

Data integration or bust

One clear issue that has emerged in the recent meltdown of the markets is the massive importance of being able to understand and communicate the value and risk of the assets in a mortgage portfolio.

Roger Chi

Both Northern Rock and Bear Stearns failed largely due to investor confidence. Both Northern Rock and Bear Stearns failed largely due to the drainage of investor confidence. When this confidence was shaken, the institutions were not able to quickly show the investors their true position because they themselves did not have the systems in place that could give them a consolidated view of the bank. This problem is especially critical for large complex structured assets such as commercial real estate where there are many details, any one of which can have a significant effect on the risk profile.

To accurately assess the portfolio, institutions must be able to quickly gather the different pieces of data that reflect each deal. The problem is that for most banks, the data is typically fragmented, in paper files, or in data-islands like desktop spreadsheet files.

Traditionally, the task of generating a portfolio risk report has been put in the hands of a small group of analysts. They trawl through the disparate data sources, gathering information, and calling around the departments to make requests for information when data is missing or seems corrupt. In the last resort, when the data is too difficult to find, they make

simplifying assumptions. This process can take several days or even months. The resulting reports are out of date and are not fully trusted by management, the investors or the regulators. This article first describes the typical current data environment and then some of the approaches can be used to quickly consolidate the data for reporting and risk assessment.

Data set

Risk assessment for commercial real estate has some of the most demanding requirements for data. The full assessment of a deal can use hundreds of data inputs including the financing structure, interest rate hedging, covenants, property characteristics, tenant quality and lease structures. Once such a clean consolidated data set has been created for the portfolio, it can be used for many purposes such as reporting stress tests to regulators and investors.

At first glance, it may seem that all relevant data should be readily accessible. Most institutions will have a central banking system where primary loan data is held including outstanding loan balances, interest structures, and repayment schedules. However, if the operations span multiple geographies, each office may have its own "central" banking system that only captures the local deals.

In addition, detailed loan data such as covenants aren't usually held in the central banking system, existing only in the paper loan documents usually. Similarly, portions of loans that are held by other banks, for example, junior or senior tranches are probably not captured by any of the in-house systems and reside only on paper or in spreadsheets.

Information on interest rate hedges and derivatives is typically in a separate treasury system. Collateral information for properties and, more importantly, for the income-generating units and tenants are typically only found in (often inconsistent) spreadsheets, on paper, or not kept track of at all.



Core principles

Given that the data integration problem is multi-faceted, any potential solution must be a very flexible system. In tackling these problems we identify seven core principles:

- 1) Provide a central structured repository for data
- 2) Where data exists in other enterprise systems, bring it in to the central repository
- 3) Where data exists in desktop files, provide the capability to easily upload them into the data structure
- 4) Where data is in paper files, input the data manually with validation
- 5) Use a “data blending” algorithm to selectively update sections of the data structure
- 6) Design the data input and cleansing task to be incremental
- 7) Immediately provide reports that are of value to the users

Data consolidation

The central data repository acts as a focus for the data consolidation exercise. It must encompass every piece of information that impacts the underlying risk of a deal. The structure of the data must reflect the underlying structure of the assets, for example, it must be able to expand to cover properties, each with multiple tenants, and each of those with multiple lease review conditions. So far as possible the structure should be exhaustive, but there will always be new data items to be added so a flexible paradigm such as XML is used to structure the data. With the new database tools, XML also provides a solid basis for data extraction and reporting.

There are two main reasons to use the data in existing enterprise systems. The first is to improve user acceptance avoiding the need to waste time re-entering data. Most of the people keying in this data are primarily employed for other high value tasks and additional data entry keeps them from focusing on their primary business. The second reason is for optimizing data integrity and minimizing the operational risk associated with re-keying. If the data in the system were the same as used to feed the reports reviewed by credit committee,

then there would be a good chance of it being complete and correct. Any data that is available from the live banking systems should be pulled dynamically into the central repository. This can be done with data transformation services or other tools.

Desktop files such as spreadsheets can be brought in to the central repository by giving tools for users to browse, select and easily upload them into the data structure. By avoiding re-typing, large amounts of fragmented data can be easily consolidated with relatively little resistance from those who have the data. Ultimately there will always be some data that is not yet in electronic form. This must be manually entered, but using a system of checks that make it clear when fields are missing data, or that validate any new data being entered.

Updating data

Having gathered the data, there needs to be a systematic merge process for updating it. Here a program can be used to automatically blend the enterprise system data with the more static user-input data. Such a blender is difficult for programmers to create because it must flexibly overwrite certain parts of the data structure without deleting manually entered data. It must also be flexible enough to adapt as the enterprise systems grow over time to capture new information.

The last two points are less about technology and more about the practicalities of implementation in a dynamic organization. The data consolidation exercise should be designed to be incremental so that the organization does not have to wait until everything is complete before getting useful results. For example, the data structure should be designed so that users can first enter gross information such as total annual income, then later, for the most important deals, they can go into greater detail, entering detailed

lease structures for the biggest tenants. When time permits, eventually they can enter the detail for each individual tenant.

With such an approach, management can get a quick picture of the overall portfolio and identify hotspots that are worth more detailed investigation. Users will be much more motivated to fill in complete, accurate data if they immediately get something useful in return out of the system, such as a credit committee report populated with deal information that saves them precious time. In this way, there is benefit for both the portfolio managers and those who have the data on individual deals.

The data consolidation exercise should be designed to be incremental so that the organization does not have to wait until everything is complete before getting useful results.

This market is punishing those institutions that cannot show that they have a clear understanding of their portfolio. Data integration can literally save institutions, and if it is done correctly, the data integration task can be much less painful than the alternative. MFG ■

Roger Chi, Risk Integrated, a consulting and technology firm specializing in risk measurement for commercial real estate.