



THE ANALYTICS COE: THE KEY TO MONETIZING BIG DATA VIA PREDICTIVE ANALYTICS

BUSINESS ANALYTICS SERVICES
KNOW MORE. DO MORE.



BIG DATA & BUSINESS ANALYTICS

AUTHOR:

DOUG FREUD

DIRECTOR, DATA SCIENCE

In early 2012 Ann Winblad, the legendary investor and senior partner at Hummer-Winblad, was asked on CNBC's Squawk box what is the new big thing. Her response: "Data is the new oil". She pointed out that the data alone is not in of itself valuable, and that technologies that add value around big data as the new hot spot for venture investing. Data is like crude. There are certainly companies that buy crude, but there is a significantly bigger market once the crude is refined into gasoline. The analogy holds true for data. Selling big data has little value, but if one can add insight via analytics then we have the opportunity to monetize it. There is a sea change in the corporate world, and IT organizations can seize upon the opportunity to reinvent themselves and transform from a support function into a profit center.

What has changed?

There are really 2 things that have changed, and created a very different mind-set. The first is that Hadoop has revolutionized the way that most companies are thinking about managing data. The cost of managing data with expensive relational based systems has dropped by orders of magnitude using No SQL based systems. In addition, one can even use Open Source Hadoop, and only pay for the hardware. Data that was either archived or thrown away can now be stored and managed with increasing flexibility and at a price point that is significantly more tolerable. While this is an important development the sea change is that the C level is now asking themselves can we improve customer experience by analyzing this data or are there secondary uses for this data (i.e. using it for something other than the original purpose for collecting it). The secondary purposes can include selling insight to third parties meaning IT can now be its own profit center. Ten years ago if someone wanted to store a new data domain into the EDW the IT organization would ask for a business case. Today the mindset appears to be store it, and we will figure out at a later date how we can monetize it.

While there is a remarkable opportunity there are also many significant challenges that organizations will have to overcome before being able to monetize big data. The relational dB world is never going away, and still is the best approach for performing the basic blocking and tackling that most organizations need to run their business. The No SQL market is immature, and it will be a while before the winners emerge. As if No SQL market is not confusing enough there are other important data management technology trends like in memory platforms and graph based dB that are also gaining momentum. In short the IT organization will need to invest in skills and technology around managing big data if they want to transform themselves into a profit center. In addition, IT will need to help analysts figure out how to work with both big data stored in No SQL systems combined with data stored in their current EDWs. While this is in itself a significant challenge it is certainly not the only issue. Storing and managing data is a necessary condition, but in order to monetize big data analytics is required to refine it into something that adds value.

The Predictive Challenges

As if the new world of managing big data is not complicated enough there are other considerations specific to predictive analytics that are problematic with respect to big data. These include a severe shortage of data scientists, what we used to call statisticians, who are capable or experienced in building predictive models. Late last year McKinsey pointed out the key to being competitive is the ability to analyze big

data, but estimates that in the 2014 US alone there is a shortage of 175,000 data scientists and 1.5 million managers or analysts who know how to make decisions based on the analytics. The shortage of analytical skills will not disappear quickly, and while offshoring and outsourcing can relieve some of the issues most senior management understands that long term it must build these competencies in house if they really want to monetize big data. In addition to the shortage of human talent is that the analytical tools that corporations use for predictive analytics are in a state of flux.

The major players in this market SAS, IBM (SPSS), are now being joined by SAP (KXEN acquisition) and Oracle with interesting capabilities. In addition there is disruption via Open Source languages like R and Python both of which are gaining increasing traction. University students are increasingly cutting their teeth on open source languages, and while there are undeniable advantages to these open source approaches there is a limited population of people who ever become truly proficient in writing code. While the low or no cost approach of open source is at first glance appealing the lack of a sufficient supply of data scientists will cause organizations to look at solutions from the major vendors especially ones that will scale across Hadoop infrastructure or in memory platforms like Hana that are ideal for solving real-time challenges. In short there is a significant opportunity to monetize big data, but there are some substantive challenges and investments that companies must make if they want transform big data into insight or revenue.

A Journey of a Thousand Miles begins with a first step: An IT based Analytics Center of Excellence

According to Wikipedia, “a **center of excellence** refers to a team, a shared facility or an entity that provides leadership, evangelization, best practices, research, support and/or training for a focus area.” The key to monetizing Big Data via Predictive Analytics is for IT to build an Analytics COE that understands the needs and challenges of the data scientist. The IT organization will be in the best position to understand the challenges associated with a fast moving Data Management landscape, and how to create an ecosystem that scales. If each individual line of business tries to do this on their own they are unlikely to succeed. The organization needs to make a commitment to developing the proper infrastructure, and the way to start this is with a well-funded and empowered COE.

The COE can't solve the shortage of data scientists, but can help out in the following ways:

- Create an environment and collaborative community that attracts Data Scientists. The best Data Scientists are most interested in building and deploying predictive models. If there is no community or a perceived lack of investment it will be difficult to attract and or retain Data Scientists. Data Scientists are not interested in doing everything themselves. They want and need a scalable, collaborative, and flexible infrastructure that only IT can support
- Invest in tools that enable Business Analysts to build predictive models. Since there is a shortage of Data Scientists IT needs to find the tools and infrastructure that empowers traditional business analysts with the ability to build predictive models. Numerous vendors are working on solutions that will work with Big Data, and the analytics COE should be involved in the identification, selection,

and implementation of these tools. Investment in such solutions will provide a compelling career path for traditional business analysts, and these tools should also be made available to third parties (e.g. consultants and system integrators). The COE should use these tools to create a Self-Serve Predictive capability. It is only a matter of time before clever vendors implement a semantic layer for creating the analytical data set, which is almost always the most difficult part of building and eventually deploying predictive models

Without an internal Analytics based COE an organization is unlikely to have the wherewithal, expertise, or momentum to monetize big data. IT typically invests significant resources around the EDW in support of BI and traditional analytics without understanding that the needs of the Data Scientist are fundamentally different. The end result is that the clever data scientist finds the EDW staging area and removes the data before it is rendered unusable into an environment that does not scale. In order to get their jobs done the Data Scientists create their own shadow IT, and there is a convoluted set of processes that often involve third party suppliers (e.g. Marketing Service Providers) in order to deploy. While this approach might work for a support function like marketing campaigns it is destined to fail if the results of predictive analytics are itself the product.

How HCL can help?

HCL's Business Analytics Services is a leader in advanced analytics and business intelligence with an Analytics Center of Excellence consisting of 50+ data scientists and analysts with advanced degrees in applied statistics. The Analytics COE provides expert implementation services:

- Architecture Assessment
- Strategy and Roadmap
- Tool Evaluation
- Architecture and Design
- Model and Rule Development
- Model Deployment and Integration
- Maintenance and Enhancement Services
- Advanced Visualization

For more information, please contact HCL at ets.bis@hcl.com.

HCL EXECUTIVE BIOGRAPHY



DOUG FREUD

Director, Data Science

Doug has over eighteen years of enterprise application development, sales, marketing and professional services experience.

Doug is an expert at helping large organizations change their business processes in order to make “data” based decisions. He has extensive experience implementing Churn, Cross-Sell/Upsell, and Propensity models.

ABOUT HCL

About HCL Technologies

HCL Technologies is a leading global IT services company working with clients in the areas that impact and redefine the core of their businesses. Since its emergence on the global landscape, and after its IPO in 1999, HCL has focused on 'transformational outsourcing', underlined by innovation and value creation, offering an integrated portfolio of services including software-led IT solutions, remote infrastructure management, engineering and R&D services and business services. HCL leverages its extensive global offshore infrastructure and network of offices in 31 countries to provide holistic, multi-service delivery in key industry verticals including Financial Services, Manufacturing, Consumer Services, Public Services and Healthcare & Life Sciences. HCL takes pride in its philosophy of 'Employees First, Customers Second' which empowers its 91,691 transformers to create real value for customers. HCL Technologies, along with its subsidiaries, had consolidated revenues of US\$ 5.4 billion, for the Financial Year ended as on 30th June 2014. For more information, please visit www.hcltech.com

About HCL Enterprise

HCL is a \$6.5 billion leading global technology and IT enterprise comprising two companies listed in India – HCL Technologies and HCL Infosystems. Founded in 1976, HCL is one of India's original IT garage start-ups. A pioneer of modern computing, HCL is a global transformational enterprise today. Its range of offerings includes product engineering, custom & package applications, BPO, IT infrastructure services, IT hardware, systems integration, and distribution of information and communications technology (ICT) products across a wide range of focused industry verticals. The HCL team consists of over 95,000 professionals of diverse nationalities, who operate from 31 countries including over 505 points of presence in India. HCL has partnerships with several leading global 1000 firms, including leading IT and technology firms. For more information, please visit www.hcl.com



www.hcltech.com

Hello there! I am an Ideapreneur. I believe that sustainable business outcomes are driven by relationships nurtured through values like trust, transparency and flexibility. I respect the contract, but believe in going beyond through collaboration, applied innovation and new generation partnership models that put your interest above everything else. Right now 95,000 Ideapreneurs are in a Relationship Beyond the Contract™ with 500 customers in 31 countries. **How can I help you?**

Relationship[™]
BEYOND THE CONTRACT

HCL