United Nations ECE/CES/2010/1/Add.1

Economic and Social Council

Distr.: General
19 May 2010

English only

Economic Commission for Europe

Conference of European Statisticians

Fifty-eighth plenary session
Paris, 8–10 June 2010
Item 3 of the provisional agenda
Coordination of international statistical work in the UNECE region

Statistical dissemination and communication: in-depth review by the Bureau of the Conference of European Statisticians

Note by the secretariat

Summary

The present note is a revised version of the document that was used as the basis for the in-depth review by the Bureau of the Conference of European Statisticians in October 2009. It has been updated to take into account the comments made in the discussion. The document has been prepared in consultation with the Steering Group on Statistical Dissemination and Communication.

I. Background

1. The Bureau of the Conference of European Statisticians regularly conducts in-depth reviews of selected statistical areas. The aim of the review is to give an overview of international statistical work in the specified area, identify problems (such as lack of activities, duplication of work, uncoordinated activities, etc.), and seek solutions to addressing the challenges.

2. In February 2008, the Bureau selected statistical dissemination, communication and publications as a topic for an in-depth review. Furthermore, the Bureau decided to experiment with a new approach: to discuss the topic at the plenary session of the Conference of European Statisticians prior to the review by the Bureau in order to obtain input from all members of the Conference.

II. Current activities in statistical dissemination and communication

3. There are several international groups that meet regularly to discuss issues relating to statistical dissemination and communication. This workincludes the development of guidelines and resources to assist statistical organizations in the effective management of dissemination and communication. Current groups operating in the region of the United Nations Economic Commission for Europe (UNECE) are listed below.

A. United Nations Economic Commission for Europe Work Session on Statistical Dissemination and Communication

4. The Steering Group on Statistical Dissemination and Communication was established in October 2004 (CES/BUR.2005/12/Add.4)[1]. The Bureau reviewed the work of the Group in October 2007 and supported the future work plans (CEC/CES/BUR/2007/OCT/4). The objectives of the Steering Group are:

(a) To facilitate the exchange of experience in statistical dissemination and communication;

(b) To develop a set of guidelines summarising good practices in national and international statistical organizations.

5. The focus is on subjects related to the organization of dissemination and communication, and not on technical questions dealt with by other groups (e.g. statistical metadata, output databases, interchange standards). Since its inception, the Group has produced three publications:

[1] The current membership of the Steering Group is: Gabrielle Beaudoin (Statistics Canada), Colleen Blessing (US Energy Information Administration), Eileen Capponi (OECD), Frances Comerford (Central Statistical Office, Ireland), Colleen Flannery (US Census Bureau), Heath Jeffries (Office for National Statistics, UK), Michael Levi (US Bureau of Labor Statistics), David Marder (Office for National Statistics, UK), Andrey Maslyanenko (Federal State Statistics Service, Russia), Kenneth C. Meyer (US Census Bureau), Terri Mitton (OECD), Leon Østergaard (Statistics Denmark) (Chair), Gina Pearson (US Energy Information Administration).

(a) *Communicating with the Media: A guide for statistical organizations*, published in 2004;

(b) *Making Data Meaningful Part 1: A guide to writing stories about numbers*, first published in English in 2006 and in Russian in 2009;

(c) *Making Data Meaningful Part 2: A guide to presenting statistics*, published in English in 2009 and in Russian in 2010.

6. These publications have a wide distribution and are used as a basis for training staff in statistical organizations. A new edition of the guide *Communicating with the Media* is planned to be published in 2010.

B. Eurostat

7. Eurostat organizes regular meetings of the Dissemination Working Group, bringing together experts from Member States to discuss issues related to statistical dissemination.

8. Eurostat's main role is to process and publish comparable statistical information at European level. However, Eurostat itself does not collect data. This is done in Member States by the responsible statistical authorities. National statistical organizations are the main - although not the only - partners.

9. Eurostat's role is to consolidate the data and ensure they are comparable, using harmonized methodology. Eurostat is the main provider of statistics at European Union level and the data it issues are harmonized as far as possible. Eurostat cooperates intensively with Member States, as well as with other international institutions, to define common statistical concepts and a common methodological basis for the statistics.

C. International Marketing and Output Database Conference

10. The International Marketing and Output Database Conference (IMAODBC) is an annual conference coordinated by a small, independent network of national and international statistical organizations (representing approximately 25 organizations). The responsibility for hosting and organizing the meeting is rotated amongst the group.

11. The potential for overlap between IMAODBC and the UNECE Work Session on Statistical Dissemination and Communication was raised in discussion at the September 2006 UNECE meeting. This issue was investigated through discussions with participants attending both meetings in 2006. Findings indicate that in their current form, the two groups are sufficiently different and address different needs. The IMAODBC provides a less formal environment and focuses on dissemination issues that are more technical in nature, in particular the relationship between technical solutions and marketing and dissemination strategies. UNECE activities concentrate on the management and organization of dissemination and communication and, unlike IMAODBC, the Group aims to produce guidelines to assist statistical organizations in this work.

12. Despite the differences, the two groups will need to remain conscious of the potential for overlap and avoid any duplication of effort. Current strategies to address this include:

(a) The aim to have some commonality between members of the UNECE steering group and the IMAODBC organizing committee;

(b) The participation of UNECE in the annual IMAODBC meeting when possible, providing updates on its activities in this area;

(c) The scheduling of UNECE meetings to avoid being close in time to the IMAODBC, so the agenda can build on issues previously discussed.

D. Organisation for Economic Co-operation and Development Turning Statistics into Knowledge seminars

13. The Organisation for Economic Co-operation and Development (OECD) has organized several seminars on this theme, including in Washington in July 2009, Stockholm in 2008 and Rome in 2007, as well as an international exhibition on "Innovative tools to transform information into knowledge" held during the second OECD World Forum on "Statistics, Knowledge and Policy" in Istanbul in 2007.

14. The purpose of these seminars is to contribute to the development of tools to help people transform statistics into knowledge and decisions. A first condition for statistics to be used this way is that relevant statistics become known and understood by wider audiences. They focus on a broad range of tools, particularly dynamic graphics. They also cover the use of videos, as explored by GapMinder and others, and participative approaches, as seen in some Web 2.0 initiatives. Although innovative tools are themselves of great interest, the focus of the seminars is on innovative applications of tools, e.g. so-called story-telling applications. Experts in statistical methodology, cognitive science and communication are therefore welcomed as active participants.

15. In parallel, there is increasing interest among scholarly publishing and research librarian communities in data discoverability and data management. International groups with a particular interest in publishing and managing statistical and other data include:

(a) International Council for Scientific and Technical Information (ICSTI): a forum for interaction between organizations that create, disseminate and use scientific and technical information;

(b) International Association for Social Science Information Service and Technology (IASSIST): an international organization of professionals working in and with information technology and data services to support research and teaching in the social sciences;

(c) Government Documents Round Table of the American Library Association (GODORT), which promotes library service and librarianship;

(d) German National Library of Science and Technology (TIB), which is organising a consortium of European technical libraries to establish a joint data registration agency.

III. Issues and challenges relating to the dissemination and communication of statistics

A. Managing communication

16. Communication of statistics is about a dialogue between users and producers of statistics. This differs from dissemination, which is about spreading statistical data by making them available through publications, output databases and other means. Communication activities include:

(a) Developing and implementing a communication and publishing strategy;

(b) Monitoring reputation;

(c) Monitoring and servicing the media;

(d) Preparing news releases;

(e) Maintaining website(s);

(f) Marketing publications;

(g) Building relationships with stakeholders and users;

(h) Conducting user satisfaction surveys;

(i) Implementing design and/or style standards;

(j) Educating colleagues in effective communication;

(k) Editing and improving language;

(l) Maintaining the intranet and other channels of internal communication.

17. These activities may be the responsibility of one organizational unit but may also be shared between two or more organizational units within national and international statistical organizations.

18. Existing approaches to managing communication activities within statistical organizations may be the result of conscious decision-making, but research suggests that, in many cases, it is more likely a result of the historical development of statistical organizations. Statistical communication has evolved considerably over recent years. Traditionally, statistical organizations focused on dissemination and one-way communication of information through few media channels (e.g. newspapers, radio and television). It was not until the 1990s that they acknowledged the need to do more than just disseminate data and hesitantly began employing communication professionals. Widespread use of the Internet has significantly changed methods of communication and dissemination, as well as increased numbers and diversity of end-users.

19. There are advantages and disadvantages in centralizing communication functions into a single unit within a statistical organization. Communication is a specialized field and establishing a single communication unit makes it possible to attract and hire professionals and specialists who can work together, combining their expertise to increase benefits for the organization. Their effectiveness will depend upon the level of responsibility and influence such a unit is given in the management structure. Furthermore, centralizing communication functions can create a sizeable team, increasing the possibilities for running important but not necessarily urgent development projects, which may not otherwise receive attention in a devolved system due to lack of resources and higher priorities.

20. A disadvantage of a centralized unit is that the team may be quite large and the activities mainly operational in nature. As top management is only interested in the strategic aspects of communication, there is a tendency to place a centralized communication unit at the same level as other operational units. In this case, the unit may not report directly to top management, hindering its ability to act strategically. Different solutions have been sought to address this problem. It is also necessary to ensure effective cooperation between communication experts and statisticians (who may take a more conservative view on how to communicate their data).

21. Discussions at the 2008 and 2009 UNECE Work Sessions on Statistical Dissemination and Communication revealed that the statistical systems of individual countries have a range of approaches when it comes to managing communication functions. The placement of these functions within the organizational structure will impact on their effectiveness. Governance processes should ensure that communication of statistical data receives the same level of attention from management as dissemination activities and the

associated technological infrastructure. More work is needed on measuring the impact of different communication practices.

22. User feedback suggests that improving communication should be a priority for many statistical organizations, particularly in the current financial crisis. More effort is needed to reach mainstream users, such as media and policy makers, not just hard-to-reach groups. Increasing competition from other data producers means that there is an increasing need to professionalise communication of statistical organizations.

B. Ethics and independence of statistical organizations

23. Ethics and independence are fundamental issues for statistical organizations. This is reflected in the Fundamental Principles of Official Statistics and the European Statistics Code of Practice. Statistical organizations face challenges to develop dissemination and communication strategies that reflect their independence, such as disseminating data on minority groups and policies on pre-release embargoes and access to information by the media.

24. For example, data on ethnic minorities or specific regions may be needed to meet policy demands, but carry a risk of giving or reinforcing negative perceptions of these groups or regions. A possible solution could be to avoid these issues by not identifying these groups, such as by not collecting sensitive variables or by disseminating only broad aggregates. A more balanced approach would be to consider sensitive aggregates in terms of the normal confidentiality and quality criteria, fine-tune them to eliminate structural effects and add an impartial commentary. According to the fourth fundamental principle of official statistics, statistical organizations should react to and try to correct any erroneous interpretation of data. This requires systematic monitoring of the media and a policy on how and when to react. Further discussion on this point is needed, and exchanging experiences and collecting good practices would be useful.

25. Good communication practices in areas such as pre-release access, the use of release calendars, and dealing with erroneous use and misunderstandings are vital to improve the credibility and independence of statistical organizations. It is important, on the one hand, to minimize errors by disseminating statistics of a high quality and, on the other hand, to develop policies and procedures for reacting when errors are discovered. Experience in the UNECE region has shown that, in order to maintain credibility, it is important to be transparent and to communicate loudly and clearly to the public when errors occur. The importance of data quality is increasingly stressed in staff training, quality checklists and error reporting.

26. There are some differences amongst CES member countries in their policies on pre-release embargoes, with some national statistical organizations banning all pre-release access and others supporting controlled pre-release as a mechanism to help the media or policy makers prepare more effective communication for the public. The differences may be related to different stages in the development of relationships with user groups. When pre-release access is provided, it is important to inform the public on the rules and procedures (who can get access, why, how early in advance, etc.) in a clear and transparent way.

27. It is not easy to ensure that all organizations producing official statistics follow the Fundamental Principles and strict pre-release policies. The Fundamental Principles prescribe that official statistics should be made available on an impartial basis, but this leaves room for different country practices. Providing pre-release access to certain media may result in more interest in statistics from these media in the short term. In the long term, however, it may be more useful to focus on improving communication and providing a

clear message simultaneously to all users. The statistical community would benefit from an exchange of experiences in pre-release of data and the collection of good practices. It would also be important to ensure that good practices are followed across the whole statistical system.

C. Emerging tools for data visualization and communication of statistics

28. Emerging tools and techniques are providing new opportunities for visualizing data. Many national statistical organizations now offer access to statistical databases through their website, allowing users to query and download statistical information. This is often complemented by a suite of visualization tools enabling users to create tables, charts or maps online, without having to download the data and work in another application. There may be concerns about the consequences of giving this level of control to the user, who could produce nonsensical charts or maps, or inappropriate correlations. To minimize potential problems, it is important to provide critical metadata in a clear and obvious way, offer support and monitor any misuse. The impact and value for money of these new visualization tools should also be better assessed.

29. Further work is needed to make data more understandable and easier to find and access by taking advantage of the new developments in web search technology. The use of Statistical Data and Metadata eXchange (SDMX) and other metadata standards is important in this respect.

30. New web technologies are changing the way statistical organizations communicate with their users. Blogs, wikis and social networks provide new communication channels that have the potential to reach a wider global audience. Other kinds of interactive websites have emerged in the recent past, which allow users to upload data and create graphs for sharing and discussion with other users. These and other websites that permit users to comment or post data and related information are adding a new dimension to statistical data dissemination and communication, increasing the number of intermediary providers of statistics. Although the impact and success of these tools has been limited to date, online communities and user-generated content are emerging areas of statistical dissemination and communication in which the international statistical community should be actively engaged, at least in the monitoring of the impact on statistical organizations. The use of more interactive technologies also increases the risk of attacks by special interest groups on ideological rather than scientific grounds, so would require careful management and monitoring. International exchange of experiences and good practices in this field will become increasingly useful.

31. Effective presentation of data is increasingly being considered an integral part of the statistical production process. As a result, there is growing demand for guidelines to help managers, statisticians and media relations officers create tables, charts and other presentations that bring statistics to life. In response to this demand, a series of guides on *Making Data Meaningful* is being developed within the framework of the UNECE Work Session on Statistical Dissemination and Communication:

(a) *Part 1: A Guide to Writing Stories about Numbers* has been extremely popular since it was first published in 2006;

(b) *Part 2: A Guide to Presenting Statistics* was published in 2009. It is a 50-page publication covering the following topics:

(i) Getting the message across;

(ii) Overview of the visualization of statistics;

(iii) Guidance on the presentation of tables, charts and maps (main part of the guide with a chapter on each);

(iv) Emerging visualization techniques;

(v) Accessibility issues;

(c) *Part 3* will be an updated edition of the guide *Communicating with the media*, which was first published in 2004.

32. These guides provide a useful reference for the wider statistical community. Although approaches vary from country to country due to practical and cultural differences, international guidelines provide a valuable complement to individual organization standards. Associated training materials could help statistical organizations in the implementation of these guidelines.

33. Statistical organizations are gradually moving away from printed publications, but it may be useful to maintain some paper outputs. Short, cross-cutting publications, with more emphasis on graphics, are becoming increasingly popular with users. In some countries, combining data with economic analyses has helped to make statistics more attractive and popular. Exchanges of experiences help to spread good practices in this respect.

D. Communicating with hard-to-reach groups

34. This issue was discussed at the May 2009 UNECE Work Session on the Communication and Dissemination of Statistics. Statistical organizations are faced with the challenge of reaching many target audiences for a variety of purposes, such as conducting population or business censuses and surveys of all kinds. Without a doubt, each country has segments of their population with whom it is difficult to communicate and secure active participation.

35. Solutions on communicating with hard-to-reach groups range from technological to educational. Rapid advances in technology provide potential solutions (e.g. blogs, podcasts and social networking) for communicating with some audiences. Outreach to hard-to-reach groups may include:

(a) Educating children to reach parents;

(b) Working with minority media;

(c) Providing multiple language information and interpretation;

(d) Partnering with local, specialized groups to ensure cultural sensitivity and to encourage understanding and participation by that community.

36. It should be possible to identify good practices and examples by continuing to share experiences between statistical organizations through regular meetings on dissemination and communication.

E. Education programs for improving statistical literacy

37. Improving statistical literacy relies on identifying user groups and developing strategies to increase users' understanding of statistical concepts and how to use statistical information, and to enhance knowledge within the statistical organizations of the issues faced by users in understanding and using statistics. Statistical organizations need to improve their explanations, provide definitions for specialists, as well as for everyday people, and enhance public education in order to raise confidence in official statistics.

38. Important issues for statistical organizations are management of limited resources and increasing pressures on staff. Statisticians not only have to produce data, but also need to make them appealing. Staff employed for statistical expertise may not have the skills to write like journalists and it may be unrealistic to expect them to. The shift from an audience of experts to widespread general interest in statistics has contributed to increasing pressures on organizations to address different levels of statistical literacy.

39. Strategies and issues about educating the public and training staff within statistical organizations include:

(a) Bringing expert journalists into the office to explain to staff what the media need from them;

(b) Organizing meetings with journalists to enhance their understanding of statistics, in order to avoid misinterpretation in the media;

(c) Partnering with educators to develop and co-produce materials for their students to use in courses of study, in order to help young people learn the skills necessary to understand and use statistics;

(d) Creating web resources;

(e) Allowing users to customize their web page views and functions;

(f) Communicating through radio or television programs;

(g) Interacting in a systematic way with the education system;

(h) Recruiting people with an educational background to help develop education programs;

(i) Investing time in outreach and building relationships with users, focusing on groups that can provide access to others (e.g. teachers associations, academicians and the media).

40. At the 2006 UNECE Work Session on Statistical Dissemination and Communication, it was recommended that the international statistical community establish activities to improve statistical literacy. There is a need to move forward from discussion of the issue to concrete action.

F. Pricing and licensing issues

41. The approach to providing statistical information free of charge or at a cost was discussed at the 2009 UNECE Work Session on the Communication and Dissemination of Statistics. Reports from participating organizations revealed that many offer a combination of chargeable and free-of-charge information. Usually, the statistical material which is directly retrievable on the Web is non-chargeable, whereas printed publications, customized tabulations and services, microdata and sensitive variables are usually chargeable services. Charges for customized services allow organizations to control the number of requests and manage the subsequent impact on resources. Some organizations avoid offering tailored services due to a lack of resources, but may use or be exploring existing technologies for establishing self-help systems so users can extract their own customized tabulations or microdata files.

42. A survey of the European Statistical System (ESS) has revealed a high degree of convergence in dissemination practices across national statistical organizations. Eurostat is developing a ‘Dissemination Policy Charter’, which includes a proposal for seven principles relating to access to statistics and microdata, privacy, release, use and

transmission. These principles will provide useful guidelines, but national legislation may prevent countries from applying them in full.

43. The scope and frequency of demands for statistical information have increased considerably since data have become more accessible through the Internet. At the 2008 Work Session on Statistical Dissemination and Communication, Professor Hans Rosling (Gapminder Foundation / Karolinska Institute, Sweden) gave a keynote presentation on 'Communicating Statistics in the Information Age'. He noted that a variety of different terms and conditions, often restrictive, are applied to statistical data by different organizations. Adopting a common access license for official statistics may facilitate response to this demand by providing a consistent approach across countries. There are a number of licensing issues to be resolved, such as standardising references to the original source, protection of the original data and metadata, conditions for re-use and sharing of data and metadata.

44. The recent experience of the Australian Bureau of Statistics (ABS) in adopting Creative Commons, an internationally recognized licensing framework, provides an example for the statistical community. The ABS developed this new approach to remove unnecessary barriers for users wishing to re-use large amounts of their data. Furthermore, the advent of new Web 2.0 technologies, which allow innovative re-use and sharing of official statistics, has necessitated action. In mid-2008, the ABS decided to adopt Creative Commons, a "simple, open and internationally recognized licensing framework for its statistical information" for content on its website (ABS, UNECE Work Session on Statistical Dissemination and Communication, 13-15 May 2009). According to its website (http://creativecommons.org/), Creative Commons is:

> *"[…] a non-profit corporation dedicated to making it easier for people to share and build upon the work of others, consistent with the rules of copyright. Creative Commons provides free licences and other legal tools to mark creative work with the freedom the creator wants it to carry, so others share, remix, use commercially, or any combination thereof […]."*

45. There are six types of *Creative Commons* licenses that can be attributed to content to specify different levels of allowable reuse, sharing and adaptation:

(a) Attribution Non-commercial No Derivatives;

(b) Attribution Non-commercial Share Alike;

(c) Attribution Non-commercial;

(d) Attribution No Derivatives;

(e) Attribution Share Alike; and

(f) Attribution.

46. A full description of the terms and conditions under each license is available from the Creative Commons website. The Attribution license, "allowing users to remix and build upon creator's work, even commercially, as long as they credit the creator for the original creation", may be best suited to meet the demands of the statistics user community for free and open access to statistical data. However, the different levels of Creative Commons licensing would accommodate more stringent controls where necessary. It should be noted that some international organizations may have difficulty in applying Creative Commons licensing, as they are not subject to individual country legislation.

47. Adopting this internationally recognized licensing framework for official statistics could provide a consistent and instantly recognizable method for managing the use of statistical information. This should improve the visibility and potential re-use of statistical

information, increasing discoverability through search engines and reducing the resources needed to maintain license administration systems within the statistical organization.

G. International data dissemination

48. The collection and dissemination of official statistics by international organizations have a significant impact on national statistical systems. Where possible, international organizations source data directly from the dissemination systems of national and supranational statistical organizations, but until electronic data and metadata exchange standards are widely implemented, national statistical systems have to manage the burden of manually providing data for international reporting.

49. Producing statistics at the international level reveals a range of complexities including:

(a) Limited comparability due to different methods or lack of international standards;

(b) Data quality problems, such as lack of data availability for particular indicators;

(c) The identification and management of discrepancies between national and international data.

50. Initiatives that aim to integrate statistics published by international organizations into a single database or portal raise further complexities, as differences and discrepancies become more apparent and users are challenged to correctly understand and interpret similar data from different sources. In the longer term, however, integration of data from different sources can help to improve harmonization by exposing unnecessary differences and increasing pressure to resolve them.

51. An example of these issues is evident in the collection and dissemination of data to monitor progress made towards the Millennium Development Goals (MDGs). Data gaps and discrepancies between national and international datasets were present before MDGs and are likely to continue into the future because, in some instances, national and international data respond to different information needs. However, a substantial reduction of discrepancies is possible if action focuses on areas where improvements are most likely to be made.

52. International reporting systems are mainly based on pre-existing data compilation and data production activities of national and international organizations. Coordination between these organizations is still insufficient for a variety of reasons, including:

(a) Lack of communication between different national data producers;

(b) Differing national priorities (for example, some indicators are not regularly produced by national statistical authorities since they are not relevant in the national context);

(c) The data collection systems of international organizations are not always able to capture all data available in countries;

(d) Lack of synchronization between dissemination of data by national statistical organizations and collection activities by international organizations.

H. Improving discoverability and linking data into the scholarly network

53. Datasets are a significant part of the scholarly catalogue and are being published with increasing frequency, either formally or informally. Many publishers are beginning to link to datasets from their journals, and authors are trying to cite them in their articles. Librarians would like a way to manage datasets alongside other publications. Authors cite data in a variety of ways, often to the organization name, organizational, departmental websites or broken links. Authors and publishers are clearly unsure about how they should cite data sources. Librarians are also making attempts to catalogue datasets yet often do not manage to lead the users directly to the datasets.

54. There is no accepted system for how datasets should be cited and catalogued. Now that datasets are becoming more widely available and so many publishers are getting involved, there is a need for a bibliographic system to help authors cite, publishers manage and librarians catalogue datasets.

55. A recent study by Inger & Gardner[2] on how readers navigate to scholarly content indicate that specialist bibliographic search engines, library web pages and journal homepages are used more frequently than general Web search engines when a reader already has a reference or citation and wishes to read the article online. Clearly, if data providers rely on a post-it-and-Google-will-find-it approach, they will miss out on a great deal of traffic from readers who are using alternative routes to discover content.

56. Many datasets are being updated on a rolling basis, adding new data as and when received. Occasionally, revisions are made to the entire dataset which changes the historical data. All of these changes are noted and explained in the statistical metadata. A citation, however, is supposed to link the reader back to the same object which the citing author used. In the case of a dynamic dataset, linking back to a dataset as it was when an author used it to write a paper is clearly impossible. This poses a significant challenge.

57. OECD has written a white paper proposing some standards for citing and bibliographic management of datasets and data tables. There is increasing interest among scholarly publishing and research librarian communities in data discoverability and data management.

[2] Inger S., Gardner T, How Readers Navigate to Scholarly Content, 2008 http://www.sic.ox14.com/howreadersnavigatetoscholarlycontent.pdf