Inside Reference Data

October 2013 waterstechnology.com/ird

Enterprise Data Management

Special Report

Sponsored by

EAGLE INVESTMENT SYSTEMS

MarkLogic®

Editor's Letter

Flying Vs

The deployment of big data resources for the purpose of enterprise data management (EDM) has made inroads for EDM, but is not completely driving it, as MarkLogic's Ken Krupa tells us in the Virtual Roundtable in this report. The "three V" concept—meaning volume, velocity and variety as descriptions of data big flows that have proved challenging for EDM—has been magnified, as Krupa puts it.

The magnified "three Vs" are driving adoption of EDM programs, because they erode data quality and increase the importance of accurate data provenance and boundary information. EDM, in effect, is another means to pursue centralization of data (see "Centralization Takes Center Stage," the column in our regular issue this month). Centralization breeds efficiency in the EDM realm and addresses the data industry driver of creating clean golden copies of data.

Aside from lowering costs, EDM users are also seeking a better take on the impact of bad data, on how a single security's price may affect a fund, and on understanding how business decisions are affected by data, according to Eagle Investment Systems' Robert Brachowski. That includes gaining intelligence on opportunities and risk. EDM systems should include execution management and data inquiries capabilities, says J.P. Morgan's Ludwig D'Angelo. Aligning data correctly is another function for the EDM model to address, says Citi's John Carney. One static data governance strategy is not enough to follow a complex data governance policy, according to Brachowski.

While the data available for EDM may be taking off on the power generated by the "three Vs," many in the industry see EDM's strength in possibilities such as data alignment and centralization, not necessarily in addressing the "three V" whirlwind. The participants in the Roundtable in this report consider how EDMs themselves continue to exist with the range of challenges now involved.

Yours sincerely,

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

Inside Reference Data

The new bumper issue

Inside Reference Data app now available – FREE for all *Inside Reference Data, WatersTechnology Data* and *Premium Package* subscribers.

Download all the content from the latest monthly print issue, build your archive of issues and view them offline.

To download and find out more visit
waterstechnology.com/static/ird-app

An incisivemedia website
incisivemedia.com

Inside Market Data Inside Reference Data Buy-Side Technology Sell-Side Technology
waterstechnology

Contents

FEATURES

NEWS

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

EDM Driven By More Than Savings, Says Alternative Investment Fund Exec

NEW YORK—Business needs and opportunities are increasingly supporting the expansion of the practice of enterprise data management (EDM), an alternative investment fund official told delegates at the North American Financial Information Summit.

"This goes hand in hand with being able to satisfy their inventory requirements and the increasing complexity of the requirements," said Michael Korby, senior vice president of front-office development at Highbridge Capital Management in New York.

EDM is also proving to be a way to address legacy IT systems and tighter budgets, according to Korby. "C-level executives and managers are really getting into moving the agenda forward, pushing an agenda on regulation and the aftermath of systemic risk," he said. Highbridge has been seeing greater emphasis on understanding this data management process and performing risk calculations, said Korby.

Michael Shashoua

EDM Council Readies FIBO Specs, Works With Bank of England and Microsoft

LONDON—The EDM Council and the Object Management Group (OMG) have released specifications for the Financial Industry Business Ontology (FIBO), their semantics standards initiative.

The specifications will align FIBO with OMG and EDM Council operations, and follows on the organizations' work at the FIBO Technical Summit in San Francisco, which brought together financial industry firms to focus on FIBO.

The Bank of England has begun working with the EDM Council to align FIBO with its liquidity reporting templates.

In addition, the Data Maturity Management model, a related framework for data management completed by the EDM Council in 2012, was recently successfully piloted by Microsoft, according to Mike Atkin, managing director of the EDM Council.

Michael Shashoua

Bank of America, UBS Find Value in DMM Model

NEW YORK—Major firms are beginning to derive value from the Data Management Maturity (DMM) model created by the partnership of the EDM Council and the Software Engineering Institute of Carnegie Mellon University, according to officials.

Bank of America Merrill Lynch based its data governance program last year on the DMM model, and as a result of setting distinct principles for defining content, controlling and moving data, capturing metadata, documentation and data lineage, was able to set a common language for data management, according to John Bottega, chief data officer in the technology and operations unit at the bank.

For UBS Securities, meanwhile, the DMM model and "shareability" enables tracing of a data set to solve data governance issues, according to Ethan Smith, executive director, global head of instrument data and Americas regional head in the data service group at the firm's Stamford, Conn. office.

The party receiving data could make assumptions about it, but the DMM model limits that, Smith says.

Michael Shashoua

Tech Mahindra Signs UBS as First MDS User

Data services provider Tech Mahindra has launched the Managed Data Services (MDS) platform and signed UBS Fund Services, a Luxembourg-based unit of UBS, as its first user, says Jonathan Clark, head of financial services at Tech Mahindra.

Tech Mahindra built MDS based on UBS's in-house Instrument Data Management (IDM) platform, which the bank had been looking to monetize, says Clark.

Aegon Chooses Markit EDM

Aegon Asset Management has implemented Markit's Enterprise Data Management (EDM) platform in the US and the Netherlands, following the success of its work with the company in the UK.

Aegon, which has €250 billion in assets under management, has been a customer of Markit EDM since 2005, through its UK arm, Kames Capital. Markit's EDM platform will be used to create a data hub in each region and will allow data from these hubs to be consolidated.

Aegon officials say Markit EDM will help achieve a consistent view of good-quality data.

Virtual Roundtable

Enterprise Data Management

Inside Reference Data gathers together leading data management professionals to discuss the latest developments in enterprise data management

What is driving the adoption of EDM programs?

John Carney, director, securities and banking technology team, Citi: EDM programs in banks today are driven primarily by two things: cost efficiency—the drive to remove any duplication in applications and the data stores that supply them; and data quality—the drive to consolidate, validate and create "golden copies" of critical business data.

Robert Brachowski, product manager, reference data, Eagle Investment Systems: The first driver is the continued underlying need to improve data quality for key initiatives such as risk management and exposure reporting. The deeper investment managers can drill into their portfolios, the better they understand their direct and indirect exposure to various investments, which strengthens risk management at the firm and helps enable better business decisions.

Increasingly, data management is at the core of everything investment managers are doing, from accounting to performance measurement and reporting. Centralized data management creates tremendous efficiencies and allows multiple systems to directly integrate with and use the same set of data. Another driver of EDM adoption is industry regulation such as the Dodd-Frank Act and Solvency II, which push companies to increase the types of data that they support and the frequency at which the data is required.

Ludwig D'Angelo, executive director, JP Morgan CIB: The drive to adopt EDM programs is centered around three opportunities or challenges: the need to respond to regulatory requirements with agility, the need to improve risk management capabilities, and the need to improve the operational processes dependent on reference data.

Good EDM platforms allow fast time-to-market implementation to incorporate new attributes and data sources. The focus on comprehensive risk management is another key driver of EDM initiatives. The fully structured view of data across a financial organization—client, instrument, counterparty, pricing and transaction data—allows risk management from any perspective. The other grass-roots driver of EDM programs is centered on operational improvements. There is ample evidence of the impact of poor data quality of governance on clearing, settlement, collateral, and reconciliation functions within financial services organizations.

Ken Krupa, chief architect, MarkLogic: This may sound a bit clichéd, but borrowing from the big data "three Vs" concept [high volume, velocity and variety] actually does apply. This is not the same as saying "big data is driving EDM," but instead saying that the problems of variety, volume and velocity that have always been a challenge to EDM have been magnified in recent years. Most or all of the problems that EDM seeks to address have become considerably more of a challenge; data quality hasn't gotten easier, master data management/reference data is more important to get right, as are data provenance, establishing authoritative boundaries, and so on. The regulators are certainly having their say too. Look no further than the legal entity identifier (LEI) initiative. In short, EDM is becoming more mainstream and critical, and much less academic.

Do you see distinctions among the types of EDM offering available? What are they? And are there strengths and weaknesses to the different types?

Carney: There are many EDM software products and tools, all of which solve a

MarkLogic

Ken Krupa,
Chief Architect
ken.krupa@marklogic.com

different part of the problem. A genuine EDM solution that improves data quality and can drive costs down requires all pieces of the jigsaw to be implemented in some form: an enterprise business glossary, an enterprise framework, a logical data model, a physical database, data profiling, metadata management, Extract Transformation and Load (ETL), data lineage reporting and business rules for run-time data quality checks.

Brachowski: There are many EDM offerings that specifically focus on reference data management, where they move data around by leveraging an ETL strategy. This is not a long-term EDM strategy. More often than not, other solutions are required to manage all of their data and help with regulatory and exposure reporting. We view EDM as an innovative solution with robust capabilities to manage the full continuum of data, including reference data, portfolio data and complex portfolio structures.

Investment managers need operational data management to support their risk management processes and exposure reporting requirements. In this type of EDM offering, an industry-proven data model is the first step toward truly achieving data governance. With extensibility, this creates a long-term strategy to get the most value out of the data. This type of solution creates efficiencies that will help investment managers grow their assets.

D'Angelo: While there are distinctions—subtle or otherwise—between EDM offerings, the core capabilities include single, transparent, extensible data models with a defined metadata layer, flexible APIs, easy integration of new feeds and sources, a robust rules engine for data validation at a single source or multi-source level, and a customizable workflow module for exception management that is adaptable to any operating model. The instrument model has distinctions between EDM offerings built around a single provider with data enrichment from other sources and those that are true multi-vendor, hierarchical data platforms with the flexibility to define source data at the asset class or individual attribute level.

Krupa: Yes. Given that EDM can be slightly all-encompassing in scope, there may be distinctions within functional areas or even cross-cutting concerns. For instance, some solutions might have a stronger play with legal entity and/or securities reference data management, while others may come across more strongly on data governance. For implementation, a distinction might be made between malleability/configurability and a more rigid and tightly integrated solution. In each case, the answer is very context-specific, and, because the domain of EDM is so broad, it is difficult to generalize something as a strength or weakness. To some, a turnkey integrated solution is ideal. To others, a highly configurable framework is better.

What has been the impact of the use of EDM in the industry, and what will its impact continue to be?

Carney: EDM as a concept has been around for more than 20 years, but only recently has come to the attention of senior business leaders as a priority. One reason for this is a realization that the complexity, volume and richness of data and content—video, for instance—are increasing dramatically year over year in financial services and other data-intensive businesses. "Big data" is a reality. For banks to keep up with this and stay competitive, they need an enterprise-wide data management strategy aligned with client priorities, regulatory requirements and investment/growth objectives.

Brachowski: The impact of EDM has been getting businesses to focus on the data and prove that they are executing their business strategy, which has been motivated in many cases by the market volatility in recent years. EDM is the focus that will help firms gain control of their data to be fully knowledgeable about the health of their operations. Evidence of this focus is the increased transparency within a firm's investment portfolios to enable a deeper understanding of how a portfolio is functioning.

As we look to the future, business intelligence becomes more critical for firms wishing to compete globally. They can use these tools to monitor themselves and provide transparency to their client base. The investment management firms that have invested in and adopted EDM and business intelligence tools will be at the forefront of entering new markets, investing in new, complex investments and taking their firms to the next level.

D'Angelo: The current and future impact of EDM in the financial services industry will again be aligned with what is driving the adoption—specifically, the capability to respond to regulatory requirements, the continued need for improvement in comprehensive risk management and the

John Carney, Citi

continued requirement to drive down operational exceptions. The financial crises of 2007 and 2008 have taught us the importance of good data in risk management by giving us an indication of what we could not easily accomplish.

Krupa: Aside from allowing entire careers to be made on subsets of EDM alone (as with reference data management), EDM has highlighted a part of IT that often does not have a direct sightline to revenue. Yet the reality is that most, if not all, things EDM can eventually be tied to top-line and bottom-line concerns (and, yes, there's that regulatory bit too). The impact of EDM will continue in this regard but will also drive more collaboration within the industry. The EDM Council is just one example. The forces making EDM more difficult to manage are pushing ever-greater collaboration between traditional competitors within financial services out of necessity.

What is the value of the Data Management Maturity model (DMM) and what impact will it have?

Carney: The DMM is a useful tool for large organizations to assess their EDM maturity levels. As banks and other data-intensive organizations move from seeing data as something that supports their applications and business processes to seeing it as core to their success—a strategic asset—industry standard models and tools like the DMM will grow in popularity and importance as a litmus test for EDM maturity and data quality.

Brachowski: The DMM is designed to help companies assess the state of their data governance and define roadmaps for implementation of the vision. Those that have yet to define a rigorous data governance process would benefit the most and see the greatest impact from using the DMM. Firms that are striving to create a well-defined data governance process will leverage the DMM to assess their data oversight. Any available measure of data governance could become a competitive differentiator for these firms, particularly in volatile markets.

D'Angelo: The key value of the DMM is in allowing firms to benchmark themselves against industry best practices in order to focus their investments and attention on those aspects of data management that will provide the biggest return. Good data management coupled with the right EDM platform will allow firms to leverage data for virtually any use in an organization. When the assessment is executed

properly and without bias, it allows firms to understand how they compare and to quickly identify areas requiring attention. The assessment will help identify the interdependencies between the core data management functions, introduce a common language around data management functions and lay the groundwork for the formation of a data management strategy. As firms have consolidated and centralized their data management organizations, potential synergies have no doubt emerged, but the DMM looks to understand the functional, and will be profound in its ability to benchmark organizations.

Krupa: Whether you're big on methodologies or not, codification and agreeing on what is important in a problem space is hard to argue against. Some enterprises may embrace it wholeheartedly, while others will pick and choose. From the solution and/or vendor side, a well-thought-out barometer such as the DMM will lend some normalization to describing what the various solutions should provide over time. For a customer looking to choose between providers, mapping the problems to a DMM maturity level and using the associated vocabularies will come in handy, whether or not they formally adopt a DMM.

What should the priorities be for aligning data under an EDM model?

Carney: The starting point should be to create an enterprise data model that serves as a reference point to identify common data across an organization. Only by doing this can duplication of data and overlap be found across many separate data stores. It provides a common enterprise definition for every critical business element that individual data stores can map to. The mapping process identifies identical data elements that appear to be different in separate data stores because they have different definitions or applications. Citi is in the process of rolling out its reference model—the Citi Reference Data Model.

Brachowski: The priority for aligning data has been and always will be to align with the business strategy and mitigate risk to create a flexible and sustainable platform that can respond to changes. To position for success and manage change, companies must fully understand both their business and data strategies and invest in the solutions they can leverage.

"Centralized data management creates tremendous efficiencies and allows multiple systems to directly integrate with and use the same set of data"

Robert Brachowski, Eagle Investment Systems

With a comprehensive data governance policy in place, firms can establish the foundation for a consistent and efficient process for storing, managing and distributing their data. A data governance strategy alone is not enough—investment managers also require an innovative reference data solution that is flexible and provides a layer of indirection between the end-user and the data in order to implement and maintain the policies. Finally, the reference data solution must be aligned to deliver across a broad spectrum of requirements to all the business groups that rely on the data.

D'Angelo: I don't know if it is correct to assume a single set of priorities for aligning under an EDM model. Each organization will be different. The DMM assessment is a great way to provide a view of the current state of data management in your organization.

Ludwig D'Angelo, JP Morgan CIB

After that, articulating a comprehensive data strategy for the organization would be the initial milestone. The strategy should include principles and guidelines for data governance, quality and distribution, and should also include data stewards for each data type, who will have responsibility for the development and execution of the data model and governance execution.

The final component in this phase is the design and implementation of the operating model to support the strategy. At this point, you have the current-state assessment and the future-state strategy, so all that is left is the execution—essentially building a roadmap to move the organization onto the new architecture and operating model. This phase will require both the greatest resources and, typically, the most time, depending on the size of the organization.

Krupa: Some might say that focusing on governance and the quality of internally generated—that is, transactional—data might be the best thing to prioritize. Others might say that reference data from external sources should be brought under control first. It's a close call, but reference data would be a narrow first choice. The tie-breaker is that internally generated data has a tendency to naturally formulate authoritative boundaries and ownership—asset class X is traded by desk Y, for instance—while reference data from external sources is often duplicated in an enterprise, resulting in multiple "masters".

What are EDM users asking for from EDM model makers or EDM service providers?

Carney: This varies depending on the size of the organization. For example, asset management firms have a relatively narrow scope and limited IT resources, so they need low-cost, flexible solutions that can satisfy multiple data management requirements such as mapping, feed management, ETL and business rules. Global banks have a much more disparate and complex application set and legacy to deal with, plus enormous volumes across many clients. In such environments, the full stack of data solutions is usually required, spanning standard data models, metadata management, enterprise ETL, business rules and more.

Brachowski: EDM users are seeking a better understanding of the impact of bad or changed data, often before it is used elsewhere. For instance, at a micro level, pricing users want to know how a security price is going to impact a fund before the fund calculates its daily market value. At a macro level, if there are issues with the data—such as data not being received on time, or bad data being received—clients want to be able to understand what business events and decisions might be impacted by that data. Finally, investment managers want a better understanding of their data through business intelligence tools that allow them to reveal emerging opportunities and risks, and to consume it anytime and anywhere using mobile technologies.

> "The forces making EDM more difficult to manage are pushing ever-greater collaboration between traditional competitors within financial services"
>
> *Ken Krupa, MarkLogic*

D'Angelo: Most of the requirements from clients boil down to flexibility in integrating new data sources, flexibility and ease in developing the validation rules for the data, an extensible and transparent data model and an API that will allow easy integration of applications to the data. The other obvious requirement is an adaptable workflow tool that will support exception management and data inquiries, as well as the metrics and key performance indicators that sit above it.

Krupa: It varies as much as the space itself. I haven't seen a critical mass around any particular category of offering. But it's difficult to find a user that doesn't list data quality as an important requirement. It is how EDM solution providers offer capabilities around data quality that represent the "devil in the detail" for customers.

Sponsored Statement

Semantics and EDM in an 'As-Is' World

While Enterprise Data Management (EDM) is a bit of an all-encompassing (and often confusing) term, one of the common threads across the many EDM functions is domain modeling. In financial services IT, attributes which describe things such as securities, legal entities, trades, cash flows, and so on, are central to executing commerce at the line-of-business level but also, and perhaps more importantly, to enable the business to operate efficiently as a whole. Functions such as position-keeping, P&L reporting, and risk and regulatory reporting would be nearly impossible without some aspects of EDM in place.

Over the years, the diversity of financial instruments, particularly within over-the-counter derivatives, has particularly challenged EDM, often stressing the capabilities of relational technology. This caused disruption in some cases, such as NoSQL technologies, where more agile and self-describing models proved a better fit for complex structured instruments. However, being able to more easily support complex models within a database only goes so far in addressing all of EDM's challenges. After all, a canonical model is needed at some point to support cross-asset-class operations. The many challenges of enterprise reference data—such as data quality and data provenance—are still not quite addressed by a richer storage model.

More recently, the concept of semantics has made its way into the financial services lexicon. Semantics is another formalized modeling tool that allows us to assert facts about entities or to describe relationships between entities in a unit of expression known as a triple. A triple follows a subject-predicate-object format such as "John Doe" "Lives In" "Brooklyn" and/or "Brooklyn" "Is Part Of" "New York City." When querying across triples, inferences may be made that connect facts and relationships to each other so the machine can conclude that "John Doe lives in New York City." From a concepts viewpoint, this is not particularly new. What is novel is the adoption of standards (e.g. RDF, SPARQL, etc.) and the technolo-

gies that allow the efficient work at scale.

Having the ability to represent parts of the business model as triples is valuable, even without considering advanced inferencing capabilities. Take data provenance as an example. Being able to "decorate" transactional data with origin information such that it is co-resident with the data itself is powerful. Imagine not having to peer across multiple transactional system boundaries or cross-reference with an ETL and/or metadata store to track a trade's full lineage. Even modeling is enhanced. In much the same way that a self-describing schema allows for flexible entity modeling (i.e. storage as-is), the ability to assert triples between individual entities allows for more powerful relationship modeling. Using triples, relationships can be late-binding and do not have to be modeled all upfront as in the relational world. Nor must they be constrained by applying to whole classes of entities, but instead can be bound to individual objects—sort of an "as-is" for relationships.

Then there are the inferencing capabilities. Some colleagues and I attended a customer's workshop on EDM-related topics recently. The chief architect mentioned how definitions within the canonical model were often not relevant at various points in a trade's lifecycle. He cited interest rate swaps and how sometimes it's not the classification as an interest rate swap that's relevant, but instead it's certain characteristics of the instrument that matter, regardless of what the desk or EDM Czars name it. Post-meeting, my colleagues and I discussed this notion a bit further. We thought about how characteristics such as interest rate sensitivity throughout an asset's lifetime are important, regardless of the asset's name. And when considering that interest rate sensitivity is a tangible characteristic of Treasuries as well, the point hit home; two things with different names yet one important fact in common. An interesting semantics discussion indeed.

So while EDM will continue to be challenged by the rapid pace of innovation in the financial services space, there is hope that the emerging paradigms and capabilities are providing more tools to tackle the thorny problems. Semantics is yet another one of these and the industry is taking notice. The efforts of thought leaders in the EDM Council, as well as semantics-related initiatives such as FIBO, illustrate the trend. But as with all new innovations, it will not be the latest new paradigm on its own that saves the day; instead, it will be a matching of the correct tools to the problems at hand and making sure they work in a well-choreographed way.

Ken Krupa is chief architect at MarkLogic

Q&A

EDM Sunset

Inside Reference Data speaks to Virginie O'Shea, senior analyst at Aite Group

Virginie O'Shea

How is increasing complexity in data management driving the use of EDM?
There is rising sentiment in the market to move away from endless reconciliation to a much more strategic approach to data provisioning and support. The ability to proactively aggregate and supply data to key internal functions (such as risk management, portfolio management and trading), external clients and regulators is a high priority. Aite Group research from earlier this year indicates that just under half of the 26 asset managers surveyed had begun data governance and data management programs to wed key data sets together, and a further 23% were considering launching such a program.

What are the prospects for standards and models such as FIBO and DMM being effective for EDM?
There is definite appetite for adoption of industry standards where they make sense and fit with the wider goals of the organization—the use of an industry-standard legal entity identifier, for example (though we are still some way off achieving this!). There is a feeling, though, that any kind of standard or benchmark needs to be set by practitioners/end-users of the data rather than by IT departments or academics. There are also subtleties and nuances between different demographics within the financial services community and these should be reflected in any maturity model.

Is EDM still evolving as a concept?
Data management is constantly evolving, but EDM as a term may have had its day in the sun. There are far more models in the market than the model associated with a rigid EDM structure—some firms are still championing the term internally, but others have turned toward the idea of a data fabric and federated models.

Can you describe some of the effects of greater use of EDM by firms?
The effects of tackling data management in a more strategic fashion can be measured in the resulting appropriateness and "fitness for purpose" of the data at the end-user level. It should be measured on the basis of the quality of the individual data sets but also how well that data set can be combined with other data for analytical or reporting purposes.

Markets change. Be ready for it.

Innovative technology solutions to help investment managers respond quickly and efficiently.

Innovative software. International expertise. We adapt to regulatory, market and industry shifts to help you stay competitive, compliant and efficient. And with Eagle ACCESS[SM], our secure private cloud, we can help alleviate the management of technology and resources.

Data Management / Investment Accounting / Performance Measurement and Attribution

Contact us today to learn how Eagle can help your business:

Americas	+1 781 943 2200
Asia Pacific	+65 6432 0581
Europe	+44 20 7163 3827
Mainland China	+86 10 58771605

www.eagleinvsys.com

BNY MELLON

EAGLE INVESTMENT SYSTEMS

Eagle Investment Systems LLC, a leading provider of financial services technology, is a division within BNY Mellon Asset Servicing and is a wholly owned subsidiary of The Bank of New York Mellon Corporation.

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