

Inside Reference Data

December 2007 Volume 2 Number 8

www.irdonline.com

The Marvel of Metrics

While experts struggle to attach a value to the cost of non-cleansed data, industry networks are looking to find new ways to measure risk and performance by using tools generating metrics to help justify investment in reference data initiatives. **Tine Thoresen** reports on how an international bank has piloted a risk model to measure and compare data quality

When managers pose the question “what have you done for me lately?” data management professionals often complain it is a struggle to explain how they have added value. The problem tends to be that they are not armed with the kind of concrete figures available to the sales team.

But this situation is about to change for one international bank, which has used an operational risk model for performance measurement. After a project aimed at identifying breaks in processes, the firm will have a data quality index to present to senior executives.

Peter Hughes, managing director of UK-based system provider Arc Best Practices, who has a background in operational risk, has, in collaboration with Allan Grody and Robert Mark, conducted research into measuring firms’ exposure to data-related risk. The international bank was the first to pilot Arc’s data version of the risk model, with the outcome being that the concept worked.

The Arc Model

The model is based on a total calculation of risk exposure – manual processes, automated processes and how they interact. To identify the exposure, a firm would first look at the total size of what it was dealing with. The



“Raw data is transformed into useful information for managers and executives so they can understand trends in risk and performance and receive alerts when exceptional events occur”

Peter Hughes, Arc Best Practices

next step is to establish factors that can lead to risk. “In data management, causal factors would be quality management, people, business recovery, policies and procedures, technological risk, process controls and monitoring information,” says Hughes.

These would all be put into a template and mapped with a score attached. The question is: what do I rely on most for data quality? according to Hughes. If quality management has a higher impact than policies and procedures, quality management would have a higher score.

In terms of people, the template would look at overtime and temporary staff, for example. If each person in the team does more than 80 hours overtime per month, the score would be 0 out of 100, as stressed people make mistakes, explains Hughes. If they do less than 20 hours overtime per person per month they are less likely to make mistakes and the data quality will be higher, meaning the score would be 100.

After attaching a score to all the factors, the project would go on to look at the reference source and business-critical fields. “If the industry code is a business-critical field, you put a weighting on it,” he says, adding that this would identify if the negative outcome would



be severe or critical. A wrong payment code could lead to losses of hundreds of millions of dollars, while a wrong title of an individual is not necessarily business-critical.

In other words, each of these business-critical fields will be mapped to the process it is used to support – processes that have a value. What values relate to that business-critical field? This will help firms attach a value rating to each field.

The outcome of the process will now be the total size of the repository. The scores can be put into a calculation tool, with the metrics then being presented on a dashboard. By multiplying the business criticality with the value that those critical data fields have, the exposure to risk can be measured in a weighted value of the firm's susceptibility to data-related losses.

"The raw data is transformed into meaningful and useful information for managers and executives at all levels of the firm so that they can understand trends in risk and performance and receive alerts when exceptional events occur," says Hughes. The raw data is first transformed into a common value-bearing metric and then consolidated and aggregated to present alerts and views of status and condition at different levels of the firm. The dashboard was developed by Arc technology partner Business Objects and presents the views on one screen. It allows users to click on trends or indicators to get more information and 'drill-down' into the data through the various levels.

In addition, the pilot firm completed one more step. "We ran a check from our master

file to our system of variance of similar proportions to ensure that yes, Arc is right," says a director at the bank, explaining that early implications suggest the information that came through from Arc is "pretty bona fide."

Overall, the whole process took three to four months and the bank will have the results before the end of the year. There were around five staff working on the project on an ad-hoc basis and "it wasn't as labour-intensive as you would think."

This result means firms can start comparing data quality both internally and with their peers. The first bank that has piloted this model has compared repositories in two different divisions, with the outcome being a measurement of the quality gap. "The only way to convince senior executives is if you can give them the metrics," says Hughes.

This will also be the outcome of the pilot for the international institution. A director at the firm says he would expect to produce an executive dashboard to indicate the flaws in business processes and give an indication to senior board members on where there are inadequacies in the processes.

Measuring in Real Time

But it is not enough to measure it once. Following proof of concept, firms would license the method from Arc and train staff to use it, meaning the client could produce metrics on an ongoing basis. The basic setup would be to regularly monitor causal factors. "I would suspect it would be an ongoing usage," says the director at the pilot institution. He

says the first step was setting the template, but once it has been created, the firm would maintain it "because there has to be ongoing measurement to see how these metrics change."

Yet the usage of Arc will depend on a firm's risk exposure. "When you start getting that information, chances are senior executives will press to get that information more frequently," remarks Hughes. Large organizations could potentially automate the metrics, measuring data quality in real time, but this would be a big and expensive project, he says.

And to achieve automation, there is still some work left to do. There is currently no software available for the method to generate metrics in real time. "At some point in the future, we will look to build software with someone," says Hughes.

But the model could potentially also have even more usages in the future. Measuring performance of data vendors, an area Arc has not explored yet, could be of interest at some point. The problem with applying the model to vendors would be that they supply data without being concerned about its usage. "Probably, we would have had to do a bit more thinking [to apply the model to vendors]," he says.

Originally, the Arc model for measuring risk in live operating environments was ready around five years ago, but Hughes says the industry was not interested at that point. While credit risk, determining the probability of default, has been measured for decades, few have invested in measuring operational risk in data management.

EDM Council

The interest now has been triggered partly by support from members of the Washington DC-based industry association EDM Council. Mike Atkin, managing director, EDM Council, says measuring data management risk is in concept everybody's objective. "They want a way of quantifying the business implications of poor data management or good data management on their operations," he says, adding that senior executives want to know "what's at stake if we do nothing, if we act now, if we have bad data or if we have bad data processes."

"If you look at the various components of how data relates to business practices, you've got to break it all down to a granular level and craft best practice, because that's how you measure yourself"

Mike Atkin, EDM Council

While this has been done in other parts of the financial industry, using a risk model such as Arc for measuring operational risk is new to data management. It is not common for firms to measure risk and performance in data management, but many would assess risk, according to Hughes. Assessment, generally relying on opinions rather than calculations, could reveal that a firm considers its data to be high, medium or

low quality. "If you can measure something it is much more meaningful and powerful," says Hughes.

Atkin says the EDM member firms are all active in the three areas of metrics that he thinks are important—measuring data quality, performance implications of bad data and operational metrics—but they are only doing pieces of it. At the moment, some firms are working on a workflow view, following data through its complete cycle and trying to quantify implications step by step, and most firms are doing internal operational metrics, looking at service-level

expectations, turnaround time, setup time and time to market. "They're all trying to figure it out," says Atkin.

Yet measuring risk of data management can be a difficult task, and members have asked the EDM Council to facilitate discussion on the topic. "The first thing we're going to do is get them all in a room," he says, remarking that this should be a good starting point for firms to share their experiences.

But Atkin has also been instrumental in amending the Arc model. He started a conversation with ARC's Hughes about the possibility of extending the Arc methodology for operational risk in the data management industry. Atkin invited several contacts to review and discuss the software and later helped create the best-practice statements and the weighting scales. "If you look at the various components of how data relates to business practices, you've got to break it all down to a granular level and craft best practice, because that's how you measure yourself," he says.

Market Index

And when everyone agrees on best practices, the model can be used to measure the market. If a number of firms decide to use the Arc model, the EDM Council could create an index, enabling it to report back to the industry on where the group is together on data quality and risk. "I think it all depends on evaluation by the financial institutions whether they agree with my assessment and then getting a number of them to participate in this kind of activity," says Atkin.

In essence, he says the model is about "quantifying the business case in real terms at a senior management level that would be trusted and valid." Firms typically want to find out how much risk they are absorbing by not doing EDM, by not having good data quality and by not having good data practices. The Arc software is another way of putting supportable evidence on the table. It allows you to have a metric that is justifiable and can enable firms to better understand which projects to prioritize. "You've got to have a way of quantifying things to make those kinds of decisions," says Atkin, adding that although everyone in the industry thinks it is important, professionals want to make sure this can be verified to executive management.

And since one firm has expressed satisfaction with its piloting of the Arc model, several members of the EDM Council are now seriously interested in the software, according to Atkin. He says he has not seen other similar systems, but if they exist, the EDM Council is not "beholden in any way to Arc" and it is an open question whether Arc is the right mechanism or not.

The only complaint from the pilot user was that "the method is very much geared for operational risk, and we had to refine it to use it for our particular purposes." He says it could be a tool for the future and time will tell if it is the right model to use.

But so far Arc has passed all the latest tests. "This, in my opinion, is worth investigating. This one looks viable," concludes Atkin.

Interview With Allan Grody, Executive Consultant at Arc

What is driving the increased interest in measuring risk in data management?

Until recently, regulators have been interested in two areas of risk—market risk and credit risk. Along comes Basel II and another bucket arrives—that bucket is called operational risk. Data creates a lot of risk in financial institutions and people were not able to measure the size of exposure to faulty data.

What is operational risk?

Operational risk is a function of the interaction of manual processes, applications and data. When this interaction fails, it creates operational risk, when it works it produces an efficient, relatively riskless straight-through processing environment .

Why is it important to measure risk in data management?

For too long people have been saying we've got a problem, but for senior management they ask: How big is the problem? Risk managers have asked operational people: What is the probability that you're going to have a loss? But in operations, they've focused on fail rates.



They are now being forced to think in risk management terms.

How can firms solve this problem?

The Arc tool comes up with a risk unit and assigns a value to the risk unit. The value is

tied to both the process of that operation and the financials of the firm.

These risk units can be aggregated up to the top of the firm as well as associated with granular activities and loss events at the departmental level.

What can firms achieve by using the tool developed by Arc?

Here is a way in which we can value a central global database against siloed warehouses and literally get a temperature reading on both. People saw that EDM made sense, but they couldn't measure its value. This can help create a business case.