Point of View

Unlocking Value in the Fragmented World of Big Data Analytics

How Information Infomediaries Will Create a New Data Ecosystem

Authors
Bill Gerhardt
Kate Griffin
Roland Klemann

June 2012

Cisco Internet Business Solutions Group (IBSG)

Unlocking Value in the Fragmented World of Big Data Analytics

How Information Infomediaries Will Create a New Data Ecosystem

Data Is "the New "Oil"

As far back as 2006, market researcher Clive Humby declared data "the new oil."[1] Today, as data transcends into the ever-expanding realm of Big Data, a growing community of experts is beginning to agree. They see in Big Data the same transformative, wealth-creating power that 19th-century visionaries once sensed in the crude black ooze trapped underground.

Just as oil once fired dreams a century or more ago, data is today driving a vision of economic and technical innovation. If "crude" data can be extracted, refined, and piped to where it can impact decisions in real time, its value will soar. When key insights can be mined from it through analytics—revealing complex behaviors, patterns, and events as they happen, if not *before*—then data will realize its inherent power. And if data can be properly shared across an entire ecosystem and made accessible in the places where analytics are most useful, then it will become a true game changer, altering the way we live, work, play, learn, buy, sell, travel—you name it.

But with Big Data comes big challenges. The problem today is that the ever-increasing deluge of information—terabytes to petabytes to exabytes—threatens to swamp us in a gusher of unfiltered, unstructured, unprocessed, and seemingly unmanageable information. As things stand now, the data ecosystem is highly fragmented. Between those who create data and those who could potentially extract value from it sits a labyrinth fraught with complexity, disparity, and miscommunication. If analytics are to be the new "refinery," some of that fragmentation will need to be addressed with greater connectivity, trust, and efficiency.

Who can and will take on this responsibility? The Cisco® Internet Business Solutions Group (IBSG) believes that this overarching role will be filled by a new class of players: Data Infomediaries. These entities will step in and tie data originators to data beneficiaries, empowering the data originators and facilitating every step along the data-value chain. In the process, they will create opportunities for all players to harmonize, realize, and monetize the vast potential of Big Data.

Crucial to this process will be the transport and storage infrastructures for Big Data. Players who can see far and wide—across networks and data centers, collecting and controlling data in motion—will be in the best position to link disparate Big Data players into a smoothly functioning whole.

In this paper, we will examine some of the wider implications of Big Data, its transformative promise, and some of the pain points that arise in trying to tame it. Above all, we will explore the opportunities for Data Infomediaries to energize Big Data's potential as a truly transformative force, "the new oil" of our age.

[1] http://ana.blogs.com/maestros/2006/11/data_is_the_new.html

The True Nature of Big Data

In seeking better-informed decision making, many organizations are running data warehouses and employing traditional data analytics—to reduce churn, bolster campaign effectiveness, and counter fraud, to name a few applications. Big Data represents a revolutionary step forward from traditional data analysis, characterized by its three main elements: variety, volume, and velocity.

- The **variety** of data comes in two flavors: structured and unstructured. Structured data enters a data warehouse already tagged and is easily sorted. The vast majority of today's data, however, is unstructured, and fed by sources such as Facebook, Twitter, and video content. It's random, difficult to analyze, and *enormous*.
- The sheer **volume** of Big Data overwhelms the normal data warehouse. For example, Facebook reports that its users register 2.7 billion likes and comments per day. For many, this magnitude of data is intimidating: they can't keep up with it, much less sort it, analyze it, and extract value from it.
- All of that data can be challenging to manage when flooding in at a **velocity** that, for many players, far outpaces their processing ability. In order for Big Data to be a game changer, it needs to be analyzed at a rate that matches the blistering speed at which information enters data warehouses. In microseconds, decisions must be made as to whether a particular bit of data deserves to be captured, and whether it has relevance when combined with other data. For example, the face of a known criminal among thousands of images might trigger a "stop"; a pattern of credit fraud might set off warnings; and indications of growing customer churn might inspire a coupon offer. Anytime there is an important anomaly in the data, it needs to be pointed out *before* the data is stored so that real-time action can be taken.

Empowered by the Network

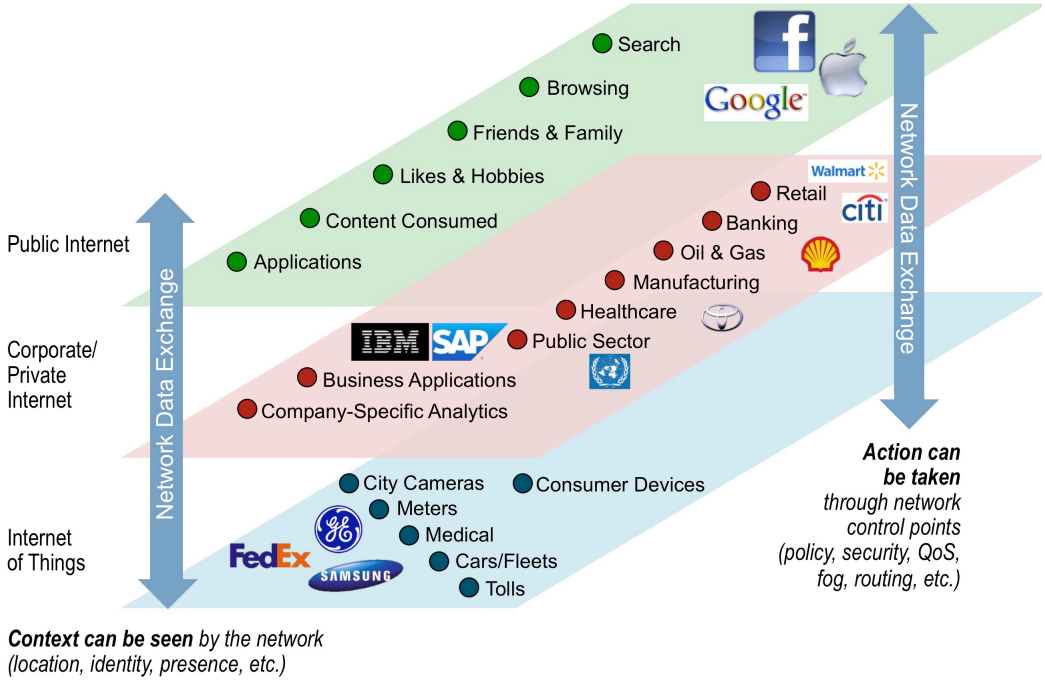
By its very nature, network traffic is Big Data. In just one slice of the network—the mobile network—there are 6 billion mobile subscriptions in the world, and every day, 10 billion text messages are sent.[2] International Data Corp., the global market intelligence firm, estimated that in 2011, all of the data created in the world amounted to 1.6 trillion gigabytes.

The amount of data crossing networks will continue to explode. By 2020, 50 billion devices will be connected to networks and the Internet.[3] But the network is far from a "dumb pipe" simply transporting those reams of data. Beyond supporting the sheer volume, the network can play two valuable roles in increasing the Big Data's potential for enterprises (see Figure 1).

First, the network can collect data and provide context at a high velocity. It can corral data from devices, private networks, and public wide-area networks, often in "network time." This is the point where value is extracted and acted upon as events occur, in real or near real time. With the data that the network collects, it can determine context. That context might include matching location, identity, or presence information to the data. Moreover, actual content of the data that is collected can have value. Knowing not only which sites people browsed, but also what they did on those sites, can provide crucial insights.

[2] International Telecommunication Union, 2012.
[3] "The Internet of Things: How the Next Evolution of the Internet Is Changing Everything," Cisco IBSG, 2011.

Figure 1. The Network Connects Data Domains.



Source: Cisco IBSG, 2012

Second, the network makes it possible for enterprises to take immediate action on their data insights. Through a series of control points in real time, the network can reach the customer or impact the customer's experience. These control points exert changes in policy, privacy, preferences, security, quality of service, and rerouting or local processing.

Having both the context and control aspect of the network in hand, those organizations with an end-to-end mastery of the network can see wide and deep.

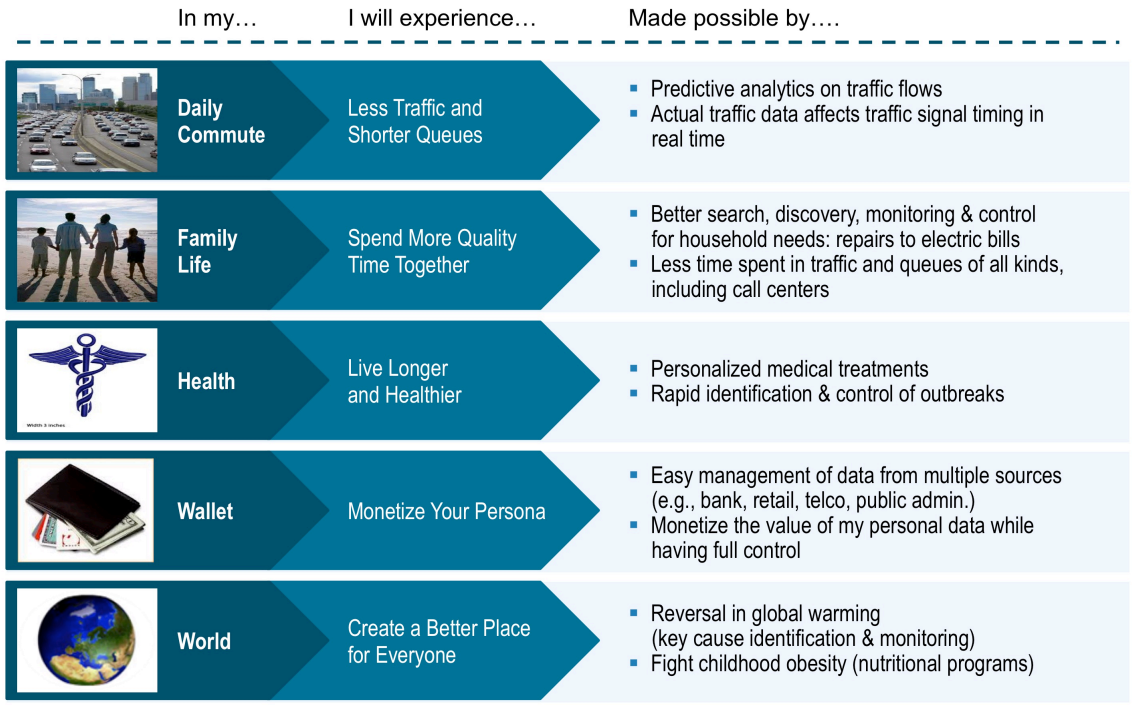
How Big Data Will Change Your Life

If Big Data's potential is fully realized, it will have a transformative impact across a broad spectrum of everyday tasks and activities, both complex and common (see Figure 2).

Traffic jams are just one example. They may seem a necessary evil, here to stay. But imagine if they were vastly reduced or even eliminated. The savings in stress and time (which, of course, equals *money)* would be enormous. Quality of life would increase, as commute times are cut and family time increased. The answer is not flying cars but rather predictive, real-time analysis on traffic flows, feeding immediate changes to traffic signals, digital signs, and routing, before backups begin.

What about those paper receipts from retailers and banks that are cluttering your wallet? Imagine if they were replaced by electronic records. Businesses could enrich these records through contextual and comparative information. Individuals could manage, share, monetize, and utilize the data through, for example, budget management and health advice applications.

Figure 2. Relevance of Big Data.



Source: Cisco IBSG, 2012

Medical advances enabled by Big Data could include treatments personalized to a patient's unique health issues. Epidemics of contagious diseases or food contamination could be predicted, tracked, and curtailed. Already, Google Flu Trends has aggregated data on searches for health information to predict outbreaks across the globe.

On a planetary scale (or beyond), Big Data can monitor and analyze vast amounts of information on anything from climate change, astrophysics, and energy consumption to geopolitics and socioeconomics. In all cases, Big Data analysis can lead to a deeper understanding of key causes behind crucial events while anticipating future trends.

Making Big Data Business-Friendly

Big Data can also change how we interact with businesses. Moment to moment, consumers' interactions with the world around them create an often-ignored by-product: massive amounts of personal data. These include searches on weather, price comparisons, purchases, and thousands of other daily choices and actions. By analyzing the data generated by all of this activity, Big Data offers an opportunity to revolutionize the way consumers and sellers interact.

Some businesses are already moving to take action. Here are a few examples:

- Harrah's enhanced revenues by 8 to 10 percent by analyzing customer segmentation data,[4] while Amazon at one time reported that its analytic recommendation engine drove 30 percent of its revenues.[5]
- Wal-Mart's inventory management implemented radio frequency identification (RFID) technology to connect real-time information between suppliers and its Retail Link data warehouse. In the process, it reduced out-of-stocks by an estimated 16 percent. [6]
- FedEx achieved real-time visibility with shipping and consumer data across more than 46,000 distribution and supply chain locations.
- Kaiser Permanente reduced operating costs by 40 percent through integration of its electronic medical records and its inpatient and outpatient management system. Integration of its clinical and cost data led to the discovery of Vioxx's adverse effects and subsequent withdrawal of the drug from the market.[7]

Consequently, Big Data has already become top of mind among CIOs.[8] Yet, Cisco IBSG envisions an even higher level of data analysis and interaction along the data value chain. We have identified eight distinct classes of use cases derived from our analysis of a few hundred potential applications of Big Data across many vertical industries:

1. **Revenue Assurance.** Integrating intra-company data could hone identification and prevent fraud before it occurs. Fraud-intensive industries such as healthcare would particularly benefit.
2. **Risk Mitigation.** Every day, networks carry petabytes of critical information for enterprises, governments, and consumers, opening an ever-increasing risk of intrusions and security attacks. Data federation across wider geographic and network footprints would enable identification of suspicious patterns while signaling the need for immediate action.
3. **Customer Lifecycle.** Businesses can zero in on instances of customer frustration and offer an immediate response, thereby improving the consumer experience and lessening churn. Any service-based industry that values its customer relationships would benefit.
4. **Market Execution.** Big Data enables better market services through analytics, creating improved opportunities for cross-selling and up-selling. Banking and Internet commerce stand out as potential beneficiaries.
5. **Product Innovation.** Consumer input is critical in product development, and today many companies are already clamoring to know more about the likes and dislikes of their customers. Integrating noncompany sources of data such as social-network feeds would provide a more holistic view of how consumers feel about a product, potentially revealing the need for a new product *before* it is imagined or on the drawing board.

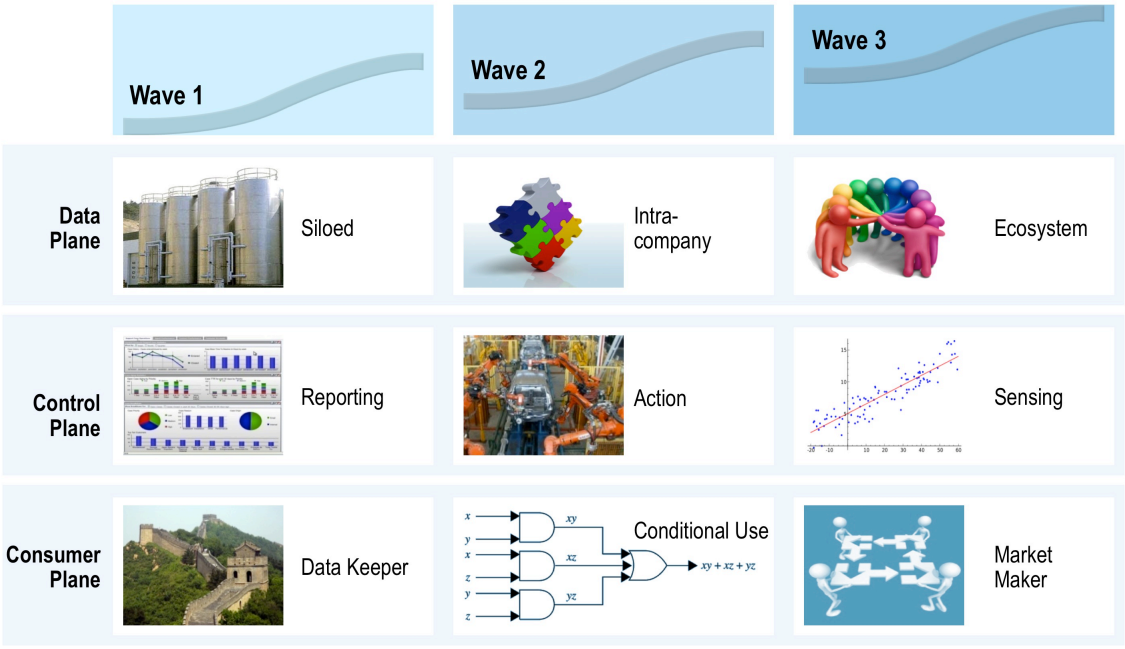
[4] Gartner, 2005.
[5] McKinsey & Company, 2011.
[6] "Does RFID Reduce Out-of-Stocks?", University of Arkansas, 2005; "Opening Up the Big Box: Measuring the Wal-Mart Effect," *The Economist,* February 23, 2006.
[7] McKinsey & Company, 2011.
[8] A 2011 IBM survey found that 83 percent of executives chose Big Data as their key topic over mobility solutions, virtualization, and cloud computing.

6. **Business Models.** New pricing algorithms will create novel monetization opportunities and more efficient interactions with consumers. Retailers in particular could benefit from new pricing models tied to the location and behavior of consumers.
7. **Operations Management.** Big Data could help almost any organization run better and more efficiently. A service provider could improve the day-to-day operations of its network. A retailer could create more efficient and lucrative point-of-sale interactions. And virtually any supply chain would run more smoothly. Overall, a common information fabric would improve process efficiency and provide a complete asset view.
8. **Advanced Advertising.** The more enterprises learn about customer behavior, the easier it is to turn that information back into advertising. eMarketer has found that with advanced, data-based targeting, the effectiveness of ads doubles. Cost-per-thousand (CPM) lifts of more than 50 percent are possible. To further exploit these advantages through Big Data analytics, trust mechanisms will need to alleviate privacy constraints and ensure that targeted ads are aligned with the consumer information crossing the network.

Catching the Waves of Big Data Evolution

The full potential of these use cases will be realized in three waves of industry evolution (see Figure 3). In each wave, the analytics activity can be described on the Data Plane, the Control Plane, and the Consumer Plane. In the Data Plane, we look at the sources of data used. In the Control Plane, we look at how the data is being used, how insight is being created, and what happens once a finding or determination occurs. In the Consumer Plane, we look at consumers' relationship to the data analytics process, bearing in mind that consumers often are the creators of the data—and only sometimes the owners of the data.

Figure 3. Big Data Industry Evolution.



Source: Cisco IBSG, 2012

- Today, most industries are squarely entrenched in **Wave One.** The data is mostly siloed within a particular business unit within an organization. The data is used primarily to create ex-post reports. After the fact, data is compiled, but the impact of these reports on any decisions is indirect. As for the consumers, they have little control over—or awareness of—their own data. Their main worry is about the security and integrity of their data, and a vague feeling of potential data abuse prevails.
- In **Wave Two,** the Data Plane moves from siloed to intra-company. Data can move across individual silos within the organization and out into a common repository. Once there, the much broader set of collected data can be analyzed and shared for more trenchant decision making. And, on the Control Plane, new insights result in actions. Instead of simple insights on one data variable, multiple variables and considerations may create an alert that results in a change (for example, a product in inventory is ordered or a coupon offer is sent to a customer). From a consumer perspective, there is a growing sense at this stage that granting access to personal data may have its rewards. Consumers can choose among voluntary opt-in options to determine how their data is used.
- By **Wave Three,** data is being collected, shared, and processed across an entire ecosystem. Enterprises can leverage not only their own data but also a myriad of third-party data, including weather patterns, economic indicators, "anonymized" retail sales, social media traffic, and so forth. At this stage, organizations are using the data in increasingly complex processes to perform predictive analysis, sensing events *before* they happen. Once such insights are delivered, the organization may undertake automated, real-time actions even without human intervention (e.g., rerouting traffic, re-ordering stock). In Wave 3, the consumers embrace the idea that there are opportunities to create value from their personal data. They set policies for use of their data, and trust that those policies are enforced throughout the ecosystem. Increasing transparency across the ecosystem makes consumers more comfortable, as they see how their data is being used. Many engage voluntarily in a marketplace where their data is a product with value—resulting in tangible benefits for consumers.

All of these ideas are spreading and, as Cisco IBSG is seeing, already having an impact. With the transition from Wave One to Wave Three, the rewards get bigger—but the pain points can become ever more acute.

Pain Points on the Road to a Big Data-Enabled World—How To Fix Them

The ultimate goal with Big Data is to collect, store, find meaning, and extract value from data. As things stand now, the Big Data ecosystem appears fractured. But there is plenty of potential to orchestrate all aspects of the ecosystem, allowing the creator and the beneficiary to work in harmony, with a much higher level of trust.

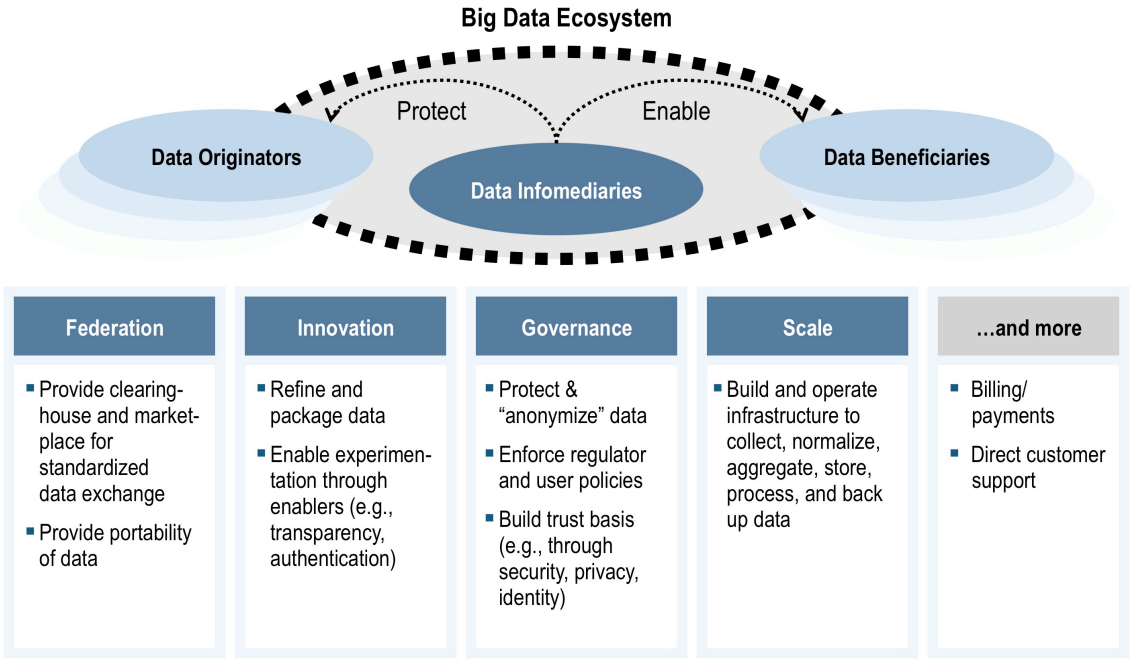
Let's look at the key success factors for other complex ecosystems such as financial transaction clearing, cellular roaming, electronic data interchange, automated road toll attendants, and electronic health records. Our research suggests that complex ecosystems need outside entities to perform value-added functions to enable Federation, Innovation, Governance and Scale.

In some ecosystems, "intermediaries" have arisen to fulfill these functions and enable their industries to prosper: TransUnion in banking, Syniverse Technologies in mobile roaming,

and RosettaNet in supply chain management. In contrast, other industries, such as automated road toll in Europe and electronic health records in the United States, have struggled. In those cases, no such entities were willing or able to take on an infomediary role, and the result has been a lack of progress.

Enter the "Data Infomediary." Cisco IBSG believes that within the Big Data analytics ecosystem, an intermediary or infomediaries will become a necessity if the industry is to reach its ultimate potential and satisfy all of the Big Data hype. We think the Data Infomediary must play four roles (see Figure 4 and below):

Figure 4. Role of Data Infomediaries.



Source: Cisco IBSG, 2012

1. **Federate: Enabling a Viable Ecosystem.** Data needs to be taken from the place where it was created to the place where it has value, while alleviating privacy and security concerns along the way. Today, however, there is no viable marketplace in which to share and benefit from a variety of data sources. Nevertheless, innovators are discovering the value of varied data. Target, for example, has carefully data-mined the consumer habits of its pregnant shoppers, since before and after childbirth there are great opportunities to establish new brand loyalties. As *Forbes* magazine suggested, Target could potentially share this valuable database of information with other companies, under the right circumstances and within the proper marketplace.[9] And a number of insurance companies are augmenting their own claims data with third-party databases that include social media and industrywide sources for anomaly detection targeted to identify new patterns of fraud.

[9] "How Companies Learn Your Secrets," Charles Duhigg, *The New York Times,* February 16, 2012; "How Target Figured Out a Teen Girl Was Pregnant Before Her Father Did," Kashmir Hill, *Forbes,* February 16, 2012.

Yet in each case today, enterprises still need to develop the means to use external data by themselves. Data Infomediaries, on the other hand, could identify the value, locate the source, work out the legal/business deals, and clean and standardize the data format so that it works with in-house systems. The Data Infomediaries would facilitate the sharing of data across the ecosystem, creating a clearinghouse and marketplace for the exchange of data within and across industries. They would also vastly simplify the convoluted process facing companies first looking to benefit from Big Data, especially smaller companies that are not among the select few already investing heavily in Big Data. To facilitate this, Data Infomediaries could create shared standards and identify who wants to buy the data and how they can create value from it.

2. **Innovate: Deriving Value From a Sea of Chaos.** For the end user, acquiring value in a tsunami of data can be like finding the proverbial needle in a haystack. Many organizations are challenged even to experiment with large amounts of data, much less to create value from that gleaming, isolated tidbit of insight. By taking on the roles of identity management and authentication, the Data Infomediary makes it easier for companies to experiment. Once authenticated, that data would be portable and ready to be shared ubiquitously. From the consumer's perspective, data that is stored on Facebook, for example, could be moved to Myspace or Amazon without having to be re-created each time.

3. **Governance: Internal and External Alignment of Resources.** Someone needs to bring the currently fragmented Big Data ecosystem together. With data passing through varied way stations—both internal and external—a lack of orchestration and governance is blocking the path to value. To counter this, the Data Infomediary would align resources—internally and externally, publicly and privately—to create harmony under a common cloud structure. This approach would protect data from sabotage, while assuring government agencies that it is being used properly. Given recent scrutiny of Facebook, for instance, the privacy issue looms large. The infomediary could contribute to setting industrywide standards, ensuring that Facebook's data is not perceived as something destructive but rather as something of value, to be used according to consumers' wishes. By gaining the trust of consumers and Big Data players alike, Data Infomediaries would make even more data available for analysis, spurring further real-time action.

4. **Scale: Dealing With Computational Requirements.** The price of handling Big Data will be prohibitive for many organizations, creating the need for third-party services to do their heavy lifting. Without access to large-scale capabilities—whether internal or external—bottlenecks in managing data soon arise. And even if internal storage is scaled to handle petabytes, concurrent requests for data analysis can still max-out a system. The Data Infomediary would play a crucial role in coping with the huge volume of data. The infomediary would offer more scalable and economic methods of storage, and aid in the all-important processing of data. The Apples, Amazons, and Facebooks of the world are already investing in these resources on their own and are gaining the ability to analyze data in real time, as it is created.

A number of players can step into Data Infomediary roles—for example, service providers, IT firms, banks, enterprises, Web 2.0 companies, government-sponsored agencies, or new

start-ups. The ability of those players to collectively remove the aforementioned pain points will make or break Big Data's future success.

Down the Big Data Rabbit Hole: A Call to Decision Makers

Business leaders are beginning to wonder just where to enter into the realm of Big Data and how far to venture down the Big Data rabbit hole. Here are some of the options that decision makers will need to consider:

- **Experiment in Wave 1.** The organization can continue to collect data selectively for individual, siloed applications, such as pricing, and extract value from it when possible.
- **Step Up to Wave Two.** Here, an organization can choose to shift into a wider dimension, from tactical thinking to strategic data empowerment. The strategy may still be limited to internal operations, but a commitment and investment are made to differentiate the service experience through data analysis and proactive resolution.
- **Think Externally.** At this stage, an organization realizes that those reams of information now being processed so much more efficiently and insightfully might just be of value to others. So, one party within an ecosystem could sell its location data to another party and get a transactional revenue rate for its services. Wal-Mart, for example, might choose to sell its retail data to Virgin Media; Amazon might rent some of its extra cloud computing capacity to small and medium-sized businesses that would not otherwise have access to such an infrastructure.
- **Become a Data Infomediary.** Certain organizations may aspire to become that most-trusted player in the industry, the Data Infomediary. At this stage, they can contribute to removing industry barriers, bridging domains, offering new services, and providing protection and security.

Getting There from Here: How To Win with Big Data

As Cisco IBSG maps out the future evolution of Big Data, it will define the strategic trade-offs that should be considered by organizations to maximize their potential advantages. In particular, Cisco IBSG believes that the rapidly expanding importance of Big Data presents great opportunities for potential Data Infomediaries. And as the market edges toward a key inflection point, moving from internally focused analytics to a federated/ contextual model, the time to act is now.

Organizations can still transform their data analytics, even if they opt to stay within an internally focused strategy. But many will benefit from a more externally focused strategy and perspective, sharing data and insights across an ecosystem. Others, particularly network service providers (with their end-to-end access to the network), may step up to the role of the Data Infomediary, sowing harmony, trust, innovation, and efficiency across the labyrinth of the Big Data ecosystem.

From Cisco IBSG's perspective, it is not a question of whether or not most organizations should invest in the future of Big Data, but rather how far they should go. Like those moguls-to-be who once weighed the costs and potential benefits of developing an oil business, today's decision makers are challenged to maximize value from the coming age of Big Data, at whatever level they choose.

As in the late 19th-century, the size of the potential prize makes it hard to ignore.

Acknowledgements

The authors would like to acknowledge the invaluable contributions of Neeraj Arora, Marc Latouche, Thomas Renger, Rob Rowello, and Evgenia Ryabchikova to the development of this paper.

More Information

Cisco IBSG (Internet Business Solutions Group) drives market value creation for our customers by delivering industry-shaping thought leadership, CXO-level consulting services, and innovative solution design and incubation. By connecting strategy, process, and technology, Cisco IBSG acts as a trusted adviser to help customers make transformative decisions that turn great ideas into value realized.

For further information about IBSG, visit http://www.cisco.com/ibsg

Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV Amsterdam,
The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at **www.cisco.com/go/offices.**