OrbusSoftware

When Will Big Data Deliver On Its Big Promises?

Big Data has been a buzzword in business circles for quite some time now (in fact, the term was first used in 1998), with plenty of popular business books and articles expounding on how it will be another transformative event and shake up the entire business landscape. As with many such trends, the reality has not quite matched up to the hype, with less than half of employees stating that corporate information helps them to get their work done.

For enterprise architects with a focus on data modeling, extracting value from data can be a major part of their purpose. Architects should have a clear understanding of what exactly big data means, and what separates it from more mundane data analysis and business intelligence, as well as what issues are hampering the field, what successes are possible, and how it will develop in the future.

What Big Data is, and what it is not

One of the biggest problems with business trends and buzzwords is that they come to mean all things to all people, and simple ideas end up encapsulating vast spaces. Big Data suffers from much the same fate, sometimes being used as a stand-in for general ecommerce or AI developments. Gartner offered the most widely used definition in 2001:

> **Big data is data that contains greater variety arriving in increasing volumes and with ever-higher velocity.**

This has led to many describing the three Vs of big data, although you will also see references to 4 or 5 Vs (the other two Vs being value and veracity). For a more practical example, many of the large Silicon Valley firms are essentially big data firms. Netflix collects massive volumes of data from its millions of worldwide subscribers, observing data points such as what people are watching, how they are watching, where they are located, when they watch certain genres, when they pause, when subtitles are used, which features are used, and a host of other data. Netflix will then use this to drive decisions on renewing or approving new content or introducing or retaining UX features.

Though Netflix is a Big Data company, keep in mind that not everything they do with data would come under Big Data. A lot of what Netflix does will be typical market research and data analytics. For example, they might adjust subscription prices for a certain group to observe changes in demand. Though this could involve large volumes, it is not big data.

Much of what organizations currently do, whether a large tech enterprise or something else, would be better described as business intelligence. Big Data really does have to be big and as such is mostly the purview of enterprises that have the capacity to collect petabytes (or more) of data.

Why has Big Data not lived up to expectations?

First and foremost: hyperbole. There are some predictions and suggestions for the field that were simply never possible or achievable in their time frames, as proponents of the field perhaps got carried away with promotion. Most business leaders and technology managers are unlikely to have been taken in by these outlandish claims, but is worth addressing nonetheless.

Another barrier is technological – the simple issue that processing vast amounts of data and identifying the signal from the noise is still a big challenge. This is an area that becomes less of an issue with every passing year, as hardware and software catch up with the demands of data.

On a similar level, many organizations lack the right people to handle Big Data and generate value from it. This does not just mean data scientists either, but the various supporting roles (enterprise architects being one such example) that are necessary to manage Big Data.

A particular issue arises from one of the three Vs – Variety. There are now so many information sources and data types that accessing, understanding and presenting data is much more difficult. This is arguably an issue that applies just as much to business intelligence as it does Big Data, though the problem is particularly acute for the latter. Big Data often involves data that is difficult to quantify and requires a lot of statistical know-how to even get anything usable from. Enterprise architects will be familiar with this issue, as they are plagued by problems with database design and legacy storage.

Perhaps the biggest problem Big Data has faced has been finding opportunities to exploit. Many, many avenues of interest identified from data will turn out to be dead ends, producing no discernible benefit, and the rush towards Big Data has played a role in exacerbating this, with businesses seeking a use from their shiny new tools or data sources, rather than identifying a use case and seeing how Big Data can help.

Where do firms find Big Data useful?

The preceding paragraphs would seem to paint quite a gloomy picture for Big Data, but there are always two sides to every story. Though Big Data has not quite lived up to expectations, enterprises have still found benefits spread across a wide range of business functions, with research from Gartner finding 12 main outcomes spread across Sales & Marketing, Customer Service and Supply Chain operations. Toyota are provided as an example, with their motor sales division improving market understanding through Big Data. Through the combination of internal, on premise data from their sales team, cloud based data from their CRM system, and external data sourced from social media and third party rating sites, the company were able to build a 360 degree customer view. This enabled them to better identify customer preferences and decision making, which in turn fed into their marketing strategies.

For Customer Service, Brocade, a US-based computer hardware firm, were able to deliver large reductions in time spent on customer cases. By mining posts on Facebook, Brocade were essentially able to identify product issues before they were even brought up with the company, enabling them to proactively provide solutions for customer issues. This led to case solution times reducing by 40%.
In Supply Chain operations, Delphi Automotive, a UK manufacturing firm, integrated data from across their global supply chain in order to identify risks ahead of time. Utilizing everything from inventory and manufacturing data to supplier data, shipping histories and sales forecasts, the firm was able to drastically reduce crisis response time and inventory analysis time, while improving risk management activities and supply chain visibility.

Potential Developments

We've seen so far that Big Data has had a number of problems that have prevented it from reaching its full potential, even as many firms are able to make use of it to drive decision making. Nonetheless, as technology and our understanding of data improve, there is doubtless room for some of the major promises of Big Data to materialize.

Perhaps the biggest development on the horizon for Big Data is machine learning, the practice of creating an artificial intelligence via training against thousands, millions or even billions of pieces of training data. A lot of publicity has been generated by the likes of Alpha Zero, which was able to triumph in Go games against the best human players, and GPT-3, a text engine that has demonstrated the ability to write working code from descriptions, accurately answer a variety of questions and even write poetry and other prose. The commercial applications of such technologies are still not entirely clear, but large tech companies are betting billions on their success. It is also worth pointing out that there are likely a number of enterprises already using machine learning for some limited applications, and this will likely expand as they develop.

Away from technological advances, one area that may move rapidly is human skill. Being able to use the tools associated with Big Data and apply the correct statistical techniques to extract valuable information depends only on human ingenuity, and a single major advance could transform the situation overnight.

One area that has not had big implications yet, but threatens to do so, is the prospect of negative publicity through Big Data. Concepts such as algorithmic bias and mass surveillance have started to be explored in academic publications, though there has yet to be any major consumer blowback against specific businesses. However, there has certainly been governmental interest in this area, with the likes of Google having faced a number of cases in Europe over the past few years. Indeed, arguably the implementation of GDPR came about due to Big Data collection, which has certainly had a major impact on firms operating in the EU. Fights over privacy, data collection and bias are only likely to intensify in the coming years, and may even become more politicized as the fight over TikTok has demonstrated.

How will this affect Enterprise Architects?

In a sense, the slow emergence of Big Data is good news for enterprise architects as massive volumes and varieties of data could easily overwhelm data modeling efforts. More time is available for improving infrastructure, removing legacy databases and technology systems, and learning how to best model Big Data for the understanding and use of a business.

From a practical standpoint, data modeling efforts can perhaps be freed from the relentless demands of more and more data; mastering business intelligence and analytics is likely to be far more valuable in the short term than chasing gains from Big Data, unless you are a large tech firm. For conceptual and logical data modeling in particular this can be a lifesaver, giving room to effectively plan rather than react to masses of new information, while the physical side is still some years away from being fully developed.

Away from data modeling, architects have a responsibility to prepare digital transformation efforts that take into account developments in machine learning, and guard against accusations of bias or intrusion into private lives. The latter point is perhaps the most challenging, as different regions pursue different strategies on data collection – the EU being litigious, the US laissez-faire, while entry into nations such as China or India may require certain data to be collected.

© Copyright 2020 Orbus Software. All rights reserved.

No part of this publication may be reproduced, resold, stored in a retrieval system, or distributed in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior permission of the copyright owner.

Such requests for permission or any other comments relating to the material contained in this document may be submitted to: marketing@orbussoftware.com

Orbus Software

Floor 4, 60 Buckingham Palace Road, SW1W 0RR

+44 (0) 20 3824 2907

enquiries@orbussoftware.com

www.orbussoftware.com