HARLAND CLARKE®

The Big Picture on Big Data: *What You Need to Know About the Future of Data-Driven Marketing*

By Dr. Charles Stryker

So much is being written about "Big Data" and the huge business opportunities that accompany this emerging science. IBM, Oracle, EMC, Amazon, and others are pumping billions of research dollars into massive computing systems to handle the processing of Big Data environments. With all of this momentum, Big Data promises to be a critical component of business success, and the science around it will continue to expand.

Looking Back to Look Ahead

To understand the business opportunities associated with Big Data today, it's necessary to consider its history. It's difficult to imagine life before computers, but the fact is that before 1970, there was no computerized information being used for business decision-making. It was not until the mid-1970s that the power of computing became sufficient to create practical, albeit limited, computer-based data assets. Before the mid-1970s, there were no computer-based credit bureaus, no computer-based marketing lists, no data cleansing programs or any of the technology-based processes that we take for granted today.

When the computer age hit, many early adopter companies such as Dun & Bradstreet, Thomson Reuters, LexisNexis and Associated Press began the process of automating their data assets. It was also in this early period that most of the current providers of information services were launched, including Acxiom, InfoGroup, Epsilon and Merkle. These early companies, as well as many others, led the way and collectively created the information service industry.

The next milestone in the evolution of this industry occurred in April 1993. That's when Mosaic shipped the first commercial internet browser. To put the world of Big Data into perspective, between the dawn of the information services industry (mid-1970s) and the shipment of the first browser (April 1993), about 1 billion gigabytes of data describing consumer and business behavior was stored. Due to the increasing power of computing and the ever-increasing volume of data being collected digitally, by 2003, 5 billion gigabytes of this type of data had been collected and stored. Much of this new data resulted from the dramatic growth of the internet and early products created such as ad networks, product ratings and reviews, and email marketing.

We now collect and store over 25 billion gigabytes of data about the behavior of businesses and consumers each day, and this volume is still growing exponentially. This massive data explosion — from 5 billion gigabytes total to 25 billion gigabytes every day — has occurred in the last 15 years due to many factors, including the advancement of the smart phone, development of the search engine, expansion of social media, and the sophistication of the tools available to structure unstructured data. Within this reality lies the Big Data business opportunity.

Exploring the Opportunity

The rapid growth in the amount of data captured and stored about human behavior has created a highly inefficient marketplace for this valuable asset. The decision systems of most businesses have no means to examine all the data available and determine how data-derived insights can inform — or transform — their decision-making process. Similarly, many companies that own the data have no means to understand how brands can benefit from new data sources. This problem is further complicated by the need for a powerful analytics engine that can extract the data insights needed to support a particular business decision.

Let's also remember that Big Data environments require constant vigilance: Just when companies think they have the process mastered, 25 billion gigabytes of new data every day must be examined and integrated.

This magnitude, combined with the inefficient marketplace, the dynamic nature of the data, and the significant processing challenges, has resulted in corporations of all sizes not yet realizing the benefit of using new data sources to transform their decisioning processes. These companies operate seeing only a very small fraction of the data that could potentially be used to inform their next steps.

The Big Data Ecosystem

To expand on this concept, it is useful to understand where the 25 billion gigabytes originate each day. The Big Data ecosystem can be broken down into six categories:

1. Published Material. Most of the world's published material and news is available via search. If you want to read or see news articles or broadcasts that relate to Apple Computer, are positive in nature, and mention the name Tim Cook, the response to this search request would be timely and at scale.

The same concept applies to scientific journals. Before the Big Data revolution, scientists and researchers would subscribe to a small number of journals to read a small number of research articles in each one. Today, scientists and researchers can specify the types of articles they're interested in, and gain access only to those articles with that specific reference across all journals. This approach allows scientists to be more informed and, at the same time, benefits their scientific discovery.

In terms of marketing, this ready access to published material has already worked its way into marketing channels. For example, financial institutions can monitor online and print sources, various PR outlets and corporate websites to identify when businesses announce that they plan to expand, hire more employees, downsize, go offshore, and so on. These news items may indicate a business prospect that's in the market for new financing.

2. Search Data. For consumer purchases that are researched before buying — cars, travel, college, a new home, refinancing — the search function represents one of the most compelling data sources to identify consumer interest. There are many points on the internet where search data can be observed. For example, if a consumer searches for a cruise using Google, he or she would likely land at one of the major travel sites such as Expedia, Kayak or Orbitz to investigate the best alternatives. These sites then know that a particular consumer is cruise shopping, and is therefore in the target market for travel-related goods and services.

The same concept applies to the auto industry. When consumers are thinking about buying a new car, they search for information using sites such as Cars.com, Edmunds, YahooAuto and others. These sites then know that these searchers are in the market for a car. They package the data appropriately, and make it available to others, including financial institutions. Clearly, financial institutions seeking new consumers for auto loans would find this to be a rich and timely source of new data from which to execute their consumer acquisition strategies.

Home refinancing follows the same principles, but with sites such as Lending Tree, LoanDepot and Bankrate. These companies then create a data asset to be packaged and sold to appropriate markets. In fact, many categories of search data fit this model and are available to be packaged as needed for financial institutions.

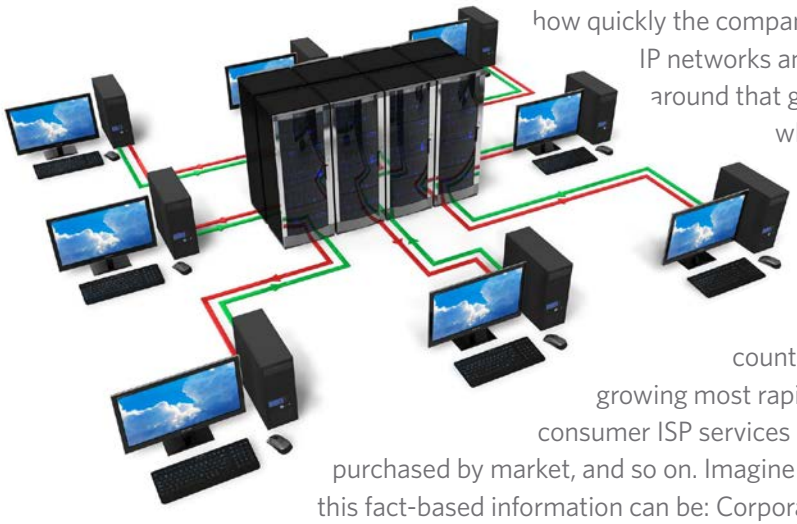
3. Social Data. Online interaction with friends and communities can be a very powerful data source. Although Facebook and Twitter dominate the social data category, there are literally hundreds of other interesting social data companies.

For example, there are companies that allow consumers to conveniently share information with their friends when they see an article or a promotion of interest. Because we can observe who is sharing information, with whom they are sharing it, and who is responding, we can identify people who are interested in certain topics like cars, music and financial planning.

We can also identify the true influencers. Once we know who is interested in cars, and can identify the most influential people to carry a message to that audience, it's possible to develop specifically targeted marketing campaigns.

4. Web Mining. The tools available to capture information from the open web have improved dramatically over the past few years. These tools not only mine content on web pages but also have the capacity to mine details about web infrastructure.

For example, there are companies that interrogate IP addresses around the world to discern information about businesses and consumers. If you identify all of the IP addresses associated with a specific company, you can understand how quickly the company is growing its IP networks and the details around that growth — in which countries the company is growing capacity; which cities in those countries are growing most rapidly; the consumer ISP services being purchased by market, and so on. Imagine how useful this fact-based information can be: Corporate growth triggers demand for a variety of products and services including financial services.



Web mining is also used for other important marketing applications. For example, websites are tracked to understand executive changes, mergers and acquisitions, hiring details and new product launches. This data tends to be global in nature and timely in that the information can be classified and reported in near real time.

5. Crowdsourcing. Millions of people around the world make a living by performing microtasks for a fee. One of the major categories for the use of crowdsourcing is data entry. For data-sourcing requirements that require human intervention, crowdsourcing represents a cost-effective and compelling method to secure critical data.

A number of companies have built applications on top of this crowdsourcing capacity. For data entry, you specify the data items and the related process required, then place that job specification in the crowdsourcing environment. Individuals then respond and you hire them to physically search the internet to collect the data elements you have specified.

To visualize this process, think about a person looking at a split screen, with the form for the data entry project on one side, and a web page on the other. Data is copied and pasted, or reentered from the web page into the form. The individual typically gets paid on a per completion basis.

Today, these techniques are used in marketing campaigns to capture data from websites where data is difficult to collect via web mining techniques — specifically from PDFs, JPEGs, and other formats where web crawling is not particularly effective.

6. Traditional Transactions. This category continues to grow as companies capitalize on the opportunity to create new revenue streams by monetizing their data assets. Generally these companies have massive amounts of data that describe details about consumer and business activities. Companies such as cell phone and credit card providers, travel-related ticketing and review sites, ad serving companies and many other similar transaction-based businesses create this data as a byproduct of their core business.

Traditional transaction data tends to be extraordinarily valuable, yet is not available in the general marketplace. Because this type of data is captured as part of a particular business model, it is subject to restrictions: No company would allow use of its data in a fashion that would compromise its core business.

For example, if a cell phone provider wanted to market its information describing the geolocation of subscribers, the company would not want to make that data available to competitive providers. On the other hand, there are many use cases for this data that would not compromise the core business model and would be suitable as a path to generate a new revenue stream. Hundreds of companies are launching new divisions specifically for the purpose of monetizing their core data assets.

The Bottom Line

The Big Data ecosystem is developing rapidly, as are the related technologies of data hosting, data hygiene, analytics, and data visualization and delivery. As this industry matures, we will not only have the data necessary to improve the quality of our decisions, we'll also have a cost-effective set of tools that will allow us to find the critical insights contained in these new data assets and put them to use.

It is an exciting time for companies that embrace Big Data and all of the remarkable advantages it can offer to inform their decisioning.

How can big data can transform your business? Harland Clarke offers solutions that take full advantage of Big Data. Our history in the industry and our expertise in data gathering and analytics give us a unique perspective on Big Data ecology and how it can work for financial services clients. To find out more, call us at **1.800.351.3843**, email us at **contactHC@harlandclarke.com** or visit **harlandclarke.com/MarketingServices**.

Dr. Charles Stryker is Chairman & CEO of the Venture Development Center (VDC). Founded in 1996, VDC is the preeminent advisory firm assisting companies in all aspects of their data-related strategy and business development activities.