The Data Requirements for Systemic Risk

FSTC Innovator • January 25, 2010

Michael Atkin, Managing Director, EDM Council Inc

The recent turmoil in the financial markets has focused global attention on the importance of managing systemic risk. The rationale is clear. Our global financial system is complex and interrelated and the fragmented nature of our regulatory environment presented an obstacle to effective oversight.

There has been no shortage of analysis on why we experienced financial turmoil. And while an assessment of underlying causal factors is important, it is not the subject of this article. Instead, my focus is on the data needed by regulators, market authorities and industry to enable them to engage in complex process of systemic risk analysis. My hypothesis is straightforward. Unique and precisely defined data are the building blocks for all forms of regulatory oversight and the key to creating the linkages and interconnections that are necessary for evaluating systemic relationships. We must get the underlying building blocks right – and we must have trust and confidence that they are both fit-for-purpose and immediately comparable if we are to have trust and confidence in our risk analytics and resulting policy decisions.

Precise data is one of the critical factors of input into regulatory oversight, is the prerequisite for managing complexity and unraveling essential inter-linkages. And I'll maintain that this "data content infrastructure" is neither complete nor precise enough to support the objectives of systemic oversight.

Data Precision versus Systemic Risk Analysis

It is important that we first understand the difference between performing systemic analysis and ensuring that the data factors of input into systemic analysis are comparable. Systemic analysis is a complex problem and the data needed will vary according to objective. Andrew Lo of MIT presented a simple mnemonic for capturing the range of information that a systemic risk regulator will need to monitor and described it as the "four L's" of leverage, linkages, liquidity and losses across the global financial system.

And to keep tabs over these four L's, regulators and firms will sometimes need volumes of detailed data from myriads of firms so they can make judgments about when firms or market segments are overleveraged, when asset bubbles are growing, when exposures are becoming correlated, and so on. The regulators might wish to examine risk concentrations, unusual escalations in asset prices, the velocity of transactions, variances of credit spreads, or the difference between mark-to-market values and those produced from theoretical models. And they might want to look at how such things are interconnected for clues about changes in stability or problems with counterparty risks and exposures.

The truth of the matter is that the monitoring of systemic risk is not a "thing," it is a process. The goal is to enable regulators and market authorities to make assessments of the interrelationships and interdependencies of many factors so as to determine the impact of specific events on the overall financial system and to gauge the likelihood of cascading failure.

If we contrast that definition of systemic risk (a daunting task) to the data requirements that support systemic risk, the problem becomes more manageable. We don't have to define every scenario that a regulator might want to evaluate or understand every factor that might go into a macroeconomic model. That is (appropriately) the task of the regulatory process.

What we have to do is ensure that the data the regulators need for their models and analysis is defined in a consistent manner, identified using precise and unique identifiers and accessible via a common reporting format. That way the regulators can collect what they think is relevant for their specific analysis, specify it at any degree of granularity, be assured that it represents what they expect it to represent, and have confidence that it is comparable to what every other firm is reporting or using as the basis for their models.

Implementing this "data content" infrastructure is at the foundation of supporting systemic risk analysis. It starts with the unique and precise identification of instruments and legal entities and includes the use of a common semantic language to describe the financial products and deals that drive our industry. And these "tags and identifiers" need to be assigned at the beginning of the financial transactions process – at issuance – so that we free ourselves from the reconciliation, cross-referencing and transformation process that characterize the way the financial industry manages data.

It's all About the Data

Precise data content is the lifeblood of the financial industry. Financial institutions and their data suppliers acquire the underlying content from hundreds of sources including prospectuses, term sheets, corporate filings, tender offers, proxy statements, research reports and corporate actions. They load the data into master files and databases so that they have access to information on prices, rates, descriptive data, identifiers, classifications, fundamental terms and conditions, credit information, etc.

Participants in the financial transactions process use this information to derive yields, valuations, variances, trends and correlations. And they feed the raw material and other derived works into pricing models, calculation engines and analytical processes as it gets passed from process-to-process among multiple global participants.

In order to understand and comply with all their legal, contractual and reporting obligations, financial industry participants must understand all the ownership hierarchies, counterparty and supply chain relationships. They have to keep up with all the changes and corporate events that occur every minute of every day for millions of entities around the world. The data is linked to accounting, trade execution, clearing, settlement, valuation, portfolio management and risk systems. And the results of all of these activities must be reported to clients, shareholders, analysts, regulators and market authorities.

The business of the financial industry is data precision intensive and all of this content is the core factor of input into every business process within every financial institution. It must be complete, accurate, and consistent (at the nuance level) in order for firms to have trust and confidence that it is 'fit for purpose' for all models and applications. And it must be comparable so that regulators and central banks can rely on it to provide oversight over complexity and guard against overly risky business practices. Without clean, accurate and comparable data, both financial institutions and regulators will continue to spend an inordinate amount time and money cleaning, cross-referencing, mapping and reconciling the data before they can use it to reliably analyze, connect, automate and monitor the complexities associated with our global and interconnected financial system.

Why Data Gaps Exist

Technological innovation has always been a driving force within the financial industry, but it is proving to be a double-edged sword. For years, data management was purely a technology issue. Data was collected and stored in centralized mainframes. Access was limited and ownership was clear. In the time of the mainframe, the 'single version of the truth' existed and was owned by the firm. But as technology evolved and PCs hit the desktops – access to and creation of - data became the task of many. Instead of data being managed as a centralized corporate asset, every department claims ownership - creating multiple overlapping silos of data. As a result, data is disparate and inconsistent. The explosion of 'data freedom' has evolved into the dilemma of 'data anarchy.' The ability to look at an enterprise view of performance or risk is now extremely difficult. Data management is no longer just a technology problem. It is now a business, organizational and cultural problem – with cascading consequences due to the interconnected nature of financial processes.

In other words, there is a serious data quality gap in the global financial system. It is the result of years of mergers, acquisitions and internal realignments. It is exacerbated by business silos and inflexible IT architectural structures. And it is hard to address because it is intertwined into systems and applications as it gets passed from process-to-process among multiple global participants.

But data gaps are not just a problem brought about by financial institutions. The process is also confounded by an inefficient and fragmented chain of information supply. Here's an illustration just to put the issue into perspective. Much of the data that resides in the master files of financial institutions is factual. It originates in some form of legal document and was precisely created by lawyers and accountants. So it was perfect when it was created. Then multiple vendors independently acquire the content, transform it to fit into their source systems and rename it using internal terms, definitions and nomenclature. Independent business units of financial institutions then source it from multiple vendors and transform it to match their internal environments using their own tags and identifiers. As a result we have content that is non-comparable and not precise enough to promote trust and confidence that it is fit for purpose for data intensive applications.

The Data Building Blocks

So let's look at the nature and status of this data foundation. Virtually every business process within financial entities and every regulatory analytical objective is constructed from four broad categories of data:

1. Instrument reference data including precise identifiers at the instrument level, comparable documentation of the legal structure of financial instruments and a definition of the contractual obligations associated with the complete transactions process.

   The problem with instrument reference data is twofold. First there are gaps in the identification of financial instruments. And while the industry, working through their national numbering agencies, has done a good job of identifying listed issues, we have problems with identifying multiply listed instruments and we are missing common identifiers for derivatives, short term paper and loans (among others). The second challenge is due to fragmentation of the data collection process and the lack of standard naming conventions for instrument reference data and the corporate actions that are used to maintain the data.

   These two problems are solvable and work is underway to address these shortfalls. For the instrument identification issue, national numbering agencies, industry utilities and

vendors need to work together under regulatory supervision to implement the missing identifiers and to reconcile the commercial challenges associated with their usage. For the data definition challenge, groups such as XBRL (for financial reporting), the EDM Council (for reference and market data), ISO 20022 (for transactions processing and corporate actions) are making significant progress on data semantics. There is clearly work to do to integrate these dictionaries into an industry wide standard at the semantics level and regulatory engagement is required to compel issuers to tag their legal and contractual documents at issuance. All one has to do is observe the work of the Securities and Exchange Commission in managing their interactive disclosure initiative as a model moving forward.

2. Entity reference data including all legal/business entity identifiers, documentation of ownership structures and definitions of all business and counterparty relationships throughout the transaction process.

   The problem with entity reference data is also associated with the lack of precise identification – this time in the identification of the business and legal entities that have a role in the financial transactions process. And while the industry, working through SWIFT and other transactions processing utilities, has done a good job in identifying entities that are involved in the exchange of money, some such as issuers and guarantors don't have standard identifiers.

   The solution to the entity identification problem is close at hand. Both ISO and SWIFT have recently been cooperating on the definition of the missing entities and on the process of identifier assignment. Regulatory support and encouragement for this essential requirement would go a long way toward expediting implementation by global financial institutions.

3. Pricing and valuation data including exposure (and comparability) of the calculation methodologies for off exchange and thinly traded instruments. There is little problem with exchange traded price transparency. The challenge relates to the consistency and comparability of calculation methodologies, probability estimates, implementation of "as of" dating conventions and transparency on the derived data components used to drive model based pricing.

4. Positions and transactions data to provide a consistent summary of settled and unsettled contractual agreements where the financial entity has a legally binding role. Specifying the positions and transactions data to be collected is the purview of regulators. The only real challenge is to recognize this information as business sensitive and to manage it as such.

Conclusion

Systemic analysis is a complex process that will be enhanced by ensuring that the underlying data factors of input are trusted as comparable and fit-for-purpose. Systemic analysis is a regulatory objective while data consistency and comparability are shared among regulators and industry. The pathway to addressing these data challenges are standards for the unique and precise identification of instruments, entities and data attributes combined with better management of the data chain of supply.

The good news is that these requirements are known and agreed to by almost all data practitioners, activity is underway and it is possible to fix the data dimension of systemic oversight in the short run without unreasonable cost. The bad news is that the financial industry will not likely solve these issues on their own due to a combination of short term orientations, lack of understanding at the top of the organizational chart, the difficulties in gaining alignment among multiple stakeholders and the silo nature of financial entity operations. Regulatory leadership and compulsion is required to implement the data content infrastructure that we all need to assess and manage systemic risk.