

OECD Expert Group on Statistical Data and Metadata Exchange

Conclusions from the Meeting 20-21 April 2005

The second meeting of the OECD Expert Group on Statistical Data and Metadata Exchange was held on 20-21 April 2005. The purpose of the Expert Group is to identify ways of improving OECD's data collection and data sharing activities with member countries, and to promote the development and use of SDMX standards for this purpose. The Expert Group meetings are the most important opportunity available for member countries to participate in discussions on SDMX standards and their evolution.

The Expert Group meeting was held back-to-back with the OECD/UN/Eurostat meeting on Management of Statistical Information Systems, to facilitate the participation of delegates. Both meetings were held in Bratislava at the invitation of the national statistical office of the Slovak Republic.

There were 40 participants from statistical agencies of OECD member countries in the Expert Group meeting. A number of partner international organisations were also represented: UN/ECE, Eurostat, IMF, and ECB. Two consultants from the SDMX standards team were present and contributed to the meeting – notably presenting the beta version of a toolkit intended to facilitate adoption of SDMX standards.

The following points represent the main conclusions of the meeting:

1. Discussion at the Expert Group meeting confirmed that **XML and Web services are now generally accepted technology standards for management and sharing of statistical data and metadata**, and are being deployed by most agencies – for both internal and national purposes. As XML and Web services are technology foundations for SDMX, it was therefore surprising that few agencies had yet sought to implement version 1 of the SDMX standards, which had been released in September 2004.
2. Participants **reaffirmed the need to establish internationally agreed standards for the exchange of statistical data and metadata, and recognised the potential of SDMX in this regard.**
3. At the same time, discussion revealed that there are **a number of concerns that need to be addressed before SDMX standards will be widely embraced.** There is, in particular, a widely held view that SDMX needs to be presented in a more understandable way, and that national statistical agencies need help in establishing a sound business case for investment in SDMX. Many delegates alluded to perceptions that SDMX standards, which some characterised as being “abstract”, would be difficult to implement. Many delegates were puzzled by what they deemed to be unnecessarily complex terminology. The “key family” concept was repeatedly cited as an example of terminology that had the effect of quickly cooling initial interest in SDMX standards. Some alternatives were put forward, such as “data structure definition” or “cube definition”. But participants from the SDMX standards team noted that making terminology changes could be difficult, now that SDMX standards had been submitted to ISO.
4. Participants from the SDMX standards team gave a demonstration of an **SDMX toolkit**, aimed at facilitating the implementation of SDMX standards. The beta version of the toolkit was released to SDMX sponsor organisations in early April. Wider distribution of the toolkit awaits resolution of a certain number of legal issues. Delegates welcomed the toolkit initiative, and felt that the idea of a toolkit was a good one. Many noted, however, that to be really useful, the toolkit should include some concrete examples -- taking as a point of departure, for example, a small data set, and illustrating how it would be put in SDMX-ML format for transfer, and what the corresponding web services would be.
5. Many participants expressed the view that **a concerted “marketing effort” was needed** to help jump start wide adoption and use of SDMX standards. Simplifying terminology and providing a toolkit would be important parts of this marketing effort. Many delegates noted that this marketing effort would

also need to include elements that would help statistical agencies establish the “business case” for adopting the standards.

6. There was general agreement that SDMX was not just a technical exercise, but also carried with it a number of organisational, statutory and legal issues. It would soon be necessary to **address a number of outstanding governance and support issues** (i.e., who “owns” the standards and will ensure their on-going evolution, who manages and communicates “key family” concepts for various categories of statistics, who maintains the Metadata Common Vocabulary, what are the minimal compliance requirements to be able to say “SDMX conformant”, etc.).

7. While SDMX version 1.0 is already available for use, preparation of SDMX version 2.0 is underway. Important extensions include:

- Hierarchical data structures
- Standard for reference metadata.
- Standard for classification of statistical themes
- Registries

Version 2.0 will be backward compatible with Version 1.0 standards. A draft of SDMX version 2.0 is expected to be published for comment in the middle of the year. Comments from members of the Expert Group on the draft standards are essential. A final version 2.0 is expected during the autumn of 2005.

8. There was an initial discussion about "notification" mechanisms that could be used to notify recipients that updates or new estimates are available, so the corresponding data can be "pulled" from the hosting service. A number of participants felt that RSS (Real Simple Syndication) could be very useful in this respect, and should be further investigated.

9. Representatives from national statistical agencies indicated they would welcome initiatives that would ease the burden of reporting statistics to international organisations. But some made it clear that this alone would not be enough to justify their investment in SDMX. **SDMX must also be, and be demonstrated to be capable of responding national data-sharing needs as well.** To this end, there was a strong wish to have the standards extended to cover, in addition to multidimensional tables, micro data and observation-level data. Some members indicated they wished to ensure compatibility regarding terminology and concepts with ISO 11179. Some countries reported that they are already looking into the feasibility of using SDMX as a vehicle for exchange of statistics either (a) among partners in their countries' national statistical system or (b) internally within their organisation.

10. The need for SDMX “proof of concept” pilots, involving “real world” applications, was a recurring theme. The “alignment projects” NAWWE (National Accounts World Wide Exchange) managed by OECD, and SODI (SDMX Open Data Interchange) managed by Eurostat, were seen as important examples. Some delegates felt such implementations would be a key facet of the required marketing effort, by demonstrating the applicability and usefulness of SDMX in small, but practical and easily-understandable data-sharing applications. Successful pilots would moreover provide implementation examples that others could emulate, generate confidence needed to justify investment in large-scale implementation, and provide a source of feedback for enriching SDMX standards. **Several countries and international organisations expressed their willingness to work with OECD on small-scale practical implementations of SDMX standards for data and metadata exchange. This work is to be initiated by OECD,** with a view to providing “how to” examples of implementing SDMX that others could emulate, and helping establish confidence in the ability of the standards to respond to “real world” needs. Experience gained will be shared with the SDMX team with a view to further enriching the standards.

Agenda, papers, presentations and list of participants are available at <http://www.oecd.org/statistics/exchangeexpertgroup/2005>