

49716

UNITED NATIONS
ECONOMIC
AND
SOCIAL COUNCIL



Dist.
LIMITED



E/ECA/PSD.3/8
12 November 1983
Original: FRENCH

ECONOMIC COMMISSION FOR AFRICA

Third session of the Joint Conference
of African Planners, Statisticians
and Demographers

Addis Ababa, 5-14 March 1984

REPORT ON THE DEVELOPMENT OF ECA'S STATISTICAL DATA BASE

Contents

Paragraphs

A. INTRODUCTION	1 - 6
B. CONCEPTUAL ORGANIZATION OF PADIS-STAT	7 - 27
I Field of application	7 - 11
II Structure and contents	12 - 18
III Resources involved	19 - 24
IV System installation plan	25 - 27
C. ESTABLISHMENT OF LEVEL III	28
I Management elements	29 - 30
II Logical structure of data	31
III Record structure	32 - 34
IV Core description	35 - 37
V Physical installation	38 - 42
VI Initiation of the bank	43 - 47
D. FUTURE DEVELOPMENT PROSPECTS	48 - 54
I Action concerning Level III	49 - 50
II Activities concerning the two other levels	51 - 54
E. CONCLUSIONS	55 - 58

Annexes

- Comprehensive diagram of PADIS-STAT
- Subjects to be stored in Level III
- Diagram of Level III

A. INTRODUCTION

1. The PADIS-STAT project is one of the five components of the Pan African Documentation and Information System (PADIS) project currently being conducted at ECA. It relates to the setting up of a bank containing socio-economic statistical data on African countries. The objective is to store a sizeable and optimum amount of statistical information and to provide speedy and multiple access to such information. To help in carrying out this important task of PADIS a survey mission organized and financed by the Government of France was asked to test in a number of African countries, the project feasibility and then to consider and submit a set of proposals for organizing the management of the system. Following the conclusions of the mission and those formulated by the ad-hoc inter-divisional working group set up to study the practical problems relating to the implementation of the project, the Executive Secretary of ECA launched in December 1981, the initial regional operations of PADIS-STAT (circular note 8/81 dated 30 December 1981).

2. The Statistics Division was rightly designated project management authority with the task of using the material computer facilities of PADIS. The Division set about this task with much determination despite its lack of experience in the subject and the lack for one year of a qualified computer expert that could assume responsibility for the data-processing aspects. After 18 months of work, the Division can now submit an initial report which is encouraging, all things considered. That report is the present paper.

3. Creating a data bank is a complex process in which many considerations come into play. Such a bank can be created only when the process is broken down into various stages or levels. This report therefore has three major chapters.

4. Chapter I will describe the conceptual level which explains how the information is organized into a coherent whole and the data processing capabilities required by the project.

5. Chapter II will be devoted to the establishment of level III. It will describe the manner in which logical and physical structures have been organized and go on to analyse the physical installation of level III.

6. The last chapter will discuss the prospects for short-term development which will be confined to the sole aspects of material implementation.

B. CONCEPTUAL ORGANIZATION

I. Field of application

7. By definition, a data bank is a mass of information stored in a computer memory and concerns a specific scientific, economic or cultural field. This means that a data bank is hooked into an information system of which it should provide as faithful an image as possible. From this point of view, regional PADIS-STAT is an integrated computerized system of macrodata relating to African economic and social statistics. Its basis is the ECA statistical production system. Its conceptual organization is derived from that system. It will therefore seem helpful, before considering its

organization, to recall briefly the main characteristics of the statistical production system. In this way, it will be easier to gain a clearer view of the main organizational and functional aspects of the PADIS system that will be considered later.

8. ECA's statistics production system has the general objective of responding as satisfactorily as possible to the need for regional statistics to provide a picture of the economic and social life of African countries. The main aspects of the system are:

Data collection;
Data processing, analysis and interpretation;
Data dissemination.

9. To carry out these activities, the system is divided into three sections, each of which is responsible for the production of data on a specific area:

The General Economic Statistics Section is responsible for national accounts, industrial statistics, prices and other economic statistics with the exception of foreign trade;

The Demographic and Social Statistics Section;

The Data Bank Section which is responsible for foreign trade, co-ordination of data processing and information.

10. The data collected by the system are general and are mainly received from national statistical systems. In other words, African countries are the main information sources of the system. Some of these data are currently stored on magnetic devices. This is the case of external trade, national accounts and price statistics. The other statistics are still stored in manual files. The system is currently engaged in:

Setting up data bases;
Issuing publications;
Preparing statistical surveys;
Providing information required on an ad hoc basis

11. The system described is not a production system in the true sense of the word but rather a system for putting into standardized form national statistics at the regional level. It follows that the quantity and quality of statistics prepared by the ECA production system depends on the quantity and quality of statistics produced at the national level. Consequently, the usefulness and development of regional PADIS-STAT will depend first of all on the quality of Africa's statistics production and the effectiveness with which such statistics are disseminated. In this field however, much effort remains to be made by the countries and also by ECA in terms of technical assistance to support and develop national activities.

II. Structure and contents of PADIS-STAT

12. Regionally, PADIS-STAT is a system primarily designed to:

- 1) Provide ECA statisticians and economists with direct and speedy access to African statistical data, data processing and analysis services.
- 2) Automate the publications of ECA's statistical information system;
3. Provide assistance services to African countries that plan to set up statistical data banks.

13. In this way, the system will no doubt have a significant impact on the exchange of information among ECA member States and on the distribution of international data.

14. Regional PADIS-STAT has three levels of which levels I and III may be considered sub-banks organically hooked up with the master bank which could be termed level II (see annex I).

a) Level I

15. This level involves a dissemination operation. It should actually make it possible to consult speedily and automatically edit country report files. For each member country, it will have a number of master data from which a quick picture of the economic and social situation of the country can be drawn. Such data will essentially be structural indicators of population, social and economic infrastructures and some aggregates of national accounts. Provided over a short period of two or three years, the data can be retrieved from level II and III.

b) Level II

16. For several years now, ECA has had a statistical data pool (ECA statistical data production system) into which information was fed by member States and certain international statistical systems. Level II will contain the information pool covering long periods. It will be a multiple data filing system like the pool composed of eleven autonomous data producers with each file covering a statistical field. In level II, each field will have its own logical structure. As indicated earlier, some of the data files are already on magnetic tape. However, they will have to be logically structured and organized into data bases that can be interrogated on line. Even though it can be opened for reference and editing, level II will primarily provide a broad basis of data for producers of the ECA Statistics Division and also a source of data supply to the two other levels. In this connection, it will have to comprise for each data file a sub-data file, and statistical processing instruments for calculating medians, variances, trends, regressions, seasonal variation adjustments and other operations.

c) Level III

17. Level III is basically geared to dissemination and analyses. It will contain only annual time series that meet the two main concerns expressed in annex II. It will mainly be a tool for reference and analytical processing primarily to be used by ECA's substantive divisions. It should enable the divisions to retrieve data from part of the base, conduct necessary analysis and produce results in a direct form that can be used for reports.

18. This three-level data processing system briefly described will have to be installed and developed at ECA headquarters. It derives its origin from the proposals made by the French mission entrusted with testing the feasibility of the project. Obviously, this will be an extensive bank of data on the two main components of effective African economic and social observation. These components are statistical (production, analysis and dissemination of statistical data) and analytical (descriptive analyses, models, projections and the like). It goes without saying that the installation of such a system will be a long-term endeavour requiring substantial data processing skills and capacities. In this regard, it should be pointed out that the material and human resources that the Statistics Division currently has for implementing the project is woefully inadequate in view of the technical requirements. This imbalance between means and objectives is highlighted further on in this chapter.

III. Resources involved

19. The choice of organization based on processing and data alone can lead to a dead end. The organization will have to be installable. This means that the type of equipment plays an important role. In the present paper, an evaluation is made of the resources needed to establish and operate the PADIS-STAT regional system in a useful manner.

(a) Equipment resources currently available:

20. The current equipment resources of PADIS are the following:

1. Equipment

a central HP 3000/III processing unit having a 2 mega byte memory (maximum capacity of an HP 3000/III);

a command console;

two magnetic disk units of 120 M-bytes each and one unit of 400 M-bytes; two tape drives;

two printers having a capacity of 300 lines per minute, 136 columns and 96 characters;

64 interactive terminals;

a graphic display unit;

a graph plotter, three-dimensional and four columns.

2. Software

An MP III (multi programming executive) which optimizes operations by time sharing and batch processing. The system can accommodate up to 64 users operating in time sharing;

Image 3000: This is a Hewlett-Packard data base management system;

Minisis: This is software developed by the Canadian international development research centre and which can be used on HP 3000 for the management of bibliographical information;

SPSS: A statistical package acquired in the United States and used for data analysis;

Other software: EDIT, FCOPY, SORT/MERGE, QUERY, VIEW, KSAM, etc.

21. It should be pointed out here that of this equipment, the Statistics Division has permanently at its disposal only one terminal and shares the computer's resources with other divisions.

(b) Required additional resources

1. Equipment

22. In order to ensure that the bank is properly built and developed the Statistics Division also needs the following equipment:

Two tape drives;
Two 400-M-byte magnetic disk units;
Magnetic tapes reels (each 2160 metres) a total of 150 tapes
A curve plotter;
A microfiche reader/reproducer;
Four interactive terminals;
Two line printers.

2) Staff resources

23. The data processing staff currently helping to establish the data bank is composed of a single programmer. This staff member has already too much to do and cannot carry out current duties in addition to project development and extension activities. It is absolutely necessary to provide the division with two statisticians with knowledge in computer science to conduct those activities.

24. It is clear that the set objectives will be attained only when adequate resources are mobilized. It is therefore urgent to find solutions to the drastic inadequacy of resources earmarked for implementing PADIS-STAT. Bilateral technical assistance could play an important role in this area.

IV. System installation plan

25. A strategy for installing the system was defined in late 1981 by an advisory committee made up of the substantive divisions of ECA particularly concerned by the project. These were the Statistical Division, the Socio-Economic Research and Planning Division, the Population Division, the International Trade and Finance Division and PADIS. The strategy had three phases of implementation. Phase one was confined to the establishment of Level II from the statistical data pool existing at ECA. Levels III and I would be implemented in the subsequent phases. The initial disc spaces required were estimated at 540 mega bytes for Level II and 269 M-bytes for Level III. No plans were made for the establishment of Level I.

26. The strategy has begun to be implemented. The study of logical grouping of producer files has actually begun. The Price Statistics file has even been finally structured and partly loaded. However, work has disrupted by the transfer to New York of the computer expert responsible for the establishment of Level II during Phase I. A replacement was secured only a full year after. Furthermore, serious problems of disc space arose when an attempt was made to transfer on to discs producer data files previously stored on magnetic tape (with the present format) for the archives.

27. Owing to these constraints, it was not possible to continue with the implementation to proceed first with the establishment of Level III which was least demanding in disc space. The following chapter will outline the progress achieved in this endeavour.

C. ESTABLISHMENT OF LEVEL III (see annex III)

28. In establishing Level III, the first objective is to fashion as soon as possible, with the available data processing equipment, a tool capable of generating and restoring in table form the information contained in the Statistical Yearbook for Africa in conversational or batch mode. This tool would subsequently be completed so that it can effectively perform the functions for which it was designed. At the moment, the scale models of Level III have been made. An initial programming has made it possible to test the reliability and performance time of basic operations. The idea is to present in this chapter, the basic procedures that were followed in making models and to identify the main orientations of future action. For the purposes of this paper, Level III will be referred to as a data bank.

I. Management elements of the data bank

29. The data bank is relational to one level and its basic entity is time series. Time series are successive values acquired with time by a given entity. To time series are attached characteristics called attributes. On account of the nature of these, the base was organized into a main data file linked by pointers to six tables of variables. The principal data file which is in fact the base contains logically structured time series. The data file is estimated at over 3000 series when the system is fully developed. The tables of variables allow base elements categories other than time series to be identified and listed by category. These are the series table, the type

series table, the reporting country table, the partner countries table and the units table. Furthermore, the bank contains three other tables that define authorizations for access to information, for example, the passwords table. It seems advisable to note here that a documentation table is being prepared and will be incorporated into the base. It is important to separate these data on different files for the following reasons:

- Economy of memory space;
- Improved research performance;
- Flexible updating.

30. All these data files are stored on discs and are therefore directly accessible on a permanent basis. For the time being, the time series data start from 1970.

II. Logical structure of the principal data file

31. The structure of a computerized data file should be considered from two points of view: the logical data file as used by the application programmer and the physical data file as stored in the computer. It was mentioned earlier that the data base concerns a single level of operation which means that the principal data file has but one logical level and consequently the number of zones of its logical recordings is constant. A time series having attributes or elements directly attached to it is considered a record. The table below sets out the format of a logical record. The approach is based closely on the data model adopted by UNCTAD for its data bank.

Table 1 - Logical record structure of the principal data file

Zone number	Title	Length in bytes	Type	Remarks
CODESER	Series identifier code	6	alpha numerical	based on the UN classification system
MODE	Mode of observation	2	alpha numerical	determines the type of series for e.g. P - production C - consumption
REPCODE	Reporting country code	4	numerical	e.g. Algeria nomenclature of UN countries
PARCODE	Partner country code	4	numerical	" "
UNIT	Unit of observation	2	alpha numerical	
BASEYR	Base year	2	numerical	e.g. 70
PRDTYPE	Reporting period of the series	2	alphabetical	e.g. A - Annual M - Monthly

Table 1 - Logical record structure of the principal data file (continued)

Zone number	Title	Length in type	Type	Remarks
STARTYR	Starting year of data values	2	numerical	
VALUE	Observation values	9 sign	numerical	
SOURCE/REM	Sources and remarks	2	alphabetical	identifies the source and comments on the observation
UPDATE	Date of update	6	numerical	e.g. year - month-date
CRID	Creator identifier	2	alphabetical	automatically generated by the system
EXDATE	Origin of the series	2	alphabetical	e.g. DW STAT - date of deletion of the private data entry

III. Record structure

32. The physical record is the set of characters transferred by the system in the course of an input/output operation. All physical records or data set blocks have a fixed length. For the principal data file, the adopted blocking factor is four. With regard to tables, the blocking factor varies in accordance with the length of records.

(a) Major physical data file

33. The storage procedure for the data file is the sequential process of two-way chaining (forwards and backwards chaining). The data file is therefore a ring or a set of blocks linked together by a closed continuous chain with the last element pointing to the first. In such a structure, each record has two pointers (PTR) for each of six "factored" tables or 12 pointers or 24 words. Furthermore, each block is preceded by a one-word bit map indicating whether a particular physical record on the block contains data or is empty. The maximum length depends on the blocking factor. In this case this amounts to 16 words per record. The size of a block is therefore 49 words as indicated in table 2 below.

Table 2 - Physical record structure

1 word bit map	24 word pointer	83 word data record	24 word pointer	83 word data record	24 word pointer	83 word data record	24 word pointer	83 word data
----------------	-----------------	---------------------	-----------------	---------------------	-----------------	---------------------	-----------------	--------------

(b) Variable tables

34. Each table is linked to the master data sets by two five-word record pointers. The physical record structure of a table contains data records, pointers and, similar to the block of the master data set, a bit map whose length depends on the blocking factors. For example, the physical record of the "series" table has the following block size:

data: 15 words
pointers: 10 words
control zone: 2 words

total: 502 words taking into account the blocking factor of 20 in this case

IV. Core description

35. The various processes which a data bank undergoes can be divided into two categories: functional processing and service processing. Service processing arises from technological constraints and do not apply to such applications as copying, temporary memorization, annexed data file creation and so on. Functional processing or utilization covers traditional processing for primarily physical creating and control, maintenance in terms of addition, selection and extraction. The core of a data bank is the set of programmes and subprogrammes that carry out functional processing. The core of the bank we are concerned with is made up of the following basic software:

Image 3000, HP management software used for setting up and maintaining the bank;

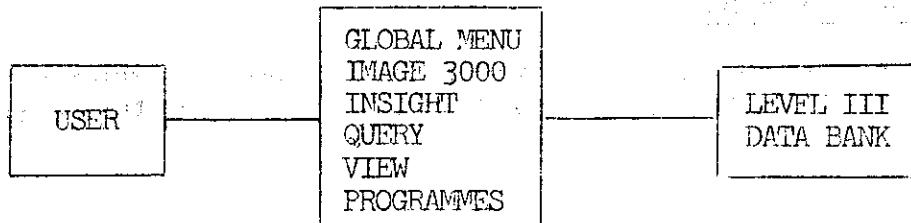
Query 3000, HP software providing access, establishment, deletion and retrieval in report form of selected information;

View, HP software that can be used for formatting is used jointly with entry 3000, for batch processing;

Insight software is supplied by Computing Capability Corporation (CCP). It is used with Image 3000, View 3000 and written programmes to allow the batch processing and linking up of Image 3000 to an application system in the transactional mode. Its major function is therefore to provide dialogue from a terminal;

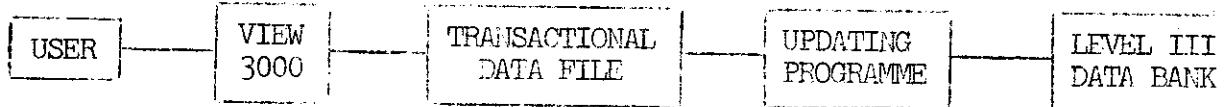
Various written programmes.

36. These programmes are used for processing in real time and by batch mode. Access can be acquired interactively in accordance with the global menu illustrated below:



37. All the retrieval, updating, tabulation, software interface procedures are developed. To carry out batch processing, View 3000, Entry 3000 and written programmes are used. In this type of operation, updating, deletion, creation and other activities, are carried out by a transactional data file. A list of changes made in this way can be provided on request using written programmes.

Batch processing



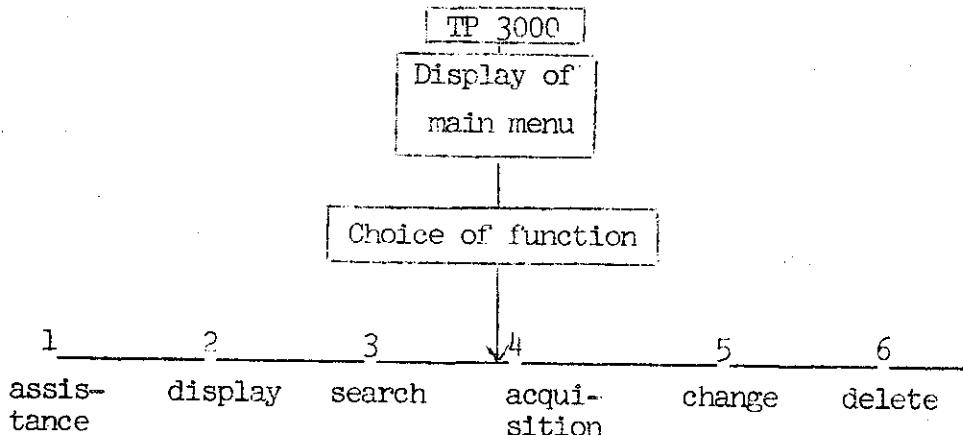
V. Physical installation

38. The screens and logical information transfer unit currently installed for data acquisition and retrieval are defined below.

1. Main menu screen

39. This is the basic screen for the system. It allows producers using the password to select the desired option from among the six proposed. The screen in some way reflects basic management operations namely: creation, updating, research and indexing. It can be obtained by using Insight HP 3000 programming and its SISTSP 10 specifications. In this way, the programme provides a visual display of the principal menu from which the choice of function can be made. The functions provided at the moment are assistance on-line, indexing, selection assistance, research, acquisition, modification and deletion.

MAIN MENU SCREEN



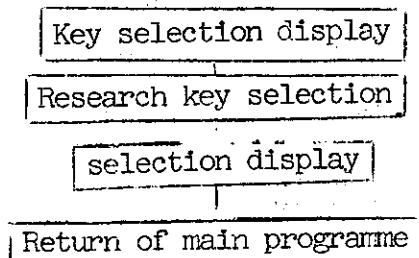
2. Screen - posting

40. The purpose of the screen is to allow one or several series to be selected from a visual display of the keys (selection material) from which the selection is to be

made. At least three keys are needed to call for a series. The keys are the series identifier, the country identifier and the series type identifier. If the value of the key is not entered, it means that all series linked to that key will be retrieved without exception. For example, if the series type of key is not used during the selection, all the series will be collected regardless of type. The model also checks the validity of the keys entered and displays any wrong selection of key. The END option indicates the end of the operation by playing back on to the visual console, the screen posting while waiting for a new selection. If the option is reused at that moment to indicate the end of selection, the main menu screen will reappear.

41. The screen is used both for the creation and updating functions of a series because the series use the same fields.

POSTER SCREEN



3. Tables and reports generation screen

42. The screen is selected from the general menu. It later becomes a menu itself allowing the user to select the type of report required. Using this model, all the tables of the Statistical Yearbook for Africa can be constructed. The screen was an initial objective set up with the help of Query 3000 and some procedures of extraction and tabulation. It is also applied in the selection instruction of series with a view to processing by some other appropriate software such as SPSS.

VI. Initiation of the bank

43. As was said in the first part of the report, the statistics assembled by the Statistics Division are stored either on paper media (manual files, Statistical Yearbook) or on computer magnetic tape. Two creation modes derived from the two current storage modes are implemented simultaneously with screen introduction and entries by magnetic tape. The initiation of the bank occurs in two phases:

Installation of parameters;
Loading of the base.

1. Installation of parameters

44. This stage of initiation consists, owing to a useful editing of the text, of installing the tables of reference whose parameters will be used in carrying out the functions. In this regard, above 1000 series with 24-character description have been coded and entered. Similarly, declaring country codes, partner country codes and series-unit type codes have been prepared and entered. Some of the tables have been left open because they will certainly need some arrangement.

2. Loading

45. Loading from existing computer data files makes it possible to save data acquisition and control time but requires programmes for reformatting. Such programmes have been developed relative to national accounts and foreign trade. In this way, all the time series covering the period 1971 to 1981 of national accounting has been fed into the bank. The same operation is underway for trade statistics.

46. The manual files will be entered in sequence or simultaneously depending on the facilities available. The work should be done under the responsibility of the Statistics Division staff since the administration of the bank provides entry formats only. It began with agricultural, transport and industrial statistical series.

47. The foregoing has been a review of the major activities carried out in the installation of Level III. It can be said that at the moment, ECA has a prototype of Level III. This does not mean, however, that the bank can be fully used now. There are still a number of constraints to be overcome before the bank becomes fully operational. These constraints are technical in terms of disc space availability for complete loading of data in machine time for final structuring. In addition to these constraints, there still remain some specific activities to be carried out to bring installation operations to a successful end. These activities are described in a part of the following chapter.

D. FUTURE DEVELOPMENT PROSPECTS

48. On account of rapid advances being made in data processing technology and the equally rapid growth of needs, it should be noted that any data bank has to adapt sooner or later. The idea here is not to anticipate such adaptation but to define the activities that need to be undertaken in order to establish a genuinely regional data bank. In this context, two types of future action can be distinguished:

- Action strictly concerning Level III; and
- Action concerning the two other levels (mainly Level II).

1. Action concerning Level III

49. The activities aim essentially at completing the installation of this level. They concern the bank's three components, namely: the feeding system, the management system and the dissemination system. Regarding the bank's feeding system, the problem to be solved is that of putting data of some international organizations on magnetic

tape. Such recording would be subject to the carrying out of specific programmes to put the existing data into desirable form on the tapes. In this regard, long and painstaking work will have to be done.

50. Regarding data-processing tools for management, the activities to be undertaken are:

(a) The steady development of the procedures already in place. Specifically the procedures will have to be well integrated through improvement and gradual enrichment. Indeed the first month of operation of the bank will no doubt reveal any malfunctions that will have to be corrected quickly.

(b) Installation of specific modules. It has already been indicated that the objective of establishing the bank is to provide a means of economic and econometric analyses as well as a means of issuing publications which happen to be in this case the African Statistical Yearbooks. In order to provide these services, additional resources, particularly calculation, will need to be introduced. The types of calculation possible at the moment concern averages, percentages, ratios, and yearly trend indices. The expansion of this field will have to be envisaged along with that of estimation procedures (interpolation and extrapolation) and modules of statistical and econometric analyses. For the moment, no plans have been made to integrate such modules into the system for fear of making its management unwieldy. Consequently, the integrated functions made available to the user will be standard research, selection editing and feeding of output files. The complex functions of calculation will have to use general interfaced software such as SPSS.

(c) Documentation is what has been termed metabase or, in other words, the information data base concerning the data base itself. Such a documentation is used in the everyday operations of the bank. A distinction must be drawn between:

(i) Core documentation (implementation of basic operations, safety procedures, definition language and data manipulation);

(ii) Acquisition documentation (acquisition formats, acquisition procedures and check lists);

(iii) Documentation of application programmes;

(iv) Documentation of data administrator (data dictionary and its updating);

(v) Documentation of non-computer staff users (language and enquiry procedure). Such documentation is drafted as decisions are taken. A substantial part of such documentation is already in the bank.

II. Action concerning the two other levels

51. As already indicated, the chapter dealing with the conceptual organization of PADIS-STAT, Level II will be the major link of the system. It should actually contain

all the statistics produced by ECA over a long period of years. Since it is geared more to production than dissemination, it will have to be organized on a multiple-filing basis or in other words be made up several sub-bases with each statistical subject containing one sub-base.

52. The activities to be undertaken in the establishment of this level coincide with conventional procedures for setting up a data bank except for one thing. In this case, the process begins from structured material which is largely computerized. The following operations will specifically have to be conducted:

- Storage on disc memory of computerized data files existing in the Statistics Division. This initial stage will have the advantage of avoiding future information reacquisition and make it possible to provide deferred time services while waiting for the on-line installation of the level;

- General substantive analysis. The various logical data files will have to be constructed taking into account the constraints imposed by magnetic tape feeding from the outside. The physical data sheets will have to be constituted and the application programmes that will produce required results defined;

- The loading of producer data files into the base and the retrieval of data contained on magnetic tapes sent in by international organizations and countries.

53. It goes without saying that these operations concerning installation of Level II imply the prior existence of the needed data-processing capability in terms of disc space, sufficient quantities of tapes, tape drives and the like, in addition to qualified data processing personnel.

54. In conclusion Level I should be mentioned for the record. Beyond the determination of content and type of storage (series or tables) it does not seem to pose any problems other than that of resources.

E. CONCLUSION

55. In this report, an attempt has been made to spell out the various stages of the process leading to the installation of a prototype of regional PADIS-STAT Level III. In fact, it might be more appropriate to speak of an initial version of Level III rather than a prototype. Indeed as has been observed, the basic operations of managing this level have been carried out. Currently, by using the various modules installed, enquiries can be made from the bank, series can be selected and the results of such selection edited. Certainly, other significant activities remain to be carried out. For example, some flexibility would possibly have to be brought into the system for editing and putting into place the necessary utilization programmes, in particular calculation and econometric analysis programmes. Programmes will also need to be made to enter outside information stored on magnetic tape. After that, the installation of the two other levels will have to be tackled. It should, however, be noted that even

though the difficulties are great, the developments are technically possible and efforts are accordingly being made. On the other hand, the data bank will be deeply influenced by the extent to which the two substantive problems are resolved.

56. The first problem concerns the resources involved, in particular, equipment resources. Indeed, the data bank might well never become operational if it is not provided with the necessary processing capability in terms of disc space, sufficient number of terminals, printing capacity and so on. The Statistics Division, which is the project management authority, has a long way to go in securing this capability. ECA's regular budget will not immediately be able to address this problem globally. For this reason, bilateral technical assistance would need to be sought.

57. The second problem touches on the feeding and documentation of the bank. As was pointed out in chapter I of this report, the countries of the region are the main information source of PADIS-STAT. Indeed, PADIS-STAT basically relies on ECA's statistical production system which in turn is totally dependent on African statistical systems. Then again, ECA faces the following difficulties in putting together statistical information from African countries:

- Undue delays in the publication and sending of material between ECA and the member States;
- Most of the information is stored on paper media.

In order to confer a certain quality and usefulness on the data bank, these difficulties must be solved. In this regard, ECA will try to use as much as possible all data sources such as national publications, international sources and information gathered in the course of missions. However, the ideal solution would be for States to record their statistical information on magnetic tapes or diskettes that can easily be reproduced and transmitted. The States will need for that purpose to have their own editing equipment.

58. Regarding documentation, the data fed into the bank will have to be referenced before it can be interpreted and used. This means that the bank should contain a data base on the information it is required to disseminate. Such documentation can be put together only with the assistance of the national bureaux of statistics that produce the information. Elements of such documentation are already available in print. The first approach would be to assemble and develop them systematically.

ANNEX I
STATISTICAL INFORMATION SYSTEMS

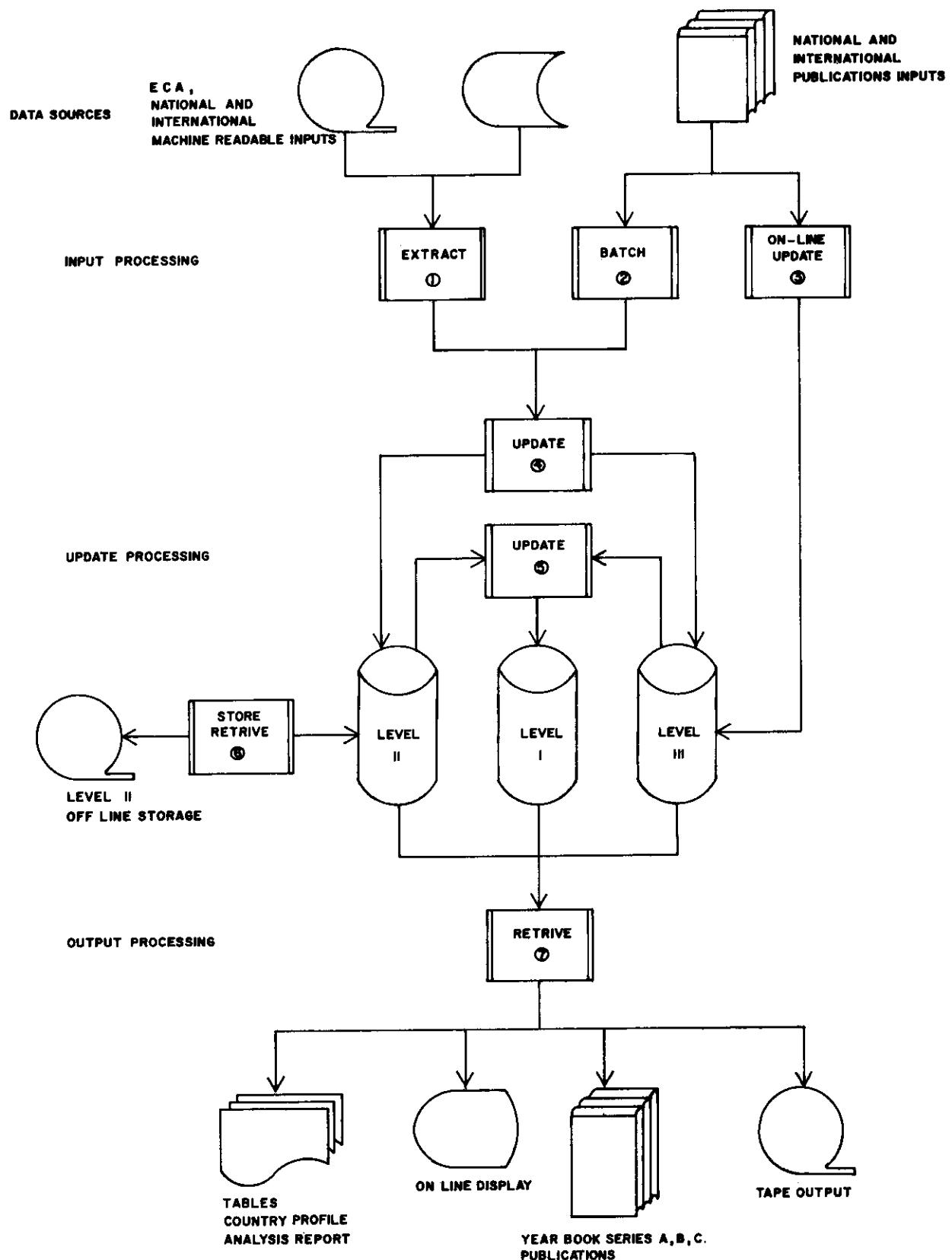


Figure I
STATISTICAL INFORMATION SYSTEMS

Present contents of PADIS-STAT Level III

1. Demographic and social statistics

Population and employment
Education
Health

2. Economic statistics

National Accounts
Balance of Payment
Public Finance
Price
Foreign Trade
Agriculture, Forestry and Fishery
Mining and Quarrying
Manufacturing
Construction
Electricity, water and gas
Transport and Communication
Tourism