

Market data normalisation for low latency environments: Addressing fragmentation challenges at scale



Introduction

Fragmented market data remains one of the most persistent sources of cost and operational drag in trading. Its impact is felt most by firms scaling across asset classes, modernising platforms, or under pressure to improve efficiency while maintaining market integrity.

In this expert talk, Pramod Nayak and Zeyu Zhao from LSEG's Low Latency Group explore how firms are addressing market data fragmentation through large-scale normalisation.

- Why market data fragmentation is a structural challenge, not a tooling problem
- How hidden complexity imposes a “technical tax” on trading and investment operations
- Why data normalisation goes far beyond field mapping
- How LSEG treats normalisation as strategic infrastructure, supporting more consistent, high-fidelity data across global markets to meet diverse latency requirements and use cases.

Authored by:



Pramod Nayak

Director, Product Management, Low Latency Group

LSEG

[Connect on LinkedIn >](#)



Zeyu Zhao

Team Lead, Business Analysis, Low Latency Group

LSEG

[Connect on LinkedIn >](#)

The hidden costs of fragmented market data

For trading firms, fragmented market data can add cost, risk, and operational complexity, often without being directly visible in the P&L. To support in-house consumption of exchange-generated market data feeds, firms need to manage data relationships across hundreds of venues. Each one of these exchanges carries its own data rules, formats, policies, and change management processes for the data it produces – creating additional operational complexity.

The result is data fragmentation, leading to a metaphorical “Tower of Babel” effect within a firm’s market data operations. Differences in data dialect and how data is handled can create potential disconnects, with downstream impacts on workflows and analytics.

Fragmentation typically manifests across three dimensions:

Terminology issues

Inconsistent field names (such as LastPx vs. TrdPx) is one challenge.

Another is semantic drift, for example when US SIP quote size shifted from lots to shares. These differences can create challenges in data identification and deployment.

Rule inconsistencies

Exchanges implement unique state machines and behaviours.

Variations in auction rules can lead to mistimed executions and flawed liquidity analysis.

Timing differences

Feed latency and out-of-order messages can generate “phantom opportunities” and confuse automated systems.

For firms managing exchange market data in-house, these challenges are continuous. They affect decision-making not just at the point of execution but across front, middle and back-office workflows.

The technical tax on trading and investment operations

Over time, fragmentation imposes a cumulative “technical tax” on trading and investment operations, resulting in hidden costs that erode profitability. This emerges through:

- **Increased operational and compliance risk:** Failing to keep up with exchange-driven changes can increase the likelihood of data errors which can contribute to operational, compliance, and trade execution risk. The costs that can emerge due to these risks crystallising can be large.
- **Impaired decision-making:** Poor or inconsistent data quality can result in traders and investors acting on misleading signals, with potential implications for operational efficiency and profitability.
- **System instability:** Incorrect or poorly sequenced data can impact system stability and increase the likelihood of remediation effort and downtime.
- **Operational maintenance overhead:** Each exchange update requires analysis, testing, and deployment. At global scale, the annual volume of changes becomes a meaningful and high-touch operational challenge.
- **Innovation contraction:** Ongoing data issues often mean that data scientists and quants spend their valuable time manually cleaning data sets before use, reducing their availability to innovate.

Taken together, fragmentation turns market data from an enabler into a constraint.

Why normalisation is harder than it sounds

At first glance, data normalisation sounds straightforward. In practice, it is anything but. While many of the same data normalisation principles apply across real-time data workflows, they become more complex and time-sensitive in low latency environments. Normalisation typically requires deep knowledge of:

- Exchange microstructure behaviours
- Protocol evolution over time
- Market-specific boundary conditions
- The way trading intent is communicated, rather than simply encoded

As firms expand across asset classes and regions, syntactic differences become a direct barrier to speed, scalability, and determinism.

This is why LSEG treats a robust normalisation layer as core infrastructure, not a downstream transformation step. Within LSEG's Low Latency Group, we structure this challenge through a normalisation maturity curve, comprising three distinct but interdependent phases:

Phase One: Syntactic alignment

The core objective of syntactic alignment is to ensure that all market events follow the same message structure. This helps to create seamless integration of data across different systems, and includes standardising messaging formats, harmonising field definitions, and enabling schema-level interoperability. Areas of focus in this phase include:

- **Event envelopes** – Establishing a traceable, logically sequential universal message header, including key fields such as Feed, Sequence Number, and Timestamps.
- **Unified terminology library** – Defining standardised data elements, such as price adjustments, time in UTC, and eliminating inconsistent field naming issues.
- **Normalised message types** – Defining “standard verbs” covering market actions – such as AddOrder, Trade, and ProductStatus – to ensure consistent parsing.

This phase allows firms to align platforms, applications, data governance, and frameworks to a single set of terms – a shared language.

Phase Two: Semantic calibration

This phase focuses on meaning rather than structure. We apply normalisation to business logic to ensure similar events are interpreted consistently across exchanges.

For example, TradingStatus.AUCTION carries a consistent and comparable business meaning across all markets, to eliminate dialect differences among the exchanges. So, on the Tokyo Stock Exchange, lifecycle management helps to isolate auction-specific orders that remain unexecuted after the opening auction. This helps reduce the risk of these orders from flowing into continuous trading, supporting orderly market behaviour by limiting crossed order books and maintaining market integrity.

Other steps that we take in this phase include:

- **Managing a time truth strategy** – Prioritising ExchangeTimestamp with ReceiptTimestamp as a supplement, so that data is processed in a more correct chronological order.

- **Decoding business intent** – Normalising SaleCondition and unifying semantics such as “whether to update statistical indicators” as well as supporting multi-value validation, to improve data accuracy.
- **Resolving boundary conditions** – Establishing unified handling strategies for odd lots and imbalance reports to support data integrity.

This phase ensures that exchange activities are identified and processed consistently, not just syntactically, but logically.

Phase Three: Stateful reconciliation

In this final phase, fragmented data streams are reconciled into a coherent, high-fidelity view of the market. This involves:

- **Consistency contracts** – Establish technical specifications designed to support the construction of a fully consistent order book, designed to support the comparability of data.
- **Order book blueprint** – Define a unique “scope” that includes InstrumentID, Venue, Session, and BookType, supporting the uniqueness and consistency of the order book.
- **Principle of data segregation** – By default, exclude “irregular” or “contaminated” data from normal datasets to preserve data purity and prevent data pollution.

The outcome is a state-consistent order book that is designed to support trading, risk management, and historical analysis and backtesting.

Managing this level of complexity across hundreds of venues requires sustained investment and specialist expertise. LSEG makes this investment, so clients do not need to replicate that footprint themselves.

Climbing the normalisation maturity curve

By the time market data reaches a trading or analytics platform, many of the most complex decisions have already been made. The practical value of LSEG’s normalisation efforts lies in the engineering work that can be reduced within a client’s environment.

As firms connect to more exchanges, complexity does not just increase linearly. Differences in market behaviour, protocol design, and data expression quickly multiply, requiring internal teams to build and maintain exchange-specific logic. LSEG addresses these challenges, helping to manage complexity upstream so clients can operate against a consistent model.

Here are some instances that demonstrate how this occurs in real situations:

1. Abstracting exchange-specific data models

- Different exchanges often represent the same concept in fundamentally separate ways. For example, CME MDP 3.0 and ICE iMPact both distribute implied liquidity, but do so using very different approaches. CME uses a separate implied book and complex data handlers, while ICE integrates implied data and requires more sophisticated strategy logic, requiring more complex interpretation at the strategy level.
- Without normalisation, firms must implement and maintain separate data handlers and strategy logic for each approach. LSEG abstracts these differences by clearly identifying implied liquidity and delivering it in a more consistent and well-defined structure. This provides a more consistent approach for firms to ingest and use data across venues using one implementation model, rather than bespoke logic per exchange.

2. Removing venue-specific structure from market data

- a. Some markets introduce additional complexity through their structure rather than their protocols. In Canada for example multiple parallel order books can operate within a single venue. Left raw, this forces downstream systems to carry detailed knowledge of how each book behaves and how trades and status events should be interpreted.
- b. LSEG normalises these structures by clearly defining order book scope, standardising order book status and trade semantics, and converting exchange-specific terms into a common data model. The result is that clients can work with a more coherent view of the market, without embedding venue-specific assumptions into every application.

3. Establishing consistent sequencing where exchanges do not

- a. Not all exchanges provide reliable or complete timestamps. This can create significant challenges for sequencing events, replaying markets and maintaining a consistent historical view.
- b. When this occurs, LSEG applies an automated receipt timestamp at the point of capture, helping to support a consistent sequencing framework. Clients can then process and analyse data in the correct chronological sequence, without building their own reconciliation or ordering logic.

Taken together, these examples show how expert normalisation operates in practice. In effect, this approach reduces the need for clients to build exchange-specific handlers, reconcile inconsistent market logic, or maintain parallel data models across regions.

Engineering teams can implement once, apply unified logic globally, and operate with greater confidence in how market behaviour is represented consistently (even as exchanges evolve), reducing long-term maintenance effort and helping free teams to focus on trading and analytics rather than data plumbing.

Interested in learning more? Speak with our team to explore how our normalisation infrastructure helps shape and simplify your downstream workflows. [Request details >](#)

Client examples and approaches described reflect specific implementations and design practices. Outcomes may vary depending on client environment, configuration, venue behaviour, regulatory considerations, and downstream system design. LSEG services are designed to support operational resilience, consistency and usability of market data, but do not guarantee accuracy, completeness, sequencing, availability, or specific system outcomes.

© 2026 LSEG. Reproduction or redistribution of LSEG content, including by framing or similar means, is prohibited without the prior written consent of LSEG. LSEG is not liable for any errors or delays in LSEG content, or for any actions taken in reliance on such content. LSEG Data & Analytics logo is a trademark of LSEG and its affiliated companies.

lseg.com