McKinsey Global Institute




May 2011

Big data: The next frontier for innovation, competition, and productivity

The McKinsey Global Institute

The McKinsey Global Institute (MGI), established in 1990, is McKinsey & Company's business and economics research arm.

MGI's mission is to help leaders in the commercial, public, and social sectors develop a deeper understanding of the evolution of the global economy and to provide a fact base that contributes to decision making on critical management and policy issues.

MGI research combines two disciplines: economics and management. Economists often have limited access to the practical problems facing senior managers, while senior managers often lack the time and incentive to look beyond their own industry to the larger issues of the global economy. By integrating these perspectives, MGI is able to gain insights into the microeconomic underpinnings of the long-term macroeconomic trends affecting business strategy and policy making. For nearly two decades, MGI has utilized this "micro-to-macro" approach in research covering more than 20 countries and 30 industry sectors.

MGI's current research agenda focuses on three broad areas: productivity, competitiveness, and growth; the evolution of global financial markets; and the economic impact of technology. Recent research has examined a program of reform to bolster growth and renewal in Europe and the United States through accelerated productivity growth; Africa's economic potential; debt and deleveraging and the end of cheap capital; the impact of multinational companies on the US economy; technology-enabled business trends; urbanization in India and China; and the competitiveness of sectors and industrial policy.

MGI is led by three McKinsey & Company directors: Richard Dobbs, James Manyika, and Charles Roxburgh. Susan Lund serves as MGI's director of research. MGI project teams are led by a group of senior fellows and include consultants from McKinsey's offices around the world. These teams draw on McKinsey's global network of industry and management experts and partners. In addition, MGI works with leading economists, including Nobel laureates, who act as advisers to MGI projects.

The partners of McKinsey & Company fund MGI's research, which is not commissioned by any business, government, or other institution.

Further information about MGI and copies of MGI's published reports can be found at www.mckinsey.com/mgi.

McKinsey Global Institute

May 2011

Big data: The next frontier for innovation, competition, and productivity

James Manyika
Michael Chui
Brad Brown
Jacques Bughin
Richard Dobbs
Charles Roxburgh
Angela Hung Byers

Big data—a growing torrent

$600 to buy a disk drive that can store all of the world's music

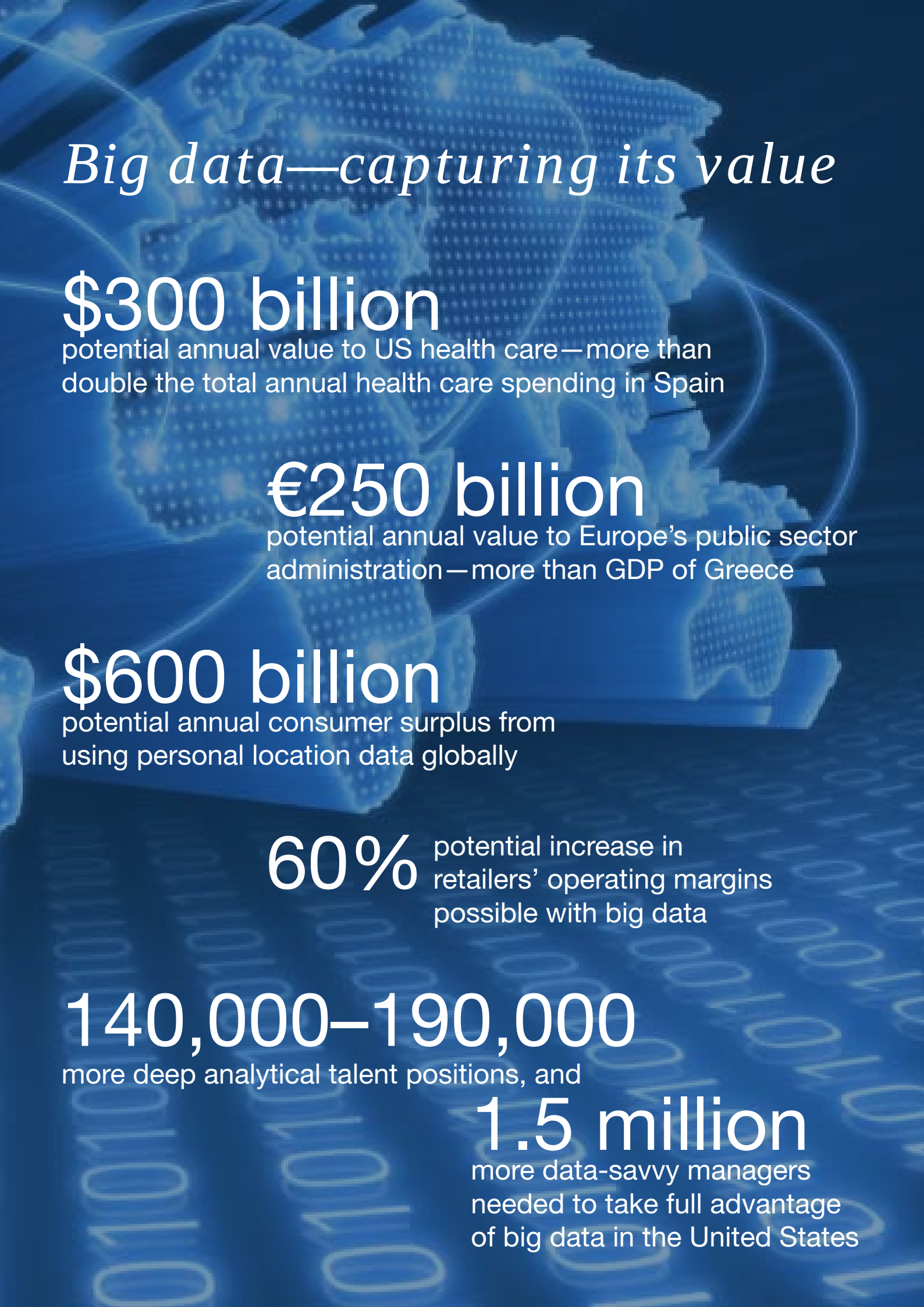
5 billion mobile phones in use in 2010

30 billion pieces of content shared on Facebook every month

40% projected growth in global data generated per year vs. **5%** growth in global IT spending

235 terabytes data collected by the US Library of Congress in April 2011

15 out of 17 sectors in the United States have more data stored per company than the US Library of Congress


Big data—capturing its value
$300 billion
potential annual value to US health care—more than double the total annual health care spending in Spain
€250 billion
potential annual value to Europe's public sector administration—more than GDP of Greece
$600 billion
potential annual consumer surplus from using personal location data globally
60%
potential increase in retailers' operating margins possible with big data
140,000–190,000
more deep analytical talent positions, and
1.5 million
more data-savvy managers needed to take full advantage of big data in the United States

Executive summary

Data have become a torrent flowing into every area of the global economy.[1] Companies churn out a burgeoning volume of transactional data, capturing trillions of bytes of information about their customers, suppliers, and operations. millions of networked sensors are being embedded in the physical world in devices such as mobile phones, smart energy meters, automobiles, and industrial machines that sense, create, and communicate data in the age of the Internet of Things.[2] Indeed, as companies and organizations go about their business and interact with individuals, they are generating a tremendous amount of digital "exhaust data," i.e., data that are created as a by-product of other activities. Social media sites, smartphones, and other consumer devices including PCs and laptops have allowed billions of individuals around the world to contribute to the amount of big data available. And the growing volume of multimedia content has played a major role in the exponential growth in the amount of big data (see Box 1, "What do we mean by 'big data'?"). Each second of high-definition video, for example, generates more than 2,000 times as many bytes as required to store a single page of text. In a digitized world, consumers going about their day—communicating, browsing, buying, sharing, searching—create their own enormous trails of data.

Box 1. What do we mean by "big data"?

"Big data" refers to datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze. This definition is intentionally subjective and incorporates a moving definition of how big a dataset needs to be in order to be considered big data—i.e., we don't define big data in terms of being larger than a certain number of terabytes (thousands of gigabytes). We assume that, as technology advances over time, the size of datasets that qualify as big data will also increase. Also note that the definition can vary by sector, depending on what kinds of software tools are commonly available and what sizes of datasets are common in a particular industry. With those caveats, big data in many sectors today will range from a few dozen terabytes to multiple petabytes (thousands of terabytes).

In itself, the sheer volume of data is a global phenomenon—but what does it mean? Many citizens around the world regard this collection of information with deep suspicion, seeing the data flood as nothing more than an intrusion of their privacy. But there is strong evidence that big data can play a significant economic role to the benefit not only of private commerce but also of national economies and their citizens. Our research finds that data can create significant value for the world economy, enhancing the productivity and competitiveness of companies and the

1 See "A special report on managing information: Data, data everywhere," *The Economist*, February 25, 2010; and special issue on "Dealing with data," *Science*, February 11, 2011.

2 "Internet of Things" refers to sensors and actuators embedded in physical objects, connected by networks to computers. See Michael Chui, Markus Löffler, and Roger Roberts, "The Internet of Things," *McKinsey Quarterly*, March 2010.

public sector and creating substantial economic surplus for consumers. For instance, if US health care could use big data creatively and effectively to drive efficiency and quality, we estimate that the potential value from data in the sector could be more than $300 billion in value every year, two-thirds of which would be in the form of reducing national health care expenditures by about 8 percent. In the private sector, we estimate, for example, that a retailer using big data to the full has the potential to increase its operating margin by more than 60 percent. In the developed economies of Europe, we estimate that government administration could save more than €100 billion ($149 billion) in operational efficiency improvements alone by using big data. This estimate does not include big data levers that could reduce fraud, errors, and tax gaps (i.e., the gap between potential and actual tax revenue).

Digital data is now everywhere—in every sector, in every economy, in every organization and user of digital technology. While this topic might once have concerned only a few data geeks, big data is now relevant for leaders across every sector, and consumers of products and services stand to benefit from its application. The ability to store, aggregate, and combine data and then use the results to perform deep analyses has become ever more accessible as trends such as Moore's Law in computing, its equivalent in digital storage, and cloud computing continue to lower costs and other technology barriers.[3] For less than $600, an individual can purchase a disk drive with the capacity to store all of the world's music.[4] The means to extract insight from data are also markedly improving as software available to apply increasingly sophisticated techniques combines with growing computing horsepower. Further, the ability to generate, communicate, share, and access data has been revolutionized by the increasing number of people, devices, and sensors that are now connected by digital networks. In 2010, more than 4 billion people, or 60 percent of the world's population, were using mobile phones, and about 12 percent of those people had smartphones, whose penetration is growing at more than 20 percent a year. More than 30 million networked sensor nodes are now present in the transportation, automotive, industrial, utilities, and retail sectors. The number of these sensors is increasing at a rate of more than 30 percent a year.

There are many ways that big data can be used to create value across sectors of the global economy. Indeed, our research suggests that we are on the cusp of a tremendous wave of innovation, productivity, and growth, as well as new modes of competition and value capture—all driven by big data as consumers, companies, and economic sectors exploit its potential. But why should this be the case now? Haven't data always been part of the impact of information and communication technology? Yes, but our research suggests that the scale and scope of changes that big data are bringing about are at an inflection point, set to expand greatly, as a series of technology trends accelerate and converge. We are already seeing visible changes in the economic landscape as a result of this convergence.

Many pioneering companies are already using big data to create value, and others need to explore how they can do the same if they are to compete. Governments, too, have a significant opportunity to boost their efficiency and the value for money

3 Moore's Law, first described by Intel cofounder Gordon Moore, states that the number of transistors that can be placed on an integrated circuit doubles approximately every two years. In other words, the amount of computing power that can be purchased for the same amount of money doubles about every two years. Cloud computing refers to the ability to access highly scalable computing resources through the Internet, often at lower prices than those required to install on one's own computers because the resources are shared across many users.

4 Kevin Kelly, Web 2.0 Expo and Conference, March 29, 2011. Video available at: www.web2expo.com/webexsf2011/public/schedule/proceedings.

they offer citizens at a time when public finances are constrained—and are likely to remain so due to aging populations in many countries around the world. Our research suggests that the public sector can boost its productivity significantly through the effective use of big data.

However, companies and other organizations and policy makers need to address considerable challenges if they are to capture the full potential of big data. A shortage of the analytical and managerial talent necessary to make the most of big data is a significant and pressing challenge and one that companies and policy makers can begin to address in the near term. The United States alone faces a shortage of 140,000 to 190,000 people with deep analytical skills as well as 1.5 million managers and analysts to analyze big data and make decisions based on their findings. The shortage of talent is just the beginning. Other challenges we explore in this report include the need to ensure that the right infrastructure is in place and that incentives and competition are in place to encourage continued innovation; that the economic benefits to users, organizations, and the economy are properly understood; and that safeguards are in place to address public concerns about big data.

This report seeks to understand the state of digital data, how different domains can use large datasets to create value, the potential value across stakeholders, and the implications for the leaders of private sector companies and public sector organizations, as well as for policy makers. We have supplemented our analysis of big data as a whole with a detailed examination of five domains (health care in the United States, the public sector in Europe, retail in the United States, and manufacturing and personal location data globally). This research by no means represents the final word on big data; instead, we see it as a beginning. We fully anticipate that this is a story that will continue to evolve as technologies and techniques using big data develop and data, their uses, and their economic benefits grow (alongside associated challenges and risks). For now, however, our research yields seven key insights:

1. DATA HAVE SWEPT INTO EVERY INDUSTRY AND BUSINESS FUNCTION AND ARE NOW AN IMPORTANT FACTOR OF PRODUCTION

Several research teams have studied the total amount of data generated, stored, and consumed in the world. Although the scope of their estimates and therefore their results vary, all point to exponential growth in the years ahead.[5] MGI estimates that enterprises globally stored more than 7 exabytes of new data on disk drives in 2010, while consumers stored more than 6 exabytes of new data on devices such as PCs and notebooks. One exabyte of data is the equivalent of more than 4,000 times the information stored in the US Library of Congress.[6] Indeed, we are generating so much

5 See Peter Lyman and Hal Varian, *How much information? 2003*, School of Information Management and Systems, University of California at Berkeley, 2003; papers from the IDC Digital Universe research project, sponsored by EMC, including *The expanding digital universe*, March 2007; *The diverse and exploding digital universe*, March 2008; *As the economy contracts, the digital universe expands*, May 2009, and *The digital universe decade—Are you ready?*, May 2010 (www.emc.com/leadership/programs/digital-universe.htm); two white papers from the University of California, San Diego, Global Information Industry Center: Roger Bohn and James Short, *How much information? 2009: Report on American consumers*, January 2010, and Roger Bohn, James Short, and Chaitanya Baru, *How much information? 2010: Report on enterprise server information*, January 2011; and Martin Hilbert and Priscila López, "The world's technological capacity to store, communicate, and compute information," *Science*, February 10, 2011.

6 According to the Library of Congress Web site, the US Library of Congress had 235 terabytes of storage in April 2011.

data today that it is physically impossible to store it all.[7] Health care providers, for instance, discard 90 percent of the data that they generate (e.g., almost all real-time video feeds created during surgery).

Big data has now reached every sector in the global economy. Like other essential factors of production such as hard assets and human capital, much of modern economic activity simply couldn't take place without it. We estimate that by 2009, nearly all sectors in the US economy had at least an average of 200 terabytes of stored data (twice the size of US retailer Wal-Mart's data warehouse in 1999) per company with more than 1,000 employees. Many sectors had more than 1 petabyte in mean stored data per company. In total, European organizations have about 70 percent of the storage capacity of the entire United States at almost 11 exabytes compared with more than 16 exabytes in 2010. Given that European economies are similar to each other in terms of their stage of development and thus their distribution of firms, we believe that the average company in most industries in Europe has enough capacity to store and manipulate big data. In contrast, the per capita data intensity in other regions is much lower. This suggests that, in the near term at least, the most potential to create value through the use of big data will be in the most developed economies. Looking ahead, however, there is huge potential to leverage big data in developing economies as long as the right conditions are in place. Consider, for instance, the fact that Asia is already the leading region for the generation of personal location data simply because so many mobile phones are in use there. More mobile phones—an estimated 800 million devices in 2010—are in use in China than in any other country. Further, some individual companies in developing regions could be far more advanced in their use of big data than averages might suggest. And some organizations will take advantage of the ability to store and process data remotely.

The possibilities of big data continue to evolve rapidly, driven by innovation in the underlying technologies, platforms, and analytic capabilities for handling data, as well as the evolution of behavior among its users as more and more individuals live digital lives.

2. BIG DATA CREATES VALUE IN SEVERAL WAYS

We have identified five broadly applicable ways to leverage big data that offer transformational potential to create value and have implications for how organizations will have to be designed, organized, and managed. For example, in a world in which large-scale experimentation is possible, how will corporate marketing functions and activities have to evolve? How will business processes change, and how will companies value and leverage their assets (particularly data assets)? Could a company's access to, and ability to analyze, data potentially confer more value than a brand? What existing business models are likely to be disrupted? For example, what happens to industries predicated on information asymmetry—e.g., various types of brokers—in a world of radical data transparency? How will incumbents tied to legacy business models and infrastructures compete with agile new attackers that are able to quickly process and take advantage of detailed consumer data that is rapidly becoming available, e.g., what they say in social media or what sensors report they are doing in the world? And what happens when surplus starts shifting from

7 For another comparison of data generation versus storage, see John F. Gantz, David Reinsel, Christopher Chute, Wolfgang Schlichting, John McArthur, Stephen Minton, Irida Xheneti, Anna Toncheva, and Alex Manfrediz, "The expanding digital universe," IDC white paper, sponsored by EMC, March 2007.

suppliers to customers, as they become empowered by their own access to data, e.g., comparisons of prices and quality across competitors?

Creating transparency

Simply making big data more easily accessible to relevant stakeholders in a timely manner can create tremendous value. In the public sector, for example, making relevant data more readily accessible across otherwise separated departments can sharply reduce search and processing time. In manufacturing, integrating data from R&D, engineering, and manufacturing units to enable concurrent engineering can significantly cut time to market and improve quality.

Enabling experimentation to discover needs, expose variability, and improve performance

As they create and store more transactional data in digital form, organizations can collect more accurate and detailed performance data (in real or near real time) on everything from product inventories to personnel sick days. IT enables organizations to instrument processes and then set up controlled experiments. Using data to analyze variability in performance—that which either occurs naturally or is generated by controlled experiments—and to understand its root causes can enable leaders to manage performance to higher levels.

Segmenting populations to customize actions

Big data allows organizations to create highly specific segmentations and to tailor products and services precisely to meet those needs. This approach is well known in marketing and risk management but can be revolutionary elsewhere—for example, in the public sector where an ethos of treating all citizens in the same way is commonplace. Even consumer goods and service companies that have used segmentation for many years are beginning to deploy ever more sophisticated big data techniques such as the real-time microsegmentation of customers to target promotions and advertising.

Replacing/supporting human decision making with automated algorithms

Sophisticated analytics can substantially improve decision making, minimize risks, and unearth valuable insights that would otherwise remain hidden. Such analytics have applications for organizations from tax agencies that can use automated risk engines to flag candidates for further examination to retailers that can use algorithms to optimize decision processes such as the automatic fine-tuning of inventories and pricing in response to real-time in-store and online sales. In some cases, decisions will not necessarily be automated but augmented by analyzing huge, entire datasets using big data techniques and technologies rather than just smaller samples that individuals with spreadsheets can handle and understand. Decision making may never be the same; some organizations are already making better decisions by analyzing entire datasets from customers, employees, or even sensors embedded in products.

Innovating new business models, products, and services

Big data enables companies to create new products and services, enhance existing ones, and invent entirely new business models. Manufacturers are using data obtained from the use of actual products to improve the development of the next generation of products and to create innovative after-sales service offerings. The emergence of real-time location data has created an entirely new set of location-

based services from navigation to pricing property and casualty insurance based on where, and how, people drive their cars.

3. USE OF BIG DATA WILL BECOME A KEY BASIS OF COMPETITION AND GROWTH FOR INDIVIDUAL FIRMS

The use of big data is becoming a key way for leading companies to outperform their peers. For example, we estimate that a retailer embracing big data has the potential to increase its operating margin by more than 60 percent. We have seen leading retailers such as the United Kingdom's Tesco use big data to capture market share from its local competitors, and many other examples abound in industries such as financial services and insurance. Across sectors, we expect to see value accruing to leading users of big data at the expense of laggards, a trend for which the emerging evidence is growing stronger.[8] Forward-thinking leaders can begin to aggressively build their organizations' big data capabilities. This effort will take time, but the impact of developing a superior capacity to take advantage of big data will confer enhanced competitive advantage over the long term and is therefore well worth the investment to create this capability. But the converse is also true. In a big data world, a competitor that fails to sufficiently develop its capabilities will be left behind.

Big data will also help to create new growth opportunities and entirely new categories of companies, such as those that aggregate and analyze industry data. Many of these will be companies that sit in the middle of large information flows where data about products and services, buyers and suppliers, and consumer preferences and intent can be captured and analyzed. Examples are likely to include companies that interface with large numbers of consumers buying a wide range of products and services, companies enabling global supply chains, companies that process millions of transactions, and those that provide platforms for consumer digital experiences. These will be the big-data-advantaged businesses. More businesses will find themselves with some kind of big data advantage than one might at first think. Many companies have access to valuable pools of data generated by their products and services. Networks will even connect physical products, enabling those products to report their own serial numbers, ship dates, number of times used, and so on.

Some of these opportunities will generate new sources of value; others will cause major shifts in value within industries. For example, medical clinical information providers, which aggregate data and perform the analyses necessary to improve health care efficiency, could compete in a market worth more than $10 billion by 2020. Early movers that secure access to the data necessary to create value are likely to reap the most benefit (see Box 2, "How do we measure the value of big data?"). From the standpoint of competitiveness and the potential capture of value, all companies need to take big data seriously. In most industries, established competitors and new entrants alike will leverage data-driven strategies to innovate, compete, and capture value. Indeed, we found early examples of such use of data in every sector we examined.

8 Erik Brynjolfsson, Lorin M. Hitt, and Heekyung Hellen Kim, Strength in numbers: *How does data-driven decisionmaking affect firm performance?*, April 22, 2011, available at SSRN (ssrn.com/abstract=1819486).

Box 2. How do we measure the value of big data?

When we set out to size the potential of big data to create value, we considered only those actions that essentially depend on the use of big data—i.e., actions where the use of big data is necessary (but usually not sufficient) to execute a particular lever. We did not include the value of levers that consist only of automation but do not involve big data (e.g., productivity increases from replacing bank tellers with ATMs). Note also that we include the gross value of levers that require the use of big data. We did not attempt to estimate big data's relative contribution to the value generated by a particular lever but rather estimated the total value created.

4. THE USE OF BIG DATA WILL UNDERPIN NEW WAVES OF PRODUCTIVITY GROWTH AND CONSUMER SURPLUS

Across the five domains we studied, we identified many big data levers that will, in our view, underpin substantial productivity growth (Exhibit 1). These opportunities have the potential to improve efficiency and effectiveness, enabling organizations both to do more with less and to produce higher-quality outputs, i.e., increase the value-added content of products and services.[9] For example, we found that companies can leverage data to design products that better match customer needs. Data can even be leveraged to improve products as they are used. An example is a mobile phone that has learned its owner's habits and preferences, that holds applications and data tailored to that particular user's needs, and that will therefore be more valuable than a new device that is not customized to a user's needs.[10] Capturing this potential requires innovation in operations and processes. Examples include augmenting decision making—from clinical practice to tax audits—with algorithms as well as making innovations in products and services, such as accelerating the development of new drugs by using advanced analytics and creating new, proactive after-sales maintenance service for automobiles through the use of networked sensors. Policy makers who understand that accelerating productivity within sectors is the key lever for increasing the standard of living in their economies as a whole need to ease the way for organizations to take advantage of big data levers that enhance productivity.

We also find a general pattern in which customers, consumers, and citizens capture a large amount of the economic surplus that big data enables—they are both direct and indirect beneficiaries of big-data-related innovation.[11] For example, the use of big data can enable improved health outcomes, higher-quality civic engagement with government, lower prices due to price transparency, and a better match between products and consumer needs. We expect this trend toward enhanced consumer surplus to continue and accelerate across all sectors as they deploy big data. Take the area of personal location data as illustration. In this area, the use of real-time traffic information to inform navigation will create a quantifiable consumer surplus through

9 Note that the effectiveness improvement is not captured in some of the productivity calculations because of a lack of precision in some metrics such as improved health outcomes or better matching the needs of consumers with goods in retail services. Thus, in many cases, our productivity estimates are likely to be conservative.

10 Hal Varian has described the ability of products to leverage data to improve with use as "product kaizen." See Hal Varian, *Computer mediated transactions*, 2010 Ely Lecture at the American Economics Association meeting, Atlanta, Georgia.

11 Professor Erik Brynjolfsson of the Massachusetts Institute of Technology has noted that the creation of large amounts of consumer surplus, not captured in traditional economic metrics such as GDP, is a characteristic of the deployment of IT.

savings on the time spent traveling and on fuel consumption. Mobile location-enabled applications will create surplus from consumers, too. In both cases, the surplus these innovations create is likely to far exceed the revenue generated by service providers. For consumers to benefit, policy makers will often need to push the deployment of big data innovations.

Exhibit 1

Big data can generate significant financial value across sectors

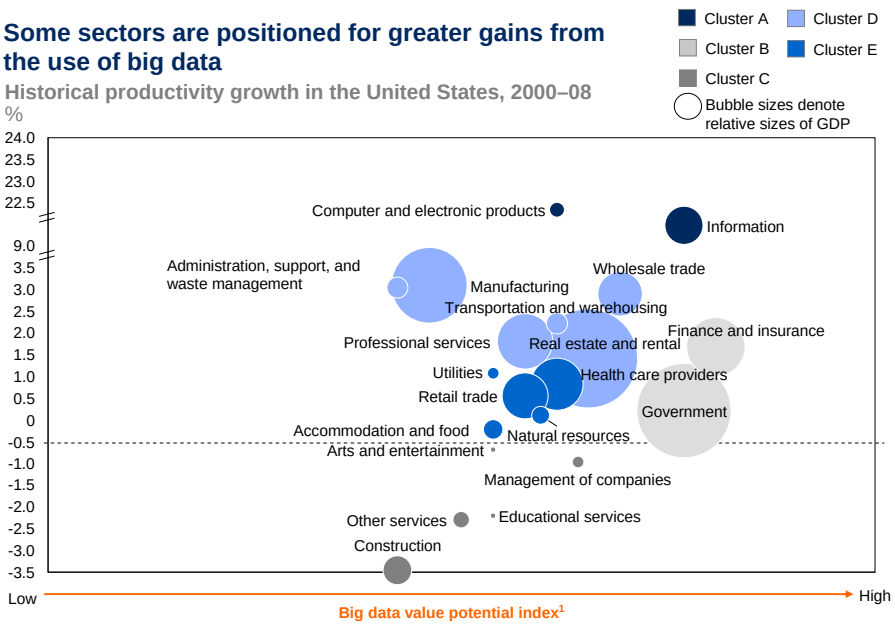
SOURCE: McKinsey Global Institute analysis

5. WHILE THE USE OF BIG DATA WILL MATTER ACROSS SECTORS, SOME SECTORS ARE POISED FOR GREATER GAINS

Illustrating differences among different sectors, if we compare the historical productivity of sectors in the United States with the potential of these sectors to capture value from big data (using an index that combines several quantitative metrics), we observe that patterns vary from sector to sector (Exhibit 2).[12]

12 The index consists of five metrics that are designed as proxies to indicate (1) the amount of data available for use and analysis; (2) variability in performance; (3) number of stakeholders (customers and suppliers) with which an organization deals on average; (4) transaction intensity; and (5) turbulence inherent in a sector. We believe that these are the characteristics that make a sector more or less likely to take advantage of the five transformative big data opportunities. See the appendix for further details.

Exhibit 2



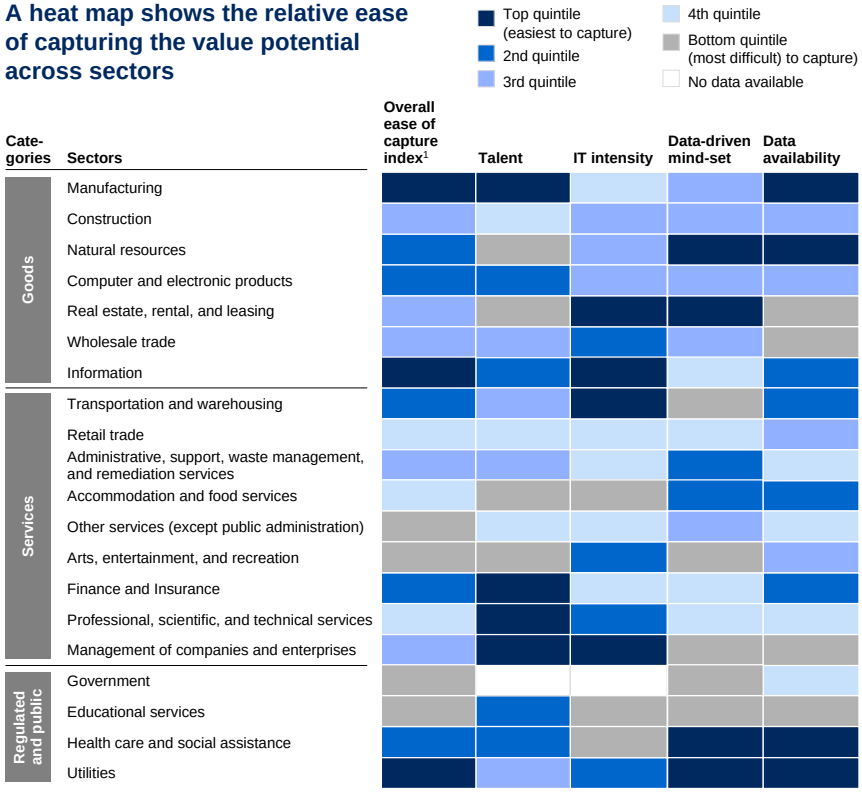
1 See appendix for detailed definitions and metrics used for value potential index.
SOURCE: US Bureau of Labor Statistics; McKinsey Global Institute analysis

Computer and electronic products and information sectors (Cluster A), traded globally, stand out as sectors that have already been experiencing very strong productivity growth and that are poised to gain substantially from the use of big data. Two services sectors (Cluster B)—finance and insurance and government—are positioned to benefit very strongly from big data as long as barriers to its use can be overcome. Several sectors (Cluster C) have experienced negative productivity growth, probably indicating that these sectors face strong systemic barriers to increasing productivity. Among the remaining sectors, we see that globally traded sectors (mostly Cluster D) tend to have experienced higher historical productivity growth, while local services (mainly Cluster E) have experienced lower growth.

While all sectors will have to overcome barriers to capture value from the use of big data, barriers are structurally higher for some than for others (Exhibit 3). For example, the public sector, including education, faces higher hurdles because of a lack of data-driven mind-set and available data. Capturing value in health care faces challenges given the relatively low IT investment performed so far. Sectors such as retail, manufacturing, and professional services may have relatively lower degrees of barriers to overcome for precisely the opposite reasons.

Exhibit 3

A heat map shows the relative ease of capturing the value potential across sectors



1 See appendix for detailed definitions and metrics used for each of the criteria.
SOURCE: McKinsey Global Institute analysis

6. THERE WILL BE A SHORTAGE OF TALENT NECESSARY FOR ORGANIZATIONS TO TAKE ADVANTAGE OF BIG DATA

A significant constraint on realizing value from big data will be a shortage of talent, particularly of people with deep expertise in statistics and machine learning, and the managers and analysts who know how to operate companies by using insights from big data.

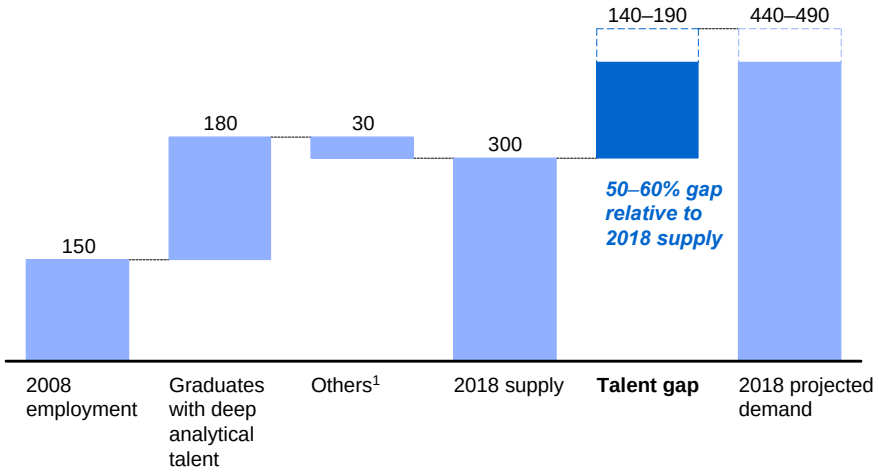
In the United States, we expect big data to rapidly become a key determinant of competition across sectors. But we project that demand for deep analytical positions in a big data world could exceed the supply being produced on current trends by 140,000 to 190,000 positions (Exhibit 4). Furthermore, this type of talent is difficult to produce, taking years of training in the case of someone with intrinsic mathematical abilities. Although our quantitative analysis uses the United States as illustration, we believe that the constraint on this type of talent will be global, with the caveat that some regions may be able to produce the supply that can fill talent gaps in other regions.

In addition, we project a need for 1.5 million additional managers and analysts in the United States who can ask the right questions and consume the results of the analysis of big data effectively. The United States—and other economies facing similar shortages—cannot fill this gap simply by changing graduate requirements and waiting for people to graduate with more skills or by importing talent (although these could be important actions to take). It will be necessary to retrain a significant amount of the talent in place; fortunately, this level of training does not require years of dedicated study.

Exhibit 4

Demand for deep analytical talent in the United States could be 50 to 60 percent greater than its projected supply by 2018

Supply and demand of deep analytical talent by 2018
Thousand people



1 Other supply drivers include attrition (-), immigration (+), and reemploying previously unemployed deep analytical talent (+).
SOURCE: US Bureau of Labor Statistics; US Census; Dun & Bradstreet; company interviews; McKinsey Global Institute analysis

7. SEVERAL ISSUES WILL HAVE TO BE ADDRESSED TO CAPTURE THE FULL POTENTIAL OF BIG DATA

Data policies. As an ever larger amount of data is digitized and travels across organizational boundaries, there is a set of policy issues that will become increasingly important, including, but not limited to, privacy, security, intellectual property, and liability. Clearly, privacy is an issue whose importance, particularly to consumers, is growing as the value of big data becomes more apparent. Personal data such as health and financial records are often those that can offer the most significant human benefits, such as helping to pinpoint the right medical treatment or the most appropriate financial product. However, consumers also view these categories of data as being the most sensitive. It is clear that individuals and the societies in which they live will have to grapple with trade-offs between privacy and utility.

Another closely related concern is data security, e.g., how to protect competitively sensitive data or other data that should be kept private. Recent examples have demonstrated that data breaches can expose not only personal consumer information and confidential corporate information but even national security secrets. With serious breaches on the rise, addressing data security through technological and policy tools will become essential.[13]

Big data's increasing economic importance also raises a number of legal issues, especially when coupled with the fact that data are fundamentally different from many other assets. Data can be copied perfectly and easily combined with other data. The same piece of data can be used simultaneously by more than one person. All of these are unique characteristics of data compared with physical assets. Questions about the intellectual property rights attached to data will have to be answered: Who "owns" a piece of data and what rights come attached with a dataset? What defines "fair use" of data? There are also questions related to liability: Who is responsible when an

13 Data privacy and security are being studied and debated at great length elsewhere, so we have not made these topics the focus of the research reported here.

inaccurate piece of data leads to negative consequences? Such types of legal issues will need clarification, probably over time, to capture the full potential of big data.

Technology and techniques. To capture value from big data, organizations will have to deploy new technologies (e.g., storage, computing, and analytical software) and techniques (i.e., new types of analyses). The range of technology challenges and the priorities set for tackling them will differ depending on the data maturity of the institution. Legacy systems and incompatible standards and formats too often prevent the integration of data and the more sophisticated analytics that create value from big data. New problems and growing computing power will spur the development of new analytical techniques. There is also a need for ongoing innovation in technologies and techniques that will help individuals and organizations to integrate, analyze, visualize, and consume the growing torrent of big data.

Organizational change and talent. Organizational leaders often lack the understanding of the value in big data as well as how to unlock this value. In competitive sectors this may prove to be an Achilles heel for some companies since their established competitors as well as new entrants are likely to leverage big data to compete against them. And, as we have discussed, many organizations do not have the talent in place to derive insights from big data. In addition, many organizations today do not structure workflows and incentives in ways that optimize the use of big data to make better decisions and take more informed action.

Access to data. To enable transformative opportunities, companies will increasingly need to integrate information from multiple data sources. In some cases, organizations will be able to purchase access to the data. In other cases, however, gaining access to third-party data is often not straightforward. The sources of third-party data might not have considered sharing it. Sometimes, economic incentives are not aligned to encourage stakeholders to share data. A stakeholder that holds a certain dataset might consider it to be the source of a key competitive advantage and thus would be reluctant to share it with other stakeholders. Other stakeholders must find ways to offer compelling value propositions to holders of valuable data.

Industry structure. Sectors with a relative lack of competitive intensity and performance transparency, along with industries where profit pools are highly concentrated, are likely to be slow to fully leverage the benefits of big data. For example, in the public sector, there tends to be a lack of competitive pressure that limits efficiency and productivity; as a result, the sector faces more difficult barriers than other sectors in the way of capturing the potential value from using big data. US health care is another example of how the structure of an industry impacts on how easy it will be to extract value from big data. This is a sector that not only has a lack of performance transparency into cost and quality but also an industry structure in which payors will gain (from fewer payouts for unnecessary treatment) from the use of clinical data. However, the gains accruing to payors will be at the expense of the providers (fewer medical activities to charge for) from whom the payors would have to obtain the clinical data. As these examples suggest, organization leaders and policy makers will have to consider how industry structures could evolve in a big data world if they are to determine how to optimize value creation at the level of individual firms, sectors, and economies as a whole.

□ □ □

The effective use of big data has the potential to transform economies, delivering a new wave of productivity growth and consumer surplus. Using big data will become a key basis of competition for existing companies, and will create new competitors who are able to attract employees that have the critical skills for a big data world. Leaders of organizations need to recognize the potential opportunity as well as the strategic threats that big data represent and should assess and then close any gap between their current IT capabilities and their data strategy and what is necessary to capture big data opportunities relevant to their enterprise. They will need to be creative and proactive in determining which pools of data they can combine to create value and how to gain access to those pools, as well as addressing security and privacy issues. On the topic of privacy and security, part of the task could include helping consumers to understand what benefits the use of big data offers, along with the risks. In parallel, companies need to recruit and retain deep analytical talent and retrain their analyst and management ranks to become more data savvy, establishing a culture that values and rewards the use of big data in decision making.

Policy makers need to recognize the potential of harnessing big data to unleash the next wave of growth in their economies. They need to provide the institutional framework to allow companies to easily create value out of data while protecting the privacy of citizens and providing data security. They also have a significant role to play in helping to mitigate the shortage of talent through education and immigration policy and putting in place technology enablers including infrastructure such as communication networks; accelerating research in selected areas including advanced analytics; and creating an intellectual property framework that encourages innovation. Creative solutions to align incentives may also be necessary, including, for instance, requirements to share certain data to promote the public welfare.

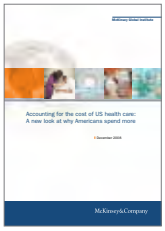
Relevant McKinsey Global Institute publications

August 2010
Clouds, big data, and smart assets: Ten tech-enabled business trends to watch

Advancing technologies and their swift adoption are upending traditional business models. Senior executives need to think strategically about how to prepare their organizations for the challenging new environment.

March 2010
The Internet of Things

More objects are becoming embedded with sensors and gaining the ability to communicate. The resulting new information networks promise to create new business models, improve business processes, and reduce costs and risks.

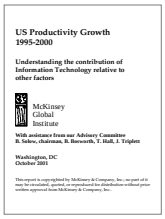


December 2008
Accounting for the cost of US health care: A new look at why Americans spend more

The United States spends $650 billion more on health care than expected, even when adjusting for the economy's relative wealth; MGI examines the underlying trends and key drivers of these higher costs. MGI finds that outpatient care, which includes same-day hospital and physician office visits, is by far the largest and fastest-growing part of the US health system. The next largest contributors are the cost of drugs, and administration and insurance.

October 2002
How IT enables productivity growth

Looking at three sectors in detail—retail banking, retail trade, and semiconductors—MGI finds that while IT enabled productivity gains in each sector, its impact was complex and varied. IT applications that had a high impact on productivity shared three characteristics: They were tailored to sector-specific business processes and linked to performance levers; they were deployed in a sequence that allowed companies to leverage their previous IT investments effectively; and they evolved in conjunction with managerial innovation.



October 2001
US productivity growth 1995–2000: Understanding the contributions of information technology relative to other factors

MGI's study of US productivity growth aimed to determine the causes of the sudden increase in the growth rate of labor productivity in the United States between 1995 and 2000. This increase was interpreted by key economists and policy makers as evidence of a "new economy," where IT applications would lead to faster US economic growth. The report lays out the nature and importance of the role played by six key sectors in the acceleration of productivity growth and its causes, with a focus on the role of IT.

www.mckinsey.com/mgi

eBook versions of selected MGI reports are available on MGI's Web site, on Amazon's Kindle bookstore, and on Apple's iBookstore.

Download and listen to MGI podcasts on iTunes or at www.mckinsey.com/mgi/publications/

McKinsey Global Institute
May 2011

www.mckinsey.com/mgi