

Exploring Big Data in Small Steps, Starting with Social Media Analytics

By Todd Nash

Last year was a significant year for many scientists and Antarctic enthusiasts around the globe, marking the 100th anniversary of the discovery of the geographic South Pole by Roald Amundsen in December 1911 and by Robert Falcon Scott several weeks later.

Success in reaching the South Pole was the culmination of years of their own trials, as well as previous explorations by others who paved the way with new geological discoveries, polar survival techniques and scientific research. Expedition leaders, scientists, experienced sea captains and their crews persevered under extreme conditions to create a better understanding of this undiscovered continent on numerous levels, including its wildlife, vegetation, impact on the global climate and its past and present relationships with other land masses.

The New Age of Exploration: Discovering Best Practices and Architectures of Big Data Analytics

Why did Amundsen and Scott succeed where others failed? Because they utilized comprehensive planning based on experience and replicated methods that were successful (dog sleds and skis, yes ... horses and cars, no).

Now we're in the throes of the "Age of Internet Exploration" and one thing is for certain: everyone in the business intelligence field has an opinion about how to harness the power of big data analytics. Vendors, educators and analysts all have a point of view and a vested interest in the topic. While 60 mph winds and minus 35 degree Celsius temperatures are thankfully not a factor, like the intrepid explorers of the past, innovation in BI depends on solid footing from the start.

Strategy, Architecture, Integration and Realization

In our practice, we discovered early on that if an enterprise tries to start a big data analytics initiative without a clear framework, they're virtually guaranteed to fall down a vast

crevasse of wasted energy, resources and time. We believe that the solution is twofold:

1. Be sure to invest your time and resources upfront in a comprehensive big data framework that is based on proven best practices.
2. Consider breaking down this big undertaking into manageable tasks, starting with a small-scale project that has limited investment and risk to your enterprise. Then, as you gather data and prove the value of the concept, expand your program based on the lessons you've learned along the way. First, a well-defined big data framework strategy for your big data initiative should clearly identify and outline all of the goals, business processes and players involved in accomplishing those objectives:

- Identify opportunities
- Build the business use case
- Determine governance and ownership
- Determine the solution architecture
- Develop a big data roadmap

This framework stresses achieving qualitative, not just quantitative, solutions that will help your organization make more informed strategic decisions. It's one thing to measure a business based on a snapshot transactional metric, such as "How many widgets did I sell last month at this location?" The real business value of your big data program becomes obvious when your firm can draw pragmatic, actionable insight as to why your firm sold so many widgets at a particular location.

Social Media Analytics as a Starting Point

Building on your initial victories and lessons learned from smaller implementations will help determine what provisions, methods, manpower and tools will function best when integrating future big data processes. For many firms, the first step in the big data journey is an effective social media analytics initiative.

Social media analytics provides a measurable means of gathering, processing, analyzing and delivering business intelligence from social media channels. Many organizations have real concerns about embracing social media software, dreading to open a Pandora's box of damaging information and negative sentiments that lack internal or external controls. These apprehensions are indeed legitimate; however, as big data and BI consultants who work with customers in boardrooms and data centers on a daily basis, part of our mission is to reinforce the idea that current knowledge is future power. To that end, greatest success is achieved by developing social media analytics components to channel and control that power responsibly, providing its proof of value within a larger strategy.

Antarctic explorer Ernest Shackleton of Endurance was the first, and arguably most brazen, polar expedition leader to successfully leverage social media (e.g., newspapers, presentations and tenuous connections with scientific societies and notable academics) to engage the public, persuade wealthy private patrons and raise funds for and then exploit his expeditions. In 2013, the social media landscape has evolved far beyond these traditional channels to include countless data resources, including but not limited to:

- Facebook, Twitter, LinkedIn, Google+, etc.
- Review sites, like Angie's List, Yelp, Urbanspoon, TripAdvisor, etc.
- Blogs and news sites that include/encourage comments
- Video and photo sharing sites, like YouTube, Flickr, etc.
- Search engines, such as Google, Bing, Yahoo and others

All of these unstructured data points can yield critical perceptions about our brand, our products, where these items are purchased, who is using them and how they feel about them. To get accurate answers to these issues, an effective social media analytics consulting approach should be based on the following four components, or layers, that build on one another.

Awareness - It stands to reason that if you don't know where you are, you won't know which way to go. The starting point is to understand consumer awareness of products and services. Measurable, disparate data is taken from the following sources: 1) social media channels; 2) search activity; and 3) regions/location (today, almost everything can be associated to a location). To gather this data, use a combination of Web crawlers and application programming interfaces to extract social media data. Typical graphics used to illustrate awareness may include charts depicting year-over-year brand mentions by region, conversations by social media channel per day and trending hashtags.

Brand & Image	How do we help protect our name, brand, image and reputation?
Transparency	<ul style="list-style-type: none"> • How much information should we disclose? • When, where, who and how should we engage with bloggers and provide content?
Uncontrollable Environment	<p>Are we enabling the wrong behaviors?</p> <ul style="list-style-type: none"> • "Playing around" on social sites during work hours. • Unauthorized responses and posting of information.
Consistent, Immediate Participation	<p>How can we possibly control our exposure?</p> <ul style="list-style-type: none"> • Content is streaming from anyone, from anywhere, across thousands of sites, in multiple languages, all the time.
Size, Scope and Scale	<ul style="list-style-type: none"> • The amount and type of content is growing rapidly. • Content is streaming across multiple channels and multiple types of devices, and is available globally via social media.
Legal Restrictions & Ramifications	<ul style="list-style-type: none"> • How do we moderate content? • Who wrote what, when, and what are the consequences?

Analysis - The next step involves understanding what's being said about products and services. Commonly referred to as "sentiment analysis," this information is positive, negative or neutral. To determine this, use a rules-based engine that filters through all types of unstructured data and then calculates a quantifiable value to it. This analysis can show a breakdown of sentiment by channel; top trending social media channels in relation to number of opinions; and whether these channels are trending up or down relative to to positive/negative sentiment, for example.

Influence - This component helps organizations understand which social media contributors are significant and which are not. For instance, if someone has many friends or followers, their sentiment will have a greater impact than someone who does not. This information can lead to intelligence showing which regions have the highest and lowest concentrations of influencers, as well as number of mentions versus actual users in a given region. Enterprises can effectively benefit from these details by pinpointing trouble spots of negative influence, or by improving market penetration and product placement geographically, for example.

Evaluation - A key result of an effective SMA strategy is the capacity to measure and evaluate a PR campaign or marketing promotion that has been deployed. Say a business tries to convince a negative influencer to retry a product or service, or incent a positive influencer to "talk" more about a product or service. This promotion will ask for a response on the part of the targeted influencer, such as clicking a "like" button or offering a coupon in exchange for a "like." Customer response rates in relation to a product or series of products, as well as the sentiment attached to these response rates, can then be measured and mapped out by social media channel and region. And then the whole process circles back to the beginning to see if the promotion is having an impact.

Enterprises that hope to thrive in the Internet age must listen and be more responsive to their customers, potential

customers and even dissatisfied customers, in order to address consumer needs more effectively. By taking advantage of the instantaneous, voluminous nature of social media analytics as the opening gambit within a fully defined framework, a business can rapidly validate and strengthen its business case, while simultaneously deriving associated business value and gaining new insights. The benefits of embarking on a low-cost, low-risk social media analytics program include micro-target marketing, brand protection, customer loyalty and promotion feedback. Adding new processes (that may include weather, breaking news or economic indicators, for example) will only increase the insight and understanding from social media and expand the impact over the life of your business.

It's worth noting that the heroic age of Antarctic Exploration involved impossibly large goals, including the race to the Geographic Southern Pole, the race to the South Magnetic Pole (it moves!) and the ill-fated transcontinental crossing attempted by Ernest Shackleton and the crew of the Endurance from 1914-1917. It's the collective sum of the less-lofty ambitions, however, carried out by scientists and adventurer-academics that offers a more valuable lesson and has an enduring impact on our understanding of the natural world

today. Explorers learned that the practice of establishing small depots of fuel, food and necessary supplies along expedition routes was absolutely essential to the survival.

Building repeatable processes is a good tip to remember as we explore best practices and architectures of big data analytics. Like the Antarctic pioneers of the past, we need to explore all forms of big data in layers and build repeatable processes into each layer. These proven processes will create direct routes into the data's meaning and play a critical role in discovering its value for your enterprise.

Todd Nash is a founding Principal at CBIG Consulting, a professional services firm that helps clients leverage their data assets through big data analytics to produce timely, effective business strategies and tactical decisions. Nash's responsibilities include managing, developing and implementing big data analytics architectures, business intelligence/data warehousing solutions, and cloud-based analytics services. CBIG's subject matter expertise extends to a variety of industries including consumer packaged goods, telecommunications, manufacturing, insurance, banking and health care. You can reach Todd at todd.nash@cbigconsulting.com.



www.cbigconsulting.com