Distr.
GENERAL

WP.33
7 April 2010

ENGLISH ONLY

UNITED NATIONS ECONOMIC COMMISSION FOR EUROPE (UNECE)
CONFERENCE OF EUROPEAN STATISTICIANS

EUROPEAN COMMISSION
STATISTICAL OFFICE OF THE EUROPEAN UNION (EUROSTAT)

ORGANISATION FOR ECONOMIC COOPERATION AND DEVELOPMENT (OECD)
STATISTICS DIRECTORATE

Meeting on the Management of Statistical Information Systems (MSIS 2010)
(Daejeon, Republic of Korea, 26-29 April 2010)

Topic (ii): Software sharing and shared maintenance

Data Documentation, Data Dissemination, and Data Exchange at the National Institute of Statistics (NIS)

Prepared by Saint Lundy, ICT Department of NIS, Cambodia

I. Introduction

1. Data Documentation, Dissemination, and Exchange with all types of datasets including surveys, censuses and other administrative data among line ministries have still been a problem at NIS (National Institute of Statistics) for a long time until new Toolkit software, the National Data Archive Software (NADA) software, and the Statistical Data and Metadata Exchange (SDMX) have been introduced to management and staff there. So far, ICT department of NIS have produced many datasets for the benefits of policy makers, development partners, investors, researchers, and other stakeholders to use for their own purposes, but the way to document, persevere, and disseminate data have posted a question to them on how qualify and credible those datasets encompass. Toolkit, NADA, SDMX are software and system which aim at better data documentation, dissemination, and exchange have provided a great help for NIS to improve its data in accordance with international standards so that NIS can catch up with other countries which have already documented, disseminated, and exchanged their data in a standard fashion.

II. Software and System for Data Documentation, Dissemination, and Exchange

A. Toolkit and NADA Software

2. By using Toolkit and NADA software, NIS can complete its task to produce an internationally accepted data. This will help researchers to find the data they are interested in with relatively little start-up time, to understand what the data are measuring and how the data have been created, and to assess the quality of the data. It gives, therefore, a reputation and credibility to NIS for research data it has conducted. There are two components of Toolkit software. The first component is Metadata Editor which allows the user to add survey metadata and create the ddi.xml and as Nesstar study document. Medadata Editor has some existing templates that give users a choice on whether to use existing templates or to create their own templates if they wish.

Medadata Editor has many elements that make users easy to document their data. Those elements, for instance, are identification (title, series information, ID number), version, overview (abstract, kind of data, unit of analysis), scope, coverage, sampling, data collection, data processing, data access, missing data, and a lot more elements which give detail information of dataset producers (NIS) have created. Moreover, Metadata Editor has PDF Generation Report that can create a brief PDF report to all information stored in the above mentioned elements. The second component of Toolkit software is CD-ROM Builder. It allows users to generate HTML output from the study that can be published on a CD or the Internet, primarily for disseminating microdata but could also be for archiving. In regard to NADA, it is a search engine that allows users to import the ddi.xml and search for variables and view metadata on the internet. It is to host into the website after finishing data preparation and documentation.

B. SDMX System

3. In addition to Toolkit and NADA software, NIS will have a plan to implement the Statistical Data and Metadata Exchange (SDMX) standards for exchanging data and metadata, which it has requested to the European Commission (EC) to support on Technical Assistance (TA) so that NIS is able to exchange data with international organizations and agencies seamlessly. The SDMX standards allow national and international organizations to exchange data and metadata through the use of technology, in an efficient and transparent way. The standards will allow NIS to fulfil their responsibilities towards users and partners, including international organizations by using their online databases to give access as soon as the data are released.

IV. Conclusion

4. As the aforementioned of the importance of using Toolkit software, NADA software, and SDMX to document, disseminate, and exchange data, it is very necessary that NIS will try to apply these new software into its working system as good as it can. This will help strengthen NIS capacity in producing statistical data relevant for policy design, monitoring and evaluation, by:

- Better documenting, preserving, anonymizing and disseminating existing microdata
- Better exploiting existing datasets (quality assessments, further analysis)
- Better strategizing and aligning survey programs and statistical outputs to priority data needs.