

Still Waiting for Big Data?

By Cameron Price

The hype around big data is astounding. Don't get me wrong, the advantages of having a big data capability are easily quantified, but the roadmap to get there is riddled with danger.

Many organizations view big data as a gold mine of incremental revenue, from products to consultancies, but is big data really an incremental requirement? I would like to think that there is a black and white answer here, but unfortunately it depends on your strategy and business case. I will come back to this later.

To describe the new big data paradigm, we talk about the V's of data, essentially volume, variety and velocity. (These are the core concepts, but there are many more.) But what do these really mean?

Volume. We are experiencing an exponential increase in data available to the organization. Just about everything we do or interact with now generates data of some nature, and that data can be captured for utilization later.

So how is this different from what we're accustomed to? We have already faced the challenge of increasing data volumes, so why is this such a big deal?

The difference is that the cost to acquire and store this data, at its atomic level, has rapidly diminished. We no longer need to concern ourselves with defining narrow selections of data, or summarizing data to meet infrastructure constraints, either on premise or in the cloud. This then leads to myriad business questions that maybe weren't considered before but can now be answered.

Velocity. The required speed at which data must be analyzed is increasing to ensure that operational requirements can be met.

Again, how is this different? In fact it's not, it's just that the nature of the analysis is changing. For example, a customer conducts a credit card transaction. That information is captured, and sometime in the future, maybe days later, the marketing department may decide to interact with that customer due to the nature of that transaction.



But in this new world, marketing would like to interact with the customer during the transaction, establishing a real-time requirement to the data capture, analysis and action. Velocity has increased.

Variety. The types of data that can be captured and analyzed have greatly increased, especially as some data sources become digitized. For example speech, video, text streaming, sensor data, etc.

Again, this is not really new. There has always been the need to store a variety of data sources. The change really has to do with the combination of volume and velocity. Consider the aforementioned credit card example. In the past, the customer may have been at a store conducting a transaction-- a binary piece of information that can be used.

In the new world, the world of omni-channels, the customer likely conducted the transaction online, and marketing would like to know the lifecycle and series of actions the customer took before the purchase. Did the customer watch a review on YouTube, research the product on an aggregator site, or receive a review from a friend (or many friends) via Facebook, etc.

So the variety of sources has increased for us to be able to understand our customer, how they behave, and why (in this example).

The Real Problem in the Industry at the Moment is Maturity

The business intelligence industry has matured over a number of decades, slowly enabling business users to take control of their own destiny. But in the big data arena, we have a collection of tools that are not really integrated and are very technical in nature. These largely alienate the business user, and IT is again relied upon to provide a solution, resulting in increased complexity, cost and reduction in speed-to-market -- the exact opposite of the very benefits big data is purporting to provide.

So how do we solve this? In the rush to develop a big data capability, we need to make sure that we don't brush aside all the lessons that the organization and industry has obtained over many years, but instead leverage those lessons so that the journey to the development of a big data capability can be as efficient and cost-effective as possible.

The age-old problem still exists: What business problem or business question are we trying to answer? I am amazed how many people are developing big data capabilities that their business has no idea how to use or what to do with the capabilities when developed.

Let's continue the discussion about incremental investment. At a lowest common denominator, what does big data provide to us? It is an ability to store and process large quantities of data, in order to gain insight in a time frame that was previously not possible.

Organizations have tremendous investments in information management architectures, enterprise data warehouses, etc., and the thought of needing to build an additional big data architecture is daunting, both in time and cost.

But organizations do not need to follow this path. You can utilize big data approaches, and design patterns to deliver a big data capability while replacing your existing infrastructure. This will provide a significantly lower total cost of ownership over time, significantly increased benefits to the business, while retaining the maturity of the information management architecture.

Moving Forward

To begin with this approach, focus on three main principles in the revised architecture:

- 1.Acquire.** Acquire everything from your source, not slices like before, even if you are not going to use the information right away. Historically this was not possible due to infrastructure costs, but in a big data architecture, that constraint is removed. The result is an extremely rich data set that can answer an ever-changing business.
- 2.Manage.** Governance and security is still key to all information. This is an area within big data that is immature, and the knowledge of existing information architecture frameworks should be clearly understood and applied.
- 3.Use.** Usage is key. How are we going to use this information? Volume, velocity and variety have all increased, but so has usage. This is putting pressure on traditional business intelligence structures, like data modeling and ETL, especially when your data model could change in a matter of days.

So how do we do this? It is imperative that robust frameworks are used to ensure rapid delivery of such an architecture to avoid unnecessary costs and potentially short-lived solutions.

It is also imperative that the people guiding the initiative have rich experience in business intelligence, with insights into future trends. This will mitigate any risks on the complex big data path.

Still waiting for Big Data? You already have it!

Cameron Price, Principal and General Manager of CBIG Consulting's Asia-Pacific region, is a well-known authority in the area Big Data and Business Intelligence, focused on building organisations that manage and leverage information as a competitive differentiator, for driving innovation and growth. Cameron brings a unique combination of thought leadership, strategic planning, business development skills, and entrepreneurial flair, coupled with deep technical understanding within the Big Data/Business Intelligence field.



www.cbigconsulting.com