



# 2023 Report on Quality and Use of Data

Webinar 26 April 2024

[ESMA12-1209242288-852 2023 Report on Quality and Use of Data](#)

# Data reporting regimes – Factsheet 2023

## EMIR

- 14.7 Bn life cycle events
- 30.6 Mn average open positions
- 180k reporting counterparties

## SFTR

- 620 Mn life cycle events
- 3 Mn average number of open SFTs in 2023
- 22k reporting counterparties

## MiFIR transactions

- 6.9 Bn transactions reported
- 20 Mn transactions reported daily
- 6.7k executing firms

## MiFIR transparency

- 1.9 Bn transactions reported<sup>1</sup>
- 219 Mn transactions published by APAs<sup>2</sup>

## Securitisation

- 515 outstanding securitizations
- 676 Bn EUR aggregated CPB
- 301 STS public deals

## Funds

- 461 MMFs and 125 managers
- 35,689 AIFs and 5,532 AIFMs

# 2023 Report on Quality and Use of Data

## Table of contents

- 1 **Executive summary**
- 2 **Data quality and use in the context of the ESMA Data Strategy**
- 3 **Key projects with significant data use**
  - 2.2 *NCAs, ECB and ESRB*
  - 2.3 *ESMA*
- 4 **Key data quality developments**
  - 4.1 *EMIR data and SFTR data*
  - 4.2 *MiFIR data*
  - 4.3 *Securitisation data*
  - 4.3 *Funds data*
- 5 **Conclusions and next steps**
- 6 **Annex**
  - 6.1 *Data Quality indicators - methodology*
  - 6.2 *Automation of APA post-trade data download*



# Data quality and use in the context of ESMA data strategy



## Mobilising ESMA data assets to serve its strategic and thematic objectives



15 June 2023  
ESMA50-157-3404

### ESMA Data Strategy 2023-2028

*Enabled by  
the ESMA  
Data Platform*

#### Objective A Enhanced data hub

- Share data with national competent authorities by on-boarding them to the ESMA data platform

#### Objective B Access to data of public interest

- Publication of statistics, working towards interactive dashboards
- European Single Access Point (ESAP) will significantly increase scope of information centrally available at ESMA

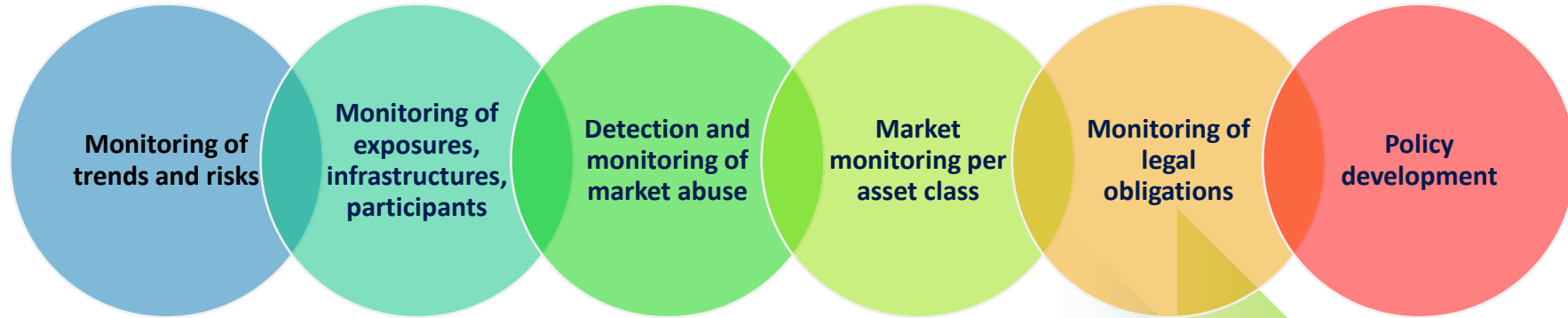
#### Objective C Data-driven Supervision

- Making available to supervisors the data, information and tools enabling data-driven supervision (e.g. develop a proof-of-concept for the detection of market abuse using AI techniques)

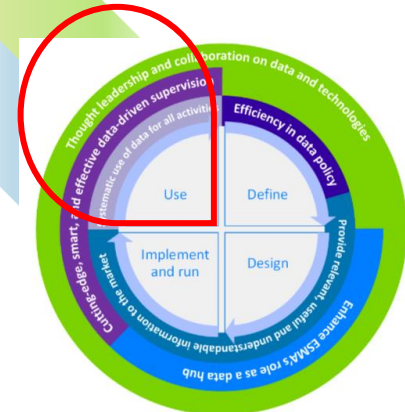
#### Objective F Systematic data use

- Migration of all datasets and analysis to the ESMA Data Platform
- Allows fast, efficient and collaborative data use by ESMA users (e.g. combination of sources)

# Use of financial data by EU regulators



- **Data plays a critical role** in the identification and monitoring of risks to the integrity, orderly functioning and stability of financial markets.
- **Data directly supports** ESMA, NCA, ECB and ESRB decision making in the areas of **policy making, supervision and monitoring of financial markets risk analysis.**
- The report provides a broad overview of use-cases implemented by data users. Where possible, links to publicly available the report included.

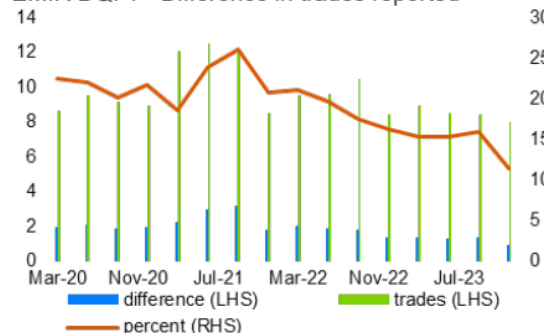


# Key development impacting data quality : EMIR

## EMIR DQIs and dissemination of the results to NCAs

Chart 6

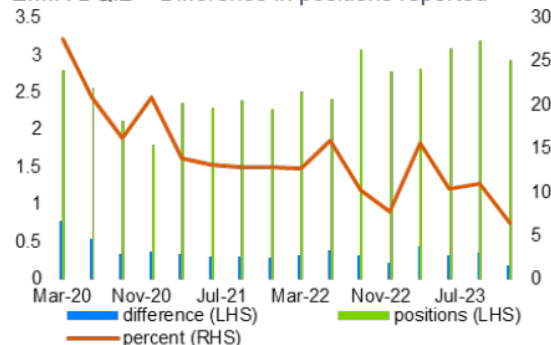
EMIR DQI 1 - Difference in trades reported



Note: EMIR DQI 1 - LHS (Millions of records): Total outstanding pairable trades and the sum of absolute differences in reporting between two parties. RHS (%): Difference as a percent of outstanding trades. Sources: EMIR Trade State Report and ESMA calculations

Chart 7

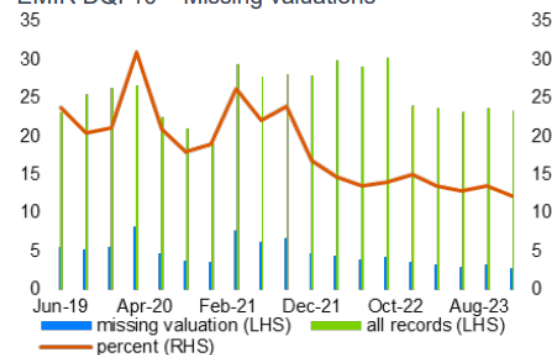
EMIR DQI2 – Difference in positions reported



Note: EMIR DQI 2 - LHS (Millions of records): Total outstanding pairable positions and the sum of absolute differences in reporting between two parties. RHS (%): Difference as a percent of outstanding positions. Sources: EMIR Trade State Report and ESMA calculations

Chart 10

EMIR DQI 10 – Missing valuations

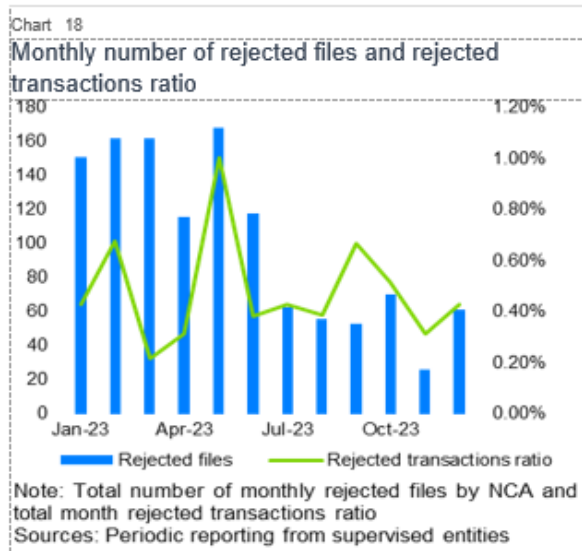


Note: EMIR DQI 10 - LHS (Millions of records): Missing valuation vs all records. RHS (%): Percent of missing records valuation out of all records. Sources: EMIR Trade State Report and ESMA calculations

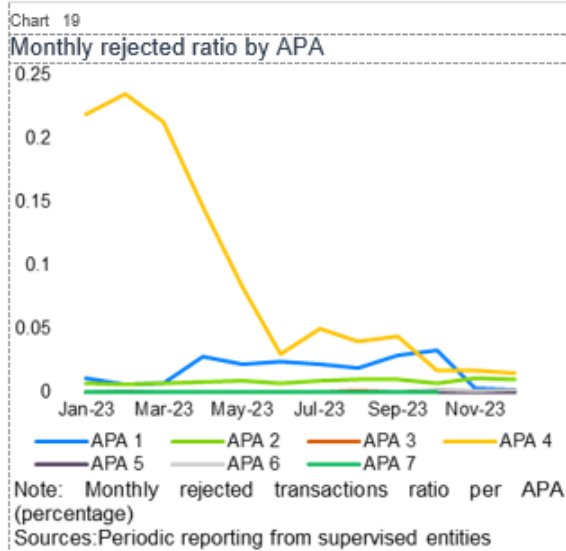
- Downward trend in the difference in the number of trades (DQ1) and positions (DQ 2) reported by two counterparties trading with each other
- After significant drop until 2021, trades reported with missing valuation continue to decrease albeit at a lower pace (DQ 10)

# Key development impacting data quality : MiFIR Supervision of ARMs and APAs

## ARMs – Transaction data



## APAs – Transparency data



- Periodic information from supervised APAs and ARMs to monitor data quality (e.g. ingestion volumes, rejection rates)
- Low rejection ratios from both APAs and ARMs (-46% for ARMs)
- One APA successfully addressed reporting issues over the course of 2023.
- Significant improvements in the average time for executing entities to submit transactions to APAs compared to 2022.

# Key development impacting data quality

## MiFIR transparency data

### Manual on post-trade transparency

- New level 3 tool with concrete examples
- Practical guidance to enhance data quality

### Transparency publications

- Periodic publications of aggregated figures for the application of MiFIR transparency rules

### Data Quality Engagement Framework

- Updated in 2023, quarterly checks shared with NCAs and DRSPs
- Risk-based approach to address most impactful issues

### Data quality updates

- Completeness: stable, good
- Accuracy : better for equity than non-equity instruments
- Issues under resolution: classification of bonds, fractional shares, OTF reporting practice

## Securitisation

Chart 27

Completeness of data, based on the usage of ND options

	ND1 = 0%	0% < ND1 <= 10%	10% < ND1 <= 30%	ND1 > 30%
ND2-4 = 0%	47.5%	10.9%	1.6%	0.2%
0% < ND2-4 <= 20%	19.7%	15.3%	1.6%	0.0%
20% < ND2-4 <= 40%	1.4%	0.8%	0.0%	0.0%
ND2-4 > 40%	1.0%	0.0%	0.0%	0.0%

Note: ND1: Percentage of fields entered as 'ND1', ND2-4: Percentage of fields entered as 'ND2', 'ND3' or 'ND4'.

ND1: where the required information has not been collected because it was not required by the lending or underwriting criteria at the time of origination of the underlying exposure; ND2: where the required information has been collected at the time of origination of the underlying exposure but is not loaded into the reporting system of the reporting entity at the data cut-off date; ND3: where the required information has been collected at the time of origination of the underlying exposure but is loaded into a separate system from the reporting system of the reporting entity at the data cut-off date; ND4: where the required information has been collected but it will only be possible to make it available at a date taking place after the data cut-off date

Sources: Securitisation repositories, ESMA

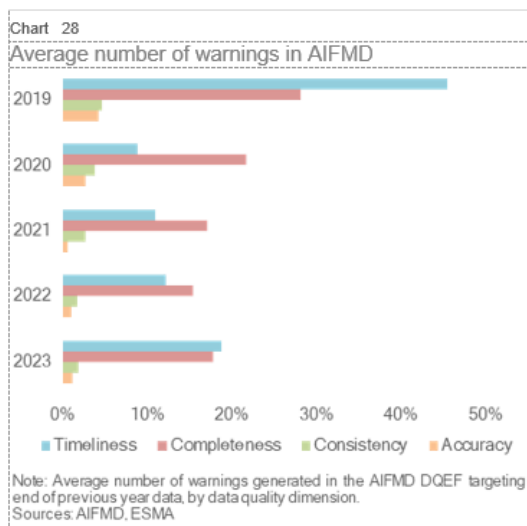
- Daily monitoring of simple, transparent and standardised (STS) register
- Monitoring over-reliance on “No Data” (ND) Option
- Consultation on possible changes to the securitisation disclosure template



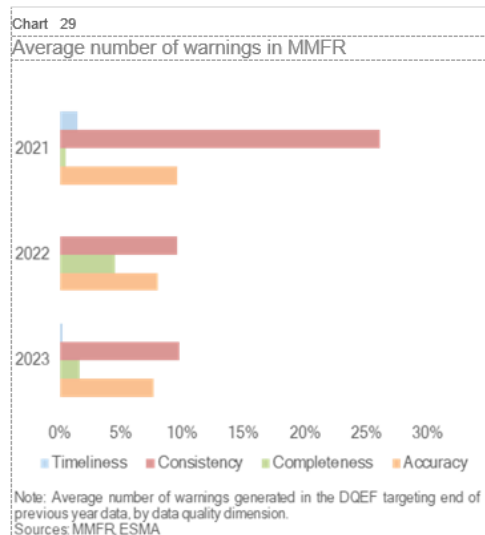
# Key development impacting data quality : Funds

## AIFMD and MMFR DQIs and dissemination of the results to NCAs

### Alternative investment funds



### Money market funds



- Long term positive impact on quality since introduction of Data Quality Engagement Frameworks
- Data quality did not improve in 2023 compared to 2022 and information on leverage (for AIFMD) needs further improvement
- Actions undertaken: new validation rules, guidance on reporting fields.
- Upcoming: risk-based approach, increase frequency of data quality cycles.

# Providing more transparency on data quality actions

## Extract of data quality indicators – EMIR

DQI	Description	Dataset
1. Difference in number of outstanding trades	A difference in the number of outstanding derivatives at trade level between a given pair of counterparties as reported by the two sides. Trades where other counterparty is non-EEA or is identified with a client code are excluded.	TSR
2. Difference number of outstanding positions	A difference in the number of outstanding derivatives at position level between a given pair of counterparties as reported by the two sides. Trades where other counterparty is non-EEA or is identified with a client code are excluded	TSR
3. Difference in number of reports with AT=N	A difference in the number of reports with action type 'New' submitted during the previous month by the two sides. Trades where other counterparty is non-EEA or is identified with a client code are excluded	TAR

### ✓ Methodology of the data quality indicators

- EMIR
- MiFIR Transparency (FITRS)
- AIFMD
- MMFR

### ✓ Indicators calculated for each jurisdiction and distributed to national competent authorities

## Extract of codes to download APA data

EU APA Name	Sample code
Bloomberg Data Reporting Services	<pre>import requests url = "https://www.bloombergapae.com/download?key=BAPE-POST2-&lt;YYYYMMDD-hh:mm:ss&gt;.csv" response = requests.request("GET",url) print(response.text)</pre>
CBOE Europe	<pre>import requests url = "https://markets.cboe.com/europe/equities/trade_data/apa/hour/rts13_public_trade_data_apa_&lt;YYYY-MM-DD_HH&gt;.csv" response = requests.request("GET",url) print(response.text)</pre>
MarketAxess Post-Trade	<pre>import requests url = "https://&lt;PERSONALTOKEN&gt;.cloudfront.net/TRADES/TRNL_APA/&lt;YYYY-MM-DD&gt;.csv" response = requests.request("GET",url) print(response.text)</pre>

### ✓ Automation of APA post-trade transparency data download

- Public data on transactions to be published free of charge after 15-min under MiFIR
- Codes provided in Python
- Each APA (Bloomberg, CBoE, MarketAxess, Nasdaq, Tradeweb and Unavista) has its own method of providing data

## Next steps on data quality

- ESMA and the NCAs will continue to work on the data quality framework in 2024:
  - Increase frequency of data quality cycles (e.g. for funds)
  - Facilitate sharing of data quality results with NCAs
  - Systematically adopt a risk-based approach for all datasets
  - Adaptation of the data quality framework in the context of the EMIR Refit go-live (29 April)
  - Provide input related to data quality in the context of the MiFIR review
- Continued cooperation with NCAs and stakeholders such as Repositories, DRSPs and reporting entities is crucial to achieving further improvements to data quality under the relevant reporting regimes.

# Q&A