



**Economic and Social
Council**

Distr.
GENERAL

ECE/CES/2009/7
30 March 2009

Original: ENGLISH

ECONOMIC COMMISSION FOR EUROPE

STATISTICAL COMMISSION

CONFERENCE OF EUROPEAN STATISTICIANS

Fifty-seventh plenary session
Geneva, 8-10 June 2009
Item 5(b) of the provisional agenda

COORDINATION OF INTERNATIONAL STATISTICAL WORK IN THE REGION OF THE
UNITED NATIONS ECONOMIC COMMISSION FOR EUROPE

IN-DEPTH REVIEW OF STATISTICAL DISSEMINATION, COMMUNICATION AND
PUBLICATIONS

**ISSUES RELATING TO THE DISSEMINATION AND COMMUNICATION OF
STATISTICS: DISCUSSION AT THE CONFERENCE OF EUROPEAN
STATISTICIANS (CES) 2009 PLENARY SESSION AS INPUT TO AN IN-DEPTH
REVIEW**

Note by the secretariat

Summary

The Bureau of the Conference of European Statisticians selected at its February 2008 meeting statistical dissemination, communication and publications as a topic for in-depth review (ECE/CES/2008/2). Furthermore, the Bureau decided it would be useful to discuss the topics selected for in-depth review by the Bureau at a Conference's plenary session prior to discussion at Bureau meetings to get input from the members of the Conference. This approach will be tested at the 2009 plenary session with the topic on statistical dissemination, communication and publications. In February 2009, the Bureau approved the outline of the note on statistical dissemination and communication and requested the secretariat of the United Nations Economic Commission for Europe together with the Steering Group on Statistical Dissemination and Communication to prepare the note for the 2009 plenary session (ECE/CES/2009/2). Following the discussion at the Conference, the topic will be reviewed in-depth by the Bureau in October 2009.

GE.09-

I. BACKGROUND

1. The Bureau of the Conference of European Statisticians (CES) conducts an in-depth review of up to five statistical subject areas each year. These reviews include an analysis of related issues, consideration of existing standards and the need for further development in the area. Consideration is also given to the activities currently being conducted by international organizations in the area, in order to identify potential overlaps and gaps.
2. In the past, these in-depth reviews have only involved Bureau members and observers. In October 2008, the Bureau decided that reviews would be enhanced with input from other CES members and that this input would best be gathered through a discussion of the review topic at the annual CES plenary session. The Bureau decided to test the approach at the 2009 session with the topic of statistical dissemination and communication.

II. AIMS OF THE DISCUSSION AT THE CONFERENCE OF EUROPEAN STATISTICIANS PLENARY SESSION

3. The discussion seeks input from all CES members on issues related to statistical dissemination and communication. CES members are requested to identify challenges for statistical organizations in the area of dissemination and communication of statistics and to determine what the international statistical community and international organizations can do to assist national statistical organizations. The focus should be on issues that are not yet - or not sufficiently - addressed at the international level and on identifying possible areas of duplication of work.
4. The outcomes of the discussion will serve as input to the in-depth review, to be conducted by the CES Bureau in October 2009, which aims to find solutions to address the issues raised.
5. Discussion at the CES meeting in June will be based on this issue paper, prepared by the secretariat of the United Nations Economic Commission for Europe (UNECE) in consultation with the Steering Group on Statistical Dissemination and Communication and experts in this field.

III. CURRENT ACTIVITIES IN STATISTICAL DISSEMINATION AND COMMUNICATION

6. There are several international groups that meet regularly to discuss issues relating to statistical dissemination and communication. This includes the development of guidelines and resources to assist statistical organizations in the effective management of this work. Current groups operating in the UNECE region include:

(a) UNECE Work Session on Statistical Dissemination and Communication: the Steering Group on Statistical Dissemination and Communication was established in October 2004 (CES/BUR.2005/12/Add.4)¹. The objectives are to:

- (i) Facilitate the exchange of experience in statistical dissemination and communication;
- (ii) Develop a set of guidelines summarising good practices in national and international statistical organizations.

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 two publications:

- (iii) *Making Data Meaningful: A guide to writing stories about numbers*, published in print and online in January 2006;
- (iv) *Communicating with the Media: a guide for statistical organizations*, published in 2004.

These publications have been widely distributed and used as a basis for training staff in statistical organizations. A third guide on the presentation of statistics will be published in mid-2009;

(b) Eurostat organizes regular meetings of the Dissemination Working Group, bringing together experts from Member States to discuss issues related to statistical dissemination. 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. 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 intensively cooperates 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 (IMAODBC): 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.

The potential for overlap between IMAODBC and the UNECE Work Session on Statistical

¹ The current membership of the steering group is: US Energy Information Administration; Central Statistical Office, Ireland; US Census Bureau; US Bureau of Labor Statistics; Office for National Statistics, United Kingdom; Federal State Statistics Service, Russia; US Census Bureau; Central Bureau of Statistics, Israel; Statistics Denmark (Chair); Statistics Canada; and Eurostat.

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 agencies in this work.

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:

- (i) Aim to have some commonality between members of the UNECE steering group and the IMAODBC organizing committee;
- (ii) UNECE participates in the annual IMAODBC meeting when possible, providing updates on its activities in this area;
- (iii) UNECE meetings are scheduled to avoid being close in time to the IMAODBC, so the agenda can build on issues previously discussed;
- (d) OECD Turning Statistics into Knowledge meetings: a seminar jointly arranged by OECD and Statistics Sweden on Innovative Approaches to Turning Statistics into Knowledge was held in Stockholm (Sweden) in 2008. This seminar was a continuation of the seminar in Rome on Dynamic Graphics held in 2007 and of the first International Exhibition on “Innovative tools to transform information into knowledge”, organized during the second OECD World Forum on “Statistics, Knowledge and Policy” held in Istanbul in 2007.

7. The purpose of these seminars was 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, available and understood by wider audiences. The Stockholm seminar focused on a broader range of tools than merely dynamic graphics, while these remained at the heart of the matter. It also included 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 was on innovative applications of tools, e.g. so called story-telling applications. Experts in statistical methodology, cognitive science, and communication were therefore welcomed as active participants.

8. 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 are organising a consortium of European technical libraries to establish a joint data registration agency.

IV. ISSUES AND CHALLENGES RELATING TO THE DISSEMINATION AND COMMUNICATION OF STATISTICS

A. Managing communication

9. 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.

10. These activities are often shared between two or more organizational units within national and international statistical organizations.

11. 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.

12. 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 get attention in a devolved system due to lack of resources and higher priorities.

13. A disadvantage of a centralized unit is that the team may be quite large and the activities mainly operational in nature. As top management are 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. 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).

14. Discussions at the 2008 UNECE Work Session 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.

B. Ethics and independence of statistical organizations

15. 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.

16. For example, data on ethnic minorities or specific regions 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, e.g. 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 statistics, statistical agencies 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.

17. The communication of quality and credibility is a key issue for 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 communicate loud and clear to the public when errors occur. The importance of data quality is increasingly stressed in staff training, quality checklists and error reporting.

18. 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. The statistical community could benefit from guidelines on pre-release access and examples of good practice.

C. Emerging tools for data visualization and communication of statistics

19. Emerging tools and techniques are providing new opportunities for visualizing data. Many national statistical organizations now provide 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 graphs, tables 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 graphs or inappropriate correlations. Potential problems can, however, be minimized by providing critical metadata in a clear and obvious way, offering support and monitoring any misuse.

20. New Web technologies are changing the way statistical organizations are communicating 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.

21. 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, graphs and other presentations that bring statistics to life. In response to this demand, a series of guides on *Making Data Meaningful* has been developed within the framework of the UNECE Work Session on Statistical Dissemination and Communication. Published in 2006, the first guide, *A Guide to Writing Stories about Numbers*, was – and still is – extremely popular. A second guide, *Style Guide on the Presentation of Statistics*, is currently being finalized for publication in the last half of 2009. It is a 40-page publication covering the following topics:

- (a) Getting the message across;
- (b) Overview of the visualization of statistics;
- (c) Guidance on the presentation of tables, graphs and maps (main part of the guide with a chapter on each);
- (d) Emerging visualization techniques;
- (e) Online accessibility issues.

22. 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.

D. Communicating with hard-to-reach groups

23. This issue will be discussed at the May 2009 UNECE Work Session on Statistical Dissemination and Communication. National 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.

24. 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.

25. 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

26. 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.

27. 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.

28. 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).

29. 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 actions.

F. Developing an international access license for statistics

30. The scope and frequency of demands for statistical information have increased considerably since data has 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 and argued that data should be freely available as a public good. Rosling suggested that the most important task for statistical organizations is to make data more accessible by providing them directly to end-users or through intermediaries such as those developing advanced visualization tools. 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.

31. 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....”

32. 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.

33. 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.

34. 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 reuse 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

35. 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.

36. Producing statistics at the international level reveals a range of complexities including: limited comparability due to different methods or lack of international standards; data quality problems, such as lack of data availability for particular indicators; identification and management of discrepancies between national and international data. 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.

37. 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, substantial reduction of discrepancies is possible if action focuses on areas where improvements are most likely to be made.

38. The international reporting system for MDG indicators is mainly based on pre-existing data compilation and data production activities of international agencies. Coordination between national and international 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 agencies 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 agencies.

39. MDG data dissemination is coordinated through the Inter-Agency Expert Group (IAEG) on MDG Indicators, for which the United Nations Statistics Division (UNSD) acts as secretary. The UNECE participates in the IAEG meetings and works closely with other United Nations Regional Commissions and the United Nations Statistical Division to address the issues outlined above.

H. Improving discoverability and linking data into the scholarly network

40. Datasets are a significant part of the scholarly catalogue and are being published more and more frequently, either formally or informally. Many publishers are beginning to link to them from their journals and authors are trying to cite them in their articles. Librarians would like a way to manage them alongside other publications. Authors cite data in a variety of ways, often to the organisation name, organisational, 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.

41. 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.

42. A recent study by Inger & Gardner² 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.

43. 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.

44. 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>