Solution Spotlight

Big Data Analytics – Your Project Reality: The Good, the Bad, and the Ugly
“Big data” analytics is hot. Read any IT publication or website and you’ll see business intelligence (BI) vendors and their systems integration partners pitching products and services to help organizations implement and management big data analytics systems. The ads make you think it’s easy—that all you need for a successful deployment is a particular technology. But Big data projects are easier said than done. In this E-Guide, explore the good, the bad, and the ugly sides of a big data analytics project.

‘Big Data’ Analytics Projects Easier Said than Done – but Doable
By: Lyndsay Wise

“Big data” has become one of the most talked-about trends -- and yes, buzzwords -- within the business intelligence (BI), analytics and data management markets. A growing number of organizations are looking to BI and analytics vendors to help them answer business questions in big data environments. Unfortunately, gaining visibility into pools of big data is easier said than done. And with vendors marketing a wide variety of technology offerings aimed at addressing the challenges of big data analytics projects, businesses may be hard-pressed to identify the one that best meets their needs.

So, what is big data -- really? A recent story by the IT publication eWeek offered the following take on it, based partly on Gartner Inc.’s definition of the term: “Big data refers to the volume, variety and velocity of structured and unstructured data pouring through networks into processors and storage devices, along with the conversion of such data into business advice for enterprises.”

That hits the mark in terms of data management and the analytics part of the equation, but it misses the essential aspect of the business challenges surrounding big data: complexity. For instance, big data installations often
involve information -- from social media networks, emails, sensors, Web activity logs and other data sources -- that doesn’t fit easily into traditional data warehouse systems.

And in many cases, all of that disparate data needs to be pulled together in order to make sense of it on a broader level. That can have big implications for business rules, table joins and other components of big data analytics systems. The complexity of big data is what really makes it different from more conventional data when it comes to storage and query management, and it’s the main reason why analytical database and data analytics software vendors have had to step up their game to help companies deal with big data.

Understanding big data is the first step in assessing your technology needs and putting a big data analytics plan in place. The second is understanding the market and the current trends that are affecting organizations looking to derive business value, and competitive advantages, from increasingly large and diverse data sets.

**Big agendas for big data analytics projects**

Many businesses have always had large data sets, of course. But now, more and more companies are storing terabytes and terabytes of information, if not petabytes. In addition, they’re looking to analyze key data multiple times daily or even in real time -- a change from traditional BI processes for examining historical data on a weekly or monthly basis. And they want to process more and more complex queries that involve a variety of different data sets. That might include transaction data from enterprise resource planning and customer relationship management systems, plus social media and geospatial data, internal documents and other forms of information. Increasingly, companies also want to give business users self-service BI capabilities and make it easier for them to understand analytical findings.

With vendors marketing a wide variety of technology offerings aimed at addressing the challenges of big data analytics projects, businesses may be hard-pressed to identify the one that best meets their needs.
All of that can play into a big data analytics strategy, and technology vendors are addressing those needs in different ways. Many database and data warehouse vendors are focusing on the ability to process large amounts of complex data in a timely fashion. Some are using columnar data stores in an effort to enable quicker query performance, or providing built-in query optimizers, or adding support for open source technologies such as Hadoop and MapReduce.

In-memory analytics tools may help speed up the analysis process by reducing the need to transfer data from disk drives, while data virtualization software and other real-time data integration technologies can be used to assemble information from disparate data sources on the fly. Ready-made analytics applications are being tailored to vertical markets that routinely have to deal with big data -- for instance, the telecommunications, financial services and online gaming industries. Data visualization tools can simplify the process of presenting the results of big data analytics queries to corporate executives and business managers.

Organizations that fit into the categories described above in relation to their data and analytics needs should begin by considering the following issues and questions, among others, before creating an implementation plan and finalizing their big data infrastructure choices:

- The required timeliness of data, as not all databases support real-time data availability.
- The interrelatedness of data and the complexity of the business rules that will be needed to link various data sources to get a broad view of corporate performance, sales opportunities, customer behavior, risk factors and other business metrics.
- The amount of historical data that needs to be included for analysis purposes. If one data source contains only two years of information but five are required, how will that be handled?
- Which technology vendors have experience with big data analytics in your industry, and what is their track record?
Contents

‘Big Data’ Analytics Projects Easier Said than Done — but Doable

The Wrong Way: Worst Practices in ‘Big Data’ Analytics Programs

- Who is responsible for the various data entities within an organization, and how will those people be involved in the big data analytics initiative?

Those considerations don’t constitute in-depth requirements planning, but they can help businesses get started on the road to deploying a big data analytics system and identifying the technology that will best support it.

About the Author:
Lyndsay Wise is president and founder of WiseAnalytics, an independent analyst firm based in Toronto that focuses on business intelligence, master data management and unstructured data management.

The Wrong Way: Worst Practices in ‘Big Data’ Analytics Programs

By: Rick Sherman, Contributor

“Big data” analytics is hot. Read any IT publication or website and you’ll see business intelligence (BI) vendors and their systems integration partners pitching products and services to help organizations implement and manage big data analytics systems. The ads and the big data analytics press releases and case studies that vendors are rushing out might make you think it’s easy -- that all you need for a successful deployment is a particular technology.

If only it were that simple. While BI vendors are happy to tell you about their customers who are successfully leveraging big data for analytics uses, they’re not so quick to discuss those who have failed. There are many potential reasons why big data analytics projects fall short of their goals and expectations. You can find lots of advice on big data analytics best practices; below are some worst practices for big data analytics programs so you know what to avoid.

“If we build, it they will come.” This repeats the classic mistake made when organizations develop their first data warehousing or BI system. Too
often, IT as well as BI and analytics program managers get sold on the technology hype and forget that business value is their first priority; data analysis technology is simply a tool used to generate that value. Instead of blindly adopting and deploying something, big data analytics proponents first need to determine the business purposes that would be served by the technology in order to establish a business case -- and then choose and implement the right analytics tools for the job at hand. Without a solid understanding of business requirements, the danger is that project teams will end up creating a big data disk farm that really isn't worth anything to the organization, earning them an unwanted spot in the “data doghouse.”

**Assuming that the software will have all the answers.** Building an analytics system, especially one involving big data, can be complex and resource-intensive. As a result, many organizations hope the software they deploy will be a silver bullet that magically does it all for them. People should know better, of course -- but still they have hope. Software does help, sometimes dramatically. But big data analytics is only as good as the data being analyzed and the analytical skills of those using the tools.

**Not understanding that you need to think differently.** Often, people keep trying what has worked for them in the past, even when confronted with a different situation. In the case of big data, some organizations assume that “big” just means more transactions and large data volumes. It may, but many big data analytics initiatives involve unstructured and semi-structured information that needs to be managed and analyzed in fundamentally different ways than is the case with the structured data in enterprise applications and data warehouses. As a result, new methods and tools might be required to capture, cleanse, store, integrate and access at least some of your big data.

**Forgetting all the lessons of the past.** Sometimes enterprises go to the other extreme and think that everything is different with big data and they have to start from scratch. This mistake can be even more fatal to a big data analytics project’s success than thinking that nothing is different. Just because the data you’re looking to analyze is structured differently doesn’t mean the fundamental laws of data management have been rewritten.
Not having the requisite business and analytical expertise. A corollary to the misconception that the technology can do it all is the belief that all you need are IT staffers to implement big data analytics software. First, in keeping with the theme of generating business value mentioned above, an effective big data analytics program has to incorporate extensive business and industry knowledge into both the system design stage and ongoing operations. Second, many organizations underestimate the extent of analytical skills that are needed. If big data analysis is only about building reports and dashboards, enterprises likely can leverage their existing BI expertise. However, big data analytics typically involves more advanced processes, such as data mining and predictive analytics. That requires analytics professionals with statistical, actuarial and other sophisticated skills, which might mean new hiring for organizations that are making their first forays into advanced analytics.

Big data analytics is only as good as the data being analyzed and the analytical skills of those using the tools.

Treating the project like it’s a science experiment. Too often, companies measure the success of big data analytics programs merely by the fact that data is being collected and then analyzed. In reality, collecting and analyzing the data is just the beginning. Analytics only produces business value if it is incorporated into business processes, enabling business managers and users to act upon the findings to improve organizational performance and results. To be truly effective, an analytics program also needs to include a feedback loop for communicating the success of actions taken as a result of analytical findings, followed by a refinement of the analytical models based on the business results.

Promising and trying to do too much. Many big data analytics projects fall into a big trap: Proponents oversell how fast they can deploy the systems and how significant the business benefits will be. Over-promising and under-delivering is the surest way to get the business to walk away from any technology, and it often sets back the use of the particular technology within an organization for a long time -- even if many other enterprises are
achieving success. In addition, when you set expectations that the benefits will come easily and quickly, business executives have a tendency to underestimate the required level of involvement and commitment. And when a sufficient resource commitment isn’t there, the expected benefits usually don’t come easily or quickly -- and the project is labeled as a failure.

Big data analytics can produce significant business value for an organization, but it also can go horribly wrong if you aren’t careful and don’t learn from the mistakes made by other companies. Don’t be the next poster child for how not to manage a big data analytics deployment.

About the Author

Rick Sherman is the founder of Athena IT Solutions, which provides consulting, training and vendor services on business intelligence, data integration and data warehousing. Sherman has written more than 100 articles and spoken at dozens of events and webinars; he also is an adjunct faculty member at Northeastern University’s Graduate School of Engineering. He blogs at The Data Doghouse and can be reached at rsherman@athena-solutions.com.
Free resources for technology professionals

TechTarget publishes targeted technology media that address your need for information and resources for researching products, developing strategy and making cost-effective purchase decisions. Our network of technology-specific Web sites gives you access to industry experts, independent content and analysis and the Web’s largest library of vendor-provided white papers, webcasts, podcasts, videos, virtual trade shows, research reports and more—drawing on the rich R&D resources of technology providers to address market trends, challenges and solutions. Our live events and virtual seminars give you access to vendor neutral, expert commentary and advice on the issues and challenges you face daily. Our social community IT Knowledge Exchange allows you to share real world information in real time with peers and experts.

What makes TechTarget unique?

TechTarget is squarely focused on the enterprise IT space. Our team of editors and network of industry experts provide the richest, most relevant content to IT professionals and management. We leverage the immediacy of the Web, the networking and face-to-face opportunities of events and virtual events, and the ability to interact with peers—all to create compelling and actionable information for enterprise IT professionals across all industries and markets.

Related TechTarget Websites

➤ SearchContentManagement
➤ SearchDataManagement