Transforming data management in the financial markets industry

*Outsourcing as an approach to solving data management issues*
Meeting the data challenge

Data management in the financial markets industry has become increasingly challenging. Financial institutions, exchanges, and market participants have undergone a significant and fundamental transformation. Today, more regulation exists. Instruments have become increasingly complex. Tracking has become more automated. And transactions have become more programmatic. As a result, serious data management concerns have arisen in the operations and information technology (IT) groups of the industry.

Data has become mission-critical and a key enabler of the business model, touching operations practically everywhere in these firms: in the front-, middle- and back-offices. Changes in the financial markets industry have left many firms to deal with costly inefficiencies—from fragmented or suboptimized IT processes to the triple threat of high costs in data licensing, technological infrastructure and human resources.

A data management environment in a financial markets company incorporates several types of data. Reference data contains descriptive information about securities, corporations or individuals. Market data refers to real-time or historical information about prices. Derived data is computed by various calculators and models, and made available to a wide range of applications.

This paper illuminates perspectives on data management drivers, issues and solutions for financial markets companies.
Understanding the drivers behind data management challenges

Fundamental changes in the financial markets industry have created a significant impact on data management platforms. Some of the key drivers of this change are:

- **New instruments**: In the quest to offer compelling products to clients, broker/dealers have created many innovative financial instruments. Currently, there are more than eight million instruments, each requiring that a firm maintain detailed, timely and accurate information. Derivative issues are only one example of financially engineered securities that did not exist just a few years ago. These new financial products and their complex terms have become a challenge for the executives who manage financial information.

- **Changes in market mechanisms**: The mechanisms of trade execution have been altered by the shifting composition of market participants. For example, there has been a rapid increase in the number of hedge funds, and the emergence of mega “buy-side” firms, many of which use program trading and other algorithmic execution models. Decimalization and program trading have led to a reduction in trade size with a corresponding increase in volume. These factors have put a strain on data management platforms because they require the delivery of high volumes of data, with low latency, to black-box trading systems.

- **Regulation and compliance**: Regulation and compliance are also key drivers in the march toward a better data management platform. The emergence of Basel II, Sarbanes-Oxley and other key risk and compliance considerations has forced firms to place a high priority on the production of accurate, timely data to feed internal risk management systems. As a result, institutions must now meet a more stringent fiduciary responsibility to provide correct data to regulatory agencies. Faulty information can result in dire consequences and catastrophic financial exposure.

- **Data aggregators’ expanding role**: The industry’s demand for a wide range of security attributes and pricing information has given rise to an entire sub-industry populated by vendors that specialize in financial data capture and distribution. These vendors are playing an increasingly significant role in managing and providing data. However, managing multiple sources of data creates cost and consistency issues that must be fixed.
These business challenges have created significant stress on existing data management platforms, resulting in marked changes in technology and business processes. An effective data management platform is a critical underpinning for addressing two additional key industry priorities: straight-through-processing (STP) and risk management/regulatory compliance.

Some of the other critical issues facing this industry include governance of access to both proprietary and public data, data synchronization across the industry, management of multiple providers, and the challenge of adapting to new financial instruments in a timely manner. The following diagram highlights many of the business and technological issues resulting from the industry’s changes.
Understanding the challenges of reference data management

As previously stated, reference data describes securities, corporations and individuals. Examples of reference data elements range from basic security descriptors, such as the name of the issuer, to detailed fields, such as optional payments for a call or sinking-fund schedule. End-of-day prices are typically included in this grouping.

IBM's definition of reference data also includes party data and corporate actions. Examples of party data are client names, addresses, delivery instructions and issuer data. Corporate actions data contains stock splits, dividend announcements and other similar information.

Historically, firms have built and maintained their own security and client master databases in isolation from other market participants. As these organizations expanded, organically or through acquisition, data silos emerged that matched each line of business. Most of these data platforms are similar in style and content within and across firms. Typically, they are maintained through a combination of automated data feeds from external vendors, internal applications, and manual entries and adjustments. It is not uncommon for these data platforms to contain aging infrastructures and disparate, highly decentralized data stores.

The industry has seen the concerns of data quality exacerbated by the increasingly international scope of business, as issues of cross-border sovereignty, regulation and currency introduce new data elements, as well as variations in existing data.

Often, the approach to cleansing data is to compare data from multiple vendors or sources. However, this process is quite difficult due to the lack of standards across the industry. Each data vendor has proprietary ways of representing its data, due to the lack of industry standards governing the representation of data. In addition, financial markets firms utilize a variety of formats, including vendor- or exchange-specific and proprietary definitions.
The probability for data error is high: problems with financial data can occur at the original source, at the vendor or in any number of in-house processes. When compounded by the multiple parties involved in the end-to-end execution of a financial transaction, issues of data quality and standardization have had tremendous detrimental impact on the financial markets industry's ability to provide accurate and timely data (see figure below).

Faulty data also impacts a firm's ability to fully implement an STP capability. Inaccurate data will delay the processing of a trade, or worse, cause a “fail” to occur. In the TowerGroup's Reference Data study, it was shown that in the area of trades processing, on average, 16.4 percent of trades were rejected from automated processing routines. TowerGroup further found that 45 percent of the exceptions (trades rejected from automated processing routines) were due to faulty (incomplete, nonstandardized or inaccurate) reference data. In fact, failed trades resulting from inaccurate reconciliation cost the domestic securities industry in excess of US$100 million per year. Although reference data comprises a minority of the data elements in the trade record, problems with the accuracy of this data contribute to a disproportionate number of exceptions, clearly degrading STP rates.
Besides data quality concerns, further improvement is needed around the governance, administration and cost control aspects of the data resource environment.

The fragmented acquisition and maintenance of reference data causes low and inconsistent data quality, difficulty in distributing data on a timely basis, and problems in governance of the data resource—leading to higher cost and risk. The process for acquiring, consolidating and cleansing data must be rationalized, leading the way to a single “golden” centralized store of data that can be accessed by any authorized application or user.

**Identifying market data management challenges and objectives**

Market data is dynamic in nature, and is often referred to as streaming real-time data. Market data is used to establish a security's worth at specific points in time and provides the foundation for developing analyses based on historical pricing. Examples of market data are stock quotes, including bid/ask, last and volume from the exchanges, foreign exchange, and the historical time series representation of this data.

Changes in the nature of exchanges and market participants have created drivers that impact the underlying technologies for market data acquisition and distribution. Increased volume and the demand for low latency are the paramount concerns. The amount of market data available is growing at geometric rates.

Today’s markets are more fragmented—meaning that traditional liquidity centers are split into multiple market centers, requiring larger communications bandwidth to attain high-speed connectivity. Securities are now traded (and quoted) in decimals, executing in smaller average lots with a higher number of bids and offers, resulting in ever-greater volume.

Program trading, once the domain of large broker/dealers, has extended to other financial participants, especially index funds and hedge funds. These algorithmic, black-box trading systems electronically execute trades in real time, often in smaller lots, driving the market quotation and reporting volume to even higher levels. The Tabb Group estimates that, “during peak cycles, top tier prime brokers could be hit with close to 150 trades per second and more than 10 times as many orders per second, imposing a tremendous expense on the applications that must update and disseminate this data across the enterprise.”
To meet these challenges, many IT organizations have received corporate mandates and funding to transition to a market data infrastructure that can achieve certain basic objectives:

- Provide capacity to meet expanding trading volume.
- Reduce latency. Ensure that data is delivered to the end users with the shortest delay possible. This implies millisecond connections to exchanges (and data vendors) and high-bandwidth in-house redistribution to systems and individuals on a global basis.
- Provide for high availability. The infrastructure must be highly reliable, with consideration to BC/DR (business continuity and disaster recovery) provisions.
- Become adaptive to change. IT organizations must provide global market data distribution services that can easily adapt to new financial products on a global basis and quickly respond to changes in trading volume.
- Provide the ability to integrate new products from software vendors and data aggregators.

**Understanding the importance of managing derived data for risk and compliance management**

Derived (analytic) data is created through the use of various calculators and analytic tools. Derived data adds value to raw data by converting it into formats for use in market comparisons, assessments and corporate decision-making. Cash flows for structured bonds, yield curves, risk calculations, performance data, pricing volatilities, correlations, indexing and attribution data are all examples of analytic data.

Areas highly dependent on derived data are risk management, program trading and portfolio analysis. Risk measurement and monitoring, in particular, has become a paramount concern to data management, driven by the compliance requirements of Basel II, Sarbanes-Oxley and other regulatory laws.

Financial engineers are constantly evolving and enhancing their trading models, and the output is often used by many different downstream applications. These analytical routines use both reference and market data; thus, raw data must be almost instantly available at the time of the calculations.
The technology cost to generate analytic data can be high, so developers often store and re-use analytic data, which creates the requirement for highly accessible and well-organized analytic data stores. Typically, this information is in time-series format, covering intra-day ticks or historical index data, over many months, and sometimes years.

There is a wide range of analytics that are based on industry-accepted standard algorithms, and firms strive to generate these computed numbers quickly, accurately and cost-efficiently. There is also a subset of derived data that firms consider to be highly proprietary—with strong competition to develop the better model. Due to the proprietary nature of this analytic data, consideration must be given to restricting access to only authorized users.

Transforming the data management landscape with a managed services approach

Now is the time for firms to take action and rationalize their data environments. Most financial institutions created their current data platforms based on highly decentralized models, which, at the time, appropriately facilitated a rapid and global expansion of new business. These firms have “in-sourced” the skills, expertise and technology to address their data needs. The decentralized and sometimes proprietary approach has been capital-intensive, with no guarantee that the services acquired were appropriate, either then or now. Many participants now recognize that a firm’s data environment may no longer provide a point of significant differentiation.

To meet today’s business challenges, IBM believes that firms should choose to invest in improving their data management platforms. This will allow for a shift to a centralized environment where economies of scale, consistency and firm-wide synchronization of data will bring significant benefit. The process will require the collaboration of all lines of business, especially information technology, operations and finance. It will require significant reengineering with changes to existing manual workflows associated with the data collection, storage and distribution processes.

Transforming a firm’s traditional data management system and related processes is a complex undertaking, requiring significant data management expertise and experience. Developing and executing a data strategy must be performed with a focus on eliminating the risk associated with change, reducing investment cost and providing an appropriate return on investment (ROI) over the entire life cycle of the project. For reference data, the initial focus will be on data acquisition and cleansing. For market data, low latency and high reliability will be the key considerations.
A managed services approach with a centralized topology provides a logical and cost-effective means of creating a data platform to support a firm’s current and future business requirements. This approach shifts the burden of technology and staffing to an outsourced environment, allowing the institution to focus on its core business competencies.

However, transforming to an outsourced model can take time. Many firms will choose a phased approach along the continuum to a managed services end-state and may find it attractive to continue with an “in-sourced” approach. Clients can still achieve savings and increased functionality by filling short-term gaps—and retain a pathway to a longer-term outsourced model.

A managed services approach can be extended to a model for the sharing of outsourced components across firms. A relevant example would be a shared generic data-cleansing capability. It could be offered to multiple participants, potentially resulting in lower costs and higher-quality data. Such an offering would not alter the licensing and ownership of the content, but would become a shared mechanism for acquiring, scrubbing and consolidating data from various sources. IBM endorses this concept and we are seeking firms to join us in this discussion.

Outsourcing may very well become the de-facto method by which financial markets firms address data management needs and, at the same time, enable the reallocation of resources to apply to revenue-generating activities.

IBM has extensive experience with data management engagements and can help clients develop a roadmap to address their needs today—with options to transition to an outsourced solution in the future. IBM offers knowledgeable business consultants, access to our infrastructure solutions, and software experts to help clients address their data management needs. Some of the initial activities can involve projects such as our assessment and strategy review, technology infrastructure improvements, and the evaluation and redesign of manual workflows to improve the data-cleansing activities.

For more information
To learn more about IBM Business Consulting Services, contact your IBM sales representative, or visit:

ibm.com/bcs