



Economic and Social Council

Distr.: General

5 April 2016

English only

Economic Commission for Europe

Conference of European Statisticians

Sixty-fourth plenary session

Paris, 27-29 April 2016

Item 4 of the provisional agenda

Work of the High-level Group for the Modernization of Official Statistics

Outcomes of the activities of the High-Level Group for the Modernisation of Official Statistics in 2015 and plans for 2016

Note by the High-Level Group

Summary

The note provides a summary of the results of the projects overseen by the High-level Group for the Modernisation of Official Statistics during 2015 and plans for 2016. The purpose of the Group is to guide strategic developments in the ways official statistics are produced.

The Conference is invited to review the work done and provide advice on future work.

I. Introduction

1. The High-level Group for the Modernisation of Official Statistics (HLG-MOS) was created in 2010 by the Bureau of the Conference of European Statisticians. It comprises the Heads of eleven national and international statistical organizations, and has a mandate to reflect on and guide strategic developments in the ways in which official statistics are produced.
2. Each year HLG-MOS organizes a workshop, inviting representatives of various groups and projects related to modernization of official statistics. These workshops help to ensure coordination of activities. They also review progress and determine the key priorities for the following year. In 2014, the HLG workshop decided that further implementation of the Common Statistical Production Architecture and developing the use of Big Data sources for official statistics were the two highest priorities in 2015. The HLG launched international collaboration projects to address these priorities, which ran during the calendar year 2015. This paper summarises the main results of these projects, as well as other activities overseen by the HLG-MOS.
3. At the request of the United Nations Statistical Commission, the HLG-MOS also provides regular progress reports to that body. The 2016 report is available at: <http://unstats.un.org/unsd/statcom/47th-session/documents/2016-5-UNECE-HLG-on-modernisation-of-official-statistics-E.pdf>.

II. Common Statistical Production Architecture (CSPA)

4. The Common Statistical Production Architecture (CSPA) was created during 2013 in an international collaboration project under the HLG-MOS. It provides a standard framework for developing the components of statistical production, in a way that they can be much more easily shared within and between organizations than has previously been the case. It builds on key international standards developed under the HLG-MOS, such as the Generic Statistical business Process Model (GSBPM), and the Generic Statistical Information Model (GSIM).
5. The 2015 project focused on providing practical support to countries and international organizations implementing the CSPA. The final outputs of this project included:
 - CSPA v1.5 – An updated version of the CSPA specification, featuring enhanced guidance for implementers
 - The CSPA Logical Information Model v0.9 – Providing a bridge between the conceptual nature of the Generic Statistical Information Model (GSIM) and the practical requirements of CSPA service designers, builders and implementers
 - CSPA Global Artefact Catalogue (including the new Investment and Capability catalogues) – Tools to find information about what other organisations can offer, or are planning.
 - Updated high level information on CSPA – Brochures, presentations, a video, and other training material.
6. These outputs, together with a wide range of supporting materials, can be accessed via the CSPA wiki - <http://www1.unece.org/stat/platform/display/CSPA>.
7. A new “CSPA Implementation Group”, comprising representatives of 16 statistical organizations, was created in January 2016, to oversee and manage CSPA activities.

This group is responsible for the maintenance of CSPA and the CSPA Global Artefact Catalogue, assessment of candidate services, provision of implementation advice and further development of the CSPA LIM. Organizations wishing to be involved in this work should inform the UNECE Secretariat (support.stat@unece.org).

III. Big Data

8. This project brought together over 40 experts from national and international statistical organizations around the world to identify and tackle the main challenges of using Big Data sources for official statistics. It ran from January to December 2015, and focused on three main areas:

(a) **Experiments** - Task teams explored the potential of several Big Data sources to produce official statistics. This activity involved accessing data from the following sources, processing them and assessing their feasibility as sources for official statistics. Detailed results are reported for each source:

- Wikipedia page views
- Trade data from the UN ComTrade Database
- Social media data from Twitter
- Data scraped from enterprise websites

(b) **Sandbox** - The project team developed a sustainable model for the continuation of the “Sandbox” environment as a resource for the official statistics community, including identifying new use cases beyond Big Data.

(c) **Training and outreach** - Various events took place during the project. Workshops were held in Brussels in March 2015 and Ljubljana in September 2015. Presentations were given to various international fora, including the UN Conference on Big Data, the Eurostat Conference on New Techniques and Technologies in Statistics, and various UNECE expert meetings.

9. All outputs from the Big Data project are available at: <http://www1.unece.org/stat/platform/display/bigdata/2015+project>

IV. Other HLG-MOS activities in 2015

10. In addition to the projects outlined above, the HLG-MOS oversees the work of four modernisation committees and a Task Team on Communicating Modernisation. The key outputs of these groups in 2015 included:

- (a) Modernisation Committee on Organisational Frameworks and Evaluation
- An on-line set of Guidelines for Managers in Statistical Organizations: <http://www1.unece.org/stat/platform/display/GFM>
 - Generic competency profiles for Big Data teams, and Big Data team leaders: <http://www1.unece.org/stat/platform/display/bigdata/Competency+Profiles>
- (b) Modernisation Committee on Production and Methods
- A set of Generic Statistical Data Editing Models (GSDEMs): <http://www1.unece.org/stat/platform/display/kbase/GSDEMs>

- A study on the state of the art in the use of machine learning techniques in statistical organisations:
http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.50/2015/Topic3_Canada_paper.pdf

(c) Modernisation Committee on Products and Sources

- Inventory of Big Data projects in statistical organizations (note: at the request of the organizations involved, most projects are only visible to people with UNECE wiki accounts):
<http://www1.unece.org/stat/platform/display/BDI/UNECE+Big+Data+Inventory+Home>

- Surveys on outreach, promotion, branding and communication of official statistics, to support the work of the Task Force on the Value of Official Statistics

(d) Modernisation Committee on Standards

- A new Generic Activity Model for Statistical Organizations (GAMSO) - <http://www1.unece.org/stat/platform/display/GAMSO>
- Draft quality indicators for the phases and sub-processes of the Generic Statistical Business Process Model (GSBPM): <http://www1.unece.org/stat/platform/download/attachments/112133421/QI.pdf>

(e) Cross Cutting Task Team on Communicating Modernisation

- A new brand and logo for HLG-MOS activities:



- A Modernstats YouTube channel, featuring videos about modernization, presentations and webinars:
https://www.youtube.com/channel/UC1GxankkLEdqd_Hkj73VQYA

- Communication packs for key modernization standards and activities:
<http://www1.unece.org/stat/platform/display/hlgbas/HLG+in+brief>.

V. Projects for 2016

11. The two major international collaboration projects identified as key priorities for 2016 are:

- **Data integration** – Exploring how to better make use of a growing range of potential data sources, and integrate them with traditional sources to produce official statistics more efficiently - <http://www1.unece.org/stat/platform/display/hlgbas/Data+Integration>
- **Implementing Modernstats Standards** – Developing tools to apply the range of standards and models needed for statistical modernization, and to support statistical organizations at different stages in the modernization, including a Modernisation Maturity Model and a Modernisation Roadmap - <http://www1.unece.org/stat/platform/display/hlgbas/Implementing+Modernstats+Standards>

VI. Other on-going activities

12. In addition to the two projects, activities are underway in the following areas, overseen by the modernisation committees:

- Identifying best practices in risk and change management in statistical organizations;
- Identifying best practices in capability development and staff engagement;
- Overcoming organizational barriers to international collaboration;
- The use of mobile devices for collecting and disseminating official statistics;
- Sharing information on open data strategies;
- Developing a metadata / modernization glossary;
- Further developing quality indicators for the phases and sub-processes of the GSBPM, with a focus on non-survey data sources;
- Review of the Generic Activity Model for Statistical Organizations (GAMSO);

13. A new CSPA Implementation Group was created in early 2016, to oversee and support the implementation of the Common Statistical Production Architecture. This group currently has members from 17 national and international statistical organisations. It provides the mechanism to assess CSPA compliance of statistical software, maintain catalogues of software and other resources, and support practical implementation. This work is very closely aligned with the relevant projects within the European Statistical System “Vision 2020” programme, which has adopted CSPA as a core standard, and is currently developing new CSPA-compliant software tools for sharing between statistical organisations.

14. Any statistical organization wishing to get involved in any of the above activities is invited to contact the UNECE Secretariat (support.stat@unece.org).

VII. The sandbox and the statistical modernization community

15. As the work on modernizing official statistics matures, more national and international statistical organizations are looking for ways to align with the various HLG-MOS initiatives. Two opportunities for greater involvement at the organization level are now available:

- **The Sandbox** – This is a shared computing environment hosted by the Irish Centre for High-End Computing. It provides space to share (non-confidential) data, test new software tools, develop methods, and share training materials, in a collaborative environment. It was originally configured to support collaborative experiments on Big Data sources, but is now available for use for any other activities that would benefit from a shared computing environment. It is operated on a non-profit basis, supported by subscriptions of participating organisations, which jointly form the governance body overseeing future developments. For more information, please see the sandbox prospectus: <http://www1.unece.org/stat/platform/download/attachments/112133045/Sandbox%20Prospectus%20Final.pdf>
- **The Statistical Modernisation Community** – The HLG-MOS has recognized the interest of several organisations in becoming part of a community where the rights and obligations of partners are clearly specified, particularly in regards to

jointly developed outputs. This community is a collaborative partnership, open to all interested statistical organisations who publicly endorse a “Statement of Intent”. It was presented at a lunch-time side-event during the 2016 meeting of the United Nations Statistical Commission, where several countries expressed interest in joining. For more information, please see:

<http://www1.unece.org/stat/platform/display/smc>

VIII. Ensuring continuing relevance

16. A key topic discussed at the HLG-MOS meeting in New York in early March was how to ensure that the activities under the HLG-MOS remain as relevant as possible to statistical organisations. It is very unlikely that all activities will be priorities for all organisations. The challenge is to ensure that overall relevance remains as high as possible, and that activities continues to be strongly demand-driven. Different mechanisms are being put in place to ensure that this is the case:

- A survey of investment intentions and organisational priorities will be carried out during the second quarter of 2016. This will build on a pilot exercise during 2015, using a data collection tool developed by Statistics New Zealand. The results will be used to inform overall priorities, and to identify clusters of organisations with similar interests and priorities.
- A strategic planning framework will be created. This will include a list of priority areas identified in a top-down way by the HLG-MOS and its Executive Board. This will provide the basis for a call for proposals for potential international collaboration projects and activities in 2017 and beyond.

IX. Action requested from the Conference

17. The Conference is invited to:
- (a) Comment on the results achieved in 2015;
 - (b) Advise on future priorities;
 - (c) Advise on mechanisms to maximise the relevance of future HLG-MOS activities.
18. Delegates are also invited to express their interest in joining the Statistical Modernisation Community.
-