1

Why Big Data?

There is an enormous amount of data. The increase in unfiltered data that has accumulated so rapidly includes an increase in needless data, which musdt be removed to allow more efficient and unbiased analyses. This requires an ability to extract correct and useful information from the data. Thus, by correctly distinguishing the "gems" from the "pebbles," Big Data analysis would assist an enterprise in obtaining a wider view when starting with a comparably narrow view. Because Big Data bases its significance in the expansion of thought, it is not about volume, velocity, or variety of data but rather about an alternative perspective and viewpoint with respect to the data. If you want to see a forest, you should not leave the forest you should climb to the top of a mountain. Likewise, to obtain meaningful insight from Big Data, we should attempt to broaden our perspective from a bird's eye view. The higher the altitude, the wider is the vision that can be obtained. To see the outside that was never observed from the inside, a different perspective is required to see the forest, and that is where Big Data steps in.

1.1 Big Data

There has been a significant influx of interest in Big Data. Gartner, one of the top marketing analysis institutions in the world, has selected Big Data as one of the top 10 strategic technologies [1] in both 2012 and 2013; in 2014, it selected Big Data and

---

Fundamentals of Big Data Network Analysis for Research and Industry, First Edition. Hyunjoung Lee and Il Sohn.

Actionable Analytics as the core strategy technology for smart governance [2]. Further, every January at Davos, global political and economic leaders gather at the World Economic Forum to discuss world issues, At the so-called Davos Forum 2012 [3], Big Data was again selected as one of the 10 technologies that have emerged as crucial for future developments. Although we are currently confronted by a financial crisis and partial recovery, along with issues related to climate change, energy, poverty, and security, the selection of Big Data seems to indicate that solutions to global issues require a broad range and amount of data, and the technology to effectively manage and extract useful data is expected to provide much-needed insight into resolving some of these potentially catastrophic global issues.

Of course, when we first encounter Big Data, we focus most of our attention on the word "Big" and become engrossed with the image of a giant being. In reality, however, Big Data is more closely associated with enormity and numberlessness. The term Big Data was defined and widely disseminated by Meta Group (now Gartner) analyst Doug Laney in 2001 to address issues and opportunities in the three dimensions of the rapid data expansion, including data volume, velocity of input/output data, and variety of data type [4]. The concept of Big Data attracting widespread interest in the 2000s can be correlated with the global proliferation of the Internet and the need to analyze the enormous amount data that it generates. The importance of analyzing massive data and converting them into useful information cannot be overstated. Next, a dimension dealing with "value" should be added to the existing three dimensions of data. If Big Data is large, expressed in real time similar to streaming, and includes unstructured data such as text, images, and videos, combining these different types of data and creating value are important. Thus, the amount of reserves is important, whereas the size of the mine is unimportant. The researcher does not need data; he or she needs information. Big Data addresses the size of the data; fundamentally, however, it is more important to analyze and produce meaningful data.

To be considered as Big Data, the data volume must be large in the data set. Although there is no specific size limit that defines Big Data, typically the data set would be a few terabytes for small data sets to as much as a few petabytes for large data sets. Table 1.1 indicates the current data sizes, with the prefixes of peta-, exa-, zetta-, yotta-, bronto-, and geop- used to express the amount of data [5]. If we were to express the amount of data in the books contained in the Library of Congress (in Washington, DC), the total would be about ~15 TB. Through 2012, the human race has accumulated a wealth of data totaling 1.27 ZB. Thus, 1 GpB would suggest an amount of data that is difficult to fathom and would describe an enormous amount of data that are created and distributed.

Another aspect of Big Data is the data velocity and the rate of data accumulation. Twenty years ago, it was expensive not only to install a high-speed data communication network but also to pay its monthly fees. However, now it is relatively easy to use wired and wireless network connections to transfer 1 Gbps (100 Mb/s is possible, at least in Korea) from the home, office, or even the street. Thus, the creation and distribution of data are occurring in the blink of an eye. Recently, natural disasters and various bulletins have first been reported not by the news but by microblogs such

Table 1.1 Data size.

Data	Size		Means				
Bit (b)	1 b	1	Binary digit (1 or 0)				
Byte (B)	8 b	2^3	English letter (1 character)				Basic data units
Kilobyte (kB)	1024 B	2^{10}	1 page				A sheet of paper with 1200 characters
Megabyte (MB)	1024 kB	2^{20}	873	Pages	4	Books	Single digital photo: 3 MB Single MP3 song: 4 MB
Gigabyte (GB)	1024 MB	2^{30}	894,784	Pages	4,473	Books	1–2 hours movie: 1–2 GB
			341	Digital pictures	256	MP3 audio files	
Terabyte (TB)	1024 GB	2^{40}	916,259,689	Pages	4,581,298	Books	Entire volume of books in the library of Congress: 15TB
			349,525	Digital pictures	262,144	MP3 audio files	
			1,613	CDs	233	DVDs	
			40	Blu-ray discs			

(*Continued*)

Table 1.1 (*Continued*)

Data	Size		Means				
Petabyte (PB)	1024 TB	2^{50}	938,249,922,368	Pages	4,691,249,611	Books	Amount of data Google processes
			357,913,941	Digital pictures	268,435,456	MP3 audio files	Every hour: 1 PB
			1,651,910	CDs	239,400	DVDs	
			41,943	Blu-ray discs			
Exabyte (EB)	1024 PB	2^{60}	960,767,920,505,705	Pages	4,803,839,602,528	Books	Amount of data contained in 100 million copies of a weekly magazine in the US
			366,503,875,925	Digital pictures	274,877,906,944	MP3 audio files	
			1,691,556,350	CDs	245,146,535	DVDs	
			42,949,672	Blu-ray discs			
Zettabyte (ZB)	1024 EB	2^{70}	983,826,350,597,842,752	Pages	4,919,131,752,989,213	Books	The amount of data existing until 2012: 1.27 ZB
			375,299,968,947,541	Digital pictures	281,474,976,710,656	MP3 audio files	
			1,732,153,702,834	CDs	251,030,052,003	DVDs	
			43,980,465,111	Blu-ray discs			

Yottabyte (YB)	1024 ZB	2^{80}	1,007,438,153,012,190,978,921	Pages	5,037,190,915,060,954,894	Books	It would take 11 trillion years to download 1 YB from a high-power broadband
			3843,307,168,202,282,325	Digital pictures	288,230,376,151,711,744	MP3 audio files	
			1,773,725,391,702,841	CDs	257,054,773,251,740	DVDs	
			45,035,996,273,704	Blu-ray discs			
Brontobyte (BB)	1024 YB	2^{90}	1,031,616,699,404,483,562, 415,936	Pages	5,158,083,497,022,417, 812,079	Books	Considering the size of the data that can be collected in real time sensor data of the IoT (internet of things)
			393,530,540,239,137,101,141	Digital pictures	295,147,905,179,352, 825,856	MP3 audio files	
			1,816,294,801,103,709,697	CDs	263,224,087,809,782,414	DVDs	
			46,116,860,184,273,879	Blu-ray discs			
Geopbyte (GpB)	1024 BB	2^{100}	1,056,375,500,190,191,167, 913,919,337	Pages	5,281,877,500,950,955,839, 569,596	Books	Largest data amount that can be fathomed
			402,975,273,204,876,391, 568,725	Digital pictures	302,231,454,903,657,293, 676,544	MP3 audio files	
			1,859,885,876,330,198,730,317	CDs	269,541,465,917,217, 192,562	DVDs	
			47,223,664,828,696,452,136	Blu-ray discs			

as Twitter. Moreover, smart meters in industrial plants, home appliances such as smart TVs and refrigerators, and driverless cars have become increasingly connected via the Internet, allowing the real-time data acquisition of enormous amounts of data, which will continue to accelerate.

Big Data should not be determined just in terms of the size and speed caused by the continuously accumulating variety of data variety. Previously, most of the data we had worked with were well formatted and easy to manage. In other words, the data were well ordered and in a particular form, resulting in structured data. For example, sales, inventory data, or defect ratios during processing are typical data that we observe. However, new types of data cannot be categorized in existing formats and are unstructured. Videos, music, images, location information, text, and so on are data that do not conform to the usual formats; they are unstructured data. These types of data have different sizes and content, which are difficult to organize but have been increasing significantly; thus, they require new processing methods to acquire meaningful data.

Considering Big Data's size, speed, and variety, a typical collection is approximately a few thousand terabytes, produced, distributed, and used at a rate ranging from a few seconds to hours, and this collected Big Data can be in the form of either a structured or an unstructured configuration, which makes it almost impossible to manage and analyze the data by using existing methods. In addition, although Big Data involves an enormous amount and rate with various forms of data, essentially what is most important is to identify meaningful data by analyzing Big Data. Therefore, Big Data includes sets of data that are difficult to manage and analyze using existing methodologies; it also includes human resources, organizations, and related technology to manage and analyze the compounding data that comprise Big Data. Through this process, the resulting creation is the value of Big Data analysis.

In today's connected society, there are infinite amounts of data. However, the existence of larger amounts of data does not necessarily mean that those data should be analyzed. If there are increasing amounts of data, there are also increasing amounts of useless and meaningless data, and special abilities are required to filter the useless raw data into meaningful data sets. Big Data places significance on expanding perspectives and thought and is related not only to the amount, speed, or size of data but also to a sophisticated viewpoint and forecasting. We do not need to leave the forest to see it; instead, we can move up the mountain. Similarly, to gain insight from Big Data, our perspective must be moved above the current perspective. Higher ground allows us to see farther and wider than we can from ground level. To view the outside, which cannot be accomplished from within the group, we must see from a different perspective; to see even more, we must use Big Data. Analysis of Big Data allows our field of vision to be significantly increased from existing ranges.

1.2 What Creates Big Data?

From the past to the present, we have accumulated approximately 1.27 ZB (until 2012), and it is estimated that by 2016, global Internet protocol (IP) traffic will reach approximately 1.3 ZB [6]. The reason for this enormous influx of data cannot be fully comprehended. One reason for this explosive amount of data production is the development of storage devices. With the development of characters, historical records that were once written on plant rinds, animal leather skins, tree fragments (such as bamboo poles), stone plates, clay tablets, and so on, were significantly enhanced with the invention of paper manufacturing and printing. The vast amount of human activities and knowledge that once disappeared without a trace began to be neatly accumulated by characters printed on paper. However, paper-stored information was larger in volume than in amount. With the coming of the twentieth century, analog storage devices such as photographic films, phonographic records, cassette tapes, and videotapes appeared, and information that was once contained on paper was now stored in smaller-volume films or tapes that could hold significant amounts of data. During the 1980s before the digital era arrived, humanity had accumulated approximately 2 620 000 TB of data, 90% or more of which were contained in films and tapes [7]. After 1990, the digital revolution occurred, in which characters, voices, pictures, and images were digitized and there was a dramatic increase in the ability to store data. The floppy disk was the first computer storage devices; later devices included the hard disk and the flash memory. Today we can save and observe tens of GB daily on our smart phone devices to retrieve books, pictures, music, and images on the go at any time. To date, if humanity's accumulated data were recorded on CDs stacked on top of one another, they would span a distance of six times the distance from Earth to the moon. Technological advancements have continued to lower the price of storage. In 1980, it would have cost $213 000 to store 1 GB of data on a hard-disk drive. In 2013, it cost $0.03 to store that same amount of data on a hard-disk drive [8]. This plummeting cost has had a significant impact on the sharp increase in data accumulation (Figure 1.1).

Another reason for the significant growth of data can be attributed to improved connectedness. In the 1960s, computers were rare and valuable, but with the advent of personal computers in the 1980s, their scarcity value dropped. Today, the availability of portable personal computers and devices has rapidly expanded, along with that of smart communication devices such as smart phones and digital pads connected to the Internet, rendering the scarcity value of computers obsolete. Today, many people's smart phones can definitively outperform some of the personal computers produced a few years ago, and the smart phones' improved performance allows users to control various types of equipment through them. In reality, as smart televisions and refrigerators become available in the marketplace, from automobiles to consumer appliances and various equipment, the connectedness of the computer and wireless communication devices is increasingly common and expanding. As a result, the concept of the IoT (Internet of Things), which describes the structural aspect of Internet connectivity technology that gathers data from sensors of objects, has been developed. According to Gartner, in 2009, there were 900 million units of objects that used IoT technology; by 2020, this number is expected to grow to 26 billion units. Cisco speculates that from 2013 to 2022, IoT technology will have an economic value of approximately $14 trillion.

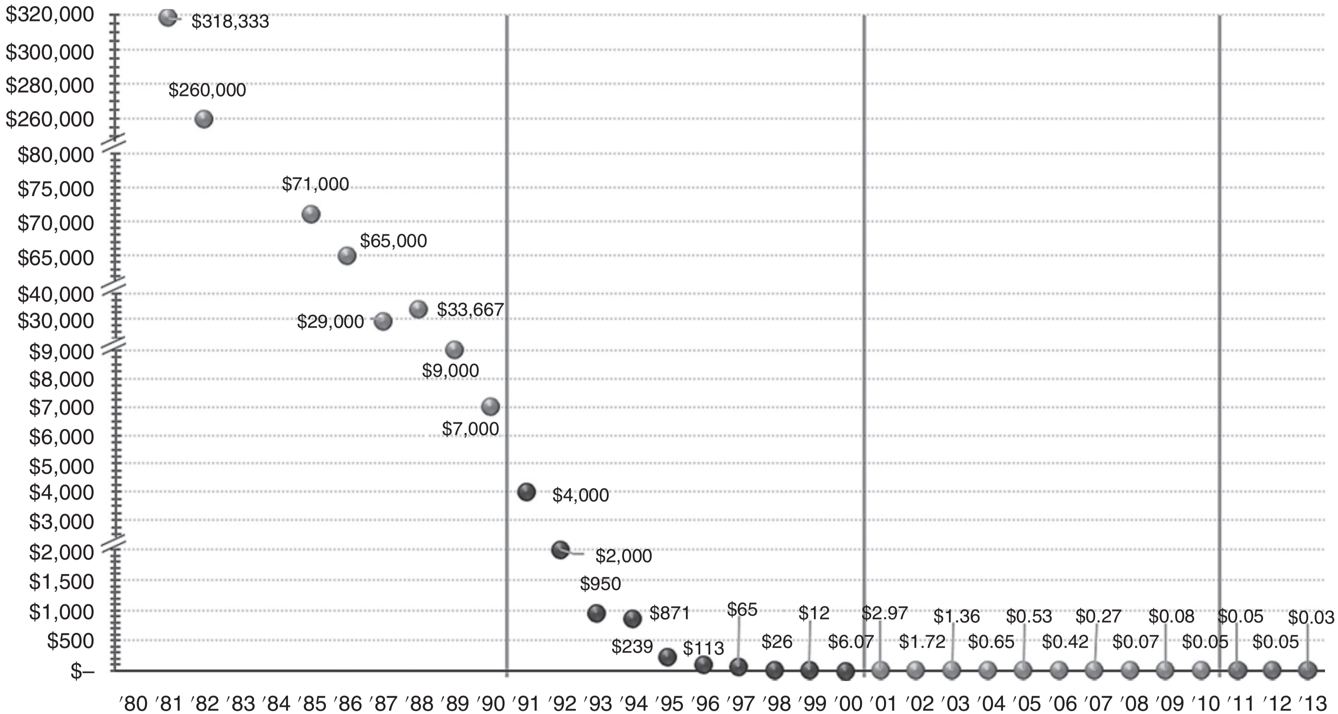

Figure 1.1 Hard-disk drive average cost per gigabytes (unit: US$).

The importance of Big Data cannot be driven solely by the exponential increase in data. Even now, various types of large data are significantly increasing and their speed is becoming faster. What is essential is the extraction of useful information from Big Data. To extract useful information, data management and analysis techniques are needed. Before the 1990s, the size of the digital image database was tens of thousands of images. However, today's online image-sharing sites such as Flicker, Picasa, and Pinterest have digital image databases whose size exceeds our imagination and can include more than tens of billions of images for each database. If image data processing technology did not evolve, even with the development of computer performance, it would not be able to manage the enormously increasing amount of data. Fortunately, the digital image processing technology of analyzing and indexing images has advanced at the pace of the increasing amount of data, and it has become possible to manage and search billions of images per second, allowing the management of vast amounts of data. Similarly, significant advancements in the analysis of various unstructured data and related technology have occurred. These technological developments allowed for recognition of the value of Big Data and are now a topic of interest to many.

Of course, advancements in the technological environment do not always attract the market's attention. Although the technology has long been available, it must match the market environment to become popular with the general public. Numerous types of technology were pioneers of their era and were expected to gain widespread popularity, but they silently disappeared without even entering the market after consumers realized that they were not needed. In that respect, the question of how to apply Big Data to corporate activity became extremely important.

1.3 How Do We Use Big Data?

How can we improve the business environment by using Big Data? Changing our perspective is the first step to improving the business environment. Analyzing Big Data and distinguishing the "gems" from the "pebbles" are not accomplished by evaluating what we know but rather by discovering what we do not know. Evaluation is verifying what we have known from the beginning, but discovery is identifying the right questions to ask through a creative and repetitive exploration process. Discovery is the creation of real value from Big Data. This provides a company with an idea to increase its corporate value [9].

The second step to improving the corporate business environment is to discover the various problems and possible solutions existing within various corporate activities. This involves changes in the corporate thought process and decision-making methods using Big Data. With Big Data, it is possible to find an important, hidden, previously unknown truth. Finding meaningful truth in Big Data that cannot be ascertained through cognitive human abilities is Big Data's problem-solving process. Forecasting is another area that is actively being pursued within Big Data. The future exists in the present. The future is not disconnected to the past or to the present. Things chosen in the past exist in the present, and things selected in the present will move to the future. Because the past, present, and future are not segmented, if the present can be ascertained, future doors can be opened. The analysis of the various past forms of Big Data can make it much easier to glimpse possible outcomes and future circumstances. From enormous amounts of data, people can discover new knowledge and events, and through this discovery, they can connect the present and the future. Thus, forecasting is enabled if data are correctly analyzed. A classic example is that of market basket analysis, which involves analyzing customers' purchasing behavior and merchandise related to cross- and up-sold items, the sale of which are optimized by selective placement. By analyzing customer purchasing behavior, it was found that people who purchase diapers also had a propensity to buy beer, and before a hurricane, both flashlights and sweet snacks are widely bought. These findings facilitate the optimal placement of products for consumers to purchase more effectively. By analyzing purchasing behavior, beer is placed next to diapers and strawberry cake is placed next to flashlights, resulting in increased sales. An important element of Big Data analysis that has increasingly been highlighted is the visualization of results. Visualization is a technique and method that allow analyzed data to be easily understood at a glance. A corporation's employees must clearly understand what has been discovered through the Big Data analysis; this need renders visualization extremely useful.

Data exist in every corporation, and the availability of data provides an opportunity to improve corporate productivity. By implementing different perspectives on existing data, the third step in improving the corporate business environment is to improve the practical use of existing data and information technology to improve work productivity. Improving productivity has often been used and could sound like a cliché. However, productivity enhancement is an important issue in the corporate environment. The methods of using Big Data for improved corporate productivity can be classified into two categories. First, by using sensor technology, data on the movement of materials can be identified and managed, resulting in lower labor, inventory, and logistics costs. Second, through Big Data analysis, unnecessary efforts in the work flow of the value chain can be minimized, allowing restructuring of the work process that optimizes work productivity. So far, the productivity enhancements in industry have been accomplished by substituting for the human work force with machines and computers; however, in the Big Data world, the raw materials, product, machines, and so on, are outfitted with various sensors and tags that enable real-time interactions and the accumulation of data to enhance productivity. Unlike in the past, the data have become more robust, quickly produced, and accumulated to allow a different dimension on increasing productivity. Through real-time data production and acquisition from various sensors, much more detailed and micro forecasting and control are possible, allowing precise management. For example, there is a bullwhip effect in SCM (supply chain management). Incorrectly assessing the supply of raw materials results in a greater error in production output forecasting at the next step of the supply chain, resulting in an increased gap between actual sales amounts. Therefore, through accurate forecasts and controls using real-time data, the gap between supply and demand can be minimized and can lower inventory and logistics costs, which can increase productivity. In reality, the next-generation power management system – labeled Smart Grid – which is the convergence of information technology (IT) with existing power grid systems, is installed with various sensors and meters in the power system that provide information on real-time power-consumption trends, allowing increased productivity and efficient management of power production. Through mutual communication, it is possible to perform detailed control and remote equipment examination of power flow. Automatic restoration during disruptions is also possible, allowing increased energy efficiency and providing optimal distribution of power according to consumption and increasing the ability to react during emergency situations. In the United States, raw materials and manufacturing service companies have moved from their previously used traditional methods of making repairs after equipment failures to a “state-based monitoring” method in which engineers actively monitor the state of equipment before failure. Waiting until failure creates an unexpected problem in productivity, and the production line is halted during a repair. However, by monitoring the temperature, vibration, production amount, and so on, and sending the data to the system for analysis, engineers can preemptively prepare for the problem and provide solutions that will not affect the production line to a great extent. Structured data are useful, but unstructured data can also assist in increasing productivity. The production and distribution of company documents can be analyzed and, by improving document use productivity, can be enhanced. Making the location of documents or images easier can minimize the cost and time to retrieve information, which is another application characteristic of Big Data.

The fourth step in improving the corporate management environment is to provide an objective viewpoint to a corporation’s decision makers. In attempting a new endeavor, several decisions must be made, during which many conflicts arise. If objective, observed data exist, the decision maker can overcome prejudice and weak points, thus resulting in rational consensus within the organization. Decision making will always be a manager’s essential duty, and the failure of many corporations is a result of a failure in decision making. The intuition of an experienced person is very important, but sometimes intuition can be an obstacle to making rational decisions. If intuition can be supplemented by a Big Data analysis, a more appropriate result can be deduced. For example, in the oil and gas industry, oil and gas fields are searched for and developed by installing large-scale sensor networks in the earth’s crust to identify with precision the possible locations and structure of oil fields. This results in lower development costs and improved oil transportation costs. To effectively use Big Data in the corporate decision-making process, an appropriate handling flow for data processing is necessary. Because Big Data itself does not assist in the decision-making process, an optimal decision-making process flow must exist within the corporation to effectively use Big Data. The final step to improve the corporate management environment is to create new value and to ensure that this created value is connected to a new business plan. The smart utilization of data allows the creation of a new business management paradigm. The ultimate goal of using Big Data is to identify what was missed in the past and to identify the hidden value of the dynamic client or create a new value and provide for the client. In the energy industry, networked sensors and an automated feedback mechanism are used to alter energy consumption patterns. By installing smart meters at power facilities in the industry, real-time energy consumption and costs can be verified, allowing the identification of peak times of energy use during work hours and providing optimal control of overloading in the power grid. Through these efforts, energy-intensive processes could be rearranged to a lower energy load time. Example is that of recent advancements in automotive navigation systems. These systems now surpass the mere combination of an electronic map and global positioning system (GPS) with the addition of accumulated information from road sensors that is transmitted to the navigation system, which is then transmitted to the central server and analyzed for traffic conditions, after which the customer is resupplied with the optimal and shortest distance to the destination of choice. Through these technological innovations and a Big Data analysis, the vehicle's fuel efficiency is improved and customers are provided with more efficient use of their time than being stuck in bumper-to-bumper traffic. It seems that in the near future, a navigation system could be developed to identify a driver's schedule and moods and to provide the optimal directions before the driver requests them. In the background of these evolutions, we cannot stress enough the importance of detailed data collection and analysis.

However, to maximize the effect of Big Data application, technology to effectively process Big Data is required. Simply, what problem should pass through a particular process, and who will manage that process to achieve the highest productivity. We typically concentrate not on the problem itself but on how quickly we can solve it. That said, the most important aspect is to identify the actual problem. To discern information about a problem, analysis of Big Data may be helpful. Generally, there are three steps for Big Data analysis. The first step is observation. Analyzing Big Data does not require collection of huge amounts of data. The data become meaningful only when they are analyzed. Therefore, through observation, we must initially decide what data to accumulate. Next, a quantifying process is needed. This is necessary to allow a systems-based observation that goes beyond simple data accumulation. In Big Data analysis, a broader and variable method of quantification is required. Finally, a process for deductive reasoning is needed. For example, the amount of data has significantly increased with the proliferation of smart phones. However, we also must use deductive reasoning about why people use smart phones.

Additionally, an increase in human resources for Big Data analysis is needed. According to McKinsey, in 2018 in the United States alone, there will be a deficiency of 14 000 to 19 000 data scientists who have the ability to deeply analyze Big Data and a deficiency of 1.5 million managers with the know-how to use Big Data to make informed decisions for greater productivity [10].

Today, there are infinite amounts of data. In the Big Data warehouse, there are clues for solving immediate problems. However, solutions do not involve Big Data sets exclusively. As previously mentioned, a greater amount of data will require an increasing amount of filtering. Therefore, by finding a gem within the enormous data size, we can effectively analyze Big Data to broaden corporate insight of the corporation at the next level.

1.4 Essential Issues Related to Big Data

Big Data's future is not necessarily uniformly bright. Big Data seems to show potential for use that excites many of its followers; however, there is much skepticism regarding the use of Big Data in such a wide range. Understanding these issues correctly and how we respond to them will determine whether Big Data will be a useful company asset or an enormous hassle.

The most widely disputed issues related to Big Data are personal privacy and protection issues. Because "Big Brother"[1] is often associated with Big Data and the use of Big Data has increased rapidly, we begin to wonder what boundaries of personal information and range of rights must be fulfilled before its use; these issues have found their way into the courtroom, with heated debate on both sides.

Another issue from a different perspective on Big Data is data abuse. With larger amounts of data, information can be fabricated to manipulate a person's true identity. The 2002 movie "Minority Report" depicted a cutting-edge crime prevention system, "Pre-crime," which enabled crime to be prevented using forecasting. The problem was that the probable event was considered to be an actual event, which a government agency then used as evidence to "suppress" so-called future perpetrators. This story demonstrates that total trust in Big Data can actually lower a company's productivity. We must realize and acknowledge that Big Data analysis provides a probability and likelihood that we can use to direct our actions.

Data acquisition and sharing are additional issues that have been contested. These issues involve intellectual property rights in open data, which seem to be addressed differently by the laws of different countries. Generally, open data used "as is" without processing is considered illegal, but once the data have been processed, creating new value, there seems to be a tendency to consider it legal. However, there remain gray areas in which different countries have different notions of how much processing is required to create new value and how to define the concept of new.

Because the United States considers freedom of speech and information distribution as more important than personal rights, there are many legal precedents that enable the distribution and use of personal data if they provide a public service, even if they infringe on personal rights. Therefore, if Big Data–related companies follow correct procedures for consent and provide anonymity, then data can be used without significant issues. Conversely, the European Union (EU) has a more conservative approach to human-rights protections and has actively promoted a law enforcing the "right to be forgotten [11]." In January 2012, the EU decreed that individuals had the right to remove personal information possessed by Internet companies without legal consent and finalized its data protection amendments, and in May 2014, the EU's highest court, the Court of Justice, approved the right to be forgotten. Many are well aware of the "Google Spain" trial, in which the Court of Justice ordered Google to delete a search result that consisted of a newspaper article on a 16-year-old incident

---

[1] Big Brother is a character of George Orwell's novel *1984*. Big Brother is taken a role that is monitoring the citizen through telescreens and deliver to the absolute. In the real world, Big Brother has become meaning of social control.

in which the plaintiff experienced foreclosure due to delinquent federal tax payments. The contents of the foreclosure auction were legal and open to the public. In the foreclosure auction process, all information must be provided pursuant to constitutional rules and regulations. If this procedure and information availability are not respected, the court official or government office may be considered to have unjustly executed the foreclosure, unjustly violated the rights of the debtor, or possibly acted to illegally accumulate wealth. Nevertheless, the Court of Justice judged that the search for information on past foreclosure information is unlawful and does not reasonably justify the acquisition of that data; it then ordered Google to remove the content. However, there exist many still-unresolved legal regulations addressing the complex and various types of Big Data being produced and created.

References

1 Gartner (2013) Gartner Identifies the Top 10 Strategic Technology Trend for 2013, http://gartner.com/newsroom/id/2209615 (accessed 22 May 2005).

2 Gartner (2014) Gartner Identifies the Top 10 Strategic Technology Trends for Smart Government, http://www.gartner.com/newsroom/id/2707617 (accessed 22 May 2005).

3 World Economic Forum (2012) Big Data, Big Impact: New Possibilities for International Development, http://www3.weforum.org/docs/WEF_TC_MFS_BigDataBigImpact_Briefing_2012.pdf (accessed 22 May 2005).

4 Laney, D. (2001) 3D Data Management: Controlling Data Volume, Velocity, and Variety, Meta Group(Gartner), http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf (accessed 22 May 2005).

5 The Economist (2010) All Too Much: Monstrous Amounts of Data, The Economist (Feb. 25 2010), http://www.economist.com/node/15557421 (accessed 22 May 2005).

6 Cisco (2014) The Zettabyte Era: Trends and Analysis, 2014 June 10, http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/VNI_Hyperconnectivity_WP.pdf (accessed 22 May 2005).

7 Hilbert, M. and López, P. (2011) The world's technological capacity to store, communicate, and compute information, *Science*, **332**(6025), 60–65.

8 Komorowski, M. (2014) A History of Storage Cost, http://www.mkomo.com/cost-per-gigabyte-update, reproduced (accessed 22 May 2015).

9 Hahm, Y.K. and Chae, S.B. (2014) *Big Data, Changing the Business*, Seoul: SERI.

10 McKinesy Global Institute (2011) Big Data: The Next Frontier for Innovation, Competition, and Productivity, http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation (accessed 22 May 2015).

11 Right to be forgotten is a concept that an individual has the right to request that his or her personal data to be removed from accessibility via a search engine European Commission (2014) Factsheet on the "Right to be Forgotten" Ruling, 2014 May 2013, http://ec.europa.eu/justice/data-protection (accessed 22 May 2015).