ATKearney

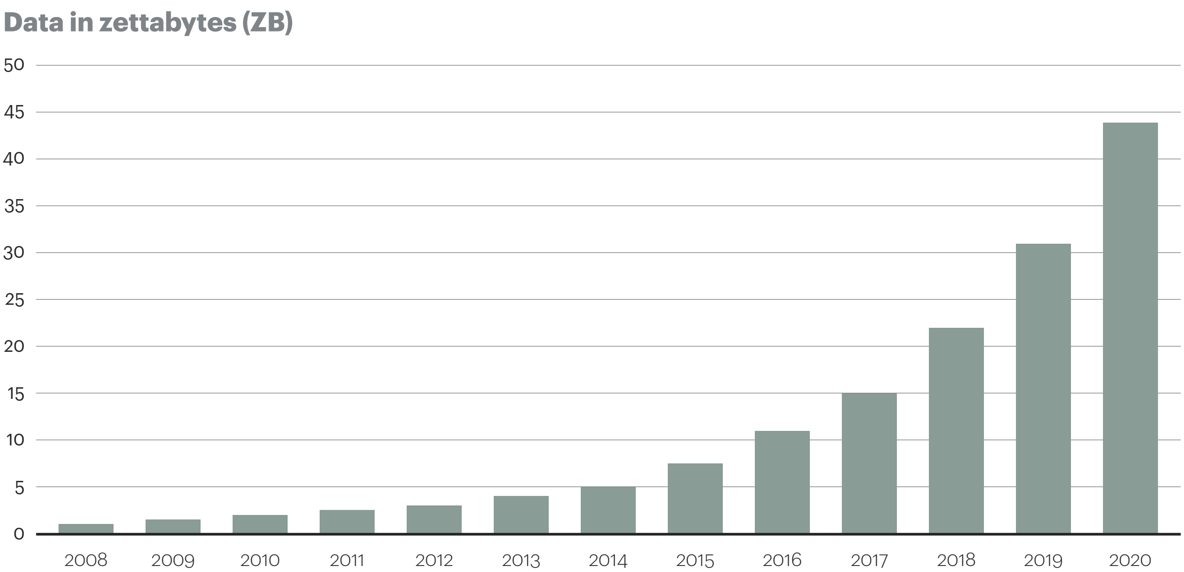
Big Data and the Creative Destruction of Today's Business Models

Big data—enormous data sets virtually impossible to process with conventional technology—offers a big advantage for companies that learn how to harness it.

Renowned 20th-century economist Joseph Schumpeter said, “Innovations imply, by virtue of their nature, a big step and a big change … and hardly any ‘ways of doing things’ which have been optimal before remain so afterward.”

Schumpeter’s words describe big data’s potential creative destruction of today’s business models. The world is inundated with data every minute of every day, and it’s not slowing down. An estimated 2.5 zettabytes of data were generated in 2012 alone, and trends indicate that the volume of business data will grow significantly every year (see figure 1).[1] Consider these statistics:

Figure 1
Data is growing at a 40 percent compound annual rate, reaching nearly 45 ZB by 2020



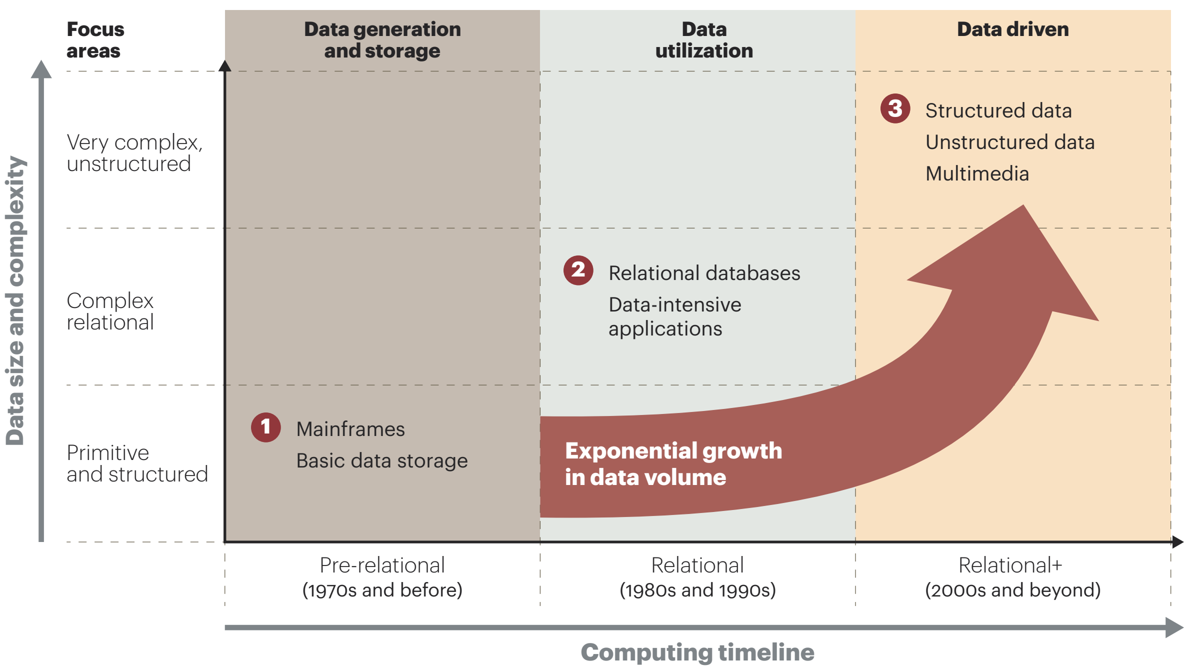
Source: Oracle, 2012

- Every day, 2.5 quintillion bytes of data are created, with 90 percent of the world’s data created in the past two years alone.[2]
- Data production will be 44 times greater in 2020 than in 2009.[3]
- The volume of business data worldwide is expected to double every 1.2 years.[4]
- Wal-Mart processes one million customer transactions per hour, stored in databases estimated to contain more than 2.5 petabytes of data.[5]
- The enormous data influx is straining IT infrastructures. In a recent survey, 55 percent of executives said data is slowing down their IT systems.[6]
- Poor data management can cost up to 35 percent of a business’s operating revenue.[7]

[1] One zettabyte is approximately equal to a thousand exabytes or a billion terabytes.
[2] “Bringing Big Data to the Enterprise,” IBM, 2012
[3] “A Comprehensive List of Big Data Statistics,” Wikibon Blog, 1 August 2012
[4] “eBay Study: How to Build Trust and Improve the Shopping Experience,” KnowIT Information Systems, 8 May 2012
[5] “Big Data Meets Big Data Analytics,” SAS, 2011
[6] “Global Survey: The Business Impact of Big Data,” Avanade, November 2010
[7] “‘Big Data’ Facts and Statistics That Will Shock You,” Fathom Digital Marketing, 8 May 2012

Big data promises to be transformative. As computing resources have evolved, advancing to better handle data size and complexity, companies stand to reap many more benefits from big data and analytics (see figure 2). Little wonder that big data is a hot topic in corporate boardrooms and IT departments, with many leading firms doing more than talking. According to a recent A.T. Kearney IT innovation study, more than 45 percent of companies have implemented a business-intelligence or big data initiative in the past two years.[8] Further studies estimate more than 90 percent of Fortune 500 companies will have at least one big data initiative underway within a year.[9] The effective use of this tidal wave can deliver substantial top- and bottom-line benefits (see figure 3 on page 4). Building capabilities in this area will not only improve performance in traditional segments and functions, but also create opportunities to expand product and service offerings.

Figure 2
The evolution of big data



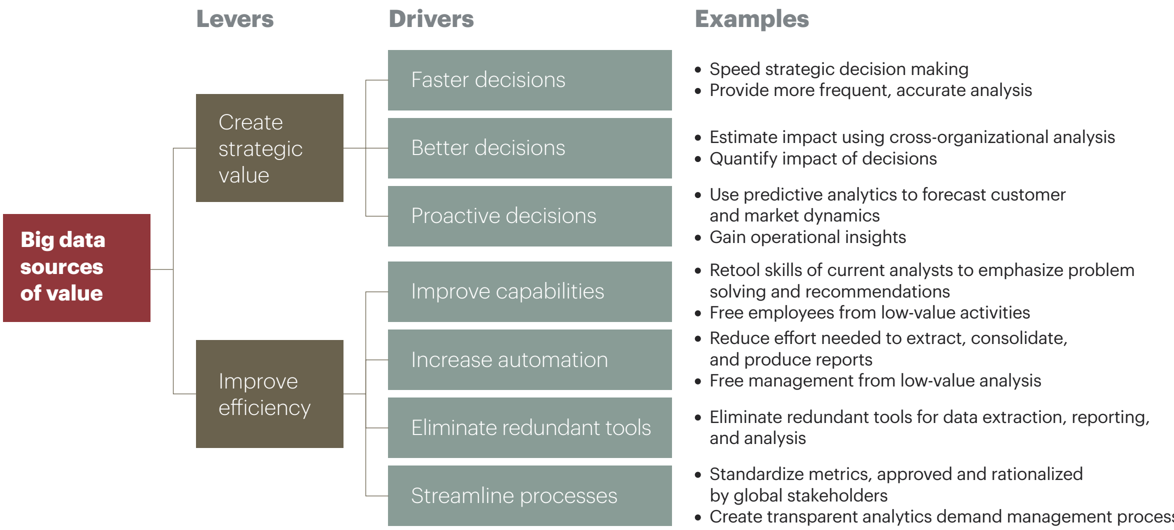
Source: A.T. Kearney analysis

In some industries, such as financial services, big data has spurred entirely new business models. For example, algorithmic trading now analyzes massive amounts of market data on a minute-by-minute basis, identifying opportunities to capture value almost instantly—a process unheard of a decade ago. In the retail sector, big data expedites analysis of in-store purchasing behaviors in near real time. With such quick insight into demand shifts, stores can adjust merchandise, stock levels, and prices to maximize sales.

Across industries, big data can produce large data sets coupled with massive processing capabilities to spur growth and reveal cost-reducing opportunities. While every industry uses

[8] See "IT Innovation Spurs Renewed Growth" at www.atkearney.com.
[9] "Big Data Market Set to Explode This Year, But What is Big Data?," Smart Planet, 21 February 2012

Figure 3
Harnessing big data can boost top- and bottom-line results



Source: A.T. Kearney analysis

different approaches and focuses on different aspects from marketing to supply chain, almost all are immersed in a transformation that leverages analytics and big data (see figure 4). Governments are not excluded from this transformation. Big data and analytics are used to improve services to citizens, optimize taxpayer funds, and better protect nations with advanced weaponry, such as unmanned vehicles and aircraft.

Figure 4
Industries are using big data to transform business models and improve performance in many areas

Illustrative

Retail	**Manufacturing**
• Customer relationship management • Store location and layout • Fraud detection and prevention • Supply chain optimization • Dynamic pricing	• Product research • Engineering analytics • Predictive maintenance • Process and quality analysis • Distribution optimization
Financial services	**Media and telecommunications**
• Algorithmic trading • Risk analysis • Fraud detection • Portfolio analysis	• Network optimization • Customer scoring • Churn prevention • Fraud prevention
Advertising and public relations	**Energy**
• Demand signaling • Targeted advertising • Sentiment analysis • Customer acquisition	• Smart grid • Exploration • Operational modeling • Power-line sensors
Government	**Healthcare and life sciences**
• Market governance • Weapon systems and counterterrorism • Econometrics • Health informatics	• Pharmacogenomics • Bioinformatics • Pharmaceutical research • Clinical outcomes research

Source: A.T. Kearney analysis

Big Data Defined

Big data typically refers to the three Vs—volume, variety, and velocity—of structured and unstructured data pouring through networks into processors and storage devices and the conversion of that data into usable business information (see figure).

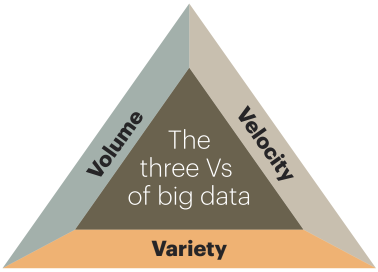
Volume. The amount of data being generated today can only be processed and managed with specialized technology. Harnessing the full capability of technology advances is the key to unlocking big data's potential.

Velocity. The rate at which big data is processed to provide usable results

Variety. The great range of data types and sources from which to draw insights

In essence, big data is the assemblage of infrastructure, data sources, software, and skills that support the three Vs, allowing companies to undertake more relevant and timely analyses than is possible with traditional business-intelligence methods.

Figure
Big data is defined by velocity, volume, and variety



Source: A.T. Kearney analysis

Yet few organizations have fully grasped what big data is and what it can mean for the future (see sidebar: Big Data Defined). Those that implement big data initiatives tend to do so as an experiment—to test the waters rather than build an important capability. And while we believe no organization should miss the opportunities that big data offers, we also believe that the greatest rewards go to those with a clear vision for how it can transform their organization, capabilities, and industry.

The hardest part is knowing how to get started. Four questions should be asked at the beginning of this transformation journey:

- Where will big data and analytics create advantage for our company?
- How should we organize to capture the benefits of big data and analytics?
- What technology investments can enable the analytics capabilities?
- How do we get started on the big data journey?

Where Will Big Data and Analytics Create Advantage for Our Company?

Understanding where big data can drive competitive advantage is essential to realizing its value. For many companies, the insights drawn from big data have already resulted in profitable, sustainable growth in three areas: customer intimacy, product innovation, and operations.

Customer intimacy

Big data puts the customer at the heart of corporate strategy. Organizations are inundated with customer data from interactive websites, online communities, and government and third-party data banks. Information on social-media platforms such as Facebook is particularly telling, with users sharing nearly 30 billion pieces of content daily. At the same time, it is now possible to bring together social-media feeds with disparate sources, including weather data, cultural events, and internal data such as customer contact information. Further, advanced analytical tools allow for faster, more effective, and less costly processing and create the potential to rapidly develop new insights.

But you know this. So let's look at some real-world examples of how big data is being deployed. A U.S.-based retail bank looks at social-media activities to identify at-risk customers, while an Asian bank analyzes customer-call audio logs to compile true sentiments. This kind of opinion gathering not only captures firsthand feedback and avoids the inherent bias of customer surveys, but also quickly develops customer and performance targets. No more waiting for monthly reports to determine strategic tactics.

Similarly, U.S. retailer Macy's uses big data to create customer-centric assortments. In the past, Macy's analyzed data points (sell-through rates, out-of-stocks, or price promotions) within the merchandising hierarchy. Using big data, however, the company can analyze these data points at the product or SKU level at a particular time and location and then generate thousands of scenarios to gauge the probability of selling a particular product at a certain time and place—ultimately optimizing assortments by location, time, and profitability.

... **45 percent** of companies have implemented a business-intelligence or big data initiative... **90 percent of Fortune 500 companies** will have at least one big data initiative underway this year.

For online businesses, big data helps customize offerings in real time. Amazon has been doing this for years by displaying products in a "Customers Who Bought This Item Also Bought" area. In offline advertising, big data assists in ad placement, helping determine which TV programs will deliver the biggest impact for different customer segments. Today's technology goes beyond pattern matching: Bluefin Labs uses advanced technology to "watch" videos and "listen" to audio to learn which customer segment prefers which ads in which TV-show genres. Defining marketing tactics at the micro-segment level means higher returns on advertising investments.

Big data was even a factor in the 2012 U.S. presidential election. By collating vast amounts of data on polling, fundraising, volunteers, and social media into a central database, the Democratic Party was able to assess individual voters' online activities and ascertain whether campaign tactics were producing offline results.[10] A target-sharing program was used to drive exceptionally high turnout, which proved key to winning the election.

Product innovation

Not all big data is new data. A wealth of information generated in years past sits unused—or at least not used effectively. A case in point: When ZestCash, a money lender to lower-income

[10] "Corporations Want Obama's Winning Formula," *Bloomberg Businessweek*, 21 November 2012

borrowers, began using cell-phone records as a proxy to ascertain credit risk, the company improved its margins by 20 percent.

New big data opportunities are cropping up worldwide to either collate and organize data from many sources or gather offline information. U.S.-based Acxiom has 1,500 data points on 500 million customers gathered from public records, surveys, and 50 trillion annual data transactions.[11] The flip side of this trend is the threat to businesses specializing in accessing hard-to-find data: What was once proprietary information—real estate data, for example—can now be gleaned from public and private sources.

Some data, once captured, can enable long-established companies to generate revenue and improve their products in new ways. GE is planning a new breed of "connected equipment," including its jet engines, CT scanners, and generators armed with sensors that will send terabytes of data over the Internet back to GE product engineers. The company plans to use that information to make its products more efficient, saving its customers billions of dollars annually and creating a new slice of business for GE.[12]

Not all big data is new data. A wealth of information generated in years past sits unused—or at least not used effectively.

Crowdsourcing and other social product innovation techniques are made possible because of big data. It is now possible to transform hundreds of millions of rich tweets, a cacophony of unstructured data, into insights on products and services that resonate with consumers. At the heart of this work is the ability to use sophisticated machine-based computational linguistics to conduct temporal sentiment analysis on a company's product portfolio and its customers. The resulting output informs product marketing and product innovation strategies.

Data, and the related analytics, is also becoming a standalone product. Technology and analytics firms have emerged to provide rich insights from data—for example, compiling and analyzing transaction data between retailers and their suppliers. Retailers that own this data, and more importantly the analytics, can use it to improve operations, offer additional services to customers, and even replace third-party organizations that currently provide these services, thus generating entirely new revenue streams.

Finally, imagine the potential big data brings to running experiments—taking a business problem or hypothesis and working with large data sets to model, integrate, analyze, and determine what works and what doesn't, refine the process, and repeat. Facebook runs thousands of experiments daily with one set of users seeing different features than others; Amazon offers different content and dynamic pricing to various customers and makes adjustments as appropriate.

Operations

Supply chain data offers a variety of information-rich interactions, including physical product movements captured through radio frequency identification (RFID) and microsensors. Airbus'

[11] "Mapping and Sharing the Consumer Genome," *The New York Times*, 16 June 2012
[12] "GE Tries to Make Its Machines Cool and Connected," *Bloomberg Businessweek*, 6 December 2012

Value Chain Visibility program, for example, uses RFID readers, motion sensors, and conveyors to monitor processes, materials, and asset movements in real time. The scope extends across suppliers, manufacturing sites, customers, and in-service partners and results in reduced inventory, improved productivity, and lower costs.

A retail chain uses detailed SKU inventory information to identify overstocks at one store that could be sold in another. Its previous rapid-reverse logistics approach could only identify the top 100 overstocked SKUs, while its big data approach takes the entire dataset—several terabytes of operational data—and creates a comprehensive model of SKUs across thousands of stores. The chain quickly moved hundreds of millions of dollars in store overstocks to various other stores and has since built a predictive model of distribution to limit overstocking altogether. The chain is now capturing pricing, promotion, and loyalty-card data to create even deeper insights into what, when, and why their customers buy.

Many times, the most effective business model for big data analytics is **a combination of decentralized services for business intelligence and standalone shared services for analytics.**

Bill Ruh, head of GE's newest R&D center in Silicon Valley, says he wants to "marry big data with some of GE's biggest businesses." In an interview with *Bloomberg Businessweek*, he says GE can use big data to help airlines that buy GE jet engines monitor their performance and anticipate maintenance needs and do the same for companies that lease commercial vehicles from GE Capital. "If I can begin to see that something is starting to deteriorate and get out there and fix it before it breaks, that's a foundational change," Ruh says. "In the end, what everybody wants is predictability."[13]

Such opportunities are reserved for those who understand that data is an asset to be capitalized upon. Allowing customer information to stagnate in a storage area is a wasted opportunity. And the value of all data is not so much the information captured but how the information is viewed from the customer's perspective. This means considering not only the impact on profits, but also how changing customer preferences affect the market place. For example, when considering the effect of a product promotion, collecting data on competitors' promotions, specifically on substitute products, can reveal how customer preferences have evolved.

How Should We Organize to Capture the Benefits of Big Data and Analytics?

Although big data processes large, diverse data sets to reveal complex relationships, humans are the crucial ingredient for interpreting and converting the data and relationships into insights.

[13] "GE's Billion-Dollar Bet on Big Data," *Bloomberg Businessweek*, 26 April 2012

These interpretations are increasingly being made by a new generation of analysts—the 21st century quants of the business world who are often referred to as data scientists. Data scientists are a new breed of people with advanced statistical and mathematical knowledge combined with business knowledge. They work closely with business managers to derive insights that lead to more strategic decisions. While business managers perform business-intelligence tasks as part of their day-to-day responsibilities, data scientists are almost exclusively focused on deriving business insights from data. For more information, see *The Science of Big Data* with Fermilab physicist, Rob Roser.[14]

With data scientists on the front lines of the transformation, organizational structures and analytical processes must be designed with these expert teams in mind.

Designing business models

As organizations evolve, so must their analytics capabilities, moving from basic and anticipatory to the more mature predictive analysis (see figure 5). Basic analytics provide a historic view of business performance: what happened, where it happened, how many times it happened. Anticipatory analytics identify unique drivers, root causes, and sensitivities. Predictive analytics perform business modeling and simulations and try to predict what will happen.

Figure 5
Analytic capabilities and processes need to evolve to fully realize big data's potential

Analytical process maturity	Techniques	
Basic • Provides a static, historical view of business performance • Draws on basic scorecards and static reports	Query and drill down	Where is the problem?
	Ad hoc reporting	How many? How often? Where?
	Standard reporting	What happened?
Anticipatory • Creates transparency into past and potential future performance drivers • Uses systems and processes to perform a range of descriptive analytics	Segmentation analysis	What are the unique drivers?
	Statistical analysis	Why is this happening?
	Sensitivity analysis	What if conditions change?
Predictive • Offers dynamic, forward-looking insights with quantified trade-offs • Requires high-quality integrated data and complex mathematical capabilities	Optimization	What is the best that can happen?
	Simulation	What would happen if ...?
	Predictive modeling	What could happen next?

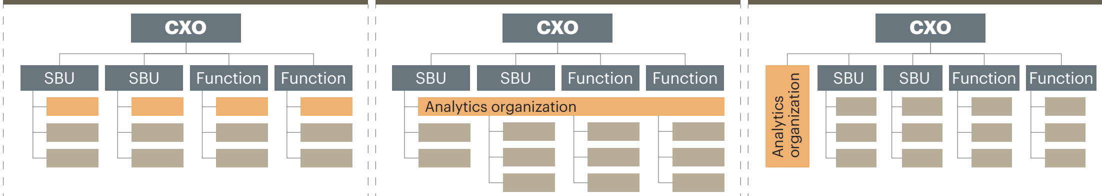
Source: A.T. Kearney analysis

Three operating models—decentralized services, embedded shared services, and standalone shared services—can be valuable tools for evolving from a data-and-information focus to a business insight-and-foresight focus. Each model has its pros and cons (see figure 6 on page 10).

[14] See "The Science of Big Data" at www.atkearney.com.

Figure 6

Three organization models for analytics groups



	Decentralized	Embedded	Standalone
Pros	• Analytical needs aligned to business and functions • Team members close to the business, issues, and customers • Analysts on hand to provide immediate tactical inputs	• Standardized processes and methodologies developed under one roof, with better coordination among departments • Functional expertise resident in existing shared services • Pre-existing shared service processes and structure	• Standardized processes and methodologies developed through a cross-enterprise mandate • Independent viewpoints shared organization-wide (unbiased by attachment to any business) • Direct reporting to CXO-level ensures faster, better organizational adoption • Best practices and insights shared organization-wide
Cons	• Incomplete cross-functional view and focus on local issues only • Limited strategic view of pressing issues organization-wide • Removed from top management and decision making • Duplication, redundancies, and inability to standardize or leverage scale	• Less transparent allocation of analytic resources • Focus on priorities or agenda of the function under which it resides • Difficult to build cross-functional alignment and consensus	• Perception that group lacks functional expertise or "practical know-how" of businesses or functions • Push back from business or function concerned about possible reallocation of resources

Organizational business or function · Analytic activities and capabilities · Business process

Source: A.T. Kearney analysis

In a **decentralized services** model, each business or function has its own analytics group, which enables and encourages rapid decision making and execution. But usually there is no dedicated role for strategic planning or best-practice sharing, which can result in duplicate resources and infrastructure. This model increases focus, but the lack of an enterprise view can undermine opportunity.

The **embedded shared-services** model is a centralized model that rolls under an existing function or business unit and serves the entire organization. It can speed execution and decision making, and its structure, support processes, and standards increase efficiency and IT expertise. But its role is performed more as customer rather than delivery partner.

The **standalone shared-services** model is similar to the embedded model but exists outside organizational entities or functions. It has direct executive-level reporting and elevates analytics to a vital core competency rather than an enabling capability.

Many times, the most effective business model for big data analytics is a combination of decentralized services for business intelligence and standalone shared services for analytics. With business intelligence embedded within business units, the entire organization quickly develops performance insights while helping drive the organizational culture toward hypothesis-based, data-driven decision making as opposed to the "gut feeling" approach.

A centralized analytics function ensures a view of the entire business and a predictive, rather than reactive, view of opportunities for the company to pursue.

Who works best with big data?

Although many of the roles integral to big data already exist in most organizations, their scope, visibility to executive management, and necessary technical skills could be more sharply defined. People who fill these roles have certain characteristics, such as a willingness to experiment and the ability to "data model" the future.

Procter & Gamble CEO Bob McDonald, for example, is convinced that "data modeling, simulation, and other digital tools are reshaping how we innovate," which has changed the skills his employees need. P&G created what McDonald calls "a baseline digital-skills inventory that's tailored to every level of advancement in the organization." Business intelligence (BI) managers have the skills to provide input into business-unit and corporate strategy, and business analysts know how to take a hypothesis-based approach to problem solving and adopt a cross-departmental view of data to identify opportunities firmwide. Similarly, data analysts work with unstructured data in developing complex statistical models, and data managers redesign IT architecture so that it provides a company-wide view, incorporating larger data manipulations (of unstructured data) and modeling activities into their decision-making process.[15]

Big data is not likely to exist as a standalone technology for long. As it becomes part of the broader analytics ecosystem, the distinction between big and traditional data is disappearing.

As mentioned earlier, data scientists are a hot commodity because they are proficient in analyzing large datasets to recognize trends, build predictive models, and identify business improvement opportunities. Unlike data managers and data analysts, data scientists usually have advanced degrees (preferably PhDs) in computer science, statistics, applied mathematics, or other relevant fields. And ideally, they have programming skills needed to procedurally manipulate and analyze large data sets. They are good communicators and often have a natural ability to translate business questions into meaningful analysis.

What Technology Investments Can Enable the Analytics Capabilities?

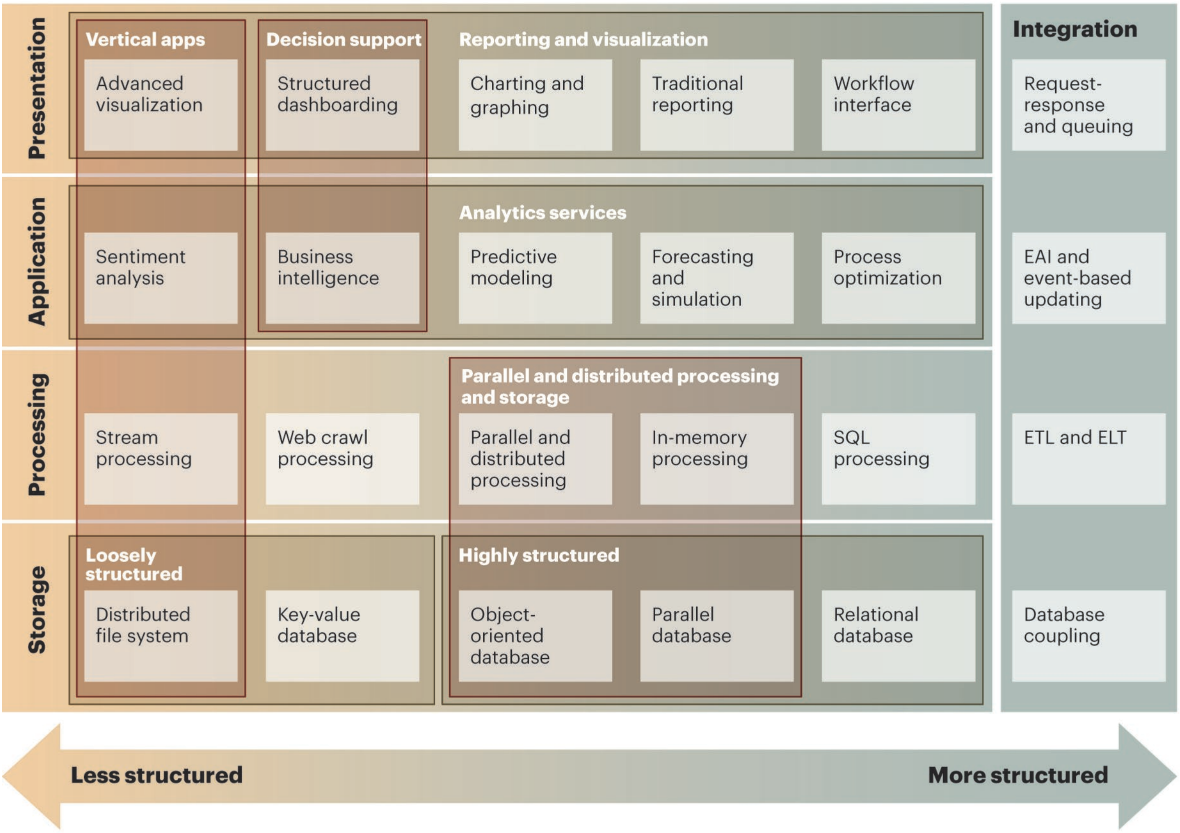
One of the challenges in delivering big data capabilities is transforming the IT architecture at an appropriate cost. This requires leveraging the old with the new. With all the options that exist, making the correct technology investments (to optimize the return on analytics) requires

[15] "Data Is Useless Without the Skills to Analyze It," *Harvard Business Review*, 13 September 2012

a complete understanding of the big data architecture needed to drive the organization's analytic vision. This return on analytics increases as the complexity of data changes—from structured to unstructured, from "clean" internal data to "messy" external data, and from one-way data flow to two-way data sharing with customers and partners. The architecture that enables analytics and big data must be able to efficiently handle all of these data types.

Figure 7
Big data architecture

Illustrative



Notes: EAI is enterprise application integration. SQL is structured query language. ETL is extract, transform, and load. ELT is extract, load, and transform.
Source: A.T. Kearney analysis

A big data technology architecture has five layers similar to those in a standard architectural stack: presentation, application, processing, storage, and integration (see figure 7). The presentation layer provides the primary interface for process workflow and management, dissemination of data reporting through dashboards, and data-visualization tools. Housing the business logic, the application layer is where the analysis, modeling, and business intelligence are executed. The processing and storage layers process and store large volumes of structured and unstructured data in real time or near real time. Finally, the integration layer interfaces with the various enterprise data sources throughout an organization and moves them in, out, and across the big data architecture. The layers, each of which contains various functions that act on data, work in tandem to produce the desired results. Each layer is home to specific data-centric jobs and functions. These are grouped to reflect the market segments for new big data products:

- **Vertical applications** or product suites (Hadoop, for example) provide data and processing while incorporating reporting and visualization tools.
- **Decision support** provides traditional BI dashboards and systems that are fortified by big data applications and visualization tools.
- **Reporting and visualization tools** represent the big data results in an easy-to-understand manner.
- **Analytics services** are the storage, processing, and applications that reinforce big data-enabled solutions, such as predictive modeling, forecasting, and simulation.
- **Parallel distributed processing and storage** enable massively parallel processing (MPP) and in-memory analytics for more-structured data.
- **Loosely structured storage** captures and stores less-structured data.
- **Highly structured storage** captures and stores traditional databases, including their parallel and distributed manifestations.

Technologies to know

The market is flooded with innovative products for managing, processing, and analyzing big data, including MPP, columnar databases, and NoSQL databases for managing structured data.[16] Notable vendors offering MPP solutions—Greenplum, Vertica, Aster Data, Netezza, and Splunk, for example—provide hardware and software combined into a single appliance. To process unstructured data, such as text messaging, social media, mobile technology, and embedded sensors, companies are turning to Hadoop, an open-source framework. The power of Hadoop lies in its ability to process hundreds of terabytes or petabytes (even zettabytes) of data.

Most companies opt for commercial Hadoop products by Greenplum, MapR, Cloudera, IBM, and others for the ease with which they can aggregate and mine disparate unstructured data sets. For example, using Facebook and Twitter data, it is now possible to gauge customer satisfaction in real time, and the resulting feeds can be combined with other data, such as geolocation, Web clickstream, and point-of-sale numbers to create more big data-enabled applications.

A new ecosystem of software applications has emerged to support these big data applications (see sidebar: Trends in Big Data Technology on page 15). They serve a variety of purposes from forecasting and simulation to predictive modeling and are supported by workflow engines, dashboards, batch reporting, and charting and graphing tools. Additionally, a host of visualization tools has arrived in the marketplace, including multi-node network diagrams, geo-spatial graphing, and multidimensional charts.

Numerous proofs of concept are now available. Here, the goal is to maintain separate but parallel environments, allowing unfettered testing in an environment that does not interfere with existing systems. There are many products worthy of evaluation, but Hadoop is on most test lists because of its central role in cost-efficient processing and its ability to handle less-structured data.

Big data is not likely to exist as a standalone technology for long. As it becomes part of the broader analytics ecosystem, the distinction between big and traditional data is disappearing—big data is becoming part of an overall advanced analytics strategy.

[16] "MapReduce and MPP: Two Sides of the Big Data Coin," *ZDNet*, 2 March 2012

How Do We Get Started on the Big Data Journey?

For every successful big data implementation, there is an equally successful change management program. This was the case in our recent work for a traditional big-box retailer. The company had not seen positive same-store sales for years, and the market was getting more competitive. A member of the executive team complained that "online retailers are eating our lunch." Poor economic conditions, changing consumer behaviors, new competitors, more channels, and more data were all having an impact. There was a strong push to move aggressively into e-commerce and online channels. The retailer had spent millions of dollars on one-off projects to fix the problems, but nothing was working. Several factors were turning the company toward competing on analytics, from competitors' investments and a sharp rise in structured and unstructured data to a need for more insightful data.

Figure 8
Getting started on the big data journey

1. Identify where big data can be a game changer

- What key business and functional capabilities are required?
- What IT capabilities are needed to support and grow the business?
- Where are the major gaps in capabilities to support the business?

2. Build future-state capability scenarios

- What are the options for future business capabilities and technologies?
- How do the options compare for capabilities, costs, risk, and flexibility?
- What functional, analytical, and technology decisions are needed to support these capabilities?

3. Define benefits and road map

- What is the investment's payback period?
- What is the implementation road map?
- What are the key milestones?
- What skills are needed? Where are the talent gaps?
- What are the risks?
- What is the third-party engagement strategy?

Source: A.T. Kearney analysis

Transforming analytical capabilities and big data platforms begins with a well thought-out, three-pronged approach (see figure 8):

Identify where big data can be a game changer. For our big-box retailer, new capabilities were needed if the business had any chance of pulling out of its current malaise and gaining a competitive advantage—the kind that would last despite hits from ever-changing, volatile markets and increased competition. The team engaged all areas of the business from merchandising, forecasting, and purchasing to distribution, allocation, and transportation to understand where analytics could improve results. Emphasis was placed on predictive analytics rather than reactive data access. So instead of answering why take-and-bake pizza sales are declining, the retailer focused on predicting sales decline and volume shifts in the take-and-bake pizza category over time and across geographic regions. In another example, the business wanted to move from reacting to safety issues to predicting them before they occur. The retailer planned to use social media data to "listen" for problems, which would not

Trends in Big Data Technology

Even as numerous standalone companies remain in the big data market, the market is changing daily as more data management players scoop up their vendors—EMC acquired Greenplum, IBM purchased Netezza, and HP bought Vertica. Partnerships are also forming to augment existing capabilities with unstructured-data processing, the most notable being Microsoft's teaming up with Hortonworks and Oracle's alliance with Cloudera.

Hardware companies are vertically integrating software solutions to meet short-term demand, but more significant integrations and larger investments will be required in the long term. Existing BI vendors are likely to consider a play for big data solutions that combine the best technologies for structured and unstructured data.

In semi-structured data, Hadoop is being deployed on commodity hardware at minimal cost, much like Apache and Linux a decade ago.[17] Hadoop, however, is not enterprise-ready and requires highly skilled professionals for implementation and support. Hadoop-based appliances are now emerging with Greenplum, IBM, and others working to provide standard products for semi- and unstructured data processing. For example, eBay's new search engine, Cassini, is based on big data technology and can support 97 million active buyers and sellers, 250 million queries per day, and 200 million items live in more than 50,000 categories.[18]

Eventually, enterprise-level information architectures will meet big data storage, processing, and application requirements, and existing storage and processing systems will likely accommodate big data. Applications will have far greater analytical capabilities. For example, future demand-forecasting systems will expand beyond factors such as promotions and price reductions to also evaluate competitor pricing, demand disruptions, stock-outs, and delivery notifications to enable real-time price adjustments.

only make the company more customer-centric but also provide a shield to future crises. The plan was to set up a business-information organization with four goals in mind:

- Deliver information tailored to meet specific needs across the organization.
- Build the skills needed to answer the competition, today and tomorrow.
- Create a collaborative analytical platform across the organization.
- Gain a consistent view of what is sold across channels and geographies.

Build future-state capability scenarios. The retailer was eager to develop scenarios for future capabilities, which were evaluated in terms of total costs, risks, and flexibility and determined within the context of the corporate culture. For example, is the business data driven, or is the company comfortable with hypothesis-based thinking and experimentation? Both are the essence of big data. We also identified trade-offs for each scenario, including comparison of capabilities, migration priorities, and timeline estimates. For example, which is most effective: a global data topology at headquarters or a local-regional-global combination? For a big data go-forward architecture, what are the trade-offs in using Hadoop versus Cassandra? These were assessed in the context of crucial opportunities, such as leveraging leading-edge technologies and providing a collaboration platform, integrating advanced analytics with existing and go-forward architecture, and building a scalable platform for multiple analytic types. This technology would enable five key capabilities and serve as the basis for future benefits:

[17] "Hadoop Could Save You Money Over a Traditional RDBMS," *Computerworld UK*, 10 January 2012
[18] "eBay Readies Next Generation Search Built with Hadoop and HBase," InfoQ, 13 November 2011

- Predict customers' purchasing and buying behaviors.
- Develop tailored pricing, space, and assortment at stores.
- Identify and leverage elasticities, affinities, and propensities used in pricing.
- Optimize global sourcing from multiple locations and business units.
- Devise models to suggest ways to reduce energy use and carbon emissions.

Define benefits and road map. Armed with these capabilities, the next questions revolved around resources. Did it make financial sense to assign internal resources? Or would it be more cost-effective to have external resources provide the big data analytics, at least initially? Naturally, the decision would depend on the company's capabilities. Technology needs were planned from two perspectives: data and architecture. A data plan was charted from acquisition to storage and then to presentation using a self-serve environment across both structured and unstructured data. Systems architecture, which may involve a Hadoop-based integration, was planned in light of the existing IT architecture, which was heavily reliant on relational data warehouses leveraging Teradata and Oracle platforms. A road map outlined a multi-million dollar investment plan that would deliver a positive payback in less than five years. The company is now positioned to realize four key benefits from its big data strategy:

- Deliver consistent information faster and with less expense.
- Summarize and distribute information more effectively across the business to better understand performance and opportunities to leverage the global organization.
- Develop repeatable BI and analytics instead of every group reinventing the wheel to answer similar questions.
- Generate value-creating insights yet to be discovered through advanced analytics.

Embracing the Creative Destruction of Business Models

Is big data the 21st century equivalent to the Industrial Revolution? We think so. Companies are increasingly experimenting with and implementing ways to capture big data's potential for both short- and long-term advantage. The crucial success factors are to first think of data as an asset—as the foundation upon which to build propositions and business models—and then to diligently build out the capabilities necessary to capitalize on big data's potential. And perhaps most importantly, embrace the creative destruction of today's business models.

Authors

Christian Hagen, partner, Chicago
christian.hagen@atkearney.com

Khalid Khan, vice president, Chicago

Marco Ciobo, principal, Melbourne
marcello.ciobo@atkearney.com

Jason Miller, principal, Chicago
jason.miller@atkearney.com

Dan Wall, principal, New York
dan.wall@atkearney.com

Hugo Evans, director, San Francisco
hugo.evans@atkearney.com

Ajay Yadav, consultant, Melbourne
ajay.yadav@atkearney.com

The authors wish to thank Matt Rahn, Bharath Thota, Anand Seshadri, Dushiant Kochhar, Jim Lee, Mohit Mohal, and Michael Wei for their valuable contributions to this paper.

A.T. Kearney is a global team of forward-thinking, collaborative partners that delivers immediate, meaningful results and long-term transformative advantage to clients. Since 1926, we have been trusted advisors on CEO-agenda issues to the world's leading organizations across all major industries and sectors. A.T. Kearney's offices are located in major business centers in 39 countries.

Americas	Atlanta Calgary Chicago Dallas	Detroit Houston Mexico City New York	San Francisco São Paulo Toronto Washington, D.C.
Europe	Amsterdam Berlin Brussels Bucharest Budapest Copenhagen Düsseldorf Frankfurt Helsinki	Istanbul Kiev Lisbon Ljubljana London Madrid Milan Moscow Munich	Oslo Paris Prague Rome Stockholm Stuttgart Vienna Warsaw Zurich
Asia Pacific	Bangkok Beijing Hong Kong Jakarta Kuala Lumpur	Melbourne Mumbai New Delhi Seoul Shanghai	Singapore Sydney Tokyo
Middle East and Africa	Abu Dhabi Dubai	Johannesburg Manama	Riyadh

For more information, permission to reprint or translate this work, and all other correspondence, please email: insight@atkearney.com.

The signature of our namesake and founder, Andrew Thomas Kearney, on the cover of this document represents our pledge to live the values he instilled in our firm and uphold his commitment to ensuring "essential rightness" in all that we do.