



EUROPEAN CENTRAL BANK

EUROSYSTEM

## Item 3.2

# Data quality assurance of granular reporting

---

## Remigio Echeverría



4 November 2020

Dialogue with the banking industry on ESCB statistics

# Overview

- 1 Technical Workshops with the banking industry and NCBs
- 2 Main principles for a harmonised data quality assessment
- 3 Common approach for data quality assurance of granular reporting

# 1. Technical Workshops with the banking industry and NCBs

## Structure and Content

### 1) Overview of respective data collections



Main features of SHSG and AnaCredit data collection and presentation of the current data quality validation process

### 2) Reporting issues & review of conceptual aspects



Discussion & clarifications on main data & reporting issues encountered during the first year of reporting

### 3) Common approach for an integrated quality assurance of granular reporting

Discussion on ..



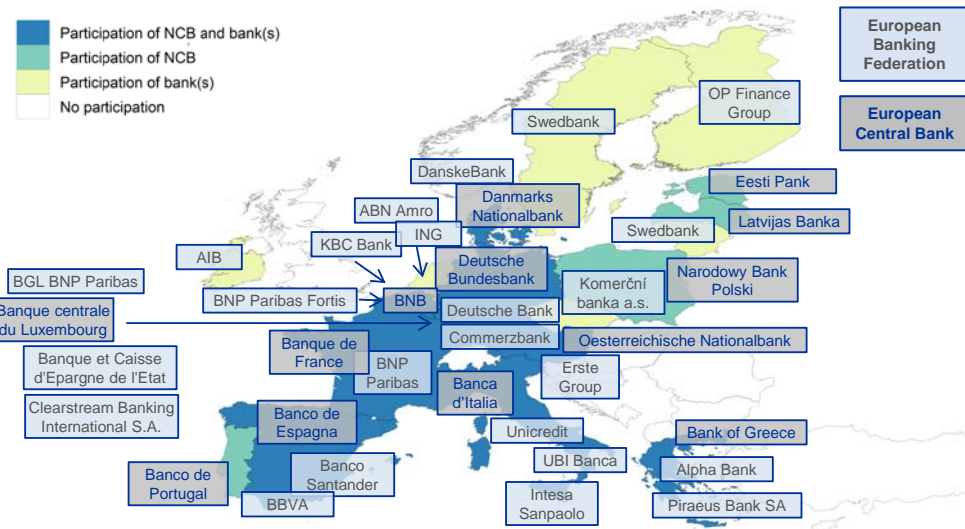
- basic concepts and principles that can be applied across granular data collections
- concrete sets of data quality summary metrics and a corresponding prioritisation of data quality work

#### ➤ **Main Outcome:**

Strong support for an integrated and harmonised data quality assessment framework across granular collections

## Wide participation among banks & NCBs

- 11 Nov 2019: Workshop with focus on SHSG (23 external participants)
- 24 Jan 2020: Workshop with focus on AnaCredit (32 external participants)
- Discussions on common approach for granular data at both workshops



## 2. Main principles for a harmonised data quality assessment



### Common understanding and perspective of data quality

Encompassing view on data quality relevant to all parties involved (reporters, producers, users)

Focus on the reported dataset as a whole instead of record-by-record

Aim for harmonised approach across different granular data collections



### Limited set of clear and simple data quality metrics (DQis)

Set of key indicators to allow easy and fast overview on the state of data quality/validation

Clear link between metrics and required data validation work as well as its prioritisation

Take into account the relative impact of individual findings and potential data issues

### Integrated and harmonised data quality assessment framework



### Effectiveness

Ensure reported data is ...

- in line with defined requirements
- reported in time to allow for timely processing
- consistent, both internally (i.e. within a single reporting and across time) and externally (across datasets)

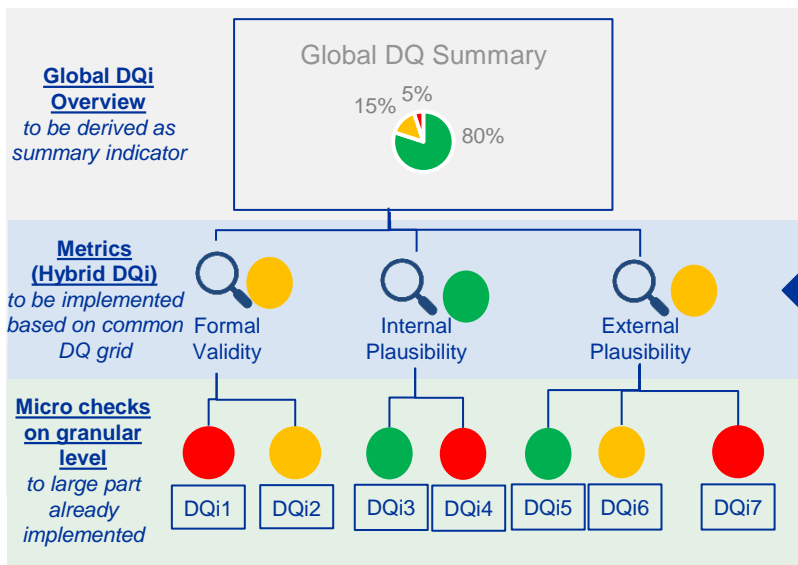


### Efficiency

Clear definition and communication of data checks, validation metrics and prioritisation of data quality work ...

- allows anticipation of a large part of required data validation efforts on reporter side
- minimises need for ad-hoc queries
- supports flexibility to further develop metrics based on gained experience and feedback

### 3. Common approach for data quality assurance of granular reporting



**Common DQ grid for the integrated quality assessment of granular reporting**

		Structure	Stability
Formal Validity	per RA*	Provision of formally compliant data input in line with the reporting deadlines	-
Internal Plausibility	per RA	Consistency across attributes	Time consistency of aggregate metrics
	across RAs	Consistency with data reported by other RAs	Changes in relative position compared to other RAs
External Plausibility	Benchmark comparisons	Consistency with statistics and supervisory reporting	Consistency of ratios over time

\* Reporting agent (reporting banking group for SHSG and credit institution for AnaCredit)

#### Way forward

- Stepwise implementation of new metrics and data quality procedures to better prioritise DQ work
- Enhance information flows between ECB/NCBs and reporting agents with respect to data quality issues
- Continue discussion in technical workshops to obtain feedback from reporters and establish a fruitful collaboration

**Thank you very much!**

Questions are welcome