

**Committee for the Coordination of Statistical Activities**

**SA/2009/14**

**Fourteenth Session**

**11 August 2009**

**Bangkok, 9-11 September 2009**

**Items for information: Item 3 of the provisional agenda**

=====

**SDMX - Progress report**

Prepared by Eurostat



EUROPEAN COMMISSION  
EUROSTAT

Directorate B: Quality, methodology and information systems  
**Unit B-6: Reference databases and metadata**



**Paper for the meeting of the Committee for the Co-ordination  
of Statistical Activities**

**Bangkok, 9-11 September 2009**

**SDMX  
Progress report**

Author: A. Götzfried (B6)  
M. Pellegrino (B5)

## 1. PURPOSE

The purpose of this document is to report on the progress made with regard to the SDMX standards and guidelines as well as with regard to the implementation of these standards and guidelines in statistical domains at sponsoring or other statistical organisations.



*SDMX consists of technical and statistical standards, and guidelines, together with an IT architecture and IT tools, to be used for the efficient exchange and sharing of statistical data and metadata.*

New and additional opportunities arising for international statistical organisations for an increasing use of the SDMX standards and guidelines will also be emphasized.

## 2. PROGRESS ON THE SDMX TECHNICAL AND STATISTICAL STANDARDS

The SDMX standards and guidelines are managed by seven international organisations (namely the Bank of International Settlements, the European Central Bank, Eurostat, the International Monetary Fund, the OECD, the United Nations Statistical Division and the World Bank). These organisations act as sponsors of SDMX.

In 2008 and 2009 larger work progress was achieved on the SDMX technical and statistical standards and guidelines, in particular with regard to the statistical standards and the availability of SDMX compliant IT applications. This work progress refers in particular to:

- (1) The approval of the technical SDMX standards (version 2);
- (2) The SDMX Content-oriented Guidelines which were officially released in January 2009;
- (3) The increasing availability of SDMX registries or other SDMX IT tools (such as SDMX converters or visualisation tools) produced by sponsoring or other statistical organisations.

**As also confirmed by the SDMX Global Conference which took place in January 2009 in Paris, the main focus of SDMX will change from the creation phase of the SDMX standards and guidelines to the implementation of these standards and guidelines in statistical subject matter domains. The latter is to be done by sponsoring or other statistical organisations.**

For further information please see also <http://www.sdmx.org>.

**For 2009/2010, the SDMX sponsors put emphasis on the following actions:**

- **Capacity building** with the following details: the creation of a standard structure for a SDMX workshop, enforced contacts with PARIS21 and the release of a SDMX e-learning package.
- **Improvements of the SDMX website:** better presentation and a better inventory of SDMX implementation activities (**see also the letter written by the SDMX chair addressed to the CCSA Member organisations in annex 1 of this document**), launching of an SDMX user forum, set-up an SDMX IT tool inventory.
- SDMX internal organisational topics related to the SDMX secretariat or the SDMX sponsors.

**What are structural metadata and reference metadata?**

**Structural metadata** are those metadata acting as identifiers and descriptors of the data, such as names of variables, dimensions of statistical cubes or titles of tables. Structural metadata must be associated with the data, otherwise it becomes impossible to identify, retrieve and browse the data.

**Reference metadata** are metadata that describe the contents and the quality of the statistical data (conceptual metadata, describing the concepts used and their practical implementation, methodological metadata, describing methods used for the generation of the data, and quality metadata, describing the different quality dimensions of the resulting statistics, e.g. timeliness, accuracy). While these reference metadata exist and may be exchanged independent of the data and its structural metadata, they are often linked ("referenced") to the data.

**3. SDMX IMPLEMENTATION AT SDMX SPONSORING OR OTHER STATISTICAL ORGANISATIONS**

**3.1 In general**

**In order to take full advantage of the benefits of SDMX, the following main SDMX implementation actions should be undertaken for each statistical domain:**

- review of the existing statistical business process used for preparing, transmitting and disseminating data and metadata;
- creation or improvement of SDMX-compliant data structure definitions (DSDs), created on the basis of already existing data structure definitions, legal acts or questionnaires;
- analysis of existing code lists and use of harmonised SDMX code lists in the respective data structure definitions (as far as already available);
- adaptation of reference metadata standards to SDMX standards and content-oriented guidelines (an example in this respect is the Euro-SDMX Metadata Structure, ESMS, used by the European Statistical System);

- adaptation of the domain-specific IT applications in order to render them SDMX-compliant.

A pre-condition for the SDMX implementation in statistical domains is certainly that the datasets and the corresponding statistical processes are stable. Therefore, statistical domains which are in development or which are facing major changes of the data structures are often put on lower priority.

These main implementation actions will mostly be launched at the level of international organisations with a considerable impact on the statistical business processes in the respective Member states. The actions mentioned above do not necessarily need to be done at the same moment. They can be stretched over a certain time period.

### **3.2 At the level of the European Statistical System**

At the level of the European Statistical System, Annex 2 of this document provides a short overview on the current status of the SDMX implementation in statistical domains. Around half of the datasets which Eurostat currently receives use already SDMX-compliant data structure definitions. The senior management supported SDMX recently again in also asking for acceleration in implementation within the European Statistical System.

These existing data structure definitions were developed for the GESMES/TS standard which is now part of SDMX. The common SDMX Information Model supports data transmission using both the GESMES (EDIFACT) syntax and the XML syntax (SDMX-ML). Both existing and newly created data structure definitions are valid for both syntaxes. New or revised data structure definitions will also use more and more harmonised code lists which will get successively available.

Eurostat has also produced and made available a number of IT tools facilitating the SDMX implementation. These tools are developed as portable applications which can be installed in other organisations. The most relevant, available for free download as Open Source Software via the SDMX website, are:

- The Euro-SDMX Registry, which provides a user interface as well as web services for interacting with the structural metadata in use within Eurostat and with statistical partners. The registry will enable national statistical offices and international organisations to obtain Data Structure Definitions (DSD), Code lists and other structural metadata, such as Metadata Structure Definitions (MSD) in use for data and metadata exchange. The Euro-SDMX Registry is already populated with all the SDMX artefacts currently available in Eurostat, including Code lists, DSDs, Concept Schemes, Organisation Schemes, Category Schemes, etc.
- The Data Structure Wizard (DSW) – a desktop application for working with SDMX structural metadata. The DSW can be used both off-line and on-line; in the on-line mode, users can interact with any SDMX-compliant registry.
- The Business Cycle Clock version 2.0 - an interactive application for dynamic visualization of short-term economic indicators, which can be fed by data in SDMX-ML format.

Other tools have been developed for internal use at Eurostat, in contact with Member states, and they often are closely tied to Eurostat's IT infrastructure. Among these, there are IT tools aimed at facilitating the delivery of SDMX reference metadata to Eurostat.

#### 4. FURTHER INVOLVEMENT OF INTERNATIONAL ORGANISATIONS

The further co-operation on SDMX at the level of international organisations mainly refers to the following:

- **At the level of the SDMX sponsors:** the international sponsoring organisations will further manage the SDMX technical and statistical standards and improve them whenever necessary (e.g. with regard to the release of more harmonised code lists). International and other statistical organisations will be able to contribute via the SDMX web site or the SDMX user forum.
- **At the level of implementation of data structure definitions:** international organisations need to proceed with the creation of SDMX-compliant data structure definitions in statistical domains; an enforced co-operation between international organisation will get necessary when common data sets and data collections are concerned (e.g. in energy or education statistics).
- **At the level of reference metadata:** International organisations should use the SDMX technical and statistical standards and guidelines for their respective reference metadata standards. The exchange of SDMX-based reference metadata between international organisations should be enforced, also in order to reduce the reporting burden on the respective Member states.

#### 5. CONCLUSIONS

The creation of the SDMX technical and statistical standards and guidelines progressed considerably in 2008/2009. Highest emphasis should now be put on the use and implementation of the SDMX standards and guidelines at the level of international and national organisations.

<b>ANNEX 1:</b>
-----------------

**Letter of the SDMX Chair from July 2009**

17 July 2009

**Development of SDMX Implementation Roadmap**

As you know, implementations and interest in using the SDMX framework of standards and guidelines is growing around the world, particularly in national and international organisations engaged in statistical compilation and dissemination.

To facilitate these developments, efforts are now underway to collect and publish an on-line and regularly updated roadmap on the SDMX website, in order to inform others in the global SDMX user community of achievements and work in progress.

Your institution is cordially invited to contribute input to this SDMX Implementation Roadmap. A template to be used in providing your information is attached. When completed, you are kindly asked to send it to the SDMX Secretariat (secretariat@sdmx.org).

It would be most appreciated if you can report on the following information by the end of August 2009:

- subject-matter domain groups<sup>1</sup> in which you are encouraging the use of the SDMX framework, e.g. by facilitating the development of data and metadata structure definitions
- projects or initiatives in which you are taking on your own or collaborating with other international organisations to reduce national reporting burden by the using the SDMX framework to encourage data-sharing arrangements, e.g., encouraging the provision of data and metadata on national websites using SDMX standards and guidelines so that international organisations and users can pull data from the site on demand.

Thank you for your cooperation and support for the further adoption of SDMX.

Pedro Díaz Muñoz

SDMX Chair

---

<sup>1</sup> See attached two documents: (1) "inventory of groups" and (2) the SDMX Subject-Matter Domain guideline.

## Annex 2: SDMX implementation within the European Statistical System (June 2009)

Theme	Description	SDMX implementation status	Data structure definitions
<b>Demographic and social statistics</b>			
1.01	Population	2 projects under way: a) demography rapid indicators (test of pull approach) b) Census Hub pilot project, aimed at delivering in SDMX format the data for the 2011 census round across the EU c) general demography statistics	Yes for (a) and (b).
1.02	Labour market	New project starting June 2008, continued in 02/2009	Several existing SDMX-compliant DSDs. New DSD(s) to be developed.
1.03	Education and lifelong learning	DSD development with UNESCO, OECD; likely area for future data-sharing.	Work in progress (to be pushed ahead)
1.04	Culture	Not tackled yet	
1.05	Health and safety	Agreed to analyse joint questionnaire Eurostat-WHO-OECD to prepare for DSD, follow-up in 04/2009	GESMES 2.1 DSD exists for statistics on accidents at work.
1.06	Distribution of incomes and living conditions	Follow-up from the side of reference metadata and quality reports needed	
1.07	Social protection	Not tackled yet	
1.08	Consumer protection	Not tackled yet	
1.09	Food safety	Not tackled yet	
1.10	Crime and criminal justice	Not tackled yet	
1.11	Discrimination	Not tackled yet	

1.12	Other work in the field of demographic and social statistics	Not tackled yet	
<b>Macroeconomic statistics</b>			
2.01	Annual economic accounts	New SDMX-compliant DSD agreed between Eurostat-ECB-OECD and in use since 2007; metadata to be followed-up	Yes
2.02	Quarterly accounts	New SDMX-compliant DSD agreed between Eurostat-ECB-OECD and in use since 2007; metadata to be followed-up	Yes
2.03	Monetary and financial statistics	New SDMX-compliant DSD agreed between Eurostat-ECB-OECD and in use since 2007; metadata to be followed-up	Yes
2.04	Government finance statistics	New SDMX-compliant DSD agreed between Eurostat-ECB-OECD and in use since 2007; metadata to be followed-up	Yes
2.05	Monitoring own resources	Not tackled yet	
2.06	Prices	a) SDMX-based project for exchange of PPP data b) For HICP reference metadata/quality reports to be followed-up; discussions on new DSDs were started.	Yes for (a) PPP and (b) HICP.
2.07	Trade in goods	Whole domain will get SDMX compliant; metadata to be followed-up;	GESMES 2.1 DSD exists.
2.08	Trade in services and balance of payments	New BOP DSD agreed Eurostat-ECB-IMF (replacing existing DSD). There is also an existing SDMX-compliant DSD for FATS. Metadata to be followed-up; bilateral meeting to be envisaged.	Yes (not in all domains)

<b>Business statistics</b>			
3.01	Annual business statistics	Need to consider new actions on SDMX implementation. DSD needs to be updated or preferably replaced by SDMX-compliant DSD (GESMES/TS);	GESMES 2.1 DSD exists but not used.
3.02	Short-term business statistics	Some indicators are in SODI project; ESMS metadata produced already; to be followed-up with the unit.	Yes
3.03	Energy	New project aiming at SDMX-based data sharing with IEA. Data collection via eDAMIS Web Forms uses SDMX; more input on the DSD and on metadata needed.	Work continues
3.04	Transport	Metadata need to be followed up with the unit; broader bilateral meeting needed with the unit in term	GESMES 2.1 DSDs exist for road, rail, air and maritime statistics.
3.05	Tourism	New data structure not stable at this stage; no immediate action on the DSD needed. Metadata could be tackled.	SDMX-compliant DSD exists.
3.06	Business registers	EuroGroups Register project uses SDMX (pilot data collection);	Under development.

**Statistics on agriculture, forestry and fisheries**

4.01	Agricultural production	Data collection via eDAMIS Web Forms uses SDMX; follow-up on metadata needed;	ad hoc DSDs corresponding to Web Forms. Definitive DSDs to be developed.
4.02	Agricultural structures	Not tackled yet	
4.03	Land use and land cover	Not tackled yet	
4.04	Agricultural accounts and prices	Not tackled yet	
4.05	Forestry statistics	Not tackled yet	
4.06	Fisheries statistics	Expressed interest in 02/2009; follow-up in terms of metadata needed.	
4.07	Reform of agriculture, forestry and fisheries statistics	Not tackled yet	

**Multi-domain statistics**

5.01	Statistics for business cycle analysis	SDMX does not seem to be needed	
5.02	Lisbon strategy and sustainable development	SDMX does not seem to be needed	
5.03	Environment statistics and accounts	SDMX used for waste statistics; real DSD needed; contact with the unit after reorganisation needed (also with regard to environmental data centres)	Yes, for waste statistics
5.04	Regional and urban statistics	Follow-up on metadata needed (Mars)	
5.05	Geographical and local information	SDMX does not seem to be needed	
5.06	Science, technology and innovation	Some discussions in 2008; follow-up in 02/2009 needed (also with regard to metadata)	
5.07	Information society	Some discussions in 2008; follow-up in 02/2009 needed (also with regard to metadata)	