



# 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**

## Statistical Data and Metadata Exchange 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](#).



## **Report of the Statistical Data and Metadata Exchange sponsors**

### **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 information technology architecture and information technology tools, to be used and implemented by the official statistics community.
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 Statistical Commission has received 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. International Labour Organization joined the Statistical Data and Metadata Exchange 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. ILO had already made significant contributions to the SDMX initiative before expanding its collaboration as a sponsor organization, developing widely used tools and actively participating in working groups and events, including an extended term leading the Statistical Data and Metadata Exchange Technical Standards Working Group. With this new role, ILO joins the Bank for International Settlements, the European Central Bank (ECB), the European Statistical Office (Eurostat), the International Monetary Fund (IMF), the Organisation 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 life cycle 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 under way that are aimed at facilitating the 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 organizations using easily available, open source, off-the-shelf tools.

7. 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 through an application programming interface (API). The first prototype of this application was presented at the ninth SDMX Global Conference, held in Bahrain in October 2023. This was followed by a presentation by EPAM, the vendor that IMF is working with to develop the product, at the SDMX + AI workshop, held in Paris in March 2024, co-organized by OECD and the Bank for International Settlements. 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, IMF released StatGPT 2.0 containing data sets from eight partner organizations, with more than 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. IMF aims to make the tool generally available to statistics publishing institutions, although a sustainable funding model will be required for the large language model operating costs, which scale with usage.

8. In another promising data discovery development, a United Kingdom of Great Britain and Northern Ireland-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, they aim to disambiguate users' natural language queries, extract relevant information and formulate structured queries. Despite some inherent large language model 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 long-standing goal of the sponsors. A Bank for International Settlements-led initiative to publish an SDMX conversational chatbot is aimed at significantly advancing this effort and improving the learning curve for those studying the standard. This large language model-powered chatbot utilizing 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, focused 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. Statistical Data and Metadata Exchange 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 the Italian National Institute of Statistics (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 microdata 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 and international organizations, inter alia, 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 organizations, as well as other users considering the 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 organizations view data as a public good; the Bank for International Settlements 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 Bank for International Settlements has partnered with international organizations (OECD, IMF, Eurostat), national statistics offices (the National Institute of Statistics and Economic Studies (INSEE) of France, central banks (Bank of Italy, National Bank of Belgium), and private sector organizations (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 Bank for International Settlements 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 Bank for International Settlements helps assure relevance through the applicability of tools to practical statistical use cases, and ease of adoption and use through ease-of-deployment approaches and learning resources.

16. These efforts have resulted in increased collaboration and coordination among 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. UNdata**

17. The UNdata modernization project has substantially enhanced the capacity for integrating and disseminating statistical data across the United Nations system in response to the Data Strategy of the Secretary-General for Action by Everyone, Everywhere. Guided by the road map for innovating United Nations data and statistics, this initiative directly supports the SDMX Technical Standards Working Group by developing tools that improve the interoperability of SDMX-compliant data infrastructures. In particular, the UNdata modernization project has created comprehensive mappings to harmonize code lists and concepts used by various United Nations entities, aligned with major international statistical classifications, thereby contributing to the interoperability of SDMX implementations across organizations and statistical domains. In addition, the alignment of UNdata with the Cross-Domain Interoperability Framework will further contribute to the application of the FAIR principles within the SDMX data ecosystem.

18. The focus of the UNdata modernization project on connecting diverse data through knowledge graph methodologies complements the SDMX Technical Standards Working Group's objectives by facilitating the translation and linkage of SDMX-formatted data with open data formats such as Data Documentation Initiative Cross-Domain Integration (DDI-CDI), schema.org and Simple Knowledge Organization System (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 the SDMX 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 Transformations and Expressions 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 information technology intervention. These characteristics allow VTL to be easily integrated in various steps of the statistical process within organizations adopting standards such as SDMX, DDI or Data Point Model (DPM).

20. The Task Force for the Validation and Transformation Language, created in 2013 at the initiative of the SDMX secretariat and responsible for developing and maintaining the language, released VTL version 2.1 in July 2024. The new release contains additional standard time operators and some bug fixes in the language

specifications. 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 present, there are four VTL open-source engines available, implemented by private companies and international organizations. VTL has been adopted by several central banks, national statistical offices and international organizations. A number of initiatives are under way. 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. The latest release of the technical standard, scheduled for the end of 2024, addresses the challenges of managing highly multidimensional data sets in official statistics. Statisticians have struggled to efficiently model data sets with high dimensionality using the SDMX information model or other environments such as a relational database. Such data sets often require the addition of new breakdowns or dimensions over time, which is expensive owing to the disruptive effect on existing data exchange processes. SDMX 3.1 mitigates these issues by allowing fixed subcubes to be defined on data structure definitions which define analytical cubes and data table schemas that remain valid even as the data structure definition adds dimensions. The innovation reduces the need for mixed dimensions and other workarounds, enhances interoperability and enables agile data structure definition 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 visualization 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. Representational state transfer application programming interface (REST API)**

24. The representational state transfer application programming interface (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 API version 1 specification simplified data discovery, retrieval and validation through various queries. SDMX-REST API version 2, 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 such as geospatial code lists and semantic versioning. Further incremental enhancements are scheduled for the fourth quarter of

2024, including “time travel”, sorting and pagination and segmented constraints. Looking ahead, the road map includes new APIs for business functions such as data validation and mapping, and additional alternatives to REST, such as 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 Standards Working Group, Eurostat has developed the Technology Compatibility Kit, which offers a toolkit for testing SDMX REST API implementations for compatibility with the SDMX standard. The Technology Compatibility Kit 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 Technology Compatibility Kit. The latest official release of the Technology Compatibility Kit can test the SDMX markup language (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 Technology Compatibility Kit that include additional features (e.g. the capability of working with the latest specifications of the API or the coverage of additional output formats such as SDMX-JavaScript Object Notation (SDMX-JSON) are regularly made available to the SDMX community.

## **VII. Official Statistical Data and Metadata Exchange Services**

### **A. SDMX Global Registry**

26. The SDMX Global Registry<sup>1</sup> 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 standardized cross-domain concepts such as frequency, unit of measure and observation confidentiality which underpin the drive to improve data set comparability through global harmonization 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 the relocation of the service onto Bank for International Settlements infrastructure.

### **B. SDMX Global Discovery Service**

28. The SDMX Global Discovery Service (SGDS) is a new official SDMX service, to be launched in the second quarter of 2025, aimed at enhancing the discovery and sharing of SDMX artefacts and data sets 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 such as 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 the FAIR

<sup>1</sup> See <https://registry.sdmx.org>.

principles is a uniform resource name (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 is aimed at exposing the new API to potential clients, mainly other services that require the discovery of SDMX endpoints. In addition, the user interface will be made available to the SDMX community under [sdmx.io](https://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 (Technology Compatibility Kit) 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 Economic Commission for Europe (ECE) High-Level Group for the Modernization of Official Statistics has released a report in which it identifies elements of the SDMX and DDI standards that are relevant in each phase and subprocess of the Generic Statistical Business Process Model (GSBPM), from the specification of statistical needs to the dissemination of statistics.<sup>2</sup> 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 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 the 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 organizations 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 classification maintenance. Complemented by advances in data

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



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, jointly with partners, has been working on the following content-oriented guidelines, of which some 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 aimed at statistical data modellers working with multidimensional data, primarily in SDMX – but with a scope that can go beyond (e.g. data warehousing in general). The patterns proposed are anchored in existing and well-established international standards in the fields of science and engineering.
- Guidelines for SDMX hierarchies. The aim of the guidelines 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. The 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. The revision adds SDMX 3.0 features and new annotations to the controlled vocabulary, thereby increasing interoperability.
- Guidelines for working with vintages in SDMX. In the 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 life cycle), 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 ontologies. As data standards mature, there is a growing need for interoperability; this could be viewed simply as a way to express information from one standard in the other. While this could be used to widen the audience for the same information, it poses a deeper question, as to whether 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 unique resource identifiers for SDMX objects.
- Microdata for SDMX 3.0. New guidelines 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 and Reporting Tool (SMART) and LMIS Excel Add-in. In the past 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 15 projects (13 national and 2 regional systems) are under way. 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. 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 seven countries in the implementation of SDMX-enabled National Summary Data Pages, three adherents to the Enhanced General Data Dissemination Standard, two subscribers to the Special Data Dissemination Standard, and two subscribers to the Special Data Dissemination Standard Plus. As at the end of 2024, there are 112 countries disseminating data in SDMX format through their National Summary Data Pages.

37. As part of the revision of the *International Merchandise Trade Statistics 2010* and the *Manual on Statistics of International Trade in Services 2010* by the task team on international trade statistics of the Committee of Experts on Business and Trade Statistics, a guidance note on SDMX for trade in goods and services<sup>3</sup> 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 data structure definitions taking account of forthcoming versions of the *International Merchandise Trade Statistics* and the *Manual* and strategies for the implementation of SDMX in the *International Merchandise Trade Statistics*, 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. The task team will continue to implement the plan outlined in the guidance notes in 2025 and beyond.

38. Global SDMX data structure definitions in the area of macroeconomic statistics continue to be maintained and used for data exchange. In particular, in 2024 the recently finalized Classification of Individual Consumption According to Purpose (COICOP) 2018 classification was implemented in several of these global macroeconomic 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 data structure definition for residential property prices is also under development, coordinated by the Bank for International Settlements.

39. As part of the work of the Committee of Experts on Environmental-Economic Accounting, the Statistics Division, in collaboration with OECD, began regular global data collection for air emission and energy accounts in accordance with the System of Environmental-Economic Accounting 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 System of Environmental-Economic Accounting data structure

<sup>3</sup> Available at [https://unstats.un.org/wiki/download/attachments/311230515/GNV.15\\_SDMX-IMTS-MSITS.pdf?version=1&modificationDate=1735849291423&api=v2](https://unstats.un.org/wiki/download/attachments/311230515/GNV.15_SDMX-IMTS-MSITS.pdf?version=1&modificationDate=1735849291423&api=v2).

definitions. Similar questionnaires have also been made available by Eurostat for European Statistical System countries. The Statistics Division disseminates the data through its SDMX application programming interface. In addition, in 2024 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 System of Environmental-Economic Accounting data for 42 Member States.

40. The Working Group on Statistical Data and Metadata Exchange for Sustainable Development Goal Indicators has continued to maintain the global data and metadata structures and dataflows for Goal indicators. SDMX data exchange with the Goal custodian agencies has significantly increased; more than 40 per cent of the global Goal database is now transmitted as SDMX data sets. Voluntary provision of Goal data sets by the Member States has also continued, with data exchange having been established with some 40 countries. The global Goal data set, as well as metadata for the Goal indicators, are disseminated through the Statistics Division SDMX application programming interfaces. The availability of metadata in a machine-readable format made it possible to link Goal indicators to their metadata at the Global Goal Database website and display the data alongside metadata. The Working Group has also contributed to the development of the popular Open SDG<sup>4</sup> platform, maintained by the Office for National Statistics of the United Kingdom, which supports the global Goal data and metadata structures and which is used by some 18 Member States to disseminate their Goal indicators. Experience gained in the Goal 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 Goal 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 1,700 downloads and SMART nearly 1,100. A new Excel add-in tool, 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. ILO has developed a simplified data modelling methodology for the implementation of labour market information systems, 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, data structure definitions, 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 the Special Data Dissemination Standard and the Special Data Dissemination Standard

<sup>4</sup> See <https://open-sdg.org/>.

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 common approach to prioritizing the development of a single, unified platform. FMR is a free and open-source platform managed by the Bank for International Settlements.

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 proved 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 great deal 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 data structure definition for several dataflows. It is widely used by sponsors and the SDMX community at large and has proved particularly useful in the initial stage of conceptualizing data models. The global data structure definitions 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 serialization, uncoded dimensions, multiple concept scheme references, custom concept roles and 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 Statistical Information System Collaboration Community (SIS-CC) (see para. 48) saw the need to ease the sourcing and visualization 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 has been 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 organizations worldwide as a data modelling platform and controlled repository for statistical metadata, helping to improve metadata governance, maintainability, standardization and harmonization. In addition to deployment within institutions, FMR also powers several public-facing SDMX structural metadata services, including the SDMX Global Registry and IMF Central. 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 to ease the management of structural metadata in the SIS-CC .Stat Suite, widely used for statistical data dissemination. Now, the Bank for International Settlements FMR team, in collaboration with colleagues from the OECD and the 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 materialized as SDMX artefacts for use in practice. The project is currently securing co-investors to fund the build stage, with a minimum viable project 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 that 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 that 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 put into supporting key SDMX 3.0 functionalities, such as array values and multiple measures. Eurostat 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 life cycle for official statistics in more than 50 organizations around the globe. It is developed and maintained by the Statistical Information System Collaboration Community (SIS-CC), a global community of more than 20 members and partners under the leadership of OECD. The .Stat Suite is a platform to manage the data life cycle 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 past two years, the community has initiated the work, together with the Bank for International Settlements and Eurostat, on a reference SDMX implementation. In addition, a number of projects have gone online, launching the .Stat Suite in several statistical offices and international organizations, including Thailand, El Salvador, Uruguay, Greece, Malta, Luxembourg, Madagascar, Maldives and the Food and Agriculture Organization of the United Nations. The .Stat Suite is also serving as a core component for the SDMX data modelling initiative, led by OECD. In support of its continued push to adopt open source, the Community has implemented support for the open source database MariaDB, which 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 ninth SDMX Global Conference, held in Bahrain from 29 October to 2 November 2023, 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 coronavirus disease (COVID-19) pandemic, was co-organized by the SDMX sponsors and the Information and eGovernment Authority of Bahrain, with the support of the Economic and Social Commission for Western Asia (ESCWA) and the 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 highlighted the significance of SDMX in the world of data management and governance and emphasized 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 twelfth SDMX Experts workshop was held in Amsterdam from 7 to 11 October 2024. 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-organized by the SDMX sponsors. A total of 11 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 that 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 more than 600 posts and 42,000 views since the launch. The number of users currently stands at more than 430 and continues to grow. The Forum has proved 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 modernization**

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 modernized under the leadership of the secretariat and, in particular, ECB, which funds the site's operations. The modernization 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 the control of the website's editorial team, led by Eurostat, which is anticipated to keep the content fresh and comprehensive. The next step in the modernization 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](https://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 organizations, regardless of their stage of development, in leveraging these modern technologies and methods to improve their statistical production capabilities. In short, sdmx.io also emphasizes how to reduce risk, improve quality and increase maturity of data production methods at all stages of the data life cycle. These goals are met through the provision of targeted short e-learning courses and regular hands-on webinars, which address practical official statistics use cases. Sdmx.io hands-on webinars consistently attract 100 to 150 participants, and the webinar recordings are made publicly available for consultation by those unable to participate in the live events.

55. ILO conducts an annual one-week residential training course on labour market information systems implementation at the International Training Centre in Turin, Italy, with some 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 IMF and the African Development Bank (AfDB) on the implementation of SDMX is benefiting from the launch of the AfDB Open Data Platform (ODP) 2.0. ODP 2.0 allows countries to model their data sets 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. Data sets in ODP 2.0 become part of a global set of publicly available official statistics. This joint work on modelling will ensure that this set aligns with international standards and is AI-ready.

57. The Working Group on Statistical Data and Metadata Exchange for Sustainable Development Goal Indicators has developed and published an online training course on SDMX for the Goals, with funding provided by the United Kingdom. The training, aimed at those with a basic knowledge of SDMX and wishing to familiarize themselves with the global Goal data structures and dataflows, has improved the learning curve, facilitated implementation of Sustainable Development Goal-based data exchange and dissemination in the Member States, as well as Sustainable Development Goal 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 organizations, including the United Nations regional commissions, other SDMX sponsor agencies and regional development banks. Online training in SDMX for Sustainable Development Goal indicators<sup>5</sup> has been provided to the Sustainable Development Goal 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 the collector. Basic online training in SDMX was delivered to the State Statistical Service of Ukraine, which was attended by more than 50 participants and has facilitated the implementation of SDMX-based dataflows at the State Statistical Service; further training events are planned. In partnership with OECD and the Asian Development Bank, basic training course in SDMX was delivered to three South-east 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 United Nations regional commissions, including the Economic Commission for Africa (ECA) and ESCWA.

59. The Asian Development Bank (ADB), in cooperation with the Statistics Division, the Economic and Social Commission for Asia and the Pacific (ESCAP)

<sup>5</sup> See <https://learning.officialstatistics.org/course/view.php?id=96>.



and the Statistical Institute for Asia and the Pacific, launched the SDMX foundation course, an online course which ran in April and May 2022 and attracted more than 500 participants from around the globe. The course was reopened in February 2023 and is available for enrolment.<sup>6</sup> 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 reopened in July 2024 and is now available on demand.<sup>7</sup>

60. ECA, in partnership with the Statistics Division and the Office for National Statistics of the United Kingdom, delivered training in modelling and dissemination of Goal indicators to seven African Member States as part of a European Union-funded project on the establishment of data dissemination of Sustainable Development Goal indicators. All seven countries have now established prototype Goal dashboards based on the Open SDG dissemination platform, with several of these expected to be officially launched shortly.

61. ESCWA has launched several key initiatives for capacity-building and data exchange of Sustainable Development Goal indicators. In collaboration with the Statistics Division and other institutions including the Arab Institute for Training and Research in Statistics and the League of Arab States, since 2020 the Commission has organized seven regional workshops on SDMX for Goal reporting, aimed at empowering Member States to implement the latest advancements in SDMX automation and data exchange processes. The ESCWA SDMX Converter for the Goals, developed to facilitate data exchange of Goal indicators, has significantly streamlined the reporting of national Goal data and is currently used by 17 Arab countries, making this information readily available on the Arab Goal Monitor. As a result, data dissemination has surged by at least 40 per cent, data quality has improved, and the reporting burden on the Member States has been substantially reduced. In addition, ESCWA has produced a comprehensive guidebook on the SDMX Converter which provides 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 Arabic, introducing 250+ new statistical terms to promote international standards within the statistical community.

62. ESCAP continues to invest in the implementation of SDMX in its region through capacity-building projects targeted at establishing national data portals. 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 the 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. In addition, 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 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

<sup>6</sup> See <https://elearn.adb.org/course/view.php?id=486>.

<sup>7</sup> See <https://elearn.adb.org/course/view.php?id=520>.



individual learner goals. As at October 2024, more than 1,000 learners have been certified with registered users from 152 countries.

64. 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 organizations. The training is designed to build and enhance the SDMX capacity of organizations that are already familiar with the use of SDMX and data modelling practice as part of their daily work and is aimed at intermediate-level data practitioners. The event is supported by a number of sponsors and partners, including the Statistics Division, ESCWA, ILO, UNESCO, the Central Bank of the Republic of Türkiye, the Bank for International Settlements and the Statistical, Economic and Social Research and Training Centre for Islamic Countries.

65. In 2023–2024, the United Nations Environment Programme (UNEP) delivered SDMX training under the Cross-Cutting Capacity Development programme funding by the Global Environment Facility to Benin and Botswana. The training was carried out in two phases, the first covering the Framework for the Development of Environment Statistics and the second covering environment information systems, at which the SDMX information model and a data structure definition 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 the Framework in August 2024.

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

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

---