DOCUMENT RESUME

ED 046 434 LI 002 450

AUTHOR	Hayes, R. M.; Maron, M. E.
TITLE	Institute of Library Research Annual Report, July 1969 - June 1970.
INSTITUTION	California Univ., Los Angeles. Inst. of Library Research.
PUB DATE	30 Jun 70
NOTE	41p.
EDRS PRICE	EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS	*Annual Reports, *Computer Programs, Electronic Data Processing, Information Services, *Information Systems, *Library Research, *Library Services, University Libraries
IDENTIFIERS	*California, Institute of Library Research, Library Automation

ABSTRACT

This "Annual Report," July 1969 to June 1970, is the fourth progress report of the Institute of Library Research. The activities of the Institute encompassed the completion of several projects and the initiation of other projects during this period. One recorded milestone of singular importance was the establishment in June 1970 of the Library System Development Project. The goal of this intramurally funded activity is to design, develop, and test computer based modules and subsystems for implementation on the campuses of the University of California. Each project can be identified as helping to resolve needs and issues in one or more of the following six areas of concern: (1) values in library and information service, (2) library and information systems, (3) methodology research, (4) information services development, (5) social and professional issues, and (6) mechanization of library processes. (Author/NH)

INSTITUTE OF LIBRARY RESEARCH

ANNUAL REPORT

July 1969 - June 1970

R. M. Hayes

M. E. Maron

30 June 1970

Institute of Library Research

University of California

Los Angeles and Berkeley, California

TABLE OF CONTENTS

ABSTRACT

This "Annual Report", July 1969 to June 1970, is the fourth progress report of the Institute of Library Research. The activities of the Institute encompassed the completion of several projects and the initiation of other projects during this period. One recorded milestone of singular importance was the establishment in June 1970 of the Library System Development Project. The goal of this intramurally funded activity is to design, develop, and test computer based modules and subsystems for implementation on the campuses of the University of California.

ANNUAL REPORT

JULY 1969 - JUNE 1970

INTRODUCTION

Projects undertaken during the fifth year of full-scale operation of the Institute of Library Research continue the thrust to understand and to solve the problems in better library service to needs for information. Consequently each project can be identified as helping to resolve needs and issues in one or more of the following six areas of concern:

(1) values in library and information service

(2) library and information systems

(3) methodology research

(4) information services development

(5) social and professional issues

(6) mechanization of library processes

Values in Library and Information Service

Many areas of endeavor perform tasks or services that can be measured in terms of need satisfaction. Other types of activity are less blessed and find great difficulty in assigning values to various system facets (which include function, configuration, operation, and output). Library and information services are often difficult to assess. This Institute program has been concerned with the social and scholarly environment of the library and other information activities, and with

many aspects of value determination. What goals are to be served? How can the quality of library collections be assessed? How are books and information used? What are the criteria for the design of library systems? What is the value of library services in education, research, and management, and how can it be measured? What is the relationship between the value of services and the response time of them? What are the problems in gaining acceptance of new services, both by the library and its users, and how can they be solved? What are the values to be gained by in-service education of librarians? Specific studies in this area during the past year have included criteria for evaluation of library effectiveness and the benefits of mechanization.

Library and Information Systems

This program has been concerned with the nature of library and information systems, the definition of their components and of the interrelationships of those components at various levels--local, university-wide, regional, and national. Characteristic projects in this area have included the analysis of present library organizations, experimentation with new or alternative forms and types of services, and the planning, conceptual development, design, testing and evaluation of new systems.

Methodology Research

This program has been concerned with the development of basic methods by which libraries and information systems can be designed and evaluated. Characteristic projects have studied the tools for development of vocabulary, for file indexing and organization, for cost accounting in information activities, for the measurement of file effectiveness, for

definition of user needs, for evaluation of equipment and physical facilities, for the organization of catalogs, for the mechanization of union catalog preparation, and for the evaluation of library mechanization programs.

Information Services Development

This program has been concerned with the development of a general purpose capability for providing computer-based information services within the library, which can be utilized by any university activity desiring to establish an information center, together with research into the problems involved in such a capability. A Center for Information Services will mobilize the intellectual facilities (knowledge of computer programs for language data processing, of man-machine communication problems, of analysis and simulation based on data files) of the campus for experimentation with advanced services that are being encompassed by university libraries.

Social and Professional Aspects

This program has been concerned with the professional and social aspects of librarianship. It is intended to support and extend the programs of the library schools through joint faculty appointments and the employment of students as research assistants. The Institute has experimented with the development of new curricula; it has sponsored conferences, symposia, seminars, and special courses in cooperation with University Extension. In addition, funds have been separately sought to offer advanced post-graduate training programs for practicing librarians and fellowships to bring experienced persons to the Institute for work on specific research projects. It has undertaken a research project concerning problems in education for librarianship.

Mechanization of Library Processes

This program has been concerned with the development of a long term plan and creation of an organization for the automation of university-wide library functions for the University of California campuses. This plan and associated proposal originated with a request from President Hitch that the Institute of Library Research develop a program of mechanization for the University's libraries. In response, a Library System Development Program document was prepared and submitted to the University through the Program Review Committee of the University of California Library Council. This submission resulted in the establishment of the Library System Development Project housed at Santa Barbara and funded by the University to conduct library automation feasibility studies during fiscal year 1970-71.

ADMINISTRATION

The Institute of Library Research has two main branches, one at Los Angeles and one at Berkeley. The Task Force Project, which has been centered at Berkeley since its inception, was restructured in June 1970 as the University of California Library System Development Project and located at the University of California, Santa Barbara. Figure 1 is an organization chart showing this structure. The administration, under the directors and managers, is organized along project lines with heavy use of student research assistants. Figure 2 is a staff list, by project, for this reporting period.

FIGURE 1

UNIVERSITY OF CALIFORNIA

INSTITUTE OF LIBRARY RESEARCH

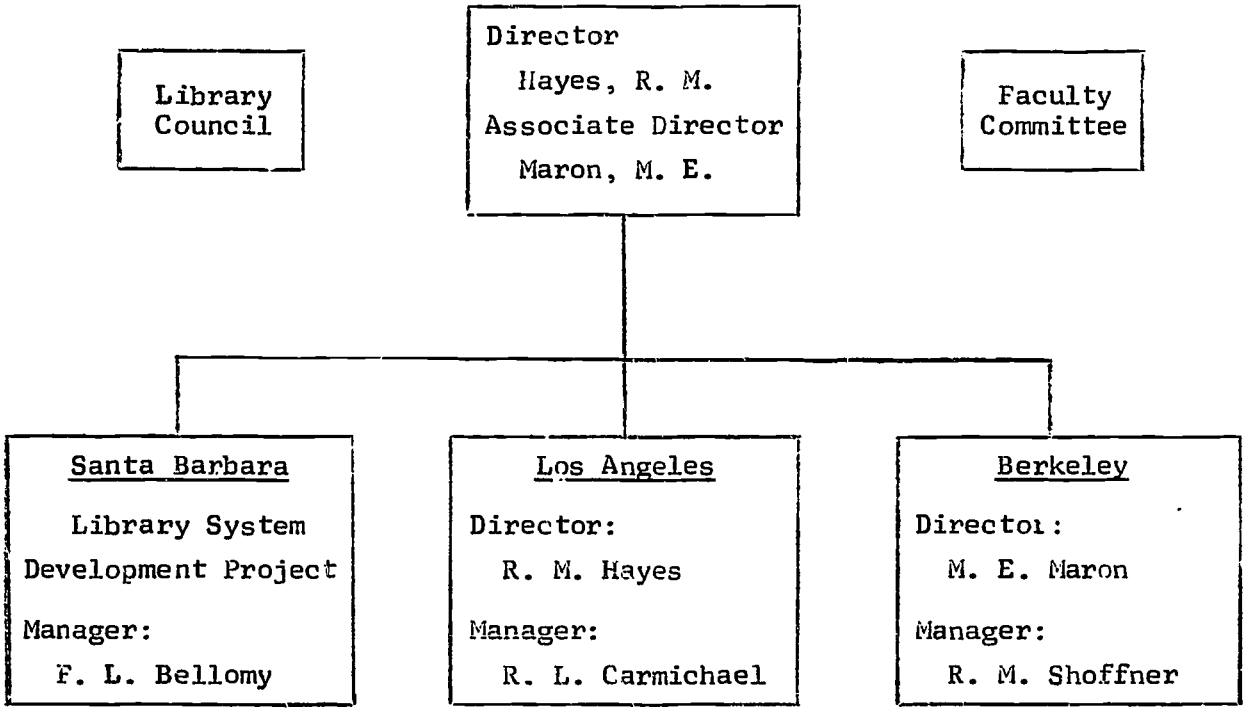

FIGURE 2

FILE ORGANIZATION: ON LINE SEARCH OF BIBLIOGRAPHIC HOLDINGS
(Office of Education)

Arjun Aiyer	Programmer
Patricia Barkley	Editor
Irene Barr	Research Assistant
Mary Benson	Research Assistant
John Courtright	Programmer
Jay Cunningham	Associate Specialist
Ruth Dunham	Junior Specialist
Margaret Ely	Programmer
Jane Farthing	Research Assistant
Martin Fried	Programmer
William Harrelson	Coder
Lois Harzfeld	Research Assistant
Jorge Hinojosa	Junior Specialist
Luke Howe	Junior Specialist
Allan Humphrey	Specialist
Stephen Johnson	Programmer
Ashok Kulkarni	Research Assistant
Katsuhiko Kurano	Coder
Donald O'Neal	Research Assistant
Chakravarthi Ravi	Programmer
John Reinke	Programmer
Steven Silver	Senior Programmer
Ralph Shoffner	Specialist
Stephen Smith	Associate Specialist
Keith Stirling	Research Assistant
Irene Travis	Assistant Specialist

SEMANTIC FOUNDATIONS FOR QUESTION-ANSWERING SYSTEMS
(Public Health Service)

John Olney	Principal Programmer
Kenneth Kress	Senior Coder
Bruce Stein	Research Assistant

CENTER FOR INFORMATION SERVICES PHASE II
(National Science Foundation)

Kevin Reilly	Associate Researcher
Robert Carmichael	Senior EDP Systems Analyst
Stuart Beal	Senior Statistician
Peter Watson	Librarian
Ida Riordan	Senior Programmer
Aeint de Boer	Programmer
Georgette Silva	Assistant Researcher

FIGURE 2 (Continued)

CENTER FOR INFORMATION SERVICES PHASE II (Continued)
(National Science Foundation)

May Lieu	Coder
Hilda Borko	Coder
Ricky Bosted	Research Assistant
Betty Butler	Research Assistant
Nancy Brault	Librarian
Richard Weisbrod	Research Assistant
Cyril Treister	Research Assistant
Linda Miroff	Programmer
Betsey Beamish	Librarian
Ray Bell	Librarian
Anthony Hall	Librarian
Lorraine Mathies	Librarian
Kay Forrest	Librarian (UC Riverside)
John Knapp	Librarian (UC Berkeley)
Thomas Burgess	Spvr. EDP Systems Analyst
Bruce Briggs	Principal Programmer
James Adams	Senior Programmer
James Lewis	Principal Programmer
Peter Donahoe	Senior Programmer
Jerry Pine	Assistant Programmer
Barbara Blankenship	Senior Programmer
Neil Ludlam	Senior Programmer
William Jordan	Senior Programmer
Robert Hayes	Principal Investigator
William Kehl	Co-Principal Investigator
Robert Vosper	Co-Principal Investigator

UNIVERSITY OF CALIFORNIA UNION CATALOG SUPPLEMENT
(University of California)

Deborah Barrett	Assistant Specialist
Ruth Dunham	Junior Specialist
Margaret Ely	Programmer
Martin Fried	Programmer
Elizabeth Gibson	Research Assistant
Thomas Hargrove	Assistant Specialist
William Harrelson	Keypunch Operator
Ugertha Hodges	Senior Coder
Luke Howe	Junior Specialist
Joanne MacDonald	Bibliographer
Linda Nakamura	Research Assistant
John Reinke	Programmer
Victor Rosenberg	Assistant Researcher
Janet Rowell	Laboratory Assistant
Maxine Sansing	Senior Coder

FIGURE 2 (Continued)

UNIVERSITY OF CALIFORNIA UNION CATALOG SUPPLEMENT (Continued)
(University of California)

Donald Sherman	Specialist
Keith Stirling	Research Assistant
Nancy Voigt	Junior Specialist
Harriet Zais	Junior Specialist

OPERATIONS TASK FORCE PROJECT (University of California)

Patricia Barkley	Editor
Ruth Dunham	Junior Specialist
Ann Giglioli	Junior Specialist
William Harrelson	Coder
Luke Howe	Junior Specialist
Gary Matkin	Research Assistant
Rodney Randall	Programmer
Ralph Shoffner	Specialist
G. Edward Evans	Professor
Thomas Burgess	Spvr. EDP Systems Analyst
Robert Carmichael	Senior EDP Systems Analyst
Ida Riordan	Senior Programmer
Aeint de Boer	Programmer
Fred Bellomy	Library Systems Analyst

INFORMATION PROCESSING LABORATORY (Office of Education)

Irene Barr	Research Assistant
John Courtright	Programmer
Edith Darknell	Junior Specialist
Constance Farrow	Junior Specialist
Lois Harzfeld	Research Assistant
Allan Humphrey	Specialist
Katsuhiko Kurano	Coder
Joseph Meredith	Associate Specialist
Edmond Mignon	Assistant Specialist
Rodney Randall	Programmer
Chakravarthi Ravi	Programmer
Stephen Smith	Associate Specialist
Howard White	Junior Specialist
Helen Yen	Junior Specialist

NEEDS FOR RESEARCH IN LIBRARY AND INFORMATION SCIENCE EDUCATION
(Office of Education)

Harold Borko	Professor
Kevin Reilly	Associate Researcher
Nancy Brault	Librarian
Carol Bedoian	Research Assistant
Cynthia Stoltz	Research Assistant

FIGURE 2 (Continued)

EDUCATION OF MEDICAL LIBRARY SYSTEMS ANALYSTS
(National Library of Medicine)

Guy Dobbs	Professor
Kevin Reilly	Associate Researcher
Ida Riordan	Senior Programmer
Linda Miroff	Programmer
Steven Silver	Senior Programmer

EFFECTIVENESS CRITERIA FOR MEDICAL LIBRARIES
(National Library of Medicine)

G. Edward Evans	Professor
Harold Borko	Professor
Stuart Beal	Senior Statistician

PHYSICAL FACILITIES RESEARCH PROJECT (Federal Library Committee)

G. Edward Evans	Professor
Peter Kamnitzer	Professor
Thomas Vreeland	Professor

RESEARCH PROJECTS

Research projects of the Institute during this reporting period included experimentation with data bases stored on magnetic tape, activities directed toward basic problems in on-line information storage and retrieval, and text processing to derive the semantic foundations for question-answering systems.

Organization and Search of Bibliographic Holdings Records in On-Line Computer Systems (Phase II completed during year)

Administrative.

Principal Investigator: Professor M. E. Maron

Project Manager: Allan J. Humphrey

Sponsoring Agency: Office of Education

Period: 7/1/67 to 6/30/68 (Phase I);
7/1/68 to 6/30/70 (Phase II)

Amount: $179,000 Phase I (completed)
$309,873 Phase II (completed)

Purpose. The purpose of this project was to develop and implement a research facility that may be used to investigate problems of file organization and interactive search for on-line retrieval systems accommodating large files of bibliographic records in the MARC II structure format. Various data bases, ranging in size from 500 to 94,000 records, have been created and stored on a direct access storage facility (disk). Major areas of research, development, and experimentation are: code compression for most effective storage representation of the master files, multi-level inverted index structures for file access, input conversion, analysis of file characteristics, facilities for evaluation of complex Boolean search arguments and dynamic processing

of queries, and methods for handling certain classes of predictable user input errors in search keys.

Reports.

Cunningham, Jay L., Allan J. Humphrey and Ralph M. Shoffner. A Study of the Organization and Search of Bibliographic Holdings Records in On-Line Computer Systems: Phase II. Final Report. Institute of Library Research, University of California, Berkeley. 1970. (In preparation).

Shoffner, Ralph M. "Economics of National Automation of Libraries." Technical Paper No. 9. Institute of Library Research, University of California, Berkeley. December 19, 1969. (Also in Library Trends 18:448-463. April, 1970).

_____. "The Organization, Maintenance and Search of Machine Files." In Cuadra, Carlos, ed., The Annual Review of Information Science and Technology. 3:137-167. Chicago. Encyclopedia Britannica. 1968. (Also in Appendix VII, Cunningham, et al. A Study of the Organization and Search of Bibliographic Holdings Records in On-Line Computing Systems: Phase I Final Report. ED 029 679).

Silver, Steven S. and Joseph C. Meredith. The DISCUS Interactive System Users' Manual. (Working Paper). Institute of Library Research, University of California, Berkeley. 1970. 189 p. (Jointly sponsored by File Organization and Information Processing Laboratory Projects.)

Silver, Steven S. A Format Manipulation System for Automatic Production of Natural Language Documents. (Second Edition) Institute of Library Research, University of California, Los Angeles. December 10, 1969.

Semantic Foundations for Question-Answering Systems (Initiated during year)

Administrative.

Principal Investigator: Professor Harold Borko

Project Manager: John Olney

Sponsoring Agency: National Library of Medicine

Period: 5/1/70 to 4/30/71

Amount: $21,185

Purpose. The purpose of this project is to computationally derive linguistic data to be used for foundational work on the semantics of languages suitable for question-answering systems and for fact retrieval.

Reports.

Olney, John. Semantic Foundations for Question-Answering Systems: Research Plan. (Draft). Institute of Library Research, University of California, Los Angeles. October, 1969.

DEVELOPMENT PROJECTS

The development work of the Institute continued to play an important role both in introducing new dimensions of mechanization to the libraries of the University and in formalizing a structure within which coordinated system mechanization could take place. Task Force activities culminated in the University of California Library System Development project. With the earlier feasibility study of mechanized data bases as a base, a study was initiated to specify the design of a Center for Information Services (CIS). This National Science Foundation funded project will eventually become a part of the future automated library system. Another project, the University of California Union Catalog Supplement (UCUCS), is, in itself, an exercise in mechanization; furthermore, the bibliographic data filled magnetic tapes generated to produce the supplement will eventually reside in the CIS data base. The contents of these and similar data tapes also will help service other mechanized or automated library functions.

Development of a Center for Information Services Phase II: Detailed System Design (Phase IIA initiated at start of fiscal year)

Administrative.

Principal Investigator: Professor R. M. Hayes

Co-Principal Investigators: Mr. Robert Vosper
Mr. William B. Kehl

Project Manager: Thomas K. Burgess (October-March)
Robert L. Carmichael (March-June)

Sponsoring Agency: National Science Foundation

Period: 7/1/69 to 12/31/70 (Phase IIA)

Amount: $208,400

Purpose. The National Science Foundation, Office of Science Information Services, awarded this grant to the Institute of Library Research of the University of California for the next step in the development of a "Center for Information Services". The purpose of the Center is to provide a capability, in the University Library, to acquire magnetic tape data bases of a variety of kinds and, in cooperation with the Campus Computing Network, to provide various kinds of mechanized information services from them to the University community. The purpose of the work under the present grant (Phase IIA) is to develop specifications for the Center, programs for the use of multiple data bases, and experimental experience with mechanized information services.

Reports.

Hayes, R. M. Mechanized Information Services in the University Library. Invited paper presented at Annual Meeting of the Association for Research Libraries in Detroit, June 27, 1970.

Hayes, R. M., W. B. Kehl, and R. Vosper. Quarterly Progress Report. The Development of a Center for Information Services Phase II: Detailed System Design and Programming. Institute of Library Research, University of California, Los Angeles.

First: 1 July 1969 to 30 September 1969 (issued October 1969).
Second: 1 October 1969 to 31 December 1969 (issued January 1970).
Third: 1 January 1970 to 31 March 1970 (issued April 1970).
Fourth: 1 April 1970 to 30 June 1970 (issued July 1970).

Reilly, Kevin D. "Task-Oriented Programming Systems for Data Base Information Centers". In Proceedings of the Fifth Annual National Information Retrieval Colloquium. Philadelphia. May 3-4, 1968.

Reilly, Kevin D. and Peter G. Watson. "The Educational and Social Impact of the University Library-Based Center for Information Services". In Proceedings of the Seventh Annual National Information Retrieval Colloquium. Philadelphia. May 7-8, 1970.

Silva, Georgette M. Preliminary Specifications for the C.I.S. Text Processing System. (Draft). Institute of Library Research, University of California, Los Angeles. July 1969.

Watson, Peter G. Center for Information Services, Phase II - Detailed System Design and Programming. First Report on the Acquisition Function. Institute of Library Research, University of California, Los Angeles. December 1969.

Watson, Peter G. "Center for Information Services, Phase II - System Design and Programming. First Report on Public Service. Institute of Library Research, University of California, Los Angeles. April 1970.

Watson, Peter G. and David R. Smith. "Center for Information Services." UCLA Librarian 22 (10). October 1969.

Weisbrod, Richard L. (Human) Information Overload and the Design of Information Systems. (Draft). Institute of Library Research, University of California, Los Angeles, June 4, 1970.

Working Papers.

Briggs, Bruce. Center for Information Services: Status Report on Phase II (Design and Programming). (Draft). June 30, 1970.

de Boer, Aeint H. Overall Approaches to the C.I.S. Reference Retrieval Software. (Draft). March 1970.

de Boer, Aeint H. Selection of Programming Languages for C.I.S. (Draft). December 1969.

de Boer, Aeint H. Summary of Available Files. (Draft). March 1970.

de Boer, Aeint H. The Feasibility of OS/360 for C.I.S. (Draft). December 1969.

Hayes, Robert M. and Peter G. Watson. ISSUES, Pro and Con. (Draft) February 1970.

Reilly, Kevin D. API (American Petroleum Institute) File: Read Routine and Other Information. December 1969.

Reilly, Kevin D. U. S. Census of Households, 1/1000 Sample. Read Routine and Other Information. December 1969.

Silva, Georgette M. Read Routine for the California Code and State Constitution Sample File (ASPEN Tape). (Draft). March 1970.

Silva, Georgette M. Read Routine for the Chemical Abstracts Condensate File. (Draft). July 1969.

Silva, Georgette M. Read Routine for the Communications of Behavioral Biology File. (Draft). August 1969.

Silva, Georgette M. Read Routine for the Computerized Engineering Index File. (Draft). July 1969.

Silva, Georgette M. Read Routine for the Nuclear Science Abstracts (NSA) Entry and Selector Files. (Draft). November 1969.

University of California Union Catalog Supplement

Administrative.

Project Director: R. M. Shoffner

Project Manager: D. Sherman

Sponsoring Agency: University of California

Period: 11/1/68 to 12/31/71 in three phases

Amount: $100,000 Phase I (Completed)
$553,896 Phase II

Purpose. During the period 1963-1967, the libraries of the University acquired over 1,200,000 titles which were new to the individual nine campuses. The purpose of this project is to produce a book form Union Catalog by converting the bibliographic records of these new holdings to machine form and using computer techniques

for its production. It is estimated that the Catalog will contain over 800,000 unique entries.

To achieve maximum value and re-use of the machine record, the project is committed to producing a MARC-compatible record, both as to structure and content. The machine record is created by an approach termed Automatic Field Recognition. In this system, no pre- or post-editing or tagging schemes are required; instead, computer algorithms work with the natural format of the catalog card to identify MARC-defined bibliographic data elements.

The first phase of the project is to develop and test a prototype production and programming system to verify the design concept. This was done during 1969-1970, and a small 5,000 entry book catalog was produced. The second phase of the project is to perform the conversion and system processing on the entire file at an estimated production rate of 20,000 MARC II records per week. This phase is scheduled to begin in August 1970. A third phase is also currently planned whose goal is to upgrade the structure and error-free quality of the data, and also improve the graphic arts quality of potential output products. Work on this third phase is planned to begin in October, 1970.

Reports.

Sherman, Don. Initial Progress Report on Automatic Field Recognition. Technical Paper No. 5. Institute of Library Research, University of California, Berkeley. August 18, 1969.

Barrett, Debbie and Don Sherman. FIX: A Language for Correcting MARC II Records. Technical Paper No. 6. Institute of Library Research, University of California, Berkeley. November 1, 1969.

Howe, Luke. COBOL Processing of UC MARC Tapes. Technical Paper No. 7. Institute of Library Research, University of California, Berkeley. November 7, 1969

Knapp, John. Sort Key Edit Program - SKED. Technical Paper No. 11. Institute of Library Research, University of California, Berkeley. January 13, 1970.

Ely, Margaret and Nancy Voigt. Documentation for Automatic Field Recognition Program. Technical Paper No. 12. Institute of Library Research, University of California, Berkeley. Revised May 18, 1970.

Barrett, Debbie. Documentation for FIX Language and Program. Technical Paper No. 13. Institute of Library Research, University of California, Berkeley. (in preparation).

Operations Task Force Project

Administrative.

Project Director: R. M. Shoffner

Sponsoring Agency: University of California

Period: 3/1/65 - continuing

Amount: Budgeted from year-to-year

Purpose. The libraries of the University of California have been faced with many urgent problems for which generally applicable solutions must be found as soon as possible. The Operations Task Force Project was established to provide a state-wide agent for development of methods compatible with the needs of each campus, but applicable university-wide. Examples of the problems with which the Operations Task Force Project has been concerned include: (1) coordination of campus-based programs for mechanization, (2) inter-campus circulation studies, (3) methods for producing supplements to the university-wide book catalog. A major effect has been devoted this year to development,

in cooperation with the librarians and library systems staff of the nine campuses, of a concerted program for library mechanization.

Reports.

Shoffner, R. M. Operations Task Force Project Status Report and Budget. Institute of Library Research, University of California, Berkeley. September 26, 1969.

Hayes, R. M. and R. M. Shoffner. Mechanization in the Libraries of the University of California. Institute of Library Research, University of California, Los Angeles. October 31, 1969.

Program Review Committee of the University of California Library Council. Library Systems Development Program. Institute of Library Research, University of California, Berkeley and Los Angeles. April 16, 1970.

Sherman, Don. Proposed University of California MARC Format. Technical Paper No. 1. Institute of Library Research, University of California, Berkeley. May 21, 1969.

Sherman, Don. University of California MARC Format. (Supersedes Technical Paper No. 1). Technical Paper No. 2. Institute of Library Research, University of California, Berkeley. June 15, 1969.

Buginas, Scott J. Information Retrieval Activities within the University of California. Technical Paper No. 3. Institute of Library Research, University of California, Berkeley. August 1, 1969.

EDUCATION PROJECTS

A major component of the program of the Institute of Library Research is the advancement of education in librarianship, information science, and documentation. In part, the extent of the commitment of the Institute to education is represented by the heavy use of students on projects as research assistants. They have come from business administration, from mathematics, from engineering, from English, and from other departments, as well as from the two library schools.

Although most Institute activities directly or indirectly provide knowledge or tools that advance education, the Institute has been responsible for a number of projects designed to meet an educational requirement.

An Information Processing Laboratory for Education and Research in Library Science

Administrative.

Principal Investigator: Professor M. E. Maron

Project Manager: A. J. Humphrey, D. Sherman (Beginning 1/1/70)

Associate Project Manager: J. C. Meredith

Sponsoring Agency: Office of Education

Period: 6/15/67 - 12/14/68 Phase I
12/15/68 - 7/5/70 Phase II

Amount: $141,763 Phase I
$200,000 Phase II

Purpose. It is the objective of this project to design, implement, operate, and evaluate a computer-based laboratory that is to assist in meeting the educational and research needs of students in librarianship. The laboratory functions in an on-line, interactive setting provided by a network of remote terminals (both mechanical and cathode ray tube) linked to a central digital computer. Software developed to date emphasizes the demonstration and teaching of formal methods of intellectual access to stored information in addition to use of the computer in instructing students in topics of traditional librarianship.

Phase I of the project emphasized planning and initial program development. The major Phase I activities included implementation of a Time Sharing Monitor, Associative Search System, and Computer Assisted Instruction courses in Subject Cataloging. Phase II of the project emphasized operational aspects of the Laboratory and integration with Library School curricula. During the 1969-1970 academic year over 200 students in the Berkeley School of Librarianship utilized the Laboratory's three video terminals. The Laboratory's inventory of programs includes: Reference Search, Associative Search, MARC II Retrieval, and a high level computer language for writing Computer Assisted Instruction modules.

A request for support for a third phase of the project has been submitted to the Office of Education.

Reports.

Maron, M. E. et al. An Information Processing Laboratory for Education and Research in Library Science: Phase II. (In preparation.)

Meredith, J. C. The CAI/Author/Instructor. Technical Paper No. 4. Institute of Library Research, University of California, Berkeley. July 10, 1969.

Meredith, J. C. and D. Ferguson. Student Feedback as a Tool in CAI Frame Development. Technical Paper No. 8. Institute of Library Research, University of California, Berkeley. November 20, 1969.

Cooper, William. A Definition of Relevance for Information Retrieval. Technical Paper No. 10. Institute of Library Research, University of California, Berkeley. December 12, 1969.

A Study of the Needs for Research in Library and Information Science Education (Continuing)

Administrative.

Project Director: Professor H. Borko

Sponsoring Agency: U. S. Office of Education

Period: 4/24/69 to 4/30/70 (extended to 10/31/70)

Amount: $54,109

Purpose. The purpose of this study is to determine the needs for research in education for librarianship, including consideration of library and information science. It was submitted to the Office of Education with the sponsorship of the American Association of Library Schools, as a means for establishing guidelines and recommended priorities for such research.

Reports.

Reilly, K. D. Author's Guide and Style Manual for the Study of the Needs for Research in Library and Information Science Education. Institute of Library Research, University of California, Los Angeles. June 15, 1969.

Borko, H. First Progress Report: A Study of the Needs for Research in Library and Information Science Education. Institute of Library Research, University of California, Los Angeles. July 24, 1969.

Borko, H. et al. Research Needs in Library and Information Science Education. (Draft). Institute of Library Research, University of California, Los Angeles. July 1970.

Education of Medical Library Systems Analysts (Continuing)

Administrative.

Project Director: Professor R. M. Hayes

Other Faculty: Professor H. Borko

Sponsoring Agency: National Library of Medicine

Period: 7/1/70 to 6/30/75

Amount: $520,641

Purpose. The increasing demand for greater information services to medicine has led the National Library of Medicine to undertake a number of revolutionary steps--mechanization of publication of Index Medicus, mechanized retrieval services through its MEDLARS system, sponsorship of the regional medical library, development of the Abridged Index Medicus (AIM-TWX) on-line bibliographic retrieval system, and planning for development of a national medical information system. The result is a great and unmet demand for personnel who combine knowledge of librarianship with expertise in systems analysis and application of computers. The purpose of this project is to educate students with this combination of knowledge. The degree program for Master of Science in Information Science (Documentation) provides the general framework. It is supplemented by internships (as research assistants) in the Institute of Library Research.

From the education and related Institute projects have come a number of computer programs, program systems, and special data bases developed and used to further student expertise. Examples include: the DISCUS project that provides computer assisted instruction (CAI) and other on-line capabilities; a cost accounting on-line interactive program that uses gross library system parameters to develop operating cost estimates for both manual and (partially) automated operating environments; a collection of experimental bibliographic ERIC data

base processing programs that can create, list, or search an inverted index or search a sequential file using Boolean descriptors or accession numbers; and an on-line format manipulation system (FMS) that permits text to be entered, modified, deleted, or corrected at a remote keyboard/display console.

Reports.

de Boer, Aeint H. A Modular Program for Reference Retrieval from Bibliographic Data Bases in a University Research Library. Masters Thesis. Institute of Library Research, University of California, Los Angeles. 1970.

Borko, H. and R. M. Hayes. Education for Information Science: Documentation. (Report Number 3). Institute of Library Research, University of California, Los Angeles. February, 1970.

Reilly, Kevin. Education in Information Science: The Applied Disciplines in a Systems-Oriented Curriculum. (Report No. 4). Institute of Library Research, University of California, Los Angeles. June 30, 1970.

PUBLIC SERVICE

The Institute of Library Research has continuously serviced the public sector through projects designed to serve libraries of the State of California, through discussions with people from throughout the world, and through extensive participation of its staff in professional activities. Many visitors, foreign and domestic, found their way to the Institute during the past year for discussions or study. Institute staff members, themselves, visited many locations and gave many talks to diverse audiences. A continuing effort exists to acquire

projects whose products can be characterized as having demonstrable utility.

California State Library Processing Center (Completed during year)

Administrative.

Principal Investigator: Dean R. C. Swank

Project Director: R. M. Shoffner

Project Manager: D. Sherman

Sponsoring Agency: State Library of the State of California

Period: 6/21/67 to 7/31/70

Purpose. To develop detailed specifications for a computer based Processing Center to serve a network of public libraries throughout California. The specifications include system design, program layout, cost analysis and plan for staffing, software development, and initial production. Conversion of retrospective holdings and generation of book form catalogs from the nucleus of the system.

Reports.

Sherman, D. and R. M. Shoffner. California State Library: Processing Center Design and Specifications (3 volumes). Institute of Library Research, University of California, Berkeley. April, 1969.

Dunham, R. and M. Fried. California State Library: Processing Center Design and Specifications: Serials (volume 4). Institute of Library Research, University of California, Berkeley. September 1969.

Hargrove, T. L. and Keith Stirling. California State Library: Processing Center Design and Specifications: Cost Analysis (volume 5). Institute of Library Research, University of California, Berkeley. July 1970.

Effectiveness Criteria for Medical Libraries (Completed during year)

Administrative.

Project Director: Professor G. Edward Evans

Other Faculty: Professor H. Borko

Sponsoring Agency: National Library of Medicine

Period: 8/1/69 to 1/31/70 (Extended to 3/31/70)

Amount: $6,974.

Purpose. Phase I of this study of criteria relates to the effectiveness of medical library service and provides an evaluation of existing criteria, as based on a review and analysis of results as reported in earlier studies. A specific example, that of "Document Delivery Test", (DDT) as reported in the University City Science Center's Final Report-National Survey of Medical School Libraries, was closely examined. Data relating to DDT was analyzed for its reliability and correlation with other criteria.

Reports.

Evans, G. E. Effectiveness Criteria for Medical Libraries--Phase I. Institute of Library Research, University of California, Los Angeles. May 28, 1969.

Evans, G. E. and Harold Borko. Effectiveness Criteria for Medical Libraries. (Final Report). Institute of Library Research, University of California, Los Angeles. April, 1970.

Physical Facilities Research Project (Completed during year)

Administrative.

Project Director: Professor G. Edward Evans

Other Faculty: Professor R. M. Hayes
Professor Peter Kamnitzer
Professor Thomas Vreeland

Sponsoring Agency: Federal Library Committee

Period: 10/1/70 - 9/30/71

Amount: $2,450

Purpose. To develop a research design for development of techniques for design of physical facilities for Federal Libraries in general, including an investigation of current practices concerning the design of physical facilities for libraries and the methods of equipping and furnishing these facilities.

Reports.

Evans, G. E., Scheele, S. and Vreeland, T. Methods of Library Building Design--A Research Plan. Institute of Library Research, University of California, Los Angeles. June 30, 1969.

Evans, G. E. An Addendum to Methods of Library Building Design--A Research Plan. Institute of Library Research, University of California, Los Angeles. October 22, 1969.

APPENDIX

COMPUTER PROGRAMS

A number of computer programs have been developed and used by Institute personnel. Some programs of broad interest have been incorporated into campus computing networks and are so utilized. Examples would include the Interactive Assembler Language Interpreter (INTX) program and the Format Manipulation System for Automatic Production of Natural Language Documents (FMS) program, both of which operate under control of the URSA on-line user system provided by the Campus Computing Network (CCN) on the University of California Los Angeles campus.

A number of programs have been developed for special research and test purposes related to the Center for Information Services project at Los Angeles, and the University of California Union Catalog Supplement project at Berkeley. Many of these programs are relatively simple, both as to purpose and complexity, and most of them are already on file as debugged, operable versions. Certain features of the Berkeley programs, as noted in the following descriptions, are being revised or developed; certain Los Angeles programs may be revised when incorporated in the information processing system being developed for the Center for Information Services project.

COMPUTER PROGRAMS - LOS ANGELES

LA-1. PROGRAM: INTX

AUTHOR: STEVE SILVER

STATUS: IN DEVELOPMENT

DESCRIPTION: INTX is a system for assembling, loading, and interactively interpreting 360 assembly language source programs under the UCLA URSA time-sharing system. It permits URSA customers to produce completely interactive programs without any risk of system damage--a feature that is not currently guaranteed by true URSA processors. The system is fully operational now but minor tuning and development will continue.

LA-2. PROGRAM: FMS

AUTHOR: STEVE SILVER

STATUS: COMPLETE

DESCRIPTION: FMS is designed for high speed, cheap generation of natural language documents for mixed text and command input. It is very useful for producing final copies of technical reports requiring extensive minor alternation before publication. The system needs some work before the current version can be released as safe.

LA-3. PROGRAM: DISCUS

AUTHOR: STEVE SILVER

STATUS: IN DEVELOPMENT

DESCRIPTION: DISCUS is a computer aided instruction system oriented towards CRT graphic terminals. It is a two part system: a compiler that runs on any 360 and an executer whose I/O package must be tailored to the individual time-sharing system. It is currently under development but is expected to be completed very shortly.

LA-4. PROGRAM: FAMULUS

AUTHOR: JERRY PINE

STATUS: COMPLETE

DESCRIPTION: FAMULUS is designed to process personal reference collections maintained by researchers. However, its basic structure renders it suitable for a large number of other applications. For this purpose it can be regarded as a general-purpose text-handling system.

FAMULUS will maintain many types of information files which can be broken into units or records with sub-categories or fields which can be identified. In a personnel information file, the data on each person comprises one record. The record may have up to 10 distinct fields in which are entered name, date of birth, job title, etc. In bibliographic files, the citation is the record, and fields are used for author, title, date, keywords, abstracts, etc.

LA-5. PROGRAM: WORDLST

AUTHOR: LINDA MIROFF

STATUS: COMPLETE

DESCRIPTION: WORDLST is a PL/I subroutine which breaks a given character string into individual words. The user specifies characters which are to act as word delimiters; these characters should not be embedded within words. WORDLST stores the individual words without any delimiters in an array. A number giving the relative position of the word in the string is stored in a parallel array. In this way other routines have access to the words as they appear within the string. A list of words, called an exclusion list, may be used to exclude non-content words. Words less than a given length may also be excluded.

LA-6. PROGRAM: CONTEXT

AUTHOR: LINDA MIROFF

STATUS: COMPLETE

DESCRIPTION: CONTEXT is a PL/I subroutine which does most of the "work" involved in building a KWIC index. It finds the contexts of given words in a given string. It can be used in conjunction with an input routine, the WORDLST routine, a sort routine, and an input routine, to produce a KWIC index.

The context of a word consists of as much of the surrounding sentence as can be fitted in the space allowed. The number of characters in the context line is set by the user. When all of the sentence cannot be fitted in, parts of the sentence furthest from the word are truncated. The word always appears in the middle of the line. Wrap-around (from left to right or from right to left) is performed to make use of extra room on either side of the word. When there is a space a "-" indicates the beginning of the context; a "+" is used at the end of the context to indicate that part of the sentence is missing.

LA-7. PROGRAM: CONTITL

AUTHOR: LINDA MIROFF

STATUS: COMPLETE

DESCRIPTION: CONTITL is a special version of CONTEXT, to be used where the text always consists of titles or single sentences. CONTITL produces KWIC records in the same manner as CONTEXT but does not recognize sentence endings or beginnings. The output produced is like that produced by CONTEXT.

LA-8. PROGRAM: DEMO1

AUTHOR: GEORGETTE SILVA

STATUS: COMPLETE

DESCRIPTION: DEMO1 reads text from magnetic tape, breaks it into individual words, sorts them alphabetically and in order of decreasing frequency, and lastly, prints each word surrounded by context in order of occurrence in the text. All these tasks are carried out in internal memory. The program performs these tasks on a fairly methodological level. The sort, for example, is simply what is termed a "bubble sort". The text must be limited to 10,000 characters. The latter condition is not serious, however, since this is more than adequate to cope with short texts such as abstracts. The main virtue of this program is that it consists of one deck, and carries out all the tasks in one run. It may, therefore, be useful as a demonstration program for beginners.

LA-9. PROGRAM: TEXTMT

AUTHOR: GEORGETTE SILVA

STATUS: COMPLETE

DESCRIPTION: TEXTMT is somewhat more sophisticated than DEMO1 and has a few additional capabilities: it does character conversion by table lookup, retrieves index terms in context, and uses a binary tree sort for the dictionary and frequency sort. These operations are carried out in internal memory, and the amount of text processed at any given time is still limited.

LA-10. PROGRAM: INDX

AUTHOR: GEORGETTE SILVA

STATUS: COMPLETE

DESCRIPTION: The INDEX program is designed to cope with texts in prose, poetry, and dramatic forms, as well as with transcriptions of oral discourse. The program reads input text from cards and produces records each containing one textual word or punctuation mark, and index information showing where the word/punctuation mark occurred in the text. Each word/punctuation mark is indexed as to volume, chapter, paragraph, sentence, and word within sentence, and word within sentence location. These records are stored in a temporary file which forms the input to program GAMMA.

LA-11. PROGRAM: GAMMA

AUTHOR: GEORGETTE SILVA

STATUS: COMPLETE

DESCRIPTION: Program GAMMA produces a word index and word frequency list from the output of INDX. The word index is in alphabetical order by word and includes frequency count. The word frequency list is ordered alphabetically within descending frequency. GAMMA invokes the IBM sort/merge to do the necessary sorting and prints the final output.

LA-12. PROGRAM: CONCORD

AUTHOR: GEORGETTE SILVA

STATUS: COMPLETE

DESCRIPTION: CONCORD reads input text from cards in stream form and produces concordance records (in KWIC format) on each word in the text which is not in an exclusion list. The records are written in a sequential data set which forms the input to program ALPHA.

LA-13. PROGRAM: ALPHA

AUTHOR: GEORGETTE SILVA

STATUS: COMPLETE

DESCRIPTION: ALPHA sorts the output records produced by CONCORD into ascending sequence by word and identification. It invokes the IBM sort/merge and prints the concordance, underlining the keyword of each line of text.

LA-14. PROGRAM: ASPEN (California Code and State Constitution) Read Routine

AUTHOR: GEORGETTE SILVA

STATUS: COMPLETE

DESCRIPTION: This routine reads magnetic tape records of the Aspen file (containing full-text images of 298 documents from the California Code and State Constitution), identifies each field within a record, and provides a detailed printout for each record type.

LA-15. PROGRAM: COMPENDEX Read Routine

AUTHOR: GEORGETTE SILVA

STATUS: COMPLETE

DESCRIPTION: This routine reads magnetic tape records of the Computerized Engineering Index file (containing bibliographic information), identifies each field within a record, and provides a detailed printout for each record type.

LA-16. PROGRAM: CHEMICAL ABSTRACTS CONDENSATE Read Routine

AUTHOR: GEORGETTE SILVA

STATUS: COMPLETE

DESCRIPTION: This routine reads magnetic tape records of the Chemical Abstracts Condensate file (containing bibliographic information), identifies each field within a record, and provides a detailed printout for each record type.

LA-17. PROGRAM: COMMUNICATIONS OF BEHAVIORAL BIOLOGY (CBB) Read Routine

AUTHOR: GEORGETTE SILVA

STATUS: COMPLETE

DESCRIPTION: This routine reads magnetic tape records of the Communications of Behavioral Biology file (containing bibliographic information), identifies each field within a record, and provides a detailed printout for each record type.

LA-18. PROGRAM: NUCLEAR SCIENCE ABSTRACTS (NSA) ENTRY Read Routine

AUTHOR: GEORGETTE SILVA

STATUS: COMPLETE

DESCRIPTION: This routine reads magnetic tape records of the Nuclear Science Abstracts Entry file (containing bibliographic information), identifies each field within a record, and provides a detailed printout for each record type.

LA-19. PROGRAM: NUCLEAR SCIENCE ABSTRACTS (NSA) SELECTOR Read Routine

AUTHOR: GEORGETTE SILVA

STATUS: COMPLETE

DESCRIPTION: This routine reads magnetic tape records of the Nuclear Science Abstracts Selector file (containing condensed bibliographic information), identifies each field within a record, and provides a detailed printout for each record type.

LA-20. PROGRAM: BINCDF

AUTHOR: STUART BEAL

STATUS: COMPLETE

DESCRIPTION: A FORTRAN function for computing the binomial cumulative distribution.

LA-21. PROGRAM: PARAMETER P TEST

AUTHOR: STUART BEAL

DESCRIPTION: A FORTRAN program which tests the P parameter of the binomial distribution.

LA-22. PROGRAM: INTRPRT

AUTHOR: AEINT DE BOER

STATUS: COMPLETE

DESCRIPTION: The INTRPRT System is composed of a language, an assembler, an URSA subselector, an URSA interpreter, and a few test programs.

a) INTERASM--This is an assembler language batch program to assemble programs written in the INTRPRT language. It is a small, fast, fairly conventional two pass assembler. Assembled code is written directly as a partitioned data set ready for interpretation.

b) INTRPRT--This is an URSA processor functioning as a subselector for the INTRPRT interpreter. It displays the members of a PDS (created by INTERASM) as programs available for execution and allows one to be selected for execution.

c) INTRPRT--(Written by Ron Reuben, included for completeness)--This URSA processor is the INTRPRT interpreter. It loads the program selected via INTRPRT and executes it interpretively. It includes good error messages and good facilities for debugging. It currently needs maintenance to conform to an incompatible change in URSA.

d) STUDENT--This is a sample program written in the INTRPRT language. It searches a file of student records by student name or student number and displays selected records.

e) FACULTY--This is a sample program written in the INTRPRT language. It searches a file of faculty biobibliography records by a combination of up to nine control fields and displays selected portions of selected records.

LA-23. PROGRAM: ERIC

AUTHOR: AEINT DE BOER

STATUS: COMPLETE

DESCRIPTION: The ERIC file search system will search, maintain and provide listings from the ERIC file.

a) MULFSCH--This is a batch program to search an inverted index to an ERIC file or a sequential ERIC file by a Boolean combination of descriptors or a list of accession numbers. It is designed to be extended to search other files.

b) ERICMIF--This is a batch program to create an inverted index to an ERIC file. It reads a sequential ERIC file, extracts the descriptors and accession number, and writes a direct access file of descriptors with corresponding accession numbers. This file can be used by MULFSCH, ERICIFL, and ERICDFL.

c) ERICIFL--This is a batch program to list an inverted index to an ERIC file. The listing includes descriptors and frequency of use and can be either one up or two up. The one up listing can include the corresponding accession numbers as an option. The listing is ascending by accession number within descriptor.

d) ERICDFL--This is a batch program to list an inverted index to an ERIC file. The listing includes descriptors and frequency of use and is ordered by descending frequency of use.

e) ERICMLF--This is a batch program to copy a sequential ERIC file deleting the abstracts.

LA-24. PROGRAM: LIMP/360

AUTHOR: AEINT DE BOER

STATUS: COMPLETE

DESCRIPTION: LIMP/360 is a batch program to solve linear programming problems. It is particularly useful to solving batches of small LP problems, e.g., class problems. It is intended to complement MPS/360.

LA-25. PROGRAM: CENSGOV

AUTHOR: AEINT DE BOER

STATUS: COMPLETE

DESCRIPTION: CENSGOV is a batch program to verify a set of documented universes for the 1960 1/1000 sample census tape.

LA-26. PROGRAM: UTILITY PROGRAMS

AUTHOR: AEINT DE BOER

STATUS: COMPLETE

DESCRIPTION:

a) SPM is a short assembler language subroutine to enable a FORTRAN programmer to Set the Program Mask portion of the PSW.

b) STIMER (a modification of a routine borrowed from Stuart Beal) is a short assembler language subroutine to enable a PL/I programmer to measure elapsed task CPU time and elapsed real time.

c) PGMPRNT is a batch program to print programs in a format suitable for inclusion in a thesis. Includes provisions for a figure title, figure number, number of parts, statement numbers and page numbers.

COMPUTER PROGRAMS - BERKELEY

B-1. PROGRAM: PRE-AFR

TYPE: INPUT CONVERSION

STATUS: COMPLETE

DESCRIPTION: This program acts as an interface program between any input device and the Automatic Field Recognition program. It performs limited validity checking and tranlates the entire character stream of input records to EBCDIC.

B-2. PROGRAM: AFR

TYPE: INPUT CONVERSION

STATUS: COMPLETE

DESCRIPTION: This program now takes catalog records keyboarded with a minimum of paragraph indention signals and constructs a complete MARC II structure record, with a limited amount of content identification.

B-3. PROGRAM: INVENTORY

TYPE: INPUT CONVERSION

STATUS: COMPLETE

DESCRIPTION: This is a general-purpose housekeeping routine which will check on missing and duplicate master record numbers, update a file of control records describing input batches according to processing status, extract a random sample of MARC input records if desired for proofing, and print records in which serious defects are detected by the AFR algorithms.

B-4. PROGRAM: FIX

TYPE: INPUT CONVERSION

STATUS: COMPLETE

DESCRIPTION: This is a general-purpose MARC editing/correction tool. It can be manually or programmatically operated; has capability at present to update a record structurally down to the indicator and subfield delimiter level. It can incorporate any content change to a record from one character to build a whole record or replace a whole record. Other forms of record maintenance are planned.

B-5. PROGRAM: FILOR

TYPE: FILE MAINTENANCE

STATUS: IN DEVELOPMENT

DESCRIPTION: This program presently creates a direct access master file on disk. Update capability is under development.

B-6. PROGRAM: ZODIAC

TYPE: FILE MAINTENANCE

STATUS: COMPLETE

DESCRIPTION: This program generates fixed length records containing index keys representing MARC II tagged

fields, for use in building inverted files for on-line search.

B-7. PROGRAM: PAX

TYPE: FILE MAINTENANCE

STATUS: IN DEVELOPMENT

DESCRIPTION: From the records output by ZODIAC, duplicates are collapsed into single ISAM keyed records, and a multi-level linkage structure with the direct access master file is established. An update facility is under development. These programs are expected to have utility for authority file operations in either a batch or on-line environment.

B-8. PROGRAM: CIMARON

TYPE: FILE SEARCH

STATUS: COMPLETE

DESCRIPTION: This program provides on-line, conversational facilities to search inverted files, using partial key matching and combinations of Boolean operators to connect a multi-key input expression. Display is currently implemented on ILR CRT terminals with keyboard interaction only.

B-9. PROGRAM: SKED

TYPE: OUTPUT PREPARATION

STATUS: COMPLETE

DESCRIPTION: This program prepares MARC records for sorting by extracting and editing to a limited extent a 256-character sort key which is then prefixed to the master record. A high degree of user control is provided on the format and content of the sort key.

B-10. PROGRAM: BIBLIST

TYPE: OUTPUT PREPARATION

STATUS: COMPLETE

DESCRIPTION: This program prepares sorted MARC records for display by line printer and will ultimately service the requirements of computer-based graphic arts quality output. It has two parts, a column-formatting generation routine and a page-formatting routine. Again, a high degree of user control is provided.

B-11. PROGRAM: CONVTAPE

TYPE: UTILITY

STATUS: COMPLETE

DESCRIPTION: This program translates MARC records from an arbitrary character set to EBCDIC and vice versa. (e.g., to and from ASCII.) It also performs re-blocking as specified by user.

B-12. PROGRAM: CALLSORT

TYPE: UTILITY

STATUS: COMPLETE

DESCRIPTION: Extracts call numbers from MARC 090 and formats them for use in a sort key. Interfaced with SKED to create a shelflist file.

B-13. PROGRAM: FILEND

TYPE: UTILITY

STATUS: COMPLETE

DESCRIPTION: Program to print and/or core dump the first and/or last block of data in a large file, either tape or disk, at user option.