



AUBURN UNIVERSITY

**DEVELOPMENT OF ALABAMA
RESOURCES INFORMATION SYSTEM
A R I S**

**Annual Report
Contract NAS8-30654
1 July 1974 - 1 May 1976**

SCHOOL OF ENGINEERING

**Departments of Industrial Engineering
and Mechanical Engineering**

**Co-Directors
B.E. Herring and R.I. Vachon**

Under Contract With

**NATIONAL AERONAUTICS & SPACE ADMINISTRATION
NAS8-30654**

**ALABAMA DEVELOPMENT OFFICE
ALA-AU-X996-1060-5**

1 MAY 1976



(NASA-CR-144342) DEVELOPMENT OF ALABAMA RESOURCES INFORMATION SYSTEM, ARIS : Annual Report, 1 Jul. 1974 - 1 May 1976. (Auburn Univ.) 83 p HC \$5.00 = N76-26670 =
CSCS 08F Unclass
G3/43 42285

AUBURN UNIVERSITY

DEVELOPMENT OF ALABAMA
RESOURCES INFORMATION SYSTEM

A R I S

Annual Report
Contract NAS8-30654
1 July 1974 - 1 May 1976

SCHOOL OF ENGINEERING

Departments of Industrial Engineering
and Mechanical Engineering

Co-Directors

B.E. Herring and R. I. Vachon

Under Contract With

NATIONAL AERONAUTICS & SPACE ADMINISTRATION
NAS8-30654

ALABAMA DEVELOPMENT OFFICE

ALA-AU-X996-1060-5

1 MAY 1976

TABLE OF CONTENTS

- I. INTRODUCTION
 - II. DATA ACQUISITION
 - III. REVIEW OF EXISTING SOFTWARE
 - IV. COMPUTER ACCESS
 - V. PROTOTYPE SYSTEM DESIGN
 - VI. CENSUS PROCESSING
 - VII. SOFTWARE INSTALLATION AND USER TRAINING
 - VIII. DATA CELL SIZE AND DATA BASE CONTENT
 - IX. DATA ENCODING METHODS
 - X. INITIAL DATA RECEIPT
 - XI. PREPROCESSOR DEVELOPMENT
 - XII. INTRODUCTORY GUIDE TO ARIS
- APPENDICES

"The preparation of this report (map, plan, etc.) was financed in part through a Comprehensive Planning Grant from the Department of Housing and Urban Development, HUD P-1060; administered by the Office of State Planning, Alabama Development Office; Office of the Governor."

I. INTRODUCTION

This report provides a major summary of activity and events related to the development of ARIS - Alabama Resources Information System - during the period 1 July 1974 through the present. The last major report was entitled "ARIS - Alabama Resources Information System" and dated 30 June 1974. This earlier report was devoted to the description and discussion of the proposed system, i.e. general system design.

Work since that initial design effort has been directed toward the detailed design and implementation of a working, user-oriented set of computer procedures which permit the storage, retrieval, and manipulation of geographic and demographic data. An initial version of such a system now exists and has been used with a limited quantity of Alabama data.

The primary portion of this report addresses a series of significant events, encountered thus far during the detailed design and implementation phases. A listing of the several events, and related status, are included in Figure 1 (ARIS Milestones). The events are presented as section headings and discussed in the narrative which follows.

II. DATA ACQUISITION

Appendix A contains a listing of data which has been acquired for potential inclusion in ARIS data bases. Much additional data is held by the Alabama Development Office (ADO) and awaits a suitable means of encoding. Certain other data, such as current land use and elevation above mean sea level, is in the process of being acquired by either ADO or the ARIS team. Some regional data is being acquired by system users or potential system users for inclusion in data bases related to their areas of planning or research concerns. The usefulness of the ARIS processing procedures has captured the interest of many. Data is being gathered, for example, by a research group concerned with the feasibility of agricultural irrigation. The Birmingham Regional Planning Commission, in a trial use of ARIS, finds ARIS to be a means of gainfully utilizing the geographic information they have gathered and encoded into the far more restrictive SYMAP system. ARIS permits modeling and "what if" studies to be performed rather than merely displaying the information via computer output methods.

Early land use data, digitized by others from mosaics of low altitude photography, were found to be suspect. Inconsistencies, which were observed, included:

- a) An error of over one kilometer was calculated in the southern Mobile County area.
- b) An error was visually present in the vicinity of Dauphin Island and Big Creek Lake.
- c) An error of over two kilometers was calculated to be present in southern Baldwin County.

<u>EVENT</u>	<u>STATUS</u>
Data Acquisition	Began in May 1974 and is continuing
Review of Existing Software	May 1974 through September 1974
Computer Access	Attempts began in June 1974; decision reached in January 1975 but with lingering problems
Prototype System Design	Began in August 1974; abandoned in February 1975 in favor of acquired software from ESRI
Census Processing	Initiated in October 1974 with the population processing segment completed in July 1975 and the housing segment nearing completion
Software Installation and User Training	Completed by vendor in June-July 1975
Data Cell Size and Data Base Content	Continuing
Data Encoding Methods	Continuing
Initial Data Receipt	January 1976
Preprocessor Development	Began in September 1975 and continuing
Introductory Guide to ARIS	Completed

FIGURE 1. ARIS MILESTONES

- d) In addition to the noted positional errors, a number of errors appeared to exist in the recorded land use.

The precise cause of the errors has never been completely determined. Probable causes were:

- a) Uncontrolled encoding from the photographic source material
- b) Key punch errors
- c) Error in human judgment during the analysis of the source material
- d) Improper or non-existent cross-checking of the photographic source material with respect to the most current geologic maps
- e) Unresolved optical distortions at the edges of the photographs utilized in forming the mosaics.

Since those data are being replaced by more current high altitude photography and professional digitization, the validity of the older material may be relatively unimportant. However, it could have been used to show land use change and it does illustrate the need for exercising care in data collection and encoding.

III. REVIEW OF EXISTING SOFTWARE

This effort began with a visit to Department of the Interior (DOI) operations in Reston, Virginia and in Washington, D.C. Information was acquired concerning the polygon grid system being used by DOI to encode and display earth resources data. Also obtained was a thumbnail sketch of the efforts of DOI and the Oak Ridge Six Weeks County Study (a small scale earth resources information system).

The ARIS team also acquired several items of software for possible inclusion in the resultant information system. These included CALFORM and SYMAP, two earth resources data processing programs, as well as a series of display and correction programs from the Bureau of the Census, a population simulator from the National Center for Health Statistics, and plotting software from Tektronics, Inc. A list of these items is included in Appendix B.

By early September 1974, it had become evident that ARIS was going to be a relatively new effort. No one else seemed to have a processing procedure that permitted the storage, retrieval, and manipulation (or modeling) of geographic and demographic data. By late September, however, information was obtained indicating others were well on the way to implementing a set of procedures which appeared to be capable of meeting the needs that had led to the start of the ARIS effort. That information related to the proprietary Maryland Automated Geographic Information system (MAGI) developed by the Environmental Systems Research Institute (ESRI) of Redlands, California, under HEW/NASA funding. A listing of the software procedures incorporated in MAGI are included in Appendix B.

Investigations and slow decision-making iterations occurred during the period October 1974 through May 1975. Periodic communications during that period included the following items:

- A. Contacts have been made with personnel with the Maryland Department of State Planning and ESRI. From these contacts, it appears the MAGI software is an implementation of programs and routines previously developed by ESRI and not a resultant of the HEW/NASA contract.
- B. Initial impressions of MAGI are favorable; if nothing else, it represents a viable, current solution to resource and planning information storage and retrieval. The source language used for the programming effort was FORTRAN, thus it can reasonably well be implemented on a fairly broad range of computers.
- C. Since the AUTOMAP packages are a major investment in terms of both funds and project personnel time, care must be taken to ensure that they meet ARIS requirements to everyone's satisfaction.
 - 1) Initially, there is a need to review the reasons for utilizing ESRI's AUTOMAP system as a part of ARIS. It is felt that the ARIS/AUTOMAP system will meet graphic information retrieval requirements for Alabama in the future. This is supported by the fact that the ESRI packages are easily expandable and can be modified without redeveloping the entire system. The modular approach used by ESRI in developing their AUTOMAP system makes the addition of new program modules or subroutines relatively easy to accomplish. New routines can be used to meet the changing requirements of information retrieval for future state systems.
 - 2) Additional reasons for the recommendation for acquisition of the AUTOMAP system are based on ESRI's extended experience in the land use and graphic information retrieval field. They have been concentrating in this area for several years and have spent over two years on just the present version of the AUTOMAP II package. Their earlier experiences with forerunners of the present system led to a vastly improved design for AUTOMAP II. These improvements include a reduced operating cost per unit of information, reduction of hardware requirements (which also reduces operating cost), improvement of error detection and correcting procedures, and ease of use by non-technical users. A lower operating cost allows greater use of the operational system. Reduced hardware requirements indicate ARIS/ESRI-AUTOMAP can be implemented on a wider range of existing state-owned, medium and large computers.
 - 3) Due to experience, ESRI has been able to reduce some of the errors in earlier systems. Simpler approaches to problems reduce the probability of user errors. Other map generating packages are considered to be inefficient because a new base map is generated for each map printed. On the other hand, AUTOMAP prepares a data file containing one base map that is utilized for all maps printed. The use of this master file ensures continuity of data input and information retrieval. The file also eliminates the cost of redeveloping each map produced by the system

by reusing the same master map outline. Due to a simple set of retrieval commands and an easy to understand configuration, the user should have few problems operating the AUTOMAP system.

- 4) Finally, due to experience and greater time, ESRI has been able to develop many sub-systems which could not be developed by the ARIS group working alone. Certain portions of the AUTOMAP mapping routines would require at least a year of concentrated effort to develop independently. By using already existing AUTOMAP systems, ARIS programming time can be better used on other sub-systems.
- D. Wide use of AUTOMAP software in over twenty installations insures that ESRI will continue to support the packages in the future. This is important because unsupported software tends to develop errors which only the original authors can repair. Wide use may also promote AUTOMAP user interaction. This interaction could provide new ideas for ARIS planners and users. ESRI could offer extensions or updates to its software.

IV. COMPUTER ACCESS

The initial plans for ARIS called for investigating the possibility of using the Marshall Space Flight Center Datacraft computer and the MERITS software as an information system base. An early report to MSFC contained the following:

The terminal on loan to Auburn from MSFC has been linked through the acoustic coupler and phone lines with the Datacraft computer at MSFC. Technical help from MSFC will be required to render the "joy stick" operable and to render the "hard copy unit" serviceable. The trouble with the hard copy unit may be a need for nothing more than new paper. Also a manual or manuals are required for the terminal units. Presently projected long distance line costs appear unreasonable for the development phase of the project. FTS availability to link the terminal to MSFC computers must be determined.

Terminal repair was accomplished under a contract issued by the ARIS team. Five months after the initial concerns were voiced relating to telephone access and manuals, neither problem had been resolved.

Beginning in July 1974, the large IBM System 370 computer had received an increasing use. This began moderately with display of the early land use data and has continued as additional components of ARIS have been programmed or acquired software activated. By January 1975, the ARIS team was committed to the use of large scale computers for the storage, retrieval, and processing of the various data bases. An entry in the January 1975 report underscores this point. It is as follows:

ESRI developed AUTOMAP with IBM equipment in mind. The ARIS programming staff wishes to minimize the program modifications to AUTOMAP. Due to major differences in the FORTRAN computer language, transfer to the DATA-CRAFT equipment could involve many problems. FORTRAN incompatibility was experienced by the Maryland (MAGI) AUTOMAP users on non-IBM equipment.

If the ESRI packages were to be modified to fit particular NASA computers, then they might have to be re-modified to run on the State's computers. Particular problems the programmers wish to avoid are the differences in operating systems and hardware word configuration. AUTOMAP stacks sub-routines and calls them in overlays compatible with IBM's O.S. (operating system) configuration. The different operating system for the DATACRAFT would require re-developing these overlays. Computer word differences exist between different manufacturers. IBM utilizes a word of four eight-bit bytes, while DATACRAFT has a 24-bit word basically broken into three eight-bit bytes. These reasons indicate that staying with IBM equipment is the optimum path at the present point of the project.

ARIS development has remained tied to a major IBM 360/370 system. By May 1975, the following was noted:

Computer operations will soon become a major item of expense. Until this time, only minimal computer time has been expended for the design and checkout of some preliminary retrieval and data base management software, error searching in the obsolete aerial photo-based land use data, and development of the census data subsystem.

A request for remote access capability was also voiced in the following:

Remote access, by ADO, to a large IBM System 360/370 will become a significant need by June 1975. ADO will need the interactive terminal, promised in April 1974, beginning in the same month of June.

This request for interactive equipment was voiced continuously for several months. The need has continued until this writing but is currently receiving relief through State means. A terminal has been purchased with contract extension funds and ARIS team software development is being directed toward use of a planned Auburn acquisition of interactive capability on a large computer or similar capability on a large State computer.

V. PROTOTYPE SYSTEM DESIGN

During the period of a search for a potentially feasible and existing set of processing procedures to use in the ARIS project, detailed system design effort was undertaken. A description of that effort is included as Appendix C to this report. While that effort was largely wasted as concerns continued development, without the insight gained, the MAGI procedures would not have made as good an impression. In fact, they would not have been recognized as potentially meeting much of the needs for ARIS.

VI. CENSUS PROCESSING

An early mention of the specific task set occurred in early October 1974. Initial efforts made toward meeting this request were reported as follows:

The Alabama Development Office requested development of a program to produce certain tabular summaries from portions of the 1970 census data acquired during the early stage of the ARIS development contract. It was felt that such a routine could evolve into a useful portion of ARIS. Prior to extraction of a subset of the census data and program writing, DAULIST, a Bureau of the Census program set for census data manipulation, was checked for possible applicability. It appears DAULIST will perform the task. Unfortunately, portions of the DAULIST tape were recorded in EBCDIC mode and other portions were recorded in BCD mode. This blunder coupled with a major renovation of Auburn's computer center and associated computer unavailability delayed completion of this task.

Little real progress was made toward satisfying this request during the next couple of months. The requirement was again mentioned in the report of January 1975. Results to that time were as follows:

Early in the fall quarter of 1974, ADO requested programs to produce tables from certain of the 1970 census data. As program design progressed, it was determined that the DAULIST-4 program (purchased from the Bureau of the Census) would fulfill ADO's needs. As of this time, completion of this special task has been delayed by difficulties with purchased program and data as well as lack of time. Much personnel time has been expended in working out problems with the land use data supplied by the University of Alabama. Little time remained for the development of procedures to meet the special request.

Difficulties with purchased material consumed remaining available time. Problems encountered included: (1) a mixed mode [partly binary coded decimal (BCD) and partly extended binary coded decimal interchange code (EBCDIC)] method of encoding the DAULIST-4 program on the tape received from the Bureau of the Census, and (2) three of the 28 tapes containing the 1970 census data, purchased from the University of Alabama, contain significant quantities of errors. Since the University of Alabama has converted to a computer system which is not compatible to Auburn's, expeditious University of Alabama assistance is not possible. Steps are presently being taken to acquire corrected data from a different source.

ADO's request for various census tables included the stated desire for this tabular system to be executed after a processing request was made for a remote terminal. Output from the system was to be displayed on that same remote terminal. At that time, the assumption was to incorporate this procedure on the Huntsville DATACRAFT machine and for ADO to be provided with a terminal to the DATACRAFT. Use of DAULIST-4 on the DATACRAFT is not feasible since DAULIST-4 is written in the COBOL programming language and the DATACRAFT has no COBOL compiler. Conversion of DAULIST-4 from COBOL to FORTRAN is not especially feasible because of the size and complexity of that program, the lack of an inexpensive communications link to Huntsville, the size of the census data base, and the size, relative speed, and word structure of the DATACRAFT machine.

It is felt the special request should be fulfilled by use of DAULIST-4 in its current COBOL configuration utilizing an IBM 360/370 system initially in a batch mode but with the possibility for later remote (terminal) entry capability. While it is possible to write a new program to search, compare, and tabulate the desired types of census data, this effort would require several months of additional effort. Additionally, should it be decided to reprogram using FORTRAN, it might be desirable to program for the DATACRAFT. This latter possibility would require a suitable communication link or temporary duty in Huntsville. This would also encounter the earlier noted problems related to use of the DATACRAFT.

After the January 1975 report was distributed, a more definitive requirement was received from ADO. This was incorporated into a specific contract extension in March of 1975. Appendix D contains a reprint of the specific requirements to be met by the computer procedure. By April of 1975, the following communication was made:

Excellent progress is being made on the development for ADO of a subsystem capable of manipulating and retrieving certain population information from the 1970 U.S. Census of Population 4th Count Summary. A prototype procedure is now working, which aggregates information from two or more census districts and certain data items. Work is continuing on the design of the procedure required to perform a research/selection function over many census districts. The purpose of the latter task is to permit mechanized searching for the census districts having certain desired population attributes.

This subsystem was on its way to becoming operational.

Reports in May, June, and July of 1975 continued to indicate progress and finally completion of CENSLIST. These reportings follow:

In May - Completion is nearing on the development for ADO of a subsystem capable of manipulating and retrieving certain population information from the 1970 U.S. Census of Population 4th Count Summary. A minor peculiarity in the data, consisting of apparent randomly placed sets of blanks, has been detected and accommodated for in the software. A procedure now works which aggregates information from two or more census districts and certain data items. In the near future, personnel from ADO will be invited to visit and try their hand at using this procedure to retrieve any material contained in that census file. It is hoped this trial will serve as the demonstration of that portion of the assignment.

Work is nearing completion on the design of the procedure to perform a research/selection function over many census districts. A set of user instructions for use of the census subsystem is being designed.

In June - The procedure for aggregating information from two or more census districts on certain specified data items is felt to be operational. Heading information, which did cause some trouble, is now functioning correctly.

The problem with random blank fields in the census data base has been corrected by careful copying of all items except for the blanks. The search/selection function is also essentially complete. As soon as ADO personnel can arrange to make a visit, they are invited to come and try their hand at using the procedures to retrieve any material contained in the 1970 U.S. Census of Population 4th Count Summary.

By July - CENSLIST, the package which goes beyond the Census Bureau's DAULIST, is now operational. In addition to performing the DAULIST function of building display tables at the Census County Division (CDD) level. CENSLIST has the additional capability to aggregate data over all desired CCDs and print DAULIST tables. CENSLIST will also scan all census districts to find and list those having a specified population attribute with a value greater than or equal to some desired target level. The manual which describes the use of CENSLIST is completed and will be transmitted shortly to ADO.

Since the development of CENSLIST, work has been directed towards the preparation of a somewhat similar procedure to list, aggregate and search through the housing data base. As of this writing, a procedure called HOUSLIST is nearing completion.

VII. SOFTWARE INSTALLATION AND USER TRAINING

As noted earlier in this report, after hearing about and then reviewing Maryland's MAGI system, the ARIS team recommended acquisition of that software. It was felt that software, modified and supplemented as necessary, would be operable sooner and less expensively than would be possible through complete local development of comparable material.

Some indecision still existed by April 1975. Because of that, the following was included in the April report:

Dr. Herring and Mr. Wysong have scheduled a trip for late April to the Environmental Systems Research Institute (ESRI) in California to evaluate new segments of the ESRI processing procedure and their data entry methods. Consideration may be given to sub-contracting (by ADO) part or all of the data encoding activity, as is done by the State of Maryland.

The above trip resulted in highly satisfactory findings. The May 1975 report included the following:

Dr. Herring and Mr. Wysong travelled to the facilities of the Environmental Systems Research Institute in Redlands, California, during the end of April and the beginning of May. It is their opinion that ESRI is very competent in the field of geographic information systems software. ESRI has worked with many governmental units in the development and use of geographic, including geologic, land use, and transportation network, data files. ESRI has set up systems for the State of Maryland, several counties in California, Newcastle County in Delaware, and several places in foreign countries.

A purchase order has been initiated for acquisition of the ten software packages listed in the letter proposal to NASA from Auburn University dated 17 March 1975. Installation of these packages on the Auburn University computer is expected to commence on 16 June 1975.

A small delay of one week was projected before software installation began. The June 1975 report noted this delay and included an invitation for review of the software capability. That report segment was:

Mr. Ray Postma of the Environmental Systems Research Institute in Redlands, California, will arrive in Auburn on June 22 to begin installation of the ESRI software. It is suggested that Thursday, June 26, would be a good day for any interested NASA or ADO personnel to visit Auburn for an opportunity to observe trial runs of the software. Data bases to be used for installation and checkout will contain geographic data from other parts of the United States and/or Japan.

A training session on the use of this land use software is tentatively scheduled for a week early after the 15th of July. ADO personnel have the responsibility to determine the precise set of dates which will fit with their workload and that of other interested parties.

Software installation occurred on schedule. A small interface problem occurred with respect to Auburn's CALCOMP plotter installation. These items were noted as follows in the July 1975 report:

Mr. Ray Postma of the Environmental Systems Research Institute in Redlands, California, was in Auburn from 22 June until 27 June installing the purchased software. All packages except for GRIDPLOT are currently operational. The GRIDPLOT package utilizes standard CALCOMP software which does not exist at Auburn. Work is being performed both at Auburn and in California to overcome this temporary difficulty. Operational software is most impressive with respect to speed of processing and the small portion of CPU space required.

A very successful training session was held at Auburn during the week of 21 July 1975. Many exercises were attempted by the participants and most were processed through the computer before the seminar ended. Fourteen individuals from State or regional planning agencies attended at least part of the seminar. A further success was also noted - the plotter interface problem was located and resolved.

VIII. DATA CELL SIZE AND DATA BASE CONTENT

Both of these problems have oscillated between being noted resolved and without solution as well as most possible points between these extremes. Data cell size was initially defined to cover one kilometer. During August of 1974, Mr. Jack Dangermond, President of ESRI, met with interested individuals. The major topics covered were data base content and grid cell size. Decisions made included:

- A. To determine the data base content, a series of general information need areas were established. These categories, shown in Table I, were called Public Safety, Renewable and Non-Renewable Resources, Unique Natural Features, Rare and Endangered Species, Biological Productive Areas, and Air Resources. Each general need was first expanded to determine the related sub-topic(s), e.g., flooding versus public safety. Each sub-topic was then investigated to determine the specific information source(s) which would satisfy the requirements implied by that sub-topic. Table I provides a summary of that exercise.
- B. Once the set of specific information sources were determined, it was possible to analyze each source for usability. This analysis is summarized in Table II. For each information source, analysis consisted of determining whether the information from that source: (1) was currently in a state to permit digitizing; (2) required mapping or further interpretation; (3) should be digitized directly with the use of a grid overlay scheme, should be digitized by conversion from polygons, should be digitized by use of scan methods, or should be digitized from coded line and point data.

During this period of detailed analysis, the question of areal grid sizes was resolved. Since the ESRI/ARIS software and data base structure is extremely flexible and since precise regional and urban data needs are not fully known, it was decided to initially keep the overall State data base separate from that established for each of the several regions and urban areas. To insure the compatibility of geographic coordinates among these three major data bases, regional and urban use of the software will be predicated on regional and urban use of State supplied geographic coordinates and necessary State coordination of the use of the software. This will provide the assurance of data base compatibility and successful software operation simultaneously with desirable flexibility and organizational independence. Statewide grid sizes will be 1/16 of a kilometer in area. This cell size is readily summarizable to either 1/4 or a kilometer or a full kilometer (or larger) for information retrieval activity requiring less detail. For comparison with acres, areas of 1/16 and 1/4 of a kilometer are approximately 15 acres and 60 acres respectively.

Since the August 1975 meeting, revisions have again been made in the cell size. Statewide data will be encoded for inclusion in a 1/4 kilometer grid cell. Regional or local interest data will hopefully be encoded into a cell size summarizable to either the 1/4 or full kilometer grid cell. However, strong insistence on this is not likely to be made. For example, the Birmingham Regional Planning Commission (BRPC) earlier encoded much data into 40 acre size cells. This cell size was chosen because of limitations on the number of cells imposed by the particular software utilized during their early efforts in establishing a data base. It is likely this 40 acre cell data base will be reformulated for entry into the ARIS system to permit modeling and other multi-data base operations to be performed by BRPC personnel.

ARIS project personnel have been pleased with a statement made by U.S. Department of the Interior in a booklet released in January 1976, entitled State

GENERAL CATEGORY	SPECIFIC STUDY AREA	SPECIFIC INFORMATION	
PUBLIC	FLOODING	FLOOD PLAINS & CLIMATE	
	UNSTABLE SOILS	(MASS MOVEMENT, SUBLIMATION, FROSTION)	SOILS, SLOPE, GROUND-WATER TABLE, GEOLOGY, RAINFALL, VEGETATION, SURFACE HYDROLOGY
	HURRICANE		CLIMATE
	UNSTABLE RIVERS		SURFACE HYDROLOGY
	UNSTABLE SEASHORES		COASTAL HYDROLOGY
RENEWABLE & NON-RENEWABLE RESOURCES	WATER RESOURCES	SURFACE & SUBSURFACE WATER QUALITY & QUANTITY, RECHARGE BASINS, CAPACITY OF BASINS	
	FORESTS	SOILS, VEGETATION TYPE, HIGHWAYS, BIO-HAZARDS	
	CROP LANDS	SOILS, SLOPE, EXISTING COVER	
	GRAZING LANDS	SOILS, SLOPE, EXISTING COVER, BIO-HAZARDS	
	MINERAL RESOURCES	GEOLOGIC MAP, MINERALS MAP, COVER, ACCESS TO TRANSPORTATION	
REPRODUCIBILITY OF THE ORIGINAL PAGE IS POOR	WATERSHED	SOIL TYPE & SLOPE	
	SURFACE WATER STORAGE AREAS	SLOPE, ELEVATION CHANGE, EXISTING COVERAGE, WATERSHED?	
	UNIQUE NATURAL FEATURES	AREAS OF RECREATIONAL VALUE, AREAS OF SCIENTIFIC INTEREST, AREAS OF SCENIC BEAUTY	MAPS OF FEATURES
	RARE & ENDANGERED SPECIES		
	BIOLOGICAL PRODUCTIVE AREAS	MARINE AQUATICS	WATER QUALITY
FRESHWATER			
AIR RESOURCES	LAND	VEGETATION	
	SURFACES OF POLLUTION		

TABLE 1. GENERAL INFORMATION REQUIREMENTS FOR PROBLEMS FACING THE STATE OF ALABAMA.

TABLE II. General Information Availability for the State of Alabama

Information	Ready to Digitize	Needs Mapping or Interpretation	Type of Encoding	Notes
State Soils	Yes		Polygon	
County Soils	Partial	Yes	Grid	Do as needed
Land Use (state)		Mapping & photo interpretation	Polygon	
Land Use (regional)		Mapping & photo interpretation	Polygon	Same as state but rerun
Ground Water Table		Yes	Polygon Digitize	Include - depth to base, piezometric quantity, water quality, yield
Slope	Yes		Scan	1) Check with USGS, 2) check with Geological Survey, 3) see E.S.C.
Geology (state)		1/2 year	?	Evaluation of state & county uses before going ahead
Geology (local)	Yes	?	Coded	
Topographic			Grid	Acquire USGS computer tapes
Flood Plains	Yes	Yes	Coded?	Code at 1/16 sq. km. and aggregate
Mineral Resources (state)	Yes	Yes	Polygon	Add significant elements to state map
Mineral Resources (local)	6-8 mos.	Yes	Coded	
(S. Sub Basins)	Yes		Polygon	
Water Sheds	?	?	Code	

TABLE II. (cont'd)

Information	Ready to Digitize	Needs Mapping or Interpretation	Type of Encoding	Notes
E - Districts	No	Yes	Polygon	Remap base map
Ownership (county)		Yes		(by sequential number - state, federal & local gov't.)
Sewered and Planned Sewered Areas		Yes	Coding	
Transportation		Yes	Coded	Combine and evaluate
Energy	Yes		Digitize?	Add coordinates
Historical & Archeological		Yes	Coded	Needs more investigation
Planning Maps by Local Gov't.	--	--	--	
State Critical Areas (by State Planning Org.)		Yes	Code by unique D.	Development of general categories
Surface Hydrology				Integrate wet-land mapping Classify stream network

Resource Management Programs, Primer, Critical Areas and Information/Data Handling. This statement noted "The Maryland Geographic Information System (MAGI) has the most comprehensive capability in existence for statewide manipulation of spatial data." ARIS contains precisely the same software to store, retrieve, and manipulate spatial data.

IX. DATA ENCODING METHODS

Table II lists recommended general encoding methods for various types of data. For the most part, these methods, such as polygon and grid, refer to the organization of the raw data before conversion to a computer readable medium. The actual work is performed by either machine methods or manual effort. Machine methods include line tracing on a map with a stylus/tablet device or with a drafting machine arrangement. While other devices, such as a light pen and cathode ray tube display are possible, the former two approaches are definitely workable. MSFC personnel have agreed to prepare the software necessary to permit digitization by the stylus/tablet method. This will permit relatively low cost entry of a large quantity of Alabama data.

Manual encoding of data includes two major steps. The first is a transformation of material from a map to some sort of coding form. This is typically performed by use of a mylar overlay superimposed onto a stable base map. The second step includes the conversion of the notations on the coding form to a machine readable medium, e.g., key punched. Appendix E contains a discussion of encoding procedures presently being used for ARIS.

In many cases, machine digitization provides the best approach. However, this represents a large startup cost and because of its throughput capability, relatively low utilization over time by any one major user. For these reasons, a regional facility to serve several states would provide a more economical solution than for each user to develop an in-house digitizer.

X. INITIAL DATA RECEIPT

ARIS software and procedures received initial real checkout with Alabama data during the last part of January 1976. Floodplain data for Walker County was encoded using material similar to that contained in Appendix E. A map shown as Figure 2 was the result of this trial effort.

XI. PREPROCESSOR DEVELOPMENT

Early, detailed analysis of the ESRI software yielded one problem. While that software was designed with the system user in mind, successful operation of the programs requires the services of an individual who is very knowledgeable in the use of a computer. The vast majority of ARIS users are not expected to have this background. For this reason, a preprocessor having two major functions is under detailed design. The first function is to serve as an instructional aid to improve user understanding of the ARIS software, data base content, and their appropriate interactions. This component is being

REPRODUCIBILITY OF THE
ORIGINAL PAGE IS POOR

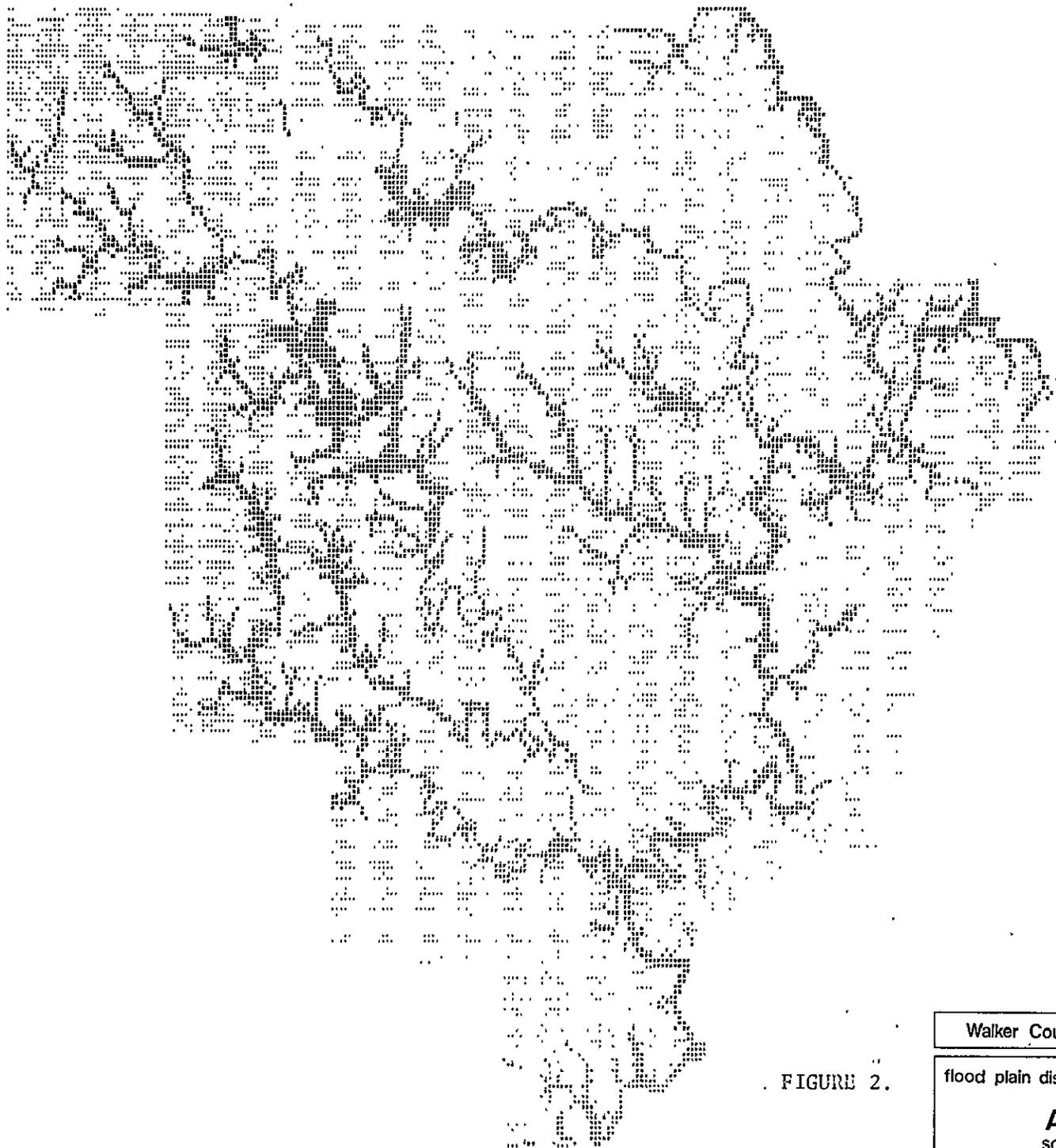


FIGURE 2.

Walker County, Alabama
flood plain display prepared by
ARIS software
Jan, 76

structured to serve as both a system introduction to new users and as a refresher for the most experienced. The second function of the preprocessor is to provide an interactive remote job entry initiator to optimize decision-maker utilization of the ARIS processing software packages and data bases. Specific action of the initiator is to deal with user error detection and provide guidance in the task of selecting the appropriate commands necessary to initiate a given processing sequence. Smooth progress on this task is being hampered by the present non-existence of computer configuration expected to be used when the system is fully implemented. The ARIS team is preparing software to utilize an interim configuration. Resultant software will be adaptable to the final equipment configuration.

XII. INTRODUCTORY GUIDE TO ARIS

A Guide to ARIS, a booklet depicting the content, configuration, and uses of ARIS has been completed. This booklet utilizes a combination of words and pictures to present the message to potential users of the system. It has been several months in production but is considered to provide a complete introduction to the ARIS concept. A rough draft copy of the booklet is included with this report.

REPRODUCIBILITY OF THE
ORIGINAL PAGE IS POOR

APPENDIX A

APPENDIX A

The following publications have been received, sorted, and filed from Geologic Survey of Alabama:

Special Maps

1. Map 10 - Profile Showing Geology Along U. S. Highway 331 Montgomery County, Alabama
2. Map 13 - Geologic Map of Limestone County, Alabama
3. Map 23 - Geologic Map of Morgan County, Alabama
4. Map 25 - Geologic Map of Madison County, Alabama
5. Map 26 - Geologic Map of Escambia County, Alabama
6. Map 36 - Oil and Gas Exploration Wells in Alabama
7. Map 37 - Geologic Map of Greene County, Alabama
8. Map 40 - Geologic Map of Pickens County, Alabama
9. Map 50 - Mineral Resources of Escambia County, Alabama
10. Map 60 - Generalized Geologic Map of Marshall County, Alabama
11. Map 83 - Mineral Resources of Baldwin County, Alabama
12. Map 89 - Mineral Resources of Mobile County, Alabama
13. Map 93 - Geology of Mobile County, Alabama
14. Map 94 - Geology of Baldwin County, Alabama
15. Map 96 - Water Availability of Baldwin County, Alabama
16. Map 99 - Subsurface Geologic Map of Southwest Alabama

Environmental Geology of an Area in West-Central Alabama - Mosher and Keener (Geological Survey of Alabama)

Bulletins

1. Bulletin 83 - Ground Water Resources of Pickens County, Alabama

2. Bulletin 87 - 7-Day Low Flows and Flow Duration of Alabama Streams
3. Bulletin 74 - Geology and Ground-Water Resources of Escambia County, Alabama
4. Bulletin 68 - Geology and Ground-Water Resources of Montgomery County, Alabama (Part A and B)
5. Bulletin 93 - An Oil and Gas Evaluation of North Alabama
6. Bulletin 86 - Geology and Ground-Water Resources of Greene County, Alabama
7. Bulletin 89 - Water Laws of Alabama
8. Bulletin 85 - Ground Water in Marshall County, Alabama
9. Bulletin 84 - Surface Water in Southwestern Alabama
10. Bulletin 76 - Geology and Ground-Water Resources of Morgan County, Alabama
11. Bulletin 50 - Well Logs of Alabama
12. Bulletin 62 - Gound-Water Resources of the Huntsville Area, Alabama
13. Bulletin 58 - Ground-Water Investigations in the Mobile Area, Alabama
14. Bulletin 57 - Well Logs of Alabama, 1940-1945
15. Bulletin 10 - Reconnaissance Report on the Fayette Gas Field
16. Bulletin 31 - Analysis of Alabama Coals
17. Bulletin 47 - Bauxite Deposits of Alabama
18. Bulletin 99 - Subsurface Geology of Southwest Alabama

Circulars

1. Circular 32 - Flow Characteristics of Alabama Streams
2. Circular 24 - Ground-Water Conditions in the Huntsville Area, Alabama, January 1960 Through June 1961
3. Circular 23 - Ground-Water Levels in Alabama in 1959 and 1960

4. Circular 53 - The Petroleum Industry in Alabama, 1967
5. Circular 47 - Geology of the Alabama Coastal Plain
6. Circular 20-E - Clays of Southwestern Alabama
7. Circular 37 - A Compilation of Ground-Water Quality Data in Alabama
8. Circular 35 - Subsurface Geology of the Gilbertown Oil Field
9. Circular 36 - A Compilation of Surface Water Quality Data in Alabama
10. Circular 39 - Alabama's Mineral Industry
11. Circular 46 - The Geology and Mineral Resources of Limestone County, Alabama

Special Reports

1. Special Report 17 - Natural Resources of the Tennessee Valley Region in Alabama
2. Special Report 18 - Ground-Water Resources of the Cretaceous Area of Alabama
3. Special Report 24 - Surface Water Resources and Hydrology of West-Central Alabama
4. Special Report 15 - Oil and Gas in Alabama
5. Special Report 11 - Resources of Southern Alabama

County Reports

1. County Report 3 - Geology and Ground-Water Resources of Madison County, Alabama
2. County Report 11 - General Geology and Ground-Water Resources of Limestone County, Alabama

Soil Map

1. General Soil Map of Alabama (Pre-publication Copy) 1:500,000
REPRODUCIBILITY OF THE ORIGINAL PAGE IS POOR

Topographic Maps*

1. Base Map of Alabama, 1:1,000,000
2. The Following 1:250,000 Topographic Map Series:
 - Blytheville (1956-64)
 - Columbia (1956-66)
 - Tupelo (1956-66)
 - Gadsden (1956-66)
 - Rome (1958-66)
 - West Point (1953-70)
 - Birmingham (1953-63)
 - Atlanta (1953-64)
 - Meridian (1953-65)
 - Montgomery (1953-69)
 - Phenix City (1955-63)
 - Hattiesburg (1953-70)
 - Andalusia (1953-65)
 - Dothan (1953-65)
 - Mobile (1953-62)
 - Pensacola (1957-70)
 - Tallahassee (1954-66)
3. The Following 15 Minute Quadrangle Topographic Maps:
 - Caledonia (Miss.) (1960)
 - Columbus (Miss.) (1960)
 - Tuscaloosa (1924)
 - Eutaw (1927)
 - Epes (1929)
 - Plantersville (1959)
 - Billingsley (1959)
 - Elmore (1959)
 - Wetumpka (1959)
 - Benton (1957)
 - Autaugaville (1957)
 - Montgomery (1958)
 - Mount Meigs (1958)
 - Letchatchee (1957)
 - Choctaw Bluff (1943)
 - Deer Park (1942)
 - Citronelle (1943)
 - Tensaw (1943)
 - Perdido (1942)
 - Wilmer (1942)
 - Creola (1941)
 - Bay Minette (1941)
 - Dyas (1942)
 - Century (Fla.) (1941)
 - Jay (Fla.) (1942)

*Ala. State Topographic 1:500,000, Ala. Shaded Relief 1:500,000, and Ala. Base 1:500,000 maps are on order but has not arrived.

15' Quadrangle Topographic Maps (Cont.)

Munson (Fla.) (1948)
 Crestview (Fla.) (1949)
 Hurley (1941)
 Mobile (1940)
 Robertsdale (1941)
 Muscogee (1941)
 Grand Bay (1958)
 Weeks Bay (1941)
 Foley (1941)

4. The Following 7½ Minute Quadrangle Topographic Maps:

Union Hill (TVA 1948)
 Salem (TVA 1948)
 Elkmont (TVA 1948)
 Ardmore (TVA 1958)
 Toney (TVA 1958)
 Fisk (TVA 1948)
 New Market (TVA 1948)
 King Cove (TVA 1948)
 Rogersville (TVA 1952)
 Cairo (TVA 1948)
 Ripley (TVA 1948)
 Athens (TVA 1958-67)
 Capshaw (TVA 1958)
 Jeff (TVA 1958-67)
 Meridianville (TVA 1948-64)
 Maysville (TVA 1948)
 Hollytree (TVA 1948)
 Hillsboro (TVA 1948)
 Jones Crossroads (TVA 1947)
 Tanner (TVA 1937-63)
 Greenbrier (TVA 1938-63)
 Madison (TVA 1948-64)
 Huntsville (TVA 1948-64)
 Moontown (TVA 1948)
 Paint Rock (TVA 1948)
 Trinity (TVA 1948-63)
 Decatur (TVA 1950-63)
 Mason Ridge (TVA 1949)
 Triana (TVA 1948-64)
 Farley (TVA 1947-64)
 New Hope (TVA 1947)
 Grant (TVA 1947)
 Swearengin (TVA 1947)
 Langston (TVA 1947)
 Danville (TVA 1960)
 Hartselle (TVA 1949)
 Somerville (TVA 1949)
 Center Grove (TVA 1947)

7½' Quadrangle Topographic Maps (Cont.)

Newsome Sinks (TVA 1947)
 Guntersville Dam (TVA 1948)
 Mt. Carmel (TVA 1948)
 Columbus City (TVA 1947)
 Grove Oak (TVA 1946)
 Massey (TVA 1960)
 Falkville (TVA 1949)
 Eva (TVA 1949)
 Lawrence Cove (TVA 1949)
 Hulaco (TVA 1949)
 Arab (TVA 1948)
 Guntersville (TVA 1948)
 Albertville (TVA 1947)
 Painter (TVA 1947)
 Brooksville (TVA 1958)
 Douglas (TVA 1958)
 Boaz (TVA 1958)
 Rodentown (TVA 1959)
 Altoona (TVA 1958)
 Fernbank (1967)
 Millport (1967)
 Kennedy (1967)
 Newtonville (1967)
 Ethelsville (1966)
 Reform (1967)
 Palmetto (1967)
 Brownville (1967)
 Reform SW (1966)
 Carrollton (1967)
 Gordo (1967)
 Elrod (1967)
 Aliceville North (1970)
 Cunningham (1970)
 Kirk (1970)
 Romulus (1970)
 Aliceville South
 Pleasant Ridge (1970)
 Mantua (1970)
 Ralph (1970)
 Fosters (1969)
 Warsaw (1970)
 West Greene (1970)
 Gainesville (1970)
 Boligee (1970)
 Mitchell Dam NW
 Richville (1971)
 Rockford SW (1971)
 Equality (1971)
 Our Town (1971)
 Forkland (1947)

7½' Quadrangle Topographic Maps (Cont.)

Mason Bend (1947)
 Red Hill
 McDowell (1948)
 Demopolis (1946)
 Tallassee
 Carrville
 Jefferson
 Montgomery North (1958)
 Shorter
 La Place
 Whitfield
 Myrtlewood N
 Montgomery South (1958)
 Lapine
 Grady (1968)
 Ansley (1968)
 Needmore (1968)
 Dixie (1971)
 Bradley (1971)
 Huxford (1960)
 Barnett Crossroads (1960)
 Wallace (1960)
 Brewton North (1960)
 Atmore (1960)
 Flomaton (1960)
 Pollard (1960)
 Brewton South (1960)
 Kushla (1940-53)
 Chickasaw (1940-53)
 Hurricane (1940-53)
 Tanner Williams (1957)
 Spring Hill (1940-53)
 Mobile (1939-53)
 Bridgehead (1939-53)
 Stapleton (1940-53)
 St. Elmo (1957)
 Theodore (1939-53)
 Hollingers Island (1939-53)
 Daphne (1939-53)
 Silverhill (1940-53)
 Kreole (1958)
 Grand Bay (1958)
 Coden (1956)
 Bellefontaine (1956)
 Point Clear (1956)
 Lillian (1970)
 West Pensacola (1970)
 Grand Bay SW (1958)
 Isle Aux Herbes (1958)

7½' Quadrangle Topographic Maps (Cont.)

Heron Bay (1958)
 Little Dauphin Island (1958)
 Perdido Bay (1970)
 Petit Bois Island (1958)
 Petit Bois Pass (1958)
 Fort Morgan NW (1958)
 Fort Morgan (1958)

Note: All remaining topographic quadrangles available have also been received from Geologic Survey of Alabama.

Monograph

1. Monograph 10 - Forests in Alabama

Soil Surveys from U. S. Department of Agriculture for the following counties:

1. Baldwin County
2. Greene County
3. Madison County
4. Montgomery County
5. Marshall County
6. Autauga County
7. Elmore County
8. Sumter County
9. Pickens County
10. Mobile County
11. Limestone County
12. Morgan County
13. Elmore County

Earth Resources Technology Satellite Photography

1. 12 Sheets of Band 5 ERTS Photography of Alabama
2. 12 Sheets of Band 7 ERTS Photography of Alabama

Land Ownership Maps

1. Baldwin County, Alabama (1974), 4 sheets
2. Conecuh County, Alabama (1974), 3 sheets
3. Pike County, Alabama (1974), 2 sheets

APPENDIX B

APPENDIX B
ACQUIRED SOFTWARE

Bureau of the Census

DAULIST 1 Display Program
 " 2 " "
 " 3 " "
 " 4 " "
 FIXDIME Correction Program

Laboratory for Computer Graphics and Spatial Analysis

CALFORM Version I Computer Mapping Program
 SYMAP Version V Computer Mapping Program

National Center for Health Statistics

POPSIM, FORTRAN IV Computer Program

Tektronics, Inc.

Plot 10/Terminal Control System
 Plot 10/Advanced Graphing II
 Plot 10/Preview Routines for Calcomp Plotters

Software Utilized in MAGI

- a. FILGEN - a program for efficiently storing grid cell information from cards or tape on new computer files for re-use by other grid analysis programs.
- b. MAP AND GRID MERGE - a program used to merge various grid cell data files of single and multiple variables from single or multiple geographic areas into one consistent, coordinated data file usable by the various analysis and mapping programs.
- c. SUBROUTINE MODEL - a program and subroutine procedure for developing overlay models of geographic data. The MODEL program is merely a framework within which the user may develop specific types of models (e.g., suitability analysis, landscape process models, etc.)
- d. SEARCH - an analysis subroutine used to calculate geographic distances between various types of phenomena stored in the data bank. It can also be used to summarize the frequency of a given phenomenon within a specified distance from a given cell. The program produces new data files which can be mapped or used in various modeling efforts. For example, for any given grid cell, one could calculate the distance to the nearest highway, commercial land use, park, etc.

- e. GRID GRAPHIC PROGRAM - a program for producing graphic output for geographic areas in the form of printer graphics. The program also produces frequency plots, level calculation, and a series of scale variations. The program is considerably flexible in terms of producing not only graphics but quantitative data as well.
- f. GRID PLOT - a plotter program that interfaces with the "GRID GRAPHICS" program. Its function is to produce maps of the data that have higher quality graphics.
- g. WINDOW LIST - a program designed to give numerical printouts of single cells or multiple groupings of cells wherein it is desirable to present a "profile" of environmental and cultural data in listing form. The program can also be used to create a window data file of a portion of the state (i.e., a single county, watershed or other geographic area wherein a specialized computer mapping or analysis effort might be desired).
- h. BASEMAP - a program creating the Base Map Image File which subsequently becomes a permanent file for re-use by the other programs. This saves an enormous amount of computer time because the grid matrix is made only once.
- i. AREAMAP - A program designed to read in the Base Map Image file and produce thematic choropleth maps on a standard chain line printer. There are a variety of simple user options available in this system such as scaling, shading, legend association, and a matrix file of the absolute values associated with each grid cell.
- J. CONTOUR MAP - a program to produce contour and proximal maps. It contains additional options that can be employed for specific applications.

APPENDIX C.

APPENDIX C
PROTOTYPE SYSTEM

After discussing the desired capabilities of the ARIS-I System with representatives of the Alabama Development Office, several new design criteria and constraints were established. These criteria and constraints are an extension of those mentioned in the ARIS study completed in June and are presented as follows:

- a) The system must accept a variable data base to accommodate a wide range of user needs. ARIS must be capable of enlarging the number of independent data items which it can correlate and retrieve. This variable design makes use of a dictionary to reference what data items are available to the user in ARIS.
- b) The system language must respond to a wide range of user language employed to retrieve data. After studying several retrieval languages and a simple natural language, no language was found that fully met the ARIS requirements. These retrieval languages were all based on a process of retrieving bibliographic data. Since ARIS is concerned with retrieval of a large variety of data, some new language is desirable. This problem was solved by providing ARIS with a few simple internal commands which can then be assembled by an ARIS-I Systems Programmer, to provide a large set of commands for the ARIS-I user to program. These commands are of a variable nature and can be changed to fit ARIS-I user needs. The actual retrieval commands will be designed after a discussion with some potential users of ARIS-I.
- c) The system cannot use a sequential file storage and search procedure. The data records cannot be accessed fast enough for a realistic retrieval of all 67 counties if a sequential file storage and search procedure is used. Thus, a procedure based on direct access records was established and this procedure is compatible with the Datacraft and IBM Systems. The direct procedure permits relatively rapid access, efficient storage, and simple design of the data base.
- d) ARIS file design has become the basis for the ARIS retrieval package. Program language constraints require the use of a key number assigned to each file record. A read command is issued for a particular record in a file by passing the key number for that record as a parameter. The main ARIS file is based on dividing the State into one kilometer grids and assigning one direct access record and key number to each grid. In each of these main file records are the key numbers for a second set of records which contain specific information about that one kilometer grid. An example is shown in Figure 1. The main record is for a fictitious kilometer grid outside of Mobile in Mobile County. This grid contains a portion of U.S. Highway 31, State Highway 42, and Three Miles Creek as shown in Figure 1a. The main record corresponding to grid 9901 is shown in part b of the figure. The grid key number allows for retrieval of the information concerning this grid as rapidly as possible. Once grid 9901 has been retrieved by the ARIS System,

further information can be retrieved concerning elements in grid 9901. By taking the key numbers found on the main record and using them to access records found in the second file, ARIS-I retrieves information concerning items found in grid 9901. Figure 1c contains examples of the records corresponding to the key numbers in Figure 1b.

The system must make use of pre-written program packages with little or no modification. Program packages, such as POPSIM and DAULIST, must be made available to the ARIS user. Due to the limited nature of this project, there will be no time or personnel to integrate fully these packages into ARIS.

The basic outline for the ARIS file structure will be expanded to fit more data items as they are needed. Further sections of the ARIS-I file design are not discussed here but will be covered in future reports as they are developed.

After completion of the file design, the next major consideration is data input for the ARIS-I data base. A manual system design is complete and an automated system is under study. Figure 3 contains a flow diagram of the manual data input procedure. Before discussing the manual procedure, it will be necessary to explain the basic data categories. The first of these is point data such as wells, particular buildings, or communication towers. The second category is line data such as rivers, roads, and railways. The last is area data such as forests, lakes, and towns. The differences in these types of data will make a difference in the manual coding and punching of data.

The manual data input as shown in Figure 2, begins with selection of a map with the appropriate information. A clear overlay grid is placed over the map and grease pencils are used to transfer the data to the grid. A code is selected from the ARIS-I DATA INPUT USERS MANUAL which represents the data being input. This code is the key number discussed in the file design section of this report. For example, if the location of Three Mile Creek is being added to the ARIS data base, then the code is 32. This information, along with a key word signifying that the data input is in the line form, is punched on a control card. Then the UTM coordinates, through which the creek flows, are punched onto data cards in a predefined format. The control card and data cards are coupled with necessary job control cards and fed to the ARIS computer. The correct ARIS program then writes the data onto the ARIS data base. Error correction will be dealt with in future reports as it is perfected.

Automated data input is based upon placing a clear map over the graphic terminal display and encoding data directly into the ARIS computer system. A graphic terminal is necessary for this procedure. Figure 3 outlines the automated procedure along with its error correction ability. Automated data input is superior to the manual form because it is faster and less error prone. The Tektronix graphic display programs will be utilized for experimentation in automated data input but ARIS-I will be based on manual system input.

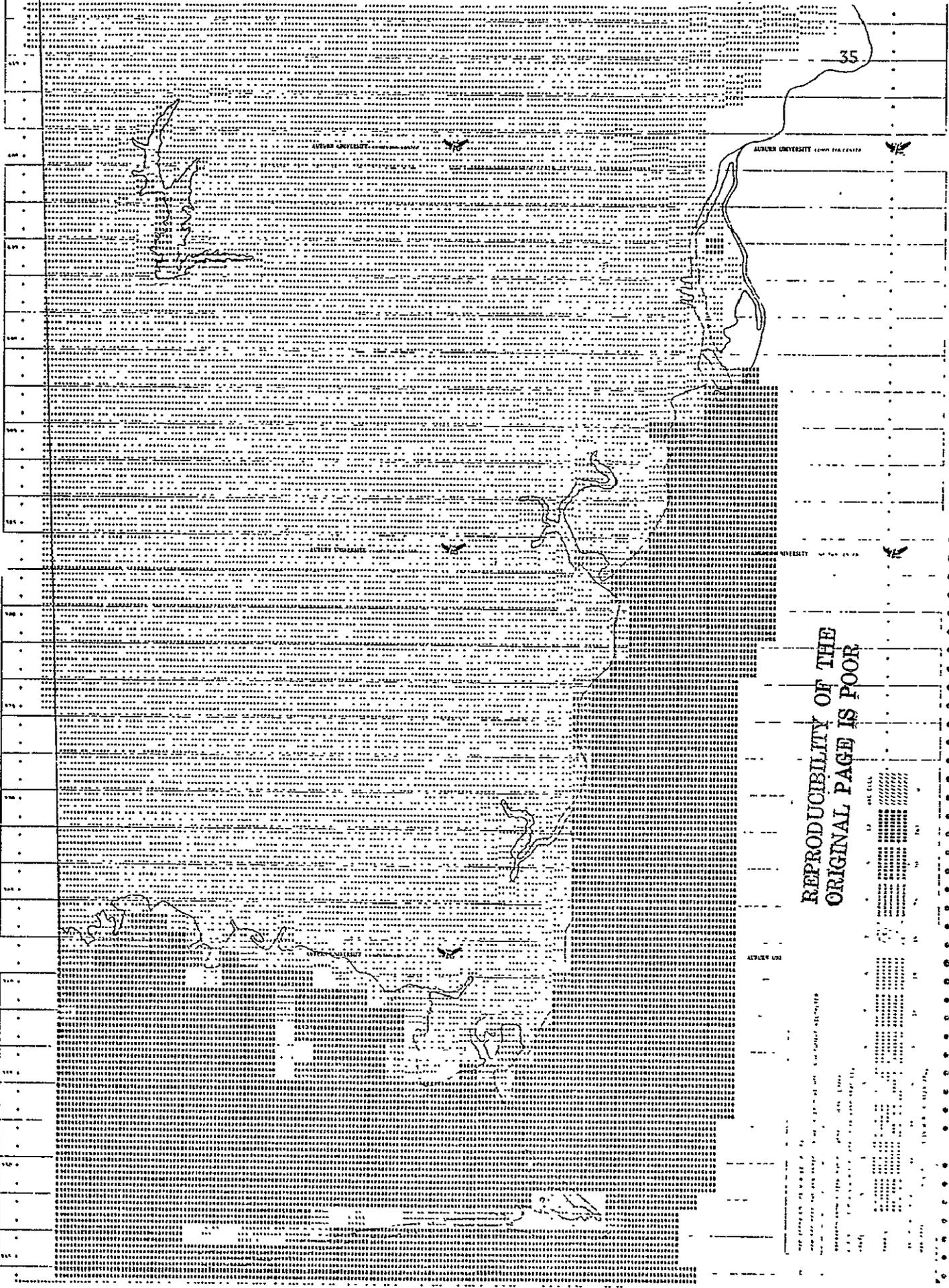
The fully operational ARIS-I system as designed is shown in Figure 4. It incorporates existing software packages, like POPSIM and DUALIST, along with the ARIS retrieval package and special input/output programs into the system. The division of the ARIS-I system into three distinct parts - special input/output packages, basic retrieval package, and canned program packages - allows each component to be developed separately and efficiently. More advanced ARIS systems will compact these three components into one programming system to allow greater freedom of retrieval. Each of the ARIS-I components will now be reviewed in their present design stage.

Figure 5 outlines the ARIS-I retrieval package component. As the user signs onto the system, he is informed of all data available in the ARIS-I data base that he can reference. The initial segment of the procedure, called the SCANNER, converts the external user request to an internal form more acceptable for processing by ARIS. This internal code is then passed to the ARIS program supervisor which calls the appropriate FORTRAN subroutines and executes them as required by the scanner. As more requests are passed into ARIS-I by the user, the supervisor executes more program segments until the necessary data is retrieved. As mentioned earlier in this report, a variable instruction set and data list is desirable. The data files containing a list of the commands and data available to the user, as well as the information required by the internal components of ARIS, is available to the user through the scanner. Work files for temporary data storage or data transfer between ARIS-I components are provided.

The second ARIS component deals with special graphic input/output to the ARIS-I data base. Pre-written program packages, such as SYMAP, CALFORM, and TEKTRONIX GRAPHICS II, will both print output maps and provide for inputting data from maps. This portion of the system has not been fully designed. Plans for the second component include tutor programs to aid the user with the more complicated portions of the graphing programs.

The final ARIS component has not been designed. This component will contain predefined processing packages to supplement the ARIS-I retrieval component. POPSIM and DUALIST are some of the programs already received to form the basis of the third ARIS component. Instructions for the use of these packages will be provided in the system.

Limiting factors for the ARIS-I system are largely machine dependent. Time, core size, and language restrictions will reduce the total capability of ARIS.



REPRODUCIBILITY OF THE
 ORIGINAL PAGE IS POOR

FIGURE 1. Display of South Mobile County and data for 1 kilometer grid squares showing discrepancies between grid data and physical boundaries.

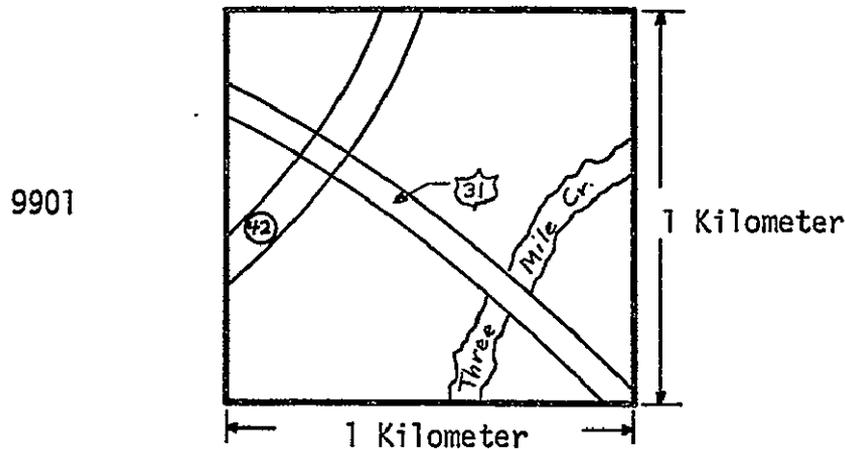


FIGURE 1a. A fictitious grid in Mobile County.

9901	01	97	65	32	00	Further data in grid 9901 such as elevation, population, etc.
Grid Number	Land Use Code	Record Key for US 42	Record Key for Ala 31	Record Key for Three Mile Cr.	Blank Record Key for more Data.	

FIGURE 1b. Main direct access record for fictional grid in Fig. 2a.

97	01	US 42	Further information on Highway 42 such as completion date, tonnage, and number of lanes.
Record Key	Highway Code	Name	
65	01	Ala 31	Further information on Highway 31 such as completion date, tonnage and number of lanes.
Record Key	Highway Code	Name	
32	01	Three Mi. Cr.	Further information on Three Mile Creek such as water flow, ownership.
Record Key	Stream Code	Name	

FIGURE 1c. Secondary direct access records for elements found in Fig. 2b.

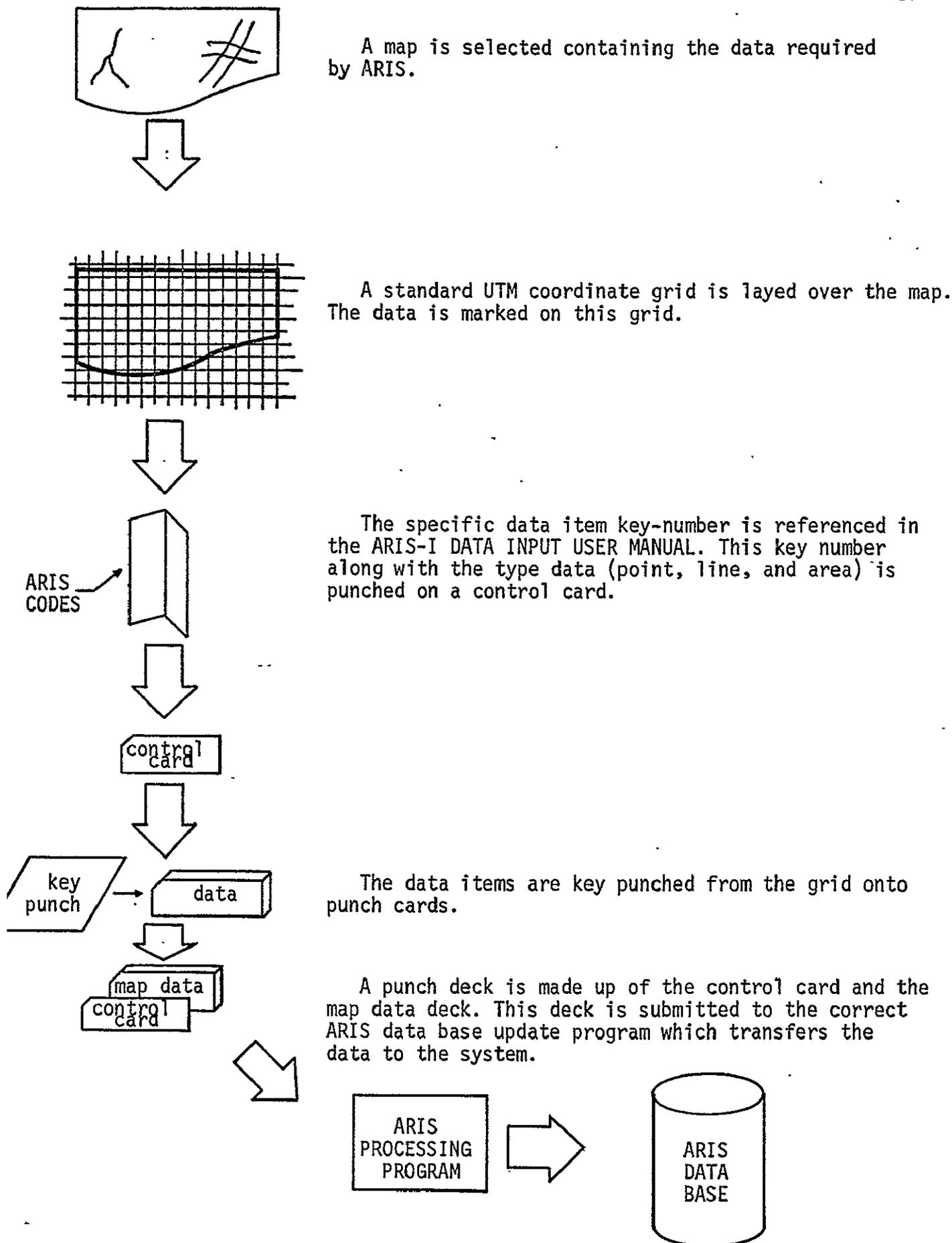


FIGURE 2. Manual data input procedure for ARIS-I.

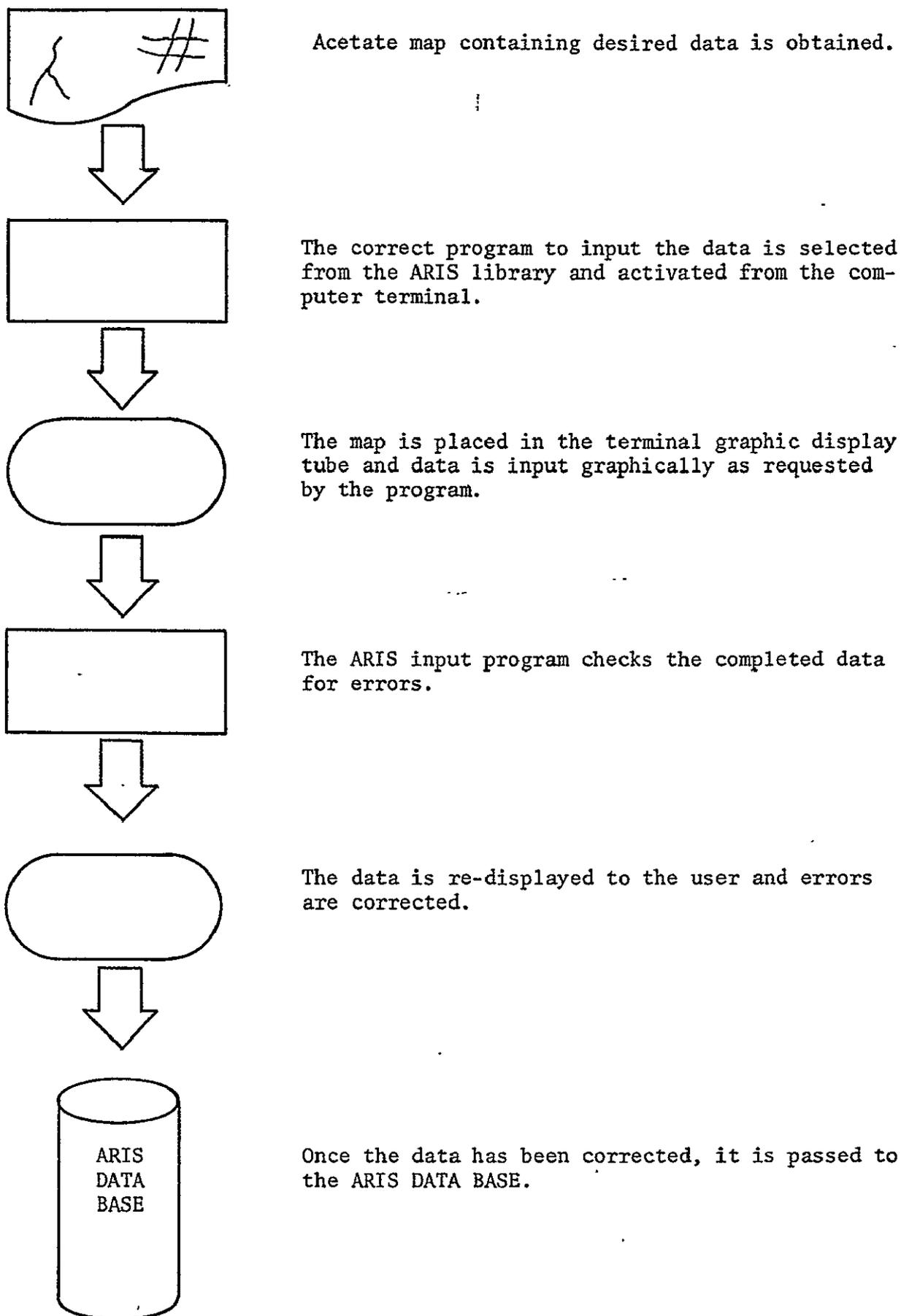
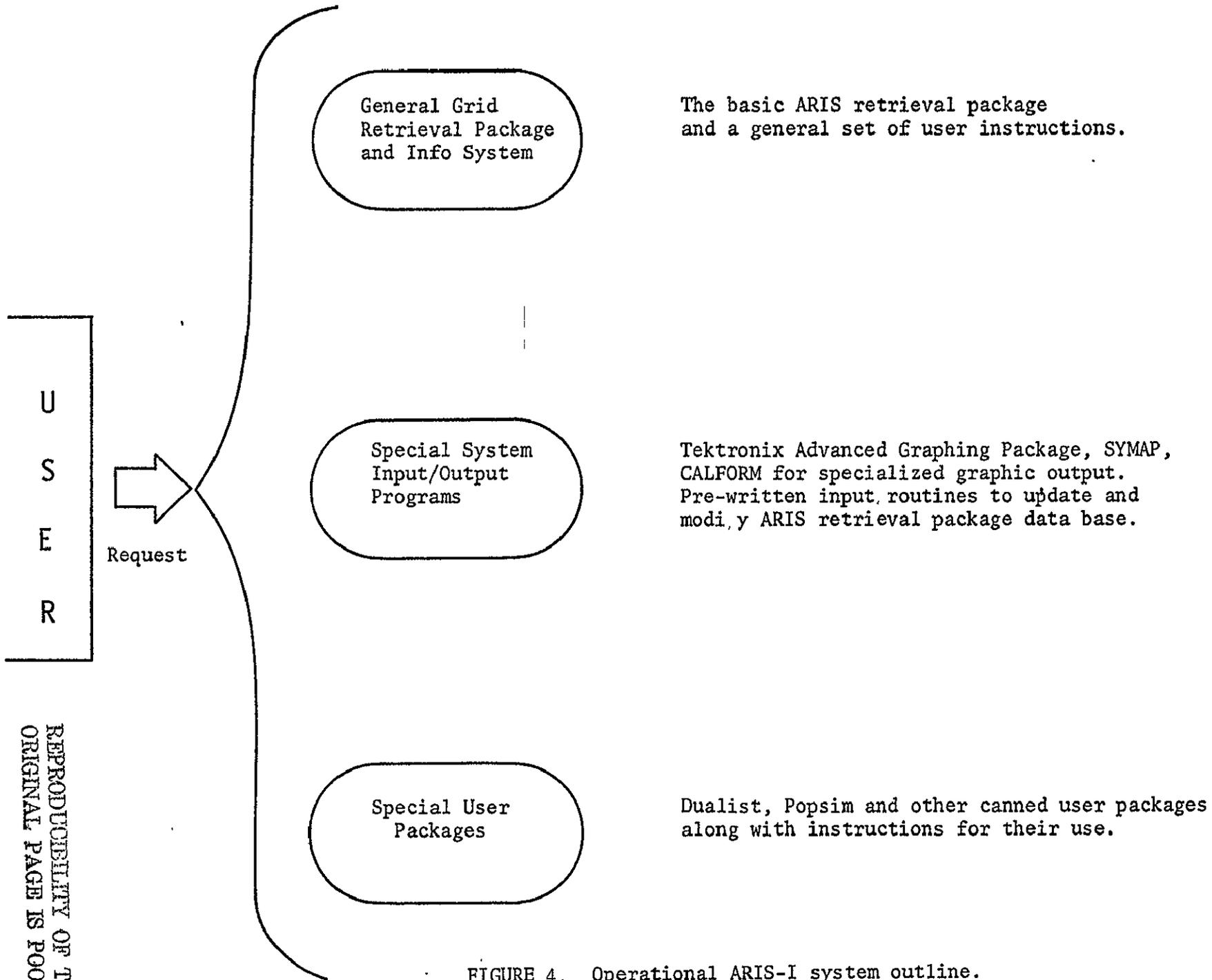
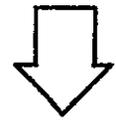
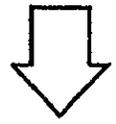
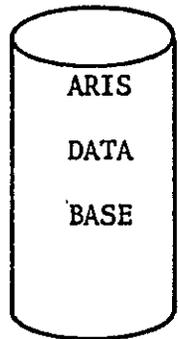


FIGURE 3. Flow diagram of projected procedure for automated data input to ARIS DATA BASE.

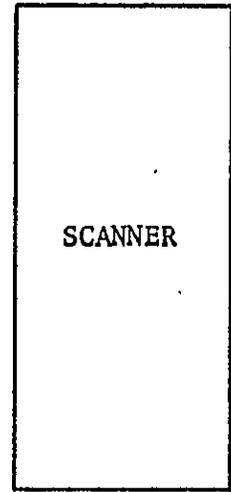


REPRODUCIBILITY OF THE ORIGINAL PAGE IS POOR

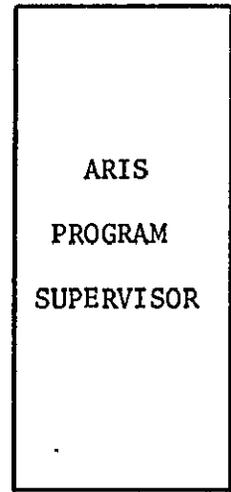
FIGURE 4. Operational ARIS-I system outline. Divides ARIS-I into three separate work areas.



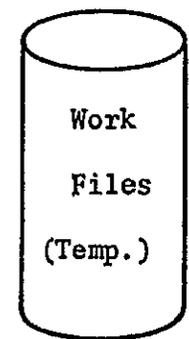
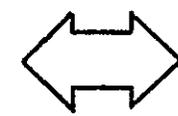
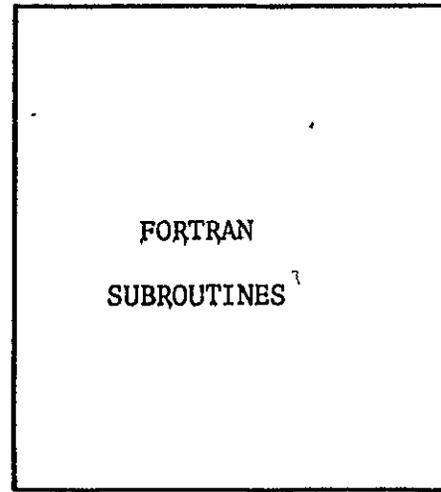
USER
INPUT
REQUEST



INTERNAL
REQUEST
CODES



FORTRAN
SUBROUTINE
CALLS



Scanner interprets input command and searches command and data list for internal codes. Passes these codes to next step or returns error to user.

Program supervisor pursues the internal request generated by scanner into Fortran subroutine calls.

Fortran subroutines execute request passed from program supervisor.

FIGURE 5. ARIS-I retrieval package outline.

APPENDIX D

APPENDIX D

CONTRACT REQUIREMENTS RESULTING IN CENSLIST

a. The Program Functions

The program developed to access the data file must perform the following functions:

- (1) Display in standard tabular form with appropriate geographic codes, titles, row and column headings, the data items in any Tabulation in the data file for any CCD.
- (2) Aggregate across any number of specified CCDs, the data items in any Tabulation. Display should be in standard tabular form and contain the geographic codes of those CCDs comprising the aggregation.
- (3) Search the data file for those CCDs in which each data item listed below, expressed as a percent of the base population, exceed a value assigned by the individual user:

<u>DATA ITEM</u>	<u>BASE POPULATION</u>	<u>CENSUS USER'S GUIDE TABULATION NUMBER</u>
1. 0-18 Years of Age	(All) Persons	17
2. 19-44 " " "	" "	17
3. 45-64 " " "	" "	17
4. 65- " " "	" "	17
5. Different State	Persons 5 Yrs. Old & Over	28
6. Work Outside County of Residence	Persons at Work during the Census Week	35
7. High School, 4 yrs. (or more)	Persons 25 Yrs. Old or Over	42
8. College, 4 Yrs. (or more)	Persons 25 Yrs. Old or Over	42
9. With Vocational Training	Persons 16-64 Yrs. Old with less than 3 yrs. College	43
10. In Labor Force	Persons 16 Yrs. & Older	54
11. In Labor Force Unemployed	Persons 16 Yrs. & Older in Labor Force	54
12. Prof., Tech. & Kindred	Employed Persons 16 Yrs. & Over	58
13. Craftsmen, Foremen & Kindred	Employed Persons 16 Yrs. & Over	58
14. Operatives, Except Transport	Employed Persons 16 Yrs. & Over	58

15. Service, Expt. Pvt. H.H.	Employed Persons 16 Yrs. & Over	58
16. Prof., Tech. & Kindred	Employed Females 16 Yrs. & Over	59
17. Craftsmen, Foremen & Kindred	Employed Females 16 Yrs. & Over	59
18. Operatives, Except Transport	Employed Females 16 Yrs. & Over	59
19. Service, Expt. Pvt. H.H.	Employed Females 16 Yrs. & Over	59
20. Agri., Forestry & Fish.	Employed Persons	62
21. Mining	16 Yrs. & Over, Total	62
22. Construction	16 Yrs. & Over, Total	62
23. Manufacturing	16 Yrs. & Over, Total	62
24. Transportation, Communication & Utilities	16 Yrs. & Over, Total	62
25. Wholesale & Retail Trade	16 Yrs. & Over, Total	62
26. Finance, Insurance & Real Estate	16 Yrs. & Over, Total	62
27. Services	16 Yrs. & Over, Total	62
28. Public Administration	16 Yrs. & Over, Total	62
29. Families with income less than \$3,000	(All) Families	75
30. Families with income less than \$7,000	(All) Families	75
31. <u>Mean</u> Family Income	Aggregate Dollars Family Income/(All) Families	1 & 75
32. Families with income less than poverty level	(All) Families	82

The display will contain the title of the data item searched for, the value assigned, the geographic codes of those CCDs meeting or exceeding the assigned value, and the value of the data item for each CCD.

EXAMPLE:

Persons 65 Years Old and Over as Percent of All Persons CCDs with 8.5% or More

<u>CCD CODE</u>	<u>PERCENT</u>
001 005	8.7
017 020	9.6
023 015	9.5
105 005	8.5

b. Users

It is intended that this system be used by persons with little or no knowledge of computer operations. Thus, the program in final development must be simple to use via remote batch entry terminal. Installation of the product programs on an interactive remote terminal is to be accomplished after ARIS hardware is acquired and under funding of the ARIS contracts.

Programming

- a. PL/I programming language should be used for the source program language and ISAM (Index Sequential Access Method) file structure method be employed in the programming elements of the contract.
- b. Unless NASA immediately provides suitable computer support and connect time, the IBM 370/150 computer at Auburn University will be used for the development, testing, and demonstration of the product programs.

End-Products

End-products of this contract shall consist of:

- a. The data file and program as described U-2 and U-3 on computer tapes and in operable condition as evidenced by: (i) minimum of three runs each of the three functions described in U-3(a)(1), (2) and (3); and (ii) a demonstration of the system at a time to be agreed upon by both parties.
- b. A narrative report detailing the steps taken to complete the project and any particular problems encountered.
- c. A user's manual describing in simple (layman's) language the procedure for using this data retrieval system on a remote terminal.

APPENDIX E

Classify the Variables

There are two basic types of encodable data - quantitative and qualitative. Normally, quantitative data is recorded with the actual quantitative value most symbolic of the cell. For example, topographic elevation would be coded as the actual elevation (estimated in feet or meters) at the center of the cell.

Qualitative data is usually recorded by identification with a representative numerical code. For example, the variable "soils" may be sub-divided into the following sub-variable classes, represented by the indicated code numbers:

Variable - "Soils"

<u>Sub-Variable</u>	<u>Code</u>
Clay	1
Alluvial	2
Rocky Terrace	3
Sloping brown forest soil	4
Flat brown forest soil	5
Bedrock	6

In this example, a given grid cell with clay soil would be recorded as a 1 for that particular cell.

Encode the Variables

It is useful to visualize geographically coded information in three dimensions (x,y, and z), where x represents the row of the data cell, y represents the column of the data cell, and z represents the geographically disposed information within the data cell. In other words, the point (x,y) represents a two-dimensional location on a map surface. The third dimensional indicator, z, represents some geographical characteristic associated with that location. Instead of coding x,y and z coordinates for each cell, a more efficient scheme was devised whereby x and y are inherently defined by their respective row and column order or sequence in the matrix, and the z sub-variable value is recorded according to its respective code value. Remembering that keypunching must follow, the data must be coded in correct order and follow a specific format necessary for the data being coded. The format depends on the number of digits used to represent the data variable. Coding forms used for this purpose may vary according to the desired formats. Two examples are shown as Figure 1, and are used to encode variables represented by one and two digit identifiers, respectively.

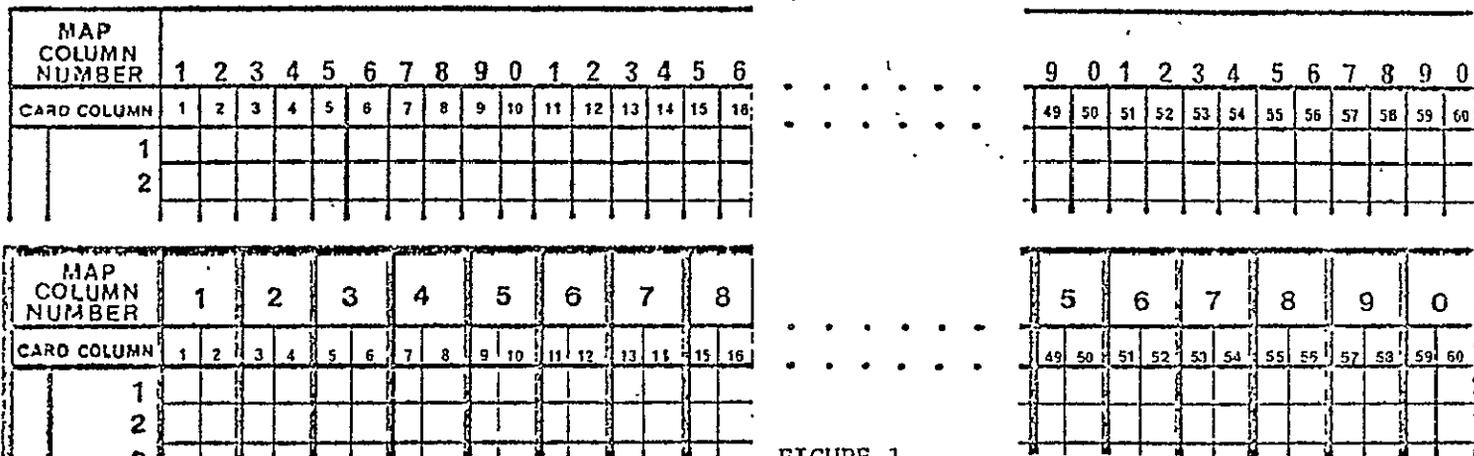


FIGURE 1

The coding forms normally represent 60 card columns of data, and 20 card columns (usually 12 are used) of an identification sequence (to be explained later).

Depending on the format and the size of the study area, each coding form represents all or part of a Basemap Data Matrix (grid). The coding form should be used so the cell representations on the form correspond respectively to the cells on the Basemap. Thus, in Figure 1, if a data variable represented by single digits is being coded, the coding sheet could contain 60 columns and 30 rows of Basemap data. If the Basemap contained a greater number of cells (either rows, columns; or both), additional coding sheets would be necessary. For instance, if a Basemap was 90 cells by 90 cells (with single digit identifiers), Figure 2 illustrates how the coding sheets would be utilized.

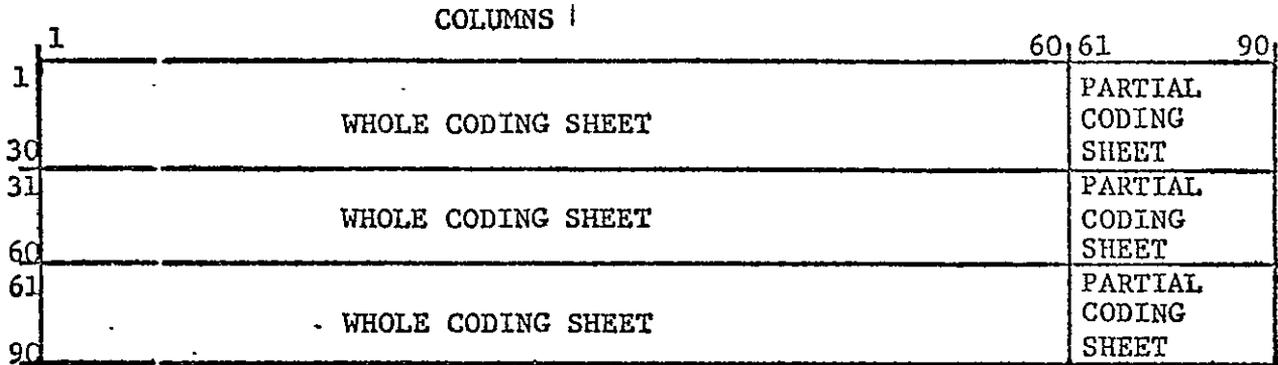


FIGURE 2

If the same 90 by 90 Basemap represented a variable identified with two-digit numeric codes, the same number of rows would fit on a coding sheet, but only 30 columns of Basemap data would fit on the 60 card columns of the coding sheet. Thus, three coding sheets would be required horizontally as well as vertically.

Encoding geographic information from a Basemap begins with the cell in the upper left corner of the rectangular Basemap matrix, or of a partitioned portion of the Basemap matrix, as shown in Figure 2, (i.e., the cell in the 1st column of the 1st row) and proceeds across the row (i.e., the cell in the 2nd column of the 1st row; the cell in the 3rd column of the 1st row, etc.) until all cells in the row of the matrix or of the partitioned portion of the matrix are coded to the extreme right edge of the matrix area or sub-area being coded. The cells in the second row are next coded in the same way, etc. As mentioned in V-C, cells outside the study area but within the matrix are encoded as 0, or left blank. This tells the computer that the cells are "outside" the study area. If the Basemap is large and has been partitioned into sub-areas, as in Figure 2, when the upper left sub-area has been coded and the first coding form is finished, a second coding form is begun either for the vertically or horizontally adjacent sub-area. This process is continued until the entire Basemap is coded. The coding sheet sequential order is arbitrary, but it should be uniform for a project and must be remembered, since after keypunching the cards must be sorted into Basemap row order.

As an example, the study area (outlined in heavy black) shown as Figure 3 has been defined and a grid overlay has been superimposed. If this data file were used to represent the variable "soils", sub-variable (1) represents "clay", sub-variable (2) represents "alluvial", and sub-variable (3) represents "rock terrace" as indicated in "D" of this section.

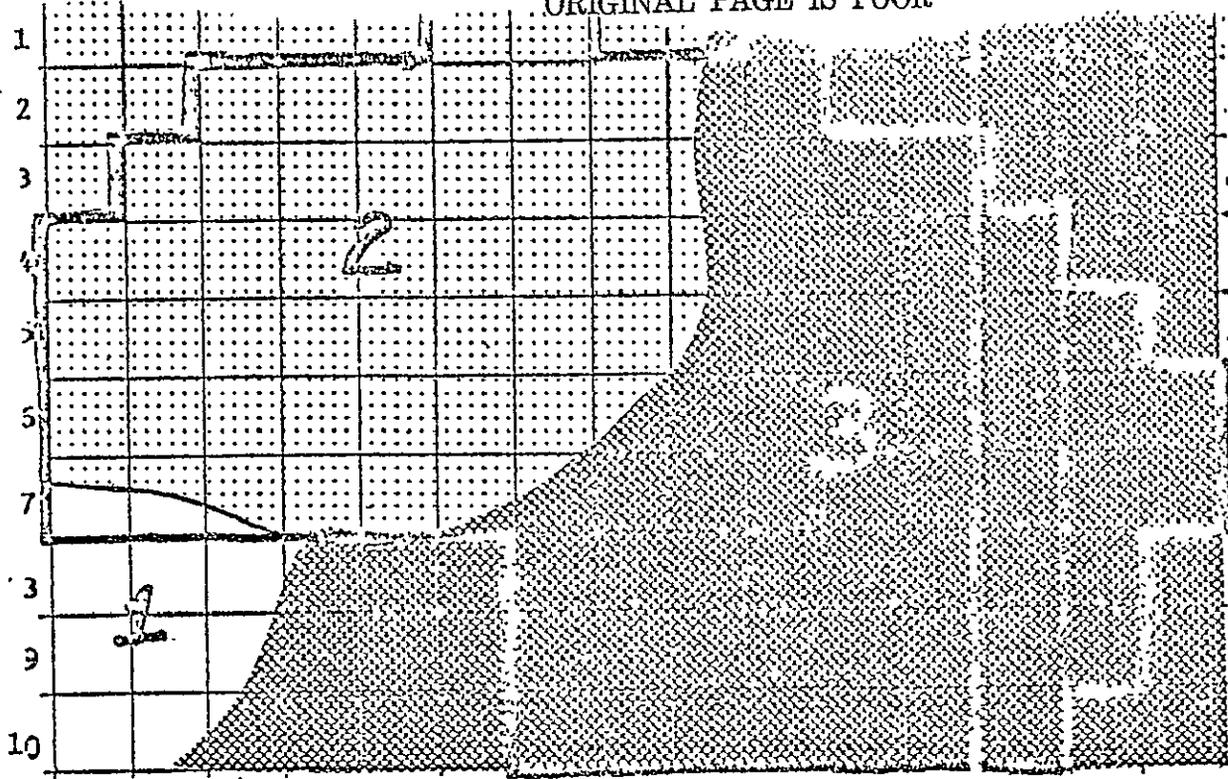


FIGURE 3

Using the single digit GRID encoding form, the first two rows can be encoded as shown in Figure 4. Note that non-study cells in the first row are left blank.

MAP COLUMN NUMBER	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5
CARD COLUMN	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1						2	2			3															
2			2	2	2	2	2	2	3	3															
3		2	2	2	2	2	2	2	3	3	3	3													
4	2	2	2	2	2	2	2	2	3	3	3	3	3												
5	2	2	2	2	2	2	2	2	3	3	3	3	3	3											
6	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3										
7	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3										
8							3	3	3	3	3	3	3	3											
9							3	3	3	3	3	3	3	3											
0							3	3	3	3	3	3	3												
etc.																									

FIGURE 4

Note that in (2,9) (i.e., row 2, column 9) where the cell contains two soil types, the soil type representing the greatest area within the cell has been chosen as that soil type most representative of the entire cell.

In some cases, coding just the dominant characteristic may be inadequate. In other words, one or more sub-dominant characteristics (secondary, tertiary, etc.) may be desirable to retain. The same procedure is followed, only the coding form may be expanded to incorporate a minor change which allows for this. Figure 5 illustrates this regarding the first two rows of the sample Basemap.

F. Keypunching

As mentioned, the encoding format must be consistent to aid keypunching, the format having been determined per variable, depending upon how many digits are necessary to numerically represent the sub-category. Figure 7 illustrates several typical formats.

NUMBER OF DIGITS	FORMAT
1 digit (such as the "1-6" soil types in Section D)	(60F1.0)
2 digits	(30F2.0)
3 digits	(20F3.0)
4 digits	(15F4.0)
5 digits	(12F5.0)
6 digits	(6F10.0)

FIGURE 7

As previously mentioned and illustrated, only the first 60 card column locations are normally used to numerically represent data. Usually, 12 of the remaining 20 columns are reserved for an identification sequence. An example of an identification sequence is shown in Figure 8.

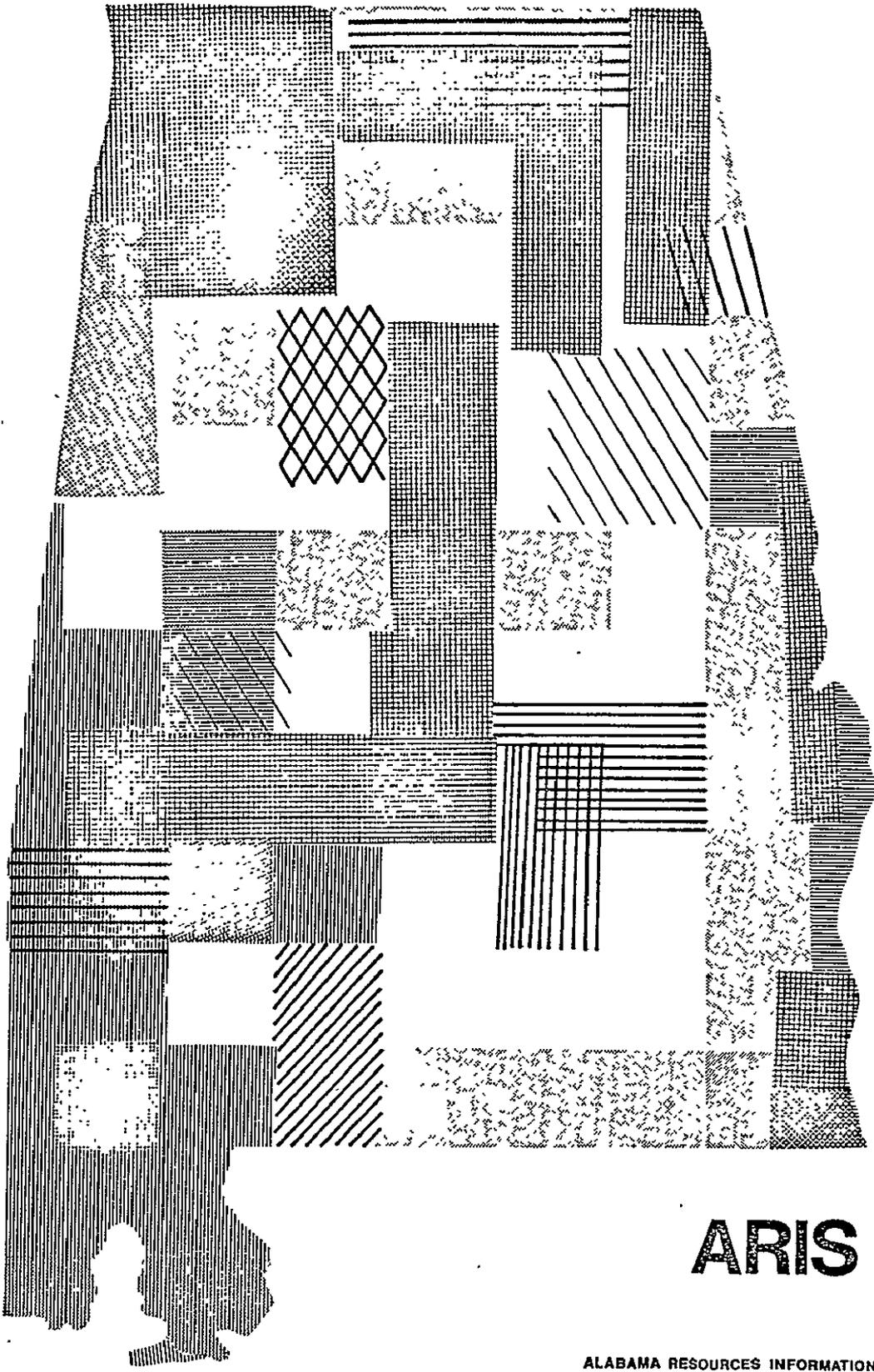
		PROJECT				VARIABLE				CARD NO.						
01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17
X		3	2	1	S	O	I	L		0	0	7	0	2	X	
IDENTIFICATION INFORMATION																

FIGURE 8

It is always wise to include a project identification number, the name of the variable being coded, the number of the row being coded, and, if the row requires more than one card, the number of the card being coded for that row. The identification sequence illustrated is shown as an example only, and may be changed according to the user's specification to help identify any particular project. The identification information is not read and transmitted to the computer; it is for quick reference by the user for editing, sorting should a deck be accidentally dropped, etc.

Notice that column 76 indicates this to be the second card used to transmit geographically coded data for row number 7 (column 74). If single digit numeric identifiers were used (as in Section D) to represent the different "soils" (variable identified in 67-70), the original data grid must here contain more than 60 cells (columns) per row, since 60 data sub-variable identifiers per row can be contained on one keypunched card. This relationship is explained, as previously indicated, through choosing applicable data formats and submitting them to the computer.

APPENDIX F



ARIS

ALABAMA RESOURCES INFORMATION SYSTEM

REPRODUCIBILITY OF THE ORIGINAL PAGE IS POOR **PRELIMINARY**

ARIS

ARIS

ALABAMA RESOURCES INFORMATION SYSTEM

PRELIMINARY

In the short history of automated geographic analysis and computer mapping, there have been a number of programming efforts which have been directed toward analysis and display of spatial data with a computer system. ARIS is a more user-oriented adaptation of the best available marketed software. The advantages of graphics produced by the computer over the products of hand cartography fall into three main categories:

1. Increased manipulation capabilities of geographic data.
2. Faster production of cartographic displays.
3. Realization of a reduced cost per unit of information provided.

Cartographic analysis and displays that will most likely benefit from the use of computers have one or more of the following characteristics:

1. Many maps or analysis outputs for the same geographic area are needed at one time.
2. Maps or analysis for the same geographic area are needed on a recurring basis over a period of time.
3. Uniformity of a map and printout appearance for the display of many types of data is desired.
4. Utilization of highly accurate data is desired, without the necessity of high cartographic resolution on maps.
5. A great number of maps need to be prepared in a short period of time.
6. Complex geographic analysis is required, involving the rescaling of data from a single map or a combination of several data variables for a common geographic area.

This booklet is intended to introduce the content of ARIS as it is presently structured, to indicate the planned evaluation of its capability, and to illustrate the method of interaction with the system by planners and others having a need to make decisions based on geographic alternatives.

The basic ARIS software is operational on a large scale IBM System 370 computer. All geographic and cartographic programs are written in the FORTRAN IV language, thus permitting a certain level of system portability to other large computer installations. Software dealing with census-based population and housing data tabulations and searches has been written in the PL/I language to permit use of the high speed information processing available with the indexed-sequential access method (ISAM) of structuring data files. This software and data structuring scheme is less portable.

GEOGRAPHIC DATA BASE

PRELIMINARY

The primary method in ARIS for acquisition, storage, and utilization of geographic data is through the use of regularly-sized units of area, called grid cells. Important to the way the software functions, is the structure of the data base; the method of data storage for software retrieval and utilization. Areal information, or variables, to be stored, such as topography or soil type, are geographically oriented, while computer storage is sequentially oriented. A system, or scheme, must exist which relates these non-similar orientations. This is accomplished by the use of a locational reference in the data.

Suppose a large geographic area was subdivided into many small sub-areas. Each sub-area could then be more definitively described than the original area, because less variation would be likely to occur. A further reduction in the size of a sub-area would further increase the possibility of unique definition. At some point, say one acre, a specific data category, such as "soil type" within the overall variable "soil" could be identified as that soil type most representative of the soils within the acre. This soil type could be considered to be discretely representative of the variable soil within the smaller areal unit. If many of these areal units were placed side by side, both horizontally and vertically, with each representing a discrete soil type, a large area could be represented regarding the variability of soils. Similar capture of other geographic variables, such as slope, surface water, ground water, etc. (each separately), permits the complete geographic description of an area to be utilized by computer software. This system of contiguous areal units to represent a variable, such as soils, slope, or surface water, is the design basis for the ARIS geographic data system.

REPRODUCIBILITY OF THE
ORIGINAL PAGE IS POOR

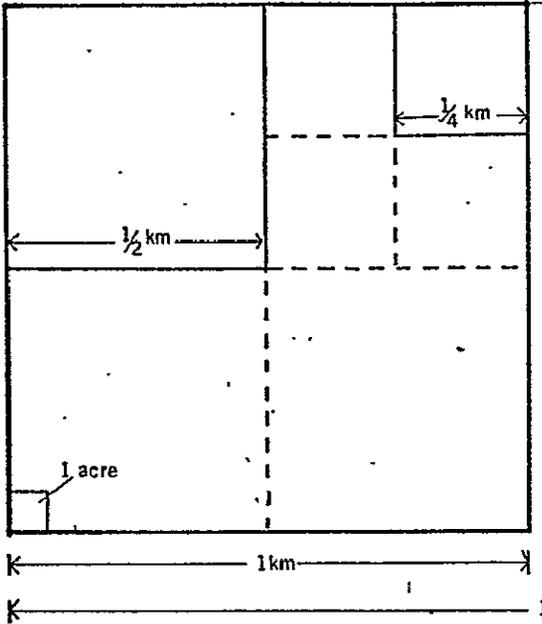
PRELIMINARY

Since the earth has a generally rounded surface and since most cartographic representations of its surface are flat, another mismatch exists. Projections of surface features, boundaries, and their locations to a flat surface receive distortion from the true situation. The Universal Transverse Mercator (UTM) system for locating geographical attributes minimizes projection distortion with respect to the State of Alabama - Alabama is nearly in the center of UTM zone sixteen. For this reason, grid cells based on the UTM system have been chosen for use within ARIS. It is expected that each data base relating to the entire State or used for state-wide analysis will encompass an area of one-fourth of a square kilometer (one half of a kilometer on a side). At the same time, those data bases having regional importance are expected to encompass an area of one-sixteenth of a square kilometer (one quarter of a kilometer on a side). All grid cells are to be related to a single unique extreme northwest point on a UTM grid overlaying the entire State. This will permit all separate data bases to be interrelatable with minimal boundary conflict.

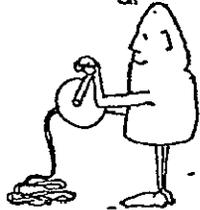
The next two pages provide pictorial reviews of grid cell selection, data capture, and a review of English to Metric conversions.

REVIEW OF METRIC TO ENGLISH CONVERSIONS
AND THE ARIS USE OF THE UTM GRID

PRELIMINARY

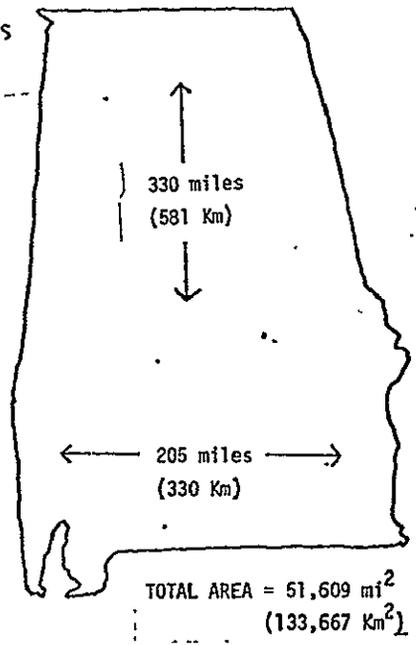


1 Km = 1000m = 3280 ft.
1 Km² = 247 acres
1/2 Km = 500m = 1640 ft.
1/4 Km = 250m = 820 ft.
1 Acre = 4047m²



The Universal Transverse Mercator (UTM) grid was chosen for ARIS. It is based on a one kilometer square grid. These diagrams provide conversions from the one kilometer grid square to the English system of measurement.

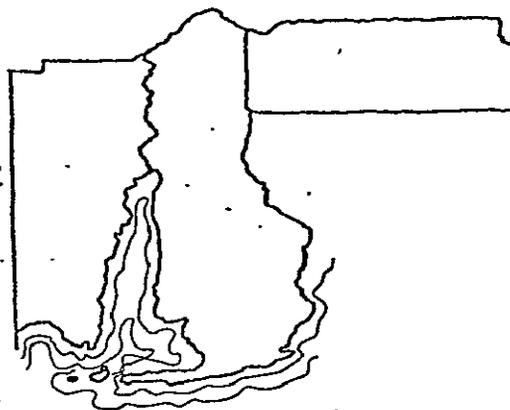
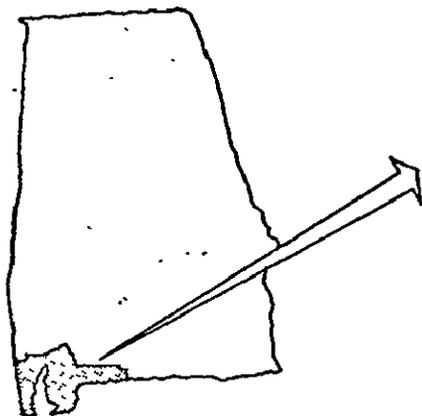
EXTREME MEASUREMENTS OF STATE



GRID CELL CONCEPT OF DATA CAPTURE AND STORAGE

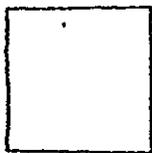
PRELIMINARY

In ARIS, one of the primary techniques for data capture and storage is the GRID system.



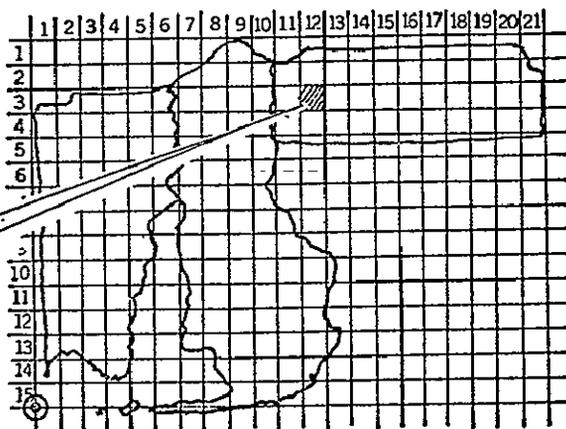
1. A study area is selected and stable base maps containing needed information are obtained.

2. A uniform grid is laid over the study area and fixed to a base point, related to the Alabama base point, so that all future data can be aligned to the same grid within the UTM system.



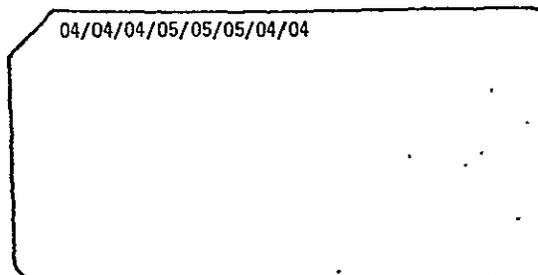
CELL (3,12)

This cell is taken from row 3 and column 12.



REPRODUCIBILITY OF THE
ORIGINAL PAGE IS POOR

3. Codes are recorded indicating the data for a single variable for each cell. Recording begins for the cell in the upper lefthand corner and moves along the row to the right, as one would read. When a row is completed, the process repeats, beginning with the left end of the next row.



- SOIL NUMBERS shown recorded for cells (1,1) through (1,8).

4. Finally, data when stored, is accessed by the software and related by its cellular (both geographic and computational) location. The software can relate the data for a specific cell from several data bases to aid in geographical analysis. For example, to find those locations having a specific soil type, certain slope characteristics, and quantity of surface water.

PRELIMINARY 21

PURPOSE OF ARIS

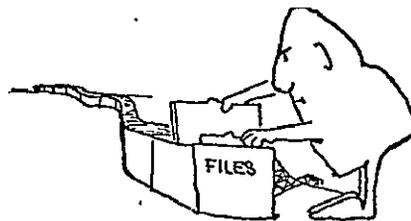
One of the situations faced by planning and other decision-making personnel is that an abundance of existing information dealing with state and local problems may exist but usually is:

1. Unknown or unavailable,
2. In a form that is inconvenient or difficult to utilize, or
3. Too much data from which to select those data items relevant to the problem.

All of these concerns are compounded by the overall problem of data manipulation cost. This is also known as the per unit cost for information provided.



1. Decision-maker does not know a portion of his required information already exists in another department.



2. Then there is always some information in "list" form when the decision-maker needs this information in graphic form or the information is at the wrong scale if it is available on a map.

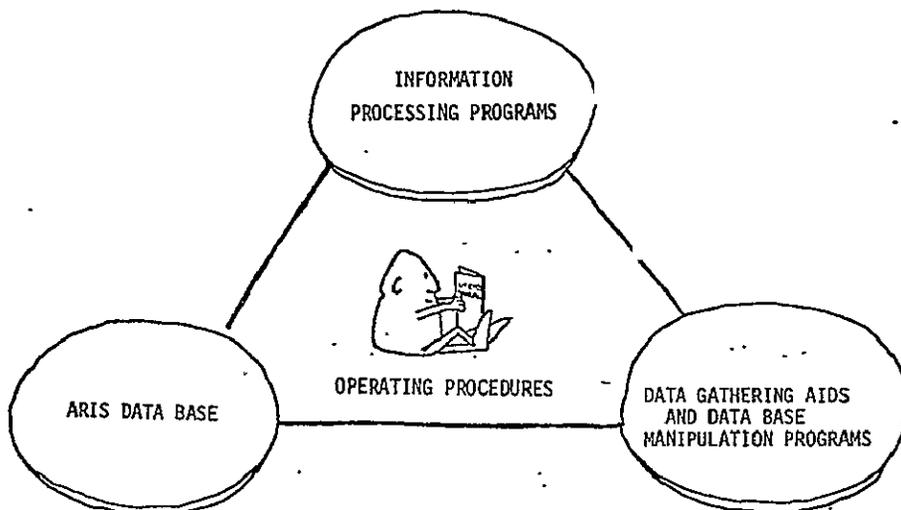


3. Finally, there is the mass of presently available information concerning everything and in which nothing can be found.

PRELIMINARY

REPRODUCIBILITY OF THE ORIGINAL PAGE IS POOR

ARIS is a three-phased approach to reduce or eliminate these difficulties. This can be and is being accomplished by the development of an automated geographic analysis and computer mapping system which requires a minimum of user effort but provides a maximum of assistance to the user. The system is built around a computer accessible, uniformly structured, and self-indexing data base. A decision-maker, using both the ARIS software and a complete description of the resources area (data base), can develop the proper relationships which lead to possible alternative approaches to the problem. Iterations consisting of varied sets of views of the geographic and other resource information, coupled with an analysis of the importance and relationship of the information components can lead to feasible solutions. ARIS is designed as a system comprised of a data base and a set of computer programs which, along with operating procedures, will work with the data to provide information in a form that is valuable to the decision-maker.



PRELIMINARY

ARIS HAS THREE MAIN COMPONENTS:

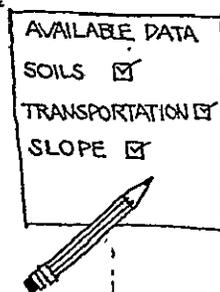
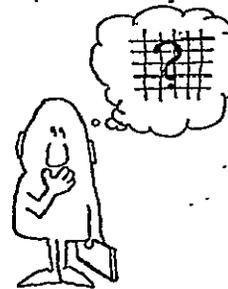
1. Data bases
2. Data gathering aids and data base manipulation programs
3. Information processing programs (These programs are directly responsible for handling user requests for information and/or analysis)

Along with these components are provided operating procedures for using or connecting the components.

Now that the geographic nature and the structure of the ARIS DATA BASE have been established, a simple discussion of how an ARIS user might solve a problem can be introduced.

FLOW DIAGRAM OF ARIS PBOBLEM

1. User identifies a problem that may be solvable, or at least become better understood, by using the ARIS geographic/grid or CENSLIST/HOURLIST search approach.



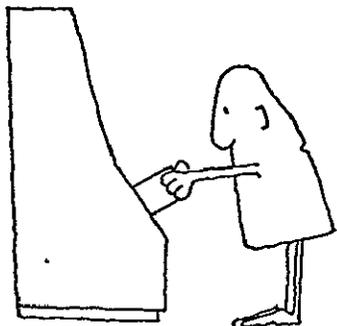
~~PRELIMINARY~~

User lists data to be used as criteria for solving the problem and checks for availability in the ARIS data base.

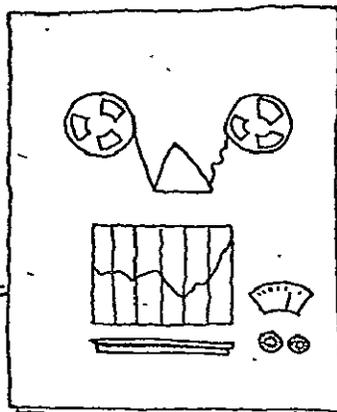
3. User establishes an algorithm (describing the relationship between the involved variables) for obtaining solution to the problem from available data. At this point, the decision-maker may determine: (1) that other variables are pertinent and return to step 2, (2) that necessary data is not available and must be acquired and inserted prior to continuing, (3) that the geographic/grid approach is not suitable and terminate current efforts, or (4) that ARIS supplies a needed tool and continues the process.

- i. check soils file for type 4
- ii. if true, test for available transportation
- iii. search surrounding cells for small slope value

MAP
STOP

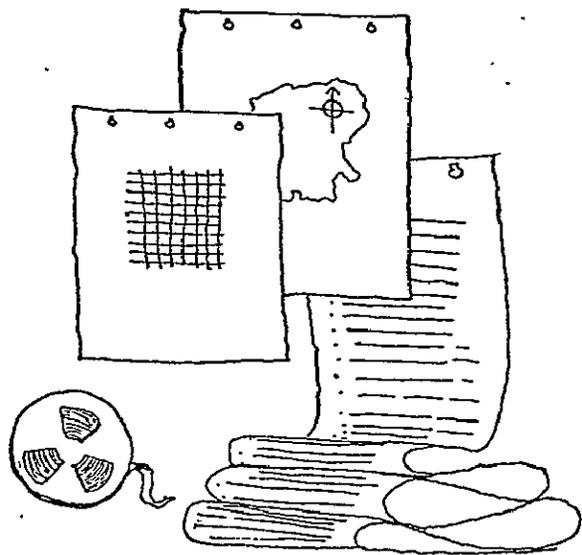


4. User enters the problem into the computer by use of punched cards or an interactive computer terminal.



5. Problem is fed to computer.

Output is developed and returned based on the user's requirements. This may include printed graphics, plotted graphics, and printed lists, or computer files.



ARIS development is divided into three phases:

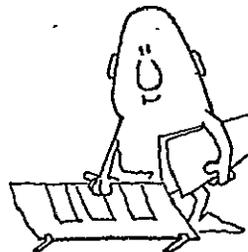
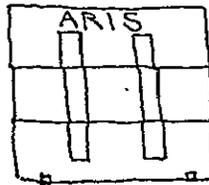
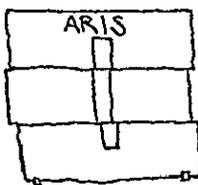
ARIS I - Presently implemented system

ARIS II - Planned additions to complete the basic ARIS system

ARIS III - Future considerations to improve and extend the ARIS capability

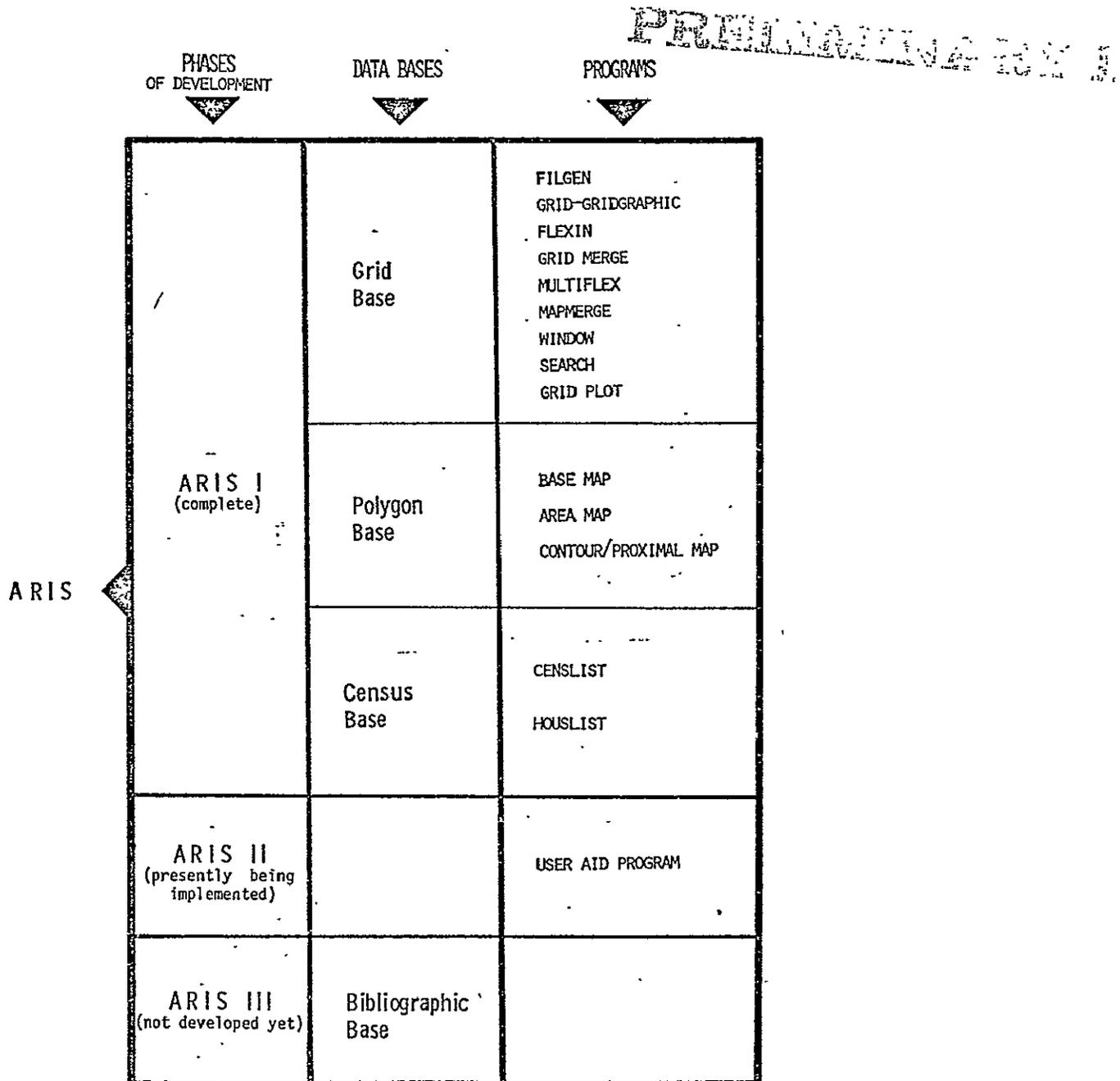
The configuration of each ARIS phase is explained in the following pages.

PRELIMINARY



REPRODUCIBILITY OF THE ORIGINAL PAGE IS POOR

ARIS STRUCTURE AND PHASES OF DEVELOPMENT



ARIS is built by Data Bases and programs. The diagram above indicates the way these are organized and their present state of development according to the phase they correspond to.

ARIS I

	PHASES OF DEVELOPMENT	DATA BASES	PROGRAMS
ARIS	ARIS I (complete)	Grid Base	FILGEN GRID-GRIDGRAPHIC FLEXIN GRID MERGE MULTIFLEX MAPMERGE WINDOW SEARCH GRID PLOT
		Polygon Base	BASE MAP AREA MAP CONTOUR/PROXIMAL MAP
		Census Base	CENSLIST HOUSLIST
	ARIS II (presently being implemented)		USER AID PROGRAM
	ARIS III (not developed yet)	Bibliographic Base	

~~PROHIBITED~~

This phase consists of a set of separate programs which deal with census data or geographic data. Included are: computer programs developed at Auburn to access census population and housing data (CENSLIST/HOUSLIST) and acquired programs which work with GRID or POLYGON data bases (ARIS - GEOGRAPHIC).

REPRODUCIBILITY OF THE ORIGINAL PAGE IS POOR

Grid Base

This type of data base has been previously explained in the "Grid Cell Concept of Data Capture and Storage"- It consists of nine programs:

1. FILGEN - a program for efficiently storing grid cell information from cards or tape into new computer files for re-use by other grid analysis software.

2. GRID GRAPHIC - is used for producing graphic output for geographic areas in the form of printer graphics. The program can also produce frequency plots, level calculations, and a series of scale variations.

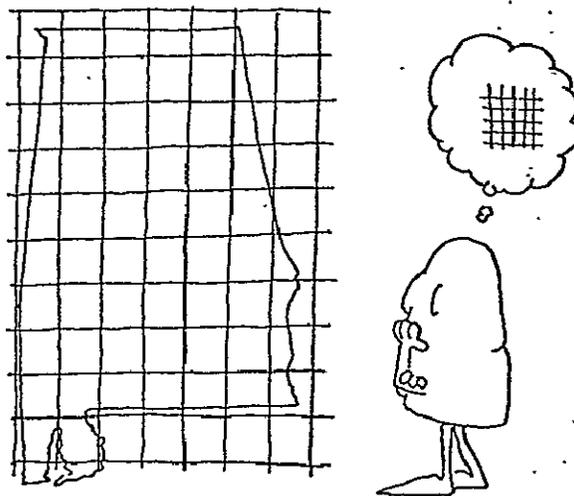
3. FLEXIN - this GRID GRAPHIC option can be used to model data derived from a single variable file (SVF). Simple weighting factors or more sophisticated algorithms can be applied to specific data prior to it being stored or mapped.

4. GRID MERGE - can merge several single variable data files simultaneously or separately into an existing or newly created multivariable data file (MVF). All data to be merged must cover the same geographic area.

5. MULTIFLEX - is used to read, map, and model multi-variable files in the same manner GRID GRAPHIC and FLEXIN are used to operate on single variable files.

6. MAP MERGE - is capable of merging data files for two adjacent (contiguous) map areas if

ARIS I (complete)	Grid Base	FILGEN GRID-GRIDGRAPHIC FLEXIN GRID MERGE MULTIFLEX MAPMERGE WINDOW SEARCH GRID PLOT
	Polygon Base	BASE MAP AREA MAP CONTOUR/PROXIMAL MAP
	Census Base	CENSUS LIST HOUSE LIST



multi-variable data files already exist for these areas. It can also perform like GRID MERGE and merge SVF or MVF to existing multi-variable files.

7. WINDOW LIST - a program to provide numerical printouts of single cells or groupings of cells wherein it is desired to present a profile of environmental and cultural information in a listed form. The program can also be used to isolate a subset data file from a larger data base wherein specialized mapping or analysis might be needed.

8. SEARCH - provides an analysis subroutine to compute geographic distances between various types of phenomena stored in the data base. It can also be used to summarize the occurrence frequency of a given phenomenon within a specified distance from a particular grid cell. This program produces new data files which can be mapped or used in various modeling efforts. For example, for any given grid cell, one could calculate the distance to the nearest highway, commercial land use, park, etc.

9. GRID PLOT - a plotter program that interfaces with GRID GRAPHIC to produce maps for which higher quality graphics is required.

ES&S/ENVIRONMENTAL

REPRODUCIBILITY OF THE
ORIGINAL PAGE IS POOR

Polygon Base

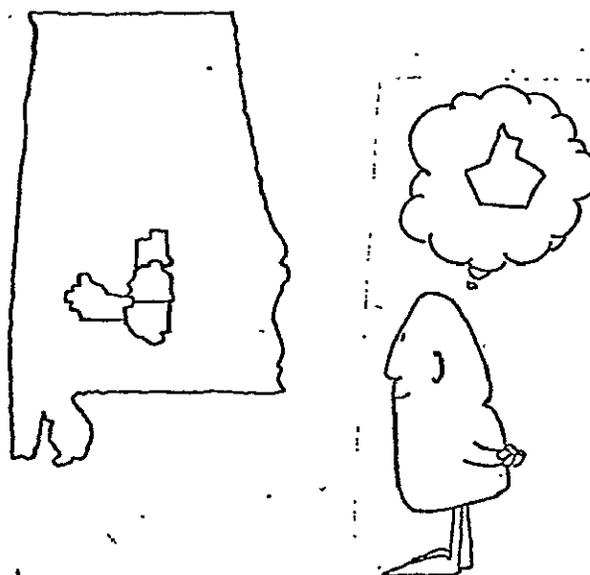
Data can be obtained for irregularly shaped areas formed by the grouping of square grid cells. These groupings can be manipulated in much the same manner as the square grid cells. A polygonal data base can be more accurate than cellular data, but computer manipulation thereof is far more time-consuming and hence, more costly. Programs:

BASEMAP - a program to create the Base Map Image File to be utilized by the next two programs, AREAMAP and CONTOUR MAP. This is the first step when dealing with polygon data.

AREAMAP - is designed to input the Base Map Image File to produce thematic choropleth maps on a standard chain line printer. A variety of available user options permit scaling, shading, legend association and creation of a matrix file, containing the absolute values associated with each polygon cell.

CONTOUR/PROXIMAL MAP - is designed to input the Base Map Image File to produce contour and proximal maps. It also contains special user options to enhance the value of the output information.

ARIS I (complete)	Grid Base	FILSER GRID-GRIDGRAPHIC FLEXH GRID PERSE MULTIFLEX MAP-ERSE MINCON SEARCH GRID PLOT
	Polygon Base	BASE MAP AREA MAP CONTOUR/PROXIMAL MAP
	Census Base	CENSLIST HOUSLIST



Census Base

Census data is available from the system in a list form derived from the Census Bureau data tapes.

The ARIS census programs, CENSLIST/HOURLIST, have been developed to acquire meaningful information from a restructured set of U.S. Bureau of the Census data files. The files have been restructured by eliminating the many blank data elements found in the basic census files and then realigning the data into an Indexed Sequential Access Method (ISAM) data base. The realignment permits very high speed processing to be performed. CENSLIST and HOURLIST prepare tabular output consisting of: (1) mere listings of portions of the data base, (2) aggregations of data from two or more Census County Divisions (CCD), or (3) the result of reaching for those CCD's meeting certain criterion, e.g. CCD's in which fifty percent or more of the population is eighteen years of age or younger.

ARIS I (complete)	Grid Base	FILGEN GRID-GRIDGRAPHIC FLEXIII GRID HEBS MULTIFLEX MAP-LRGE MAPCOR SEARCH GRID PLOT
	Polygon Base	BASE MAP AREA MAP OUTLINE/PROMPTAL MAP
	Census Base	CENSLIST HOURLIST

Table (Blank) County, Census Division 5,
Total Population By Age and Sex

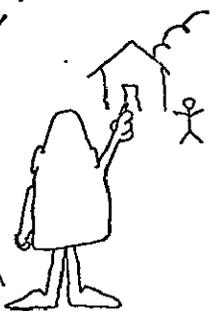
<u>Male</u>	<u>Female</u>	<u>Under 3</u>
109	101	3-4
60	41	5
31	56	6
108	119	7-9
---	---	---



REPRODUCIBILITY OF THE ORIGINAL PAGE IS POOR

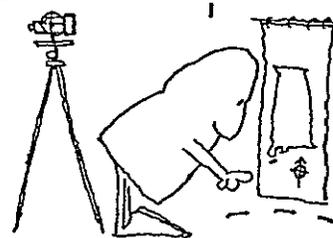
A PICTORIAL DESCRIPTION OF ARIS I USE

ARIS user must first select type of data needed to solve a problem.



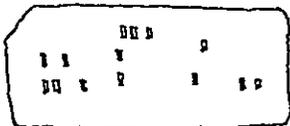
CENSLIST/HOUSLIST

User requires information concerning census population data (CENSLIST) or housing data (HOUSLIST).

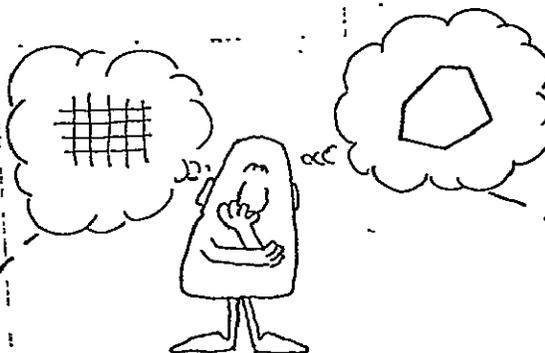


ARIS GRID/POLYGON COMPUTER SOFTWARE

User needs to manipulate geographic data.



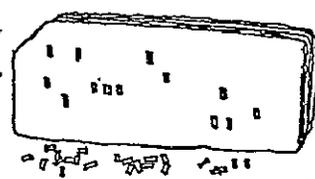
User defines and keypunches a request onto computer cards.



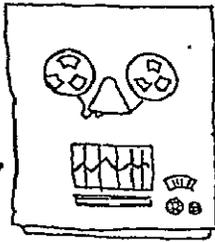
User selects to work with GRID or POLYGON data.

GRID - Cells based on a square metric grid.

POLYGON - Polygon based processing programs

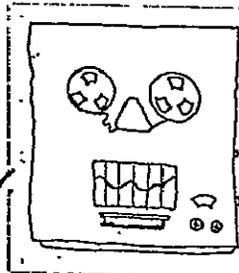


User keypunches request to ARIS GRID/POLYGON programs.



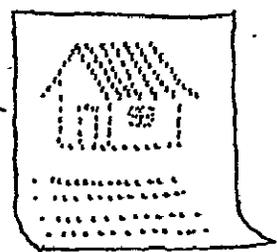
ARIS CENSUS DATABASE

Cards are fed to a computer for processing

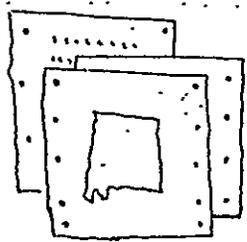


ARIS GRID/POLYGON DATA BASE

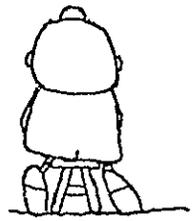
OUTPUT FORMS FOR GRID/POLYGON COMPONENTS OF ARIS I.



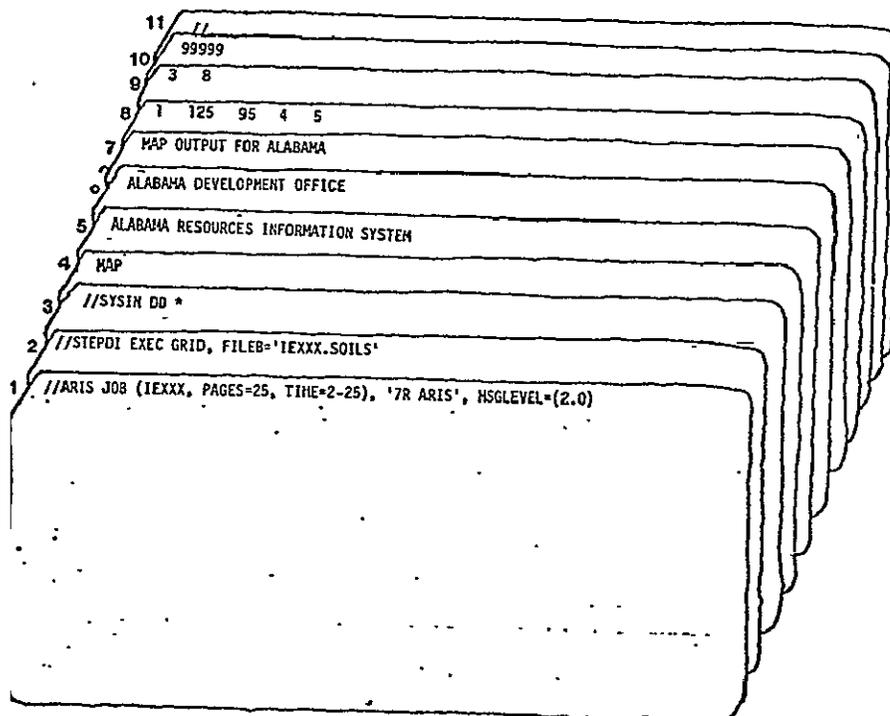
User receives results in the form of computer printed tables.



Printed Computer Output Maps
Calcomp Computer Graphics
Computer Printout



A TYPICAL ARIS I PROGRAM JOB SUBMISSION



The above job submission card deck is typical for use of ARIS I. The user must not only prepare all eleven cards, but have some understanding of the meaning of the material to be included therein. While cards 1, 3, and 11 are relatively typical system job control instructions, the remainder are not. In card 2, the user must specify both the ARIS program to use and the correct cataloged name of the data base. Card 4 is used to specify output type (MAP) and cards 5, 6, and 7 provide a heading or title for the output. Card 8 is used to describe the output. The first item represents the card type (1), the width of the map in grid cells (125), the height of the map in grid cells (95), the height of the cell in printer lines (4 - one half of an inch), and the width of the cell in printer positions (5 - one half of an inch). Card 9 denotes that a mapping function is to be performed (3) which has eight levels of detail. Card 10 serves to signal an end of input data condition to the processor.

ARIS II

	PHASES OF DEVELOPMENT	DATA BASES	PROGRAMS
ARIS	ARIS I (complete)	Grid Base	FILEN GRID-GRIDGRAPHIC FLEXIN GRID MERGE MULTIFLEX MERGE WINDOW SEARCH GRID PLOT
		Polygon Base	BASE MAP AREA MAP CONTOUR/PERCENTAGE MAP
		Census Base	RENSLIST HOUSLIST
	ARIS II (presently being implemented)		USER AID PROGRAM
	ARIS III (not developed yet)	Bibliographic Base	

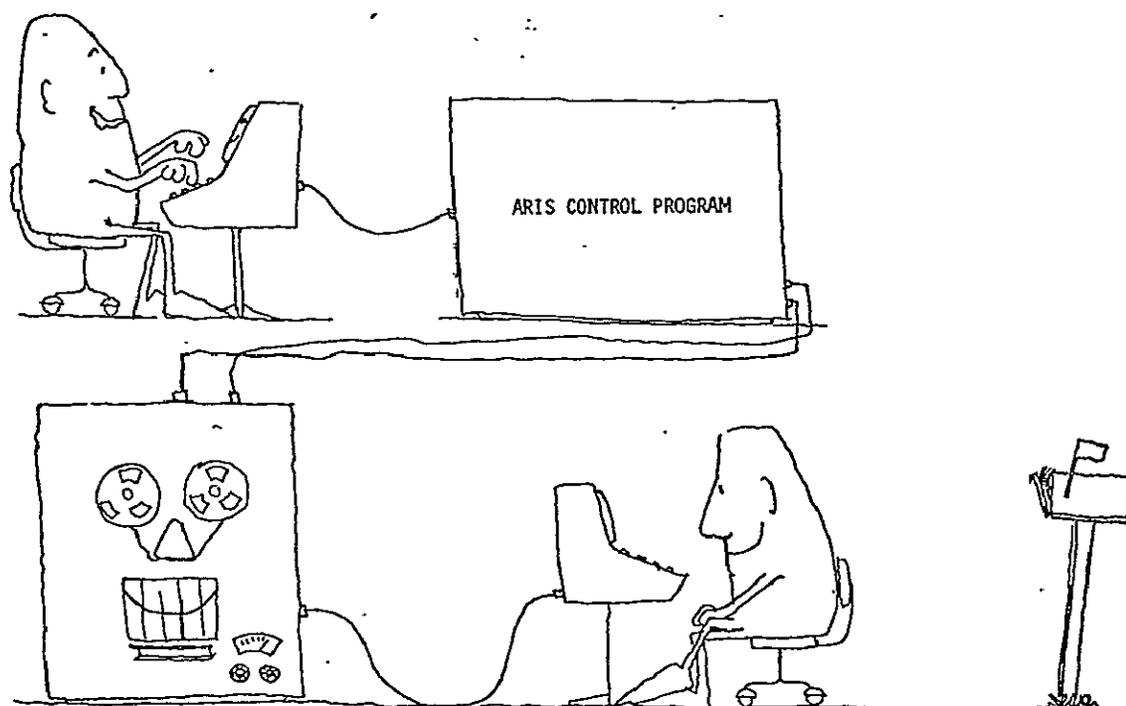
REPRODUCIBILITY OF THE ORIGINAL PAGE IS POOR

ARIS II

USER AID PROGRAM

Phase II of the ARIS system consists of preprocessor software which performs some of the effort expended by the user in ARIS I. The user will merely identify his requirements to a program job supervisor who will take command of the individual program modules in ARIS. This process reduces the amount of effort required by the ARIS user by eliminating problems of job control language and tedious job step instructions.

1. An ARIS user will interact with the preprocessor, via a terminal, to describe the problem to be solved, the algorithm to be used, and the involved data base(s).
2. After completing the dialogue with the preprocessor, the ARIS user can turn control over to this program which will build the input program request.

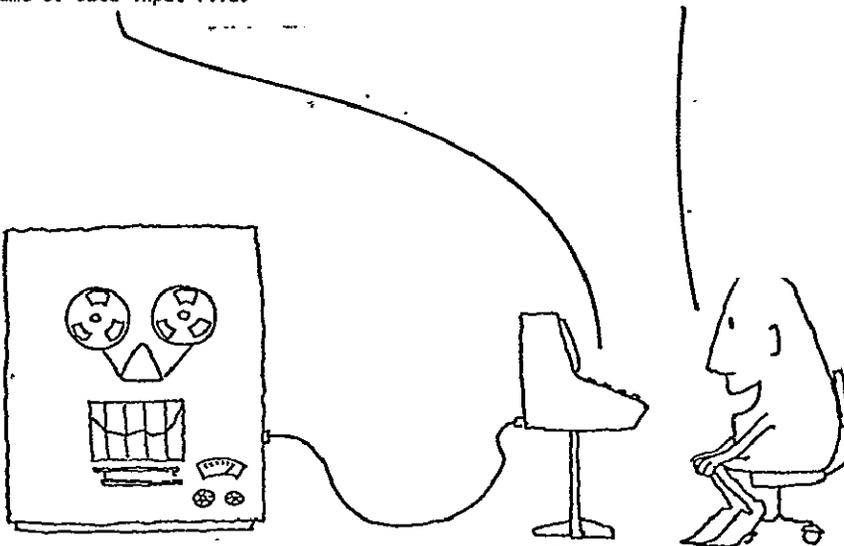


3. A computer is used to process the request formed by the ARIS preprocessor.

4. The ARIS user receives his output at a later time by mail, courier, in person, or in some cases, by way of the computer terminal.

A possible ARIS II dialog between the decision-maker and the preprocessor to a program job submission which leads:

- | ARIS PREPROCESSOR | DECISION-MAKER |
|--|---|
| Program Selection Step | |
| a. Which function do you want to execute? | |
| 1. Censlist 2. Houslist | |
| 3. Map 4. Filgen | |
| 5. Search 6. Gridmerge | |
| b. Title of Map (3 lines)? | b. Alabama Resources Information System
Alabama Development Office
Map Output for Alabama |
| c. Number of columns on map? | c. 125 |
| d. Number of rows on map? | d. 95 |
| e. Map output cell size? | e. 4, 5 |
| f. Any other options? | f. Number of levels = 8 |
| g. Name of data input file? | g. Soils. 75 |



REPRODUCIBILITY OF THE
ORIGINAL PAGE IS POOR

ARIS III

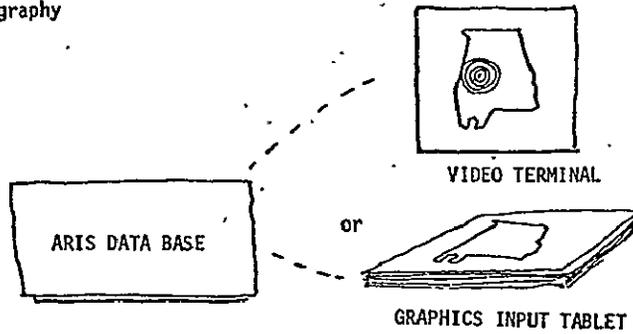
PHASES OF DEVELOPMENT	DATA BASES	PROGRAMS
	Grid Base	FILGEN GRID-CRIDGRAPHIC FLEXION GRID MERGE MULTIFLEX MAPMERGE WINDOW SEARCH GRID PLOT
ARIS I (complete)	Polygon Base	BASE MAP AREA MAP CONTOUR/PROXIMAL MAP
	Census Base	CENSLIST SCHOOLIST
ARIS II (presently being implemented)		USER AID PROGRAM
ARIS III (not developed yet)	Bibliographic Base	

ARIS III Possible future additions to improve system capability

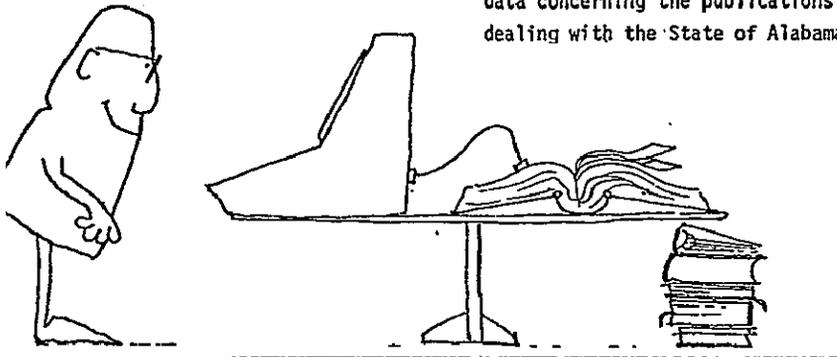
Bibliographic Base

Bibliographic data will include a computer indexed general bibliography for the State of Alabama.

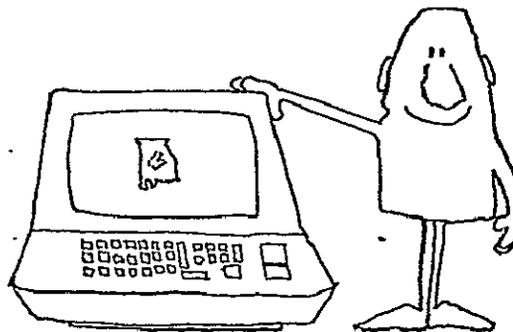
Digitizing data base input through



A series of computer programs for storing and retrieving bibliographic data concerning the publications dealing with the State of Alabama.

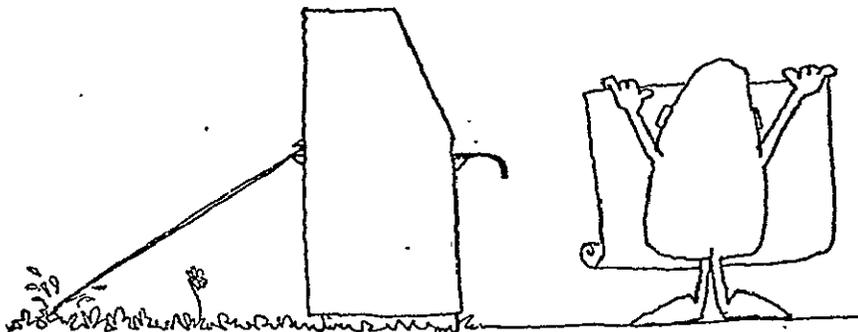


Use of ADVANCED GRAPHING II computer programs to directly store and interpret geographic information about the State. This system component adds greater accuracy for studying small areas of non-grid information.



ARIS has been designed to provide a minimal effort system for collecting, storing, and utilizing geographic information on land resources, economic activity, population and other material which becomes more meaningful and useful when given a spatial dimension.

Its primary task is to retain and operate on geographic data stored in a "grid" form. This "grid" form is based on overlaying a study area with evenly-spaced grid lines to form equal-sized cells. Information about these cells is then recorded in a uniform manner available to the computer to form a data base component.



REPRODUCIBILITY OF THE
ORIGINAL PAGE IS POOR