DOCUMENT RESUME

ED 026 330 24 SP 002 179

By-Coffin, Robert

Appendix Y. The Integrated Communications Experiment (ICE) Summary.

Northwest Regional Educational Lab., Portland, Oreg.
Spons Agency-Office of Education (DHEW), Washington, D.C. Bureau of Research.
Bureau No-BR-8-9022
Pub Date Oct 68
Contract-OEC-0-8-089022(010)
Note-13p.; Appendix Y in A Competency Based, Field Centered, Systems Approach to Elementary Teacher Education; Final Report. Vol. III, p231-243.
EDRS Price MF-$0.25 HC-$0.75
Descriptors-*Computer Oriented Programs, *Electronic Data Processing, Information Retrieval, Information Storage, *Information Systems, Program Descriptions, Programing, Recordkeeping, *Teacher Education
Identifiers-*ComField Model Teacher Education Program

This appendix describes the Integrated Communications Experiment (ICE), a comprehensive computer software capability developed for the ComField Project. Each major characteristic of the data processing system is treated separately: natural language processing, flexibility, noninterference with the educational process, multipurposeness, comprehensive retrieval, centralization of records and distribution, statistical analysis, long-range data management, editing capacity, automatic coding and uncoding, data accuracy, uniform data preparation, and computer independence. Included also are descriptions of the three major programing systems--file generation and translation, file management and editing, and file retrieval--which are considered a minimum set of software functions which would satisfy the total demand of a record system in the natural language. Two flow charts are appended: (1) schematic configuration showing one possible ComField system using a comprehensive storage/retrieval system such as ICE and (2) the ICE system including document description, translation, information retrieval, and file maintenance and editing. This document and SP 002 155-SP 002 180 comprise the appendixes for the ComField Model Teacher Education Program Specifications in SP 002 154. (JS)

APPENDIX Y--THE INTEGRATED COMMUNICATIONS EXPERIMENT (ICE) SUMMARY

Submitted for a Consortium of
Institutions and Agencies by the
Northwest Regional Educational Laboratory
400 Lindsay Building
710 S. W. Second Avenue
Portland, Oregon 97204

THE INTEGRATED COMMUNICATIONS EXPERIMENT (ICE) SUMMARY

Robert Coffin

In the majority of data processing facilities, each new application necessitates the investment of considerable time and expense before the originator of a project is able to benefit from the collective effort of professional computer personnel. Too often, this expense and delay has resulted in disillusion and disappointment on the part of educators who have little appreciation for the technical implications of demands which they consider perfectly reasonable from their own points of view. Traditional problems such as grade and attendance reporting, payroll accounting and other administrative functions have won general acceptance largely because the benefits are tangible and easily justified in terms of costs. However, these programs were hopelessly inadequate for processing information pertaining to the more intricate and critical problems of constructive evaluation of student performance, maintaining libraries of curriculum alternatives for individualized instruction, and analysis of subjective student data as collected by teachers, counselors, and special service workers. These problems, diverse in nature and continuous in time, had to be bound together in a common approach in such a manner that problems and situations, even though unanticipated, could be accommodated without large, additional expenditures of time and money. The landmark decision to develop a comprehensive computer software capability was made in the Department of Educational Research in 1965; and under financial support of the Hill Family Foundation of St. Paul, the Integrated Communications Experiment (ICE) was designed and started in the spring of 1966. The conceptual model for the ICE system existed as successively operating medical record system in the Portland area*, consequently the designers of the ICE system benefited from a huge investment in a generalized information storage and retrieval system developed in a related environment.

An underlying promise of the ICE system was that the filing system, from data collection to processing and final analysis

---

*The fundamental concept herein presented is based upon work conducted at the Oregon Regional Primate Research Center upon Contract to the National Institution of Health, Institute of Child Health and Development. (Contract PH43-64-1163).

had to be implemented so that the user, teacher, supervisor, or researcher could continue to record information in his own language. This decision was forced by the simple expedient that to gain the acceptance and cooperation of the user, it is not possible to require him, except in rare instances, to code information or adopt any recording method which is fundamentally alien to their traditional habits. In general, the language of the user dictates the form of the information input whether entirely verbal, entirely quantitative, or a mixture of quantitative and verbal data. To this objective, the ICE system allows the user to freely define his document to suit his own subject area. An easy-to-learn notation is used to describe a document for machine processing. An example of a typical input document is the Remedial Reading Record shown on Page "5". The fixed information, called the Header, serves to identify the record in the master file. The remainder of the record, the variant, is arrayed in outline form to reflect the breadth of subject matter of intent to a particular user in a specialized application. From the view of programming economy, it is important to note that any document arrayed in a similar manner may be accommodated in the ICE system without further programming effort. At the present time, several dozen information documents have been designed and are in use in the Portland Public School System.

A comprehensive retrieval system is currently being implemented to provide the user with a variety of types of output. User requests tend to fall into three broad categories. The most frequent requests are for sets of raw records which have a specific set of characteristics and serve to answer questions of transient importance -- raw records may be reduced to take the form of reports which cover particular cross-sections of the master file. A further refinement is statistical analysis which may be obtained by linkage of the ICE master file to existing statistical programs.

Important auxiliary functions offered by the ICE system aid in overcoming many of the disadvantages of variable content, variable length record keeping. Capacity for editing of misspellings has provided a means for improving the quality of text, removal of ambiguities and standardization of descriptive entries. Other editing and file maintenance routines give the computer personnel control over the data files in order to insure that data is recorded and reproduced with maximum accuracy.

The following characteristics of the ICE system are presented as the minimum conditions to satisfy the total system requirements for ComField.

Flexibility

Any records system which will withstand the tremendous pressure of time must have a capacity to change through time. Inherent in this statement is the fact that a record system can be neither rigid in subject matter that it accommodates nor limited in the diversity of that subject matter. The user must have the ability to vary the content of records as the need dictates. Further, there may be several kinds of information contained in the same file. In rigidly defined data collection systems, it is possible to define an entry on a particular card which is assumed to have certain limits and it is not likely to change. However, in the educational environment there are very few appropriate subjects in which rigidly defined data collection systems are meaningful. Experience has shown that information changes dynamically, thereby making any rigid planning obsolete in a very short period of time.

Economic considerations of the rigid versus the flexible system are very clear. Natural Language Processing permits more dynamic expression and communication of information within practical economic limits. In the educational environment it is not possible nor desirable to have a large staff of highly salaried, skilled programmers in order to facilitate continual change. If a general record system could be implemented which would allow a wide variation in content and complete capacity to growth through time, and if such a system could be implemented and turned over to a small data processing group for operation, then the scope of an automated records system becomes quite feasible.

Noninterference with the "Educational" Process

It is not realistic to assume that data processing personnel can impose a drastic change upon educational personnel without having a great deal of friction develop which would lead to a state of noncooperation and decline of the records effort. Perhaps the most important feature offered by Natural Language Processing is that of reducing the problem of interfering with the "educational process." This is to say, it is possible merely to substitute one piece of paper for another in order to implement the data collection scheme. The teacher, special personnel, or administrator need not undergo a lengthy process of orientation, familiarize himself with a particular coding scheme, or write in any manner which is foreign to him. Data processing may then become a harmonious part of the long established educational system. Natural Language Processing can accomplish this end.

Multipurposeness

It has been considered impractical to use information only once. If information is in the natural language of its user, a multiplicity of use can be easily achieved. Further, it has been considered impractical to maintain separate record systems for separate purposes within small institutions. A record system must serve groups of diverse interests in a unified manner which does not interfere with the normal processes of each group. Multipurpose use of records is mandatory in the educational environment. Natural Language Processing offers a unified approach to record keeping and at the same time can serve many different people.

Natural Language Processing does not exclude the use of coding schemes since any coded system is a subset of the natural language. Coding schemes are encouraged whenever the user feels they will provide more effective use of information. Natural Language Processing provides the user in the educational environment flexible, multipurpose storage and communication of information with the least interference with the "educational" process. There are twelve principal features that should be considered in the programming structure of a natural language system. These features are viewed as desirable for any filing system.

Flexibility

The basic programming structure must contain the ability to translate and process variable length, variable content records in order to provide the system with flexibility. The integrity (content, length) of a record must be maintained throughout all processing. The programming system must be able to accept new subject areas that were not anticipated in the initial systems design. Change cannot be predicted, so the ability to change must be an integral part of the system.

Compatibility with the "Educational" Process

In order to gain general acceptance, the programming system should preserve the integrity of existing source documents and reports. The most important single feature that can be offered in terms of gaining acceptance of the system is simply the fact that it is possible to substitute one piece of paper, which is much like the one it replaces, for another. Further, there must be the ability to design a source document which can accommodate the material which is to be recorded without introducing any bias

into the information. The amount or detail of information should be left up to the user.

Comprehensive Retrieval System

In order for a records system to serve many different groups of personnel, it is necessary to have a broad and comprehensive retrieval system which anticipates general requirements of each group. Just as it is not possible to anticipate the breadth nor scope of the information content, neither is it possible to anticipate the breadth and scope of retrieval functions which will be required of the system. Therefore, a very generalized and flexible retrieving system is required to handle the many different demands for information.

Centralization of Records

The system must provide for centralization of record keeping in a smooth flowing data collection and processing scheme which allows for management of bottlenecks, data flow, and currency of subject information. Centralization of records is desirable, provided that it can be achieved within practical limits. The system eliminates many limiting factors characteristic of hand systems, but does not replace complete systems planning. Most often records are decentralized because it is too cumbersome to generate and maintain a unified, centralized system.

Centralized Distribution

A common complaint of centralized record systems is the inability for adequate distribution of records. In centralized hand record systems, records are continually being handled causing wear and jeopardizing the quality of the file. The availability of records is never certain and one must put forth special effort to obtain them. In a computerized system, the source documents s are never removed for purposes of distribution; consequently, the user is assured of continued high speed availability.

Statistical Analysis

From the view of the researcher, an individual record may be of particular interest, but more broadly he is interested in populations of records which share certain features. It is therefore

necessary to supply to the user the ability to extract in meaningful arrays the basic documents in a form suitable for statistical analysis.

Long-Range Data Management

The initial system planning must include provision for records which have accumulated through time and which are of potential value, but which must be considered inactive. This, plus other data management considerations (sort, merge) must be provided for by the basic programming system.

Editing Capacity

To a large degree the success of a filing system will depend on the ease and naturalness with which errors can be corrected. An editing system should be incorporated to first identify and categorize all types of errors, and secondly to provide a convenient and natural mechanism to edit these errors. It is necessary to maintain a constant vigilance on the quality of material that comes into the files. Provisions should be made to identify not only misspellings, but also ambiguities and word equivalences.

Automatic Coding and Uncoding

The capacity to perform generalized editing has a secondary benefit which enables an investigator to arrive at a satisfactory coding structure for his own information after he has had some experience in collecting and recording his data. It is not necessary for the user, who may be naive in terms of machine analysis and data collection, to decide upon a coding structure. Rather, it is possible for him to collect data in his own natural stride and at a later time identify a possible structure which will enable him to rapidly and concisely analyze his information.

Data Accuracy

Basic to data accuracy is a fundamental concept that one and only one recording of the source information can take place. In many systems it has been necessary to rerecord and recopy by hand various portions of recorded information before it is prepared for the machine. This results in an unacceptably high error rate and corresponding low data quality. The programming system must preserve absolute accuracy of data.

Uniform Data Preparation

In order to reduce error in data preparation, the data collection sheets should be structured in such a way as to maintain consistent appearance to the personnel preparing the data, and once learned, a new subject area with its corresponding data sheet need not be a separate object of instruction. Without this uniformity the preparation is difficult, tedious, and increases the probability of error.

Computer Independence

Natural Language Processing as a philosophy of record keeping is absolutely machine independent; that is, the system planners may consider the scope and the nature of the problem relative to any suitable computer configuration. Of course, at the time it becomes an operating system, the philosophy and system features must be reduced to a set of machine instructions which then become machine dependent. If the implementation is consistent with the philosophy of machine independence, the user can expect to benefit from vertical and horizontal machine communication.

The following list of major programming systems and their brief description are viewed as a minimum set of software functions which would satisfy the total demand of a records system in the natural language.

File Generation and Translation System

This set of programs allows for the initial translation of records into the master file. During the translation process each record is subjected to a series of tests which detect gross errors by comparing the incoming record with an on-line dictionary which contains an image of each type of source document. Erroneous records are not merged with the master file, but are printed together with a diagnostic statement for subsequent correction and reprocessing. The system also must provide capacity for the definition of new source document images, and modification of existing images.

Each record must have an identifying or header portion which determines its unique position in the master file. One scheme might be to structure each individual's file as a chronology. While the header portion is fixed, the other portion is variable. The translator must perform interpretive operations on the

variable portion of the record. The interpretation of a record is controlled by the dictionary image which gives the translator its flexibility. The translator can be designed to accept any type of input media (cards, paper tape) and it is the purpose of the translator to check any error considerations characteristic of each media.

File Management and Editing System

Once a file has been created there is a problem of maintenance. The life and value of a file may depend on a flexible management system which must have the ability to delete records from a file, to dynamically combine files and to condense inactive records for economical storage. It is further desirable to perform editing functions within a file. The editing system provides the ability to edit errors and examine text. As mentioned above, spelling errors and inconsistent language detracts from the value of information. This programming system will improve quality and longevity of files.

File Retrieval System

This complex set of programs must provide the user with adequate communication to the master file to effectively identify and generate information. The retrieval system functions like a compiler in that it interprets a retrieval statement and compiles a program which screens each record as it is received into memory. A retrieval statement must be checked for validity against the dictionary tape. Basically there are three classes of retrievals. Retrieval of original records will occur most frequently in which the user must have the ability to retrieve all, or part, of an individual's file. Secondly, there is a need to create subfiles which will be used in connection with report generation programs. The report generated by these programs can be very helpful in analyzing the information contained in the master file. It is also possible to freeze the status of the master file by creating subfiles. Finally, it is desirable to link the master file with other programming systems which deal with fixed length records such as FORTRAN or COBOL, and thereby gives the user the ability to link the master file with any existing programs.

Other programming systems can be added to make effective use of the master file. If the master file is contained on a random access storage device, it is desirable to have an automatic indexing system. Further, it is highly desirable to have a textual

analysis system which analyzes content and syntax of word groupings.

In summary, Natural Language Processing has several advantages which make it desirable for use in the educational environment. Only through extensive systems considerations and planning will the educator realize any benefit. The programming systems provide an effective tool. It is the responsibility of the educator to develop the content and skill necessary to use the tool.

PORTLAND PUBLIC SCHOOLS
Department of Special Education
REMEDIAL READING RECORD CARD

Student Identification	Sex	Form No.	Mo.	Day	Year	Time (24 Hrs.)	Edit Code	Card No.	Tag
		0 1 0 C 1 4					0 1 0		A

A

1. Name (Last First) 2. Birthdate (/ /)

3. School ( ) 4. Grade ( )

B

1. Date Entered (/ /) 2. Times/Week ( ) 3. Length of Period (Min.) 4. Date Dismissed (/ /) 5. Days Attended ( )

C

Mental Tests	a. Form	b. Date	c. I.Q.	d. Verbal	e. Perf.	f. Tester
1. ( )	( )	( )	( )	( )	( )	( )
2. ( )	( )	( )	( )	( )	( )	( )

D

Reading Tests	a. Form	b. Date	c. C.A.	d. V.	e. C.	f. R.G.	g. M.A.	h. R.A.	i. Ret.	j. Gain
1. Preliminary	( )	(/ /)	( )	( )	( )	( )	( )	( )	( )	( )
2. Progress	( )	(/ /)	( )	( )	( )	( )	( )	( )	( )	( )
3. Final	( )	(/ /)	( )	( )	( )	( )	( )	( )	( )	( )

E Recommendation
1. Dismiss
2. Retain
3. Reevaluation

F Reason for Dismissal
1. Condition Corrected
2. Improved
3. Moved
4. Graduated
5. Other ( )

G. Pertinent Information (

)

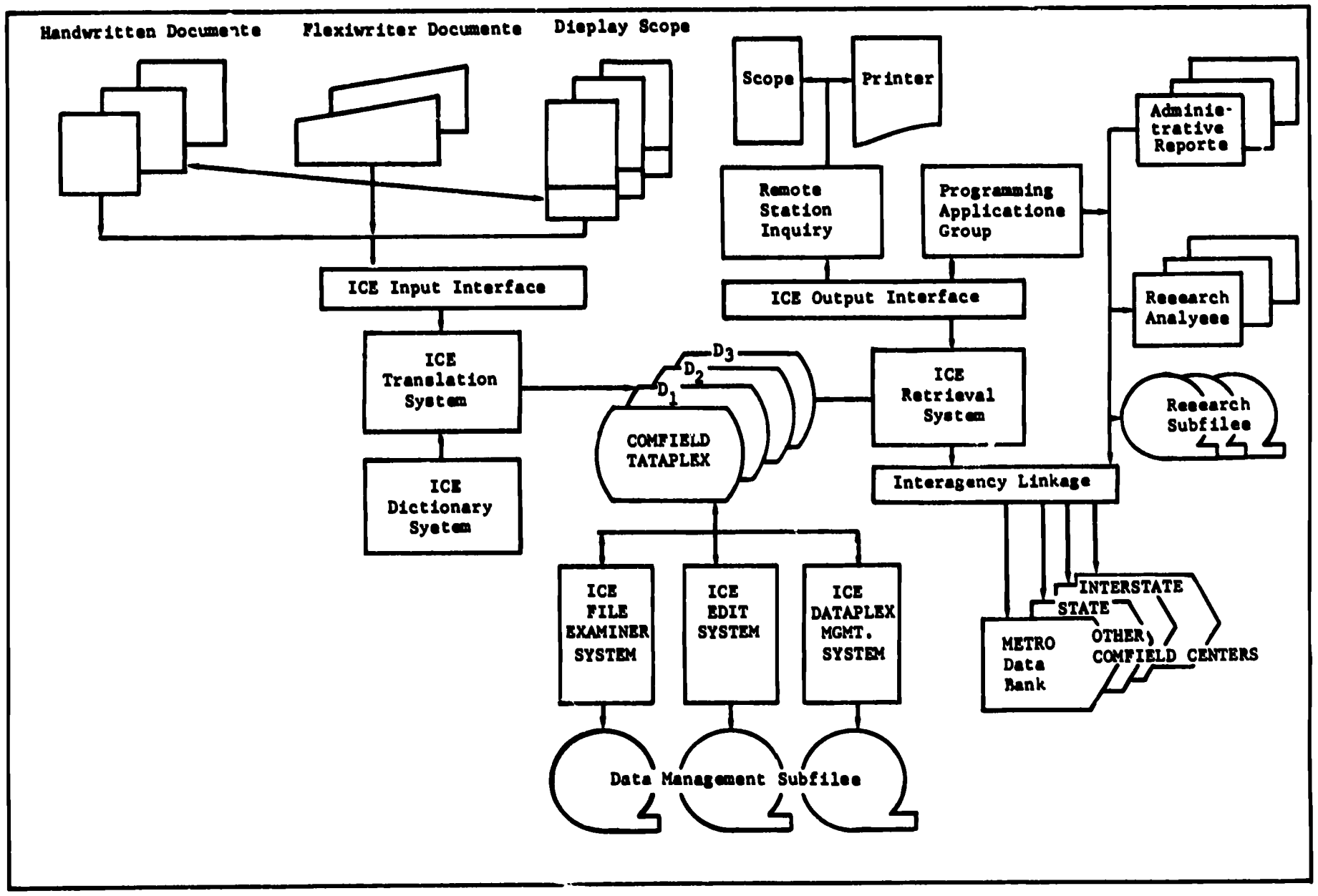
H. Summary Statement (

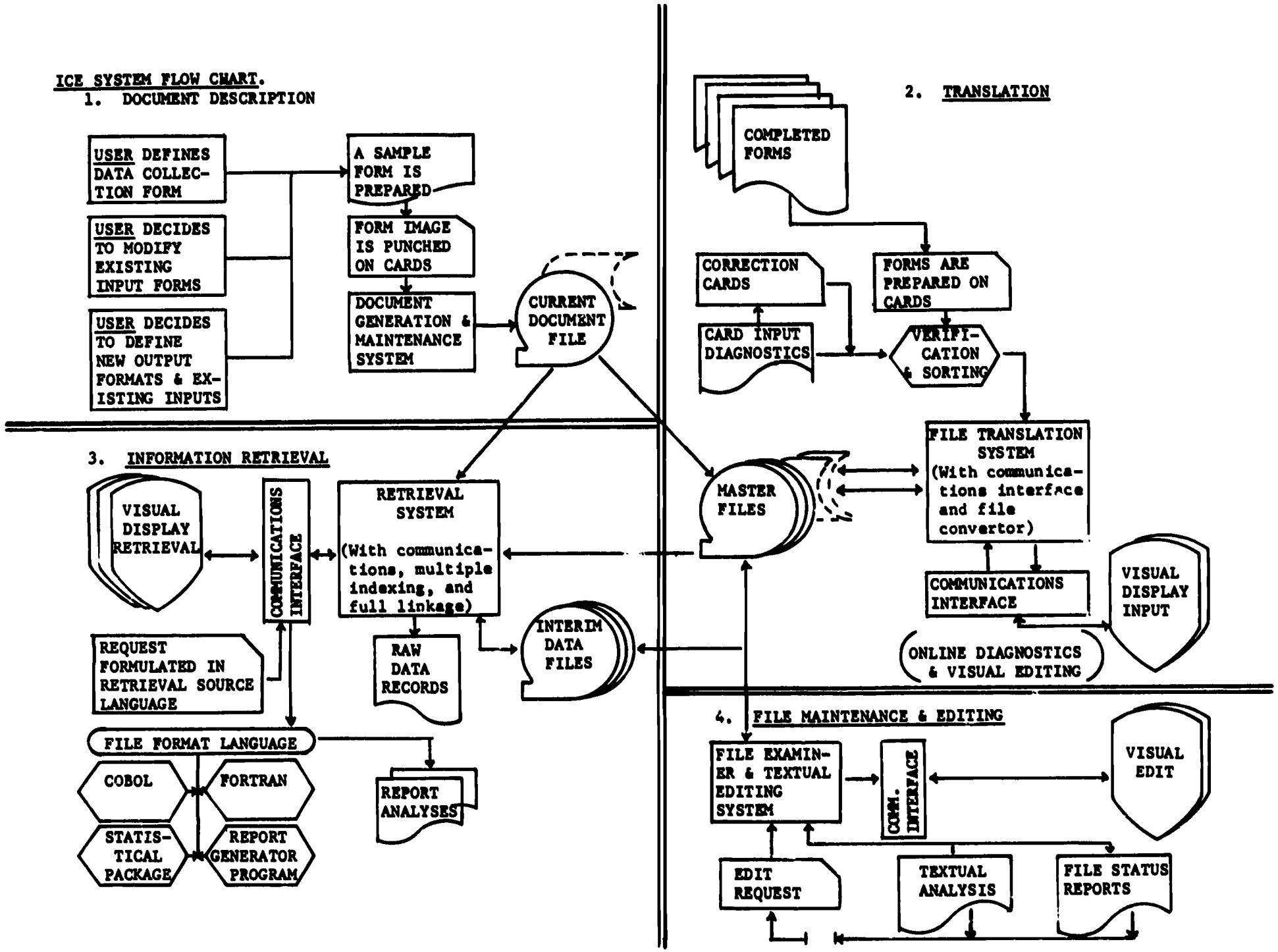
)

I

1. Teacher ( ) 2. Date (/ /)
3. Teacher ( ) 4. Date (/ /)

SCHEMATIC CONFIGURATION SHOWING ONE POSSIBLE COMFIELD SYSTEM USING
A COMPREHENSIVE STORAGE/RETRIEVAL SYSTEM SUCH AS ICE




ICE SYSTEM FLOW CHART.
1. DOCUMENT DESCRIPTION
USER DEFINES DATA COLLEC-TION FORM
USER DECIDES TO MODIFY EXISTING INPUT FORMS
USER DECIDES TO DEFINE NEW OUTPUT FORMATS & EX-ISTING INPUTS
A SAMPLE FORM IS PREPARED
FORM IMAGE IS PUNCHED ON CARDS
DOCUMENT GENERATION & MAINTENANCE SYSTEM
CURRENT DOCUMENT FILE
2. TRANSLATION
COMPLETED FORMS
CORRECTION CARDS
FORMS ARE PREPARED ON CARDS
CARD INPUT DIAGNOSTICS
VERIFI-CATION & SORTING
FILE TRANSLATION SYSTEM (With communica-tions interface and file convertor)
3. INFORMATION RETRIEVAL
VISUAL DISPLAY RETRIEVAL
COMMUNICATIONS INTERFACE
RETRIEVAL SYSTEM (With communica-tions, multiple indexing, and full linkage)
MASTER FILES
COMMUNICATIONS INTERFACE
VISUAL DISPLAY INPUT
ONLINE DIAGNOSTICS & VISUAL EDITING
REQUEST FORMULATED IN RETRIEVAL SOURCE LANGUAGE
RAW DATA RECORDS
INTERIM DATA FILES
FILE FORMAT LANGUAGE
COBOL
FORTRAN
STATIS-TICAL PACKAGE
REPORT GENERATOR PROGRAM
REPORT ANALYSES
4. FILE MAINTENANCE & EDITING
FILE EXAMIN-ER & TEXTUAL EDITING SYSTEM
COMM. INTERFACE
VISUAL EDIT
EDIT REQUEST
TEXTUAL ANALYSIS
FILE STATUS REPORTS