

FINsights:

CTP Whitepaper Series

#5: Data Quality

A taxonomy of the key issues and some suggestions

14 December 2022

Executive Summary

In this paper we try to create a singular view of the issues relating to ‘data quality’ based on analysis of **135,000,000 transaction records across all asset classes** – bonds, equities and derivatives. We have previously highlighted key issues relating to bondsⁱ in a [series of whitepapers](#), as well as working with the Dutch Authority for the Financial Markets’ (AFM) regulatory sandbox and contributing to [the Principles](#)ⁱⁱ they published and the [evidence-based policy formulation](#) with AFMⁱⁱⁱ.

1. Transaction data issues fall into two categories:

- **Consistency** - different venues seem to have either adopted, permit or accept certain conventions or automated processing of MiFID transaction data and we also found that **application of flags**^{iv} also lacks conformity with the current rules.
- **Coherence** - when data is examined at a granular level, some of the ‘golden’ data (**currency, price and quantity**) is being reported in a manner that prevents effective aggregation leading to incoherent results.

2. Reference data issues fall into two different categories:

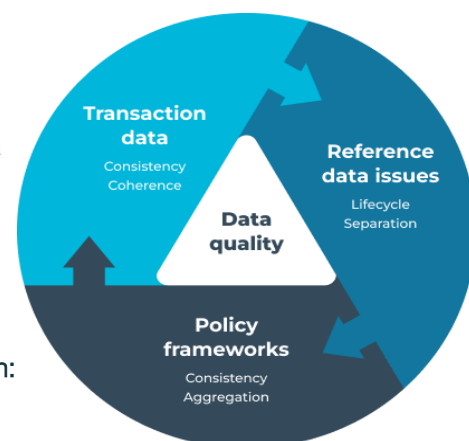
- **The lack of lifecycle data** means that participants get a ‘snapshot’ of reference data at a point in time rather than the full time series of data, which is critical for cross- referencing.
- The separation of EU and UK FIRDS requires parallel ingestion of overlapping, but separate, reference data sets.

3. There are effects of policy framework decisions directly affecting data aggregation:

- **The current deferral process isn’t consistently applied across the EU’s 27 jurisdictions**, which leads to inconsistent data across the EU as a whole.
- **The records for aggregated transaction leave gaps in the data** – records reported under the supplementary deferral flag regime (for example, COAF) do not, from most venues, contain actual trade dates or allow precise reconciliation with the underlying (initial) volume deferred records.

In general, we do **not** see the issues we have identified as being **specific to any single venue**, rather elements that, together, affect overall data quality, when records from multiple trading venues and relevant reference databases are aggregated and cross-referenced.

This paper seeks to summarise, with straightforward examples, some of the challenges faced by market participants – and retail investors - in relation to data quality.



Consistency

1. 'Notional Amount' and 'Quantity' fields ... on different venues

RTS2 requires multiple transaction fields to be provided, yet this creates inconsistency:

Examples: zero coupon 10-year (February 2031) BUND 'PERC'

Fields specified	Trade #1	Trade #2	Trade #3	Trade #4
ISIN	DE0001102531			
Venue of Publication (Data Group)	#3	#2	#3	#2
Notional Amount	1,350,000	25,000	2,233,845.9	3,200,000
Notional Currency	EUR	EUR	EUR	EUR
Price	104.91	101.31	1.01079	102.51
Price Currency	EUR	EUR	EUR	EUR
Price Notation	PERC	PERC	PERC	PERC
Quantity	1,350,000	2,500,000	2,210,000	1

The FIRDS records show a nominal value of €0.01¢:

Currency of nominal value	EUR
Nominal value per unit/minimum traded value	0.01

with the only record that is accurate being Trade #2

2. 'Notional Amount' and 'Quantity' fields ... on the same venue

example: zero coupon 10-year (August 2026) BUND via Data Group #3

Fields specified	Trade #1 Data Group #3	Trade #2 Data Group #3	Trade #3 Data Group #3
ISIN	DE0001102408		
Instrument ID type	ISIN		
Notation of Quantity in Measurement Unit	-	-	-
Notional Amount	1,560,000	1,000,000	2,543
Notional Currency	EUR	EUR	EUR
Price	104.06	104.05	104.014
Price Notation	PERC	PERC	PERC
Quantity	1,560,000	1	1
Venue of Execution	XOFF	SINT	XOFF
Venue of Publication	XXXX	XXXX	XXXX

Once again, the FIRDS records show a nominal value of €0.01¢, so each of these fields are incorrect.

3. 'Mone' misunderstanding

The definition is: “where price is reported in monetary terms, it shall be provided in the major currency unit.v”

Examples: zero coupon 10-year (February 2031) BUND 'MONE'

Fields specified	Trade #1	Trade #2	Trade #3	Trade #4
ISIN	DE0001102531			
Venue of Publication (Data Group)	#3	#3	#3	#2
Notional Amount	2,044,620	7,700,000	5,330,000	2,750,000
Notional Currency	EUR	EUR	EUR	EUR
Price	102.231	101.9655	105.181	102.42
Price Currency	EUR	EUR	EUR	EUR
Price Notation	MONE	MONE	MONE	MONE
Quantity	20,000	7,700,000	53,300	1

- Rather than a monetary amount being displayed, these records seem, in general, to represent **percentages**.

In most cases, the ability to reconcile the transaction records was further complicated by the fact that the reporting of **'QUANTITY'** - “the number of units of the financial instrument, or the number of derivative contracts in the transaction^{vi}” – did not follow a consistent pattern:

- In 80.6% of transaction records for this bond, it reflected the minimum denomination of the BUND (€100) combined with the **'PRICE'** as in Trade #1.
- In other cases (17.4%), the **'NOTIONAL AMOUNT'** was equal to the **'QUANTITY'** as in Trade #2.
- While, in last cohort (2.0%), there seemed to be no connection whatsoever between **'PRICE'**, **'QUANTITY'** and **'NOTIONAL AMOUNT'** that could be ascertained as in Trades #3 and #4.

4. Flags ... different venues with different practices for the same ISIN

We examined ETF transactions for an industry association and noticed that reporting was classified as either equity or fixed income. Where the instruments were classed Exchange Traded Certificates ('ETC'), they tended to be reported as fixed income. A deeper dive into the sub-asset class led us to look at those ETCs with larger volumes. However, we noticed that the pattern of reporting with flags related to an ETC – one of the most traded bonds with some **6,312** trades in 2021 – was inconsistent across venues:

By venue of publication	Number	ILQD % of total*
Data Group #1	3,512	0%
Data Group #2	357	71.2%
Data Group #3	2,443	55.6%

* ILQD flagged

- Note: the patterns were not solely linked to **'SINT'** or **'XOFF'** transactions

5. Flags ... incorrect application of the RTS 2 framework

A number of the **supplemental deferral flags**, as defined^{vii}, should only be reported on a mutually exclusive basis.

Non-equity flags specified in Table 3 of Annex II of RTS 2

- i. Descriptive flags: 'BENC', 'ACTX' and 'NPFT'. Descriptive flags can be combined with each other, with the exception of ACTX that cannot be combined with NPFT, as well as with flags under ii) and iv).
- ii. Post-trade deferral flags: 'LRGS', 'ILQD' and 'SIZE'. The application of the deferred publication is an option and not an obligation. Post-trade deferral flags should only be used in case of the effective use of the deferred publication. In case of the use of supplementary deferrals under iv), these flags should be used after the supplementary deferral period has lapsed and all the details of the transactions on an individual basis are published. These flags can be combined among each other, except 'LRGS' + 'SIZE', and with flags under i), iii) and v).
- iii. Package transaction flags: 'TPAC' and 'XFPH': These flags are mutually exclusive. In case of the use of supplementary deferral under iv), these flags should be used after the supplementary deferral period has lapsed and all the details of the transactions on an individual basis are published. These flags can be combined with flags under i), ii) and v).
- iv. Supplementary deferral flags: 'LMTF', 'DATF', 'VOLO', 'FWAF', 'IDAF', 'VOLW' and 'COAF'. These flags are mutually exclusive. **They cannot be combined with flags under i), ii), iii) and v).** For components of a package transactions, only the supplementary deferrals providing for volume omission under Article 11(3)(a) and (b) of MiFIR should be used. In case a package transaction benefitted from a deferral, all components should use the applicable flags under ii) and iv) (except DATV, FWAF and IDAF) regardless of whether those components would have qualified for such a deferral if they had been traded outside a package.
- v. Full details flags: 'FULF', 'FULA', 'FULV' and 'FULJ'. They should be reported once the deferral time period lapses and all the details of the transactions on an individual basis are published. These flags are mutually exclusive and should be combined with the post-trade deferral flags under ii). These flags can be combined with flags under i) and iii).

in some recent analysis we conducted for sovereign bonds, we saw that some **3%** (by number of records) and **2%** (by traded volume) combined supplemental deferral flags with other categories

Coherence

6. Transaction records that make no sense ... the trade with the 'lowest' price

Fields specified	Bond details
ISIN	DE000A3E5KG2
Instrument ID type	ISIN
Notional Amount	45,624,176
Notional Currency	EUR
Price	-372,176.23
Price Currency	EUR
Price Notation	MONE
Quantity	431
To be Cleared	FALSE
TradeTimeDate	2021-06-XX
Venue of Publication	XXXX
Flags	FULJ, ILQD, LRGS

We can see that the bond is:

- A 5% TUI AG semi-annual pay, convertible bond
- Issued in April 2021 and maturing April 2028
- Trades in minimum units of €100,000

From a 'market' perspective:

- Around the trade date (+/- 1 day), the price was in a range between 103.1% – 104.6%
- The bond is traded on a number of exchanges

7. Transaction records that make no sense ... the trade with the 'highest' price

Fields specified	Bond details
ISIN	HU0000523071
Instrument ID type	ISIN
Notional Amount	165,000,000
Notional Currency	HUF
Price	164,979,375
Price Currency	HUF
Price Notation	PERC
Quantity	16,500
To be Cleared	FALSE
Trade Time Date	2021-04-XX
Venue of Publication	XXXX
Flags	LRGS, ILQD, SIZE

The transaction is a 'PERC' with a monetary amount mistakenly entered:

- A domestic Hungarian 364-day bond
- Issued in April 2020, maturing 21st April 2021
- The bond seems to be offered via the domestic savings banks and Volksbank associations.

From a 'market' perspective, we can make the following observations about this bond:

- Although reported as 'PERC', the price is not expressed as a percentage, rather a monetary amount – the corrected 'PERC' price is **99.9875%**.

8. Transaction records that make no sense ... 'PRICE' < 1

Fields	Bond #1	Bond #2	Bond #3
ISIN	XS1683139692	XS2250998049	US445545AH91
Instrument Group	BOND	MEDIUM TERM NOTE	BOND
Notional Amount	12,000,000	150,000	500,000
Notional Currency	US\$	EUR	US\$
Price	0	.9984	-0.43736
Price displayed on relevant exchange during 2021	98.941 - 101.121%	100.00 - 102.68%	105.30 - 109.70%

- In some cases, such as Bond #2, the 'PRICE' representation can be adjusted to a 'PERC' – based on an assumption.
- Whereas in the other cases, the explanation is not as clear. For Bond #3 (a Republic of Hungary 5.375% 2023 bond) there are 118 transactions affected by this form of reporting.

9. Transaction records that make no sense ... 'PRICE' > 150

Fields	Bond #1	Bond #2	Bond #3
ISIN	XS1925432400	FR0014003J12	DE000A190ND6
Instrument Group	MEDIUM TERM NOTE	ASSET BACKED SECURITY	MEDIUM TERM NOTE
Notional Amount	1,225,320	40	500,000
Notional Currency	EUR	EUR	EUR
Price	306,330	101,480	100,345
Price displayed on relevant exchange during 2021	96.15 - 98.89%	-	100.255 - 101.235%

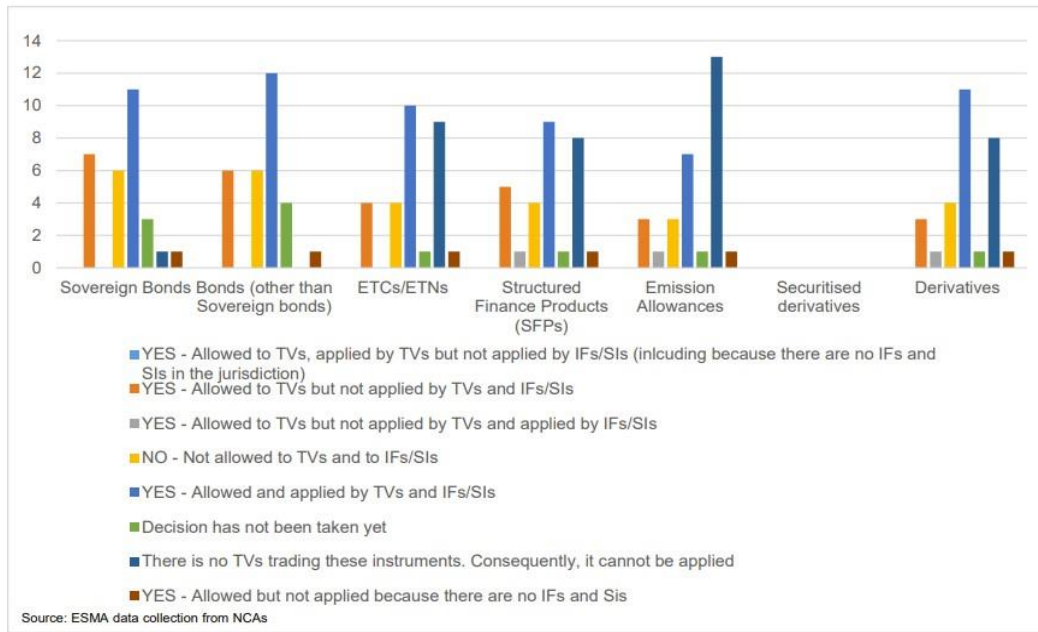
- Once again, based on a broad assumption, the representation of 'PRICE' for Bonds #2 and #3 might be easily adjusted to 101.48% and 100.345% respectively - although it is not clear who would intervene to ensure remediation.
- However, the Bond #1 is not easy to explain nor remediate

Policy

1. Different application of the same rules

Even though the rules that are laid out by ESMA are clear and singular, there are still discretions applied at a national level across the EU-27. A recent report^{viii} by ESMA highlights the complexity of deferral regimes across the EU-27:

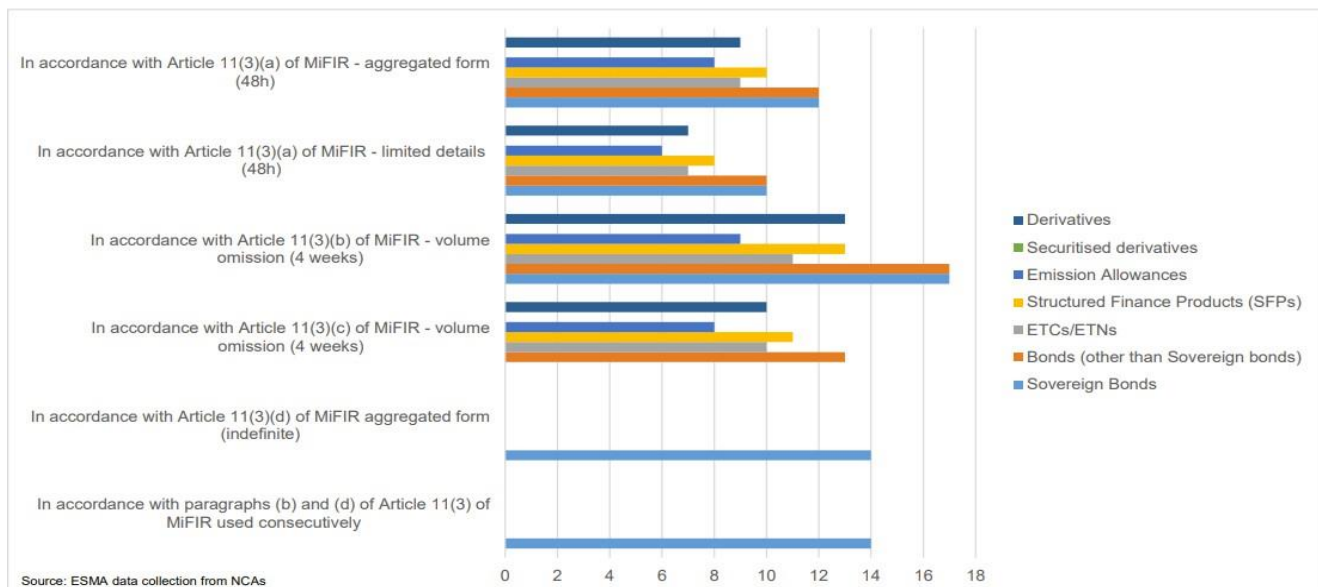
FIGURE 46: NUMBER OF NCAs APPLYING THE NON-EQUITY DEFERRAL REGIMES FOR TRANSACTIONS IN ILLIQUID INSTRUMENTS, 2021



2. Permutations on a theme ...

Furthermore, the uneven approach taken for supplemental deferrals (COAF, IDAF) for sovereign bonds, in particular, creates complexity vis-à-vis standardised data:

FIGURE 50: NUMBER OF NCAs APPLYING THE SUPPLEMENTARY DEFERRAL REGIME FOR TRANSACTIONS IN ILLIQUID FINANCIAL INSTRUMENTS, PER ASSET CLASS, 2021



FIRDS

1. Tapping of sovereign bonds FR0011883966 as an example

We can observe the 'tapping' of a French OAT, as a generic example, of how, especially for liquidity calculations over specific periods, the issuance amount changes:

Date	Status	Amount Outstanding (€)	Currency	Issue/Redemption Price (%)	Yield (%)
6-May-14	Initial Issuance	4,965,000,000	EUR	98.18	2.64
10-Jun-14	Re-opening	7,757,000,000	EUR	99.54	2.54
8-Jul-14	Re-opening	11,015,000,000	EUR	101.57	2.38
9-Sep-14	Re-opening	15,210,000,000	EUR	107.79	1.92
10-Nov-14	Re-opening	17,941,000,000	EUR	109.25	1.81
12-Jan-15	Re-opening	22,405,000,000	EUR	116.61	1.3
9-Feb-15	Re-opening	24,360,000,000	EUR	122.19	0.94
8-Jun-15	Re-opening	28,387,000,000	EUR	111.01	1.66
7-Sep-15	Re-opening	31,023,000,000	EUR	111.59	1.61
8-May-17	Re-opening	33,208,000,000	EUR	117.29	1.07
12-Nov-18	Re-opening	35,119,000,000	EUR	117.19	0.92
11-Feb-19	Re-opening	38,876,000,000	EUR	119.5	0.7
5-Aug-19	Re-opening	41,642,000,000	EUR	128.21	-0.1
8-Jun-20	Re-opening	45,360,000,000	EUR	124.65	0.02
10-Aug-20	Re-opening	49,233,000,000	EUR	127.56	-0.27
11-Oct-21	Re-opening	52,189,000,000	EUR	122	-0.05
6-Dec-21	Re-opening	54,855,000,000	EUR	122.93	-0.19
25-Apr-22	Re-opening	57,749,000,000	EUR	110.65	1.12

2. Tapping of sovereign bonds using issued amount as a denominator

While the details maybe accurately updated, the 'history' is overwritten as the records currently in FIRDS show:

Notional currency 1	EUR
Total issued nominal amount	57749000000

However, for any market participant using FIRDS as a reference dataset, looking at liquidity for 12 months to end April 2022 they should be seeing the issued amount change over that period - from €49,233,000,000 to €57,749,000,000.

3. Bonds using FIRDS data for maturity dates

Similarly, we find that maturity dates, based on venue selection, can cause differing data to be created on the FIRDS system itself.

Trading venue	BTEE	TSAF
Financial instrument short name	FRANCE EMPRUNT /DBT 20300525	FRANCE EMPRUNT /DBT 20300525
Request for admission to trading by issuer	No	No
Date of admission to trading or date of first trade	2014-04-28 06:03:29.0	2018-05-29 00:00:00.0
Termination date	2019-08-30 17:15:00.0	2030-05-25 00:00:00.0

4. Bonds 'outsized' transactions records

For our recent report with AFMEix, we noted that there was a small number of unusually large transactions where the notional amounts reported were a number of times greater than the (original) issuance size. These amounts represented some **27.5% (€2,316.2bn)** of the gross amount and we excluded these from the analysis as we could not reconcile the 'skew' associated with **outsized single records** that had been reported:

Table 1: Relevance of 'Irregular' records

Description	Traded Volume (€bn)	% total	# records	% total
'Gross' Traded Volume - original sovereign bond data set	8,414.9	100%	1,799,521	100%
'Net' Traded Volume - minus outsized trades	6,098.8	72.5%	1,799,514	100%

The details of those 'abnormal' records are as follows:

Table 2: 'Irregular' or outsized records

Description	Flag	Total volume (€bn)	# ISINs	Jurisdiction	Times Issuance size
Record #1	COAF	1,764.4	1	UK	425x
Record #2	COAF	204.2	1	Germany	9x
Record #3	COAF	166.4	1	Germany	7x
Record #4	COAF	88.0	1	Slovenia	29x
Record #5	COAF	60.1	1	UK	1x
Record #6*	FULJ	40.0	2	EU (as issuer)	2x
Record #7	COAF	13.1	1	South Korea	35x
Total		2,316.2			

*a record for an EU bond seemed replicated and one of the records was omitted – a single record of €20bn remains in the dataset

5. Aggregation and Indefinite Deferral Flags

Unlike the MiFIR reporting regime for corporate bonds where aggregated trades must be disaggregated within 4 weeks, the regime for sovereign bonds permits national competent authorities ("NCAs") to allow indefinite deferral of disaggregation.

This materially reduced FINBOURNE's ability to make conclusions from the data, due to a variety of flags reported related to the aggregations and deferrals. A summary of how the various flags are defined is included below.

Proposal for bonds

Transaction data fields

1. MiFID II post-trade transaction reports

- There are up to **38** fields in MiFID transaction reports, yet in many cases, FINBOURNE believes that a smaller number of fields is required to provide the relevant essential **bond** information.

Of MiFIR's RTS 2 **41** Business Fields, only **8** specifically reference (directly or indirectly) bonds and in order to improve the effectiveness of any **Fixed Income CT**, we would propose that **the following 6 or 7 trade fields (as well as venue and timestamp) are needed for bonds:**

1. Identifier

Details	PROPOSAL
Instrument identification code	can cross-referenced to FIRDS

2. Price

Details	PROPOSAL
Price	'Price' should remain the clean price
Price notation	For Fixed Income, the 'Price' should ONLY be expressed as 'PERC'
Price currency	Transactions should clearly express if differs from 'Notional Currency'

3. Notional

Details	PROPOSAL
Notional currency	could be a FIRDS field populated directly?
Notional amount	should be a 'clean' amount

4. Quantity

Details	PROPOSAL
Quantity	For 'unitised' instruments (e.g. structured bonds or ETNs/ETCS) could be added for information but for other fixed income products, not relevant

This was feedback received in the Call for Evidence on RTS 1 and RTS 2 by ESMA and 1 considered as part of the recent RTS 1/RTS 2 Consultation Paper:

3.3.1.3 Fields "Price", "Price currency", "Price notation" and "Quantity"

142. In the CfE several stakeholders requested that APAs publish the price in the post-trade reports in EUR. ESMA considers that the information on the currency in which the trade was made is sufficient to compare post-trade reports and does not consider it necessary to add this new requirement.

Concentrating on a smaller number of fields and industry and creating a set of 'market practices' or 'standards' for trade reporting could exponentially increase the accuracy, as well as usefulness, of transaction data.

2. Value of a consolidated tape

A single tape allows direct comparisons of transactions across the array of EU trading venues, which, if enhanced with the appropriate technology can support the development of, and improvements to, a single, accurate view of transactions. This will help create visibility of credible data for the market, regulators, policy makers and the general public.

FIRDS data

1. Data fields

There is a degree of complexity in **FIRDS record fields** that could be re-assessed. There are **113** fields of which **14** related to debt instruments, yet **84** relate to derivatives.

2. Legacy currencies

In previous analysis we have conducted,^{vi} we found transactions remained in the 'system' in 3 legacy currencies - **Italian lira**, **Dutch Guilder** and **German Deutschemark** (long dated bonds issued pre-EMU). A one-off exercise to amend these records would be useful.

Policy

1. Application of deferrals at national level

While the merits (or de-merits) of policy itself are not subjects on which a technology firm is in a strong position to opine, the uneven application of these flags in and of themselves, creates 'lumpy' data that is inconsistent and limits the effectiveness of any efforts to create standardised and comparable transaction data across the market as a whole.

Greater efforts should be made to harmonise approaches across the EU-27 NCA's. These efforts would need to be carried out in full consultation with wider market participants and should, where feasible, be informed and led by analysis of existing, dis-aggregated trade data.

Selected trade flags glossary

Acronym	Definition	Explanation
AMND	Amendment Flag	When a previously published transaction is amended.
DUPL	Duplication Flag	When a transaction is reported to more than one APA.
CANC	Cancellation Flag	When a previously published transaction is cancelled.
COAF	Consecutive Aggregation Flag	Supplementary deferral allowing for an indefinite period without full post-trade transparency
IDAF	Indefinite aggregation flag	Transactions for which the publication of several transactions in aggregated form for an indefinite period of time has been allowed.
PNDG	Pending Flag	When a transaction or price is currently not available but due to become available.
LGRS (LIS)	Post-Trade Large in Scale Transaction Flag	Transactions where deferred publication is permitted on the basis that they are large in scale compared with normal market transactions.
TPAC	Package Transaction Flag	A trade composed of several components/legs.
ILQD	Illiquid Instrument Flag	Transactions executed under the deferral for instruments for which there is not a liquid market.
FULJ	Full Details Flag	Individual transactions which have previously benefited from aggregated publications because of their status as non-equity instruments that are not sovereign debt.
FULV	Full Details Flag	Transactions for which limited details have been previously published about why an individual transaction has been given an extended time period of deferral for four weeks.
FULF	Full Details Flag	A transaction whereby limited details have been published about the value and average daily turnover of the transaction.
FULA	Full Details Flag	Individual transactions for which aggregated details have been previously published.
LMTF	Limited Details Flag	A supplementary deferral which requires additional information.
DATF	Daily Aggregated Transaction Flag	A supplementary deferral which requires additional daily aggregation information.
VOLO	Volume Omission Flag	A supplementary deferral allowing for an extended period without full post-trade transparency.
VOLW	Volume Omission Flag	Transactions for which limited details are published and for which the publication of several transactions in aggregated form for an indefinite period of time will be allowed.

DISCLAIMER

This document has been prepared by FINBOURNE Technology Limited (“FINBOURNE”) on an ‘as is’ basis. It provides general background information regarding FINBOURNE’s activities and is supplied for information purposes only. Nothing in this document should be regarded as an invitation, inducement or recommendation to engage in investment activity (a financial promotion) as defined in section 21 of the UK Financial Services and Markets Act 2000 (“FSMA”) and the information contained in this document is not intended to be an offer to buy or sell any interest in any investment.

Information set forth herein is only a summary of certain information as at the time this document is provided. FINBOURNE does not make any representation or warranty (express or implied) as to the accuracy or completeness of the information provided and it shall have no liability in relation to the content or its use. You should not place any reliance on any statements contained herein and such statements are subject to change by FINBOURNE and uncertainty and contingencies outside FINBOURNE’s control. The information, materials and opinions contained herein are not intended to constitute legal or other professional advice and should not be relied on, nor treated as a substitute for specific advice relevant to particular circumstances.

This document is the property of FINBOURNE and any reproduction, dissemination or re-distribution of this document or the information herein without FINBOURNE’s prior written consent is forbidden.

© 2022 FINBOURNE Technology Limited. All rights reserved.

-
- i Insights | FINBOURNE
 - ii AFM invites businesses to respond to its ‘Principles for Choice Architecture’ | AFM
 - iii AFME/FINBOURNE study shows need for longer deferrals for large, fixed income trades | AFME and MiFIR 2021 Sovereign Bond Trade Data Analysis and Risk Offset Impact Quantification | AFME
 - iv Questions and Answers ‘On MiFID II and MiFIR transparency topics’ version page 16 5 September 2022 | ESMA70- 872942901-35
 - v Technical Standards - FCA Handbook ANNEX II Details of transactions to be made available to the public Table 2 List of details for the purpose of post-trade transparency
 - vi as above
 - vii see (iv) above
 - viii Annual Report – 2022 On the application of waivers and deferral page 84 ESMA 22 November 2022 https://www.esma.europa.eu/sites/default/files/library/esma70-156-6093_annual_report_2022_waivers_and_deferrals.pdf
 - ix MiFIR 2021 Sovereign Bond Trade Data Analysis and Risk Offset Impact Quantification | AFME
 - x Consultation Paper On the review of RTS 1 (equity transparency) and RTS 2 (non-equity transparency) 9 July 2021 | ESMA70-156-4236
 - xi Insights | FINBOURNE



About

FINBOURNE's solutions deliver an interconnected network of functionality and data that enables the investment community to better serve clients in a constantly evolving market.

Its investment management solutions and cloud-native data management platform ensure that investment and operations teams can increase revenue, reduce costs, and better manage risk across the investment life cycle.

Get in touch

To discover more about FINBOURNE and learn how our solutions can drive growth, increase control and improve data access, contact us below

finbourne.com/info@finbourne.com

+44 (0)20 3880 1307

FINBOURNE Technology

North America: 666 3rd Avenue, New York, 10017

UK: 1 Carter Lane, London, England, EC4V 5ER

Singapore: 790, Level 7 Capital Square,
23 Church Street, 049481