The reference data revolution:
How firms can leapfrog legacy technology

The
Realization
Group

EUROMONEY
TRADEDATA

The reference data revolution: How firms can leapfrog legacy technology

It's a classic dilemma, where companies find themselves with one foot in the past and one foot in the future. Financial services firms are confronted with the challenge of managing reference data in legacy systems and simultaneously introducing new technology that can better identify and manage risk, improve operations and support new strategic priorities. The technology for delivering and handling reference data has advanced significantly – slow, clunky batch files are no longer the only way for firms to keep their reference data accurate and up to date. Yet many companies are stuck with embedded legacy systems that can lead to costly data problems, resource-draining maintenance issues and missed business opportunities. Is there a blueprint for delivering and handling market reference data that can allow firms to take advantage of new technological solutions without massive change management programmes?

*In this Financial Markets Insight report, we speak with industry experts on the frontline of the reference data battle. How can companies move ahead to ensure they have accurate, timely and accessible market reference data? To find out, we asked experts from data suppliers, technology providers, industry figures and some of the people directly responsible for keeping their firms' reference data pristine and easy to access. We hear from **Mark Woolfenden**, Managing Director at Euromoney TRADEDATA, **Matthew Cheung**, CEO of ipushpull, **Krishna C. Nadella**, Global Head of Solutions at Symphony Communication Services, **Anupam Bansal**, Head of Change and Product, Instrument Reference Data at JP Morgan Chase, **Choon-Teck Lee**, Business Project Manager at UBS, and **John Macpherson**, Deputy Chair of the Investment Association Engine panel. They offer a unique collection of perspectives on best practice, while also describing a world where leapfrogging into the future need not be as daunting as people might think.*

Introduction: The importance of glue

Let's face it, reference data is generally not thought of as among the more exciting aspects of a company's trading operations. **Mark Woolfenden** of Euromoney TRADEDATA jokingly referred to it as "the poor cousin to market data". But in the same breath, he said, it is also "the glue".

That glue has business impacts in a variety of ways. In terms of cost, issues with reference data can cause broken trades and force companies to devote resources to identify and address the sources of problems; or they can require slower-than-desired workflows. In terms of revenues, how a company deals with reference data can make the difference between launching a new service in a timely fashion and not being able to launch one at all.

Many financial companies do recognise they have an issue when it comes to reference data. One recent survey of senior executives who oversee data operations across the sell-side showed that more than half of all Tier 2 and Tier 3 banks (53%) reported that their current set-up was inefficient. Tier 1 firms were in a better position as they have the scale and resources to devote more time and investment into tailoring their systems to their needs. But even some of them faced issues, with 14% reporting inefficient operations.

At the same time, it can be difficult for companies to articulate the value of reference data systems. "Proposing a new system has to be paid for," Woolfenden says. "And when everybody's still chasing margins and profit, you have to have a really good case."

To make that case, firms need to understand the full dimensions of the problems that old-fashioned reference data processes cause. They need to consider the growing demands that are being put on their systems. And they need to understand that addressing these issues ultimately is not only about data, but also about workflow.

But there are solutions. A rising number of trading companies are finding success by moving away from traditional methods of distributing data, using an app-based approach to data consumption and more fully embracing the cloud.

> *"Proposing a new system has to be paid for. And when everybody's still chasing margins and profit, you have to have a really good case"*
>
> **– Mark Woolfenden, Euromoney TRADEDATA**

Batches and patches

A good number of the problems start with the way reference data gets delivered: in batches. **Matthew Cheung**, who runs real-time data sharing and workflow platform ipushpull, says that historically, data solutions have been based on once-a-day batch files. A data vendor takes a snapshot of a database and sends it via FTP to the client, typically at the end of the day.

"The bank is getting this massive file. They have to write their own code and scripts to take it and put it somewhere else and do stuff with it. It's a very, very big, clunky process that hasn't really changed in 20 or 30 years," Cheung says.

The systems can be slow. About the best a company could hope for, in terms of intraday updates, would be snapshots every half hour. "When you start trying to get faster than that, it just starts falling apart, because it was never built to do that quickly," he notes.

The speed factor can then morph into an accuracy issue. Reference data is changing all of the time – not at the microsecond or nanosecond level like market data, but much more frequently than once a day. That means that at any given moment, some portion of the reference data that is sitting in a bank's database may become out of date and the situation may not get rectified for hours.

Concerns about accuracy do not stem only from the speed of data updates. A large amount of reference data is unstructured or not standardised, making it potentially more prone to data errors. "There's a lot of context switching between different applications," Cheung said.

Cheung says that virtually 100% of people who handle reference data use Excel in their day-to-day work. A typical scenario: "I need to extract something in my system via CSV, open it in a spreadsheet, do something with it, save it back as a CSV, upload it back into another system, copy and paste some of that into a chat with someone internally or the counterpart. All the time you're doing that, you're application context switching."

> *"The bank is getting this massive file… It's a very, very big, clunky process that hasn't really changed in 20 or 30 years."*
>
> **– Matthew Cheung, ipushpull**

The number of spreadsheet mistakes only compounds when there is context switching. "You've got an operational risk of you just making a mistake or missing something. Plus, there is the mental fatigue, when you're jumping from one thing to another," Cheung said, noting there have been studies showing that human efficiency suffers when people need to switch tasks routinely. "All of that plays into the idea of moving towards a more integrated workflow, where the data is at your fingertips when you need it in the right application."

System complications

A big obstacle to moving towards an integrated workflow, however, is legacy technology. Adjusting a process that, on the surface, makes complete sense, can quickly become problematic. Just how big an issue is legacy technology for a bank looking to make changes to its reference data systems?

"Massive," says **Anupam Bansal**, who is responsible for Instrument Data and serving several internal lines of business at JP Morgan Chase. These businesses include Markets, Security Services, Asset and Wealth Management and corporate functions such as Risk and Compliance.

The assumption is that once there is a source identified for the data, it can be brought into the bank and seamlessly consumed by all. "One of the biggest challenges we face is if you have legacy systems which have limitations."

"What it means is that you then end up with band-aids or manual processes to massage and manipulate the data to make it consumable by the legacy systems. But this accommodation could then end up creating breaks for other systems and applications".

"One of the biggest challenges we face is if you have legacy systems which have limitations"

– Anupam Bansal, JP Morgan Chase

The maintenance of legacy systems and technology is also a major challenge. The resources that can manage the legacy technology become hard to find and any change to a legacy system becomes an expensive proposition and eats into the budget meant for strategic change.

Meanwhile, Bansal said his team has encountered in big increase in demand for instrument data over the past few years as this function has become formalised. Instrument data has become a central utility and it is now a component of all initiatives, including new products, new business, new or changing regulations, compliance requirements or risk reporting demand. The requirement for intraday and real time data and completeness has also spiked as data has been identified as a key element of end-to-end processing and lack of it will break the process. Given all these demands, having to additionally accommodate requirements from legacy systems becomes a huge overhead.

Aside from this increased demand, the complexity and scale of certain demands also pose challenges. "If you look at MiFID II, the scope and size of the regulation meant that even our vendors were struggling to interpret the requirements and determine what the product offering should be."

Not every regulatory change is as far-reaching as MiFID II but Bansal said similar situations can arise for other regulatory initiatives or any initiative. Lack of clarity in terms of data requirements will impact the time to market for a solution. Add to that the complexity of ensuring the solution will work for legacy systems and it gets really painful.

Over at UBS, **Choon-Teck Lee** is responsible for sourcing reference data. He too sees legacy technology as a major drain.

"A lot of legacy systems still exist. And then sometimes client vendors merge, and they inherit a newer system, and they want us to move to that system and we're not ready." Lee said vendors may be happy to wait, but sometimes it can drag on for a few years. He sees cloud transformation as the answer. "If everything goes up to the cloud, I hope that this problem might be reduced, or at least become more manageable over time."

Go with the flow

The increased demand that data experts such as Bansal noted has been a long time in the making. **Krishna C. Nadella** of Symphony Communication Services linked the evolution of reference data practices to major changes in the way capital markets operate. One such change was the rise of big data, which gained momentum in the 2010s.

Companies began to rely increasingly on data scientists and created roles such as chief data officer. "These were people who were coming out of business schools, who had more of a computer programming background," Nadella said. They were people who understood how to find trends and make the most of data. For them, the key issue was access. Companies wanted to move beyond a terminal-centric way of operating. They were saying: "Give me the data. I don't need the confines of the house it lives in. Let me house it how I see fit."

John Macpherson, deputy chair of the Investment Association Engine panel, said data science was a term that barely got used in the industry until about seven years ago. "Now it seems that if you haven't got data science in your title or your job description, you're missing something." He noted that reference data, if firms drill down into it, can provide immense amounts of colour which can influence trading decision-making.

What many firms still do not always fully appreciate, however, is the way that reference data is not something static that sits apart from other parts of the operation. There are constantly people working with it both in the back and front office. Put simply, making it accessible within workflows represents a new way of working for most companies.

"That's the new flavour of the decade," Nadella said. "We have individuals who understand workflow, but they also understand data and they understand how it all fits together. And so, they don't see front office, they don't see middle office, they don't see back office. What they see is, how is my workflow affected? And how is this ecosystem operating at peak efficiency?"

According to Nadella, those firms that see data and workflow as part of the same coin will have an edge in capital markets in the coming years. "It's the workflow that's the real focus in this decade. How are we attacking the clients' workflow to allow them to be able to do the same job function but in a more efficient automated fashion, and ensure that the data is able to flow?"

Many firms do recognise that they need data science skills in their organisations. "But the resources are limited," Macpherson added. "I think there's definitely a limitation in terms of the right people with the right experience."

> *We have individuals who understand workflow, but they also understand data and they understand how it all fits together. And so, they don't see front office, they don't see middle office, they don't see back office. What they see is, how is my workflow affected?"*
>
> **– Krishna Nadella, Symphony Technology Group**

To improve reference data workflow, a company first needs to understand where the blockages occur. Key determinants for what constitute pain points start with the form of the data and the delivery method.

In terms of form, Nadella noted that some clients want only raw, unstructured data and find that any structure that has been added compromises it. He used the analogy of an uncut diamond. A good jeweller will take away as little as possible. "Ultimately, you still want the integrity of that diamond," Nadella said.

In terms of delivery, Euromoney TRADEDATA commissioned recent research that showed sell-side brokers were being asked to provide reference data in real time. "It's really about getting the delivery right and access to data repositories or sources of data that can be instantly available as and when they needed in real time."

That is exactly what Cheung said ipushpull does. Effective data distribution, he says, is about being able to make data available at the right time and place, for the right person and in the right application. It involved access controls, permissioning, audit and various enterprise-heavy features, while also giving the end-user a way to interact with that data.

*"Now it seems that if you haven't got data science in your title
or your job description, you're missing something."*

– John Macpherson, Investment Association

Bypassing legacy technology

Are there ways that companies can address their data pain points and create business-centric workflows without completely overhauling systems and bringing in skillsets? This is where technology provides such as ipushpull and Symphony, and data suppliers such as Euromoney TRADEDATA, are seeking to build new operational solutions.

Some companies have system configurations that go back decades. At the same time, they are constantly introducing new components into their systems to ensure their processes are up to date without undertaking major overhauls.

"A lot of that may have been also forced by the pandemic, and the need to really work remotely. The division has been reinforced by what has happened in the past 12 to 18 months," Woolfenden said. "Various companies, such as ourselves, who represent key parts of the trading, clearing, settlement lifecycle trading operation, are having to gear up but also to think holistically how it how it all works, and how our solutions complement each other."

In fact, financial firms need not think about a major system overhaul to modify their reference data operations and workflow. "The agility of the technology has enabled firms to move incrementally now, and to patch around big legacy systems to get the functionality that if you were designing it from scratch, you would get with a lot fewer miles of pipes in the works," Woolfenden said.

He likened it to a natural process where a predatory plant wraps itself around a tree and eventually there is little left of the tree, which slowly dies away while the plant surrounding it continues to live. "You can plumb and replumb around the big legacy systems until such time as they are not working, or rather, they're working on a very minimalist basis, and then the switch is less onerous."

This is typically done via APIs and distributed data models. API-based solutions are not just about increasing efficiency. Cheung noted that when firms build licensed reference data into products, API technology can automate specific data sets to feed into downstream systems that generate products. "All this is interoperable. That's the way the industry is moving. We've seen more momentum in that in the last 12 months than years before."

Data security is also enhanced. The data itself may sit in the cloud, where it is encrypted. And then all access is permissioned and can be audited.

In terms of the value attached to more streamlined, workflow-based data systems, the main benefit comes from not having to devote hours to discover why something did not work. "That's just wasted time. And if you, if you added all of that up for everyone who works in back office in all the banks, that adds up to a very sizable amount," says Cheung.

A good example of where reference data and workflow come together is in the resolution of broken trades. For instance, firms could build a workflow that allowed secure chats to be created without context switching to address broken trades, with an audit trail incorporated.

By giving the end user direct access to the data, with whatever controls the firm wants to establish, it means the user has a richer set of information to work with. Cheung said app-based technology – as opposed to the old way of receiving reference data – meant firms could build workflows around the data. For example, changes in reference data could set off alerts that appear in chat windows.

The new direction of travel

What is key, however, is that firms recognise that app-based delivery and consumption of data, hosted in the cloud, is the way of the future. Nowhere is this truer than for the Tier 2 and Tier 3 firms, where the pain points are most acute.

"The second and third tier firms get squeezed because the overhead on certain costs of participating in the market is fixed. There are some big overheads, in setting out your reporting processes and compliance processes, that a bigger firm can amortise and spread the cost more easily," Woolfenden said. "You're carrying a disproportionate overhead relative to your order book, and that's a classic margin squeeze. Therefore, they have to be really clear in creating and sustaining a lowest cost operating model."

At the same time, he noted that second and third tier firms are less wedded to legacy technology and in general think in terms of being agile. There will still be a cost to change, but not at the level where they need to justify a massive switch. They can think in terms of Excel add-ins through to APIs and GUIs.

The vision, what Woolfenden called "the direction of travel", is all about eliminating wastage in the marketplace. For many firms, that means making a break with the past and embracing data on demand. "What don't we need to use from the historical way that the industry provided data services? That's what occupies most of our focus," Woolfenden said.

For more on-demand content please visit
www.therealizationgroup.com/insights/financial-markets-insights/

Financial Markets Insights from The Realization Group, is a series of interviews with thought leaders in financial and capital markets. The purpose of the series is to provide exclusive insights into industry developments, through in-depth conversations with C-level executives and key experts from banks, exchanges, vendors and other firms within the financial markets ecosystem. For more information, please visit
www.therealizationgroup.com/insights/financial-markets-insights/

The Realization Group is a full service marketing and business development services company specialising in the capital markets. Our team contains industry practitioners from both the trading and post trade disciplines and we have expertise equally in the on-exchange and OTC trading environments. We apply our comprehensive set of marketing programs and wide-ranging media and business networks to complement the business development requirements of our client organisations. For more information, please visit
www.therealizationgroup.com