



Economic and Social Council

Distr.: General
19 December 2024

Original: English

Statistical Commission

Fifty-sixth session

New York, 4 - 7 March 2025

Item 5(h) 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 (SDMX) sponsors

Note by the Secretary-General

In accordance with Economic and Social Council decision 2024/312 and past practices, the Secretary-General has the honour to transmit the report of the Statistical Data and Metadata Exchange (SDMX) sponsors, which is submitted to the Commission for information.

* E/CN.3/2025/1.

I. Introduction

1. The aim of the Statistical Data and Metadata Exchange (SDMX) initiative, established in 2001, is to create and maintain technical and statistical standards and guidelines together with an IT architecture and IT tools, to be used and implemented by the community of official statistics.

2. The SDMX technical standard has evolved significantly since its inception, with SDMX 2.1 launched in 2011 being the first version to gain substantial traction. The release of version 3.0 in 2021 brought essential support for microdata, expanding the standard beyond its traditional focus on aggregate statistics and time series. SDMX 3.0 additionally introduced semantic versioning for metadata artefacts which strengthens metadata management and exchange by defining the contract for change. Robust rule-driven metadata versioning reassures metadata users that the structures are stable, promotes metadata reuse and supports the Findability, Accessibility, Interoperability, and Reuse (FAIR) principles by enabling persistent references to immutable artefact versions. In recent years, SDMX has increasingly been used beyond its original application of facilitating data reporting and exchange between countries and international agencies, with dissemination, data and metadata management in support of data governance, and other use cases gaining in importance and prominence.

3. The UN Statistical Commission has received SDMX progress reports since 2002. The Commission recognised 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 organisations to increase the use and implementation of SDMX.

II. International Labour Organisation joined the SDMX Sponsor Agencies

4. The International Labour Organization (ILO) has become a Sponsor Organization for the SDMX initiative, accepting a formal invitation from the Sponsors Committee and signing the SDMX Memorandum of Understanding. The Memorandum of Understanding sets out the governance and framework for cooperation between the Sponsors, including intellectual property rights and resource sharing.

5. The ILO had already made significant contributions to the SDMX initiative before expanding their collaboration as a Sponsor Organization, developing widely used tools and actively participating in working groups and events including an extended term leading the Technical Working Group. With this new role, the ILO joins the Bank for International Settlements (BIS), the European Central Bank (ECB), the European Statistical Office (Eurostat), the International Monetary Fund (IMF), the Organization for Economic Co-operation and Development (OECD), the United Nations and the World Bank, who have sponsored the SDMX initiative since 2007.

III. Artificial Intelligence opportunities and impacts

6. The pace of artificial intelligence (AI) application to SDMX has accelerated substantially since the emergence of ChatGPT at the end of 2022 bringing opportunities to revolutionize every stage of the statistics lifecycle from initial data collection to final publication, discovery, and analytics. Despite advances in the usability and content of institutional web data portals, practical natural language data discovery has remained elusive until recently, but the situation has begun to change. Several initiatives are underway that aim to facilitate querying of SDMX Application Programming Interface (API) using natural language. This will make it possible to use Artificial Intelligence with any data published by national or international organisations using easily available, open source, off-the-shelf tools.

7. The IMF has been developing StatGPT – an AI-based SDMX query building assistant. The main task of StatGPT is to properly decipher and extract all the necessary parameters from a natural language prompt, and use this information to construct SDMX query parameters to return statistical data via an application programming interface (API). The first prototype of this application was presented at the SDMX Global Conference in Bahrain in October 2023. This was followed by a presentation by EPAM, the vendor the IMF is working with to develop the product, at the SDMX + AI workshop in Paris in March 2024 co-organised by the OECD and BIS. At the conclusion of that meeting it was agreed to leverage the work done on StatGPT and make the application available to the larger SDMX community, with data from partner organizations to facilitate its testing. In September 2024

the IMF released StatGPT 2.0 containing datasets from 8 partners organizations and over 100 participants carrying out the testing. Feedback, usage statistics, cost, and the results of a survey will be compiled into a report for the SDMX community who will meet to determine potential co-investment opportunities. The IMF aims to make the tool generally available to statistics publishing institutions although a sustainable funding model will be required for the LLM operating costs, which scale with usage.

8. In another promising data discovery development, a UK-based AI and search technology company, SEASE, has collaborated with SDMX sponsor organizations to enhance the accuracy of responses to data consumers' questions. By combining traditional search engine technology with Large Language Models (LLMs), they aim to disambiguate users' natural language queries, extract relevant information, and formulate structured queries. Despite some inherent LLM limitations, this approach demonstrates a relatively simple solution for improving the accuracy of results from natural language questions on statistical data.

9. Making the SDMX standard more accessible has been a longstanding goal of the Sponsors. A BIS-led initiative to publish an SDMX conversational chatbot aims to significantly advance this effort and improve the learning curve for those studying the standard. This LLM powered chatbot utilising the retrieval augmented generation (RAG) technique allows users to ask questions about the SDMX technical standards and statistical guidance. Plans are in place to expand the source information with capacity-building materials, SDMX software tool manuals and other relevant content. Development and testing are ongoing, focusing on ensuring the accuracy and relevance of the chatbot's answers.

10. Exciting advances have already been made but the SDMX AI journey is only just beginning. That significant investment is being committed by both institutions and private companies demonstrates the confidence that AI can bring concrete value to the statistics value chain. The collaborative efforts and innovative solutions being developed highlight the community's commitment to leveraging AI for the benefit of all stakeholders.

IV. SDMX business case review

11. The first SDMX Business Case was released in mid-2020. This marked a milestone achievement, as for the first time the benefits of SDMX and aspects of its use were collected in a single concise document. However, rapid developments since have quickly made the Business Case obsolete. This pertains, in particular, to the fact that usage of the standard has extended far beyond its original purpose of data exchange between countries and international agencies. In many regions of the world, the principal use case of SDMX is data dissemination facilitated by powerful platforms such as the SIS-CC .Stat Suite and IStat StatKit, as well as data and metadata management in support of data governance facilitated by the Fusion Metadata Registry (FMR). In addition, entirely new use cases such as micro-data exchange and geospatial-enabled data are supported in more recent versions of the standard.

12. In view of these developments, in early 2024 the SDMX Sponsor Agencies established an SDMX Business Case Review Task Team. The Task Team will reach out to stakeholders at national statistical offices, central banks, academia, international organisations, and others, in order to catalogue use cases beyond data exchange and compile users' experiences with the implementation of SDMX for the various use cases. This information will be used to update the Business Case, thus facilitating informed decision-making on the part of national and international statistical organisations as well as other users considering implementation of SDMX dataflows. The Task Team is expected to deliver its results in 2025.

V. New initiatives

A. sdmx.io open-source tools ecosystem

13. Many official statistics organisations view data as a public good; the BIS has taken this idea one step further and committed to software (used in the production or use of statistical data) as a public good. This idea is encapsulated in an initiative labelled sdmx.io.

14. Aligned with the vision of the sdmx.io initiative, the BIS has partnered with international organisations (OECD, IMF, Eurostat), national statistics offices (the French National Institute of Statistics and Economic Studies, INSEE), central banks (Bank of Italy, the National Bank of Belgium), and private sector organisations

(Meaningful Data, HMS, Making Sense) to improve tools used for national and international statistics production and use. The outputs and outcomes from these initiatives have contributed to creating new or improving upon a rich ecosystem of high quality open-source software tools produced by public and private sector entities.

15. To support tool producers, the BIS offers assistance in the creation and sustainability of open-source tools and encourages adherence to good practices on tool interoperability with minimal feature overlap. To support tool adopters, the BIS helps assure relevance via applicability of tools to practical statistical use cases, and ease of adoption and use via ease-of-deployment approaches and learning resources.

16. These efforts have resulted in increased collaboration and coordination amongst tool producers which in some cases have led to new tools being produced which address gaps in the data production ecosystem (gingado, FMR Workbench, SDMX Dashboard Generator) or alternatively, to multiple tools being consolidated into a single best-in-class tool (pysdmx, FMR, Matrix Generator) co-produced and maintained by shared resources and good governance models.

B. UN Data

17. The UN Data Modernization Project has substantially enhanced the capacity for integrating and disseminating statistical data across the UN system in response to the Secretary-General's Data Strategy. Guided by the Roadmap for Innovating UN Data and Statistics, this initiative directly supports the SDMX Technical Working Group by developing tools that improve interoperability of SDMX-compliant data infrastructures. In particular, the UN Data modernization project has created comprehensive mappings to harmonize code lists and concepts used by various UN entities, aligned with major international statistical classifications, thereby contributing to the interoperability of SDMX implementations across organizations and statistical domains. Additionally, the alignment of UN Data with the Cross-Domain Interoperability Framework (CDIF) will further contribute to the application of the FAIR principles within the SDMX data ecosystem.

18. The UN Data Modernization Project's focus on connecting diverse data through knowledge graph methodologies complements the SDMX Technical Working Group's objectives by facilitating the translation and linkage of SDMX-formatted data with open data formats like DDI-CDI, schema.org, and SKOS. This work is expected to enable seamless navigation between structured statistical data and broader metadata contexts, enhancing both accessibility and analytical capabilities. The project's emphasis on standardized, machine-readable mappings underpins SDMX's aim of efficient, scalable data sharing and promotes the use and impact of authoritative statistical resources.

VI. Technical Advances

A. Validation and Transformation Language

19. The Validation and Transformation Language (VTL) is a standard language for defining validation and transformation rules including a set of operators, their syntax and semantics, for any type of statistical data. VTL builds on the Transformation section of the SDMX information model by taking the common parts of the Generic Statistical Information Model (GSIM), SDMX and Data Documentation Initiative (DDI) standards for the representation of concepts and data. The logical formalization of validation and transformation rules provides a "technology-neutral" language that can be used directly by business users, without any IT intervention. These characteristics allow VTL to be easily integrated in various steps of the statistical process within organisations adopting standards such as SDMX, DDI, or Data Point Model (DPM).

20. The Task Force for the Validation and Transformation Language, created in 2013 under the initiative of the SDMX Secretariat and responsible for developing and maintaining the language, released the VTL version 2.1 in July 2024. The new release contains additional standard time operators and some bug fixes in the language specifications. Furthermore, the User Manual and the Reference Manual are expected to be available by the end of 2024 in a markdown format, improving readability and maintainability, and with more examples added.

21. At the moment, there are four VTL open-source engines available, implemented by private companies and international organisations. VTL has been adopted by several central banks, national statistical offices, and

international organizations. A number of initiatives are underway. VTL is expected to be integrated in the Fusion Metadata Registry, and several developers of VTL tools are joining the SDMX.io initiative.

B. SDMX 3.1

22. This latest release of the technical standard scheduled for the end of 2024 addresses the challenges of managing highly multi-dimensional datasets in official statistics. Statisticians have struggled to efficiently model datasets with high dimensionality using SDMX's information model or other environments such as a relational database. Such datasets often require the addition of new breakdowns or dimensions over time, which is expensive due to the disruptive effect on existing data exchange processes. SDMX 3.1 mitigates these issues by allowing fixed sub-cubes to be defined on Data Structure Definitions (DSD) which define analytical cubes and data table schemas that remain valid even as the DSD adds dimensions. The innovation reduces the need for mixed dimensions and other workarounds, enhances interoperability, and enables agile DSD development by avoiding disruption to established data collection pipelines.

C. SDMX 3.2

23. Projecting forward to 2025, a further incremental enhancement to the SDMX information model is foreseen with support for custom structural metadata classes. To date, the information provides a collection of standard classes for artefacts including concepts, code lists and data structure definitions. From SDMX 3.2, users will be able to create their own classes. The initial use case for this 'custom structure definition' feature is presentational metadata which typically describes how data should be formatted when presented on websites and data discovery portals. Presentational metadata is often specific to institutions and the visualisation software making it hard to define a standard metadata class that satisfies all use cases. Custom structure definitions are expected to find application beyond presentational metadata. Examples already exist for publication tables, reporting templates and balance equality rules where non-standard structural metadata classes are in use which otherwise follow the SDMX patterns for metadata definition, storage and exchange.

D. REST API

24. The REST API is one of the three core components of the SDMX technical standard, along with the information model and transmission formats. Originally introduced in SDMX 2.1, the REST v1 specification simplified data discovery, retrieval, and validation through various queries. SDMX-REST v2, released as part of SDMX 3.0, introduced non-backward compatible enhancements, including improved data queries, maintenance operations, and reference metadata queries. It also supported new information model features like geospatial code lists and semantic versioning. Further incremental enhancements are scheduled for Q4 2024, including "time travel", sorting and pagination, and segmented constraints. Looking ahead, the roadmap includes new APIs for business functions such as data validation and mapping, and additional alternatives to REST, like gRPC, for specific use cases such as microservices.

E. Technology Compatibility Kit

25. As part of its contribution to the work of the SDMX Technical Working Group, Eurostat has developed the Technology Compatibility Kit (TCK), which offers a toolkit for testing SDMX REST API implementations for compatibility with the SDMX standard. The TCK allows users and developers of SDMX software to measure the compliance and coverage of an SDMX RESTful endpoint against the available SDMX REST API specifications through a battery of tests that is built dynamically by the TCK. The latest official release of the TCK can test the SDMX-ML output of REST API queries for SDMX structures, schemas and data against SDMX REST API specifications up to version 1.4 of the API. Test releases of the TCK that include additional features (e.g. the capability of working with the latest specifications of the API or the coverage of additional output formats like SDMX-JSON) are regularly made available to the SDMX community.

VII. Official SDMX Services

A. SDMX Global Registry

26. The SDMX Global Registry¹ is a central repository of globally agreed data structure definitions for several domains including national accounts and the Sustainable Development Goals. It is also the definitive source of standardised cross-domain concepts such as frequency, unit of measure and observation confidentiality which underpin the drive to improve dataset comparability through global harmonisation of concepts.

27. The Global Registry supports the latest version 3.0 of the SDMX standard. Planned developments centre on support for SDMX 3.1 and subsequent versions of the standard, and relocation of the service onto BIS infrastructure.

B. SDMX Global Discovery Service

28. The SDMX Global Discovery Service (SGDS) is a new official SDMX service to be launched in Spring 2025 aimed at enhancing the discovery and sharing of SDMX artefacts and datasets as more SDMX registries and data dissemination services are deployed and used in production. The need for a robust discovery service is evident given the essential role of the Global Registry which maintains core artefacts like top-level Agency Schemes, and the various international and national registries maintaining widely used structural metadata.

29. A typical use case envisaged is discovery of the authoritative source of SDMX artefacts for a specified agency. Given the ID of the SDMX agency, the SGDS will return the SDMX REST structure API endpoint from where the agency's artefacts can be retrieved. A key additional feature which enables SDMX support for FAIR principles is a URN naming service. This translates an artefact's persistent URN to a URL from which the artefact can be retrieved.

30. The SGDS provides a REST API for convenient programmatic access but also a simple user interface for browsing and searching SDMX artefacts. The service is currently in test and evaluation at <https://gds.sdmx.io>. Stakeholder feedback has already led to design improvements, including the rationalization of messages to align more closely with SDMX practice and enhancing client-oriented requests for better discoverability.

31. The SGDS aims to expose the new API to potential clients, mainly other services that require the discovery of SDMX endpoints. Additionally, the UI will be made available to the SDMX community under sdmx.io for comments and feedback. The development of processes for onboarding new agencies and their services is also a priority. Furthermore, a Test Configuration Kit (TCK) will be added to test and configure SDMX service entries, including supported resources and formats.

VIII. Interoperability

A. General work on interoperability

32. The Supporting Standards Group of the UN Economic Commission for Europe (UNECE) High-Level Group for the Modernisation of Official Statistics (HLG-MOS) has released a report that identifies elements of the SDMX and Data Documentation Initiative (DDI) standards that are relevant in each phase and sub-process of the Generic Statistical Business Process Model (GSBPM), from the specification of statistical needs to the dissemination of statistics². Both SDMX and DDI standards have widespread adoption, each having its own advantages and communities of users. The motivation for using GSBPM to contextualise SDMX and DDI artefacts in this way, is to help experts in one of these standards to easily see which artefacts from the other standard might be relevant for a given stage of the statistical production process. This is an important prerequisite for those who use both standards, and who wish to make them interoperate with each other, for example to construct a data pipeline. SDMX and DDI have many similarities between their artefacts, and have evolved in recent years to start to overlap in certain respects in the roles that they can perform. The report also provides an introduction to SDMX and DDI standards in the context of GSBPM, and guidance on making them interoperate

¹ See <https://registry.sdmx.org>

² See <https://unece.org/sites/default/files/2024-10/Implementation%20Standards%20in%20the%20context%20of%20GSBPM.pdf>

with each other and with VTL. As well as being a useful part of a developer's toolkit, it is hoped that this work may be a foundation for further work to include other open implementation standards.

33. In 2024, the SDMX Sponsor Agencies established the SDMX Classifications Task Team, which aims to facilitate creation, publication, and maintenance of official statistical classifications as code lists at the SDMX Global Registry. While many classifications are already published as SDMX code lists and used in global data exchange as well as dissemination data structures, the coverage is incomplete. The availability of official classifications should both simplify development of dissemination data structures by national and international statistical organisations, and significantly improve their interoperability. A key component of the Task Team's work is the development of standardized procedures for the maintenance of statistical classifications aimed at minimizing the burden and putting in place sustainable processes for the same. Complemented by advances in data modelling in recent releases of SDMX as well as forthcoming implementation of the FAIR principles and linking SDMX to semantic web, this initiative is expected to lead to substantial improvements in the usability of statistical data, reduction of the cost of combining data from multiple sources ("data wrangling"), and making the data easily consumable by artificial intelligence applications.

34. The SDMX Statistical Working Group (SWG) jointly with partners have been working on the following content-oriented guidelines, some of which have been published and some are in progress:

- Unit of measure guidelines. The importance of unit of measure in statistical data modelling is to assure consistency and accuracy of the data. Measurement units are used to quantify data and to offer an immediate mechanism to assess the comparability and computation scope of the data. This content-oriented guideline for Units of measures is destined to statistical data modelers working with multidimensional data, primarily in SDMX – but with a scope that can go beyond (e.g. data warehousing in general). The patterns proposed in this paper are anchored in existing and well-established international standards in the fields of science and engineering;
- Guideline for SDMX Hierarchies. The purpose of this document is to illustrate the use case for hierarchies in SDMX, provide examples and recommend best practices for its implementation. Hierarchies are commonly used to represent various relationships and classifications. They play a crucial role in data management systems for tasks such as data modelling and data dissemination;
- Revision to the guidelines on the creation and maintenance of SDMX code lists. This revision adds new features from version 3.0 of SDMX, and other practical issues and recommendations, such as breakdowns with multiple variants;
- SDMX annotations guideline and controlled vocabulary revision. This revision adds SDMX 3.0 features and new annotations to the controlled vocabulary, thereby increasing interoperability;
- Guideline for working with vintages in SDMX. In these guidelines, vintage is defined as a published set of data at a particular moment. Vintage data allows academics to reproduce others' research, build more accurate forecasting models, and analyse economic policy decisions using the data available at the time;
- Implementing SDMX at the enterprise level. A new guideline on implementing SDMX at the enterprise level (multi-domain and across the statistical lifecycle); which could be used by agencies to kick-start their enterprise-level implementation, save resources and time and achieve the objectives of SDMX adopters;
- Investigating the benefits of interoperability between SDMX and Linked Open Data (LOD) ontologies. As data standards mature, there is a growing need for interoperability; this could be viewed simply as a way to express information from standard into the other. Whilst this could be used to widen the audience for the same information, it poses a deeper question, can the two standards complement each other by leveraging the capabilities of the other. This work will include an SDMX ontology and potential changes to SDMX, such as providing URIs for SDMX objects;
- Microdata for SDMX 3.0. A new guideline will describe best practices for microdata modelling, sharing, interoperability and confidentiality.

B. Subject matter domain-specific developments

35. In the area of labour statistics, the International Labour Organization (ILO) continued to support countries in the implementation of Labour Market Information Systems (LMIS) by developing a series of SDMX capacity building activities and by providing a toolkit of open-source software tools based on SDMX, namely the SIS-CC .Stat Suite, SDMX Constructor, Statistical Metadata-driven Analysis & Reporting Tool (SMART) and the LMIS Excel Add-in. In the last two years, four countries (Uruguay, Chile, Botswana and South Africa) have launched their systems, another four are in the final stages of initial data upload, and fifteen projects (13 national and 2 regional systems) are underway. A labour global data structure definition developed by OECD, ILO, Eurostat, the European Central Bank and the World Bank has been released and is available in the Global Registry for labour-related data and metadata exchange.

36. The IMF has continued to facilitate data dissemination using SDMX through the implementation of its Data Dissemination Standards initiatives. In the past two years, the IMF assisted seven countries in the implementation of SDMX-enabled National Summary Data Pages (NSDP)—three adherents to the Enhanced General Data Dissemination Standard, two subscribers to the Special Data Dissemination Standard (SDDS), and two subscribers to the SDDS Plus. As of end-2024, there are 112 countries disseminating data in SDMX format via their NSDPs.

37. As part of the revision of the International Merchandise Trade Statistics 2010 (IMTS 2010) and Manual on Statistics of International Trade in Services 2010 (MSITS 2010) by the task team on international trade statistics of the Committee of Experts on Business and Trade Statistics, a guidance notes on SDMX for trade in goods and services was prepared, and it went through a global consultation in February 2024. The guidance note consists of a list of minimum variables, an update of the DSD taking account of forthcoming versions of IMTS/MSITS, and its implementation strategies, including capacity building. The positive outcome of the global consultation supported the use case of SDMX for trade statistics, highlighting its potential benefit for data sharing, standardization, and efficiency. With this, the task team will continue implementing the plan outlined in the guidance notes in 2025 and beyond.

38. Global SDMX Data Structure Definitions in the area of Macro-Economic Statistics continue to be maintained and used for data exchange. In particular, in 2024 the recently finalised COICOP 2018 classification was implemented in several of these global Macro-Economic Statistics Data Structure Definitions. The domains currently covered include National Accounts, Balance of Payments, Foreign Direct Investment, Consumer Price Indices and Environmental-Economic Accounts. A global DSD for Residential Property Prices is also under development under the coordination of the BIS.

39. As part of the work of the Committee of Experts on Environmental-Economic Accounting, the Statistics Division, in collaboration with the OECD, began regular global data collection for air emission and energy accounts according to the System of Environmental-Economic Accounting (SEEA) in 2023. The global data collection complements the data collection for European Statistical System countries by Eurostat. The data collection utilizes SDMX-enabled Excel questionnaires which map to the global SEEA data structure definitions. Similar questionnaires have also been made available by Eurostat for European Statistical System countries. The Statistics Division disseminates the data through their SDMX application programming interface. In addition, this year the Statistics Division began exchanging air emission and energy accounts data with Eurostat and OECD using SDMX application programming interfaces. As a result, the Statistics Division now disseminates SEEA data for 42 Member States.

40. The Working Group on Statistical Data and Metadata Exchange for Sustainable Development Goal (SDG) Indicators has continued to maintain the global data and metadata structures and dataflows for SDG indicators. SDMX Data Exchange with the SDG Custodian Agencies has significantly increased; over 40% of the global SDG database is now transmitted as SDMX datasets. Voluntary provision of SDG datasets by the Member States has also continued, with data exchange having been established with about 40 countries. The global SDG dataset as well as metadata for the SDG indicators, are disseminated the Statistics Division SDMX application programming interfaces. The availability of metadata in a machine-readable format made it possible to link SDG indicators to their metadata at the Global SDG Database web site and display the data alongside metadata. The Working Group

has also contributed to the development of the popular Open SDG³ platform, maintained by the UK Office for National Statistics, which supports the global SDG data and metadata structures, and which is used by about 18 Member States to disseminate their SDG indicators. Experienced gained in the SDG data exchange led to the Working Group making significant contributions to the development of SDMX 3.0 and SDMX 3.1. The Working Group has also developed an e-learning course on SDMX for SDG Indicators and coordinated capacity building efforts in the area.

IX. Software tools and applications

41. ILO continued to support its SDMX toolkit composed of the SDMX Constructor for structural metadata editing and SMART for data compilation, recoding and reformatting. The Constructor has more than 1700 downloads and SMART nearly 1100. A new Excel add-in tool called "LMIS Excel Add-in" has been released, which allows any SDMX dataflow downloaded from any SDMX compliant API to be downloaded into Excel. When connected to a .Stat Suite platform, this Excel data can be edited and pushed back into the data warehouse.

42. The ILO has developed a simplified data modelling methodology for the implementation of LMIS, using a shared workbook with one sheet for entering the list of concepts, another sheet for the code lists linked to each concept by its ID, and a third sheet for defining the dimensional tables to be published, defined by the indicator and breakdowns. A function in the SDMX Constructor allows the user to import these worksheets and create all the SDMX artefacts (namely concept scheme, code lists, DSDs, dataflows and content constraints) required for the initial data upload in a single step.

43. The IMF SDMX Central continued to support member countries in the conversion, validation, and registration of SDMX data in their National Summary Data Pages. Since 2016, IMF has been providing a cloud-based platform and web services free of charge, promoting SDMX as an efficient exchange medium for countries that adhere to the Enhanced General Data Dissemination System and subscribe to Special Data Dissemination Standard and Special Data Dissemination Standard Plus. SDMX Central plays an essential role in the automated data collection processes between IMF and its member countries. In 2023, IMF transitioned SDMX Central to the Fusion Metadata Registry (FMR) bringing important benefits, including a strategic collaboration with other SDMX Sponsor agencies and a community-based approach to prioritize developments. FMR is a free and open-source platform managed by the BIS.

44. The SDMX Matrix Generator is an Excel-based tool to visually model and design SDMX data artefacts and generate the SDMX-ML markup for implementation. It is an enhancement of the Generic SDMX Design Matrix which has proven to be successful as a collaborative design tool for non-SDMX experts. The primary goal of the tool is to embody the SDMX modelling guidelines and be able to create the artefacts without a lot of SDMX technical knowledge, and to put the focus on the statistical aspects of the data model. These aspects have led to its extensive use in modelling training. It enables flexible modelling solutions, such as using one DSD for several dataflows. It is widely used by sponsors and the SDMX community at large and has proven particularly useful in the initial stage of conceptualising data models. The global DSD projects use the SDMX Matrix Generator to design their structural metadata collaboratively. New features added by request are category scheme and agency scheme creation, easy constraint definition, scaling to modelling surveys and large code lists, decomposition serialisation, uncoded dimensions, multiple concept scheme references, custom Concept roles, many performance and stability improvements.

45. With Microsoft Power BI becoming one of the most widely used Business Intelligence tools on the market today, the SIS-CC Community saw the need to ease the sourcing and visualisation of statistical data by developing a Power-BI connector for SDMX. The connector is certified by Microsoft, making it integrated in all Power-BI installations worldwide. This is the first time that SDMX is integrated into a private sector company's tools suite, which has significantly increased the SDMX installation base.

46. The Fusion Metadata Registry (FMR) is a mature, free-to-use SDMX "structural metadata registry" developed and maintained by the Bank for International Settlements. It is widely employed in organisations worldwide as a data modelling platform and controlled repository for statistical metadata helping to improve

³ See <https://open-sdg.org/>

metadata governance, maintainability, standardisation and harmonisation. In addition to deployment within institutions, FMR also powers several public-facing SDMX structural metadata services including the SDMX Global Registry, IMF Central, and others. Since its earliest release in 2007, FMR has provided a user interface for statisticians and data officers to interactively create and maintain SDMX metadata artefacts. That capability has been extended to FMR Workbench, a variant of the original FMR that works with any SDMX metadata repository that exposes the standard REST API. A key driver for the FMR Workbench was the needs of the National Bank of Belgium (NBB) to ease the management of structural metadata in the SIS-CC .Stat Suite, widely used for statistical data dissemination. Now, the BIS FMR team in collaboration with colleagues from the OECD and National Bank of Belgium have designed an additional FMR user interface specifically for the statistical domain modelling use case. This ‘FMR data modeller’ tool implements the well-established ‘matrix’ approach allowing complete statistical domains to be described as abstract models. These can subsequently be materialised as SDMX artefacts for use in practice. The project is currently securing co-investors to fund the build stage with an MVP expected by early 2026 subject to financing. FMR version 11 was released in 2021 with support for the information model and core features of SDMX 3.0 and has since been developed through 19 minor iterations. FMR 12, the next major release, is due at the beginning of 2025 and adds support for ‘metadata time travel’ which allows structural metadata to be retrieved as it was at any arbitrary time in the past. A key use case for metadata time travel is validation of SDMX data using the structural metadata as it was at the time the data were created – essential for data collectors and reporters in domains where the metadata changes rapidly. FMR 13 scheduled for mid-2025 will support semantic versioning of artefacts and the foreseen SDMX 3.1 enhancements for modelling data with large numbers of dimensions.

47. Eurostat has continued to upgrade the SDMX tools it develops and maintains. Eurostat upgraded its SDMX Converter application and the various components of its SDMX Reference Infrastructure to support SDMX 3.0 data structures and data formats. Particular effort was spent on supporting key SDMX 3.0 functionalities, such as array values and multiple measures. Eurostat’s tools are all available as open source.

48. The .Stat Suite is a mature, open-source free-to-use SDMX native platform powering the data lifecycle for official statistics in more than 50 organisations around the globe. It is developed and maintained by the Statistical Information Systems Collaboration Community (SIS-CC), a global community of more than 20 members and partners under the leadership of the OECD. The .Stat Suite is a platform to manage the data lifecycle for official statistics (design, collect, process, disseminate), explore data and develop various reporting and dissemination experiences, and builds on best practices in statistical data modelling. The .Stat Suite leverages the Eurostat SDMX Reference Infrastructure, and FMR Workbench. In the last two years, the community has initiated the work, together with the BIS and Eurostat, on the Reference SDMX Implementation. In addition, a number of projects have come online, launching the .Stat Suite in several statistical offices and international organisations, including Thailand, El Salvador, Uruguay, Greece, Malta, Luxembourg, Madagascar, Maldives, Food and Agriculture Organisation, and others. The .Stat Suite is also serving as a core component for SDMX data modelling initiative led by the OECD. In support of its continued push to adopt open source, the Community has implemented support for the open source database MariaDB that will bring about substantial cost savings for implementers.

49. The SDMX Secretariat recently launched a new web page on SDMX tools for developers and implementers, providing a comprehensive list of SDMX tools offered by sponsors and members of the wider SDMX community.

X. Communications, outreach and stakeholder engagement

A. Global events

50. The 9th SDMX Global Conference held from 29th October to 2nd November 2023 in the Kingdom of Bahrain, attracted significant worldwide interest with more than 550 people registering to participate. Approximately 230 attended in person with others watching the plenary sessions remotely on live stream, and on catchup. The event, the first to be held in person since the COVID pandemic, was co-organised by the SDMX Sponsors and the Bahrain Information & eGovernment Authority (iGA) with the support of UN Economic and Social Commission for Western Asia (ESCWA) and Bank of Italy, and successfully brought together a broad range

of stakeholders from national, regional, and international agencies, academia, and the private sector. The theme of “Empowering Data Communities” explored how SDMX tools and technologies can be leveraged to enhance collaboration and knowledge sharing. A total of nine plenary sessions were held over three days covering a broad range of topics from AI to advances in data governance using SDMX. The event ended with two parallel SDMX capacity building streams for beginners and more advanced practitioners. The concluding panel discussion highlighting the significance of SDMX in the world of data management and governance and emphasised the need for continuous adaptation and collaboration to meet evolving data needs and challenges. The panel agreed that a user-centric approach and focus on innovation and data governance provide a clear direction for the future of SDMX.

51. The 12th SDMX Experts workshop was held from 7 to 11 October 2024 in Amsterdam, the Netherlands. The event was attended by 87 SDMX experts in person, with others following the event remotely via live stream. The event was hosted by Eurostat and co-organised by the SDMX Sponsors. A total of eleven plenary sessions were held over four days and were complemented by an advanced capacity building session on the fifth day as well as by several side meetings of different SDMX governance groups and task forces. The event highlighted how the use cases for SDMX have grown over time far beyond the original focus of the standard on data exchange and showcased the potential SDMX has in enabling AI-assisted data discovery, retrieval and analysis.

B. SDMX User Forum

52. The SDMX User Forum is a platform for knowledge sharing, problem solving and collaboration. It was launched in November 2022 in a collaborative effort of national statistical offices, the SDMX sponsor agencies, and regional commissions of the United Nations, as part of the Statistics Division-led Global Network of Data Officers and Statisticians. The Forum has since become one of the most active communities of the Global Network, with over 600 posts and 42,000 views since the launch. The number of users currently stands at over 430 and continues to grow. The Forum has proven to be a popular and important meeting place for SDMX practitioners, where support is provided by a community of experts, events and initiatives are announced, and training resources are offered. Participation in the Forum is open to all users, irrespective of their level of experience and SDMX knowledge.

C. SDMX website modernisation

53. The official SDMX website hosted at <https://sdmx.org> has successfully provided the entry-point for those new to the standard and resources, news and guidance for the community in general. The website is now being modernised under the leadership of the Secretariat and particularly the ECB who fund the site’s operations. The modernisation is being tackled incrementally starting with improving the presentation of the software tooling and capacity building catalogues and streamlining their maintenance. Significantly, e-learning and tool producers can now submit their own additions and updates under control of the website’s editorial team lead by Eurostat, which is anticipated to keep the content fresh and comprehensive. The next step of the modernisation programme is a refresh of the landing page with updated styling, colour palette and content designed to engage the audience and promote the standard. Concept designs have been agreed with the mock-ups converted into a preview implementation ready for testing in early 2025.

XI. Capacity-building

54. The sdmx.io vision extends beyond the creation and sustainability of an ecosystem of best-in-class tools for the production and use of official statistics, it also includes the provision of knowledge resources to assist official statistics organisations, regardless of their stage of development, to leverage these modern technologies and methods to improve their statistical production capabilities. In short, sdmx.io also emphasises how to reduce risk, improve quality, and increase maturity of data production methods at all stages of the data lifecycle. These goals are met through the provision of targeted short eLearning courses and regular hands-on webinars which address practical official statistics use cases. [Sdmx.io](https://sdmx.io) hands-on webinars consistently attract 100-150 participants and the webinar recordings are made publicly available for consultation by those unable to participate to the live events.

55. The ILO conducts an annual one-week residential training course on LMIS implementation at the International Training Centre in Turin, IT, with about 30 participants from several countries. The course includes an introduction to SDMX, SDMX data modelling and data preparation for extract-transform-load (ETL) on a .Stat Suite platform. In addition, an annual online training course "Introduction to SDMX" is offered.

56. In the Africa region, collaboration between the IMF and African Development Bank (AfDB) on the implementation of SDMX is benefiting from the launch of the AfDB's Open Data Platform (ODP) 2.0. ODP 2.0 allows countries to model their datasets before dissemination, which will increase the usability and openness of these data. To complement the provision of the ODP 2.0 platform, assistance for developing capacity for data modelling will be jointly provided. Datasets in ODP 2.0 become part of a global set of publicly available official statistics. This joint work on modelling will ensure this set aligns with international standards and is AI-ready.

57. The Working Group on SDMX for SDG Indicators has developed and published an online training course on SDMX for the SDGs, with funding provided by the United Kingdom. The training, aimed for those with basic knowledge of SDMX and wishing to familiarize themselves with the global SDG data structures and dataflows, has improved the learning curve, facilitated implementation of SDG-based data exchange and dissemination in the Member States as well as SDG Custodian Agencies, and complemented instructor-led online and offline training provided by the Statistics Division and other international agencies.

58. The Statistics Division has continued to provide online and in-person capacity building in partnership with international organisations including the UN regional commissions, other SDMX sponsor agencies, regional development banks, and others. Online training in SDMX for SDG indicators⁴ has been provided to the SDG Custodian Agencies, which has resulted in a marked increase in the amount of data transmitted as SDMX, reducing the reporting burden on both the reporter and collector. Basic online training in SDMX was delivered to the State Statistical Service of Ukraine (SSSU), which was attended by over 50 participants and has facilitated the implementation of SDMX-based dataflows at SSSU; further training events are planned. In partnership with OECD and the Asian Development Bank, basic training course in SDMX was delivered to 3 Southeast Asian countries in September 2024; this is expected to be the first in a series of increasingly advanced training events. Further training events were provided in partnership with the UN regional commissions including UNECA and UNESCWA.

59. The Asian Development Bank (ADB), in cooperation with the Statistics Division, the Economic and Social Commission for Asia and the Pacific (ESCAP) and the Statistical Institute for Asia and the Pacific, launched SDMX Foundation, an online course, which ran in April and May 2022 and attracted more than 500 participants from across the globe. The course was re-opened in February 2023 and is available for enrolment⁵. In collaboration with the Division, ESCAP, and SIAP, ADB has also launched the SDMX Tools e-Learning course, initially run from November to December 2023. The course was re-opened in July 2024 and is now available on demand⁶.

60. The UN Economic Commission for Africa (UNECA) in partnership with the Statistics Division and the UK Office for National Statistics (ONS), delivered training in modelling and dissemination of SDG indicators to about 7 African Member States as part of a European Union-funded project on the establishment of data dissemination of SDG indicators. All the 7 countries have now established prototype SDG dashboards based on the Open SDG dissemination platform, with several of these expected to be officially launched shortly.

61. The UN Economic and Social Commission for Western Asia (UNESCWA) has launched several key initiatives capacity building and data exchange of SDG indicators. In collaboration with the Statistics Division and other institutions including the Arab Institute for Training and Research in Statistics (AITRS) and the League of Arab States, since 2020 the Commission has organized seven Regional Workshops on SDMX for SDG reporting, aimed at empowering Member States to implement the latest advancements in SDMX automation and data exchange processes. The ESCWA SDMX Converter for SDGs, developed to facilitate data exchange of SDG

⁴ See <https://learning.officialstatistics.org/course/view.php?id=96>

⁵ See <https://elearn.adb.org/course/view.php?id=486>

⁶ See <https://elearn.adb.org/course/view.php?id=520>

indicators, has significantly streamlined the reporting of national SDG data and is currently used by 17 Arab countries, making this information readily available on the Arab SDG Monitor. As a result, data dissemination has surged by at least 40%, data quality has improved, and the reporting burden on the Member States has been substantially reduced. Additionally, UNESCWA has produced a comprehensive guidebook on the SDMX Converter that offers technical guidance on data exchange and validation, further supporting the Member States in enhancing their data infrastructure. The Commission has also enriched the region's resources by providing a detailed SDMX glossary in the Arabic language, introducing 250+ new statistical terms to promote international standards within the statistical community.

62. The UN Economic and Social Commission for Asia and the Pacific (ESCAP) continues to invest in the implementation of SDMX in its region through capacity building projects targeted at establishing national data portals. The Maldives launched its first dissemination platform in October 2023 following two hands-on workshops organized by ESCAP and attended by staff from the national statistics office and other data-providing entities. The modernization of the business process has allowed for streamlined dissemination of data from producers to the public, taking advantage of free and open source tools, such as the .Stat Suite and Fusion Metadata Registry. Building on this experience, ESCAP partnered with the Asian Development Bank and the Pacific Community to launch a joint initiative for Samoa and Fiji, at the request of the two national statistics offices, aimed at applying SDMX to the dissemination processes and building data portals. Additionally, the project fosters stronger technical cooperation among the two countries ensuring the long-term sustainability of the initiative in the region.

63. The .Stat Academy, a product of the SIS-CC under the OECD leadership, provides free online self-paced training to support capacity building in the .Stat Suite and Data modelling in SDMX for data toolers and data producers. The .Stat Academy learning paths have been designed to build knowledge and skills over time, geared towards individual learner goals. As of October 2024, more than 1000 learners have been certified with registered users from 152 countries.

64. The OECD is scheduled to host a regional SDMX capacity building event in December 2024 at its newly established regional Centre in Istanbul, Türkiye with more than 45 participants from 22 organisations. The training is designed to build and enhance SDMX capacity of organisations that are already familiar in the use of SDMX and data modelling practice as part of their daily work and aimed at intermediate level data practitioners. The event is supported by a number of Sponsors and partners including the BIS, UNSD, ILO, SESRIC, Central Bank of the Republic of Türkiye, UNESCO, and UNESCWA.

65. In 2023-2024, the United Nations Environment Programme (UNEP) delivered SDMX training under the Cross Cutting Capacity Development (CCCD) programme funding by the Global Environment Facility (GEF) to Benin and Botswana. The training was carried out in two phases, the first covering the Framework for Development of Environment Statistics (FDES) and the second covering Environment Information Systems where the SDMX information model and a Data Structure Definition (DSD) for environment statistics were introduced. UNEP is scheduled to deliver SDMX training in December 2024 for Statistics Indonesia with a focus on environment indicators identified in the country during the first workshop on FDES in August 2024.

XII. Action to be taken by the Statistical Commission

66. The Statistical Commission is invited to take note of the report.