



# Economic and Social Council

Distr.: General  
14 December 2020

Original: English

---

## Statistical Commission

### Fifty-second session

1–3 and 5 March 2021

Item 4 (f) of the provisional agenda\*

**Items for information: common open standards for the exchange and sharing of data and metadata**

## Report of the Statistical Data and Metadata Exchange sponsors

### Note by the Secretary-General

In accordance with Economic and Social Council decision 2020/211 and past practices, the Secretary-General has the honour to transmit the report of the Statistical Data and Metadata Exchange (SDMX) sponsors. The report contains a summary of achievements, recent developments and planned activities by the initiative in areas such as the Sustainable Development Goals, the System of Environmental-Economic Accounting and macroeconomic statistics. An overview of capacity development in the area is also included. The Statistical Commission is invited to take note of the report.

---

\* E/CN.3/2021/1.



# Report of the Statistical Data and Metadata Exchange sponsors

## I. Introduction

1. In 2001, the Bank for International Settlements, the European Central Bank, Eurostat, the International Monetary Fund (IMF), the Organization for Economic Cooperation and Development (OECD) and the United Nations joined together to develop more efficient processes and standards for exchanging and sharing data and metadata. The World Bank joined the initial group of sponsor organizations in 2003. The commitment of the sponsor organizations to the Statistical Data and Metadata Exchange (SDMX) initiative was reinforced in 2007 through a memorandum of understanding, which established an SDMX secretariat and set the framework for the creation of two working groups: the Statistical Data and Metadata Exchange Technical Standards Working Group and the Statistical Data and Metadata Exchange Statistical Working Group.

2. The SDMX initiative is aimed at creating and maintaining technical and statistical standards and guidelines, together with an information technology architecture and tools, to be used and implemented by the SDMX sponsors and other organizations dealing with SDMX. Combined with modern information technology, the SDMX standards and guidelines should improve efficiency when statistical business processes are being managed. After more than 10 years of SDMX experience, the standards and guidelines are mature enough to be implemented broadly in various statistical domains and globally by statistical organizations.

3. The Statistical Commission has been receiving SDMX progress reports since 2002. The Commission recognized and supported the SDMX standards and guidelines in 2008 as the preferred standard for the exchange and sharing of data and metadata, requesting the SDMX sponsors to continue their work and to encourage national and international statistical organizations to increase the use and implementation of SDMX.

## II. Statistical Data and Metadata Exchange Roadmap 2025

4. The SDMX sponsors approved a new road map to continue the development of SDMX in the next five years. Through Roadmap 2025, the strategic objectives set forth in Roadmap 2020 will continue to be pursued, specifically (a) strengthening the implementation of SDMX; (b) facilitating its adoption for various types of uses, including data sharing and data dissemination; (c) leveraging SDMX to modernize statistical processes and standards and information technology infrastructure; and (d) improving and broadening communications.

5. The need to address a broadening set of requirements and improve the usability of the standard is recognized in Roadmap 2025. The planned launch of SDMX version 3.0 at the SDMX Global Conference in 2021, as well as its progressive adoption in the coming years, will play a key role in this regard. SDMX is set to become the central piece in the integration of the data workflow, from data collection, validation, reshaping, mapping, computation and aggregation all the way to dissemination. The standard will stretch beyond traditional aggregated data to better represent microdata and administrative data, geospatial data and other types of non-traditional sources, including big data. Lastly, the standard will be streamlined and adjusted on the basis of experience gained during the growing number of SDMX system implementations and SDMX-based exchanges.

### III. New developments: Statistical Data and Metadata Exchange 3.0

6. The SDMX version 3.0 delivery project was initiated in March 2020 with the objective of releasing version 3.0 of the standard by late 2021. The most recent major update, version 2.1, which was released in 2011, was followed by a number of incremental improvements, notably, support for the JavaScript Object Notation (JSON) and comma-separated values (CSV) exchange formats, as well as the development of the Validation and Transformation Language (VTL). By 2020, the need for a new major update became apparent.

7. Version 3.0 will have about 20 new features and updates, which were selected prior to the start of the delivery project, following consultations on requirements and priorities with the sponsor agencies and public stakeholders. The features and updates selected, which align with the strategic goals of the SDMX Roadmap, include the following:

(a) Support for microdata exchange: modifications to the data structure definition and other information model artefacts to allow microdata and unit record data to be described and exchanged. Data sets with multiple measures will be supported, as will multiple values for attributes;

(b) Improved reference metadata: simplification of referential metadata, including an easier and more flexible way to add additional information, such as footnotes, to statistical data sets;

(c) Standardized geospatial data exchange: extensions to the SDMX information model to provide formal support for data with geospatial references, including points, lines, polygons and grids, in consultation with the Expert Group on the Integration of Statistical and Geospatial Information;

(d) Semantic versioning of structures: a change in the way that structural metadata artefacts are versioned, from the current two-number system (e.g., “1.2”) to the widely adopted three-number semantic versioning model (e.g., “1.2.5”);

(e) Enhancement of several artefacts, such as constraints, structure sets and code lists, including hierarchical code lists;

(f) Deprecation of the Simple Object Access Protocol (SOAP) application programming interface in favour of an improved and easier-to-use RESTful (conforming to the representational state transfer architectural style) application programming interface;

(g) Deprecation of the obsolete EDIFACT (Electronic Data Interchange for Administration, Commerce and Transport) format for data and structures, and standardization of a single message format for XML (Extensible Markup Language) data exchange.

8. The project is being delivered by the Statistical Data and Metadata Exchange Technical Standards Working Group, with advice from the Statistical Data and Metadata Exchange Statistical Working Group. The SDMX sponsor organizations and its secretariat are steering the SDMX 3.0 project, with the Bank for International Settlements providing financial coordination and oversight. Good progress has been achieved. A draft of the version 3.0 technical specifications is being prepared for public consultation by mid-2021. The final specifications, incorporating feedback received from the public consultation, will be released in late 2021 and announced at the SDMX Global Conference.

## IV. Ongoing statistical and technical work

9. The Working Group on Statistical Data and Metadata Exchange for Sustainable Development Goal Indicators was established by the Inter-Agency and Expert Group on Sustainable Development Goal Indicators with the mandate to develop a solution for the exchange and dissemination of indicator data and metadata. The Working Group released the first official data structure definition for the indicators in June 2019. Data exchange has since been established with about 22 countries and five international agencies. An SDMX application programming interface has been established for the Global Sustainable Development Goal Indicators Database. A draft metadata structure definition was released in late 2019. Following a request by Member States, Sustainable Development Goal metadata concepts were aligned with existing metadata standards, such as the Single Integrated Metadata Structure, the global metadata concept scheme and the SDMX glossary. A Sustainable Development Goals reference metadata exchange pilot was carried out in the second half of 2020. The pilot was greatly facilitated by metadata authoring tools, developed by the Statistics Division of the Department of Economic and Social Affairs of the Secretariat in cooperation with the World Bank, which can retrieve rich text metadata from a Word template and convert it to SDMX. The Sustainable Development Goals Lab, a component of the Sustainable Development Goals global platform, facilitates the uploading of SDMX data and metadata, as well as comparisons between global and national Sustainable Development Goal data sets, while metadata help to explain differences between the two. Production SDMX metadata exchange for Sustainable Development Goal indicators is expected to be established in 2021. An SDMX application programming interface, which is already available for Sustainable Development Goal reference metadata, is expected to be populated with all available global metadata in 2021.

10. An inter-agency working group comprising Eurostat, the International Trade Centre, OECD and the United Nations Conference on Trade and Development, with the Division serving as the secretariat and chair of the working group, was established to support the implementation of SDMX standards in international merchandise trade statistics in 2013. The data structure definition of SDMX standards in international merchandise trade statistics was released in 2016 and has since been updated regularly. In late 2019, version 1.2 of the three data structure definitions of SDMX standards in the international merchandise trade statistics was published in the SDMX Global Registry with the Division as the maintenance agency, which will facilitate its implementation by countries and other stakeholders.

11. In 2018, the Statistics Committee of the European Central Bank created an interim task force on development of shared SDMX tools with the objective of fostering collaboration and the sharing of SDMX software tools, primarily within the European System of Central Banks. The objectives included assessing SDMX training needs in the European System of Central Banks and developing SDMX training, promoting awareness of and adoption of VTL, supporting the harmonization of SDMX converters and exploring synergies in the development of tools with other groups. In the course of its activities, the task force developed the open-source VTL engine and editor tool. In December 2020, it concluded its work by reporting to the Statistics Committee the successful achievement of its primary objectives.

12. IMF has continued to facilitate data dissemination using SDMX through the implementation of its Data Dissemination Standards initiatives. In the past two years, IMF assisted 27 countries with the implementation of SDMX-enabled National Summary Data Pages: 19 adherents to the Enhanced General Data Dissemination Standard, 1 subscriber to the Special Data Dissemination Standard (SDDS) and 7

subscribers to SDDS Plus. As at the end of 2020, 91 countries are disseminating data in SDMX format through their National Summary Data Pages.

13. A project on the implementation of SDMX for the global exchange of education statistics is under way. The project involves Eurostat, OECD and the Institute for Statistics of the United Nations Educational, Scientific and Cultural Organization.

14. In order to harmonize and standardize the reporting and exchange framework for official labour statistics, a set of global data structure definitions and related artefacts have been published by international organisations involving the European Central Bank, Eurostat, the International Labour Organization, OECD and the World Bank. Work has begun on a new version of the Global data structure definitions.

15. Several other SDMX global exchange projects begun in past years have come to fruition. In particular, Global SDMX data structure definitions have been published and are being used for data exchange in the domains of price statistics and environmental-economic accounting. For Environmental-Economic Accounts, five global data structure definitions have been constructed for air emission, energy, water, material flow and land accounts. Currently, the Division, OECD and Eurostat are exchanging data using SDMX, with a view to building an SDMX-based global dissemination portal for Environmental-Economic Accounts. A further expansion of data exchange among the international organizations is also planned.

16. The Statistical Data and Metadata Exchange Technical Standards Working Group and the Statistical Data and Metadata Exchange Statistical Working Group are responsible for the management and enhancement of the SDMX technical and statistical standards and the related SDMX information technology applications. The Statistical Data and Metadata Exchange Technical Standards Working Group has focused primarily on the development of SDMX 3.0. Nevertheless, work has continued on VTL. VTL can be used in conjunction with SDMX to express and share validation and transformation rules applicable to data. The SDMX 2.1 specification was updated in August 2020 to include VTL support. A newer version of the VTL 2.0 grammar was published in August 2020 to fix bugs that had been detected by the VTL user community since the publication of version 2.0.

17. The Statistical Data and Metadata Exchange Statistical Working Group has been working on a number of guidelines for and improvements to the standard, including the following:

- The business case for SDMX, which is being translated into multiple languages
- A new version of the SDMX glossary and the global metadata concept scheme derived from it
- A guideline on and controlled vocabulary for standard SDMX annotations, including classification-related metadata
- A checklist for SDMX data providers that provides users with a best practices process for implementing SDMX for reporting
- A number of new cross-domain code lists, and revision of existing code lists
- The work of the Statistical Data and Metadata Exchange Technical Standards Working Group on SDMX 3.0 feature specifications

## V. Capacity-building

18. Demand for capacity-building has grown substantially in recent years. Accordingly, SDMX sponsor agencies have stepped up their capacity-building

activities. The activities are carried out as part of global events, such as the SDMX Global Conference and the SDMX Expert Group Meeting, through regular capacity-building activities for their constituencies, and other avenues.

19. Eurostat has set up regular courses on SDMX for both beginners and information technology developers. The courses first and foremost target European Union member States, but are open to participants from other countries. In the framework of its international cooperation activities, Eurostat has also financed SDMX capacity-building activities in Eastern Europe, North Africa and South America. Eurostat also maintains the SDMX InfoSpace, which provides access to information on the SDMX standard in three languages (English, French and German), SDMX tools developed by Eurostat, upcoming training initiatives and information on progress of SDMX implementation in the European Statistical System.

20. The Division has carried out a number of capacity-building activities in areas such as the Sustainable Development Goal indicators, international merchandise trade statistics, economic statistics and others. With regard to the Sustainable Development Goals, the Division is running a project in support of Sustainable Development Goal monitoring, funded by the Foreign, Commonwealth and Development Office of the United Kingdom of Great Britain and Northern Ireland, for 20 countries across Africa and Asia. Of those, data exchange has been established with about 13 countries. Data and metadata exchange is expected to be established with all 20 countries by 2021. As part of the project, the Division has carried out SDMX training in, and provided technical assistance to, these countries. A series of international capacity-building events on Sustainable Development Goal data and metadata sharing have been organized as part of the Statistics Division-Foreign, Commonwealth and Development Office project on Sustainable Development Goal monitoring and the tenth tranche of the United Nations Development Account, with partners, including the regional commissions, the United Nations Children's Fund, OECD and the African Development Bank. Owing to the coronavirus disease (COVID-19) pandemic, in 2020 capacity-building activities have been carried out in the virtual format, including large-scale training co-organized by the Division and ESCWA. The Division-Foreign, Commonwealth and Development Office project has also been conducting online webinars and follow-up workshops with its 20 project countries on their SDMX implementation for the Sustainable Development Goals. In February 2020, the Division, the Asian Development Bank and the Association of Southeast Asian Nations organized a joint regional workshop on economic statistics SDMX to train substantive and technical experts in the region on the System of National Accounts (SNA) and international merchandise trade statistics SDMX data structure definitions and tools. The workshop led to the establishment of a new Asian Development Bank technical assistance project on strengthening the digital information of the national statistical data systems, scheduled to begin in 2021. The project will provide technical assistance to selected Asian Development Bank member developing countries to implement SDMX, as well as to develop an online training course on SDMX.

21. The European Central Bank organized an SDMX training session in October 2020, targeted primarily at European System of Central Banks participants, followed by a second training session organized in December 2020. The sessions, which were held in webinar format owing to the COVID-19 pandemic, covered an introduction to SDMX, tools and VTL, including the introduction of a real-life SDMX project at the National Bank of Iceland. Overall, the webinars were very well received by the attendees. Similar events are planned for 2021.

## VI. Software tools and applications

22. Eurostat has continued to upgrade the SDMX tools that it develops and maintains. In particular, its SDMX converter and SDMX structural validation services have extended their support of various formats, such as Excel, fixed-length record formats and SDMX-CSV. The SDMX converter and the associated structural validation service have also been made available as a web application and therefore can now be accessed and used by any interested stakeholder without the need for local installation. The Eurostat tools are available as open source.

23. The Fusion suite of software tools consists of Fusion Registry 10, a commercial SDMX statistical data warehouse, and Fusion Metadata Registry. Fusion Metadata Registry is a mature and fully fledged SDMX registry available free of charge that is supported by the Bank for International Settlements. Previously called Fusion Registry Community Edition, it was renamed and enhanced in late 2020 as part of an ongoing development and maintenance programme. Fusion Metadata Registry supports the standard application programming interfaces and also provides a graphical interface, helping users with the tasks of authoring and maintaining SDMX structural metadata. Additional functions include data validation, data transformation and data reporting using Excel templates, with auditing and the rollback of changes recently added. Support for SDMX 3.0 is planned for 2021. The Bank for International Settlements continues to be fully committed to funding the maintenance of and improvements to Fusion Metadata Registry. However, a more ambitious approach is being examined to develop Fusion Metadata Registry together with other tools created by the SDMX community into a cohesive and complete SDMX and Generic Statistical Business Process Model statistics data platform with a cloud option, technical support and training services. The new products and services, largely funded by third-party grants and donations, will be free, easy to use and explicitly designed to help least developed and developing countries to produce and publish world-class official statistics with the social and economic benefits that they bring.

24. The Statistical Information System Collaboration Community, led by OECD, is a group of statistical organizations with common goals and similar needs in terms of data dissemination and production that is interested in mutualizing costs and sharing knowledge and is committed to fostering common standards. The Collaboration Community develops open-source digital solutions to support the official statistics for the world at large. One such solution, the .Stat Suite, is an SDMX native standard-based, componentized, open-source platform for the efficient production and dissemination of high-quality statistical data. The road map for the product is progressively covering the full macro data life cycle, enabling organizations to bring further value as trusted quality data producers, empowering them with quality assurance by design and advanced workflow and reporting mechanisms, high-performance calculation and import and export tasks to enable efficiency gains and high-quality user experiences in data operations. The Collaboration Community also works in close partnership with Eurostat, enabling the reuse of open source and common modules from the SDMX reference infrastructure package. The .Stat Suite is developed under a continuous user-driven product delivery model, with a state-of-the-art software delivery mechanism, known as DevOps, to combine transparency, agility, a multiplicity of contributors, simplicity of deployment in the cloud, software quality and security. In 2020, the .Stat Suite was rolled out at a number of organizations (the United Nations Children's Fund, the Economic and Social Commission for Asia and the Pacific, the Pacific Community, the International Labour Organization, the Federal Competitiveness and Statistics Authority of the United Arab Emirates (FCSA), the Australian Bureau of Statistics (ABS) and the National Bank of Belgium), as well as having been successfully configured on the

United Nations Global Platform by the Division, OECD and the Economic and Social Commission for Asia and the Pacific. It is currently being used by Cambodia as its official statistical data warehouse, with further avenues being explored for making the .Stat Suite on the United Nations Global Platform available for use by other national statistical offices.

25. The European System of Central Banks task force on shared SDMX tools further enhanced its VTL engine and editor tool, including a web services application programming interface that enables integration in any external application by means of web services calls, a command line interface and new operators, as well as new configurable properties, such as the possibility of picking and choosing the desired environment, such as R, SDMX or CSV, and the configuration of an SDMX registry or web service as a metadata repository. An R Shiny version of VTL Studio, made available online,<sup>1</sup> allows access to the tool without a local installation. Since January 2020, the European Central Bank has been running an information technology project that implements a new statistical data production and compilation environment using big data technologies, such as the full Hadoop stack, and leveraging SDMX tools such as Fusion Registry. SPACE will provide functionality to ingest, gain access to and manipulate SDMX data with programming languages such as Python and R, implementing business processes logic on top of the big data platform. The Network of Experts on Machine Learning, a substructure of the Working Group on Statistical Information Management of the Statistical Committee of the European System of Central Banks, brings together experts from national central banks and national competent authorities to share knowledge and exchange experiences with advanced analytics and related tools. An ongoing network pilot project currently evaluates various approaches to outlier and anomaly detection in time series data. The integration of those solutions into the SPACE platform will enable business users to carry out modern data quality assurance processes with SDMX data received from European System of Central Banks institutions.

26. The World Bank has developed a tool for computer-assisted translation of Sustainable Development Goal metadata from English into other languages that provides output in human and machine-readable format. The Division has expanded the tool to facilitate the conversion of a Word template into SDMX-compatible, machine-readable format.

27. The IMF SDMX Central tool continued to support member countries in the conversion, validation and registration of SDMX data in their National Summary Data Page. Since 2016, IMF has provided website and web services free of charge, promoting SDMX as an efficient exchange medium for countries that adhere to the Enhanced General Data Dissemination System (e-GDDS) and subscribe to SDDS Plus. SDMX Central plays an essential role in the automated data collection processes between IMF and member countries.

## VII. Events

28. The seventh SDMX Global Conference,<sup>2</sup> on the theme “SDMX for you”, was held in Budapest from 16 to 19 September 2019. The conference was hosted by the Hungarian Central Statistical Office and was followed by a two-day capacity-building session. A total of 219 participants from 97 countries attended the event. The conference highlighted both rapidly increasing rates of implementation in all regions of the world and a growing demand for global data structure definitions to facilitate

---

<sup>1</sup> See <https://vpinna80.shinyapps.io/vtlStudio>.

<sup>2</sup> The conference report and materials are available at [https://sdmx.org/?sdmx\\_events=7th-sdmx-global-conference-summary-report-now-available](https://sdmx.org/?sdmx_events=7th-sdmx-global-conference-summary-report-now-available).

data exchange. In reports presented at the conference, it was demonstrated that implementation of SDMX for data exchange and dissemination, as well as its integration into business processes, was supported by the availability of software tools, which can be adapted to various scenarios, depending on the available infrastructure, from the simple to the most sophisticated. The conference demonstrated the increasing interest of the private sector, with representatives of six commercial companies giving presentations.

29. The tenth SDMX Expert Group Meeting,<sup>3</sup> originally scheduled for September 2020, has been postponed to 25 to 28 January 2021. It will be hosted by IMF in a virtual format owing to the COVID-19 pandemic.

## **VIII. Action to be taken by the Statistical Commission**

30. The Commission is invited to take note of the present report.

---

---

<sup>3</sup> Further information on the meeting is available at [https://sdmx.org/?sdmx\\_events=update-10th-sdmx-expert-group-meeting-held-virtually-in-january-2021](https://sdmx.org/?sdmx_events=update-10th-sdmx-expert-group-meeting-held-virtually-in-january-2021).