Irving Fisher Committee on
Central Bank Statistics

BIS

IFC Report
No 17

SDMX adoption and use of open source tools

Results of an IFC survey of central banks

February 2025

Contributors to the IFC report[1]

Bank for International Settlements (BIS)	Bilyana Bogdanova
	Brian Buffett
	Bianca Ligani
	Ismail Mustafi
	Stratos Nikoloutsos
	Rafael Schmidt
BIS & Irving Fisher Committee on Central Bank Statistics (IFC)	Bruno Tissot

This publication is available on the BIS website (www.bis.org).

ISSN 1991-7511 (online)

ISBN 978-92-9259-817-4 (online)

[1] Respectively, Deputy Head of Departmental Research Support and Head of Macroeconomic and Statistical Support (Bilyana.Bogdanova@bis.org), Business Analyst (Brian.Buffett@bis.org), Macroeconomic Analyst (Bianca.Ligani@bis.org), Data Scientist (Ismail.Mustafi@bis.org), Principal Software Engineer (Stratos.Nikoloutsos@bis.org), Head of IT, Monetary and Economic Department (Rafael.Schmidt@bis.org), and Head of BIS Statistics and Head of the IFC Secretariat (Bruno.Tissot@bis.org).

The views expressed are those of the authors and do not necessarily reflect the views of the BIS, the IFC or its members or the other institutions mentioned in this report.

We thank Douglas Araujo and Olivier Sirello for comments and suggestions and Márcia Cavalinhos and Ilaria Mattei for editorial support.

Contents

Executive summary ... 1

1. Introduction ... 2

2. SDMX – state of adoption ... 4

3. Using SDMX across the various phases of the data lifecycle ... 6

Box A: SDMX + AI: unlocking the potential of NLP to enhance data access ... 8

Box B: Enhancing the interoperability of SDMX tools for improved organisational efficiency ... 10

Box C: SDMX metadata-driven approach: the BIS experience ... 11

4. SDMX tools: state of adoption and remaining gaps ... 12

Box D: The BIS Open Tech initiative and the sdmx.io ecosystem ... 14

5. SDMX by statistical domain ... 15

Box E: Global data structure definitions – an application to property prices ... 19

References ... 20

Annex 1: IFC survey questions ... 23

Annex 2: List of IFC jurisdictions that responded to the survey ... 25

Executive summary

A key international initiative of relevance to central bank statistics is the Statistical Data and Metadata eXchange (SDMX) standard. **SDMX is an ISO standard that facilitates the seamless exchange, production and dissemination of statistical data and metadata** between and within organisations. This standard provides an integrated approach enabling interoperable implementations for information systems dealing with the exchange and processing of statistical data and their related metadata.

This report presents the results of a comprehensive **survey on the adoption of and interest in SDMX among central banks**. The objectives were threefold: to assess the current state of SDMX implementation, measure its adoption rate and gain insights into the usage and requirements of SDMX tools across various statistical processes. The survey results, summarised in this report, provide a detailed understanding of central banks' experiences and needs related to SDMX.

A main highlight is that there has been a **notable and significant increase in SDMX adoption** among central banks in the recent decade. The standard is now implemented by about eight tenths of respondents, compared with two thirds 10 years ago. Looking across the main statistical domains, adoption of SDMX has been particularly significant in five key areas: government finance, financial accounts, monetary and financial statistics, financial markets and securities and balance of payments.

One key factor driving SDMX adoption has been that **its information model is tailored specifically for metadata-driven data processing**. This enables central banks to standardise their statistical processes, resulting in better governance and operational efficiency. It also offers greater flexibility and agility, as modifications to the data model can be easily made through metadata configuration alone, without disrupting the processing setup or implementation, thereby supporting more adaptable data management.

Second, the **SDMX standard enables interoperability**, which is a crucial element in supporting central banks to become more data-driven. A key reason for this is that the possibility to integrate and connect different tools seamlessly is essential to manage and exchange data effectively, which in turn helps institutions enhance their data management strategies.

A third, related lesson is the importance of statistical standards such as SDMX in improving the **findability**, **accessibility**, **reusability and openness of data**. This is particularly relevant for central bank statistics, not least in order to maximise their impact and ensure that they adequately address growing multidimensional user needs, particularly those expressed by policymakers.

A fourth important factor underscored by the survey is that there is a **rich ecosystem of publicly available tools and software** that can be used to facilitate SDMX implementation in all stages of the data lifecycle, especially in data production, exchange and dissemination. In particular, most central banks reported using multiple open source tools and expressed satisfaction with both the tools and the associated support from the SDMX community from which they can benefit.

However, a main obstacle to both the initial adoption and the transition to new versions of SDMX remains the **lack of adequate resources and internal expertise**. This makes adapting internal IT tools to the standard difficult. Moreover, important gaps persist in the available software despite the increased global offering of open source tools, hindering the implementation of SDMX across the various stages of the data lifecycle.

A second key challenge is that data reporting requirements are barely keeping up with changes to SDMX, and as a result individual central banks are **operating different versions of the standard** depending on the data compiling agency and statistical area. This limits the identification of potential synergies across statistical domains, raising the risk of inefficient use of resources. These shortcomings

underscore the need for statistical compilers, including central banks, to stay up to date with the latest version of the SDMX standard and to rely on associated tools that are adequately updated.

Looking ahead, the survey shows that central banks can proactively address existing challenges and take greater advantage of the possibilities offered by SDMX to enhance the way they produce and access statistics. First, SDMX can serve as a catalyst to **harness artificial intelligence (AI) capabilities** in central banks, enabling more automated data processing, improved data quality and streamlined data discovery.[2] Second, there are increasing opportunities provided by **international cooperation and knowledge exchange**, allowing central banks to benefit from both the experience reported by other institutions and the availability of adequate learning resources when developing their own SDMX toolkit. A case in point relates to sdmx.io,[3] the recently launched platform which addresses common SDMX-related challenges by offering open source statistical software, learning materials and practical guides.

1. Introduction

In an era of rapid technological advancements and evolving data sources, statisticians face pressure to increase the efficiency, interoperability and timeliness of their processes. At the same time, one needs to ensure the quality of the information produced, in line with the Fundamental Principles of Official Statistics (UN (2014)). Fortunately, by facilitating the seamless exchange, production and dissemination of statistical data and metadata between and within organisations, the SDMX[4] ISO standard (ISO (2023)) triggered and supported by international organisations can be a **key element supporting the production of official statistics as a global public good** (ADB (2024), Tissot (2021)).

Indeed, a statistical standard like SDMX can play an instrumental role in **securing the accessibility, findability, integrity and openness of official statistics**. In a context marked by fast-paced technological advancements, a robust information model such as the one offered by SDMX can help statisticians to augment their data with additional metadata, facilitating for instance the correct training of AI systems (DOC (2025)). In addition, better interoperability with other statistical standards will facilitate the sharing of data, software, resources and knowledge.

The **benefits of the SDMX standard are particularly relevant for central banks** that rely on the availability of structured and well documented statistics on the economy and the financial system to conduct their policies. Their data collection, compilation, sharing and dissemination exercises have become even more critical since the 2007–09 Great Financial Crisis (GFC) and in response to the need to address pressing information gaps (IFC (2023a)). In this endeavour, they have been able to leverage the SDMX standard to streamline and facilitate their statistical work, particularly as regards IT-related aspects (Tissot (2017)). These benefits have been recognised in particular in the context of the Data Gaps Initiative (DGI) endorsed by the G20 after the GFC.[5]

SDMX comprises two primary components: a robust information model and a detailed technical standard. The information model provides a comprehensive framework for describing fundamental concepts related to statistical data, metadata and data exchange processes, making it applicable to any multidimensional data set regardless of the domain. The technical standard offers detailed specifications for developing IT tools that support the SDMX information model throughout the

[2] It has been argued for instance that the advent of generative AI holds great potential for the editing of metadata in official statistics (Sirello (2024)).

[3] See www.sdmx.io.

[4] See www.sdmx.org.

[5] In particular, recommendation 14 of the third phase of the G20 DGI focuses on enhancing micro data-sharing with the aim of working towards an international micro data standard in coordination with SDMX (IMF et al (2023)).

data lifecycle. Furthermore, SDMX implementation is facilitated by the sharing of international best practices and common specifications, such as domain-specific data models and cross-domain codelists (SWG (2023), sdmx.io (2023)).

Significant progress has been achieved since the original launch of the SDMX initiative by a number of sponsor international organisations in 2001.[6] This was already highlighted by the general survey organised in 2015 and targeted at the international community of official statistics. On that occasion, the IFC decided to examine in more detail the specific situation of central banks, with a focus on three key areas: (i) their experience with the SDMX standard; (ii) their level of satisfaction; and (iii) their expectations for potential developments in the standard (IFC (2016)).

It was found that **SDMX adoption was already widespread among central banks 10 years ago**, with two thirds of jurisdictions utilising the standard and the majority finding it beneficial. Central banks were primarily focusing their efforts on data dissemination, including reporting, and were already exchanging or sharing data with international organisations using the SDMX standard.[7] At the same time, central banks reported important challenges when adopting SDMX, such as insufficient human resources and training capacities as well as difficulties in obtaining support from subject-matter statisticians. These constraints were reflected in the fact that data exchange was mainly concentrated in statistical domains covered by well established, internationally agreed data structure definitions (DSDs), such as balance of payments, national accounts or foreign direct investment.[8]

Since then, an increasing number of central banks across the world have adopted and made effective use of the SDMX standard. Moreover, **a key priority has been to expand the availability of global DSDs in various domains** across macroeconomic and finance statistics, international trade statistics, business statistics and prices (including property prices). Furthermore, the availability and quality of SDMX resources, in particular related software tools, have significantly improved. A concrete milestone was reached with the release of the new SDMX version 3.0 in 2021, which introduced important features for handling geospatial and micro data, thus opening new venues for expanding the implementation of the standard beyond traditional macroeconomic aggregates.[9]

To assess progress, **the IFC decided in 2023 to update its 2015 survey for evaluating the state of SDMX implementation among central banks**, measuring the degree of adoption of the standard and related barriers, and assessing related software tools' offerings and usage. It also decided to delve into the full statistical data lifecycle and statistical domains, allowing for a more granular analysis compared with the exercise performed a decade earlier.

This **report outlines the main conclusions of the new IFC survey**. Section 2 summarises central banks' adoption of the SDMX standard and the key barriers faced in this endeavour. Section 3 describes the usage of SDMX throughout the data lifecycle, from collection to dissemination. Section 4 documents the adoption and use of open source SDMX tools. Section 5 reviews the use of SDMX by statistical domain.

6 As of January 2025, the SDMX sponsor organisations are the BIS, the European Central Bank (ECB), Eurostat, the International Labour Organization (ILO), the International Monetary Fund (IMF), the Organisation for Economic Co-operation and Development (OECD), the United Nations (UN) and the World Bank Group (WB). For a more recent stocktaking of SDMX initiatives and governance, see ECOSOC (2025).

7 For instance, countries participating in the key IMF data standards initiatives, such as SDDS (Special Data Dissemination Standard) and SDDS Plus, can use the automatic exchange and sharing of statistical data and metadata in SDMX to facilitate machine-to-machine transmission; see IMF (2013) and Dissemination Standards Bulletin Board (DSBB).

8 These DSDs provide the necessary concepts for describing and identifying the data; for further details on the definition of DSD and SDMX terminology more broadly, see SWG (2020).

9 For additional information on the major changes and new functionalities introduced with SDMX 3.0, see TWG (2021).

2. SDMX – state of adoption

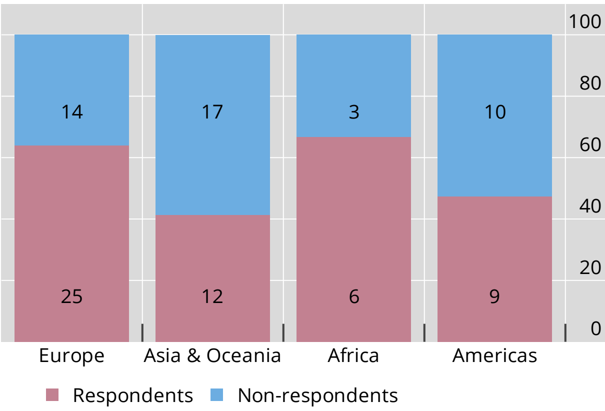
The 2023 survey received responses from **52 jurisdictions**, representing about half of IFC members[10] with broad geographical coverage (Graph 1.A).

Central banks' adoption of the SDMX standard

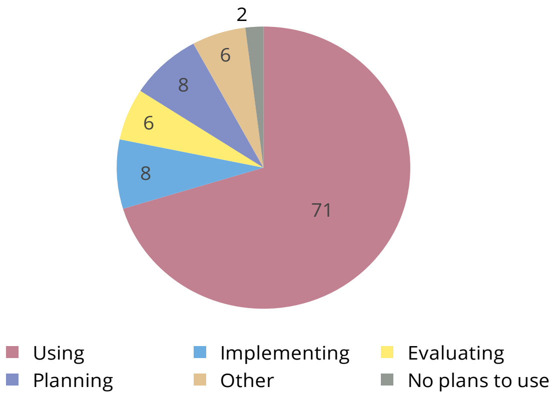
In per cent

Graph 1

A. Geographical coverage[1]



B. Intention to use SDMX



[1] The figures on the bars are expressed as number of respondents.

Sources: IFC survey of SDMX tools (2023); authors' calculations.

The survey confirms a **high degree of SDMX adoption, with 79% of responding central banks *using or implementing* SDMX**. A further 14% are *evaluating* or *planning to implement* the standard (Graph 1.B). This compares with 64% and 13% respectively in 2015 (IFC (2016)).

For those central banks not using SDMX, the **key barriers to adoption** are reported to be the *need to adapt internal tools* and *lack of resources* (Graph 2.A). Having adequate IT tools (internal, open source or commercial) thus appears to be a key priority, notably in Africa and Asia. In addition, central banks emphasised the importance of training and technical assistance to facilitate adoption of the standard. A related lesson of the survey is the need to increase SDMX knowledge, deemed *extremely or very relevant* to ensuring its actual implementation (see Section 3).

The survey also shed light on **the diversity of SDMX versions adopted by central banks**, partly reflecting legacy issues and the significant evolution of the standard observed over the past decades. SDMX 2.1, released in 2011, is currently the dominant version of the standard in use, with SDMX 2.0 still experiencing a relatively high level of usage (Graph 2.B).

More recently, **the standard achieved a significant milestone with the introduction of the 3.0 specification, released in 2021** (TWG (2021)). This version introduced major features to address evolving business needs[11] and support a greater variety of data types, such as granular and geospatial information. It also added important changes to the SDMX information model, notably for enhancing

[10] For further details on IFC membership, see bis.org.

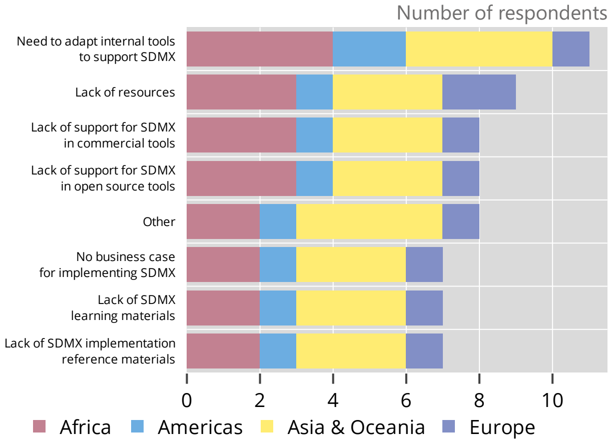
[11] To better respond to users' and producers' needs, the SDMX sponsor organisations regularly review their strategic priorities through "Roadmaps" over a period of five years. The latest edition, SDMX Roadmap 2021-2025, is available at sdmx.org.

metadata, as well as to the SDMX application programming interface (API) and SDMX data (exchange) formats.

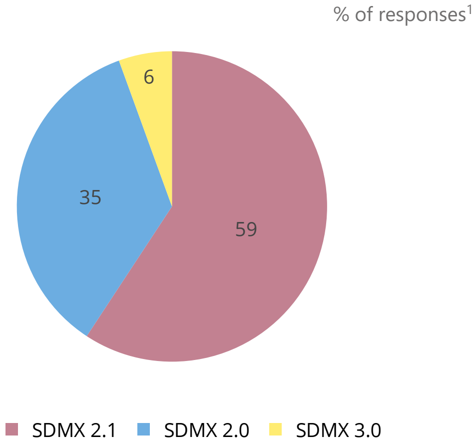
Barriers to SDMX adoption and versions in use

Graph 2

A. Barriers to adoption



B. Versions in use



[1] Some respondents were using more than one version.

Sources: IFC survey of SDMX tools (2023); authors' calculations.

Yet, the survey shows that **SDMX 3.0 has so far been adopted by only 6% of responding central banks** (Graph 2.B). This reflects a number of factors, namely the *lack of resources and insufficient SDMX support in IT tools* (Graph 3.B). Moreover, the *lack of compelling business justification* for upgrading to the new version was frequently cited as a major deterrent. Central banks also indicated that they would delay upgrading until recipient organisations explicitly required SDMX 3.0-compliant data submissions, thereby justifying the investment. Nevertheless, of those central banks using older SDMX versions, about two thirds intend to migrate to SDMX 3.0 in the future (Graph 3.A). An important reason for this, reported in one fourth of the cases, is the ability to deal with micro data with the new version.

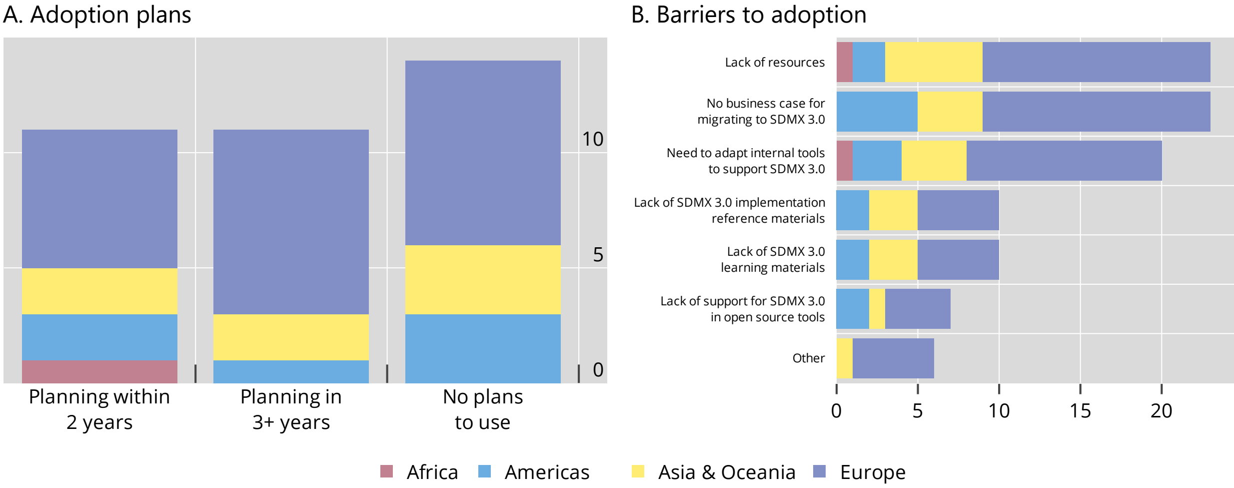
The survey also assessed **the adoption rate of other standards for dealing with statistical data and metadata**. A key one is **XBRL** (eXtensible Business Reporting Language), primarily used for regulatory financial reporting.[12] Notably, 52% of responding central banks are either using or implementing XBRL, with a further 6% evaluating or planning to implement it. The main XBRL applications identified comprise the ingestion of data received from banks or other reporting agencies and the submission of supervisory data.

Another – closer to SDMX – standard is VTL (Validation and Transformation Language), which is a standard complementary to SDMX that focuses on advanced data validation and transformation capabilities. Although the supply of VTL tools is currently limited, development and integration with SDMX is under way and important proof-of-concept use cases are being explored. A significant part of central banks appear to be interested in VTL, with 16% indicating that they are using it or plan to use it within the next two years.

[12] For example, the European Banking Authority (EBA), the ECB and the European Insurance and Occupational Pensions Authority (EIOPA) are using XBRL along with the Data Point Model for supervisory reporting by banks. For additional details, see EBA et al (2024).

Central banks' adoption of version 3.0 of the SDMX standard

Number of respondents Graph 3



Sources: IFC survey of SDMX tools (2023); authors' calculations.

3. Using SDMX across the various phases of the data lifecycle

The SDMX standard features a comprehensive information model which covers the key concepts around statistical data, metadata and data exchange processes. This information model can be applied to describe any multidimensional data set regardless of the domain. Moreover, it encompasses not only technical specifications such as data formats, structures and functionalities, but also statistical guidelines such as concepts, methodologies and classifications. These "artefacts" of the SDMX model can be mapped to cover *all* phases of the statistical lifecycle (SWG (2016)), in particular the whole set of business processes needed to produce official statistics, as described in the Generic Statistical Business Process Model (GSBPM) (Muñoz (2020), UNECE (2024b)). To shed light on this topic, **the survey comprehensively examined the key phases of the data lifecycle**, from design to collection, reporting, processing and dissemination. The aim was to analyse the perceived relevance and advantages of the SDMX standard for each of these phases and to identify both gaps and potential synergies in tools and learning resources.

An overwhelming 81% of jurisdictions reported SDMX as *extremely or very relevant* for the data reporting phase (Graph 4.A). This reflects the fact that, for most central banks using SDMX, the initial decision to adopt the standard was driven by the need to report data to recipient (international) organisations. This should not come as a surprise given that major global initiatives for developing global data exchanges have historically relied on SDMX, especially in the areas of the Sustainable Development Goals (SDGs), the International Merchandise Trade Statistics (IMTS), the IMF data standards (eg SDDS and SDDS Plus) and the G20 DGI (UNECE (2020), ECOSOC (2025), IMF et al (2023)).

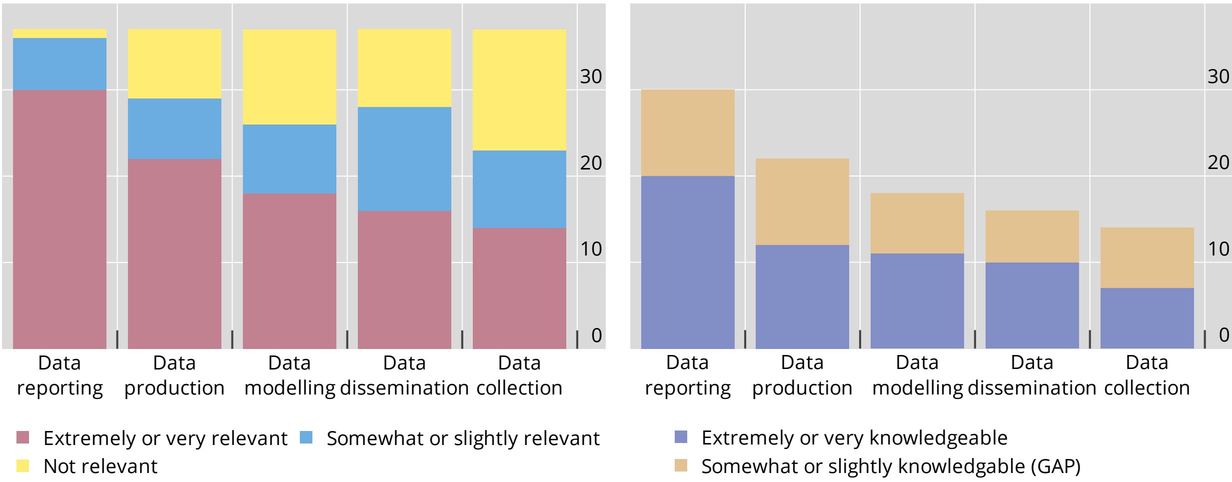
However, **central banks have over time developed a more comprehensive approach**, leveraging SDMX to support the various other stages of the statistical production process. Several factors can explain this more diversified use of SDMX across the statistical chain. One is that the standard has over time shown a high degree of agility and flexibility, with the development of adequate resources, guidelines and specifications (UNECE (2024b)).

Relevance and knowledge of SDMX in statistical processes

Number of respondents — Graph 4

A. Relevance by statistical process

B. Knowledge of respondents rating SDMX as extremely or very relevant, by statistical process[1]



[1] The main knowledge gap to be addressed refers to respondents who rated SDMX as extremely or very relevant but are only somewhat or slightly knowledgeable.

Sources: IFC survey of SDMX tools (2023); authors' calculations.

Looking more specifically at the various phases of the data lifecycle, 59% of the surveyed institutions using SDMX rated the standard as ***extremely or very relevant* for data *production***. To a lesser extent, relevance was also rated as relatively high for data *modelling* and data *dissemination*. Turning to data *collection*, a more limited yet still significant part of the respondents (38%) rated the standard as – at least – very relevant. As noted above, this lower score may be explained by the presence of competing standards for collecting statistics: examples include XBRL for supervisory statistics, ISO 20022 for payments data (particularly for cross-border transactions; see CPMI (2025)), the Data Documentation Initiative (DDI) for describing micro data and dealing with record-level information eg from surveys, and the various standards such as ISO 19115 and 19111 Geographic Information used for geospatial data not least to deal with climate change issues (UNECE (2025)).

One consequence of the different degrees of SDMX adoption across the data lifecycle is that **central banks often work with multiple standards and formats across the different production stages**, converting their data sets into SDMX only at the final stage before reporting. This suggests that, in order to fully leverage the benefits of the SDMX information model, central banks would need to integrate the standard more comprehensively throughout all stages of the data lifecycle, rather than adopting it solely for the purpose of final data reporting.

Lastly, **SDMX can also be used beyond statistical processes**, thanks to its versatile and flexible information model. One important area is data science, where the standard can effectively support the implementation of machine learning techniques (IFC (2023b)) for instance to conduct data quality checks (Fajt Mayer et al (2024)). SDMX can also support AI-based research activities, for instance by improving data accessibility (see Box A) and facilitating the assembling of vast data sets for training models.[13] Lastly,

[13] For instance, Araujo (2023) shows the benefits of SDMX for economic analysis through an open source library that retrieves standardised data ready to be used.

the standard can support innovative ways for communicating official statistics (IFC (2024a)), for instance by backing the dissemination of open data in line with best practices (eg the FAIR principles; see Gregory (2024), Wilkinson et al (2016)) as well as outreach initiatives to non-expert audiences – eg through its smooth integration with tools such as data portals that enable easy data queries and APIs; see Laaboudi et al (2024).

Box A

SDMX + AI: unlocking the potential of NLP to enhance data access

Eric Anvar (OECD) and Rafael Schmidt (BIS)

The intersection of SDMX and AI, in particular natural language processing (NLP), has the potential to transform the way organisations access statistical data. To shed light on this issue, the Organisation for Economic Co-operation and Development (OECD) and the BIS have worked together to evaluate the main use cases of NLP, for instance to enhance data access, explore related IT solutions and identify areas for collaborative development.①

One topic of key interest to the SDMX community is how to improve data dissemination, in particular when dealing with multiple SDMX sources. Ongoing projects suggest that their access can be greatly facilitated through an AI-enabled "universal data concierge", specifically by leveraging generative AI techniques and drawing on the rich SDMX metadata. One central cooperative project in this context is to improve **the accessibility of SDMX data via natural language search and discovery techniques.** This is very important for producers of official statistics, as accessing their data remains difficult for end users. Many features can be developed in this endeavour, in addition to the natural language and conversational exploration of data, for instance by scoring the relevance of search results, enriching metadata or leveraging generative AI for code generation.

Ideally, **an international cooperative approach to delivering such SDMX+AI data accessibility features should involve LLMs** to generate or extract information and start with the production grade of a "minimum viable product". The solution should be inclusive, open source and vendor-agnostic. Most importantly, the capacity of LLMs to reliably interpret SDMX metadata is crucial to ensure a meaningful user experience and the possibility of exploring multiple SDMX sources with equal accuracy.

Fortunately, **there are already a number of international initiatives in this area**. One is the IMF's StatGPT project to modernise its statistics processing and dissemination platforms, which could be extended to other interested stakeholders.② Another possibility would be to develop an open source approach, and the SDMX community will continue to actively explore this issue in the near future.

① See SIS-CC (2024b). ② See UNECE (2023) and Kroese (2024).

Looking ahead, how can the implementation of SDMX be further expanded across the various phases of the data lifecycle? The above developments suggest that **a key objective would be to ensure greater interoperability across the various existing standards**, noting that each of them can provide specific benefits depending on the stage of the statistical chain being considered (UNECE (2024c)). This calls for making available tools that allow for the conversion from one standard into another, not least to streamline data exchange across different reporting systems. And, indeed, 25% of responding central banks indicated that XBRL-to-SDMX conversion is a *moderate priority*.

In addition to the development of conversion tools, **a key objective is to develop a shared understanding across the various data stakeholders on the key aspects required for data exchange and sharing**, such as semantics, representation, roles and governance models (UNECE (2024a)). This would address the lack of SDMX knowledge highlighted by the survey (Graph 4.B) and can be achieved through both comprehensive mapping exercises between standards and the development of cross-domain interoperability frameworks (Gregory et al (2024)).

Furthermore, **advancing global SDMX adoption calls for a focused approach targeting the specific business needs associated with each main phase of the data lifecycle**. Fortunately, the standard allows for the adoption of a modular, flexible approach instead of the development of single, monolithic "end-to-end" IT solutions. This makes it possible to identify and prioritise learning needs across the entire data lifecycle and ensures that all IT tools provided by the SDMX community have the necessary level of interoperability (see Box B).

Usage of open source SDMX tools by stage of data lifecycle

Tools used by at least one respondent central bank for at least three years — Graph 5

Data modelling	Data collection	Data production	Data dissemination	Data reporting
API SDMX	Bulk Data Reporting	Bulk Data Reporting	.Stat Suite	.Stat Suite
Bulk Data Reporting	FMR	FMR	API SDMX	API SDMX
Euro SDMX Registry	pandaSDMX	pandaSDMX	Bulk Data Reporting	Bulk Data Reporting
FMR	rsdmx	Python, Pandas	FMR	pandaSDMX
pandaSDMX	SDMX Connectors for Statistical Software	SDMX Connectors for Statistical Software	pandaSDMX	Python, Pandas
SDMX Converter	SDMX Converter	SDMXSource	SDMX Converter	SDMX Converter
SDMX Global Registry			SDMX Global Registry	SDMX Global Registry
SDMX-EDI			SDMXSource	SDMX-EDI and SDMXRI
SDMXSource				SDMXSource

Sources: IFC survey of SDMX tools (2023); authors' elaboration.

There appears to be a **particular demand for action in the area of data modelling**, a fundamental stage in the data lifecycle that should naturally precede and facilitate any other stage (OECD and PARIS21 (2021)). The implementation and use of global SDMX DSDs should therefore be an important priority. It can ease the modelling phase and, through knowledge-sharing, allow institutions with limited expertise to implement SDMX. Important information is already available to facilitate this task via the SDMX Global Registry,[14] a powerful repository that serves as a central reference point and source for structural metadata such as global DSDs, cross-domain concepts and codelists (see Box C). This registry supports statistics producers by providing standardised DSDs, in turn helping to streamline data exchange and improve comparability across different statistical domains. It also aids users by offering easy access to metadata and reference materials.

Lastly, **building and sharing common knowledge is essential**. As can be seen in Graph 5, there is already a rich ecosystem of open source IT tools available to support SDMX adoption and implementation throughout all the steps of the data lifecycle. These open source tools, often complemented with in-house or commercial tools, provide all the architectural components necessary to support statistical processes, covering the main GSBPM phases.[15] These resources should be complemented by learning facilities, a need that may have been underestimated compared with the high demand for tools expressed by the survey respondents. From this perspective, developing a more extensive knowledge exchange would be instrumental to addressing barriers to SDMX implementation.[16]

[14] See registry.sdmx.org/.

[15] A list of available tools for SDMX implementers and developers is also published on sdmx.org.

[16] As a key example, the SDMX user forum was launched in 2023 to connect SDMX users across the globe. In addition, the SDMX community and the BIS have established centralised learning resource pages on sdmx.org and sdmx.io that comprise training materials from various providers and numerous existing tools.

Box B

Enhancing the interoperability of SDMX tools for improved organisational efficiency

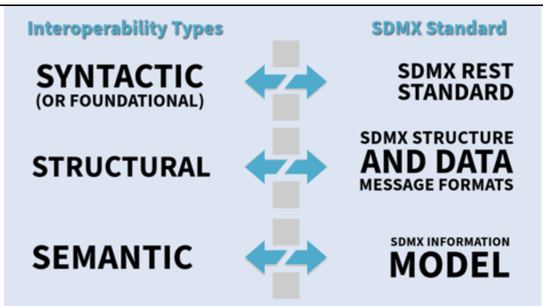
Fadhila Najeh (FAO) and Jonathan Challener (OECD)

In the world of official statistics, various tools have been developed over the past decade to collect, process and disseminate data. However, these tools often lack the ability to communicate with one another, creating challenges to the effective management and exchange of data. **Achieving interoperability is thus crucial for organisations to become more data-driven and enhance their data management strategies**. To address this issue, the SDMX standard provides a solution that enables different tools to seamlessly work together and exchange data.

There are **three main types of interoperability** that SDMX focuses on: syntactic, structural and semantic (Graph B1). Syntactic interoperability enables different systems to communicate and share data. Structural interoperability defines the data exchange format and standards used to format messages. Semantic interoperability allows systems to connect and exchange data in a way that is mutually understandable.

SDMX interoperability types

Graph B1



Source: Authors' elaboration.

The SIS-CC① conducted an assessment in 2022 to better understand the current state of SDMX tools and their ability to work together. The related report② highlighted the need for clear goals to achieve effective interoperability and integration and recommended establishing interoperability criteria as official SDMX guidelines.

The assessment identified the **following key interoperability requirements** that can play a critical role in achieving seamless integration and improving organisational workflows:

1. **Effective handling of versioning and upgrades** is essential for ensuring the smooth operation and compatibility of tools and systems.
2. **Robust access control mechanisms** must be implemented to safeguard data and maintain privacy and security. This includes controlling user access rights and permissions to ensure appropriate data governance.
3. **Comprehensive documentation** is crucial for facilitating the understanding, implementation and maintenance of tools and systems. This documentation should include clear instructions, guidelines and explanations for users.
4. Establishing **effective governance frameworks** is necessary to ensure proper management and oversight of tools. This calls for defining roles and responsibilities, establishing policies and procedures and implementing quality assurance processes.

The report also emphasised the **importance of tool providers adhering to good practices** in addressing these interoperability requirements.

In conclusion, **improving the interoperability of SDMX tools is essential for enhancing organisational efficiency in managing and exchanging official statistics**, and the SDMX standard provides a solution to enable different tools to work together seamlessly. While significant steps have already been taken to address existing interoperability challenges, adherence to recommended good practices and collaboration with tool providers can contribute to further progress, achieving greater interoperability and efficiency in tool implementation, maximising the benefits of SDMX and advancing organisational data management practices.

① Statistical Information System Collaboration Community; see siscc.org. ② See SIS-CC (2024a).

Box C

SDMX metadata-driven approach: the BIS experience

Glenn Tice, Matt Nelson, Stratos Nikoloutsos and Xavier Sosnovsky (BIS)

A key characteristic of the SDMX information model is that it provides a formal metadata specification and is therefore designed specifically for metadata-driven data processing. Such a **metadata-driven approach** can help institutions to systematise their statistical processes, in turn delivering better governance, agility and efficiency.

The term "metadata" may mean different things in the realm of statistics and data engineering. It can be defined as the "information that is needed to be able to use and interpret statistics"① or, more simply, as "data about the data". Metadata are often understood as *reference metadata* (sometimes called "explanatory metadata"), ie the description of the contents and quality of the statistical data from a semantic point of view. A second concept is *structural metadata*, ie the information used to identify, formally describe or retrieve statistical data.

In the SDMX context, metadata will typically refer primarily to ***structural*** **and** ***processing*** **metadata** describing, respectively, the structure of data sets (in terms of their concepts, dimensions, measures and attributes) and the rules to be applied for data transformation and quality control.

Formalising and externalising metadata allows for statistical processes to become more standardised and flexible.② This is because one can deal with changes in metadata (such as data structures, codelists or constraints) without directly affecting the underlying processing logic. This allows for smoother adaptation to evolving requirements and strengthens governance by providing direct control over data management. In addition, consistency between data structures and processing logic, an important element in environments characterised by large and complex metadata and/or rapidly changing processing rules, can be more easily enforced.

While the SDMX information model is designed specifically for metadata-driven processing, its practical implementation requires a dedicated repository as the controlled central source of metadata truth. For simpler use cases, that metadata repository could consist of SDMX structure files on a shared drive. For more complex, enterprise-wide cases, it is more effective to set up an **SDMX metadata registry** to enforce consistency and integrity, especially when metadata managers make changes in parallel. For this purpose, institutions have the possibility to use the **Fusion Metadata Registry (FMR)**.③ This is a mature, free and open source SDMX metadata registry made available on BIS Open Tech, a platform for sharing statistical and financial software as public goods. Lastly, developers, analysts and data scientists can use the BIS's open source **pysdmx Python library,** which provides an efficient way to retrieve and employ SDMX metadata to drive data processing and analytical jobs.④ Together, the widespread use of the Python language, the rich SDMX information model and FMR's centralised metadata management capabilities form a solid foundation on which metadata-driven statistical processing solutions can be built.

As regards the BIS's own statistical work, an important initiative has been to develop **a common monetary and economic data lake** (called MEDAL) that is entirely metadata-driven and leverages SDMX 3.0 and FMR. Its **Data Portal** (data.bis.org), the BIS's main statistical publication platform providing users with an intuitive interface for discovering, evaluating and analysing data, also follows the SDMX metadata-driven approach.⑤ In addition to processing and structural metadata, the platform also leverages presentation metadata. For instance, many data sets can be visualised as dynamic "publication tables" – like traditional pivot tables, but more flexible in their content and composition. The descriptions for these tables are managed as SDMX metadata artefacts, providing agility, control and efficiency, not least by avoiding the need for statisticians to handcraft and maintain custom data tables.

① See Eurostat's glossary for statistics. ② See Froeschl (1999). ③ FMR powers the SDMX Global Registry (registry.sdmx.org) and the IMF's SDMX Central (sdmxcentral.imf.org) and serves as a metadata registry, data modelling platform and data processing engine supporting the statistical processing systems of various institutions worldwide. ④ FMR and pysdmx are available at sdmx.io. ⑤ See Lambe and Park (2024).

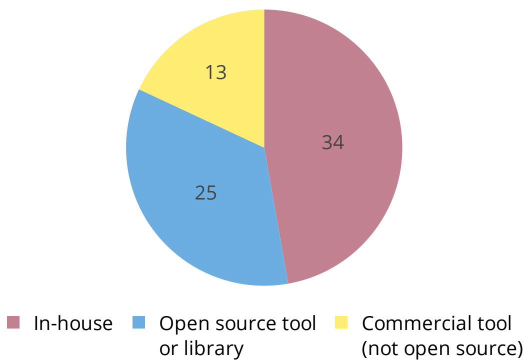
4. SDMX tools: state of adoption and remaining gaps

The survey provided **insights into the prevalence and utilisation of SDMX tools within central banks**, with a particular emphasis on the adoption and implementation of open source tools.

Categories of SDMX tools used by central banks | Graph 6

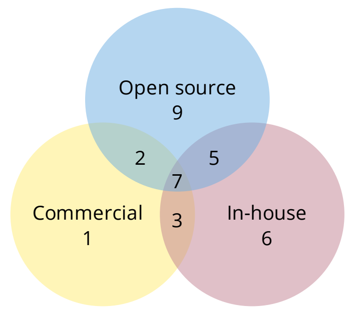
A. Unique tools in use

Number of tools



B. Categories of tools

Number of respondents



Sources: IFC survey of SDMX tools (2023); authors' calculations.

Central banks use a very large number of different SDMX IT tools (72 were mentioned by the reporting institutions). These resources are either developed in-house, commercial or open source (Graph 6.A) and can be used in combination (Graph 6.B). This variety illustrates one of the strengths of SDMX, ie its modular and loosely coupled design that encourages the use of fit-for-purpose tools which can be combined to respond to specific needs.

Open source tools appear to be the most widely used SDMX tools, accounting for 51% of installations, followed by in-house tools (36%) and commercial tools (13%). The usage of these tools varies depending on their type, with in-house tools typically limited to a single organisation, whereas open source and commercial tools are used by multiple organisations. While a large proportion of open source SDMX tools are used by only one or two central banks (Graph 7.A), a significant number (25) are in active use by at least one central bank (Graph 6.A).

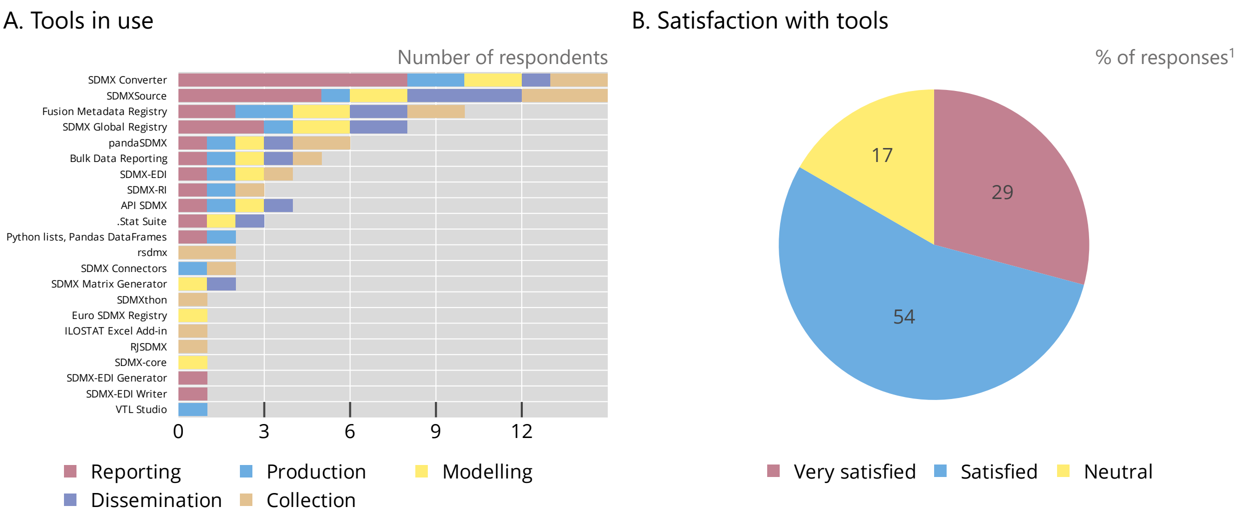
A notable highlight of the survey is that **the vast majority (83%) of central banks reported being *very satisfied* or *satisfied* with the wide range of open source SDMX tools available** (Graph 7.B).[17] Moreover, these resources appear to be increasingly utilised in all stages of the data lifecycle. One key reason for this success is that as statistical producers gain hands-on experience with open source tools, their usage can expand from simple applications to mission-critical and customer-facing applications (Buffett (2014), IFC (2023c)). In addition, open source software also offers significant benefits to users, including better customisation, stronger resilience and closer adequation between the IT features developed and business needs (Araujo et al (2023)).

Another key message is that, despite increased usage, there remain significant barriers to the adoption of SDMX tools. Of particular importance are the *cost of adapting internal software tools* and the *lack of support for SDMX in (traditional) open source tools*. These barriers, combined with the previous findings that many existing SDMX tools are used by only a few institutions, could indicate **an inefficient**

[17] Satisfaction with in-house or commercial tools was not assessed.

and fragmented software offering. One possible solution is therefore to consolidate the software supply, for example under common platforms. Key examples are global initiatives such as sdmx.io, which offers an ecosystem of IT tools for SDMX, and the repository Awesome list of official statistics software.[18] For its part, the open source software producer community has taken various steps to better understand and address user needs by coordinating efforts to foster SDMX tool adoption (see Box D). A number of in-house implementations may also be beneficial to the entire SDMX community and would therefore deserve to be shared as open source tools to fill existing gaps and reduce the duplication of efforts across institutions.

Open source SDMX tools – usage and satisfaction Graph 7



[1] Respondents could indicate different satisfaction levels with different tools.

Sources: IFC survey of SDMX tools (2023); authors' calculations.

Cyber security is another reported key barrier to the adoption of SDMX tools, especially those that are open source. Addressing this calls for ensuring compliance with industry standards in order to build trust and ease adoption processes, as open source tools, thanks to their transparency, can be more secure and receive quicker patches.

While reducing the above-mentioned barriers is a key step towards increasing the adoption of SDMX tools and in turn promoting efficiency and innovation, another key ingredient is having the related, necessary knowledge and skills within an organisation. This will depend on the **availability of related learning resources and support services** from which they can benefit.

To shed light on this issue and identify improvement opportunities, **the survey specifically measured satisfaction with such learning resources and support services**, as well as overall satisfaction with open source tools. As noted above, central bank users appear to be largely satisfied with their usage of open source SDMX tools, and a key reason for this is the value of the learning resources and support they receive. This applies to both pre-implementation and post-implementation support materials (Graph 8), with essentially no *dissatisfied* responses reported. Lastly, while the survey also showed that

[18] The repository can be accessed at github.com. For additional details on the repository, see Verbruggen (2023).

central banks' learning needs vary greatly across the various tools, the demand for more *learning videos* and *recorded workshops* emerged as a common request.[19]

Box D

The BIS Open Tech initiative and the sdmx.io ecosystem

Antonio Olleros (Meaningful Data), Brian Buffett and Stratos Nikoloutsos (BIS)

One notable part of the BIS's medium-term strategy, Innovation BIS 2025,① which has relied on important investment in next-generation technology, has been to provide **software as a public good**. This is testimony to the fact that, in an increasingly interconnected global economy, open source software is becoming the norm across numerous industries.

In this context, a key step was the launch in 2022 of **BIS Open Tech**, a platform which fosters international partnerships and knowledge exchange and provides a hub for the sharing of cutting-edge statistical and financial tools. This platform hosts software developed according to international best practices and standards. This material can be reused and further developed in a wide variety of environments. Beyond the BIS Open Tech pages presenting these tools, the BIS also offers source code repositories for sharing software transparently with stakeholders and improving the collaboration between contributing counterparts. Hence, this global initiative is about more than developing innovative software solutions, since its aim is to enhance financial and statistical capabilities through collaboration and shared knowledge.

At the heart of the BIS Open Tech initiative are the **open source tools for SDMX provided by the BIS** in close collaboration with various international and national organisations as well as the private sector. As one of the main founding members of SDMX, the BIS has also developed sdmx.io, a use case-oriented platform that provides open source SDMX solutions for working with data and metadata. Examples include:

- the **FMR open source SDMX metadata registry** and its associated workbench for metadata management and sharing;
- the **.Stat Suite** platform for data exchange, processing and dissemination;
- **Python libraries** (in LinkageX, ie pysdmx, gingado),② which are powerful data science tools for harnessing SDMX data and metadata information;
- the **SDMX Dashboard Generator**, which provides a platform for dynamic, metadata-driven data visualisation using SDMX sources; and
- the **SDMX TCK (Test Compatibility Kit)**, which supports the development of application programming interfaces (APIs) by assessing their compatibility with the SDMX API standard.

In addition to offering software, the sdmx.io ecosystem is based on a governance model designed to ensure software interoperability. It also provides a comprehensive array of learning resources (such as educational materials and practical guides) designed to assist users. Hence, sdmx.io is not so much about software development as it is about cultivating an inclusive community that thrives on shared knowledge and problem-solving.

Therefore, sdmx.io is well equipped to respond to the needs of a wide range of stakeholders, including:

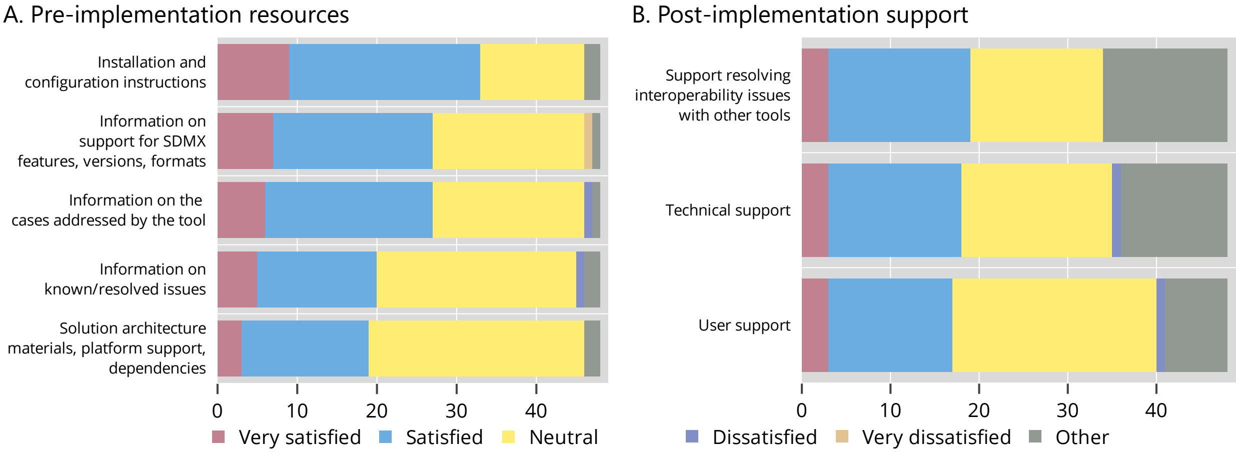
- **statistics organisations**, by providing a set of high-quality tools they can use internally for a robust implementation of SDMX-based statistical processes;
- **users of statistics**, by offering tools that abstract SDMX complexities and allow them to seamlessly use the data disseminated (while such tools may have been already available before the launch of sdmx.io, they were often underused because of the difficulty of identifying the adequate IT solution); and
- **software vendors**, by encouraging the creation of high-quality open source tools as public goods, financed by interested stakeholders (eg international organisations).

① See BIS (2020). ② See Araujo (2023).

[19] See www.sdmx.io/resources/elearning/ and sdmx.org/learning/.

Degree of satisfaction with the resources and support for open source SDMX tools

Number of respondents | Graph 8



Sources: IFC survey of SDMX tools (2023); authors' calculations.

5. SDMX by statistical domain

A key focus of the survey was to examine SDMX adoption by type of statistics. A first insight is that **most of the central banks' statistical production (using SDMX or not) covers five key domains**, namely *monetary and financial statistics*; *financial market and securities statistics*; *balance of payments and international investment positions*; *payment system statistics*; and *supervisory* and *prudential statistics*. Additionally, central banks also produce statistics less actively in six domains: *government finance statistics*; *financial accounts*; *international trade*; *GDP and non-financial accounts*; *prices* and *labour market statistics* (Graph 9.B).

A second insight is that, when a central bank actually produces statistics for a specific area, the SDMX standard is actively used, albeit to varying degrees across the 11 statistical domains (Graph 9.A).

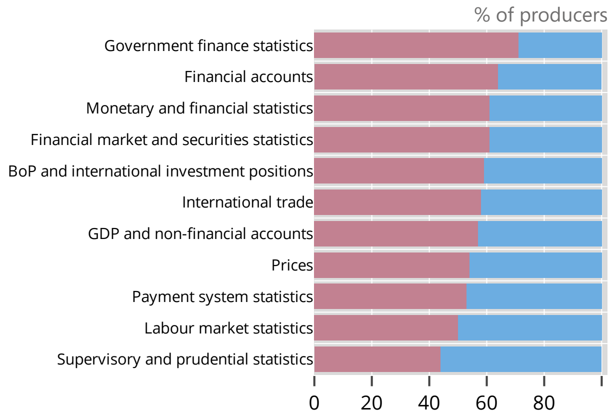
Third, and consistent with the situation observed a decade ago, **SDMX is used to an important extent in the macroeconomic statistics predominantly produced by central banks, benefiting from the existence of global DSDs**. These areas notably include *monetary and financial statistics, financial market and securities statistics* and *balance of payments and international investment positions* – around 60% of the respondents produce statistics in these three key domains using SDMX. Interestingly, SDMX is highly used by the relatively lower number of central banks producing data in the two domains of *government finance statistics* and *financial accounts* – reflecting the fact that these two areas have well established global DSDs facilitating SDMX adoption.

Lastly, **SDMX is relatively less used (by respectively 53% and 44%) for two key statistical domains produced by central banks**, namely *payment system statistics* and *supervisory and prudential statistics*.

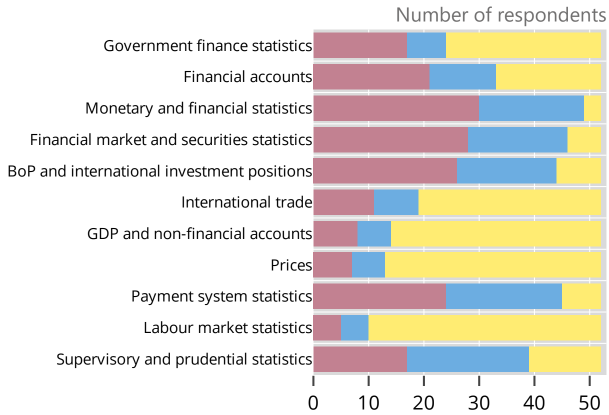
SDMX usage by statistical domains

Graph 9

A. SDMX usage when a statistical domain is produced by the central bank



B. Degree of statistical domain production



Using SDMX   Not using SDMX   Not produced

Sources: IFC survey of SDMX tools (2023); authors' calculations.

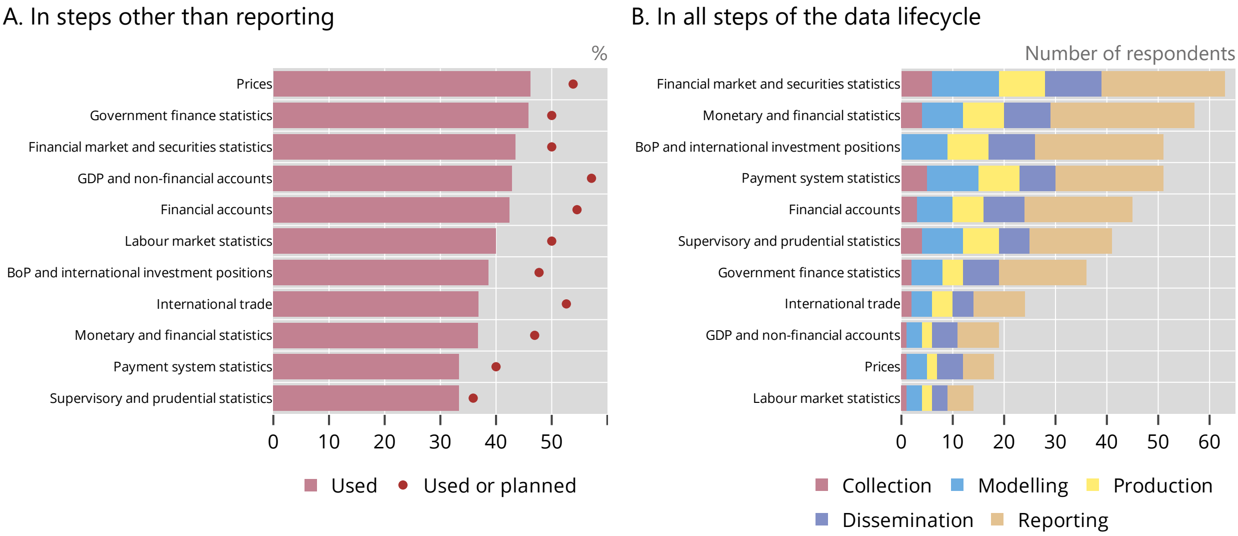
Apart from the 11 main domains mentioned above, **one area of increasing relevance to central banks is environmental data** (IFC (2024c)). A number of key initiatives have been developed by the SDMX community to better deal with them, especially to produce global DSDs for data reporting that are aligned with the System of Environmental-Economic Accounting.[20]

The survey also shows that the degree of SDMX adoption observed across all statistical domains differs across the various specific stages of the data lifecycle. In particular, it confirms the common trend already analysed above, ie that **SDMX is predominantly used for data reporting**, a feature that characterises all the various statistical domains (Graph 10.B).

Nevertheless, reported plans for the coming years show that **central banks intend to extend the implementation of SDMX to other phases of the data lifecycle**, especially data modelling, production and dissemination (Graph 10.A). This is particularly the case for statistical domains with a lower degree of adoption, such as *international trade*, *GDP*, *prices* and *labour market statistics* – areas that have already established global DSDs that can greatly facilitate *domain-specific* adoption.

[20] Specifically, some international organisations, such as the OECD and UN, started to collect data on air emissions and energy accounts; see ECOSOC (2025).

SDMX adoption plans in statistical domains produced by central banks[1] Graph 10



[1] The figures in per cent refer to the respondents who are producing data in the given domain.

Sources: IFC survey of SDMX tools (2023); authors' calculations.

From this perspective, **making more global DSDs available** appears to be key to promoting the use of the standard by central banks across data domains. This calls for filling existing gaps and in particular elaborating DSDs for *payment system statistics* and *supervisory and prudential statistics*, ie domains where most central banks produce data but for which SDMX adoption is less widespread. Similarly, important efforts have been put into developing a global DSD for property prices, an area of increasing interest to central banks since the GFC (see Box E).

A second action point is to **develop tools and learning resources tailored to specific statistical domains**.[21] The demand appears to be particularly high for those areas where SDMX is most widely adopted, such as *balance of payments statistics*, *monetary and financial statistics,* and *financial market and securities statistics* (Graph 11).

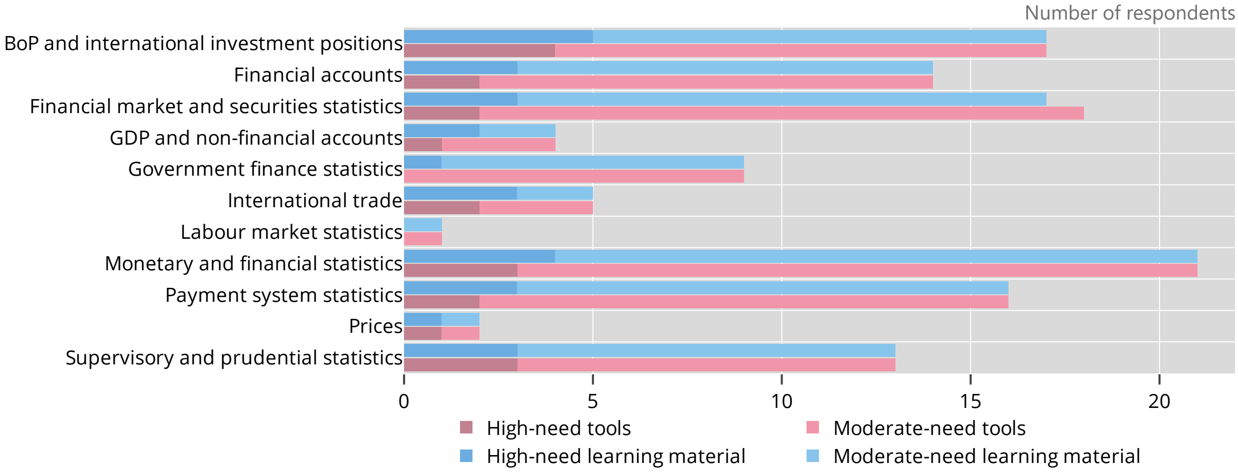
Third, SDMX usage should be **expanded to cover new data types** other than macro time series or aggregate statistics. The avalanche of big data, with the proliferation of alternative sources and new user needs (such as the demand for climate-related indicators), clearly calls for further implementation of SDMX in the realm of micro and geospatial data (IFC (2021)). The features introduced in version 3.0 of the standard move in this direction, not least by adding the required technical specificities and balancing the required flexibility with the need for standardisation. Yet, additional statistical guidance, capacity building and use case-sharing at the international level are needed to help central banks transition towards the new SDMX version and harness its benefits.[22]

[21] This question was asked only to those respondents that reported producing statistics in the statistical domain of interest.

[22] For example, the Cross-Working Group on Microdata – established in 2024 – is expected to deliver a content-oriented guideline on how to use SDMX for granular data, such as financial big data and granular information; see ECOSOC (2025).

Need for domain-specific tools and learning resources

Graph 11



Sources: IFC survey of SDMX tools (2023); authors' calculations.

Box E

Global data structure definitions – an application to property prices

Robert Szemere and Bianca Ligani (BIS)

In the SDMX information model, a data structure definition (DSD) outlines how data are structured and described, allowing compilers and users to share this information in a timely and efficient manner. A DSD brings together three types of structural descriptor: (i) **dimensions**, which identify and describe the data; (ii) **attributes**, which provide additional information, such as whether the data are estimates; and (iii) **measures**, which represent the phenomenon to be measured. As a result, a DSD identifies the data through the dimensions values that form their unique **series key**.

While statistical agencies can define "local" DSDs for managing their own data, **a number of "global" DSDs have been designed for major statistical domains**, such as national accounts, balance of payments and consumer prices. The primary advantage such global DSDs offer to compilers is that having standardised templates reduces the burden of submitting the same data to different institutions. It also ensures proper governance of statistical concepts, particularly for cross-domain ones such as in national accounts and external statistics.① Turning to data users, a global DSD provides a common language for exploring and retrieving data. As the data modelling is done by the SDMX community, institutions with less technical knowledge or reduced capacity can leverage a globally shared data model and avoid custom implementations. Ultimately, global DSDs support the creation of comparable statistics across the globe by providing harmonised structures to national statistical offices, central banks and international organisations. The free dissemination of harmonised data also enhances their accessibility, interpretability and sharing.

However, the creation of a global DSD requires a joint effort from those institutions supporting the SDMX standard. Implementation also poses **governance** challenges, as a maintenance protocol and periodic reviews need to be established.② To this end, the international Macro-Economic Statistics Ownership Group (MES OG) coordinates the design, implementation and maintenance of global DSDs. As major changes would require joint approval of all parties involved, attention is put on designing comprehensive, "future-proof" DSDs that can accommodate expansions in the data universe without requiring revisions, helping to reduce maintenance costs.

One interesting and recent illustration has been the joint development by the BIS, Eurostat, the IMF and the OECD of a global DSD for **property price indicators**. One key driver of this effort is the fact that public statistical authorities in many countries have started their compilation relatively recently, in many cases after the GFC and in response to the associated financial stability analytical needs.③ In particular, important compilation work has been conducted in the context of the second phase of the G20 DGI actively supported by the BIS.④ Yet, currently available data are still heterogeneous in terms of sources, compilation methods and coverage. The release of the new DSD will therefore help to further enhance the compilation of property price indicators, especially by allowing for a comprehensive classification of sub-national geographical areas and property types. In this respect, a number of SDMX 3.0 features (such as hierarchical codelists) are particularly useful.

More generally, the **integration of global DSDs not only streamlines reporting processes but also fosters a collaborative environment**. For instance, the ongoing standardisation work on property prices will ultimately reduce the reporting burden which is often borne by central banks. Additionally, users will benefit from standardised property price data, with easier comparison and analysis across multiple sources.

① See Diz Dias et al (2024). ② See SWG (2023). ③ See Tissot (2014). ④ See FSB and IMF (2018).

References

Araujo, D (2023): "gingado: a machine learning library focused on economics and finance", *BIS Working Papers*, no 1122, September.

Araujo, D, S Nikoloutsos, R Schmidt and O Sirello (2023): "Central banks as users and providers of open-source software", *IFC Bulletin*, no 59, Box 1, pp 8–10, October.

Asian Development Bank (ADB) (2024): *Enhancing data management through the Statistical Data and Metadata eXchange standard: A special supplement of the key indicators for Asia and the Pacific 2024*, August.

Bank for International Settlements (BIS) (2020): *Annual Report 2019/20*, June.

Buffett, B (2014): "Factors influencing open source software adoption in public sector national and international statistical organisations", *EUROSTAT Working Paper*, February.

Committee on Payments and Market Infrastructures (CPMI) (2025): "BIS CPMI takes further steps to promote ISO 20022 harmonisation for enhanced cross-border payments", January.

European Banking Authority (EBA), European Central Bank (ECB) and European Insurance and Occupational Pensions Authority (EIOPA) (2024): *Memorandum of Understanding on the establishment of a common Data Point Model (DPM) governance framework – "DPM Alliance"*, March.

Fajt Mayer, T, G Raimondo Quaratino and J Cordero Romero (2024): "Variational autoencoders for multivariate time series outlier detection", presentation at the 12th IFC Biennial Conference on "Statistics and beyond: new data for decision making in central banks", 22–23 August.

Financial Stability Board (FSB) and International Monetary Fund (IMF) (2018): *Second phase of the G20 Data Gaps Initiative (DGI-2): third progress report*, September.

Froeschl, K (1999): "Metadata management in official statistics – an IT-based methodology approach", *Austrian Journal of Statistics*, vol 28, no 2, pp 49–79.

Gregory, A (2024): "Open and FAIR: trends in scientific publishing and the implications for official statistics", *Statistical Journal of the IAOS*, vol 40, no 1, pp 91–6, February.

Gregory, A, D Bell, D Brickley, P Buttigieg, S Cox, M Edwards, F Doug, L Gonzalez Morales, P Heus, S Hodson, C Kanjala, Y Le Franc, L Maxwell, L Molloy, S Richard, F Rizzolo, P Winstanley, L Wyborn and A Burton (2024): *WorldFAIR (D2.3) Cross-Domain Interoperability Framework (CDIF) (Report synthesising recommendations for disciplines and cross-disciplinary research areas)*, May.

IMF (2013): *The special data dissemination standard: guide for subscribers and users*.

IMF, Inter-Agency Group on Economic and Financial Statistics and FSB Secretariat (2023): *G20 Data Gaps Initiative 3 workplan – People, planet, economy*, March.

International Organization for Standardization (ISO) (2023): *ISO 17369:2013 – Statistical Data and Metadata eXchange (SDMX)*.

Irving Fisher Committee on Central Bank Statistics (IFC) (2016): "Central banks' use of the SDMX standard", *IFC Report,* no 4, March.

——— (2021): "Use of big data sources and applications at central banks", *IFC Report*, no 13, February.

——— (2023a): "Data-sharing practices", *IFC Guidance Note*, no 3, March.

——— (2023b): "Data science in central banking: applications and tools", *IFC Bulletin*, no 59, October.

——— (2023c): "Central banks' use of time series products", *IFC Guidance Note*, no 4, December.

——— (2024a): "Communication on central bank statistics", *IFC Bulletin*, no 60, April.

———(2024b): "External statistics in a fragmented and uncertain world: addressing novel analytical challenges", *IFC Bulletin*, no 62, August.

——— (2024c): "Empowering carbon accounting: from data to action", *IFC Report*, no 16, October.

Kroese, B (2024): "Exploring official statistics using generative AI and SDMX", presentation at the 72nd plenary session of the Conference of European Statisticians, Geneva, 20–21 June.

Laaboudi, C, M Karlberg and M Islam (2024): "Open data dissemination at Eurostat: State of the art", *Statistical Journal of the IAOS*, vol 40, no 1, pp 81–90, February.

Lambe, E and T Park (2024): "The BIS Data Portal project – delivering the next generation platform for BIS statistics", *IFC Bulletin*, no 60, April.

Muñoz, J (2020): "SDMX and GSBPM", presentation at the UNECE ModernStats World Workshop, virtual, 27–30 October.

Organisation for Economic Co-operation and Development (OECD) and PARIS21 (2021): *Data flow analysis framework: guidelines for analysing data flows in national statistical offices*, January.

SDMX Statistical Working Group (SWG) (2016): *SDMX content-oriented guidelines – Introduction*, February.

——— (2020): "SDMX Glossary – version 2.1", *SDMX Guidelines*, December.

——— (2023): "A reference framework for SDMX structural metadata governance – version 1.0", *Statistical Guidelines*, March.

SDMX Technical Working Group (TWG) (2021): "Summary of major changes and new functionality – version 3.0", *SDMX Standards*, October.

sdmx.io (2023): "SDMX Structural Modelling – Transform and map statistical logical models to the SDMX standard", sdmx.io/resources/elearning/essential-sdmx-structural-modelling/.

Sirello, O (2024): "Editing metadata with generative AI", background paper prepared for the UNECE Expert Meeting on Statistical Data Editing, Vienna, 7 October.

Statistical Information System Collaboration Community (SIS-CC) (2024a): "Enhancing SDMX tools interoperability for improved organisational efficiency", March.

——— (2024b): "SDMX+AI: unlocking the potential of NLP to enhance data access", March.

Tissot, B (2014): "Monitoring house prices from a financial stability perspective – the BIS experience", International Statistical Institute Regional Statistics Conference, Kuala Lumpur, 16–19 November.

——— (2017): "SDMX, a key standard for central banks' statistics", proceedings of the 61st ISI World Statistics Congress, Marrakech, 17 July.

——— (2021): "SDMX, a key element supporting official statistics", presentation at the 10th Expert Group Meeting on Statistical Data and Metadata eXchange, Washington DC, 25 January.

United Nations (UN) (2014): *Resolution adopted by the General Assembly on 29 January 2014: 68/261 – Fundamental Principles of Official Statistics*, January.

United Nations Economic and Social Council (ECOSOC) (2025): *Report of the Statistical Data and Metadata eXchange (SDMX) sponsors*, 56th session of the United Nations Statistical Commission, New York City, 4–7 March (forthcoming).

United Nations Economic Commission for Europe (UNECE) (2020): *Guide to sharing economic data in official statistics*.

——— (2023): "Large language models for official statistics", *HLG-MOS White Paper,* December.

——— (2024a): *Data governance framework for statistical interoperability*, High-Level Group for the Modernisation of Official Statistics, March.

——— (2024b): *Statistical implementation standards in the context of GSBPM*, High-Level Group for the Modernisation of Official Statistics, October.

——— (2024c): "A path towards the use of common standards to support the INtegration of GEospatial and STatistical information", *Working Paper Series on Statistics*, no 11, August.

——— (2025): *Data for climate action: how national statistical offices can contribute*, January.

US Department of Commerce (DOC) (2025): *Generative artificial intelligence and open data: guidelines and best practices*, AI and Open Government Data Assets Working Group, January.

Verbruggen, M (2023): "Moving towards open-source technologies: strategic and managerial perspectives", presentation at the 71st plenary session of the Conference of European Statisticians, Geneva, 22 June.

Wilkinson, M et al (2016): "The FAIR Guiding Principles for scientific data management and stewardship", *Scientific Data*, vol 3, no 1, 160018, December.

Annex 1: IFC survey questions

1. **Please describe the statistical infrastructure in your central bank.**
    1.1. Are there any factors limiting the use of open source tools in your organisation?
    1.2. Does your organisation use SDMX and/or XBRL?
        1.2.1. Does your organisation see the need for an SDMX-to-XBRL or XBRL-to-SDMX converter?
    1.3. Are there any factors preventing your organisation from using SDMX?
2. **Assessment of SDMX versions**
    2.1. Which versions of SDMX are in use at your organisation?
    2.2. Does your organisation plan to migrate to SDMX 3.0?
    2.3. Are there any factors preventing your organisation from transitioning to SDMX 3.0?
    2.4. Rate the relative level of sophistication in your organisation's use of SDMX.
    2.5. Does your organisation plan to use SDMX for micro data modelling? Why or why not?
    2.6. Does your organisation use or plan to use Validation and Transformation Language (VTL)?
    2.7. What are your organisation's primary use cases for Validation and Transformation Language (VTL)?
    2.8. Indicate the tools used in your organisation to support the implementation of Validation and Transformation Language (VTL).
    2.9. Provide a brief explanation of your organisation's essential and high-priority use cases, if any, which are not being met by existing open source VTL tools.
3. **For major statistical processes and sub-processes, identify the relevance of SDMX for your organisation and the level of SDMX knowledge in your organisation.**
    3.1. Rate the level of SDMX knowledge in your organisation.
    3.2. Rate the level of SDMX experience in your organisation in the following process areas.
    3.3. Identify the tools and/or libraries used in your organisation to support the implementation of SDMX.
    3.4. Identify the tools used in your organisation to support the implementation of SDMX in the following process areas.
    3.5. Indicate your organisation's implementation of SDMX for each area of financial statistics and by major process.
4. **Assessment of SDMX tools**
    4.1. Rate your organisation's level of satisfaction with open source SDMX tools.
    4.2. Indicate your organisation's implementation of open source SDMX tools or libraries by major process.

5. **Assessment of documentation and support materials for open source SDMX tools**
    5.1. Rate your organisation's level of satisfaction with documentation/material to support open source tool selection and implementation.
    5.2. Rate your organisation's level of satisfaction with post-implementation support available for open source SDMX tools.
    5.3. Rate your organisation's priorities for tool-specific learning resources.
6. **Assessment of SDMX capacity building priorities**
    6.1. Rate your organisation's priorities for process-specific tools and learning resources.
    6.2. Indicate your organisation's need for domain-specific tools and learning resources.
    6.3. Is there a need for SDMX courses to provide a certificate of completion to employees in your organisation?
7. **Overall**
    7.1. Please rate your overall satisfaction with open source SDMX tools.
    7.2. Additional comments/suggestions regarding open source SDMX tools?

Annex 2: List of IFC jurisdictions that responded to the survey

1. Argentina
2. Austria
3. Bank for International Settlements
4. Belgium
5. Canada
6. Chile
7. Costa Rica
8. Croatia
9. Denmark
10. Ecuador
11. European Union
12. Finland
13. France
14. Germany
15. India
16. Indonesia
17. Ireland
18. Israel
19. Italy
20. Japan
21. Kazakhstan
22. Korea
23. Latvia
24. Malaysia
25. Mauritius
26. Mexico
27. Moldova
28. Montenegro
29. Morocco
30. Mozambique
31. Namibia
32. Netherlands
33. North Macedonia
34. Norway
35. Peru
36. Philippines
37. Portugal
38. Romania
39. Saudi Arabia
40. Singapore
41. Slovakia
42. Slovenia
43. South Africa
44. Spain
45. Suriname
46. Sweden
47. Switzerland
48. Thailand
49. Tunisia
50. Ukraine
51. United States
52. Vietnam