

Big Data for HCM: Big Idea or Big Hype?

By Yvette Cameron, Constellation Research, Inc.

It is said that the world's information is doubling every two years; that 90 percent of the data in the world today did not exist two years ago. Where is this data coming from? Everywhere. From Internet browsing, social networking and mobile devices. From sensors and GPS signals. From user-generated content such as videos and pictures. From the tens of billions of sales, service, support transactions in which we engage and the *quintillion* bytes of data generated from analyzing all those transactions.

And yes, even HCM applications are contributing to the growth in data. But are we taking advantage of all the data available to us today to drive better business results, and are we ready for the next wave of data?

Enterprise data warehouses are the common solution for managing large data sets – internal data from many different systems, and often the external data sets such as benchmarking information. But the truly big sets of data – such as the volume of data coming out of social networking tools or even internal e-mail systems, are often too large to be included in the typical data warehousing solution.

Welcome to Big Data; the massive volume of structured and unstructured data that our “traditional” business systems cannot effectively process. Structured data is the data we have been working with for years – data in our financial applications, HCM technologies, spreadsheets and even XML and other interfaced data. Unstructured data includes documents, text and IM messages, e-mails, videos, social media, voice and any other data that doesn't follow a pre-defined data model.

Attributes of Big Data

Any initiatives involving big data must support and capitalize on its key characteristics, often referred to as the “4 V's of Big Data.”¹ Applying an HCM lens to these attributes helps to uncover both the opportunities and challenges in big data initiatives.

1. **Volume:** As the most commonly understood attribute, volume speaks to the quantity and size of data to be analyzed. Consider HCM transaction data that is captured across learning and talent applications: while these data sets can be a large volume of data, they are not typically “big data” large. Now consider the volume of e-mail transactions, instant messages

(IMs), social conversations in the organization, as well as the data analysis from network utilization. Specialized tools are necessary to store and query the vast quantities of data to get to actionable intelligence. Tools such as SAP Hana, IBM Watson, and Apache Hadoop are purpose-built for this capability.

2. **Variety of data:** Workforce analysis is traditionally based on data from internal systems (HR, payroll, learning, finance, talent management, etc.). The intriguing opportunity for organizations now is to expand the data sets to broader views that directly and indirectly affect employee acquisition, retention and engagement strategies. Instant messenger and e-mail, social media engagement, mobile and location services, user-generated content, customer interactions, regional and global pay practices, local competitive data and hiring rates: All these could be accessed and leveraged in a broader analytics initiative. The addition of unstructured data such as speech, text and language complicates the data landscape but also enriches the analysis opportunities. A common challenge for organizations is understanding what information to gather: Hoarding data from every source can prove too overwhelming while limiting the data sources too severely will cause you to miss the important correlations that only come from large scale analysis.
3. **Velocity:** In the commercial world, we evaluate petabytes of information to detect fraud at the time of a credit card transaction or to surface highly personalized and contextual marketing content as consumers visit a Web page. Is this same real time analysis necessary to detect, for example, projected employee turnover? Or to secure the best candidates? Whether the urgency is there today or not, any lag time in information and decision support will become less tolerable over time as consumers, who are also our employees, are increasingly accustomed to personalized, real-time, highly relevant and contextual information.
4. **Veracity:** Can the data be trusted? Consider these questions: Is the quality of data reliable? Are interpretations of unstructured data accurate? Is spam sent via social media effectively weeded out, or at least discounted from the overall predictive results? How will data streams, taken out of context, affect conclusions? Getting to the truth of the data is one of the challenges big data tools look to address.

With an understanding of the characteristics of big data, organizations can begin their journey toward greater work-

force insights, decision support and predictive capabilities. Three questions that should guide your investments are relatively straight forward:

1. **What questions am I trying to answer? What problems am I trying to solve?** You don't begin a big data initiative with technology. In fact, you don't set out to embrace big data unless it is required to get to the information you need. Always begin with an understanding of the problems that need to be solved, the questions needing answers. The challenge, of course, is that "you don't know what you don't know." A process of continual refinement can be expected as you prepare your strategies for getting to the information you need.
2. **What data is available to me today?** HR, payroll, time and labor, scheduling, financials...all the systems we use today provide the necessary foundation for workforce analytics. Many practitioners report ongoing challenges in getting to data within their workforce technologies, much less across the entire enterprise. Work closely with IT, line of business leaders and others as needed to free the data already captured within the business as a starting point in your analytics journey.
3. **What additional data sets are available?** With the advent of the Internet, social networking, mobile, infographics and more, organizations need to look beyond traditional enterprise systems to what additional data is available that could influence or inform workforce intelligence. Our "online digital identity" is often a much richer expression of ourselves than the profiles housed in the enterprise HRIS. Sites such as LinkedIn, Facebook, Amazon, Netflix, Pinterest, WordPress, Twitter and many others capture profile data, engagements, comments, likes and dislikes, all of which develop a more comprehensive profile of the individual. Other internal systems also add color to workforce understanding. E-mail and instant messenger streams, for example, can provide significant insight into worker sentiment.

What can Big Data do for HCM?

Consider this example: With high attrition rates in its call centers, Xerox turned to big data analytics to improve its hiring processes.² Hiring processes to date had focused on recruiting staff with extensive call center experience. With big data, Xerox discovered that job experience did not correlate to employee attrition. Instead, it found that behavioral indicators, not past experience, were more indicative of employee longevity. By focusing on applicants with "creative" rather than "inquisitive" behavioral traits, and defocusing on experience, Xerox was able to cut its attrition rates *in half*.

Another well-known example of analytics transforming results is "Moneyball"—the story of a baseball team that went beyond traditional skills analysis to true performance analysis, using big data to uncover new but critical areas of focus.

HR can drive significant outcomes by embracing big data analysis within its HCM initiatives. Below are several examples of how traditional data capture and analysis processes can be transformed to support better talent processes across the organization.

1. **Know your talent! New approaches in talent profile development.** LinkedIn CEO Jeff Weiner recently revealed his vision for LinkedIn in the coming decade:

*"Our ultimate dream is to develop the world's first economic graph. In other words, we want to digitally map the global economy, identifying the connections between people, jobs, skills, companies, and professional knowledge – and spot in real-time the trends pointing to economic opportunities. It's a big vision, but we believe we're in a unique position to make it happen."*³

Weiner goes on to discuss how the linkage of job opportunities with the skills required for those jobs, coupled with the companies across the globe offering those opportunities and overlaid with the professional profiles of people across the globe, will transform the talent acquisition landscape as we know it. His planned approach is to bring the data together, and then "get out of the way and let the connections happen and flow seamlessly as it needs to."

With the aggregation of data across a wide variety and massive volume (global) of data, Weiner's vision is transformational. Similar thinking can and should be applied in today's organizations; fortunately, the technologies are emerging to help make this possible. For example, Oracle, as part of its Social Relationship Management suite, has signaled a growing focus on the social profile of employees. SAP and IBM have each begun to develop a broader view of the employee incorporating both social and formal attributes to drive better planning and outcomes. Workday is another recent entrant in the big data space, with social insights expected to come from Salesforce Chatter. The tools are emerging; the challenge is for HR to lead the use of these tools in their application to broaden workforce analytics.

2. **New approaches in recruiting.** The recruiting process is frequently accompanied with one or more assessments. Big data analysis now allows us to combine the hundreds of thousands of responses across cloud-based survey engines with actual workforce data such as performance, longevity and other scores to uncover better predictors of success. Recall from the earlier example that personality attributes trumped experience at Xerox when worker retention in the call center role was the driving business issue – critical data in the hiring process that was unknown before the use of big data.

Other areas of recruiting are benefitting from big data as well. New technologies provide a broader view of active and passive candidates through analysis of social data; they mine employee social networks to

match connections to opportunities – and more. Candidate Relationship Management is benefitting from the principles of another “CRM” – Customer Relationship Management – as information, service and communications are amplified by social analytics. Best practices in the sourcing and acquisition stage of the employee life cycle are highly correlated with best practices from the early stages in the customer life cycle.

3. **Identifying top talent.** A few years ago a large consumer products company looked at the correlation of its top talent as identified through the annual review process, and “top talent” as defined by high influence and peer feedback in its internal social network. Interestingly, there was little correlation between the two. This example points to the need to rethink our definitions and identification processes for high potentials, top performers and “key players” in today’s era of collaborative business. Social media highlights the influencers in the network and makes visible their impact across the company through feedback, badges, content contributions, peer rankings of that content, and more. Expanding our data sets from a review of goals and peer feedback to a broader analysis of social activities is one of the many ways our talent management processes must evolve.

These are just a few of the many opportunities to come from applying big data to traditional HCM processes. Much of the transformational thinking to redefine processes with big data has already taken place within the customer relationship management (CRM) space. Examining the lessons learned and leading practices from CRM approaches can serve as an instructional guide for HCM leadership as it begins the next stage of its analytics journey in the months or years to come.

Learning from our CRM Counterparts

Listening to social signals – the conversations, chats, tweets, etc. – and deriving new insights, predictive capabilities and new marketing strategies is becoming commonplace. Companies have focused on understanding and optimizing their customer relationships, but the same cannot be said about employee relationships. Approaches to employee surveys to assess workforce engagement and overall sentiment, for example, are still challenged by survey quality, point-in-time snapshots and the inherent challenges that come with asking employees to respond to workplace surveys on workplace tools; can the resulting data be trusted?

It’s time to apply the lessons learned from customer relationship management (CRM) to *employee relationship management*. It’s no longer OK to know more about our customers than we know about our employees. Now, through big data, we no longer have to settle for this imbalance. Consider the following examples.

- **Social monitoring** is a critical component of managing the customer life cycle. Data streaming in

from online social media channels, customer support sites, e-mails, surveys, and even VOIP or telephone conversations is analyzed to determine customer sentiment. Analysis at the individual level enables customer issue resolution; at the aggregate level, it provides predictive modeling and trend analysis.

Just as companies monitor social media to understand who is talking about their company and what is being said, so, too, should companies monitor social channels and potential other internal communications to understand employee sentiment and identify emerging risks or opportunities. Through data normalization and the masking of identifying information, data privacy can be ensured while organizations can glean new and unique insights into top performers, identify best practices, improve customer success and drive higher revenues. For example, one organization adopted next-generation sentiment analysis tools, and was able to correlate top performers to e-mail responsiveness and sentiment: Top Performers sent 50 percent more e-mail externally, responded 40 percent faster to clients, and sent messages with 55 percent more positive tone than the bottom tier. (see sidebar, page 31) This insight was only made possible through *big data analytics* applied to internal communications and workforce data.

- **Social campaigns** in CRM involve brand advertising, content-based campaigns, engagement and communication paired with the results of social monitoring, transaction analysis, community trending and other analysis. Applied to the employee life cycle, social campaigns become the way in which we better source and engage with future employees. Engaging the individuals who have positive company sentiment, for example, is a good first step in developing high potential candidates. “Campaigns” can also be aimed internally to improve employee engagement, collaboration and development – either individually or in aggregate – with the content of such campaigns driven out of internal and external data analysis.

Getting Started

Big data or small, getting to better workforce intelligence is a growing business imperative. To begin or advance your company’s journey toward better business intelligence, perhaps powered by big data, a few key tenants should underpin your analytics strategies:

1. **Embrace data exposure.** It is in HR’s nature to silo and protect data due to privacy concerns, competitive considerations and more. However, we need to work through these barriers with our legal, business and technology partners so that we can unleash the data in our enterprises for greater analytics insight.
2. **Big Data projects require executive sponsorship,** in large part because they require a culture of measurement; without this culture, driven from the

top, big data projects for workforce initiatives ultimately fail.

3. **Don't let data issues stop you from beginning the journey.** Technologies today are analyzing massive amounts of "dirty" data – messages abbreviated in "unnatural" ways to fit a 144 character limit, for example – and yet they are able to draw important, valid conclusions from those streams. Consider temporary swat teams to help clean-up internal data, but even in traditional workforce analytics, today's technologies are ready to accept some level of inconsistency.

Endnotes

1 R "Ray" Wang of Constellation Research speaks of two additional attributes, "Viscosity" and "Virality," addressing the friction, flow rates and peer-to-peer dissemination of data. For more information, see his blog post "Monday's Musings: Beyond the Three V's of Big Data – Viscosity and Virality," <http://blog.softwareinsider.org/2012/02/27/mondays-musings-beyond-the-three-vs-of-big-data-viscosity-and-virality/>

2 http://professional.wsj.com/article/SB10000872396390443890304578006252019616768.html?mod=WSJ_hp_LEFTWhatsNewsCollection, *Wall Street Journal*.

3 (Source: <http://www.linkedin.com/today/post/article/20121210053039-22330283-the-future-of-linkedin-and-the-economic-graph>)

About the Author



Yvette Cameron, *Workforce Solutions Review* Advisory Board member, is vice president and principal analyst at Constellation Research, Inc. and founder of NextGen Insights, LLC, an independent consulting and marketing communications firm. A proven executive leader and visionary, she has more than 25 years' experience developing, marketing, evangelizing and implementing market leading HCM technology solutions. Prior to joining Constellation, she led product and marketing strategy teams at Saba, Oracle, PeopleSoft, SAP and JD Edwards. She has more than 10 years of experience as an HR practitioner, where she honed her skills in change management and the development of effective workforce process and technology strategies. Today she serves on advisory boards for the Institute for Human Resources for Social Media and Employee Communications and for enterprise analytics provider, Thrive Metrics. She can be reached at y@constellationrg.com or through her blog at www.nextgeninsights.com.

Case Study on Workforce Sentiment Analysis

Summary: A large electronics distributor leverages large-scale text and sentiment technology to drive improved decision-making and business performance.

A large electronics distribution company was looking for an efficient way to monitor service levels with key customers, as it pertains to responsiveness and coordination of their sales teams. It was experiencing increased customer attrition and wanted to take a more proactive approach to customer retention.

Challenge

- The company wanted real-time visibility into frequency and content of customer interactions with front-line employees and across the company, instead of after the fact results.
- Some metrics were already tracked on their CRM system, but content-based insights, response levels and relationship trends (i.e., when customers start trailing off) were not available.

Solution

- Client engaged with Thrive Metrics, a new provider of analytics technology for large-scale text and sentiment technology.
- Real-time customer management dashboard with communication metrics, client coverage and service level metrics were unavailable in the CRM system.
- Customized alerts to screen out lower priority topics and track meaningful changes in client service and relationships; react faster to events that really matter to the business.

Results

- Improved internal coordination with strategic customers and streamlined contacts from 20 to just a few.
- Incorporated proactive monitoring of sales teams across customers (prioritization, selling activity, top topics) with other performance metrics (i.e., number of calls, financial performance, etc.), realigning priorities with business objectives.
- Improved customer experience with reduced response times from sales and support staff - from average of 18 hours to 9-10 hours.
- Established best practices based on communication trends of top performers – Top Performers sent 50% more e-mail externally, responded 40% faster to clients, and sent messages with 55% more positive tone than bottom tier.