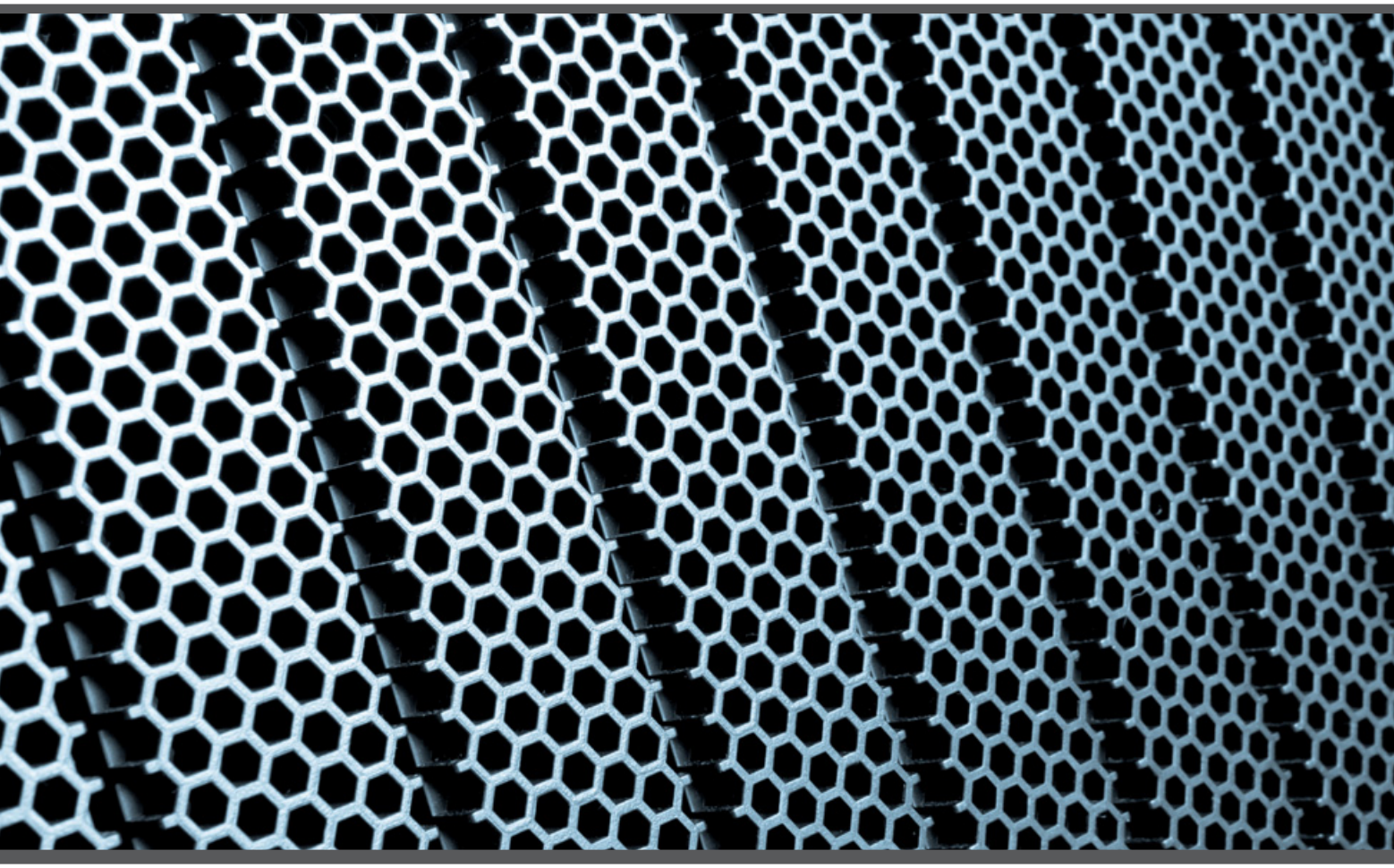
waters

trading technologies for financial-market professionals

Big Data
Special Report



Sponsored by:

cloudera®

intel®

MarkLogic®

Editor-in-Chief Victor Anderson
victor.anderson@incisivemedia.com
tel: +44 (0) 20 7316 9090
US Editor Anthony Malakian
anthony.malakian@incisivemedia.com
Deputy Editor, Sell Side James Rundle
james.rundle@incisivemedia.com
US Staff Writer Jake Thomases
jacob.thomases@incisivemedia.com
US Staff Writer Tim Bourgaize Murray
timothy.murray@incisivemedia.com
European Staff Writer Marina Daras
marina.daras@incisivemedia.com
Head of Editorial Operations Elina Patler
elina.patler@incisivemedia.com

Contributors
Max Bowie, Editor, Inside Market Data
Michael Shashoua, Editor, Inside Reference Data

Global Commercial Director Jo Webb
jo.webb@incisivemedia.com
tel: +44 (0) 20 7316 9474
US Commercial Manager Bene Archbold
bene.archbold@incisivemedia.com
US Business Development Executive Joseph Donato
joseph.donato@incisivemedia.com
Business Development Executive Shayna Hersh
shayna.hersh@incisivemedia.com
Account Manager Tom Riley
tom.riley@incisivemedia.com
Senior Marketing Manager Claire Light
claire.light@incisivemedia.com
Marketing Manager Natasha Gordon-Douglas
natasha.gordon-douglas@incisivemedia.com
Design Lisa Ling
lisa.ling@incisivemedia.com

Multiple Subscriptions/Site Licences
Monazer Rashid tel: +44 (0)20 7316 9606 (UK)
Andre Anderson tel: +1 646-736-1870 (US)

Subscription Sales
Hussein Shirwa tel: +44 (0)20 7004 7477 (UK)
Dennis Bond tel: +1 646 755 7374 (US)
waters.subscriptions@incisivemedia.com

Group Publishing Director Lee Hartt
Chief Executive Tim Weller
Managing Director John Barnes

Incisive Media Head Office
32-34 Broadwick Street
London W1A 2HG,
tel: +44 (0)20 7316 9000

Incisive Media US
55 Broad Street, 22nd Floor
New York, NY 10004
tel: +1 646 736 1888

Incisive Media Asia
14th Floor (Unit 1401-3)
Devon House, Taikoo Place
979 King's Road Quarry Bay
Hong Kong
tel: +852 3411 4888

Incisive Media Customer Services
Haymarket House
28–29 Haymarket
London SW1Y 4RX
tel (UK): +44 0870 240 8859 (UK)
tel: +44 (0)1858 438421 (International)

Waters (ISSN 1068-5863) is published monthly (12 times a year) by Incisive Media. Printed in the UK by Wyndeham Grange, Southwick, West Sussex.

The Spice of Life?

Few *Waters* readers will be familiar with the British poet William Cowper, much less his magnum opus, *The Task*, a collection of blank-verse poetry presented in six books. But most will know the idiom: "Variety's the very spice of life," taken from book two, *The Timepiece*. This cliché, no doubt, rings true for most people, although if you're a chief data officer at a financial services firm, or you have a data-centric job description, variety is the stuff of nightmares.

In the capital markets, data can be segmented into two constituents: structured and unstructured. When it comes to most things in life, structure is good. It allows us the anticipate situations and plan ahead. It fosters familiarity, consistency and continuity, and it presents us with a certain logic and pattern that we find comforting. It is what got Amedeo Avogadro and Leonardo Fibonacci out of bed in the mornings, and continues to draw the planet's finest brains to the pursuit of the sciences. In short, structure is good.

Conversely, a lack of structure is essentially chaotic, which, when it comes to data management in the financial services industry, is not a good thing. Processing and storing large data volumes from disparate sources being generated and disseminated at ever-increasing velocities, is no walk in the park, but as participants in this report's virtual roundtable explain, the technology and know-how is available to pretty much all market participants, allowing them to stay on top of their data from this respect. But when it comes to processing unstructured data—described by MarkLogic's Amir Halfon in his editorial on page 3 as "information, which is either hidden in free-form text, or scattered across incompatible schemas"—things become markedly more complicated. However, unlike many big data discussions we've listened to or read in the past, participants in this report not only look at the operational and technology challenges of effectively addressing the unstructured data conundrum—they also propose the techniques and technologies available to market participants, allowing them to process huge data quantities and crucially analyze it, thus deriving value from one of their most precious commodities. ■

Victor Anderson
Editor-in-Chief

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

Emerald Advisers Takes Indata iPM Cloud Service

Asset-management firm Emerald Advisers has migrated to Indata's Intelligent Portfolio Management (iPM) cloud platform, according to Indata.

Leola, Penn.-based Emerald selected the platform after a competitive bidding process, saying that the cloud architecture and data management capabilities of iPM secured its selection.

"Data integrity and control is important and iPM's capabilities are advantageous and easy to use. iPM's advantage extends beyond data management to also provide more efficient integration with outside systems," says Scott Rehr, COO at Emerald Advisers. "The system's trading, modeling and reporting capabilities are robust and the cloud delivery option allows us to focus on our investment process without having to be concerned with the reliability or responsiveness of aging IT infrastructure."

Tech Mahindra Launches Managed Data Services

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

Tech Mahindra built MDS based on UBS' own in-house Instrument Data Management (IDM) platform, which the bank had been looking to monetize, adds Clark. Tech Mahindra successfully bid to acquire the platform after showing UBS that IDM was similar to, and aligned with, its own development plans and strategy for data services, according to Clark.

Fluent Enlists Solace for Data Messaging Layer

Fluent Trade Technologies, a New York-based provider of hosting, connectivity, data storage and distribution systems and feed handlers, is using Ottawa-based Solace Systems' messaging middleware appliance as the data distribution layer for its solutions, enabling transfer of market and trading data between its front- and back-office systems, and between big data warehouses and client interfaces, to support its customers' trading strategies.

Panopticon Enlists NuPont Growth Advice

Swedish visual data analytics provider Panopticon has enlisted market data and technology consultancy NuPont to advise on how to bolster its presence in the capital markets and spearhead a marketing push to achieve greater growth among financial clients.

Jerry Brunton, former global head of marketing at BT Radianz, who joined NuPont as a senior consultant in January, will act as advisor to Panopticon—which was acquired recently by Chelmsford, Mass.-based data management software developer Datawatch in a deal worth $31 million—to create a marketing plan that will help achieve the vendor's growth ambitions.

Brunton was initially contracted by Panopticon to provide general advice to the vendor's management team on how to achieve growth in the capital markets, though following the acquisition—which will see Datawatch integrate Panopticon's real-time visualization software into its Information Optimization Platform (IOP)—Datawatch is also seeking to grow its big data visualization business in markets where Panopticon already has a footprint, such as telecoms, energy and retail.

NuPont will also help Panopticon develop a "more robust" partner program to enable the vendor to communicate more effectively how its technology functions with third-party technologies, such as Thomson Reuters' Enterprise Platform for Velocity Analytics, SAP's HANA appliance software, and OneMarketData's OneTick database.

Big Data Is Not Just About Size

Data is at the center of most challenges facing financial institutions today, with business drivers such as new regulations, aggregated risk management, and deep customer insight all having critical data management implications. By Amir Halfon

Big data has become a common way to describe data sets at the center of many financial firms' activities, and while some challenges are associated with large volumes, it isn't really the size of the data that's the issue. At this point, we know how to handle large volumes: Use shared-nothing architectures that scale horizontally on commodity hardware. The trickier problem has to do with a different "v" of big data—variety. There are countless examples of business value locked up in data that does not fit neatly into rows and columns. The most frequently cited is social media, with its ability to offer deep customer insight and sentiment analysis. And there are many others within the company's firewall: gleaning information from on-boarding documents for Fair and Accurate Credit Transactions Act (Fatca) and anti-money laundering (AML) compliance, getting a better handle on credit risk by analyzing ISDA agreements, and lowering cost-per-trade by consolidating the processing of diverse asset classes with varied and complex structures.

Handling this Information

How can we effectively handle all this information, which is either hidden in free-form text, or scattered across incompatible schemas? Hierarchical structures such as XML and JSON come to mind, as they can accommodate various degrees of structure, organized in a way that mirrors intuitive human perception. Indeed, many organizations use XML to handle these business challenges and have reaped some benefits, but found themselves constrained by the underlying relational database management system (RDBMS) platforms that actually managed the data.

The RDBMS approach to handling hierarchical information requires that data be "shredded" into tables—a customer/derivative trade/legal document, with all its hierarchical attributes, is shoehorned into an entity–relationship (ER) model that satisfies referential integrity. Don't get me wrong—I love relational modeling and I have spent years doing it, but Third Normal Form has its limitations when it comes to diverse data—just consider the typical first step when analyzing normalized data: De-normalize it!

There is an alternative to shredding though: NoSQL databases—a wide set of technologies that transcend the boundaries of relational schemas. The acronym is unfortunate since SQL is actually one of the best features associated with an RDBMS. The problem with RDBMSs is not SQL but the prerequisite of a schema definition for data ingestion and analysis, which hinders business agility. We've all seen cases where the business needs have been delayed (sometimes by months) while data models, transformations and analytical schemas were developed. NoSQL databases enable real business agility.

However, one architectural principle has prevented a wide adoption of NoSQL technologies within the enterprise: BASE. It stands for Basically Available, Soft state, Eventual consistency (a play on ACID transactions—Atomicity, Consistency, Isolation, Durability, which are associated with relational databases). BASE has several advantages when it comes to non-transactional systems, as it relaxes consistency to allow the system to process requests even in an inconsistent state. Social media sites are a perfect example of this—few people mind if their Facebook status or latest tweet is inconsistent within their social network for a short period of time; it's much more important to get an immediate response than to have a consistent state of users' information.

Different

Financial and other enterprise systems are different, though. Imagine, for instance, a corporate merger action, occurring at the same time a firm is trading the affected instrument: The post-trade processing systems would certainly have to be consistent with the reference data system, or costly exceptions would ensue.

So how do we avoid schema woes without giving up ACID transactions, as well as other enterprise qualities such as fine-grained entitlements, point-in-time recovery, and high availability, all of which we've come to expect for mission-critical systems? The answer lies within a different category of database, Enterprise NoSQL, which has been designed and built with transactions and enterprise features from the ground up, just like relational databases. But unlike RDMBSs, an Enterprise NoSQL database models the data as hierarchical trees rather than rows and columns. These trees are aggressively indexed in-memory as soon as the data is ingested, and then used for both element retrieval and full text search, unifying two concepts that have traditionally been separate—the database and the search engine.

An Enterprise NoSQL database also offers full SQL access, thus combining the benefits of both worlds—the business agility associated with NoSQL and search, and the data integrity and sophisticated querying associated with a traditional RDBMS. Big data for many companies has meant big challenges, but with Enterprise NoSQL, lengthy development cycles associated with schema creation/modification, ETL (extract, transform, and load) workflows are no longer concerns. ■

Amir Halfon is CTO of the financial services division of MarkLogic Corporation.

The Vagaries of Variety

The concept of big data has been around the capital markets for some time now, to the extent that most financial institutions understand its implications to their infrastructure. And while the sheer volumes and velocity of that data no longer present firms with insurmountable operational and technology challenges, the same cannot be said for the variety of data that firms must deal with on a day-to-day basis in order to derive value from it.

Q What are the specific business applications for big data across the buy side and sell side? Which business processes are most affected by the continued growth of data volumes, in addition to its complexity and variety of sources?

Mark Atwell, distinguished engineer, head of software platforms engineering, RBS Markets: The current regulatory agenda affecting us—and many of the banks—is driving up the amount of data we consume, for example, with risk measures and scenarios, but not to stratospheric levels. But this data is highly structured. The need to do aggregation over these relatively large, but not yet huge, data sets is somewhat offset by a hybrid approach. We produce aggregate views with drill-down to the underlying data, but believe that the need to do high-performance aggregations in different ways isn't quite as dynamic as some suppose.

So, consequently, we don't really see the variety of data.

Amir Halfon, CTO, financial services, MarkLogic: There are a few key applications that we see: compliance with new regulations that require rapid, on-demand access to varied data; risk management and trade processing across complex asset classes; deep customer insight that brings together data from diverse internal and external sources; and pre-trade decision support that includes new sources of information beyond market data.

The emphasis changes between the buy and sell side, but these are the core applications and related business processes that are most affected by big data challenges. It is interesting to note that among these challenges, high volumes are not necessarily the most demanding—MarkLogic is seeing data variety and diversity as much more important in terms of pushing the envelope across the industry.

Mark Record, partner, Capco: Data is the "final frontier." It is big, largely "untamed," and often misunderstood. In all financial services organizations of scale, data proliferates. Yet the ways in which it is managed vary substantially from one institution to another. Even more significantly, approaches vary from one part of the same organization to another. At the most fundamental level, there is substantial variation in how senior managers actually view the data itself.

Ownership of data also varies across and between organizations with some firms defining clear ownership and others not assigning it at all. Capco's 2012 data management survey found that 35 percent of respondents did not assign data ownership to anyone.

Mike Olson, chief strategy officer and co-founder, Cloudera: Buy-side firms need the best information they can get about markets and the factors that affect valuations. Trend and sentiment data from social media, for example, are a powerful complement to traditional transaction and market-share financial metrics when predicting future growth. Sell-side analysis wants precisely the same view of the future, of course, but wants even deeper insight into broad financial markets and their likely evolution. They need to answer the question: Is it better to sell or to hold?

cloudera®

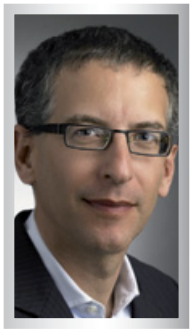
Mike Olson
Chief Strategy Officer and Co-Founder
Cloudera
Tel: +1 650 843 0595
Email: mike.olson@cloudera.com
Web: www.cloudera.com

In both cases, the combination of complex and traditional data types overwhelms legacy data management and analytic tools, and demands the flexibility of a native, big data-capable platform. Beyond the simple problem of variety, though, the best analysis demands the most powerful analytic tools. We see much more sophisticated modeling in use today, based on techniques like machine learning and natural-language processing over that variety of data. Those models are more accurate and deliver more reliable results with long-term, detailed data—a decade's worth, rather than a quarter's worth—so volume is a critical issue.

The value to be gained from big data and the new generation of platforms—Apache Hadoop in particular—varies by market. We see most of that value captured by innovative business applications aimed at vertical-specific problems. In financial services, a more accurate assessment of portfolio risk, and a shrewder look at sources of fraud and illegal behavior in transactions flows, are possible. Health care companies can use more data and better tools to design drugs, or a comprehensive analysis of symptom, treatment and outcome data across large patient populations to deliver better care more cheaply. Retailers need the best information they can get about customer preference and behavior in order to engage more profitably across channels—web, in-store and otherwise.

We expect big data to become a critical piece of the infrastructure for enterprises generally.

Joe Opdahl, director of marketing for big data analytics and storage solutions, Intel Corporation: Applications we see in the financial services sector range from increasing operational efficiencies to more targeted marketing efforts and spawning whole new business lines. Sell-side firms, along with retail banking—credit cards, transactions, mobile, and so on—benefit from access to more data, both structured and unstructured, to better understand their customers. The new wealth of this data is analyzed, which enables a more pointed marketing strategy, resulting in more consistent, or even predictive, sales growth.



MarkLogic®

Amir Halfon

CTO Financial Services
MarkLogic
Tel: +1 877 992 8885
Email: sales@marklogic.com
Web: www.marklogic.com

Buy-side firms can have faster access to larger historical datasets to further increase statistical confidence in existing indicators and strengthen trading-signal analysis. This, of course, also applies to the investment arms of insurance companies as they iterate their actuarial models to utilize these big datasets in near real time.

Data lifecycle management deserves a mention as a traditional business process that is heavily affected by big data. Data volumes are less of an issue here—storage costs in general have continued to drop by as much as 90 percent over the last decade. However, addressing data back-up and disaster recovery with big data is rising as a new challenge. Complexity increases as firms try and compare and correlate high-volume data stores cost-effectively when they discover the lack of consistent format, context, and data quality. We are seeing many of the "new" big data firms offering these types of data operations products in the marketplace today, to solve these data "variability" issues often by employing the older extract, transform, and load (ETL) model.

We are witnessing the power of social media to move markets—we've seen a single tweet alter asset valuations by billions in a very short time. Trading based on sentiment gauged from internet feeds is being explored by funds and proprietary trading shops. They often utilize natural language processing (NLP) algorithms that are compute-intensive.

Q How do firms go about planning and implementing a solid, pragmatic big data strategy? What do they tend to overlook when embarking on such projects?

Opdahl: Firms should start by exploring how big data will provide a competitive advantage to their existing business strategies and consider what new lines of business service value they could offer with a more comprehensive or unique perspective of their existing or targeted markets. Another key aspect of a solid big data strategy is the talent needed to implement big data. Having the right people with the right skills is crucial for cultural acceptance of quantified insights and creating the right business processes to support the underlying infrastructure for data collection, processing, and dissemination.

Finally, one should look at data sets and tools. It is likely that firms have more data than they know what to do with, and they do not have to look far to find recommendations for tools that connect and update these into a more modern, cost-effective, and scalable tool infrastructure. Intel has worked hard to accelerate technologies and solution partners to market into the big data space.

In general, the steps organizations take to adopt a traditional, major IT technology are similar to those they would use to deploy big data analytics. IT departments can be quick to focus on the new and shiny tools or technologies. However, the rest of the organization may initially be suspicious of new indicators and insights, which direct precious resources and funding away from traditional approaches into new and uncomfortable directions. That is why it is important to have leaders and line-of-business sponsors with clear and well thought-out business cases for adopting big data insights. A word of caution here: We have also seen recommendations for firms to throw out existing methodologies, which can be a hasty mistake. Often, new big data analytical insights are built on an existing understanding from long-accepted analytic insights, and grow from there.

Mark Atwell
RBS

Atwell: We are looking to establish a viable value proposition by performing speculative searches across our data. We are not alone in this. There is a lot of noise around big data, but it seems to be more of a band-wagon movement at the moment. It's one that the vendors are pro-actively promoting, looking to badge anything as big data or big data integration—rather like cloud before it.

Our ability to produce exactly the same results repeatedly means that atomicity, consistency, isolation, and durability (ACID) properties are still very important to us. We do perform analysis over transient data, such as access logs, but this is primarily for rapid fault analysis and remediation.

Record: Firms need a constructive, cross-organizational and practicable approach to effective data management. This approach must be flexible enough to accommodate changing market and operational conditions, supported by appropriate technologies and systems, and driven by a common data culture. Analytical tools should be available to people who liaise with the customer, not just those in the head office. Increasingly, there is a need to analyze real-time data as opposed to data that is weeks or months out of date. Customers should be given the tools to analyze their own data and discover new buying options.

Olson: The critical step is to identify, clearly and crisply, the business problems to be addressed by big data. The most successful new projects start with just one or two well-defined use-cases, and build out the platform and analytic infrastructure to deliver on them.

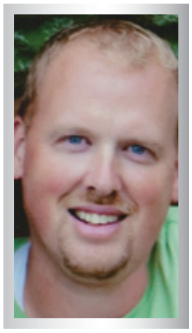
With that experience, teams gain the insight they need to expand their use of the platform. What problems are amenable to big data techniques, and which aren't? What combination of current and new tools are required to gain value? Virtually every enterprise we have worked with has quickly grown from its initial, focused deployment to broader use across more business problems, but it's done so deliberately, gaining skills along the way.

No enterprise deploys big data solutions in a vacuum. New platforms absolutely must integrate with the platforms and the tooling already in use. The biggest risk in big data deployment is to ignore that crucial fact. Success demands a comprehensive, integrated approach. New big data infrastructure is only a piece of the puzzle in the datacenter, and must mesh with the rest of the enterprise's mission-critical systems and applications.

Halfon: A pragmatic approach to big data is predicated on a fine balance between an holistic architectural vision, and a specific set of tactical business initiatives that drive adoption and funding. Such an approach takes into account the overarching data management requirements, new technology capabilities, and organizational aspects such as reusability and shared services, while avoiding the pitfalls of becoming an ivory tower on one hand, or a collection of "shadow IT" science projects on the other. It's those two pitfalls that can be easily overlooked, causing big data efforts to devolve into either uncoordinated one-off projects, or vision-only initiatives that are not anchored in tactical, well-funded business projects and actual use-cases.

Q What are the technology and operational challenges that need to be considered when dealing with big data? What technologies are available to firms looking to address this big data challenge?

Record: At present, all firms will have a data policy, but most employees won't be aware of it. It is now important to ensure that they do. Recent changes in the financial industry have been substantial and data is a vector of organizational change. As organizations change, they must adapt to the way they manage data. Data policies are often outdated by the time they are finalized and this must be prevented. Data ownership must also be considered because if it is unclear who owns the data, there is potential for people to change it, thus damaging its quality. Few banks are currently undertaking big data initiatives and there are a number of reasons why. Unstructured data is hard to quantify and firms can't justify expenditure on new systems to monitor this. In addition, privacy is an issue especially as banks try to rebuild trust with their customers. Banks' appetites for such initiatives have also been suppressed due to an occupation with historical business-intelligence investments that attempt to mine legacy data, improve analytics, and support the immediate needs of analysts, marketing and sales.



Joe Opdahl
Director of Marketing for
Big Data Analytics & Storage Solutions
Intel Corporation
Email: Joseph.B.Opdahl@intel.com
Web: ww.intel.com

Opdahl: Scalability of infrastructure comes to mind. If IT infrastructure is geared toward high-performance computing (HPC) emphasizing low-latency processing of structured data, then scaling can be expensive. To address this challenge, look at Hadoop and NoSQL as well as thin provisioning of data stores.

Business-intelligence systems are most often the final stop for the data and insights before they are consumed by the decision-makers at firms. It is imperative that these business-analytics systems are well integrated with Hadoop and other big data elements. Open-source solutions such as Revolution R are gaining popularity.

With respect to security and compliance, as new data sets are connected, firms need to be confident that their competitive insights are not lost or corruptible. Here, Intel offers infrastructure and database cell-level technologies, introduced to the market through AES-NI, Project Rhino, and McAfee.

As far as the democratization of data is concerned, how do firms connect their data to new datasets via public clouds or other data application programming interfaces (APIs)? The inclusion of big data APIs in public datasets makes these much more accessible. Firms may think of creating, opening and monetizing their own private data sets, and companies like Mashery Solutions can help in this regard.

Atwell: A number of incumbent reporting and visualization technologies are still SQL-based—as indeed is our staff. It is noticeable that many of the big data vendors and other purveyors are now, in effect, retrospectively adding these features.

Nonetheless, genuine operational concerns include sensitivity and security—and once again, this is an area that some of the vendors are looking to address.

Halfon: The key challenges revolve around the maturity of the technology on one hand, and the organization on the other. Many technologies in the big data space are relatively immature, and may not provide adequate enterprise capabilities required for production deployment. Capabilities such as cross-record ACID transactions and fine-grade security authorization are particularly critical, and can be easily overlooked in the early stages of adoption. This is also related to organizational maturity, in terms of having a solid enterprise architecture that's focused on these key systemic requirements.

In terms of available technologies, they fall into three main categories: RDBMS appliances and fit-for purpose solutions, which expedite the performance of traditional RDBMS, but do not address data variety challenges; Hadoop-based solutions, which take advantage of the schema-free nature of Hadoop Distributed File System (HDFS) to handle non-relational data in a highly parallel, long-running fashion, but are not suited for interactive, operational use-cases; and Enterprise NoSQL, a mature version of NoSQL, which is schema-agnostic, and enables the management of any structure data, while still providing the enterprise capabilities required for mission-critical enterprise applications.

Olson: The biggest challenge for most enterprises is that big data platforms and tools are new. Existing business and IT staff are expert in the use of traditional relational databases and applications like enterprise resource planning (ERP), for example, but are not yet familiar with (in Cloudera's case) Hadoop and the new suite of tools and applications that work with it. That means that staff must adopt new tools and learn new skills to use the platform. The good news is that that's not rocket science—business analysts and system administrators are entirely capable of mastering new tools and learning new systems. It's important, though, to take a considered approach to the new system, especially thinking about training and services to ensure success in the deployment. The population of available tools and applications native to big data is large and growing quickly. Training for developers, administrators, analysts, and data scientists is important and easy to get.

Q Are there existing technologies—cloud computing, for example—that can be deployed in a complementary fashion alongside specific big data technologies to help alleviate the big data burden?

Olson: The most important existing technology to think about is your current data management and analysis infrastructure. Enterprises generally adopt big data as a complement to their current systems, not as a wholesale replacement. Think about where data lives, how it moves, and when and where you process and analyze it. Make sure that your systems can talk to one another.

We certainly see cloud computing in particular as important in the market, but in our view it is a separate strategic decision from adoption of a big data platform. Enterprises have current IT infrastructure and a vision for how they'll evolve those systems over time. Migration to the cloud is attractive to many. But a useful big data platform must run on the infrastructure that the particular customer uses, whether that is a privately operated datacenter behind the enterprise firewall, leased systems managed by a trusted hosting provider, or public cloud platforms like Amazon Web Services. CIOs must decide where storage and compute infrastructure will live. The big data platform must run on that infrastructure.

Opdahl: Yes. There are efforts to address the need for running business analytics and data operations, such as Hadoop, as part of a public or private cloud that eases the deployment of big data infrastructures. Examples would be Savanna from OpenStack, the VMware Serengeti project, and Mashery. All offer compelling ways to quantify secure access to firms' data and/or analytics suites making it possible to monetize proprietary datasets. Over time, we believe the lines between cloud, high-performance computing, and big data will blur, and we expect more technology, analytics and insights to be consolidated and incorporated.

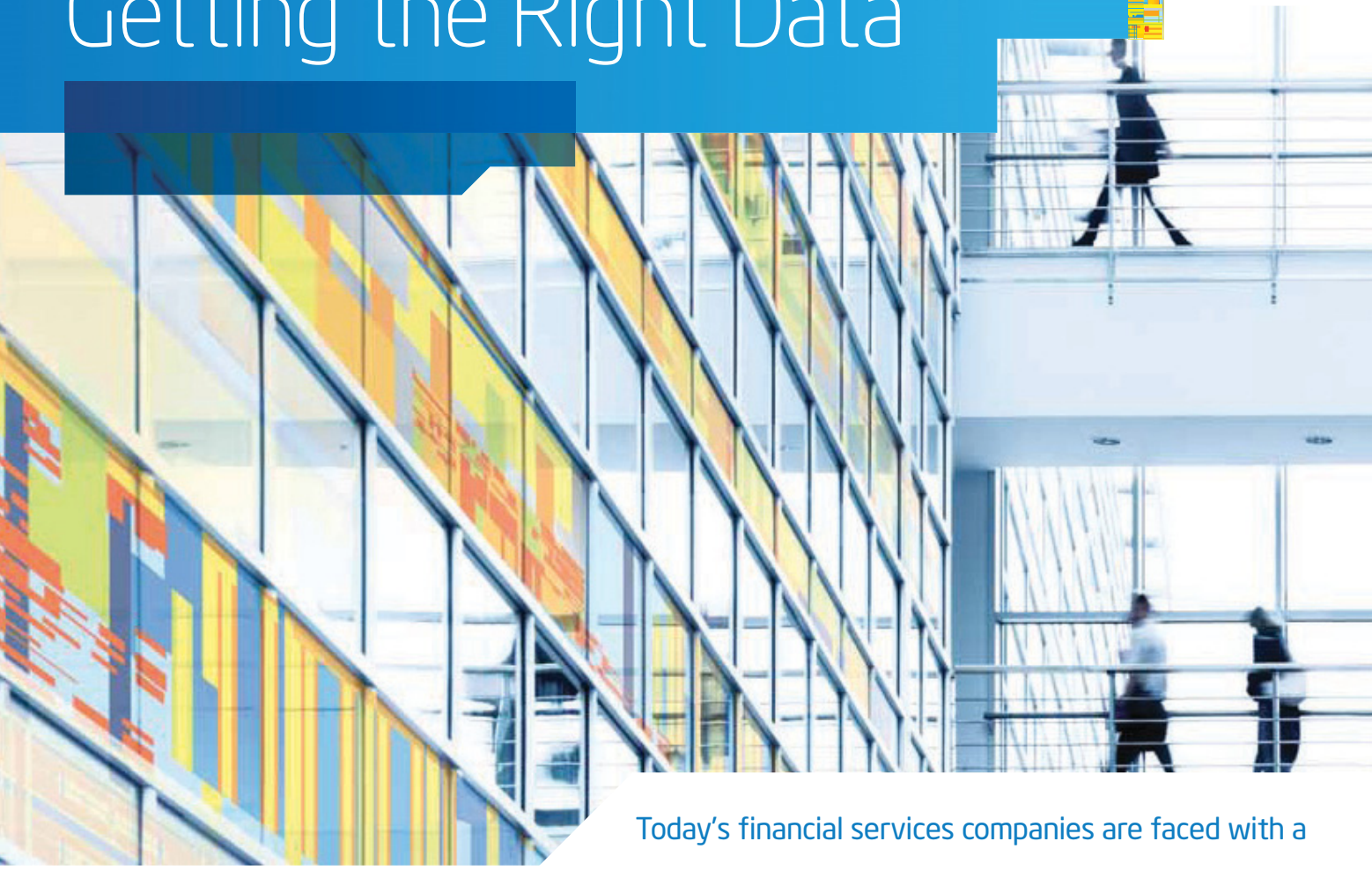
Mark Record
Capco

Record: Emerging technologies are being driven by social media, both in terms of the need to turn data into information as well as provide customers with the tools to allow them to take more control over their own data and information. We will also see more regulation in this space and greater data protection. Customer data is being mis-sold and mis-used without the customer's consent or knowledge, and, as a result, customer education will be required in the future.

Halfon: Absolutely. Especially in terms of dealing with large volumes, the cloud is a natural complement to the technologies mentioned earlier. An interesting emerging trend is using the cloud as a storage tier, accessed by the data management platform alongside other tiers such as solid-state drive (SSD), local storage, and storage area network (SAN), according to user-defined service-level agreements (SLAs). This provides an information lifecycle management mechanism that is more dynamic and more cost effective than traditional approaches. ■

Note: Dejan (Dan) Kusalovic, Charles Milo, James Fister and Parviz Peiravi, all from Intel, collaborated with Joe Opdahl on this editorial.

Risk Management Means Getting the Right Data



Today's financial services companies are faced with a stark choice: evolve or perish. Squeezed to relentlessly push the boundaries of technology, you demand the most advanced solutions to remain competitive in an industry in which success is measured in microseconds. Intel provides the innovative platforms, strategic relationships, and future-proof flexibility required to stay ahead of the curve.

ARCHITECTS OF WHAT'S NEXT

Intel in Financial Services

www.intel.com/financialservices

waters