

ANNEXES

ANNEX 1 - THE UNITED NATIONS FUNDAMENTAL PRINCIPLES OF OFFICIAL STATISTICS

The Statistical Commission,

- Bearing in mind that official statistical information is an essential basis for development in the economic, demographic, social and environmental fields and for mutual knowledge and trade among the States and peoples of the world.
- Bearing in mind that the essential trust of the public in official statistical information depends to a large extent on respect for the fundamental values and principles which are the basis of any society which seeks to understand itself and to respect the rights of its members.
- Bearing in mind that the quality of official statistics, and thus the quality of the information available to the Government, the economy and the public depends largely on the cooperation of citizens, enterprises, and other respondents in providing appropriate and reliable data needed for necessary statistical compilations and on the cooperation between users and producers of statistics in order to meet users' needs.
- Recalling the efforts of governmental and non-governmental organizations active in statistics to establish standards and concepts to allow comparisons among countries,
- Recalling also the International Statistical Institute Declaration of Professional Ethics,
- Having expressed the opinion that resolution C (47), adopted by the Economic Commission for Europe on 15 April 1992, is of universal significance,
- Noting that, at its eighth session, held in Bangkok in November 1993, the Working Group of Statistical Experts, assigned by the Committee on Statistics of the Economic and Social Commission for Asia and the Pacific to examine the Fundamental Principles, has agreed in principle to the ECE version and had emphasized that those principles were applicable to all nations,
- Noting also that, at its eighth session, held at Addis Ababa in March 1994, the Joint Conference of African Planners, Statisticians and Demographers, considered that the Fundamental Principles of Official Statistics are of universal significance,

Adopts the present principles of official statistics:

Principle 1. *Official statistics provide an indispensable element in the information system of a society, serving the government, the economy and the public with data about the economic, demographic, social and environmental situation. To this end, official statistics that meet the test of practical utility are to be compiled and made available on an impartial basis by official statistical agencies to honour citizens' entitlement to public information.*

Principle 2. *To retain trust in official statistics, the statistical agencies need to decide according to strictly professional considerations, including scientific principles and professional ethics, on the methods and procedures for the collection, processing, storage and presentation of statistical data.*

Principle 3. *To facilitate a correct interpretation of the data, the statistical agencies are to present information according to scientific standards on the sources, methods and procedures of the statistics.*

Principle 4. *The statistical agencies are entitled to comment on erroneous interpretation and misuse of statistics.*

Principle 5. *Data for statistical purposes may be drawn from all types of sources, be they statistical surveys or administrative records. Statistical agencies are to choose the source with regard to quality, timeliness, costs and the burden on respondents.*

Principle 6. *Individual data collected by statistical agencies for statistical compilation, whether they refer to natural or legal persons, are to be strictly confidential and used exclusively for statistical purposes.*

Principle 7. *The laws, regulations and measures under which the statistical systems operate are to be made public.*

Principle 8. *Coordination among statistical agencies within countries is essential to achieve consistency and efficiency in the statistical system.*

Principle 9. *The use by statistical agencies in each country of international concepts, classifications and methods promotes the consistency and efficiency of statistical systems at all official levels.*

Principle 10. *Bilateral and multilateral cooperation in statistics contributes to the improvement of systems of official statistics in all countries.*

ANNEX 2 - THE INTERNATIONAL STATISTICAL INSTITUTE DECLARATION ON PROFESSIONAL ETHICS

1. Obligations to society

1.1 *Considering conflicting interests:* Statistical inquiry is predicated on the belief that greater access to well-grounded information is beneficial to society. The fact that statistical information can be misconstrued or misused, or that its impact can be different on different groups, is not in itself a convincing argument against its collection and dissemination. Nonetheless, the statistician should consider the likely consequences of collecting and disseminating various types of data and should guard against predictable misinterpretations or misuse.

1.2 *Widening the scope of statistics:* Statisticians should use the possibilities open to them to extend the scope of statistical inquiry, and to communicate their findings, for the benefit of the widest possible community.

1.3 *Pursuing objectivity:* While statisticians operate within the value systems of their societies, they should attempt to uphold their professional integrity without fear or favour. They should also not engage or collude in selecting methods designed to produce misleading results, or in misrepresenting statistical findings by commission or omission.

2. Obligations to funders and employers

2.1 *Clarifying obligations and roles:* Statisticians should clarify in advance the respective obligations of employer or funder and statistician; they should, for example, refer the employer or funder to the relevant parts of a professional code to which they adhere. Reports of the findings should (where appropriate) specify their role.

2.2 *Assessing alternatives impartially:* Statisticians should consider the available methods and procedures for addressing a proposed inquiry and should provide the funder or employer with an impartial assessment of the respective merits and demerits of alternatives.

2.3 *Not pre-empting outcomes:* Statisticians should not accept contractual conditions that are contingent upon a particular outcome from a proposed statistical inquiry.

2.4 *Guarding privileged information:* Statisticians are frequently furnished with information by the funder or employer who may legitimately require it to be kept confidential. Statistical methods and procedures that have been utilised to produce published data should not, however, be kept confidential.

3. Obligations to colleagues

3.1 *Maintaining confidence in statistics:* Statisticians depend upon the confidence of the public. They should in their work attempt to promote and preserve such confidence without exaggerating the accuracy or explanatory power of their data.

3.2 *Exposing and reviewing methods and findings:* Within the limits of confidentiality requirements, statisticians should provide adequate information to colleagues to permit their methods, procedures, techniques and findings to be assessed. Such assessments should be directed at the methods themselves rather than at the individuals who selected or used them.

3.3 *Communicating ethical principles:* To conduct certain inquiries statisticians need to collaborate with colleagues in other disciplines, as well as with interviewers, clerical staff, students, etc. In these cases

statisticians should make their own ethical principles clear and take account the ethical principles of their collaborators.

4. Obligations to subjects

4.1 *Avoiding undue intrusion:* Statisticians should be aware of the intrusive potential of some of their work. They have no special entitlement to study all phenomena. The advancement of knowledge and the pursuit of information are not themselves sufficient justifications for overriding other social and cultural values.

4.2 *Obtaining informed consent:* Statistical inquiries involving the active participation of human subjects should be based as far as practicable on their freely given informed consent. Even if participation is required by law, it should still be as informed as possible. In voluntary inquiries, subjects should not be under the impression that they are required to participate; they should be aware of their entitlement to refuse at any stage for whatever reason and to withdraw data just supplied. Information that would be likely to affect a subject's willingness to participate should not be deliberately withheld.

4.3 *Modifications to informed consent:* On occasions, technical or practical considerations inhibit the achievement of prior informed consent. In these cases, the subjects' interests should be safeguarded in other ways. For example:

- Respecting rights in observation studies. In observation studies, where behaviour patterns are recorded without the subject's knowledge, statisticians should take care not to infringe what may be referred to as the 'private space' of an individual or group. This will vary from culture to culture.
- Dealing with proxies. In cases where a 'proxy' is utilised to answer questions on behalf of a subject, say because access to the subject is uneconomic or because the subject is too ill or too young to participate directly, care should be taken not to infringe the 'private space' of the subject or to disturb the relationship between the subject and the proxy. Where indications exist or emerge that the subject would object to certain information being disclosed, such information should not be sought by proxy.
- Secondary use of records. In cases where a statistician has been granted access to, say, administrative or medical records or other research material for a new or supplementary inquiry, the custodian's permission to use the records should not relieve the statistician from having to consider the likely reactions, sensitivities and interests of the subjects concerned, including their entitlement to anonymity.
- Misleading potential subjects. In studies where the measurement objectives preclude the prior disclosure of material information to subjects, statisticians should weigh the likely consequences of any proposed deception. To withhold material information from or to misinform subjects involves a deceit, whether by omission or commission, temporarily or permanently, which will face legitimate censure unless it can be justified.

4.4 *Protecting the interests of subjects:* Neither consent from subjects nor the legal requirement to participate absolves the statistician from an obligation to protect the subject as far as possible against the potentially harmful effects of participating. The statistician should try to minimise disturbance both to subjects themselves and to the subjects' relationships with their environment.

4.5 *Maintaining confidentiality of records:* Statistical data are unconcerned with individual identities. They are collected to answer questions such as 'how many?' or 'what proportion?'; not 'who?'

The identities and records of co-operating (or non- cooperating) subjects should therefore be kept confidential, whether or not confidentiality has been explicitly pledged.

4.6 *Inhibiting disclosure of identities:* Statisticians should take appropriate measures to prevent their data from being published or otherwise released in a form that would allow any subject's identity to be disclosed or inferred.

For more information see <http://www.cbs.nl/isi/ethics.htm>.

ANNEX 3 - TOWARDS THE NEW OECD STATISTICAL INFORMATION SYSTEM

1. Introduction

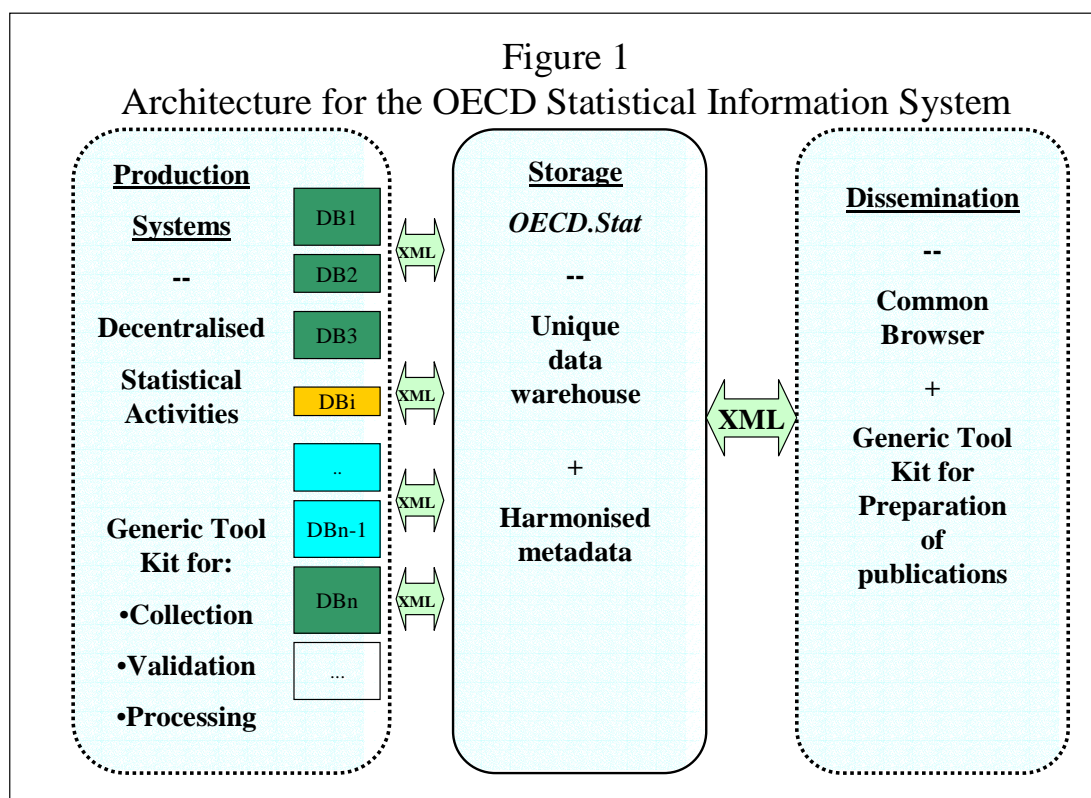
The OECD has a decentralised statistical structure where statisticians and analysts work closely together on a subject by subject basis. Because of the decentralised decision process, past decisions involving IT implementation for statistics have given priority to the efficiency of individual processes and to requirements of statisticians in individual subject matter areas. This has created a fragmented system focused on individual activities, developing optimal solutions from the producers' point of view, but sub-optimal solutions for users in the Organisation and outside.

A "vision" for a modernised OECD Statistical Information System has been developed in the context of the OECD new strategy for statistics launched in 2001. In that vision, the OECD statistical information system encompasses tools for data collection, manipulation, storage, dissemination and user tools for data discovery and retrieval. The main idea is to preserve the independence of data producers while making their data and metadata part of a coherent and seamless corporate system. Figure 1 below shows a schematic view of the architecture of the OECD statistical information system. It comprises three layers, as follows:

- a production layer for data collection, validation and processing;
- a storage layer for providing data and metadata to internal users; and,
- a dissemination layer for producing publications.

New developments in information technology offer advanced tools for managing a decentralised statistical system. In particular, IT architectures based on the use of web technologies provide the possibility for integrating data and metadata among different groups of statisticians, increasing the efficiency and overall coherence of statistical activities and products, maintaining a decentralised and flexible system. However, given the constraints of the regular production of statistics and of limited resources, the OECD needs to have a pragmatic approach and to implement the vision through incremental improvements to the existing individual production processes. Those systems need to be maintained, further developed and eventually changed in order to take advantage of new technologies or of changes in the overall technical environment. In the context of the new strategy for statistics, those tasks are conducted with users' needs and corporate considerations in mind.

- This Annex describes the progress made in that area over the last two years. In particular, new developments in the areas of data collection, data processing and database management and dissemination are described. The OECD is also using managerial levers to implement its vision of a modern statistical system and to fully realise the potential efficiency gains offered by technical changes. Finally, the Annex shows how the OECD Quality Framework and the implementation of the new infrastructures support each other to improve the overall quality of OECD statistics.



2. Improving the existing production systems

2.1 Data collection

In most cases the OECD collects its data from national statistical organisations (NSOs)³⁷ or other international statistical organisations. Data are collected via specialised questionnaires or from the data already disseminated at national level. The nature of the details and processes involved depends on a number of criteria:

- nature of data to be collected, e.g. whether or not they are part of a regular production process;
- frequency of the collection, which varies from weekly to annual or even occasional;
- regularity of the content over time;
- homogeneity of content regarding the internal structure of the NSO (e.g. subject matters);
- quantity and level of data (e.g., disaggregated versus aggregated data);
- method of collection;
- security and confidentiality requirements

³⁷ In this paper NSOs comprises all national statistical organisations, including the National Statistical Office, the Central Bank and the various ministries producing statistics.

- technology and infrastructure.

In addition to the diversity generated by those criteria, processes for data collection have evolved over the years as the result of bilateral arrangements between the OECD and national agencies or other international organisations. Of course, progress in information technology and communication has been an important driver of such evolution. Also, policies regarding arrangements vary from accepting any format and medium suggested by the data provider to imposing a fixed format and fixed medium. This Section reviews successively, progress in the areas of data collected through off-line questionnaires, through on-line databases and describes an experiment for a new way of providing data to the international statistical system.

2.1.1 Collection through off-line questionnaires

Regarding electronic questionnaires, current experience shows that good practices in data collection developed by some Directorates are applicable to other parts of the OECD. In 2002-2003, new electronic questionnaires have been designed, incorporating automatic data checks and the management of metadata. In 2003, a review of current practices will be carried out in the context of the quality reviews for several statistical activities that imply large data and metadata collections. The results will be used for further improving the tool for preparing new questionnaires, which is part of the **statistical toolkit** developed for modernising individual statistical production processes. This initiative should reduce the cost of data collection and minimise the burden on data providers. Also, since 2003 all new questionnaires should be submitted to the Chief Statistician before being sent. A central collection of those questionnaires is going to be developed to make them available to all statisticians in the Organisation.

Improving internal co-ordination inside the Organisation (through the OECD Glossary of Statistical Terms, the statistical toolkit, etc.) should improve the harmonisation of definitions of variables across questionnaires sent to countries. Organising information on data collection consistently across collections will further increase their corporate value and avoid the risks of duplication, or inconsistencies in data collection.

2.1.2 Extractions from on-line databases

For the collection of data that are already part of the regular production of NSOs (e.g. in the area of short-term statistics), the OECD has been very flexible and has preferred to adjust to formats readily available to individual data providers in order to gain on timeliness. New ICT has permitted very spectacular gains in efficiency in this domain, for both the OECD and its data providers.

The latest technique developed and implemented by the OECD in the last two years is called "Web Queries". This technique is based on the ability of Microsoft Excel to save, on the user's PC, queries made to a database through the web. When it is possible to use that technique, there is no extra development needed on the NSO side. On the contrary, web queries add at no cost a functionality that had been developed by some NSOs for their output database systems. Also, preselections are stored on the user side, which does not increase the burden on the producer side and does not require giving users the right to write into the NSO database. On the OECD side, one click on a button automatically retrieves the data needed from the NSO database. Also, since Web queries can be parameterized, modifications are easily made. The required conditions for Web queries to be used to retrieve data from one NSO database are that data can be seen on a web browser. The technique allows going through password protection. Unfortunately, some database systems in NSOs have extra layers, used for security or for commercial purposes, that prevent the technique to work.

2.1.3 *Data sharing and NAWWE*

The example of Web queries described above shows how efficient data collection is for the OECD when the data collected are published by the data provider on the Web. The “data sharing model” pushes the idea one step further in suppressing the collection by international organisations. In such a model, data available from international organisations’ web sites are not extracted from international organisations’ databases but directly retrieved from NSOs web sites on users’ request. Of course, for the model to work a central data catalogue must be maintained and some statistical as well as IT standards must be used by NSOs. This concept was presented at the 2002 Conference of European Statisticians meeting in Paris³⁸ and is further developed in the context of the Statistical Data and Metadata Exchange initiative (SDMX)³⁹. The 2002 OECD meeting of National Accounts experts proposed that an experiment be started to test the implementation of the data sharing model in the National Accounts area amongst NSOs and the OECD. The project is called National Accounts World Wide Exchange (NAWWE). This section describes the progress on the experiment so far.

The idea of the NAWWE project is to implement a model in which data are not transferred across organisations but, rather, published on the web in such a form that users can access them by simply using the country and variable references. A constraint in the project is to start from the Excel tables already produced by NSOs for transmitting national accounts data to International Organisations.

Another objective of the NAWWE project is to have the data collected by international organisations to be the data officially disseminated by NSOs. The two advantages of this model are that the burden of reporting to international organisations is minimised, and data quality is maximised for the international statistical community since the data they use are those officially disseminated and not specially compiled for and transmitted to international organisations.

In technical terms, the approach is based on XML and the associated technology standards. In succinct terms, NSOs would post data on their web sites using a standard XML based format. The files can be password protected or not. Then a directory of the corresponding XML files URLs for each country is made available centrally on the web, for example on the OECD site. The OECD would make available a “web service” permitting simple extraction of data from the system using parameters, the country identifier and the national accounts standard SNA93 code. In this model, each NSO can also use the same XML file as central source of data for its own dissemination for all formats and media. The NSO can easily mix, in their own tables, data from other countries and from international organisations for international comparisons.

OECD developed an Excel macro that allows NSOs to produce standard XML files from the existing Excel files they send to International Organisations. The first draft XML schema proposed by SDMX in April 2003 has been used to produce the NAWWE XML files. A simple Web application has been developed by the OECD which exploits the XML files from several countries together. The system simulates a database where in fact there is no database. In the next step the OECD would like to deploy the system amongst the statistical community in order to generate a discussion on the political implications of the data sharing model.

³⁸ See the IMF and OECD, “Progress Report on New Developments in Data and Metadata Collection for International Organisations”, Conference of European Statisticians, Paris, June 2002. <http://www.unece.org/stats/documents/ces/2002/8.e.pdf>

³⁹ See www.sdmx.org

2.2 *Databases and production systems*

Data at the OECD is stored in databases, dataset by dataset. Because of the decentralised decision process, datasets have been implemented over the years in a variety of systems. Major elements in the database software infrastructure used across OECD statistical areas today are:

- MS SQL-Server 2000 for data and metadata storage and cataloguing;
- MS OLAP component for multi-dimensional data-manipulation;
- ORACLE Express for multi-dimensional data-manipulation, 4GL programming and some data storage;
- FAME for time-series data manipulation, graphics, 4GL programming and data storage;
- SAS for analysis of disaggregated data;
- MS-Excel/Access for common data manipulation and some data storage.

This diversity of systems has resulted in a variety of different implementations all of which require continued maintenance and, sometimes, further developments. In fact, at present, because of a lack of resources, it is impossible to abandon existing solutions and to develop a completely new production system. It has been decided to continue maintaining existing production systems and to develop an intermediate layer between original production systems and final users, to allow the latter to navigate across datasets using a common software package. This new system is called *OECD.Stat* and will be described in section 3 of this Annex.

In the meantime, recent evolutions in the software market, notably the phasing out of the current version of Express software by Oracle, have made it necessary to replace a range of statistical production systems at the OECD. In the context of the new strategy for statistics, a generic toolkit for statistical database management has been developed (*OECD.Statworks*). This toolkit has been assembled into an application and uses standard OECD software tools (MS SQL Server 2000 relational database, MS OLAP, Excel, .NET development environment). The generic SQL Server database is also open to access via existing statistical software (SPSS, SAS etc) and 3rd party analytical tools as required. This toolkit ensures the full integration of statistical data with the corporate statistics warehouse *OECD.Stat*.

2.3 *Systems for data dissemination*

On the dissemination side of the process a lot of progress has been made to increase the global value of OECD statistics. Most electronic dissemination is standardised on a unique software package⁴⁰ and data format, used on the web as well as on CD-ROMs. However, the implementations still vary and the production processes used to create the necessary input files have been developed independently to each other. The content of data files is not harmonised, with consistent data and metadata presentations. Furthermore, because data files are independent, some data series need to be duplicated and some apparent inconsistencies can occur because of, for example, differences in data vintage and of lack of metadata.

To address these issues several actions have been undertaken, to develop corporate tools for improving the efficiency and the coherence of statistical activities. In particular, it is worthwhile to mention the new OECD statistics portal. The first OECD Statistics Portal (www.oecd.org/statistics) was launched in September 2001, in the context of a major revision of the OECD web site. The Statistics

⁴⁰ The package presently used by the OECD for data dissemination is Beyond20/20 TM.

Portal provided, for the first time, an OECD web site exclusively dedicated to statistical information. The Portal has been revised and a new version has been made public at the beginning of July 2003.

This new version has been designed taking into accounts results of usability tests conducted since the launch of the previous version. The main conclusions of those tests were that the navigation and the jargon were still complex for most users. In the new version navigation is simplified. In particular, the thematic classification is complemented by a short description with standard keywords for all themes. Also, the number of pages has been reduced and data files are available on the first click on the selected theme.

3. The new output database: OECD.Stat

3.1 General description

The preceding sections have described how the Organisation is working on improving its existing systems to incrementally implement the vision of a corporate statistical system. This section describes how the Organisation is putting together the information necessary to seamlessly integrate the output, to facilitate usage of statistics across the Organisation and to help collaboration between individual statistical units. The name of this corporate system is *OECD.Stat*.

The aims of the project are to:

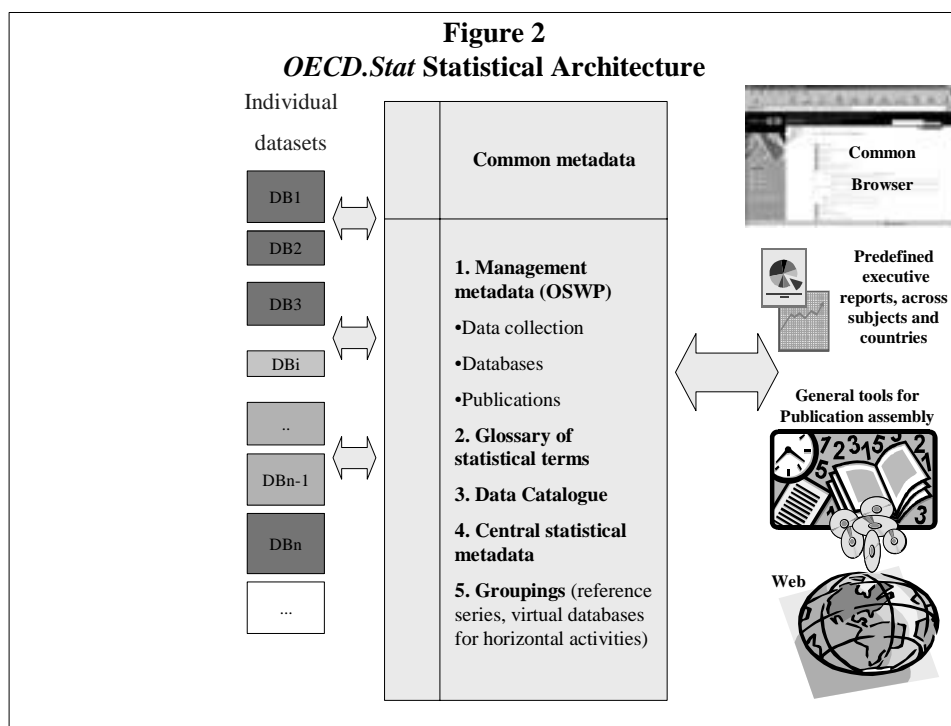
- improve accessibility to, and global visibility of, OECD statistical databases, enabling internal users to conduct quick searches and extractions from various datasets, to develop analytical and horizontal studies and create new “integrated” products;
- minimise the number of duplications of data and metadata and to improve coherence across different databases;
- provide a corporate technical framework in which the Organisation can facilitate the implementation the quality standards required by the OECD Quality Framework

In *OECD.Stat*, each Directorate contributes to a corporate “data warehouse”, where all relevant data are stored with their metadata, in order to give full and easy access to all authorised internal users. Each Directorate remains independent in determining its own statistical and analytical processing of raw and final data under the guidelines of the OECD Quality Framework, but the outputs of these processes are part of the corporate data warehouse. In practice, because of the complexity and variety of statistical activities, the system is based on a “constellation” of datasets, connected to each other and parts of a single data warehouse. The intention is to make *OECD.Stat* the unique corporate source of statistical information.

Integration of datasets in this environment is realised through metadata. The OECD Statistical Work Programme (OSWP) database, described below, and the “OECD Glossary of Statistical Terms” are parts of *OECD.Stat* and they play the role of catalysts for the integration process. The OSWP represents the top part of the metadata system, containing descriptions of datasets. The Glossary permits association of target definitions to variables in the system, improving accessibility, interpretability and coherence. Figure 2 below shows a conceptual view of *OECD.Stat*. Metadata are represented in the central box and include:

- management metadata giving detailed information on statistical activities, the OECD Statistical Program of Work (OSWP);
- a glossary of statistical terms for the harmonisation of terminology and concepts;

- a central data catalogue for the location of data in the collection of OECD datasets;



- a central metadata repository for the storage of metadata elements that are independent of individual data items; and
- information on groupings of data: publications, the set of most commonly used data series (referred to as Reference Series), and virtual databases for horizontal studies.

Metadata provide the integration and individual datasets remain independent. As stated above, each Directorate remains independent in determining its own statistical and analytical processing of raw and final data under the guidelines of the OECD Quality Framework. *OECD.Stat* is the output database. Each dataset is stored in *OECD.Stat* as a multidimensional object or hypercube. Because of the multiplicity of production systems, it has been decided to create a generic data entry module for *OECD.Stat*. The module accepts generic XML files and updates data in the data warehouse accordingly. XML files are generated independently by each production process based on a standard acceptable schema. This makes *OECD.Stat* and the existing production systems technically independent, but linked in real time to the latter.

This project is technically complex and has required experimenting with several IT solutions, and a wide analysis of data and metadata models. The sections below describe successively, various parts of the system, the way in which *OECD.Stat* and the production areas are linked, the “groupings” concept (in particular, the group of “Reference Series”), the metadata structure and the technical implementation.

3.2 *OECD Statistical Programme of Work (OSWP) and associated database*

The OECD Statistical Programme of Work (OSWP) was developed as a tool for internal coordination and external communication. The Statistics Directorate is responsible for preparing the Programme, using information provided by relevant Directorates through an electronic questionnaire for

each line of statistical activity⁴¹. The two main uses of the Programme are: for users of OECD statistics to have information on what is going on in the OECD and to enable the OECD to take decisions about existing activities (e.g. launching of new activities, or co-ordinating work on activities across international organisations). The classification used to present activities in the OSWP has been derived from the UN classification of statistical themes and is currently used in the OECD web “Statistics Portal”.

A further use of the database behind the OSWP is to facilitate access to the products and outputs of the statistical work of the Organisation and to provide information on other parts of the OECD statistical information system. Information collected by the OSWP on individual datasets is used to allow users to search for datasets by Directorate, theme, activity, related publication or keywords. Information collected for the OSWP is also used to conduct activity reviews within the context of the OECD Quality Framework. Therefore, the OSWP represents the main entry point into the OECD statistical information system.

In 2003 the OSWP database containing detailed information was made available on *OLIS.net*, the network that is normally used by all Organisations’ governmental bodies to access all OECD documents and outputs (publications, statistics, etc.). This will represent an important tool for agencies responsible for co-ordinating statistical activities, as well as for all public bodies using OECD statistics.

3.3 *Glossary of statistical terms*

Definitions of statistical data elements and concepts are an essential part of international standards in the area of statistics. Because of the diversity of the institutional, economic and cultural climate within different countries these definitions are often the result of compromises and by necessity, are fairly broad in some instances. To a large extent they can be viewed as “target” definitions requiring elaboration and adaptation to national circumstances.

Notwithstanding these limitations, it is important to promote the use of these concepts and definitions by increasing their visibility by making them readily accessible via the Internet. For this reason, the OECD has developed a Glossary of Statistical Terms, available at <http://cs3-hq.oecd.org/scripts/stats/glossary/index.htm>.

The Glossary contains both target definitions of the main variables (data elements) collected by the Organisation for use in its statistical and analytical output (e.g. for industrial production, services, unemployment) and definitions of terminology/concepts (data element concepts) used in OECD publications (e.g. accrual accounting, metadata). The 5,600 definitions included in the OECD Glossary cover a very broad range of statistical subjects. The main elements of the Glossary are:

- unique title for the definition;
- the actual definition;
- for some definitions, text providing further background on the definition, its application and relation to similar or related concepts. This field may also contain URLs to relevant documents describing appropriate use of the variable defined, etc;
- detailed source information;
- classification of each definition to a broad statistical theme;

⁴¹ A line of activity is defined as “an activity that produces at least one statistical output, such as a dataset or database available to internal or external users through Internet, Intranet, OLISNet, CD-ROM, etc., or a publication (whether classified or not) that is statistical or is an analytical publication with extensive statistical content”.

- internal cross-links to related definitions, etc., contained elsewhere in the Glossary;
- URL links to the complete source document containing the definition where this is currently located on the websites of international organisations or national agencies.

The Glossary also includes search and interrogation facilities and information describing each of the fields. One of the powerful features of the Glossary is the linkage (via URLs where available) of definitions to the full standard/source containing the definition. This enables the user to readily obtain further information on the definition. The Glossary is an essential part of *OECD.Stat*. It will permit to relate variables, to navigate the system and to improve the overall coherence.

3.4 *Groupings: reference series and horizontal products*

In the present situation, users have still to navigate the various hypercubes in order to locate their data. In order to deliver rapidly tangible benefits to analysts in the Organisation it has been decided to create, from data residing in various hypercubes, a specific dataset comprising data series that are the most frequently accessed by users who are not involved in the corresponding statistical activity. An initial list of Reference Series has been obtained through consultation with analysts throughout the OECD. The list contains the following statistics: main aggregates of national accounts, with history and forecasts; labour force and population data; exchange rates; purchasing power parities and price indices.

Analysts who are not experts in a subject matter area from which they need data have difficulties in locating those data. This is because the data they are searching are part of complex datasets containing thousands of data series. For example, GDP data are stored with all the rest of national accounts data in a complex accounting framework. In some cases, there is a risk that the wrong data is used. In the past, data corresponding to "Reference Series" were duplicated in individual databases to facilitate their use as background data to calculate ratios, per capita, etc. This was another factor for confusion and risk of inconsistencies. Reference series are also meant to define the standards in terms of associated documentation, with the objective to provide information to users who are not experts in the particular domain of the corresponding Reference series.

In technical terms, Reference Series are contained in a specific virtual hypercube taking its data from actual hypercubes corresponding to output datasets. The same technique can be used to create virtual cubes for horizontal or any analytical study.

3.5 *Metadata structure*

A very difficult part of the project is the management of metadata. Following on from the development of the OECD Glossary of Statistical Terms, which includes all main statistical definitions used by various OECD Directorates, the SDMX Common Metadata Vocabulary developed by the OECD and Eurostat, under the auspices of the SDMX project (see below), will be released in 2003. The definitions contained in the SDMX Vocabulary will be used for developing the OECD common repository of metadata, which should be completed by the end of 2003.

In the *OECD.Stat* data model, harmonized metadata are present at all levels of the data structure, including dataset level, dimensions, elements in dimensions, crossing of dimensions and their elements. Dataset information come from the OSWP, relevant target definitions are linked to the Glossary and, in the future, more corporate metadata will be added, making *OECD.Stat* the central source for all OECD statistical information. The OECD will align the metadata structure and content to those suggested by SDMX in the context of the work on metadata repositories.

3.6 *Technical implementation*

In technical terms, *OECD.Stat* is developed using Microsoft tools. MS SQL Server 2000 and its OLAP component are used as central data and metadata repository. As stated above, datasets are stored as hypercubes in that structure, with common dimensions taken from a central repository. Each hypercube is related to the relevant item from the OSWP. Elements of dimensions, in general those other than country and time are related to their corresponding entry in the Glossary. An Excel add-in has been developed to permit easy access to individual datasets and to Reference Series using Excel Pivot Tables.

4. **The reinforcement of quality through the implementation of the new architecture**

The implementation of the new statistical information system is part of the overall strategy for improving the cost efficiency of statistical activities and the quality of OECD statistics. In particular, the new system can improve several of the quality dimensions quoted in the OECD Quality Framework: relevance, accuracy, accessibility, interpretability and coherence. First, as already said, in preserving individual production processes in Directorates, the new architecture preserves the benefits of the decentralised organisation. At the same time, it improves the capabilities for cross-cutting studies by permitting navigation and selections across datasets and by permitting the creation of new virtual datasets (as is already done for Reference Series) according to user needs. At the same time, accuracy is potentially improved through more data confrontation possibilities. Therefore, the new system reinforces both the relevance and the accuracy of OECD statistical products.

Second, accessibility is the most obvious quality dimension immediately improved by the new statistical infrastructure. The architecture has been developed focusing on the needs of the final user, providing the latter with a “user friendly” tool for accessing all datasets in the system. Metadata help data discovery and retrieval.

Interpretability is improved by central tools for data documentation and by improved accessibility to metadata. The Glossary, central metadata items and a common set of metadata are the main instruments for achieving this target. Coherence of metadata is also an important factor for improving interpretability.

Coherence of data has four important sub-dimensions: within a dataset, across datasets, across countries and over time. Coherence within datasets is improved in *OECD.Stat* by the fact that data are related to the glossary which forces data concepts to be related to their official target definition and other attributes, whether internationally agreed to or not. Coherence across datasets is improved by the data confrontation allowed by the data warehouse in the improved accessibility and interpretability of data.

Finally, the OECD Quality Framework will facilitate the implementation of the new architecture which in turn will increase quality. Quality guidelines encourage the use of the new tools provided in the new statistical system and quality reviews will encourage and facilitate the full implementation of all statistical activities into the new architecture.

ANNEX 4 - A LAYOUT FOR THE NOTES SECTION OF OECD DATA SETS PUBLISHED ON SOURCEOECD

Since the end of 2001, OECD statistics have been available on-line from three sources:

- to paying clients via the SourceOECD internet service hosted by Ingenta;
- to OLISnet users via the WDS service;
- to all users via the statistics portal.

In particular, both SourceOECD and WDS are based on Beyond 20/20 technology and use the same underlying files ("IVT" format) that are currently generated for diskette and CDROM products.

Early tests on these services, carried out at OECD, highlighted apparent inconsistencies across OECD data products, particularly when using the search engine facilities. For example, (i) searches using the key words "GDP" and "Employment" can yield many possible data products for the user to choose from; and (ii) comparisons of key variables (e.g. GDP) across different data products sometimes show different values when the user might expect them to be identical. Also, identical searches on *SourceOECD* and WDS can give very different results. Finally, there was some concern about the general 'look-and feel' and ease of navigation within the Statistics section of *SourceOECD* as well as the general technical environment.

The main causes for apparent inconsistencies across OECD statistics products on *SourceOECD* and WDS were identified as follows:

- different data products with different vintages of the same time-series without the necessary documentation explaining reasons and/or giving the last update date of the corresponding file;
- different data products containing similar variable names and/or classifications without the necessary documentation explaining underlying sources and definitions and relations to other data products;
- missing or incomplete minimal metadata elements in data file (units, power of ten, etc.);
- different types of search engines used;
- misleading results from the search engines. On *SourceOECD*, this can be due to superfluous information within the abstracts for certain data products. On the WDS, this can be due to the wide scope of the searches across IVT files.

Therefore, recommendations were issued to maximise the information to users, improving the metadata associated both to individual statistical databases. In particular, it is necessary to provide a standard presentation of general notes within OECD data products (IVT files) published on *SourceOECD* and WDS. Within the framework of B20/20 technology, it makes extensive use of links to PDF, HTM and Excel files to enhance the provision of metadata. It would be highly desirable for the *SourceOECD* abstract for each product to match the corresponding description in the IVT general notes section.

The box below describes a standard layout for the notes section of OECD data sets published on SourceOECD and uses STI's STAN industrial database as an example. Note that this does not refer to the

SourceOECD abstract - the first point of contact by clients and the page interrogated by Ingenta's search engine. It would be desirable for the abstract to closely match the description in the notes section.

Basically, the information is divided into two columns with the general description on the left and links to more detailed information / metadata on the right. Notes on the various elements suggested follow.

<p>Last update : 15th May 2001</p> <p>The <i>STAN database for Industrial Analysis</i> has been developed to provide analysts and researchers with a comprehensive tool for analysing industrial performance at a detailed level of activity. It includes annual measures of output, labour input, investment and trade that allow users to construct a wide range of indicators to focus on areas such as productivity growth, competitiveness and general structural change and make comparisons across countries. A standard industry list is used that provides sufficient detail to enable users to highlight high-technology sectors.</p> <p>STAN is primarily based on member countries' annual National Accounts tables and uses data from other sources, such as national industrial surveys/censuses, to estimate any missing detail. Since many of the data points are estimated, they do not represent official member country submissions.</p> <p>Notes for previous users of STAN :</p> <p>This new version of STAN is based on ISIC Rev. 3 (compatible with NACE Rev.1) and has been expanded to cover all activities (including services) and a wider range of variables - it has effectively been merged with OECD's International Sectoral Database (ISDB) which is no longer updated. STAN is now updated on a 'rolling basis' (i.e. new tables are posted as soon as they are ready) rather than annually - thereby improving timeliness.</p> <p>Recommended citation : <i>OECD, STAN database, 2001</i></p>	<ul style="list-style-type: none"> • News • Corrections [Optional] • Full documentation • Current data coverage [Optional] • Recommended uses and limitations • Relationships to other OECD data sets • Examples of use in publications and research papers [Optional] • Compatible data sets for further research [Optional] • Useful link [Optional] • Technical help [Optional] • Contact
---	---

1. General notes

- **Last update** - useful to be highlighted as it immediately allows regular users to see if any updates have occurred since their last visit.
- **The general description** - should be clear and concise and approved by Editorial Board for Statistics to ensure that it is up to date and uses terminology consistent with other data sets.
- **Notes for previous users [Optional]** - In cases where major changes have occurred to the data sets, it could be useful to alert regular users.
- **Recommended citation** - useful to ensure that an OECD data product is cited in a consistent and recognisable manner around the world.

2. Links to detailed information

A standard set of links would provide users with easy-to-access information concerning the data set as a whole. (More detailed country and variable notes will be found in the 'dimension notes' within the data

product's 'selection' pages.) Some categories would be optional, though they should have consistent titles across data products. The links generally refer to PDF or HTM files, though in some cases Excel files may be preferable. Although some work will be required in the short-term to set up such lists and regular updates required for some of the categories, the results should make it worthwhile. Suggestions for other categories or different titles for those below are welcomed.

News - could include outlines of revisions and updates since the last version; descriptions of new tables; changes in underlying methodologies; links to recent related publications/ methodological manuals; recent major citations and any other new information that may interest users.

Corrections [Optional] - In the (unlikely) event that erroneous or misleading figures have been published, this link would point to a page outlining the nature of the error(s); an explanation; how long the error has been in the data set; and a humble apology.

Full documentation - This would be a link to a PDF file, which users can print out to read at their leisure, containing the usual detailed sources, definitions, methodological notes, estimation techniques and other detailed metadata. Some of the information will also be accessible separately from the data product notes above (such as recommended uses and limitations and relations with other data sets) and from the dimension notes (such as variable definitions).

Current data coverage [Optional] - Many users find data coverage tables helpful, particularly for quickly determining, for example, to what extent they may make aggregates across countries, or more generally, the usefulness of the data product for the type of analysis they had in mind. [HTM, PDF or Excel file depending on detail provided] *In STAN an Excel file is used with one page per country giving the period for which data is available for each variable x industry combination.*

Recommended uses and limitations - It is important to guide users with limited knowledge of the statistics presented and to help them determine whether the product meets their requirements. It could contain examples of the type of indicators that can be constructed and/or inferences that can be made; the types of analyses that can be performed; the type policy questions it can help answer; the 'shelf-life' of the data product etc. These notes could also explain when caution should be employed, what type of calculations should be avoided and the type of inferences that should not be made.

Relationships to other OECD data sets - If the data product uses other OECD databases as inputs, these should be mentioned here (and possibly linked to). If there is any possibility that duplicated data may not match - reasons should be made clear. If there are other OECD data products containing identically named (and/or defined) series, reasons for differences should be provided. Criticism of other OECD data products is to be avoided! [HTM page or PDF file] *STAN uses STD's ANA, SSIS, IIA and ITCS databases as inputs (as well national sources) and has variable and industry descriptions that to varying degrees match those in ANA, SSIS, IIA and STD's "Services: Statistics on Value Added and Employment" publication. Therefore discussion is required on the difference between National Accounts and Survey/Census data; detailed v. aggregate National Accounts; Short-term v. Annual measures; product v. industry classifications, etc.*

Examples of use in publications and research papers [Optional] - an opportunity to publicise the benefits of the data product by giving titles (and if possible links) to any publications and research papers, (not just OECD) that have made extensive use of the data.

Compatible data sets for further research [Optional] - mention (and provide links to) any other data set (OECD or otherwise) that can be used in conjunction with the data product. *The STAN family of databases also includes databases containing R&D expenditures (ANBERD), bilateral trade by industry (BTD) and Input-Output tables (I-O) linked by compatible industry lists. These would be mentioned here.*

Useful link [Optional] - provides links to websites that may enrich use of the data product. For example, methodological manuals, international standard classifications, academic research, future developments in the field, data for non-OECD countries etc.

Technical help [Optional] - could provide hints on navigating around the data product within SourceOECD. Useful if advanced features are being used and/or data organised in a 'non-standard' way. It could provide, for example, best practice for downloading large selections of data.

Contact - As well as providing contact details, a note encouraging users to provide comments and suggestions could be included.

ANNEX 5 - AN EXAMPLE FOR THE PREPARATION OF THE SELF-ASSESSMENT IN THE CONTEXT OF QUALITY REVIEWS OF EXISTING ACTIVITIES

A. An overview of the STAN⁴² database today

The STAN database is designed to provide analysts and researchers with a comprehensive tool for analysing industrial performance at a relatively detailed level of activity and carry out comparative studies across countries.

It includes annual measures of output, labour input, investment and international trade in goods which allow users to construct a wide range of indicators to focus on areas such as productivity growth, competitiveness and general structural change. The industry list provides sufficient detail to enable users to highlight high-technology, high-growth and 'knowledge-based' sectors and is compatible with those used in related OECD databases such as the other data sets in the "STAN family" – notably ANBERD (R&D expenditure by industry), Bilateral Trade Database (BTD) and Input-Output tables⁴³.

STAN is primarily based on Member countries' annual National Accounts by activity tables and uses data from alternative sources, such as national industrial/business surveys, to estimate missing detail. Trade data by industry are derived from data classified by products using a standard conversion process.

The latest version of STAN is based on ISIC Rev. 3 and covers all activities (including services) - it has effectively been merged with the OECD's International Sectoral Database (ISDB) which is no longer maintained. STAN is now updated on a "rolling basis" rather than published as an annual "snapshot".

The STAN database is managed by the Economic Analysis and Statistics (EAS) division of the Science, Technology and Industry Directorate (STI). Its development since the early 1990s has been overseen by the Statistical Working Party (SWIC) of OECD's Committee for Industry and the Business Environment (CIBE).

Detailed information including full documentation, industry and variable lists, details of data availability and sample tables, can be found at www.oecd.org/sti/stan.

The following summary and outline of main issues were derived from the first part of the quality review process, the "self-assessment" by data managers, carried out in early 2003. Comments and suggestions from SWIC delegates, OECD colleagues and a wide range of STAN users, including those in national statistical offices (NSOs), have also been taken into account.

B. Attributes of STAN identified in the Quality Review

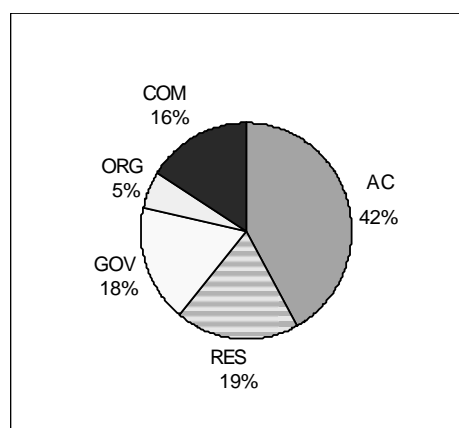
B.1 Perceived strengths

- **High demand** – STAN is widely used by OECD, national administrations and other international organisations and by a variety of academic and research institutions – see below for illustration of user profiles based on a sample of over 200 different institutions. There is high interaction with

⁴² The name STAN is derived from SStructural ANalysis.

⁴³ The other data sets in the "STAN family" were not considered in this quality review.

users via a STAN users distribution list which encourages regular feedback and ensures ongoing user needs are understood;



COM = commercial enterprises

AC = universities

RES = research institutes

GOV = government agencies

ORG = international organisations.

- **Proximity to analysts** - STAN benefits from being maintained within an analysis-oriented environment. The STAN data manager assumes the role of user when contributing to publications such as OECD's *Science, Technology and Industry Scoreboard*⁴⁴. Also, STAN is ultimately overseen by an OECD policy committee, namely CIBE, and has thus developed to address issues forced by this group such as economic growth, globalisation, structural change and competitiveness;
- **Low demands on national data providers** – The request for data for STAN, which usually accompanies the standard OECD/Eurostat Annual National Accounts questionnaire, simply asks NSOs to send National Accounts by activity tables for as many measures and years as possible at the most detailed level of activity (according to national classifications) available, in whatever format they find convenient. Much use is also made of tables provided on the Internet;
- **Timely** - STAN is released on a “rolling basis” (i.e. new country tables are disseminated as they become available) rather than as an annual snapshot. It is therefore only available via Internet through OECD's commercial service *SourceOECD* and OLISNet;
- **Variety of dissemination formats** - To meet varying user requirements, STAN is available either via a query-based application⁴⁵ or as a set of ready-to-use Excel tables. An ASCII version is available on request for easy loading into users' local data systems (e.g. SQL, SAS, Fame);
- **Good provision of metadata** – A range of information is provided including full documentation, data coverage tables, update history, country-specific notes, recommended uses and links to other datasets. Notes attached to individual data points highlight whether they are official National Accounts data or Secretariat estimates;
- **Comprehensive data checking and diagnostics** – A range of summary statistics are supplemented by detailed graphics, both on-screen and printed, of source data, selected ratios, industry shares and growth rates – comparing different data sources and/or data vintages. Feedback is provided to data providers when anomalies are discovered.

⁴⁴ See web version at <http://www1.oecd.org/publications/e-book/92-2001-04-1-2987/>

⁴⁵ See http://cs4hq.oecd.org/oecd/selected_table.asp?tableId=591&lang=eng&userid=&password=

B.2 *Known weaknesses*

- ***Difficult to determine the quality of estimated data*** – Except when NSOs subsequently provide National Accounts data previously estimated within the STAN process. This can also be an issue at the national level since few of the series in National Accounts are directly observed;
- ***Comparability problems*** – Despite the use of common frameworks (such as SNA93, ESA95), there are many problems in comparing National Accounts, and thus STAN, across countries due to the varying underlying data sources and methods used by NSOs to fill the various tables. Such issues are increasingly being addressed in broader discussions of comparability such as the measurement of economic growth and productivity across countries. Reconciling national industrial classifications to ISIC Rev.3 can also be challenging. To address these issues in the short-term, the STAN country notes highlight where national definitions may diverge from target definitions, while general notes attempt to expose areas where analysts should be prudent when making international comparisons;
- ***Data coverage problems*** – Data coverage is weak for many countries for certain measures such as investment and capital stock, output volumes and hours worked – productivity analysts are the most vocal on this point. This mainly reflects difficulties that some NSOs are having in making the necessary estimates by activity under SNA93/ESA95 recommendations. For example, with software now considered an investment good, re-estimating investment series by activity for long time periods in order to generate capital stock by industry is non-trivial. The feasibility of generating OECD estimates of capital stock by activity, the practice in the old ISDB database, is being investigated;
- ***Vintage software*** - The STAN system is based on software (dating from 1995) soon to be phased out by its vendor. Following a technical review towards the end of 2002, first steps have been undertaken to re-develop the STAN data system. Most likely, the solution will involve SAS linked to SQL data storage;
- ***Technical knowledge concentrated on one person*** – An important problem currently being addressed. The proposed technical re-development is an opportunity to simplify and codify processes for others to inherit, while continuing the close involvement of data users;
- ***Heavy data updating processes*** – Cumbersome updating processes are partly due to varying database structures within the OECD (from which STAN draws some of its source data), and partly due to the desire to minimise data providers' burden to maximise their response. Improved data flows within OECD are expected as the work to create a general data warehouse gathers pace. However, continuing to work with country-specific data submissions from NSOs could be considered an advantage since it leads to a better understanding of national practices.
- ***Complex estimation routines*** – During the STAN technical re-development the convoluted, and often difficult to understand, estimation routines (used to fill detail not provided in National Accounts tables) will be re-specified, re-coded and well documented so that the more mechanical aspects of the STAN process can be delegated to other members of staff.

Quality dimensions	Overall evaluation				
	Very weak	Weak	Satisfactory	Strong	Very strong
Relevance					X
Accuracy			X		
Credibility				X	
Timeliness				X	
Accessibility				X	
Interpretability				X	
Coherence			X		

C. Summary of main recommendations and action points

C.1 Main issues for National Statistical Offices

- **Co-ordination** – Many NSOs group all available National Accounts by activity tables together either within their on-line data services or when submitting tables to OECD. However for some NSOs, much effort is required to gather the tables for the various measures presented in STAN since their dissemination reflects internal organisational structures. Contact with up to four separate divisions within NSOs may be required - such as I-O division (value added and components), investment division (GFCF, capital stock), employment division and price division (output volumes). In fact, we often get requests from national analysts, within ministries and other government agencies, asking for STAN versions of their own countries' data. Co-ordinating the compilation of National Accounts by activity tables would not only provide easier access for productivity and structural analysts but could encourage NSOs to review the coherence across measures – for example, employment v. labour costs.
- **Encourage closer links with users** – It can be useful for statisticians to have a good knowledge of how the statistics they produce are used. This allows them to make judgements on what should be publicly released and/or what metadata are required. Often statisticians are far removed from the users of their data, particularly in an international context. For example, Eurostat's requirement to provide 'back estimates' (pre-1995) according to ESA95 has resulted in some apparently rough estimates being produced by under-resourced national accounts divisions to meet deadlines – and while these estimates may not appear in national publications, the fact that they may be widely used once submitted to Eurostat and OECD is sometimes overlooked. Occasionally, we send data providers with examples of where STAN has recently been used, particularly by their own governments (e.g. White Papers), not only to motivate them to submit updates but to highlight that once sent, their statistics do not just gather dust in the corner of an OECD database. Also, national statisticians are encouraged to review the STAN tables for their country and join the STAN distribution list so they are kept informed of how their contributions are used.
- **Details of compilation of National Accounts tables** – As mentioned earlier, much of the content of National Accounts is not directly measured and sources and methods for filling various tables differ across countries. For STAN, more precise knowledge of what combination(s) of sources are used to fill the activity tables (industrial surveys? business registers? income surveys? labour force surveys?) could help improve estimation techniques and help to make them more country specific. Better understanding of national practices could also contribute to more prescriptive chapters of future versions of SNA (for example, suggestions of types of surveys and computational methods to use to fill certain tables).
- **National STAN tables** – If NSOs compiled National Accounts by activity tables at a sufficient level of detail going back 20 years or more, OECD's role could then be one of co-ordination and second level quality control from an international perspective (which to some extent is already

undertaken). There is an increasing demand for such tables, and despite STAN shortcomings, analysts around the world have a common data set to work with. Without it, analysts would take National Accounts tables and estimate the detail they require in different ways, with varying levels of understanding and degrees of success.

C.2 *Main issues for OECD*

The suggestions below could be considered within OECD in order to improve the STAN database. Some concern STAN directly while others relate to the more general OECD statistical environment.

- **Address data duplication issues** – Not surprisingly, there is overlap between STAN and the activity tables presented in OECD's annual *National Accounts of OECD Countries* publication (tables 7 to 10 of volume II). However, while the latter is limited to 31 relatively aggregate activities (the A31 list), STAN attempts to exploit the most activity detail countries can provide. Also, tables and content can differ due to differing updating cycles and publication regimes. Efforts could be made to address this issue;
- **Review quality of source data sets** – The quality of the contents of STAN is inevitably influenced by the quality of underlying data coming from other OECD data sets. Outcomes of forthcoming quality reviews of OECD's *Structural Statistics of Industry and Services (SSIS)*, *International Trade in Commodity Statistics (ITCS)* and annual *National Accounts of OECD Countries (ANA)* should also contribute to improvements in STAN;
- **Reduce process time** – As the STAN system matured and managers became more experienced with national practices, process time decreased – a natural occurrence for most data systems (major changes notwithstanding). With the STAN technical development, it is hoped that process time can eventually be further reduced;
- **Improve online dissemination** – A query tool is not always appropriate for many analysts who need to download complete data sets, often the case for STAN. There is currently no mechanism for doing this on the commercial on-line data service (*SourceOECD*) – though complete STAN tables by country in zipped Excel format are available on OLISNet. The presentation of metadata could also be improved, for example, making it to easier to compare certain notes across countries. The current metadata structure was designed within the constraints of the software chosen;
- **Encourage NSOs to publish STAN tables** – Continuing the theme outlined above, OECD should strive to reduce the number of Secretariat estimates in STAN by encouraging national accounts divisions to provide more activity detail. For many countries the inclusion of just a few (less than 10) extra detailed activities (such as pharmaceuticals, shipbuilding, aircraft, telecommunications services), would meet the needs of a large proportion of users. As mentioned earlier, OECD's role could then be one of co-ordination and second level quality control from an international perspective. Providing advice and proposing common solutions for the more mechanical aspects of the compilation of national accounts tables could also be a function;
- **Refine quality review process** – Although having been partially involved in the design of the quality framework and review questionnaire, after participating in the process it seems that some improvements could be made to the 30-page questionnaire to make it easier for data managers to complete and interested parties to interpret. Also, since the OECD quality review summaries may provide information useful to NSOs it could be helpful to define a standard layout.