

# **USERS NEEDS : AN IMPACT ON STATISTICS**

Jana Meliskova, consultant of the Czech Statistical Office, Geneva, Switzerland

## **Summary**

Recent developments in the world economy and information technologies influence significantly the spectrum of statistical users and their needs. Globalization has placed new demands on national and international statistical agencies. Communication technologies integrate information flow on national and international level. New users and users' groups appear. It calls for harmonization , standardization and update of many statistical concepts

The general impact of this development in statistical organizations is a shift to client centric services. Users are becoming partners of statistical agencies in designing and implementing their statistical production process. However, who the users are, what are their needs and, how are those needs impacting statistical information system is a still remaining challenge for many statistical organizations.

The aim of the paper is to draw attention to this complex phenomena and to point out some reasoning why the role and needs of statistical users should be more focused by national statistical institutes (NSIs).

The paper is exploring in more detail problems in the following groups of statistical users:

- (i) multinational enterprises (MNEs) – integration with statistics, measurement of MNE(s) activities, international developments;
- (ii) national administrations- statistical and administrative concepts, standardization at the EU level;
- (iii) clients looking for geographic searching- joint Geographic Information System (GIS) projects among intergovernmental institutions- spatial data warehousing.

When writing the paper, the author benefited from practical experiences in the work of the project AMRADS ( Accompanying Measure for R&D in Statistics ) organized by the European Commission. The author reflected also considerations related to this topic in the work program of the Conference of European Statisticians.

Section 1 presents reasoning why nowadays the needs of external user should be more focused by NSIs when designing and implementing their information systems. Section 2 explores the impact of MNEs on design of statistical information system. Section 3 considers common trends in NSIs towards taking increasing advantage for the use of administrative sources for the production of national statistics. Section 4 brings some succesful examples reflecting growing demand of users for geo-spatial search facilities. Section 5 comprises some concluding remarks.

## **1. EVIDENCE THAT EXTERNAL USERS SHOULD BE FOCUSED**

### **1.1 Globalization**

Globalization has placed new demands on national and international statistical agencies to provide the information necessary to inform policy in today's increasingly interdependent world economy. It has manifested itself in the interdependence of financial markets, the increased role of multinational corporations/enterprises, the transfer of technology, the increasing dependence of domestic markets on foreign trade, and the necessary interdependence of monetary, fiscal, and regulatory policy. It has led to increased demands for harmonization in world statistical standards and related metadata.

The process calls for harmonization, standardization and update of many already approved statistical data and metadata models like system of national accounts, balance of payments, business accounting system, e-commerce regulations, international data dissemination standards and many other tasks. The needs for new models of statistical data and metadata might appear.

External needs increasingly impact architecture of national statistical information system. New groups of users are appearing. Multinational enterprises and companies are such example. NSIs are facing new tasks which can be solved only in close cooperation with all supranational and international organizations involved. The research should play an eminent role in exploring and measuring of multinational enterprises and related issues.

### **1.2 Challenge at national level**

Because of growing number of e-government initiatives, NSIs are at the national level confronted with requests of other governmental institutions to adapt their statistical information system. Integration of data and metadata at governmental level is more pertinent. Traditional survey oriented approach becomes conflicting against the cross-cutting users needs.

It calls very often for redesigning of a whole statistical information system (data, metadata) encompassing the vision, legislation, planning, implementation and consequently all phases of the production process. Prominent is namely to integrate statistics with the public administration and business information systems. It calls for methodological work on definition of statistical data and metadata. Well-orchestrated knowledge and metadata management is an important condition for success achievement. Experiences in this work are fragmented.

### **1.3 Audience of national external users**

The users' audience at the national level grows larger and much more divers. Such development provides often the means for closing the loop between the statistical data and metadata suppliers and their users. For example, many national governmental institutions, particularly state administration, business enterprises communities, health, environment, education and justice communities, require shared access to micro-data and/or request NSI to produce value-added outputs. In these cases, the respondents are both suppliers and users of statistical data and metadata. Corporate (on governmental level) metadata repositories, the design of which is based on the models relevant to the needs of divers groups of external users, proved to be an efficient solution in some countries.

Besides governmental institutions the following important users group can be distinguished: subject matter researchers, political decision makers, public officials, executives, teachers, students, librarians, archives and journalist. Irrespective of relevance to a specific group, the users seeking the information through Internet, in general, needs metadata for the following functions: to see what data are available, to assist in the search for information, to interpret the information and, if necessary, to assist post processing of the information.

The role of data archives as users of official statistics is growing. There is an increased demand from research, academia and other users for information stored and maintained in data archives. Demand for speedy delivery is pushing data archives to upgrade their information systems. Important part of such system is implementation of acquisition policy. Integral part of it should be a well-developed documentation of datasets made available to the potential users of archives.

The documentation should include a numerous metadata regarding the methodology employed, description of sampling methods, sample size, response rate etc. Accountability should be ensured. Closely related to the relevant documentation this incorporates correct use and application of the data, accurate interpretation of results and the ability to reproduce the work. Those and other important challenges like data quality and data confidentiality should be taken into the consideration. Since NSIs are important suppliers of data and metadata to archives they should ensure that above mentioned metadata will be to the archives systematically delivered.

#### **1.4 Impact of spatial technologies**

There is a large and growing demand for geospatial search activities, for example users who want to find out if there is any statistics, covering a particular object which relate to a specific city, area or region. The rapid development in mapping technologies has visibly upgraded the power of statistical analyses and their dissemination. Focused on spatial dimension and using diverse remote sensing methods, geographic information systems (GIS) make the integration of statistics borderless and statistical analysis more value-added, permitting the evaluation of statistical data across traditional subject-matter fields. Some countries conduct on governmental level projects aiming to develop corporate geospatial warehouses. Such activities call for accelerate integration with statistical metadata.

#### **1.5 Challenge at international level**

Individuals, foreign and multinational enterprises, international organizations (IOs) and other agencies (national, international) are important users of statistics. Their demand for greater consistency when interacting with NSIs is more and more imminent. In the case of IOs the metadata and data requirements (and their collection and exchange) have to be coordinated not to overburden countries with duplicate requests. In order to fulfill this task, better integration of metadata and data between IOs is needed.

A lot of metadata is available on websites of IOs. Links could be inserted from the metadata of IOs to the more detailed metadata on national websites. Coordination of access could be achieved through a single gateway for data and metadata, e.g. through a portal side. To this end, the IMF has launched (together with OECD, EUROSTAT, UNSD, UNECE, the Bank of International Settlement and the European Central Bank) a joint cooperation.

The Task Force on the project called Statistical Data and Metadata Exchange (SDMX) was created. SDMX is at present under the focused development. Needs of international users and related activities increasingly impact architecture of national metainformation systems. NSIs are facing new tasks which can be solved only in close cooperation with all IOs and other international users involved.

## **1.6 Stakeholder relations**

Together with dissemination policy, there has to be a policy for user support in statistical office. This support should cover both technical and functional issue. It has to be clear to users what level of service they can expect from NSI for each of these areas. It is essential that NSIs develop and maintain a mechanism for interacting with the individual users groups. Advisory committees and councils and various users and distributor consultation forums could constitute such mechanism. Some examples how NSI might carry out such interacting mechanism are:

- User needs can be assessed through survey and special studies.
- Advisory activities. The establishment of a website for all advisory groups would facilitate an exchange of dialogue.
- Joint program and research activities.
- Development of Extranet or virtual private networks to ensure secure exchange of information.
- Use of XML format for real time message delivery with administrative partners. This would make it possible to mutually use administrative sources within public administration and basic nation-wide registers such as population register, business register, etc.
- User consultation activity. Users need to be consulted with respect to survey content and to develop the agency product's and services.
- Impacting libraries, academia and researchers. An increasing number of products and services are moving on-line. It can provide unprecedented opportunities to involve the above-mentioned communities in program planning.
- On-line training capability to help support the dissemination of information and to provide an educational opportunity for the general public.

## **1.7 Usability testing**

Strengthening the role of external users provides new opportunities for NSIs to analyze users feedback, to monitor users' behaviour and, through this process, to maintain up-to-date data dissemination strategy. Cognitive studies and usability testing can contribute significantly in solving this task.

To conduct feasibility studies on divers communities of metadata stakeholders, exploring what they want and what their expertise are is vital to design statistical metadata repositories and websites efficiently. This is, however, a very demanding and expensive process. NSIs in many cases lack resources as well as specialists with the necessary expertise for their own research. Therefore, the role of supranational R&D in solving this task is very important.

## 2. MULTINATIONAL ENTERPRISES

MNEs have a potentially dramatic impact on a nation's economy, from the volume of goods and services they export and import to the labour income they generate for domestic workforce. Measuring their impact on national economies is critical for NSIs. They increasingly take advantage of more liberal trade environments and modern technologies to integrate their activities across the borders.

Their information systems are designed accordingly and may not correspond to the models and concept used by, and/or the territorial jurisdictions of NSIs. Such situation is potentially leading to gaps, duplications or other anomalies in the measurement of the activities of MNEs by NSIs. This can have negative consequences such as biases in national economic statistics, GDP calculations, trade and balance of payments statistics and many others.

Understanding the behaviour and impact of MNEs is important to assess the effects of globalisation. A MNE's investment may result in a significant injection of capital into the local economy of a given area, provide jobs directly or through the growth of local ancillary businesses such as banks, insurance companies. It can initiate a multiplier process, generating more income as newly employed workers spend wages on consumption. Conversely, the impact may be exact the opposite, if an MNE moves out from the area.

### 2.1 Measurement of activities

How to best ensure the accurate measurement of activities of MNEs in a given country and the impact abroad of MNEs based in the same country? This is a quite new domain for which data and metadata models should be developed. Clearly, all traditional models like business statistics, SNA, and many others would be affected. Diverse proposal on how to solve this problem were developed (e.g. by Statistics Canada, US Bureau of Economic Analysis, OECD, IMF).

The characteristic feature all of them is that such work should be organized internationally, the isolated national approach could not meet the foreseen results. MNEs, NSIs and international organizations (OECD, IMF, Eurostat) are expected to cooperate in such project. The important place in this work would have national and international research. The goal of such project would be to explore whether there are biases in national economic statistics caused by gaps and overlaps in the coverage of activities of MNEs by NSIs.

### 2.2 Standardization.

Research mentioned above would help to determine whether the impact of MNEs activities can be alleviated by the NSIs further standardising in the following areas:

- **Definitions of forms of organizations**, and of ownership and control (for example full ownership, joint ventures, portfolio participation);
- 
- **Statistical units** and how they relate to the bookkeeping practices of MNEs.  
What entities are acceptable to MNEs in terms of accounting;  
How MNEs define, measure and allocate profit (by product line, geographic or organizational entities etc.)?  
How these practices impact on the ability of NSI to define statistical units and measure value added for the various component parts of an MNE?

- **Chart of accounts :** can statistical community define a common structure of financial accounts that allows the linking of bookkeeping practices of enterprises to SNA concepts? Could this chart of account be used in conjunction with the emerging **XBRL** (extensive Business Reporting Language, a rapidly emerging standard for business and financial reporting) to evolve a statistical reporting framework compatible with business practice, commercial accounting software and SNA concepts?
- 
- **Classifications :** How MNEs define and keep track of their activities and outputs? Are these definitions compatible with the frameworks used by statisticians? Several ongoing efforts are already on the way (e.g. in industrial classifications). MNEs are major players in the production and trade of services and must be key considerations in evolving services classifications.

### 2.3 Manual on Economic Globalisation Indicators

The prominent part of the work is conducted by **OECD**. OECD has already published guidelines for MNEs, which provide voluntary principles and standards for responsible businesses. The document is called “ Manual on Economic Globalisation Indicators” and is planned to be published in 2004.

The manual map out the realm of globalisation indicators being relevant to policy makers and propose a selection of them that would warrant a systematic compilation effort. It provides methodological and statistical guidelines needed to construct the chosen indicators, incorporating as much international harmonization as possible. It provides compilers with clear, practical and consistent guidance on the collection of activities of MNE data on inwards and (secondarily) outwards investment. The document creates an important basis for further work on harmonization and standardization on MNE and its integration, including development of metadata, with existing statistical concepts and standards.

## 3. NATIONAL ADMINISTRATIONS

Governments all over the world draw upon official statistics for the use in the administration of important public programs. Whether to determine socio-economic program parameters by region, to determine national contributions to international organizations, to administer airports, to establish quota or subsidy levels in international trading agreements or to share tax revenues equitably in overlapping jurisdictions. These uses are typically among the most important mandate of an NSI.

National public program administrations are, and will be even more in future, important users of statistics. The exploitation of official statistics in government administration may offer significant savings while, at the same time, reducing the overall compliance load of businesses. A visible shift forwards can be observed in the relation of NSIs with state administration institutions and business enterprise communities.

Common trends in NSIs towards taking increasing advantage of the use of administrative sources for the production of national statistics can be clearly recognized. Rapid development in ICT makes administrative data technically available and this could bring significant cost savings for NSIs.

### 3.1 Statistical and administrative concepts of data and metadata

**At the national level**, it can be observed that the administrative sources are used in the similar statistical subject matter areas. Business registers, VAT and income tax register, social security

and customs administrations are the areas where administrative data are most frequently used. Consequently, compatible models of metadata in these areas are required.

Many countries, however, experienced some obstacles in seeking synergy between statistics and administrative sources at the national level. Data quality, data comparability, legal aspects and data confidentiality were often the troublesome areas.

Changes in national administrative regulations can cause inconsistencies in time series. The concepts of variables in administrative sources can vary according to the statistical ones. Last but not least, the re-processing of administrative sources also implies the need to cooperate closely in the harmonization of administrative and statistical concept of data and metadata. The direct cooperation of statisticians and public program administrators on the harmonization of administrative and statistical concepts is therefore eminent.

Some examples for misalignment of statistical concepts with administrative requirements. Business surveys are typically focused on statistical enterprises and establishments, whereas government programs may apply to other units such as licenses, other regular entities, or corporations. Reconciling between different units can be quite troublesome.

Another example: data collected on business survey are usually aimed (in part or entirely) at the need of the SNA. However, the SNA concepts may not fit well with administrative needs of particular government program.

o facilitate the harmonization of administrative sources, the following grouping of administrative sources could be considered: enterprise internal register (bookkeeping); service bureau register (for accountant practices); public registers developed by local and regional government and the “quasi “ public register created by the institutions carrying out special administrative functions (such a social security registration).

**At the international level**, significant obstacles are different concepts of administration reflecting institutional differences between the countries, different coverage of national statistical surveys using the administrative sources and, quality of administrative data.

Important prerequisite to achieve a better international comparability at the European level would be the international harmonization of non-statistical concepts. In the social statistics, for example, the concepts of households, occupational categories and labor qualifications can be harmonized. For business statistics, the possibility of harmonizing statistical and legal unit should be explored. In the European framework, the degree of harmonization could vary, depending on the power and objectives of EU policies, the similarity of national realities, cost effectiveness and other factors.

Harmonization of national statistical outputs should be the basis for further improvement of statistical comparability at the national level when using national administrative sources. The history of SNA shows that the harmonization of outputs can be very successful. Harmonization of output concepts should concentrate on the standardization of final results from national statistical systems while the choice of sources and estimation methods should be left to the responsibility of NSIs.

### **3.2 Reluctance in using official statistics**

The exploitation of official statistics in government program administration may offer potential for large cost economies compared to the alternatives while, at the same time, reducing the overall compliance load on businesses. Nevertheless, one may ask why governments do not more frequently avail themselves of this kind of option.

One reason may simply be that government authorities perceive as insufficiently accurate to be used to administer certain programs. Statistical sampling, by its nature, allows a degree of imprecision. Statistical revisions may also feed a perception of inaccuracy, even if the final, revised statistics are in fact highly accurate. While in some instances such perception may be well founded, in others they can reflect misunderstandings. Sometimes, administrative data collected, processed and directly used within the government programs (instead of official statistics) may be significantly less accurate than their users understand to be the case.

The dependability, responsiveness and service quality provided by the NSI can be a key factor. If they are to adopt official statistics for use in administering their programs, government officials will need a feeling of reassurance that they can count on the continuing cooperation, advice and assistance of the NSI. They need to be ensured that the statistical programs in question will not be cancelled or greatly modified without full and open advance consultation. The administrators will also want good program documentation and attentive service. It can take years to develop the sort of confidence and trust required.

Another reason for reluctance could be that, the relevant statistical programs do not provide sufficient details for purposes of administrative task. For example, the government program may need information for local communities, while the available surveys have insufficient sample size to support statistics at that level. Even in cases where sample sizes are large, or where the statistician is able to produce exhaustive breakdowns by taking advantage of comprehensive administrative data from tax sources, confidentiality restrictions may prevent the statistical office from providing information at the desired level of detail.

More common cases are probably those where the program requirements arise long after the relevant statistical programs have been designed and implemented. There is less scope to adapt the requirements of the government program to what the statistical system can provide.

Timeliness can be another factor. Government programs often need supporting data in real time, whereas the collections, processing and release of official statistics sometimes takes years to complete.

Closely associated with timeliness is the revision issue. For some government programs, revisions are unacceptable. This can be a powerful factor inhibiting the use of official statistics for administrative purposes, especially if timeliness is another critical requirement.

Clearly, there are reasons for public program manager to think carefully before adopting statistics for administrative purposes. If they are to do job effectively, the statistics must be perceived as sufficiently accurate, timely and detailed. Underlying statistical concepts and definitions must line up well against those required for the program. The administrators must have confidence that their statistician colleagues will provide impartial, dependable and responsive service and will be unlikely to revise the numbers too frequently or to a very considerable extent. And, it may be necessary as well to resolve problems arising from the requirement to keep information about individual households or businesses strictly confidential.

#### **4. GEOGRAPHIC SEARCHING**

To use the geographic location as a stable key to integrate statistical data is better fitting to the modern mobile society where the importance of administrative borders is decreasing. To specify an administrative location is the only way in traditional technologies where the instability in the



definitions of administrative units complicates the comparability and integrity of data significantly.

Spatial dimension permits the integration of statistics from different sources (such as enterprises, other non-statistical providers, academia etc.) into the data collection phase. It enforces joint projects among intergovernmental institutions and has a significant upward impact on statistical analysis.

#### **4.1 Spatial data warehousing**

An essential component in geo-data warehousing is spatial database management. It includes metadata, spatial data quality and standard tools. Although rapid technological development could have a significant impact on spatial data warehousing in the future, the main obstacles are not technological ones. Instead, reaching agreements on standards, definitions, core data sets (models), disclosure rules, etc. are the significant elements missing for upgrading the role and function of geographical databases.

Metadata standards are basic recurring items for geo-data warehousing. While content oriented statistical metadata (e.g. definitions of statistical indicators, measurement units, classification and nomenclatures, etc.) are already implemented in many countries, metadata related to geographically referenced statistical information (e.g. related to boundaries) require further development. A clearing house concept for special metadata seems to be a promising approach.

#### **4.2 Geo-data access service for EU administration and data providers**

The European Union (together with Eurostat) developed under the 4<sup>th</sup> Research and Development Program a project GEOSERVE, which aims to develop a network of geo-data access services for EU administration and major data suppliers. The project is driven by users demands. It supposes to provide geo-data service nodes and value-added functions on Internet and Intranet.

Important European users are involved in the project: surveying institutes in Germany and Finland, Ministry of Environment of North Rhine Westfalen (Germany), the Italian Statistical Institute (ISTAT), the Decision System Limited of Dublin (Ireland) and the “ Institute Géographique National “ (France). All users contribute as providers of geo-services and geo-data. Major validation sites will be set up in four countries: Germany, Finland, Italy and Greece.

#### **4.3 Register based statistics**

Point based approach where statistical micro-data are linked the X/Y coordinates is implemented in the Nordic countries (Sweden, Finland, Norway, Denmark) where statistical systems are register-based. For example the Swedish statistical system consists of four registers covering demographic, labor market, economic and geographical dimension of statistics. The metainformation related to the geography in GIS module can thus be linked to a large number of registers in the system. Another example is the GIS application on land use for cities and urban settlements in Norway. The aim of this project is to establish a standard method for analysis of environment-friendly urban settlement, population, ground property utilization and transport. The project is based on the Norway's official register of ground properties, addresses and buildings, is integrated with the National Register of Taxes and database on all serviceable roads.

#### **4.4 Decentralized statistical systems**

In the UK, for example, the Office of National Statistics (ONS) is responsible for collecting and producing statistics for a wide range of demographic, social and economic topics across much of the UK. The desire to integrate statistical analysis from across these subject-areas and from outside agencies has led to the fundamental review of how the ONS data is referenced. The idea

behind this integration is to use coordinate references as the key method of referencing observed data. When a new event is surveyed, the location of this event will be recorded and allocated a grid reference (an X/Y coordinate in the British National Grid). This grid reference can be used in the allocation of the observation to an area and thus to the production of statistical analysis.

The US Federal Geographic Data Committee developed a Clearinghouse for spatial metadata to facilitate sharing of data collected and managed by US Federal Government activities. The clearinghouse technology has been widely employed at government, private and academic level making the project a success. The US National Spatial Data Infrastructure (including the Clearinghouse) was created already in 1994. The Clearinghouse as a central electronic catalog service hosts description of spatial data sets developed guidelines for spatial metadata and the US federal agencies and their statistical departments are required to post their metadata in a searchable Node of the Clearinghouse. In this way, geo-spatial data sets can be used for multiple purposes beyond their intended initial scope.

#### **4.5 Environmental-economic accounting**

The GIS application developed by Germany focuses on the inclusion of natural resources depicting the performance of national economy and users in an ecological area sampling procedure, to obtain information on all ecological systems known. Another sample survey has been conducted on flora and fauna using visual interpretation of satellite image data and subsequent digitalization. In this way a stock of data on land cover for the whole Germany has been compiled. The space and thematic data are derived from cartographic sources. Other official statistics were also considered for inclusion in this project, starting with agriculture statistics.

### **5. CONCLUDING REMARKS**

Economical, political and technological development will oblige statistical agencies to focus more and more on the needs of national and international users. Users needs are influencing not only technology and spectrum of statistical services. They are impacting primarily the content and methodology of the whole statistical information system.

Integration of statistical information system with users information systems is indispensable. Existing statistical concepts should be revised according to the users requirements. It may lead to the shift from traditionally survey oriented approach to the subject oriented concepts. Methodological definitions of statistical data and processes should be harmonized with those of users. Highly important will be a corresponding harmonization of statistical metainformation system where increasing role will have development and implementation of national and international standards. The solution of these tasks is very often of a research nature and, therefore, the involvement of research will be necessary.

Those processes will also impact the management and organization of statistical activities. An interdisciplinary team approach is the only possible solution. In such team external users, statistical methodologists, IT experts, researchers and all other parties involved should participate.

## References

1. Beedham, H.: *External Users Needs*, , UK Data Archive, University of Essex, AMRADS Workshop on Metadata, Edinburgh, May 2002
2. Beedham, H.: *Disseminating Data Constrains and Relevance*, UK Data Archive, University of Essex, AMRADS Workshop on Metadata, Ljubljana- Slovenia, March 2003
3. Klep, J.: *Country report by Statistical Office of the Republic of Slovenia*, , AMRADS Workshop on Metadata Edinburgh, May 2002
4. Meliskova, J.: *Some Thoughts on External Users of Statistical Metadata*, , AMRADS, Workshop on Statistical Metadata, Edinburgh, May 2002
5. Sindoni, G.: *Country report by Italian National Statistical Institute*, AMRADS Workshop on Metadata, Edinburgh, May 2002
6. *Final Reports on AMRADS Workshops on Metadata*, Edinburgh May 2002 and Ljubljana, March 2003
7. *Globalization – New Needs for Statistical Measurement*, ONS-UK (CES/2003/17), Conference of European Statisticians (CES), June 2003
8. *Impact of Globalization on Traditional Statistical System* (CES/2003/34), CES, June 2003
9. *OECD' Efforts to Measure Activities of Multinational Enterprises* (CES/2003/15), CES, June 2003
10. *Reports of the UN/ECE/Eurostat Work Sessions on Integration of Statistics and Geography*, Conference of European Statisticians (<http://www.unece.org/stats>)
11. *Reports of the UN/ECE/Eurostat/OECD Meetings on Statistical Information Technology*, Conference of European Statisticians (<http://www.unece.org/stats>)
12. *Reports of the UN/ECE/Eurostat/EECD Work Sessions on Statistical Metadata*, Conference of European Statisticians (<http://www.unece.org/stats>)
13. *Report on the progress in SDMX project*, Carol Carson, IMF, CES, June 2003