E

Economic and Social Council

Distr.
GENERAL

ECE/CES/2006/4/Add.4
4 May 2006

ENGLISH ONLY

ECONOMIC COMMISSION FOR EUROPE   STATISTICAL COMMISSION

CONFERENCE OF EUROPEAN STATISTICIANS

Geneva, 3-5 April 2006

REPORT OF THE APRIL 2006 WORK SESSION ON STATISTICAL METADATA (METIS)

Note by the Secretariat

This meeting was organised jointly by the UNECE secretariat, Eurostat and OECD.

1.   The Joint UNECE/Eurostat/OECD work session on statistical metadata was held in Geneva, Switzerland, from 3 to 5 April 2006. It was attended by participants from: Albania, Australia, Austria, Brazil, Bulgaria, Canada, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Japan, Kazakhstan, Latvia, Lithuania, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, Romania, Russian Federation, Slovakia, Slovenia, South Africa, Spain, Sweden, Switzerland, The former Yugoslav Republic of Macedonia, Turkey, United Kingdom, and United States of America. Representatives of the Statistical Office of European Communities (Eurostat), European Central Bank (ECB), Bank for International Settlements (BIS), Food and Agricultural Organization of the United Nations (FAO), International Monetary Fund (IMF), Organization for Economic Cooperation and Development (OECD), World Bank, World Trade Organization (WTO), United Nations Educational, Scientific and Cultural Organization (UNESCO), United Nations Conference on Trade and Development (UNCTAD) and the United Nations Industrial Development Organization (UNIDO) also attended. Participants from the Catholic University of Leuven (Belgium) and the SDMX Standards Team attended the Work Session at the invitation of the Secretariat.

GE.06

2.   The agenda of the work session consisted of the following substantive topics centred on the Common Metadata Framework for National Statistical Offices, a document that was developed after the last meeting of METIS in 2004:

        (a)   Metadata in a Corporate Context;
        (b)   Metadata Concepts, Standards, Models and Registries;
        (c)   Metadata and the Statistical Cycle; and
        (d)   Implementation.

3.   Mr. Daniel Gillman (United States) acted as Chairman.

4.   The representatives of Eurostat, OECD and UNECE welcomed delegates. Mr. Daniel Gillman addressed the meeting, highlighting the history of metadata related activities in the context of the Conference of European Statisticians and the challenges for the future work. The importance of metadata for statistical offices is evidenced by a growing number of participants and contributions.

5.   The following persons acted as Session Organizers: Topic (i) – Mr. Bo Sundgren (Sweden); Topic (ii) – Mr. Paul Johanis (Canada); Topic (iii) – Mr. Graeme Oakley (Australia); and Topic (iv) – Mr. Max Booleman (Netherlands).

RECOMMENDATIONS FOR FUTURE WORK

6.   The participants reviewed the recommendations for future work put forward by Task Force on Common Metadata Framework. Their proposal centred on completion and future maintenance of the Common Metadata Framework for National Statistical Offices. The following general recommendations were made:
        (a)   The Framework shall be a living/evolving reference to standards, concepts, and best practices. An editorial board shall be created to ensure future maintenance of the Framework.
        (b)   The Framework should be published in an electronic form on the UNECE website, with some parts published in print (after the public and editorial group review).
        (c)   The Framework shall comprise two types of information
               •   Text that is written for the Framework as its integral part;
               •   Selected and edited contributions by countries (good practices).
        (d)   It is important to add a general introduction, and provide "packaging" to link individual parts of the Framework.
        (e)   In order to attract broader readership, each item of the framework should comprise a general overview, avoiding overly detailed technical information, with links to more technical and detailed papers.
        (f)   Linked contributions may relate to multiple parts of the Framework.
        (g)   All material included in the Framework should be accessible without restriction.
        (h)   Drafts will be made available for public review. After two rounds of review they will be left stable within the Framework for some (reasonable) time.
        (i)   The draft Framework shall be updated using contributions from METIS 2006.
        (j)   The editorial board should be open to all inputs within and outside the

framework of METIS and ensure collaboration with other relevant groups.

7.   Recommended changes to individual parts and sections of the Framework are enumerated in Section VI of the Annex.

8.   The participants also noted establishment, by the Bureau of the Conference of European Statisticians, of the Task Force on Electronic Data Reporting for Primary data Collection. Eurostat representatives expressed the willingness of Eurostat to co-sponsor the Task Force. Detailed recommendations are contained in Section VII of the Annex.

FURTHER INFORMATION

9.   The conclusions reached during the discussion of the substantive items of the agenda are contained in the Annex. A summary of the feedback received from participants is also attached. All background documents and presentations for the meeting are available on the website of the UNECE Statistical Division:
                    http://www.unece.org/stats/documents/2006.03.metis.htm.

ADOPTION OF THE REPORT

10.   The participants adopted the present report before the Work Session adjourned.

ANNEX

**SUMMARY OF THE MAIN CONCLUSIONS REACHED AT THE WORK SESSION
ON STATISTICAL METADATA (METIS)**

I.   METADATA IN A CORPORATE CONTEXT
Session Organizer: Bo Sundgren (Sweden)
Documentation: Invited papers by UNECE/Australia and Netherlands; supporting papers by
Brazil, Bulgaria and Czech Republic.

1.   At the METIS work session in February 2004, a Task Force was established to draft a
Statistical Metadata Framework. The task force has divided the framework into four parts: A –D,
the contents of which align to the topics being discussed at this work session.

2.   The first invited paper for this topic was a draft of Part A: 'Statistical Metadata System
and its role in a statistical organization'. This part describes the value and strategic relevance of
Statistical Metadata Systems (SMS) for a statistical organization (SO). It provides information
that can be used to develop a business case for an SMS. A second invited paper on the
experiences of Statistics Netherlands, and three supporting papers from Brazil, Bulgaria and the
Czech Republic, provided further information about metadata in a corporate context.

3.   The following issues were raised during the discussion on this topic:
    (a)   Dissemination of data through the Internet will usually make it easier (but in
             some cases rather more difficult) to obtain feedback from users. The
             discoverability of data online, such as through search tools, can enable users to
             highlight errors, incoherencies and inconsistencies in data. This often provides a
             strong argument for better metadata. Publishing a contact person with each
             statistic disseminated can motivate statisticians to create better metadata in order
             to avoid repetitive questions from users.
    (b)   Mistakes have been made in the past by mystifying metadata systems and treating
             them as a separate system. It is now evident that they must be fully integrated
             within the statistical information system.
    (c)   Organizations are making the transition from costly 'stovepipe' structures to more
             efficient 'process-oriented' structures. It is vital to integrate and standardise
             processes inside the SO to better facilitate this transition. Managers need good
             management skills and statistical knowledge, rather than a focus on technical
             expertise. Relevant management methods can be used to create a new culture that
             uses interdisciplinary teams.
    (d)   There are advantages and disadvantages of both gradual implementation vs 'big
             bang' implementations. Circumstances (eg available resources) may not allow an
             agency a free choice.
    (e)   Involvement of top management in SMS projects is vital. Governance models
             should be sustainable and need to avoid creating bottlenecks.
    (f)   A suggested approach to achieve success in SMS projects it to divide the work
             into small sub-projects, each with goals that are achievable in a timely manner.
             The links between the sub-projects should be clearly defined. Resulting successes

should be communicated to senior management and other stakeholders.

(g)   In addressing the question of the 'value added' of metadata, it is important to also explore the external value, particularly in relation to administrative data in a distributed statistical system.

(h)   To help organizations make a case for improved metadata systems it would help if countries could share anecdotal evidence of how an effective SMS can result in cost savings. This evidence should be collated for the METIS group to draw upon.

(i)   Managers aren't necessarily engaged by terms such '(enterprise) architectures' and 'statistical metadata systems' but they have a set of issues that need to be addressed and are interested in solutions. Groups responsible for metadata system governance should be committed to maintaining managers' interest and understanding of metadata issues.

(j)   The importance of a vision and forward planning for SMS projects was reiterated. Having these clearly developed prevents loss of momentum as the project progresses through its phases. The vision must be integrated in the business architecture of the agency and metadata systems must be integrated in the business process systems.

(k)   The importance of a global SMS architecture was highlighted. The architecture should be based on the approved vision and cover two basic areas:
   (i)   Use of SMS in existing tasks, projects and processes of the SO
   (ii)   Unique functioning of SMS (eg access portal, search interpretation).

4.   Participants were encouraged to propose enhancements to the draft of Part A, either during the meeting or via email. The changes suggested to date are:

(a)   Include an introduction to the entire framework – why it exists, who will use it and how.

(b)   Review and elaborate on the core principles in chapter four.

(c)   Add further case studies and anecdotal evidence.

(d)   Provide more practical elements by highlighting the approaches organizations have used to achieve success (ie 'tricks'). This could provide solutions to balance the section on barriers, which focuses on the challenges faced.

(e)   Include mention of the geospacial dimension including the relevant ISO standards.

(f)   Add information about achieving higher efficiency (costs vs benefits) and improving the quality of metadata.

II.   METADATA CONCEPTS, STANDARDS, MODELS AND REGISTRIES
Session Organizer: Paul Johanis (Canada)
Documentation: Invited papers by the United States and the SDMX Standards Team; supporting papers by Finland, Slovenia, Spain, United States, Romania and OECD.

5.   The two invited papers represented draft input into the Part B of the Common Metadata Framework. They considered synergies between the following standards, which were thought as the most relevant for the NSOs at the present time:

(a)   ISO/IEC 11179 – a standard for metadata registries maintained by the Technical

Committee JTC1, Information technology, subcommittee SC 32, Data management and interchange;

(b) Corporate Metadata Repository (CMR) – an extension of ISO/IEC 11179 specifically for metadata objects that are relevant to NSOs;

(c) eXtensible Business Reporting Language (XBRL) – open standards, primarily focused on exchange of financial business data;

(d) Data Documentation Initiative (DDI) – a project aimed to establish XML based metadata standard for social sciences – for content management, preservation and transport;

(e) Statistical Data and Metadata Exchange (SDMX) – standard format for aggregated data and metadata, mainly for transmission to and between international organizations, sponsored by the BIS, ECB, Eurostat, IMF, OECD, UNSD and World Bank;

(f) Neuchâtel Group – focused on common constructs, and particularly classifications (standard terminology), sponsored by a group of national statistical offices;

(g) Common Warehouse Model (CWM) – developed by the Object Management group.

6. These standards and models were presented as complimentary rather than in competition with each other. Each is suited to a particular function in NSOs. It is essential however to be able to map these standards to each other. There is a fair chance that such mapping could be automated, because while each of the standards organizes the information differently, there may be common elements of information. This was demonstrated through looking at areas of commonality between, ISO 11179 and SDMX, particularly through:

(a) comparison of basic constructs of ISO 11179 and SDMX models;

(b) a possible method to create ISO 11179 data elements from SDMX data and metadata constructs;

(c) comparing the scope of the two standards regarding the registry specifications.

7. The SDMX Metadata Common Vocabulary (MCV) was an important input to Part B of the Framework. It has a specific position among the existing standards because it aims to be a unifying element among different models and concepts for statistical metadata. The current version of the MCV comprises 384 items, including a small set of synonyms (30), which is going to be expanded in a future release. Definitions and context explanations are taken from about 70 different sources. The MCV is available, along with other content-oriented guidelines, for public comment until 31 May 2006 from the SDMX web site. Participants were encouraged to download the relevant documents at http://www.sdmx.org and participate in this review.

8. The following issues were raised during the discussion on this topic:

(a) What standards are closer to the needs of statistical offices? The fundamental need for a statistical office is a variable centred metadata repository and a reference standard in this regard. One solution can be the Corporate Metadata Repository (CMR), based on ISO 11179, but there are other options. Ideally, the internal standards within the statistical office should allow mapping to all/most of international standards listed above.

(b)   Metadata repositories: There should be a distinction between metadata sets prepared for the final users (that is outputs) and metadata repositories that are used in statistical production. Metadata repositories are considered to be a key element of the statistical information system. The CoSSI (Finland) uses a model where metadata are stored in XML files accompanying the data.

(c)   Terminology and definitions: National statistical offices aim at using international vocabularies and databases of concepts and definitions like CODED (Eurostat Concepts and Definitions Database), the OECD Glossary, or the SDMX Metadata Common Vocabulary (MCV). In this connection, translation into other languages is an important issue, in particular different flavours of established terms in target languages. Some European statistical offices have already translated the MCV and other terms extracted from CODED into their national languages. Some countries, in addition to the translation, need to make a distinction between the popular and the formally correct statistical terms.

(d)   Eurostat works towards achieving consistency of definitions. CODED uses the MCV definitions, and will also use others emanating from the SDMX content-oriented guidelines. CODED will soon have a more user-friendly interface integrated with the RAMON database.

(e)   There was an enquiry about standards for concepts management. Such standards were not included in the present drafts, but some of the current ISO standards would serve this purpose. These may be added to Part B of the Common Metadata Framework.

(f)   Quality related information: There is a possibility to develop procedures for generating such information automatically, for example from SAS. Presently, statistical offices often provide the quality related information only for selected variables.

9.   The following suggestions for Part B of the Common Metadata Framework were made:

(a)   The situation with standards evolves with time. Therefore this part of the Common Metadata Framework should be evolving and regularly updated document that provides a snapshot of the situation at the last time of update.

(b)   The MCV is presented as a flat structure, and it may benefit from introducing a hierarchical structure.

(c)   Mapping from one standard to another and from one vocabulary to another is an important theme for the Framework.

III.   METADATA AND THE STATISTICAL CYCLE
Session Organizer: Graeme Oakley (Australia)
Documentation: Invited papers by Canada, New Zealand, Slovenia and South Africa; supporting papers by Finland, Australia and Netherlands.

10.   There was no draft prepared by the METIS Task Force for this part of the Common Metadata Framework, and the discussion was based on contributions by national statistical offices. The goal of the session was to:

(a)   build up a body of material to assist in drafting Part C of the Framework;

(b)   discuss and agree on structure of this part of the Framework;

    (c)   obtain input, eg examples of implementations;

    (d)   gather content related to architectural options, system design and development;

    (e)   identify helpers for future work.

11.   The inputs contained in invited and supporting papers dealt with the roles that metadata plays in the statistical cycle:

    (a)   Central metadata repositories in support of the survey life-cycle for the whole set of statistical surveys (including administrative collections);

    (b)   Use of knowledge bases and document management systems;

    (c)   Automation in questionnaire design and more generally;

    (d)   Quality assurance, particularly automated editing and imputation;

    (e)   Machine-to machine communication;

    (f)   Delivery of quality related information to end users;

    (g)   Capturing and updating of metadata, registration and use of workflow engines in support of registration processes;

    (h)   Versioning: changes to mandates and characteristics of statistical surveys.

12.   The following issues were raised during the discussion:

    (a)   Possibility to store questions and their specifications in a database, with a view to automate their inclusion in various questionnaires. There is evidence of an overlap of questions among various forms and questionnaires, and such a centralised database might optimise the design process.

    (b)   Some offices concentrate help files from distributed statistical systems into a single knowledge base. The example referred to at the meeting used the Lotus Notes database coupled with a documents management system.

    (c)   The relations between terms "survey" and "questionnaire" or "statistical observation". One survey usually covers one statistical area, for example wholesale trade, retail trade, etc., and may use several different questionnaires that are adjusted to the type of the business that is surveyed. In this connection there is a need to clarify the terminology. The classic way makes a distinction between statistical surveys (that collect data exclusively for statistical purposes) and the use of administrative sources, while some statisticians cover both by the term "survey". One possibility may be to use the term "statistical operation", "statistical activity" or simply "collection". Also need to cover operations such as accounts compilation.

    (d)   Centralized metadata repositories need a clear definition of responsibilities. One model is that the subject-matter units are responsible for the content, while the metadata team ensures coordination and consistency/quality checks. It is also important to synchronise the schedule of update of the metadata repository with the release calendar of statistical data.

    (e)   Partnerships between statistical offices, with a view to development of metadata systems, may be preferred to outsourcing development to the private sector. The disadvantage of outsourcing can be that private enterprises rarely have a sufficient knowledge and understanding of the environment to grasp the metadata issues.

    (f)   Linking the statistical software used in various phases of surveys with metadata repositories.

    (g)   Workflows are used in some offices for production as well as in management

information systems. One of the applications is to set the timeliness based on the target release date and the expected workflow. Another application that might be implemented on the basis of workflow management is Customer Relationship Management (CRM).

(h)   Quality information relates to both the quality of the process and metadata, and the data quality. The latter is probably more important for end-users.

(i)   Discussion about costs, which were mentioned during this session, will be added to the framework as part of the review of Part A.

13.   The original plan was to organize the material collected for Part C under the following headings:

C1.   survey planning and design;
C2.   survey preparation;
C3.   data collection;
C4.   input processing;
C5.   derivation, estimation, aggregation;
C6.   analysis;
C7.   dissemination;
C8.   post survey evaluation.

14.   An alternate method might be to structure Part C primarily by cross-cutting issues, like versioning, linking, etc., and use the phases of statistical survey cycle as a secondary division. Both methods had their support among the participants. As a conclusion, the participants agreed to re-organize the structure of Part C. A possible new structure is:

C1.   Positioning SMS in the Statistical Cycle- model of processes and metadata types relevant to each;
C2.   Definitions - cover what is meant by 'survey'- what terminology to use to handle administrative data sources, account compilation etc;
C3.   Understanding the cycle- using the current C1 to C8 list of processing phases, discuss the metadata used, metadata created etc;
C4.   Architectures, Business and System Design, and Integrating Processing Tools;
C5.   Organizational and cultural issues;
C6.   Links to useful case studies, previous METIS papers, other resources.

15.   The following non-exhaustive list will be considered by the Task Force on the Common Metadata Framework as potential content for Part C:

(a)   Explanation of what the statistical cycle is and what metadata are needed at each stage;
(b)   Organizational and cultural issues, case studies and complex methods to manage the metadata throughout the entire survey cycle;
(c)   Information that might relate to architecture and good practices in transfer of metadata from central repositories to statistical tools like Blaise, PC-AXIS, etc.;
(d)   Versioning, changes of metadata with time;
(e)   Information on statistical operations, defining the statistical survey, use of

administrative data and combination of survey data with administrative sources.

IV. IMPLEMENTATION

Session Organizer: Max Booleman (Netherlands)
Documentation: Invited papers by Netherlands, Sweden and OECD; supporting papers by Croatia, Eurostat and Norway.

16. The Common Metadata Framework will include a section (Part D) on the experiences of national statistical offices that have recently implemented or re-engineered their statistical metainformation systems (SMS). There was no draft prepared for this part of the Framework prior to the work session. Countries have been asked to contribute information about their implementation experiences and these papers will become available on the Internet as part of the Framework publication.

17. There were three invited and three supporting papers for this session. Common themes from the presentations and discussion included:

    (a)   Avoiding redundancy and reusing metadata wherever possible. There was agreement that metadata should be updated in one place, but be available throughout the organization. There is an interest in sharing appropriate metadata between organizations, especially the international standards;

    (b)   The development of 'generators' to automatically run metadata from one source to another and systems that force creation of metadata before any data entry occurs;

    (c)   Development of process generators with the help of metadata;

    (d)   Strategies to transform metadata between systems, either to the next generation system or to a central (e.g. international) one;

    (e)   Frameworks for keeping track of the history of particular data, such as the link of the associated subject matter expert to concepts;

    (f)   Issues that develop following the implementation of a metadata system, such as usability, change management, increased efficiency and the results of easy access to information on the Internet;

    (g)   Any 'tips and tricks' that were raised during discussion about Parts A and C will be incorporated into Part D of the Framework.

V. OPEN DISCUSSION

18. The open discussion session aimed to identify the most important points from the four substantive topics discussed during the first two days of the meeting. The following issues were highlighted:

    (a)   Involvement of the top management makes it necessary to clearly prove the usefulness of metainformation systems. It is important to convince managers of the benefits that corporate metadata management brings in terms of efficiency, reduction of costs and quality.

    (b)   Many of the statistical offices represented are reportedly in an advanced stage of development of corporate metainformation systems. It is important to be pragmatic, envisage evolving development, and avoid creating "empty

cathedrals".

(c)   It is important to know the metadata needs of methodologists and designers. Questionnaire design is a topic of particular interest, because it helps to improve the quality and consistency of data and reduces costs. A corporate system for questionnaires is important for re-using questions that appear in several forms/questionnaires, and for surveys that cut across different areas. The automated approach requires questionnaires within the statistical office to be standardized.

(d)   It is difficult to draw a boundary between metadata specialists and methodologists, and more generally between the statistical metadata and the subject-matter knowledge, or from again other angle, the boundary between the corporate and local metainformation.

(e)   Statistical metainformation systems have to be seen in a broader context, and as an integral part of statistical information systems.

(f)   It is important to look into the issue of globalization, e.g. treatment of supranational enterprises, when the information exceeds national boundaries. Synergies between national and international metainformation systems need to be addressed, particularly international comparability and issues related to globalization, for example, information on multinational enterprises.

VI.   RECOMMENDATIONS FOR FUTURE WORK

19.   The participants reviewed the recommendations for future work put forward by Task Force on Common Metadata Framework. Their proposal centred on completion and future maintenance of the Common Metadata Framework for National Statistical Offices. The following changes and additions were recommended:

General aspects:
(a)   The Framework shall be a living/evolving reference to standards, concepts, and best practices. An editorial board shall be created to ensure future maintenance of the Framework.

(b)   The Framework should be published in an electronic form on the UENCE website, with some parts published in print (after the public and editorial group review).

(c)   The Framework shall comprise two types of information:
        (i)   Text that is written for the Framework as its integral part;
        (ii)   Selected and edited contributions by countries (good practices).

(d)   It is important to add a general introduction, and provide "packaging" to link individual parts of the Framework.

(e)   In order to attract broader readership, each item of the framework should comprise a general overview, avoiding overly detailed technical information, with links to more technical and detailed papers.

(f)   Linked contributions may relate to multiple parts of the Framework.

(g)   All material included in the Framework should be accessible without restriction.

(h)   Drafts will be made available for public review. After two rounds of review they

will be left stable within the Framework for some (reasonable) time.
(i) The draft Framework shall be updated using contributions from METIS 2006.
(j) The editorial board should be open to all inputs within and outside the framework of METIS and ensure collaboration with other relevant groups.

Part A (Editors: Jana Meliskova; Graeme Oakley,Australia); Bo Sundgren, Sweden)**:**
(a) Immediate actions to be taken:
- Incorporate comments from METIS 2006
- Circulate for public review
(b) Possible additions to the content:
- Cost / benefits studies
- Quality considerations

Part B (Editors: Paul Johanis, Canada; Daniel Gillman, United States; Alice Born, Canada)**:**
(a) Immediate actions to be taken:
- Prepare for public review
- Ensure regular maintenance
(b) Possible addition:
- More studies on the context of standards and how they can be mapped/linked
(c) Related issues outside the framework of METIS:
- SDMX has put the MCV up for a public review until 31 May 2006.

Part C (Editor: Gary Dunnet, New Zealand)**:**
(a) Immediate actions to be taken (Graeme Oakley, Australia):
Create a new structure taking into account comments from METIS 2006
- Include a 3-5 pages introduction
(b) Possible addition:
- Functions of metadata
- Country case studies

Part D (Editor: Max Booleman, Netherlands)**:**
(a) Immediate actions to be taken:
- In clued 3-5 pages introduction
- Publish country case studies and good practices
(b) Possible addition:
- Managerial and organizational issues

20. The following deadlines were agreed:
5 May 2006   Comments from the public review of Part A
May 2006   Website prototype (by the UNECE)
June 2006   Public release of the first version of website
October 2006   Progress report to the Bureau of the CES
Spring 2008   METIS 2008

VII. ESTABLISHMENT OF AN ELECTRONIC DATA REPORTING TASK FORCE
Documentation: Terms of Reference and work plan for the Task Force

21.   Electronic data reporting for primary data collection was a topic for the seminar session at the 2005 plenary session of the Conference of European Statisticians. In concluding their discussions, the members of the Conference suggested to create a Task Force on Electronic Data Reporting for Primary Data Collection. The Bureau of the Conference has considered the proposal and agreed on Terms of Reference and work plan.

22.   The themes identified for the Task Force on Electronic Data Reporting for Primary Data Collection (EDR) are:
   (a)   Content:
   * Metadata-driven generalised system for EDR;
   * common data taxonomies, including their relevance
   * electronic questionnaire design;
   (b)   Techniques:
   * multi-mode surveys with EDR as an option;
   * promoting the EDR response option;
   * XBRL
   * "pull" techniques;
   * Internet collection for population censuses;
   * security arrangements and their public communication;
   * open-source EDR and case management solutions;
   (c)   Organization:
   * linking business reporting and statistical data collection;
   * participating in a Government general approach to EDR;
   (d)   Other themes that the Task Force may find important in effort to progress the implementation of EDR

23.   The following points were made during the discussion:
   (a)   Metadata is one of the key roles in the EDR. Some of these topics are already covered by the ongoing work on the Common Metadata Framework: they just need to be profiled in order to put the EDR and metadata issues together.
   (b)   There may be benefits from general exchange of experience, but the EDR Task Force should try to concentrate on some specific issues and topics within the overall mandate. It was suggested the focus should be on the issues for which recommendations or expert input are needed.
   (c)   The EDR Task Force will have to define concrete outputs/deliverables. It is understood that the goal of the Task Force is to bring quickly tangible results and accelerate the implementations of electronic raw data collection.
   (d)   It should be taken account that the development level of EDR varies from country to country. Some countries may cope with EDR immediately, but some may not.
   (e)   Participants mentioned existing activities that should be taken into account by the EDR Task Force. For example, the US agencies are using CASIC – Computer Assisted Information Collection, which combines various EDR techniques. The UNECE group working on recommendations for population censuses could

provide input on e-Census. The MSIS work session could include a topic to explore technologies such as XFORMs. The next METIS meeting could have a focus on experiences related to metadata and the statistical cycle stage, in particular design and data capture.

(f)   Eurostat will cooperate with the EDR Task Force, and intends to encourage members of its CORD (Coordination of Raw Data Collection) Task Force to participate.

(g)   The representatives of national statistical offices will find out whether there is an interest within their offices and facilitate the contact with the UNECE Secretariat.

(h)   The idea of organizing a Work Session on EDR (UNECE jointly with Eurostat) in October/November 2006 was welcomed. It is important to issue the invitations early, in order to allow national statistical offices to prepare their participation.

* * * *

Meeting Evaluation
Joint UNECE / Eurostat / OECD Work Session on Statistical Metadata
 Geneva 3-5 April 2006

Number of participants: 85
Responses received: 58 (68%)
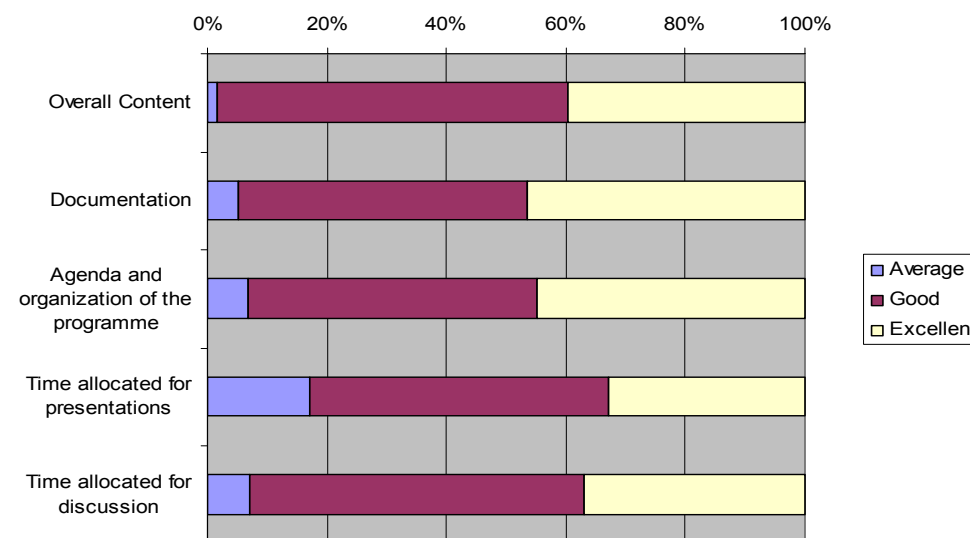
Quality of the meeting
Participants were asked to rate the following characteristics of the meeting between 1 (unsatisfactory) and 5 (excellent):

1. Overall Content
2. Documentation
3. Agenda and organization of the programme
4. Time allocated for presentations
5. Time allocated for discussion

92% of the ratings given were either "excellent" or "good", and nothing was rated below average. The time allocated to presentations was the least favourably rated, with 17% of participants considering it "average". 47% of the respondents found the documentation "excellent" and 98% rated the overall content above average.

General comments about the quality of the meeting were:
- Suggestion: use more time for the presentations. Two days is too short, while programme on 3rd day was "light".
- First 2 days very full programme; a lot of presentations. 3rd day too much time planned for the subjects. Better planning possible.
- More breaks please.
- 100% of presentations on the programme were presented! Well organized.
- Too much time allocated to discussion at the end.
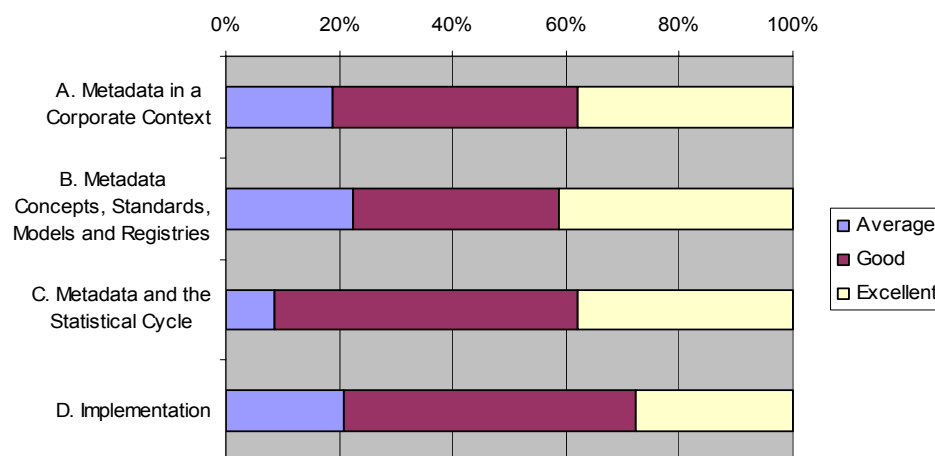- Very good work session, comparatively with METIS 2004.

Rating of Topics

Using the same scale, participants rated the usefulness of each topic for their work. Across all topics, 66% of the ratings were above average and none of the topics were considered below average.

The session on Part C, Metadata and the Statistical Cycle, was found to be the most useful topic. It was closely followed by Part A, B then D.

Four general comments about the topics were given:

- Very useful meeting.
- Shareability of work done should be given - at least as an overview.
- Unlike the Chair, I think the METIS Manual chapter on the statistical cycle should be structured according to the processing steps "everybody knows".
- Extremely useful.

Ideas for Improvement - 22 respondents provided suggestions on how the METIS work session could be improved.

Content

- I was curious about the stage each country has reached, concerning their metadata organization, registrying and exchanging with other agencies/organizations. Maybe this could be shown in a prior presentation of UNECE/Eurostat/OECD for the next meeting. Thank you!
- What is the actual metadata needed for E2E survey cycle - beyond steps of survey cycle.
- Would be great to have more papers on practical implementations of metadata related projects.
- Include a session as workshop, eg "live" demonstration of a country's system.
- Improve the diversity of contributions and comments (not more than 50% by the "key-players").
- My opinion is to include in some seminar the NSO which have a small experience in creation and improvements of metainformation systems. Because it will be useful to develop SMS for NSO who need advice, recommendations, changes in experience.
- The focus on concrete deliverables made this a very productive meeting as compared to past METIS meetings. Keep this focus in future.
- Exchange of experience is important. Standardization and its implementation.
- In fact, a statistical metadata system is a key component in a modern statistical production system. But the main problem is harmonization databases. We need more info about it.
- We need to look at common areas and see what can be shared. Willingness to share needs to be indicated.

Organization

- Draft a more realistic agenda: this time, it was clear that 3rd day was "empty". If sessions don't cover the morning of the last day, it is better to say that, or else extend the time for presentation of papers, or insert a slot for demo and free time for "tour de table"
- Better planning available time vs agenda. First 2 days very full and 3rd day left a lot of time.
- Allocate a slot for each country.
- Use more of 3rd day for content/topic discussion. Finishing half a day early is a waste.
- Combine with workshops (in parallel)
- Greater participation from all (especially smaller nations).
- Have more specific topics - the ones selected this year were too long. Metadata terminology is hard to understand without examples as everyone uses them in a different way. Presentations should have more examples and less diagrams (the one with boxes and arrows everywhere) which are very hard to understand in a PowerPoint presentation.
- The implementation topics should be given enough time, possible in the morning as well.
- More time: 26 presentations in 2 days probably exceeds most people's absorption capacity. Perhaps could have taken time for presentations from that allocated to discussion on how to develop framework documents.
- Less documents unless available earlier by at least 2 weeks.
- To increase the range of topics.
- Keep up the good work.

Further developing the Common Metadata Framework - Participants were asked to give their ideas on how to further develop the Framework:

Part A	Part B	Part C	Part D
- I would like a poster to present to top management and colleagues. It should be simple and succinct. - Develop ranges which part of corporate model is covered by metadata model. - Include crisp records of benefits derived by NSOs through implementation of SMS.	- Concrete examples, distinction and overlapping of models. - Show clearly how the various standards (1179, CMR, DDI, SDMX) can be used at each stage of the lifecycle and how they can interoperate. - Explicit mapping between standards.	- Document which (rough) parts of the model cover what part of the cycle. - I suggest that the main processes should be defined as a combination of roles, processes and artefacts. I observed that people focussed on processes and results (= artefacts) but I missed the role-aspect, which is really useful when you have to define responsibilities. - Probably the most fruitful area. - Detail, eg of metadata limited to statistical cycle. - Less focus on technological issues.	- Examples of successful and best practices. - Relate the implementation process to a maturity model (eg look at CMMI with 5 maturity levels), that will be helpful in convincing management and in benchmarking statistical organizations in relation to the framework. - Several case studies on metadata implementation. - Less focus on technological issues.

Overall improvements to the Framework

- No suggestions.
- Send emails and documents to Secretariat and copy to those responsible for further work from each session…
- Microdata archiving in connection with research requirements.
- Implement a dashboard or "wiki" for the framework and contributed ideas and comments.
- Elaborate on procedure to develop a SDMX - Key Family.
- All topics give a very good metadata framework for us.
- Invite more statistical agencies.
- I think this is an evaluation of ideas and has to progress as we evolve. Note comments during meeting on how metadata was seen at early METIS conferences. Most organizations have to foster the concept of metadata and its role inside their organizations. This has to be done mostly by linking metadata to organisational imperatives, and to other similar cross-cutting activities, or to development of an Enterprise Architecture. Therefore I have limited enthusiasm the notion that we can get the metadata picture "right" in some sense. I expect at METIS 2008 we will have somewhat different ("evolved") perspective.
- Executive Summary followed by annexes containing detailed information including more examples underlining the business case.
- To increase the part regarding implementation.

Future Work

80% of respondents found the suggestions for future work to be satisfactory. Of the remaining respondents, 10% were partly satisfied and 10% did not respond to this question. Seven people provided general comments:

- I am afraid this part of the agenda has been excessively compressed for closing of the meeting in advance!
- Is this just a "duplicate" storage area? Once off or to be maintained? Links to other relevant websites? Template for contributions? Don't underestimate the maintenance or usability. How will you respond to incorporate feedback with time… Cmat page for feedback?
- Note: I was partly absent on Wednesday morning and I lost this part of the discussion.
- More emphasis on process metadata and quality metadata.
- Although it is difficult to make concrete agreements it would be good to be more precise what is asked from the participants.
- Standardized metadata content is essential to move this process forward; exchange within each organization is fine; but exchange between organizations is the next step.
- Comparing different standards eg 11179 with 19100 series.

UNECE preparation and servicing of the meeting

95% of respondents thought the UNECE Secretariat adequately prepared and serviced the meeting. 4% were partially satisfied and the remaining 1% did not respond to this question. General comments were:

- Excellent
- It would be nice to have water in the conference room.
- Thank you very much!
- Signs were good! Coffee right outside the meeting room would be an improvement.
- Excellent preparation and service. Thank you.
- Wireless internet access was a great idea.
- Very well done and well organized.
- Great job!
- The wireless network was great. Thanks. It would have been good if presentations could have been posted on the website during the conference. Also better access to power for computers would be good. Swiss power outlets are unusual. Perhaps instead of providing lots of Swiss multi-outlet boards, you could provide boards from other areas, eg standard European, US, UK, Australia/NZ. National delegates would be happy to supply unusual power boards.
- Finally, easy to find the meeting room.
- It is a great pleasure to work with this friendly and professional staff.

General Comments

Four people provided general comments at the end of the evaluation form.

- The implementation or detail about metadata standard is partially discussed at the meeting. I think in order to understand the content of what metadata element is, there is a need to discuss statistical metadata standard.
- The chairman was OK :-)
- Documents should have running footers or running headers and a page number on every page.
- You do one of the best jobs in the statistical domain!