FMSB
FICC MARKETS STANDARDS BOARD

Spotlight Review

The critical role of data management in the financial system

May 2020

About FMSB

FICC Markets Standards Board Limited ('FMSB') is a private sector, market-led organisation created as a result of the recommendations in the Fair and Effective Markets Review ('FEMR') Final Report in 2015. One of the central recommendations of FEMR was that participants in the wholesale fixed income, currencies and commodities ('FICC') markets should take more responsibility for identifying and fixing poor market practice so that they operate in the best interest of their clients. Clear, practical guidance that delivers transparent, fair and effective practices will rebuild sustained trust in wholesale FICC markets.

FMSB brings together people at the most senior levels from a broad cross-section of global and domestic market participants and end-users.

In specialist, focused committees, sub-committees and working groups, industry experts debate issues and develop FMSB Standards and Statements of Good Practice, and undertake Spotlight Reviews that are made available to the global community of FICC market participants and regulatory authorities. FMSB has issued 18 publications since 2016. As part of its analysis on the root causes of market misconduct, FMSB is focusing on the challenges of new market structures.

Spotlight Reviews

Spotlight Reviews encompass a broad range of publications used by FMSB to illuminate important emerging issues in FICC markets. Drawing on the insight of members and industry experts, they provide a way for FMSB to surface nascent challenges market participants face and may inform topics for future work. Spotlight Reviews will often include references to existing law, regulation and business practices. However, they do not set or define any new precedents or standards of business practice applicable to market participants.

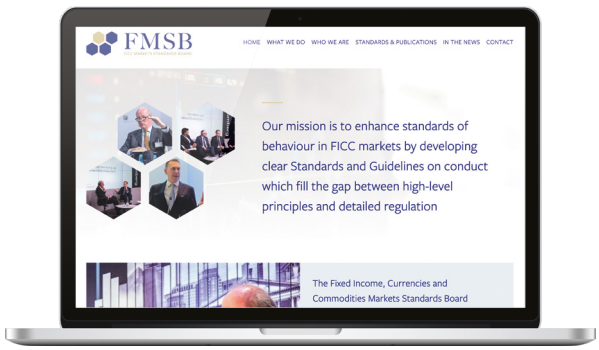
Find out more about the FICC Markets Standards Board on our website fmsb.com

The author

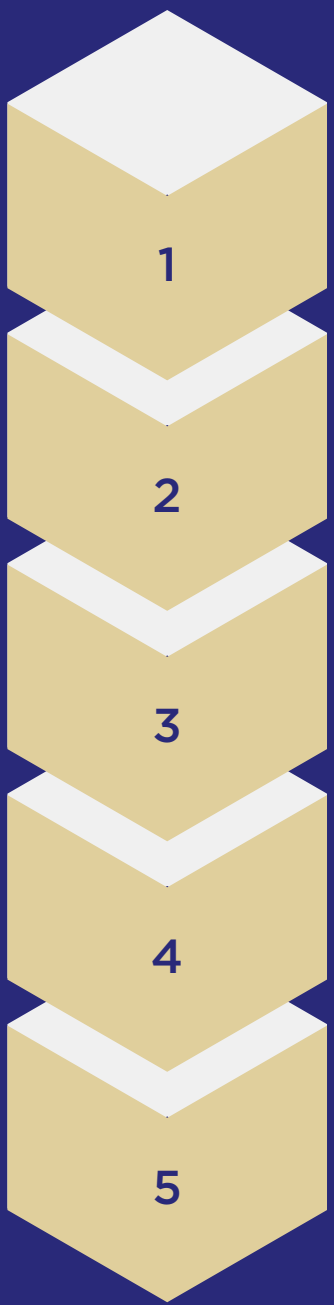
Rupak Ghose is leading the FMSB Spotlight Reviews on FICC market structure and the impact of regulatory and technological change on the fairness and effectiveness of wholesale FICC markets.

Rupak has more than 15 years' experience in FICC market structure. He was Head of Corporate Strategy for ICAP/NEX for six years providing advice on strategic direction, market structure change, customer/competitor landscape, new business ventures and M&A. He was a trusted counsellor to the PLC Board on the dramatic transformation of the business from interdealer broker to focused financial technology firm and the ultimate sale of the businesses.

Prior to this Rupak spent more than a decade at Credit Suisse as an equity research analyst where he was Top Three ranked on numerous occasions by buy-side clients in Institutional Investor and other surveys for his coverage of asset managers, exchanges and other capital markets focused companies.



Contents



Executive summary

This Spotlight Review examines the crucial role that data and the management of data play in today's wholesale FICC markets and financial systems. It aims to create further discussion on this topic and its relevance to future standards work by FMSB. This Spotlight Review will be of interest to those with responsibility for handling and using data for information and as an input to decisions from its creation through the entire business process.

Data plays a critical role in today's financial system. If markets are to remain stable and trusted, fair and effective, then the rapid growth of new technology and of data science must be balanced with ever more effective governance and control. Data runs through all the infrastructure of participants in global wholesale FICC markets, including the pricing, order and trade management, risk management, regulatory reporting, financial and corporate systems. These systems continuously process and feed data around the organisation. Each system plays a different part in storing, processing, enhancing and transmitting data. Data is duplicated, aggregated, integrated, cleansed, enhanced and acted upon in real time, in multiple locations around the world. Failure to exercise appropriate controls over data increases the significant risks to organisations and market functioning, from both a reputational and financial perspective.

The amount of data being handled, manipulated and acted on by financial markets participants has grown exponentially in recent years. Rapid changes in the uses, and users, of data – where the same term can be duplicated and presented in a myriad of ways using a variety of definitions – risks introducing high degrees of complexity and inconsistency. Increasing complexity and inconsistency could also increase the risks of market misconduct and market instability resulting from improper access to, or use of, data. Against this background, the need of firms for secure, accurate, timely, complete and consistent datasets has never been greater. To mitigate these risks there is an increasing focus on the importance of developing standards for data and on strengthening data governance.

The huge growth of electronic trading in the past two decades, coupled with increased electronic reporting requirements, has significantly increased the quantum of transaction data. This has increased transparency and improved overall efficiency of FICC markets. However, there remain issues with the clarity, consistency and ease of use of data.

A standardised approach to data governance can be effective in controlling the risks associated with sourcing, storing, processing, transmitting, selling, and using data. Different industry initiatives such as the Financial Information eXchange (FIX) Protocol and Legal Entity Identifiers (LEIs) have been successful in standardising the way in which financial data is transferred and counterparties and clients referenced from one system to another reducing risks arising from the data being 'lost in translation'. But more is needed to manage the risks arising from the ever-increasing role of data in financial markets.

For global wholesale FICC markets this Spotlight Review outlines the:

> principal areas of data risk;

> regulatory authorities' work in this field; and

> key components of data governance.

Contents

Screening for the seven sources of critical data risk

The proper functioning of trading, reporting and risk management is critically dependent on the security, accuracy, timeliness and integrity of data. Wholesale FICC market participants are facing increased risks to stability and conduct resulting from the growing challenges of data management. Standardisation and good data governance are key to mitigating these risks.

Many financial services market participants are undertaking regular comprehensive reviews of risks and controls to identify, understand and mitigate the different sources of risk. Increasingly these reviews are seeking to address enterprise-wide risks related to data. Data risk comes in a broad range of shapes and sizes and depends on the specifics of an organisation, its management of technology and its framework for governing data. Here we focus on seven sources of data risk vital to stability, fairness and effectiveness.

1. Business continuity and operational risk

Dependence on critical data sources can lead to significant loss of capability should those sources be interrupted or corrupted. Understanding those data services that are critical to the accuracy and availability of core functions is therefore key to establishing appropriate monitoring and contingency plans to cope with any potential loss of availability. Critical data sources include third-party suppliers of data or data processing. In FICC markets this may include the dependence on public reference prices provided by electronic trading platforms or essential reference data inputs.

2. Security and confidentiality risk

Ineffective controls to protect data can result in inadvertent disclosure or unauthorised access to data either internally or externally. This can result in breaches of applicable law or regulation relating to the control of personal and other commercially sensitive data (such as positions and exposures) and can lead to other abuse risks arising from access to inside information, significant loss events and regulatory action. This risk increases as more and more sensitive position and other data is transmitted across organisations, is duplicated in different systems and is sent to third parties such as regulatory authorities, trade repositories, compression providers, and other central infrastructure providers, such as central counterparty clearing houses (CCPs) and settlement services.

3. Commercial trading risk

Both humans and machines rely on accurate data in order to achieve optimal outcomes in trading, investing and risk management. Performing due diligence on sources of data is key to ensuring they are fit for purpose and have sufficient rigour and controls across their creation and usage. The risk from stale or erroneous data is greater in discontinuous low liquidity FICC markets where there may be naturally occurring gaps in trading activity. A focus on data quality should include any inputs from evaluated or indicative pricing models to identify and mitigate scenarios where unintended outcomes may result.

Systemic data risk continued

4. Aggregate exposure risk

If data pertaining to risk positions in different parts of a firm or running through different systems cannot be aggregated into a consistent centralised picture of risk exposure in a timely way, then this could give rise to significant and unanticipated firm-wide exposures. Such risks may be exacerbated where large intraday positions are frequently being taken and unwound before end of day, creating significant spikes in real-time risk at key points in the day. Accurate timestamping and time-series risk management capabilities are key to understanding the profile of aggregated risk on a granular intraday basis, but the quality and quantity of data and computer power needed are significant. In an ever-changing world with the frequent introduction of new systems and processes it can be very hard to maintain continuous transparency on intraday aggregate risks.

5. Regulatory enforcement risk

Regulatory supervisors are increasingly taking punitive action against firms consistently failing to meet obligations to report in an accurate and timely manner to regulatory authorities. If firms cannot accurately map their data to the requirements of a multitude of different reporting obligations, then these risks can result in material financial, reputational and regulatory consequences. Regulators can take enforcement action over a range of other data-related failures that lead to operational instability, lack of transparency and conduct issues.

6. Ownership and rights risk

Ambiguity and misunderstanding of commercial rights over data is an increasing risk. This is particularly an issue given the explosion in the amount of data held by third-party vendors, and the complexity and lack of consistency in many agreements. Data bought under licence from a third-party provider or captured as part of a service provided to customers, and then processed, combined or enriched, can result in significant uncertainties with regard to the rights and obligations for the holder of that derived data.

7. Security and conduct risk

Inadequate controls over permissions for access and manipulation of data could lead to opportunities for misconduct.

The extent to which these seven sources of data risk can combine to cause losses, disrupt stability, damage reputations and result in the destruction of trust in markets is prevalent in the minds of firms' boards and executives, and of the regulatory authorities. So, what steps have been taken so far?

The post financial crisis push for increased reporting by regulators

In the decade since the start of the global financial crisis there has been an unprecedented focus from financial services regulators on the transparency and reporting of data. In many cases this has brought significant benefits, increasing the fairness and effectiveness of FICC markets. However, much of the new data eco-system that resulted has not been optimised because of a lack of data quality and standardisation. This is particularly evident in the context of Markets in Financial Instruments Directive (MiFID) II where the lack of specificity in post-trade transparency requirements has led to inconsistencies in the disclosures provided by market participants, which in turn has impacted their usefulness in promoting fair and effective markets.

In January 2013, the Basel Committee on Banking Supervision (BCBS) published BCBS 239 'Principles for effective risk data aggregation and risk reporting'.[1] There are fourteen principles of which the first six cover effective risk data aggregation with guidelines on the need for strong governance, data, and IT architecture and the accuracy, integrity, completeness, timeliness, and adaptability of risk data aggregation. These guidelines are relatively high level but have provided a catalyst for the considerable investment being made by banks in this area in recent years. This has involved the implementation of broader data management strategies by banks, that go well beyond risk data aggregation. This applied to global systemically important banks (G-SIBs) from January 2016 but it has relevance and is of interest across all financial services firms. The introduction in the BCBS 239 paper states:

"One of the most significant lessons learned from the global financial crisis that began in 2007 was that banks' information technology (IT) and data architectures were inadequate to support the broad management of financial risks. Many banks lacked the ability to aggregate risk exposures and identify concentrations quickly and accurately at the bank group level, across business lines and between legal entities. Some banks were unable to manage their risks properly because of weak risk data aggregation capabilities and risk reporting practices. This had severe consequences to the banks themselves and to the stability of the financial system as a whole."

Regulators have put focus internally on their own data strategy

In June 2019 the United States Office of Management and Budget published a memorandum establishing a Federal Data Strategy for government agencies and asked for comments and suggestions from the private sector. On 7 January 2020 the Bank of England (the 'Bank') issued a discussion paper[2] in which it stated:

"The Bank, through its role of defining reporting across the financial sector, plays an important part in shaping how firms approach their own data. Transforming our data collections would offer an opportunity to support wider improvements to the quality and to usability of financial sector data, for example if this initiative can provide a lever to drive the development and adoption of data standards. Common data standards, widely used, represent a public good with wider benefits than just reporting efficiency, and can support private innovation."

In this discussion paper, the Bank highlights the idea of developing industry-wide standards for common data inputs, and best practice in data architecture with off-the-shelf application programming interfaces (APIs) able to transmit data from a common data utility. This is a long-term vision. However, good data governance is a key stepping stone to achieving these overall goals. The Bank also notes that the heterogeneity of both private sector firms' data and of the Bank's own data needs creates a costly and inflexible process for its data collection. The Bank states that for any given product or transaction, different financial services firms frequently describe equivalent data differently and this makes it hard for the Bank to write specific reporting instructions in many cases. They also cite the bespoke nature of much of the Bank's data requests.

"When the Bank implements a new request, firms may need to go back to underlying source systems, even where that request is similar to an existing collection. To take a simple example, the Bank collects various data on banks' activities in relation to Small and Medium Enterprises (SMEs) for both statistical and regulatory purposes. In these data requests, the Bank employs three different definitions of SME to meet different objectives, including the need to harmonise with wider (non-bank) statistical reporting and harmonised EU capital reporting. Meanwhile, none of these definitions may be the same as the one banks use when targeting and servicing small business customers. For each request, the firm may need to establish a new process to query data in underlying systems on customers' turnover, assets or employees to flag which are 'SMEs'."

Contents

Benefits lie in increased standardisation across different international regulators

The increase in data required by multiple regulators has crystallised the need for stronger data governance and standardisation of how data is handled within firms. There are disparate models for data reporting between regulators and government agencies in any one jurisdiction, as well as significant variance between countries or regions, which lead to inconsistencies, increased costs and requiring detailed mapping. For example, in the context of swaps, derivatives exposure is measured on a notional or risk-adjusted basis. There would be significant improvements in risk monitoring and efficiencies in terms of cost, for both firms and regulators, from more streamlined and rationalised regulatory reporting.

Bringing industry standards into the spotlight

Industry standards have already emerged in many parts of the FICC markets, driven by the desire to reduce costs and mitigate operational risks. LEIs have allowed for the creation of a common, global taxonomy. The Global LEI system is designed to uniquely and unambiguously identify participants in financial transactions. The International Organisation for Standardization (ISO) 17442 standard[3] defines a set of attributes or legal entity reference data that are the most essential elements of identification. The LEI code rests on four principles: i) it is a global standard; ii) a single, unique identifier is assigned to each legal entity; iii) it is supported by high data quality; and iv) it is a public good, available free of charge. A legal entity must publish the obtained LEI as well as the related LEI reference data, e.g. official name, address, country of formation, and date of assignment of the legal entity. Subsidiary LEIs create future opportunities, for example, the ability to include each firm's legal entities in a multi-dealer trade compression exercise.

Given its large size, significant inconsistencies and high processing costs, the over-the-counter (OTC) derivatives market has been a major area of focus for new standards. As well as supporting LEIs the G20 leaders agreed in 2009 that all OTC derivatives contracts should be reported to trade repositories.[4] The 2012 CPSS-IOSCO report 'OTC derivatives data reporting and aggregation requirements'[5] also outlined that "a product classification system would allow regulators to perform data aggregation to monitor exposures to, or positions in, various groupings of products." A Unique Product Identifier (UPI) would give the authorities what they require, or may require in the future, for analysing OTC derivatives products reported to trade repositories.

A recent example of an industry standard that could deliver operational efficiencies and reduce risk is the ISDA Common Domain Model (CDM).[6] This covers how events and processes that occur during the life of a derivatives trade (which may be many years or even decades) are captured and represented within market participants' systems. A standardised set of digital representations reduces the need constantly to cross-check and reconcile trade information and enables firms to develop automated solutions that can be interoperable and scalable. CDM is also being used to support the Financial Conduct Authority (FCA) and the Bank in initiatives such as the digital regulatory reporting (DRR) pilot for derivatives. DRR is a UK initiative exploring the use of technology to help firms meet their regulatory reporting requirements and to improve the quality of information reported.

Contents

Combining eight key components to promote effective data governance and standardisation

There is no one-size-fits-all approach to data governance and controls. However, firms can gain long-term benefits from a focus on the important building blocks and areas of developing best practice in the design, construction, and maintenance of data architectures.

One of the first considerations for any market participant is the degree to which it should pursue a centralised approach to enterprise data architecture for coordinating the movement, enhancement, integration, quality, and availability of data. In principle a centralised approach would integrate fragmented data and streamline the data architecture in the most efficient way. However, for large firms full centralisation can also be very expensive as well as unnecessarily cumbersome at the individual business unit level. An alternative 'hybrid' approach may be a preferable strategy. For instance, a consistent centralised data architecture is important across overlapping businesses that have common clients, counterparties and data needs. It may be less relevant when looking at very different business lines with no overlapping client base. It is also important for firms to build a data governance strategy that can deal with inconsistencies across an ever-wider array of external data sources including other firms, vendors and regulators.

A data governance theme of increasing cross-industry importance is how to control appropriate internal and external use and sharing of data. Data governance must ensure the consistency, timeliness, security and delivery of data. Eight key components to promote effective data governance and standardisation are discussed below.

1. Data lifecycle	Data follows a similar path to a trade lifecycle through creation, storage enrichment, and disposal or retention known as the 'data lifecycle'. The details of this process vary from firm to firm and depending on the products and jurisdictions involved, but a clear understanding of this lifecycle is fundamental to good governance and risk control.
2. Data policies	The foundation of any data strategy is an assessment of data risks, data quality and data policies which govern the partnership between data provider and data consumer. Data risks include rules around data privacy and re-use. Data quality is key to ensuring fit-for-purpose data, to allow speed of execution. Data policies and standards must ensure the data is correct and appropriate for the user, with the latter dependent on the data provider also understanding what and how the data will be used by the data consumer.

3. Data taxonomy

In an ideal world, a common taxonomy of terms and definitions within firms, between firms, with vendors, and across jurisdictions (including between market participants and regulators) would be a goal. But this is an extremely challenging undertaking given the huge disparities in the amount, consistency and ease of use of their data and the data capture and technology infrastructure systems that are deployed to process data. The need for common taxonomies also applies through the markets value chain as there are inconsistencies in data labelling between vendors and their client firms, and between different vendors. Standardisation would bring benefits to business processes leveraging the data internally, as well as external uses such as regulatory reporting.

Many firms have disparate business units often managed independently, or never fully integrated after a merger or acquisition resulting in inconsistent data labelling. There may also be differences between the various functional departments, e.g. front office trade systems versus risk and compliance systems. Increasingly firms are looking to have one controlled vocabulary of data definitions across different business units that can be used by stakeholders to describe their data in a common way to enable a clear understanding of data needs. However, the complexity of legacy infrastructure often makes it more realistic for a firm to start by implementing a consistent taxonomy for new initiatives.

A common example of different data labelling and descriptions, which is applicable across most firms, is customer information. There may be multiple data sources within different parts of a firm that purport to create the same customer data, potentially with inconsistencies in how the customer is described. The same customer may be labelled 'XYZ Corp' in one case, 'XYZ Corporation' in another or merely 'XYZ' and if these records are not linked downstream systems can be misled. There is also the more basic need for consistencies in definitions, e.g. customer information may be known as party data, customer data, client data, or something else in different parts of the same firm.

4. Mapping data sources

An understanding of where data is stored and maintained and how it relates to other data and systems is essential to an effective data governance architecture. The data map should include descriptions of the business meaning of the data, its uses, its quality, the applications that maintain it, and the database technology in which it is stored. Documentation of a data source must describe the semantics of the data so that subtle differences in meaning are understood. This data map is most powerful when documented using the data taxonomy (described previously) in order to understand the movement of a given type of data across a firm. If the data map is created using disparate terminology, it cannot be stitched together accurately.

5. Data movement and lineage

The movement and transformation of data from one system to another can make it difficult to label data accurately or assign data ownership. A data lineage capability provides visibility into where the data comes from and where it is going to within a firm and provides perspective on whether data derives from the authoritative source. This is particularly important in financial markets today when trade data used in artificial intelligence and machine learning applications may have many potential sources. It is important to understand the frequency of movement, how data is transformed as it moves, any aggregation or calculations and the golden source of data.

Ultimately, data should be sourced from a designated System of Record (this is defined as a source of data with a right to change it) that serves as the recognised authority to originate that type of data. Alternatively, data could be sourced from an Authoritative Data Source (ADS) if that ADS has been authorised to re-distribute but not materially change the data. If the authoritative source changes the data, it becomes the System of Record for that data. Where a business derives data in a specific manner (e.g. using averages or a formula to fill in data gaps) there is a need for a methodology outlining how this derived data is created.

6. Data classification

Data classification enables good governance. For instance, who can access the data. It must also define the granularity of data that should be made available to different internal and external stakeholders. Data classification techniques need be multi-layered to work in large complex firms given the scale of their data assets. Access controls need to be supplemented with relevant tagging of sensitive data. One of the benefits of granular access control is that if an administrator modifies a database by adding columns that are not relevant to a particular user, they shall not be visible to that user.

7. Data leakage detection

Data leakage detection and other technology controls can be used to ensure the rules are followed. However, to be effective, these must be supported by robust monitoring processes such as access recertification. The rules and subsequent controls should be applied in a risk-based manner (i.e. the level of protection should be commensurate with the risk posed). To summarise the rules on data protection, it must be clear who is consuming the data, what it is being used for and whether the data is sensitive or not.

8. Data quality

Controls over data quality are a key part of data governance. Firms must not only demonstrate the effective operation of data quality controls but also ensure consistency in processes across different business units. This can involve making sure the most appropriate data is being used for the relevant purpose (e.g. exposure data versus risk data). Moreover, it is important that there are appropriate controls around any data transformation from one form to another in order to ensure data is accurate, timely and complete. When relying on third-party data with a different control environment, firms should ensure that relevant processes are in place to ensure data quality.

A world moving towards a more centralised data strategy

There is a clear need for robust data governance and management strategies across firms who are active in rapidly changing wholesale FICC markets. Market participants are increasingly moving to more centralised data strategies, with significant scope for efficiency gains in terms of longer term cost reduction, risk reduction, and allowing better use of data to drive value for commercial purposes. At the same time, in the near term this can be a costly and complex exercise with centralisation creating most value in overlapping parts of a group in terms of common customers, counterparties and business models. Therefore, the magnitude of centralisation and the specific structures required will vary from firm to firm depending on size, complexity and business models.

The seven sources of data risk outlined in this Spotlight Review do not impact all firms equally, so firms need to undertake careful analysis of their own organisations, systems and processes in order to consider how best to combine the eight key components to promote effective data governance and standardisation discussed above. Moreover, there are some areas that the market can address collectively to avoid duplication and inconsistency of approach. This would help to build transparency and consistency – regardless of organisational and international boundaries – which are so important to ensuring the fairness and effectiveness of global FICC markets.

Data standardisation involves complex problems and potentially significant changes to infrastructure that will not be resolved quickly. As an industry, we must ensure that efforts at standardisation deliver the intended benefits, improving efficiency and effectiveness for market participants and end-users in global wholesale FICC markets.

End notes

1 www.bis.org/publ/bcbs239.pdf

2 'Transforming data collection from the UK financial sector' available at www.bankofengland.co.uk/paper/2020/transforming-data-collection-from-the-uk-financial-sector

3 ISO 17442:2019 'Financial services – Legal entity identifier (LEI)' is available at www.gleif.org/en/about-lei/iso-17442-the-lei-code-structure

4 'Implementing OTC Derivatives Market Reforms', 25 October 2010, available at www.fsb.org/wp-content/uploads/r_101025.pdf

5 OTC derivatives data reporting and aggregation requirements" (Data Report) – CPSS-IOSCO. See www.iosco.org/library/pubdocs/pdf/IOSCOPD366.pdf

6 'What is the ISDA CDM?', 2018, available at www.isda.org/a/z8AEE/ISDA-CDM-Factsheet.pdf

luminous

Design and production
www.luminous.co.uk

125 Old Broad Street, London EC2N 1AR
secretariat@fmsb.com
+44 (0)20 3961 6150
fmsb.com