10 Observations on Big Data In Late 2015

📅 12/11/2015   👁 222   0 👍 Like

Tom Davenport **(http://www.grroups.com/blog_authors/detail/Tom-Davenport)**
Professor at Babson College

Opinions expressed by Grroups members are their own.

(http://www.grroups.com/blog_authors/detail/Tom-Davenport)

See all posts from Tom Davenport (http://www.grroups.com/blog_authors/detail/Tom-Davenport)

Top Authors

Kumar   La

1	2	3	4	5	6	7	8	9	10
0	0	0	0	0	0	0	0	0	0

10 Observations on Big Data In Late 2015

The term "big data" has become nearly ubiquitous. Indeed, it seems that every day we hear new reports of how some company is using big data and sophisticated analytics to become increasingly competitive. The topic first began to take off in late 2010 (at least according to search results from Google Trends) and, now that we're approaching a five-year anniversary, perhaps it's a good time to take a step back and reflect on this major approach to doing business. This article describes 10 of my observations about big data.

1.The term "big data" is entering a decline.

According to Google Trends analyses, the words "big data" have been trending downward for the last six months or so. This decline is somewhat deserved, given that the term never had a precise meaning in the first place. It encompassed data that was too big to fit on traditional servers, too fast-moving to be segregated and analyzed in a data warehouse, and/or too unstructured to fit into the row-and-column format of relational databases. It was never clear whether the "and/or" in the previous sentence should really have been just "and" or just "or" (my preference). And, it was never really established what "big data" didn't include, because few (if any) ever identified themselves as working with "small data."

2. The term "big data" might be in decline, but the concept is still extremely valid (and will likely remain so).

The pace of data creation continues to accelerate and, according to IDC, the global market intelligence firm, less than 0.5 percent of all global data is analyzed in any way. Meanwhile, new data types are becoming commonplace all the time, while increasing numbers of organizations are attempting to adopt data- and analytics-driven decisions. In short, the importance of data and analysis to human beings will not be fading away anytime soon, if ever.

3. "Big data" really means "all data"—or just "data."

Large-volume, relatively unstructured data were never inherently more valuable than smaller, easier-to-analyze transaction data. Indeed, both types are equally important. To understand customers and predict their behavior, for example, a company needs to employ small data (such as sales records of what a customer has bought in the past) as well as big data (such as information from customer clickstreams and social media). In other applications as well, making a hard-and-fast distinction between big data and everything else really makes no sense. As such, most executives who refer to big data really mean all data.

4. The technology to manage big data is handy (and inexpensive) for all types of data.

People tend to think of "big data" as only the information itself, but an entirely new class of technology (mostly open source tools) has arisen for managing that data. These types of programs, including Hadoop (http://www.grroups.com/what-is/Hadoop) and related technologies, can be helpful in storing and manipulating— and, in some cases, analyzing—all kinds of data. They also are substantially less expensive to acquire than traditional technologies, and organizations of all types and sizes can benefit from their price-performance.

5. Nevertheless, companies have been holding onto traditional tools.

Although open-source tools for big data are growing rapidly, I have yet to find a company that has unplugged its traditional data warehouse or canceled its subscription to commercial analytics tools such as the SAS software suite (from SAS Institute) or SPSS Statistics (acquired by IBM). Instead, new tools are being combined with the old in hybrid architectures. In the future, most companies will have multiple ways to store and analyze data, and the tool chosen will vary by data type and application. I do expect, however, that the growth curves for commercial software vendors in the analytics space will flatten out unless they develop radically improved technologies.

6. The skills to deal with big data are becoming much more widely distributed.

The primary challenge to implementing big data and the associated technologies was human resources: finding people, who came to be known as data scientists,[1] with the necessary skills. But, the shortage of these skills is abating rapidly. There are now more than 100 university programs (most of them master's degrees) in data science and analytics, with several universities offering more than one. Before long, the supply of data scientists will fulfill the demand.

7. Data products remain a critical objective.

One of the important business developments associated with the big data movement was that companies began to offer their customers products and services based on data and analytics. Originally these firms were born digital companies such as Google, Facebook, and LinkedIn. Later, such traditional industrial corporations as General Electric, Monsanto, and NCR began joining the action. Monsanto, for example, now offers "precision planting" advice to farmers. Several large banks also have pursued similar approaches, and many companies today are examining their data assets and wondering how they might monetize that resource or provide additional value to their customers.

8. Big data starts to move beyond simple descriptive statistics.

In the early days, big data often equaled small math, as companies spent most of their time simply getting unstructured data into a form that could then be analyzed. Such efforts produced relatively simple visual analytics— good for facilitating executive understanding of results, but not useful for portraying multivariate statistical relationships. But, as companies have gotten better and more efficient at structuring data, they've been able to incorporate that information into more complex models and algorithms.

9. New sources of large-scale, unstructured data continue to emerge.

Just as companies have begun to master information such as text and genetic data, the Internet of Things data and drone imagery have become increasingly available in massive quantities. Each new type of data demands not only new approaches to analysis but also new approaches to structuring and storing that information. Moreover, all these new data types have to be integrated with other existing data. As such, companies are not likely to run out of work dealing with data anytime soon.

10. The resource in shortest supply continues to be management awareness and understanding of the business potential of big data.

Hardware, software, and programming or analytics skills are all becoming increasingly available, but many managers still don't know the true value of big (and small) data. Although there are abundant university programs teaching big data and analytics skills, there are far fewer executive education programs that provide a sophisticated understanding of big data capabilities. Because senior management sponsorship is necessary to proceed aggressively with this resource, educating those individuals is arguably the most important thing that organizations can do.

"Big data" still might need to be defined more precisely, but without a doubt the concept has been gaining considerable traction at many companies. Yet even as significant progress is being made with regard to this important business resource, the growing use of big data and analytics has not been without its challenges. As has often been the case in the past, the technology itself may now be outpacing the ability of organizations to deploy and manage it effectively.
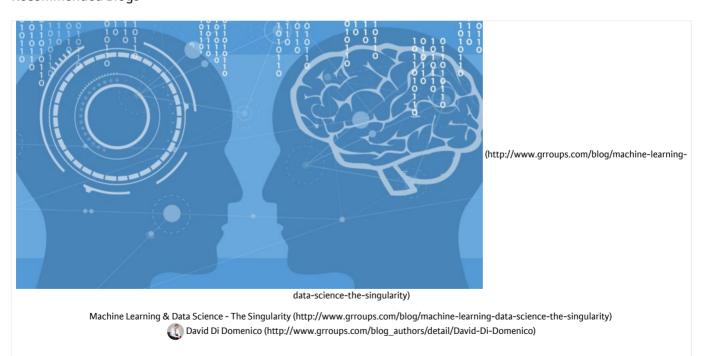
> About Author: Tom Davenport (@tdav) is a Distinguished Professor at Babson College, a research fellow at the Center for Digital Business, director of research at the International Institute for Analytics, and a senior adviser to Deloitte Analytics.
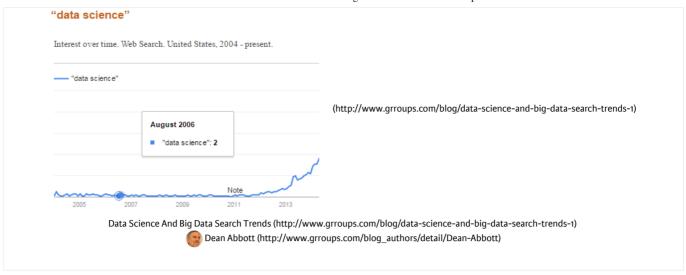
* This article was originally published on Babson College (http://www.babson.edu/executive-education/thought-leadership/premium/Pages/10-observations-on-big-data.aspx), November 2015.◻
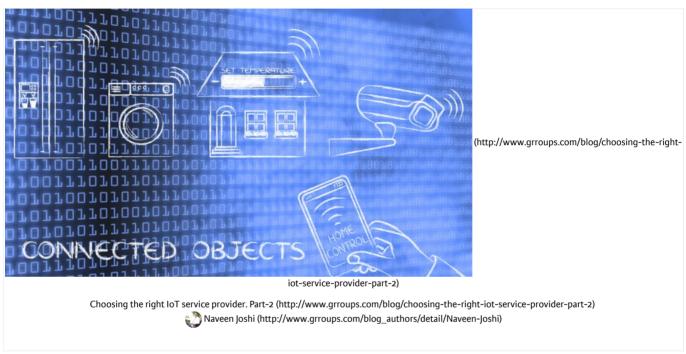
📅 12/11/2015   👁 222   0 👍 Like

<< Prev Blog (http://www.grroups.com/blog/advanced-regression-models-robust-regression)

Next Blog >> (http://www.grroups.com/blog/advanced-regression-models-robust-regression)

Recommended Blogs

(http://www.grroups.com/blog/machine-learning-data-science-the-singularity)

Machine Learning & Data Science - The Singularity (http://www.grroups.com/blog/machine-learning-data-science-the-singularity)
David Di Domenico (http://www.grroups.com/blog_authors/detail/David-Di-Domenico)

(http://www.grroups.com/blog/data-science-and-big-data-search-trends-1)

Data Science And Big Data Search Trends (http://www.grroups.com/blog/data-science-and-big-data-search-trends-1)

Dean Abbott (http://www.grroups.com/blog_authors/detail/Dean-Abbott)

(http://www.grroups.com/blog/choosing-the-right-iot-service-provider-part-2)

Choosing the right IoT service provider. Part-2 (http://www.grroups.com/blog/choosing-the-right-iot-service-provider-part-2)

Naveen Joshi (http://www.grroups.com/blog_authors/detail/Naveen-Joshi)