DOCUMENT RESUME

ED 116 534 HE 006 890

AUTHOR	Thomas, Charles R.
TITLE	Data System Design.
INSTITUTION	Organisation for Economic Cooperation and Development, Paris (France). Centre for Educational Research and Innovation.
PUB DATE	Oct 73
NOTE	26p.; Paper presented at the Professional Seminar on Programme Budgets for University Management and Planning (Paris, France, October 1-5, 1973)
EDRS PRICE	MF-$0.76 HC-$1.95 Plus Postage
DESCRIPTORS	Computers; Conference Reports; Educational Planning; *Electronic Data Processing; *Higher Education; *Information Systems; *Management Systems; Operations Research; Seminars; Systems Analysis; *University Administration

ABSTRACT

Some of the major elements of administrative information systems design as applied to higher education are described. Differences between the application of computer technology in the commercial environment and the educational environment are discussed. The major steps in systems development from problem definition through implementation are defined, including the differences among production, research, and art-form types of systems development. Traditional horizontal and vertical systems integration techniques are presented and the complexity of interrelating operational data systems in higher education are noted. To give the administrator an appreciation of the need for management of the technical side of system development, modular file structures and table-driven software techniques are used as examples. The relatively new technique of data base management systems is also covered. One definition of management information systems is presented, and the need for closely coordinated and supporting operational data systems is stressed. (Author/LBH)

PROGRAMME ON INSTITUTIONAL
MANAGEMENT IN HIGHER EDUCATION

PROGRAMME SUR LA GESTION
DES ÉTABLISSEMENTS D'ENSEIGNEMENT SUPÉRIEUR

DATA SYSTEM DESIGN

CHARLES R. THOMAS

I.

PROFESSIONAL SEMINAR

PROGRAMME BUDGETS FOR UNIVERSITY MANAGEMENT AND PLANNING

PARIS 1ST-5TH OCTOBER 1973

SÉMINAIRE PROFESSIONNEL

BUDGETS DE PROGRAMME, GESTION ET PLANIFICATION DES UNIVERSITÉS

PARIS 1-5 OCTOBRE 1973

CENTRE FOR EDUCATIONAL RESEARCH AND INNOVATION
ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

DATA SYSTEM DESIGN

by

Charles R. Thomas
Executive Director
College and University Systems Exchange
Boulder, Colorado

I

Prepared for:

ORGANIZATION FOR ECONOMIC COOPERATION AND DEVELOPMENT
Program on Institutional Management in Higher Education
Paris, France October 1-5, 1973

DATA SYSTEM DESIGN

The application of computer and systems technology to administration in higher education is generally thought to be quite similar to industrial management. While much of the technology and many of the tools are the same, there are significant differences which make higher education data systems design unique.

Most commercial applications of computer technology involve a reasonably small number of interrelated systems with well-defined relationships between systems. In addition, transaction volumes are measured in hundreds of thousands or millions.

In higher education administration, there are many separate systems which traditionally have not had well-defined working interrelationships. The transaction volumes are also relatively small. For example, an institution of 10,000 students may have forty or fifty different data systems, each with a separate file, with no single file containing more than 100,000 records. This environment is further complicated by the fact that some of the data systems will be automated and others, which must be integrated, will be manual systems.

Combine the above factors with the generally independent nature of the operations of the various segments of a university administrative organization, and data systems design takes on additional dimensions.

SYSTEMS DEVELOPMENT.

From the general standpoint, all application data systems perform three major functions: (1) they collect and edit data; (2) they maintain files; and (3) they produce reports. A batch processing system performs these functions periodically, whenever it is appropriate, on the basis of timing or the accumulation of a sufficient number of transactions. An on-line system performs these same three functions; however, one transaction may constitute a batch and the processing time is immediate. Without going into technical detail of data system design, it should suffice to indicate that a wide variety of techniques are available to accomplish the above three major system functions at varying levels of sophistication.

The generally accepted steps involved in the development of a data system are:

Definition. This first step describes the application, or problem to be solved, in general terms. It may also include the reasons the task is to be considered, and the expected result of the solution.

Design. This step documents the proposed system, including descriptions of the input data and its coding structures, the processing steps, and the output documents. It includes a general narrative of the system flow and shows the interrelationships with other systems. In addition, procedural narratives are written for each of the steps in the design. Also included is a cost-benefit analysis showing the development costs and production costs of the proposed system separately. Both quantitative and subjective benefits should be a part of the cost-benefit analysis and should be documented in advance. Another section of the system design should include a description of the previous process and the costs associated with that operation. Management and user approval should be obtained at an early stage, prior to the expenditure of major resources for detail system design.

Programming. At this stage, computer programs are prepared in accordance with the system design. In administrative data systems, a business oriented language such as "COBOL"[1] should be used, primarily for the self-documenting features. Obscure and super-sophisticated programming techniques should be avoided in administrative systems so as to simplify later modification.

Testing. Testing of an administrative system involves a form of risk analysis. While a complete systems test should be made from data creation through final output, it is generally not possible to test all possible combinations of all possible transactions. The testing process should include a detailed analysis of the processing steps and, if possible, parallel runs with the previous system. In many cases it is not possible to run the old and new systems simultaneously. In these cases, it is possible to create a sample of previous data and run a test of the new system with the sample data, comparing the output to a previous set of known results.

Documentation. Unfortunately, this step is the least completed step in most data processing installations. Good documentation should include a one-page executive summary of the system, user department manuals, system and program documentation and production documentation. Much of the information in good documentation is a by-product of earlier system design steps. In addition to the problem of completing the initial system documentation, most installations generally find it difficult to maintain current documentation of existing systems.

Validation. The validation step in administrative systems insures that the system, as designed, accomplishes the desired results. This step may generate a re-cycle back through the design step for some portion of the system.

Implementation. This step involves the installation and operation of the system on a production basis. Personnel in both the user's office and the data processing office must be trained in the production procedures.

System development for administrative systems is conducted in a "PRODUCTION" mode. There are, however, both "RESEARCH" and "ART-FORM" types of system development. It is interesting to look at the characteristics of the system development steps for production, research, and art-form systems in table form.

CHARACTERISTICS OF SYSTEMS DEVELOPMENT

	PRODUCTION	RESEARCH	ART-FORM
DEFINITION	KNOWN	VAGUE	IRRELEVANT
DESIGN	STRAIGHT-FORWARD	ESSENCE	CREATIVE
PROGRAMMING	PROFESSIONAL	EXPEDIENT	ASTHETIC
TESTING	COMPLETE	AS REQUIRED	DEMON-STRATION
DOCUMENTATION	COMPLETE	SOME	RARE
VALIDATION	EXTENSIVE	SOME	NONE
IMPLEMENTATION	ALWAYS	SOMETIMES	RARE

FIGURE 1

As can be seen from this table, there are considerable differences in approach between the various types of system development. Unfortunately, in many cases, administrative systems are developed in the "RESEARCH" mode or, even worse, in the "ART-FORM" mode. When systems developed in either of the latter modes are implemented in an administrative operation, a considerable amount of difficulty will be encountered, both in the production operation and in attempts to modify the system to meet changing administrative needs.

One key to the development of administrative systems in higher education is a DATA ELEMENT DICTIONARY. Exhibit A is a page from the Data Element Dictionary developed by the National Center for Higher Education Management Systems at the Western Interstate Commission for Higher Education in Boulder, Colorado, U.S.A. While the Data Element Dictionary developed by NCHEMS at WICHE is divided into sections dealing with major areas of data, that is, Student, Staff, Finance, Courses, and Facilities, a specific data element dictionary in an institution will be structured around files. An institutional data element dictionary should be published at two levels: (1) the management level, with general descriptions of data elements and code structures, and (2) the technical level, containing more detailed information necessary for data processing personnel.

Another major key to integrated data system design in institutions of higher education is the development of a uniform code structure for the organizational heirarchy of the institution. The data elements in the uniform code structure are the primary codes used to link files. For example, the financial administrator in an institution must have a code for every organizational unit in the institution. The academic adminisstrator may deal only with those organizational units which offer courses, a sub-set of all organizational units. In many institutions, these two officers maintain a separate code for organizational unit. When it becomes necessary to link financial data with student data, many unnecessary steps must be performed to resolve discrepancies between the two coding structures. Code conversions and other techniques must be used to perform even the simplest form of data matching.

UNIFORM CODE STRUCTURE EXAMPLE

- INSTITUTIONAL ACCOUNT CODE -

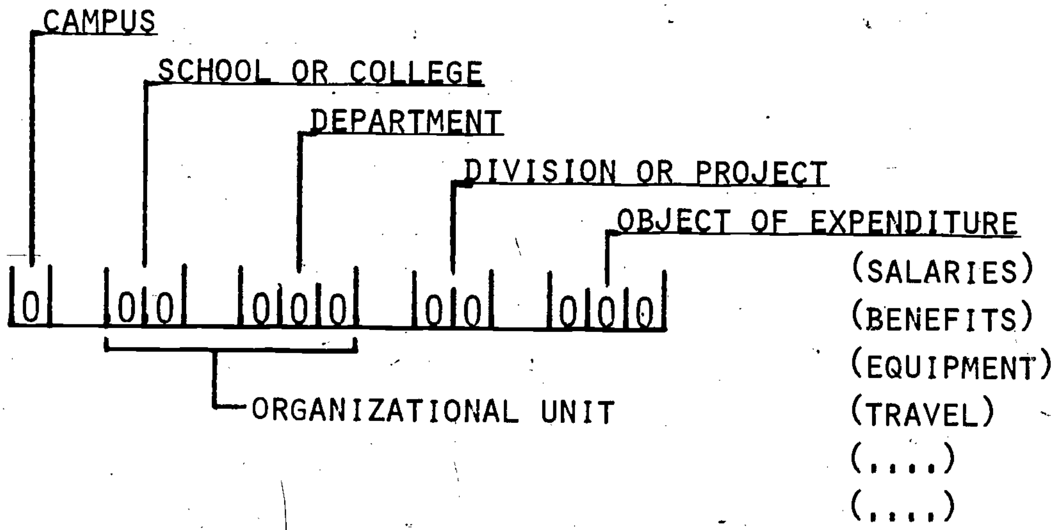


FIGURE 2

Figure 2 shows an example of a uniform code structure for a U.S. institution. Notice that two sub-elements of the institutional account code identify the organizational unit. The sub-elements of "school or college" and "department" codes should also be used in the student records system to record information about courses offered by the organizational units, and in all institutional files identifying organizational units.

Examples of linking codes that should be included in the uniform code structure are:

1. Organizational unit code.
2. Facility code (building and room).
3. Course code.
4. Person identifier (faculty and student).

Both the data element dictionary and the uniform code manual should be published, maintained, and distributed by a non-parochial organizational unit in the administration. This could be an office of management systems, administrative studies, or some other organization unit not having responsibility for any one of the major operational data systems.

SYSTEMS INTEGRATION

Traditionally, operational data systems in institutions of higher education have been designed somewhat independently of each other. The first steps of systems integration usually occurred within one of the major administrative offices and used classical horizontal systems integration techniques. For example, passing information from an admissions system to a student records file. Exhibit B illustrates the classical horizontal systems integration technique.

As the demands for more statistical and management information increase, it becomes necessary to develop vertical integration of systems. In this process, the day-to-day operating systems produce summary information which supports operating and policy management decisions. Exhibit C depicts the classical vertical systems integration technique. Exhibits B and C illustrate the commercial model of horizontal and vertical systems integration. While both of these models apply to some degree, the data systems design problem is not quite so simple in an institution of higher education.

Exhibit D shows the network of major operational data systems in an institution of higher education. Also shown on this network, by the connecting lines, are the major relationships between the operational data systems. In addition to these major relationships, there are many additional minor connecting points between systems.

In general, the systems shown on the left of the diagram in Exhibit D are those related to the student area. The systems shown on the right of the diagram in Exhibit D are the fiscal systems. The four systems in the center of the diagram are those that support the management structure. The data element dictionary and uniform code systems were discussed earlier. Many of the products of NCHEMS at WICHE fall into the boxes labelled "INSTITUTIONAL ANALYSIS AND PLANNING" and "MANAGEMENT SYSTEMS". Notice that the "INSTITUTIONAL ANALYSIS AND PLANNING" and the "MANAGEMENT SYSTEMS" are linked to all of the operational data systems.

In an article entitled, "Data Management and Interrelated Data Systems for Higher Education," Mr. John Chaney stated:

> "Two very important concepts underlie the interrelating of data systems to support institutional planning and management: (1) institutional analysis files are based on operating data systems; and (2) operating data bases are linked into a network by a predetermined set of uniformly coded data elements."[2]

If planning and management systems are to reflect reality on a regular basis, they must be driven by data from the operating data systems produced as a by-product of day-to-day operations. Such an information systems network should be designed by each institution before embarking upon data systems design for any of the sub-systems.

It may not be possible to design and implement all of the systems in the network. The designers and managers should be cognizant of the fact that the network exists, and design the subsystems in such a way as to fit with the future sub-systems that may be designed and implemented at a later date.

SYSTEMS DESIGN

In the past, design analysts have insisted that the specifications for a system be cast into concrete before starting on the system. Should a change in the specifications occur, serious delays and major expenditures would be predicted by the design analyst. As the state-of-the-art in system design progressed, a number of techniques were developed which allow the preparation of more generalized software to accomodate changing conditions when they occur, rather than if they occur.

To illustrate some of the techniques of systems design that provide more flexibility, modular files and table-driven software will be discussed. There are, of course, many similar techniques available to the professional data systems designer.

Conventional systems design will normally place most of the data elements for a file into a single record, since current software will handle very large records. By grouping data elements into separate modules, dependent upon some grouping criteria, it is possible to develop a system of programs which will operate on specific modules of a file, ignoring all other modules. Such a system design then allows the introduction of new modules into the file without disturbing the operation of previous programs.

The concept of modular file design is quite simple and many times is disdained by professional technicians. The result of a non-modular file design is the requirement that all programs for a system be changed if the file design changes. When such a change occurs, not only must all programs be changed, but the entire system must be retested at considerable expense.

Table-driven software is another generalized software design technique which can provide for change at minimal cost. An illustration of this technique is the use of "tax-tables" in a payroll calculation program. Most programmers will include tax-tables as a set of constants inside the source program, a technique that requires a considerable amount of effort and testing should the tax-tables change. A more professional program designer will require that the program read the tax-tables as an external set of data, eliminating any programming problems should changes be required in the tax-tables.

Both the modular file design and table-driven software techniques are relatively simple, in fact, so simple that many technicians consider them to be unsophisticated and not worthy of their technical expertise. In such situations, it is sometimes necessary to apply strict management supervision of technical staff to achieve the desired flexibility.

A recently developed software technique referred to as "Data Base Management Systems" promises to provide great help to systems designers and users. The CAUSE Data Base Management Systems Project defines a Data Base Management System (DBMS) as "A generalized software package which is user-oriented, and can perform tasks related to file definition, creation and update, and reporting/query. Generic names labelling such systems are: data management systems, generalized information retrieval systems, information management systems, file management systems, and information network and executive systems."[3]

The data base management system provides the system designer with a single set of routines that can be utilized for the major functions of all applications without having to write separate programs for each function for each system.

A common characteristic of data base management systems is that they are dictionary-driven; that is, user reference to data stored with the system is symbolic, and a dictionary of the symbols is provided to each user. Internally and transparent to the user, then, are the necessary technical routines required to store and manipulate the data inside the computer.

Perhaps the capabilities and complexities of a data base management system can best be explained by examining the emerging position of the "Data Base Administrator."[4]

The functions and responsibilities of a data base administrator vary widely; however, a consensus of the major functions provide insight to the use of this new technique.

The first function of the data base administrator is the maintenance of the data base. Traditional batch processing applications place files onto unique storage media, such as a magnetic tape or a magnetic disk. When a particular application is not active, the files can be stored in a vault, office, or cabinet, physically separate from other operations. In the data base management system environment logical files are not physically separated onto unique storage devices. Typically, many files are combined and share the same physical device in order to support the interrelating of files as described earlier. Technical considerations require that files be reorganized frequently, involving physical movement within the computer. Normally such reorganization is totally transparent to the user organization. The data base administrator controls all of the activities associated with maintaining the data base inside the computer storage devices.

In addition to the user data files, the data base administrator must also maintain many files used by the data base management system; for example, the data element dictionary and a directory of all the files. Much technical information is necessary for the data base management system to isolate the user from the physical computer processing.

It should be obvious that the same sophistication that makes a data base management system very flexible and powerful can also create an environment in which major catastrophes can readily occur if proper control is not exercised.

Probably the most widely recognized duty of the data base administrator is the control of the security aspects of the system. Normally users will be assigned passwords or passkeys or other methods of defining each user's authorization for access to data elements and files. In the batch processing environment, user access to files is controlled by individually documented security procedures. For example, personnel in the student records processing area seldom have access to the payroll files. With all information stored by and made available by the same set of routines, there must be some method of describing and controlling which users may work with which files. The data base administrator is normally responsible for developing and maintaining the policies and procedures governing the security of the data within the data base management system.

Each user is naturally concerned about the integrity of his data, since he is the "owner" of that particular file. The stability of the data base management system which serves each user is critical, particular in on-line systems allowing terminal access to the files. Procedures for backup and recovery from disasters must be carefully established and operated. Systems that log activities and maintain audit trails for examination must be maintained for all of the files within the purview of the data base management system.

The uniform code structure mentioned earlier is also related to data integrity. Inconsistencies between files within the institution become more critical. For example, frustrations will occur if the fiscal officer utilizes student record data and is unaware that the registrar has used a different code for organizational units. One of the major difficulties in developing a set of uniform codes is that each user is quite willing to compromise, so long as everyone uses <u>his</u> code as the uniform code.

<u>Quality control</u> is another major function which concerns the data base administrator. Certainly most administrators are familiar with the old computer acronym, "GIGO," (garbage in, garbage out). Unfortunately, the production of information by a computer sometimes lends an aura of accuracy that may be totally unwarranted. In addition to the quality of data, the data base administrator must also be concerned with the quality of the operation of the system, so that users will not become distrustful of the computer operation.

Because most data base management systems support many users simultaneously, sometimes in an on-line environment, the data base administrator must constantly be aware of the level of <u>performance</u> of the system. He is normally concerned about providing a balanced set of capabilities to all users and must take corrective action should one application suddenly begin using all of the resources of the system, resulting in degradation of the operation for other users.

The major functions of the data base administrator provide an in-depth look at the discipline required by the use of a data base management system. Surprisingly, many of these same functions are found in successful installations utilizing traditional batch processing application techniques. It is interesting to note that the relatively new technique of data base management systems absolutely requires a set of disciplines that are found to be quite helpful to traditional data processing operations

Although a great amount of material has been written at the technical level, there are few publications which adequately describe data base management systems to the uninitiated. The College and University Systems Exchange Data Base Management Systems Project will be publishing, in December 1973, such a document for beginning data base management systems users.

MANAGEMENT SYSTEMS

It is only in recent history that the tools of management science have been applied to the administration of higher education. In many institutions, implementation of simulation modelling, or some other management science technique, has focused major attention on the operational data systems support for management systems. The National Center for Higher Education Management Systems at the Western Interstate Commission for Higher Education is currently stimulating a great amount of interest in management information systems in higher education in the United States.

Walter J. Kennevan provided a very good definition of a management information system:

"A management information system is an organized method of providing past, present, and projection information related to internal operations and external intelligence. It supports the planning, control, and operational functions of an organization by furnishing uniform information in the proper time frame to assist the decision process.[5]

The key words in Mr. Kennevan's definition are "organized method," "related to internal operations," "uniform information," "proper time frame," and "assist the decision process." To relate those key words to the previous discussion of data systems design, the following points are made.

The "organized method" can be related to the development of a network of operational data systems for the institution before embarking upon the design of any single data system.

"Related to internal operations" and "uniform information" are key words specifically relevant to the earlier quote by Mr. John Chaney.

The "proper time frame," of course, is a very relevant factor in any data system design. I am sure many users of data systems have received voluminous computer printouts of information one week after major decisions were made without the benefit of information.

Finally, all of the management information possible cannot make a better decision. Organized information, however, can make a better decision possible.

Every operational data system should be designed with two major goals. One, to support the day-to-day operational process; and two, to provide data and support for the management systems of the institution.

SUMMARY

The purpose of this presentation has been to acquaint the administrator with some of the major elements of administrative information systems design as applied to higher education. Differences between the application of computer technology in the commercial enviroment and the educational environment were discussed. The major steps in systems development from problem definition through implementation were defined, including the differences between production, research, and art-form types of system development. Traditional horizontal and vertical systems integration techniques were presented and the complexity of interrelating operational data systems in higher education were noted.

To give the administrator an appreciation of the need for management of the technical side of system development, modular file structures and table-driven software techniques were used as examples. The relatively new technique of data base management systems was also covered. One definition of management information systems was presented and the need for closely coordinated and supporting operational data systems was stressed.

Greater detail could be developed within each of the sections of this presentation. Hopefully, this overview will prompt administrators to take an active interest in the administrative data systems designed for their own institution.

EXHIBIT A

006
ELEMENT NUMBER

National Center for Higher Education Management Systems

DATA ELEMENT DICTIONARY
Finance Related Data Elements

ELEMENT TITLE Object Classification

DEFINITION

The classification of expenditures according to the nature of the cost incurred.

CODES OR RECORDING INSTRUCTIONS

The Cost Finding Principles Project currently is considering the following categories:

Code	Category	Definition
10	Faculty Salaries	Compensation including fringe benefits to those individuals that the institution considers its faculty (part-time and full-time faculty as well as graduate assistants).
20	Nonfaculty Salaries	Compensation including fringe benefits paid to all employees of the institution except those considered faculty.
30	Supplies and Services	All current operating expenditures other than compensation for personal services (including fringe benefits), expenditures for capital equipment, and stipends.
40	Capital Equipment	Those items of property that have an acquisition cost of $200 or more and an expected service life that exceeds one year.
50	Stipends	Financial assistance awarded to students (both undergraduate and graduate). Includes scholarships, fellowships, and traineeships. Recipients of stipends are not required to render service to the institution as a consideration of their awards, nor are they required to repay them.

USES CFP, IEP, PM, HEFM, RRPM, FAA

COMMENTS

Institutions will usually have more detailed categories within the above groupings.

DATE ISSUED September 1972

EXHIBIT B

HORIZONTAL SYSTEMS INTEGRATION

* * * * * * * * * * * * * * * * *

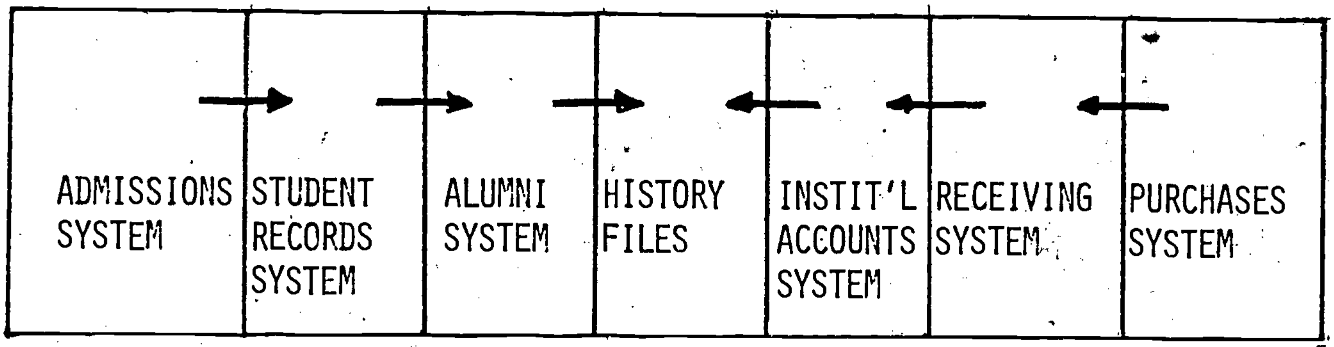


EXHIBIT C

HORIZONTAL AND VERTICAL SYSTEMS INTEGRATION

* *

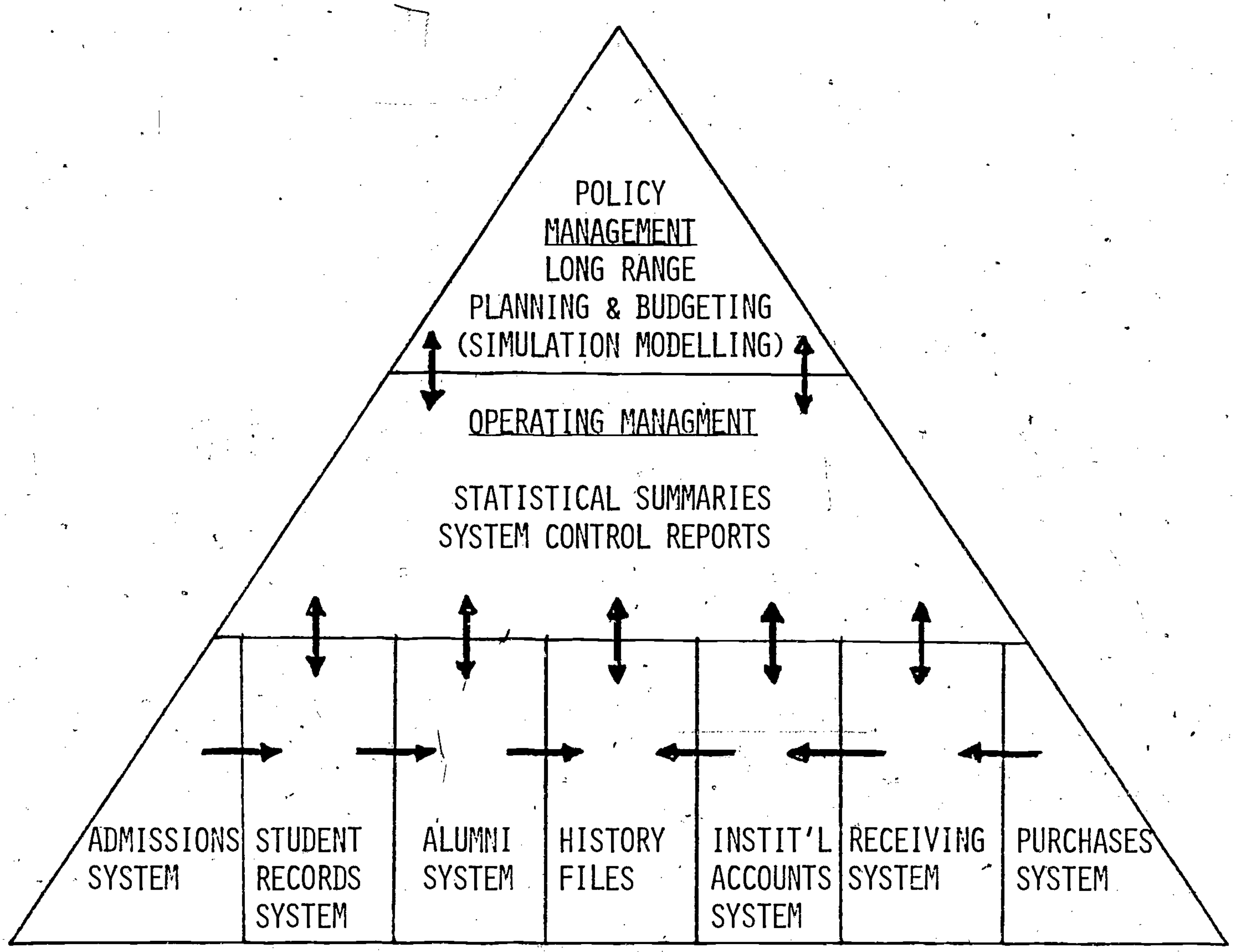
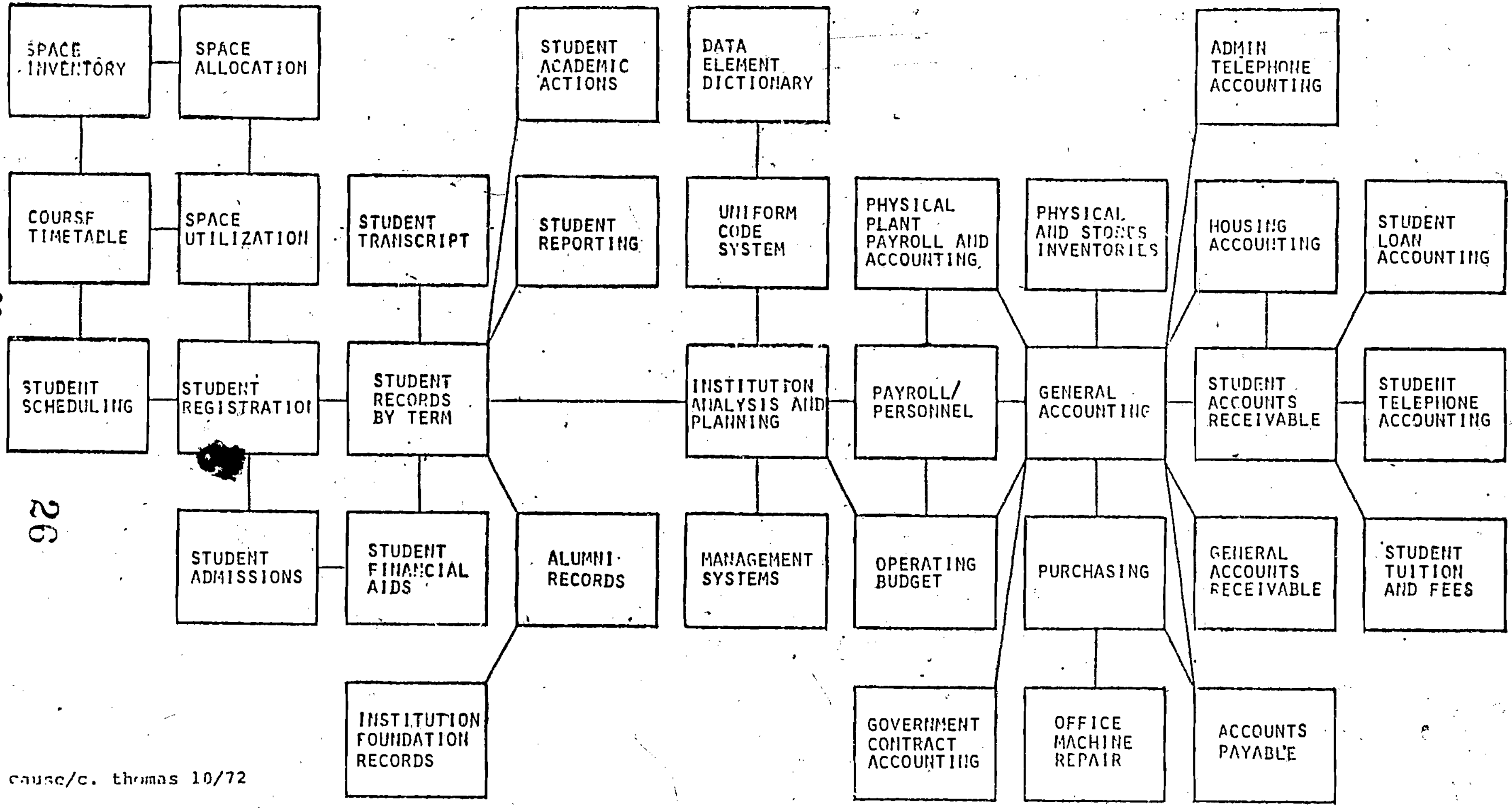


EXHIBIT D

HIGHER EDUCATION INFORMATION SYSTEMS NETWORK



cause/c. thomas 10/72