

**UNITED NATIONS
ECONOMIC COMMISSION FOR EUROPE
CONFERENCE OF EUROPEAN STATISTICIANS**

**EUROPEAN COMMISSION
STATISTICAL OFFICE OF THE
EUROPEAN UNION (EUROSTAT)**

**ORGANISATION FOR ECONOMIC COOPERATION
AND DEVELOPMENT (OECD)
STATISTICS DIRECTORATE**

Work Session on Statistical Metadata
(Geneva, Switzerland, 6-8 May 2013)

REPORT OF THE WORK SESSION

I. Introduction

1. The Joint UNECE/Eurostat/OECD work session on Statistical Metadata was held in Geneva, Switzerland, from 6-8 May 2013. It was attended in person by participants from Azerbaijan, Bulgaria, Canada, China, Denmark, Estonia, Finland, France, Hungary, Italy, Republic of Korea, Latvia, Mexico, Republic of Moldova, New Zealand, Norway, Portugal, Russian Federation, Serbia, Slovenia, Spain, Sweden, Switzerland, Turkey, the United Kingdom and the Ukraine. Representatives of the Bank for International Settlements (BIS), the Statistical Office of European Union (Eurostat), the Food and Agriculture Organization of the United Nations (FAO), the International Labour Organization (ILO), the International Monetary Fund (IMF), the Organisation for Economic Co-operation and Development (OECD), the United Nations Conference on Trade and Development (UNCTAD), the United Nations Educational, Scientific and Cultural Organization (UNESCO) and four independent experts (from Rapanea Consulting Limited, Metadata Technology UK, Metadata Technology North America, and the University of Michigan/DDI Alliance) also attended. Representatives of Australia, Canada, France, Italy, the Netherlands and the United States of America participated remotely via video link.
2. The work session considered the following substantive topics:
 - (i) Metadata standards and models;
 - (ii) Metadata in the statistical business process;
 - (iii) Case studies and tools.

Details of the presentations and discussions in each of these sessions can be found in the Annex to this report.

3. Work session participants elected Ms Alice Born (Canada) to chair the work session.
4. The work session was organized by a steering group consisting of Mr Klas Blomqvist (Sweden), Mr Max Booleman (Netherlands) (until February 2013), Ms Alice Born (Canada), Mr Trevor Fletcher (OECD), Mr Daniel Gillman (United States Bureau of Labor Statistics), Mr August Götzfried (Eurostat), Mr Alistair Hamilton (Australia), Mr Håkan Linden (Eurostat), Ms Jenny Linnerud (Norway), Ms Isabel Morgado (Portugal), Mr Juan Muñoz López (Mexico) and Mr Ebbo Petrikovits (Czech Republic), with secretariat support from Mr Steven Vale and Ms Fiona Willis-Núñez (UNECE). Members of the steering group acted as session organizers.

II. Recommendations for future work

5. Participants discussed ideas for future work on statistical metadata, standards and related topics to be undertaken within the coordinating framework of UNECE under the new Modernisation Committee on Standards (introduced during the work session). Proposals included those listed below, as well as others which are already included as work packages within the two projects of the High Level Group for Modernisation of Statistical Production and Services, and which are therefore not listed separately here.
 - Update Part B of the Common Metadata Framework (incorporate an inventory of standards)
 - Develop an official statistics profile for DDI and a common approach for software vendors (e.g. Colectica)
 - Explore the use of BPMN/BPEL for implementations based on GSBPM
 - High priority: make decisions on the future governance and harmonisation/consistency between GSIM and Neuchâtel
 - Update GSBPM/metadata case studies based on revised template
 - Further work on the relationships between GSIM and GSBPM
 - Development of process /technical metadata standards
 - Explore potential and implications of linked open data – relate to GSBPM etc.

III. Closing of the meeting

6. In her concluding remarks, the chair noted that the work session showed evidence of great progress and hard work in recent years, especially in the development of GSIM as a new reference model. The way in which it was developed is an example of the benefits of international cooperation. It is hoped that it will receive formal approval by the Conference of European Statisticians. Meanwhile, much work is being undertaken to refine it, and the work session has provided useful feedback for this exercise. The work session also illustrated the efforts to ensure strong links and coherence across the field, e.g. through the work on mapping standards to one-another. Country case studies demonstrate a movement towards sound metadata management in statistical organizations, which is a positive step following long periods in which this was more a goal than a reality. The chair stressed the importance of metadata experts within the evolving international work on modernisation, and encouraged all participants to become involved in the various initiatives discussed throughout the work session.

IV. Adoption of the report

7. The participants adopted the draft report before the work session adjourned.

V. Further information

8. Details of the specific topics considered in this work session are presented in the Annex. All background documents and presentations for the meeting are available on the website of the UNECE Statistical Division (<http://www.unece.org/stats/documents/2013.05.metis.html>).

Annex – Details of Papers and Discussions

1. The topics on the agenda of the meeting are described in more detail below:

(i) Metadata standards and models: Progress on Part B of the Common Metadata Framework

Session organizers:

Alice Born, Statistics Canada (alice.born@statcan.gc.ca); Alistair Hamilton, Australian Bureau of Statistics (alistair.hamilton@abs.gov.au); Juan Muñoz López, Instituto Nacional de Estadística, Geografía e Informática, Mexico (juan.munoz@inegi.org.mx).

2. A paper by the UNECE gave an outline of the recently-developed Generic Statistical Information Model (GSIM). It explained the rationale for developing such a model, and gave an overview of the nature of the model, which is a reference framework of information objects. It is deliberately designed to be a conceptual framework independent of any specific design or implementation platforms. Its potential benefits and uses include improved communication within and between organizations through the use of common terminology and clear structured representations of relationships between information objects throughout a statistical organization. The 'layered' nature of the GSIM and the sprint approach used for its accelerated development were described. It was announced that a review and refinement process would be launched at a later point in this Work Session.
3. A paper by an independent expert from Metadata Technology North America presented the preliminary findings of continuing work to map GSIM to SDMX and DDI in a collaborative process involving two virtual working groups (GSIM-SDMX and GSIM-DDI). The main purposes of this work are to determine mapping tables between the standards, and to identify gaps in each of them, which can in turn be fed back into the stewards of each of the standards. Findings will also be used to inform the refinement of GSIM itself. The presentation stressed some key principles, including that mappings should not be 'forced' as neither SDMX nor DDI can or should be expected to cover everything in GSIM (not least because they were developed prior to GSIM). Some examples of findings were given with reference to the business, concepts, production and structures groups of GSIM. Some areas are strongly supported by one or both, while some leave gaps that hopefully will be filled in due course. GSIM expected to provide a good basis for convergence of standards in the future.
4. A paper from Italy described the standards and methodology used by ISTAT to create and manage integrated and reusable metadata, tackling the issues of how to model structural metadata for the data dissemination phase, and how to define integrated and reusable classifications. The approach is based on both SDMX and GSIM. Since 2010, ISTAT has been using a dissemination system called dati.istat.it, with a single exit point and a unified metadata system. The approach aims to integrate and re-use metadata throughout the statistical production process, and particularly during the transition from micro data to macro data. The system was illustrated with an example of a classifications problem. Some potential enhancements for SDMX were proposed.
5. A paper by Eurostat introduced a current cross-cutting project on information models and standards, part of their Vision Infrastructure Programme (VIP), a transformation programme for the modernisation of the production systems in the European Statistical System. The aims of the project are to provide a set of agreed-upon standards to support the modernisation of statistical production, and to increase coherence between standards. The starting point was the inventory of metadata standards from Part B of the Common Metadata Framework. The focus is on the combination of DDI and SDMX based on analysis of use cases. To facilitate this, metadata should be stored and indexed in a repository so that they can be expressed either as SDMX or DDI as needed basis. A key message was that the strengths and weaknesses of SDMX and DDI can be capitalised upon by using them in combination— they should be viewed as allies rather than adversaries.

6. A paper from the DDI Alliance reported on the strategic priorities for the Alliance in 2013. It began by introducing the DDI, its components and its benefits. It then focused on priorities for the future including the working methods to be used to fulfil these objectives. Given the rapid changes in the 'data ecosystem', DDI intends to respond in several ways: restructuring the organization, developing the 'next generation DDI', and enhanced outreach to its target communities. The further development of DDI will be informed by work on GSIM (especially the GSIM-SDMX-DDI mapping work), and will be designed to better interact with other models and standards. Specific goals and new content areas were described. The Alliance will prioritise involvement in future collaborative work with NSIs.

7. Participants discussed these papers and more generally the topic of GSIM and its implementation using DDI and SDMX, and the relationships between these standards. Areas covered in the discussion are outlined below.
 - Participants expressed interest in the findings of the mappings from GSIM to DDI and SDMX
 - The modellers expect that the second plug & play sprint (Rome, June 2013) will contribute to a more in-depth mapping between the standards
 - It is anticipated that BPMN & BPEL will be the correct implementation standards for GSBPM
 - SDMX sponsors are working on guidelines and harmonisation of cross-domain codelists.
 - The standards organizations are interested in working on an expression/transformational language such as the one used by the Bank of Italy.
 - Both GSIM and DDI are designed to manage the change of concepts over time.
 - Concern was expressed that the user cases for the mapping to SDMX & DDI are strongly influenced by a plug & play architecture for which the standards were not originally designed
 - Eurostat is assisting areas such as national accounts, which have not previously had data and metadata exchange standards, to adopt SDMX.
 - Participants expressed concern about scope creep of the standards that could extend them beyond their useful purpose and make them technically more complicated. SDMX was initially designed for exchange and DDI was initially designed for data documentation.
 - The standards organizations reassured the participants that no attempt is being made to create one super-standard. Neither standard will or should cover 100% of the GSIM.
 - The GSIM so far is seen as rich for Concepts & Structures and potentially rich for Production, but Business information objects need more testing.
 - It is as important to identify any gaps in the GSIM as it is to identify gaps in DDI and SDMX.
 - From an Enterprise Architecture viewpoint SDMX and DDI could be seen as linking the applications layer to GSIM, in the information layer, and GSBPM, in the business layer.
 - Different audiences need different views of the GSIM.
 - It was stated that the GSIM - DDI and GSIM - SDMX mappings are to be completed by June 2013.

8. A paper from Canada and Norway on behalf of an international informal task force described the ongoing work of the task force in reviewing and revising the Neuchâtel Terminology Model for Classifications. An introduction to the original model was followed by an explanation of the rationale for revisiting it, and a description of the working methods used. The task force was composed of volunteers identified through METIS and the UN Expert Group on International Statistical Classifications, and is composed of members of nine NSIs and four international organizations. Countries known or thought to be implementing the Neuchâtel model were consulted by means of a questionnaire. Revisions under consideration include definition changes; object deletions; attribute deletions and additions; clarifications and additional examples; database restructuring; and mapping to GSIM. Some specific examples of proposed changes were presented, including the criteria for distinguishing between classification versions and variants and the definitions of these. The next steps will include settling on the way forward for each proposed change; finalising the revised model; completing the mapping work in consultation with the GSIM implementation group; and deciding upon the stewardship of the revised model for future maintenance.

9. A paper from Italy presented the approach taken at ISTAT to the implementation of the ESS standards for reference metadata and quality reporting. The presentation described ISTAT's project to fulfil the ESS and ESQRS standards, using and enhancing a reference metadata system called SIDI/SIQual to create standard metadata output files which are in compliance with the SDMX technical standards. The architecture and contents of SIDI/SIQual were described. The project's aim is to enhance this system so that it enables fulfilment of the international reference metadata and quality reporting requirements. The design and implementation phases of this project were outlined and discussed. Also a proposal for the coordination of the implementation process is being defined to be then approved by the top management.
10. A paper by an independent expert from Metadata Technology UK described an approach to designing a metadata repository. The paper outlined the major issues facing a designer of such a repository, and suggestions for how these can be overcome, emphasising a standard generic model of metadata so that solutions proposed could be applied to the cases of different organizations and needs. The aim is to develop a repository for reference metadata as commonly used in quality frameworks, based on the SDMX metadata model. A number of design issues were presented and discussed, and it was emphasised that knowing user needs was important for informing the design. Development of a generic solution based on a standard (or multiple standards) is more difficult, but ultimately more valuable, than development of solutions specific to a given organization.
11. A paper developed jointly by representatives of the United States, France and the Food and Agriculture Organization of the United Nations (FAO) introduced the eXtended Knowledge Organization System (XKOS), a successor to and extension of the Simple Knowledge Organization System (SKOS). The rationale for its development was explained, having its basis in the increasing trend towards linked open data and the absence of appropriate standards to deal with this trend. It was developed through a series of DDI-RDF (resource description framework) workshops in Dagstuhl, Germany. The principles of linked open data were outlined and SKOS was introduced as the basis for explaining the development of XKOS. Some of the new classes and refined properties of XKOS were presented, and the system was demonstrated with reference to the French Classification of Activities. Next steps will include seeking feedback from the metadata community, submitting the proposal to W3C, publicizing and of course implementing it.
12. Participants discussed these papers and more generally the topic of metadata standards for classifications and metadata management. Discussion covered the following.
 - Participants felt that GSIM and the Neuchâtel model for classifications should be consistent.
 - Statistical organizations need a rich conceptual model to manage statistical classifications.
 - The stewardship of revisions of the Neuchâtel and GSIM models is not yet clear. A number of participants proposed that Neuchâtel terminology should be part of GSIM, while others suggested that Neuchâtel model should remain on its own. The discussion was to be continued on Tuesday as part of the session on the HLG Frameworks and Standards work packages, in particular on the GSIM revision work package.
 - Participants felt the joint efforts of both METIS metadata specialists and classification experts from the UN EG on classifications enhanced the Neuchâtel revision process, and should be retained in completing the revision.
 - Coding of data is important in any implementation in a statistical organization, but are not part of the Neuchâtel model. They are however a part of the GSIM.
 - Current work on a metadata repository may result in some proposed changes to SDMX.
 - Could/should XKOS be positioned as a DDI/W3C standard? Both would be good.
 - Both Italy and Denmark have quality and reference metadata stored in their own national systems and are endeavouring to output these in the standard structures required by Eurostat. Denmark commented that they use Colectica and that Colectica will include output to DDI.
 - Some participants wished to know how well GSIM covers quality frameworks and reference metadata. From the GSIM point of view a relevant quality measure could be attached to any information object. Quality is not viewed as an information object in its own right. Reference metadata can be identified through the mapping of GSIM to SDMX.

13. A paper by Australia described ongoing work on statistical and spatial frameworks and the use of geographic standards in statistical organizations, noting that the combination of geospatial data with other data is becoming increasingly commonplace, and emphasizing the potential for added value (e.g. in policymaking) that this trend can present. The paper discussed the emerging focus on standards for spatial data and metadata. A Global Geospatial Information Management (GGIM) Initiative and United Nations Statistical Commission Programme Review are facilitating progress towards an international framework for statistical and geospatial standards. A national example from Australia was presented, and some questions were posed to participants to provoke ideas about future work in this area.
14. A paper from Mexico presented an overview of the use of geo-referenced statistical data in the Mexican statistical and geographic office (INEGI). It began with an explanation of the meaning of geo-referencing and the ways that it can be done. Some examples were given of systems developed by INEGI: Mexico in Figures, the National Statistical Directory of Business Units, and the National Inventory of Dwellings. It was emphasised that standards are needed to ensure the accuracy and coherence of an information infrastructure for geo-referencing statistics. The rise in geo-referenced statistics will be valuable for all users from governments to individuals.
15. Participants discussed these papers and the use of geospatial standards in statistical offices in general. Session organizers posed specific questions about the use of geospatial standards in participants' organizations, with a view to informing work that is about to start as one of the work packages of the High Level Group (HLG) project on Statistical Frameworks and Standards. Participants were invited to indicate their own willingness to take part in this work package, or to pass on the invitation to contribute to their colleagues. Specific areas discussed were as follows.
 - In Mexico, geospatial visualisation of non-sensitive business information by NSIs is welcomed by businesses that cannot afford this infrastructure themselves.
 - Use of geospatial standards requires close cooperation between NSIs and geospatial agencies.
 - Australia, Mexico, New Zealand, Canada, Sweden & Portugal are engaging with geospatial communities
 - NSIs could learn from the geospatial community how to attract the interest of commercial companies.
 - NSIs need to cope e.g. by mesh blocks with frequent changes of administrative boundaries.

(iii) Case studies and tools

Session organizers:

Jenny Linnerud, Statistics Norway (jenny.linnerud@ssb.no); Klas Blomqvist, Statistics Sweden (klas.blomqvist@scb.se).

16. A paper from Estonia presented the approach taken at Statistics Estonia to metadata management and the statistical business process. The paper began by describing the metadata-driven integrated statistical information system (iMeta) and the business process model used in the organization. It explained that the system is based upon the Neuchâtel models for classifications and for terminology and on MMX. It detailed the contents of the system and its general architecture, and showed some examples from the user interface. It then outlined current projects— these are aimed at implementing ESS standards for describing statistical activities and disseminating data; and at the technical development of iMeta. The paper outlined some future plans: by July 2013 it is planned that ESMS-based reference metadata will be disseminated online; it is planned that an output database for concepts, definitions and methodology will be released in the future on the Statistics Estonia website; and further developments of iMeta are proposed. The paper concluded with some lessons learned: the involvement of potential internal users of any metadata system in the development of that system is vital; appointing a unit to be responsible for metadata management is also necessary; support from management is crucial for success since the work is labour-intensive.
17. A presentation by New Zealand described the metadata infrastructure and standardisation at Statistics New Zealand as part of the 'Statistics 2020' Transformation Programme. Following the format of the UNECE template for metadata case studies, the presentation outlined the organization's approach, followed by next steps, challenges and lessons learned. The transformation programme is a ten-year undertaking involving three phases and over 160 projects. The vision for metadata infrastructure is one in which information flows efficiently through the statistical business process, enabling more agile statistical production. Exchange, discovery and re-use should be facilitated by the metadata system. The problems posed by the pre-existing situation were noted— the main 'pain point' is prior to dissemination. The key element of the solution is re-use: metadata should be re-used wherever possible. Amongst the lessons learned, it was noted that it is not always necessary for all actors to understand the details of the models being used (GSIM, SDMX, DDI)— messages need to be tailored to audiences according to their roles, and this goes hand-in-hand with the importance of 'selling the benefits' of change to those who are required to undertake new or additional tasks, even if they are not the immediate and direct beneficiaries.
18. A paper from Canada introduced an updated metadata case study from Statistics Canada, emphasising the benefits and challenges of implementing GSIM. The paper presented Statistics Canada's implementation of a metadata strategy based on GSIM, and within an integrated service oriented architecture (SOA) approach. Many statistical production and infrastructure services use metadata, so the challenge is to make the necessary metadata available, and exchange metadata between services in a standardised way. A canonical approach to information models, based on SOA, replaces the more traditional point-to-point approach. A proof of concept is under development using DDI. Future work includes determining how to best implement GSIM, and manage any future changes to that model, as well as investigating the possibility of an event-driven SOA approach.
19. A presentation from Bulgaria was given in the form of a case study following the UNECE metadata case studies format. A longer version of the case study will be made available on the meeting website since this one was abbreviated to fit into the allotted time of the meeting agenda. The presentation first gave a general overview of the approach to metadata in the Bulgarian NSI, followed by a section on their classifications system which is based on the Neuchâtel terminology model. The Bulgarian office follows the ESMS requirements for reference metadata as well as the IMF's SDDS format. Internally, the office has an intranet which contains systems for classifications, a metadata common vocabulary and an electronic glossary of statistical terms. It was noted that Bulgaria does not yet have a common metadata strategy, model or system. The goals, scope and technological considerations for the development of a classifications system were presented, as well as planned

future activities including involvement in international collaborative efforts; development of an SDMX 2.1-compliant dissemination database; and initiatives surrounding culture change to encourage support for SDMX and standardization in general in the Bulgarian statistical office.

20. A paper from Spain reported on the situation with regard to metadata at the Spanish National Statistical Institute (INE). INE has been involved in Metis and has been working in the area of metadata for ten years. During this time the office has moved from a situation in which metadata were treated as something 'outside the process' to one in which metadata is an integral part of the statistical production process which in turn is based upon GSBPM. Their integrated metadata system (SIM) promotes efficiency, re-use and improved governance. The presentation demonstrated some examples of the reference, structural and process metadata system in action on the INE website and explained the advantages offered by the integrated system. The change from vertical to thematic information system has been a positive one. In future there will be a focus on dissemination of metadata, and inclusion of administrative sources in the metadata system.
21. A second paper from New Zealand, in collaboration with Metadata Technology North America, concerned the implementation of a new classification management system as part of the Statistics 2020 programme. It will be a replacement for the legacy Classification and Related Standards (CARS) system. It will be a concepts-based system, allowing more complex relationships. It also includes more automated authorisation routines, better search capabilities and greater re-use of objects. The conceptual model is based on the Neuchâtel model, SKOS and GSIM. It is designed to be reusable by others, and is comprised of five packages; core, classifications, coding, conceptual and concordances.
22. Participants discussed the findings of the case studies presented, and their own national experiences. Matters discussed included the following:
 - New Zealand stated that Linked Open Data (LOD) is part of the future roadmap of their new metadata program. The portal is similar to the UK open portal, and will improve metadata across government.
 - Some countries, including Spain, have integrated Eurostat's ESMS and ESQRS into their metadata systems
 - There was a question on the use of business modelling standards for modelling in developing the classification management system, such as BPMN and BPEL. It was stated that it was necessary to use these standards for building the system.
 - Regarding challenges in metadata management, the following is required: easy use of metadata standards, users need to see the concepts, allow users to create their own versions on the fly. The Classification Management Systems (CMS) is only part of the metadata management system; CMS should also be integrated into a wider context – variables, and how the classifications are used in surveys.
 - In response to these comments, New Zealand and the developer stated that there is a new approach to building the CMS based on both the concepts and the classifications; there are plans to integrate with a broader metadata system – linked more closely with survey systems; NZ's approach is based on federated metadata repositories (i.e., question library, classifications, and survey)
 - There was a question to Estonia on whether their MNX system could be mapped to GSIM.
 - It was acknowledged that NZ's CMS will be able to handle all types of changes in classifications (i.e., merge of classes, one-to-many relationships)
 - Denmark was pleased that New Zealand was using Colectica for managing metadata.
 - There was a question of using and mapping metadata to a DDI format. Canada is working on GSIM-DDI mapping, and has completed a ISO/IEC 11179 to DDI mapping tool with consultant – value domains, value meanings, classification option model. Statistics Canada is also using GSIM directly for building conceptual models. In New Zealand's case: DDI-GSIM in Colectica; CMS will be based on metadata standards and Colectica has agreed to modify the Colectica model; integrating classifications into Colectica: CMS will hold the classifications, and Colectica

will not be the classification store since Colectica does not have the functionality needed for managing classifications.

- New case study template: No objections to the template revision and it was accepted and approved by the METIS participants.
23. A paper from the Bank for International Settlements (BIS) presented an innovative approach to building a data portal based on SDMX. Users increasingly want a single access point for data, whilst providers need standard formats. The project created a series of "sandbox" environments, one for the BIS, and one each for fourteen national central banks. Each sandbox contains a registry, a database and a SDMX web service. A data portal was created, synchronised with the sandboxes, to provide a single access point. A video demonstrated these tools in action. This project has shown building blocks for the SDMX vision working in practice, connecting different data sources, and proving that the SDMX vision is real.
24. A paper by Switzerland discussed the use of SDMX in stocking, disseminating and managing a Swiss economic classification in the Swiss Federal Statistical Office (FSO). FSO now uses a new statistical metadata system (SMS), including an SDMX registry, to maintain this classification. This has required some changes for the classification unit and the implications of these change was elaborated upon. The experience of using the SMS as a whole was good and the system met the needs of the situation. The challenge for the office was not the technical implementation but addressing conceptual questions of translation and mapping of classification concepts into the SMDX-language.
25. Participants discussed these papers, and more generally the use of SDMX in statistical organizations. Matters covered included the following.
- Different approaches to implementing SDMX were discussed, however, what steps do we need to go forward?
 - Can SDMX support classification management?
 - Concepts, classifications, codelists and categories – DSDs
 - Harmonization, content – technology
 - It was stated that SDMX and DDI were never designed for classification management (not a use case) and Neuchâtel model covers this.
 - Eurostat stated that 2/3 of the job in creating SDMX DSDs and MSDs, etc. is determining the content and 1/3 is technical 1/3.
 - It was recognized that classifications are the backbone of the organizations.
 - There were concerns that we were developing/using too many models for classifications – Neuchatel, XKOS, GSIM, NZ CMS. The co-organizer proposed that they have different uses and should be integrated. This approach is not “all bad” – the differences are between conceptual and implementation models and standards. However, they should be harmonized with GSIM conceptual model.
26. A session on the High-Level Group for Modernisation of Statistical Production and Services (HLG) introduced the origins, recent changes, past, present and future work of this group; its relationships to Metis; and the plans for governance which will impact upon the future organization and activities of Metis. A modernization committee on standards will oversee the activities related to metadata. The implications of the new HLG governance structure for Metis were discussed by participants.
27. Presentations were given to introduce the two 2013 projects of the HLG: Common Statistical Production Architecture ('Plug and Play') and Frameworks and Standards. The work packages and working methods of each of these projects were outlined, and participants were invited to become involved in any of the areas presented, or to inform colleagues of the call for involvement. The future-focused nature of the projects was noted and participants expressed support for both of the projects.
28. Finally a process for reviewing and revising both the GSBPM and GSIM was launched. It was stressed that stability of the models (especially GSBPM, being more well-established) is a major priority and hence a strong case will have to be made for any changes. Changes to documentation are

more likely to be incorporated than changes to the principal GSBPM phase/sub-phase diagram, for example. Participants were informed that feedback may be submitted via the UNECE wikis or by email to the secretariat. A methodology and timeline for considering and implementing proposed changes was presented—the task will be complete by the end of 2013. Participants noted that it is important that a good business case is needed before making major changes to the models, as stability in the models is very important, especially to those already heavily invested in implementation.

(ii) Metadata in the statistical business process

Session organizers:

August Götzfried, Eurostat (august.goetzfried@ec.europa.eu); Isabel Morgado, Statistics Portugal (isabel.morgado@ine.pt).

29. A paper from Australia described the metadata-driven business process in the Australian Bureau of Statistics. It explained why it is important for processes to be metadata-driven: benefits include reduced cost, improved quality and timeliness, enhanced efficiency and agility. 'Metadata-driven business process' was defined as 'the systematic and consistent use of metadata to determine the inputs, outputs and behaviour of a statistical business process'. For processes to be considered truly metadata-driven, metadata should be used systematically, consistently and actively. The paper outlined some of the work done by ABS in this area, starting in 2002, and continuing to the current Information Management Transformation Programme (IMTP) and 'ABS 2017' programme. Two case study examples were given: Metadata Registry and Repository, Statistical Workflow Management Systems (MRR/SWM) and eforms. Key messages from the presentation were: metadata-driven business processes present tangible business benefits; metadata-driven business process is consistent with the vision of the HLG; metadata driven business process may require significant changes to the way metadata is managed in an NSI. The paper highlighted the need to 'evangelise' to process owners to convince them of the importance of metadata-driven business processes.
30. Participants discussed the topics raised in this paper:
- It was noted that it will take a long time to complete process re-engineering. This was acknowledged and ABS have created a dedicated team to work on this. Advantages which will help to achieve this goal include support from top management, having the governance structures in place to achieve the goals and being able to leverage the work of international collaboration groups (such as METIS).
 - SDMX is being used at the dissemination phase of GSBPM, and DDI for specification and creation of infrastructures and editing. The goal is to use both SDMX and DDI throughout the full GSBPM.
31. A paper from Portugal on behalf of an informal task force on metadata flows in the GSBPM presented the outcomes of work undertaken by the task force. The goal was to identify and define the flows of information objects between the sub-phases of the GSBPM. The presentation described the origins of the task force and the methodology it has used to develop its outputs. The task force was formed after the release of GSIM 1.0, and decided to use 21 information objects as identified in the communication layer of GSIM, to explore relationship between GSBPM and GSIM. It aimed not only to identify the relationships but to make them understandable to target audiences, primarily managers. A by-product of the work was that it contributed to the review of both GSBPM and GSIM, by identifying gaps, areas where clarification is needed, etc. Some examples of the concrete output were presented in detail: the full work will be made available on the UNECE wiki and on the website of this meeting. Conclusions included the realisation that terminology can cause problems if understood differently by different users; working methods make a difference to the nature of outputs— a 'sprint' approach might be more conducive to rapid production of quality outputs for similar work in the future; refinements of GSIM might include a higher-level layer for non-specialists; and definitions of some of the GSIM information objects could be revised to take into account the relationships identified in this work. Ideas for future work were proposed.
32. Participants were invited to discuss the findings of the previous paper, and to consider some questions posed regarding the work produced; its level of understandability; and the future scope of the work. Discussion expanded to cover views and feedback on GSIM in general. Issues raised included the following:

- It is useful to show how combining GSIM with GSBPM in real use cases can provide understanding and illustrate the HLG vision for modernisation. This should be used to drive future work.
 - It is important to show how GSIM and GSBPM can be transformed in practical cases.
 - It was suggested that higher level GSIM information objects might be needed for a better understanding of the model.
 - Guidelines for use of GSIM at a national and international level would be useful
 - There is a need to identify which information objects are most important from a business user point of view, and to test and validate information objects with different target audiences
 - Some GSIM information objects and terminology need to be clarified, and some NSIs felt that it would be useful if they could specify their own information objects—although this was subject to debate
 - A key to success is to get the business users to understand the information objects and business processes at an organization level
 - NSIs risk losing control of input data – requirements differ between NSIs and International Organizations. This reinforces the importance of the role of metadata to give meaning to inputs.
 - Having organizations develop complete end-to-end processes would be possible but would make it difficult to reuse information between processes.
33. A paper by Eurostat introduced the Euro process metadata structure (EPMS), whose aim is to standardise, integrate and document the statistical business process in the European Statistical System (ESS). The links between EPMS and the ESS standards for reference metadata and for quality reports were explained. While the latter two are standards for national statistical offices, EPMS is a Eurostat-only standard. The EPMS details the statistical concepts related to data processing, data validation, dissemination and IT applications. It is envisaged that its use will facilitate the streamlining of processes, for example by reducing the duplication of data validation work in national offices and then in Eurostat; offering opportunities for shared services; and permitting the rationalisation of IT applications. To further enhance process integration with Member States, a standard process metadata structure for national statistical business processes will also need to be defined.
34. A paper by the IMF reported on a pilot project to use the GSBPM and GSIM at the IMF. Their statistics department intends to use both when upgrading or re-engineering processes, as part of its "streamline, standardize, automate" programme. The pilot involved identification of the sub-processes and information objects used in the department's Integrated Monetary Database (IMD). The presentation outlined how the project was conducted; its findings; how it is intended that the results will be used to improve efficiency in collection, processing and dissemination. It was found that use of the GSBPM would help to standardise documentation across products and make it easier to compare processes, increasing transparency and re-use. GSIM presented more difficulties, and the attempt at mapping to the IMD used only ten of the 150 GSIM objects. The paper concluded with some proposals for changes to the models, especially GSIM, which it was felt does not yet meet all its objectives. GSIM would better achieve its objectives if it were restructured to have a less complex layer for communication, education and collaboration purposes.
35. A paper from the ILO described the ILOSTAT's new statistical information system, which is designed to be standards-based and metadata-driven. An all-encompassing new approach to data compilation entails the use of new and updated tools, as well as revised procedures for data collection, processing and dissemination. As the new system was designed, several types of metadata were identified. The paper outlined the relationships between each of these kinds of metadata and the GSBPM, explaining which kind or kinds of metadata arise from each GSBPM phase; it then examined how each is dealt with in ILOSTAT.
36. Participants discussed metadata in the statistical business processes of national and international organizations. Matters discussed included the following:

- In Eurostat the re-engineering of statistical processes is carried out when either data structures or IT systems change. This provides the opportunity to standardise content (via SDMX) and technology (using corporate IT systems)
- There was discussion of the importance of a higher level of GSIM for communicating existing business processes to complete the “as is” picture. There is a need for a common language for all statistical organizations. A high-level overview of GSIM is needed to get buy-in and ensure progress.
- Eurostat explained that metadata structures will be defined for ESS but a timetable has not yet been defined.
- It was suggested that GSIM could be too complex for countries with less developed statistical systems but can provide a good framework for communication.
- Comments on the lack of understanding in terminologies within GSIM were noted and will be considered within the HLG project for refinement of the model.
- It was suggested that GSIM be adapted to have core and extended levels for more easier adoption by organizations
- FAO asked whether terminology problems could be addressed by common vocabularies as within the SDMX CWG
- DDI agreed that a DDI profile would be useful for NSIs in the data collection process and offered to provide assistance in setting this up.

37. In concluding the session, the session chair noted that:

- Metadata driven processes were seen to be organized in different ways in organizations
- A number of pre-conditions exist for metadata-driven processing. These include: standards, IT and governance. It is a long-term process.
- GSBPM can be seen as a successful and widely-used model. GSIM is emerging and needs further refinement.
- Work is needed to communicate GSIM to different audiences at different levels.
- There is a need to validate GSIM and GSBPM with business users on their own processes.