OCCASION

This publication has been made available to the public on the occasion of the 50th anniversary of the United Nations Industrial Development Organisation.

DISCLAIMER

FAIR USE POLICY

CONTACT

Please contact publications@unido.org for further information concerning UNIDO publications.

For more information about UNIDO, please visit us at www.unido.org

UNITED NATIONS INDUSTRIAL DEVELOPMENT ORGANIZATION
Vienna International Centre, P.O. Box 300, 1400 Vienna, Austria
Tel: (+43-1) 26026-0 · www.unido.org · unido@unido.org

RESEARCH AND STATISTICS BRANCH

WORKING PAPER 06/2008

UNIDO Data Quality: A quality assurance framework for UNIDO statistical activities

UNITED NATIONS
INDUSTRIAL DEVELOPMENT ORGANIZATION

RESEARCH AND STATISTICS BRANCH
WORKING PAPER 06/2008

UNIDO Data Quality: A quality assurance framework for UNIDO statistical activities

Shyam Upadhyaya
Chief Statistician

Valentin Todorov
Information Management Officer

UNITED NATIONS INDUSTRIAL DEVELOPMENT ORGANIZATION
Vienna, 2009

TABLE OF CONTENTS

LIST OF ACRONYMS AND ABBREVIATIONS V

1 BACKGROUND 1

2 UNIDO'S STATISTICAL ACTIVITIES 3

2.1 Statistical data production 4

2.2 Development of industrial statistics methodology 17

2.3 Technical cooperation in industrial statistics 20

2.4 Support to UNIDO programmes 22

3 QUALITY DIMENSIONS IN UNIDO CONTEXT 25

3.1 Relevance 25

3.2 Accuracy 26

3.3 Completeness 27

3.4 Timeliness 34

3.5 Comparability 35

3.6 Coherence 40

4 QUALITY ASSURANCE FRAMEWORK OF UNIDO 43

5 ROLE OF METADATA FOR QUALITY ASSURANCE 45

5.1 UNIDO metadata classification 48

5.2 UNIDO metadata system 50

5.3 Statistical data and metadata exchange (SDMX) 60

6 ANNEX I - FUNDAMENTAL PRINCIPLES OF OFFICIAL STATISTICS 63

REFERENCES 66

List of acronyms and abbreviations

ECE	Economic Council for Europe
EU	European Union
GDP	gross domestic product
INDSTAT	industrial statistics database
IRIS	international recommendations for industrial statistics
ISIC	international standard industrial classification of all economic activities
MVA	manufacturing value added
NACE	Nomenclature statistique des Activités économiques dans la Communauté Européenne (Statistical classification of economic activities in the European Community)
NISP	National Industrial Statistics Programme
NSO	National Statistical Office
OECD	Organisation for Economic Co-operation and Development
PPP	purchasing power parity
SDMX	statistical data and metadata exchange
SNA	system of national accounts
STA	Research and Statistics Branch
UN	United Nations
UNIDO	United Nations Industrial Development Organization
UNSD	United Nations Statistics Division

1 Background

The question of data quality has always been important for statistics. At the early stage, assessment of data quality was centred on the notion of the reduction of variance around the mean that would increase the accuracy of results. In the last century, however, the quality issue was first raised with respect to the statistical process control pioneered by Walter A. Shewhart in 1930s and carried on by W. Edwards Deming in 1950s. Further development on the quality of survey statistics went on the assessment of the *fitness for use* and the *fitness for purpose* of statistics produced. More recently, it has been widely recognized that in addition to accuracy there are other dimensions of data quality that correspond to the kind of statistical activity undertaken by one or another institution. Data quality, in the present context, refers to the totality of features and characteristics of data that has a bearing on their ability to satisfy the purpose and needs of users.

Several international organizations and national statistical offices (NSOs) have identified the quality dimensions relevant to their activities, which have been specified in their quality assurance framework and implemented successfully. International and European conferences on quality in official and survey statistics as well as Conferences on data quality for international organizations have been instrumental in encouraging and promoting awareness of quality assurance among data producers. UNIDO, for its part, has been an active participant at these meetings and has learned and adopted the best national and international practices of quality assurance. UNIDO is also a signatory to the Fundamental Principles of Official Statistics adopted by the United Nations Statistical Commission in 1994.

Statistical activity in UNIDO started with the establishment of the industrial statistics database in 1979 and its main objectives were to provide an accurate assessment of structure and growth of industrial sector and meet the internal statistical needs of the Organization. Accuracy and cross-country comparability were, however, the main quality challenges that UNIDO encountered right from the start. Intensive efforts were, therefore, made to develop a dataset that was

comparable over time and across countries. For many years, UNIDO's database was dependent on external data sources, especially that of the United Nations Statistics Division, which used to collect the data from NSOs. In 1993, the United Nations Statistical Commission, at its twenty-seventh session, mandated UNIDO for the collection, maintenance and dissemination of worldwide key industrial statistics in partnership with the Organisation for Economic Co-operation and Development (OECD). Ever since, UNIDO has been directly interacting with NSOs in various ways, such as collection of industrial data, dissemination of international industrial statistics and technical cooperation for strengthening the institutional capacity for undertaking industrial statistical operations. Each of these activities has underlying quality dimensions.

UNIDO's technical assistance to developing countries and countries with economies in transition is aimed at either creating a new industrial database or improving the existing statistical system. Through its technical support, UNIDO promotes the quality assurance of industrial statistics produced by the recipient country. Data quality of survey results depends greatly on quality of the frame (business register) and the efficiency of using administrative data (such as registration and tax data, company reports and so on) in combination with the survey results. Accordingly, UNIDO's statistical projects are designed to assist NSOs in producing statistical data that are accurate, complete and internationally comparable.

UNIDO Statistics Unit, in the past, produced a number of methodological and working papers describing its activities on quality assurance. It has now become necessary to sketch a generalized framework covering major quality aspects of statistics produced by UNIDO. The present paper depicts the key quality dimensions applicable to UNIDO's statistical activities and serves as a framework of data quality assurance for the Organization.

2 UNIDO's statistical activities

Statistical activities of UNIDO are defined by its responsibility to provide the international community with global industrial statistics and meet internal data requirements to support the development and research programme of the Organization. Currently, UNIDO maintains an industrial statistics database, which is regularly updated with the data, collected from NSOs[1] and OECD (for OECD member countries). UNIDO also collects national accounts-related data from the National Accounts Main Aggregates Database of UNSD, the World Development Indicators of the World Bank and other secondary sources. Such data are primarily used to compile statistics related to manufacturing value added (MVA); its growth rate and share in gross domestic product (GDP) in various countries and regions. UNIDO disseminates industrial data through its publication of the *International Yearbook of Industrial Statistics* and CD products. Data on major indicators are also posted on the UNIDO website as well as the UNIDO Intranet under the Statistical Country Briefs.

International comparability of statistics can be ensured through common standards on methods and classifications used in data compilation of reporting countries and territories. Therefore, UNIDO closely interacts with international organizations, especially with UNSD, for the development of industrial statistics methods and classification standards. Many developing countries and countries with economies in transition lack the technical capacity needed to introduce the latest methodological developments to their statistical system. Accordingly, UNIDO extends technical assistance to these countries and ensures its compatibility with international standards.

---

[1] In a number of countries, responsibility for industrial statistics lies with the Ministry of Industry or related line ministry. Reference to NSOs in this document is made to any national institution that is responsible for industrial statistics in the reporting country without making particular differentiation between a national statistical agency, a line ministry or other institution.

Statistical activities of UNIDO are carried out by the Statistics Unit of the Research and Statistics Branch (PCF/RST/STA) – hereafter referred to in this paper as STA. These activities can be divided into four groups: (1) statistical data production; (2) development of industrial statistics methodology; (3) technical cooperation in industrial statistics and (4) support to UNIDO programmes.

2.1 Statistical data production

2.1.1 UNIDO statistical databases

UNIDO maintains four different databases. The industrial statistics database (INDSTAT) constitutes a core data bank. The other three databases include the industrial demand and supply balance (IDSB) database, the MVA database and the employment-size class database. Details of each database are as follows:

Industrial statistics database (INDSTAT)

The industrial statistics database comprises data for 175 countries and territories spanning the period 1963 to the latest available year. The current database consists of historically formed databases, which are generally comparable over time and across countries but differ at the level of detail. The previous database maintained by UNSD covered 21 data items. At the time of transferring the responsibility to UNIDO, the number of variables was, however, reduced to eight, and the level of detail was changed from three- to four-digit level of ISIC. Following the endorsement of the third version of ISIC,[2] many countries switched over to ISIC Rev. 3 in 1990s, which prompted UNIDO, in 1997, to create a new database using ISIC Rev. 3. These databases are recognized by ISIC version and the level of detail is as follows:

---

[2] The first revision of ISIC was issued in 1958 following its approval by the UN Statistical Commission at its tenth session. The second revision was in 1968 approved by Commission at its fifteenth session. The third revision of ISIC was considered and approved by the Commission at its twenty-fifth session in 1989 and issued in 1990. In 2007, the Statistical Commission endorsed the plan for implementing the fourth revision which was approved at its thirty-seventh session in 2006.

INDSTAT3 Rev. 2: database by ISIC Rev. 2 at three digit level
INDSTAT4 Rev. 2: database by ISIC Rev. 2 at four digit level
INDSTAT4 Rev. 3: database by ISIC Rev. 3 at four digit level

The first database, INDSTAT3 Rev. 2, is the largest in size and covers a longer time period. While a small number of countries still supply their data in accordance with ISIC Rev. 2, for others, STAT converts data to ISIC Rev. 2 in order to maintain historical time series. Other measures of the size of these databases are given below:

Table 1. Size measures of the UNIDO Industrial statistics database

	Database by ISIC Rev. 2		Database by ISIC Rev. 3
	INDSTAT3	INDSTAT4	INDSTAT4
Number of ISIC categories	29	81	151
Number of variables	8	7	7
Number of data items	14	13	13
Number of countries and territories	181	116	111
Earliest reference period	1963	1981	1990

More recent data are reportedly using ISIC Rev. 3, hence the size of INDSTAT4 Rev. 3 is likely to grow. Although the database covers 175 countries and territories, many countries are not in a position to provide data regularly. The database is therefore updated only for some 70 countries each year.

Furthermore, since historical data are split because different ISIC revisions are used, STA has recently initiated the compilation of a single data series using ISIC Rev. 3 at the two-digit level for the entire period -- 1963 to the latest available year. This dataset has been developed by converting Rev. 2 data for past years and combining Rev. 3 data to two-digit level of ISIC Rev. 3. It thus provides a comparable set of long-term, time-series data, which is highly demanded by researchers.

The INDSTAT database contains annual figures according to industrial sectors (ISIC), country and year for the following variables:

1. Number of establishments
2. Number of employees
3. Number of female employees
4. Wages and salaries
5. Gross output
6. Value added
7. Gross fixed capital formation
8. Index numbers of industrial production

Selection of these eight variables was based on factors, such as internal and external demand of data users, keeping in mind the international division of labour in statistics, in order to avoid duplication and redundancy, reduce the reporting burden of NSOs, and importantly, the resource availability in UNIDO. Data for these indicators are readily available in most national industrial census and survey results. Thanks to their relation, several other relative variables, such as number of employees per establishment, wages and salaries per employee, value added output ratio and others, can be calculated on the basis of such data, which are not only important for economic analysis, but also very useful for checking the internal consistency of the database.

Industrial demand and supply balance (IDSB)

The IDSB database pertains to the manufacturing sector and data are classified by ISIC, country and year. The data are derived from output data reported by NSOs together with UNIDO estimates for international trade, based on the United Nations commodity trade data. This database contains annual time-series data in current US dollars on the following eight items (see **Table 2**), which are constructed in supply and use components.

Table 2. Content of the IDSB database	
1. Domestic output	5. Domestic consumption
2. Imports from the world (3+4)	6. Exports (7+8)
3. Imports from developed world	7. To developed world
4. Imports from developing world	8. To developing world
Total supply (1+2)	Total use (5+6)

Data in the IDSB database are compiled at four-digit level of ISIC for some 80 countries. However, the reporting period and data items covered differ from country to country. Databases for ISIC Rev. 2 and Rev. 3 are separately maintained. Details of coverage in the latest 2008 edition of the database are given in **Table 4**.

MVA database

The MVA database includes data for GDP and MVA at current and constant prices for 170 countries and territories from 1990 onwards. This database is entirely compiled from the secondary data sources, such as World Development Indicators (the World Bank) and the National Accounts Statistics database (UNSD). Supplementary data as per requirement are also obtained from OECD, Asian Development Bank, African Development Bank, Economic and Social Commission for West Africa, other regional organizations, as well as national data sources.

This database is used to estimate the GDP and MVA growth rates, MVA share in GDP, MVA per capita and the MVA structure by world regions. For the latest two years, STA produces estimates of these indicators based on a statistical method for forecasting (usually called now-casting since estimates are provided up to the current year based on historical trends).

Employment size-class database

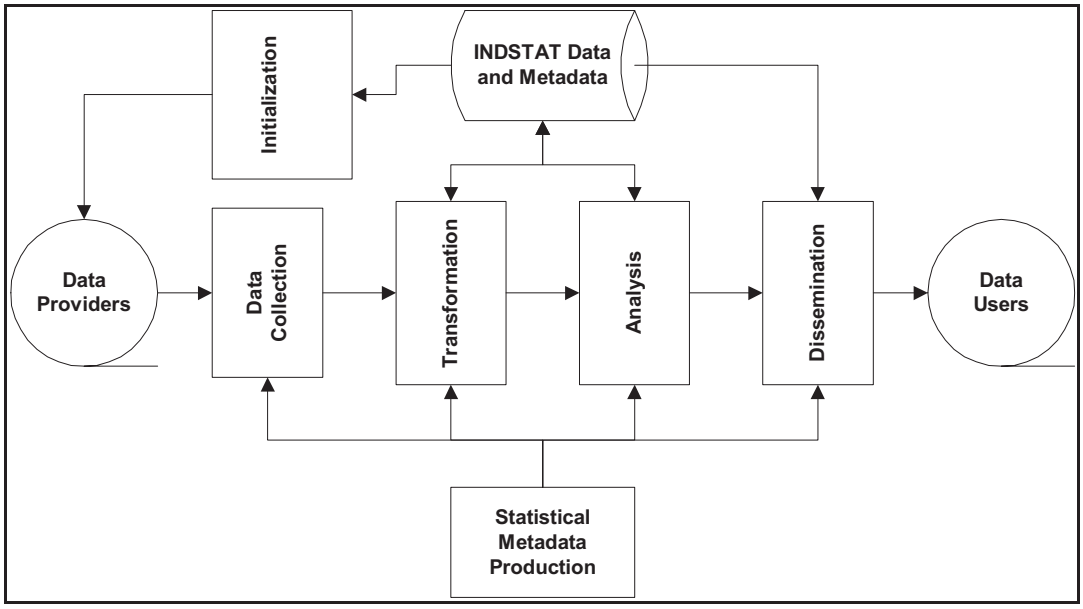
This database is organised by ISIC Rev. 2 and Rev. 3. The Rev. 3 database was started in 2004. Hence, it is still in the process of development. It provides major indicators of industrial statistics by different size classes in terms of number of employees. This database is developed to serve the UNIDO programme for small

and medium enterprise development and to fulfil the demand of external users interested in size-class data.

2.1.2 UNIDO statistical production process

The statistical activities in UNIDO follow a well defined work flow which starts with initialising the process by pre-filling the questionnaire and distributing it to member countries, collecting the data when the completed questionnaire is returned, validating, transforming and processing the data, performing data analysis and finally disseminating the produced statistics. This life cycle is annual and its main steps are presented schematically in **Figure 1**.

Figure 1. The main steps in the UNIDO statistical production process



During each step of the life cycle, statisticians have at their disposal the tools for the integrated statistical development environment (ISDE). **Figure** 2 presents an overview of the system and maps the tools used in the statistical production process. A short description of the system is given in 2.1.3 and the metadata-related part is presented in section 5.2.

Figure 2. Overall structure of the ISDE and its relation to the statistical production life cycle

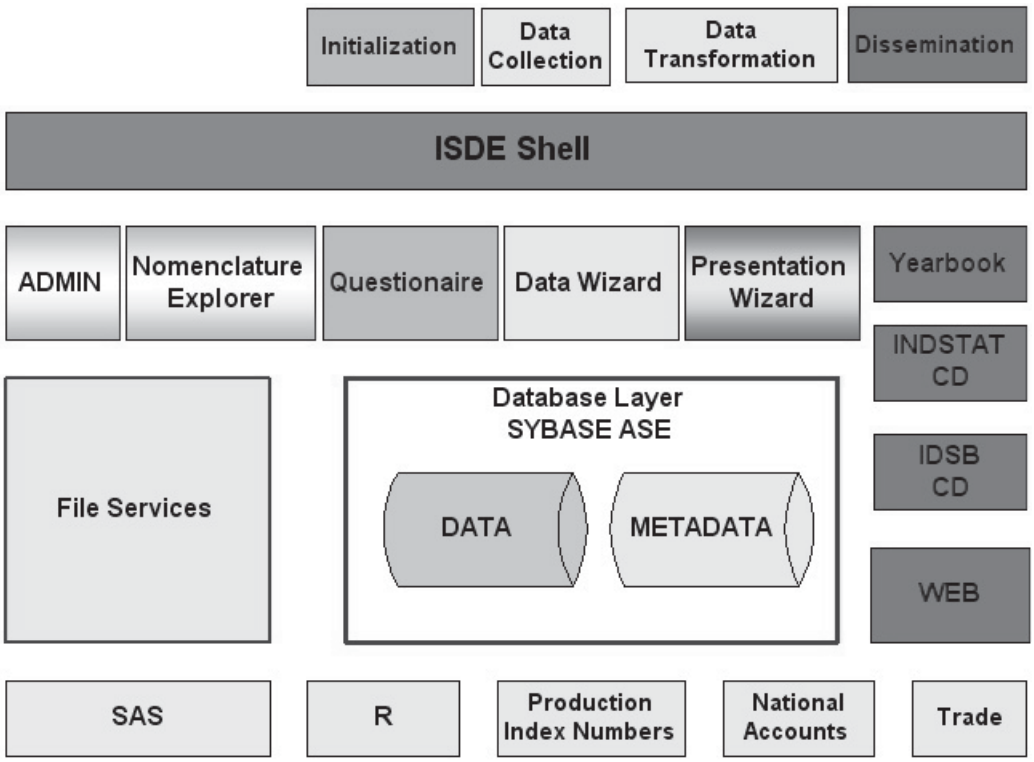


Table 3 shows the mapping between the survey life cycle phases adopted by the METIS group and the various phases in the UNIDO production process. The main difference is that there are no explicit preparation phases like 'need' or 'design and develop', but consistently at the beginning of the statistical production cycle, the current status is analyzed and, if necessary, the questionnaire as well as the process are updated. Further, there is no 'archive' phase, since as soon as the data are processed completely, they are stored permanently in the UNIDO statistical databases, rendering dedicated archiving unnecessary.

Table 3. Mapping of the UNIDO cycle phases to those developed by the METIS group	
METIS	**UNIDO**
Need	Need [optional]
Develop and design	Develop and design [optional]
Build	Initialization: pre-fill and distribute questionnaires
Collect	Data collection
Process	Transformation/processing
Analyze	Analysis
Disseminate	Dissemination
Archive	-
Evaluate	Evaluation

Initialization step

At this step, pre-filling the outgoing UNIDO general industrial statistics questionnaire with previously reported statistical data and metadata for their possible revision by the NSO is performed. The questionnaire is created in Excel format in three languages - English, French or Spanish - as appropriate for the particular country. Pre-filling is automated using available data and metadata.

Data collection

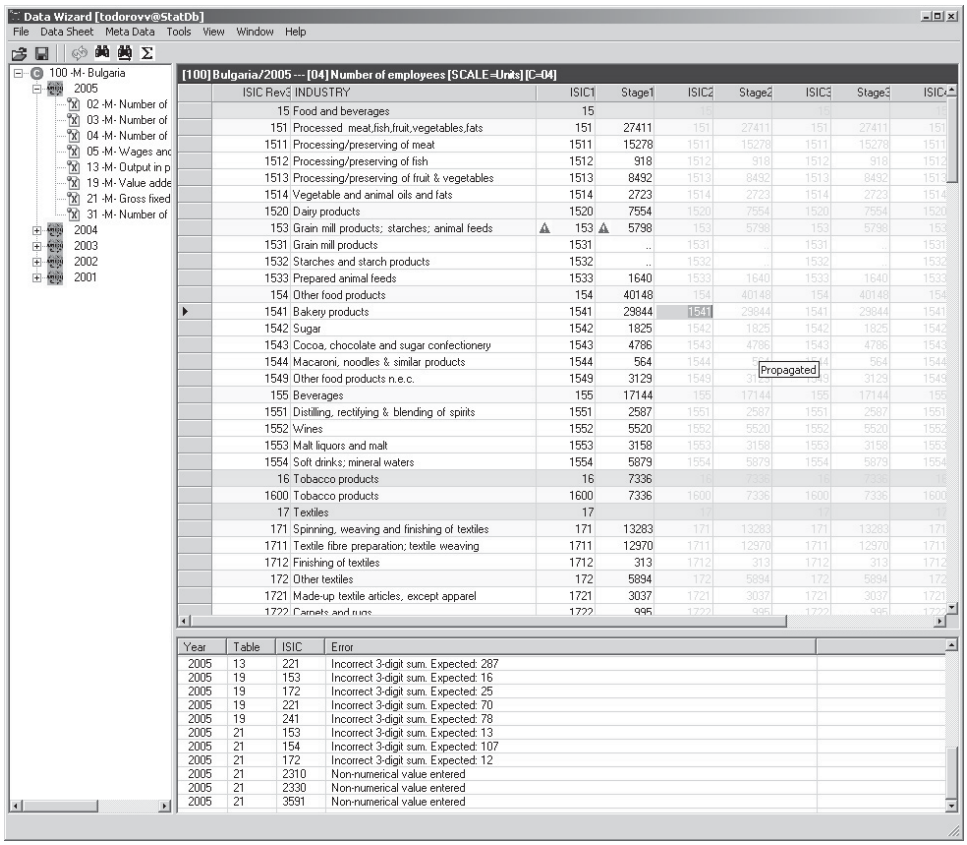
UNIDO collects data from UN member States in collaboration with OECD. Data for OECD countries are collected in the form of a joint OECD/UNIDO questionnaire and transmitted to UNIDO. Data from non-OECD countries are collected directly from the national statistical offices (NSOs) or other national sources using the general industrial statistics questionnaire. This questionnaire covers seven of the eight above-mentioned variables. Data for the production index are collected by UNSD and transmitted to UNIDO for compilation.

Data transmission to UNIDO from most national sources is done electronically. In a few cases, data are entered manually either from national publications on industrial census, surveys, NSO reports, or data sheets.

After receiving the completed questionnaires, the data are entered in the system for validation and further processing. The Excel file is read in automatically and the statistician uses a range of tools to validate, analyze, correct, etc. (see **Figure 3**).

While processing a particular questionnaire with all data and metadata included, it can be stored in the interim in XML format. The metadata can then be edited or new data can be entered.

Figure 3. Data and metadata in collection phase



The chart in **Figure 4** shows the data flow from sources to products in the UNIDO database system.

Data transformation

Data collected from primary sources are further transformed to a ready-to-use dataset. Data transformation is done in five stages, which not only constitute an operational framework for UNIDO statisticians, but also provides additional description of statistics (generated metadata which is attributed to each data item) to

users. While details on these stages are presented in UNIDO (1996), pp 6-8, a brief summary of the same is given below:

i) Detection and, if possible, correction of obvious reporting errors. Data are stored in original form (stage 1 data). These data are used for pre-filling following editions of the questionnaire for each country;

ii) Inconsistent data are corrected using supplementary information from national publications (stage 2 data). Stages 1 and 2 data are considered official;

iii) Data are adjusted to eliminate departure from the level of ISIC aggregation using national and international sources or supplementary data (stage 3);

iv) Missing data are estimated by UNIDO statisticians who apply related proportion or interpolation as and when applicable (stage 4);

v) Provisional estimates are made for the latest year (stage 5).

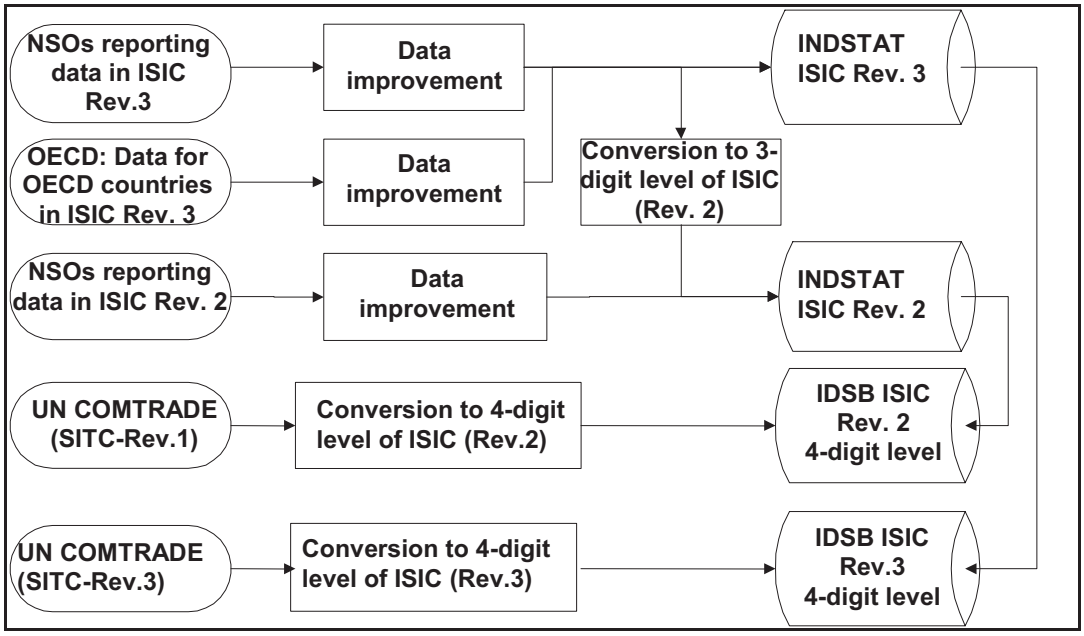
Updating the database is a continuous process. Incoming data from national sources undergo screening and scrutiny as soon as they are obtained. After transformation, as per above-described staging procedure, cleaned data are stored in the database. Only the major updates of stage 5 are carried out annually at the end of the calendar year. Apart from the statistical data, as already mentioned, the statistical metadata provided also requires updating. For example, country names and national currencies change. In the 1990s, following the break up of USSR and Yugoslavia, a number of new sovereign States emerged in Euro-Asian region. During the same period, 12 European Union (EU) member States adopted a common currency, the euro, replacing previous national currencies. In recent years, in Turkey and Ghana, appreciation of their currency necessitated the recalculation of past data series. More recently, new country datasets were created for Timor-Leste and the Republic of Montenegro. Besides, only recently two more countries (Bulgaria and Romania) joined the EU, and two countries (Malta and Cyprus) adopted the euro as the national currency. These and similar changes are duly updated in the database as necessary.

The processing phase also involves the re-basing of constant prices for national accounts time-series and revision of base weights for production index series. Re-basing of price series is usually undertaken every five years by moving the base year to a more recent year. Base weights for the production index normally refer to the same year. The core data for base weights are obtained from UNSD, which undergo some additional processing and transformation before computing indices.

Data dissemination

Data collected by UNIDO from NSOs and further transformed according to the quality requirements in the transformation phase constitute the major source of data for several recurrent publications produced by PCF/RST/STA. The metadata collected from NSOs together with the data, goes through the same transformation process as the data, and is complemented by metadata generated by the transformation process. All resulting metadata, including the necessary structural metadata, are used in the dissemination process – for details on metadata see section 5. The data and metadata stored in the database are used for the production of recurrent (annual) statistical publications: *International Yearbook of Industrial Statistics* (a hardcopy commercial publication); CD-ROM sales products of *Industrial Statistics (INDSTAT) Databases* and *Industrial Demand-Supply (IDSB) Databases* in different industrial classification schemes and industry aggregation also provided as Web-based service, Web *Country Statistics*.

Figure 4. UNIDO databases and data sources



The ***International Yearbook of Industrial Statistics*** is the main statistical product of UNIDO, which has been the most important medium for data dissemination for many years. The latest *Yearbook* released in 2008, covered data from 1995 to the latest year available. Country data was updated for 74 countries and is compiled on the basis of stages 1 and 2, as described earlier.

Another medium used for data dissemination is CD products, which might include data from all stages described earlier. The demand for CD products from national and international institutions, academia and researchers keeps increasing every year. For information on purchasing procedures and licensing, readers should refer to www.unido.org/statistics. The latest release of CD products in 2008 is shown in **Table 4**. For many years, UNIDO used to produce an additional CD featuring industrial statistics at the three-digit level of ISIC Rev. 3 contained in the INDSTAT3 database. Following the conversion process of the entire database in 2007, Rev. 3 INDSTAT3 database has been discontinued. However, with the establishment in 2008 of a new INDSTAT2 database, historic data between 1963 and the latest available years combined for some 180 countries.

Table 4. UNIDO CD products and their data coverage – edition 2008

CD product	Classification level	Number of countries	Period covered
INDSTAT 4 2008	4-digits of ISIC rev-2	116	1977-2005
	4-digits of ISIC rev-3	117	1985-2005
IDSB 2008	4-digits of ISIC rev-2	81	1981-2006
	4-digits of ISIC rev-3	80	1990-2006
INDSTAT 2 2008	2 digits of ISIC rev-3	Not yet released for publication	

Another form of data dissemination is providing statistics by selected variables from the different UNIDO databases to each member State. These are posted on the UNIDO website http://www.unido.org/statistics under the item Country Statistics. Country data on the website are presented for several years, together with figures for the world and regions for comparison over time as well as in relation to the region.

Apart from the recurrent publications listed above, industrial statistics data can be disseminated in response to **ad-hoc queries** mainly from internal but in some cases, also from external users.

2.1.3 The integrated statistical development environment

The overall structure of the ISDE is presented in **Figure 2**. The system utilizes a three-tier architecture built on net technology. The data and metadata are stored in a centralized database, and the user interacts with the system through the ISDE shell -- a desktop application which serves as a container for other ISDE applications. The commonality of the system is achieved by using component libraries that can be shared. The development of the entire ISDE has been carried out in-house in the context of migration from the mainframe to a client/server platform. For the migration, a step-wise approach was adopted because the goal was not only to migrate the system, but also to develop a completely new one, the requirements of which are not yet completely specified (owing to limited resources), and more importantly to ensure that the established UNIDO data services are not disrupted. As an example of the "migration" to "new development", it must be noted that while the I*nternational Yearbook of Industrial Statistics* was produced from the main frame as a camera-ready line printer output which was glued together with many MS

Word and MS Excel documents, the output of the client/server system is an automatically–generated, page-numbered PDF file of some 700 pages.

For building the system, the international standard (ISO/IEC[3] 11179) was taken into consideration. Some key metadata concepts from SDMX MCV are utilized. Inter-component data and metadata exchange are done in XML format. These are usually temporary files and typical examples are queries, created by the "query builder" and used by the "presentation wizard". The questionnaires are likewise locally stored temporarily during the first phase of validation, etc.

Table 5: List of tools used for the development of the system

Application/tool	Description
Sybase ASE 12.5	"Adaptive Server Enterprise" - the relational database management software manufactured and sold by Sybase, Inc. There are two separate databases running on the Sybase server – a test and production one. Another couple of test/production databases is used for web publications, but it is completely outside of the statistics unit.
Development and maintenance tools	
Erwin 4.1	Data modelling and database maintenance
SQL Programmer 12:0	Database maintenance
NET Framework 2.0	Most of the applications are written in C#
Crystal reports	A general purpose reporting tool, used for the production of the Industrial Statistics Yearbook as well as other publications. The version bundled with MS Visual Studio was used
MS Visuals Studio 2005	The main tool used for the development of the Client/Server libraries and applications
SVN 1.4.6 and tortoise SVN	Subversion - Source control system and (Windows) interface to it
XML Spy 2.0	Advanced XML editor
Other statistical tools	
ISDE File services	Used for interactions between ISDE and other tools like SAS and R. This is a shared network drive onto which the ISDE users have access.
SAS	Used for processing the National accounts data, the Production Index numbers as well as for serving any ad hoc requests for data
R	Currently used only for very specialized tasks, very high graphical potentials
Legacy systems/tools used	
VB 6.0	There are several legacy tools written in VB 6.0 which are not yet ported to the Microsoft .Net Framework (migration pending)

3 International Organisation for Standardization (ISO) and International Electrotechnical Commission (IEC).

Database layer

The database consists of two identical but physically separated databases – test and production databases – running on Sybase ASE RDBMS under Linux.

Access to data and metadata from client applications is performed through component libraries. These allow replacing, for example, the Sybase database by an MS SQL server or Oracle without any modification of the applications.

Component libraries

The object-oriented component libraries are developed also in C# and are used to unify many common tasks like database access, file access, printing, access to common data structures, etc.

Client applications

The client applications are developed using MS Visual studio in C#. They connect to the database and interact with each other using component libraries developed also in C#.

Other tools

Table 5 lists other tools integrated in the ISDE system.

2.2 Development of industrial statistics methodology

The primary responsibility for the development and publication of UN recommendations for statistical methods, including industrial statistics, lies with the UN Statistics Division, which reports to the UN Statistical Commission. UNIDO has been actively working with UNSD for the development and revision of international recommendations related to industrial statistics. UNIDO also plays an important role in promoting international statistical standards on concepts and definition of

industrial statistics, classification and data compilation method through its active interaction with national and international statistical institutions, as described below.

Feedback to NSOs through annual data compilation programme

UNIDO statisticians regularly interact with NSOs on a wide range of questions on methods and practices of industrial statistics through the annual data compilation programme. The general industrial statistics questionnaire developed by UNIDO is sent to NSOs with supplementary methodological notes that serve as a reference for comparison of national data with international standards. Upon receipt of the completed questionnaire, UNIDO statisticians make additional queries to detect any inconsistency or deviation from the standard classification in the reported data. These queries may relate to a simple computing error or conceptual differences. UNIDO's feedback to NSOs on reported data not only helps them to locate and correct inconsistencies but also to ensure clarity on various concepts on definitions used and their compliance with international standards.

Technical assistance to developing countries and countries with economies in transition

Technical cooperation programmes serve not only as assistance to a national statistical system in need but also as an opportunity for UNIDO to implement recent developments in statistical methods and best practices in the national statistical system. The basic administrative and legal frameworks, under which industrial activities are carried out widely, vary from country to country. It has a direct impact on the formation of target population of statistical observation which requires the decision on an efficient method of statistical inquiry. On the other hand, development priorities of countries, which are also different, demand very specific statistical information. Implementation of international statistical standards under specific circumstances often raises various methodological questions and practical challenges.

In the process of technical assistance project implementation, UNIDO employs highly qualified international experts who produce analytical reports, supported with a set of empirical results from statistical observations, and makes important methodological recommendations. These reports describe the different types of country experiences and contribute enormously to developing methodologies for industrial statistics.

Participation of UNIDO in activities of the international statistical community

UNIDO has been regularly represented at annual sessions of the Statistical Commission where the main agendas of international statistical standards are developed and reviewed. In recent years, the UN Statistical Commission conducted a programme review on international industrial statistics and revised the international recommendations for industrial statistics. The programme review was carried out by the Government of Japan in 2005-2006 at the request of the Commission. UNIDO contributed substantially to the review process as well as to the preparation of the final report. In 2005-2008, UNSD conducted three expert group meetings specifically to revise the international recommendations for industrial statistics (IRIS) and the manual of index numbers of industrial production. UNIDO was actively involved in the entire process of this revision. Upon approval of IRIS 2008, the UN Statistical Commission, at its thirty-ninth session, adopted the programme of implementation that required the cooperation of UNIDO for conducting regional capacity-building workshops for NSOs.

As a member of the Committee for the Coordination of Statistical Activities, UNIDO interacts with various UN and other international and regional agencies engaged in statistical activities and holds discussions on outstanding issues of statistical methods and practices. Conferences of the regional statistical committees are usually attended by a large number of NSO representatives. Thus, UNIDO's presence has always been very fruitful for exchanging views on recent questions of industrial statistics and country experiences. Participation of UNIDO at other international and regional conferences and meetings is based on the intended target group and the subject matter of the discussions. Preference is normally given to

meetings that focus on industrial statistics for official statistics, industrial survey methods, data quality and metadata system. Data and other information presented at such meetings are used to improve the statistical methods of relevance to UNIDO statistical activities, including technical assistance.

STA also provides advisory services to external data users who often contact statistical staff with questions pertaining to statistics contained in UNIDO publications. External users include students, researchers, development economists, journalists and other data users.

2.3 Technical cooperation in industrial statistics

In the late 1980s, the microcomputer data-processing system steadily penetrated the statistical system of developing countries. Following the 1983 World Programme of Industrial Statistics, and also to meet the growing demand for a specialized data-processing package for industrial statistics, UNIDO designed a package, referred to as the national industrial statistics programme (NISP) aimed at providing technical assistance to developing countries to collect a minimal range of industrial statistics. The conceptual background of the programme is contained in the international recommendations for industrial statistics (IRIS)-83 which distinguishes the recommendations for the industrial statistics programme (minimal range of data) and its extension (full range of data). IRIS-83 also lists data items for industrial statistics programme in accordance with priority ratings from 1 to 3. NISP covers all first-priority data items as well as some second-priority data items. Depending on the technical capacity of the national statistical system, NISP can be adjusted to include a larger coverage of data items.

National statisticians with limited programming knowledge were able to handle this software package very easily. NISP with its own package for data processing and database management for industrial statistics was effective in many countries. UNIDO therefore undertook NISP projects in a large number of countries in Asia and Africa. In subsequent years, the NISP software was upgraded to NISP Plus and NISP Windows. As the commercial data-processing software and manpower for its

operation became increasingly available in the local market, UNIDO discontinued updating NISP software in 2002 giving more attention to the statistical aspects of the programme.

In the 1990s, following the break up of the United Soviet Socialist Republics and the dissolution of CMEA,[4] countries that were previously using MPS[5] decided to change their statistical system to SNA standard. A number of countries requested UNIDO to provide methodological support for converting their industrial statistical system to one compatible with international standards based on SNA. UNIDO implemented the NISP project in Cambodia, Moldova, Mongolia, Lao People's Democratic Republic and Viet Nam, and established a new system of industrial statistics. In recent years, NISP, as a statistical programme, has also changed as the nature of the demand for industrial statistical system in the present context is significantly different from those contained in the World Programme of Industrial Statistics 1983. UNIDO is currently in the process of formulating a new programme for technical assistance commensurate with the revised version of IRIS. Until its completion, UNIDO technical cooperation in industrial statistics continues to be carried out during the implementation of development projects, which can be included as a component of the UNIDO integrated programme or as a stand-alone project.

Each project is designed to meet the needs of a country, but a typical UNIDO project for industrial statistics is intended to:

- support the NSO in creating a computerized business register with an efficient updating mechanism
- assist statistical operation with data collection methods and tools (questionnaire, sampling design etc.)

---

4 Council for Mutual Economic Assistance (otherwise abbreviated as COMECON) was an international economic organization comprising seven East European countries, Cuba, Mongolia and Viet Nam that existed between 1956 and 1991. COMECON countries followed a different statistical standard and classification based on the material production system (MPS).

5 MPS, also known as the System of Balances of the National economy, divided the activities of institutional entities, such as government, household and organizations, into production and non-production spheres and the value of the work done in the non-production sphere was not included in outputs or any other production measures.

- create a data-processing and menu-driven reporting system of principle indicators of industrial statistics
- improve the database management
- undertake statistical analysis of the survey results
- provide a monthly/quarterly survey design for production indices based on the updated weight from recent survey results
- train national staff in the latest statistical methods and standards of industrial statistics and data-processing through on-the-job training by UNIDO experts, study tours and training abroad and in-country group training courses.

Currently, ongoing projects cover different areas of industrial statistics, such as assistance to industrial census and surveys, registry updating system to improve industrial statistics, development of national indicators of industrial statistics, development of institutional capacity for annual manufacturing surveys and development of methodologies for statistics of products information and communication technology sector of manufacturing industry.

Results from the statistical operation conducted under UNIDO technical assistance are compatible with SNA methods and are internationally comparable. In accordance with the global mandate of UNIDO to promote industrial development, technical cooperation in industrial statistics aims to produce reliable and internationally comparable data required for formulating industrial policy and monitoring its implementation. Improvement of national industrial statistics system contributes to extension of the coverage as well as improvement of the quality of UNIDO database and thereby serves the needs of national as well as international data users.

2.4 Support to UNIDO programmes

As mentioned earlier, statistical activity in UNIDO was initiated in the late 1970s to meet its internal demand for industrial statistics. Support to UNIDO programmes

with timely, reliable and internationally comparable statistics remains one of the prime objectives of UNIDO statistical activities. The database provides recently available macroeconomic as well as business structure statistics for different countries and regions as per requirement of UNIDO programmes and units. As STA has obtained access to databases of other international organizations, internal demand for statistics, if necessary, is also met from external sources.

UNIDO research programme

Research and Statistics have merged as a single branch in the UNIDO organizational structure as their activities are interrelated. The Research Unit is the main user of statistics inside UNIDO, while STA receives feedback from research with respect to content and quality of statistics supplied. STA also provides data for the Industrial Development Report, which is the flagship publication of UNIDO, and presents the perspectives of Organization on industrial development worldwide based on empirical results and economic analyses. Furthermore, STA provides required data for in-house production of in-depth reports on specific issues, articles, working papers and other documents, which fall within the realm of the UNIDO research programme.

In 1999, STA developed a system of indicators for measuring the industrial development level of a country. Among others, the system includes:

1. MVA per capita
2. Share of MVA in GDP
3. Share of higher technology production in MVA
4. Growth of MVA
5. Manufacturing productivity
6. Diversification of manufacturing output

These indicators were first compiled from the UNIDO database in 1999, which permitted updating from time to time. Recently, a proposal was made to include a new indicator in the above system pertaining to the share of information and communications technology products in total manufacturing (RST/STA; 2006).

While the indicators in the above list are important in their own capacity and also convenient in compilation directly from the UNIDO database, a composite index of competitive industrial performance – CIP index (UNIDO; 2002) – was developed adding the trade data from UN COMTRADE database for two more indicators, namely, the share of manufactured exports in total exports, and the share of medium- and high-technology products in manufactured exports. The CIP index also replaced the share of high-technology production in MVA with the medium- and high-technology share in MVA. The CIP index and its component indicators are included in the UNIDO scoreboard database. Recently, the Research Unit has produced an updated version of Industrial Development Scoreboard using data derived from the industrial statistics database and industrial demand/supply balance database.

Other data requirements of the Organization

Formulation of a technical assistance programme covering the different service modules of UNIDO demands prior study of the economic context of the recipient country based on internationally comparable statistics. Such a study will reveal the current economic status of a member State and its position in the region. STA compiles country briefs which are available online. Additional information is supplied upon request from the database as well as from the collection of national statistics publications available in STA.

3 Quality dimensions in UNIDO context

Based on the nature of statistical activities carried out, a set of quality dimensions has been identified for the quality assurance framework of UNIDO. This framework is targeted to ensure that the statistical activities of UNIDO are **relevant** and data compiled and disseminated are **accurate**, **complete** within the defined scope and coverage, **timely**, **comparable,** in terms of internationally recommended methods and classification standards, and internally **coherent** with variables included in the datasets.

Quality dimensions defined here apply to UNIDO statistical activities. While each NSO may define its own quality assurance framework, UNIDO makes maximum effort to ensure that data produced from the statistical operation undertaken under the UNIDO technical cooperation are accurate, internationally comparable and coherent.

3.1 Relevance

One of United Nations Millennium Development Goals accords highest importance to poverty reduction. In response to this global challenge, UNIDO has defined its mission – Reducing poverty through sustainable development. To fulfil this goal, the Organization has set priority areas for technical cooperation: productive activities, trade capacity and energy and environment. Enhanced productive activities generate employment and self-employment, increase earnings and reduce income poverty. However, in the current age of globalization, producers face global competition in accessing markets for realization of goods and services of their production. Accordingly, UNIDO also helps to develop the trade capacity of producers in developing countries to enable them to sell their products. Expansion of production has a high impact on the environment. Therefore UNIDO emphasizes the efficient use of energy, and helps entrepreneurs to acquire cleaner production technologies. The role of industrial statistics in this process is indispensable, as the strategy of sustainable development can only be formulated based on detailed empirical data on

industrial structure and growth. Statistics from the UNIDO databases provide the clear link between production and trade, indicate the amounts and rates of wages and salaries in various sectors of manufacturing, that is, earnings from manufacturing activity, and relate productivity to the overall industrial performance.

Industrial sectors, especially manufacturing, play the determinant role in the economic growth of developing countries. In the developed world, where the share of the service sector is higher, manufacturing still remains important, thanks to its strength in high technology that supports the growth of non-manufacturing sectors. In developing countries, manufacturing has been the leading sector with the highest growth potential. However, in the current epoch of globalization, the real growth potential of national economies lies not necessarily in the entire manufacturing sector but in some of its specific branches. Therefore, there is the tendency of deeper specialization of production and trade, which is closely associated with the notion of comparative advantage, competitiveness, productivity and structural changes. Monitoring and analyzing this process requires detailed and reliable statistics on the industrial structure. Therefore, demand for internationally comparable data on detail industrial structure, which UNIDO has been producing, has increased significantly in recent years.

3.2 Accuracy

Accuracy is an indispensable quality dimension of data reported by UNIDO. Accuracy of data is examined at all stages: data collection, transformation and reporting. At the collection stage, UNIDO sends the standard general industrial statistics questionnaire with pre-filled data by ISIC codes and description of the preceding period reported by NSO. Methodological notes as well as a metadata sheet are attached to the questionnaire. NSOs fill out the questionnaire, provide data on the last available year and, if necessary, correct previously reported data.

At the transformation stage, STA screens the reported data in two phases: first, possible abnormalities of data are identified, and second, abnormalities are redressed to the extent possible. In this process, errors arising from wrong ISIC or country

codes, rounding up of figures and the like are controlled. Data are further checked through arithmetic and logical control of errors by computing the mean and ratio of key indicators. Data inconsistencies that could not be revealed at this stage are further verified with related national and international publications (stages of data transformation as described in the previous chapter). The UNIDO database is updated only after accuracy is attained to the extent possible.

However, it is important to note that any intervention by UNIDO might have limited effect in ensuring the accuracy of data if supplementary information is not available. UNIDO depends on national data sources and cannot make changes in reported data without consulting the reporting organization. As the primary responsibility for data accuracy lies with NSOs, UNIDO constantly interacts with them through its data compilation programme and technical cooperation projects.

3.3 Completeness

One of the important quality dimensions of UNIDO industrial statistics is completeness, which is measured in relative terms of coverage at different levels as described below:

- Country coverage: Number of countries included in the database
- Activity coverage: Coverage of manufacturing activity of the reporting country in industrial data supplied to UNIDO
- Unit coverage: Number of observations and response rates
- Data items coverage: Census value added versus total value added

While complete coverage of all types of manufacturing activities in every country and territory worldwide would be an ideal target, UNIDO aims at maximum possible coverage taking the following objective measures into consideration.

Country coverage

In terms of the total number of countries presented, the UNIDO database has a fairly good coverage. Currently, the INDSTAT database covers 181 countries and territories, including full coverage of Europe and North and South America. The database also has quite good representation of Asia and the Pacific region, except for a number of island nations in the South Pacific. Obviously, the role of the manufacturing sector in the economies of countries in this region is quite limited. However, the database has the lowest coverage of sub-Saharan Africa. Out of some 45 countries in the region, no data are available for 18 countries, while for some 10 other countries the database has not been updated for several years.

Most countries in sub-Saharan Africa, for which data are available, are those that received technical assistance from UNIDO in the past. It therefore emphasizes the importance of UNIDO's technical assistance in sustainable development of the national industrial statistical system, and also as an effective way for achieving completeness in its database coverage.

Activity coverage

For very objective reasons, industrial census or surveys carried out in many countries do not cover the manufacturing sector in its totality. The survey method that applies cut-off size excludes smaller units as described below.

Cut-off size designated for industrial census/surveys

Following the recommendations of the World programme of Industrial Statistics 1983, as well as IRIS-83, most developing countries apply a cut-off size to exclude smaller units from the industrial data collection programme. Generally, a fairly updated register is maintained for larger establishments, which provides the frame for a regular industrial survey. Although much smaller in the number of units, larger establishments produce a substantial part of total MVA. It would be a very time- and resource-consuming exercise to maintain the register for a large number of small and economically unstable units. Therefore, the World Industrial Statistics programme

recommended, “Owing to the limitations in financial and human resources for such work in a particular country, the census coverage may be limited to the larger establishments. In practice, such establishments might be defined as those with five or more persons engaged” (UN, 1981).

The cut-off size is also applied in developed statistical systems. However, in most such countries, the cut-off size is maintained at a fairly low level. Only very small units that are exempted from tax registration are excluded from the business register. Hence, the contribution of units below the cut-off size to the total value of major indicators is negligible. However, most developing countries apply the cut-off point in regular industrial surveys at around the size class of 5-10 persons engaged. For example, the annual industrial survey of India covers all establishments with power equipment with 10 or more person engaged and all other establishments with 20 or more persons engaged, while Argentina uses the cut-off point with 10 or more employees, in general.

In the national statistical system there are some estimates from the occasional inquiry or sample surveys representing smaller establishments. These data are not adequately utilized to adjust the gap in the annual survey data. Data provided to UNIDO are mostly derived from the regular industrial survey results and do not represent those establishments below the cut-off point. Therefore, users of UNIDO database are often reminded that for many developing countries, data represent a significantly large portion of manufacturing, but the proportion of coverage varies slightly.

Unit coverage

Often in statistical surveys it is not be possible to observe all eligible units due to non-coverage or non-response. Non-coverage relates to the problem of identification of unit, while non-response refers to failure of observation of units that have been identified. Non-coverage is higher in developing countries due to the poor quality of the frame. In developed countries, the quality of the frame is highly reliable, but the response rate is low.

Obviously, UNIDO does not have control over the rate of unit coverage. However, it collects information irrespective of whether proper adjustments or estimations have been made by NSOs for non-response or non-coverage. Data obtained from OECD are adjusted for non-response. In a number of developing countries, where the quality of the frame is poor, it is difficult to establish the total number of units. In such cases, NSO cannot compute or report the rate of non-coverage or non-response. Here UNIDO can help to improve the data quality in national statistical sources through technical assistance and interaction in the process of compilation of country data for the UNIDO database.

Data items coverage

The industrial survey programme in many developing countries is based on the census concept, which covers those data items that are necessary to compute the census value added, but not sufficient to derive the total value added.

Census value added versus total value added

The difference between these variables arises from the different coverage of data items depending on the statistical unit chosen for industrial surveys. The statistical unit can be an establishment or an enterprise. Both approaches carry certain advantages and disadvantages.

IRIS-83 recommended establishments as a statistical unit for industrial statistics inquiry. Because, - "… it is the most detailed unit for which the range of data required is normally available. The data gathered, in order to be analytically useful, need to be grouped according to such characteristics as kind of activity" (UN, IRIS 1983). Therefore many countries use establishments as a statistical unit in their industrial census or surveys. When an establishment is a part of the single establishment enterprise, it does not make a difference whether the data are collected at the establishment or enterprise level. However, in the case of a multi-establishment enterprise, transactions in certain type of "non-industrial" services are made at enterprise level. Therefore, establishment-level data are not covered and the

resulting figure adds up to the census output, census input and census value added, as shown later.

The census value added gives a close approximation of the total value added figure when the difference of the revenue from, and the cost of, the non-industrial services is not so big. Revenue from non-industrial services include receipts for transport services rendered to others, other than delivery of own products, storage of goods and warehousing, right to use patents, trademarks, copyrights and the manufacturing and quarrying rights, technical "know-how" and receipts for similar kind of services. Costs include those of advertising, legal, accountancy, consulting, planning, research and development services, patent and license fees (but not the value of outright purchases of patents and licence), costs of business travel and meetings, contribution to business and professional associations, cost of communication, entertainment and other similar services.

Table 6. Composition of output and input of manufacturing	
Output components	**Input components**
A. Sale of goods • Own produced goods from main and allied activities • Goods produced by another establishment using materials supplied by reporting establishment • Transfer of goods to another establishment of the same enterprise • Receipts from the sale of goods in the same condition as purchased, *less* cost of these goods	**K. Cost of materials and supplies** • Materials and supplies purchased • Materials and supplies transferred from another establishment of the same enterprise • Cost of packing materials, tools etc. • Cost of fuel and electricity
B. Value of own produced fixed assets • Value of machinery, equipment and other items of capital goods produced and retained for the use of establishment • Value of work done on own account	
C. Change in stock • Finished goods • Goods for resale • Work on progress	**L. Change in stock** • Materials and supplies • Fuel
D. Receipts from industrial services • Receipts for contract and commission work done for others from their own materials • Receipts for repair and maintenance services • Other receipts from industrial services	**M. Cost of industrial services** • Cost of contract and commission work • Cost of repair and maintenance • Cost of other industrial services
E. Census output = A+B+C+D	**N. Census input = K-L+M**
F. Receipts for non-industrial services	**O. Cost of non-industrial services**
***G.* Gross output = E+F**	**P. Intermediate consumption = N+O**

The following relations are derived from the above composition:

Census value added = Census output – census input

Total value added = Gross output – intermediate consumption

Total value added = Census value added ± difference of the receipt from and cost of non-industrial services.

When the reporting unit is an enterprise, the costs and receipts of non-industrial services are reported for the whole enterprise. For a multi-establishment enterprise, these values are proportionally distributed to obtain a complete measure of gross output and value added at establishment level. In such cases, data are complete

although an approximation is involved when the total costs or receipts are distributed proportionally to establishments.

The argument in favour of the census value added is that the difference between the total and the census value added might be important to measure the level, but not for structural analysis. More importantly, data based on the census concept might be more accurate and complete, because approximate allocation of non-industrial costs and receipts tend to introduce the measurement error to estimates. Despite these advantages, the concept of census value added is not used any more in the system of national accounts. Countries are increasingly extending the coverage of data items to derive the accurate measure of value added through a data collection programme that involves establishments as well as enterprises in annual industrial surveys.

Due to the limitation of coverage described above, the country data for value added obtained from the UNIDO database is equal or less than total MVA reported in GDP estimates. It has been found that data from more than one third of reporting countries fully cover the sector. The percentage of reporting countries by level of coverage is given below.

Table 7. Coverage of MVA in national industrial data reported to UNIDO, 2005

Level of coverage	Number of reporting countries in percentage to total	Value added covered by reported data in percentage of total MVA	Cumulative number of reporting countries in percentage to total
Fully covered	36.00	100	36.00
Mostly covered	37.33	More than 75	73.33
Fairly covered	17.33	More than 50	90.67
Poorly covered	9.33	All observations	100.00

For more than 70 per cent of countries, industrial data reported to UNIDO include 75 per cent of total MVA. For this part, the database provides sufficiently reliable and relatively complete statistics for economic and structural business analysis. The database contains poorly covered data for less than 10 per cent of the countries that have a weaker statistical capacity in place. UNIDO identifies these countries and addresses the problem by extending technical assistance in building the institutional

capacity for industrial statistics or providing methodological support through training programmes or other forms of interaction with NSOs concerned.

3.4 Timeliness

The UNIDO database is constantly updated with incoming data from its sources to ensure the timeliness of its products. STA works with a calendar year schedule to ensure that the dataset for the next edition of the *International Yearbook of Industrial Statistics* is completed by November each year. A new edition of UNIDO publications and CD products are released at the beginning of each year.

Normally there is a time lag of three years between the reference year of the most recently reported data and their publication in the *International Yearbook of Industrial Statistics*. This time lag is because the time required for data flow from the national survey to the final international publication. For example, the annual industrial survey of ***country A*** for reference year 2005 is conducted in 2006. NSO completes data collection, processing and dissemination by the end of 2006. NSO of ***country A*** transmits the data to UNIDO by September 2007for publication in the 2008 edition of the *Yearbook* at the beginning of 2008. The time gap between production of data by NSOs and publication by UNIDO is spent on some provisional estimates made by UNIDO statisticians using statistical methods for forecasting based on data reported for earlier years. UNIDO maintains the dataset of base weights and production indices. Based on the projected growth at sector level, MVA estimates for recent years are updated annually at the two-digit level of ISIC. In cases where data for individual items are missing for one or another period, estimations may involve both interpolation and extrapolation.

Estimated figures are published in relative terms (such as MVA per capita) and indicate the structure (share of MVA in GDP) or growth (MVA average annual growth 2000-2005). Such estimates, on the one hand, satisfy the immediate demand of data users (support timeliness) and, on the other, they avoid any inconsistency with the actual value produced by NSOs (maintain accuracy and coherence).

3.5 Comparability

International comparability is one of the main challenges that UNIDO faces in compilation of regional and global indicators of industrial statistics. In order to ensure international comparability, it is necessary that national data comply with the various UN recommendations related to industrial statistics, especially:

- the System of National Accounts (SNA-93 and later updates)
- IRIS 1983 and 2008
- ISIC of all economic activities -- Revisions 2, 3 and 4.

Data incomparability might arise from deviation of national data from international recommendations on different concepts and standards. UNIDO pays special attention to international comparability of data, in terms of scope and coverage defined for industrial statistics, industry classification standard and the valuation method of principle indicators.

Scope and coverage of industrial statistics

The industrial sector, as defined in IRIS 2008, comprises all establishments located within the territorial boundaries of the reporting country that are engaged primarily in the following activities, as classified in ISIC Rev. 4:

Table 8. ISIC Rev. 4 activities

	Industrial activities	ISIC Rev. 4
1	Mining and quarrying	Section B
2	Manufacturing	Section C
3	Electricity, gas, steam and air-conditioning supply	Section D
4	Water supply, sewerage, waste management and remediation activities	Section E

Activities in sections D and E were combined into one ISIC category in earlier revisions. Some countries may include construction in industrial sector. However, UNIDO does not collect data for construction. Data on mining and quarrying, electricity, gas and water are collected through a separate questionnaire, and then

transmitted to UNSD for further compilation. Air-conditioning supply and waste management and remediation activities have recently been included.

The UNIDO industrial statistics database covers only the manufacturing sector. It is imperative for data comparability that country data do not deviate from the recommended definition of the manufacturing sector in ISIC. Occasionally, a reporting country might include some activities that do not belong to the manufacturing sector, or some activity that belongs to manufacturing is omitted. This type of deviation makes data incomparable. Therefore, at the data transformation phase, STA sends queries to NSOs if any deviation is detected. Data are subsequently corrected in order to attain accuracy at activity classification (leading to comparability). However, as regards coverage of the manufacturing sector, there might be differences owing to the designated cut-off point as described earlier. If one country dataset covers its manufacturing sector completely, but another does not, then these two datasets become incomparable. UNIDO tackles this problem by maintaining an additional database for total MVA and GDP by country. The total of sector value added reported in the industrial statistics database is a significant portion of MVA distributed at detail level of industrial activities. Thus users are recommended to use the industrial statistical database for structure analysis, and national accounts database for macroeconomic analysis.

Classification of economic activities

There are different classification standards used in different parts of the world. The United States, Canada and Mexico use North American Industry Classification System (NAICS) – latest version NAICS 2002. The EU and a number of East European countries use NACE Rev. 1.1. Most other countries use ISIC Rev. 3 and a small number of countries, including China, use ISIC Rev. 2. Some countries have national versions of classifications largely based on one of these international standards. Although, all these classifications claim to be comparable with each other, in a number of cases there is no one-to-one correspondence using four-digit level codes.

Data received by UNIDO from OECD are converted to ISIC Rev. 3, which cover all NAICS and most NACE user-countries. Other NACE user-countries send their data based on ISIC Rev. 3. Despite this, incomparability arises, on the one hand, from deviation in national versions of classification from ISIC and, on the other, the simultaneous use of ISIC Rev. 2 and Rev. 3 by member States, consequently making it necessary to maintain two databases – one for each ISIC version – in UNIDO.

Any national deviation in classification from international standards is corrected by STA in order to make data internationally comparable at the transformation phase. In the case of Rev. 2 and Rev. 3, UNIDO has addressed the problem in two ways. First, since the past 10 years, it maintains two separate databases for ISIC Rev. 2 and Rev. 3 to keep the original data. Second, STA has developed a conversion programme to bring the data from one set to another. Recently the fourth version of ISIC has been endorsed and many countries are expected to introduce it very soon.

UNIDO has therefore created yet another database, INDSTAT2 using ISIC Rev. 3 at the two-digit level, which combines historical data from separate ISIC Rev. 2 and Rev. 3 databases into a single database of ISIC Rev. 3 at the two-digit level. It has also created a comparable time-series of industrial statistics between 1963 and latest year.

Computation method of principle indicators

SNA remains the main source of computation methods for various indicators of industrial statistics as a part of economic statistics. MVA, as a statistical measure of contribution of manufacturing sector to GDP, is computed based on national accounts practice. At the same time, it has also been recognized that industrial statistics indicators are important in their own right as they serve the purpose of industrial performance analysis. For example, compensation of employees as a component of value added would have been a better measure of total returns to labour in the manufacturing sector. But for industrial performance analysis, the average based on wage rate per employee by ISIC is a far more important indicator. Similarly, the number of persons engaged is a comprehensive measure of

employment, which includes all paid employees as well as working proprietors, active business partners and unpaid family workers. However, for productivity analysis the number of employees is a more suitable indicator than the number of persons engaged. There is therefore, some trade-off between comparability and data demand, which is a usual dilemma in data quality between *fitness for use* and *fitness for purpose*.

Some possible cases of deviation and their effects are shown in Table 9.

Table 9. Deviations and their effects

	Required indicator	Most likely deviation	Effect on comparability
1	Number of enterprises	Number of establishments	Number of establishments is usually more than the number of enterprises, because an enterprise may own more than one establishment but an establishment does not own the enterprise
2	Number of persons engaged	Number of employees	Number of persons engaged also includes working proprietors, active business partners and unpaid family workers
3	Compensation of employees	Wages and salaries	In addition to wages and salaries compensation of employees includes any payment made to social security funds on behalf of employees
4	Census output	Gross output	Census output is less than gross output as shown on page 21
5	Census value added	Value added	Value of these indicators might be different as shown on page 21

In some countries, industrial data are compiled merely for national accounts estimation without due attention paid to sector analysis. In such cases, the NSO is likely to report data for an indicator that looks similar to the one mentioned in the UNIDO questionnaire, which is essentially different. This could lead to data incomparability. Therefore, the definition and composition of the list of indicators are contained in the methodological notes attached to the questionnaire sent to NSOs. STA makes necessary checks to locate any deviation from the requested indicators. Some differences, such as census and total value added, are objective, and therefore, cannot be fully controlled, but in other cases, data might be corrected after additional inquiry with NSOs or through other corrective measures.

Valuation of output measures

SNA 93 as well as IRIS 2008 recommends that gross output should be measured at basic prices. It is not possible to measure value added directly, but value added at

basic prices is derived from output by deducting intermediate consumption at purchasers' prices as follows:

Value added at basic prices	=	Output at basic prices	–	Intermediate consumption at purchasers' prices

Some countries report gross output at producers' prices. Subsequently value added derived from this measure is also at producers' prices:

Value added at producers' prices	=	Output at producers' prices	–	Intermediate consumption at purchasers' prices

The difference between the two arises from commodity taxes and subsidies. Gross output at producers' prices includes both commodity and non-commodity taxes (except value added tax or any deductible taxes) and excludes commodity and non-commodity subsidies (such as rental or labour subsidies). Both producers' and basic prices are actual transaction prices, which can be directly observed and recorded. In some cases, data are reported at factor cost, which excludes all kind of taxes, but includes all subsidies.

For a cross-country comparison of output measures (gross output and value added) it is necessary that their valuation in countries under observation is identical. UNIDO encourages NSOs to provide data at basic prices, but accepts data reported in any kind of valuation. Difference in valuation methods affects comparability of output measures, especially in those ISIC branches where commodity taxes are normally high and vary across countries, such as manufacture of alcoholic beverages, tobacco products, manufacture of transport and communication equipment. For highly levied production sectors, data at basic prices are more comparable and give a precise picture of level and structure.

A more difficult aspect of valuation for cross-country comparison is the price level of manufacturing products. Currently, cross-country comparisons are made in

United States dollars. However, it is apparent from the comparison of the purchasing power parity (PPP) that exchange rates under- or overvalue the United States dollar in relation to other currencies in highly varied proportions. At the same time, conceptual problems arise from using adjusted exchange rates with PPP for cross-country comparison of MVA. PPP is an aggregated rate of consumer prices based on the composition of consumption expenditure. GDP at national level can be calculated both from production and consumption. Consumption expenditure, which is the main component of GDP, is measured at consumer prices. However, the same does not apply to sector output like MVA. Moreover, the survey conducted under the international comparison programme to obtain cross-country price relatives includes a limited number of manufactured products in its basket. Therefore, it is necessary to obtain a separate conversion rate based on price relatives of manufacturing products. As long as such relatives are not available, comparison made in United States dollars or in any other currency is subjected to limitations.

3.6 Coherence

The UNIDO database is a multi-country database comprising historical series for several years, thus making coherence an important quality dimension. This implies internal consistency of data in different aspects. First of all, it is necessary that the terms and concepts used in one dataset have exactly the same meaning in another dataset, unless the difference is explicitly mentioned. The terms, manufacturing, employees, wages and salaries, value added etc. have the same meaning for all countries and for years included in the database, unless any deviation is reported.

Secondly, even in the case of deviation, data can still be reconciled. For example, if country provides data using ISIC Rev. 2 for earlier years and ISIC Rev. 3 for later years, the two datasets are different. However, some key common elements exist in both datasets that allow combining them. UNIDO checks for coherence of data in the following aspects:

Coherence in country data

Data on the manufacturing sector for the given reference year for each country are collected from a defined population and constitute an independent dataset. For example, amount of the wages and salaries (*W*) of a manufacturing branch *s* of country *X* paid in a year *t* and the number of employees *n* of the same sector belongs to the same population, which allows one to calculate the average wage rate (AWR) of type:

$$AWR_{st}^{X} = \frac{WS_{st}^{X}}{n_{st}^{X}}$$

Since the variables in the dataset refer to the same population (same set of production entities) they are coherent. Thus users can derive additional analytical variables from the dataset. It does not imply, however, that users can take one variable from the UNIDO database and another from an external source to obtain any meaningful result, unless data in both instances refer to the same population and same reference period.

Country data can be aggregated to a higher level or disaggregated to a lower level of ISIC, as the statistical terms used have the same meaning for all statistical units of all manufacturing activities.

Coherence across countries

Data obtained from different countries are checked to ensure that there is no significant deviation from the UN recommendations on concepts and definitions. Thus the statistical variables used in different country datasets are compatible with each other. From the above example, wages and salaries, number of employees and the manufacturing sector, have the same definition and meaning in country X and country Y. Thus, AWR_{st}^{X} and AWR_{st}^{Y} are comparable figures (after conversion to a single currency), if a user wishes to find out whether country X has a higher or lower average wage rate than country Y.

Coherence across countries also implies that data of different countries can be aggregated to higher levels to compute required variables at regional level. UNIDO publishes data by different country groupings, such as the EU, Commonwealth of Independent States, developed countries, developing countries etc. It also allows one to compute the share of a country in the total.

Coherence over time

This implies that concepts and methods as well as classification standards applied to data are common over time. Again using the above example of average wage rate, all terms involved have a common meaning over a period so that the growth of the average wage rate can be computed as:

$$\frac{AWR^{X}_{st+1}}{AWR^{X}_{st}}; \frac{AWR^{Y}_{st+1}}{AWR^{Y}_{st}}.$$

When changes occur in concepts or classification that affects comparability, the data series might break. In such cases, STA identifies elements that are consistent over time, works out a data conversion method and creates comparable series for all years. Such situations occurred in the past, especially due to the change in ISIC (Rev. 2 to Rev. 3) and change in currency (from several national currencies to the euro and from the rouble to several national currencies in countries of the Commonwealth of Independent States).

4 Quality assurance framework of UNIDO

The quality dimensions described earlier are applicable in the current context of UNIDO statistical activities. Thanks to efforts made by STA, the quality of statistics produced by UNIDO has been duly recognized by data users worldwide. Many international organizations, NSOs and line ministries as well as a large number of researchers approach UNIDO every year with requests for statistics from the UNIDO database and also with methodological queries. It has therefore become necessary to establish a formal framework that can serve as a guiding instrument for quality assurance of statistical activities of UNIDO in general, and its statistical products in particular.

The following scheme depicts the quality dimensions applicable to UNIDO and the corresponding statistical activities.

Table 10. Overall quality assurance framework (May 2008)

Quality dimension	Targeted statistical activity	Tasks for quality assurance
Relevance	*Maintenance of global industrial statistical database and support to UNIDO programme*	1. Industrial statistical activities of UNIDO fulfil the mandate of UN Statistical Commission as well as support the UNIDO programmes 2. The Organization allocates required human and financial resources to maintain the quality of its statistical activities
	Development of statistical methods	3. The organization, through its work on databases as well as through interaction with NSOs, assesses the relevance and effectiveness of existing statistical methods and standards, develops and proposes new methods through participation in international meetings
	Technical cooperation	4. UNIDO formulates a programme of technical cooperation in industrial statistics that underlines its relevance, specifies the field of UNIDO expertise and describes the most likely components of the project. STA updates the programme from time to time 5. Prepares a list of countries for which a systematic data gap is reported for a longer period of time and makes recommendations for technical assistance
Accuracy	*Data collection and transformation*	1. Process for editing data and checking for consistency for several data collection stages are clearly defined 2. Data capturing programmes are equipped with built-in tools for screening errors and inconsistencies 3. Unexplained errors and inconsistencies detected from the primary checking process are verified with the original data sources 4. Data transformation stages are assessed, improved and applied to ISIC Rev. 3 database

Quality dimension	Targeted statistical activity	Tasks for quality assurance
Completeness	Data collection and production	1. UNIDO extends the country coverage of its database as far as possible 2. Conducts regular inquiries through its metadata questionnaire to ensure full coverage through industrial surveys of different countries
	Technical cooperation	3. Collects information on the treatment of non-coverage and non-response 4. Encourages countries to lower the cut-off point of industrial surveys or to provide the data for entire sector whenever possible 5. Makes suggestions to NSOs on a sound survey design to improve the coverage of industrial surveys
Timeliness	STA working schedule	1. STA prepares an annual work schedule to ensure the timely publication of its main statistical products
	Provisional estimates methods	2. Encourages and assists NSOs in transmitting the requested data on time 3. STA produces provisional estimates for latest years and assesses the efficiency of the method employed
Comparability	Data transformation and compilation	1. ISIC Rev. 3 remains the main industry classification standard which STA refers to for comparability 2. STA compiles a dataset of historical series by ISIC Rev. 3 to complete the data conversion from ISIC Rev.2
	Data dissemination	3. Data in all UNIDO statistical products are presented by ISIC Rev. 3
	Technical cooperation	4. A list of countries using different statistical units and different output measures is maintained and updated 5. NSOs are recommended to produce output measures at basic prices
Coherence	Data transformation and compilation	1. Statistical tools from the descriptive statistics, such as ratio proportions and mean, are widely used to control the internal coherence of country data 2. Relative variables and mean values are compared across countries and over time and any inconsistencies revealed are corrected 3. Any deviation of published data in UNIDO statistical products from the standard methods is reported in metadata system to explain or to alert users

Quality assurance is an ongoing process. The current framework serves as a departure point from a stage where some quality standards are established. As the Organization aims at continuous improvement in the quality of its products, the framework will be updated regularly to include new dimensions. Quality assurance however depends on the human and financial resources allocated to statistical activities. The current framework is designed on the assumption that additional manpower will be available to STA shortly.

5 Role of metadata for quality assurance

Usually when referring to quality and quality dimensions only statistical data are considered. But since metadata undergo the same life cycle as statistical data, the same quality dimensions can be equally applied to metadata. On the other hand, metadata is the most important driver that can leverage each of the quality dimensions.

The availability of standardized statistical metadata is central to inter-operability and can be a powerful tool. It enables the user to discover and select relevant data quickly and easily. Poor quality metadata, for its part, can cause a dataset to become essentially invisible within a repository or archive and can therefore remain unused. Clearly high quality metadata has an important role to play in realizing the goals set for dissemination of statistical data. Much effort has already gone into developing standardized approaches to metadata, the most remarkable being the achievements of the METIS group: see http://www.unece.org/stats/cmf/. Together with the OECD and Eurostat, the UNECE organizes a working session on statistical metadata every two years. These meetings provide a valuable opportunity for national and international statistical organizations to present and discuss important developments in this field.

UNIDO has also been a forerunner among international organizations in using statistical metadata for the purpose of quality assurance. The experience of UNIDO in metadata management was presented in several articles and conference presentations – see Fröschl et al. (2002), Yamada (2004) and Todorov et al. (2008).

The simple definition of metadata is “data that defines and describes other data” – see for example, OECD (2008) and, respectively, statistical metadata are “data about statistical data”. According to the definition given in UNECE (1995) statistical metadata provide information on data and on processes for producing and using data. Metadata describe statistical data and - to some extent - processes and tools involved in the production and usage of statistical data. Although this definition is

easy to remember and understand, it does not make much sense if taken out of context – that is, in order to identify particular data as metadata one needs to specify the purpose of its use.

For whom and why is metadata necessary? First of all these are the users of the data but not less important are the needs of the producers of the data and finally although often neglected, metadata is necessary to control the proper functioning of the software tools used in the statistical production process.

Statistical metadata documents a dataset so that users can find, understand and evaluate whether the data are appropriate for their intended use, that is, the user can judge the **relevance** of the data with regard to the problem at hand. Statistical metadata also provides information on **accuracy** (precision, reliability) of the statistical data (background, purpose, content, collection, processing, and related information). It is also important for the user to know where and how the data can be found, that is, **availability** of statistical data. This aspect becomes extremely important in today's age of Internet and information technology, since metadata can facilitate the resource discovery. All this information allows researchers to find, understand and manipulate statistical data in the proper way. The availability of such metadata extends the number and diversity of people who can successfully find and use statistical data.

Good and well-managed metadata reduces the time lag (reuse of software, content, procedures, etc.) and thus contributes to the **timeliness** of statistical data. An even more positive impact on the timeliness of statistical data is due to the simultaneous management of data and metadata in one integral production process. The dimension of timeliness, of the UNIDO quality framework, benefits from the almost automatic, metadata-driven procedures for collecting data. The outgoing UNIDO general industrial statistics questionnaire, with previously reported statistical data and metadata for their possible revision by the NSO, is pre-filled automatically by the system, using the metadata available in the system. The questionnaire is created in Excel format in one of the three languages (English, French or Spanish) as appropriate for the particular country. The incoming completed questionnaire by the

NSO with the new and updated data and metadata is read automatically in the system and provides for further validation and transformation by the UNIDO statistical staff.

Good statistical data must be well defined in order facilitate **comparability** within each of the defined dimensions (such as countries, branches of industry, years) which is actually one of the main challenges to the UNIDO statistical production process, especially for least developed countries. Proper use of these statistical data can be ensured only through accurate metadata. Without metadata, the user might misinterpret the difference in country coverage or classification as a change in the measured economical phenomenon. It is therefore necessary to manage metadata on: (i) different and changing classification systems; (ii) the computation and processing methods of the principle indicators (for example, number of establishments or enterprises, number of employees or persons engaged); (iii) the difference in valuation of output measure (basic prices, producer prices or factor prices).

Applying data estimation using supplementary information as well as econometric techniques also increases the timeliness and **completeness** of the data. Keeping track simultaneously of the sources and methods used through automatically generated metadata is of primary importance. Essential for documenting the completeness and the imputation or estimation techniques is the staging framework (a metadata stage is generated which is attributed to each data item) described earlier in this document (see 2.1.2).

Following the **International recommendations for Industrial Statistics,** the development of metadata was accorded high priority and their dissemination is considered an integral part of the dissemination of industrial statistics. Moreover, it is recommended that in consideration of the integrated approach to compilation of economic statistics, the development of a coherent system and a structured approach to metadata across all areas of economic statistics be adopted, focusing on improving their quantity and coverage. Further, the dissemination of statistical data and metadata using web technology and the Statistical Data and Metadata Exchange (SDMX) standards is recommended as a way to reduce the burden of international

reporting (The SDMX technical standards content-oriented guidelines provide common formats and nomenclatures for exchange of and sharing of statistical data and metadata using modern technology).

5.1 UNIDO metadata classification

Metadata are classified according to their usage and role in the statistical production process. The main types of metadata according to these criteria are as follows:

- **Structural (or definitional) metadata** – Structural metadata refer to metadata that act as identifiers and descriptors of the data. Structural metadata are needed to identify, process, retrieve, navigate and interpret statistical datasets – these are, for example, variable names and dimensions of the datasets. The structural metadata exist prior to the data and are created and maintained independently of the data. These are used to define the data structures. Examples of structural metadata are country names and codes, currency names and codes and their relation to the country, definitions of indicators, classifications, such as ISIC Rev. 2, ISIC Rev. 3, etc. Through these core data also some basic metadata elements like metadata classes, stages, sources and methods, etc. are defined. Historically, this metadata type was first established (imported from the mainframe, re-factored and formalized) in ISDE. Structural metadata are maintained by the statistical staff using the tool Nomenclature Explorer, and strictly follow the rules on user authorization and ownership.

- **Implicit metadata** – Implicit metadata are a special class of metadata arising throughout the specific usage of other metadata. Typical examples are the ISIC combinations. For example, several industry categories can be combined and reported together by a given country for a given indicator and time periods. In the questionnaire completed by NSOs, such a combination is expressed in the following manner:

1511	Processing/preserving of meat	1234 a/
1512	Processing/preserving of fish	… a/
1513	Processing/preserving of fruit and vegetables	… a/
...		

Notes: a/ 1511 includes 1512 and 1513.

The codes, 1511, 1512 and 1513, are combined and reported as a single number '1234'. The combined industries are linked by the footnote a/. This is resolved by the system as a dummy ISIC code 1511A defined as "1511 includes 1512 and 1513" which is used throughout the production process and appears accordingly in the publications as well as in the pre-filled questionnaire.

In a similar way, other country-specific classification discrepancies, like industry codes at the three-digit level that **exclude** one or more specific four-digit industry codes can be solved. Implicit metadata can be used also for defining synonyms – for example, '040' is the country code for Austria and is the same as that **substituted by** the ISO code 'AUT' or the International Monetary Fond code '122'. Or for specifying aggregation, for example, the aggregation code 'EU' is composed by the codes of the single countries. The keywords **substitute**, **included**, **excluded** used in the above described context are called operators

- **Operational metadata –** Operational metadata are generated by the process of data transformation and attributed to the respective data items. As described in the data transformation phase, each data item is stored in the database with a stage indicator reflecting its credibility. The transformation process generates also "source" and "methods" metadata, describing the source of the data item and methods applied for its generation.

- **System metadata –** Such metadata are used to drive automated processing throughout the individual phases of the life cycle. These can be layout definitions for the *Yearbook* (for each country, for each edition of the *Yearbook*) as well as country lists, etc., used in the automatic generation of the PDF output; installation and packaging lists,

directories, templates, etc., for creating the CD product. These metadata are specific for the application where they are used and do not relate to the data. Therefore, although stored in the centralized repository, they are maintained by each application separately and are called “properties” of the respective process, namely, *Yearbook* properties, questionnaire properties, etc.

- **Reference (or descriptive, methodological) metadata** describe the contents and quality of statistical data and thus form the main bulk of metadata. They are received from primary data reporters, using the UNIDO questionnaire, and then are further processed together with the data. During this process, additional metadata can be input by the UNIDO statistical staff. Reference metadata includes the following subcategories: (i) “conceptual” metadata, which describes the concepts used and their practical implementation, and helps users to understand what the statistics are measuring; (ii) "methodological" metadata, which describes the methods used for generating the data (for example, sampling, collection methods, editing processes); (iii) "quality" metadata, which describes the different quality dimensions of the resulting statistics (for example, timeliness, accuracy). Reference metadata can be attached to all possible levels ranging from the complete dataset down to individual data items. This is done by assigning the same dimensions as those of data to metadata.

5.2 UNIDO metadata system

The conceptual development of the UNIDO metadata subsystem was initiated in 1999 with the aim of automating information production (data and metadata) using the latest management technology. Keeping in mind the inherent structural complexity of the datasets involved, only a comprehensive metadata-based system re-design approach was considered promising. Thus, the project favoured an **integrated data and data documentation (metadata) framework** emphasizing that, while allowing scrutiny of data documentation (statistical metadata) both

individually and jointly with statistical data, any statistical data access always entails the retrieval of associated metadata without demanding specific inquiry measures or actions. This way a rather tight interrelation of data and metadata is both enforced and assured by purely technical means. However, as its major precondition, this principle presupposes a homogenous representation of all bits of data documentation in order to provide uniform data and documentation access procedures.

It is a good practice in data management, in general, to capture the data at the place and moment where/when they originate. Furthermore, minimal human effort should be involved in doing this. The same is valid for metadata. Also the collection of metadata should ideally be an integral part of the process of creation of the data to which it relates. It is well known that when creating metadata manually as a separate process, following the data capture is prone to error and time consuming. Thus, the creation of different types of metadata items, which are necessary to satisfy the different needs of users (users of the UNIDO databases) and producers (primary data reporters and UNIDO statistical staff), is integrated in the most suitable phase of the statistical production process.

Moreover, as it was imperative that a change in data representation should not disrupt established UNIDO data services, **a smooth migration policy** was called for, leaving interface requirements of downstream systems and data usage almost untouched. While this entailed enormous effort, an expected side-benefit of re-designing the INDSTAT system is its potential applicability to focus on operational data management areas in need of refashion. The concrete design and implementation of this subsystem was realized as a part of an integrated data and metadata system, namely, **Integrated Statistical Development Environment (ISDE)**. ISDE was developed in a stepwise manner in the context of a migration project of the complete UNIDO statistical databases from an IBM mainframe to a client/server platform. Details on the migration project itself, its current status and relation to the newly developed statistical applications and information and communications technology infrastructure are provided later.

An essential requirement for the UNIDO metadata system is that all metadata must be available in three languages (English, French and Spanish). This allows pre-filling each questionnaire in the preferred language for the country and later to process it accordingly.

The integrated system is based on a formal framework, described in detail in Froeschl et al. (2002) and Froeschl and Yamada (2000). The proposed information system architecture comprises two cubes -- one for statistical data and the other for metadata -- interrelated by a set of shared dimensions. Such a data cube resembles a multi-dimensional (cross-sectional) statistical table with each cell holding the value of some indicator (aggregate value, statistical datum) broken down with respect to a couple of cross-classifications (table dimensions). To satisfy the requirements of the UNIDO data structures, the "concept" of a data cube is generalized significantly in the following ways:

- Cross-classifications are used as a *formal* device for any kind of data segmentation, including dimensions for spatial and temporal breakdown as well as dimensions for separating different data-processing stages and even different types of indicators. Each dimension (or *cube edge)* has its particular semantics and must thus be treated differently, especially from the processing point of view.

- The content of a cube cell distinguishes between statistical data and statistical metadata, where formally metadata include any kind of information directly associated with a data cube cell. The subject-matter subdivision of cell-related metadata is made possible through a symmetrical extension of the formal cross-classification concept by "metadata dimensions" which can break down the metadata into different categories (not classes).

The formal framework is narrowed down in the context of the INDSTAT database to form a data cube composed of five edges representing, in turn,

- a temporal breakdown dimension (in years);
- a geographic breakdown dimension (countries);

- a breakdown of data in terms of industry (ISIC Rev. 2 and 3);
- a formal breakdown of data according to processing stages;
- another formal breakdown of data, distinguishing between the economic statistical indicators maintained within INDSTAT

The preparation of appropriate statistical metadata as background information in support of INDSTAT databases requires concrete and well-documented metadata inputs from primary data compilers. Thus, UNIDO requests NSOs to provide, together with available statistical data, such descriptive information through its industrial statistics country questionnaire. The key items for which the Organization needs to obtain meta-information include:

1. name of the reporting agency
2. inquiry on which data are based
3. data-reporting system (major deviation from ISIC)
4. reference period:
   - calendar year
   - fiscal year (Please specify)
5. Reference unit (type of the statistical unit)
   - establishments
   - enterprises
   - other (Please specify)
6. Survey scope - type of reference units covered
   (information on coverage and cut-off size)
7. Employed method of data collection
8. Employed method of enumeration
   (Direct interview, mail or web-surveys)
9. Response rate
10. Have data been adjusted for non-response?
11. Concepts and definitions of variables on which data are reported
    (Details of each indicator)
12. Titles of related publications
13. URLs of related electronic publications on the Internet.

Additionally, it is possible to attach to each data item in the questionnaire one or more metadata items (footnotes in the older UNIDO terminology), like "missing because of confidentiality reasons" or combinations of ISIC codes like "1511 includes 1512", etc – see Figure 5.

Figure 5. Metadata in the collection phase

	A	B	C	D	E	F	G	H
1	UNIDO: General Industrial Statistics Questionnaire							2005 Edition
2	Country: ALBANIA		UN Code: 008					Table 05
3				Wages and salaries paid to employees				
4				(thousands of Albanian Leks)				
5	ISIC Rev. 3	INDUSTRY	Note	2001	2002	2003	2004	2005
153	3694	Games and toys		...	...	...	...	...
154	3699	Other manufacturing n.e.c.		...	...	...	...	...
155	3710	Recycling of metal waste and scrap		...	...	...	...	15491q/
156	3720	Recycling of non-metal waste and scrap		...	...	...	...	...q/
157	D	Total manufacturing		5173538p/	5988617p/	8703238p/	9004800p/	9228711p/
158	REMARKS							
159	a/ 1511 includes 1512.							
160	b/ 1531 includes 1532.							
161	c/ 154 includes 1533.							
162	d/ 171 includes 172, 1730, 1810 and 1820.							
163	e/ 191 includes 1920.							
164	f/ 2010 includes 202.							
165	g/ 210 includes 221, 222 and 2230.							
166	h/ 2310 includes 2320 and 2330.							
167	i/ 241 includes 242, 2430, 251 and 2520.							
168	j/ 2610 includes 269.							
169	k/ 2710 includes 2720, 273, 281, 289, 291, 292 and 2930.							
170	m/ 3000 includes 3110, 3120, 3130, 3140, 3150 and 3190.							
171	n/ 3610 includes 369.							
172	p/ Sum of available data.							
173	o/ 241 includes 2310, 2320, 2330, 242, 2430, 251 and 2520							
174	q/ 3710 includes 3720							
175								

Returned questionnaires from NSOs provide detailed information on the data supplied to UNIDO. This allows data to be checked thoroughly in terms of several quality dimensions, such as accuracy (method of data collection), completeness (coverage of the survey in relation to total manufacturing and non-response treatment), comparability (classification method, valuation) and so on. Metadata is further transformed, in a similar manner as data transformation, in order to bring them in international context and to explicitly indicate any deviation in national data from international standards. The deviation may relate to statistical methods or classification standards.

Data for OECD member countries, collected through joint OECD/UNIDO questionnaire and transmitted to UNIDO (Excel format) are entered into the system in a similar way and are ready for further validation and processing. These questionnaires do not contain metadata, therefore the necessary metadata are extracted from other OECD publications - currently OECD (2003) Industrial Structure Statistics, vol. 1, Core Data.

First of all, UNIDO statisticians make every effort to correct or adjust the indicated deviation to make the data internationally comparable. However, an excessive intervention could adversely affect the accuracy of originally reported data, or lose consistency with national data sources. Therefore, UNIDO prefers to take a more responsible approach by providing users with supplementary information, that is, metadata indicating the limitation of statistics produced by the Organization in terms of accuracy and international comparability. This approach ensures the accuracy and coherence of data with another important quality dimension of comparability of international statistics.

UNIDO statisticians are currently working on further improvement of its metadata system. In order to make greater use of metadata in quality assurance as well as in broader statistical analysis, it is important to distinguish between structural metadata and reference metadata. The structural part of the metadata would allow some grouping of countries concepts and methods used in their data collection. For example, a number of countries use cut-off point or report value added at producers' prices. Such information is very helpful to users who want to make cross-country comparative analysis of industrial performance. Reference metadata, for its part, can be used to analyze overall assessment of methods and standards employed by NSOs.

The metadata collected from NSOs and OECD together with the data undergo the same transformation process as the data and is complemented by metadata generated during the transformation process. All resulting metadata, including the necessary structural metadata, are used in the dissemination process:

- To define the dissemination products – for this purpose structural metadata, like country names and codes, currency names and codes, classifications, etc. are used;

- To guide the dissemination process – for example, the selection of data to be published in the different products depends on the degree of confidence they deserve, as identified by the stage (metadata generated in the transformation process);

- To provide users with information they may need to interpret the disseminated data.

Figure 6, Figure 7, Figure 8 and Figure 9 give examples of metadata presented in the different dissemination products.

Figure 6. Metadata in the dissemination phase: different types of metadata visible in the data viewer of the CD product INDSTAT4

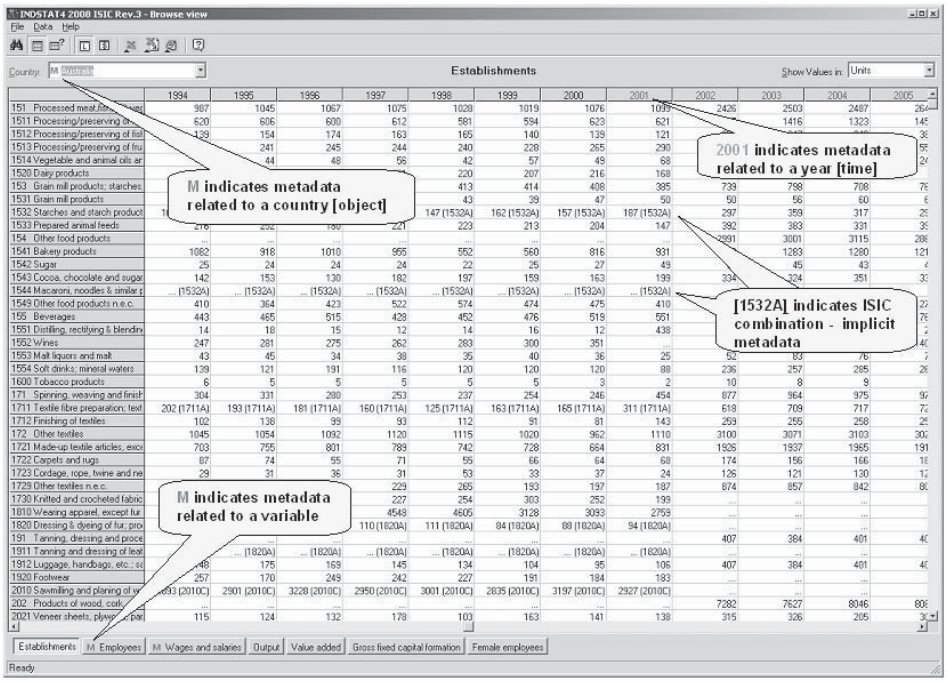


Figure 7. Metadata in the dissemination phase: different types of metadata visible in the Web Country Statistics

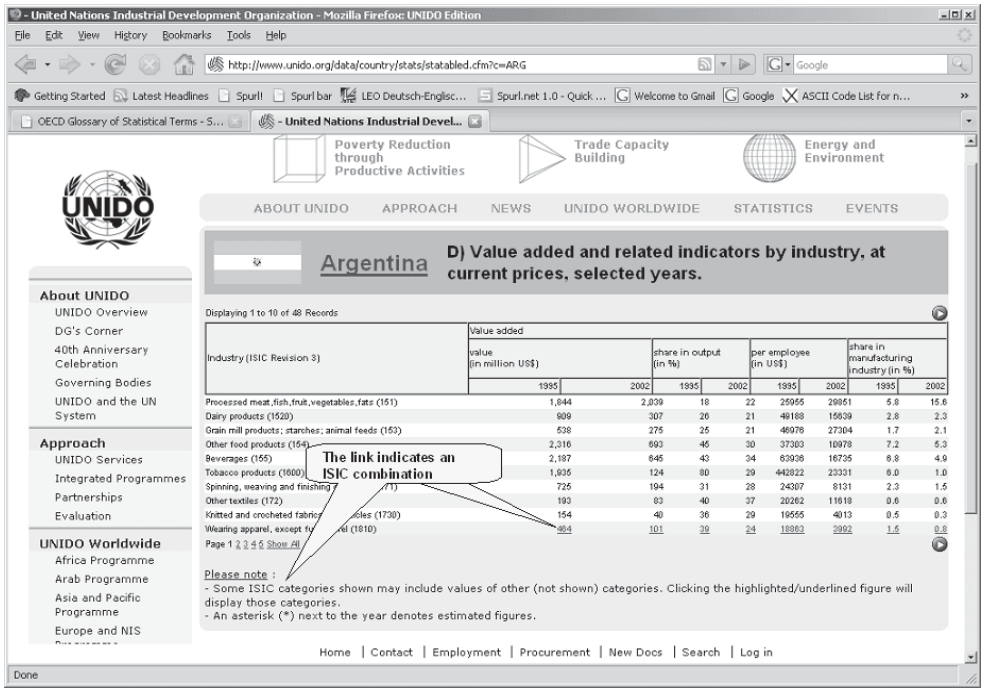


Figure 8. Metadata in the dissemination phase: methodological metadata shown in the Yearbook

Australia

Supplier of information:
Australian Bureau of Statistics, Canberra.
Industrial statistics for the OECD countries are compiled by the OECD secretariat, which supplies them to UNIDO.

Basic source of data:
Annual manufacturing survey.

Major deviations from ISIC (Revision 3):
Data have been compiled on the basis of the Australian and New Zealand Standard Industrial Classification (ANZSIC). This classification is not fully compatible to the ISIC (Rev. 3), especially at a high level of disaggregation.

Reference period:
Fiscal year ending 30 June of the year indicated.

Scope:
All establishments with paid employees.

Method of data collection:
Questionnaires are distributed by mail.

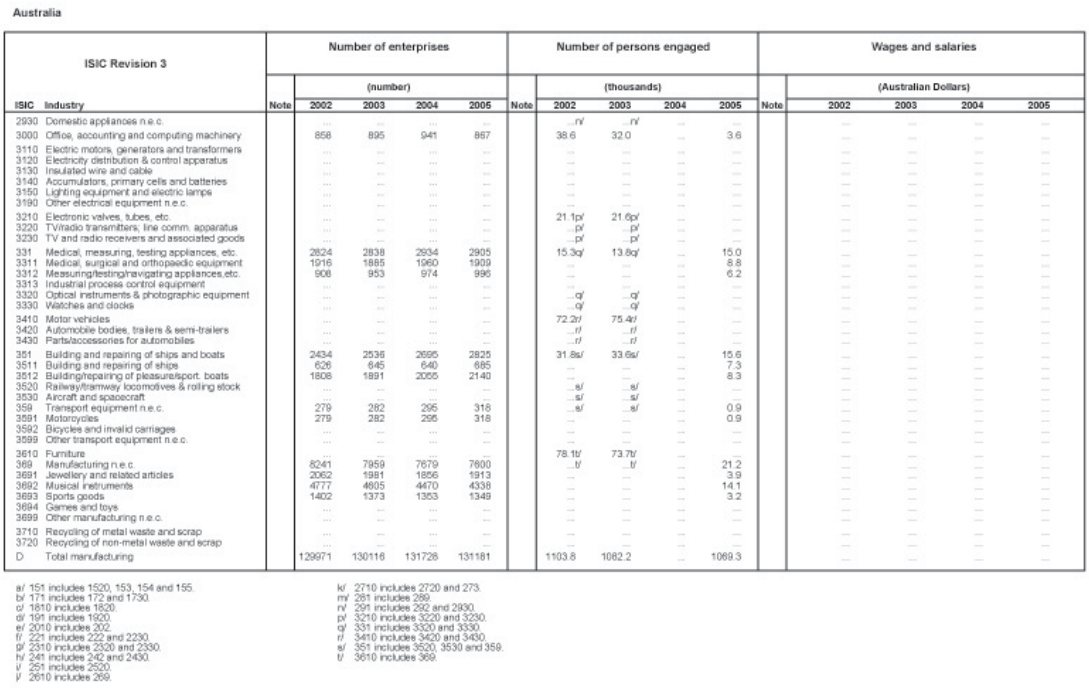
Type of enumeration:
Not reported.

Adjusted for non-response:
Yes.

Concepts and definitions of variables:
Number of persons engaged is not converted to full-time equivalent.
Wages and salaries includes salaries and fees of directors and drawings from profits by working proprietors and partners but excludes payments to consultants, contractors and persons paid solely by commission. Payments and benefits in kind are excluded.
Output is measured as the sum of the value of sales of goods or services produced by the unit, receipts for industrial work done or industrial services rendered to others, resales and net change in stocks of goods purchased for resale in the same condition as received, net change in stocks of finished goods and in stocks of work in progress manufactured by the unit, value of electricity sold and value of capital assets manufactured for own use. Advertising incomes are also included. Valuation is made at basic prices.
Value added is output less value of industrial materials and services for processing, value of fuels and purchased electricity, cost of industrial services received, and payments for contract and commission work. Valuation is made at basic prices.

Related national publications:
Manufacturing Industry, published by the Australian Bureau of Statistics, Canberra.

Figure 9: Metadata in the Dissemination Phase:
Metadata for data elements shown in the Yearbook



Australia

ISIC Revision 3		Number of enterprises				Number of persons engaged					Wages and salaries				
			(number)				(thousands)					(Australian Dollars)			
ISIC	Industry	Note	2002	2003	2004	2005	Note	2002	2003	2004	2005	Note	2002	2003	2004	2005
2930	Domestic appliances n.e.c.		...	...	...	...		../r/	../r/	...	...		...	...	...	...
3000	Office, accounting and computing machinery		858	895	941	867		38.6	32.0	...	3.6		...	...	...	...
3110	Electric motors, generators and transformers		...	...	...	...		...	...	...	...		...	...	...	...
3120	Electricity distribution & control apparatus		...	...	...	...		...	...	...	...		...	...	...	...
3130	Insulated wire and cable		...	...	...	...		...	...	...	...		...	...	...	...
3140	Accumulators, primary cells and batteries		...	...	...	...		...	...	...	...		...	...	...	...
3150	Lighting equipment and electric lamps		...	...	...	...		...	...	...	...		...	...	...	...
3190	Other electrical equipment n.e.c.		...	...	...	...		...	...	...	...		...	...	...	...
3210	Electronic valves, tubes, etc.		...	...	...	...		21.1p/	21.6p/	...	...		...	...	...	...
3220	TV/radio transmitters; line comm. apparatus		...	...	...	...		../r/	../r/	...	...		...	...	...	...
3230	TV and radio receivers and associated goods		...	...	...	...		../r/	../r/	...	...		...	...	...	...
331	Medical, measuring, testing appliances, etc.		2824	2838	2934	2905		15.3q/	13.8q/	...	15.0		...	...	...	...
3311	Medical, surgical and orthopaedic equipment		1916	1885	1960	1909		...	...	...	8.8		...	...	...	...
3312	Measuring/testing/navigating appliances,etc.		908	953	974	996		...	...	...	6.2		...	...	...	...
3313	Industrial process control equipment		...	...	...	...		...	...	...	...		...	...	...	...
3320	Optical instruments & photographic equipment		...	...	...	...		../q/	../q/	...	...		...	...	...	...
3330	Watches and clocks		...	...	...	...		../q/	../q/	...	...		...	...	...	...
3410	Motor vehicles		...	...	...	...		72.2r/	75.4r/	...	...		...	...	...	...
3420	Automobile bodies, trailers & semi-trailers		...	...	...	...		../r/	../r/	...	...		...	...	...	...
3430	Parts/accessories for automobiles		...	...	...	...		../r/	../r/	...	...		...	...	...	...
351	Building and repairing of ships and boats		2434	2536	2695	2825		31.8s/	33.6s/	...	15.6		...	...	...	...
3511	Building and repairing of ships		626	645	640	685		...	...	...	7.3		...	...	...	...
3512	Building/repairing of pleasure/sport. boats		1808	1891	2055	2140		...	...	...	8.3		...	...	...	...
3520	Railway/tramway locomotives & rolling stock		...	...	...	...		../s/	../s/	...	...		...	...	...	...
3530	Aircraft and spacecraft		...	...	...	...		../s/	../s/	...	...		...	...	...	...
359	Transport equipment n.e.c.		279	282	295	318		../s/	../s/	...	0.9		...	...	...	...
3591	Motorcycles		279	282	295	318		...	...	...	0.9		...	...	...	...
3592	Bicycles and invalid carriages		...	...	...	...		...	...	...	...		...	...	...	...
3599	Other transport equipment n.e.c.		...	...	...	...		...	...	...	...		...	...	...	...
3610	Furniture		...	...	...	...		78.1t/	73.7t/	...	...		...	...	...	...
369	Manufacturing n.e.c.		8241	7959	7679	7800		../t/	../t/	...	21.2		...	...	...	...
3691	Jewellery and related articles		2062	1981	1856	1913		...	...	...	3.9		...	...	...	...
3692	Musical instruments		4777	4805	4470	4338		...	...	...	14.1		...	...	...	...
3693	Sports goods		1402	1373	1353	1349		...	...	...	3.2		...	...	...	...
3694	Games and toys		...	...	...	...		...	...	...	...		...	...	...	...
3699	Other manufacturing n.e.c.		...	...	...	...		...	...	...	...		...	...	...	...
3710	Recycling of metal waste and scrap		...	...	...	...		...	...	...	...		...	...	...	...
3720	Recycling of non-metal waste and scrap		...	...	...	...		...	...	...	...		...	...	...	...
D	Total manufacturing		129971	130116	131728	131181		1103.8	1082.2	...	1069.3		...	...	...	...

a/ 151 includes 1520, 153, 154 and 155.
b/ 171 includes 172 and 1730.
c/ 1810 includes 1820.
d/ 191 includes 1920.
e/ 2010 includes 202.
f/ 221 includes 222 and 2230.
g/ 2310 includes 2320 and 2330.
h/ 241 includes 242 and 2430.
i/ 251 includes 2520.
j/ 2610 includes 269.
k/ 2710 includes 2720 and 273.
m/ 281 includes 289.
n/ 291 includes 292 and 2930.
p/ 3210 includes 3220 and 3230.
q/ 331 includes 3320 and 3330.
r/ 3410 includes 3420 and 3430.
s/ 351 includes 3520, 3530 and 359.
t/ 3610 includes 369.

From a technical point of view, the metadata system is part of the ISDE and provides end-to-end metadata services throughout the statistical production provides end-to-end metadata services throughout the statistical production process. It was developed in the context of the migration from the mainframe to a client/server platform. Figure 2 presents the overall structure of ISDE and its relation to the statistical production life cycle. The client part of the system is presented to the user as a desktop application -- the **ISDE shell --** that serves as a container for client/side applications. These applications are described briefly below.

- **ADMIN** – provides administrative services, like user and authorization management, logging and auditing of the system, backup and restore management;
- **Nomenclature Explorer** is the tool used for maintaining the core definitional metadata, which is not related to particular data items but rather serves to define the structure of the data and metadata. These first two applications are outside the life cycle;

- **Questionnaire** is the application for managing the pre-filling and distribution of questionnaires to member countries (that is, used in the *Initialization* phase);

- **Data Wizard** is the main data and metadata maintenance tool used in the *data collection* and *transformation* phases of the life cycle. It provides services for:
    i. reading in the data and metadata from the completedExcel questionnaire
    ii. initial validation of the read-in data and storage in the database (at stage 1)
    iii. maintenance of the metadata
    iv. screening
    v. aggregation and further data validation and transformation

- **Presentation Wizard** is mainly a visualization tool, which can be used in the dissemination phase for answering ad hoc requests. Because of its versatile functionality it is widely used also in the *data transformation* phase;

- **Publication applications -** these are applications used in the dissemination phase for generating different publication products

    i. **Yearbook** –a complex set of applications is necessary for the production of the *International Yearbook of Industrial Statistics*, including aggregation, layout, PDF file generation according to pre-defined templates and other tools. The final result is a publication -- PDF file of some 700 pages;

    ii. **INDSTAT CD** – used to produce the INDSTAT type CD products;

    iii. **IDSB CD** – used to produce the IDSB type CD products;

    iv. **WEB** – used to generate the necessary data and metadata for updating the WEB dissemination database (this database is outside the ISDE system, and is managed by the computer section);

- **Other applications** –this category includes any other applications used in the process, like SAS, R, tools for compilation of production index

numbers and national accounts data (which are beyond the scope of this report) and others.

5.3 Statistical data and metadata exchange (SDMX)

The SDMX initiative is an international project carried out by several international organizations, including the United Nations, specialized agencies, OECD and Eurostat. (Since 2008 Eurostat holds the chair of the SDMX initiative). SDMX has been endorsed by the UN Statistics Commission as the preferred method for use by the international statistical system. SDMX aims at defining standard formats, information technology architecture and content-oriented guidelines for the national and international exchange of statistical data and metadata.

The SDMX Technical Standards Version 2.0 developed and reviewed with the goal to replace, within the context of the International Organization for Standardization (ISO), of the previous version (ISO/TS 17369:2005 SDMX), which provides technical specifications for the exchange of data and metadata based on a common information model. Its scope is to define formats for the exchange of aggregated statistical data and the metadata needed to understand how the data are structured. The major focus is on data presented as time series, although cross-sectional XML formats are also supported. Version 2.0 Technical Standards though backward are compatible with the earlier Version 1.0 efforts, which focused on XML- and EDIFACT-syntax data formats. The latest work broadens the technical framework to support wider coverage of metadata exchange as well as a more fully articulated architecture for data and metadata exchange.

As it takes years for a large number of national and international agencies to adopt common standards, SDMX started with a limited number of agencies that have required technical facilities in place and are familiar with sharing data. The practical utilization of SDMX standard is still in its infancy, not only in UNIDO but also in most international organizations. Some prominent pilot projects (not a complete list), from which lessons can be learned include:

a) SDMX Open Data Interchange (SODI) which is a data-sharing and exchange project within the European Statistical System. The project started with a pilot exercise involving National Statistical Institutes of France, Germany, the Netherlands, Sweden and the United Kingdom. The statistical institutes of Denmark, Italy, Norway and Slovenia joined the pilot exercise in 2006, while Finland and Ireland joined in 2007.

b) FAO CountrySTAT, which is based on the application of data and metadata standards of FAOSTAT and SDMX, is a web-based system being developed since May 2004 using PX-Web at FAO Headquarters. It was successfully tested in the statistical offices of Kenya, Kyrgyz Republic and Ghana during 2005. Many other developing and developed countries have shown an interest in and are adopting it: http://www.fao.org/es/ess/countrystat/

c) Data exchange between OECD and IMF:I Exchange Rates data from IFS.

There are many ways to use SDMX to exchange data characterizing this activity in simple terms. For example, a primary distinction can be made on whether the data are being sent by one counter party to another (called a "push" scenario) or whether the data are posted in an accessible location, and then obtained when needed (called a "pull" scenario). In the push mode, which is the traditional data-sharing mode, different means, such as e-mails and file transfers, are used to exchange data. It shows how UNIDO and many international agencies collect data from NSOs and international organizations. To use SDMX for data-reporting or data collection, which are actually two aspects of the same task, that is, the task of data exchange, at least two counter parties are required, one or more of which provides data to another. Any of these counter parties must adopt the same technical standards, have a common data structure and use common vocabulary. The lack of necessary technical faculties could be a serious stumbling block for developing countries involved in the process.

As a first step towards SDMX utilization, UNIDO is currently developing a data and metadata exchange procedure based on the web service provided at OECD.Stat. OECD.Stat is the central repository where validated statistical data and metadata are

stored, and is intended in due course to become the sole coherent source of statistical data and related metadata for the OECD statistical publications. This will allow one to automatically retrieve and process data for all OECD countries, which is currently done by transferring Excel files.

6 Annex I - Fundamental Principles of Official Statistics

URL: http://unstats.un.org/unsd/methods/statorg/FP-English.htm

Preamble

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, had 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 democratic 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.*

References

Bruce E. Bargemeyer and Daniel W. Gillman (2000) "Metadata Standards and Metadata Registries: An Overview", paper presented at the *International Conference on Establishment Surveys II*, Buffalo, New York.

Carson, Carol (2000) What Is Data Quality? A Distillation of Experience, URL: http://www.thecre.com/pdf/imf.pdf, IMF, 2000.

Cathryn S. Dippo and Bo Sundgren (2000)" The Role of Metadata in Statistics", paper presented at the *International Conference on Establishment Surveys II*, Buffalo, New York.

Fröschl, K.A., Yamada, T. and Kudrna, R. (2002) Industrial Statistics revisited: From Footnotes to Metadata-Information Management, *Austrian Journal of Statistics*, **31** 1, pp. 9-34.

Fröschl, K.A. and Grossmann, W. (2001) Deciding Statistical Data Quality. In: New Techniques and Technologies for Statistics/Exchange of Technology and Know-how, Pre-Proc. NTTS/ETK 2001, Vol. 1; Hersonissos (Kreta), 2001, pp. 567–575.

Fröschl, K.A. and Yamada, T. (2000) The UNIDO Industrial Statistics Information Exchange Architecture (An Integrated Statistical Data and Data Documentation Framework). Working Paper No. 19, Work Session on Statistical Metadata (METIS), Conference of European Statisticians (Washington D.C., 28–30 November 2000), UN/ECE, 2000.

Proceedings of the European Conference on Quality in Survey Statistics; (24-26 April 2006), Cardiff, UK, URL: www.statistics.gov.uk/q2006

OECD (2003), Quality framework and guidelines for OECD statistical activities, URL: http://www.oecd.org/dataoecd/26/42/21688835.pdf

SDMX User Guide: Getting started 2006, URL: www.sdmx.org

Sundgren, B. "*Information Systems Architecture for National and International Statistical Offices" Guidelines and Recommendations*. Conference of

European Statisticians Statistical Standards and Studies No. 51, United Nations 1999.

Sundgren, B. (2008) *"*Classification of Statistical data*"*, Working Paper No. 7, Work Session on Statistical Metadata (METIS), Conference of European Statisticians (Luxembourg, 9-11 April 2008).

Todorov, V. and Upadhyaya, S. (2008) Case Study – United Nations Industrial Development Organisation (UNIDO). Working Paper No. 20, Work Session on Statistical Metadata (METIS), Conference of European Statisticians (Luxembourg, 9-11 April 2008).

United Nations and others (1993) System of National Accounts.

UNECE (2006) Proceedings of the Work session on statistical metadata, Geneva (3-7 April 2006),
URL: http://www.unece.org/stats/documents/2006.03.metis.htm

UNECE (2000) Terminology on Statistical Metadata.

UNECE (1995) Guidelines for the Modelling of Statistical Data and Metadata.

UNSC and UNECE (1995) Guidelines for the modelling of statistical data and metadata.

UNIDO (2006) Measuring ICT sector through Industrial statistics, Working paper by Shyam Upadhyaya and Laureta Kazanxhiu, PCF/RST/STA.

UNIDO (2002) *Industrial Development Report 2002/2003*, Vienna.

UNIDO (1999) Measure by Measure, Building UNIDO's System of Industrial Development Indicators (SIDI); Statistics and Information Networks.

UNIDO (1996) Industrial Statistics Database: Methodological notes; 1996.

United Nations (2008) International recommendations for industrial statistics,
URL: http://unstats.un.org/unsd/statcom/doc08/BG-IndustrialStats.pdf

United Nations (1983) International recommendations for industrial statistics; Series M No 48, New York.

United Nations (1981) Recommendations for the 1983 World Programme of Industrial Statistics, Series M No 71, New York.

Upadhyaya, S. (2008) Industrial Statistics Programme of UNIDO for Countries of Developing and Transitional Economies UNIDO, Working paper and presentation on EGM on Industrial Statistics, 15-17 July, New York.

Yamada, T. (2004) Role of Metadata in Quality Assurance of Multi-country Statistical Data. Paper presented at the Conference on data Quality for International Organizations, Wiesbaden, Germany, 27-28 May 2004.

Printed in Austria
V.08-58885—January 2009—200

UNITED NATIONS INDUSTRIAL DEVELOPMENT ORGANIZATION
Vienna International Centre, P.O. Box 300, 1400 Vienna, Austria
Telephone: (+43-1) 26026-0, Fax: (+43-1) 26926-69
E-mail: unido@unido.org, Internet: http://www.unido.org