BIG DATA 2.0
New business strategies from big data

BY VIVEKANAND GOPALKRISHNAN, DAVID STEIER, HARVEY LEWIS, JAMES GUSZCZA, AND JOHN LUCKER > PHOTOGRAPHY BY MATT LENNERT > INFOGRAPHICS BY ILOVEDUST
BIG DATA 2.0
“The possibilities are numerous once we decide to act and not react.”

—George Bernard Shaw

The possibilities presented by big data are numerous. But, to paraphrase George Bernard Shaw, how should business leaders act rather than merely react? In seeking to derive maximum value from their businesses, organizations should act before they are faced with challenges. Under the appropriate circumstances, big data can lead to insight within emerging trends that are invisible in small data, and thereby provide the means for businesses to be more proactive.
THE CURRENT STATE OF BIG DATA FOR BUSINESS

Room to grow for big data

Indeed, in a competitive environment, big data may compel such action and dictate against mere reaction. But assuming that an organization has carefully weighed the benefits of big data applications against the corresponding costs, which among the numerous big data-enabled possibilities is best? Big data presents three types of possibilities:

- Answering existing questions in existing businesses, with a focus on improved efficiency and operations
- Answering new business questions in existing businesses, with a focus on opportunities for growth
- Answering new questions in new businesses, with the goal of reshaping the competitive landscape.

While acknowledging that companies may be at many points along their data journey, research suggests the primary focus of big data applications has been on the first category of applications, and that the time is ripe to look more closely at the second two. A recent survey of more than 100 CIOs across multiple industries and geographies showed that big data—including its applications to business and the technologies to realize knowledge discovery—is expected to be one of the top three most disruptive technologies in 2013, third only to cloud deployment and mobile enablement. As defined in Clayton Christensen's *The Innovator's Dilemma*, a disruptive technology should create a new market that will ultimately overtake an existing one. In Christensen's terminology, the bulk of today's big data applications are generally *sustaining*, in that they merely improve existing products with the goal of making more profits from higher-end customers.

From big data 1.0 to big data 2.0

“If you do not change direction, you may end up where you are heading.”
—Lao Tzu

New infrastructures or data sources can unlock some value in big data by answering existing business questions. This happens when conventional ways of creating business value from data become untenable due to significant growth in the available data. For example, Rackspace started out with a small customer base in its email hosting service. Over time, the number of its customers grew rapidly to a million users, resulting in a considerable paper trail of 150 GB per day of logs recorded in various formats. The situation created a bottleneck in Rackspace’s capacity to respond to troubleshooting requests using its data.
Tasks that had previously taken minutes to complete now took hours because the company was still reliant upon a legacy system that could not scale to meet demand. Rackspace could only resume reaping business value from its email-hosting service by migrating to a Hadoop stack-based big data infrastructure.

Beyond enabling faster answers, big data may enable better answers as well. For example, telecommunications companies can improve the value of customer churn analyses by augmenting existing customer data with new data from customer interactions on social networks.\(^4\)

Yet upon closer inspection, in neither of these applications of big data do the fundamental strategies and ways in which businesses are run change. The companies’ goal of understanding customer churn, for example, still remains largely the same, only augmented with attributes gleaned from social media data. This relatively conservative approach seems characteristic of today’s big data applications. A survey conducted by the Economist in 2010 asked, “What new opportunities do you see for your organizations as the result of the availability of increased amounts of data?” Most respondents cited “increasing operational efficiency” as the top effect (51 percent). In striking contrast, “innovating new services/products” came in only at number four (24 percent).\(^5\)

This focus on efficiency may not be surprising given the state of the economy at the time of the survey, when businesses were more focused on cutting costs. However as the economy improves and the focus shifts from cuts to growth, different approaches may be warranted.

To create disruptive innovations, companies should adopt new paradigms and find new ways of creating and stimulating growth. Recall the way that content-creator driven Web 2.0 technologies disrupted the content-consuming Web 1.0 world, contributing to dramatic changes in the ways businesses interact with their customers, innovate in products and services, collaborate, and channel their marketing efforts.\(^6\) Similarly, big data 2.0 strategies can open up new markets, providing leading businesses with small windows of opportunity within which to reap significant benefits that their competitors have yet to identify.

**NEW BUSINESS STRATEGIES TOWARD DISRUPTIVE BIG DATA**

Our review of existing research and discussions with business leaders suggest three types of big data strategies for creating disruptive innovation.

The first is a **customer strategy** of leveraging access to customer interaction data to shape—instead of merely sense—customer behavior. Access to this data can give companies the opportunity to generate new profits by anticipating and directing the market toward unrealized demands. This strategy can be combined with a
EVOLUTION OF BIG DATA BUSINESS STRATEGIES—A TAXI COMPANY EXAMPLE

Big data 1.0 strategy

*Technology for scalability:* A Singapore-based taxi company, ComfortDelGro, at one point served taxi booking requests only via phone conversations with human operators. Over time, as the number of its customers grew, this booking channel became congested, presenting the need to invest in big data. The company spent over $60 million\(^7\) to ramp up an advanced taxi booking system based on an automatic dialing system and smartphone application. At the back end is a data infrastructure that supports storage and processing of hundreds of thousands of trips,\(^8\) 15,000 taxis, and billions of data points on real-time GPS location data.\(^9\) This move has scaled up the company’s capacity to serve 20 million taxi bookings per year.
Big data 2.0 strategies

*Shaping customer behavior:* ComfortDelGro had collected years’ worth of data on taxi usage and the fluctuating demand for taxis through the course of the day. Responding to a persistent increase in taxi demand at certain times of days and weeks, largely produced by Singapore’s population and tourism growth, the company revised its fares by applying various surcharges at different times and locations. This move actually helped shape the customers’ booking patterns in a way that allows the company to meet their demands more consistently.

*Creation of a new product or service:* A taxi company with access to real-time customer and taxi locations, and in combination with its stores of historical booking records, can technically forecast optimal routes that are not crowded during different times (at different times of day or on the weekend, for example). Based on this capability, the company can offer a new service in the form of real-time rerouting recommendations, on top of its usual taxi operations. Not only can this service improve the ability of its taxi drivers to anticipate both demand and traffic, it can also be sold to any other drivers in the city.

*Ecosystem view of data:* A reliable traffic auto rerouting service can be made from an ecosystem view of data shared among taxi operators and national land and environment authorities. These entities have complementary data and interests in this context. While the land authority owns temporal snapshots of traffic density across the city, the taxi operator stores sparser but more detailed trajectories of traffic movement collected from its active fleets. These data, combined with real-time weather and road-flooding information from the environment authority can enable better forecasts of route congestions. This service benefits the three entities simultaneously. Reduction in citywide traffic congestion is one of the primary goals of the land authority. Less congested roads translate to higher revenues for taxis. And finally CO₂ emission is reduced with less traffic jams, something that is in the interest of the environment authority.
product strategy, where unrealized demands come in the form of new products or services that monetize big data. As important as these strategies are, relying only on them may not lead to persistent benefits. This leads us to the ecosystem strategy, a third approach whereby companies aim to participate in or even shape a new industry-oriented community in which the sharing of data enhances the overall business propositions of its members.

Businesses in some domains, however, are beginning to actively shape, rather than remain content to merely sense, their customers’ behavior. This involves coupling a comprehensive view of customers, encompassing their behavior, preferences, and competitive actions, together with real-time location data that may be, for example, triangulated from cell tower or wireless hotspot signals.

Customer strategy: Shaping customer behavior

“The most significant idea for big data is that it allows you to see around corners and react,” according to Michael Cavaretta, technical leader of predictive analytics and data mining at Ford Research and Innovation. Such a reactive stance has been widely adopted by data-driven companies even in their dealings with customers. Until quite recently, the primary way for companies to sense their customers’ behavior was by engaging market research firms. Companies sought to respond to the needs of their customers based primarily on what they learned from these studies. Today, even as the expression of market sentiment has gradually shifted to social media, the essentially reactive approach of sensing customer behavior remains dominant.

Businesses in some domains, however, are beginning to actively shape, rather than remain content to merely sense, their customers’ behavior. This involves coupling a comprehensive view of customers, encompassing their behavior, preferences, and competitive actions, together with real-time location data that may be, for example, triangulated from cell tower or wireless hotspot signals. Doing so enables companies to make highly customized offers using the most suitable channels and at the appropriate moments.

Organizations such as Netflix and Amazon tap into such data to identify the tastes and preferences of their respective customer bases and use this information
to provide helpful and relevant real-time offers. These organizations thereby influence their customers’ purchase behavior. In the case of Netflix, the recommendations are not just limited to new movies but also include older titles to help reduce licensing costs.\textsuperscript{13} Retailers, too, gain access to customers’ preferences through customer loyalty cards and by leveraging real-time “check-in” data from Foursquare, for example, and then they try to influence customer shopping patterns by sending promotions via mobile apps.\textsuperscript{14}

We recently worked with a financial institution to critically assess its loan and borrowing risk exposure by pooling a large variety of data from macroeconomic indicators, such as consumer- and home-price indices, as well as from national-level charge-offs (how many loans are written off as uncollectible). This broad approach shaped the attitudes of financial institutions towards risk assessment by elevating the stress-testing bar to a more realistic level.

Yet implementing this strategy involves particular challenges. The chief concern is that of individual privacy: Matters pertaining to personal and sensitive information—even if such information has not been derived from personal data—are best handled judiciously and transparently. From an execution standpoint, companies should also try to anticipate the resulting changes in customer behavior. Given the uncertainty involved in estimating how many customers ultimately become influenced by recommendations, this is hardly a trivial task. In some cases, companies do not have sufficient visibility and control of their supply chain to allow the additional variability into customer demand that real-time offers would introduce. The conclusion is that companies should continuously monitor their customers to learn how much “shaping” is appropriate.

**Product strategy: Creation of new products and services**

There are organizations in the data value chain that are situated in busy spots of
data traffic, finding themselves in strategic positions to benefit economically from the data they have access to. These include organizations in the communications, media, and entertainment industries. Such organizations, by virtue of their significant interactions with customers over digital channels, are becoming repositories of vast amounts of valuable customer data.

Many of them have leveraged this data for insights that support the day-to-day business of serving their existing markets and clients. Banks have been collating 360-degree views of their customers from their profiles, transactions, and online and phone banking interactions to improve customer satisfaction. For example, out-of-cash ATM incidents can be minimized and the pricing of products and services can be refined. Others, though, have innovated and designed entirely new business models by monetizing their data and targeting new markets. Telecommunication companies for instance receive a fine-grained and real-time view of their large customer base encompassing location, usage, social networks, and other features through the smartphones being used. They can put this data to use to offer new services, such as location-based marketing. For example, on top of their regular phone services, the three local telcos in Singapore—M1, Starhub, and Singtel—in collaborations with Singapore Press Holding and other retailers have been offering location-based advertising (LBA) SMS to customers. The number of SMS’s being sent, and potentially the rate of consumer responses, then translate into additional revenues for the telcos.

Big data can therefore be used to provide lifestyle-related services to customers in real time. These strategies can help telcos improve customer retention as well as create additional revenue streams. The same concepts can be put to use in other areas as well. For example, insurance companies can create new products and services rather than just sell standardized insurance policies. This could be done by aggregating customer risk appetite, policy adoption, and claim history data over a period of time as inputs to a new service line of regulatory reporting that is more realistic than the conventional approach.

This product strategy is not limited only to established companies and their spin-offs. As the provision of new products or services typically caters to uncharted markets, this strategy offers big opportunities for new entrants as well. In the retail space, for example, real-time price comparison services such as GetPrice in Australia and PriceRunner in the UK provide customers with enhanced information while also opening up new channels for more targeted online advertisements. In the health care space, Castlight Health, founded in 2008, uses big data to inform patients of health care costs by providing them with pricing not usually available to customers. The community site PatientsLikeMe offers a free forum and friendly environment within which patients can seek others who have similar diagnoses,
medications, and even lab test results. It generates revenue by reselling data to pharmaceutical manufacturers, all the while remaining transparent with users about its use of their data ratings, reviews, and opinions.

Of course the creation of new products and services from big data is not without its challenges. New entrants should be sensitive to the legal and ethical aspects of data usage, especially when involving personal customer data or deriving information from big data of a personal nature for profit-making purposes. Legislation related to data is under constant scrutiny by policymakers around the world, and a growing body of case law in various jurisdictions is being developed. The regulatory landscape for data-monetization and profit-making opportunities will likely remain dynamic in the near future.

As big data technology advances apace, data protection and privacy legislation will likely continue to struggle to encompass all its possible uses. Therefore, businesses that adopt new customer and product strategies using big data have, at a minimum, an obligation to keep customers aware of how their data is being used and to provide them with sufficient information to enable informed choices. Doing so can allow benefits to be realized by both parties, and such transparency may help maintain regulatory and moral discipline as well as strengthen reputation, customer loyalty, and the brand.

Reliance solely on new products and services for data monetization can pose certain longer-term risks as well. Without being an integral part of an established system, it is possible that the new market can be disrupted by other new developments. From a data perspective, this calls for a viewing of data through the prism of an ecosystem within which data originators, beneficiaries, competitors, and regulators alike thrive and mutually benefit from the sharing of data and insights.

**Ecosystem strategy: Ecosystem view of data**

It often happens that a single organization does not have access to enough complementary views of its customers to create a new and compelling product or service. In such cases, organizations can fill in the blanks using complementary sources of data from other organizations within the ecosystem. Such ecosystems are built upon appropriate collaboration strategies implemented in such a way as to benefit all parties involved, from enterprises to consumers. The ecosystem view can take many forms. At one end of the spectrum is collaboration between companies viewed traditionally as competitors; at the other end are citywide collaborations between various public sector agencies to enable better delivery of services. In addition to the immediate short-term benefits of collaboration, the ecosystem strategy helps generate long-term benefits by ensuring risk diversification.

In the area of insurance, we are already seeing these types of data collabora-
tion emerge. Identifying and preventing fraudulent motor insurance claims, for example, improves profitability but also helps to keep premiums low for drivers. Members of the Association of British Insurers (ABI) in the United Kingdom are now combating this problem by sharing claims data from millions of customers and centralizing its analysis in the Insurance Fraud Bureau—a not-for-profit organization funded by the ABI. The information from its database, called the Insurance Fraud Register (IFR), is already having a significant impact on the estimated number of fraudulent claims made every year. According to the ABI, “The IFR will help insurers identify whether individuals have committed insurance fraud so that they may take appropriate action. The information may be used at any point in the lifecycle of the product from point of sale, at renewal, or when a claim is made and at any other point.”

Several music-related organizations, including publishers, music service providers, and songwriter associations, are also helping to shape the music industry’s digital future by creating a “Global Repertoire Database”—a single, authoritative view of all songs, for all territories and uses. This data will be usable by all organizations in the music publishing value chain to enable efficient and accurate licensing of music and subsequent payment of royalties for its use. The database marks a step change in approach to rapidly evolving online business models for music service provision, consumption, and licensing.

Although the evidence base is as yet relatively small, the strongest of these intra-industry strategies tend to focus on using big data to address a specific shared risk—regulatory, commercial, or technical—while still allowing participating companies to compete for business from customers as normal. This approach leads to the smallest set of potential conflicts that would otherwise tend to tear the consortium apart and echoes Evan Rosen’s view that collaboration among competitors makes sense when such associations create value for both parties, when they begin with structure and clarity, and when they involve nondifferentiating processes.

†Note that precedents exist for the “ecosystem strategy” of data pooling described here. For example, insurance companies routinely write commercial and professional liability contracts for relatively small numbers of heterogeneous risks. For such heterogeneous and low frequency/high severity insurance products as Medical Professional Liability (Medical Malpractice) or Workers Compensation insurance, few insurers collect enough data to reliably estimate the cost of providing insurance to the multitudinous types of risks seeking insurance. (For example, there are hundreds of medical specialties and types of businesses.) It is therefore not merely advantageous but essential for insurers to pool data through insurance rating bureaus. The bureaus in turn use this pooled data to actuarially estimate the expected loss frequencies and severities by risk classification.

The major US credit bureaus provide another long-standing instance of the ecosystem data strategy. These bureaus collect small amounts of borrowing and payment information from each of a number of financial institutions. In return, they provide more complete profiles (credit scores, cluster assignments, collections of variables, and so on) of individuals’ financial positions and financial stability.

It is reasonable to anticipate many more such arrangements emerging as individuals leave behind ever more digital "traces" of their daily activities.
Going cross-industry, big data provides telcos and financial institutions with the potential to generate more insight through collaboration than they could individually—particularly as retail payments and mobile technologies converge. By leveraging their respective customer data, they can create a truly differentiated mobile banking platform by collaboratively analyzing their combined data.

Government can also have a role in the ecosystem. Many businesses can benefit from the inclusion of additional data, such as real-time weather and transportation-related sensors, which is collected typically by the public sector, and which would be prohibitively expensive for any single company to replicate. Businesses are incentivized to work with government agencies to share investment costs in data collection where they also have a stake in the downstream impact of the service. For example, when planning transportation of shipments, a company can benefit from integrating its internal freight and booking data with external real-time port data from the sensors and radars deployed by port authority. It is in the interest of the port authority as well to ensure the logistic efficiency and the safety of personnel and ships, hence its willingness to invest in the sensor infrastructure.

Government–industry collaboration is also seen in the emerging ecosystem of open data, which is currently spearheaded by various governments around the world (with huge quantities of open government data now available, for example, on data.gov, data.gov.uk, data.gov.sg, and many other national data portals). Open data typically does not include personal data, which means that it is easier for organizations to handle and mix the data with their own proprietary sources. For example, the Practical Law Company is one of several commercial and academic institutions working hand-in-hand with the UK government to build an accurate and up-to-date repository of UK legislation, published as linked open data on the website legislation.gov.uk. In collaborating with the government and other supporters to maintain the open data, the Practical Law Company saves considerable time and expense it would otherwise have incurred on maintaining its own data set, but is now able to use the common, authoritative repository to compete for legal services on the open market.

While data sharing and collaborative analytics can bring tremendous benefits, there are considerable hurdles to overcome. The greatest challenge lies in sharing personal customer data that has been collected for other purposes. Recent surveys suggest that customers are wary of companies that share even completely anonymized data. And in many countries, data protection legislation explicitly precludes organizations using or sharing personal data without the explicit and informed consent of their customers. Larger ethical questions also arise as different pools of data are shared between different parties. What was once unconnected, and therefore implicitly private information, is now joined up and exposed to the scrutiny
BIG DATA 2.0
of advanced analytics. There are fears, too, that collaborators could learn too much from each other’s data, thus compromising their status of competing on a level playing field. Other potential pitfalls include collusion on the part of the collaborators to either fix prices or other similar trade practice violations. Operationally, the lack of a common data schema among even structured datasets across organizations—let alone unstructured data—makes collaborative analysis challenging. In this respect, governments can have a key role to play as catalysts for promoting standards through their open data initiatives.

Whichever strategy is chosen, businesses ought to develop the appropriate kinds of big data resources, from empowering and harnessing the desired talent, to investing in appropriate technology infrastructure to sustain operations. Companies should be mindful of the economics of big data and weigh the costs of the infrastructure and expertise required to store, sort, and analyze large volumes of data against the potential benefits big data can bring to business.

WHAT CAN BUSINESS LEADERS DO?

Each of the three strategies discussed offers big data opportunities in the right business context. Business leaders can ask themselves a number of questions to determine if they are positioned to tap into the positive disruptive potential of these strategies.

- **Consumer strategy**: Big data provides businesses the additional opportunity to shape customer behavior, with offers that address needs that customers themselves may not have recognized. To determine whether an organization is ready for such a proactive use of big data, it should answer a number of questions. What customer decisions are being made, and what are the processes involved in making those decisions? Is there an opportunity to influence these decisions with new data? If so, where does the necessary data reside? Does the appropriate infrastructure exist to serve it up cost-effectively and in a timely fashion (including real-time, if necessary)?

- **Product strategy**: Businesses also need to assess whether they are in a position to create new products and services that will give them a competitive edge. This involves answering questions about the value and the volume of
the data they possess. Do they have any unique assets? Would these assets in some combination address a market need? Will it be new or existing markets that will respond to the new products and services, and if new ones, what channels can be used to penetrate new markets? Will the investment into new products or services come at an opportunity cost to their existing business operations?

- **Ecosystem strategy**: A business should analyze whether it is in a position to derive maximum value from strategy shifts in isolation or whether it would be better off collaborating with other businesses to form a unique, enhanced interpretation of that data. Do businesses have complete visibility of all other players in their business value chains? If so, business leaders should determine whether datasets or business insights owned by these players can complement their own businesses. Furthermore, business leaders should also establish whether it is possible to share data without losing their competitive edge.

Not every business will be in a position or have the necessary capabilities to implement all three strategies simultaneously, or they may need to implement only one or two of the strategies to improve performance in aspects of business they are targeting. Whichever strategy is chosen, businesses ought to develop the appropriate kinds of big data resources, from empowering and harnessing the desired talent, to investing in appropriate technology infrastructure to sustain operations. Companies should be mindful of the economics of big data and weigh the costs of the infrastructure and expertise required to store, sort, and analyze large volumes of data against the potential benefits big data can bring to business.

Is big data revolutionary? While there is certainly increased awareness and a growing number of tools related to big data, for the majority of businesses, the transformation has not yet been disruptive. There will be new players that will emerge strongly and new markets that will open in the near future as people leverage the benefits of big data and advance new business strategies that go hand in hand with big data. Those businesses that ignore the hype and focus on how big data might be used to ask new business questions and meet market needs will be well positioned to use the data to drive sustainable competitive advantage.
Vivekanand Gopalkrishnan is a director with Deloitte & Touche Financial Advisory Services Pte Ltd Singapore and leads Deloitte Analytics Institute Asia.

David Steier is a director with Deloitte Consulting LLP and leads the Deloitte Analytics Solutions Group.

Harvey Lewis is a director with Deloitte UK and leads the research programme for Deloitte Analytics in the UK.

James Guszcza is the National Predictive Analytics lead for Deloitte Consulting’ LLPs Actuarial, Risk, and Analytics practice.


Endnotes