

Tunis, 11-12 September 2008

Items for discussion and decision: Item 3 of the Provisional Agenda

Modalities of Data Sharing among International Organizations: Proposal for a CCSA Code on Data Sharing

Note prepared by ECE¹

I Introduction and General Remarks

1. As a result of a first paper on this issue prepared by ECE that was discussed at the CCSA meeting in Madrid in September 2007, ECE was mandated, in cooperation with ECB and the World Bank, to outline good practices for the CCSA meeting in September 2008.
2. There are two basic approaches that could be followed to facilitate the access to, exchange and sharing between members of the CCSA of aggregate, non-confidential statistical data that are publicly released. The first is to list a number of general principles that, once adopted, would be implemented by each member individually. The second approach is to agree on a code of good practice, and in addition to list those datasets from CCSA members that are accessible under the conditions of the code by all other CCSA members. Because not all members are able to make available their full range of data products under the proposed conditions, the second approach is developed further and proposed here to allow an in-depth discussion. It is difficult to see what the first approach could add over and above the provisions and good practices that were already included on this topic in the *Principles Governing International Statistical Activities*², adopted by the CCSA in 2005.
3. Each CCSA member is free to select those statistical data sets publicly available in electronic form that shall be put on the *List of datasets subject to the CCSA Code on Data Sharing*, and also to withdraw from the list datasets that are already listed. As an example: in the case of one single comprehensive on-line database, the member is free either to list this database as a whole or only well identified specific datasets. At a first stage it is assumed that each member is able to put on the list at least one dataset.
4. The code is based on the following principles:
 - a. As a general principle there are no payments between members of the CCSA for exchanging publicly released data³. The limitations to the use and redistribution of datasets on the list are described below.
 - b. The use, including re-dissemination in electronic form, is as free from conditions as possible and subject only to a clear acknowledgement of the source, and reference to the essential metadata attached to the original data. To re-sell data obtained from

¹ ECE acknowledges the comments by ECB, which are reflected to a large extent in this draft, as well as the general comments by the World Bank with respect to an earlier draft by ECE that was circulated in July

² http://unstats.un.org/unsd/methods/statorg/Principles_stat_activities/principles_stat_activities.htm

³ In exceptional cases, for example when repackaging is necessary, additional costs could be covered by the recipient.

another CCSA member or to include them in a commercial product, the CCSA reseller must obtain explicit consent from the original provider.

5. With the exception of the requirement that adjustments of data received from another organization have to be made transparent, there is no attempt to introduce, as part of this code, any requirements of quality in terms of validation, imputation of missing values, or content of metadata to be associated to the statistical products. Such issues should be left to other guidelines, because their scope goes far beyond data exchange between international organizations. Since most of the members of the CCSA are bound by the Principles Governing International Statistical Activities, they will be guided by the relevant principles contained there, and possible further separate specifications and guidelines that may be developed and adopted by the CCSA in future.
6. There is a clear relationship between the code proposed here, and the SDMX code adopted by the CCSA as standard in February 2007. For this reason, data exchange through the present code encourages once more the use of the various relevant elements of the SDMX family of standards by producer organizations when they share data with other CCSA members.
7. The present code does not aim at replacing bilateral or multilateral agreements on data sharing or reporting between international organizations concerning the collection and validation of national data (e.g. through joint questionnaires).
8. The draft is inspired by the *Creative Commons Legal Codes*⁴, and especially the variant called *Attribution-NonCommercial*. This set of licenses was presented at the recent OECD/Statistics Sweden seminar on Innovative Approaches to Turning Statistics into Knowledge (26/27 May 2008). However, the draft presented below is much shorter and simpler, deliberately avoiding legal jargon and precise definitions, so as to allow for flexibility and learning by doing. The proposed code is intended as “gentlemen’s agreement”, applicable to selected datasets, and in no way as legally binding contract.

II The Draft CCSA Code on Data Sharing

Non-confidential statistical datasets in electronic form produced by members of CCSA (producer organizations) are provided free of charge to other CCSA members (user organizations), for use under their own responsibility in accordance with the following conditions:

- (a) The user organization can, in principle, re-disseminate in any format as statistical data under its own name, the whole or part of a dataset contained in the *List of Datasets subject to the CCSA Code on Data Sharing (List)*, including in combination with data from other sources, provided that the producer organization is clearly acknowledged, the essential metadata provided by the producer organization are directly associated, and other metadata are provided through appropriate references.
- (b) The user organization can use its own methods to adapt data contained in any dataset on the *List*, provided the methods of adaptation are made transparent to users, notably when disseminated to third parties.
- (c) User organizations are strongly encouraged to include their statistical datasets in the *List*, particularly if they contain information that is derived from other products on the *List* or are otherwise obtained free of charge.

⁴ <http://creativecommons.org/about>

- (d) Datasets contained in the *List*, or statistical data based on datasets on the *List*, should not be disseminated for commercial purposes by CCSA members unless explicitly authorized by the original producer organization.
- (e) The user organization can use all or part of a statistical dataset on the *List* for any non-statistical product, or any statistical product in other than electronic form, produced and where applicable disseminated under its own name, if conditions (a) and (b) are observed.
- (f) Besides other dissemination formats that may be used, participating organizations are encouraged to make the maximum possible use of the SDMX standards. The use of these standards facilitates efficient processing (application to application), promotes conceptual discipline and best practices in defining and organizing data and metadata, and facilitates international harmonization and efficient use of resources⁵.

III Organizational Implications

- 9. If the code is adopted, the CCSA website should include a part devoted to this code, where the list of datasets for which the code is applicable is constantly updated.
- 10. The provisions of this code do not intend to create conflicts with the general rules on dissemination that may be applicable for a CCSA member, or with general copyright statements and disclaimers that are valid for an organization as a whole, but rather to serve institutional co-operation based on the sharing of statistical data with no or minimal restrictions. Since the present code applies only between members of the CCSA, it is hoped that any general rules de facto allow exceptions within a community like the CCSA members should the present code not be fully in line with the general rules of an organization. For this reason, it would not be advisable to ask for an opinion on this code from legal advisers, but to consider this issue as a professional one that is part of the responsibility of the chief statistician of each organization.

⁵ SDMX standards were adopted by the CCSA as data exchange standards in February 2007 – See: [Report of the 9th Meeting of the Committee for the Coordination of Statistical Activities \(New York, 26 February 2007\)](#)