Inside Reference Data

October 2011 www.waterstechnology.com/ird

Enterprise Data Management

Special Report

Sponsored by

Staying ahead of the game

Inside Reference Data | Awards 2011

Best Corporate Actions Data Provider

Best Reference Data Initiative (Vendor)

WINNER

Get complete data for your automatic processing needs from SIX Telekurs. Know the latest stock splits, company mergers and dividend payments. Our award-winning corporate actions will keep you ahead of the game.

SIX TELEKURS

www.six-telekurs.com

Editor's Letter

EDM Comes of Age

This issue's hefty Virtual Roundtable features robust discussion by representatives of investment management firms and service providers both large and small. Participants confronted the influence of the "big data" buzzword in the EDM space, and present a picture of when and why firms choose to use EDM rather than more mature data management strategies.

EDM projects are becoming enterprise-wide, says Tony Brownlee of Kingland Systems. "The support for EDM in the past three to four years has shifted from divisional or department and become much more top-of-the-house focused." David Thomas of Barclays Capital echoes this, identifying implementation efforts and their relative speed as dependent on the maturity of data management teams. Those whose strategy depends on significant system improvement will need more time, and are clearly creating stronger services, standards and controls, he says.

"The foundation of any data management strategy and its implementation continues to be a strong governance and operating model," he says. "This enables easier stakeholder communication and buy-in and is critical to funding and implementation decisions." Barclays has invested a great deal over the past two years in data management platforms, and has seen first-hand the results of doing so: faster implementation of system changes and quicker responses to stakeholder needs and changing environments, according to Thomas.

However, challenges remain, and the panelists discuss ways to achieve an accurate data management plan that is fast enough, and definitely faster than those "mature" data management strategies. One thing they agree on: Everyone will have to do more with less.

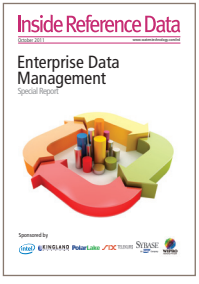
Yours sincerely,

Michael Shashoua
Editor, *Inside Reference Data*
Email: michael.shashoua@incisivemedia.com
Tel: +1 646 490 3969

Contents

FEATURES

NEWS



An incisivemedia publication

recycle
When you have finished with this magazine please recycle it.

Incisive Media
32–34 Broadwick Street,
London, W1A 2HG

EDM Council Adds Corp Actions to Repository

LONDON—The EDM Council has completed ontology work necessary to add corporate actions to its Semantics Repository, and plans to liaise with the Object Management Group (OMG) to focus on making it a financial industry business ontology (Fibo) standard.

The Council planned to present the final corporate actions ontology work to subject matter experts in September, to then compare with the OMG's processing and messaging of corporate actions. "Although the ontology is now fully complete, corporate actions will not become one of the Fibo ontologies for a while," says London-based Mike Bennett, head of semantics and standards at the EDM Council, and recently appointed vice chair of the Financial Domain Task Force, one of the groups that oversees creation and approval of OMG standards. "We are looking at how to make use of other OMG standards of process modeling as part of the corporate actions standardization model."

Meanwhile, the Council has completed a proof of concept that will be presented to the CFTC using Fibo, "to demonstrate that instruments can be described semantically with absolute precision," says Washington, DC-based Mike Atkin, MD at the EDM Council.

The full version of this story appeared in *Inside Reference Data* September 2011.

Carla Mangado

SEI, EDM Council Prep Final Version of Data Maturity Model

LONDON—The EDM Council has completed the first draft of its Data Management Maturity (DMM) model and expects to release a final version to the industry by 2012.

In 2009, The EDM Council and the Software Engineering Institute (SEI) of Carnegie Mellon University launched an initiative to create a DMM model with an auditable methodology. Currently, it consists of six core categories including strategy, operations, IT and platform, data quality and policies and procedures; 15 components, including areas such as data management goals; 36 process areas; 18 standards and procedures areas and 200 capability measures.

Washington, DC-based Mike Atkin, MD of the EDM Council, says: "We are going through each section and adding some explicit goals for each."

The full version of this story appeared in *Inside Reference Data* July 2011.

Carla Mangado

UniCredit Dives Into Data Governance Strategy

MILAN—UniCredit plans to establish a greater role for data governance within its firm and involve the business side of the company in proceedings to a greater extent, according to Milan-based Alberto Ricciotti, head of global modeling and warehousing in the CFO data governance department.

The path to create a data governance strategy is far from smooth, and enterprise-wide support is considered a key factor for its success. Ricciotti explains that the road to securing buy-in from top management remains one of the more complex challenges. "The first barrier you are inevitably going to face is finding the sponsor for the project," he says, adding it is essential for data governance efforts not to be viewed in isolation.

In 2008, UniCredit unveiled an initiative to roll out and manage a data warehouse, which went live two years later, to support data processes within the CFO environment. "When building up data governance, you have to think it is not the only solution. You need management to sponsor your view and support the data governance structure. Inevitably you need to gain user community confidence and support," explains Ricciotti. "You need to put in place in your data governance team a long-term approach."

The full version of this story appeared in *Inside Reference Data* July 2011.

Carla Mangado

Perpetual Selects Cadis EDM Platform

Perpetual Investments, which manages investment funds exceeding $27 billion in fixed income, domestic and global equities, has chosen Cadis as its central data hub. The Cadis EDM platform will be used to manage securities, pricing and the processing of positions and transactions, replacing Perpetual's internally built system.

Mark de Courcey, general manager, investment administration at Perpetual, says Cadis has improved the company's ownership and management of its data.

EDM Council, OMG Team For Semantics Standardization

The EDM Council and international standards organization Object Management Group (OMG) have set up a semantics joint working group to establish the strategic, technical steps and deliverables necessary to maintain the semantics repository as an OMG standard. The group has been divided into three workstreams covering the disposition of the standards, the implications of semantic modeling and re-use of existing standards, and the necessary technical modeling framework.

Virtual Roundtable

Enterprise Data Management: Speed, Harmony and Compliance

Inside Reference Data gathers leading industry professionals to discuss the crucial issues affecting today's EDM environment

What types of enterprise data management projects do you see firms working on at the moment?

David Thomas, Global Reference Data, Operations, Barclays Capital: In the current market environment, I see data management teams following all other teams and simply working on mandatory regulatory projects and RTB activities. Although there may be some funding available for system or process improvements, this is almost certainly going to be considerably reduced in 2012.

Clearly, Dodd-Frank related work on the LEI, UPI and USI will continue as a priority, and I expect Fatca to become more of a focus in Q4 2011 and into 2012.

Daryan Dehghanpisheh, Director of Financial Services Sector, Intel: From Intel's vantage point, the "big theme" is "Big Data." This has the potential to become another industry buzzword like "cloud," although the benefits of these projects are almost immediately seen by management and profit centers of financial firms. Big Data projects are being evaluated for both market data and reference data elements. The nirvana is a near real-time capability to detect and predict trends and opportunities for profitable positions and trades. The challenge, though, is that most of today's infrastructure from servers and storage through application stacks isn't built to handle the size or speed that financial firms deal with. Some solutions deal with one or the other challenge, but there doesn't seem to be a full solution to handle both.

Tony Brownlee, Managing Director, Data Solutions, Kingland Systems: We see firms that have less mature data management strategies focusing on planning and data governance projects. Those firms that already have more mature data

management strategies in place are more focused on enterprise data management projects. The big change we see is that these projects are now really becoming enterprise-wide. The support for EDM in the last three to four years has really shifted from divisional or department and has become much more top-of-the house focused. In addition to EDM in general, there is a lot of focus on the LEI.

The recurring theme for all data management projects is that the key drivers are regulation, risk and operational efficiency. I'm yet to find a project without at least one of these drivers.

John Randles, CEO, PolarLake: Many firms are focused on getting their data management infrastructure ready for a new era of regulation, risk management and transparency. Whether it is price processing, security master, corporate actions or feed on-boarding, firms are concerned that existing infrastructures cannot scale to today's demands, not to mention new regulations and business requirements coming down the track.

Michael McMorrow, Head of Data Management Services, Enterprise Information, AIB Bank: The main theme I see at the moment relates to enterprise data "consistent reliability"—both with regard to data definitions and data quality. Data enters the organization through a plethora of channels, systems and structured/unstructured formats. Generally, data is particularly fit for the purpose of the first operation/process that uses it and has probably sponsored its acquisition. The challenge of the past few years has been the demand to see data across the organization as a holistic set from which to drive enterprise management information/business intelligence reporting for internal and external stakeholders. This is driving more concentrated data management focus on ensuring the super-set of data is integratable (based on consistent data definitions) and dependable (based on understanding the highest level of data quality demanded by all enterprise users, which may be a higher level than demanded by 'first user'). These overarching, front-to-back data management initiatives are technically and politically complex and require sponsorship by top-tier management.

Daryan Dehghanpisheh,
Director, Financial Services Sector, Intel
Tel: +1 905 842 7543
www.intel.com

Tony Brownlee,
Managing Director
of Data Solutions,
Kingland Systems
Tel: +1 641 355 1000
www.kingland.com

Ivo Bieri, Head of Strategic Business Development, SIX Telekurs: Centralizing data and harmonizing data across different areas continues to be a big issue for firms. Recent developments in the legal entity identification space have fuelled further interest in data management, and even before all the details about the new LEI are released, firms are looking at preparing for the implementation, mainly for risk purposes.

In addition, data management projects are focused on timeliness. One area that may be particular for our business is the opening of new securities. There is continued pressure to decrease the turnaround time for new openings of new securities.

This is also the background for our recent launch of VDF Flash, which delivers data about new securities in a matter of seconds. Firms require faster delivery of data, and they want the data to be updated more regularly. Tomorrow is no longer good enough.

Peter Blenninger, head of software development and quality management, Baader Bank: Internal [security] master and cross-referencing between different asset classes, such as corporate actions in underlyings and effects of options and futures.

Sinan Baskan, Vice President of Global Financial Services Solutions, Sybase, an SAP Company: Firms have two main areas of focus right now for data management projects. One is financial market data repositories and reference data consolidation. The other is risk and trade transaction data for accelerated enterprise risk and internal governance purposes. Firms are concentrating on these areas to help keep them out of difficult situations with the regulatory bodies and to eliminate blind spots on their balance sheets. These types of projects will improve transparency and help them meet compliance regulations and survive the increased scrutiny. The overarching goal is to avoid anything like what happened at AIG and Lehman Brothers. What's interesting is that firms are realizing when they make these improvements, their internal processes are becoming more efficient and they are able to identify profitable transactions earlier.

Jennifer Ippoliti, Practice Director, Data Management, Wipro Technologies: The theme of the moment is business entity and client data. It would be difficult to find a firm that is not working on some sort of entity data initiative right now. Why? There are four main drivers. Some firms are preparing for Fatca or LEI implementation. Others are going through extensive integration initiatives due to acquisitions. Some firms are working toward achieving a single view of the customer for risk and regulatory purposes, while others are investing in their customer data to facilitate cross-selling and improving the customer's experience.

Data quality, especially of entity data, is also a top priority for financial services firms this year, because short-term investments in data quality tools and projects can deliver big returns down the line. Expected results of data quality projects include reduced rework and smaller exception queues, more productive operations teams, and reduction of hardware and software license costs, as well as more accurate data.

What can be done to facilitate a speedy implementation of a data management strategy?

Thomas: Clearly the implementation effort and relative speed are largely dependent on where the data management teams are on the maturity curve. Teams that have a strategy with a dependency on significant system improvement or replacement work will have a longer lead time than those with a more robust platform and are focused on developing stronger services, standards and controls. Barclays has invested significantly (both cost and resources) over the past two years to ensure the data management platforms across product, pricing and client data are scaleable, flexible and enable fast development timelines. This work has enabled faster implementation of system changes and quicker response to stakeholder requirements or changing control environments.

The foundation of any data management strategy and its implementation continues to be a strong governance and operating model. This enables easier stakeholder communication and buy-in and is critical to funding and implementation decisions. Like many industry participants, the remit and scope of data management systems in Barclays has moved to a group-wide service model, and the governance complexities associated with this need to be addressed with a robust model.

Dehghanpisheh: Achieving the ROI from Big Data analytics requires unique human capital: someone comfortable with business intelligence development, programming, data mining and visual analytics.

PolarLake

John Randles, CEO, PolarLake
Tel: +353 1 4491010
www.polarlake.com

Outside of particular skill sets, the people have to be highly capable problem-solvers who can easily shift their thinking from one approach to the next. There is a particular advantage in having those who are both developers and domain experts.

The first thing to do is take stock of what requires improved speed and what doesn't. Moving from historical analysis to prediction with data sources means applications and software must go from spotting conditions pre-specified to spotting patterns automatically. That activity starts with quickly finding what you are already looking for, which becomes quickly finding the things you weren't looking for, identifying connections between anomalies and providing deep context to explain the behavior the system is identifying. This is as tough as it sounds, because it involves a paradigm shift in how you think about the problems the firm is trying to solve.

To speed this process, Intel worked with a customer to build a platform for Big Data analytics, using systems based on current deployments and experimental approaches to refine components where the customer had less experience.

We measured results of the experiments in terms of performance, and there were clear winners in approaches. We found a way to load and store 200 billion rows of data at faster than 100gb per hour, and also found a way to retrieve analytic results from the data in less than 30 minutes.

The data sources stored in the platform include pattern query data of 3 billion rows a day, logs of 900 million rows a day, and web traffic logs of 300 million rows a day. Any of these could be considered a Big Data challenge.

Brownlee: One of the first considerations is for firms to understand data quality limitations up front. If they skip looking at data quality, it bites them later on and leads to delays and re-work, which no-one wants.

Second, firms should now be aware the tools are out there that can really help implement a data management strategy. Software has really evolved, and particularly in the Master Data Management (MDM) space, the solutions can accelerate implementation by well more than a year versus typical in-house build strategies.

Randles: Agreeing clear and precise business objectives and KPIs in areas such as compliance, risk management and operational improvement for data management projects are the prerequisites for any speedy successful program. Once these are agreed, the obvious danger is being led on a three- to five-year walk in the desert by a vendor's own agenda and abilities. These endless projects can end up with little to show after mammoth cost and effort, resulting in a demoralized team and cynical business users. Working with old technologies will keep EDM projects working to these timelines. So my advice for speedy implementations is firms should not work on technology and a data management methodology developed more than five years ago, as it was designed for a different problem in a different era. If they do, they will be missing out on some of the biggest advances in data management technologies since the relational database was developed, such as the application of semantic web technologies.

McMorrow: The potential scale of effort related to data management strategies can frequently appear to be daunting. Organizations are rightly fearful of their ability to land mammoth projects over extended timelines due to real risks such as evaporating momentum, stakeholder churn and ultimately unrealized return-on-investment. Beware Homer Simpson-style fatalism: "Trying is the first step to failure." This may seem obvious, but there are three simple questions to ask when planning a data management strategy: 1) Where is the world at? 2) Where are we at? 3) Where are we hurting the most? This surfaces the priority items to work on speedily (with natural buy-in by business stakeholders) but also shapes each solution towards the 'strategic ideal.' So, you must define "The Big Strategy" but deliver in small steps based on priority need. We've been saying this sort of thing for years, but, the main practical challenge is making sure the steps really do aim at "The Big Strategy" and don't degrade into cheap-cheerful-but-disjointed 'tactical' initiatives.

Bieri: When I look at the data management initiative we are running at the moment, I think an important factor to succeed is to think big but start small. If you can show tangible success early on, you will get buy-in from the wider organization and be able to continue the project.

There is a risk that the business changes during the course of the project—priorities might change, or the management team might change, for example. If you're two years into a project and you're unable to prove tangible success, the business may pull the plug on the initiative.

Ivo Bieri, Head of Strategic Business Development,
SIX Telekurs
Tel: +41 58 399 5418
www.six-telekurs.com

Blenninger: Invest in people who have a long experience in raw data management and have profound knowledge in reference data, and give them time to work on the projects.

Baskan: In terms of program management, start by identifying a small project and rebuild it completely. Don't develop a big data management plan with 10 subprojects; start small. Create a successful prototype that addresses a subset of the overall problem. For instance, if the equity derivatives infrastructure needs to be revamped, a first step is to launch a fully funded pilot that rebuilds the next-generation infrastructure from live data capture to settlements and clearing, including risk controls and aggregated reporting. A pilot like this makes more sense than re-building an infrastructure across the trade life cycle for all asset classes and geographies. The scope is more manageable and when complete, the pilot can be used as a blueprint for repeated success in other lines of business. The return on investment for this type of project can be infinite, especially if a firm avoids penalties or even worse, closure.

Firms are increasingly aware of the importance of ensuring golden copy data can be consumed by downstream systems. What are the most important factors to keep in mind when assessing mechanisms for delivering cleansed data to consuming applications?

Thomas: Control, control and control! The increasing trend towards centralized data sources and stronger distribution layers within the architecture adds further pressure on the data management teams to get quality and accuracy "right first time." Implementation of a framework that focuses on system-based, proactive controls should ensure high levels of accuracy are achieved at the point of capture. This removes or reduces the effort required for reactive reconciliation or exceptions processes, and should also prevent the requirement of multiple "amendment" messages flowing down to consuming systems.

Maintaining a standardized, single distribution layer is also an advan-

WIPRO
Applying Thought

BUILDING THE BANKABLE FUTURE

SIMPLE.SAFE.INTUITIVE.

Wipro Banking and Financial Services

Wipro's Banking practice has partnered with over 50 of the world's leading banks including 9 of the top 10 US banks and 2 of top 5 European banks along with leading banks in Asia Pacific. Our expertise across retail banking, enterprise banking, investment banking regulatory risk and compliance, mobile banking, channels, payments, lending & mortgages and wealth management address transformational changes for our customers. Our solutions on cloud. mobile, social computing and analytics enable our customers to build a bankable future.

DO BUSINESS BETTER

NYSE:WIT | OVER 120,000 EMPLOYEES | 54 COUNTRIES | CONSULTING | SYSTEM INTEGRATION | OUTSOURCING

SYBASE An SAP Company

Sinan Baskan, Vice President, Global Financial Services Solutions, Sybase, an SAP Company
Tel: +1 1800SYBASE5
www.sybase.com

tage. With an increasing focus on reference data across organizations, system connectivity can become more complex. Mandating a single distribution mechanism and forcing downstream consumers to connect to this single point eliminates the need for tailored feeds and the significant effort and resources required to manage them. With the integration of Lehman's business platforms and group-wide focus, the centralized client database in Barclays now feeds 300 downstream consuming systems using a single distribution layer—managing multiple feeds and requirements from such a diverse consuming system base would be impossible.

Brownlee: This is an area where technology can help. A lot of people think MDM is all about creating a golden copy but it's probably more about distributing data to downstream systems, including legacy systems.

The fact the golden copy data is better doesn't necessarily mean it will work in legacy systems, though. To make it work, firms would often have to account for unique data models or rules inherent in the legacy systems—the data is what makes these systems work, and must be accounted for.

It is also about looking for the right expertise. There are vendors in the industry—and we're one of them—who know how these projects are done and how to "see around the corner" to avoid problems and pitfalls.

Randles: The most important factor when looking at different delivery mechanisms is looking at how well thought-out this mechanism is, or if it is an afterthought to an old EDM platform or a generic piece of ETL technology. The afterthought to a platform is quite dangerous, as it trivializes real challenges that lie ahead, and typically distribution is not part of the DNA of the supplier. This is often only discovered very far into a project, when it can be too late to repair. Generic technology can lead to large expensive builds with the client having to maintain the solution afterwards themselves. Look to a mechanism that supports a superset of all your known distribution requirements today and one that has been proven at scale. Also learn

from the vendors about what is now best practice and don't just try to replicate what you have been doing historically.

McMorrow: The term "cleansed data" implies it has been dirty in some way to start with. Your energy must be on ensuring only clean data enters your organization and that you subsequently don't make it dirty. The more complicated aspect is when you manufacture new data—for instance, when you transform original base data under some business rules that map to some prescribed target data definition. The golden copy concept is useful for holding original data in a single location for efficient access by downstream consumers but is vital for holding certified, official, single-instance results of transformations to eliminate the risk of conflicting transformations being performed independently by downstream consumers.

Bieri: The architecture is really important. If you have a serialized data environment, you lose depth of data in every step. If you have a data warehouse, you can avoid this. Introducing such a centralized environment means you may have to change more interfaces, which can take time, but it pays off later.

In a serialized environment, you face increased dependencies, and growing data volumes will create a bottleneck, which will affect everyone using the data. Instead, a 'data warehouse' design that serves many downstream systems in parallel can provide firms with high performance and access to the full depth of data.

Michael McMorrow, AIB Bank

Blenninger: Do it yourself. Choose modern tools, and work together with the vendors of the consuming applications.

Baskan: A common mistake made when generating golden copy is to overlook how dynamic the data capture process is. Because this process is so dynamic, the "golden copy" is better described as a state of data or a stable image. It is better to think of it as a process of building and maintaining a "record of reference" in support of a specific application that is "true" in a specific time window. The time window may be in minutes or in days or weeks, but the temporal relationships between the pieces of information across a time line must be maintained and managed for the application cycle.

Ippoliti: Data delivery to downstream systems is one of the most challenging

Jennifer Ippoliti,
Practice Director, Data Management, Wipro Technologies
Tel: +1 732 421 2781
www.wipro.com

aspects of data management. At a minimum, consider these four factors:

- **Format:** Will the data be delivered as a flat file? A message stream? Will you deliver the full universe of data, or just the deltas? Can it be rolled back if a job fails or restarted mid-file if a file becomes corrupt?
- **Frequency:** Overnight batch or intraday? If the latter, how often? How often do the downstream systems update (the golden copy should be on the same schedule, or close).
- **Performance:** The system must be able to handle large volumes of information. Can it deliver the data quickly? Will your overnight batch take more than one night to run? Do you have wiggle room if a vendor file is delivered late?
- **Flexibility:** The golden copy can potentially be disseminating data to thousands of applications using diverse technologies, some proprietary and/or legacy. Can your golden copy system accommodate all its customers and their vastly different requirements?

With maturing data management operations, many organizations are focusing on alignment of reference data processes across silos and regions. What should be the priority for a firm trying to bring it all together? What is the best place to start?

Thomas: I would start with ensuring you know two things:

1. What data is held (and where)?
2. Who "owns" or maintains the data?

A consolidated data dictionary—even if this is captured across a number of systems—will identify the breadth of the data services offered and will also start to identify overlap or duplication. Understanding the ownership or governance around each of the data sets (and down to attribute level) will effectively provide the current state governance model – even if this is "de facto" and not formalized. This can then be followed up with documentation of who consumes the data sets and the risks or issues associated with them.

Analysis of this consolidated information set should clarify issues and strengths and should drive the longer term strategy for consolidation into a global, cross-functional service.

Brownlee: The best place for firms to start is with data governance. Data governance can be at a policy-level. It can be as simple as setting standards that will be used across the silos and regions and independently starting to use them. Then you can take the policy-level down to the authority-level. You can assign data stewards with responsibility for data in specific silos or regions. Without this level of governance, people often assume it is someone else's responsibility.

A robust data governance framework can help a firm advance its data management strategy pretty effectively. Ideally, you would have an orchestrated series of reference data process, and you can become a lot more orchestrated just by improving governance. By starting with these key steps, you can improve data quality within 12–18 months.

Randles: A great place to start is on feed on-boarding. This can allow firms to create a real operational, live inventory of all consumed data, and add value very quickly. It can also be done rapidly (a matter of weeks) providing access to a new breed of consumers such as the front office and risk analysts. Subsequent projects like price processing and security master can be easily extended on from that initial feed on-boarding phase. This is also an invaluable inventory management resource for market data managers, typically responsible for the commercial arrangements for all third-party data.

McMorrow: The best place to start is to define your real 'problem/opportunity statement' and to assess what moves can realistically be made in the context of your particular organization structure (political and operational).

For instance, a move towards rigid alignment depends on assessing the effort versus the benefits of changing existing systems and processes within your organization. If all are fundamentally operating in a similar way with similar data, then there are probably significant benefits to be gained from rationalizing operations under central functions, or systematically aligning divisional functions. Conversely, there may be enough business, system and data differences across the silos and regions for rigid alignment to require major re-engineering spend and/or cause significant business disruption.

The full set of reference data across your organization is extensive and probably ballooning à la Moore's Law. A main question to ask, regardless of centralized versus divisionalized processes, is should you have a one-size-fits-all approach for aligning reference data across your organization. Do you need to apply exactly the same standards, processes and

David Thomas, Barclays Capital

alignment to all reference data within the terabytes/petabytes stored across your suite of operational data stores? One approach to consider is to formally categorize your 'enterprise critical reference data' and to apply the most rigorous and integrated data management focus to this set. This should be a relatively small subset of data items and should be of feasible scale for the organization to enforce advanced, aligned data management strategies (data definition, data governance, data profiling, data scoring, data retention, data lifecycle processes, rich metadata etc). If you adopt this approach you must have a formal framework in place for ongoing review of the 'enterprise critical reference data' list and for ongoing oversight of corresponding adherence to the data management strategies.

Bieri: I cannot provide a general recipe. It's very individual, and typically people start by addressing a pain point. We see a lot of firms focusing on counterparty data, and others on enhancing STP rates for corporate actions and instrument reference data, but I can't see one solution that will fit all organizations.

Blenninger: Start with something easy. That is probably the area where your company is already very successful.

Baskan: Firms should be consolidating their data repositories and capturing and filtering data as close to the entry point as possible. The challenge is that the data sources are increasing. There is more market data, asset pricing information, trading venues, even newsfeeds from Dow Jones and Reuters that must be captured to accurately assess trading and risk decisions, as well as post-trade and positions management. This larger volume of data increases both the software platforms that need to be maintained and the number of user communities that need to be serviced.

Ippoliti: The key to success in process alignment is to have a dedicated, independent team focused on the activity. An unbiased approach is key when dealing with stakeholders in different divisions or regions who are resistant to change. Start by establishing that team, and understand that it will be a full-time job. Have them do a current state assessment: how bad is the problem? Then, consider the underlying architecture and data flows. Some processes may not be standardizeable because the applications and systems they are based on are different. Finally, pick a few quick wins that

will get you the success and buy-in you need to carry the program forward.

One approach many firms have found successful is to outsource or offshore data management operations. This forces firms to take a hard look at their processes. A detailed process assessment is a must prior to outsourcing or offshoring, and typically includes a gap analysis and identification of "outlier" processes. Training an offshore team is a complex undertaking, so firms must minimize the number of processes and exceptions to processes that are being moved. Therefore standardization becomes an important by-product —as well as a major benefit—of moving processes offshore.

When assessing processes for standardization, don't overlook processes that take place outside the core reference data world, but which impact reference data—consider risk officers, credit analysts, loan operations and other back-office areas as possible areas for standardization and improvement.

What more can vendors do to help meet changing client requirements in the EDM space?

Thomas: I would recommend three key components that I would look for when evaluating a vendor service:

1. Relevance or linkage to existing industry best practice or data elements—LEI being the obvious current example
2. Clear articulation of the value the service adds to a data management offering—over and above an internally driven alternative
3. Concise documentation of the control framework they have in place and how this is maintained

My recent engagement with vendors in the reference data environment indicates they are ahead of the game and recognize these components are critical to their services.

Dehghanpisheh: The results [of the experiments mentioned earlier] concluded that identifying patterns using Big Data analytics can be further helped with hardware improvements, a unique set of design principles, and better tools than are currently available from many commercial vendors. What Intel has found in designing solutions is:

1. Parallel tasks wherever possible
2. Avoid disk I/O as a bottleneck (keep it in memory)
3. De-couple the event signal from the processing of the event

Our involvement with the project demonstrated a successful architecture for a Big Data problem. Combining optimized hardware and software with a novel design for loading and querying can create a solution to mine 200 billion rows for pattern analysis. This represents

two months of data critical to proactive risk management for the customer. The design can be duplicated for other Big Data problems with similar success.

While the promise of Big Data analytics is there, Intel believes this work points to the need for innovation with both software and hardware vendors. Many of the requirements of Big Data analytics are outside the BI tools of today.

A platform-independent implementation of parallel database querying needs to be developed so companies can better integrate current investments with emerging Big Data investments. Solid-state drives should play an important role in Big Data analytics, but their current capacity is a limiting factor.

Intel is helping lead the way in providing a way to effectively manage parallel computing from the client applications to the server applications. Without these improvements, Big Data analytics will be slowed by limited features, prohibitive costs, un-optimized technology, and progress by trial and error. This work is important to Intel because Big Data analytics focuses on every element of a system. We believe Intel architecture can be a key differentiator in selecting components of a Big Data analytics solution.

Brownlee: As a vendor, we always have to continue to invest in products and capabilities. We are investing in data quality and the data management maturity model—both areas industry association the EDM Council is doing a great job driving. Vendors need to work with the industry and their clients, and they need to invest. Our current activities are aimed at improving our offerings, while helping the industry in general. Vendors that survive the long haul and have been around for 20 years, such as Kingland, know this is important.

Randles: Across all aspects of data management, the key issue in 2011 is time to market. The rate of change of business requirements, the focus on data management, the upcoming deluge of regulation and the unrelenting focus on cost means speed is of the essence. Data management operations, more than ever, will need to do more and more with less. This focus on cost is in an environment that has never had such unrelenting scrutiny from internal business sponsors and regulators. Innovation and the application of new technologies, which are well aligned to the business requirements, is the approach that will deliver more with less.

McMorrow: The enterprise data management space is huge and covers a diverse network of functionality. While it can be attractive from an administration

and licensing perspective to deal with a single vendor for the complete stack, it is implausible for any vendor to be the best-in-breed on all fronts. Some enterprise data management functionality is quite new, with room for further evolution. Niche vendors tend to be best at driving new thinking forward with their dedicated focus and their flexibility to explore a concept independently.

Clients require the freedom to choose offerings from multiple vendors and expect vendors to operate openly together. A client may choose one vendor for its core ETL tool, a different vendor for complex data profiling/scoring and another vendor for presenting results using advanced visualization techniques. Meanwhile, the client may be running a POC [proof of concept] with a start-up vendor on how to extract benefit from the explosion of unstructured 'Big Data.' And the client may plug in and out as vendors emerge or stagnate. Vendors should therefore design solutions with genuine 'open interfaces' to support multi-vendor client architectures—enterprise data management is a big pie, and will continue to grow in size and effectiveness as niche vendors develop particular slices of this pie.

Bieri: Vendors can provide faster reference data, and add data as it becomes available. I also think it is important for a vendor to focus on their speciality, and continue to invest in that because that's what customers expect from you. This means both enhancing your products and providing documentation and support—to clients and intermediaries.

For SIX Telekurs, it is reference data and data used to meet regulatory requirements, such as tax information and classifications. In the next 12 months, these are the areas we'll continue to invest in, including data requirements for Fatca, and withholding tax requirements in our Swiss home market. As these and other regulations are coming into force in 2013, we have to move quickly and with focus.

In addition, vendors should avoid creating proprietary standards and instead support already established and widely accepted emerging standards. Vendor adoption of identifiers, classification codes and similar stan-

> "I think it is important for a vendor to focus on their speciality, and continue to invest in that because that's what customers expect from you"
>
> *Ivo Bieri, SIX Telekurs*

dards definitively helps clients, particularly in terms of cross-referencing and data scrubbing.

Blenninger: Vendors can provide customers with more knowledge on the semantics of different sources.

Static and reference data is something that costs money to get right, but even more if it is done wrong. Actually, poor data quality in our branch is caused by "not talking to each other." Vendors could help bring different customers together to solve the same problem only once.

Baskan: Vendors are building innovative technology components for financial services firms; however, they can do much more in helping with knowledge transfer, technical training and implementation assistance. These are the areas where firms need more support.

More options for technology from the cloud should also be made available, because these types of services help internal IT organizations increase their expertise, capabilities and capacity. Firms that want to consolidate services and decrease their footprint while still being flexible and dynamic are interested in developing private clouds. Public clouds are viable options for initial testing, prototypes and early-stage development. Firms can work with a cloud provider at the start of projects and then migrate the technology back to the internal team.

Ippoliti: The implementation of the standard LEI is a phenomenal opportunity for vendors of entity (or issuer) data to provide services to their clients. Larger firms with deep pockets may have the budget to map and cross-reference their existing entity data to the LEI, but smaller firms may not, and will be dependent upon vendors, on whose identifiers they already rely, to provide that mapping. Vendors whose entity data is thoroughly validated, de-duplicated, and mapped to LEI will be the beneficiaries of the LEI standard, and may even steal clients away from their competitors.

In addition, vendors have an opportunity to create new value-added services on top of the data that will be available from the facilities manager. The most obvious of these opportunities is the creation of legal hierarchy (ownership structure) trees. Once business entities can be uniquely identified, the creation of legal hierarchy trees will become substantially easier. Most vendors today are only providing links to the immediate and ultimate parent, leaving the bulk of the work of tree creation to their clients. Offering the full hierarchy would be of tremendous value to clients.

WIPRO
Applying Thought

BUILDING THE BANKABLE FUTURE

SIMPLE.SAFE.INTUITIVE.

Wipro Banking and Financial Services

Wipro's Banking practice has partnered with over 50 of the world's leading banks including 9 of the top 10 US banks and 2 of top 5 European banks along with leading banks in Asia Pacific. Our expertise across retail banking, enterprise banking, investment banking regulatory risk and compliance, mobile banking, channels, payments, lending & mortgages and wealth management address transformational changes for our customers. Our solutions on cloud. mobile, social computing and analytics enable our customers to build a bankable future.

DO BUSINESS BETTER

NYSE:WIT | OVER 120,000 EMPLOYEES | 54 COUNTRIES | CONSULTING | SYSTEM INTEGRATION | OUTSOURCING

Sponsored Statement

The Lack of Innovation

Intel has recently been a sponsor and has attended several Incisive Media events, notably around reference and market data. After a several-year hiatus, my team and I are re-engaging with the field in a bigger way. Intel has always been involved in small, direct engagements with the largest firms and data companies in some capacity. Now, we are scaling our efforts.

My colleagues and I have walked away from the recent conferences with three major opinions:

- Data continues to be the life-blood of a firm, and management challenges of that data grow as rapidly as the size of the data does. Big Data is the Big Problem, but also a Bigger Opportunity. Both vendors and end-users will be able to use Big Data in new ways for new profits.
- There are highly innovative solutions around hardware and software technologies to address the myriad issues around data management. But the fragmentation of solutions means that the need for more open standards is greater than ever.
- There is a general lack of game changing business models for data management in our space.

In several panels we've participated in and in discussions on the conference floor, it's been obvious to us that the current *laissez faire* approach to the business model side of data management continues. It perplexes me and my team that no-one is "shaking up the industry" with regard to how end-customers and data providers value a firm's data. This

In several panels we've participated in and in discussions on the conference floor, it's been obvious to us that the current laissez faire approach to the business model side of data management continues

applies to a firm's relationship with exchanges, data providers and even some software/application vendors.

In two forums, we asked: "Can anyone tell us of any other industry where the data contributors have to pay to access that same data?" So far, most people have unequivocally stated "No." It begs the question: What will it take to reshape the business model of data management in the financial services segment?

in Business Models

As consumers and businesses, we all have what Intel calls a "data footprint." This is the concept of a unit of measurement inclusive of all the data a person or entity generates and consumes. In the consumer space, this has been harnessed and utilized in ever more creative ways. From Siri (the latest Apple iPhone innovation around speech recognition and data integration technology), to Google, which allows us to search and index the massive amount of data across the virtual universe we all create, consumer technology continues to set the new high bar for data management. As consumers, we don't pay to access the data we and others generate. The value is in how the data is analyzed and then delivered to us. But the raw data itself is rarely, if ever, assumed to have a monetary value greater than $0.

Why is it the trade data created by firms on exchanges is then billed back to those firms by the exchanges themselves as well as the data aggregators? Why is the value of the data a firm creates sold back to the same firm at a premium?

As stewards of Moore's Law, Intel will continue to focus on bringing platforms to the market that help generate and analyze data in ever more impactful ways. The ecosystem we work with will continue to supply solutions to more effectively search, index, manage and use the data footprint of financial firms.

However, the notion of paying a premium to access data that has already been generated seems to be the model into the future. I suspect this will change in the next 3–5 years, as someone enters the market and radically reshapes the valuation models for a company's data footprint.

The colliding forces of regulation and competition will hopefully mean someone steps forward with innovations that go beyond hardware and software and change the business model.

The simplest way to think about this might be a simple question. What happens if Google ever bought a large data vendor? Do you think Google would charge you to access that data? It's doubtful. The value would be in the analysis they can provide behind that data. The day when firms no longer pay a premium for their own data is coming, perhaps sooner than we think.

Daryan Dehghanpisheh is director of financial services sector, Intel Corporation

Sponsored Statement

Centralized Intelligence

The days of isolated islands of data are coming to a close. Firms are moving toward shared operating environments, where reference data is seen as a common utility. Data is always a powerful asset, and when multiple types of data are linked intelligently, data value increases exponentially. Firms are investing in technology and internal resources to help them centralize functions, improve efficiencies, integrate data assets and serve the business. Organizations that can realize these benefits quickly will have a considerable advantage over those still struggling to understand—and manage—their enterprise data.

Firms have two main areas of focus right now for data management projects. One is financial market data repositories and reference data consolidation. The other is risk and trade transaction data for accelerated enterprise risk and internal governance purposes. Firms are concentrating on these areas to keep them out of difficult situations with the regulatory bodies and to eliminate blind spots on their balance sheets. These types of data management programs will improve transparency, help firms meet compliance regulations, as well as survive increased scrutiny. What's interesting is that firms are realizing when they make these improvements, their internal processes are becoming more efficient and they are able to identify profitable transactions earlier.

The secret to implementing a successful data management strategy is to think big but start small. In terms of program management, start by identifying a manageable project and rebuild it completely. Don't develop a big data management plan with 10 subprojects. Create a successful prototype that addresses a subset of the overall problem. For instance, if the equity derivatives' infrastructure needs to be revamped, a first step is to launch a fully funded pilot that rebuilds the next-generation infrastructure from live data capture to settlements and clearing, including risk controls and aggregated reporting. A pilot like this makes more sense than re-building an infrastructure across the trade life cycle for all asset classes and geographies. The scope is more manageable and when complete, the pilot can be used as a blueprint for repeated success in other lines of business.

In any enterprise data management project, firms must start with clean, accurate data—a golden copy—that can

be consumed by downstream systems. A common mistake, however, is to overlook how dynamic the data capture process is. It is better to think of the golden copy as a process of building and maintaining a "record of reference" in support of a specific application that is "true" in a specific time window. The time window may be in minutes or in days or weeks, but the temporal relationships between the pieces of information across a time line must be maintained and managed for the application cycle.

Where firms capture data also has an impact on clean, accurate data. Firms should consolidate their data repositories and capture and filter data as close to the capture point as possible. The challenge is that the data sources are increasing. There is more market data, asset pricing information, trading venues, even newsfeeds from Dow Jones and Reuters that must be captured to accurately assess trading and risk decisions, as well as post-trade and positions management. This larger volume of data increases both the pieces of software needing to be maintained and the number of user communities needing to be serviced.

Implementing an enterprise data management strategy is far from simple. Vendors are building innovative technology components for financial services firms but they are lacking in services that help with knowledge transfer, technical training and implementation assistance. These are the areas where firms need more support.

More options of technology from the cloud should also be made available, because these services help internal IT organizations increase their internal expertise, capabilities and capacity. Firms that want to consolidate services and decrease their footprint while still being flexible and dynamic are interested in developing private clouds. Public clouds are viable options for initial testing, prototypes and early-stage development. Firms can work with a cloud provider at the start of projects and then migrate the technology back to the internal team. All these strategies can help financial services firms better manage their enterprise data—and unlocking that valuable data will quickly lead to better business decisions and improved profits.

As director of business development at Sybase, Sinan Baskan is responsible for developing solutions for lines of business in the financial services sector. He has held various positions in the product engineering, professional services and marketing organizations at Sybase. Previously, he worked at HSBC Corporate Investment Bank on risk analytics

Q&A

The Regulatory Challenge

Inside Reference Data speaks to Philip Lawton, senior analyst on Aite Group's institutional securities and investments team

Philip Lawton, Aite Group

How is the intensification of regulation affecting the enterprise data management (EDM) space?

Modeling risk aggregation to meet Basel III capital adequacy and stress testing requirements entails intensively managing the volume and quality of enterprise-wide data. This is a terrific challenge for large, complex organizations—especially those with multiple accounting and risk management systems (often legacy systems). In the US, a regulatory development to watch is the initiative sponsored by the Treasury Department's Office of Financial Research to standardize legal entity identifiers.

What have been the most significant technological advances in the EDM space since the crash of 2008?

Technologists continue to focus on users' central issues—managing the volume and quality of data at an ever-faster pace. We see advances in automated data scrubbing, for instance, in corporate actions. Also, buy-side firms' urgent need for centralized information to support risk management and investment decision-making on the fly is impelling systems vendors to provide integrative solutions for continuous, multi-source data flows.

Is it preferable for end-user companies to buy off-the-shelf technology or to develop technology in-house?

I may have had too many discouraging experiences with internal systems development and maintenance projects to offer an unbiased opinion. In my view, superior systems vendors have specialized technological expertise and implementation project management skills that are hard to assemble and retain in financial services companies' IT departments. I would not sponsor an in-house initiative without researching what's available.

How do you expect the role of EDM to evolve as turbulence continues?

Market intelligence and complex event processing systems are already employed in high-frequency or low-latency trading operations; there are indications large firms are also progressively incorporating them into risk management processes. It is also to be anticipated that regulators' efforts to monitor systemic risk will place greater demands on market participants to report transactions in real time, for example, to feed the SEC's proposed consolidated audit trail system.

Intelligent Performance

intel inside
Xeon

Increasing performance when and where you need it

Incisive Media, 32–34 Broadwick Street, London, W1A 2HG