International Conference on "Statistics for Sustainable Finance", co-organised with the Banque de France and the Deutsche Bundesbank

14-15 September 2021, Paris, France, hybrid format

How integrated reporting by banks may foster sustainable finance?[1]

Antonio Colangelo, European Central Bank (ECB),
and Jean-Marc Israël, formerly ECB

[1] This presentation was prepared for the conference. The views expressed are those of the authors and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.

How integrated reporting by banks may foster sustainable finance?[1]

Antonio Colangelo, Senior Team Lead, Analytical Credit and Master Data Division, ECB
Jean-Marc Israël, former Head of the Analytical Credit and Master Data Division, ECB

Abstract

Central bank statistics have steadily developed over the past decades, especially in the wake of the 2008 financial crisis. Their use has steeply intensified in response to the crisis and the need for an in-depth understanding of economic and financial developments and a close monitoring of the impact of (monetary, economic and macroprudential) policy measures. A major breakthrough in the statistical data collection of the European System of Central Banks (ESCB) has been the move towards granular reporting of securities in the context of Securities Holdings Statistics since 2012, (further enhanced in recent years) and of loans through the AnaCredit project completed in 2018. Through the multiple dimensions available in the datasets and their combination with other data sources (eg registries on entities and commercial data sources), granular data have proven very valuable for a wide range of analyses, including on assessing risks arising from climate change. Going forward, the ESCB is working on reengineering and integrating its statistical reporting with the establishment of the Integrated Reporting Framework and the Banks' Integrated Reporting Dictionary. These projects will further enhance the availability and usability of granular data to policy makers and markets at large, and could align to other ongoing initiatives like the development of taxonomies to classify sustainable economic activities. They will accompany and contribute to the digitalisation and standardisation of financial market activities while also supporting and monitoring in how far they develop into sustainable finance.

Keywords: sustainable finance, climate change, central bank statistics, banks reporting, granular data, AnaCredit, integrated reporting framework (IReF), banks' integrated reporting dictionary (BIRD)

JEL classification: G21, E50, Q56, C81

[1] The views expressed are solely those of the authors and do not necessarily reflect the opinion of the European Central Bank. The authors would also like to thank Catherine Ahsbahs, Davide Continanza, Violetta Damia, Corinne Devillers, Francis Gross, Paul Hiebert, Angelo Saponara, Florian Schuster and Frauke Skudelny for useful discussions.

Contents

1. Introduction

The 2008 financial crisis showed limits of traditional central bank statistics to monitor and analyse economic and financial developments in interaction with a highly concentrated and interlinked global financial system. In a world of rapid and deep interconnectedness across markets and actors, policy makers need to unveil the heterogeneity across different economic areas (countries or groups of countries), industry sectors, market segments. The G20 Data Gaps Initiative[2] also gave an impetus for central bank statistics to further steadily develop towards timelier and disaggregated statistical information that allows drilling down from aggregates to micro developments.

The need for more granular data became prominent, at individual institution or even an individual instrument level on, eg, single securities and loans exposures. Granular data on loans through credit registers – and recently through AnaCredit in the ESCB – and holdings of securities, duly enriched with information on the characteristics of the borrowers and the lenders, can be directly used for panel data analyses. Granular data also allow for dynamically building aggregates that can provide researchers with methodologically sound time series, also when developing new analyses or new macroeconomic or macroprudential models. They also allow assessing with accuracy tail factors and distribution measures in a given population under review. Micro-prudential supervision has only given further impetus to the need for more detailed information. More recently, granular data have become key to analyse the impact of climate change and, more specifically, to assess physical risks that might impact the value of financial assets (also in their role as collateral) as well as transition risks that would arise from mitigation policies and changes in public sentiment. The increasing relevance of granular data, together with the interconnectedness of financial markets, also highlights the benefits of data harmonisation and standardisation, for regulators and for the financial industry.

The paper explores how the availability of granular data determined a shift in the analytical toolbox of central banks, and how this can help going forward in the context of assessing the risks of climate change and the pace of implementing sustainable finance. Section 2 looks into the main features of granular data, with particular reference to Securities Holdings Statistics and the Analytical Credit dataset (AnaCredit). Their use for policy analysis is possible also thanks to an extended use of identifiers and through linkages to the ESCB Register of Institutions and Affiliates Data (RIAD). Section 3 puts the paper into perspective, describing the wide range of policy analyses run by central banks and other policy makers and by researches. The section also looks into more detail in the analytical needs arising from climate change, discussing current uses and existing challenges. Section 4 introduces the Integrated Reporting Framework (IReF) and the Banks' Integrated Reporting Dictionary (BIRD), two on-going ESCB initiatives aimed at increasing the efficiency of data collection from banks. While these initiatives aim at alleviating the burden on the banking system, they will improve the information set of policymakers in terms of content and usability of the data in particular to assess sustainable finance. The section also offers a longer-term perspective, providing some perspectives to the future of reporting in

[2] See the IMF website for detailed information on these initiatives.

the digital age, notably to classify sustainable economic activities. Section 5 concludes.

2 Recent developments in processing granular data

Traditionally, economic and financial statistics have focused on describing the economy through aggregated indicators, implicitly considering economic areas as homogeneous entities. Describing the complex behaviour of economic agents by means of simple or weighted averages led to a substantial loss of information and to ignore complex relationships between agents. Furthermore, defining and developing macroeconomic statistical indicators require long lead-time, and the resulting aggregates provide little flexibility to support a wide range of policies.

The financial crisis showed the limits of such paradigm and catalysed the development of granular datasets, which allow measuring economic phenomena at the level at which they occur. These better supports policy makers in analysing the heterogeneity of the economy and financial markets. The new paradigm enabled central bank statistics to 'move beyond the aggregates' towards the collection of granular data to serve diverse policy needs in a timely, flexible and cost-effective manner. Indeed, granular data allow high flexibility and timeliness in adjusting to new user demands. Also thanks to new performing technologies, it is now possible to extract, classify and aggregate data at any point in time — also building consistent time series over several years — to respond to new user requests, without collecting new information either via an ad-hoc survey or via a new (aggregate) statistical requirement. An additional value for analysis of the data is the possibility for analysts to drill down from aggregated to institution-level or contract-level data so as to better understand new developments.

The section provides a high-level overview of Securities Holdings Statistics, AnaCredit and RIAD, which currently represent the core granular statistical datasets of the ESCB.

2.1 Securities Holdings Statistics

Securities Holdings Statistics provides since 2013 security-by-security quarterly information on securities held in the euro area, broken down by instrument type, sector and residency area of the debtor and further additional classifications.

Two modules are currently covered in the dataset. The sector module provides information on the holdings of listed ISIN securities by investors resident in the euro area (eg households in Germany or banks in France) and holdings of euro area securities by investors resident outside the euro area and deposited with a euro area custodian. The group module covers individual holdings by the largest banking groups with head offices in the euro area.

The data refer to four instrument types: short- and long-term debt securities, listed shares and investment funds shares/units. The collected information focusses on the holder side (eg amount held of a particular ISIN), while the data are enriched during the compilation process with reference data from the ESCB reference database

on securities (ie the Centralised Securities Database[3]), which contains rich information on individual securities such as the type of security, the name and sector of the issuer, the maturity and issue dates, price, outstanding amount/market capitalisation. Using ISIN codes, securities issues statistics can also be derived.

Additional attributes and/or dimensions can be obtained by linking Securities Holdings Statistics with other databases using a common identifier (the ISIN code and the name of the issuer are the most common options).[4]

2.2 Credit registers and AnaCredit

As regards loans, data on credit and credit risk were traditionally collected in several countries to support a variety of needs, often as a means to help banks in assessing the creditworthiness of their clients or potential new borrowers. In other cases, micro-prudential supervisors intended to better assess credit exposures of the supervised banks. In most cases, the wealth of information collected on credit and credit risk were scarcely shared across national authorities, and even less so with other competent authorities.

In line with initiatives by the World Bank (2011) to encourage credit registers, AnaCredit marks a paradigm shift for central bank statistics. The initiative was launched in 2013 to construct a pan-European granular database on loans provided by euro area credit institutions to legal entities that would model credit intermediation on an instrument-by instrument, counterparty-by-counterparty and protection-by-protection basis (as well as the relationship among these three building blocks). Since 2019 it provides a large set of harmonised and comparable information to Eurosystem users, with an unprecedented level of detail (ie 88 attributes that were defined based on the needs of a variety of business areas). The standardised attributes allow defining a wide range of indicators across multiple dimensions, with the possibility to back-cast them, also supporting an assessment of modelling and their results, eg in the context of stress testing.

As discussed by Collazo and Watfe (2017), the variables and measures included in AnaCredit can be combined in multiple ways to measure economic phenomena at different levels in a flexible way and construct networks between agents when enriched with reference data on entities (see below Section 2.3 on RIAD). This helps, among others, to analyse concentration of risk at different levels, eg within banking and corporate groups, or according to sectors of economic activity. Distribution of individual measures can be derived across the whole population, as well as across strata or segments. For instance, the tail(s) of the distribution reveal important insights, eg about risky counterparties. These analyses were hardly possible beforehand at euro area level. Even where national credit registers were in place, significant methodological differences existed in terms of scope, underlying definitions, collection basis (eg loan by loan or borrower by borrower) and reporting thresholds. The successful implementation of AnaCredit also relies on the establishment of the BIRD dictionary (see also Sections 4.2 and 4.3), which paved the way for further harmonising its implementation across euro area countries.

[3] See ECB (2010).

[4] See also Fache Rousová and Rodriguez Caloca (2014) for a more detailed description of the dataset.

2.3 The role of a repository of legal entities

The value of granular data on exposures and risks relies on the possibility to analyse the heterogeneity of the economy. Whether the granular data are collected from reporting agents or are purchased from commercial data providers, the first crucial step is the effective identification of the entities and, in the case of legal entities, the entity's ownership structure, eg to identify the actual risk bearer (usually the headquarters).

Traditionally, the identification of counterparties in the euro area has relied on national heterogenous systems. At the international level, with the support of G20, the Financial Stability Board and many regulators around the world, an initiative has been launched in recent years to develop a global Legal Entity Identifier (LEI). The current coverage of the LEI is still suboptimal and relationships across entities are still scarce.[5] In general, the LEI aims at providing a snapshot of factual information. For instance, the relationships are based on the accounting group perimeter of the legal entities, and no information is available on other variables of interest for policy purposes like the sector of economic activity (i.e. the NACE code), the statistical sector of classification or other relationship data.

This role is normally fulfilled by repositories of legal entities, which require an institution complementing facts with additional informational that require judgement. In the euro area, RIAD has been established as a common repository of master data on entities. Whereas it initially helped identifying reporting agents under ECB statistical regulations, its coverage has extended to ca. 10 million entities throughout the EU that are counterparties to loans covered in AnaCredit. RIAD can provides for the unique identification of legal entities, linking the existing available national sources, the LEI where available and information collected from reporting agents. Further to information on characteristics of the entities such as the statistical institutional sector and the NACE code, RIAD also supports information on links with other institutions to facilitate a broad range of policy analysis, both at the aggregated level (e.g. foreign direct investment relationships for the balance of payments) and at the exposure level.

3. Strengthening the analytical frameworks of policy makers with granular data

This section describes how granular data provided, in particular but not only, by Securities Holdings Statistics, AnaCredit and RIAD have strengthened the analytical framework of traditional central banking policy areas.[6] The use of granular data to assess climate-related risks is then discussed, reviewing the opportunities that granular datasets opened to support these analytical needs and the existing challenges that would need to be tackled going forward.

[5] See also GLEIF (2021)

[6] See also Israel and Tissot (2021)

3.1 Monetary policy conduct and operation

The monetary analysis toolset has much enriched in recent years thanks to the use of granular data along two main dimensions: zooming-in from aggregated data for the banking sector to bank-specific information (eg risk exposures and individual balance sheet information) and including data on individual loans and securities held duly enriched with information on the characteristics of the borrower (eg probability of default and rating). The resulting large datasets are then used for panel data analyses with a wide range of applications, like studying demand and supply effects in credit markets[7], or identifying shocks to credit demand and credit crunches. Granular credit data are also key to assess the heterogeneity in segments of lenders and borrowers that policy makers should consider when defining policies[8] and identify potential vulnerabilities as bottlenecks in some sectors may spread over other sectors, also depending on specific features of different economies.

A key aspect is the analysis of credit market conditions for small and medium-sized enterprises (SMEs), also in response to monetary policy measures. It has long been uneasy to stratify the population of enterprises. Statisticians and economists have developed international standards to accurately define industry activities – eg the International Standard of Industry Classification or the General Industrial Classification of Economic Activities within the European Communities (NACE)[9]. Crossing activity and size fosters a meaningful assessment of aggregate developments and trends in credit to the economy. This is all the more relevant that SMEs have the strongest impact on employment. Indeed, an accurate stratification of businesses requires also other criteria, in particular the size of firms (in turnover, number of employees, or total balance sheet).

These various elements are important both for defining monetary policy, especially unconventional monetary policy measures (eg ECB Targeted Long Term Refinancing Operations), and for monitoring the transmission channels of standard, as well as non-standard monetary policy measures. This fosters the possibility to analyse the "credit channel" and "risk-taking channel" of monetary policy. This also allows fine-tuning the measures more accurately and in a timelier manner.

3.2 Macroprudential policies

Whereas micro prudential supervision focuses on the soundness of individual institutions, macroprudential policies address systemic risk, which is the risk that the entire financial system may collapse as a consequence of idiosyncratic events affecting individual institutions. The assessment of the interconnectedness among financial intermediaries and market concentration are thus at the very core of macroprudential policies. As discussed in Collazo and Watfe (2017), granular datasets prove essential in this context. Besides the high level of detail, which provides flexibility to define new indicators, granular datasets (duly complemented by reference information on counterparties – eg their ultimate parent) allow users to identify linkages between agents, which are key to assess risk concentrations and

[7] For instance, see Altavilla et al (2020).

[8] For instance, see Altavilla et al (2018).

[9] See the dedicated pages on the websites of the United Nations and of Eurostat.

propagations. In addition, granular data would enable to focus on specific parts of the distributions – eg the tails.

Granular data is also instrumental for defining structural and cyclical macroprudential policies, as well for monitoring their effects and fine-tune them when needed. For example, instrument-level data on loans and holdings of securities can be used to set-up structural or countercyclical capital buffers[10] or to define limits on loan-to-value ratios (LTVs).

More recently, macroprudential (and financial stability) analyses have increasingly looked at risks arising from climate change and the greening of the economy. The use of granular data in this context is discussed in detail in the next section.

3.3 Micro prudential supervision

Micro prudential supervision aims at assessing the stability of individual supervised institutions. These activities are performed based on data on individual banks but can benefit significantly from granular risk data on loans and holdings of securities. For instance, supervisors can assess the effectiveness and accuracy of internal ratings models of banks and analyse credit exposures vis-à-vis groups of connected clients (non-financial groups, supply-demand chains, country-risk, indirect financing of firms). At the juncture of on-site inspections in supervised banks, accurate credit and credit risk data facilitate the identification of strata in the credit portfolio of institutions for a more effective and efficient review. This is particularly the case for the identification of non-performing and forborne loans, or for assessing the fairness in the way institutions classify their loans in different classes.

Granular data may also reduce the burden on banks when running stress tests, as data are already available to supervisors and may facilitate the dialogue with the banks on salient features in the results of the tests.

3.4 Using granular data to assess climate-related risks

In recent years central banks around the world have intensified their analytical work aimed at assessing climate-related risks to financial stability and monitoring of brown vs green financing. For instance, the Network of central banks and supervisors for Greening the Financial System (NGFS) was launched in 2017 to prepare approaches that could operationalise an active development in financing towards sustainability.[11]

Climate change poses unique challenges to financial stability through two main channels: i) physical risks, relating to natural disasters such as flooding and wildfires as well as heat and water stress, that might impact the value of financial assets (eg corporate and household exposures of financial intermediaries in stressed areas); and ii) transition risks, arising from mitigation policies and changes in public sentiment.

Analyses aimed at assessing and monitoring these risks intrinsically need to capture the heterogeneity of the economy and therefore call for an intensive use of

[10] For instance, see Budnik et al (2019).

[11] See also Network for Greening the Financial System (2020).

granular data to assess the impact across groups of counterparties rather than modelling aggregated developments. As described in ECB (2021a), a "*fine resolution measurement is required to trace out heterogeneous and novel physical and transition risk impacts across geographies, sectors and firms.*" For instance, to capture the localised effects of physical risks, granular data is key to correctly identify the location of counterparties and/or collateral. Similarly, for analysing the impact of transition risks, granular data can be used to assess the nature of counterparties and even individual exposures on loans and securities, to the extent they relate to activities with different climate impact.

In general, assessing climate-related financial risks requires a three-layer approach: derive economic risk factors from climate risk drivers; linking climate-adjusted economic risk factors to exposures of banks and other financial intermediaries; and translating this input into an assessment of the related financial risks based on sound analytical models. As regards the first layer. geolocational datasets from public or commercial data providers are available to capture risks of physical damage associated with physical risks. Sectoral and institution-level data on carbon footprint can also be used to define scenarios on risks arising to the green-transition.

These new user needs were certainly not the strongest drivers when the ESCB granular reporting was first set-up. Through the multiple dimensions available in the granular datasets and their combination, it becomes possible to assess in great depth the status-quo and closely monitor the extent to which the funding evolves – eg from brown to green firms. However, as explained in ECB (2021b), data are currently far from being complete and their usability is often questionable not least due to limited comparability. New and more granular data collections will thus be needed for both physical and transition risk assessments. The EU regulatory framework for sustainable finance (including the EU Taxonomy[12] for sustainable activities), should contribute to filling in some of the existing data gaps.

4. On-going ESCB initiatives and a vision for the future

Mapping climate risk drivers to financial exposures is also quite challenging at current stage. Firstly, gaps exist in the current granular data collections. For instance, AnaCredit covers loans of credit institutions above € 25,000, leaving out many banks' counterparties and particularly smaller firms. In addition, AnaCredit does not cover other non-banks financial intermediaries, which do offer loans to the economy (e.g. factoring and leasing companies). Similarly, Securities Holdings Statistics focus on listed ISIN securities and do not foresee the direct collection of data on holdings of securities from securities and derivatives dealers and captives, resulting in potential information gaps.

The existing granular statistical datasets are also based on independent legal acts and in many countries are collected in silos. This is challenging for users, who often have to reconcile the available statistical information to effectively support their analyses. In fact, the dictionaries that are supporting the national statistical collections

[12] Regulation (EU) 2020/852 of the European Parliament and of the Council of 18 June 2020 on the establishment of a framework to facilitate sustainable investment, and amending Regulation (EU) 2019/2088 (OJ L 198, 22.6.2020, p. 13).

are harmonised, but not fully standardised, meaning that concrete financial market instruments that are mapped to the statistical definitions of 'loans' and 'securities' may not be completely overlapping across countries (and across reporting agents). The reconciliation of statistical and supervisory granular datasets may be even more challenging due to the use of different dictionaries.

In line with the approach that was followed for collecting data on insurance corporations in 2016 and for pension funds in 2019 thanks to a very close cooperation between the ECB and EIOPA and across national authorities in the EU, the ESCB has embarked into visionary data integration activities for banks. These initiatives are mainly aimed at reducing the burden of reporting agents while providing users with data of better quality, and are strictly connected to the feasibility study requested under Article 430c of the updated Capital Requirements Regulation[13].

The section provides a brief overview of the on-going activities and stresses the key role of a data dictionary going forward. Some considerations are also provided for the longer-term, in the light of the opportunities that digitalisation may create.

4.1 The Integrated Reporting Framework (IReF)

The IReF integrates banks' statistical reporting requirements into a unique framework that would be directly applicable to euro area banks, without any translation into national collection frameworks. In particular, the IReF targets primarily the datasets collected under the ECB regulations on MFI balance sheet item (BSI) and interest rate (MIR) statistics, Securities Holdings Statistics and AnaCredit. As explained in Bier et al. (2018), in order to effectively integrate the existing requirements, the IReF will encompass a set of requirements with different levels of granularity that will consolidate the existing reporting lines across countries. To achieve an effective multiuse of the data, it is envisaged to collect most of the information at a monthly frequency and according to earlier transmission timelines compared to the current timeliness in the national statistical collection frameworks. For instance, if loan by loan and security by security data would be directly used to compile aggregated statistics, the reporting of this information shall take place at monthly frequency and early enough to be able to release data on monetary aggregates. Processes will overall be speeded up, with more possibilities to develop early indicators.

Collecting data as part of the same framework ensures that they are integrated, consistent and standardised at the source. Several data gaps are also being looked at. For instance, while the shared AnaCredit dataset only covers loans to legal entities above the € 25,000 threshold, the current IReF baseline scenario foresees the granular collection of data on all loans to legal entities. Similarly, banks' exposures to extra-euro area counterparties include significant positions relating to non-ISIN securities, which are currently excluded from Securities Holdings Statistics. The proposed

[13] Regulation (EU) 2019/876 of the European Parliament and of the Council of 20 May 2019 amending Regulation (EU) No 575/2013 as regards the leverage ratio, the net stable funding ratio, requirements for own funds and eligible liabilities, counterparty credit risk, market risk, exposures to central counterparties, exposures to collective investment undertakings, large exposures, reporting and disclosure requirements (OJ L 150, 7.06.2019, p. 1).

scenario for the IReF is to collect this information at the granular level.[14] These possible extensions may prove useful in the area of sustainable finance.

While the IReF is currently focusing on the integration of the existing statistical reporting of banks (ie deposit-taking corporations, in the statistical terminology), going forward it could be further strengthened. In fact, the IReF is being developed as a scalable product that could easily be extended to cover the existing reporting of other financial intermediaries or set out new reporting requirements in an integrated and standardised way.[15] For instance, should the user needs justify it (assessed via the 'merits and costs' ESCB procedure), (money market and non-money market) investment funds or non-bank financial corporations engaged in factoring and leasing could report according to the draft IReF reporting scheme. Such reporting in a standard format would be facilitated by the possible re-use of pieces of software within banking groups or via software vendors. Moreover, based on the EU Taxonomy, the IReF scheme could be enriched to cover the classification of individual loans or securities recorded on-balance sheet to help assess whether the economic activities they finance can be considered sustainable. The labelling and scoring provided by institutional and/or market sources could possibly be extended to counterparty data.

The IReF implementation will provide new impetus on closing data gaps on master data on entities. The coverage of RIAD will further extend to 'all' legal entities that are borrowing from euro area credit institutions. The sharing of master data on legal entities with reporting agents (or at least the main attributes that are not confidential) will also be instrumental to the process, making the identification process of counterparties easier and thus enhancing the quality of reference information available in the system overall.

4.2 The Banks' Integrated Reporting Dictionary (BIRD)

The ESCB's medium-term approach to data collection from banks also foresees supporting reporting agents in optimising the organisation of the information stored in their internal systems (e.g. for accounting, risk management, securities or deposits) in a single redundancy-free 'input layer', which could then provide the basis for fulfilling the statistical requirements covered in the IReF, as well as prudential and resolution reporting obligations. The BIRD also defines transformation rules to be applied to banks' input data in order to transmit data to the authorities. Those transformation rules would be specified in a single formal language facilitating their implementation by banks, leading to reduced efforts and costs. Figure 1 shows how BIRD and IReF will affect the reporting of banks.

The BIRD is being developed in close collaboration by a group composed of members from the ECB, some euro area NCBs and the banking industry. Banks occupy a central role in the process, as they have specific knowledge of their operational systems and reporting systems, while the ECB and NCBs mainly work as catalysts of the initiative and ensure the management and the maintenance of the BIRD over time.

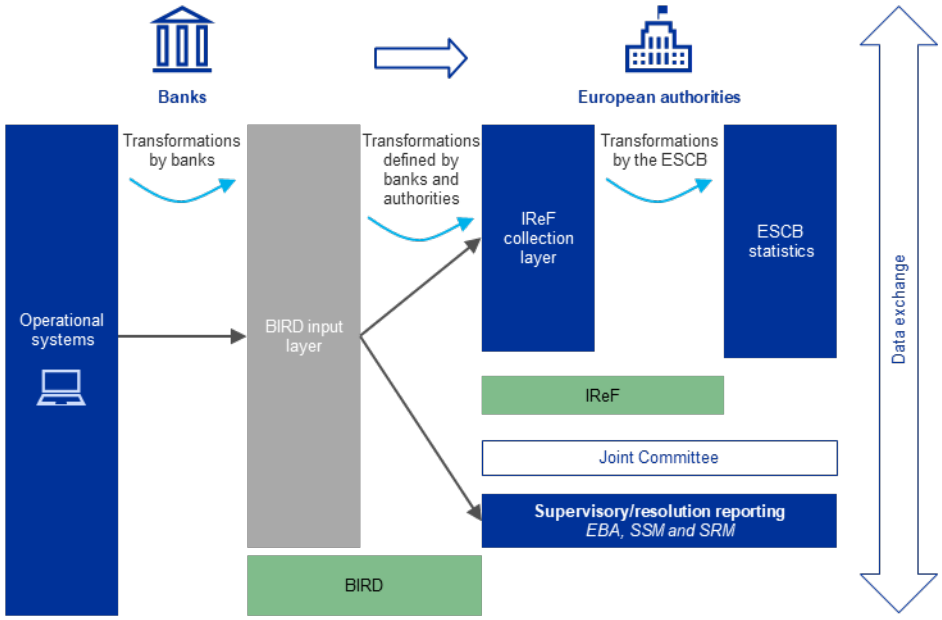
[14] These possible features of the IReF reporting were tested with the stakeholders in the cost-benefit assessment that took place from November 2020 to April 2021; see also ECB (2020c). The results of the questionnaire are currently being analyses.

[15] See also Sections 4.2 and 4.3 for some perspectives on the wider integration of statistical, prudential and resolution data.

Also due to its governance structure, the BIRD can be easily extended to incorporate other data needs of banks. Due to the granularity of input layer, for instance, it could be relatively straightforward to model additional variables of relevance for sustainable finance – eg for internal purposes, to fulfil disclosure requirements or submit them to authorities, as needed.

The BIRD and IReF are strictly related to each other. Without the IReF, the BIRD would need to undergo national adaptations to tune the input layer to the national collection systems of each country. In turn, the BIRD moves one step closer to the operational systems of banks, enabling banks to fulfil all reporting requirements in a consistent way. To the extent this approach will bear fruit, in the future authorities might be more inclined and willing to abandon aggregated (eg template) reporting in favour of a more structured and (close to) redundancy-free granular reporting. Aggregated reporting might stepwise diminish through an extension of the IReF. Only limited aggregated requirements may continue to apply as reporting agents should remain liable for certain values, eg on the side of prudential requirements, leading to an integration of statistical, prudential and resolution data.

Figure 1. Eurosystem strategy for collecting data from banks



Notes: The figure is from ECB (2020b). EBA stands for European Banking Authority, SSM for Single Supervisory Mechanism and SRM for Single Resolution Mechanism.

4.3 The role of a single data dictionary

The IReF and the BIRD will rely on a common data dictionary[16] that will ideally be developed as a joint endeavour by the ESCB, the European Banking Authority (EBA), the Single Resolution Board and the Directorate General for Financial Stability,

16 See ECB (2020b).

Financial Services and Capital Markets Union (DG FISMA) of the European Commission as a follow-up to the on-going EBA feasibility study.[17]

This common dictionary can be thought of as a repository of the definitions (and the corresponding codifications) of all the minimum distinct data items which are required to derive the existing statistical, prudential and resolution requirements. It will thus ensure the standardisation of the definitions, reducing the effort banks would otherwise make interpreting and reconciling instructions formulated in different frameworks. Even in the absence of the physical integration of statistical, prudential and resolution requirements depicted in the previous section for the longer-term, such a data dictionary would thus support the 'semantic' integration of the requirements. Most crucially, for users this will introduce much higher quality thanks to a precise and unambiguous definition of the information. For instance, with a common data dictionary the linkage between the definitions of the Large Exposure requirements with AnaCredit and Securities Holdings Statistics would be performed at the source.

While the BIRD is expected to remain voluntary for the banking industry (at least for the time being), the use of a common dictionary will represent an important incentive for adopting it.

4.4 The future of reporting in the digital age

The on-going acceleration in digitalisation activities, also in finance, creates unique challenges for statistical measurement. Machine-to-machine interactions are fast and automatic and lead to high data volumes on a global scale which cannot be controlled and measured by humans with a traditional approach – i.e. by reconciling source information to ensure harmonised output based on standards and manuals.

As elaborated in Colangelo et al (2021), while the initiatives mentioned above are moving in the right direction, only a complete change of paradigm can effectively support measurement activities of statisticians in the longer run. Statistics have to adapt to the world they measure, effectively using the traces that digital interactions leave. The economy could be seen as a network of contracts that connect a global population of parties. A data infrastructure that would hold globally standardised identification of all parties and contracts in real-time represents a key strategic objective going forward. As regards the identification of legal entities, the LEI system could represent a good basis but would need to obviously aim at a universal coverage. All contracts should then be represented in a standardised, mathematically rigorous algorithmic language. The development of standards for financial contracts is thus instrumental towards the setting-up of such infrastructure. While many initiatives are on-going in this direction, more global efforts, coordinated among the relevant authorities, need to be made.

As an example, a repository of master data on legal entities as indicated in Section 2.3 would be a derived product of such an infrastructure. All relationship data step from contracts, and those can be used to trace in (close to) real time linkages of all types among legal entities. The dictionary of BIRD and IReF would also have to

[17] The development work would be carried out by a Joint Committee to be established by the mentioned authorities; see also ECB (2020a). The Joint Committee is also depicted in Figure 1 as a point of union between statistical reporting on one side, and prudential and resolution reporting on the other.

converge to the new public-good standards of financial contracts, thus making possible automatic and live provision of data to users. Such granular structured data would truly enable analyses suited for a digital economy, eg based on machine learning and other artificial intelligence algorithms.

5. Conclusions

A major breakthrough in the statistical data collection has been the move towards granular reporting of securities (since 2012, further enhanced in recent years) and of loans through the AnaCredit project completed in 2018. A pan-European register of institutions (RIAD) also plays a key role, notably in matching lending and borrowing entities as well as, in the case of loans, protection providers. This new paradigm has provided central bank analysts and markets with a more accurate picture of banks' exposures and of businesses' indebtedness that enable them to better assess the heterogeneity of the economy. Through the multiple dimensions available and their combination, the datasets have driven recent research efforts in the area of climate change, notably through the assessment of new risks to financial stability as well as monitoring the credit flows for firms ranging from brown to green.

At the same time, the existing data availability is still suboptimal, and several challenges were identified in this paper. The on-going ESCB initiatives in the area of statistical reporting of banks will bear fruits for policy analysis in terms of data content, including a full standardisation of the dictionaries applied across the euro area and EU for statistical reporting and beyond, and in terms of connectivity with reference data on entities and other external datasets.

The overall approach will also facilitate policy work to assess and monitor sustainable finance in an effective and efficient way, while providing all users, analysts, researchers and economists with rich, time-consistent, multidimensional datasets.

References

Alestra, C., Cette, G., Chouard, V. and R. Lecat (2020): "Long-term growth impact of climate change and policies: the Advanced Climate Change Long-term (ACCL) scenario building model", Working Paper Series, No 759, Banque de France.

Altavilla, C., Boucinha, M., Holton, S. and S. Ongena (2018): "Credit supply and demand in unconventional times," Working Paper Series 2202, European Central Bank.

Altavilla, C., Laeven, L. and J.L. Peydró (2020): "Monetary and macroprudential policy complementarities: evidence from European credit registers," Working Paper Series 2504, European Central Bank.

Basel Committee on Banking Supervision (BCBS 2021a): "Climate-related risk drivers and their transmission channels," April.

Basel Committee on Banking Supervision (BCBS 2021b): "Climate-related financial risks – measurement methodologies," April.

Berg, F., Kölbel, J.F. and R. Rigobon (2019): "Aggregate confusion: The divergence of ESG ratings," MIT Sloan Working Paper, 5822-19.

Bier, W., Israel, J., Colangelo, A. and R. Bonci (2018): "Analytical credit dataset, the integrated reporting framework and the banks' integrated reporting dictionary: Do we overshoot? Or do we undershoot?," Journal of Securities Operations & Custody.

Budnik, K., Affinito, M., Barbic, G., E., Hadj, S.B., Chretien, Dewachter, H, Gonzales, C.I., Hu, J., Jantunen, L., Jomborean, R., Manninen, O., Martinho, R., Mencia, J., Mousarri, E., Naruševičius, L, Nicoletti, G., O'Grady, M., Ozsahin, S., Pereira, A.R., Rivera-Rozo, J., Trikoupis, C., Venditti, F. and S. Velasco, (2019): "The benefits and costs of adjusting bank capitalisation: evidence from euro area countries", Working Paper Series 2261, European Central Bank.

Carvalho, D. and Schmitz, M. (2021), "Shifts in the portfolio holdings of euro area investors in the midst of COVID-19: looking-through investment funds," ECB Working Paper Series, February

Colangelo, A., Gross, F. and F. Schuster (2021), "Effective measurement of the economy in the emerging digital age," Proceedings of the 63rd World Statistics Congress of the International Statistical Institute, Forthcoming.

Collazo Brananova, O., Watfe, G, (2017), "Use of AnaCredit granular data for macroprudential analysis," BIS-IFC

Deutsche Bundesbank (2021), "Securities holdings statistics," Frankfurt

European Central Bank (2020a): "The ESCB input into the EBA feasibility report under article 430c of the Capital Requirements Regulation (CRR 2)", September.

European Central Bank (2020b): "The Eurosystem Integrated Reporting Framework: an overview", November.

European Central Bank (2020c): "Cost-benefit assessment on the Integrated Reporting Framework", November.

European Central Bank (2021a), "Climate-related risk and financial stability," July.

European Central Bank (2021b), "Climate-related risk and financial stability, Data supplement," July.

European Central Bank, BIS, IMF (2015), "Handbook on Securities Statistics"

European Central Bank (2010), The "Centralised Securities Database" in brief

Fache Rousová, L. and Rodriguez Caloca, A. (2014), "The use of Securities Holdings Statistics (SHS) for designing new euro area financial integration indicators," IFC Bulletin No. 39

Global LEI Foundation (2021), https://www.gleif.org/fr/

Israel, J. and I. Lefebvre (2019), "Des agrégats à la donnée élémentaire," Revue Banque No. 830.

Israel, J. and B. Tissot (2021): "Incorporating micro data in macro policy perspective," IFC Bulletin No. 53.

Network for Greening the Financial System (2020), "Overview of environmental risk analysis by Financial Institutions"

How integrated reporting by banks may foster sustainable finance?*

International Conference on “Statistics for Sustainable Finance”

Paris, 14-15 Sept. 2021

Jean-Marc Israël
Former Head of the Analytical Credit & Master Data (AMA) Division, ECB

Antonio Colangelo
Senior Team Lead, AMA Division, ECB

* The views expressed are solely those of the authors and do not necessarily reflect the opinion of the European Central Bank nor of Banque de France

Content

1. Introduction

Traditional macroeconomic statistics…

- consider economic areas as homogeneous entities
- ignore complex relationships between agents
- have long lead-time and provide little flexibility

… while granular data

- allow to reflect the heterogeneity of economies and can be used for panel analyses
- measure economic phenomena at the level at which they occur
- allow high flexibility by users in handling data
- enable them to drill down from aggregated to institution-level or contract-level data so as to better understand new developments

"…well-established ESCB statistics will continue to provide the "big picture" of economic developments. But we should also offer a magnifying glass."

Mario Draghi, ECB President, 8th ECB Statistics Conference, 2016

2. Recent developments in processing granular data

Securities Holdings Statistics

- since 2013, security-by-security quarterly information on securities held in the euro area, broken down by instrument type, sector and residency area of the debtor, and
- further additional classifications

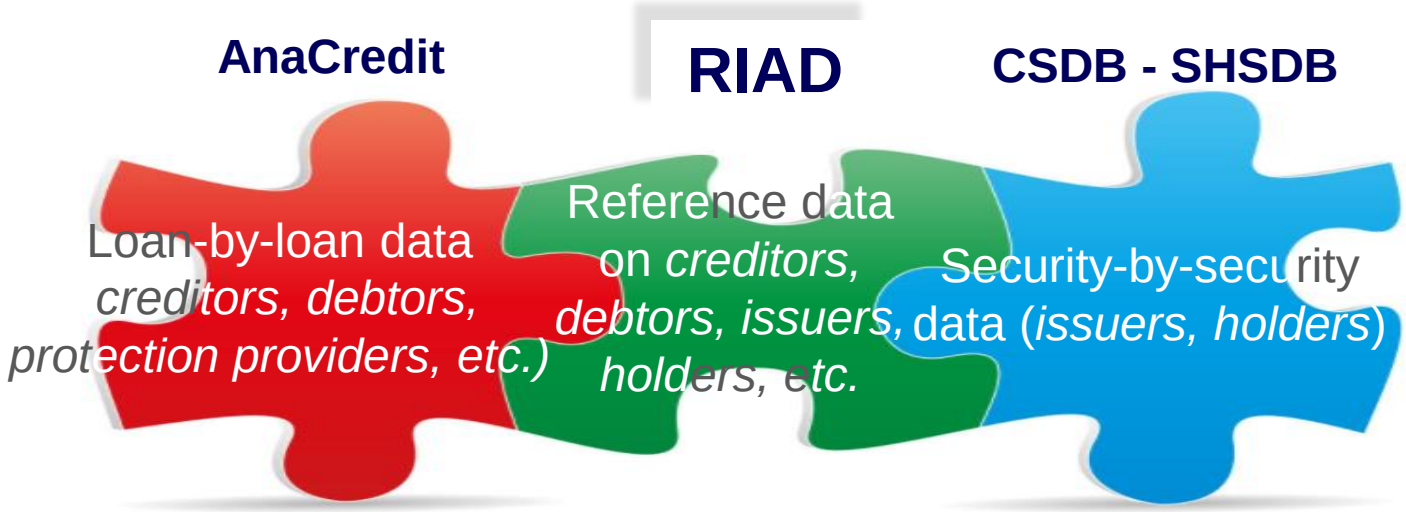
AnaCredit

- pan-European granular database on loans provided by euro area credit institutions to legal entities that models credit intermediation on an instrument-by instrument, counterparty-by-counterparty and protection-by-protection basis
  - ✓ as well as the relationship among these three building blocks
- since 2019, it provides a large set of harmonised and comparable information – ie 88 attributes defined based on the needs of a variety of business areas

2. Recent developments in processing granular data

The role of a repository of legal entities

- enrich granular data with information on borrowers and creditors as regards
  - ✓ their characteristics – eg sector of activity, residency
  - ✓ their relationships – eg the actual risk bearer
  - ✓ identifiers to link granular data with each other and other sources
- RIAD as the backbone of all ESCB granular statistics



3. Policy uses of granular data

Monetary policy conduct and operation

- study demand and supply effects in credit markets
- assess the heterogeneity in segments of lenders and borrowers -
  - ✓ eg credit market conditions for SMEs
- monitor the transmission channels of standard, as well as non-standard monetary policy measures, eg TLTROs

Macroprudential policies

- assess risk concentrations and propagations
- focus on specific parts of the distributions – eg the tails
- define structural and cyclical macroprudential policies – eg set-up structural or countercyclical capital buffers or to define limits on loan-to-value ratios

Micro prudential supervision

- analyse exposures to specific sectors or individual (groups of) counterparties
- assess the effectiveness and accuracy of internal ratings models of banks
- reduce the burden of stress tests, as data are available to supervisors

4. A closer look at sustainable finance

Climate-related physical and transition risks

- capturing the heterogeneity of the economy calls for an intensive use of granular data
- link risk factors from climate risk drivers, eg based on geolocational datasets or data on carbon footprint, to exposures of banks

Brown vs. green financing

- analyse in detail the nature of counterparties, the *credit allocation and their environmental impact*
- monitor developments *towards green financing*
- EU regulatory framework for sustainable finance, including the EU Taxonomy for sustainable activities

4. A closer look at sustainable finance

Challenges

- Ever rising need for more data, with greater detail and accuracy
- Some remaining data gaps
  - ✓ AnaCredit covers loans of credit institutions above € 25,000, leaving out many banks' counterparties, notably smaller firms, and loans from non-bank financial intermediaries
  - ✓ SHS covers listed ISIN securities and does not foresee the collection of data on holdings of securities from securities and derivatives dealers and captives
- Datasets are collected in national silos and not fully standardised dictionaries
- Still relatively limited connectivity of granular data with external datasets

5. Ongoing ESCB initiatives (1/2)

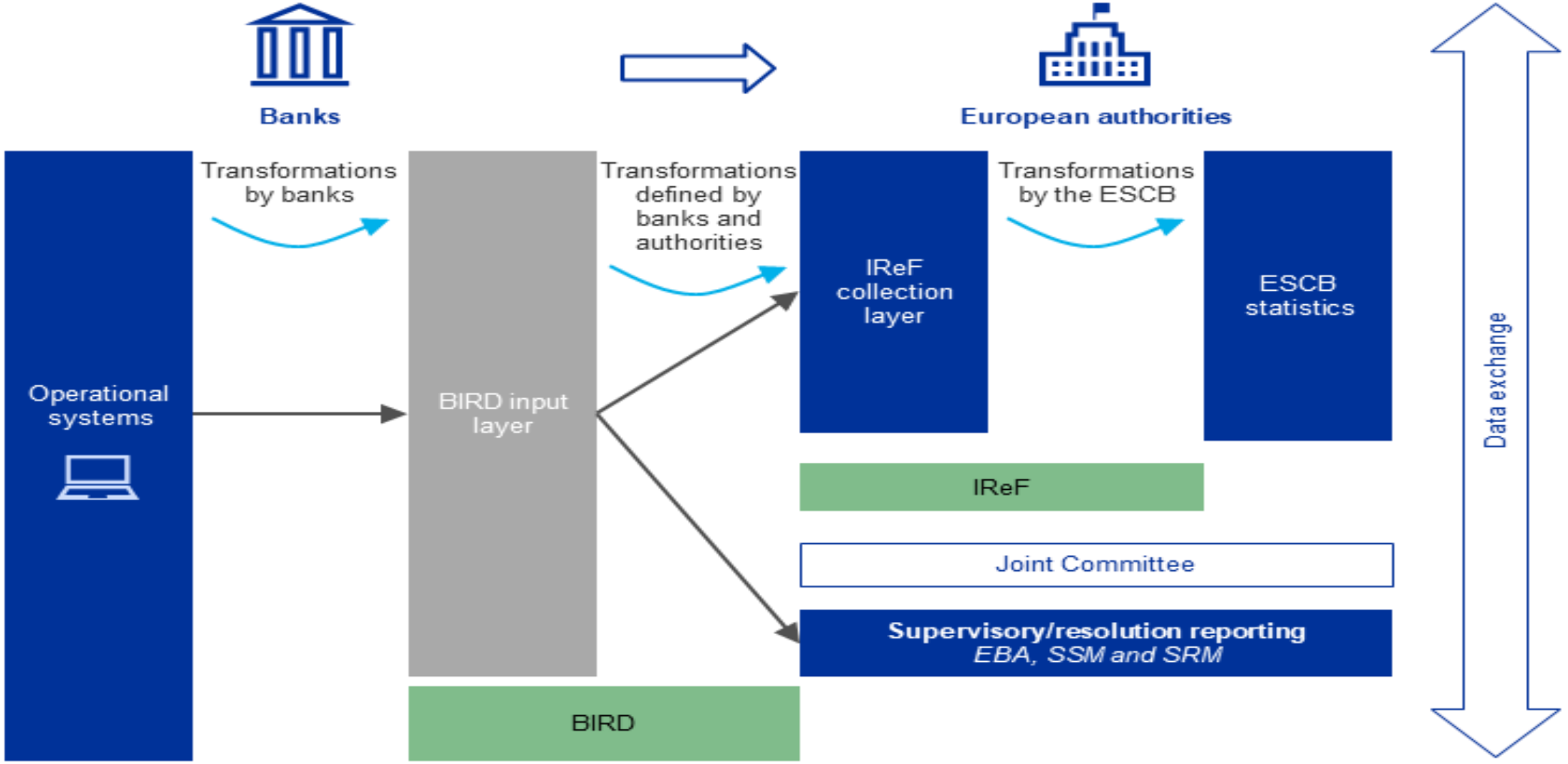
ESCB Integrated Reporting Framework (IReF)

- IReF for banks across countries and across (initially statistical) domains, with a focus on *ECB statistical requirements*
- gaps in granular data will be addressed, as well as existing gaps in master data on entities, eg extension of entities covered in RIAD to all banks' counterparties
- reporting based on a standard data dictionary; potentially shared with prudential and resolution reporting - Article 430(c) of CRR II
- scalable project that can extend granular reporting to other financial intermediaries

Banks' Integrated Reporting Dictionary (BIRD)

- organise data in banks' internal systems in a single redundancy-free 'input layer' as basis for fulfilling the statistical, prudential and resolution reporting obligations
- BIRD moves one step closer to the operational systems of banks
- in the future authorities might be reduce aggregated, eg template-based, reporting towards a more structured (and closer to redundancy-free) granular reporting

ECB broader strategy for statistics: envisaged approach



6. Conclusions

Granular reporting and data standards as a breakthrough

Datasets using ESG criteria to support green financing

Statistics to support researchers at central banks (*and beyond*) by offering services tailored to their needs

- Continuous **dialogue** with *analysts* and *researchers* to define appropriate *data marts* for pre-defined queries; also **banks** to benefit from *feedback loops*
- Ensure **methodological support** – eg consolidate exposures or debt – *correct usage of the data* and interpretation of the results

Thank you!

Questions?

Word-cloud from the AnaCredit Regulation