BIG DATA, BIG TRANSFORMATIONS

BIG DATA SUCCESS REQUIRES VISION, TALENT, AND TECHNOLOGY

Competitive enterprises that embark on big data strategies do so with the expectation that their businesses will transform. They don’t just want answers from the data they collect and analyze, they want results. Be it with small, fledgling trials or large, cross-functional efforts, these enterprises want to see clearly how big data can make a difference—with their customers, their processes, their bottom lines and, most important, with growing the business.

However, many organizations that appreciate the advantages big data can offer, and some that already have a few big data projects under their belts, remain unclear on how to get maximum results from these initiatives. In 2012, McKinsey & Company conducted a survey of 1,469 executives across various regions, industries and company sizes, in which 49 percent of respondents said that their companies are focusing big data efforts on customer insights, segmentation and targeting to improve overall performance.1 An even higher number of respondents—60 percent—said their companies should focus efforts on using data and analytics to generate these insights. Yet, just one-fifth said that their organizations have fully deployed data and analytics to generate insights in one business unit or function, and only 13 percent use data to generate insights across the company.

As these survey results show, the question is no longer whether big data can help business, but how can business derive maximum results from big data.

"A lot of enterprises understand that big data is something they need to tackle, but the challenge they experience now is how to do it effectively," says Generosa Litton, director of big data marketing with EMC. "They know the advantages, but getting to the maximum benefits is something they’re still grappling with. In order to clear that hurdle, company executives should view big data as a means for meeting business initiatives. They should align big data projects with tangible goals."

Much of the apprehension around big data projects is not about the technology that can provide insights to pressing questions or problems. Rather, it’s about having the right management structure and business processes in place to implement what’s been learned and begin to transform the business. Big data itself can’t produce business results; what’s needed is a clear vision of how the business can gain results, using big data to help get there. That’s where value is created.

The roadmap to results
Getting results from big data doesn’t happen overnight, and can’t be done in a vacuum. Enterprises must approach big data projects with forethought and preparation. They must ensure that all stakeholders are on board and invested in success, and that the right infrastructure and staffing pieces are in place. By combining the following elements in their big data strategies, enterprises are much more likely to maximize their investments, shorten time to insight and uncover business value:

VISION—Successful big data strategies start at the top with executive buy-in, sponsorship and support. Executives must set the tone for what strategic outcomes are expected from a big data project, and remain involved so that the vision stays central to the project. A big data project can require involvement from many different departments of the company; for example, marketing, finance, customer service and IT could all be involved in understanding the impact of a certain marketing campaign. Executives must put in place flexible management structures to allow big data projects to flow across functions, and resist siloing data and employees by department or other limiting factors.

METRICS—While top executives should set the vision for big data results, line-of-business managers are best suited to frame the project with metrics that offer a detailed description of success. This means moving from evasive goals, such as “boosting product sales in the Southwest,” to more specific descriptors of success, such as “increasing product sales by 10 percent in Arizona and New Mexico during the winter months.” By putting metrics around a project, companies can more easily determine and quantify its success and translate that success to the impact it has on revenue. Having quantifiable results is useful when digging further into big data to reveal additional insights or launching new big data projects in other areas of the business.

TECHNOLOGY—Any successful big data project needs the right infrastructure in place, from storage all the way up to applications:

- Big data storage environments store structured and unstructured data and can scale out to meet the size, scope and velocity demands of big data.
- Big data analytics platforms analyze structured and unstructured data without requiring data to be transferred, therefore avoiding latency associated with having to move the data around. They also support flexible yet powerful analytics to enable faster insight through self-service and collaboration.
- Big data applications are purpose-built and customized to distribute insights throughout the organization, and created with agile programming techniques that speed development and promote collaboration.

“A lot of business users haven’t been able to envision the realm of the possible; they may know what questions they are trying to answer and decisions that need to be made, but they just don’t understand what some of the new data sources can do for them.”

— Bill Schmarzo, CTO of enterprise information management and analytics service, EMC Consulting
TALENT—In addition to the right technology, big data projects require trained professionals to leverage resources and deliver results. Data scientists possess the skills and expertise that go beyond business intelligence to generate business insight.

“The data scientist is really going to help the business stakeholders establish a vision for what can be done with data,” says Annika Jimenez, senior director for analytics solutions with EMC Greenplum. “It’s someone who can leverage predictive models and machine learning to push the limits of what can be done with data.”

SERVICES AND EDUCATION—While some companies, like Facebook and Google, grew up on big data and have built their organizations to take advantage of it, most have not. To succeed with big data projects, many enterprises need a partner who can work with both the IT department and line-of-business managers to make the most of the information they collect.

“A lot of business users haven’t been able to envision the realm of the possible; they may know what questions they are trying to answer and decisions that need to be made, but they just don’t understand what some of the new data sources can do for them,” says Bill Schmarzo, CTO of enterprise information management and analytics service with EMC Consulting. “They need a partner with industry expertise who understands how big data impacts them, who can explore the realm of the possible and identify use cases.”

EMC solutions drive big data results
EMC understands all aspects of big data and works in partnership with enterprises to make their vision and goals a reality. By leveraging EMC solutions that put data in context and make it part of all decision-making processes, enterprises gain the insight needed to transform into predictive, agile organizations.

The Big Data solution from EMC—This set of technologies from EMC provides scale-out storage, a unified analytics platform and business process and application development tools to produce results.

EMC Isilon is a storage platform for big data designed to easily scale capacity over 15 petabytes within a single cluster while greatly simplifying storage and data management. The Isilon scale-out network-attached storage (NAS) platform provides a highly efficient and resilient storage infrastructure that scales in both capacity and performance, simplifying and reducing storage management to a few hours a week while optimizing resources with automated, tiered storage capabilities. As the first and only scale-out storage platform with native HDFS support, Isilon NAS is well suited to support Hadoop-based analytics of large-scale, file-based unstructured data. With its flexible, multiprotocol support and record NAS performance, EMC Isilon is an ideal storage platform to consolidate big data environments while supporting a wide range of applications and workloads, including data analytics. Isilon scale-out storage provides more than 80 percent storage utilization to help enterprises maximize ROI while reducing capital expenditures as well as operational costs.

EMC Unified Analytics Platform consists of the Greenplum Database, Greenplum HD and Greenplum Chorus. This unified analytics platform enables analysis of structured and unstructured data, providing insights to make informed decisions.
unstructured data and provides a productivity layer for data science teams. Combining analytics into one system and performing analysis against all data, the Greenplum Unified Analytics Platform (UAP) enables more accurate predictions and more precise model development. The massively parallel processing (MPP) Greenplum Database manages, stores and analyzes petabytes of structured data. It quickly loads data and process queries using the computing power of all the nodes in the cluster. As a result, users can continuously optimize their predictive models, resulting in better sales, lower costs and less risk. Greenplum HD is an enterprise software framework based on The Apache Foundation’s Hadoop that lets customers process and run analytics on massive amounts of unstructured data. Greenplum Chorus is a self-service and collaborative platform that helps data science teams create sandboxes using self-service provisioning, and instantly start analyzing data and sharing insights. It also includes social network features that allow all stakeholders to participate and collaborate on big data projects.

Tokyo-based BrainPad, a platform-as-a-service (PaaS) provider of analytical applications and services designed to help social gaming companies analyze in-game data, turned to Greenplum when its existing database hit the wall. When BrainPad released its data analysis PaaS, demand far exceeded expectations. While this was good news for the business, the infrastructure wasn’t up to the task of meeting the demand. The installed database was taking more than two days to load data and upwards of a week to load large volume sets, resulting in significant performance hits for BrainPad’s customers.

BrainPad installed the Greenplum Database, which features a shared-nothing, MPP architecture to resolve I/O bottlenecks, and quickly saw greater performance for analyst queries. The time needed to load large volumes of raw data has been reduced from more than a week to two hours.

“Our customers’ data scientists can be highly responsive to changes in consumer behavior because they can get answers from the database within minutes of the data being loaded,” says Seinosuke Sato, co-founder and director of BrainPad. “Greenplum is optimized for fast query execution and supports our goal of making it easy for our customers to remain focused on improving their marketing ROI.”

EMC Greenplum Pivotal Labs helps enterprises complete the “last mile” of the big data analytics process through the development of custom, data-driven applications. These applications help enterprises that have uncovered insight with Greenplum UAP to put that knowledge to use, which is critical to the overall success of the analytics project. Because off-the-shelf applications for big data don’t exist, and the insights gleaned from big data can’t be simply plugged into existing BI applications, enterprises are faced with the challenges of building custom software. Pivotal Labs makes this development process faster and less expensive, and in the end enables development teams to deliver better software with a higher quality of code. By providing project-based consulting engagements to EMC customers, Pivotal is removing the common challenges of developing custom software and enabling data-driven enterprises to build the next generation of big data applications.

EMC Documentum xCP provides a next-generation development platform that simplifies how organizations create, deploy and embed real-time analytics in applications. Documentum xCP leverages information from a variety of big data sources that are process-centric and decision-oriented. This allows organizations to use graphical application composition rather...
than requiring them to write custom code. Documentum xCP increases business user agility and IT staff productivity by improving processes and applications with big data insights—helping them to easily adapt to evolving business requirements. By accessing and exchanging data and content across multiple enterprise systems and applications, xCP enables organizations to provide end users with a single view of all relevant information in context—regardless of its source, location or format. In addition, xCP-built applications can easily aggregate and integrate Greenplum analytics with content and process analytics to provide valuable insight throughout an enterprise for improved decision making.

**EMC is the right partner**

Beyond technology, enterprises often need to partner with experienced professionals on their big data projects to arrive at the desired results. Education and training are key to bringing their staff members up to speed.

EMC Consulting works with enterprises to unlock the business value in big data. EMC professionals prepare clients to think critically about their greatest challenges and to develop strategies that rationalize and optimize the use of growing amounts of information. EMC Consulting’s 2,000 consultants across the globe have an average of 17 years consulting experience with specific expertise in business and industry challenges to provide the strategy, execution and business acumen to deliver on the promise of big data.

**EMC Consulting’s Big Data Vision Workshop** helps enterprises define where and how big data and advanced analytics can transform an organization from both a competitive and a financial standpoint. The on-site workshop leverages group dynamics and envisioning exercises to identify big data opportunities, and discover how big data and advanced analytics can deliver results.

The Vision Workshop is designed to achieve these objectives:

- Drive agreement from all participants regarding the prioritization of identified business opportunities from a financial perspective. This could be expressed in hard dollars, return on investment or ranking versus other projects.
- Drive agreement regarding feasibility of implementation of each of the proposed business opportunities, based on technology, project and organizational parameters.
- Drive agreement on the ranking of each business opportunity against the criteria of business value and implementation feasibility.

**EMC Greenplum Analytics Labs** brings Greenplum data scientists on-site to help develop an analytics roadmap and kick-start analytics projects. By combining services, training and, in some cases, hardware and software, these unique labs partner with analysts, data platform administrators and business leadership to solve top business challenges and find new opportunities in data, all on an accelerated schedule.

**EMC Data Science and Big Data Analytics training and certification programs** are designed to build a solid foundation for data analytics with a particular focus on the opportunities and challenges presented by big data. Data science professionals/teams are needed to help shepherd their organizations to realize increased success and the full potential from big data.

**Conclusion**

Companies with the right infrastructures, talent and vision in place are well equipped to take their big data strategies to the next level and transform their businesses. They can use big data to unveil new patterns and trends, gain additional insights and begin to find answers to pressing business issues.

The deeper organizations dig into big data and the more equipped they are to act upon what’s learned, the more likely they are to reveal answers that can add value to the top line of the business. This is where the returns on big data investments multiply and the transformation begins. Harnessing big data insight delivers more than cost cutting or productivity improvement—it reveals new business opportunities.

**Next Steps**

To learn more about EMC’s big data solution and how it can help transform your business, please visit:

- [www.EMC.com/BigData](http://www.EMC.com/BigData)
- [www.EMC.com/Isilon](http://www.EMC.com/Isilon)
- [www.PivotalLabs.com](http://www.PivotalLabs.com)
- [www.EMC.com/xCP](http://www.EMC.com/xCP)
- [www.EMC.com/services/](http://www.EMC.com/services/)