Distr.: LIMITED

TRANSCOM/666
June 1993

Original: ENGLISH

ECONOMIC COMMISSION FOR AFRICA

Launching Seminar on a
Regional Transport Data Base

Addis Ababa, Ethiopia
15 - 18 June 1993

SUMMARY OF FINDINGS AND DRAFT RECOMMENDATIONS CONCERNING THE ESTABLISHMENT OF A REGIONAL TRANSPORT DATA BASE AT ECA

Introduction

1. This document presents a summary of findings and recommendations from a paper produced in the framework of the analysis of existing means and the proposed implementation of a transport data base at United Nations Economic Commission for Africa (ECA) Transport, Communication and Tourism Division (TCTD). It is intended to provide the participants of the launching seminar with background information on the anticipated implemention of this data base and invites them to make comments and suggestions in order to gain the most appropriate set up at TCTD.

2. The full report is scheduled for the end of June 1993 and will be distributed to all parties concerned. As the recommendations contain as well the suggestion of a distribution scheme of TCTD's transport data base to ECA member States, the national and institutional focal points in the region will be invited to comment on this report.

Existing equipment at ECA

3. ECA has long been depending on a single source of computer power: a Hewlett Packard (HP) 3000 mini-computer purchased in the early 80's within the framework of a UNDP-financed information systems project, the Pan African Development System (PADIS). Only recently ECA acquired ca. 300 micro-computers for office automation purposes, thus permitting decentralized computing within ECA.

4. Additionally two WANG Virtual System (VS) mini-computer systems have been purchased of which one is installed under the United Nations Transport and Communications Decade for Africa (UNTACDA) and the second is installed at Administration, Personnel Section.

5. All three mini-computer systems, the HP 3000 and the two WANG VS, can no longer be considered to be up-to-date equipment in a changing world of micro-computers which became so powerful that they even challenge the computing power of mainframe computers.

6. The installed HP 3000 is basically serving the following purposes:

   a. **Administrative computing**; this includes payroll, finance and budget tables, the maintenance of ECA's manning table, mailing addresses, etc. The programming and maintenance of this type of computing is managed by ECA's Information Systems Section (ISS) and its four programmer/analysts.

      Programmes are mainly written in-house using COBOL language.

   b. **Statistical data base computing**; initially created with PADIS financial support for the production of ECA's African Statistical Yearbook, the Statistics Division of ECA is responsible for the maintenance of several data bases containing time-series on statistical data concerning the

African continent. ECA staff outside the Statistics Division have access to this data base, and *ad hoc* request from outside ECA are handled by the Statistics Division. The main data base contains over 140,000 time-series on African social and economic data.

The data base is based on IMAGE 3000, a proprietary HP data base system, and is extended for table generation and queries with in-house COBOL programmes.

c. **Non-numerical computing;** PADIS is maintaining its bibliographical data base system using the Canadian International Development and Research Centre's (IDRC) bibliographical data base system MINISIS. This system is written entirely in SPL, a HP-unique programming language, and thus not transportable to other systems. PADIS maintains a number of data bases under this system: for instance PADdev, a bibliographical development information data base, PADexp, a data base on experts, PADins, a data base on development institution, etc.

d. **networking;** the HP 3000 is connected using a modem to ECA's alternate voice and data (AVD) line. This allows the connection to the United Nations Headquarters computer systems. At the moment selected users within ECA have access to this facility within a daily two hour connection window.

e. **user-initiated computing;** in some substantive divisions within ECA individual users have developed their own programmes in order to fulfil their work assignments. As an example, in the Socio-economic Research and Planning Division (SERPD) a data base has been programmed in FORTRAN. Other users are still using the HP 3000 as a word processing machine, despite the above described switch to micro-computers.

7. Reviewing the use of the HP 3000, it must be mentioned that there are plans to migrate from this outdated equipment to more up-to-date machinery: the on-going Integrated Management Information System (IMIS) project that will be introduced for UN Administration world-wide will be based on a modern UNIX mini-computer, connected to local area networks (LANs) within administrative units and will also providing wide area networking (WAN) facilities to all of ECA; ECA's statistics division is planning to migrate its statistical data bases to micro-computer based LANs; with the arrival of MINISIS version H, IDRC's non-numerical data base system written in the programming language C and thus implementable on a wide range of computer systems including micro-computers, PADIS is planning to move its non-numerical data bases to micro-computers.

8. The lack of modern, user-friendly software for the HP 3000, its word processor is much more cumbersome to use than any micro-computer based word processing

system, will finally force the other users off the HP 3000 as well, once further support to it from ISS has ceased.

9. UNTACDA's WANG VS system, located at TCTD, has been basically used as a word processing machine. No attempt was made to acquire or create applications beyond this facility. It must be mentioned here as well that the equipment is below today's standards in particular as far as user applications are concerned. Software and hardware for this machine are extremely expensive compared with micro-computer hard- and software. In addition it should be mentioned that the maintenance of this machine incurs high cost, way above similar requirements for micro-computer equipment.

10. The second WANG VS system, located at Personnel Section, is used mainly for historical reasons: New York Headquarters, once equipped mainly with WANG computers, could provide software for the handling of P5-actions, a major task within Personnel Section. All other computer-based work is executed on micro-computers: word processing, spreadsheet work, and small data bases. As with other administrative work this use of the WANG VS system will cease once the IMIS project is fully implemented. Personnel Section will be fully integrated within the in-house administrative LAN and its connection to Headquarters. The WANG computer will be phased out.

11. The majority of computer power within ECA is nowadays based on micro-computers. Two different models have been introduced to ECA: the standard machines (all machines are NCR) are equipped with a 80386 20 MHz processor, a 40 Mega Byte (MB) hard disk, a 3 1/2 inch, 1.44 MB diskette drive and a colour screen; the "power-user" machines are equipped with a 100 MB hard disk and an additional 5 1/4 inch, 1.2 MB diskette drive.

12. Within three ECA divisions LANs are to be set up: Administration, Statistics Division and SERPD. There will be a connection between Statistics Division and SERPD in order to allow access to their respective data. These LANs are not yet fully implemented and appropriate software solutions are still missing. In the long range it is planned to network all substantive divisions, including TCTD.

13. The existing micro-computers are loaded with word processing software, in this case WordPerfect 5.1, a spreadsheet programme, Quattro Pro 2.0, and a data base programme, Paradox 3.5. This software has been chosen by UN Headquarters as a standard throughout the system. All ECA staff has been trained in word processing and selected staff has been trained in the use of Quattro Pro and Paradox.

14. Additional software like graphics and desktop publishing software has been installed at selected user sites. For very few staff members training in desktop publishing was provided by ISS. The major users of this type of software, in the Documents and Publishing Service, ISS and PADIS, have trained themselves or had

previous experience. Graphical User Interfaces (GUIs) like MS-Windows have not been installed widely within ECA.

15. Some major problems concerning the implementation of the transport data base are still existing: (1) the "standard" machines are under-equipped as far as speed, available memory and storage facilities are concerned; (2) "power-user" machines are inadequate in speed and available memory; (3) network boards are not installed; (4) useful peripherals like scanners, CD-ROM readers and archival storage facilities like Bernoulli boxes are not available; (5) a standard GUI for use within ECA has not yet been identified; and (6) the majority of printers is composed of inkjet printers instead of faster and higher quality laser printers. This set up stems mainly from the fact that the computerization of ECA was rather geared at office automation than a more sophisticated network design. With the arrival of IMIS and dropping hardware prices for micro-computers some of these problems might be solved soon.

Present information collection at TCTD

16. At present TCTD requires basically information for use by staff preparing technical publications, documents for meetings and conferences and for conducting research and delivery of programme outputs as may be required. Data for use in data bases have not been anticipated up to now. The processes and structures for collection of the required information in the field of transport in Africa have so far not been adequately developed to satisfy the demand.

17. There is a well founded concern for the development of information collection whenever the need arises but all has been done on an *ad hoc* basis. Sources for data needed for the preparation of the various publications and surveys have so far been based on:

a. The African Statistical Yearbook maintained at ECA's Statistics Division. The primary sources of maintaining the Yearbook are the national statistical offices of ECA member States.

b. The annual reports prepared by the various national transport agencies, i.e. the ports annual reports, the railways yearbooks and airlines statistical returns etc. They are normally prepared by national offices of the various transport modes and are often supplemented by those available in ECA. This type of data often suffers from serious time lags.

c. Statistical publications of the various international organizations, i.e UNCTAD review, ICAO statistics, the World Bank annual reports etc. This category of data sources is often available in the ECA library but can sometimes be obtained directly from the publishers on request.

d. Data collected by various experts in the division while on missions. Staff from TCTD is always mandated to collect relevant country

documents and publications in transport and communications and/or socio-economic development data of the African countries visited.

e. Statistical reports available in the various subregional organizations, i.e. port management associations, SATCC, UAR, International Road Federation etc.

f. Statistics included in the various national studies in transport, i.e. sectoral master plans.

g. Statistics which are available from existing transport data bases, i.e. Advanced Cargo Information System (ACIS) developed by UNCTAD covering railways, ports and container statistics.

18. The anticipated establishment of a transport data base will help in streamlining the processes of data collection at various levels, and will eliminate the dependence on various publications and sources as given above. Some of the publications which ECA now depends on are not regularly produced and some of the regular ones reach ECA when their contents is already out of date. To overcome these problems, experts from TCTD often visit selected countries for data collection. This method is time consuming and sample data collected has in many cases not been representative depending on the number of countries visited and the quality of data available.

19. It is therefore clear that a coordinated system which can assure the proper organization and availability of transport statistics is overdue. The establishment of national data bases in the countries to feed a regional one, to be established in ECA, would be the answer to the problem.

Draft Recommendations

20. The draft recommendations concerning the establishment of a regional transport data base at ECA are basically directed toward three major areas:

- information collection;
- equipment and personnel requirements at ECA's level; and
- information distribution.

Information collection

21. Information collection is the most crucial part for the development of a regional transport data base: as the old saying in computer science goes "garbage in, garbage out". Only if relevant, timely and accurate data are available, the data base can fulfil its function for development planning in the African region properly. In this

context it has to be pointed out that up to now one major part of transport information has not yet been taken care of: non-numerical information.

22. The term "non-numerical information" describes all kind of data which are not statistical data, up to now the major source for analysis in TCTD. In the framework of the emerging regional co-operation and integration in the African region information on laws, rules and regulations governing transport in African countries are of equal importance for proper planning and implementation of transport related projects. Therefore greater attention has to be focused on such information, which definitely should constitute a major part of the regional transport data base.

23. It is recommended that the basic source of information should be on the national level. This means that national focal points for national data collection have to be established which exchange their data on a regular basis with ECA. Preferrable these national focal points should have access to computer equipment in order to transfer the information straight to ECA's data base. Necessary software for data entry and correction has to be bought or developped and installed for all participating focal points. The equipement at the national level should at least be comprised of a 80486 micro-computer, a 300 MB hard disk, a colour screen, a CD-ROM reader, a 150 MB Bernoulli box, a scanner, and a laser printer. Using this equipment would allow adequate data collection for numerical and non-numerical data. Numerical data could be typed into a data collection form and non-numerical could be transferred using optical character recognition (OCR) techniques into readable text for further use in the data base.

24. In many African countries development information systems already exist and more are supposed to be established in the near future. Unfortunately these systems are mainly sectoral oriented: bibliographical systems are linked to universities and statistical systems, if existing, to the respective central statistical authorities. These two sources are usually not linked to each other so that each effort in acquiring infiormation has to be directed towards different systems or institutions.

25. It would be ideal to combine the information of different national sources or systems into the hands of a single focal point at national level. One possibility would be the use of institutions which exchange already data with ECA on a regular basis, such as PADIS's National Participating Centres. This would eliminate the necessity to establish, equip and train for all sectoral data needs sectoral information centres, thus reducing cost and the use of valuable manpower. It is recommended that TCTD dicusses this possibility with PADIS in close co-operation with ECA's Statistics Division. It might be necessary to provide African gouvernment with some seed money in order to establish the necessary infrastructure. This money could either be provided from the envelop of United Nations Development Programme (UNDP) Indicative Planning Figures (IPFs) or by bilateral sources which could be approached by ECA, World Bank and other UN agencies.

26. As has been proven with other data collection efforts in the African region, the need of face-to-face contact will not be eliminated even if such a system would have been established. TCTD's budget has to foresee means for data collection missions into ECA member States even in the future. This is particularly true for special data needs as they may arrise from *ad hoc* request in the framework of humanitarian or emergency measures handeled by ECA. The budget should allow for at least 10 mission per annum in order to assure the quality and timeliness of data for the transport data base.

27. The already mentioned co-operation and integration on regional and sub-regional level will greatly enhance the opportunities to use sub-regional institutions like ECOWAS, PTA, etc. for data collection. These institutions should be encouraged to collect sub-regional data and give ECA access to their collection.

28. A further source of data on African transport should also be considered: Universities and research institutions outside the continent. Reports and research papers on Africa are often produced in such institutions but never find their way back to the African continent. Todays facilities using electronic mail and other telecommunication connections allow TCTD to tap these source from its staff member's desk. The use of academic oriented networks like Bitnet could greatly enhance TCTD's knowledgebase and give valuable information for the transport data base. This could either be implemented by connecting a simple micro-computer to a modem and a direct telephone line or in the overall framework of the implementation of IMIS in the near future.

29. As a last group of information sources, both UN and other international organizations and regional groupings should be mentioned. In the "Directory of United Nations Dtabases and Information Services", Fourth Edition, 1990, compiled by the Advisory Committee for the Co-ordination of Information Systems the following UN bodies are mentioned as sources for data bases on transport: ECE, ECLAC, ESCWA, ESCAP, WHO, IAEA, UN Headquarters, UNCHS, FAO, and ICAO. These organizations maintain data bases already and could be used for data exchange on a global level in particular as far as global transport is concerned. Regional groupings like the European Community (EC) or the Organization for Economic Co-operation and Development (OECD) could be used at least in order to verify primary data available in Africa.

Equipment and personnel requirements at ECA's level

30. As outline above, ECA is mainly micro-computer bound. Even after the installation of IMIS the major source of computing power will be the micro-computer. Therefore it is recommended to base the regional transport data base on a micro-computer environment.

31. Existing micro-computers could easily be upgraded with additional memory (at least 8 MB) and larger hard disks (at least 300 MB) to meet networking require-

ments. Additional peripherals like CD-ROM readers could be easily connected to the existing machines. But it is evident that the main file server for the system has to be purchased. This machine should be a powerful 80486DX 66 MHz micro-computer, equipped with at least 32 MB Random Access Memory (RAM), 1 Giga Byte (GB) or above of hard disk, and appropriate backup facilities.

32. A second powerful machine is recommended to be used for OCR, CD-ROM production (see below), desktop publishing as well as a communications gateway to other networks including the backbone network which will be installed under IMIS. This machine, a 80486DX 66 MHz with 16 MB of RAM and a 500 MB hard disk, should be connected to a high resolution scanner, a high speed laser printer (12 pages per minute or above) and a colour-postscript laser printer (for map printing).

33. As far as software is concerned it is recommended that TCTD settles for the popular Windows systems under the general network software Novell. Most of the new software is designed to run under this GUI and only recently a new version on Windows NT was released which gives true multi-tasking facilities but lets all popular Windows-applications run as before. For statistical data the new Paradox for Windows is recommended, thus capitalizing on the existing knowledgebase at TCTD. As far as non-numerical data bases are concerned the IDRC MINISIS version H is recommended, a software which could be received for free in-house as PADIS is one of the two MINISIS Resource Centres for Africa. Even though the original distributed version will not yet run under Windows, it has been indicated that users will provide the necessary programme code for the proper linkage of MINISIS to Windows. The use of this software would greatly enhance the standardization within the house and the continent at least as far as non-numerical data processing is concerned.

34. As mentioned above, all ECA staff has followed basic micro-computer training. This includes the Disk Operating System (DOS) as well as word processing. Users of the regional transport data base system would have to be trained in the handling of the data base. In order to guarantee a proper functioning of the data base and in light of the diversity of data handled, it is recommended to create two posts for the data base management: one data base manager at P5 level and one senior programmer/analyst at P4 level. This personnel could handle the introduction of TCTD staff into the use of the system as well as carry out all data base management tasks and preparation of publication (see below).

Information Distribution

35. Four layers of distribution of the regional transport data base are anticipated: within TCTD, within ECA, within the African region and the global distribution. The distribution within TCTD is established by a LAN, based on a data base server, maintained by a data base manager and a senior programmer/analyst. Each staff member of TCTD will have thus access to all relevant data in the system using his or hers micro-computer. Distribution to interested users within ECA will be established either by downloading of data from the data base to diskettes or, once the

IMIS backbone network is installed, by direct search in the data base. This later option should be finalized some times in late 1994. In case that a working system would be established before that time, the data base manager could serve as a dispatcher of requested information.

36. More advanced national focal points in African member States could have access to the data base either on-line by dialling in through a modem or by means of using a bulletin board system. This later option is in so far more recommended as telecommunication costs in Africa are still very high and on-line searches of data bases could quite take some time. In case that the bulletin board option is installed, a distant user could post his or hers query requests in a mail box, the data base management would run the queries and post in return the results in the user's mail box. From their it could be automatically downloaded to the users site.

37. The same means could also be used to exchange information. Data files produced at the national level could be send directly to the ECA and entered to the data base. Thus costly data collection missions could be avoided.

38. Only recently a new medium for distribution of larger data bases has become available to a wider public as hardware prices have dropped dramatically: Compact Disk Read Only Memory (CD-ROM). Up to now the equipment for production of CD-ROMs has been prohibitive expensive: such equipment would cost US$ 100,000 and above. Newer development in the field of multi-media computing allows now to produce small number of copies with equipment of not more than US$ 5,000. This write-once CD-ROM technology could be employed in order to distribute the entire regional transport data base to the national focal points. Regular updates containing all data received from all member States and institutions would enable the national focal points to answer *ad hoc* questions on the spot, thus giving decision makers and planners access to more up-to-date and comprehensive information on transport issues in Africa.

39. Unfortunately one will not find the necessary sophisticated equipment in all African member States. This requires the use of traditional means of information distribution: the production of printed publications. It is recommended to use at ECA desktop publishing software in order to produce master copies for traditional reproduction. TCTD should decide on the different type of printed publications based on the regional transport data base contents. Within ECA, procedures have been established to publish directly from data bases using standard data base software like dBASE IV for creating files which are immediately readable from the desktop publishing software. This will reduce the turnaround time for publication production to a minimum. The goal of timely and up-to-date information distribution could thus be easily reached.

40. Finally a word of caution: all mentioned specifications for hard- and software might be out of date at the time of the actual implementation of the regional transport data base. It is sure that in a short time the first micro-computers with the

new Pentium processor from Intel will be available on the market. New operating systems from IBM, HP, and Next might give a better value than the one mentioned above. Upgraded versions of commercially available software like Paradox for Windows might be available soon. Thus it is recommended to review carefully all hard- and software options at the actual time of implementation.

Conclusion

41. In the light of the actual status at ECA's TCTD concerning information collection and processing, a micro-computer based LAN system seems to be the most adequate way of providing ECA and its member States with necessary data on transport. National focal points will provide primary information to the system, enhanced by information received from international organizations outside the African region as well as analysis executed at TCTD.

42. The distribution of the data base information will be carried out by direct networking, and by publication of CD-ROMs and printed material.