



Economic and Social Council

Distr.
GENERAL

E/CN.3/1987/2
2 December 1986

ORIGINAL: ENGLISH

STATISTICAL COMMISSION
Twenty-fourth session
23 February-4 March 1987
Item 3 (a) of the provisional agenda*

SPECIAL ISSUES: COMMUNICATIONS BETWEEN NATIONAL AND INTERNATIONAL
STATISTICAL AGENCIES AND THE DEVELOPMENT OF DATA BASES

Issues of communication between national and international
statistical agencies and data base development

Report of the Secretary-General

SUMMARY

The present report is submitted in response to a request made by the Statistical Commission at its twenty-third session. It contains information on the data collection activities of international organizations and on the current activities related to the dissemination of international statistics (paras. 4-19). The report discusses issues in statistical data communications between national and international statistical agencies (paras. 20-29) and presents a summary of recent developments in the statistical data bases of international statistical agencies, including related communications issues (paras. 30-67). Possible measures to improve data communications between national and international statistical agencies are suggested (paras. 68-81). Points for discussion by the Commission are included (para. 82).

* E/CN.3/1987/1.

CONTENTS

	<u>Paragraphs</u>	<u>Page</u>
INTRODUCTION	1 - 3	4
I. REVIEW OF THE CURRENT STATUS OF COMMUNICATIONS BETWEEN NATIONAL AND INTERNATIONAL STATISTICAL AGENCIES	4 - 19	4
A. Data collection activities	4 - 10	4
B. Data dissemination activities	11 - 19	6
II. ISSUES IN COMMUNICATIONS BETWEEN NATIONAL AND INTERNATIONAL STATISTICAL AGENCIES	20 - 29	8
A. Demand for international statistics	20 - 21	8
B. Increased expectations from information technology for improved data communications	22 - 23	8
C. Availability of internationally comparable data	24 - 27	9
D. Necessity of standards for the exchange of data in machine-readable form	28	9
E. Special needs in international statistics for meta-data and footnotes	29	10
III. DEVELOPMENT OF DATA BASES AT INTERNATIONAL STATISTICAL AGENCIES	30 - 67	10
A. Statistical Office of the United Nations Secretariat ..	30 - 34	10
B. Economic Commission for Europe	35 - 36	11
C. Economic Commission for Latin America and the Caribbean	37 - 38	12
D. Economic Commission for Africa	39 - 42	12
E. International Labour Organisation	43 - 46	13
F. Food and Agriculture Organization of the United Nations	47	14
G. World Health Organization	48 - 50	14
H. World Bank	51 - 53	15
I. International Monetary Fund	54 - 56	15

CONTENTS (continued)

	<u>Paragraphs</u>	<u>Page</u>
J. United Nations Industrial Development Organization	57 - 59	16
K. General Agreement on Tariffs and Trade	60 - 61	16
L. Organisation for Economic Co-operation and Development	62 - 64	17
M. Statistical Office of the European Communities	65 - 67	17
IV. MEASURES TO IMPORVE COMMUNICATIONS BETWEEN NATIONAL AND INTERNATIONAL AGENCIES	68 - 81	18
A. Efforts to improve data comparability	68	18
B. International co-operation for setting standards for data communications	69 - 73	18
C. Other measures	74 - 81	20
V. POINTS FOR DISCUSSION	82	22

INTRODUCTION

1. At its twenty-third session, the Statistical Commission, in approving the provisional agenda for its twenty-fourth session, requested the Working Group on International Statistical Programmes and Co-ordination to consider, in particular, the possibility of merging sub-items 3 (a) (Communications between national and international statistical agencies) and 3 (c) (The interrelations between statistical services and the emerging network of data bases). 1/

2. The Working Group, at its eleventh session in 1985, amended item 3 (Special issues) of the provisional agenda for the twenty-fourth session of the Commission by merging sub-items 3 (a) and (c). The Working Group agreed that one document should be prepared for the item, incorporating relevant elements of the documents originally proposed by the Secretariat under the two sub-items. The document was to deal with all pertinent issues of communication between national and international statistical organizations, including classifications, methodological issues, technology issues and questions of data comparability between national and international organizations. Steps to enhance or improve communications were also to be dealt with in a range of areas including, but not limited to, technology.

3. The present report contains a review of the current status of communications between national statistical offices and international statistical agencies in the areas of data collection and data dissemination. It summarizes the main issues in data communications and discusses possible measures to improve data communication between national and international statistical agencies, taking into account views and suggestions provided by those agencies. The report also contains a summary of the current status of the development of statistical data bases at the international level in the context of their role in communication between national and international statistical agencies, based on the information provided by statistical agencies within the United Nations system and other international statistical agencies.

I. REVIEW OF THE CURRENT STATUS OF COMMUNICATIONS BETWEEN NATIONAL AND INTERNATIONAL STATISTICAL AGENCIES

A. Data collection activities

1. Data collection through questionnaires

4. International statistical agencies collect data from national statistical services to provide a wide range of statistical data for their own programmes and for dissemination to other intermediary and final users, including users in other international organizations and member States. Major efforts of international statistical agencies are, indeed, directed towards the collection, processing and dissemination of reliable and internationally comparable statistical data. For national statistical services, this effort means that, by putting their own data into an international context, it is possible to obtain comparative perspectives that cannot be obtained in any other way.

5. Most statistics currently collected from countries by various international organizations are obtained through questionnaires, except for some statistical series, such as international trade statistics, which are mainly collected on magnetic tape (see para. 8 below). In all, 21 international organizations, including regional commissions and non-United Nations organizations, regularly collect over 250 statistical series from central statistical offices and other government offices.

6. The international statistical community, concerned about reducing the reporting burden on countries, has been dealing with the co-ordination of questionnaires through the Sub-Committee on Statistical Activities of the Administrative Committee on Co-ordination and has monitored them in order to avoid unnecessary duplication. To this end, many questionnaires are used jointly by organizations of the United Nations system and other international organizations.

7. The following are examples of the exchange of completed questionnaires between the Statistical Office of the United Nations Secretariat and other organizations: copies of the relevant parts of the completed Demographic Yearbook questionnaires are sent to the regional commissions, the International Labour Organisation (ILO), the Food and Agriculture Organization of the United Nations (FAO), the United Nations Educational, Scientific and Cultural Organization (UNESCO), the World Health Organization (WHO) and the Inter-American Statistical Institute (IASI) for internal use and for inclusion in their publications; copies of the completed questionnaires for the Statistical Yearbook are sent to the Economic Commission for Europe (ECE), the Economic and Social Commission for Asia and the Pacific (ESCAP) and the Economic Commission for Africa (ECA); copies of the completed questionnaires on indexes of industrial production are sent to the Economic Commission for Latin America and the Caribbean (ECLAC) twice a year; copies of the completed questionnaires on industrial commodity production are sent to ESCAP, ECLAC and ECA and, since in 1983, copies of part of the completed questionnaires for the Yearbook of Construction Statistics have been sent to ECE. The Statistical Office of the United Nations Secretariat and the Organisation for Economic Co-operation and Development (OECD) issue and use jointly the main annual national accounts questionnaire.

2. Data collection on magnetic tape

8. The collection of statistics from countries on magnetic tape has been done by various international organizations for some years. The Statistical Office of the United Nations Secretariat, the ESCAP, ECA and OECD collect international trade statistics in this way. The Statistical Office of the United Nations Secretariat collects maritime transport statistics on tape from some 20 countries. WHO collects detailed cause-of-death statistics by age and sex from selected countries on magnetic tape. The World Bank, through the Debtor Reporting System, receives tapes from a few countries containing external debt statistics. The International Monetary Fund (IMF) receives tapes from some member States containing national statistical data concerning international reserves, money and banking, interest rates, exchange rates, wages, employment, prices, production, trade, national accounts, balance of payments and government finance. The secretariat of the

General Agreement on Tariffs and Trade (GATT) annually requests import statistics by custom tariff item and origin recorded on magnetic tapes. The Statistical Office of the European Communities (EUROSTAT) collects from countries over 600 tapes covering various fields, including external trade statistics, input-output tables, balance-of-payment data, and agricultural and social survey data. OECD collects, for example, statistics of sea-borne trade.

3. Inventory of statistical data collection activities

9. The inventory of statistical data collection activities, which was initiated by the Statistical Commission at its seventeenth session in 1972, serves as a major tool in consultations among organizations for further refining the co-ordination of questionnaires and the processing of data. Updated every two years, the list provides a basis for reviewing progress in this area. The latest inventory (E/CN.3/1985/AC.1/R.6) was prepared in 1985 for review by the Commission's Working Group at its eleventh session. This document summarized the principal changes that had occurred in international statistical data collection activities during the previous two years (additions and deletions) and gave a brief description of the statistics collected from countries on magnetic tape. With respect to co-ordination, it described the exchange of completed questionnaires among international organizations and the sharing of data other than by the exchange of completed questionnaires. It also presented descriptions of each of the questionnaires and instances of statistics collection through magnetic tapes, information about dispatch and deadline dates and an indication of the addressees.

10. At the request of the Working Group, the inventory was circulated to all central statistical offices as a means of informing them of the various requests made to countries by the statistical services of international organizations and to invite any comments that national statistical offices might have on the data collection activities of international statistical organizations. It must be recognized, however, that within various international organizations, units other than the statistical services of these organizations engage in the collection of statistical data from countries. Such data collection activities may be routine or ad hoc, narrowly focused or broad in nature. For the most part, however, these activities - which add to the reporting burden placed on countries - are not covered in the inventory.

B. Data dissemination activities

1. Activities

11. The dissemination of international statistics is a major task of the international statistical agencies and, directly or indirectly, it consumes a large proportion of the available resources. The data available from the international statistical agencies are of great value to a very wide range of users, including Governments and research institutions. The Statistical Commission considered the issue of the dissemination of international statistics at its twenty-second and twenty-third sessions (see E/CN.3/1983/4, E/CN.3/1985/2, E/CN.3/1985/4) and is doing so again at its present session (see E/CN.3/1987/4).

12. As discussed in the previous section, data are provided by national statistical agencies and processed and compiled for dissemination by international statistical agencies. At the processing stage, the data are evaluated, edited and adjusted, where necessary, to achieve as much comparability as possible in a standard format.

13. The compilation, processing and dissemination work has been characterized by efforts to improve efficiency and economy. Extensive and improved use of computers has been central to this effort. A computer-controlled photocomposition system has facilitated timely, economical and improved production of statistical publications of the international statistical agencies.

14. Dissemination of international statistics takes place through publications, microfiche, computer tapes and so forth. The Statistical Office of the United Nations Secretariat alone prepares 16 statistical publications that are issued annually or more frequently. In order to make available to users of international statistics up-to-date information on the statistical series disseminated by international organizations, the Statistical Office of the United Nations Secretariat prepared the Directory of International Statistics, which was first published in 1975. Volume I of a revised version of the Directory was published in 1982. 2/

15. Volume I of the Directory contains a listing of the statistical series compiled by the organizations of the United Nations system and some non-United Nations organizations active in international statistics. This includes references to the organizations that collect the data from countries, the publications in which the data can be found, the frequency of the series, and an indication of whether the series are maintained in machine-readable form. Volume I also contains an inventory of the machine-readable data bases of economic and social statistics by subject and by organization.

16. The Directory was also prepared to stimulate and further facilitate the efforts of international organizations towards continued co-ordination of their data collection, processing and publishing activities. The Directory was prepared through a joint effort of the Statistical Office of the United Nations Secretariat, the statistical divisions of the regional commissions, the specialized agencies and other international organizations.

2. Dissemination of data through media other than publications

17. Microfiche. The Publishing Division of the Department of Conference Services of the United Nations Secretariat produces United Nations publications on microfiche. In addition to international trade statistics, the major yearbooks prepared by the Statistical Office and the Monthly Bulletin of Statistics (1947 to 1982) are now available on microfiche.

18. Magnetic tape. The Statistical Office of the United Nations Secretariat has been providing data on magnetic tape for the past 20 years, starting with the dissemination of international trade statistics. At present, data tapes are

available for international trade statistics, demographic statistics, special population estimates, energy statistics, maritime transport statistics, general industrial statistics, index numbers of industrial production, national accounts statistics and world development statistics. A similar activity exists in many of the international statistical agencies, and tapes are freely exchanged among them. Provision of data on magnetic tape is generally made on a cost basis but Governments receive the tapes at a highly discounted rate.

19. On-line access. Direct on-line access to the data base of international statistical agencies is an important new development where technological improvements are already contributing to the provision of better services to users. For example, a data base of international trade statistics stored in the computer of the International Computing Centre (ICC) at Geneva has been providing direct on-line access to users among international organizations and some member States (see paras. 31-34 below). EUROSTAT and OECD have also started such on-line services either through commercial service bureaux or directly to users of their statistics (see paras. 62-66 below).

II. ISSUES IN COMMUNICATIONS BETWEEN NATIONAL AND INTERNATIONAL STATISTICAL AGENCIES

A. Demand for international statistics

20. At the international level, as well as at the national level, there is an increasing demand for statistical information. Discussions, negotiations and decision-making are becoming more complex and global, requiring statistical inputs that are more comprehensive in terms of both geographical and subject-matter coverage, and more integrated, timely and complex than before.

21. Particularly in the area of international co-operation for economic and social development, access to relevant statistics is very important in order to determine status and monitor progress. International conferences, seminars and other forums have almost always stressed the importance of improved statistics and have included a statistical summary or inter-country study as basic reference material or sometimes as a central point of discussion.

B. Increased expectations from information technology for improved data communications

22. A number of developments have converged to increase expectations for improved data communication. First, there is the growing demand for comprehensive international statistics. In addition, the components of an international statistical information system that provide an appropriate response to that demand have developed greatly. Over the past decade, there has been rapid growth in the number of data bases and information retrieval systems at international agencies and accelerated development of information systems at national and regional levels. These data bases and information systems very often include international statistical data whether or not they are specifically termed "statistical" or are

managed by the statistical services of agencies. Data communication networks are expanding every year and now reach almost every part of the world. The cost of communication is falling and the speed and reliability of transmission have improved. A related development is the successful implementation of a number of information systems supported by appropriate communication technology and by business or public end-user organizations. For example, it is not unusual today for a company to transmit data from one country to a branch in another and for the branch to do the necessary computing and transmit the results back.

23. It is important to note, however, that information technology is not the panacea for all problems of statistical data communication. The availability of useful statistics to planners and decision-makers will not be directly improved by advances in computing technology. Advanced information technology only provides a means by which the relevant information can reach users efficiently and effectively.

C. Availability of internationally comparable data

24. Although there has been a steady increase in the availability of internationally comparable data, there is a wide range of statistics that would be invaluable to users at national and international levels if they could be made more rapidly available in comparable form on a world-wide basis. The problems of attempting to compare national accounts and related data compiled according to national conventions have been one of the main issues in the international comparability of statistics. International tables on various economic or social statistics are characteristically accompanied by extensive footnotes which explain deviations from or exceptions to standard definitions and/or methodology.

25. It is somewhat surprising to those who are accustomed to developments in business communications or in information technology that the availability of internationally comparable data remains so limited in a number of important respects.

26. National needs obviously command first priority in national statistical offices and international needs are taken up when resources are available to reassemble data into whatever form is requested by international statistical agencies. As a result, the response to international questionnaires, as well as the subsequent availability of data, may be delayed.

27. While the increased transfer of data in machine-readable form has speeded up the communications between national and international statistical agencies, the greatest obstacle in some areas is that only limited progress has been made towards harmonization of statistical concepts, definitions and classifications used by countries in reporting statistics to international agencies.

D. Necessity of standards for the exchange of data in machine-readable form

28. The computing counterpart of statistical harmonization is the standardization of data formats. At present, when countries provide data in machine-readable form,

/...

they use their own particular format and rely on each international agency to convert the returns to its own format. For example, in the collection and management of international trade statistics, the Statistical Office of the United Nations Secretariat expends considerable time and effort in dealing with the wide range of data formats used by Member States in providing data. In fact, it was observed that among the tapes provided by countries, no two were the same. Differences exist in physical recording modes (owing to different hardware), logical structures, codes and classifications. As a result, a host of special routines had to be worked out to deal with the various idiosyncrasies of tapes generated by different countries. At the international level, despite some improvement, there is still considerable need for standardization. A simple but basic example is the use of country codes. Despite the long-time existence of the standard alpha codes of the International Organization for Standardization and the numeric codes of the Statistical Office of the United Nations Secretariat, there are many data bases and magnetic tapes produced by agencies with different codes.

E. Special needs in international statistics for meta-data and footnotes

29. One problem related to effective and accurate data communication is how to transfer so-called meta-data or data about data. The meta-data relate to the information system itself, indicating, for example, what data exist and how they are coded and classified. In official statistics, especially at the international level, there is a need to store explanatory texts along with the data and it must be possible to retrieve and publish the text automatically when the data are retrieved. This is of critical importance, when allowing external users access to an international agency's data base, to ensure that the explanatory text including footnotes accompany the data retrieved. Until now, no standards for meta-data and footnotes have been developed and additional testing and development work will be required in a number of agencies to ensure that footnotes can routinely be made available along with the associated tabular information.

III. DEVELOPMENT OF DATA BASES AT INTERNATIONAL STATISTICAL AGENCIES

A. Statistical Office of the United Nations Secretariat

30. For more than 20 years, the Statistical Office of the United Nations Secretariat has provided a variety of statistics from its data base to a wide range of users in machine-readable form. These data bases are stored mainly on magnetic tapes (see para. 18 above) but an increasing number of data sets are kept on disks of the computer at the New York Computing Service (NYCS). The management of the data on disk is performed through the United Nations Statistical Information System (UNSYS), developed by the Statistical Office. UNSIS not only manages the storage and retrieval of data but also contains a photocomposition facility. Most of the recurrent statistical publications are photocomposed utilizing the UNSIS facility. The Statistical Office is working on adding to UNSIS a facility providing on-line access from outside the Office.

31. The Statistical Office processes international trade statistics at its International Trade Section outpost at Geneva. Data are processed by the computer of the International Computing Centre and stored in the data base at the Centre in compressed form. The data base is called the Compressed International Trade Data Base (COMTRADE). It comprises all annual series in the revised Standard International Trade Classification (SITC) from 1962 to 1985, available annual series in SITC, Revision 2, from 1976 to 1985, and available quarterly series for the past three years. The magnitude of COMTRADE is currently in excess of 1,300 megabytes for 4,000 country/periods, consisting of nine million compressed records.

32. COMTRADE is a direct access data base with an efficient and generalized retrieval system. The data base management system was developed internally by the Statistical Office (Geneva Unit). The access software consists mainly of on-line or off-line (batch) user access, providing functions of retrieval, decompression and manipulation of commodity trade statistics.

33. COMTRADE is installed on the equipment of the International Computing Centre. However, a similar version of the data base is in operation at the New York Computing Service. The scope of COMTRADE is updated weekly with new and revised country periods, and the updates are sent in a compressed format to New York via the ICC-NYCS satellite communication link. Thus the latest trade data available from COMTRADE can also be accessed at the New York Computing Service with the minimum delay. Currently, eight users within the United Nations system, two users from non-United Nations international agencies and two users from the member States have access to COMTRADE at Geneva.

34. The International Computing Centre is now linked to TELEPAC, the public data communication network of Switzerland. Therefore, theoretically, COMTRADE is available to any user around the world having access to one of the multinational switching networks. However, national or international agencies not envisaging a direct link to COMTRADE continue to receive data either in the compressed format or in one of the several decompressed formats of the Statistical Office. Internally, COMTRADE is used to compile all the trade-related publications, including the International Trade Statistics Yearbook, Commodity Trade Statistics - Series D, the World Trade Annual and its Supplement and the publications on microfiche.

B. Economic Commission for Europe

35. The work of the ECE data base has continued in accordance with the plan of work outlined in the "Implementation Programme for the ECE data base system" of 29 June 1983. In 1984-1985, the Statistical Division of ECE prepared and maintained a number of computer data files on foreign trade and national accounts.

36. ECE has commenced preparations for giving access to its data base to member countries of the region and to international organizations. Direct (on-line) access to the ECE data raises a number of policy issues concerning both statistics and electronic data processing, as well as a number of technical problems:

(a) Reorganization of data base. The contents of the present ECE data base need to be reorganized according to an ECE statistical classification or nomenclature;

(b) Data base management system. As far as the provision of direct access is concerned, the major restriction of the present status of the ECE data base is its management, as well as the lack of descriptions of the data. ECE needs a new data base management system which has the capability of providing users with access to the required data through a data dictionary as well as with the descriptions of the selected data. The development of data communication facilities is also required;

(c) Access to the International Computing Centre. ECE utilizes the computer facilities of the ICC. The Centre is linked to TELEPAC. Therefore, the ECE information system could be available to any user having access to one of the public data communication networks. For this, preliminary arrangements would have to be made with the Centre with regard to registration of account numbers, names, passwords and so on. Member countries not envisaging a direct link to the Centre's computer could receive data on magnetic tape.

C. Economic Commission for Latin America and the Caribbean

37. The development of data bases of national accounts and statistics on the production of goods and on social development, in particular, has been hampered by the fact that this type of information has not been computerized. Communications with countries are also difficult because of the widely differing extent to which recommended international definitions and classifications are used and because of the differing techniques used in the production of statistics.

38. The situation is quite different with regard to communications with such subregional bodies as the Latin American Integration Association and the Central American Common Market and with such international organizations as IMF, the United Nations Industrial Development Organization (UNIDO) and FAO. These organizations provide information on tape concerning foreign trade, the balance of payments and production, organized in such a way as to permit the development of specialized data bases. These data bases are subsequently used to form data bases that provide a high degree of flexibility in the preparation of statistics for substantive projects. One example is the Foreign Trade Data Bank for Latin America and the Caribbean (BADECEL), which contains information organized in accordance with the principal international classifications, SITC, Revision 1 to 5 digits in quantities and in values expressed in United States dollars; trading-partner zones are shown, and direct on-line access to the information is provided. A data base on household surveys (BADEHOG) and one on annual series (BADEANU) have also been organized, although the latter is at an early stage of development.

D. Economic Commission for Africa

39. The data collected from member countries of ECA are processed by the ECA Statistical Division, as far as possible in accordance with international

recommendations, and important deviations are indicated when they are known to exist. In particular, international trade statistics for ECA members are processed in standard formats in collaboration with the Statistical Office of the United Nations Secretariat.

40. To facilitate the exchange of data between national statistical offices and the ECA Statistical Division, a programme on the development of a regional statistical data base (the ECA Statistical Data Base) and interconnected national statistical data bases was established. The ECA Statistical Data Base was designed to comprise three levels: level I provides country profiles of one to two pages on member States; level II contains the bulk of the statistical data, most of which will be at a low level of aggregation, and serves as a data source for level III; level III contains selected statistical time series.

41. Limited resources made it difficult to start the development of level II. Priority was therefore given to level III which contains at present more than 110,000 time series on a number of subject fields; certain inputs, however, continue to be lacking for the satisfactory development of this level.

42. Investigations were made for linking the ECA information system to other international statistical agencies across communication lines. The link through telephone lines was found to be very costly and rather unreliable in the region. The problem was worse at the country level. The malfunctioning or non-existence of telephone lines within and among countries in the region makes it difficult to envisage such possibilities in the near future, despite the potentially tremendous impact that such links could have on the statistical services, particularly in terms of the timeliness of data.

E. International Labour Organisation

43. The ILO Bureau of Statistics regularly receives data on labour statistics from member and non-member countries through specialized questionnaires, national publications and other communications. Since 1981, these data have been stored and updated in a data base (LABORSTA) on disk, using the computer at the International Computing Centre as host. LABORSTA is divided into SAS (Statistical Analysis System) data libraries by subject and is designed to operate in an interactive manner using the SAS package.

44. At present, LABORSTA is primarily used in the production of ILO statistical publications. It is used in the production of the Supplement to the Bulletin of Labour Statistics and the monthly memo to the Statistical Office of the United Nations Secretariat, as well as in some analytical studies. At present, there are approximately 3.5 million data items, of which one third are numerical data and two thirds are the codes used for defining and accessing the numerical data. They are stored in approximately 115,000 SAS observations on some 80 essentially different SAS variables. The total size of this data base is about 10 to 20 megabytes.

45. The data entry and update of LABORSTA is done interactively using SASFSP (SAS Full Screen Product) which provides an interactive menu-driven facility for data

entry, editing and retrieval. The data for the Yearbook of Labour Statistics is updated using a computerized pre-filled questionnaire. The questionnaire is sent to all countries, on a yearly basis, for updating. The data for the Bulletin are updated on a continuous basis using national publications and replies to special correspondence.

46. ILO plans within the next two years (a) to make LABORSTA available on-line for internal use (and to the international agencies connected to the International Computing Centre) using SASIAF software; (b) to make it available on-line to the public using commercial hosts; (c) to increase the volume of downloading and use of LABORSTA data; and (d) to make LABORSTA available on a master tape with an updating service, complete with documentation and data cataloguing.

F. Food and Agriculture Organization of the United Nations

47. The AGROSTAT Information System of the FAO Statistical Division is, in an advanced stage of development. It aims at disseminating on-line to internal and external users validated statistical information compiled by various units of FAO. Availability on-line will be restricted to internal users until FAO becomes part of an international computer network.

G. World Health Organization

48. Currently, WHO maintains data bases on mortality by cause of death, health resources (hospitals and health establishments, health manpower), morbidity from infectious diseases (age and sex distribution, seasonal distribution) and the global indicators (at national, regional and global levels) for monitoring progress towards Health for All by the Year 2000.

49. Each year, WHO receives mortality data according to cause of death from about 60 countries or territories. In about one third of cases, data are transmitted to WHO on magnetic tape according to a standard format. For the remainder, the data are transmitted via a standard questionnaire which is then entered into the computer data base at WHO headquarters at Geneva. In some cases, the questionnaire is completed at WHO headquarters by extracting data from statistical yearbooks or annual health publications. In the case of magnetic tapes, data are transmitted to WHO according to either the 3-digit or the 4-digit detailed list of the ninth revision of the International Classification of Diseases (ICD-9) but are aggregated at WHO into the Basic Tabulation List recommended for compiling cause of death data. The standard ICD-9 mortality questionnaire is a mirror image of the Basic Tabulation List.

50. Modern technologies such as data transfer through communication networks or direct on-line access to national and international data bases are not yet used by WHO in the field of statistics. Rather, magnetic tapes are used to exchange large data sets. In 1985, for example, 15 tapes were copied from the WHO data bank, of which 3 were sent to national health authorities, 7 to official health institutions, 3 to intergovernmental or non-governmental organizations and 2 to

private researchers. In some cases, especially for smaller data sets, data are transferred via microcomputer diskettes.

H. World Bank

51. The World Bank has a formal responsibility for collecting information on external debt from its borrowing member countries. For this purpose, it operates a Debtor Reporting System. The data are generally reported by national ministries of finance or central banks. National statistical offices are not involved in these collection efforts.

52. All countries that borrow from the World Bank or its affiliate, the International Development Association, agree to report details of every long-term loan that is obtained by a public agency or by a private agency with a public guarantee. Private sector debt not guaranteed in the debtor country is reported in aggregate. Collectively this reporting system is known as the Debtor Reporting System. Under an agreement with other international organizations affiliated with the United Nations, the information collected by the World Bank, aggregated to a level to ensure confidentiality, is shared with IMF and other international organizations. As a service to both borrowers and lenders and with the agreement of reporting countries, summary information is published annually in World Debt Tables.

53. The Bank has for several years attempted to upgrade the quality and timeliness of the data through a series of activities. These have included technical missions to member countries, the placement of technical advisers, support for the computerization of debt recording systems, and the encouragement of data reporting in machine-readable form. The issue of comparability and consistency of data and underlying concepts between the agencies has been addressed through the group of Debt Compilers made up of the World Bank, IMF, the Bank for International Settlements and OECD. Some degree of standardization has been achieved and attention has been paid to maintaining a consistent approach vis-à-vis international guidelines on external transactions (see, for example, the Balance of Payments Manual published by IMF).

I. International Monetary Fund

54. Procedures for the collection of data from national statistical services by the Bureau of Statistics of IMF continue to rely primarily on the submission of reports and questionnaires. The reporting of data and the resolution of outstanding data and methodological problems are effected on a monthly, quarterly or annual basis, as required for the Fund's operational purposes and for inclusion in the publications of the Fund. Some 20 countries are now submitting data on a regular basis on magnetic tape. In addition, for one of the publications, completed questionnaires are now submitted by a number of countries via diskettes.

55. With regard to improving currentness, and within the limits imposed by resource constraints and the need to avoid, as far as possible, placing additional

reporting burdens on member countries, efforts have centred largely on taking advantage of new technology, including the use of telefacsimile equipment. High priority data series are transmitted via cable/telex services.

56. In order to accommodate the increasing use of data transmission to and from national agencies via magnetic tape - and, further ahead, the use of computer-to-computer links - IMF is giving some priority to establishing standards for data exchange that will meet the needs of international comparability established in IMF methodologies in the fields of balance of payments, government finance statistics, and money and banking statistics, while accommodating the reporting requirements of individual countries. For this purpose, close attention is to be given to standards for data exchange established in the international agencies.

J. United Nations Industrial Development Organization

57. The UNIDO data base for industrial statistics, developed and maintained by the Statistics and Survey Unit, Division for Industrial Studies, contains information for 144 countries. The time series for each country is an annual one, spanning the period 1963 to the latest year. Data have been collected for five fields: value added, gross output, employment, wages and salaries, and production indices. For each country, year and field, this information has been compiled for up to 28 industries in the manufacturing sector. The coverage for 60 countries is regarded as complete; that is, statistics are generally available for all years, for each of the five data fields and for the 28 industries in the manufacturing sector.

58. After UNIDO receives data, the information passes through several stages, all contained within the data base. At each stage, the data is evaluated and analysed in relation to a specific source of supplementary information - namely, national industrial censuses and surveys, UNIDO field work, input-output tables and data compiled by a number of other international and regional organizations. Similarly, the methods of adjustment and estimation carried out in the later stages are dependent on the results achieved in the preceding stages.

59. Based on the accumulated experience of past years, UNIDO redesigned its data structure in 1984. The ADABAS data base management system was chosen and UNIDO has developed an on-line query language.

K. General Agreement on Tariffs and Trade

60. The GATT secretariat has prepared and is maintaining several data bases. One of the main activities is the maintenance of the tariff and trade data file, in which detailed import statistics are combined with customs tariff data. Twelve markets are included in the file.

61. Access to the basic data stored in the tariff and trade file is restricted to those contracting parties participating in the exercise. Access to the other data

bases, such as inventory of non-tariff measures, quantitative restrictions, and technical barriers to trade, is given on request to delegations but circulation of these data on a large scale has never been envisaged.

L. Organisation for Economic Co-operation and Development

62. OECD provides most of its published data on magnetic tape. In addition, a customized extraction service is offered for external trade statistics. In order to meet the demand for OECD statistics from member States, OECD negotiated a special agreement with a time-sharing company with its own private telecommunication network. Under the terms of agreement, unlimited free access to OECD data bases is provided to almost all OECD member States via one telephone number in each country. As access is gained through the company's private network, the communication cost of access is limited to the cost of a local phone call. In return for this service, OECD provides a full set of its data to the company, together with permission for it to offer the data on-line to its own clients.

63. OECD has also started to disseminate small amounts of data on-line to a limited number of government agencies in member countries. Microcomputers are being used, both for on-line data capture and for on-line dissemination. This project is still in the process of development but it is planned to expand the facilities as rapidly as possible, depending upon the availability of the appropriate hardware and software.

64. OECD started to disseminate data on diskettes at the end of 1985. Five different data bases are offered, including Main Economic Indicators (MEI). The data are provided on 5 1/4-inch diskettes suitable for IBM or IBM-compatible microcomputers and are formatted on both sides. The data are in an OECD specific format and the diskettes include a simple programme for transferring the data into DIF, SYLK or LOTUS.WKS formats. The data can therefore be readily used with software packages such as LOTUS 1-2-3, VISICALC, SYMPHONY and MULTIPLAN.

M. Statistical Office of the European Communities

65. EUROSTAT provides a wide range of statistics to its member States, international organizations, universities and other users through the provision of magnetic tapes and on-line access to its numerical data bases: CRONOS (and its associated documentary base CADOS), COMEXT, REGIO. A brief description of each follows:

(a) CRONOS contains the bulk of the time series on some 24 domains or major subject fields covering the whole of the EUROSTAT statistical system: national accounts, industrial, agricultural, social, external trade and regional statistics obtained from surveys, questionnaires and magnetic tape. Currently, this base contains approximately 1.5 million series. It is also used for production (validation, calculation, derivation and estimation) and for dissemination on-line or in the form of hard copy. As from early 1986, some 750,000 time series of the CRONOS data bases are available on commercial hosts: CISI (France), GSI-ECO (France) and Datacentralen (Denmark);

(b) COMEXT is a data base of external trade containing quarterly data in NIMEX classification. This data base is distributed by CISI;

(c) REGIO is a data base of regional figures and, like CRONOS, covers all aspects of the economy and society;

(d) CADOS (catalogue documentaire statistique) is a documentary data base containing CRONOS meta-data and definitions and comments. CADOS provides a data base management function by checking the presence of a title, documented data, the date of latest updating and so forth.

66. The CRONOS software, developed within the Commission of the European Communities itself, was designed to allow on-line access to the data base primarily by Commission departments. Currently, over 250 terminals located throughout the Commission departments at Brussels and in Luxembourg give direct access to CRONOS.

67. EUROSTAT also provides free direct access to some of its data bases on its own computer centre in Luxembourg to privileged users, such as member States, over the public networks. Experiments are currently being carried out by EUROSTAT for the collection and transmission of external trade data by new methods of telecommunications.

IV. MEASURES TO IMPROVE COMMUNICATIONS BETWEEN NATIONAL AND INTERNATIONAL STATISTICAL AGENCIES

A. Efforts to improve data comparability

68. International comparability is an important element in the usefulness of international statistics. International statistical organizations, under the guidance of the Statistical Commission, have long strived to achieve a high level of comparability of international statistics through the development of methodologies for data collection, standards, classifications and processing in various fields of statistics. These methodologies provide a basis for comparability among statistics of member States. The wide range of efforts of international organizations related to technical co-operation and the co-ordination of statistical activities and programmes also serves to promote the international comparability of data. These efforts are more fully considered in the reports dealing with methodological issues, co-ordination and technical co-operation that are before the Commission at its present session.

B. International co-operation for setting standards for data communications

69. The exchange of machine-readable data using standardized procedures might be considered the prerequisite for efficient data communication among national and international statistical agencies. This involves establishing common procedures on formats, coding, classification, handling of footnotes and so on.

70. The international statistical agencies, recognizing the implications of changes in technology on data communications, established a Technical Working Group on Statistical Data Bases in 1982 comprising members of the ACC Sub-Committee on Statistical Activities to consider technological developments and co-ordination in data base developments and data communications. The terms of reference of the Group are as follows:

(a) Summary of current data base experience in statistical organizations.

The Technical Working Group should attempt to summarize the experience to date of statistical users with the view to determining what the special requirements in a data base system for statistical work are. The investigation into present experiences with statistical data bases should result in a summarization or synthesis that highlights the special needs of statistical data;

(b) Guidelines for a statistical data base system. The Technical Working Group should attempt to formulate standards for such data base systems. The standards formulated by the Group should refer to the various types of statistical data in machine-readable form and file structures best suited to handling each type;

(c) Standards for recording data in statistical data bases. The Technical Working Group should propose some standardization in the recording of statistical data in a data base. Internationally recognized classifications should form the basis for all encoding of data but much more attention needs to be paid to the present real usage of such classifications and attempts should be made to formulate better methods of coding the exceptions to and deviations from the standard so that their presence in a data base is immediately recognized and accounted for by any remote user or by the software system itself;

(d) Interchange of data. The Technical Working Group should proceed to determine procedures for the interchange of data between data bases, taking into account not only the technical modalities but such aspects as confidentiality, reliability and maintenance of the integrity of the data in a data base;

(e) Data base and data communications information service. The Technical Working Group should institute an ongoing procedure to acquaint all statistical organizations with the availability and practicality of hardware and software suitable for statistical data bases and the exchange of information between data bases.

71. The Technical Working Group has met annually since 1983 to discuss the issues relating to data communication and statistical data bases. In 1986, the Group produced the "Inventory of computerized statistical data bases of international agencies", compiled from information supplied by the members of the Group. The inventory is intended to be updated on an annual basis.

72. The Group also discussed the issues of standardization of coding, recording formats, footnotes and so forth. The Group agreed that standardization of coding was very important but pointed out that the problem lay not in the use of different numeric or alpha codes but rather in the exact coverage of the characteristics being coded. A proposal was made that, as an experiment, all the various code sets

used by all organizations for countries and regions should be related to one central list of such entities. The central list would be maintained in such a way as to exclude duplication but permit significant variants of entities. The Group was of the view that it was too soon to suggest that organizations should take the additional logical step of using a simple standard set of codes for countries and areas for the transfer of data between organizations (it should be noted that in different fields of statistics different basic geographical reporting units are sometimes used by Governments in compiling their own statistics and in international reporting). With regard to the problem of footnotes, the Group recognized that footnotes were frequently unavailable in machine-readable form for the data received. Further, the problem was an internal one for the data base designer and more work within organizations would be needed.

73. It is expected that the Group will continue to meet annually and attempt to solve technical problems in data communications and in the development of data bases.

C. Other measures

1. Measures to improve data collection

74. The major concern of national statistical services with regard to data collection by international statistical agencies has been the reporting burden borne by countries. This problem will continue to be addressed by the international statistical agencies in their efforts to avoid the duplication of questionnaires. Suggestions to improve the collection of data through questionnaires are summarized below.

75. Timing of questionnaires. Several countries have suggested in the past that the burden of responding to questionnaires from international statistical agencies could be reduced if more time were given for the preparation of replies. The deadlines for submission of replies may have been too stringent.

76. Model questionnaires. To facilitate correct entries, it was suggested by the statistical offices of developing countries that model entries (samples) and detailed but easy-to-understand explanations should be provided with the questionnaires. In particular, they should be aimed at developing countries and should explain how to fill out questionnaires in situations where the classification of data is partially or even entirely different from the standard classifications on the questionnaire. For instance, the data coverage might be incomplete, the accuracy questionable, the time-series not continuous, the base year for indices different or the currency used for valuation different.

77. Questionnaires on diskettes. Since microcomputers are already widely used among national statistical offices, it was suggested that sending a questionnaire recorded on diskette could facilitate data entry on to the questionnaire. This would save international agencies from having to enter data and contribute to more accurate and more rapid processing of national data. The ACC Sub-Committee on Statistical Activities, at its twentieth session in June 1986, requested the

Technical Working Group on Statistical Data Bases to study the use of diskettes as questionnaires for the collection of data from countries. The study would involve identifying appropriate types of statistics, and suitable hardware and software for recording questionnaires and data entry.

2. Measures to improve data dissemination

78. It is recognized that international statistical agencies must give continuous attention to the timely dissemination of statistics to users, including the national statistical offices that contribute national data to international statistical agencies. In this connection the utilization of various media, particularly those in machine-readable form, for dissemination is being explored in order to further improve timeliness.

79. Dissemination of data on diskette. Diskettes can be used for dissemination of statistics in numerical form (or in binary form) or text form. Spread sheet software or data base software can be used to create a statistical data file on diskette. A diskette or a set of diskettes may contain a whole publication including statistical tables with their full descriptive parts, such as title, table headings, stubs, footnotes and technical notes. At present, the most widely used diskette is a 5 1/4-inch 360 KB diskette (for the IBM PC or compatible microcomputer) which may contain about 100 pages of medium-sized documents. Two other types of diskettes are also becoming popular: the 3 1/2-inch 720 KB diskette, which provides twice the capacity of the 5 1/4-inch 360 KB diskette; and the 5 1/4-inch 1.2 MB diskette (for the IBM PC/AT or compatible microcomputer) with three times the capacity; that is, nearly 300 pages of documents or statistical tables.

80. Optical disk. A technological breakthrough is taking place in the storage device for microcomputers which would dramatically improve the efficiency of data dissemination through microcomputers within a few years. A compact optical disk based on entirely new technology for data recording/retrieval can hold up to nearly 500 megabytes (or 500 million characters) on a 5 1/4-inch disk. This is more than 1,000 times the capacity of the current 5 1/4-inch 360 KB magnetic floppy disk and is equivalent to 150,000 pages of documents. At present, the optical disk available on the market is either the read-only device (CD-ROM, Compact Disk-Read Only Memories) or the write-once, read-mostly device (WORM). It is expected that, within five years the optical disk will become as easy to use and as common as the current magnetic floppy disk, thus providing for the efficient transfer between microcomputers of a large amount of recorded data.

81. On-line access. Dissemination of statistics through on-line access to the data base will provide improved services to users. A very promising recent development in this regard is the expansion of public data communication networks through which microcomputers can be utilized for the exchange of data between statistical organizations. The Statistical Office of the United Nations Secretariat is studying the possibility of using the microcomputer-based on-line communication system for the exchange of relatively small files. Data to be transmitted would include not only statistical data but also technical notes and other descriptive information.

V. POINTS FOR DISCUSSION

82. The Statistical Commission may wish:

(a) To comment on the current status of data communications between national and international statistical agencies;

(b) To comment on the issues and relevance of new technology for improvements in data communications;

(c) To comment on the development of data bases at international statistical agencies;

(d) To suggest effective ways to improve data communications between national and international statistical agencies.

Notes

1/ Official Records of the Economic and Social Council, 1985, Supplement No. 6 (E/1985/26), para. 148.

2/ Statistical Papers, Series M, No. 56, Rev.1 (United Nations publication, Sales No. E.81.XVII.6).
