files which will be temporary in nature are not included.

1. MESH Master File. This file will contain one record for each tag, "See" entry, "See Under" entry, "See Also Related" entry, and "See Also Specific" entry.

2. Annual MESH File. This file will have a format similar to the MESH Master File but will contain only main heading and "See" and "See Under" data.

3. MEDLARS Dictionary Tape (MDT). This tape will consist of the following small, separate files all contained on one reel of magnetic tape:

   - Daily MESH File. This file will have a format similar to the MESH Master File but will contain only main headings and no cross-reference data.

   - Journal Record File (JRF). This file will contain all pertinent journal information including the dates when journals are due for computer entry and their places of publication.

   - Lands File. This file will contain the language abbreviation codes and form tags for use as subheadings in listing monographs in Index Medicus.

   - Recurring Bibliographic Parameter File. This file will contain all the parameters required to describe completely the format and content of recurring bibliographies.

   - Recurring Bibliography Print Heading File. This will contain special headings used only for the printing of recurring bibliographies.

   - Recurring Bibliography Selection Criteria File. This will contain the criteria for selecting citations for inclusion in specific recurring bibliographies.

4. Compressed Citation File (CCF). This file (see Figure 5) will contain all of the unit records that the system has in storage. It will be searched to obtain the citations needed to service demand search requests. The citations will be compressed - that is, certain unit record elements will be stored in code form - primarily to facilitate economical, high-volume storage.

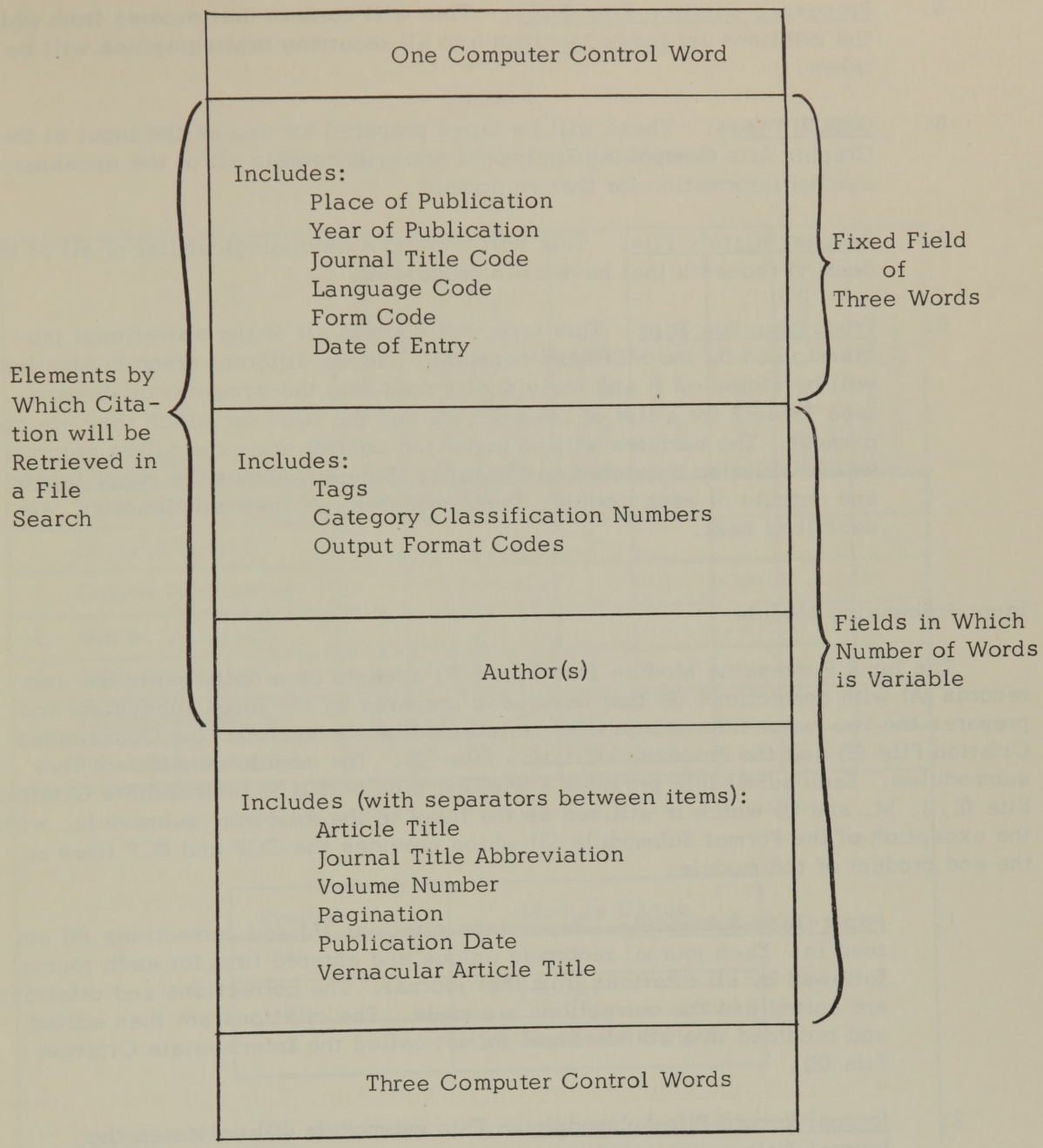


Figure 5. The Compressed Citation File: Arrangement of the Elements of One Citation

5. Processed Citation File (PCF). This will contain unit records from which the citations for Index Medicus and all recurring bibliographies will be taken.

6. GRACE Tapes. These will be tapes prepared for use as the input to the Graphic Arts Composing Equipment and will contain all of the necessary control information for that equipment.

7. Request History File. This will contain a chronological list of all of the demand requests that have been processed.

8. Production Run Tape. This tape will contain all of the operational programs used by the MEDLARS computer. Seven different program modules will be stored on it and individually read into the computer on demand (see Table 2 for a list of the modules and the files for which they are concerned). The modules will be under the control of an executive program, which will also be stored on the tape. Figure 6 shows the major inputs and outputs of each module; these modules, as they will function, are described next.

Input Processing Module

The Input Processing Module (see Figure 7) accepts on a daily basis the unit records (A) with corrections (B) that have been prepared by the Input Subsystem and prepares the two major information files maintained by the system: the Compressed Citation File (T) and the Processed Citation File (U). The module consists of five submodules. Each submodule provides a working tape output or Intermediate Citation File (E, J, M, and Q) which is utilized as the input to the following submodule, with the exception of the Format Submodule (S) which provides the CCF and PCF files as the end product of the module.

1. Paper-Tape Submodule. Paper-tape citations (A) and corrections (B) are read in. Each journal record is unique and entered first for each journal followed by all citations from that journal. The corrections and citations are sorted and the corrections are made. The citations are then edited and recorded in a standardized format called the Intermediate Citation File (E).

2. Journal Record File Submodule. This submodule (G) validates the Journal Title codes received on input and adds the correct Journal Title abbreviation to each citation. Simultaneously, any journals which are overdue are detected and reported (F). For each journal, the journal record file (I) is updated with the latest volume and issue number received along with its publication date.

3. MESH Submodule. This submodule (L) validates all of the tags received on input and encodes the citations with equivalent compressed codes which will be used for retrieval and processing. At the same time, usage data on MESH Headings and other tags are accumulated (N).

Table 2

File Usage for Each Program Module

Program Module	MESH Master File	Annual MESH File	MEDLARS Dict. Tape *	Compressed Citation File	Processed Citation File	GRACE	Request History File	Demand Search Output Tape
1. Input Processing			I	O	O			
2. Demand Search			I	I			O	O
3. Output Processing		I	I		I	O		
4. Report Generator						O		I
5. Statistical			I	I				
6. MESH Generator	X	X	X					
7. File Maintenance				X	X			

Symbol	Module Usage
I	Input
O	Output
X	Input and Output

* Contains the daily MESH file, journal record file, the recurring bibliography parameter file, and language file.

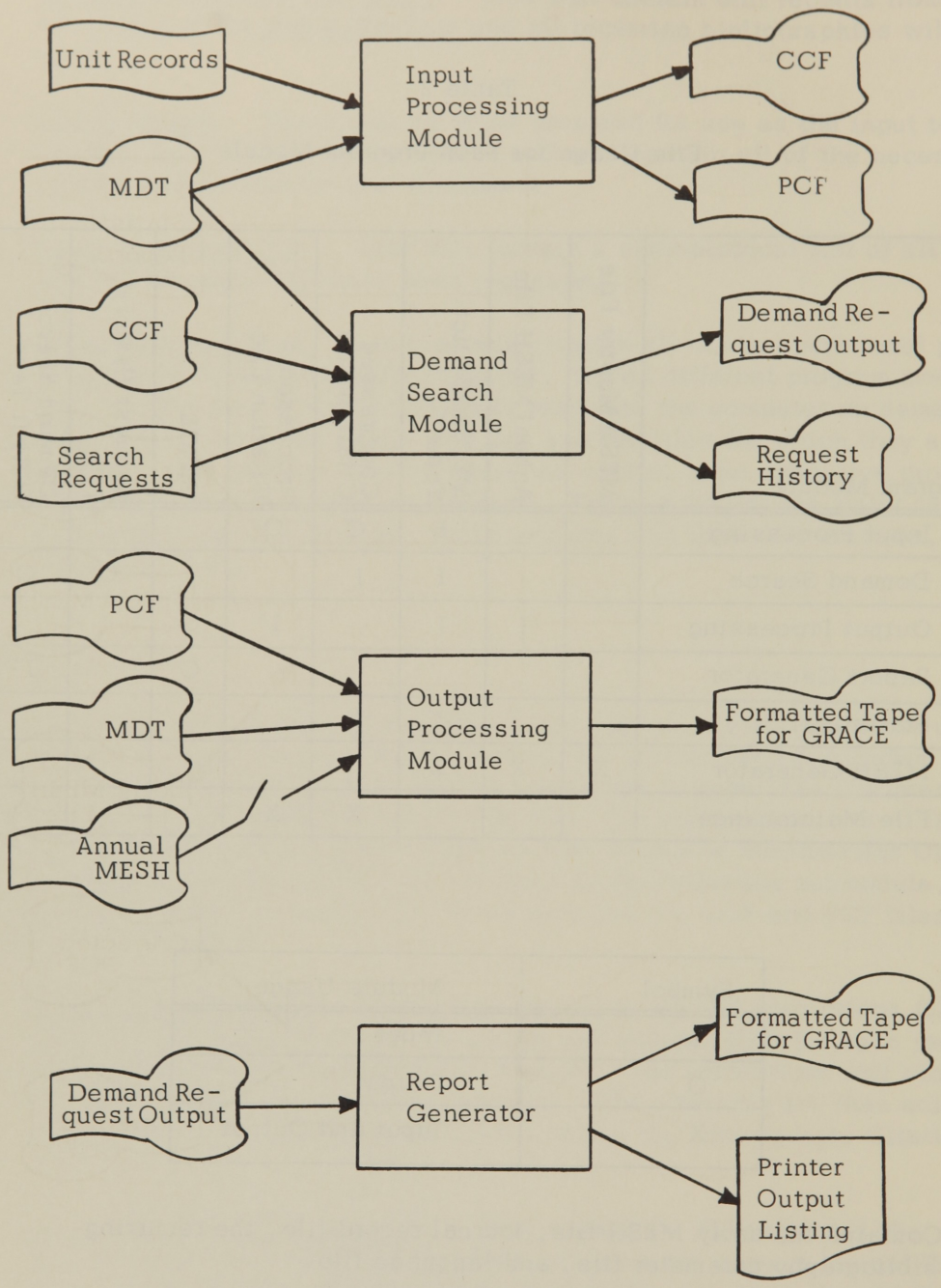


Figure 6. Major Inputs and Outputs of Each Program Module
(Page 1 of 2)

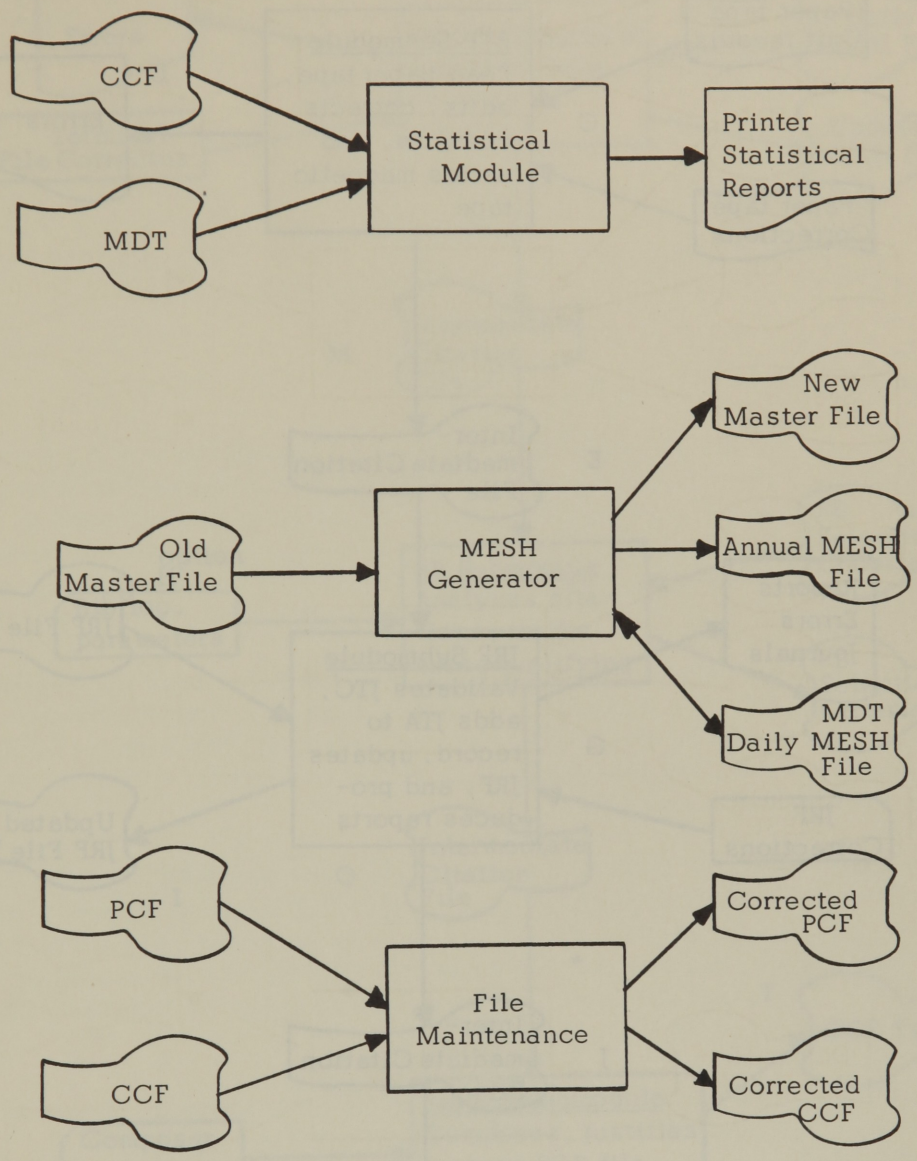
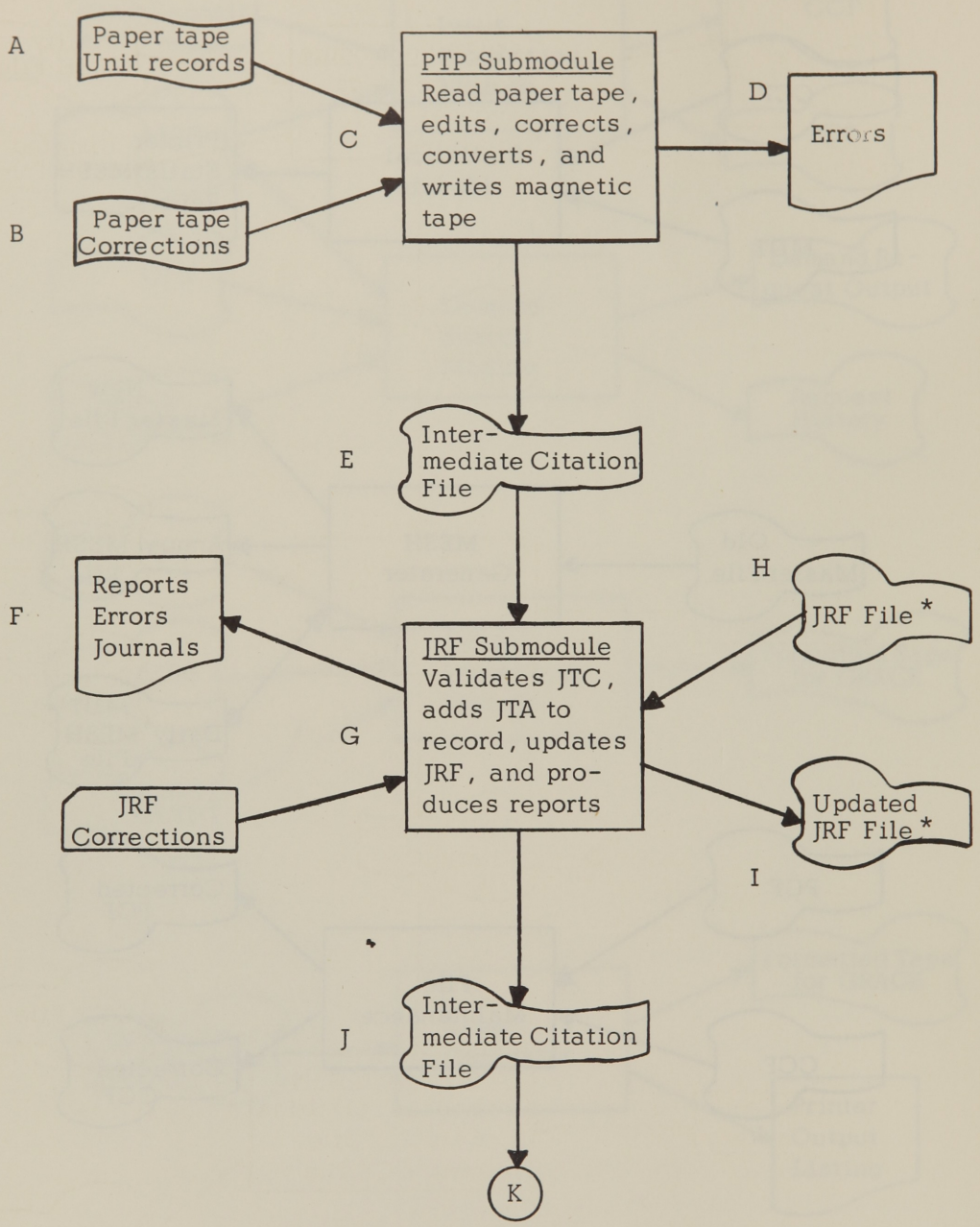
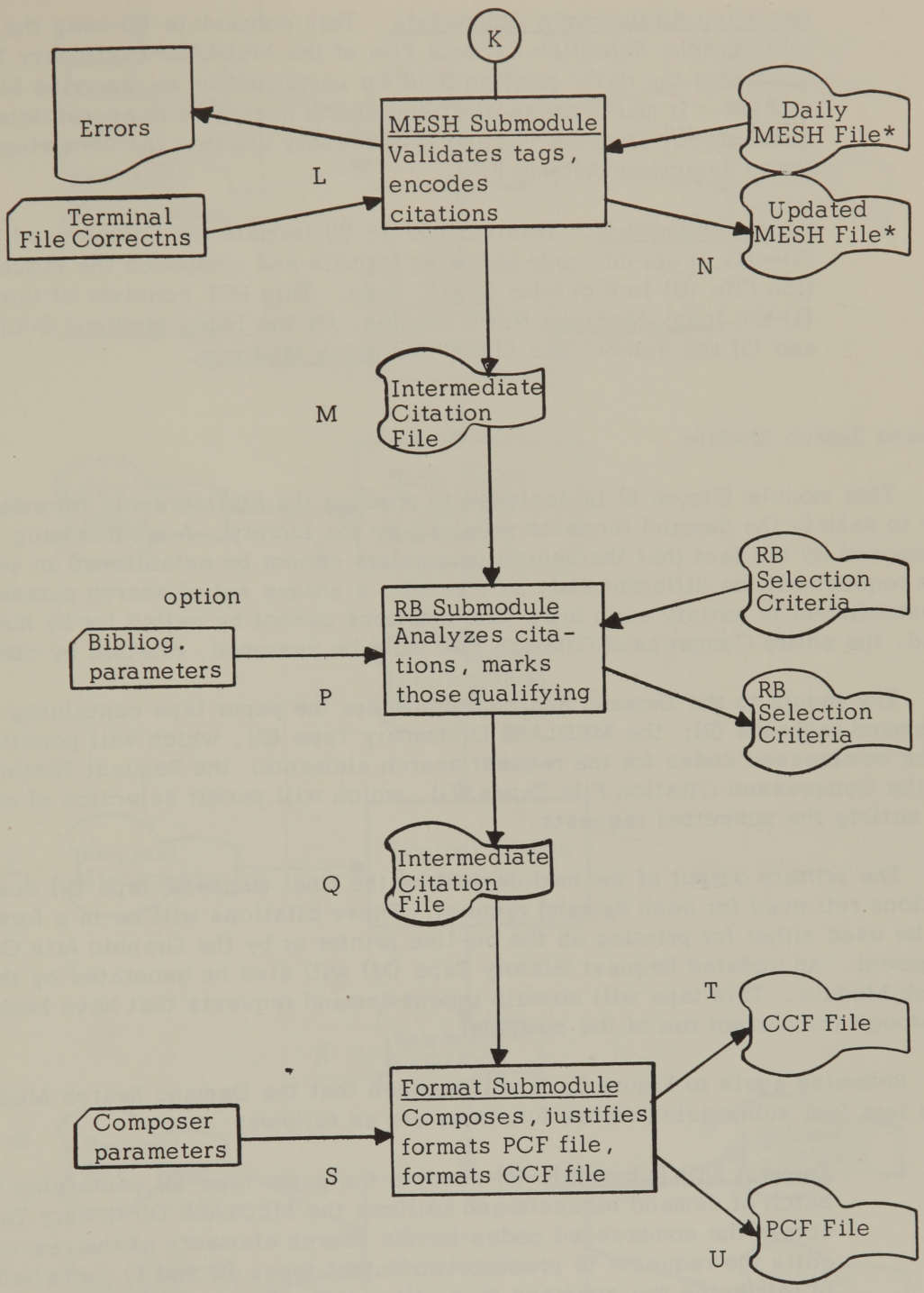


Figure 6. Major Inputs and Outputs of Each Program Module
(page 2 of 2)



* - MEDLARS Dictionary Tape

Figure 7. Block Diagram for Input Processing Module
(Page 1 of 2)



* - MEDLARS Dictionary Tape

Figure 7. Block Diagram for Input Processing Module
(Page 2 of 2)

4. Recurring Bibliography Submodule. This submodule (P) uses the Recurring Bibliography Selection Criteria File of the MEDLARS Dictionary Tape in analyzing the daily citation load for participation in recurring bibliographies. It marks those citations which qualify with appropriate recurring bibliography numbers (Q) and if necessary updates the Recurring Bibliography Selection Criteria File.

5. Format Submodule. This submodule (S) formats the Compressed Citation File (T) in six-bit code and also formats and composes the Processed Citation File (U) in eight-bit GRACE code. This PCF consists of three sections: (1) the Index Medicus Name Section, (2) the Index Medicus Subject Section, and (3) the Author "See Lines" for Index Medicus.

Demand Search Module

This module (Figure 8) is designed to produce the bibliographic information necessary to satisfy the demand requests received by the Library. A programming limitation is imposed by the fact that the search parameters cannot be established in advance. Each request will be different and, as a result, a unique set of search parameters must be established to satisfy each one. The citations cannot be called for by number; instead, the entire Compressed Citation File must be searched, citation by citation.

The inputs to the Demand Search Module are the paper tape containing the batch of demand requests (B); the MEDLARS Dictionary Tape (C), which will permit looking up the compressed codes for the request search elements; the Request History Tape (M); and the Compressed Citation File Tapes (G), which will permit selection of citations that will satisfy the submitted requests.

The primary output of the module will be the final magnetic tape (N) containing the citations retrieved for each demand request. These citations will be in a format that may be used either for printing on the on-line printer or by the Graphic Arts Composing Equipment. An updated Request History Tape (M) will also be generated by the Demand Search Module. This tape will contain recent demand requests that have been processed up through the current run of the computer.

Referring again to Figure 8, it will be seen that the Demand Search Module is divided into four submodules whose functions are as follows:

1. Request Edit Submodule (A) accepts the paper tape (B) containing the daily batch of demand requests and utilizes the MEDLARS Dictionary Tape (C) to obtain the compressed codes for the search elements of the requester. It edits the requests to produce two output tapes (D and E) formatted for use in retrieving the citations that will satisfy the requests. It also selects from the incoming search-request statements their most significant elements in terms of retrieval criteria, placing the elements in computer memory (F).

2. High-Speed Search Submodule (H) sequentially searches every record in the Compressed Citation File (G), using the table of significant elements (F) selected from the requests as retrieval criteria. It produces an

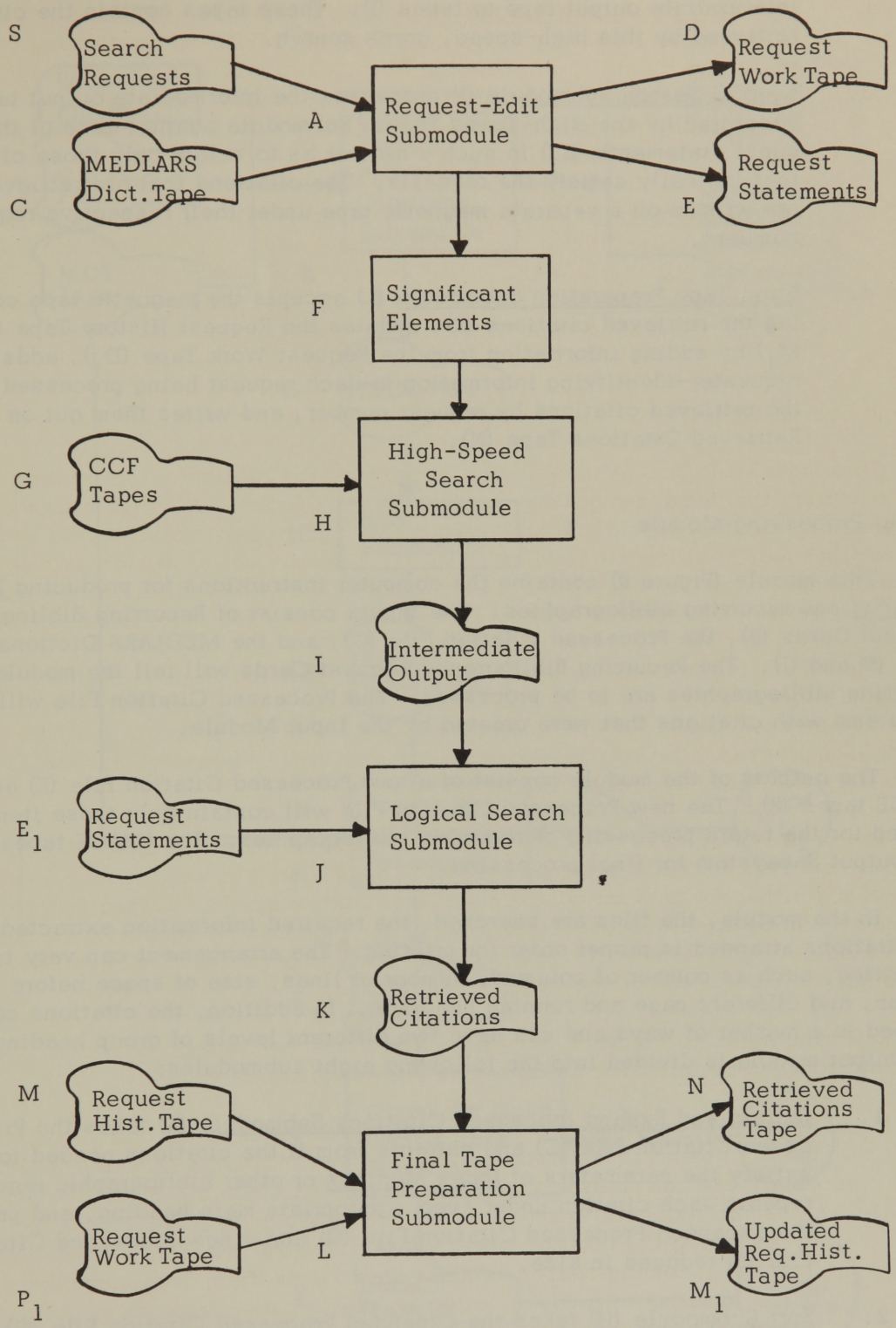


Figure 8. Block Diagram for Demand Search Module

intermediate output tape or tapes (I). These tapes contain the citations retrieved by this high-speed, gross search.

3. Logical Search Submodule (J) processes the intermediate output tape (I) generated by the High-Speed Search Submodule against each of the request statements (E_1) in such a manner as to select only those citations that logically satisfy the requests. The citations that are retrieved (K) are written on a separate magnetic tape under their respective request numbers.

4. Final Tape Preparation Submodule (L) accepts the magnetic tape containing the retrieved citations (K), updates the Request History Tape (M and M_1) by adding information from the Request Work Tape (D_1), adds requester-identifying information to each request being processed, sorts the retrieved citations by request number, and writes them out on the Retrieved Citations Tape (N).

Output Processing Module

This module (Figure 9) contains the computer instructions for producing Index Medicus and recurring bibliographies. The inputs consist of Recurring Bibliography Request Cards (B), the Processed Citation File (C), and the MEDLARS Dictionary Tape (D and U). The Recurring Bibliography Request Cards will tell the module what recurring bibliographies are to be processed. The Processed Citation File will contain items with citations that were created by the Input Module.

The outputs of the module consist of a new Processed Citation File (E) and the GRACE tapes (Z). The new Processed Citation File will contain only those items needed for the future processing of recurring bibliographies. The GRACE tapes go to the Output Subsystem for final processing.

In the module, the files are searched, the required information extracted, and the citations arranged in proper order for printing. The arrangement can vary in page formatting, such as number of columns, number of lines, size of space before each column, and different page and running headings. In addition, the citations can be grouped in a number of ways and can have two different levels of group headings. The Output module is divided into the following eight submodules:

1. Extract and Expand Processed Citations Submodule (A) takes the Processed Citation File (C) and extracts from it the citations needed to satisfy the parameters of Index Medicus or other bibliographic requests, repeats each citation under every appropriate main heading, and prepares the Expanded Processed Citation File (G) and a new Processed Citation File (E) reduced in size.

2. Sort Submodule (H) takes the Expanded Processed Citation File (G) and ranks its citations in the necessary sequence for printing.

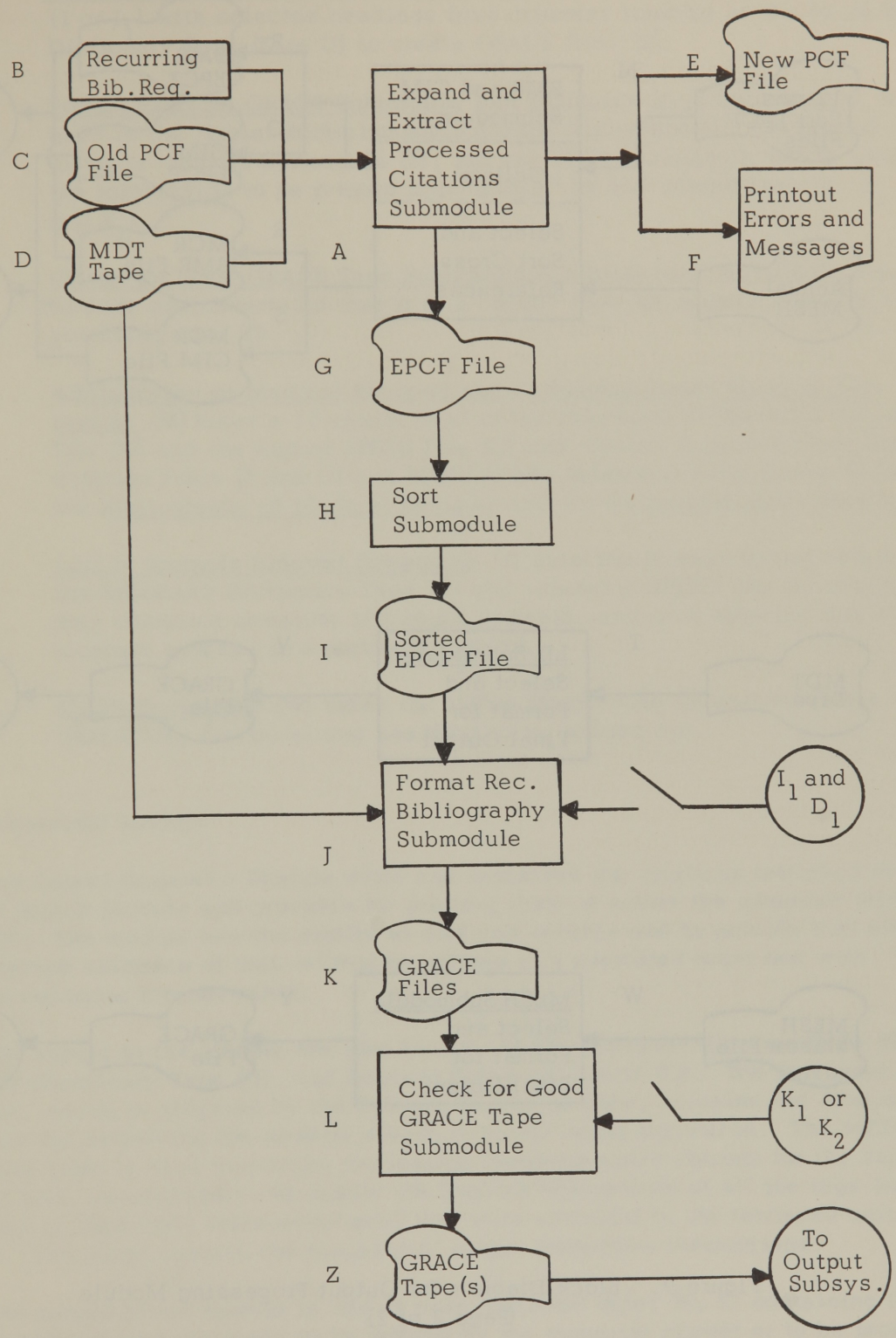


Figure 9. Block Diagram for Output Processing Module
(Page 1 of 2)

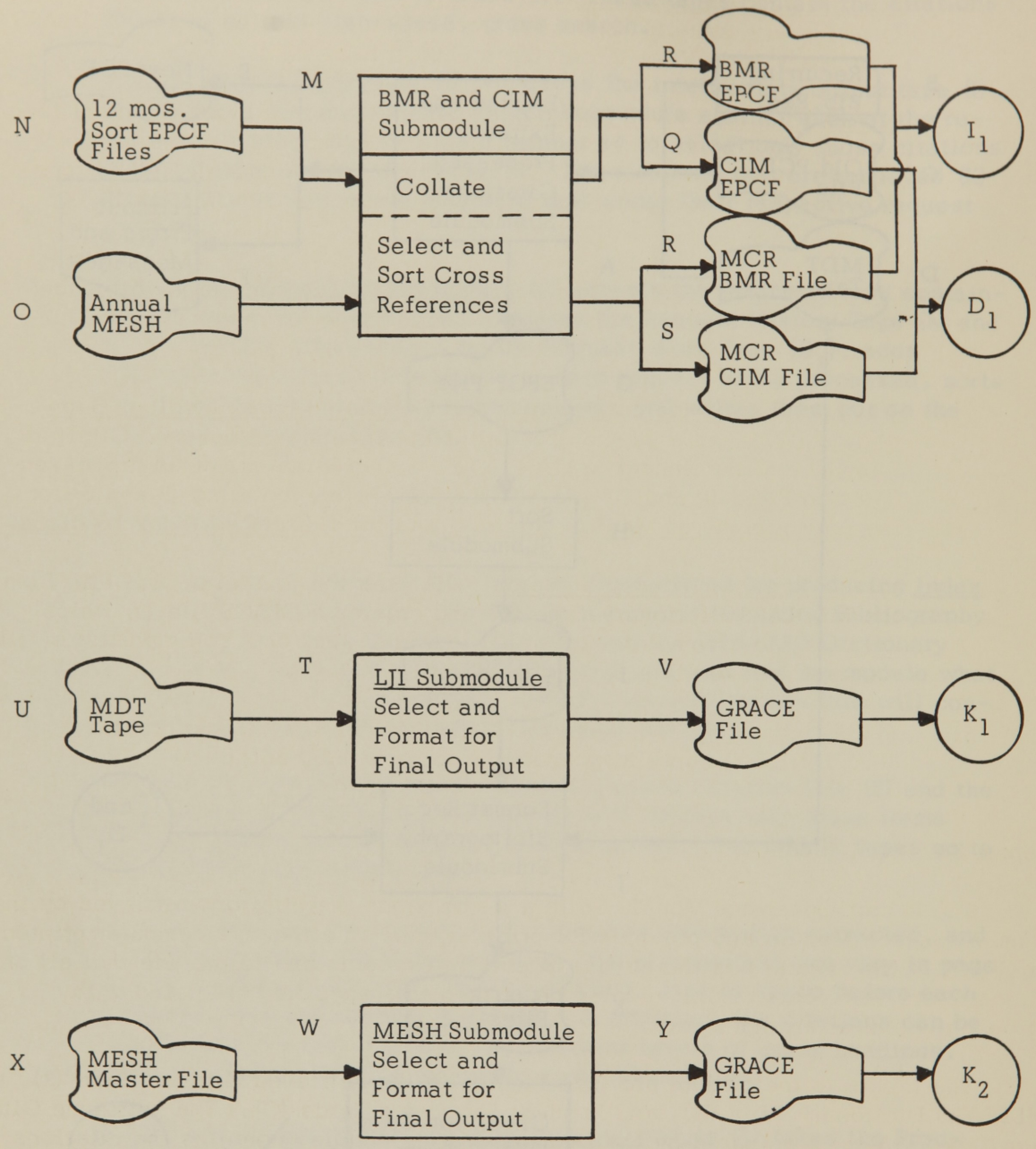


Figure 9. Block Diagram for Output Processing Module
(Page 2 of 2)

3. Format Recurring Bibliography Submodule (J) combines the sorted file (I or I_1) with selected headings from different subfiles of the MEDLARS Dictionary Tape (D or U) to create GRACE files (K).

4. Format Page for GRACE Submodule (not indicated on block diagram) is used in the Format Recurring Bibliography Submodule (J), the List of Journals Indexed Submodule (T), and the MESH submodule (W) to format the information to be printed into "pages" for the GRACE files (K, K_1 and K_2).

5. Check for Good GRACE Tape Submodule (L) reads each tape with page formats and ascertains that it is a "good" tape for release to the Output Subsystem (Z).

6. Bibliography of Medical Reviews and Cumulated Index Medicus Submodule (M) takes a 12-month total of the Expanded Processed Citation File (N) and the annual MESH File (O) and creates Expanded Processed Citation Files (P and O) and MESH Cross Reference Files (R and S) for the Bibliography of Medical Reviews and the Cumulated Index Medicus.

7. List of Journals Indexed Submodule (T) uses the Journal Record File from the MEDLARS Dictionary Tape (U) and creates a GRACE file (V) which may contain a complete list of all journals, and/or a selected list of all journals indexed in a particular period.

8. MESH Submodule (W) takes the MESH Master File (X) and creates a GRACE file (Y) containing headings to be printed out.

Report Generator Module

The Report Generator Module edits and composes the citations retrieved by the Demand Search Module and provides for printing them on either the computer printed or GRACE. The module has the ability to sort unit records and to provide that all or only selected elements of unit records be printed in a specified order and with differing columnar arrangements.

The inputs to the module are (see Figure 10) the Retrieved Citation Tape (R), the MEDLARS Dictionary Tape (H), and Request Parameter Cards (C). The Retrieved Citation Tape, which is prepared by the Demand Search Module, contains the citations necessary for satisfying the demand search requests being processed. The MEDLARS Dictionary Tape is used to acquire the foreign language abbreviations for the search requests and, as necessary, to supply the English equivalents of all the tags (which are stored in their code representations) that were assigned to the retrieved unit records. The cards contain the requesters' output formatting instructions.

The output of the module is one or more magnetic tapes (K, L) containing formatted citations that are ready to be printed by the computer printer or to be processed by GRACE for high-quality typography. A flexible printed format is available

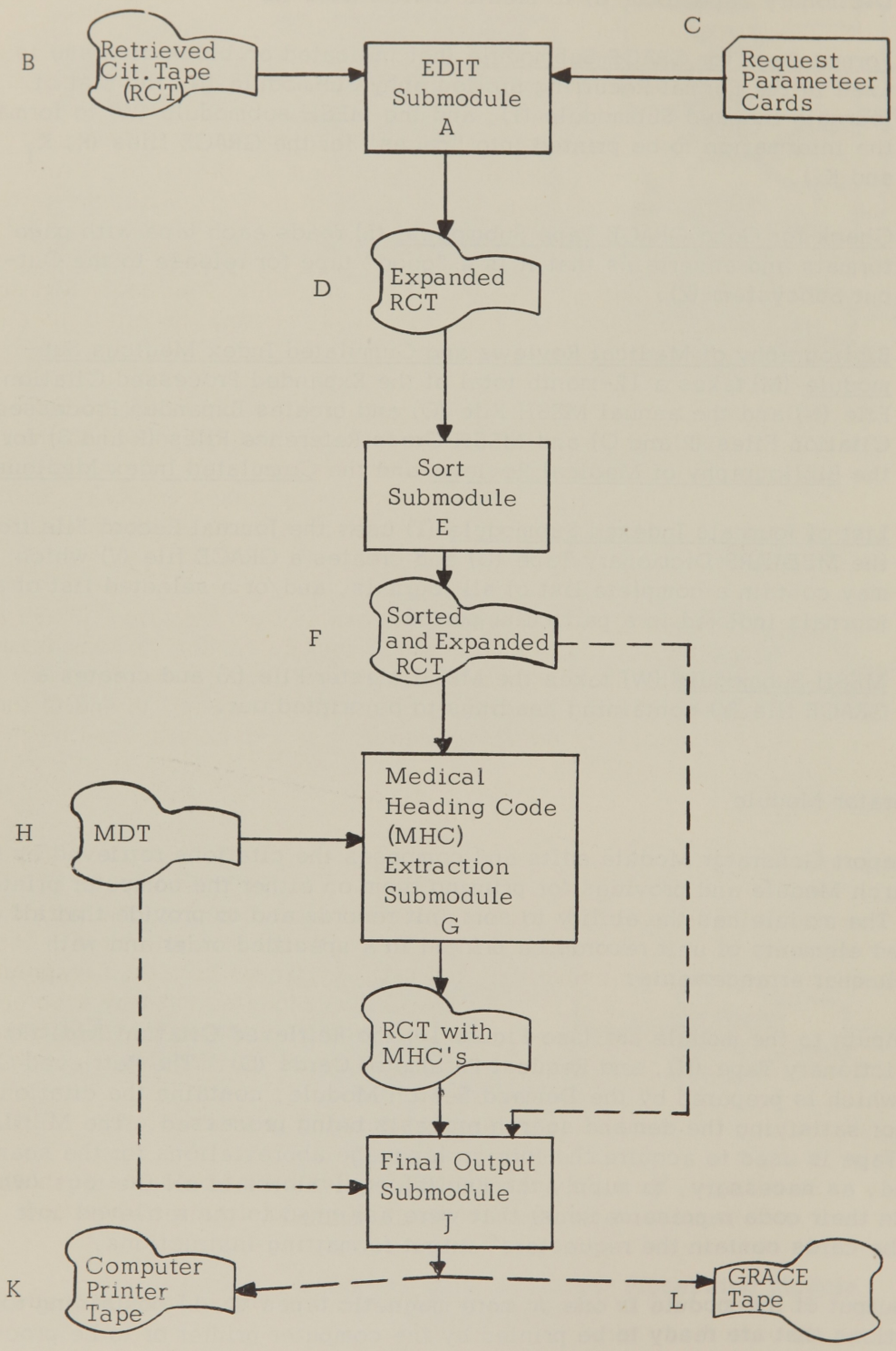


Figure 10. Block Diagram for Report Generator Module

to the requester, including the choice of either high-quality or standard typography. It is possible for a requester to specify, for example, that page headings be provided, that citations be grouped by major headings, and that the columns be a particular width with, say, three columns to a page.

The Report Generator Module is composed of the following four submodules:

1. Edit Submodule (A) accepts the citations (B) retrieved by the Demand Search Module, expands them and formats them according to the requesters' instructions (C), and writes them on tape (D) with sort keys added.

2. Sort Submodule (E) takes the output of the preceding submodule and sorts the citations according to the specified keys, writing the citations on tape (F) in their final printing sequence.

3. Medical Heading Code Extraction Submodule (G) accepts the sorted citations (F) and utilizes the MEDLARS Dictionary Tape (H) to add to the retrieved citations for certain demand search requests all of the tags assigned to the articles cited. This submodule is used to process citations for only those requests for which "extraction" of tags is specifically requested.

4. Final Output Submodule (J) composes citations from the tapes (F and I) produced by either of the two preceding submodules (for the tape from the Sort Submodule, also adds required medical subject headings (H) into complete lines of "print", providing on magnetic tape(s) (K and L) the physical arrangement for the final printed output to be produced by the computer printer or GRACE.

Statistical Module

The Statistical Module produces statistical reports on request. The period of time that a report may cover is flexible and there are optional report formats. The data from which the reports are produced are normally extracted from the Compressed Citation File, which is prepared by the Input Processing Module, but may also be taken from the MEDLARS Dictionary Tape. The design is flexible so that additional reports may be obtained through only minor modifications of the module.

The reports which this module is presently designed to produce are:

- A listing of the title abbreviations and publication dates of those journals actually indexed for Index Medicus during a specified period.

- The frequency of use of each main heading whose total exceeds a specified number.

- The number of English-language or foreign-language articles indexed, broken down by language abbreviation, and including a subtotal for each foreign language.

- The number of citations printed in the Subject Section, the number printed in the Name Section, and the number of junior author cross-references.

- The number of articles and citations appearing in Bibliography of Medical Reviews.

The Statistical Module is divided into four submodules (Figure 11). In the first submodule (A), the Compressed Citation File (B) is read and the data requested (C) is extracted from it and placed onto a work tape and into temporary storage in memory. The second submodule (D) sorts the extracted data. The third submodule (E) processes the sorted information against the Journal Record and MESH files on the MEDLARS Dictionary Tape (F), generating a tape to be printed. In a final and altogether separate submodule (G), the tape is printed on the on-line printer.

Other Program Modules

The remaining two modules on the System Program Tape are:

- The MESH Generator Module

- The File Maintenance Module

These two modules perform the system's "housekeeping" functions; they are used solely for correcting and updating the MESH and Citation files, and only indirectly on the output of the system.

OUTPUT SUBSYSTEM

The principal function of this subsystem will be to transform, by means of a Graphic Arts Composing Equipment (GRACE), automatic film processing and ancillary equipments, the coded magnetic-tape output of the Manipulation Subsystem into processed film from which a printer can produce a publication such as Index Medicus (see Figure 12). (The responsibility for actually preparing the offset lithographic plates from the film and printing the designated publication is delegated to another organization.) Publications not requiring high-quality typography will also be produced, using the computer's on-line mechanical printer.

Graphic Arts Composing Equipment

The Graphic Arts Composing Equipment will convert information from magnetic tape into intelligible images or characters on film. Its input will be electrical

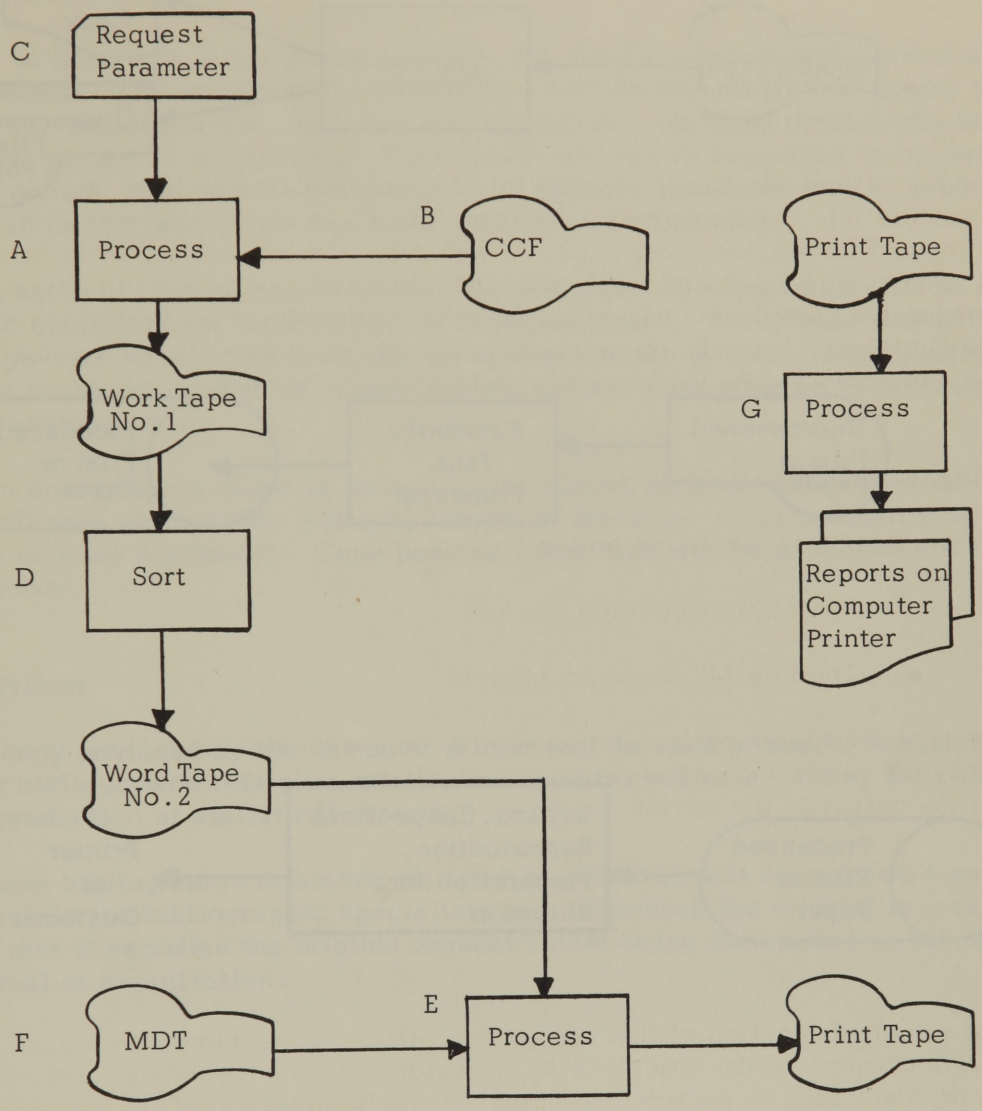


Figure 11. Block Diagram for Statistical Module

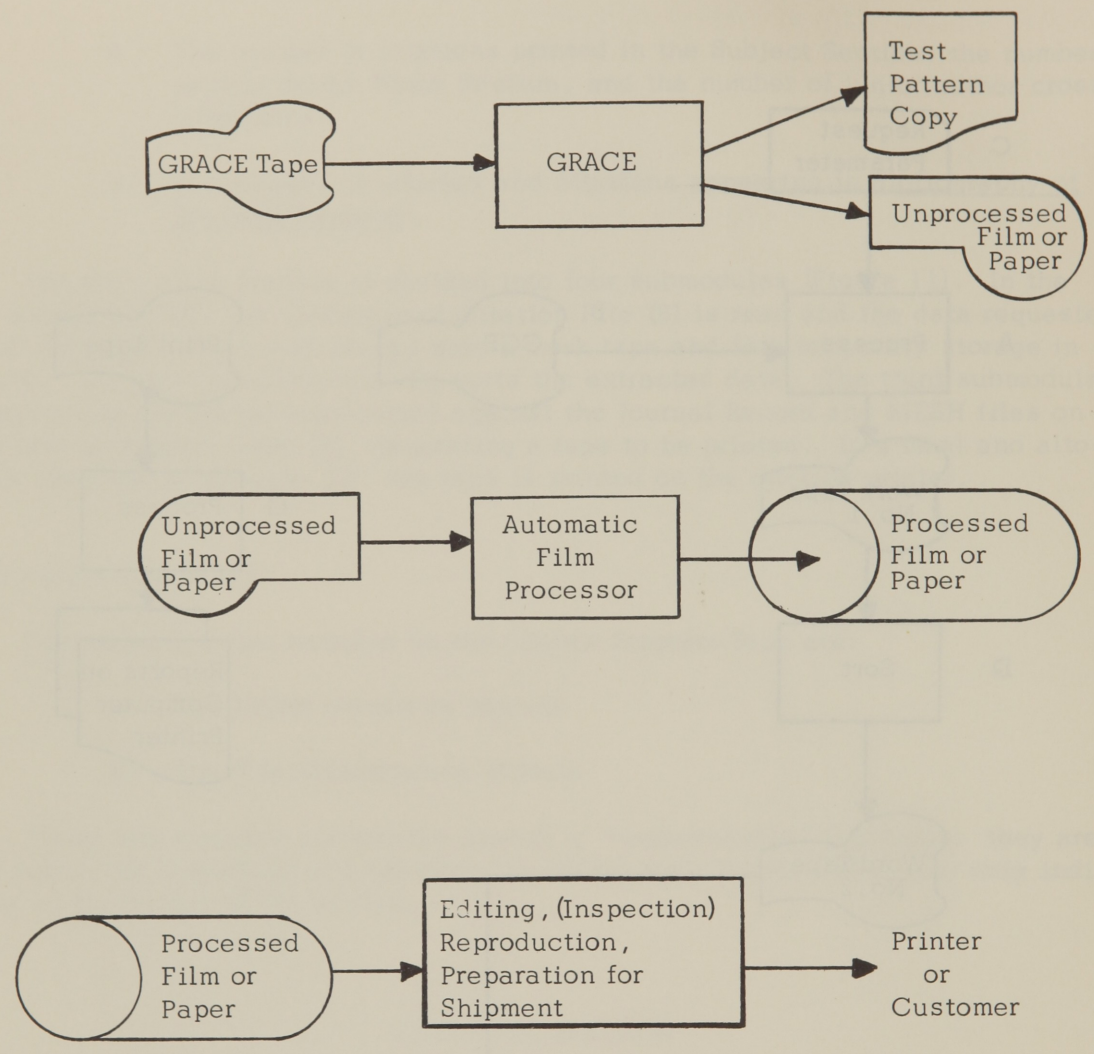


Figure 12. Block Diagram for Output Subsystem

signals from a magnetic-tape transport operating fully under its control. These signals will arrive in bursts of coded characters which represent a full line of legible type and any combination of the 226 different alphanumeric symbols that will be provided will be accepted. GRACE will set this "type" on film or paper at a rate in excess of 440 characters per second. An automatic film processor will process the film and paper that GRACE prepares.

Prior to initiation of any production run, the GRACE operator will calibrate GRACE and the automatic film processor by composing a test pattern on film (or paper) and processing the product in the automatic film processor. Iterative checks will be made until the test pattern is acceptable. These tests will not be restricted to the pre-production period; they will be performed on a regularly scheduled basis, such as after each 50 to 100 feet of film has been exposed and processed.

Each batch of film (or paper) leaving the automatic film processor will be inspected and classified as "acceptable" or "unacceptable". Acceptable film will be cut into page-size sheets and packaged for delivery to the printer. Unacceptable film will be cause for corrective action of some nature, depending on the reason why the film is unacceptable.

When photographic paper is chosen as the output medium reproduction of the output copy will take place at the National Library of Medicine using commercially available office copying equipment. Copy passing inspection will be packaged and mailed to the customer.

Computer Printer

The copy produced by the computer printer will be used primarily to satisfy the request for bibliographic reference material on specific subjects. Other uses will include the production of statistical reports.

Demand bibliographic reference material so produced will be inspected and then referred back to the bibliographic specialists who processed the original request to make sure that it satisfies the original request before being transmitted to the requesting individual or organization.

III

MEDLARS AT WORK

INTRODUCTION

The following chapters of this report discuss in technical terms what MEDLARS will do, how it will do it, and the design and management considerations underlying the system's development. This chapter projects the reader into the future to illustrate some of the operations of the system. Its purpose is to provide a clearer understanding of how the human and machine components of the system will function.

Scenes that could take place on any day at the National Library of Medicine after MEDLARS is operational are briefly described. It will be assumed that the system has been implemented precisely as it is now designed, even though relatively minor modifications may be made. Since the descriptions will contain a minimum of technical detail, the reader may wish to refer occasionally to other portions of this report for related technical information.

THE INPUT AND CONVERSION SUBSYSTEM

We will take an imaginary tour of MEDLARS' operations and begin by observing how the incoming journals are processed (see Figure 1, Journal Processing Flow Chart). To do this, we proceed to where the indexing staff is located on the Library's first floor.

JOURNAL INDEXING

A few moments ago the Head of the Indexing Unit received 60 journals from the Library's Technical Services Division. He has just finished distributing them, taking into account journal language and content, to the 18 members of the indexing staff. One of the staff is beginning to index a journal.

The indexer turns to the first article in the journal and scans it. He then inserts a "data sheet" (see Exhibit A) into his typewriter and types his indexer-code number, the article's pagination, and nine medical subject headings that he decides are descriptive of the article. Removing the data sheet from the typewriter, he takes a pencil and puts check marks in front of several of the medical subject headings, to indicate that the article should be cited under them in Index Medicus, and also in front of two of the other tags provided on the data sheet. Now he clips the data sheet to the page opposite the one on which the article begins and turns to the next article. When he has prepared data sheets for all of the journal's articles and has them clipped in place, he puts the journal in the "out" container on his desk and reaches for another journal.

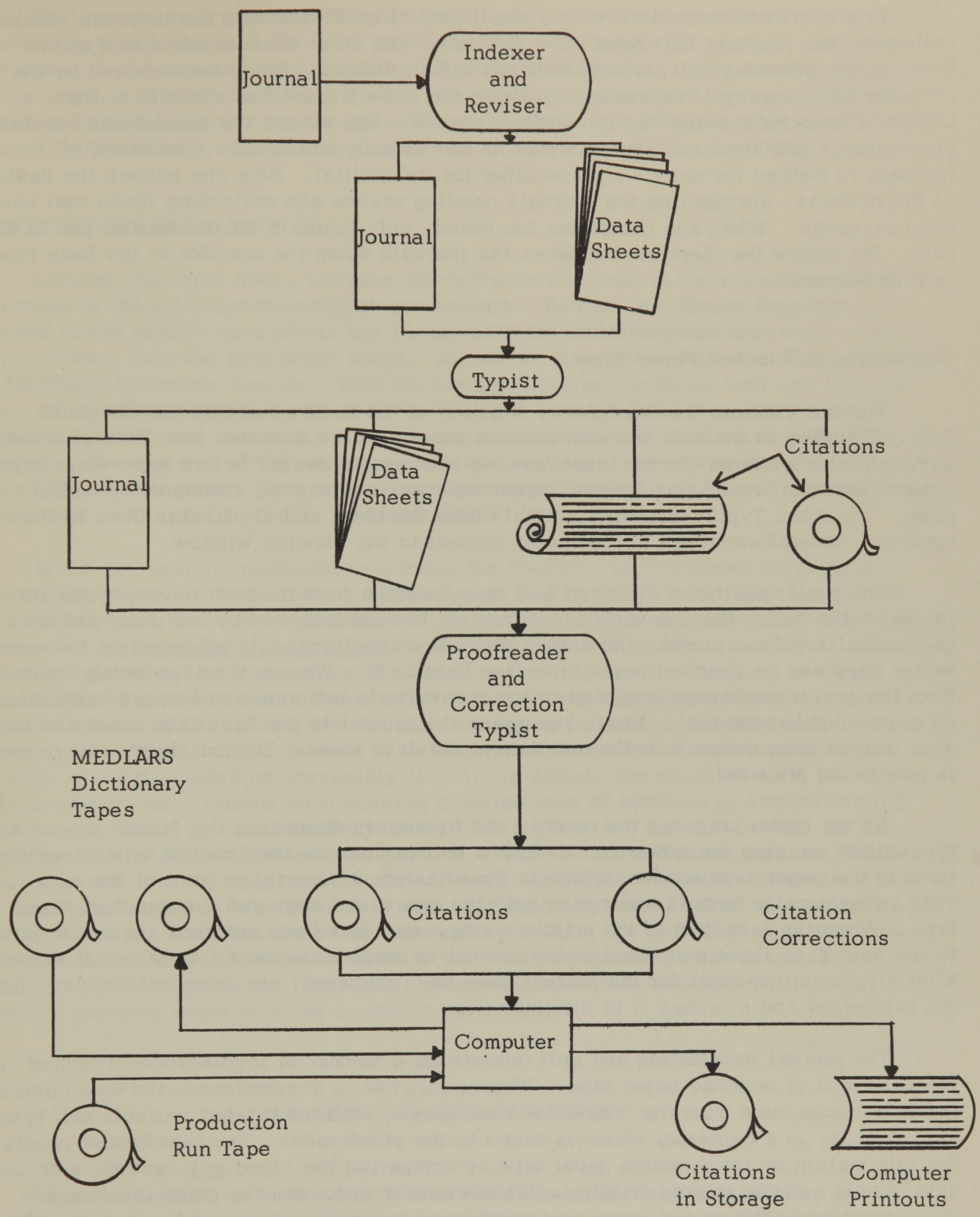


Figure 13. Journal Processing Flow Chart

In a short while, a clerk with a small cart circulates among the indexers' desks collecting the journals that have been indexed. She takes them to the office of the Head of the Indexing Unit, who determines which journals should be reviewed by the revisers for accuracy of indexing and which are to be transmitted directly to the Library's Data Processing Section without review. She places the journals on his desk and removes two stacks of journals that he has already examined. One stack of journals is marked for review and the other for transmittal. Now she passes the desks of the revisers, distributing the journals needing review and collecting those that have been reviewed. When she completes her round, only journals for transmittal are in the cart. We follow the clerk as she takes the journals down the corridor to the Data Processing Section.

Conversion to Punched Paper Tape

The clerk enters the Input Area, which is in the room adjoining the Computer Area. Standing in the hall, we can observe the activities in either area through a large, semi-circular window. In the Input Area we see typists seated before paper-tape typewriters that can "read" and "punch" paper tape while producing standard typewriter copy. The Head Typist takes the journals from the clerk and distributes them to the typists. Several are given to the typist nearest to the viewing window.

The typist positions a journal and then, reading from its front cover, types its Journal Title Code, the indexer's identification number followed by her own, and then the journal's volume number and date. These are simultaneously produced on the typewriter copy and on punched paper tape (see Exhibit B). We see the tape being emitted from the punch mechanism attached to the typewriter's left side and being accumulated on a detachable tape reel. Next, she opens the journal to the first data sheet and begins to type information from the article and the data sheet. The article's "unit record" is now being prepared.

As the typist prepares the record, she frequently depresses the "read" key on the typewriter, causing the typewriter to follow instructions on the "control tape" feeding through the paper-tape reader, which is the attachment directly in front of the punch. This automatically formats and codes both the typewriter copy and the punched paper tape. When the last item of the article's unit record has been entered, the typist turns to the next data sheet contained in the journal to prepare the next article's unit record. After all the unit records for the journal have been prepared, she tears off the copy from the typewriter and attaches it to the journal.

The journal information and unit records for a number of journals are "batched" on the same roll of punched paper tape. When a full roll is accumulated, the reel containing it is dismounted from the typewriter and placed, with the related journals and typed unit records, in a container which is taken to the proofreader. The proofreader proofs the information on the punched paper tape by comparing the typed unit records with the data sheets and the journal articles. When an error is found, the proofreader makes the correction in red on the typed unit record and prepares a "correction work sheet" which she attaches to the record. She then passes the journals, tape, and associated documents to the correction typist.

The correction typist uses the correction work sheet to prepare a supplementary "correction tape" which, after final proofing, will accompany the original tape when it is taken to the computer input station. The various journals, data sheets, typed unit records, and correction work sheets are retained in the Input Area until the computer has finished processing the punched paper tape. At such time, the journals will be sent to the Library collection and the related materials will be discarded.

REQUESTS FOR BIBLIOGRAPHIES

Leaving the Input Area, we move down the hall to look in for a moment on the activities of the bibliographic search specialists. Earlier, the Senior Searcher distributed to the search specialists the 11 requests for bibliographies that were received by the Library over the past eight hours. All are on the standard request form, "Request for MEDLARS Literature Search", that the Library furnishes to those who use its services. Most of the requests were received on such forms. Some requests, those described in letters and received by telephone, were entered on the forms by the Senior Searcher. We see one of the search specialists reading a request from a medical school for a bibliography "comparing the treatment of histoplasmosis and coccidioidomycosis with nystatin and/or amphotericin B" (see Exhibit C).

When the search specialist concludes his reading, he withdraws a "demand search data sheet" from the file on his desk and begins to fill it out (see Exhibit D). After specifying the request's number and his own identification number, he refers to his copy of Medical Subject Headings to identify pertinent medical headings. Now he lists the "elements" of this request on the form, writing them as medical headings with codes. Next, he scans several reports containing information about relevant citations stored in the system and several containing information about previous similar requests. Then, after a moment of deliberation, he writes a search equation on the form using the codes he previously listed. In effect, the equation tells the computer to search for citations with heading combinations of nystatin or amphotericin B and histoplasmosis or coccidioidomycosis. The request is, in other words, for all citations that have been tagged with any one or more of the following pairs of headings: nystatin and histoplasmosis; nystatin and coccidioidomycosis; amphotericin B and histoplasmosis; or amphotericin B and coccidioidomycosis. After also indicating that the retrieved citations are to be arranged by author, the searcher clips the form to the incoming letter and placed them in the "out" container on his desk. The material will be taken to the Input Area, where the search information will be converted to punched paper tape for computer processing.

THE MANIPULATION SUBSYSTEM

Let us return to our former position at the window to observe activities in the Computer Area (see Figure 14). This is a completely enclosed, separately air-conditioned room designed to house the Honeywell 800 computer, its peripheral equipment, and the Graphic Arts Composing Equipment (GRACE). Looking through the viewing window we see the console, or the human control for the computer, in the center

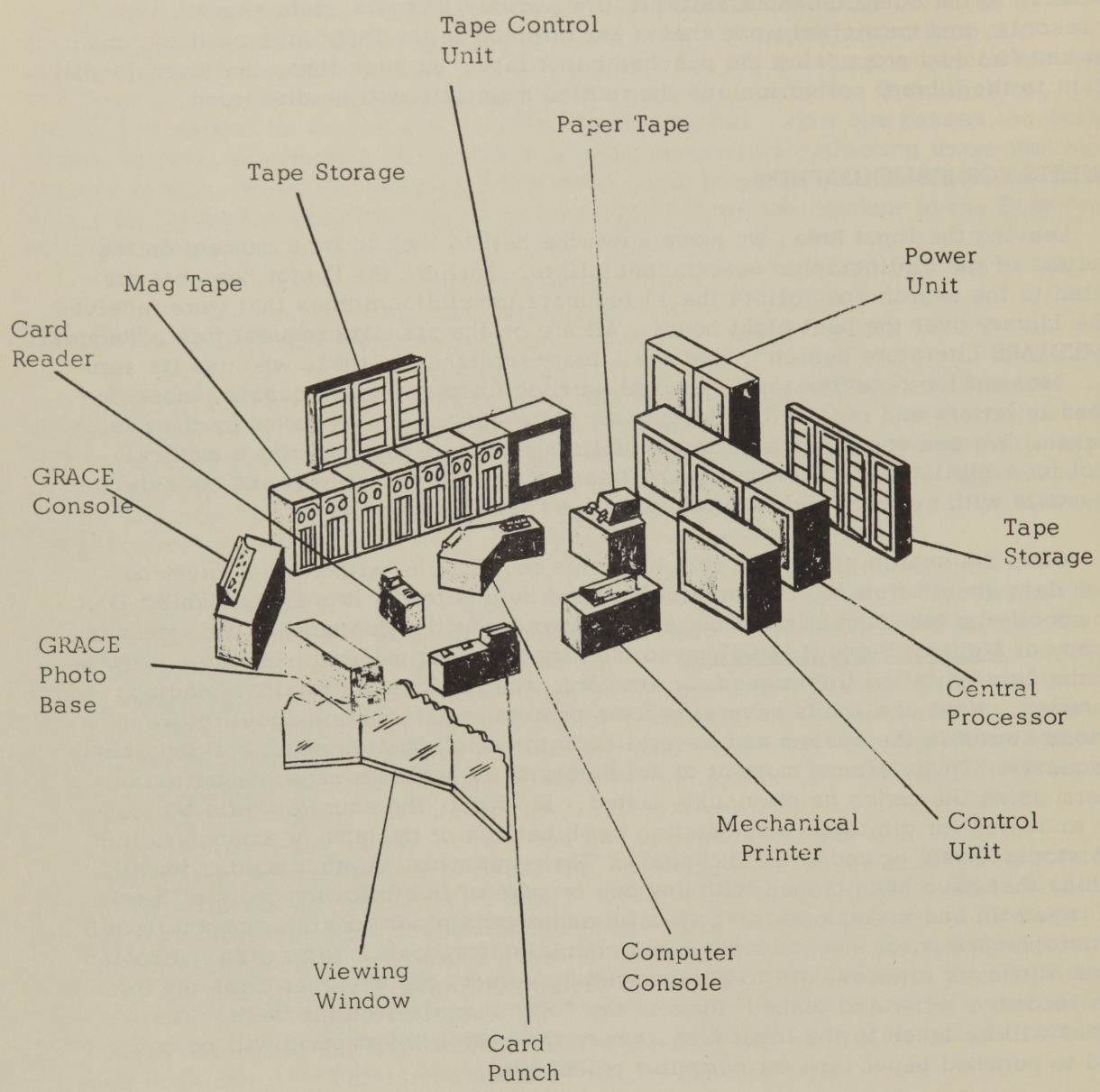


Figure 14. Computer Area

of the area. Some of the larger equipments surrounding the console are GRACE, seven magnetic tape drive units, and the computer central processor. The smaller equipments close to the console are a paper-tape reader, a card reader, a card punch, and the computer's mechanical printer. The typewriter on the console is used by both the console operator and the computer. The operator uses it to type in data and instructions for the computer's operations. The computer uses it to type out requests for instructions and notices about its operations.

INPUT PROCESSING

The console operator and his assistant have readied the necessary equipment for placing into magnetic-tape storage the citations resulting from yesterday's indexing activities. The MEDLARS Dictionary Tape was mounted on tape unit 2 and "work" tapes were mounted on tape units 3, 4, 5, and 6. The work tape on unit 3 is actually an outdated MEDLARS Dictionary Tape on which current files are to be written during the processing. The tapes on units 1 and 7 did not need to be changed. The tape on unit 1 is the Production Run Tape, containing all of the operational programs used in the MEDLARS system, and the tape on unit 7 is being used in connection with GRACE's operations.

1. Conversion to Magnetic Tape

The assistant console operator obtains 14 reels of punched paper tape, containing batched citations, and 3 containing corrections for the citations. He mounts one of the citation tapes on the paper-tape reader, threading the tape-end through the reader mechanism and attaching it to the take-up spool. Following this, he signals to the console operator that the equipment is ready. At the console, the operator types instructions to the computer, telling it in code to follow the Input Program on the Production Run Tape mounted on unit 1. Then he presses the EXECUTE button on the console; we see the reel on the paper-tape reader turn, and the computer writes into the input area of its high-speed memory the citations read by the paper-tape reader. Now the computer (a) converts the information on punched paper tape to computer language; (b) attaches journal information to each individual citation; (c) adds a "sort key" to each citation to indicate its processing order; and (d) writes the citations on tape unit 4.

The assistant console operator periodically mounts and dismounts paper tapes on the paper-tape reader until the 11 tapes containing citations have been read in. The citations being written out on unit 4 are collectively referred to as the "Intermediate Citation File", a name which they retain during processing until they are finally written on the tape on which they will be stored.

Now the assistant console operator mounts one of the correction tapes on the paper-tape reader. Although it is not a required procedure, the three correction tapes were saved until last. As in the case of citations, the computer converts the correction information to computer language, adds sort keys, and then writes the information on the Intermediate Citation File. The computer does not add journal information because it is already specified in the correction data. Soon the last of the three correction tapes has been read into the computer.

2. Correction of Citations

After the last correction tape has been read into the computer, the operator presses the EXECUTE button and immediately the reels on tape unit 4 turn as the Intermediate Citation File is read back into the computer. Partly relying on information stored in its high-speed memory, the computer (a) locates the citations for which corrections are specified and makes the corrections; (b) prepares a "main record" for each citation that includes all the information pertaining to the citation except its medical subject headings; (c) prepares a "subrecord" for each individual medical subject heading; (d) looks for logical errors in the citations (such as the absence of a required code symbol) and notifies the operation of such errors by means of "printouts" on the mechanical printer; and (e) writes the Intermediate Citation File on unit 5 in the form of main records and subrecords.

3. Processing Journal Information

Now automatically, but with the console operator monitoring the operations, the Intermediate Citation File and the Journal Record File on unit 2 begin reading into the computer. Comparing the information on the two tapes, the computer (a) determines if the journal title codes in the citation records are valid and notifies the operator of invalid codes by printouts on the mechanical printer; (b) adds journal place-of-publication codes and appropriate journal title abbreviations to the records; (c) reports, through printouts on the mechanical printer, overdue journals and journals received late; (d) updates the Journal Record File and writes it on unit 3; and (e) writes the Intermediate Citation File on unit 4.

4. Processing Subject-Heading Information

The Intermediate Citation File is again read into the computer. The computer uses information stored in its memory to expand and sort the citation records and to arrange them in alphabetical order by subject heading. Thus formatted, the Intermediate Citation File is first written on unit 5, then it is read into the computer along with the Medical Subject Heading File (MESH) on unit 2. The computer compares the information on the two tapes and this time (a) determines the validity of the subrecords by comparing their headings with those in the MESH File; (b) deletes subrecords with headings that cannot be matched and notifies the operator of such actions by printouts on the mechanical printer; (c) adds appropriate classification numbers (denoting structured subcategories of subjects to which the headings in the records relate) and inserts subject heading codes; (d) identifies and marks headings that qualify the citations for inclusion in specific recurring bibliographies; (e) updates the MESH File and writes it on unit 3; and (f) writes the Intermediate Citation File on unit 4.

5. Selecting Citations for Recurring Bibliographies

Now the Intermediate Citation File is read in, and the computer, using information stored in its high-speed memory, sorts the records back into their

original processing sequence, i.e., one record for each citation arranged under its proper journal. These are first written out on unit 5 and then read in as the Recurring Bibliography Parameter File on unit 2 is read in. Using logical search equations supplied by the latter tape, the computer (a) determines which citations contain elements that satisfy criteria for any of 50 main recurring bibliographies and any of their more particular sub-bibliographies; (b) marks those citations that qualify by attaching appropriate bibliography numbers to them; (c) updates the Recurring Bibliography Parameter File if necessary and writes it on unit 3; and (d) writes the Intermediate Citation File on unit 4.

6. Storing the Citations

As soon as the Intermediate Citation File has been written on unit 4, processing stops momentarily. The assistant operator labels and removes the two MEDLARS Dictionary Tapes from units 2 and 3 and then replaces them with the latest Compressed Citation File tape and the latest Processed Citation File tape, respectively. With the tapes mounted, the operator presses the EXECUTE button and processing proceeds. The Intermediate Citation File is read in and for each of its citations the computer, using information stored in its high-speed memory, prepares a highly compact citation containing only those elements necessary for retrieving and printing the citations for search requests. These abbreviated citations are added to the Compressed Citation File, and then the Intermediate Citation File is read in again. This time the computer uses information in its memory to format citations that will, after certain additional processing, be composed by the Graphic Arts Composing Equipment. It adds these citations to the Processed Citation File.

The addition of the citations to the two files completes the input processing operation. The citations prepared by the indexers on the previous day are now in the computer's "primary storage", that is, they are on magnetic tape ready for retrieval when needed. The entire operation required about 60 minutes of the computer's time.

Demand Search Processing

We have seen how citations are placed in storage. Since the computer's next operation is to process the 11 requests for demand (non-recurring) bibliographies that were received from the Input and Conversion Subsystem, we have an opportunity of seeing how citations are retrieved from storage. This operation has a special significance to a growing number of medical researchers, who are able to obtain, in a matter of a day or two, individually tailored bibliographies covering both American and foreign publications.

Preparing for the operation, the assistant console operator selects the reel of punched paper tape containing the request information and mounts it on the paper-tape reader. Next he mounts a work tape on tape unit 2, the current MEDLARS Dictionary Tape on unit 3, and Compressed Citation File tapes on units 5 and 6. No changes are necessary in the tapes on the other units. Meanwhile, the operator types instructions

for the computer on the console typewriter. Now, with the equipment ready, the operator presses the EXECUTE button on the console, the reel on the paper-tape reader turns, and the processing of demand search requests begins.

1. Conversion and Coding

The 11 requests are first read into the input area of the computer's high-speed memory. The computer reviews the information for logical errors, formats it into processing sequence, and then writes it on tape unit 2. Meanwhile the computer, using the console typewriter, informs the operator that the requests require a search of the entire Compressed Citation File, which consists of all the citations stored in MEDLARS' magnetic-tape memory. Accordingly, all 9 reels of tape making up the File will need to be mounted and read into the computer. Following this, the reels on tape units 2 and 3 turn as the MEDLARS Dictionary Tape and the search elements of the requests are read in together. The computer, using information from the first tape, substitutes codes for each search element and inserts the number of citations the element would retrieve from storage if no conditions were imposed. It places this information (element codes and related citation counts) in its high-speed memory.

2. Retrieving the Citations

Now the MEDLARS Dictionary Tape rewinds on tape unit 3 while the request statements on tape unit 2 are read in. The computer places the element codes in algebraic expressions and compares the expressions with the number of citations that each would retrieve from storage. Based on this comparison, it then makes a logical selection of only those codes that will most effecitvely limit the citations retrieved to the kinds and number needed to satisfy the request statements. The computer retains these "significant elements" in its high-speed memory and writes the request statements on tape unit 4. Meanwhile, the assistant console operator has dismounted the MEDLARS Dictionary Tape from unit 3 and mounted a work tape in its place.

Next, the Compressed Citation File tapes on units 5 and 6 are, in that order, read into the computer. The computer compares the significant elements in its memory with the compressed citations contained in the File and selects those citations that contain one or more of the elements. We see the assistant console operator alternately dismounting and mounting tapes on units 5 and 6 until the entire Compressed Citation File has been read in to be searched in this way. All of the citations selected from the File during the search process are written on unit 3. Then, as the assistant operator replaces the tapes on units 5 and 6 with a work tape and the current Request History Tape, respectively, the retrieved-citation tape and the search requests on tape unit 4 begin to read in. The computer compares each citation with each request statement and, if a citation is responsive to a request, appends the number of the request to the citation. Following this, the computer writes the citations on unit 5, listing them under their respective request numbers.

3. Printing the Bibliographies

Now both the retrieved-citation tape and the request-information tape on unit 2 are read in. The computer extracts formatting specifications from the latter tape and adds them to the citations, and then writes the citations on unit 6. Promptly, the request-information tape is again read in, this time as the Request History Tape is read in. The computer updates the Request History Tape by adding data from the first tape and, as this is done, writes the tapes out again.

Next, only the retrieved-citation tape is read in. This time the computer formats the retrieved citations for each request according to the specifications of the individual or group who submitted it. The formatted citations are written on unit 5, read back into the computer and sorted into printing sequence, written out on unit 6, and then again read in to be physically arranged into "printed lines", and finally written out on unit 5 in a form appropriate for the mechanical printer.

THE OUTPUT SUBSYSTEM

As the computer is printing out the last of the demand search bibliographies, we see the operator of the Graphic Arts Composing Equipment (GRACE) preparing it for its task of composing citations on film for Index Medicus. A special test tape has been mounted on tape unit 7, which is connected to GRACE, and a 100-foot roll of film has been mounted in GRACE's film magazine. We see the operator turn GRACE's operation selector switch, positioning it so that the equipment will run through the "CONFIDENCE MODE". Then he presses the START button and GRACE automatically reviews its own operational condition for indications of malfunctioning components. None of the trouble-indicator panels on GRACE is illuminated, meaning that no function errors were found. The operator repositions the selector switch, presses the START button, and this time the test tape is run through GRACE which, in turn, produces a "test pattern" on film in the magazine. The operator removes the strip of exposed film, places it in a container, and then hands the container to a clerk who takes it downstairs to the Dark Room Area.

The container is given to the operator of the automatic film processor. He removes the film strip from the container and feeds the film into the processor. The processor automatically develops, fixes, washes, and dries the film. Soon we see the fully processed film being emitted from the processor. Removing the film, the operator inspects it, first under a magnifying glass and then with a densitometer, looking for errors in either photo-composition or film development. An error would normally require an adjustment in GRACE or the automatic film processor. Finding none, he notifies the peripheral equipment operator that test run results are satisfactory and that operations may proceed.

Receiving this information, GRACE's operator removes the test tape from unit 7 and in its place mounts a reel of magnetic tape containing citations for Index Medicus.

The computer previously extracted these citations from the Processed Citation File and expanded and formatted them for use with GRACE. He pushes the START button and GRACE begins to "set" type by extracting a line at a time in coded form from the magnetic tape, manipulating and decoding the information, and photographically recording composed three-column pages of Index Medicus on the film in the magazine (see Exhibit E).

After several minutes of processing, GRACE abruptly stops, which means that it has found an error either in the information from the magnetic tape or in its own functioning. A lighted indicator panel states: FLASH LAMP FAILURE. The operator turns the selector switch, presses the START button, and once more runs GRACE through the CONFIDENCE MODE. Soon the number "211" appears on the flash lamp trouble-indicator panel. The operator replaces flash lamp number 211. Finding that this corrects the problem, he sets GRACE back into normal operation, first directing it to recompose the page which was in process when the error was detected.

Processing continues for about an hour before GRACE stops once more. This time the stop indicates to the operator that the 100-foot roll of film in the magazine has been completely exposed. The operator removes the roll of exposed film and mounts another roll in the magazine. As GRACE resumes operations, the exposed film is taken down to the Dark Room Area, where it will be developed by the automatic film processor, inspected for quality faults, cut into page-size strips, given a final review, and then packaged and sent to the printer.

CONCLUSION

This concludes our MEDLARS tour. We did not, of course, view all of the activities and operations of the system. However, those we saw were illustrative of how the system will function.

IV

DESIGN CONSIDERATIONS

SYSTEM CRITERIA

INTRODUCTION

MEDLARS will incorporate and expand the features of the existing publication system at the Library and, in addition, will add a storage and retrieval capability to this existing system. The authority list for Subject Headings would continue to be used for the publication system. A question concerning the adequacy of this list for storage and retrieval and, if adequate, how the headings would be utilized, however, had to be answered and the existing MESH List modified as necessary.

MEDICAL SUBJECT HEADINGS

When the present mechanized system was adopted by the Library in 1960, it was necessary to revise and recast the then existing subject heading authority list. This 356-page volume represented the Library's experience over many years. It contained about 4,500 main headings and 67 standard topical subheadings arranged in an alphabetical listing of all main heading terms and contained extensive cross references.

Since MEDLARS has the added capability for storage and retrieval of citations and also will require an increase in the depth of indexing, investigations were conducted by NLM during 1962 and certain revisions made to MESH. The second edition of MESH (published in January 1963) contains a total of some 5,700 subject headings of which 1,400 are new and 4,300 were used in the first edition. Thus, subject headings were made more specific especially in areas of greater interest as evidenced by the quantity of published articles.

The important innovations in the subject heading list included the appending of a categorized list, revisions in cross-reference designation and structure and the discontinuation of the use of topical subheadings. The decision to discontinue topical subheadings was made reluctantly after much investigation. They are not really adaptable to coordinate searching and with more specific main heading terms being utilized, their disadvantages outweighed their advantages. However, 24 form subheadings will be utilized with monographs in Index Medicus.

The major emphasis of MESH is on clinical aspects of medicine. The terms for para-medical and non-medical sciences concentrate more on chemistry and anthropology than on agriculture and sociology. MESH is based on the concept that subject cataloging of monographs and subject indexing of periodical articles are essentially like operations. This conciliation of formerly incompatible methods has both positive and negative aspects, but time and usage will determine the extent of the success of this technique.

HUMAN INDEXING

Although the computer and the manipulation subsystem are of central importance to the operation of MEDLARS, the importance of the human indexing task performed in the input subsystem cannot be overemphasized. MEDLARS could fail to meet its objectives if the indexing were poorly or incorrectly performed.

The state of knowledge about language and meaning is such today that the indexing task cannot be mechanized for a system of this size and complexity. However, the capabilities of the manipulation subsystem can also be used to facilitate and improve the task and art of indexing. It can be safely predicted that, as system personnel gain experience and the body of statistics regarding system performance grows, the computer and the information base available to it will become increasingly more important tools in the task of improving system performance in general and in making indexing more of a science than an art.

SEARCH REQUESTS

As already mentioned, the handling of search requests is just as important to the success of the retrieval of information for demand bibliographies as indexing. When it is remembered that these non-recurring requests for bibliographical information will reach an estimated total of 2,500 in 1964 and 22,500 in 1969, the importance of satisfactorily retrieving such a mass of information from that stored in the computer files cannot be overemphasized.

Four aspects of the search requests will determine their primary effect on the MEDLARS system: (1) the complexity of the requests, (2) the volume of requests that must be processed, (3) the amount of time allowed for the processing, and (4) the volume of unit records retrieved.

The processing of search requests in a satisfactory manner will have its maximum impact on the manipulation subsystem. Specifically, the effects will be manifested by the number, kinds, and lengths of computer programs required, the way in which the computer files are organized, the media required for storage of the file contents, and the size and speed of the computer. The effects on the other subsystems will be no less pervasive.

In addition to selecting headings that adequately describe the requester's needs, the search request specialists must be able to denote the logical interrelations among the retrieval criteria in a precise and unambiguous way. The way in which a search request is defined must not only serve the needs of the bibliographic specialists, it must also avoid complication of the typing task in the conversion part of the subsystem, so as to keep to a minimum the number of typing errors.

The effects of search requests on the output subsystem will be primarily those caused by the volume of unit records retrieved for printing and by the flexibility in format and typography that the subsystem must be able to provide.

SUBJECT HEADINGS

The unit records and the search requests will be joined primarily by the subject headings assigned to the citations. The ability of the system to retrieve those unit records that are relevant to a requestor's needs will be directly related (1) to the number and correctness of the subject headings appended to the citations and (2) to the subject headings and the logical relations among them specified in a search request.

The greater the number of headings appended to a citation or search request, the higher the probability that it will be retrieved when relevant to a request. However, the probability of retrieving non-relevant or only peripherally-relevant unit records also increases with the number of headings per citation or request. How to achieve an appropriate balance has been a matter of speculation, debate, and experimentation for many years. Direct experience with the MEDLARS installation may provide the answer to the question.

The increase in the number of subject headings per citation will have a second important effect: The indexer will specify headings under which a citation will appear in Index Medicus; the greater the number of headings thus specified, the greater the number of pages in a given issue. The decision as to how many should generally be specified will, to some extent, be a matter of policy, but in the individual case it will rest with the indexer, and also, for some portion of his output, with the reviser.

COMPUTER SELECTION

Because the heart of MEDLARS is a computer, and because the requirements that the computer has to meet are exacting, it follows that the determination of the type of computer to be used and its specific characteristics were of primary importance in designing the over-all system. The procedure followed consisted of a study of the system requirements, a determination of specific tasks, an investigation of file techniques, and an evaluation of computer systems. The final selection was based on a cost-effectiveness trade off, the latter being derived from the timing for a sample exercise.

SYSTEM REQUIREMENTS

The system requirements were the major factors affecting the detailed system design. The transformation of these requirements into specific computer tasks and the determination of the computer loads that they present, when subject to the constraints of computer execution rates and of throughput requirements, determined: (1) the number, complexity, and the length of the computer programs; (2) the organization of and information flow among the computer programs; and (3) the number and content of various files needed to implement the programs.

The major tasks of and primary loads on the system will be:

1. To accept unit records from punched paper tape, including correction tapes that may be attached;

2. To edit citations for storage in the system files;

3. To accept and edit demand search requests;

4. To retrieve, edit, and format for printing the citations to satisfy demand search requests; and

5. To retrieve, edit, and format for printing those citations required for recurring bibliographies, of which Index Medicus is the most important.

FILE STRUCTURE

The fundamental problems in the design of a file structure and of techniques to manipulate the files stem from the sizes of the files, the rates at which they grow, the need to retrieve information from them within the constraints imposed, and the need to update and to revise them. As explained in preceding paragraphs, these files will contain, in addition to compressed and processed unit-record information, program procedures and data covering journal identification, subject headings, and historical statistics.

The principal problem involved in the design of this particular MEDLARS area was whether to use magnetic tape, random-access disks, or a combination of both for storage purposes. A cost-comparison study showed that, at an estimated yearly file growth rate of 80 million characters, an additional disk unit would be required each year if all the files were to be stored on disks. This was obviously too costly, so disks were rejected as the primary storage medium and magnetic tape was selected for the purpose.

But, even though disks were not an economical primary storage medium, some combination of tape and disks could conceivably provide the most practical and effective method of accessing the large storage of information involved. To evaluate such a combination, a study was made of the two basic file-structure and retrieval techniques, i.e., a disk-oriented locator file in combination with a magnetic-tape file versus a sequential file stored only on magnetic tape.

The basic difference between the two methods is that the disk-type locator method, by utilizing information stored in the locator, partially or completely determines what citations satisfy the retrieval criteria. It then retrieves the selected citations by passing the entire citation file in review and processing the selected citation for final evaluation or selection. In the sequential storage on magnetic tape, the selection is made as the citation file is passed and each citation is examined.

In the magnetic-tape sequential method, every file element is read into the computer from the tape and processed against the request logic. If the elements satisfy the requestor's parameters, the citation is selected for output. The speed at which the file is read determines whether the system is tape- or processor-limited. If the former, then the tape read time is the total time for the demand request process. If the computer processing time is greater than tape read time, then the total time is essentially equal to the computer processing time.

The concept of the locator file and its use is very complex and contains many subtleties and variations. Such a file is more properly termed a locator index, as it is an index to a file and not the file itself.

This locator index would be stored on disks and the citations on magnetic tape. The index would contain tape and record addresses under an appropriate locator index key. The system would contain 10,000 such keys, so the size of the locator index becomes an important factor. In its simplest form, it would increase at the rate of 12 million characters per year and accumulate to 60 million in 1969. The size and constant updating of a locator for this application require that it be stored on disks to be of practical value.

With a disk locator file, the logical processing of various elements contained in the locator file would determine which citations satisfy, or partially satisfy, the requester's parameters. The analysis of typical requests and the file records satisfying those requests showed that the number of records to be retrieved and their random location required that the entire file of tapes must be read, even when using the disk-type locator method.

The final requirement in determining the design of the file structure was that of retrieving and printing the various recurring bibliographies, including the monthly and annual Index Medicus. This proved to be a simple study because, regardless of the retrieval method applied, every citation had to be retrieved for lengthy processing, including sorting, formatting, and printing, indicating that the sequential tape method would be more economical.

The conclusions drawn from the design study of file structures were:

1. Both retrieval methods would use the same citation file, but processing techniques would vary.

2. Both methods would require the reading of the entire file for the demand bibliography searches.

3. The total time required by the disk-type locator method, in a high-hit-ratio application, would be the sum of the locator processing time plus tape time.

4. The total time for the sequential (tape) method would be equal to either the tape time or the processing time, depending upon which of the two was greater. Calculations showed that the logical processing time using the sequential method would initially be less than tape time, but as the number of requests on the system increased, the processing time would increase above tape time, and eventually the locator-disk method might be more practical.

5. The selected computer system should have a high-speed internal processing capability.

As a consequence, it was decided to adopt the sequential (magnetic tape) method of retrieval for initial operation, with the provision that it could be expanded to include a locator disk index if and when it became necessary, without altering the file structure.

COMPUTER EVALUATION

To determine which of the many computer configurations available would perform the required tasks most economically, it was necessary to design a preliminary manipulation subsystem in order to determine what types of processing capability were needed.

Necessary requirements of such a computer included satisfactory performance, reasonable cost, and availability on time. In addition:

1. A solid-state (transistorized) system was desirable because it uses less electrical power, is more reliable than a vacuum-tube system, and takes up less room.

2. A demonstration of performance reliability was required, to be obtained through the performance records of similar equipment.

3. A demonstration was required, through previous field use, of the adequacy of the programs, operational routines, subroutines, diagnostic routines, compilers, and generalized sort and input/output routines used by the system.

4. The computer had to be compatible with other equipments actually in use or contemplated by the Library.

A typical equipment configuration would include at least one central processor with 40,000 characters of core storage, six magnetic-tape units, a paper-tape reader, a card reader, an on-line printer, and a card punch.

The first review covered 19 different computer designs. From these, seven were eliminated because they were obviously too large or too small. Four more were eliminated because of limitations of design or availability. Narrowing down the list of the remaining eight designs was accomplished through the use of a test problem.

TEST PROBLEM

A study was made of the types of computer tasks to be performed, and from this the types of computer processing required to perform these tasks were determined. From this analysis, a problem was designed that would be representative of this type of processing.

The problem was submitted to the computer manufacturers involved. It specified not only what had to be done but also how it should be done, since this was an evaluation of computer processing capability and not of programming techniques.

For each computer, optimum use was made of its processing capabilities in performing the task. However, programming techniques to the extent possible were eliminated as a variable. The computer operations required by the problem consisted, in the main, of character comparisons, character look-ups, and data movement. The problem solution also required both magnetic-tape handling and input/output capability.

Emphasis was placed on central processor capability because a major portion of the system loads will consist of character comparisons and processing.

The results were carefully checked and evaluated, including lengthy discussions with vendors' representatives. In a few cases, problem solutions were not consistent with the intent of the problem, and the manufacturers resubmitted the problem in the manner requested.

From the results, a relative computer-processing rating was determined for each design. In order to convert this rating into an absolute load factor for each computer, the machine having the fastest problem-solution time was used to estimate the actual total time required to meet the MEDLARS requirements. This load factor was then applied to the relative rating established for each computer by the test problem. The estimated total loads for each computer were determined by applying the ratio to the estimated running times of the tasks specified by the system requirements.

As a consequence, only three computers merited further consideration. The costs of the three systems were approximately equal. One was eliminated because it had less processing capability than the other two; the second was ruled out because the manufacturer could not meet the specified delivery date; the third and recommended computer was the Honeywell 800 manufactured by the Minneapolis-Honeywell Regulator Company

THE OUTPUT SUBSYSTEM

The specifications for MEDLARS state that Index Medicus must be continued in essentially its existing form, especially in terms of its typographical appearance. The form in question was the eventual copy resulting from the use of Justowriter machines. This requirement, coupled with the requirements for total monthly Index Medicus character loads and the time provided for composition, as well as the variety of different characters used during composition, established the three basic equipment capabilities that had to be satisfied; namely, speed, typographical quality, and total character complement. The equipment to be used had to be either one that was commercially available or one that could be modified for the purpose.

SPEED

The monthly Index Medicus printing load in characters will have an estimated average growth rate of slightly less than 10 percent per year for the period from 1964 to 1969. A monthly load of about 13 million characters is predicted for 1969. This figure was used to measure system performance.

A time limitation of 16 hours was derived from the over-all maximum time allowance for the composition of Index Medicus. Approximately 15 percent of the 16 hours was estimated as necessary for recomposition of parts of Index Medicus. Using the remaining time as a basis for computation, it was determined that the output printer must have a composition rate of 440 characters per second.

TYPOGRAPHY

It was essential that the level of typographical quality for Index Medicus be maintained at or above that which can be produced on a Justowriter or Varityper. It was also desirable to consider utilizing type face with and without serifs and necessary that characters be proportionally spaced. These requirements are not only a matter of aesthetics but also involve legibility and significant space-saving advantages.

It was also desired that the lines of type be justified on the right, solely a matter of aesthetics.

CHARACTER COMPLEMENT

Index Medicus uses some 170 different type characters on its pages, and this represents a lower limit on the number of different characters required. Furthermore, since Index Medicus will in the future require the vernacular or transliterated vernacular, it was felt that some further characters should be provided in addition to those already available. These characters are of two kinds: (1) letters, both capitals and lower case, in various sizes; and (2) numbers, symbols, punctuation marks, and diacritical marks.

EQUIPMENT EVALUATION

Of the available typesetting or printing equipments on the market, only four types were compatible to any extent with the requirements: cathode-ray tube computer printers, paper-tape-driven typewriters, photographic typesetters, and mechanical computer printers. Further study of these indicated that the cathode-ray tube computer printer would meet the speed requirement only, while the use of tape-driven typewriters and photographic typesetters would require a large number of separate units to meet the speed requirement. This narrowed the choice down to mechanical computer printers.

The two basic types of mechanical computer printers were studied in turn, and the rotating-drum type was eliminated because of excessive cost. The chain type was more reasonable in price, and it met the requirement for a font of at least 170 characters, but the fonts were essentially limited to one size without proportional spacing, and the printing speed was affected by the number of characters in the chain.

A detailed study was made of the characteristics of the chain-type mechanical computer printer to determine whether relaxing the requirements to a total of 120 characters would allow it to be used, or whether a specially designed photographic typesetter should be adopted. The results of this study indicated that the use of a chain-type mechanical computer printer would require the printing of 100 percent more pages for an issue of Index Medicus than would a specially designed photographic typesetter.

As a consequence, it was decided to pursue the development of a special design of photographic typesetter. The end product, the Graphic Arts Composing Equipment, will be an off-line device that will accept magnetic tape as its input medium and produce as its output exposed film suitable for lithographic plate-making after processing.

This equipment is being specially designed and manufactured by Photon, Incorporated, according to specifications prepared by General Electric.

SELECTION OF DATA CONVERSION EQUIPMENT

A few words are in order concerning the selection of the conversion equipment. Three basic configurations were considered as follows:

1. Use of an automatic character reader

2. Use of punched cards

3. Use of punched paper tape

Automatic character readers were first considered. The input material as received is not in the form required by the system nor can the type sizes and styles be controlled It is not indexed nor can the citation be extracted from the printed matter by any existing character readers. Thus, since the material must be at least re-typed, it is just as easy and less expensive to prepare punched cards or tape while re-typing the material as it would be to prepare the data for and to utilize a character reader.

The use of punched cards was next rejected as the primary input method for the following reasons:

1. Existing card punch machines provide for typically 47 different characters.

2. The unit records are of variable length and each record will require an average of five cards.

3. Close to 5,000 cards per day would have to be kept in the correct serial sequence as they passed through proofing and correction procedures.

Punched paper tape with proof copy was selected as the conversion medium. The paper tape can be checked for errors by visual proof-reading of the printed copy, and corrections can be made by utilizing a second tape. Paper-tape typewriters can provide 88 unique characters and by utilizing a red/black color shift, the number can be effectively doubled.

V

MANAGEMENT CONSIDERATIONS

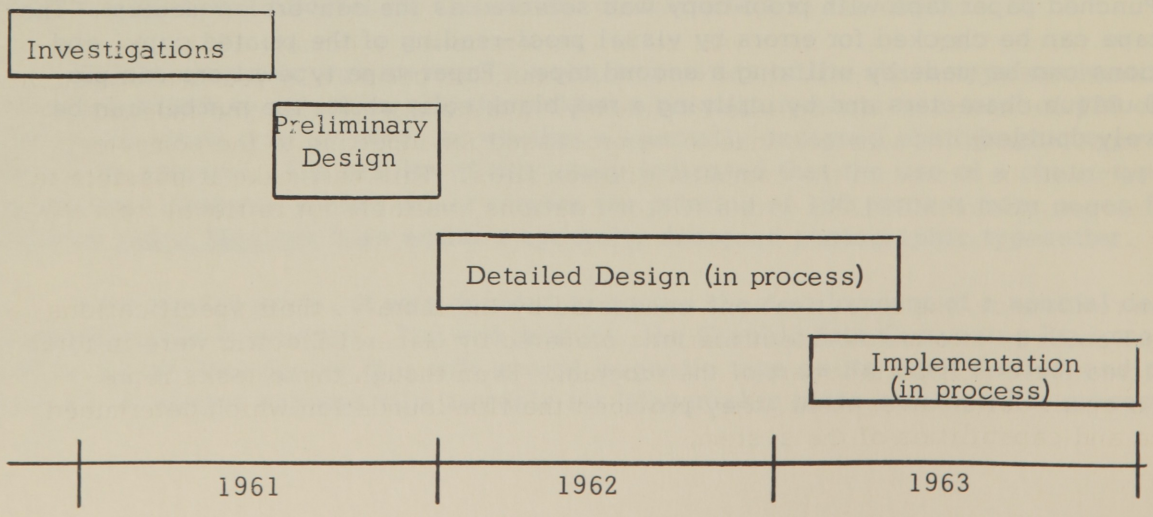
INTRODUCTION

There are many similarities among systems such as MEDLARS which utilize standard digital computers and other equipments. This section describes the over-all technical approach and schedule, summarizes the types of equipment and operational personnel required, describes some of the more important operational details, and explains some of the ways in which the system capabilities can be increased

TIME SCHEDULE

Many tens of man-years of human effort over a period of years are required to conceive, design, and implement a complex system such as MEDLARS. It is difficult to state precisely the point in time when MEDLARS was born. As the Library became involved in the studies which formed a part of the Index Mechanization Project, the whole concept of utilizing a computer as an integral part of the bibliographic retrieval system was brought into sharper focus. Further evidence is attested to by the contract signed with Dr. Robert L. Ledley in April 1959 "to conduct a study to investigate the feasibility of using electronic digital computers for the publication of the Index Medicus and also as a basis for the construction of an efficient reference and bibliographic service"

Encouraged by the support received from the Board of Regents and from the National Heart Institute, the Library made a formal proposal in the fall of 1960 to the National Heart Council to underwrite the development of a system with a computer base that would "have profound effects in the direction of substantially reducing the bibliographical tangles of the literature problem". The assent of the National Heart Council signaled the beginning of a period of activity to inaugurate the development of the system which is now in process. The MEDLARS timing schedule is shown and each step (or phase) is briefly described.



Investigations started in November 1960 with the Library obtaining the temporary services of a full-time Systems Analyst to assist in examining and defining the problem. During February 1961, specifications and a covering letter of invitation were mailed to commercial and non-profit organizations to request proposals for the study, design, installation, and trial operation of an electronic data-processing capability for the organization, storage, and retrieval of bibliographical citations to the medical literature. The objectives of this new system (MEDLARS) were divided into two categories: those considered to be of primary importance, and a secondary group of longer range significance to the Library.

Twenty-five proposals were received by the Library during April 1961 and were reviewed by the Library and the Public Health Service. The General Electric Company's proposal was finally selected and negotiations were held and a contract for the Preliminary Design effort was ready for Contracting Officer approval by mid-June. At this point a delay ensued as a result of procedural aspects of the appraisal and contracting mechanisms.

The Preliminary Design was started by General Electric in August of 1961. The team effort consisted of an average of six people from the General Electric Company. This team consisted of a system analyst, system engineers, a computer program designer, a human factors specialist, and a library consultant. This team, working very closely with selected Library personnel, studied and amplified the system requirements, determined possible system and subsystem configurations and recommended the system configuration which best satisfied the needs of the Library. (Several of these investigations are described in Chapter IV.) The Preliminary Design was completed January 31, 1962.

The acceptance of the Preliminary Design concepts by the Library meant that the system performance capabilities were defined and the Detailed Design of the system could proceed.

The Detailed Design effort (see timing schedule shown on previous page) was started in February 1962 and called for the designing and engineering of the system, the ordering of long lead-time items, and the planning for the system implementation. As of this writing, the Detailed Design is nearing completion and Implementation has begun.

It is planned that the system will be fully operational in January 1964. As mentioned in preceding paragraphs, there will be a transition period of a year during which the new and old systems will operate in parallel. During this period, the indexed journal article citations will continue to be processed for inputting to the computer, where they will be stored in the computer citation files. This will make it possible to start the operation of MEDLARS with a file of citations available for retrieval from storage.

In retrospect, the investigations conducted by the Library, their specifications, and the Detailed Performance Specifications prepared by General Electric were in toto a most creative and important part of the program. Even though these tasks represented a small dollar investment, they provided the firm foundation which determined the size and capabilities of the system.

The concept of dividing the system design into two categories by the Library (those considered to be of primary importance and those considered to be important but in the future) is a factor to be considered by those groups contemplating the design and implementation of similar systems. The category considered to be of primary importance consisted of the storage and retrieval system and the publication system. The responsibility for these two problems was vested with the Bibliographic Services Division and its Chief was named as the Project Officer of the program. Since the mechanized publication is presently in operation, future loads could be estimated with some assurance. Even though retrieval was not as well defined, its loads would initially have less impact on the system. While emphasis was placed on these two major problems, other smaller, less complex problems as well as future requirements were also considered to assure that the new system would not become obsolete but could expand to accomplish additional tasks.

This approach, which is in reality evolutionary in nature, means that the most important tasks would be investigated and solved first and then other tasks would be taken in turn. The proper design approach for a new system is indeed a difficult task and requires firm direction.

EQUIPMENT

The major items of equipment which will be utilized by the personnel operating the system consist of input paper-tape typewriters, a Honeywell 800 computer, the Photon Inc., Graphic Arts Composing Equipment, and the film processor. Neither the long and arduous equipment investigations nor the procurement procedures will be discussed here (although the computer selection is described in Chapter IV).

OPERATING PERSONNEL

The mechanized system presently in use at the Library requires about 40 man-years per year. MEDLARS will require approximately 60 man-years per year. Working time for both systems is limited to an effective six-hour day, single-shift operation, a five-day week, and a 250-day year.

The Bibliographic Services Division of the Library will have the operating responsibility of MEDLARS as its primary function.

SYSTEM OPERATING TIMES

One of the most important objectives of MEDLARS is to provide timely system outputs so as to reduce literature-screening functions external to NLM. It was realized

that an improvement in the efficiency of producing Index Medicus would not only allow for other operations to be performed by the Library but, in addition, free consumers from having to make certain independent bibliographic researches.

Under the present mechanized system, the throughput time of journal articles is 22 working days. This is the total time elapsed from the receipt by the Library of the last journal with an article to be included in a monthly issue of Index Medicus, until the delivery of the film containing that article to the printer. MEDLARS is expected to reduce this time to five working days.

The five-day throughput time and its distribution among the three MEDLARS subsystems established a major parameter affecting the selection of equipment and the design of system procedures and configurations. The allocation of the time is as follows: Input and Conversion Subsystem, 1-1/2 days; Manipulation Subsystem, 1-1/2 days; and Output Subsystem, 2 days. The possibility of at least a partial overlap in the times for the last two subsystems effectively increases the time allowed each.

The servicing of the various recurring bibliographies will be scheduled in advance. Plans call for their being available for transmittal to a requester no more than a single working day after their scheduled preparation.

The total time allotted for processing demand bibliographies is no more than two days. It is expected, however, that at infrequent times some requests for demand bibliographies will require immediate processing; that is, a throughput time of less than two hours from receipt of request to transmittal of bibliography. This may require the addition of a special priority arrangement in the system operations.

It is also expected that the response-time requirements for bibliographies of either kind may be relaxed by up to two working days during the time when a monthly issue of Index Medicus is being processed by the computer and the printer. During a corresponding processing of the Cumulated Index Medicus, the response-time requirements may be relaxed by up to seven working days.

The time required for computer operations may be of particular interest because in many ways the efficiency of the system is geared to the processing speed of the computer. Estimated average operating times to perform the usual daily tasks expected in 1964 are given in Table 3. It should be noted, however, that the times indicated are approximations based in part on factors that are subject to change or about which only preliminary information is available.

Table 3

Computer Time Requirements by Type of Usual Daily Task, 1964

Type of Daily Task	Estimated Computer Time Required (in minutes)
Input Processing	65
File Searching for Demand Requests	40*
Output Processing for Demand Requests	30
Processing of Recurring Bibliographies	90
Preparing Tape for 1/22nd of Monthly Index Medicus	25
Statistical Runs	10
Conversions	15
Assemblies	15
Debugging	30
Tape Duplication	20
Training	10
Other	10
Total Time for Daily Tasks	360

* Assumes 8 tapes.

MISCELLANEOUS OPERATING DETAILS

A number of other considerations important to the management of MEDLARS should be mentioned. These include the correction of errors, provision for automatic search expansion, and provision for request histories.

CORRECTION OF ERRORS

Errors to be corrected can be categorized as (1) those that the computer detects and (2) those detected by humans.

The computer will be programmed to check, when possible, for the presence and correctness of all elements required by citations, search requests, and programs. The computer will print out the errors for correction by the staff.

Provision will also be made for the detection and correction of indexing, conversion, and other input errors by the inspection of indexing and by proofreading.

PRE-SPECIFIED SEARCH EXPANSION

It will be possible for a requester to have the computer expand a search by writing out a complete statement of each of up to three subsearches to be executed, if necessary, in serial sequence in order to provide the required number of citations. This search expansion will reduce the time to satisfactorily process requests.

HISTORIES

Provision will be made for the storage, on a machine-readable medium, of all search requests processed by the computer. Any additional information required will be appended to each request or filed in such a way that it can be retrieved with the request. This will make it possible to analyze such historical data and generate statistics as needed.

In addition to the substantive content of a request, the record of search requests will include the requestor's name, organization, and address; the date and time of receipt of the request in its original form, the date when the processing of the request was completed, plus the total time required to process it; the type of request; and the batch number, if requests are batched for processing.

The processing of requests for data and analyses of this kind will almost always have a low-priority status.

SECONDARY OBJECTIVES

Although immediate emphasis has been placed on the storage and retrieval system and the publication system, the secondary future long-range objective requirements must be examined and considered in the initial design and implementation of the system.

These secondary objectives are: (1) the requirements for a national, decentralized, medical bibliographic system; (2) the possibility of storing and retrieving graphic images of textual material; and (3) the mechanization of other Library functions.

NATIONAL, DECENTRALIZED MEDICAL BIBLIOGRAPHIC SYSTEM

The fundamental reason for the possible need for a national, decentralized, medical bibliographic system is the expected rise in demands for service, particularly with respect to demand search requests. It is expected that this load will become so great that, unless provided for in advance, it will tie up telephone lines, build up waiting lines, and introduce intolerable delays in responding to requests.

There are at least two obvious ways in which the increased load could be handled. One would be to increase the capacity of MEDLARS to process the requests, coupled with an increase in communication lines, and possibly coupled with remote input/output facilities located at centers of large demand.

The other solution would be to decentalize the actual servicing of the requests through the use of remotely located search capabilities. The input processing (acquisition, indexing, conversion, and storage for retrieval) would all be done centrally at the National Library of Medicine. But the magnetic tapes comprising files that could be searched would then be reproduced by the Library and distributed to the remotely located search centers.

Any study of the need for and feasibility of such a system would, of necessity, require the determination of the true needs of those whom the system would be designed to serve.

STORAGE AND RETRIEVAL OF GRAPHIC IMAGES

As MEDLARS evolves and more consumers utilize the proposed system, the present Interlibrary Loan procedures and system will require modifications to handle the increased loads. During 1961 over 100,000 requests were received and 2 million pages of copy were provided. These loads, although large, are now handled by a manual microfilming system utilizing one Xerox Copy Flo machine and one Xerox 914 Copier.

An investigation of the present and future requirements in this area has resulted in recommendations regarding the next steps to be taken.

MECHANIZATION OF OTHER LIBRARY TRANSACTIONS

When MEDLARS is operational, emphasis will shift to the adding of functions to the system. It is planned that additional journals will be indexed, the serial record will be automated, and catalog cards will be prepared on the Graphic Arts Composing Equipment. Other tasks that can be accomplished by the system will be automated as time and manpower permit.

SUMMARY

Starting some time during early 1964, MEDLARS will provide the National Library of Medicine with an improved central medical information service that will benefit "consumers" who are primarily in the fields of medical research, education, and practice. It will give direct support to the Nation's medical research program by handling the vast and increasing quantities of published information produced by the medical scientific world.

To this end, MEDLARS will provide the Library with a high-speed, data-processing facility that will:

- Reduce the time required to produce Index Medicus.
- Improve the quality of and enlarge Index Medicus.
- Make possible the production of other bibliographies similar to Index Medicus in form and content.
- Make possible the prompt and efficient servicing of requests for special bibliographies on both a demand and a recurring basis.
- Keep statistics and perform analyses that will provide the information required to monitor and improve the system's effectiveness.
- Reduce the need for medical literature-screening elsewhere that would duplicate the Library's service.

The MEDLARS system will accept and process both indexed data and bibliographic search requests on a daily basis. It will have a very large storage capacity and it will be capable of high-speed searching. It will provide on-line computer printing and off-line, offset-printing-quality camera copy at a reasonable cost. It will be capable of monitoring its own operations periodically. It will be possible to adapt the system to other secondary objectives that will satisfy future requirements of the National Library of Medicine.

EXHIBIT A

Journal Data Sheet

(Preliminary)

A—anonymous
D—non-std date
P—non-std pagination

Pagination

Opt. Pagination

Senior Author Code

Author Data

Listo matic Code

Title (translation)

Title (vernacular or transliterated vernacular)

Reviser

SUBJECT ANALYSIS

Code No.	NIM	IM	Heading
515958			REVIEW — references
			AGE GROUPS
303206			INFANT, NEWBORN (to 1 mo.)
302802			INFANT (1 mo. to 2 yrs.)
116459			CHILD (2-12 yrs.)
014100			ADOLESCENCE (13-18 yrs.)
248708			AGED (66 yrs. & older)
480301			PREGNANCY
			EXPER. ANIMALS(B1,B2)-Specify
			RACE -(I)- Specify

NON IM — CHECK TAGS

Code No.		Heading
		TYPE OF ARTICLE
281907		History
216457		Experimental (lab) study
600353		Toxicologic Report
478255		Popular
		GEOGRAPHIC HEADING-Specify
		PHS SUPPORT GROUPS
496455		Public Health Service
401225		Nat. Insts. of Health
401212		NCI (Cancer)
401214		NHI (Heart)
401216		NIAID (Allergy)
401218		NIAMD (Arthritis)
401220		NICHHD (Child Health)
401222		NIDH (Dental Health)
401224		NIGMS (Gen.Med.Sci.)
401226		NIMH (Mental Health)
401228		NINDB (Neuro.-Blind)
179242		Div.of Research Facilities
092262		Bureau of Medical Services
092270		BSS-Community Health
092288		BSS-Environmental Health
		PROVISIONAL HEADINGS

EXHIBIT B

Unit Record on Typed Copy

(Hypothetical Example of Foreign-Journal Article)

J/	ANP	0911	14-3	*Aug 58	R/
C/	P	*Z12	*1		
A/	GRANONE FG: Interrelation of the biological action of ionizing radiations, the aging process and longevity. Acta Neurol (Nap) 14:A12, Aug 58 (It)				
V/	Rapporti tra azione biologica delle radiazioni ionize anti, processo di invecchiamento e duraga della vita.				
T/					
000123	*(27)				
271025					
032508					
032201					
001111					
X/	R/				

Unit Record Information on Paper Tape

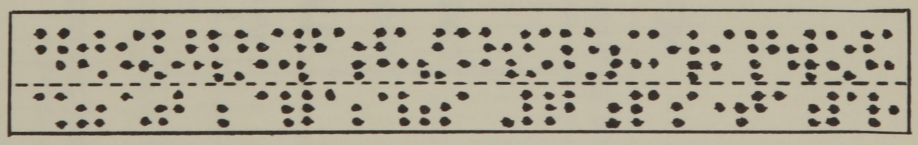


EXHIBIT C

Request for MEDLARS Literature Search

(Preliminary)

Requester's Name and Address

John Jones, Librarian
University Medical School Library
Baltimore, Maryland

Requester's Specialty:

- [ ] A. Research
- [ ] B. Clinical
- [x] C. Education
- [ ] D. Student
- [ ] E. Other ____

Affiliated with:

Community University
Baltimore, Maryland

If engaged in work connected with PHS or other grant, please specify grant number:

Problem Statement:

I would like a bibliography comparing the treatment of histoplasmosis and coccidioidomycosis with nystatin and/or amphotericin B. If possible, I would like the references arranged by author.

Number of Citations Required:

- [ ] Small (1-25)
- [x] Medium (26-100)
- Large (over 100) Specify: ____

Language(s) Preference
Specify: No preference

EXHIBIT D

Demand Search Data Sheet

(Preliminary)

DEMAND SEARCH DATA SHEET

Request Number	Searcher's No.	Number of Citations Retrieved	Number of Citations Sent	Date
640021	05	32	27	1/31/63

REQUEST ELEMENT DEFINITION

Code	Definition	Frequency	Code	Definition	Frequency
M1	NYSTATIN	19			
M2	AMPHOTERICIN B	28			
M3	HISTOPLASMOSIS	86			
M4	COCCIDIOIDOMYCOSIS	49			

Search Equations:

R: (M1 + M2) * (M3 + M4)

EXHIBIT E

GRACE CIM Page

(Preliminary)

TAENIA

PSYCHOANALYTIC THERAPY (A 01)

Psychotherapy and psychoanalysis BOULANGER JB. Can Med Ass J 6:123-6, Jun 61 (Fr)
Psycho-analysis of a stammering girl. BALKANYI C. Int J Psychoanal 42:97-109, Jan-Apr 61
Introjection, reprojection, and hallucination in the interaction between schizophrenic patient and therapist. CAMERON N. Int J Psychoanal 42:86-96, Jan-Apr 61
The compensatory work of transference. EVAN S. Psychoanalysis 48:19-29, Summer 61
A transference manifestation of male homosexuals. FINE R. Psychoanalysis 48:116-20, Summer 61
The disguised cry for help: narcissistic mothers and their children. LOVE S, et al. Psychoanalysis 48:52-67, Summer 61
The rule of abstinence in psychoanalysis SCHEUNERT G. Psyche (Stuttg) 15:105-23, May 61 (Ger)
Anxiety concerning change, as seen in the psychotherapy of schizophrenic patients—with particular reference to the sense of personal identy. SEARLES HF. Int J Psychoanal 42:74-85, Jan-Apr 61
Countertransference crisis in suicidal attempts. TABACHNICK N: Arch Gen Psychiat (Chicago) 4:572-8, Jun 61
The analytic screen: an instrument or an impediment in the psycho-analytic technique. VESZY-WAGNER L. Int J Psychoanal 42:32-42, Jan-Apr 61
Seating position as nonverbal communication in group analysis. WINICK C, et al. Psychiatry 24:171-82, May 61

PNEUMONIA, INTERSTITIAL PLASMA CELL (E 05, L 11)

On a home infection with interstitial pneumonia. ABEGG W. Praxis 49:50-3, 21 Jan 60 (Ger)
Pneumonia due to Pneumocystis carniii. KOCHMAN R. Hartfuah 58:292, 1 May 60 (Heb)
Pneumocystis carnii pneumonia. Report of a case and review of the recent literature. WILSON JF, et al. Pediatrics 25:468-76, Mar 60
Pneumocystis carnii pneumonia associated with hypogammaglobulinaemia. McKAY E, et al. Lancet 2:713-5, 31 Feb 59

PROTEUS

Biological properties of a B. Proteus vulgaris strain resistant to nicotinic acid thioanilide. MARKOV KI, et al. Z Naturforsch (B) 16B:258-9, Apr 61 (Ger)
New concepts on the lysotype of Proteus hauseri: lysotype I. Zbl Bakt (Orig) 182:49-56, May 61 (Fr)

chemistry

Gelatinase activity of B. proteus vulgaris HUBAREV EM, et al. Mikrobiol Zh (Kiev) 23:39-45, 1961 (Uk)

culture

Dynamics of changes of the oxidation-reduction potential and pH in media of pure and mixed cultures. II. Changes of the oxidation-reduction potential and pH in media of pure and mixed cultures of Proteus, Staphylococcus aureus and Bacillus pyocyaneus GIMRANOV MG. Zh Mikrobiol 32:92-8, Apr 61 (Rus)

metabolism

Research on the enzymatic action, with respect to casein, of culture filtrates of some microbial species. RICHOU R, et al. C R Acad Sci (Par) 252:3345-6, 24 May 61 (Fr)
On the mechanism of anaerobic formation of succinate by Proteus vulgaris. SASAKI S, et al. J Biochem (Tokyo) 49:421-6, May 61

PSEUDOHERMAPHRODITISM see DIPHTHERIA

PSEUDOHERMAPHRODITISM see HERMAPHRODITISM

PSEUDOLEUKEMIA see ANEMIA, LEUKOERYTHROBLASTIC

PSEUDOSCLEROSES, see under HEPATOLENTICULAR

PSEUDOTUBERCULOSIS PASTEURELLA, see under PASTEURELLA

PSETTUCOSIS LYMPHOGRANULOMA GROUP VIRUSES see under MIYAGAWANELLA

RICINUS (F02)

see also CASTOR OIL
Permeability measurements of castor-bean seed indicative of cold-test performance. THOMAS CA. Science 131:1045-6, 8 Apr 61

PSYCHIATRY (A 12)

see also related MENTAL DISORDERS
see also related MENTAL HYGIENE
see also related PSYCHOANALYSIS
see also related PSYCHOSOMATIC MEDICINE
Research in psychiatry. DIETHELM O, et al. Canad Med Ass J 6:108-22, Jun 61

Home medical care in the provinces and abroad FIESSINGER H. Rev Prat (Par) 11(4) Suppl:i-iii, 11 May 61 (Fr)
Psychiatry. MUKHERJI DR. Indian Med J 55:85-90, Apr 61
The future for psychiatry in East Africa. MARGETTS EL. E Afr Med J 37:448-56, Jun 60
Psychiatric Research Institute.. PROKUPEK J. Cesk Psychiat 57:73-5, Apr 61 (Cz)
Considerations on the current status of psychiatry in Canada. QUEROL M. Rev Neuropsiquiat 23:240-52, Jun 60 (Sp)
The psychiatric nurse and the dictor-nurse relationship. RODRIGUEZ RJ. Quart Bull Northw Univ Med Sch 45:116-9, Summer 61

RICINUS (F 02)

see also specific CASTOR OIL
Permeability measurements of castor-bean seed indicative of cold-test performance. THOMAS CA. Science 131:1045-6, 2 Apr 61

PROTEUS

Biological properties of a B. Proteus vulgaris strain resistant to nicotinic acid thioanilide MARKOV KI, et al. Z Naturforsch B 16B:258-9, Apr 61 (Ger)
New concepts on the lysotype of Proteus hauseri: lysotype I VIEU JF, et al. Zbl Bakt (Ori) 182:49-56, May 61 (fr)

chemistry

Gelatinase activity of B. Proteus vulgaris HUBAREV EM, et al. Mikrobiol Zh (Kiev) 23:39-45, 1961 (Uk)

culture

Dynamics of changes of the oxidation-reduction potential and pH in media of pure and mixed cultures. II. Changes of the oxidation-reduction potential and pH in media of pure and mixed cultures of Proteus, Staphylococcus aureus and Bacillus pyocyaneus. GIMRANOV MG. Zh Mikrobiol 32:92-8, Apr 61 (Rus)

metabolism

Research on the enzymatic action, with respect to casein, of culture filtrates of some microbial species RICHOU R, et atl. C R Acad Sci (Par) 252:3345-6, 24 May 61 (Fr)
On the mechanism of anaerobic formation of succinate by Proteus vulgaris SASAKI S, et al. J Biochem (Tokyo) 49:421-6, May 61

PSYCHOANALYTIC THERAPY (A 01)

see also related MENTAL DISORDERS
see also related MENTAL HYGIENE
see also related PSYCHOANALYSIS
see also related PSYCHOSOMATIC MEDICINE
Research in psychiatry. DIETHELM O, et al. Canad Med Ass J

They're learning to live again. FINDLEY AP. Amer J Nurs 61: 84-6, Jun 61
Psychotherapy of a withdrawn schizophrenic. Participation of a nurse. MANDELL AJ, et al. Arch Gen Psychiat (Chicago) 4:597-602, Jun 61

PSEUDODIPHTHERIA see DIPHTHERIA

PSEUDOHERMAPHRODITISM see HERMAPHRODITISM

PSEUDOLEUKEMIA see ANEMIA, LEUKOERYTHROBLASTIC

PSEUDOSCLEROSES, see under HEPATOLENTICULAR

PSEUDOTUBERCULOSIS PASTEURELLA, see under PASTEURELLA

PSETTUCOSIS LYMPHOGRANULOMA GROUP VIRUSES see under MIYAGAWANELLA
PSEUDODIPHTHERIA see DIPHTHERIA

PSEYDOHERMAPHRODITISM see HERMAPHRODITISM

PSEUDOLEUKEMIA see ANEMIA, LEUKOERYTH BLASTIC

PSEYDOSCLEROSES, see under HEPATOLENTICULAR

PSEUDOTUBERCULOSIS PASTEURELLA, see under PASTEURELLA

PSETTUCOSIS LYMPHOGRANULOMA GROUP VIRUSES see under MIYAGAWANELLA

RICINUS (F 02)

see also specific CASTOR OIL
Permeability measurements of castor-bean se d indicative of cold-test performance. THOMAS CA. Science 131:1045-6, 8 Apr 61

PROTEUS

New concepts on the lysotype of Proteus hauseri: lysotype I. VI U JF, et al. Zbl Bakt (Orig) 182:49-56, May 61 (Fr)

On a home infection with interstitial pneumonia. ABEGG W. Praxis 49:50-3, 21 Jan 60 (Ger)
Pneumonia due to pneumocystis carinii. KOCHMAN R. Hartfuah 58:292, 1 May 60 (Heb)
Pneumocystis carinii pneumonia. Report of a case and review of the recent literature. WILSON JF, et al. Pediatrics 25:468-76, Mar 60

blood

Pneumocystis carinii pneumonia associated with hypogammaglobulinaemia. McKAY E, et al. Lancet 2:713-5, 41 Feb 59

PROTEUS

Biological properties of a B. Proteus vulgaris strain resistant to nicotinic acid thioanilide. MARKOV KI, et al. Z Naturforsch B 16B:258-9, Apr 61 (Ger)
New concepts on the lysotype of Proteus hauseri: lysotype I. VIEU JF, et al. Zbl Bakt Orig 182:49-56, May 61 (Fr)

chemistry

Gelatinase activity of B. Proteus vulgaris. HUBAREV EM, et al. Mikrobiol Zh (Kiev) 23:39-45, 1961 (Uk)

culture

Dynamics of changes of the oxidation-reduction potential and pH in media of pure and mixed cultures. II. Changes of the oxidation-reduction potential and pH in media of pure and mixed cultures of Proteus, Staphylococcus aureus and Bacillus pyocyaneus. GIMRANOV MG.Zh Mikrobiol 32:92-8, Apr 61 (Rus)

metabolism

Research on the enzymatic action, with respect to casein, of culture filtrates of some microbial species. RICHOU R, et al. C R Acad Sci (Par)252:3345-6, 24 May 61 (Fr)
On the mechanism of anaerobic formation of succinate by Proteus vulgaris. SASAKI S, et al. J Biochem (Tokyo) 49:421-6, May 61

PSEUDODIPHTHERIA see DIPHTHERIA

PSEUDOHERMAPHRODITISM see HERMAPHRODITISM

PSEUDOLEUKEMIA see ANEMIA, LEUKOERYTHROBLASTIC

PSEUDOSCLEROSES, see under HEPATOLENTICULAR

PSEUDOTUBERCULOSIS PASTEURELLA, see under PASTEURELLA

PSETTUCOSIS LYMPHOGRANULOMA GROUP VIRUSES, see under MIYAGAWANELLA

RICINUS (F 02)

see also specific CASTOR OIL
Permeability measurements of castor-bean seed indicative of cold-test performance. RHOMAS CA. Science 131:1045-6, 8 Apr 61

PSYCHIATRY (A 12)

see also related MENTAL DISORDERS
see also related MENTAL HYGIENE
see also related PSYCHOANALYSIS
see also related PSYCHOSOMATIC MEDICINE
Research in psychiatry. DIETHELM O, et al. Canad Med Ass J 6:108-22, Jun 61
Home medical care in the provinces and abroad. FIESSINGER H. Rev Prat (Par) 11(4) Suppl:i-iii, 11 May 61 (Fr)
Psychiatry. MUKHERJI DR. Indian Med J 55:84-90, Apr 61
The future for psychiatry in East Africa. MARGETTS EL. E Afr Med J 37:448-56, Jun 60
Psychaitric Research Institute. PROKUPEK J. Cesk Psychiat 57:73-5, Apr 61 (Cz)
Considerations on the current status of psychiatry in Canada. QUEROL M. Rev Neuropsiquiat 23:240-52, Jun 60 (Sp)
The psychiatric nurse and the doctor-nurse relationship. RODRIGUEZ RJ. Quart Bull Northw Univ Med Sch 35:116-9, Summer 61
Special series on hospital planning. I. Psychiatry in the general hospital. SCLARE AB. Scot Med J 6:203-9, May 61
Psychiatric aspects of Sir Thomas Browne with a new evaluation of his work. SCHNECK JM. Med Hist 5:157-66, Apr 61
Acquisitions in psychiatric science 1958 and 1959. STROMGREN RE. Nord Psykiat T 14:1-15, 1960 (Dan)
Some problems in the theory of psychiatry. VISHNEVSKII RE.

Zh Nevropat Psikhiat Korsakov 61:769-72, 1961 (Rus)
Yet another college?Lancet 1:1485-6, 24 Jun 61
chotherapy of schizophrenic patients—with particular reference the sense of personal identy. SEARLES HG. Int J Psychoanal 42:74-85, Jan-Apr 61
Countertransference crisis in suicidal attempst. TABACHNICK N. Arch Gen Psychiat (Chicago) 4:572-8, Jun 61
The analytic screen: an instrument or an impediment in the psycho-analytic technique. VESZY- AGNER L. Int J Psychoanal 42:32-42, Jan-Apr 61
Seating position as nonverbal communication in group analysis. WINICK C, et al. Psychiatry 24:171-82, M y 61

GLOSSARY

COMMON ABBREVIATIONS

CCF	- Compressed Citation File
EPCF	- Expanded Processed Citation File
GRACE	- Graphic Arts Composing Equipment
ICF	- Intermediate Citation File
JAR	- Journal Article Record
JRF	- Journal Record File
LANDS File	- Language and Subheadings File
LJI	- List of Journals Indexed
MDT	- MEDLARS Dictionary Tape
MESH	- Medical Subject Headings
NJAR	- Non-Journal Article Record
PCF	- Processed Citation File
RED List	- Request Element Definition List

DEFINITION OF TERMS

ALPHANUMERIC - A contraction of "alphabetic-numeric"; characters which may be letters of the alphabet or numerals or their symbols.

BINARY - A number system using only two symbols, the digits one and zero.

BIT - A contraction of "binary digit". The smallest unit of information, corresponding to either one or zero.

CLASSIFICATION NUMBERS - Numbers added by the computer to a citation's subject headings to indicate their location within hierarchically structured categories of related subjects.

CODE - Information presented in the language that the computer can understand and handle; to code is to put into machine language.

CODING - A list of successive computer operations in code for solving a given problem. Also the act of writing a prepared list in code.

CONTROL TAPE - A removable loop of punched paper tape that feeds through the reader mechanism on the input typewriters. It contains information that automatically formats and codes typewriter output.

COMPUTER WORD - A group of 48 bits which the computer manipulates as a unit.

CONVERSION - The process of changing the representation of information to a form usable by the computer, e.g., converting it to machine language on paper tape.

CORE MEMORY - The computer's high-speed memory in which information is stored internally on magnetic cores.

CORRECTION TAPE - Punched paper tape containing corrections to citations which are to be entered into the computer.

DEBUGGING - The process of determining the correctness of a computer routine by test operation.

EDIT - To check, validate, and/or rearrange information; it may involve the deletion of unwanted data, the selection of pertinent data, or the insertion of additional data.

FILE - An organized collection of information arranged or classified for convenient reference. (Files in machine language on magnetic tape are listed in the next section of this Glossary.)

GRACE - The Graphic Arts Composing Equipment, a photocomposition device that converts information from magnetic tape into latent typographical images on photosensitive film or paper.

IM TAGS - Descriptive tags assigned to an article, monograph, or other document by an indexer or cataloger under which the article, monograph, or other document is to be listed in Index Medicus.

INSTRUCTION - A machine "word" or set of characters which tells the computer to perform a certain operation.

ITEM - The contents of a single message, such as the author's name or the place of a journal's publication.

LOGICAL OPERATIONS - Computer operations which are logical in nature, involving a logical "and", logical "or", their negations (not "and", not "or") and decisions.

MACHINE LANGUAGE - Information in the physical form that a computer can handle, such as properly coded information punched on paper tape or cards or in electrical or magnetic form on magnetic tape.

MAGNETIC CORE - A doughnut- or ring-shaped storage device in which binary information is represented by the direction of magnetization.

MAGNETIC TAPE - Tape coated with a magnetic material which can store information.

MODULE - Part of a program (see PROGRAM).

NON-IM TAGS - Descriptive tags assigned to an article, monograph, or other document by an indexer or cataloger under which the article, monograph or other document will not be listed in Index Medicus. Such tags may be used as retrieval criteria for servicing demand and recurring bibliographies.

OFF-LINE - Operating independently of the computer.

ON-LINE - Operating under the direct control of the computer.

OPERATION - A specific action that the computer will perform whenever the instruction calls for it.

PARAMETER - A quantity to which arbitrary values may be assigned, such as the number of columns requested in the printing of a demand bibliography.

PERIPHERAL EQUIPMENT - Auxiliary machines (e.g., magnetic tape handlers) which may be placed under the direct control of the computer.

PRIMARY STORAGE - The main internal storage of the computer.

PROGRAM - A sequence of steps or coded instructions to be executed by the computer to solve a given problem. A program may be divided into "modules", which, in turn, may be divided into "submodules". Also, a submodule may consist of several "routines", each of which may be made up of two or more "subroutines". The relative number of operations involved and their specificity generally determines which term applies. To program is to plan the steps necessary for the computer to solve all or part of the problem.

RED LIST - Request Element Definition List, the elements of a demand search request which are put into an equation(s) by the searcher and used by the computer in retrieval operations.

REQUEST STATEMENT - Request information in machine language, including requester's identifying information, the search equation(s), and output formatting instructions.

ROUTINE - Part of a submodule (see PROGRAM).

SEARCH ELEMENTS - Tags and other terms and their symbols as used in search equations.

SEARCH EQUATION - An equation prepared by the searcher using the symbols of search elements and logically indicating the types of citations requested. It is a symbolic representation of the retrieval criteria for a search request.

SIGNIFICANT ELEMENTS - Search elements which are of greatest value as retrieval criteria in a high-speed, gross screening or search of the stored citations to service a particular request.

SORT - To arrange items of information into a particular order.

STORAGE - A device for retaining information in machine language. "Memory" usually means storage inside the computer.

SUBMODULE - Part of a module (see PROGRAM).

SUBROUTINE - Part of a routine (see PROGRAM).

SUBSEARCH - An expanded or more general form of a search equation for purposes of retrieving more citations than the initial search equation yields.

TABLE - Synonymous with FILE.

TABULATING EQUIPMENT - Equipment using punched cards.

TAG - A descriptor of the contents of an article or a document which is assigned by an indexer or cataloger. Medical subject headings, age group designations, and language designations, for example, are TAGS.

TRANSITION PERIOD - The first year of MEDLARS operation when the existing mechanized system will be producing Index Medicus and associated publications as usual and citations will be entered into MEDLARS' files for storage and future retrieval.

UNIT RECORD - A citation plus the tags assigned to it by an indexer.

MAGNETIC-TAPE FILES

ANNUAL MESH FILE - Contains one record for each tag, "See" entry, and "See Under" entry.

COMPRESSED CITATION FILE (CCF) - A compact unit record file which is searched in servicing demand search requests.

GRACE TAPES - Tapes prepared for use as the input to the Graphic Arts Composing Equipment.

MEDLARS DICTIONARY TAPE (MDT) - Consists of six files on one reel of magnetic tape, in the following sequence:

1. LANDS FILE - Contains language abbreviation codes and form tags for use as subheadings in listing monographs in Index Medicus.

2. RECURRING BIBLIOGRAPHY PARAMETER FILE - Contains all the parameters required to completely describe the format (number of columns, lines to a page, etc.) of recurring bibliographies.

3. JOURNAL RECORD FILE - Contains all pertinent journal information.

4. DAILY MESH FILE - Contains only medical subject headings.

5. RECURRING BIBLIOGRAPHY PRINT HEADING FILE - Provides special headings for the printing of individual recurring bibliographies.

6. RECURRING BIBLIOGRAPHY SELECTION CRITERIA FILE - Contains the criteria for selecting citations for inclusion in specific recurring bibliographies.

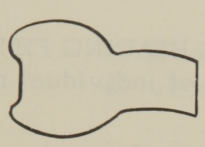
MESH MASTER FILE - Contains one record for each main heading, "See" entry, "See Under entry, "See Also Related" entry, and "See Also Specific" entry.

PROCESSED CITATION FILE (PCF) - A unit record file which contains preprocessed citations for the output required to produce Index Medicus and all recurring bibliographies.

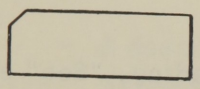
PRODUCTION RUN TAPE - Contains all the operational programs used by the MEDLARS computer.

REQUEST HISTORY FILE - Contains a chronological listing of all the demand search requests that have been processed over a given period.

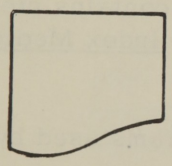
DEFINITION OF SYMBOLS



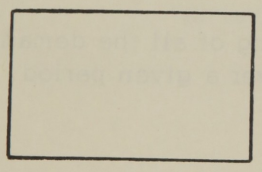
Magnetic Tape



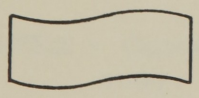
Cards



Computer Printouts



Computer Operations



Paper Tape

☆ U. S. GOVERNMENT PRINTING OFFICE : 1963 O - 684850