SDMX implementation in the statistical practice

Marco Pellegrino, Eurostat

Key words: Metadata, registry, SDMX

The SDMX initiative (Statistical Data and Metadata eXchange) is aimed at developing efficient processes for exchange and sharing of statistical data and metadata among international organisations and their member countries. In the years 2001-2008, SDMX produced a set of technical standards and statistical (content-oriented) guidelines to be used for the exchange of data and metadata. SDMX technical standards are recognised as ISO TS 17369 in the version 1, while version 2 is in the process of ISO validation. Full information on the SDMX standards, guidelines and organisation is available at http://www.sdmx.org.

The version 2.0 of SDMX introduced a series of enhancements on the previous version, in particular on metadata management (with the introduction of the "metadata structure definition" to describe the structure of a metadata set) and on the so-called registry architecture, useful for providing visibility to large amounts of data and metadata. The SDMX Registry can be seen as a central application which is accessible to other programs over the Internet (or an Intranet or Extranet) to provide information needed to facilitate the reporting, collection and dissemination of statistics.

In January 2009, SDMX sponsors disclosed the first package of content-oriented guidelines. This package recommends practices for creating interoperable data and metadata sets using SDMX technical standards. They focus on the standardisation of specific concepts and terminology that are common to a large number of statistical domains. Such harmonisation is useful for achieving a more efficient exchange of comparable data and metadata, building on the experience already gained in several implementations.

Eurostat is currently implementing the SDMX content-oriented guidelines for the harmonisation of both structural and reference metadata within the European Statistical System. With regard to reference metadata, Eurostat is using the cross-domain concepts list for the implementation of a Europe-wide standard (ESMS, Euro-SDMX Metadata Structure) to be used for metadata reporting to Eurostat and for exchanging metadata between international organisations, for instance between Eurostat, the European Central Bank and the IMF. The new structure, fully compliant with SDMX standards and guidelines and with the European Statistics Code of Practice, has already been implemented in Eurostat for metadata dissemination, and is being recommended for use within the ESS through a specific Commission recommendation. Technical specifications for the use of ESMS in accordance with the SDMX-ML standard format are also ready.

This document also reports on the development of some SDMX-compliant tools and on the general IT infrastructure at Eurostat. SDMX envisages the promotion of a data-sharing architecture using the pull mode to facilitate low-cost and a high-quality statistical data and metadata exchange: a data reporting organisation publishes data once, and lets their counterparties "pull" data and related metadata as required. The data-sharing architecture is based on the possibility of discovering easily where data and metadata are available and how to access them. In general, SDMX tools are produced in a decentralised manner and made freely available to the user community on the web (a specific section dedicated to IT tools has been added to the SDMX website). A series of tools have been produced for addressing several needs; these are available for Member States, under the EUPL Open Source Software license, as packages containing the application, documentation and source code files. Other tools have been developed for internal use at Eurostat: these are not portable, because they are more or less closely tied to Eurostat's IT infrastructure.

SDMX standards and IT architecture, together with content guidelines, are being used for a series of implementation projects at Eurostat, also involving ESS countries. One of these is the European Census Hub for the dissemination of the result of the 2011 census data in the European Union, via Eurostat web site, using SDMX solutions.