The Big Data Continuum:

From Data Scientists to Empowered Business People

By: Connie Moore

Highlights

DIGITAL CLARITY GROUP

Executive Summary

The amount of words that have been uttered and written about big data over the past four years is astounding. If big data hasn't become a big topic in your organization, it will. And it's probably sparked many conversations, even if the technology hasn't arrived yet at your firm.

At this point you may be wondering: What's the big deal? Is this hype or reality? And if it's reality, how do organizations turn big data to their advantage? What's to keep the big data swamp from swamping the organization it's supposed to help? And what's IT's role, where do data scientists fit in and how do ordinary business people get empowered by big data? Plus, how does it work in, say, marketing organizations where CMO careers live or die based on marketing's ability to slice and dice information and then make real-time decisions that impact the top and bottom line?

This report provides an introductory overview of how big data and advanced analytics tools are empowering business people to gain insight and take action in real-time situations. It describes some of the business applications for data discovery and visualization, and the business benefits of using these tools. At the same time, the report takes an honest look at how unrealistic it would be for organizations to replace their data scientists with business people who are powered by analytics so easy to use that workbenches, libraries, best practices, and guidance are no longer needed. These tools, while powerful, have not advanced to that stage, and may not get there for a long time, if ever.

Big Data Has Come A Long Way In A Short Time

The past five years have seen tremendous strides in big data tools and the analytics that consume big data. From the bottom up, there's in-memory database systems (IMDB) that operate much faster than disk oriented DBMSs by combining main memory for data storage with internal optimization algorithms. IMDBs have made significant inroads in data analytics because of greater performance and lower cost. Moving up the stack, HADOOP, an open source software project that supports the distributed processing of large data sets across clusters of commodity servers, speeds the processing of big data.[1]

Going up another notch, advanced analytics has seen big breakthroughs that parallel the advances in big data. The predecessor to advanced analytics, traditional BI, requires hard-to-find BI specialists within IT to build time-consuming reports that are difficult for business people to understand and virtually impossible to use in real-time settings. With traditional BI, a business person who needs a report submits a request to IT and a few weeks later IT provides it. Then if the business person needs different views of the data, IT gets involved again, and the business person must wait another two to three weeks for the changes. Traditional BI skills are so difficult to find that many CIOs and COOs create a centralized BI service that supports the entire organization, making it very hard to get timely turnaround.

In contrast, today's analytics cover a spectrum of products for a combination of data scientists, data analysts, and business people (like marketers) operating in real time. Workbenches for data scientists make it possible to quickly create dashboards that data analysts and business people use with little training. And these easy-to-use analytics give ordinary business people powerful ways to visualize large data sets and take actions in real-time, mobile environments. One key benefit is reducing the time to develop insights and take action. For example, traditional BI tools rely on reports that take time to build, generate, and interpret, while advanced analytics tools from companies like Adobe and Qlik take action in a fraction of a second. Plus, the more IT empowers the business to use analytics, the more the business can transform its processes.

Ironically, in a big data world, the explosive amount of data isn't the big deal. After all, some companies analyze billions of rows in two seconds. Instead, knowing the importance of specific data, while discounting all the other data as distractions is what really matters. As a result, data discovery and visualization is the most explosive segment of the analytics market. And these new and still developing tools take BI out of IT's hands by empowering business analysts and business people to slice and dice, visualize, and analyze their own data and then put it to use when making business decisions.[2]

But Big Data Needs Advanced Analytics to Make a Big Impact

When big data first arrived on the scene four or five years ago, it meant everything in the data stack. Now, there's pretty much universal agreement that big data is about the collection, storage, and delivery of data by IT to the business using data stores and underlying tools like ETL, while analytics focuses on the advanced business tools for analyzing, visualizing, and using data on a real-time basis. Plus, there's an all-important bridge between the big data specialists in IT and the business people consuming the information: it's the data scientists that create the means for big data to move from one realm to another, and the data analysts who help interpret analytics for business people.[3]

Data doesn't lie, but data doesn't tell the whole story

For big data to have a big impact, you need a team of data experts that can combine enormous data sets with a spectrum of analytics tools. Without them, big data is just a pipe dream. Worse, if analytics and a defined set of metrics are missing, then data alone can be spun different ways, depending on the person's point of view. For example, consider how eyewitnesses to a crime recall a different set of facts based on the same reality. Eyewitness reports are unreliable because witnesses interpret the same facts in many different ways. Or, think about how the prosecution and defense attorneys spin completely different stories about what happened at a crime scene while using the same set of exhibits and facts. The same thing happens in businesses; without analytics, business people unwittingly create lots of different stories using the same data. The goal in analytics-driven firms is for everyone to use one agreed-upon set of metrics for business results, so that everyone can make the same conclusions, take similar actions, and tell the same stories.

The business case for data analytics is compelling

The business case for investing in analytics is persuasive. For example, a retailer that sells outdoor products was concerned about the effectiveness of its ads. Marketing decided to aggregate all the audiences and conversion rates to determine where the most effective ad buys occurred. The contribution analysis for one ad campaign uncovered $1.7M in lost conversions of ads *per week*. Through this analysis, the company quickly figured out which ads they should turn off and which ones they should continue. Without it, they had no statistical rigor and insight into which ads to keep and which to eliminate.[4]

Smaller companies may be tempted to dismiss big data and analytics out of hand as tools reserved for bigger organizations, but should think again before making that decision. The analytics vendors claim (credibly) to have very sophisticated customers with less than $100M in revenues. One vendor cites an example of a B2B company with $50M in annual revenues that invests $2M per year in big data and analytics, as well as employing a team of eight data scientists and analysts. Larger companies invest even more in analytics. The same vendor told us they have customers spending $10M per year on analytics technology alone. These expenditures are more than justified through the realization of much greater savings and increased revenues.

Combine analytics with selling to drive revenues and satisfaction higher

As analytics have become more visual and easier to use, demand has grown for well-established and little-known applications in a wide range of industries. For example, some Asian companies now use analytics to find patterns that determine where to fish on a

daily basis, while police use analytics to find serial killers, and health care companies search for patterns in an explosion of health-related data. Or take campaign analysis, a better-known application in most marketing departments. With this application, marketers use analytics to increase the ROI on ad campaign performance, which may be too expensive or not working well, or possibly marketing is using it to try to catch a trend. Two other examples of analytics in action are:

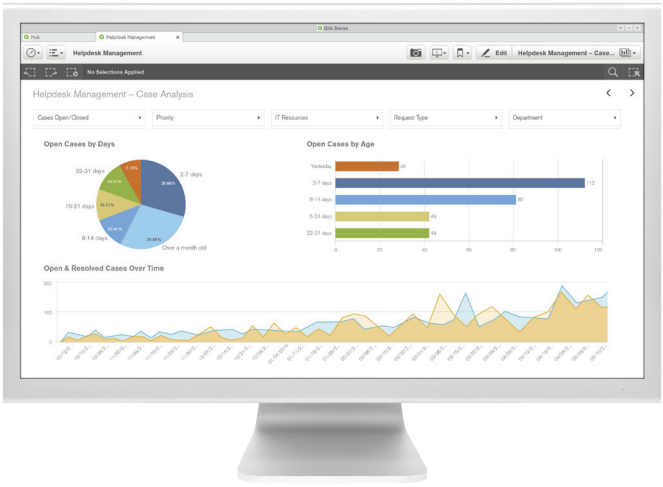
- **Real-time analytics linked to online gaming purchases.** One of the most dramatic and compelling uses of analytics is in the online gaming apps, Candy Crush and Candy Crush Saga from King, a leading interactive entertainment company for browsers and mobile devices. Candy Crush is a free game that is monetized primarily through embedded purchases that are bought inside the game. As customers play the game, King analyzes customer profiles to determine who is downloading, when they are playing, and so forth. By applying analytics in real time, they are able to cross-sell and upsell purchases in the moment by adding something to the game as people use it. In real time, King reacts to patterns and capitalizes upon the behavior it discovers.
- **Hospital support for gathering and analyzing symptoms on tablets.** In physician offices, urgent care centers, and hospitals, diagnosing a patient's illness typically involves the nurse interviewing the patient for symptoms and taking notes, then possibly the physician's assistant conducting a second interview for a deeper look and finally the physician reviewing her notes and examining the patient. In paper-based environments this involves handwritten notes and oral communications; in more automated offices

Figures 1. and 2.

Executive dashboards

Figure 3.

Helpdesk management dashboard for business users



it involves entering data that can be shared by everyone. But with analytics, the physicians, PAs and nurses can identify symptoms, visualize possible diagnoses, eliminate unlikely scenarios, winnow the selections and come to a conclusion. It works by matching and associating symptoms with patterns that are presented in real time on tablets to the physician for consideration. The doctor gets all the data about a diagnosis within context.[5]

The analytics market is moving fast, too, constantly expanding the types of applications being deployed. Currently mobility is huge, along with the trend to empower customers with information (known as consumer empowerment). Plus, the use of social data will continue to grow and help fuel customer experience management. (See Figures 1 and 2 for examples of executive dashboards that show the impact of telecom churn, including a mobile dashboard. See Figure 3 for a helpdesk management dashboard for business users.)

Big Data Needs Data Scientists, Analysts, and Users to Work

The data management people hierarchy in a typical marketing organization, from the bottom up, consists of data scientists, data analysts (also called business analysts), marketers, marketing managers, and marketing executives. Or the line could lead from data scientists to business analysts to product managers and innovation executives — it varies between organizations. No matter how it is organized, this marketing hierarchy needs a continuum of huge data sets that become more simplified toward the top tiers of the org chart (see Figure 4). For example, a data analyst may work with large data sets, while a manager or executive may have 3-5 key performance indicators (KPIs) that really matter. The continuum of data consumed up and down the firm can be correlated with the type of analytics tools and the amount of technical support needed by business people (see Figure 5). For example, analytics tools become less technical, more visual, and easier to use the higher an individual works within the organization.

But tools are only part of the equation. Can business people use analytics tools by themselves or does the organization still need to employ a bevy of data scientists and data analysts to bridge between IT and end-users? The answer is both. Analytics tools are beginning to approach the same ease of use as

Figure 4.

Data Analytics Usage in Marketing Organizations

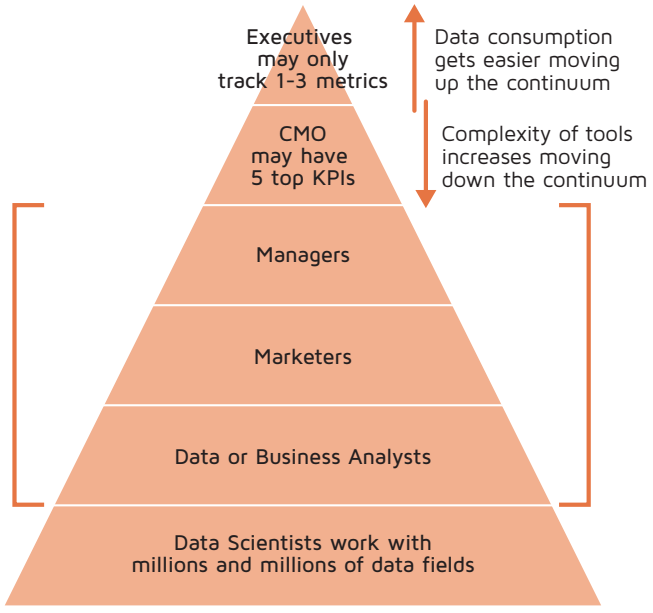
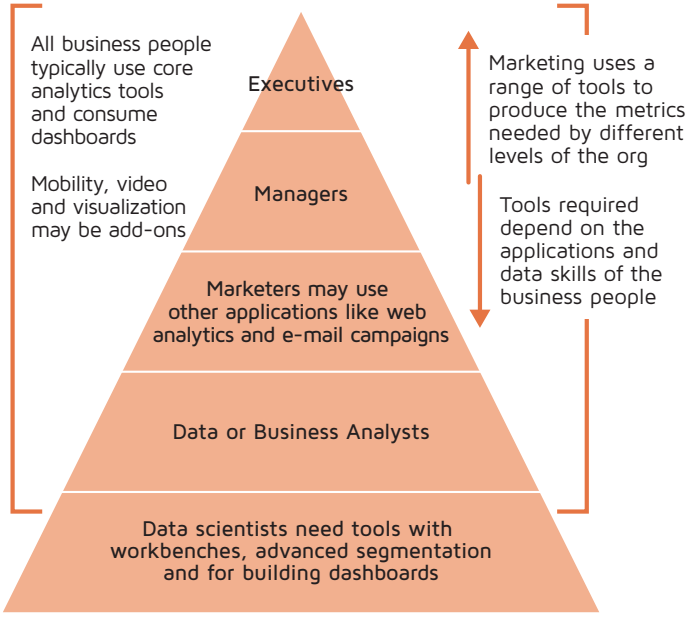


Figure 5.

Analytics Tools in Marketing Organizations



Excel — at least that's the goal. But as big data continues to churn out more data from more sources, organizations still need to employ data scientists with deep skills in math, statistics, or computer science. The answer really depends on the complexity of the application and the data aptitude of the business person. If a non-analytical marketer wants to undertake a complex analysis, he would need pre-built libraries. In this scenario, IT could run pre-built measures that are already in an analytics platform. But perhaps a more knowledgeable marketer wants to combine data sources on his own. He could consume data with an app that is built and deployed by data specialists, or he could produce his own analytics.

Think of the analytics tools this way: Imagine the difference between what an ordinary business person can do with Excel, what an ordinary accountant can do with Excel and what a statistician could do with Excel. Or what a Ph.D. in math can do with a calculator versus a regular business person or an accountant. There's no comparison. Analytics tools are similar; the background, training, and aptitude of the user determines how much and what she can do. In the meantime, the tools for non-technical users continue to advance while organizations continue to hire data scientists and data analysts.

Data scientists are necessary for the care and feeding of big data

Data scientists are sometimes referred to affectionately as "quants." Almost as rare as unicorns, these individuals usually have Ph.D's in mathematics, statistics, or computer science and are employed in the business, not IT. They are hard to find and even harder to hire. Because their skill sets are rare, many companies use digital agencies or other service providers instead of employing data scientists for marketing analysis. And the companies that do hire them must retain their scientists' interest by keeping the fascinating projects coming.

Data scientists spend their time looking for something that may be wrong and figuring out how to correct it or for something unexpected and how to capitalize on it. These talented individuals work with data quite differently from data analysts; they create equations, determine the best way to visualize data, build complex charts and libraries, develop best practices, and identify which type of data graphics would be best for the business to use. Data scientists also often build vertically oriented dashboards for business people. Because the demand for scientists outstrips the supply, dashboards may be built by vendor partners or the vendors' engineering services team. In short, data scientists do things that other people don't know how to do, creating a bridge between the big data world — where IT provides the muscle in security, privacy and data stores — and the business people that constantly clamor for more data.

Conclusion: Focus On Insights and Actions

This report set out to answer one simple question: Have advanced analytics like discovery and visualization, combined with real time support and mobility, and packaged with vertical and horizontal apps (like campaign management) reached the point that ordinary business people could perform all the analytics needed in the organization? And assuming that is the case, could organizations begin to reduce their numbers of data scientists and data analysts? The answer to that question is a resounding "no."

Analytics have made great strides in the past 3-4 years. Business people capable of using pivot tables in Excel can now most likely work with some of the more user-friendly analytics tools. But it takes a hierarchy or spectrum of individuals within an organization to distill the right information from millions of data fields into the one, three or five KPIs that top executives need. The reality is that data analysts and data scientists will be greatly valued for the foreseeable future, if for no other reasons than the amount of big data is exploding, the sophistication of the apps is advancing, and the number of business people using analytics (and needing the support of data analysts and data scientists) is rapidly expanding. Perhaps trying to determine if data scientists are still needed is the wrong question. Yes, they are definitely needed as we move more aggressively into a big data world. But the better question to determine is how your organization can capitalize on the explosion of analytics in a real-time world in which insight and action are the keys to survival. That's where to focus your attention.

Endnotes

1. According to Tech Target's Essential Guide, "Hadoop is a free, Java-based programming framework that supports the processing of large data sets in a distributed computing environment. It is part of the Apache project sponsored by the Apache Software Foundation." http://searchcloudcomputing.techtarget.com/definition/Hadoop.
2. For an executive level on-line presentation of big data, see "Big Data For Smarter Customer Experiences," OgilvyOne Worldwide, http://adayinbigdata.com.
3. For an entertaining but highly informative overview of big data, see "Steamrolled by Big Data," The New Yorker, by Gary Marcus, March 29, 2013. http://www.newyorker.com/tech/elements/steamrolled-by-big-data.
4. Deloitte conducted a research project on the use and benefits of analytics based on 100 online survey responses and deep dive interviews with senior executives at 35 companies in the United States, Canada, China, and the United Kingdom. The interviews were overseen or conducted by Thomas H. Davenport, a visiting professor at Harvard Business School. See "The Analytics Advantage We're Just Getting Started," 2013. https://www2.deloitte.com/content/dam/Deloitte/global/Documents/Deloitte-Analytics/dttl-analytics-analytics-advantage-report-061913.pdf.
5. For a description of other data analytics uses in healthcare, see "Why More Employers Are Using Data Analytics To Improve Healthcare Benefits," Forbes, 11/17/2014. http://www.forbes.com/sites/castlight/2014/11/17/why-more-employers-are-using-data-analytics-to-improve-healthcare-benefits.

.

About DCG

Digital Clarity Group

Digital Clarity Group is a research-based advisory firm focused on the content, technologies, and practices that drive world-class customer experience. Global organizations depend on our insight, reports, and consulting services to help them turn digital disruption into digital advantage. As analysts, we cover the customer experience management (CEM) footprint — those organizational capabilities and competencies that impact the experience delivered to customers and prospects. In our view, the CEM footprint overlays content management, marketing automation, e-commerce, social media management, collaboration, customer relationship management, localization, and search. As consultants, we believe that education and advice leading to successful CEM is only possible by actively engaging with all participants in the CEM solutions ecosystem.
In keeping with this philosophy, we work with enterprise adopters of CEM solutions, technology vendors that develop and market CEM systems and tools, and service providers who implement solutions, including systems integrators and digital agencies. For more information about DCG, visit www.digitalclaritygroup.com or email info@digitalclaritygroup.com.

Contact Us

Email: info@digitalclaritygroup.com
Twitter: @just_clarity
www.digitalclaritygroup.com