



## WHITE PAPER

# Big Data in the Enterprise: When Worlds Collide

Sponsored by: Oracle and Intel

Dan Vesset  
February 2014

Carl W. Olofson

## IDC OPINION

---

Much of the language we use to describe core business processes is as old as commerce itself. Businesses still produce products and provide services, acquire and retain customers, and protect their assets and comply with regulations. But the digitization of these and other core tasks has changed them dramatically:

- The channels of customer interaction have multiplied, the digitization of these interactions is pervasive, the variety of data produced by these interactions has exploded, and customer expectations have changed as a result of new, innovative consumer-driven practices. These trends are fundamentally changing sales, marketing, and customer service processes.
- Digitization has also extended to physical things – cars, buildings, bridges, appliances, clothing, and equipment – that are generating huge volumes and streams of data available for analysis, decision making, and optimized action. This phenomenon, known as the "Internet of Things," is fundamentally affecting production, logistics, asset management, and procurement processes.
- The digitization of interactions between people and things is also expanding rapidly, creating further complexity in big data and analytics requirements. At the same time, significant progress in big data and analytics technology has resulted in vastly improved price/performance characteristics for acquiring, managing, and analyzing large volumes of multistructured data.

The net result has been the availability of volumes of a wide variety of data generated at extreme velocity. What has become known as "big data" presents enormous opportunities to fundamentally change business processes, create new business models, and drive innovation. It also presents new challenges for most organizations. Big data has become big business, and the recent hype surrounding it has resulted in confusion about technology options, unrealistic expectations, and often advice that is neither actionable nor effective for most organizations.

In 2013, the broad big data and analytics technology and services market for the first time broke through the \$100 billion mark. Big data and analytics represents a range of information management and analysis technologies, yet lately most of the buzz has been around a few specific technologies such as Hadoop and advanced analytics software. These technologies have an important place in most organizations' architectures. However, alone they address only a subset of requirements.

Today, there is mounting evidence of the end of irrational exuberance in big data. 2014 is ushering in a long-term (very long-term) trend of pragmatic purchasing and deployment of a range of big data and analytics technologies and services. This pragmatism has already resulted in a realization of not only the need for coexistence of relational and nonrelational big data and analytics technologies but also the fact that together these technologies can enable completely new ways of conducting business, serving and protecting citizens, caring for patients, teaching students, and managing natural resources.

To unlock the value of big data, organizations need to extend the existing IT architecture from the tried and tested relational technology to nonrelational technology. Only a combination of both relational and nonrelational data, technology, and the analytic techniques they enable can propel organizations toward fact-based decisions driven by the analytics of large and diverse big data sets and decision management enabling consistent, compliant, and optimized responses to business events.

For IT decision makers, the increasing number of technology choices and decisions that need to be made in response to evolving business demands may seem overwhelming. However, a structured evaluation of the following requirements can help in developing a long-term big data and analytics strategy and an iterative execution plan that will leverage a combination of relational and nonrelational big data and analytics technologies for competitive advantage:

- Experimentation and performance management
- Executives and managers, business analysts and data scientists, and operational and customer-facing employees (
- Strategic, operational, and tactical decisions

## **BIG DATA IN THE REAL ECONOMY – THE GROWING EVIDENCE OF BENEFITS**

---

Big data is about collecting, processing, storing, and using diverse data on an unprecedented scale, exploiting technologies that are affordable and available to enterprises. It is about monitoring fast-moving data in motion. It is also about combining structured and unstructured data. It's about a full range of data preparation, analysis, and distribution processes that are driven by the increasingly digitized interactions among and between people, organizations, and physical things.

But more than anything else, big data is about change. It's about asking new questions, using new data, new analytics, and new metrics. Instead of viewing big data only through the lens of the three Vs of data – volume, velocity, and variety – organizations also need to view it as an opportunity to combine existing and new technologies to embrace change and drive innovation.

IDC's 2013 *Big Data and Analytics Maturity Benchmark Survey* found that 46% of organizations started using new analytic techniques in the past 12-24 months. Over the same time period, 41% of organizations started using new metrics or key performance indicators (KPIs), and 27% started using new data types. As part of our research, we compared results from high achievers (organizations where the benefits derived from big data and analytics projects met or exceeded expectations) and low achievers (organizations where these benefits did not meet expectations or were not present at all). A comparison of the results from these two groups showed that maintaining the status quo can have a negative effect on achieving desired outputs from big data and analytics solutions.

This is just one recent example that correlates investments in big data and analytics with positive outcomes. There are a growing number of others based on survey or case-based market research studies. The former includes sources such as:

- The 2011 MIT study titled "Strength in Numbers: How Does Data-Driven Decision Making Affect Firm Performance?" showed that data-driven organizations are more profitable and more productive than their competitors.
- In another 2013 IDC and *Computerworld* study, 70% of respondents indicated that they had experienced tangible benefits from their big data and analytics projects, and for 90% of them, the value of benefits met or exceeded their expectations.

Given big data and analytics solution benefits, what requirements should organizations address?

- Should they focus first and foremost on the needs of data scientists or executives?
- Should they overinvest in supporting experimentation with data or in standardization and control dictated by best performance management practices?
- Should they provide decision support for strategic decisions or automate tactical decision-making processes?

The answer is yes, yes, and yes. There is a need to do it all – in the real world – not all at once but at least under the guidance of a big data and analytics strategy that considers all these requirements.

## BIG DATA AND ANALYTICS SOLUTION REQUIREMENTS

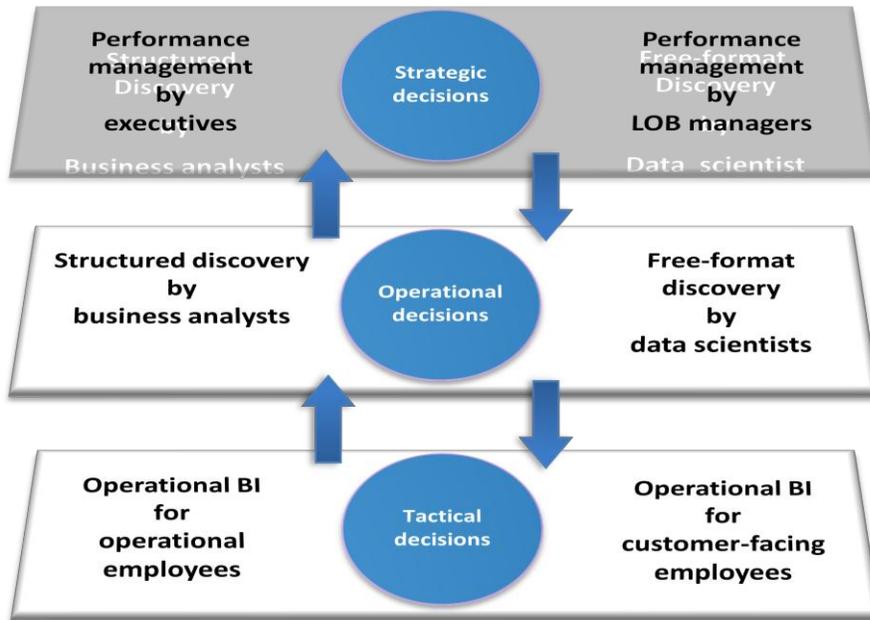
---

Most discussions of decision making assume that only senior executives make decisions or that only senior executives' decisions matter. This is a dangerous mistake. – Peter Drucker

Big data and analytics solutions have a broad reach and affect every corner of an organization from the executives to customer-facing employees. However, this breadth of applicability also means that the number of user groups, decision processes, methods of analysis, and other interactions with the data varies widely. Figure 1 shows IDC's framework for identifying and assessing the requirements of the various constituents in a typical organization and the impact big data and analytics has on these requirements. It depicts various decision-making processes, groups of decision makers, and the type of analytics or information access requirements of members of each group. There are three categories of requirements that dictate the broader set of more specific requirements.

FIGURE 1 \$

IDC's Decision Management Framework \$



Source: IDC, 2014

### Strategic Decisions

These decisions set the long-term direction for the organization, a product, a service, or an initiative and result in guidelines for operational decisions. These decisions are made by executives and line-of-business (LOB) managers and are typically supported by performance management methodologies and applications, with an emphasis on decisions based on narrowly defined KPIs and experience.

Executives and (LOB) managers typically engage in what we call performance management. These managers are interested in a few relevant KPIs that they need to access on demand and anywhere. These users typically don't perform much data analysis on their own; instead, they rely on analysts to delve deeper into questions that might arise from reviewing KPIs.

In the world of big data and analytics, the value of experience, judgment, and leadership does not disappear. However, the strategic decision-making processes are being upended by the availability of new data sources, new analytics, new metrics, and more powerful technology. For example, the ability to rapidly iterate through models of risk variables based on a broader set of internal operational data, external industry or economic data, competitive intelligence, and customer sentiment is allowing executives to make better risk-adjusted decisions about the future of their organizations. The breadth of available data and the speed with which it can be processed and analyzed are unprecedented.

In other examples, several high-profile cases have shown the negative effects of ignoring or not having right-time access to information mined from unstructured content sources such as social media or emails that often act as leading indicators of adverse information that later comes to light from analysis of historical data. Executives can eliminate reputational risk, use data-driven models to assess operational risks, and avoid regulatory risks by ensuring their organization's IT and analytics capabilities can support big data and analytics requirements for strategic decision making.

## Operational Decisions

These decisions focus on specific projects or processes. Decisions often involve the need to make changes to these processes or projects to operationalize strategic objectives. These operational decisions result in guidelines that are used to identify when tactical decisions should occur. They are supported by business analysts, data scientists, and midlevel managers who are engaged in either structured or ad hoc data discovery.

Analysts are broadly segmented into business analysts and data scientists. Both are charged with ad hoc analysis, but the former group typically does so within the structured framework of multidimensional analysis, interactive reports, or spreadsheets. The latter group performs free-format, ad hoc analysis that typically requires a mix of statistical and programming language expertise.

In the world of big data and analytics, the work of business analysts and data scientists has changed radically. The latter group now has access to intuitive, consumer-oriented visual data exploration tools. Self-service functionality has decreased or eliminated the reliance of analysts on IT for data access and manipulation. Data scientists are able to experiment on an ongoing basis using a range of structured, unstructured, and semistructured data and a combination of commercial and open source advanced analytics as well as development tools (e.g., Java). Data sampling is becoming a way of the past, and running multihour or overnight simulation is also becoming a way of the past – replaced by parallel processing of large volumes of diverse data.

One organization IDC interviewed is combining mobile device location data, time and date information, current weather data, historical customer transaction data, online browsing behavior, and customer service records to identify not only the next best action in terms of product recommendations for its customers but also when and where to send that recommendation to the customer's mobile device. In part, this company is assessing the mood of its customers by using the information it has – rainy Monday mornings in an area of poor mobile network coverage is not a good time to send out an email to a customer who recently left a bad review for one of your products.

In other examples, retailers have begun to use in-store video analytics in combination with inventory data to optimize store layouts. Even cases of the use of facial recognition exist whereby a buyer's reaction is being analyzed to a sales pitch in a retail store to help create a body of data that will enable data scientists to develop better analytics for personalized offers.

## Tactical Decisions

These decisions focus on transactions, which involve specific instances of revenue or cost. By identifying and remediating transactions that fall outside the bounds of the guidelines set by operational decisions, tactical decisions help ensure that operational objectives are met. These are supported by customer-facing or operational employees (including IT staff) and in some cases are fully automated.

In the world of big data and analytics, tactical decisions are increasingly characterized by the embedding of intelligence, derived by analysts, into operational applications. Those on the front lines are not and should not be analysts, but they are benefiting every day from recommendations, guidelines, rules, and scripts that are the outputs of tactical decision automation technologies. Today's technology is enabling organization to run transactional and analytic workloads on the same database, decreasing the amount of data movement and enabling rapid, tactical decision support. Stream processing methods are monitoring high-velocity event data and complementary rules engines applying business rules to this data and are routing information to the right people at the right time in the right place. In some cases, full automation from data monitoring to execution of an action is changing the human resources requirements of organizations.

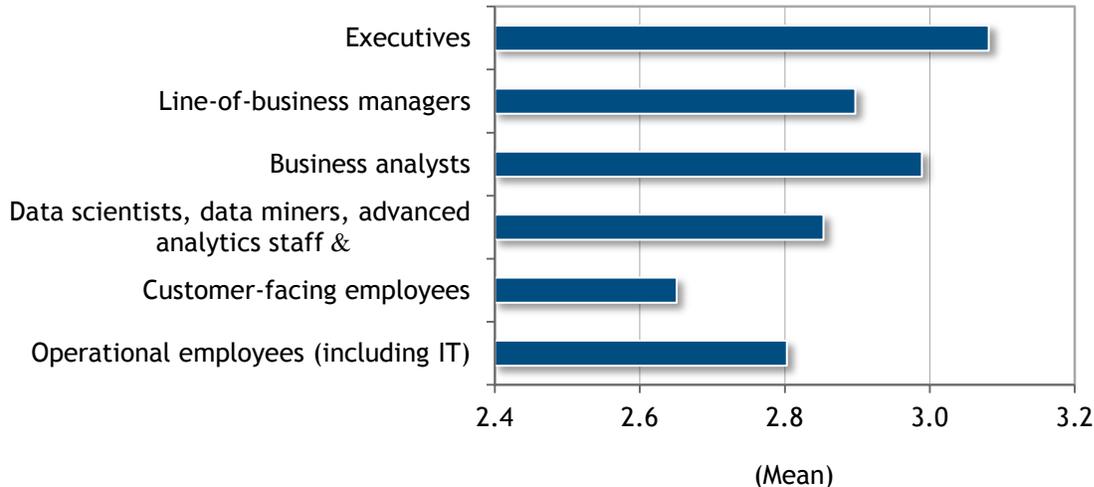
For example, a farm equipment manufacturer is monitoring the performance and location of its tractors being used in the fields, combining that stream data with data from soil samples, and weather reports to provide tactical recommendations to farmers about their operations as well as maintenance of the tractors. Similarly, airplane engine and automobile manufacturers are monitoring telematics and ongoing performance data to enable predictive maintenance that helps avoid costly catastrophic repairs. In another example, insurance companies are monitoring accident claims and extracting from text and audio content information that feeds their real-time fraud detection applications. As Figure 1 suggests, the three types of decisions don't occur in isolation. The auto manufacturers as well as logistics companies don't only monitor streaming data from vehicles, containers, or packages; they also store this granular event data for offline longitudinal analysis. In other words, data-in-motion and data-at-rest use cases are increasingly intertwined.

Although most of the focus in the big data and analytics market has been placed on data scientists in the past couple of years, it's important to take a broader view of all decision makers. Data scientists play a critical role in uncovering new insights from big data, but the output of their work is not an end in itself and needs to coexist with the analytic and decision-making processes of others within the organization. Our research, as illustrated in Figure 2, shows that the decision-support requirements of customer-facing employees are least met by the current big data and analytics solution capabilities available at their organizations. To be fair, the needs of the data scientists are not met that much better.

## FIGURE 2 \$

### Meeting End-User Requirements

Q. # To what extent does the big data and analytics technology your organization has in place meet the decision support requirements of the following user groups?



n = 330

Note: Mean scores are based on a scale of 1 to 5, where 1 = to a very limited extent and 5 = to the fullest extent needed.

Source: IDC and *Computerworld's Big Data and Analytics Survey, 2013*

Despite the fact that one of the expected benefits of big data and analytics is improved tactical decision making, most enterprises surveyed reported that those employees who would be aided the most in this regard (i.e., frontline customer-facing employees) experience the least benefit with respect to current use of big data and analytics in their organization. Most organizations still have a ways to go in fully realizing the benefits of big data and analytics through its integration into applications at every level.

## DON'T REINVENT THE WHEEL: BIG DATA AND ANALYTICS TECHNOLOGY REQUIREMENTS

In today's environment, it is not only the access to information but also the ability to analyze and act upon it in a timely manner that creates competitive advantage in the marketplace, enables sustainable management of communities and natural resources, and promotes appropriate delivery of social, healthcare, and educational services. With so many opportunities to effect positive change and drive value from big data, IT decision makers are under increasing pressure to make appropriate decisions among the many technology choices.

The user groups and their requirements outlined in this white paper highlight the need for a range of appropriate big data and analytics technology functionality as well as a combination and integration of

various capabilities. As shown in Figure 1, the enterprisewide big data and analytics requirements of ad hoc and structured discovery, performance management, and embedded operational intelligence in applications require support from a range of relational and nonrelational technology. As shown in Figure 3, the big data and analytics platform consists of three primary layers.

### Big Data Capture, Integration, and Movement

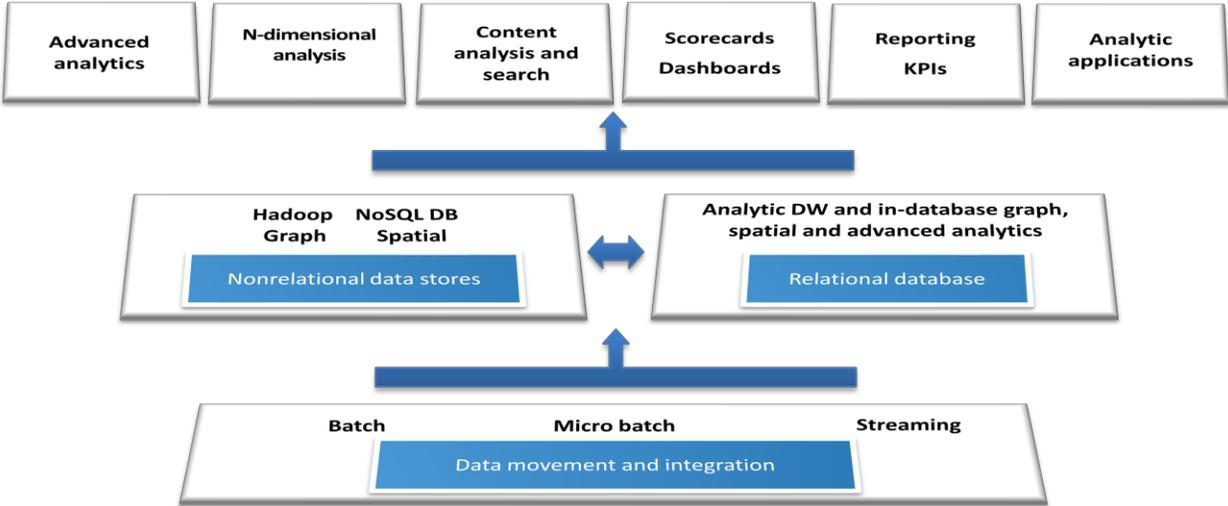
In the reference architecture, the lowest technology layer is made up of data capture, integration, and movement tools. Some of these tools provide batch data extraction, transformation, and loading, while others enable data to be streamed into target data stores. Big data and analytics requirements put tremendous scalability pressure on this layer of the platform. Moving multiterabyte data sets or processing millions of streaming events is not an anomaly for many organizations.

Additionally, this software is being utilized to move data between the various data repositories. In other words, there may be a need to capture consumer Web clickstream data, move that into a Hadoop cluster for processing, and then move a subset of the data into the data warehouse, or there may be a need to monitor streaming data from sensors while capturing and moving event patterns into a data warehouse for longitudinal analysis.

On the one hand, the big data and analytics platform must be able to provide highly reliable performance and to ensure high levels of data quality as part of the complex multistructured data transformation processes. On the other hand, the platform must be dynamically scalable to address unexpected requirements – especially from data scientists – of projects involving experimentation with new data types and sources or new combinations of existing data types and sources.

**FIGURE 3**

### Big Data and Analytics Technology Platform \$



Source: IDC, 2014

## Big Data Management and Processing

The middle layer of the big data and analytics platform is the data management and processing layer that is made up of two categories of technology: relational and nonrelational. Although most of the recent market focus has been on technologies such as Hadoop and NoSQL databases, the position of the relational database in the big data and analytics platform is secure and will remain so for the foreseeable future. On one hand, relational and nonrelational technologies address different "sweet spots" in terms of workloads. On the other hand, they must operate together in a growing number of scenarios. For example, organizations are using the relational data warehouse as the source for the "single version of the truth" – a trusted, governed, and secure source of information for performance management and structured ad hoc analysis. At the same time, organizations are using a Hadoop database or a NoSQL database (usually depending on the type of data and analysis required) for free-format discovery.

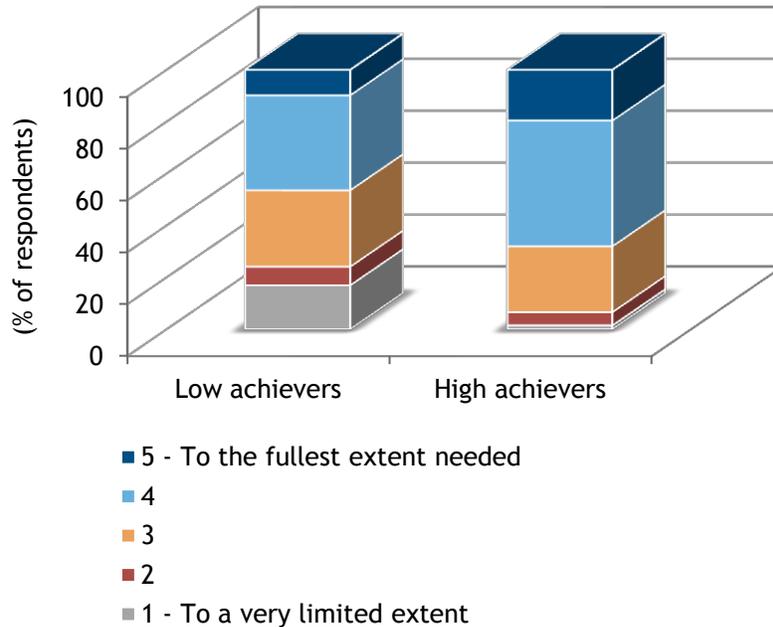
But a growing number of real-world deployment and use cases require the movement, processing, and analysis of data from both relational and nonrelational technologies. For example, a multichannel retail company utilizes its nonrelational data management technology for experimentation on a data set that is primarily made up of clickstream data and also augmented with customer reference data from the relational data warehouse. An insurance company is using audio to text translation and then text analysis along with customers' transaction data to glean insight from all customer interaction sources to enhance its fraud detection (and prevention) models. An online media company uses Hadoop to store and preprocess Web clickstream data and then moves the relevant subset of data to a relational data warehouse where the data becomes available to business analysts.

The latest IDC research shows that relational and nonrelational technologies are better together. Figure 4 shows the results from a recent IDC study of 701 large organizations that were asked about the extent to which their organization has supplemented general-purpose RDBMS technology with technologies such as Hadoop, NoSQL databases, and graph databases. When we segmented the user population into high and low achievers based on their big data and analytics project outcomes, we found that high achievers are more likely to have a big data and analytics platform that combines relational and nonrelational technology. Among high achievers, 68% have supplemented general-purpose relational technology with other big data and analytics technologies to a great extent or the fullest extent needed. A different lens on the same data shows that twice as many high achievers as low achievers have supplemented relational technology with nonrelational technology *to the fullest extent needed*.

FIGURE 4 \$

## Impact of Combining Relational and Nonrelational Big Data and Analytics Technology

Q. # To what extent has your organization supplemented general-purpose RDBMS with any specialized, fit-for-purpose, big data and analytics technologies (e.g., Hadoop, NoSQL databases, graph databases, scalable MPP databases)?



n = 701

Note: Responses were based on a scale of 1 to 5, where 1 = to a very limited extent and 5 = to the fullest extent needed.

Source: IDC's *Big Data and Analytics Maturity Benchmark Survey*, 2013

## Big Data Analytics and Applications

The upper layer of the big data and analytics technology stack includes a range of analytics and business intelligence tools and prepackaged analytic applications. Some of these tools and applications have existed for years, but there is also a new generation of tools and applications. One of the key new capabilities is the functionality to access both relational and nonrelational data. Other capabilities include incorporation of consumer-centric data visualization and interaction functionality that promotes more pervasive use of analytic and BI tools. Some tools help organizations derive meaning from unstructured text in the form of customer comments, social media interactions, documents, and emails. Combined with transactional or operational data analysis, these capabilities are enabling a new set of applications for voice of the customer analysis, warranty management, predictive maintenance, fraud detection and prevention, and many others.

In one example, a high-tech manufacturer used a combination of big data and analytics tools and applications such as real-time streaming data processing, rules management, machine learning

analytics, and data visualization to radically improve its customer interaction and product upsell and cross-sell processes. For example, it analyzed past support issues and responses and refined the parts that were sent to customers in response to service requests. In a different use case, the new big data and analytics solution provided customer service staff at the time of interaction with upsell recommendations based on previously conducted deep analysis. In a separate example, an insurance company used unstructured text analysis, predictive analytics, and rules and streaming data management to decide best content for personalized marketing campaigns. The combination of real-time and historical analyses allowed this company to figure out which customers were most likely to respond to messages containing humor versus specific cost-saving offers or other options.

## ORACLE IN THE BIG DATA AND ANALYTICS MARKET

---

Oracle, which IDC currently ranks as the world's largest business analytics vendor based on software revenue (see *Worldwide Business Analytics Software 2013-2017 Forecast and 2012 Vendor Shares*, IDC #241689, June 2013), has a broad portfolio of big data and analytics technology, supporting use cases for data at rest and data in motion and those involving structured data and unstructured content (and a mix of both). As shown in Figure 5, big data and analytics technology ranges from relational databases (with both disk and memory optimized options) to Hadoop and NoSQL data management software.

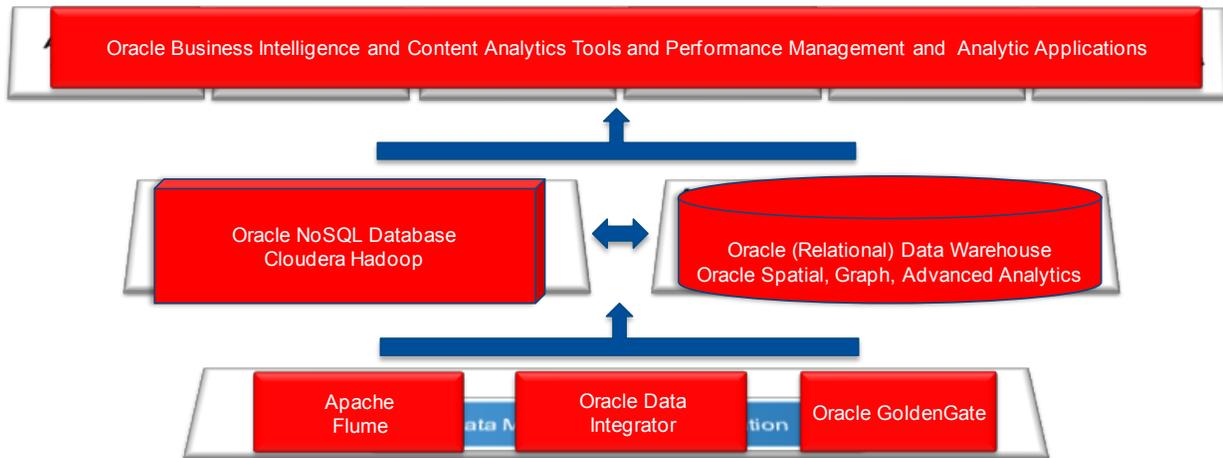
Some of this technology is available on-premise in the software-only form factor, some is available in the cloud, and some is available in the form of appliances that Oracle calls engineered systems. These include Oracle Big Data Appliance, Oracle Exadata, Oracle Exalogic, and Oracle Exalytics. In addition to its own technology, Oracle partners closely with Intel for the production of engineered systems, which are powered by Intel Xeon processor E5 and E7 families. Oracle sees these offerings as a whole that is greater than the sum of the parts. The key segments of Oracle's big data and analytics solution include:

- **Big data at rest: data reservoir and data warehouse.** Oracle Big Data Appliance, which the company calls a data reservoir, is preloaded with Cloudera Hadoop, Oracle NoSQL database, and connectors to the data warehouse. Oracle Exadata in its data warehousing configuration is well suited for both database consolidation and accelerating warehouses whose performance lags behind as a result of data growth. These two work as peers – the warehouse funneling data to the reservoir to encourage new discoveries and the reservoir enabling experimentation with new combinations of data to uncover new metrics the warehouse should serve up.
- **Big data in motion: fast capture and processing of streaming data.** Oracle Exalogic is the firm's recommended choice for real-time middleware and applications. For example, Oracle Event Processing uses rules-based processing on data streaming in from sensors and Web interactions to quickly load NoSQL databases as well as to execute data-driven actions based on real-time data in distributed in-memory caches like Oracle Coherence. One Oracle customer uses this technology for time-sensitive mobile ad targeting and delivery to consumers.
- **Big data analytics: data discovery and business intelligence.** Oracle Exalytics is the engineered system for both Oracle Endeca Information Discovery and Oracle Business Intelligence Foundation. Oracle has integrated these two products, including automatic indexing of BI metadata into Oracle Endeca, bridging the gap between relational reporting and nonrelational discovery. The BI Foundation's native integration with Hive is another indication of Oracle's commitment to deploy the relational and nonrelational worlds together.

- **Big data applications: data-driven action.** Oracle offers a wide variety of applications with embedded analytics ranging from smart meter management solutions for utilities to marketing campaign management as well as human capital management.

**FIGURE 5**

**Oracle's Big Data and Analytics Technology \$**



Note: The list of Oracle Big Data and Analytics technology depicted in the figure is for illustrative purposes only and is not meant to be exhaustive.

Source: IDC, 2014

**RECOMMENDATIONS**

The growing focus on big data and analytics solutions as a basis for competitive advantage is both an opportunity and a challenge for most organizations. The promise of better and faster data-driven decision making has pushed big data and analytic capabilities to the top of executive agendas. Many organizations, however, do not yet have the big data and analytics maturity to address the range of technology, staffing, and process requirements needed to capitalize on big data assets and to deploy analytics pervasively to optimize operational, tactical, and strategic decisions. In addition, organizations have different starting points for extending their big data and analytics capabilities. Some have existing basic business intelligence capabilities, some may have appropriate data warehousing support, and some may have deployed a Hadoop cluster but lack other big data and analytics technology capabilities. Your organization's current situation will determine next steps and ways of moving from pilot and proof-of-concept projects to broadly available and accepted solutions that drive competitive advantage or create value for all stakeholders.

IDC provides considerations derived from ongoing research with organizations worldwide in the sections that follow.

## Phase: Pilot

If your organization is at an early stage in the evolution of its big data and analytics capabilities, we recommend focusing on proof-of-concept or pilot projects where the initial value is defined through new knowledge and learning about big data and analytics technology components and their interplay. One approach could be to start with your relational data warehouse and the functionality of a NoSQL database or a Hadoop deployment. First assess the potential of such technologies in your specific case and then assess the points of integration with the data warehouse and the data integration or movement capabilities between relational and nonrelational components. At this stage, the value comes from learning what works and what doesn't in your particular environment. Initially, there may be very limited business value, but in the subsequent stages, as pilots are moved to production solutions and operationalized, the knowledge value grows and business value opportunities become visible.

For example, an insurance company wanted to improve its claims fraud detection. To do so, the company launched a project that was initially focused on capturing insights from auto accident claims made by telephone. The insurer used new technology to capture the audio files, transcribe them to text, analyze the text, and enhance its existing fraud models with the new insights gained from unstructured content analysis.

The lessons learned from this and similar cases point to the need to:

- Develop an initial department-level big data and analytics strategy or a statement of intent (e.g., need to decrease claims fraud). Budget for localized projects that have management support from department or business unit leadership. There is no such thing as a "big bang" approach to big data projects – start small and iterate.
- Launch a proof-of-concept or pilot project using existing resources. Focus on a project(s) within a specific business domain. In this case, the insurer did not attempt to revamp the whole claims process at once – it focused only on improving its existing fraud models.
- Use existing data with the recognition that it may be incomplete and may lack the necessary quality, which will require manual data preparation effort. Begin to integrate data from multiple sources to move toward high levels of trust in the information coming from the big data and analytics system. The insurance company combined audio content (transformed to text) and transactional data from its operational application based on relational technology to uncover new insights.
- Deploy new technology (on-premise or in the cloud) that is optimized for specific types of data, analytic techniques, and end-user interaction. Given the initial efforts to integrate multistructured data from multiple sources, this technology will include both relational and nonrelational tools.
- Do not spend significant time and effort to try to gain executive support before establishing initial project proof points. Look for support from colleagues with specialized big data and analytics skills. Follow up by establishing a big data and analytics team with skills in existing and newly deployed technologies. Ensure that the new technology and data allow data scientists to experiment to uncover new insights.

## Phase: Beyond Basics

If your organization has already moved beyond the basics and has one or more big data and analytics solutions in production, it's time to establish repeatable big data and analytics practices that lead to persistent use of relational and nonrelational big data and analytics technology and associated analytics. At this phase in your big data and analytics journey, business value is clearly realized but typically remains localized to individual business units.

Extending the previously started (real-world) example, the insurance company followed its initial success in creating improved fraud detection models with operationalizing them into ongoing fraud detection applications and processes. This means ensuring appropriate system performance, availability, and security, as well as a process and skilled staff for ongoing monitoring and period improvement of analytics.

The lessons learned from this and similar cases point to the need to:

- Develop cross-departmental, business-unit-level big data and analytics strategies. Budget for business unit needs. Perform a localized cost-benefit analysis for big data and analytics projects. In this case, fraud detection is only one part of an extended claims management process, and this phase of the project brought together analytic and operational staff to collaborate on operationalizing the new analytics and insights.
- Continue to expand the availability of and integrate internal multistructured data sources. Be aware that data governance policies and procedures will be difficult to implement at the single business unit level.
- Expand the availability of fit-for-purpose technology with the understanding that initial adoption will be selective.
- Assign, train, and hire staff based on the big data and analytics strategy. Augment existing skills with specialized external service providers.
- Begin to monitor and document decision processes and decision outcomes. Ensure the big data and analytics team has representatives from all stakeholder groups to facilitate collaboration. The insurance company allocates staff hours toward initial monitoring of the quality of decisions made based on new analytics as well as the ongoing monitoring of the performance of the system and its components.

## Phase: Capture Competitive Advantage

If your organization has been leveraging a portfolio of big data and analytics technology, skills, and processes, it is in a strong position to start deriving competitive advantage from its big data and analytics capabilities. To do so, your organization will need to establish management and optimization practices and methods that fully incorporate big data and analytics capabilities in the organizational culture and business processes. Enabled by demonstrated big data and analytics IT efficiency, this will lead to more predictable outcomes and transition of new product and service opportunities enabled by big data and analytics to business plans. From there, the organization will be able to move to an optimized state of continuous learning and improvement where previously unattainable business value is continuously produced.

In the next stage, the insurance company realized that improved fraud detection and prediction of potential damage reported in claims meant that some of the routine investigation work performed previously by the most experienced staff could be either eliminated or assigned to more junior staff. This freed most experienced staff to address only real exceptions not handled by increased automation.

The lessons learned from this and similar cases point to the need to:

- Develop an enterprisewide big data and analytics strategy that is championed by a C-level executive. Budget for big data and analytics projects. Make available tools and methodology for business case development and performance and outcomes measurement. The initial new fraud detection insights and related changes to claims process automation resulted in a reengineered process flow and staffing requirements, which in turn changed the company's interaction with customers. In addition, the company experienced quantifiable benefits from lowering the costs associated with fraud as well as human resources.
- Make available information about all the relevant internal and external data sources for users with the appropriate security rights. Establish metrics and methodology for data governance and metrics by which big data and analytics processes, staff, and outcomes are measured.
- Maximize the use of fit-for-purpose and workload-optimized system, automated system performance management, and dynamic scalability features of big data and analytics technology. Most, if not all, solutions will require a combination of relational and nonrelational technology. Incorporate predictive analytics into technology performance monitoring and management processes. Enable broad technology adoption by ensuring that an appropriate technology pricing structure is negotiated with IT vendors.
- Regularly provide training to all the big data and analytics technology, analytics, and business staff. Maximize big data and analytics staff centralization for functions such as data integration, systems management, and report and dashboard development while ensuring close alignment of analysts with lines of business. The insurance company was able to change the composition of staff, improve productivity, and optimize staff utilization based on experience.
- Ensure that experimentation, performance management, and operational BI processes are all supported with appropriate staffing, technology, and funding. Employ decision management techniques to enable continuous process improvement and integration of analytics into business processes.

## CONCLUSION

---

Convergence of intelligent devices, social networking, mobility, and big data and analytics is ushering in a new economic system that is redefining relationships among producers, distributors, and consumers. The flood of new (big) data, faster cycle times, and the availability of a new generation of relational and nonrelational information management and analysis technology make clear both the need and the opportunity to change how decisions are made to harness these new circumstances to achieve advantage in the market.

The big data and analytics market is evolving rapidly, and questions about the most appropriate technology for the wide range of big data and analytics use cases abound. This is the right time to raise such questions, to explore the available options, and to develop a big data and analytics strategy. However, one of these questions should not be whether to select relational or nonrelational technology for your organization's big data and analytics requirements. The answer is that you'll need both. The big data and analytics platform and the new generation of data visualization and analytics tools and applications deployed on it must be able to address:

- Structured data and unstructured content
- Data arriving in batches and streaming in continuously
- Ongoing, rapid experimentation using advanced and predictive analytics and enterprise-grade, operationalized, strictly governed, and structured information flows
- Deployment of results or outputs of the big data and analytics solution to operational systems and employees

## About IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications and consumer technology markets. IDC helps IT professionals, business executives, and the investment community make fact-based decisions on technology purchases and business strategy. More than 1000 IDC analysts provide global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries worldwide. For more than 48 years, IDC has provided strategic insights to help our clients achieve their key business objectives. IDC is a subsidiary of IDG, the world's leading technology media, research, and events company.

## Global Headquarters

5 Speen Street  
Framingham, MA 01701  
USA  
508.872.8200  
Twitter: @IDC  
[idc-insights-community.com](http://idc-insights-community.com)  
[www.idc.com](http://www.idc.com)

---

### Copyright Notice

External Publication of IDC Information and Data – Any IDC information that is to be used in advertising, press releases, or promotional materials requires prior written approval from the appropriate IDC Vice President or Country Manager. A draft of the proposed document should accompany any such request. IDC reserves the right to deny approval of external usage for any reason.

Copyright 2014 IDC. Reproduction without written permission is completely forbidden.

