United Nations ECE/CES/GE.20/2016/18

Economic and Social Council

Distr.: General
19 April 2016

English only

Economic Commission for Europe

Conference of European Statisticians

Group of Experts on National Accounts

Fifteenth session
Geneva, 17-20 May 2016
Item 4 of the provisional agenda
Measuring Global Production

Data sharing for macro-economic statistics

International activities and European perspective

Prepared by Eurostat[1]

Summary

For many years, international organisations have been collecting - and exchanging among each other - macro-economic data from countries to respond to user needs as regards the availability of data for economic analysis and decision making. Recently, several international organisations have taken a further step in making selected macro-economic statistics more readily available thanks to close collaboration through the Inter-Agency Group on Economic and Financial Statistics (IAG) and new technical possibilities. From now on, GDP and some related indicators will be identical across the respective databases of several international organisations.The paper will show how the international data cooperation is set up, what benefits it has and what future plans can be envisaged for working better together in the world of official statistics. It will show some differences between data reporting, dissemination and sharing and present activities undertaken by Eurostat and some EU Member States in those areas.The International Data Cooperation initiative can in principle also be extended to other organisations flexibly, which provides a solid basis for official macro-economic statistics of high quality in the public domain.

[1] Prepared by Daniel Suranyi, Eurostat. National accounts, prices and key Indicators, Daniel.Suranyi@ec.europa.eu

I. Technical framework and governance

A. Why SDMX?

1. The adoption of SNA 2008 / ESA 2010 and the work on the associated transmission programmes created a new momentum for the further alignment of international standards for the compilation and dissemination of macro-economic statistics. In the light of scarce resources, there is a growing willingness for closer international cooperation and an increasing emphasis on statistical business process integration. SDMX is the leading standard for exchanging and sharing data and metadata in official statistics. SDMX is sponsored by seven International Organisations and is recognised by many international bodies, such as the United Nations' Statistical Commission, the Interagency Group on Economic and Financial Statistics coordinating and monitoring the G20 Data Gap Initiative, the UN regional Economic Commissions, the European System of Central Banks (ESCB) and the European Statistical System (ESS) (1). It is thus an enabler for rationalisation of data flows, harmonisation of reporting needs and standardisation of information systems (2).

B. Governance around SDMX in Macro-Economic Statistics

2. The SDMX Sponsors group manages the SDMX standards themselves with the main objective to deliver a stronger and more global information system that can provide open and real-time access to official statistics. To achieve this, national and international statistical authorities will make available timely and comparable statistics using globally agreed data structures supported by good-quality metadata.

3. To oversee the maintenance of global data structure definitions for statistical domains related to macro-economic statistics, the SDMX Sponsors group mandated the Ownership Group for SDMX in Macro-Economic Statistics.

4. The following organisations are member of the Ownership Group: (3)

- Bank for International Settlements (BIS)
- European Central Bank (ECB)
- Eurostat (ESTAT)
- International Monetary Fund (IMF)
- Organisation for Economic Co-operation and Development (OECD)
- United Nations Statistics Division (UNSD)
- World Bank (WB)

5. The work of the Ownership Group started with the initial packages for SDMX in National Accounts and SDMX in Balance of Payments, published by the two respective Steering Groups in September 2014. The packages define statistical data models for the domains under its coverage, with a maintenance agency selected for each domain:

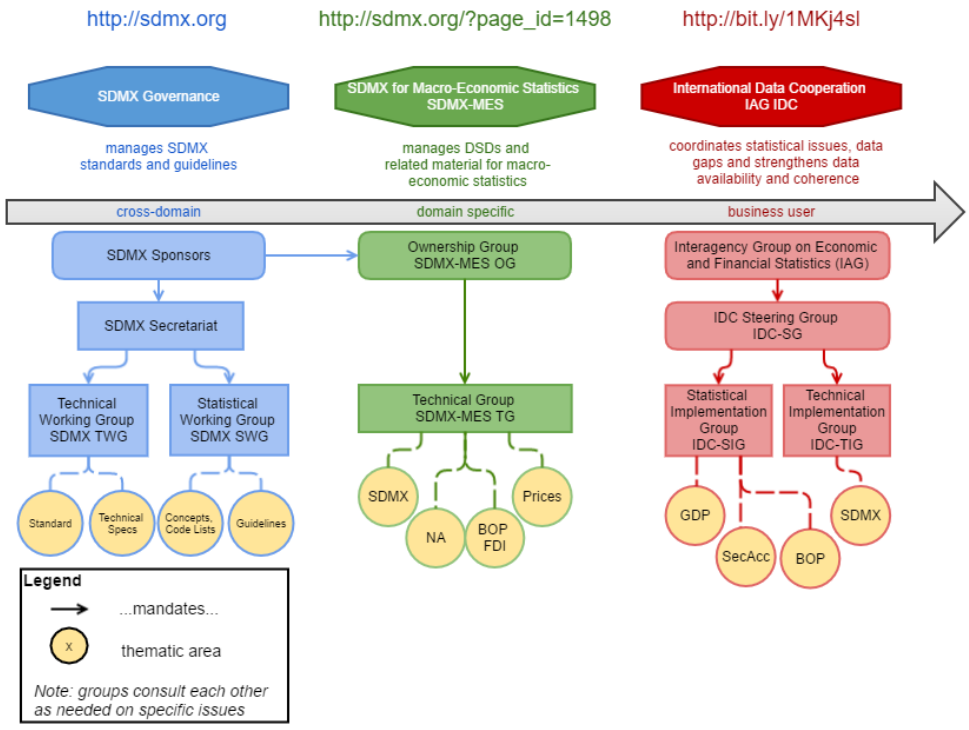
- National Accounts (including Government Finance), following SNA 2008 / ESA 2010 and SNA 1993, maintained by Eurostat;
- Balance of Payments, following BPM6, maintained by IMF;
- Foreign Direct Investments, following BMD4, maintained by OECD.

6. It is planned to include price statistics in 2016, looking into Harmonised Index of Consumer Prices (HICP), Consumer Price Indexes (CPI), Producer Price Indexes (PPI), Residential Property Prices (RPPI), Commercial Property Prices (CPPI), Import and Export Price Indexes. Purchasing Power Parities (PPPs) would however be excluded for the time being.

7. Following the governance principles outlined by the SDMX Sponsors, the maintenance agencies act upon agreement of the Ownership Group, when implementing changes to the SDMX artefacts and supporting materials. The Ownership Group mandated a Technical Group, consisting of experts with significant technical and content-related expertise in the thematic areas.

8. The IAG International Data Cooperation, as one of the major users of the data structure definitions, is closely linked to the governance structures of SDMX itself and the SDMX-MES Ownership Group. It started as a task force under the auspices of the IAG. Following the successful go-live of the first dataset on GDP and population, it has now been moved into a permanent structure. An International Data Cooperation Steering Group (IDC-SG) was mandated by the IAG to strengthen data availability and coherence by working on efficient data exchange agreements between international organisations. It is co-chaired by Eurostat and IMF. The IDC-SG created a technical and a statistical implementation group, in order to work on technical (i.e. SDMX-related) and statistical (i.e. domain-specific) issues for bringing the data exchange agreements forward.

9. The governance around SDMX in National Accounts can be visualised as follows:



10. The overview shows all 3 layers of a successful implementation: the SDMX Sponsors manage a cross-domain standard, the Ownership Group defines a domain-specific framework within that standard and finally the framework is used for exchanging data by business projects such as the IAG International Data Cooperation.

C. Where is SDMX used in Macro-Economic Statistics Today?

11. The data structures (DSDs) and related materials published by the Ownership Group are used by many organisations for several purposes such as data exchange, data reporting and data dissemination. The big advantage of using global DSDs for describing statistical data is that mapping from the data to the internationally agreed coding framework needs to be done only once and can then be reused for different purposes. For instance, having your GDP time series mapped to the global DSD for National Accounts, the same mapping can be used for reporting to international organisations, dissemination in context of SDDS+ and data exchange between national statistical actors.

12. Putting the focus on National Accounts, the table below shows a few examples for areas where the National Accounts DSDs are used:

Area	Users	Description	Comments
Reporting	National statistical institutes, central banks and other statistical organisations in Europe	ESA 2010 data, following the transmission programme and SDMX as data exchange standard.	EU legal requirement; see chapter 0, page 4.
Reporting	National statistical institutes, central banks and other statistical organisations all over the world	SNA 93/2008 data reporting to international organisations	According to bilateral agreements with OECD, IMF, UNSD…
Dissemination	National statistical institutes, central banks and other statistical organisations all over the world	Several SDDS+ data categories are covered by SDMX-MES global DSDs	National self-commitment; see chapter 0, page 7.
Dissemination & Exchange	BIS, ECB, Eurostat, IMF, OECD, UNSD, WB	International Data Cooperation in macro-economic statistics	See chapter 0, page 5.

13. In that context, the areas listed are defined as:

- Dissemination being an activity where data is provided to the public at large.
- Reporting being an activity, where a dedicated sender regularly transmits data to a dedicated receiver, following a bilateral agreement on data templates, a transmission calendar, formats and quality standards.
- Exchange being an activity where two or more organisations agree in principle to exchange certain data. Data templates, formats and quality standards are usually agreed up-front, but transmissions themselves may be ad-hoc or irregular.

II. On-Going Activities

A. SDMX in the European System of National and Regional Accounts

14. ESA 2010 is the newest internationally compatible European Union (EU) accounting framework for a systematic and detailed description of an economy. The ESA 2010 Transmission Programme presents the programme of national accounts data delivery within the framework of ESA 2010. It lays out for each table, aggregate and variable, the deadlines and frequency of transmission as well as the date of the first delivery. Particular emphasis has been given to the SDMX standard. This approach was chosen to reduce the workload of EU Member States in transmitting data to international bodies, while

reinforcing consistency as the data transmission is integrated with international partners (4). In a separate implementing regulation, the ESS has agreed to use SDMX-ML as transmission format for ESA data (5).

15. The transmission tables laid out in the transmission programme have been linked to the National Accounts DSDs published by the Ownership Group:

DSD ID	DSD Description	ESA Tables
NA_MAIN	Main Aggregates	1, 3, 5, 20, 22
NA_SEC	Sector Accounts & Government Finance	2, 6, 7, 8, 9, 11, 26, 27, 28
NA_REG	Regional Accounts	10, 12, 13
NA_SU	Supply/Use and Input/Output Tables	15, 16, 17, 18, 19
NA_PENS	Pensions	29

16. For practical reasons in data transmission, many of the tables are split in sub-tables. SDMX data flows are defined at the level of the sub-tables. As part of the package published by the SDMX-MES Ownership Group, Eurostat has published Excel visualisation templates and SDMX-ML sample files for each ESA 2010 data flow. The Excel visualisation templates can be used as input to the Eurostat SDMX Converter tool and will be maintained until September 2016. While the regulation does not specify the means by which the Member States implement the standard, they are expected to avoid using Eurostat Excel templates as parts of the production process for ESA data and achieve full SDMX implementation. Eurostat is supporting those initiatives through offering free and open source software, trainings, workshops and grants.

17. This work will increase data quality and process efficiency through built-in structural validation and also significantly reduce reporting burden through more automation and possibility of re-using the SDMX mappings for other purposes. On the basis of the common coding framework, the ESS can also agree on standardised data validation rules (for more details on benefits of SDMX for validation, see chapter 0, page 8).

B. IAG International Data Cooperation

18. As lined out above, the International Data Cooperation project started as a task force under the auspices of the Inter-Agency Group on Economic and Financial Statistics (IAG) as part of the G-20 Data Gaps Initiative. The IAG was started in 2008 by the same 7 international organisations that form the SDMX Sponsors group: the Bank for International Settlements, the European Central Bank, Eurostat, the International Monetary Fund, the Organization for Economic Co-operation and Development, the United Nations and the World Bank. It was established to coordinate statistical issues, look into data gaps highlighted by the global financial crisis and to strengthen data collection. It is chaired by the International Monetary Fund.

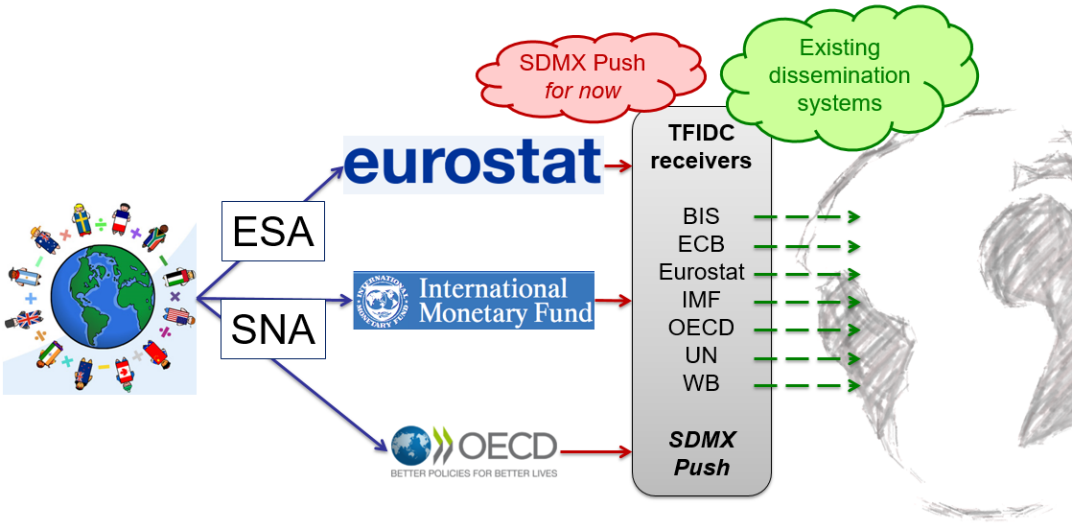
19. The Task Force International Data Cooperation was started to improve data availability and coherence. Initially two pilot projects were defined: GDP & population and sector accounts. Recently it was suggested to add a pilot on Balance of Payments. One of the main features of the cooperation is the establishment of a clear distribution of responsibilities between international organisations. Data are transmitted by national data providers to international organisations through different means and formats. Within the European regulatory framework, transmissions for National Accounts as well as Balance of Payments already follow SDMX, but that is not yet the case globally. Following national compilation, validation and transmission, data is further validated once by an international

organisation chosen as primary validator. Data is then shared and finally published through the existing dissemination systems of all international organisations involved.

20. For instance, Eurostat is responsible for the validation and subsequent sharing of the data for EU Member States, candidate countries and EFTA countries. Eurostat receives data from OECD for OECD Members, Key Partners and Accession countries, not belonging to the country group above and data from IMF for a selected range of non-OECD countries such as some of the main trading partners of the EU. These data are disseminated via specific Eurostat dissemination tables.

21. This allows users to find coherent data on the most important macro-economic indicators for most economies on the websites of all participating organisations in a single place. Some metadata is also available, which shows the most important reference information such as data source and methodology used by the respective economy.

22. The exchange process behind is based on the SDMX Push model, since not all actors involved in the process have implemented web services for global DSDs in their dissemination systems yet:



23. ESA / SNA data is sent from national to international organisations through existing reporting or exchange agreements. For each economy a “primary disseminator” is chosen. This organisation is responsible to collect data, validate it and – if not already the case – transform it into the agreed SDMX structure. It will then push the data to the TFIDC receives. The receivers then publish all data in their respective dissemination system, following again their own dissemination standards their users are used to. This setup ensures that neither initial data providers (national organisations) nor users regularly accessing the data through the website of a specific organisation have to change their working methods to use the new datasets (6).

24. The minimum quality standards are agreed upon by endorsing the checks lined out in the Eurostat and OECD protocol of co-operation in the area of National Accounts (7). In the future the IDC will also profit from having those checks defined in VTL (see chapter 0, page 8).

25. From the technical side, it is planned to move the data exchange method from SDMX Push (file transmission based) to SDMX Pull (web service based) in order to further streamline data exchange and also in order to allow other actors to join the initiative as data receivers. More details on that are explained in chapter 0 (page 9).

26. From the statistical side, more datasets will be added to the data cooperation initiative. Currently one dataset is available showing GDP, population and selected related macro-economic indicators. Work is starting to add data on sector accounts and on balance of payments.

C. SDDS Plus: EU Perspective

27. The IMF data standards initiatives were launched after the financial crisis of 1994/95 on realisation that data deficiencies and lack of transparency can contribute to market turmoil. The newest addition to the initiatives – SDDS Plus – is the highest of three tiers. It has to be noted that SDDS Plus is not a data reporting framework, but IMF members commit voluntarily to observe high standards of data dissemination from a statistical and technical perspective. Data from 27 categories are publicly disseminated by SDDS Plus countries following published release calendars and are available simultaneously to all data users (8). Data validation is under sole responsibility of the national organisation publishing the data.

28. Currently 9 countries adhere to SDDS Plus, out of which 8 are EU member states (Czech Republic, France, Germany, Italy, Netherlands, Portugal, Spain, Sweden, United States)[2]. The other EU member states subscribe to SDDS, the second tier standard[3]. Setting aside political considerations, moving from SDDS to SDDS Plus poses two implementation challenges:

- Statistical challenge: having data available for the 9 additional data categories within a certain timeframe.
- Technical challenge: creating a new national summary data page for SDDS Plus and having all data available online in SDMX-ML format. SDDS Plus supports the global DSDs available from the Ownership Group. Additionally, there is a specific "Ecofin DSD" for SDDS Plus, which also supports data categories currently not covered by global DSDs.

29. Most EU countries do not face major issues with the statistical requirements. A large part of data requested within SDDS Plus is already collected and reported to Eurostat and/or ECB under the European regulatory framework. However, several countries reported issues with setting up the technical framework and requested Eurostat and the ECB to give support in setting up an infrastructure. Basically the technical challenge can be broken down in three actions:

- SDMX mapping templates, specifying the coding used for each data point in each data category;
- A technical infrastructure creating the respective SDMX-ML files or opening SDMX web services to the data itself as well as generating specific notification messages to the IMF when data changes;
- A National Summary Data Page Plus, listing links to the SDMX files or web service endpoint and human readable data tables.

30. Eurostat, the ECB and IMF are currently working with some EU countries to facilitate the three actions for EU countries. Work has been started on creating SDMX

[2] http://dsbb.imf.org/Pages/SDDS/CountryList.aspx?sp=y
[3] http://dsbb.imf.org/Pages/SDDS/CountryList.aspx

mapping templates. These templates will identify which data is already available under which EU regulation or ECB guideline and will provide SDMX data flow coding templates for each of them. Following that, it will be investigated how the infrastructure can be setup. Some SDDS Plus adherents already use the Eurostat SDMX Reference Infrastructure[4] as a backbone for ESA 2010 as well as SDDS Plus.

III. Outlook: Working Better Together

A. Overview

31. As the SDMX 2020 Roadmap shows, SDMX for Macro-Economic Statistics and the International Data Cooperation projects have been key users of the SDMX standard in the last years. Several guidelines and templates have been based on the work done and the lessons learned within these two initiatives (1). The "Checklist for SDMX Design Projects", which will soon be published by the SDMX Secretariat, was developed by the SDMX Statistical Working Group on the basis of the SDMX-MES experience as a comprehensive guideline to support the design and creation of new SDMX artefacts and the management of such a project. The major project phases are based on the UNECE "Generic Statistical Business Process Model", managed by the High Level Group for the Modernisation of Official Statistics (9).

32. Following the business requirements in the area of National Accounts and also expressed in the strategic goals lined out by the SDMX 2020 Roadmap, the focus for the technical work stream of the International Data Cooperation initiative is now put on two areas:

- Improved data validation, based on the success of implementing the SDMX standard globally (linked to SDMX 2020 strategic goal 3.3 on data validation);
- More efficient data exchange through dissemination web services (linked to SDMX 2020 strategic goals 1.3 on data sharing and 2.1 on data dissemination).

B. SDMX and Data Validation

33. Already today, the SDMX standard allows some validation. However, this is limited to "technical checks", validating the structure of SDMX messages including ensuring correct coding and usage of the correct sub-set of the DSD within a given data flow. SDMX does not yet allow "statistical checks", defining validation rules checking the data itself.

34. From the business side, validation rules need to be:

- unambiguous
- understandable for subject-matter experts
- independent of IT tools
- standardised in order to be able to exchange them between several organisations

35. The Validation and Transformation Language (VTL) has been developed to close that gap. Using an agreed information model (e.g. SDMX), data is seen as a mathematical function, to which transformations can be applied. Following that logic, validation is only a

[4] https://webgate.ec.europa.eu/fpfis/mwikis/sdmx/index.php/SDMX_Reference_Infrastructure_SDMX-RI

special kind of transformation, where the result of the transformation applied on the dataset is compared to an expected result defined by the validation expression (10).

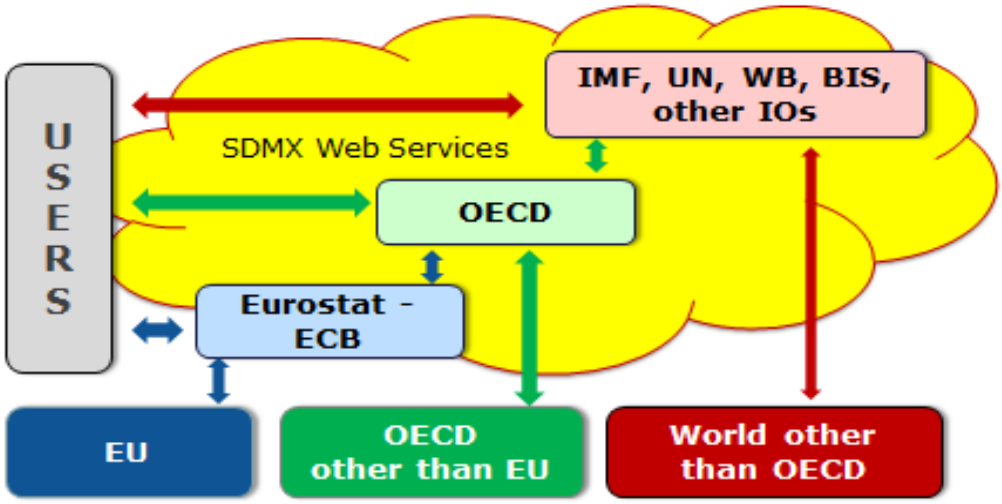
36. From the content side, an ESA validation task force is currently looking into defining validation rules for National Accounts in Europe. Since ESA follows SNA, it is expected that those rules will also be largely applicable internationally. The work of this task force will facilitate expressing the validation framework agreed upon in the IDC initiative also as VTL rules, which can then be directly shared and built into validation services of participating organisations.

37. In parallel Eurostat is currently working with the ESS on shared services for structural and content validation. National Accounts is a forerunner in that area to test the approach also for other statistical domains. In that context, the structural validation service will be based on the SDMX standard and the content validation service on VTL. The services follow the principles of sharing statistical services as laid out by UNECE in the "Common Statistical Production Architecture" (CSPA) (11).

C. Data from Producer to Consumer in Real Time

38. Combining the standardisation of data brought by SDMX, the standardisation of validation brought by VTL and technologies such as web service for data as well as shared statistical services for processing, it is expected that data will travel much faster and in better quality from the producer to the final user. The vision of International Data Cooperation is to have coherent macro-economic data of highest quality from the producer to the user in real time.

39. A possible data sharing model has been visualised as follows:



40. Economies would transmit data to the international community through a primary disseminator. Reporting burden is reduced to the minimum by having a single flow from national to international. Data is quality assured by applying agreed validation checks to it and data is available in the public domain through SDMX web services. End users could access the data for any economy through any organisation and web services would ensure that the latest data is available to anybody at any point in time (6). The International Data Cooperation initiative can in principle also be extended to other organisations flexibly, which provides a solid basis for official macro-economic statistics of high quality in the public domain.

References

SDMX Sponsors Group. SDMX Roadmap 2020. SDMX.org. [Online] 1 March 2016. [Cited: 1 April 2016.] https://sdmx.org/?sdmx_news=sdmx-roadmap-2020.

Task Force International Data Cooperation. Different portals, same data: A step towards more coherent macro-economic data. [Online] 1 June 2015. [Cited: 1 April 2016.] http://ec.europa.eu/eurostat/documents/24987/6642470/TFIDC_2015-07_CommunicationNote.pdf.

Ownership Group for SDMX in Macro-Economic Statistics. Terms of Reference of the Ownership Group SDMX in Macro-Economic Statistics. [Online] 9 February 2016. [Cited: 1 April 2016.] https://sdmx.org/wp-content/uploads/SDMX-MES_OG_TermsOfReference.pdf.

European Commission. European System of Accounts: ESA 2010 Transmission Programme for Data. Luxembourg : Publications Office of the European Union, 2014.

Commission Implementing Regulation (EU) No 724/2014 of 26 June 2014 on the interchange standard for the transmission of data required under Regulation (EU) No 549/2013. European Commission. 1 July 2014, Official Journal of the European Union, Vol. L 192/38.

Suranyi, Daniel. International data sharing and macro-economic data exchange. SDMX Global Conference (Session 7). [Online] 30 September 2015. [Cited: 1 April 2016.] http://unstats.un.org/unsd/SDMX/sdmxglobalconference2015/default.html.

Eurostat, OECD. Protocol for co-operation between Eurostat and the OECD in the area of National Accounts. [Online] 12 June 2013. [Cited: 1 April 2016.] http://ec.europa.eu/eurostat/documents/24987/6624264/OECD-Eurostat-cooperation.pdf.

Hake, Andreas. The IMF Data Standards Initiatives: Moving Towards the SDMX Data Sharing Vision. SDMX Global Conference (Session 7). [Online] 30 September 2015. [Cited: 1 April 2016.] http://unstats.un.org/unsd/SDMX/sdmxglobalconference2015/default.html.

UNECE. Generic Statistical Business Process Model (GSBPM). UNECE Statistics Wikis. [Online] December 2013. [Cited: 1 April 2016.] http://www1.unece.org/stat/platform/display/GSBPM/Generic+Statistical+Business+Process+Model.

Pellegrino, Marco and Del Veccio, Vincenzo. VTL: A new international standard for enhancing data validation and processing. SDMX Global Conference (Session 8). [Online] 30 September 2015. [Cited: 1 April 2016.] http://unstats.un.org/unsd/SDMX/sdmxglobalconference2015/default.html.

UNECE. Common Statistical Production Architecture. UNECE Statistics Wikis. [Online] December 2015. [Cited: 1 April 2016.] http://www1.unece.org/stat/platform/display/CSPA/Common+Statistical+Production+Architecture+Home.