Second Phase of the G20 Data Gaps Initiative (DGI-2)

Status report on the implementation of Recommendation II.20 “Promotion of Data Sharing”

Prepared by the Deutsche Bundesbank Staff, Eurostat Staff and IMF Staff

6 January 2020

Table of Contents

List of abbreviations

AnaCredit	-	Analytical Credit Datasets
BACH	-	Bank for the Accounts of Companies Harmonized
BEA	-	Bureau of Economic Analysis
BIRD	-	Banks' Integrated Reporting Dictionary
BIS	-	Bank for International Settlements
BoE	-	Bank of England
CB	-	Central Bank
CPMI	-	Committee on Payments and Market Infrastructures
CSDB	-	Centralised Securities Database
DISC	-	Data Intelligence Service Centre
DGI	-	Data Gaps Initiative of the G20
DGI-II	-	Second phase of the Data Gaps Initiative
DSD	-	Data Standard Definition
EBA	-	European Banking Authority
ECB	-	European Central Bank
ECCBSO	-	European Committee of Central Balance-Sheet Data
EGR	-	EuroGroups Register
EMIR	-	European Market Infrastructure Regulation
ESCB	-	European System of Central Banks
ESS	-	European Statistical System
EU	-	European Union
FDI	-	Foreign Direct Investment
FRIBS	-	Framework Regulation Integrating Business Statistics in the EU
FSB	-	Financial Stability Board
G-SIBs	-	Globally Systemically Important Banks
G20	-	Group of Twenty
GLEIF	-	Global Legal Entity Identifier Foundation
iBACH	-	individual Bank for Accounts of Companies Harmonized project
IDH	-	BIS International Data Hub
IDIS	-	Integrated Digital Information System
IFC	-	Irving Fisher Committee on Central Bank Statistics
IHSN	-	International Household Survey Network
IMF	-	International Monetary Fund
INEXDA	-	International Network for Exchanging Experience on Statistical Handling of Granular Data
IOs	-	International Organizations
IOSCO	-	International Organization of Securities Commissions
IReF	-	Integrated Reporting Framework of ESCB

ISIN	-	International Security Identification Number
ISO	-	International Organization for Standardization
ISWGHHS	-	UN Inter-Secretariat Working Group on Household Surveys
LEI	-	Legal Entity Identifier
MDIS	-	Micro-Data Integration Service
MMSR	-	Money Market Statistical Reporting
MNE	-	Multinational Enterprise
MoU	-	Memorandum of Understanding
NCB	-	National Central Bank
NSO	-	National Statistical Office
OECD	-	Organisation for Economic Co-operation and Development
OFR	-	Office of Financial Research
ONAs	-	Other National Authorities producing statistics
OTC	-	Over-the-counter
RDC	-	Research Data Centre
RIAD	-	Register of Institutions and Affiliates Data
SDD	-	Single Data Dictionary
SDMX	-	Statistical Data and Metadata Exchange
SHSDB	-	Securities Holdings Statistics Database
SWIFT	-	Society of Worldwide Interbank Financial Telecommunication
UK	-	United Kingdom
UNCTAD	-	United Nations Conference on Trade and Development
UNECE	-	United Nations Economic Commission for Europe
UPI	-	Unique Product Identifier
US	-	United States
UTI	-	Unique Transaction Identifier

Management Summary

The report highlights the progress made in each of the recommendations for promoting data sharing. General trends and selected examples were mentioned. The core messages for individual recommendations are summarized below.

Regarding the promotion and use of common statistical identifiers (recommendation 1), the importance of the Legal Entity Identifier (LEI) was underlined by all respondents. For example, the banking industry has developed options to include the LEI in current and future SWIFT payment messages and European Union (EU) legislation has made the use of unique identifiers such as the LEI mandatory in various areas. A large number of reporting countries also use the LEI to compile part of their statistics. Other identifiers used are the International Securities Identification Number (ISIN), Unique Product Identifier (UPI) and Unique Transaction Identifier (UTI). Only one reporting country currently does not intend to use common identifiers, and another foresees more intensive cost-benefit analysis before a more intensive use of common identifiers.

Many activities have been conducted to exchange experience on statistical work with granular data and to improve transparency (recommendation 2): the UNECE Task Force on exchange and sharing of economic data has published its final report, international Working Groups continue, various reports appeared and workshops took place or are foreseen for the future. At national level, cooperation between the National Statistical Office (NSO) and the Central Bank (CB) as well as Other National Authorities (ONAs) is the key driver.

The International Network for Exchanging Experience on Statistical Handling of Granular Data (INEXDA) is the most mentioned initiative to promote the exchange of experience on statistical work with granular data. Many respondents are already INEXDA members or guests.

The feedback underlined that as much information as possible should be published in an appropriate manner, while balancing user' needs and data confidentiality (recommendation 3). One way to achieve this is to develop ways that allow users to be able to use information without directly having access to but based on flexible compilation of granular data. This happens in some countries through the establishment of Research Data Centres (RDCs).

Linking different datasets (recommendation 4) is facilitated through sharing corresponding knowledge and using common identifiers. Once again, the cooperation between NSOs, CBs, and ONAs plays an important role at national level. Examples include signing of MoUs between NCBs and NSOs, the integration of bank and business data and the implementation of joint Working Groups. One country reports that work related to linking different datasets is still under consideration.

Recommendation 5 – the provision of data at the international level – was seen against the background of globalisation. Examples refer to systems or frameworks that aim at ensuring a regular collection and appropriate sharing of data, for instance the Bank for International

Settlements' (BIS) International Data Hub (IDH), the individual Bank for Accounts of Companies Harmonized (iBACH) project of the European Committee of Central Balance-Sheet Data (ECCBSO) or the implementation of the Framework Regulation Integrating Business Statistics (FRIBS) in the EU. Also the promotion of statistical standards like SDMX was commonly mentioned. But the progress made across economies is not at all homogeneous. Nonetheless, countries that are still considering their actions support recommended measures overall.

Progress for recommendation 6 (alternative ways of improved sharing of granular data) is mainly achieved on national level. Key to success is the adoption of national law and a continuously improved cooperation between national authorities. A large number of institutions have set up new Data Labs or RDCs. These units mostly offer a kind of open data room for granular data where researchers can access data on the spot. This allows access to more sensitive data in a safe environment and the use of advanced analytical tools.

Standardisation and integration of data is a priority in order to achieve the goal of collecting data only once and avoiding multiple collections (recommendation 7). The European System of Central Banks (ESCB) Integrated Reporting Framework (IReF) and Banks' Integrated Reporting Dictionary (BIRD) are key elements of this strategic approach at the ESCB.

In some cases, countries have already achieved the goal of collecting data only once in a few areas. This allows the opportunity to share granular data with supervisors and statisticians, who then aggregate the granular data reports depending on the respective purpose.

A targeted, step-by-step approach for data sharing seems warranted: (1) Work to identify a problem that can only be improved by granular microdata exchange. (2) Limit the exchange to data required to solve the problem. (3) With time expand data sharing for other purposes that require granular data from other sources.

This approach will support progress on other DGI-2 recommendations, where data sharing within and between countries will allow greater quality and consistency, particularly for the recording of financial and non-financial multinational enterprises. Notable examples of the recommendations that will benefit include:

II.4 Globally Systemically Important Financial Institutions
II.5 Shadow Banking
II.6 Derivatives
II.8 Sectoral accounts
II.9 Household distributional information
II.10 International Investment Position
II.12 Portfolio investment
II.14 Cross border exposure

1. Introduction

1.1. Background

1. The key outcome of the Workshop on Data sharing, organized by the IMF and Eurostat, in cooperation with and hosted by the Deutsche Bundesbank in Frankfurt am Main, Germany, on January 31-February 1 2017 included an agreement on common terminology, the identification of the main barriers preventing the sharing of disaggregated data and micro data, and ideas on possible approaches to overcome such barriers. The workshop concluded with seven recommendations, providing guidance at national and international levels on improved accessibility and sharing of granular data.[1]

2. These recommendations were welcomed by G20 Finance Ministers and Central Bank Governors at their meeting in Baden-Baden on 17 and 18 March 2017 and by the G20 leaders at their subsequent meeting in Hamburg.[2]

3. Given the qualitative nature of the seven recommendations to promote data sharing, the 2017 DGI Global Conference agreed to launch a questionnaire to identify best practices in data sharing and the progress achieved in implementing these recommendations, including in relation to other recommendations of the DGI-2[3]. The proposed questionnaire was agreed at the 2018 DGI Global Conference. This report summarizes the feedback received.

4. The questionnaire was distributed in August 2018. 22 answers were received: 16 from G20 economies and 6 from international organisations.

1.2. Introducing the structure of the report

5. After this introduction, the second section of the report will provide a summary of the outcomes of the questionnaire, structured by the individual recommendations. The global, regional and national developments are described in separate paragraphs. Selected examples, as well as general progress, are summarized (in alphabetical order of the countries' names). The report concludes with an assessment of the usefulness of the recommendations, highlights the key findings, and suggests a way forward.

6. The report focuses on granular data and progress made in the second phase of the Data Gaps Initiative (DGI-2). Responses to the questionnaire primarily reflecting progress in macro statistics or occurrences before 2015 are not included in the report. For practicability reasons the report can only contain a selection of contributions received.

[1] See Annex
[2] https://www.g20germany.de/Webs/G20/EN/G20/Summit_documents/summit_documents_node.html and https://www.g20germany.de/Content/EN/StatischeSeiten/G20/Texte/g20-gipfeldokumente-en.html.
[3] The most relevant other recommendations are mentioned in the management summary.

2. Recommendation 1 – Promoting the use of common statistical identifiers

Economies and international organizations, as appropriate, are encouraged to foster the use of common identifiers and make every effort to adopt the latest international conceptual frameworks.

2.1. Global developments

7. The BIS uses the LEI for identifying G-SIBs'[4] top 50-100 counterparties.

8. FSB and CPMI have organized workshops with industry experts to further promote the LEI. In response, the banking industry has developed options to include the LEI in current and future SWIFT payment messages, which should provide further incentives for non-financial corporations to obtain and maintain LEIs.

9. FSB published the Peer Review Report[5] "Thematic Review on Implementation of the Legal Entity Identifier" at the end of May 2019. The Peer Review Report summarizes the findings and gives the following recommendations, among others, to FSB jurisdictions:
a) follow-up on CPMI-IOSCO guidance and use of LEIs for the identification of legal entities in the data reported to trade repositories for OTC derivatives,
b) consider requiring the use and timely renewal of the LEI in reporting or disclosure frameworks, for the identification of all entities in major financial groups, a wider set of financial market participants and infrastructures, their counterparties, and related entities (including direct and ultimate parents), especially in a cross-border context,
c) explore ways to promote further LEI adoption, for instance by fostering nationwide implementation strategies to maximise the cross-sectoral benefits of the LEI, communicating on LEI benefits through public outreach initiatives, leading by example in obtaining LEIs for the central bank and other public sector bodies, especially issuers of public debt, and considering the potential for LEI use before introducing new identifiers.

10. The IMF is providing Capacity Development (training and technical assistance) to members to foster the use of the most modern conceptual frameworks for economic and financial statistics. In addition, the IMF Data Standards Initiatives are encouraging subscribers to publish data in SDMX format using the agreed coding structures (Data Standard Definitions (DSDs) for global use and/or the IMF EcoFin DSD).

11. For external individual loan-by-loan external debt, the World Bank has developed a common identifier from the collected country data. This common identifier has been shared with UNCTAD and COMSEC to facilitate international reporting and in-country analysis using their respective tools.

[4] See Recommendation II.4 Globally Systemically Important Financial Institutions of the second and third DGI-2 Progress report for further details.
[5] See https://www.fsb.org/wp-content/uploads/P280519-2.pdf.

2.2. Regional developments

12. EU legislation has made the use of unique identifiers such as ISIN and LEI mandatory in various areas (e.g. EU Prospectus Regulation, Securities Financing Transactions Regulation, European Markets Infrastructure Regulation, Markets in Financial Instruments Regulation). The ISIN coverage in the EU is almost complete.

13. Eurostat maintains a common identifier for European enterprises, within the EuroGroups Register (EGR), and has been very active in the promotion of the use of SDMX and related common data structures, both within Europe and outside.

2.3. National developments

14. National examples are the following:

- Australia uses the National Business Number as a common identifier for taxation and statistical purposes by a wide variety of government and private agencies.
- In Germany, common identifiers are further examined at national level. An initiative investigates the possibility for the creation of a basic register with reference data on enterprises, which includes the establishment of a cross-agency business identifier.
- A reform of the registration system for companies by promoting five certificates incorporated into one licence with only one code was issued in China. An information sharing platform has been created and connected with different statistical institutions.
- Italy uses the LEI code, when available, in both the proprietary multi-purpose statistical register and in the statistical reporting to the central bank.
- In Japan, data from trade repositories to supervisors also includes LEI. Periodic seminars are offered by supervisors and encourage financial market participants to take LEI on board. UTI and UPI are included or will follow later.
- Korea starts issuing LEI.
- In Russia, LEI and ISIN are used for compiling external sector statistics including FDI statistics and financial accounts. UTI has already been implemented in the OTC derivatives reporting requirements to the trade repositories in order to avoid duplication.
- Spain has recently included the LEI code into the deposit of annual financial statements for all corporations for the mercantile register.
- Turkey has promulgated the regulation for trade repositories and will obligate reporting companies and counterparties of derivative transactions to have an LEI.
- The United States Bureau of Economic Analysis (BEA) and Federal Reserve have both started to use the LEI within their reporting requirements. The BEA collects LEIs from FDI survey respondents. The Federal Reserve makes LEI-based information available to both regulators and the public.
- A few other G20 economies either currently do not have any intention to move towards common identifiers or judge the work on identifiers highly preliminary and

view cost benefit decisions for adoptions of identifiers inherently challenging as the cost are clear, upfront, and often significant, while the realisation of the benefits depends critically on a broad-based adoption.

2.4. Outlook

15. The G20 developed the LEI initiative with the goal of covering all parties of financial contracts. Further progress to achieve that goal is required.

16. In the context of the consultation process regarding the "Proposal for a Directive of the European Parliament and the Council amending Directive (EU) 2017/1132 as regards the use of digital tools and processes in company law", the Global Legal Entity Identifier Foundation (GLEIF) announced that future fees charged by the LEI issuing organisations to the Legal Entity requesting a LEI may possibly be waived. Instead, other types of compensation for the GLEIF could be discussed, if LEIs were issued by official business registers.

3. Recommendation 2 – Promoting the exchange of experience on statistical work with granular data and improving transparency

Economies are encouraged to have a more active dialogue and exchange of practical experiences, in particular on the accessibility of granular data, metadata, and on techniques for statistical analysis and data protection.

3.1. Global developments

17. The International Network for Exchanging Experience on Statistical Handling of Granular Data (INEXDA) was launched in January 2017 to promote the exchange of experience on statistical work with granular data[6]. INEXDA handles issues such as the accessibility of data and metadata, techniques for the statistical analysis of granular data, procedures for the confidentiality and security of data, and methods of output control. Furthermore, INEXDA provides a framework for investigating possible methods of harmonising access procedures and metadata structures, developing comparable structures for existing data and further fostering the efficiency of statistical work with granular data. Its ultimate aim is to facilitate the use of granular data for analytical, research and comparative purposes by users outside the participating institutions, within the limits set by the applicable confidentiality regulations.

18. INEXDA members are[7]: Banca d'Italia, Banco Central de Chile, Banco de España, Banco de Portugal, Bank of England, Bank of Russia, Banque de France, Central Bank of the Republic of Turkey, Deutsche Bundesbank, Eurostat and European Central Bank.

[6] https://www.bundesbank.de/resource/blob/604908/93625a234cbaa68e609186beab7d0654/mL/inexda-memorandum-data.pdf.
[7] As of May/June 2019. More institutions have signalled their interest.

INEXDA guests are institutions that have participated or have confirmed participation in at least one INEXDA meeting but have not yet signed the MoU. INEXDA guests are[8]: Banco de México, Bank for International Settlements[9], Federal Statistical Office of Germany, Oesterreichische Nationalbank, Office for National Statistics UK, and Swiss National Bank. Other central banks, national statistical institutes and international organisations are welcomed to join or participate in INEXDA.

19. The IMF regards Global Data Commons as one of the six strategic priorities of the Overarching Data and Statistics Strategy of the Fund, endorsed by the IMF Executive Board in March 2018. The global data commons are an integrated network of official websites publishing data essential for surveillance on a pre-announced schedule.[10]

20. The World Bank provides tools and support to member states to improve the documentation of their household surveys (improved metadata), and tools that facilitate searching these metadata. They also work to provide support to member states to anonymize their household survey data so that this granular-level data can be publicly disclosed. World Bank shares its experiences with others in the international community via the International Household Survey Network (IHSN) and the UN Inter-Secretariat Working Group on Household Surveys (ISWGHHS).

21. The Irving Fisher Committee on Central Bank Statistics (IFC), comprising 91 central bank members, deals with manifold issues related to the exchange of knowledge. For example, in January 2015 the IFC report on Data-sharing: issues and good practices and in December 2016 the IFC Report on the sharing of micro data – a central bank perspective were published[11]. The 2018 biennial IFC conference devoted sessions on data sharing.[12]

22. The UNECE Task Force on exchange and sharing of economic data operates under the umbrella of the United Nations. The objective of the task force is to provide guidance on national and international exchanges of economic data and to identify innovative ways of exchanging economic data in order to increase the quality, coherence and granularity of statistics and to better analyse the activities of MNEs.[13]

23. The OECD has set up the Working Group on International Investment Statistics. There, countries are encouraged to share their experiences with confidentiality practices.[14]

[8] As of May/June 2019.
[9] The Bank for International Settlements supports the INEXDA initiative without being a full member.
[10] https://www.imf.org/en/Publications/Policy-Papers/Issues/2018/03/20/pp020918imf-executive-board-supports-new-strategy-for-data-and-statistics-in-the-digital-age.
[11] https://www.bis.org/ifc/publ/ifc-report-microdata.pdf and https://www.bis.org/ifc/events/7ifc-tf-report-datasharing.pdf.
[12] https://www.bis.org/ifc/events/9thconf.htm
[13] https://www.unece.org/statistics/networks-of-experts/task-force-on-exchange-and-sharing-of-economic-data.html.
[14] http://www.oecd.org/daf/inv/oecdinvestmentcommittee.htm.

24. Finally, the G20 Data Gaps initiative Contact Group and associated events provide a platform for sharing experiences on statistical work with granular data.

3.2. Regional developments

25. At European level, the ECB in cooperation with BoE and Office of Financial Research (OFR) promotes the exchange of experience on best practices in using granular data through a joint initiative on the standardisation of granular data (e.g. through the regular "Setting Global Standards for Granular Data" workshops organised by ECB, BoE and OFR). The ECB is also regularly welcoming delegations of other central banks and statistical institutes to share its experiences in the development and management of micro databases.

26. The EU-project “European System of Business Registers” is also a platform for the exchange of microdata as well as methods, techniques and practical experiences at the national level of EU countries.

3.3. National developments

27. National examples are the following:

- Australia maintains a number of bilateral relationships to cooperate in data sharing.
- At national level, new forms of cooperation between National Statistical Offices (NSOs) and Central Banks (CBs) as well as other National (statistical) Agencies (ONAs) have been implemented.
- Some G20 economies report data sharing initiatives (e.g. concerning Business Register for statistical purposes), and support for research and Big Data projects. Some of the main tasks are the strategic development of a research data infrastructure and accreditation and evaluation of research data centres (RDCs). Also exchange of experience on record linkage with company data is practiced. In a similar way the ongoing exchange between RDCs and other RDCs or institutions, which are on the way to establish a RDC, is an example of closer cooperation efforts.
- China set up a joint data development centre between its Statistical Office and a university to open the NSO's granular data to all domestic higher educational institutions and their subordinate non-profit research institutes.
- Korea implements a Micro-Data Integration Service (MDIS), shares raw data related to statistical production, and provides separate cases of statistical analysis using raw data to government ministries, local governments and research institutes.
- The Bank of Russia maintains ongoing relationships with key neighbouring partner-countries as well as holds occasional bilateral meetings on external sector statistics with a number of other countries, mainly from the EU.
- The US BEA maintains an ongoing relationship with certain partner agencies, notably on asymmetries, and is actively engaged in the US government’s initiative to leverage data as a strategic asset. The US Federal Reserve facilitates a Data

Management Strategies and Technologies Workshop, with participants from Central Banks of many countries.

3.4. Outlook

28. Further cooperation between NSOs, CBs and ONAs is of particular importance. Some examples of potential projects were mentioned in the responses, one of them with the strategic objective of promoting and guaranteeing the dissemination of statistics through digital platforms.

29. Italy is setting up a Research Data Center (RDC). The objective is to expand and improve the elementary data available to economic researchers, both internal and external to Banca d'Italia. The project includes the creation of a lab, where confidential microdata can be used and processed under the supervision of RDC personnel, in due respect of related restrictions, and as a remote execution application a web tabulator, depending on the characteristics of the different datasets and the specific user.

4. Recommendation 3 – Balancing confidentiality and users' needs

Economies are encouraged to revisit existing confidentiality rules, practices, and approaches in light of evolving users' needs and taking into account peer practices, seeking to maximize, as appropriate, the amount of information released while maintaining the confidentiality of data deserving protection.

4.1. Global developments

30. For banking statistics, the BIS has worked closely with reporting countries to revise confidentiality rules and increase dissemination, especially by promoting new practices. For the very sensitive data collected by the International Data Hub, ways have been developed to allow other IOs / Committees to be able to use derived information without having access to the granular data itself.

31. The World Bank works with member state statistical offices to anonymize their confidential data so that public use versions can be disseminated.

4.2. Regional developments

32. To promote the public availability of supervisory data to users, the ECB supports the integration of supervisory reporting with Pillar 3 disclosure requirements, which would allow the EBA to publish in a single repository the Pillar 3 data disclosed by banks. Clear rules and criteria as well as related user awareness must be in place for data usage of confidential statistical information. Eurostat coordinates work on statistical confidentiality in the European Statistical System, based on the relevant European legislative requirements, and publishes public use files for selected microdata sets.

4.3. National developments

33. Issues on balancing confidentiality and users' needs were put on the agenda in many countries. National examples are the following:

- Data already in the public domain can be disclosed in Australia, without getting a specific permission for the disclosure of data. Nevertheless, data on individuals can only be used for purposes that were explicitly disclosed when the data were collected. The longitudinal census data set is created from data where explicit permission for retention was given on the collection instrument.
- A recurring item is a close cooperation between CBs and institutions of the National Statistical System and the exchange of statistical information between them. In Argentina, national (statistical) plenary meetings are organised covering topics such as confidentiality, open data, anonymization, statistical best practices (including information dissemination), quality management processes and certification. This goes hand in hand with a new policy on granularity of data, with the aim of reducing transaction costs associated with dissemination, discovery, access and use of data. New regulatory frameworks regarding internal data sharing to facilitate the exchange of information between agencies are being implemented.
- Canada has begun to publish considerable additional macroeconomic data based on a review and updates to confidentiality rules, which provide for additional dissemination while respecting confidentiality of all microdata.
- In Germany, passive confidentiality principles are already applied in trade in goods statistics.
- A new process to assess users' needs in order to provide access and dissemination of statistical information was established in Mexico. The reporting agency is obliged to perform a mandatory public consultation for any new product or methodological change, which allows the determination of users' needs, and then to identify the viability of those data needs. Another measure is to provide access to microdata for free, for example through a Microdata Laboratory, which is commonly used for research purposes without violating the confidentiality of the underlying information.
- The establishment of a Research Data Centre was also mentioned several times. There, researchers can use and work with anonymised (micro-)data in a secure and safe environment within the institution. By doing so, the researchers are allowed to use the non-confidential results of their findings but the underlying raw data never leave the institution.
- Russia has established a special "closed" system with restricted information for staff maintaining processes of loans' data collecting subject to different access rights for staff supervising banks, compiling macroeconomic statistics and implementing monetary and financial stability policies.
- The US Federal Reserve makes publically available a tremendous amount of bank-level or bank-holding-company-level data, both for publically traded and privately-

held institutions. Where the Federal Reserve holds sensitive (non-publishable) data, it has the legal authority to share.

- Some countries reported progress on the expanded use of dispersion indicators for the dissemination of aggregated statistics on non-financial corporations. This allows preserving the confidentiality of the individual data while adding information on the distribution of microdata.

4.4. Outlook

34. Canada intends to review existing confidentiality practices in the medium term.

35. The US BEA is studying several potential approaches to increasing data availability without compromising confidentiality.

36. For the Eurosystem, the ECB has established a Task Force with NCBs to analyse how the use of confidential statistical information could be maximised whilst managing confidentiality risks.

37. In the EU, the measurement of globalisation benefits from ongoing pilot exercises on exchange between National Central Banks (NCB) and NSOs regarding MNEs.

5. Recommendation 4 – Linking different datasets

Economies are encouraged to allow for the fullest possible use of data (for policy making, research, and statistical purposes) including through linking of different non-public datasets (possibly from different institutions) ensuring that strong governance and confidentiality protection arrangements are in place.

5.1. Global developments

38. The BIS is building an internal data lake that will help facilitating the linking across datasets, by having better metadata documentation and apply new machine learning techniques (e.g. fuzzy matching).

39. The OECD encourages members of the Working Group on International Investment Statistics to share their experiences with linking data at the national level. These discussions cover many practical aspects of data linking, including the value of common identifiers, how to link datasets that use different statistical units (e.g. establishment-level data with firm-level data) or sampling methodologies, and institutional arrangements used for such exercises.

40. The IMF encourages the three key producers of macroeconomic data – the NSO, the central bank and Finance Ministry – to share data on a routine basis, including ensuring that data supplied to policymakers are consistent across datasets.

41. The World Bank links data from various sources, such as linking satellite imagery with household survey data. Steps are taken to ensure these spatial aggregations do not reveal individual/confidential information to the general public.

5.2. Regional developments

42. The ECB's securities micro databases, Centralised Securities Database (CSDB) and Securities Holdings Statistics Database (SHSDB), are linked through the ISIN code. The process of linking of CSDB/SHSDB with the Register of Institutions and Affiliates Data (RIAD) and the future Analytical Credit Datasets (AnaCredit) via the RIAD code is in the final stage. The ECB has also developed a Single Data Dictionary (SDD), which provides a unified logical model to store data set related metadata and to map codes and classifications across data sets. The SDD covers all data sets available at the ECB. In addition, a new Data Intelligence Service Centre (DISC) was set up, which provides a single platform for accessing and linking the ECB's microdata sets (including CSDB data, ratings data, SHS data, supervisory data, RIAD data, MMSR[15] data, and EMIR[16] data) as well as macro data sets.

5.3. National developments

43. On national level, cooperation between NSOs and CBs as well as ONAs has an important role again.

- In Argentina, a Memorandum of Understanding (MoU) between the CB and NSO has enabled the sharing of supervisory data at granular level (balance sheets and market transactions) for statistical purposes. An important factor is the integration of governmental systems and data generation, avoiding administrative redundancy and easy exploitation of data.
- Australia has established a number of National Government "Integrating Agencies". Regional agencies can apply to access a variety of data sets and also provide their own data sets for integration and analytical purposes. A Data Integration Division has been established at the NSO to manage the increasing demand for the services. At present access for non-government users is limited but experimental work with a "5 Safes"[17] approach has commenced.
- France and Germany report that they try to combine business and bank data with commercial and big data for internal and external research, also in close cooperation between their respective CB and NSO.
- Germany has established a joint working group for the exchange of business data for statistical purposes between the central bank and the statistical office.

[15] MMSR – Money Market Statistical Reporting
[16] EMIR – European Market Infrastructure Regulation
[17] Managing the risk of disclosure: The five safes framework
The five elements of the framework are:
(1) Safe People, (2) Safe Projects , (3) Safe Settings , (4) Safe Data , (5) Safe Outputs.
http://www.abs.gov.au/ausstats/abs@.nsf/Latestproducts/1160.0Main%20Features4Aug%202017?opendocument&tabname=S

- In Italy, data-linking of various internal data collections and registers is adopted very frequently in order to enhance the efficiency of the system. This is supplemented by an extensive implementation of data linking with external sources.
- Mexico mentions a deeper exploitation and use of administrative records.
- The Bank of Russia employs a system that allows the enrichment of the loans database with other sources like centralized balance sheet database and business registers as well as records maintained by the NSO and ONAs.
- In Turkey, a single Big Data Hub includes microdata from various data sources.
- The US BEA links its data on the activities of US affiliates of foreign MNEs with datasets collected by other US statistical agencies, for example the Census Bureau's establishment level data.

5.4. Outlook

44. While the ECB's Single Data Dictionary (SDD) and Data Intelligence Service Centre (DISC) are currently only available within the ECB, there are plans to also provide the ESCB with access in the future.

45. The US BEA has projects underway to link its microdata to data held by the Bureau of Labor Statistics and the Census Bureau, which face the obstacle of a lack of common identifiers.

46. When developing new statistics, the need to be able to link different data sets should be taken into consideration. Further measures to ensure the unique identification of financial instruments and entities seem required to facilitate the linking of data sets (e.g. by making the use of the relevant identifiers mandatory).

47. Not only public authorities but also the financial industry could benefit from advancing standardisation, e.g. in the realm of digital representation of (financial) contracts in their analytical or operational systems.

6. Recommendation 5 – Provision of data at the international level

Economies are encouraged to have in place regional and international data exchange systems and appropriate legal frameworks to allow for sharing and exchange of granular data across borders.

6.1. Global developments

48. On the international level, data on global systemically important banks (G-SIBs) are provided to the International Data Hub (IDH) hosted by the BIS. The IDH ensures the regular data collection and appropriate sharing of these data between its member institutions.

49. Other examples of progress at global level under recommendation 5 include the promotion of the ISO-standard SDMX[18], the provision of open source tools for the international community, and the introduction of international data cooperation between international institutions (for national accounts, population and sector accounts data at present, and being extended to balance of payments data).

50. All IMF Reports on the Observance of Standards and Codes in the statistical area strongly advise authorities to strengthen the legal and regulatory bases for data sharing by all official data compilers. In addition, through the implementation of the data standards, the IMF helps countries to use the SDMX and cloud technologies to disseminate data.

6.2. Regional developments

51. At European level, the exchange of granular data is common within the ESCB on the one hand and the ESS on the other hand, and practised for particular data sets also between the two statistical systems (e.g. FDI network and EuroGroups Register (EGR)).

52. An example for the sharing of microdata concerning nonfinancial corporations is the iBACH project. The corresponding dataset contains balance sheet and profit and loss data of firms collected by the European Committee of Central Balance Sheet Data Offices (ECCBSO) through its Working Group on Bank for the Accounts of Companies Harmonized (BACH)[19].

53. With the implementation of the Framework Regulation Integrating Business Statistics (FRIBS), the EU intends to consolidate and further harmonize the legal bases for business statistics. FRIBS contains a set of guidelines that will help to better capture multinational activities and gives a legal basis for data sharing in specific domains.

54. Furthermore, in 2017, the ECB and the US Federal Reserve Board performed a first pilot exercise on exchanging security-by-security data to cross-check the sectoral classification of issuers of relevant EU and US securities.

6.3. National developments

55. Good examples on national level are:

- Secondment remains a potential vehicle in Australia. It was used by the Department of Foreign Affairs and Trade together with the OECD on a recent task.
- To make statistical data available to other international users, many countries publish the English version on the institution's website as a first step.
- Introducing SDMX classification and sharing data with international organisations is also mentioned.

6.4. Outlook

[18] SDMX - Statistical Data and Metadata exchange

[19] More information can be found via the following link: https://www.eccbso.org/wba/pubblica/database.asp

56. A key aspect will be a continued promotion and implementation of standardisation efforts. For example, the SDMX format allows all data to be edited using the same tools, while data sharing also becomes technically easier. The provision and use of open source tools are also very useful. In order to further promote the supply of data internationally, issues of data confidentiality must remain on the political agenda. Emerging countries could be further supported in implementing measures to improve data sharing.

7. Recommendation 6 – Consideration of ways of improved sharing of granular data

Economies are encouraged to consider alternative forms of access to granular, potentially micro data, taking into consideration the confidentiality restrictions and users' needs.

7.1. Global developments

57. The BIS has implemented the concept of "restricted" data for banking statistics, i.e. these data sets are not publicly available but are accessible by all the reporting countries on a mutual basis.

58. As appropriate, the IMF advises authorities to allow NSOs and central banks to access anonymized tax and other data for statistical purposes.

7.2. Regional developments

59. Within the ESCB, confidential micro data are shared, to the extent necessary, for the performance of ESCB tasks across Eurosystem central banks and national competent authorities, and where envisaged by the relevant legislation with other organisations.

60. Eurostat and National Statistical Institutes established the Micro Data Access Network to benefit from the experiences made. However, at the same time it should be noted that granting externals access to confidential micro data for banks or companies is highly complex. Alternative ways, such as anonymization, confidentiality on the fly[20] or remote execution of code work are in practice but require very careful analysis to preserve the confidentiality of the related entities, especially taking into account that with publicly available data, the entities identity could be revealed.

7.3. National developments

61. A lot of efforts are invested to improve data sharing between national authorities, (could include the central bank, statistical office, bank/financial supervisory authority or the deposit insurer as well as other government agencies and universities). However,

[20] The "confidentiality on the fly" approach is based on the random noise protection method developed by the Australian Bureau of Statistics, initially for census data. The ABS method consists in applying small perturbations to the data with the predefined probability distribution for the noise to be applied. A specific mechanism randomly draws from that noise distribution in a consistent manner. See also: https://ec.europa.eu/eurostat/cros/content/item-43-confidentiality-fly_en.

significant legislative restrictions limit the institutions' ability to share collected granular data.

62. A growing number of institutions have set up new Data Labs or Research Data Centres.[21] These units mostly offer a kind of open data room for granular data where researchers can access data on the spot. This allows access to more sensitive data in a safe environment and the use of advanced analytics tools. Another example is the promotion of private use of public information by expanding the disclosure of information by public institutions. In this case, the competent authority strives to protect the personal information of respondents of statistical surveys by sharing raw data and non-public statistics through a dedicated microdata integration service and non-identifiable measures for the raw data.

7.4. Outlook

63. In some countries the work on how to improve sharing of granular data is still being considered. Granular data are used in some business areas and promoting further use is a clear goal for the future.

64. Some countries that have not set up units for local data access plan to do so in the near future, or enhance their units with equipment for remote access, for example.

65. Germany starts work on an Integrated Digital Information System (IDIS) project. It aims on developing and identifying the best text analysis and machine learning techniques to discover relationships between data sets, researchers, publications, research methods and fields. The final objective is to build a set of tools that enable collaborative knowledge creation.

66. The US BEA is soon to enter into an agreement to participate in the US Census Bureau's Federal Statistical Research Centers program, which will allow researchers access from 29 locations across the US.

8. Recommendation 7 – Collection of data only once

Economies are encouraged to avoid multiple collections of the same data by promoting flexible uses of existing datasets to minimize the reporting burden. Increased sharing of data among relevant institutions at the national level should be a key priority.

8.1. Global developments

67. International cooperation remains crucial achieving the objectives of this Recommendation. The use of SDMX as a global standard provides benefits in this regard.

[21] Please also refer to recommendation 2 in section 3.

68. For micro level household survey data the World Bank, as the IHSN secretariat, provides a common portal for accessing all manner of public use household surveys.[22] For aggregate macroeconomic indicators published by member states, they use standardized approaches (e.g. SDMX) as much as possible to reduce the response burden.

8.2. Regional developments

69. The European System of Central Banks (ESCB) has the medium-term objective to standardise and integrate, across domains and countries, the collection of data for banks laid down in ESCB statistical frameworks. The Integrated Reporting Framework (IReF[23]) represents a key element of this strategic approach and is currently under assessment. In particular, the ESCB, in close cooperation with the banking industry, has initiated a cost-benefit analysis to assess the impact of the IReF on the supply and demand sides. The other element of the strategic approach is the development, in collaboration with NCBs and the banking industry, of the Banks' Integrated Reporting Dictionary (BIRD[24]), which supports reporting agents in efficiently organising the information available in their internal systems and fulfilling their reporting requirements, for both statistical and supervisory purposes. The BIRD is freely available to banks and all interested parties (e.g. software houses that develop application packages for financial reporting).

70. Another step on the European level is the design of the new European regulation on pension funds statistics, which has been defined with a view to minimise the reporting burden on agents. The new pension funds regulation includes only additional statistical data requirements not covered by supervisory data (which could be used for statistical purposes).

8.3. National developments

71. National examples are the following:

- The Australian NSO has a number of longstanding arrangements with other agencies to include data for statistical purposes in their data collections. It is also focussing on alternative data sources including, but not limited to, administrative data in preference to conducting new surveys.
- The Canadian central bank states that the collection of data only once is largely applicable for banking data.
- The German central bank mentions that several statistics (e.g. statistics on insurance corporations based on Solvency II, AnaCredit, upcoming statistics on pension funds) already make use of the possibility of sharing data for statistical and supervisory purposes. Data is collected once and - according to national

[22] For details see: www.ihsn.org
[23] More information can be found via the following link: https://www.ecb.europa.eu/pub/pdf/other/ecb.escb_integrated_reporting_framework201804.en.pdf
[24] More information can be found via the following link: https://banks-integrated-reporting-dictionary.eu

cooperation agreements - disseminated to supervisors and statisticians, who then aggregate the granular data depending on the respective purpose.

- Consistency of concepts, classifications and other standards are required for a high level of system integration in statistical systems.
- Japan makes use of data from a kind of “security depository centre”, instead of collecting data from individual financial institutions, in order to avoid multiple collections of data.
- In Russia, data associated with financial activity and financial sector data are collected only once. Authorities collect data according to their sphere of responsibility and share it under bilateral agreements with others. Approach of collecting data for statistical purposes only once has been announced in national project “Digital Analytical Platform” governed by the Federal State Statistics Service with other authorities involved. With the objective of alleviating the reporting burden of banks, the Bank of Russia is carrying out an analysis of the indicators presented in the current reporting forms in order to reduce their number and avoid multiple collections of the same data.

8.4. Outlook

72. Respondents mention that following the completion of this initiative, further opportunities to streamline data collections will be assessed. This will go hand in hand with a revision of statistical surveys and a renewal of the technology used.

73. A targeted, step-by-step approach for data sharing seems warranted: (1) Work to identify a problem that can only be improved by granular microdata exchange. (2) Limit the exchange to data required to solve the problem. (3) With time expand data sharing for other purposes that require granular data from other sources.

74. This approach will support progress on other DGI-2 recommendations, where data sharing within and between countries will allow greater quality and consistency, particularly for the recording of financial and non-financial multinational enterprises. Notable examples of the recommendations that will benefit include:

II.4 Globally Systemically Important Financial Institutions
II.5 Shadow Banking
II.6 Derivatives
II.8 Sectoral accounts
II.9 Household distributional information
II.10 International Investment Position
II.12 Portfolio investment
II.14 Cross border exposure

9. Selected ongoing initiatives and possible steps forward

9.1. Selected initiatives

75. According to the FSB, as of end-2018, over 1,300,000 LEIs have been issued. The LEI system was initiated by the G20 with the objective to cover all parties of financial contracts. Since this goal has not yet been achieved, G20 countries should develop (further) legislation requiring the use of the LEI to allow for universal and mandatory identification of legal entities, under a global standard, in a global, public-good infrastructure.

76. The establishment of RDCs in statistical institutions is a promising step forward in the field of granular data sharing. Some are already fully operational; others are still in the start-up or planning phase.

77. Statistical issues in the context of globalisation will continue to be a topic of interest. Data sharing between CBs and NSOs regarding MNEs is and will be of particular interest. In order to facilitate data sharing, in particular between different institutions, a key task is to connect and integrate data from different sources. This implies a substantial need for standardisation and harmonisation of data. On the European level for instance, FRIBS will help to better capture multinational activities in the field of business statistics.

78. In the framework of INEXDA, several institutions are aiming at facilitating the use of granular data for analytical, research and comparative purposes. The current focus of INEXDA is gradually shifting towards harmonising metadata, disseminating INEXDA results, exchanging experiences about data access procedures, and output control. INEXDA is growing from 5 members in January 2017 to 11 members now. Further institutions have expressed their interest to join and other statistical institutions are welcome to join.

9.2. The way forward within the scope of the G20 DGI-2

79. The evaluation of the questionnaire shows that the seven recommendations have been very helpful in creating and/or supporting developments on the national, regional and global level. For the way forward, the respondents have set different priorities:

- A regular conference (every 2/3 years) in order to keep the momentum on data sharing issues was suggested.
- Another suggestion is to give priority to promote the mandatory use of unique and licence-free identifiers on a global scale. This can go hand in hand with the further development of the LEI for non-financial corporates.
- Furthermore, data access to global digitalised MNEs and the exchange of using new techniques such as web scraping etc. are additional points. Another main issue is the need to find the right balance between confidentiality of data and information

needs, taking into account the evolution of the economic, social and technological context.

- When contemplating the establishment of new RDCs, in particular remote data access could be the interconnection between institutions' own RDCs and other "micro data laboratories".
- International statistical recommendations should promote international exchange of granular data.

80. Numerous processes in the course of implementing the recommendations are still in progress. It is suggested that representatives from INEXDA and the UNECE Task Force on Exchange and Sharing of Economic Data will be invited to present the results of their work at the next DGI Global Conference.

Annex

G20 report on the seven recommendations on data sharing

Update on the Data Gaps Initiative and the Outcome of the Workshop on Data Sharing

Inter-Agency Group on Economic and Financial Statistics March 2017

This Report was prepared by the Inter-Agency Group on Economic and Financial Statistics (IAG)[1]

The Member Agencies of the IAG are:

Bank for International Settlements (BIS)
European Central Bank (ECB)
Eurostat
International Monetary Fund (IMF, Chair)
Organization for Economic Co-operation and Development (OECD)
United Nations (UN)
World Bank

The Financial Stability Board (FSB) Secretariat also participates in the IAG meetings.

[1] The IAG was established in 2008 to coordinate international statistical work following the financial crisis.

Contents **Page**

ACRONYMS

DGI	Data Gaps Initiative
DGI-2	Second Phase of the Data Gaps Initiative
FMCBG	Finance Ministers and Central Bank Governors
FSB	Financial Stability Board
G-20	The Group of Twenty
IAG	Inter-Agency Group on Economic and Financial Statistics
IMF	International Monetary Fund
INEXDA	The International Network for Exchanging Experiences on Statistical Handling of Granular Data
ISIN	International Securities Identification Number
LEI	Legal Entity Identifier
LEID Number	Legal Entity Identifier Number MNE Multinational Enterprises
OECD	Organization for Economic Co-operation and Development
SDMX	Statistical Data and Metadata eXchange

EXECUTIVE SUMMARY

In September 2016, the G-20 Leaders welcomed the First Progress Report of the second phase of the G-20 Data Gaps Initiative (DGI-2)[2] and supported the proposed action plans for the implementation of its twenty recommendations.

This report by the Inter-Agency Group on Economic and Financial Statistics (IAG) provides an update on progress in implementing the DGI-2 work program for 2017 and summarizes the outcomes of the thematic workshop on data sharing. The report is to inform the G-20 Finance Ministers and Central Bank Governors (FMCBG), at their March 2017 meeting.

The DGI-2 work program for 2017 encompasses four thematic workshops, bilateral meetings with participating economies, quarterly videoconferences of the IAG, and the annual Global Conference for senior-level representatives in June 2017. Throughout these events, coordination in the implementation of DGI-2 is ensured among IAG member agencies, FSB Secretariat, as well as participating economies.

The first thematic workshop, on Recommendation II.20 of the DGI-2 on data sharing, has already taken place and has set out seven recommendations to promote data sharing (Annex 1). It aimed at creating a platform where participating economies exchanged practical experiences on data sharing with focus on the main obstacles preventing the sharing of granular data and the possible approaches to overcome such obstacles.

The IAG member agencies support the seven recommendations and will encourage the participating economies to overcome the identified obstacles to data sharing. The overall progress in implementing the DGI-2 during 2017 will be monitored and reported to the G-20 FMCBG for their October 2017 meeting, through the Second Progress Report of the DGI-2.

[2] The First Progress Report of the DGI-2 is available at http://www.imf.org/external/np/g20/pdf/2016/090216.pdf. It includes twenty recommendations as well as the action plans for the implementation of these recommendations agreed by the participating economies and the international organizations in the first half of 2016. An overview of the DGI-2 recommendations is provided in Annex 2.

I. Progress in Implementing the DGI-2 Work Program for 2017

1. The four thematic workshops indicated above are as follows:

a) **Data sharing,** organized by the IMF and Eurostat, in cooperation with and hosted by the Deutsche Bundesbank in Frankfurt am Main, Germany, during January 31–February 1, 2017. (Recommendation II.20 on the promotion of data sharing).

b) **Data gaps on systemic risk in the insurance sector,** to be organized and hosted by the Financial Stability Board (FSB) and the International Association of Insurance Supervisors in Basel, Switzerland on March 14, 2017. (Recommendation II.4 on global systemically important financial institutions).

c) **Institutional sector accounts,** to be organized by the Organization for Economic Co-operation and Development (OECD) and the IMF, and hosted by the OECD in Paris, France, during April 10–12, 2017. (Recommendation II.8 on sectoral accounts).

d) **Financial Soundness Indicators,** to be organized and hosted by the IMF in Washington, D.C., U.S.A., during April 26–28, 2017. (Recommendation II.2 on financial soundness indicators).

2. In addition to the thematic workshops, other key events are included in the DGI-2 work program for 2017. These include regular quarterly videoconferences of the IAG and bilateral follow-up discussions of international organizations with participating economies, leading to the annual Global Conference[3] to be hosted by the IMF during June 14−15, 2017 in Washington, D.C.

II. The Thematic Workshop on Data Sharing

A. Background

3. In the increasingly globalized economic and financial environment where entities are becoming more interconnected within and across borders, sharing of detailed information is becoming more important for better informed decision-making. In September 2015, a new recommendation (Recommendation II.20 on promotion of data sharing) was included in the DGI-2 aiming at promoting the sharing and accessibility of economic and financial statistics.

4. Despite the increasing need for data sharing, the efforts to make available and share granular information face significant difficulties. The key barriers against availability include (i) existing legal frameworks and administrative arrangements that can impede data sharing

[3] The annual Global Conferences of the DGI are attended by senior-level representatives from all G-20 economies and from the non-G-20 FSB member economies.

at national and international levels, (ii) technical and financial challenges associated with de-identification of confidential data to protect statistical confidentiality, (iii) cultural constraints.

5. Aggregated data may mask key information on drivers and effects of changing patterns in an economy, which could be observed using more granular information. Given the central role of the financial sector during the 2007/08 global crisis, more granular data on financial institutions and markets are increasingly being requested by policy makers to help straddle the divide between micro and macro analysis. Sharing of granular data would also contribute to a comprehensive and accurate view of developments in markets that are deeply interconnected across different jurisdictions.

6. At the same time, granular data provide the opportunity to improve the quality of aggregate statistics. If shared, these data allow statistical compilers to identify and resolve, inconsistencies between data compiled in different institutions and jurisdictions. Against this background, and with a view to facilitate the implementation of Recommendation II.20 of the DGI-2, an informal group of representatives from G-20 economies, and relevant regional/international organizations ("the Group"), was created in July 2016 to share experiences on data sharing with a specific focus on granular data. The Group was chaired by the IMF and Eurostat, in cooperation with the Deutsche Bundesbank.

7. Furthermore, to facilitate the implementation of the relevant targets of the action plan for Recommendation II.20, a thematic workshop on data sharing ("the Workshop") was held during January 31–February 1, 2017 in Frankfurt am Main, Germany.[4] The Workshop benefited from the contributions of almost 90 participants, from G-20 economies, international organizations, non-G-20 FSB member economies, and other non-G-20 jurisdictions that have practical experience in data sharing.

B. Key Outcomes and Recommendations of the Workshop

8. The key outcomes of the Workshop included agreement on a common terminology on data sharing, the identification of main barriers preventing the sharing of disaggregated data and micro data (including cross-border disaggregated data), and discussion on possible approaches to overcome such barriers. The Workshop, which focused on economic and financial data, concluded with seven recommendations aiming to provide guidance to national and international authorities as well as to encourage increased accessibility and sharing of granular data.

Scope of the Workshop: The discussions covered data sharing at both national level (within and among relevant institutions within the respective economies) and international level

[4] Prior to the Workshop, the Group held a meeting at the IMF on November 16, 2016 in Washington, D.C., with the participation of the representatives of the G-20 economies and international organizations.

(between different jurisdictions as well as between national authorities and international organizations).[5]

Terminology—the following types of data were identified:

***Aggregated data*:** data aggregates that have a low likelihood of identification of individual reporting units, such as those found in traditional datasets (e.g., those covered by most of the DGI-2 recommendations).

***Disaggregated data*:** data below the level of aggregated data and with a higher likelihood of identifying individual reporting units than in the aggregated data.

***Micro data*:** data on individual reporting units or specific transactions/instruments, which in most cases allow the identification of individual entities and therefore considered confidential. In addition, publicly available data on individual reporting units are considered non-confidential although they can still be subject to data sharing limitations due to commercial property rights.

Granular data: disaggregated data and micro data.

9. The seven recommendations concluded by the Workshop, which are presented in detail in Annex 1, are as follows:[6]

Recommendation 1: *Promoting the use of common statistical identifiers*

Recommendation 2: *Promoting the exchange of experience on statistical work with granular data and improving transparency*

Recommendation 3: *Balancing confidentiality and users'*

needs ***Recommendation 4:*** *Linking different datasets*

Recommendation 5: *Provision of data at the international*

level

Recommendation 6: *Consideration of ways of improved data sharing of granular data*

[5] Issues related to data sharing and cooperation among the international organizations are being addressed within the scope of Recommendation II.19 of the G-20 DGI-2 and therefore are not covered by the work of the Group.

[6] The recommendations are based on the experiences of participating economies', which were presented at the workshop in Frankfurt am Main, during January 31–February 1, 2017 and are available at http://www.principalglobalindicators.org/?sk=E30FAADE-77D0-4F8E-953C-C48DD9D14735&sId=1433357451568.

Recommendation 7: *Collection of data only once*

C. Conclusions

10. There is recognition that shifting from traditional aggregated statistics to include the flexible use of granular data would contribute to an increased availability of data for policy use, even though it may require time. The IAG member agencies and participating economies will continue to work on identifying the obstacles to data sharing and explore possibilities of overcoming such obstacles based on the seven recommendations indicated above.

11. The Workshop agreed on the following main principles:

- National authorities should, to the extent possible, consider reviewing their data sharing frameworks to maximize the amount of information which can be shared on a "need to know" basis, including at the international level. In some cases, restrictions on data sharing may not be solely due to legal reasons but also to technical, financial and cultural constraints. To facilitate a more open and transparent approach to data sharing, national authorities should consider reviewing such non-legal restrictions, wherever possible. Ways to overcome barriers against data sharing need to be considered, as relevant, taking into consideration the differences in administrative, cultural and historical backgrounds across jurisdictions.

- National authorities and international organizations should aim at building and maintaining trust between all relevant parties, including data users and data producers by striking a balance between making more data available for users while maintaining confidentiality.

- National authorities should place emphasis on facilitating the sharing of data at the national level across and within relevant institutions, which would also contribute to further sharing of data at the international level.

- Definitions of data and the use of standard statistical methods and classifications should be harmonized, where appropriate, within and across jurisdictions and accompanied with high quality metadata. To this end, further standardization of data, particularly of those that are key for policy making and surveillance, and adherence to international frameworks would contribute to reducing the reporting burden. It would also significantly improve the quality and comparability of

available information and any aggregation process. Ways to promote the adoption of common identifiers should also be considered.

- Focal contact points on respective areas of work should be identified by national authorities and shared within and across jurisdictions to facilitate access to the right

  contacts. The DGI Contact Group Members[7] could serve as first contact points for questions on data sharing and accessibility.

- National authorities should work towards adoption and implementation of the agreed latest standards for data transmission (e.g., Statistical Data and Metadata eXchange (SDMX)).

III. WAY FORWARD ON DATA SHARING

12. This report is mainly to inform the G-20 FMCBG of the seven recommendations concluded at the Workshop to promote data sharing, as set out in Annex 1. The overall progress in implementing the DGI-2 will be monitored and reported to the G-20 FMCBG for their October 2017 meeting, through the Second Progress Report of the DGI-2.

[7] DGI Contact Group Members are senior-level officials identified by the G-20 national authorities to serve as main contacts for the IAG on the DGI. These officials are first contact points on the annual monitoring reports, attend the global conferences, and coordinate with the policy departments of their respective institutions

Annex I. Recommendations Concluded at the Workshop on Data Sharing

(1) Promoting the use of common statistical identifiers

Economies and international organizations, as appropriate, are encouraged to foster the use of common identifiers and make every effort to adopt the latest international conceptual frameworks.

Internationally agreed common identifiers (e.g., Legal Entity Identifier (LEI), Unique Transaction Identifier, Unique Product Identifier, International Securities Identification Number (ISIN)) would help aggregating, linking and managing data. Minimizing registration and maintenance costs for enterprises or offering these services for free as a public service could be a way to increase the use of common identifiers, in particular the LEI for entities and the ISIN for instruments. Authorities should also consider including the LEI in their data disseminations and data collections, mandating its use, as appropriate. In this context, economies and relevant international organisations, the Global LEI Foundation and the LEI Regulatory Oversight Committee should continue working together to further investigate all ways to promote wider use of the LEI, enabling a better coverage of the non-financial sector and linking to existing identifier systems that already have very wide coverage, such as the Legal Entity Identifier Number (LEID Number) of Eurostat, the BIC-Code of SWIFT and the ISIN.

Adoption of the latest regional and international conceptual frameworks (e.g., *System of National Accounts 2008*, the *Balance of Payments and International Investment Position Manual*, sixth edition, the *Government Finance Statistics Manual 2014, the 2015 Guidelines for reporting BIS International Banking Statistics*) is vital for the standardization, consistency, and comparability of datasets regionally and internationally.[8]

(2) Promoting the exchange of experience on statistical work with granular data and improving transparency

Economies are encouraged to have a more active dialogue and exchange of practical experiences, in particular on the accessibility of granular data, metadata, and on techniques for statistical analysis and data protection.

Transparency about available granular datasets and the respective rules for accessing the data are needed for a well-informed analysis and research. Exchanging practical experiences on statistical work with existing granular data would facilitate understanding the obstacles and identifying possible approaches for an improved use of granular data for policy, research, and statistical purposes without jeopardizing the respective confidentiality regimes. It would also

[8] For the exchange of statistical information, the SDMX initiative has developed standards which have proven to be successful in practice.

create a basis for deeper investigating possibilities to harmonize access procedures and metadata structures, to develop comparable structures for existing data, and to further foster efficiency of statistical work with granular data.

The International Network for Exchanging Experiences on Statistical Handling of Granular Data (INEXDA)[9]—a cooperation project launched by the Banca d'Italia, the Banco de Portugal, the Bank of England, the Banque de France, and the Deutsche Bundesbank provides a good example of international co-operation and for intensifying and enhancing ongoing work. The Conference of European Statisticians Task Force on the exchange of economic data, particularly on the activities of multinational enterprises (MNEs) as well as the BIS Irving Fischer Committee work on data sharing are other examples for exchanging experiences in the context of data sharing.

(3) Balancing confidentiality and users' needs

Economies are encouraged to revisit existing confidentiality rules, practices, and approaches in light of evolving users' needs and taking into account peer practices, seeking to maximize, as appropriate, the amount of information released while maintaining the confidentiality of data deserving protection.

Data compilers should strike a balance between maintaining relevance and trust taking into consideration the needs and perspectives of users (both for policy making and research) when assigning confidentiality restrictions to data. Efforts should be made to release as much information as possible and appropriate, considering that certain data may have already been made publicly available and that the need for confidentiality may diminish over time**.** For example, in some cases even data that are published on the web are considered by national authorities as confidential and, therefore not made publicly available through official statistics.

Broadening the use of existing passive confidentiality rules—i.e., applying confidentiality only when reporting entities ask for it—could also be considered. The concept of passive confidentiality—has worked well in external trade statistics over decades. In addition, when assigning confidentiality levels to datasets, the potential number of reporters could be considered rather than their actual number (e.g., in the context of minimum three reporters rule). Getting consent from the data reporting entities, publication of more and comparable aggregated data and sharing of qualitative data without quantitative information as well as improvements in metadata could be among other ways of improving data sharing.

[9] The network is open to other central banks, national statistical institutes, and international organizations.

(4) Linking different datasets

Economies are encouraged to allow for the fullest possible use of data (for policy making, research, and statistical purposes) including through linking of different non-public datasets (possibly from different institutions) ensuring that strong governance and confidentiality protection arrangements are in place.

Use of common identifiers to the extent possible (see Recommendation 1) would be useful in linking different datasets. Resulting datasets may be assigned with different confidentiality levels than the underlying datasets, and may require application of other forms of anonymization as they may include further details increasing the risk of identification of reporting entities. Redistribution of granular data aggregated at the international level (i.e., through a global aggregator), linking datasets from different sources across jurisdictions, may alleviate some of the confidentiality issues. In such cases, the global aggregator may need access to micro-data for aggregation purposes (e.g., to avoid double counting) not to be shared beyond the global aggregator.

(5) Provision of data at the international level

Economies are encouraged to have in place regional and international data exchange systems and appropriate legal frameworks to allow for sharing and exchange of granular data across borders.

Understanding the effects of globalization, the behaviour of financial and non-financial multinationals (including the potentially major impacts of reorganization of cross-border business structures), interconnectedness and interdependencies require the combination and exchange of disaggregated and, possibly micro data at the international level. Efficient and secure mechanisms for data sharing between relevant jurisdictions (including with the cooperation of the MNEs concerned) and with international organizations would ensure that such effects are recorded consistently and that an accurate global perspective could be obtained. Such mechanisms would also increase the overall quality of data by helping to reduce information asymmetries and allow for the correct and timely allocation of the activities of MNEs to national jurisdictions.

Where needed, overhaul of existing or passing new legal texts about data should address data sharing and accessibility also at international level, in line with what is suggested at national and potentially regional levels (see Recommendation 7). Standards for sharing of and/or access to international granular datasets could be developed and reviewed on a regular basis based on the best practices (see Recommendation 3). Authorities are encouraged to make use of existing regional tools to the extent possible.

(6) Consideration of ways of improved sharing of granular data

Economies are encouraged to consider alternative forms of access to granular, potentially micro data, taking into consideration the confidentiality restrictions and users' needs.

Where existing and well-justified restrictions do not allow for sharing of granular/micro data, alternative forms of access[10] (e.g., remote execution, working with the data in the producing agency, guest visits to a secure environment) or different forms of anonymization (formal, factual or absolute[11]) should be considered in the context of decision-making about the dissemination of existing or new datasets to data users (for research and policy purposes).

(7) Collection of data only once

Economies are encouraged to avoid multiple collections of the same data by promoting flexible uses of existing datasets to minimize the reporting burden. Increased sharing of data among relevant institutions at the national level should be a key priority.

Support should be given to initiatives that aim to minimize the burden of data collection at the national, and potentially regional level (e.g., SDMX, open data platform), and support well justified and flexible uses of data and metadata. To this end, national authorities are encouraged to facilitate data sharing at the national level (among the public institutions such as central banks, national statistical offices, supervisors). Harmonization or alignment of definitions, statistical methods, classification systems and development of common datasets would also contribute to the reduction of reporting burden. Inter-agency cooperation is essential to avoid operational issues that may be faced in common datasets.

Every overhaul of existing or introduction of new legislations (or legal texts to the extent possible) which may have implications for data collection (including for administrative uses) should address data sharing and accessibility at national and potentially regional levels to prevent duplicated information requests by different authorities. To this end, data compilers should be involved early on in—and should have the ability to influence—the process of reporting development and other areas of work that may be relevant for their data collection systems, including administrative data. National authorities could create common statistical infrastructures with access given, as appropriate, both to respondents and national, regional and international institutions. In particular, official statistics should make use of other data sources including administrative and commercial data access (where they can be re-used) to the extent appropriate. National authorities should be given facilitated access to such data sources.

[10] In fact, some of these access forms are typically used even when sharing of granular/micro data is allowed.

[11] One of the traditionally used techniques for anonymization is aggregation.

Annex II. Overview of DGI-2 Recommendations

II.1: Mandate of the DGI
Monitoring risks in the financial sector
II.2: Financial Soundness Indicators (FSI) **II.3:** FSI Concentration and Distribution Measures (CDM) **II.4:** Data for Global Systemically Important Financial Institutions (G-SIFIs) **II.5** Shadow Banking **II.6** Derivatives **II.7** Securities Statistics
Vulnerabilities, Interconnections, and Spillovers
II.8: Sectoral accounts **II.9:** Household Distributional Information **II.10:** International Investment Position (IIP) **II.11:** International Banking Statistics (IBS) **II.12:** Coordinated Portfolio Investment Survey (CPIS) **II.13:** Coordinated Direct Investment Survey (CDIS) **II.14:** Cross border exposures of non-bank corporations **II.15:** Government Finance Statistics (GFS) **II.16:** Public Sector Debt Statistics Database (PSDS) **II.17:** Residential Property Prices (RPPI) **II.18:** Commercial Property Prices (CPPI)
Communication of Official Statistics
II.19: International Data Cooperation and Communication **II.20:** Promotion of Data Sharing