Organisation de Coopération et de Développement Economiques
Organisation for Economic Co-operation and Development

20-Apr-2004

English - Or. English

STATISTICS DIRECTORATE
STATISTICAL INFORMATION MANAGEMENT AND SUPPORT DIVISION

STD/SIMS(2004)14
Unclassified

Conclusions from the First Meeting of the Expert Group on Statistical Data and Metadata Exchange

Paris 1-2 april 2004

Lars Thygesen, Statistics Directorate
Tel: 33 1 45 24 84 02
Email: lars.thygesen@oecd.org

JT00162489

English - Or. English

CONCLUSIONS FROM THE FIRST MEETING OF THE OECD EXPERT GROUP ON STATISTICAL DATA AND METADATA EXCHANGE, PARIS 1-2 APRIL 2004

(To be submitted to the OECD Statistics Committee)

1. The Expert Group met 1-2 April 2004 at OECD headquarters in Paris, Château de la Muette[1]. The meeting was attended by experts from the national statistical offices of 21 Member countries, as well as from 5 other international organisations. The meeting included a very fruitful exchange of experience in the field of data exchange. However, the aim of this report, rather than summarising all of these valuable discussions and presentations, is to highlight the action points agreed upon, and the conclusions that would be of interest to the management of OECD, and to national and international organisations.

2. This meeting was the first opportunity for the OECD to thoroughly discuss the technical aspects of the cooperation concerning data collection and exchange with its constituency, i.e. the national statistical offices (NSOs) of Member countries. The focus was on proposals on how to enhance this cooperation in order to render the exchange more efficient, leading to higher quality and easing the work burden on NSOs, as well as on OECD. Especially, it was discussed how new technologies should be used to enhance effectiveness. There was agreement that standards for exchange would be essential in this work. The meeting aimed at strategies and solutions that could be implemented and create tangible results in the short term, as well as in the intermediate and longer term.

3. Participants stressed the distinction between the "technology" aspects (i.e., "how" to share/exchange), and the "contents" aspects (i.e. "what" to exchange and how to make the technology work in a standardised way for statistical data and metadata)... Most participants seemed to feel that the "standards relating to contents" issue was the biggest challenge as "technology" was always emerging with new opportunities for greater efficiency. Thus the interest in and support of SDMX[2], which seeks to address standards questions in the field of statistical information exchange and web dissemination. Some concern about a multiplicity of standards was expressed. The people responsible for developing SDMX standards were therefore encouraged to take other work in the field into account and make use of it.

The results aimed at were twofold:

- long-term objective – standards and procedures;
- short-term – achieve practical results from what is already available, and identify and progressively remove obstacles to this; follow-up bilaterally to improve the current situation.

[1] Documentation of the meeting (agenda, terms of reference, participants list, papers and presentations can be seen on OLIS or on www.oecd.org/std/research/exchangeexpertgroup

[2] SDMX (*Statistical Data and Metadata Exchange*) is an international cooperation initiative aiming at developing and employing more efficient processes for exchange and sharing of statistical data and metadata among international organisations and their member countries.

4. The Terms of reference [STD/SIMS(2004)1] were adopted with an amendment proposed by the Swiss delegate, and agreed by all. The amended TOR is now available on OLIS.

5. The Expert group agreed that XML[3], already being used in several NSOs for their own development work aiming at national statistical needs, would be very important in standardising and easing the exchange between NSOs and OECD (and other international organisations).

6. Many pointed out the importance of developing agreed, common XML-based solutions. For the moment, many of the organisations present were implementing their own XML standards. Nevertheless, a main message was the importance for all to get quickly familiar with XML, and not wait for the standards. Getting into XML is key – solutions can easily be transferred to other XML formats as standards emerge.

7. Web Queries is a Microsoft Excel based technique presently being used by OECD for automating some of its data collection, sending tailor-made queries to NSOs' Internet dissemination databases. It can save resources at the NSO as well as at OECD but can only be used when databases fulfil certain demands (timely updating, sufficient contents, and certain technical restrictions); thus, this technique is used in relation to 13 NSOs[4]. In the short term, most participants committed themselves to try, together with OECD, to solve some of the technical problems in order to increase the possibility of use of Web Queries.

8. However, most participants stressed that already in the relatively short term, more efficient and standardised methods should be implemented, as Web Queries require special action from OECD for each database, and needs adjusting when database interface is changed. **It was agreed that Web Services would probably be the right standard** for this. Australia had already specified a draft Web Service for its nation-internal data exchange, and the SDMX pilot project (see below) also includes the specification of a Web Service. **Many participants agreed to proceed with the specification and pilot implementation of a Web Service for exchange between OECD and NSOs in the very short term, as soon as the SDMX reviewing process (see 7 below) was finished which was scheduled to be 1 June 2004. Australia would take the lead in this task force, joined by OECD, Austria, Germany, and Norway**.

9. The SDMX was a key theme in the meeting. It was agreed that SDMX had great potential of giving good and generally accepted standard solutions to the exchange problems, and consequently SDMX was cautiously supported by the NSOs present. In the present situation, countries must send same data in multiple formats to multiple international organisations. SDMX seeks to address this. It was agreed that the NSOs should take active part in the public review process for proposed SDMX standards. The tentative schedule for this year is expected to start when the new SDMX web site is launched at the beginning of May. Version 1.0 draft standards would be published on the site and comments invited through June. A more complete set of draft standards (version 2.0) is envisaged to be published on the site in about a year, taking into account issues that would need more time to be addressed in a broader context of requirements and functionality. The OECD will send e-mail messages to the Expert Group to alert them when the draft standards are available on the SDMX web site. The Expert Group would also discuss feedback on issues relating to the standards among members. The proposed SDMX guidelines for the use of web services would be key to the proposed exchange between NSOs and OECD (see 6 above).

10. The SDMX work on specification of a standard for metadata repositories was regarded as very important to Member countries. It was necessary to make sure that the coming standard was in harmony with information actually processed in and needed by NSOs also for national purposes in order to avoid duplication of work. Again, the Expert Group should be actively involved in the review process.

[3] E**x**tensible **M**arkup **L**anguage

[4] Including three were the method is specially adapted

11. The case of the European Central Bank's extensive use of GESMES/TS for standardising and automating all statistical data collection was demonstrated as a very convincing proof of the SDMX concept. It was noted with satisfaction that the GESMES/TS data model is the foundation of the initial release of SDMX standards. GESMES/TS is, however, time series oriented, and does not necessarily respond to all needs for sharing exchange of cross-sectional and large volume data. It is important to ensure forward compatibility of GESMES/TS within SDMX.

12. The NAWWE (National Accounts World Wide Exchange) pilot project was shown as another case study, demonstrating the potential of the SDMX ideas of standard-based sharing of statistical data, and its new web site (stats.oecd.org/nawwe/) was launched. **Eurostat and ECB agreed to work more closely with the NAWWE project to see if a common XML based data sharing model could be implemented in the medium term.**

13. The Group discussed the evolution towards a "pull" model, whereby data providers (countries) make data available at a specified place and the data collector performs the action of extracting the data , rather than the, now more frequently used, "push" model, whereby countries must take the initiative to "send" data. The present use of GESMES/TS is through "push" technology. Web services as foreseen in SDMX makes use of "pull" technology.

14. The importance of a "demonstration environment" for testing proposed new standards against real statistical activities was repeatedly emphasised. That is, "proof of concept" should be provided to verify that proposed standards can work, and feedback should be provided on how the proposed standard might be enriched.

15. There was some discussion of "open source" software. A few countries had started to experience with the use of "open source" and further progress in this area would be useful to review in future meetings.

16. Participants found **that the outcomes of the meeting underlined the importance of this Group and indicated a need for future work**. It was expected that there would be important new results and proposals to discuss in 6 months, when the proposed Web Service had been piloted. Consequently it was hoped that a meeting could be scheduled for then, preferably back-to-back with another meeting already demanding the presence of several of the members. The Eurostat IT Steering Committee would be a good candidate.

17. Between meetings, members should communicate intensively by electronic means. The OECD will take the lead in fostering this communication.