

Article

Sustainable Development Plan for Korea through Expansion of Green IT: Policy Issues for the Effective Utilization of Big Data

Hyun Baek ¹ and Sun-Kyoung Park ^{2,*}

¹ IBM-Korea, Dogok 2-Dong 467-12, Gangnam-Gu, Seoul 135-856, Korea;
E-Mail: hyjimbo@gmail.com

² Hanyang Cyber University, Wangsimri-Ro 222, Sungdong-gu, Seoul 133-791, Korea

* Author to whom correspondence should be addressed; E-Mail: sun.kyoung.park@gmail.com;
Tel./Fax: +82-2-2290-0432.

Academic Editor: Marc A. Rosen

Received: 8 December 2014 / Accepted: 19 January 2015 / Published: 27 January 2015

Abstract: The South Korean government is providing full support for green IT as one of the growth engines of Korea. The purpose of this study is to derive policy issues needed for the sustainable development of Korea through utilizing Big Data by applying green IT. The analysis is done using a Delphi technique. Results show that the establishment of computing platforms that can easily share data and generate value is prioritized for the effective use of Big Data from the environment. In addition, the government-led publication of genetic information and electronic medical records for research purposes has been derived as an important policy issue for the use of bio-Big Data. Besides, a guideline concerning the standardization of machine to machine and Internet of Things communication and data security is needed to effectively use Big Data from machines/things. Moreover, a review of legislation related to the utilization of Big Data from digital media has been derived as an important policy issue. The results of this study propose the direction in which the Korean government should move for green growth through effective utilization of Big Data. The results can be also useful resources for establishing relevant policies for various countries that are accelerating sustainable development.

Keywords: green IT; Big Data; Delphi technique; green growth; sustainable development

1. Introduction

With such advancements in the development of information communication technology and the increase of data storage space, large amounts of data unparalleled in past decades are being created. If information technology is effectively applied to the analysis of Big Data, which contains countless information, it can contribute to the sustainable growth of society as a whole and of humanity. Such IT that can contribute to the increase of environmental sustainability across society is called green IT. Green IT is a compound word from green, meaning eco-friendly, and information technology (IT). At first, the meaning of green IT was recognized as eco-friendly improvement to the issues of IT itself, such as the energy efficiency of data centers or resolving environmental issues through waste recycling of IT equipment [1–5]. This is called “Green for IT” or “Green IT 1.0”. In addition to the contribution to resolving environmental issues through the greening of such IT itself, it is being recognized as a way for IT to contribute to the resolution of environmental issues, which is called “Green by IT” or “Green IT 2.0” [6]. The scope of Green IT 2.0 has been expanded to the utilization of IT in improving environmental sustainability beyond the IT sector, meaning the overall operation of enterprises and society (Table 1). For example, “eco-friendly supply chain management” was originally considered as the management issue. However, recently, through the development of green IT, the use of radio-frequency identification (RFID) in the parts of electronic devices has improved the tracking of the waste of electrical and electronic equipment (WEEE). The recycling of WEEE is essential for supply chain management, due to the international regulations represented in derivatives of WEEE and RoHS (restriction of hazardous substances in electrical and electronic equipment).

Table 1. Application area of Green IT 1.0 and Green IT 2.0 [6–9].

Categories	Applications
Green IT 1.0	“Green datacenters” “IT asset disposal and recycling services” “IT energy measurement, localized cooling, managed printing services, storage capacity optimization”, <i>etc.</i>
Green IT 2.0	“Intelligent transportation system” “Smart grid” “Teleconference, remote working, paperless working” “Eco-friendly supply chain management” “Building energy management systems”, <i>etc.</i>

When green IT is divided into two types, Green IT 1.0 and Green IT 2.0, we can see that each of the categories encompasses a wide range of areas, as shown in Table 1. Studies that classify the application areas of green IT differently according to the area of utilization have also been conducted (Table 2) [6,10]. The first category of “green IT capabilities from the (natural) resource-based view” is an area related to the production, use and disposal of IT products, which can said to belong to Green IT 1.0 (Table 2) [6]. On the other hand, the second category corresponds to Green IT 2.0, because it is associated with the reduction of greenhouse gas emissions through such methods as video-conference or web-based business services based on IT. The third category is difficult to clearly classify as either Green IT 1.0 or 2.0, because it can be interpreted as image improvement concerning

IT itself. Hilpert *et al.* (2013) largely divided the “impacts of green information systems (IS) on eco-sustainability” into three categories (Table 2) [10]; where the first category was considered to correspond to Green IT 1.0 and the second and third categories to be included in Green IT 2.0 (Table 2). In this way, the categories according to the green IT area of utilization can also be divided into Green IT 1.0 and Green IT 2.0, but each category can be said to be defined more specifically in terms of its utilization purpose.

Table 2. Classification according to the application area of green IT. IS, information system.

Application Area of Green IT	Classification
Green IT capabilities of the (natural) resource-based view [6]	(1) “To reduce the environmental impact of IT through operation and life cycles”
	(2) “To reduce the environmental impact of operations using IT through ICT-enabled business processes”
	(3) “To enhance IT reputation through green IT performance and reporting”
Impacts of green IS on eco-sustainability [10]	(1) “Reducing the negative environmental impacts of IS itself”
	(2) “Reducing the negative environmental impacts of other economic sectors, such as industry, logistics, households or agriculture, by means of IS”
	(3) “Developing innovative IS-enhanced products and processes, which profoundly change the ways of living and result in the reorganization of business practices and consumption towards being more sustainable, thus cause behavioral changes towards greater eco-sustainability”

Green IT 2.0, IT that contributes to improving environmental sustainability through the application of IT in non-IT areas, can be applied in many areas besides Big Data, but this study focuses more on value creation through Big Data analysis. Big Data, as the word implies, is a large volume of data. Big Data is often characterized by the 3 Vs (volume, variety and velocity) or 4 Vs (veracity or value added to the 3 Vs). Volume refers to the size of Big Data, which determines whether the data is considered Big Data or not. Variety is the type to which Big Data belongs. The types of Big Data include structured (e.g., transactional data), unstructured (e.g., natural language, sound or image data), semi-structured (e.g., log data in a webpage) and hybrid data. Velocity refers to the urgency of processing to meet the demands. Besides the 3 Vs, veracity or value, the quality of data is often considered as a very important element that characterizes Big Data.

We are now able to access countless data due to the development of information communication technology along with the increase in storage space. Because various characteristics of data are included in Big Data, we may be able to shed light on changes in social phenomena and the discovery of natural laws that we could not have known before, if we effectively use Big Data [11]. However, because of the particular characteristics of Big Data, such as volume, variety and velocity, it was impossible to capture, store, manage and analyze it using existing technology [12]. However, such limitations can be resolved through the development of technology, and the analysis of massive data has become possible with the development of recent information technology [7,13].

According to the results of a survey by the IBM Institute for Business Value and Saïd Business School at the University of Oxford, which targeted 1144 people from 95 countries concerning businesses and IT experts, as well as the results of interviews with other experts, most organizations are currently in the starting stages in the Big Data sector, but 28% of the respondents have answered that

they are utilizing Big Data already. The survey showed that “customer-centric outcomes” accounted for the largest portion, 49%, concerning the utilization of Big Data, and the subsequent order was “operational optimization” (18%), “risk/financial management” (15%), “new business model” (14%) and “employee collaboration” (4%) [14]. In this way, Big Data is currently being largely used in areas related to enterprises’ customer management, but Big Data in the green IT area is getting a lot of attention as an important technology that can greatly contribute to the sustainable growth of society as a whole, such as energy reduction, carbon emission reduction, disaster management and disease prediction.

In 2008, “low carbon, green growth” was proclaimed as a national vision in South Korea. The national strategy for green growth pursues three main objectives. The first objective is to effectively deal with climate change and to attain energy independence, and the second one is to create new engines for economic growth. The third one is to improve the overall quality of life and to contribute to the international community. Thus, the meaning of “green” not only includes environmental sustainability, but also encompasses improving the quality of life. Since 2009, South Korea has set green IT as a national strategy and is actively promoting it on a national level with the goal of becoming a leading nation in green IT [15,16]. In 2013, the Korean government declared the “creative economy” as a new national vision, and information and communication technology (ICT) was considered as one of the important tools to effectively utilize Big Data to implement this new vision.

In the area of “Green for IT”, which can said to be Green IT 1.0, financial support is provided by the government concerning core issues, such as the development/export strategy of the world’s best green IT products, acceleration of green IT services and establishment of a safe network that is 10-times faster. In the area of “Green by IT”, which can said to be Green IT 2.0, the same support is given surrounding core issues, such as the transition to a low carbon business environment through IT, realization of the IT-based green life revolution, greening of the IT-converged manufacturing industry, the transition to smart green transportation and logistics systems, the establishment of an intelligent power network infrastructure, the establishment of intelligent real-time environmental monitoring systems and the establishment of an early disaster response system. However, there are no active studies on what type of basic support is needed for these policies by the Korean government to effectively promote IT-related issues.

This study particularly derived policy issues that should become the priority in improving the sustainability of society through effective analysis of Big Data in the application area of green IT. The process of reaching a conclusion via a conventional conference has the disadvantage of resulting in decisions that lack subjectivity, due to influences from the surrounding environment [17]. In order to compensate for this effect in this study, individual opinions from groups comprised of targeted experts have been investigated through a survey. In order to derive priority issues among the contents of the survey, the Delphi technique was used in accordance with the survey results. Studies such as this that use the Delphi technique have the advantage of revealing the opinions of each individual more specifically than the conventional conference method. In light of this fact, this study is significant because it is the first research to uncover the data that derive the policy issues that should be first addressed for the development of green IT in Korea through the Delphi technique. Policy issues for the effective utilization of IT in Big Data analysis in Korea, which is being supported at the government level with the goal of becoming a leading nation in green IT, can be used as basic data for establishing

the direction of relevant policies in various developing countries that are making transitions from a typical economy to sustainable development in terms of growth.

2. Materials and Methods

2.1. Delphi Technique

The purpose of this study is to derive policy issues that can contribute to Korea's sustainable development through effective analysis of Big Data. In this paper, the Delphi technique was used to derive these policies [18]. This can be viewed as a process of obtaining the most reliable consensus through repetition of the expert groups' opinions, feedback, survey, *etc.* [17]. The Delphi technique was originally developed with the purpose of evaluating the expected effect in the event of a large-scale nuclear attack in the U. S. during the 1950s. This method was introduced at the time because the traditional conference method for deriving a conclusion was deemed disadvantageous, because individual's subjective judgment can be blurred by the influence of surrounding opinions. Currently, the Delphi technique is largely used when there is a lack of standardized data or as a method for objectifying opinions from individual experts. However, it is also used to prioritize or sequence opinions when there are no reference data or when predicting an uncertain future situation. In addition, the technique is widely used as a supplementary material for studies in diverse fields [19].

In order to introduce the Delphi technique in studies, there must be great care in the selection of the survey subjects. It is necessary to comprise a group of experts or of those who have extensive experience in the relevant occupations concerning the survey questions and not those who simply have knowledge of the survey contents [20–23]. For this study, 30 people who are experts in green IT and Big Data have been carefully selected as the survey subjects. However, the limitations of this study include that the opinions from the limited number of experts could not represent 100% of the current status of the issues related to Big Data in Korea. Among the survey subjects, 17 people were employees of different industries in green IT and the environment sector and 13 people were experts affiliated with universities and private/governmental research institutes (Table 3). If the number of selected survey subjects is too small when conducting the Delphi analysis, a critical disadvantage is that the results cannot be generalized. Although the appropriate number of survey subjects varies depending on the survey content, 15–20 people are commonly selected, and under 50 survey subjects are selected in most cases [24].

The Delphi analysis method converges opinions over several iterations. The first round of Delphi analysis starts with an individual open-ended questionnaire. The surveyor gathers and organizes individual opinions, and the opinions of survey respondents are gathered again. The respondents rank the answers of various other respondents, including their own, in the second round. Then, in the third round, the results of the second round are provided to the respondents, and ranking of each question is requested again from the respondents. In this way, the Delphi technique converges on individual opinions through a repetitive process. There is no set number of iterations for consensus, and the iteration continuous until there is consensus, but three iterations are sufficient in many cases [25–28]. Consensus was also reached through three iterations in this study.

Table 3. Information of the survey subjects.

Division and Organization	Age: 30s	Age: 40s	Age: over 50	Total
Experts from Industry				
Environmental Department	2	3	1	6
Information and Communication Technology Department	1	9	1	11
Experts from Universities, Private/Governmental Research Institutes				
Private/Governmental Research Institutes	7	1	2	10
School of Business Administration (University)		1		1
School of Economics (University)			1	1
Department of Environmental Engineering (University)		1		1
Total				30

2.2. Classification of Big Data for the Effective Utilization of Green IT

In order to specifically derive policy issues for the sustainable development of Korea through effective utilization of Big Data, Big Data needs to be classified in detail. In this paper, based on the original source of Big Data, Big Data were classified as being derived from nature/the environment, lifeforms, machines/things and people. The above four classifications are able to encompass large portions of Big Data sources. However, areas of complex Big Data that occur in more than two categories were considered to be beyond the scope of this study. The types of data included in each field of Big Data are described in detail below.

2.2.1. Big Data Arising from Nature/Environment (Meteorological, Atmospheric and Environmental Pollution Big Data)

In this paper, meteorological data, such as temperature, precipitation, humidity, wind direction, wind speed and solar irradiance, as well as environmental pollution data, such as atmospheric greenhouse gases and contaminant concentrations in water/atmosphere/soil, are included in the Big Data from nature/the environment. As observation in dense areas in terms of time/space is now possible with the advancement of observation technology, the quantity of the accumulated data has increased greatly. Furthermore, the development of computer hardware and software has led to the development of a computational model that predicts air pollution and weather conditions by dividing a region into grids. Thus, it has become possible to divide a model region into finer grids, which, in turn, lead to huge amounts of predicted data [29–32]. Rather than using the entire, huge amount data, it is more effective to conduct analysis by selecting data that conform to the purpose of the study, and IT is essential in dealing with environmental data that are increasing in such a way [33]. In addition, effective analysis of natural/environmental Big Data related to sea level rise, coastal erosion, floods, *etc.*, is also very useful in disaster prevention [2,34,35]. Not only is IT effectively used in the analysis of such environmental data, but it is also closely associated with Geographic Information Systems, which are used for the comprehensive analysis of each region's land use, geological constraints, generation quantities, access roads, existing electrical grids, *etc.*, which are used for selecting the optimum location of renewable energy generation facilities [36].

2.2.2. Big Data Arising from Lifeforms (Plant/Animal Genes and Disease Big Data)

In this paper, plant/animal genes and disease Big Data are included in the Big Data from lifeforms. Big Data in this field are actively being used in the medical, agricultural and food sector [37–44]. For example, in order to analyze the correlation between coronary arterial plaque burden and stenosis and the gene expression score (GES), help from IT is essential for the analysis of the huge amount of data created from the gene samples of 610 patients [45]. As the analysis of such a vast amount of genetic information is now possible through IT development, it is being applied to the treatment of diseases, among other things, and the development of IT is contributing to advancement in the medical sector.

Genetic information is not only making a contribution to the medical field, but also to the increase of food production through the improvement of seeds for crops. The useful application of this information is implemented in the agriculture and food sector, such as for the extraction of useful materials or the production of bio-energy from crops. Of course, such developments of genetic engineering have advantages, but concerns of further influences are the topic of ongoing discussions [46]. Such gene and disease-related Big Data can be used more effectively along with the natural/environmental Big Data mentioned earlier. For example, the analysis of the effect of the change in the concentration of greenhouses gases in the atmosphere on crop conditions is possible through simultaneous analysis of Big Data in these various areas.

2.2.3. Big Data Arising from Machines/Things (M2M and IoT Big Data)

In this paper, automatically collected data from machine to machine (M2M) or Internet of Things (IoT) sensors are included in the scope of Big Data from machines/things. M2M refers to giving machines communication functions in order for them to collect and communicate information on their own. Examples of such M2M are car navigation systems, smart meters, unmanned security systems, *etc.* IoT refers to the communication among nearby machines with minimal human intervention. As described above, M2M and IoT contain similar meaning. Therefore, M2M and IoT will be comprehensively expressed as M2M/IoT in this paper. Based on the development and increased propagation of wireless networks, M2M/IoT is being applied in diverse ways, from people's daily lives to industrial fields. These M2M/IoTs are applied in many areas, such as building management systems, intelligent transportation systems, smart supply chain management and smart grids [47–50].

A smart building that automatically communicates and controls energy usage, such as lighting and heating/cooling, through sensors attached to the building can be said to be an example of a building management system that utilizes M2M/IoT. In addition, areas that make useful application of M2M/IoT are intelligent transportation networks, which contribute to providing the optimum driving course for drivers by providing real-time traffic information, and smart supply chain management, which has automated supply network management by attaching radio frequency identification (RFID) to products. In addition, the smart grid, known as a two-way electrical grid, has been recently regarded as the most important application area of green IT, because it stabilizes the power supply and reduces energy waste through the convergence of IT in the overall power infrastructure [51,52]. The existing power infrastructure has a unilateral supply structure, where the provider basically supplies electric energy to the consumers. In this case, if the accurate prediction of the change in power demand is not

made, the electricity generated by the power plant can result in an excess or a shortage. Unlike the existing power grids, the smart grid establishes a two-way communication that enables real-time data exchange between the power supplier and the consumer, where power consumption is transmitted in real time between the meter and electrical appliances and between the meter and the power supplier. Through such two-way information in real time, power suppliers can predict power demand more accurately.

2.2.4. Big Data Arising from People (Digital Media Big Data)

In this paper, sharable records and traces of people left behind from the use of digital media are included in the scope of Big Data from people. Specifically, this includes social network services, blogs, search words in search engines and user location information from smart phone usage. With the recent sharp increase in such digital media, data quantity also has increased greatly. The three types of Big Data mentioned earlier, that is meteorological/environmental, genetic/disease and M2M/IoT data, have the characteristic of having a standardized form. On the other hand, data generated from digital media have the characteristic of being mostly atypical. The reason is that the natural language that people use to express and communicate is used for most of the data appearing in digital media. In this way, human language, which does not have a standardized form, is used for expression in blogs or for search words in search engines, and these kinds of data are vast, but also provide important clues for understanding social phenomena.

Due to the active utilization of digital media via mobile devices, the importance of “text mining”, which is a new academic field for discovering and analyzing meaningful phenomenon in natural language, is becoming more pronounced. Text mining can be applied in many ways; for example, it was also used to track down criminals and terrorists by analyzing past criminal records and applying the results to similar types of crimes [53]. In addition, the frequencies of search words in search engines were used to predict outbreak trends of diseases. For instance, from the fact that certain search words are closely related to hospital visits, such as due to the common cold, the frequency of search words was used to accurately predict the trend of flu outbreaks [54]. Furthermore, by analyzing atypical customer comments that appeared on web bulletin boards, such as product reviews, the information was used to find problems or ways to improve the products. Furthermore, analyses of text related to patent subjects are sometimes used to predict research trends [55,56].

Data from the location of digital devices by human owners can also be effectively used to determine the movement of a population in a city. NTT DoCoMo in Japan utilized the network system of mobile phones to construct “mobile spatial statistics” [57]. “Mobile spatial statistics” is a sort of demographic information created through the use of the current location, address, gender, age, *etc.*, of a mobile phone user. This information can be used as basic data in improving services by estimating daily fluctuations of the population per region and identifying the status of urban activities and the supply/demand of public transportation. Furthermore, this information was used to determine densely populated areas and the number of people unable to return home during large-scale earthquakes in order to effectively allocate aid, such as water, food and shelters, in each region, as well as to establish responsive measures. Thus, the analysis results of digital media and devices not only enable the prediction of flu outbreaks and crime, but can also be used as a method for effectively coping with emergency and disaster situations.

3. Results and Discussion

The purpose of this study is to derive policy issues that can improve the sustainability of society through effective analysis of Big Data by utilizing green IT. Recently, the Korean government has been active in supporting research related to Big Data. The Ministry of Science, ICT and Future Planning opened the Big Data Center in 2014. As the first government-led Big Data center, small- and medium-scale companies, universities and research institutes can easily access the facilities needed for Big Data analysis, such as workstation servers and a large amount of data storage at the center.

In order to derive such policy issues, individual opinions from groups comprised of experts in relevant fields have been investigated through a survey. The progress and analysis of the survey were conducted based on the Delphi analysis method, and the 30 participating respondents were guaranteed anonymity and separation. The first survey was conducted as an open survey, where respondents were free to write down issues or preceding issues concerning green IT vitalization based on the newly-proposed four Big Data, so that a wide range of expert opinions can be accepted. After the answers to the first survey went through a simple organization, such as deleting duplicate answers and unifying terminologies, the results of the first survey were organized into units of similar groups, which resulted in 6–8 items of expert opinions for each of the four categories. Then, in the second survey, the respondents had to choose the level of importance of all items from a scale from one to five (higher points for issues with a higher level of importance).

Finally, for the third survey, the results of the second survey, which included points assigned from 1–5 by 30 experts, were summarized and organized, where each respondent was provided with statistical data, like the average value of each item as a reference, so that the respondents could reflect on the survey once more. Additionally, the third survey did not use the method of having the respondents assign points to all of the items, but rather, asked them to select three items that they felt were most important. The three selected items were not given any points, but it was only specified whether they were selected or not. Two weeks were provided to answer each of the first, second and third survey [58]. It has been reported that study results may display the problem of becoming biased when respondents do not participate in all of the iterations of a survey [28]. In this study, 100% of the first 30 survey respondents participated in the second survey, and 90%, 27 respondents, participated in the third survey.

According to the results of the expert survey that was conducted over three phases, the findings per green IT category system are summarized as shown in Tables 3–6. The first column of each table contains the organized answers of the first open survey concerning important issues or preceding issues. The second column is the result of the second survey, which assigned points from 1–5 on all items, and the third column is the result of the final survey, which had the respondents select three items that they felt were the most important among the categories. When a particular item is selected by all of the respondents, the result of the third survey for the item is 100%, and the result is 0% if no respondents selected the item. Here, since the survey required the selection of three items by each respondent, the results of the third survey per item are added, and the total becomes 300%. Each item with a higher result from the third survey was selected as an issue with a higher priority. The result was organized, centering on the opinions of the entire survey participants, and a comparison analysis

was conducted for each opinion from the respondents, who were experts from universities and private/governmental research institutes and those currently working in the industry.

3.1. Big Data Arising from Nature/Environment (Meteorological, Atmospheric and Environmental Pollution Data)

Several opinions were proposed concerning matters that should be implemented as a priority in order to contribute to the improvement of environmental sustainability through effective utilization of Big Data from nature/the environment (Table 4). From the various opinions that have been pieced together, it can be summarized that vitalization of research through the publication of data at the government level is required in order to effectively utilize Big Data arising from nature/the environment. A detailed look at the opinions derived from the Delphi technique is provided in Table 4.

Opinions derived from the first survey results can be largely grouped into eight categories. The category that had the highest priority from the third survey results was the integration and sharing of pollution data from the private/government sector. This is the opinion that the first step for effectively utilizing environment-related Big Data is to integrate pollution data held by government and private sectors in order to create a computing platform where all citizens can easily access the data and share the generated values. For this purpose, the development of an information publication guideline should, of course, be implemented in parallel.

Second is to establish an environmental pollution forecast and warning system. Concerning ozone and PM_{2.5} among air pollutants, an alert or warning is issued in Korea when the observed concentration exceeds the standard amount, and the proper action is announced to the public in the case of such events. Such an alert system can be expanded to and publicized for water pollution, marine radiation levels, *etc.*, in order to reduce the impact to the public. Furthermore, through the analysis of data on such alerts, it will be possible to predict the movement of pollution sources and the negative consequences to aquatic ecosystems and water sources.

Third is the monitoring of environmental pollution near industrial complexes. Although the capacity for Big Data analysis is increasing with the development of IT, proper data need to be provided to effectively utilize this ability. In Korea, there have been many voices of concern regarding the impact of exhaust fumes discharged by nearby plants, including nuclear power plant facilities and various other power facilities, but there are not enough study results that can support the claims. If medical information of residents nearby industrial complexes is analyzed in parallel with environmental pollution data, this will not only contribute to the promotion of healthcare and increased health for nearby residents, but it can also be utilized as the base data for implementing industrial complex construction and other similar projects elsewhere in the future.

Fourth is the sharing of renewable energy generation data. Renewable energy displays a big difference in power generation depending on the location. Such renewable energy is the source of next generation energy, and Korea is also making strong efforts to expand its renewable energy. If the efficiency of renewable energy generation in each region at each time can be provided, it can be used to select the location of future renewable energy generation facilities and to predict generation capacity. An interesting fact is that, while experts affiliated with universities and private/governmental research institutes emphasized the importance of sharing renewable energy generation data, the experts

currently working in the industry placed more importance on the monitoring of environmental pollution around industrial complexes.

Fifth is the sharing of meteorological data subdivided into terms of time and space. Through the analysis of real-time meteorological data, it is possible to prevent meteorological disasters and other disasters, such as floods, coastal erosion and sea level rise. Sixth is the sharing of recycled waste data. Recycled waste can be used as a renewable energy source.

Seventh is the long-term plant and animal population and crop monitoring. When such data are used, it is possible to predict global climate change, as well as the influences on the ecosystem resulting from meteorological phenomena, such as El Niño and La Niña. Finally, eighth is the sharing of greenhouse gas emission data subdivided into terms of time and space. Due to global climate change, many countries around the world, including Korea, are diligently working toward reducing greenhouse gas emissions. Through the analysis of such high resolution greenhouse gas emission data, it is possible to analyze an effective method for reducing greenhouse gases.

Table 4. Policy issues for utilization of Big Data arising from nature/the environment.

	Results of First Survey	Results of Second Survey *	Results of Third Survey ***
		[All] (Industry/Others **)	[All] (Industry/Others **)
1	Integration and sharing of environmental pollution data from private/government sector	[4.4 ± 0.86] (4.4 ± 0.86/4.4 ± 0.86)	[89%] (93%/83%)
2	Establishment of environmental pollution alert system	[3.9 ± 0.91] (4.2 ± 1.00/3.8 ± 0.80)	[63%] (67%/58%)
3	Monitoring of environmental pollution near industrial complexes	[3.8 ± 0.82] (3.8 ± 0.75/3.8 ± 0.93)	[59%] (67%/50%)
4	Sharing of renewable energy generation data	[3.7 ± 1.07] (3.4 ± 0.89/4.0 ± 1.22)	[33%] (13%/58%)
5	Sharing of high resolution meteorological data	[4.0 ± 1.07] (4.1 ± 1.20/4.0 ± 0.91)	[22%] (20%/25%)
6	Sharing of recycling waste data	[3.1 ± 0.78] (3.2 ± 0.90/3.3 ± 0.63)	[19%] (20%/17%)
7	Long-term monitoring of plant and animal population and crop conditions	[3.5 ± 1.14] (3.4 ± 1.12/3.5 ± 1.20)	[7%] (13%/0%)
8	Sharing high resolution greenhouse gas emission data	[3.1 ± 1.07] (3.1 ± 1.03/3.2 ± 0.93)	[7%] (7%/8%)

* Mean ± SD of the second survey (max: 5); ** others include universities and private/governmental research institutes;

*** totals may not equal 300% due to rounding.

3.2. Big Data Arising from Lifeforms (Bio-Big Data Related to Plant/Animal Genes and Diseases)

Delphi analysis was conducted for matters that should be implemented as a priority in order to contribute to the improvement of environmental and human sustainability through effective utilization of Big Data from lifeforms (Table 5). According to the result, the first issue that received the most points was the establishment of policies and legislation for the use of data related to genes and diseases in order to effectively use Big Data. Of course, when genetic and disease data are made public, a

thorough management system for preventing personal information leakage should be established, and a review of the current laws on personal information protection should also be conducted.

Second is the sharing of medical information and the establishment of a utilization system. This is the same opinion about the joint use of data, which was required for the effective utilization of Big Data related to the aforementioned atmospheric and environmental pollution. Statistical data related to medical insurance processing by and Health Insurance Review and Assessment Service, such as the number of patients in each day and the long-term prognosis of patients undergoing surgery, can be effectively used for managing and preventing diseases. The biggest factor that has impeded utilization of Big Data related to genes and diseases was also found to be the difficulty of securing data.

Third is the analysis of various electronic medical records written in natural language. Although the classification of the diseases of each patient is included in the admission record from the aforementioned Health Insurance Review and Assessment Service, important information, such as a patient's condition and progress, are not included in such data, and such information can also be known through medical records. These medical records need much time to analyze, because they are written in natural language. However, the text mining technique, one of the data mining techniques to extract key information from natural language, can be used for early detection and prevention of diseases.

Fourth is the creation of a database through the registration of individual genetic information in order to utilize personal genetic information for criminal investigation or for finding missing children. Thus, in order to utilize genetic information, laws on personal information protection must be also reviewed. All experts currently working in the industry and experts affiliated with universities and private/governmental research institutes considered both the first and second categories to be of utmost importance. However, the group of experts currently working in the industry placed importance on the analysis of medical records in natural language and the establishment of a sharing environment, whereas the experts affiliated with universities and private/governmental research institutes assigned higher points to the creation of genetic information database and vitalization of research utilization.

Fifth is the support for establishing a system that provides real-time health information. With the development of IT, if a system is built that can automatically place in a database information on weight, body temperature, blood pressure, urine, *etc.*, that is measured at home, it can be used to predict future health issues and the like by using such basic information.

Sixth is the establishment of an environment for studying and analyzing data concerning the school sports activities of children and adolescents and the influences on their growth and health. Through this, genetic information and health-related information per each individual can be secured from a young age, which can lead to a better life for the public through the utilization of customized health information per individual at a national level. Genetic information contains vast amounts of information that can be very useful for disease prediction and prevention.

Seventh is the creation of a database for information concerning varieties of crops, the area of production, climate conditions and yield, which can be used as important data for predicting the future yield of crops and improving varieties. Eighth is to establish genetic analysis services at a reasonable price so that many experts can easily access and analyze genetic information, which may lead to the development of a wide range of solutions for a better life for all of mankind.

Table 5. Policy issues for the utilization of Big Data arising from lifeforms.

	Results of First Survey	Results of Second Survey *	Results of Third Survey ***
		[All] (Industry/Others **)	[All] (Industry/Others **)
1	Establishment of policies and legislation for the use of bio-Big Data	[4.2 ± 0.91] (4.1 ± 0.99/4.2 ± 0.83)	[96%] (100%/92%)
2	Sharing data from the Health Insurance Review and Assessment Service	[4.3 ± 0.79] (4.2 ± 0.83/4.4 ± 0.77)	[96%] (100%/92%)
3	Sharing medical records written in natural language	[4.0 ± 0.81] (4.2 ± 0.73/3.8 ± 0.90)	[48%] (60%/33%)
4	Creating a database of genetic information	[4.1 ± 0.98] (4.1 ± 0.93/4.0 ± 1.08)	[33%] (27%/42%)
5	Establishment of a real-time health information system	[3.3 ± 0.84] (3.4 ± 0.71/3.2 ± 1.01)	[15%] (7%/25%)
6	Support research on the influence of school sports activities on growth and health	[3.2 ± 0.86] (3.5 ± 0.72/2.8 ± 0.90)	[7%] (7%/8%)
7	Creating a database on yield according to crop variety, area of production, climate conditions, etc.	[3.4 ± 0.89] (3.2 ± 0.97/3.6 ± 0.77)	[4%] (0%/8%)
8	Establishment of a genetic analysis service system at a reasonable price	[3.0 ± 1.00] (2.8 ± 1.07/3.2 ± 0.90)	[0%] (0%/0%)

* Mean ± SD of the second survey (max: 5); ** others include universities and private/governmental research institutes.

*** totals may not equal 300% due to rounding.

3.3. Big Data Arising from Machines/Things (M2M and IoT-Related Data)

Issues that must be established in order to contribute to the sustainability of the environment and humans through effectively utilizing Big Data from machines/things were subdivided into seven issues (Table 6). The first policy issue is the standardization of the government-led M2M/IoT service platform. Most standardized technology of M2M/IoT has been specifically designed for technologies in the corresponding industries. However, in order to support services linked to various industries, such as smart metering, smart appliances, smart grid and smart city, standardization of the M2M/IoT service platform is required. Such standardization should not be limited to domestic services, but there should be active participation in international standardization, so that the services can be linked internationally.

In addition to this, data security technology for preventing misuse or abuse of M2M/IoT data was selected as the second policy issue. If data from M2M and IoT are indiscriminately exposed, these can be used for criminal activities through data manipulation. Therefore, such security technology is the most urgent above all.

The third issue is the establishment of an environment for promoting service development based on various M2M/IoT. In order to utilize M2M/IoT in services, individuals and enterprises should develop service platforms to facilitate data access, and furthermore, plans for promoting the development of applications that can be easily accessed through, say, mobile phones are required.

The fourth issue is the expansion of a system that can monitor the status of energy consumption and recyclable waste treatment in real time. For the effective operation of an energy management system, real-time energy consumption monitoring data with the use of M2M/IoT technology can be useful. The

application of such an energy management system can be expanded to homes, plants, buildings, *etc.* In addition, if M2M/IoT technology is used to identify the movement of products from production to disposal, it can contribute to increasing the recycling rate of resources through the collection and management of wastes.

Table 6. Policy issues for the utilization of Big Data arising from machines/things. M2M, machine to machine; V2V, vehicle to vehicle; V2I, vehicle to infrastructure.

	Results of First Survey	Results of Second Survey *	Results of Third Survey ***
		[All] (Industry/Others **)	[All] (Industry/Others **)
1	Standardization of the M2M/IoT service platform	[4.4 ± 0.62] (4.5 ± 0.62/4.3 ± 0.63)	[100%] (100%/100%)
2	Development of technology for M2M/IoT data security	[4.2 ± 0.97] (4.4 ± 0.94/4.0 ± 1.00)	[59%] (67%/50%)
3	Support for the establishment of an eco-environment for promoting various types of M2M service development	[3.7 ± 0.91] (3.9 ± 0.90 / 3.5 ± 0.88)	[52%] (60%/42%)
4	Expansion of monitoring systems for real-time energy consumption and waste resources	[4.1 ± 0.84] (4.2 ± 0.83/3.9 ± 0.86)	[48%] (27%/75%)
5	Establishment of a disaster prevention system based on automatic measuring technology using M2M/IoT	[3.7 ± 1.03] (3.6 ± 1.00/3.7 ± 1.11)	[26%] (27%/25%)
6	Support for policies on the vitalization of V2V and V2I	[3.2 ± 1.05] (3.2 ± 1.03/3.1 ± 1.12)	[7%] (13%/0%)
7	Establishment of a government-led M2M/IoT data exchange center	[3.3 ± 1.09] (3.2 ± 1.13/3.5 ± 1.05)	[7%] (7%/8%)

* Mean ± SD of the second survey (max: 5); ** others include universities and private/governmental research institutes; *** totals may not equal 300% due to rounding.

The fifth issue is to support the establishment of a disaster prevention system based on automatic measuring technology that utilizes M2M/IoT. By analyzing monitoring data concerning the level of safety and the like received through sensors attached to public facilities, such as buildings, vehicles, ships, aircraft and bridges, accidents, such as a building collapse, can be prevented in advance.

The sixth issue is to vitalize vehicle to vehicle (V2V) and vehicle to infrastructure (V2I) technology, which is possible through M2M/IoT technology. As the names suggest, V2V is the communication between vehicles and V2I is the communication between vehicles and a traffic information center. This is effective for preventing accidents by notifying the driver with information like information about traffic ahead, approaching vehicles or the possibility of a collision. Furthermore, V2V and V2I can contribute to saving energy and reducing greenhouse gas emissions through smoothing traffic flow.

The seventh issue is the establishment of a government-led M2M/IoT data exchange center, since the proper use of M2M/IoT data can be a big help for developments in the industrial and academic field. Concerning the utilization plan for data arising from machines/things, the opinions from both groups—experts currently working in the industry and those affiliated with universities and

private/governmental research institutes—are in agreement overall. However, experts currently working in the industry placed higher importance on practical issues for promoting various forms of M2M service development, such as the establishment of an eco-environment, whereas the institute/academic group placed more importance on the expansion of monitoring systems for real-time energy consumption and waste resources.

3.4. Big Data Arising from People (Digital Media-Related Data)

Issues that must be addressed first to contribute to human and social sustainability through effectively utilizing Big Data from people can be divided into six issues (Table 7). Overall, the group of experts currently working in the industry and those affiliated with universities and private/governmental research institutes assigned a similar level of importance. The first issue is the publication of government-led safety-related data. Through the analysis of location, time, age, gender and accident data, it is possible to establish an accident prevention information service that is customized to each individual, and such a system can be used for disaster preparation. For example, the location information on a personal mobile phone can be used to provide a service that can notify the user of a danger when the user is identified as moving toward an accident-prone area.

As the second issue, the review of laws related to the protection of personal information was proposed. Although personal location information is very useful for disaster prevention and measures, personal information can be leaked and cause damage to individuals. In order to prevent this from happening, the establishment of a security management system and legal system concerning the gathering and use of such information is imperative.

The third issue is the publication of criminal data subdivided into time/space and the establishment of such a utilization system. Through this, time/space information concerning a crime is not only helpful for the general public to prepare in advance, but also the result of an analysis on Big Data related to this can be used as a precautionary measure to block environments with a high possibility of criminal activity in advance.

The fourth issue is the establishment of a system for improving welfare management based on Big Data analysis. Through this system, needy people or areas without welfare benefits can be discovered, and a plan to provide convenient facilities can be established effectively based on behavioral analysis.

The fifth issue is the active support for the analysis of data, such as various search engines and social network services (SNSs). The amount of data from such digital media is increasing exponentially, and with effective utilization of the data, useful information can be obtained, such as the general public's perception of politics, the economy, society and culture, as well as consumption patterns, *etc.* In addition, this can be very useful for emergency response processes during a national disaster. However, it is difficult for general researchers to access data from such digital media. With the publication of such data and support for relevant research at a government level, useful information can be obtained for various fields, such as product development and service improvement for enterprises, the identification of research trends for schools and institutes, *etc.*

Finally, the sixth issue is the establishment of a location information system for individuals by using mobile phones and other devices used by individuals. Such location information can be used to facilitate the identification of the public's moving path or densely populated areas, which is effective

for implementing emergency response processes during national disasters, such as the outbreak of a new epidemic. In addition, it can be useful for discovering areas without welfare benefits through behavioral analysis. The above location information system can also determine the location for installing convenient facilities for the general public and be applied in predicting area-specific power demands, due to the movement of the population. Prior to the establishment of such location information system, it is necessary to secure fair competition among businesses, as well as to discuss the neutrality, openness, *etc.*, of the location information. Of course, the protection of personal information must be reviewed prior to the establishment of such individual-specific location information systems, which is the content of the second issue.

Table 7. Policy issues for the utilization of Big Data arising from people.

	Result of First Survey	Result of Second Survey *	Result of Third Survey ***
		[All] (Industry/Others**)	[All] (Industry/Others **)
1	Publication of government-led data related to safety	[3.9 ± 0.97] (3.9 ± 0.83/3.8 ± 1.17)	[78%] (87%/67%)
2	Review of legislation related to personal information protection and usage	[4.2 ± 0.91] (4.2 ± 0.95/4.2 ± 0.90)	[67%] (60%/75%)
3	Publication of criminal data and establishment of a utilization system	[4.0 ± 1.13] (4.1 ± 0.90/3.8 ± 1.41)	[67%] (67%/67%)
4	Establishment of a system for improving welfare management based on Big Data analysis	[3.6 ± 0.94] (3.5 ± 0.94/3.7 ± 0.95)	[37%] (33%/42%)
5	Publication of social media and search engine records	[3.9 ± 0.91] (3.8 ± 1.03/4.2 ± 0.69)	[30%] (33%/25%)
6	Establishment of an environment for energy reduction, the prevention of disaster/emergencies and welfare based on individual-specific location information	[3.9 ± 0.80] (3.8 ± 0.73/4.0 ± 0.91)	[22%] (25%/20%)

* Mean ± SD of the second survey (max: 5); ** others include universities and private/governmental research institutes; *** totals may not equal 300% due to rounding.

4. Conclusions

This paper discussed what issues in terms of policy must be first addressed in order to improve sustainability for society through the effective use of Big Data through green IT. In this study, Big Data were divided into four categories based on the sources, which were nature/the environment, lifeforms, machines/things and people, where the preceding issues or policy issues were derived for each category. To derive the issues, an open-ended questionnaire was conducted by targeting experts in relevant fields, and a consensus was reached by using the Delphi technique. According to the results of the Delphi analysis, which was conducted over three phases, policy issues that should be addressed as priorities can be summarized as follows.

First, in order to effectively utilize Big Data arising from nature/the environment, it is required to establish the infrastructure for the integrated management and sharing of government-led environmental and meteorological data. Since 2013, the Korean government has made the transition of multiple secured data into public data as a part of the “Government 3.0 plan”. The government-led

continuous publication of data will increase, for the public to have the chance to effectively utilize data and, further, to contribute to green growth of Korea. In addition, there should be support for the establishment of an environmental pollution alert system at a national level by using meteorological and environmental pollution data. Besides, support of the monitoring of environmental pollution focused on areas near industrial complexes is especially needed in order to contribute to improving the public health.

Second, in order to effectively utilize bio-Big Data arising from lifeforms, support for a government-level policy for the joint utilization of genetic and disease-related Big Data is needed. In 2014, a post-genome project was officially launched to develop and commercialize genomic technologies in Korea. However, government policies related to the joint use of the genetic information are still needed. Furthermore, the vitalization of research on diseases has been derived as an important issue in order to improve the quality of people's lives. Thus, medical records, such as the number of patients and the prognosis of patients undergoing surgery for each disease category, need to be open for research purposes.

Third, in order to create a base for effectively utilizing Big Data arising from M2M/IoT, the government-led standardization of the M2M/IoT service platform is required more than anything else. Especially, in order to vitalize small enterprises with insufficient technology, such standardization work should be led by the government, instead of relying on large enterprises. Currently, the standardization of the M2M/IoT service platform and related areas has been mainly led by the private sector in Korea. For example, the standardization of the smart grid was conducted through the Smart Grid Standardization Forum and National Coordinator for the Smart Grid. If the government actively participates in this, the standardization will be accelerated. In addition, there must be more effort on the development of data security technology in order to prevent the misuse of data resulting from M2M/IoT. Furthermore, the establishment of an eco-friendly environment by developing various services through the proper use of data from M2M/IoT has been derived as an issue. Furthermore, priority was given to establishing a monitoring system for energy consumption and waste resources, as well as a disaster prevention system.

Fourth, in order to effectively utilize Big Data arising from people using digital media, the publication of data related to public safety and the review of legislation related to its utilization have been derived as important policy issues. Additionally, opinions that the data should be used to prevent various crimes and to improve welfare have been derived. For example, data can be utilized to effectively implement emergency response processes during disasters through identification of people's movement behavior by establishing an individual-specific location information system using devices, like mobile phones. In 2014, the Korean government announced a disaster response system, the "Smart Big Board", based on Big Data-related technologies, including digital media analytics, as a part of the "Government 3.0 plan". However, the legislation of policies related to the publication of safety data is still needed.

Korea is making great efforts to move one step closer toward becoming an advanced country through sustainable development rather than the past form of economic development. The level of IT is already remarkably developed to properly utilize Big Data, but support from policy is currently needed. The results of this study show what type of policies need to be addressed in order for Big Data to be effectively used for Korea's sustainable development. These results can not only be effectively

used as data for making policies that will contribute to the sustainable development of Korea, but also as useful data for developing countries that are on or will be on a similar path as Korea.

Author Contributions

Hyun Baek performed the experiments, analyzed the data and wrote the paper. Sun-Kyoung Park, the thesis supervisor, provided guidance, revision and feedback for this work.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Bachour, N.; Chasteen, L. Optimizing the Value of Green IT Projects within Organizations. In Proceedings of the Green Technologies Conference, 2010 IEEE, Grapevine, TX, USA, 15–16 April 2010; pp. 1–10.
2. Kuo, B.; Dick, G. The greening of organisational IT: What makes a difference? *Australas J. Inf. Syst.* **2009**, *16*, 81–92.
3. Schmidt, N.-H.; Ereke, K.; Kolbe, L.M.; Zarnekow, R. Towards a Procedural Model for Sustainable Information Systems Management. In Proceedings of the 42nd Hawaii International Conference on System Sciences, Manoa, HI, USA, 5–8 January 2009; pp. 1–10.
4. Watson, R.T.; Boudreau, M.C.; Chen, A.J. Information Systems and Environmentally Sustainable Development: Energy Informatics and New Directions for the IS Community. *Manag. Inf. Syst. Q.* **2010**, *34*, 23–38.
5. Yi, L.; Thomas, H.R. A review of research on the environmental impact of e-business and ICT. *Environ. Int.* **2007**, *33*, 841–849.
6. Rahim, R.E.; Rahman, A.A. Green IT Capability and Sustainable Development. In Proceedings of the 1st International Conference on Green Computing, Technology and Innovation (ICGCTI), Kuala Lumpur, Malaysia, 4–6 March 2013; pp. 80–88.
7. Gartner. The importance of “Big Data”: A Definition. Available online: <https://www.gartner.com/doc/2057415/importance-big-data-definition> (accessed on 21 June 2012).
8. Gartner. Predicts 2014: Green IT and Sustainability. Available online: <https://www.gartner.com/doc/2637318/predicts-green-it-sustainability> (accessed on 13 December 2013).
9. Kim, S.; Kim, H.-K.; Kim, H.J. Climate Change and ICTs. In Proceedings of the 31st International Telecommunications Energy Conferences, Incheon, Korea, 18–22 October 2009.
10. Hilpert, H.; Schumann, M.; Kranz, J. Leveraging Green IS in Logistics Developing an Artifact for Greenhouse Gas Emission Tracking. *Bus. Inf. Syst. Eng.* **2013**, *5*, 313–323.
11. Nagar, Y.; Malone, T.W. Making Business Predictions by Combining Human and Machine Intelligence in Prediction Markets. In Proceedings of the 32nd International Conference on Information Systems: Knowledge Management and Business Intelligence, Shanghai, China, 4–7 December 2011.

12. Manyika, J.; Chui, M.; Brown, B.; Bughin, J.; Dobbs, R.; Roxburgh, C.; Byers, A.H. *Big Data: The Next Frontier for Innovation, Competition, and Productivity*; McKinsey Global Institute: New York, NY, USA, 2011.
13. Villars, R.L.; Olofson, C.W.; Eastwood, M.E. *Big Data in CEE: What It Is and Why You Should Care*; IDC Analyze the Future: Framingham, MA, USA, 2011.
14. IBM. *Analytics: The Real-World Use of Big Data: How Innovative Enterprises Extract Value from Uncertain Data*; IBM Global Services: Somers, NY, USA, 2012.
15. Castro, D. *Learning from the Korean Green IT Strategy*; The Information Technology & Innovation Foundation: Washington, DC, USA, 2009; pp. 1–4.
16. Presidential Commission on Green Growth. *Road to Our Future: Green Growth: National Strategy and the Five-Year Plan (2009~2013)*; Presidential Commission on Green Growth: Seoul, Korea, 2009; p. 13.
17. Hsu, C.-C.; Sanford, B.A. The Delphi Technique: Making Sense of Consensus. *Pract. Assess. Res. Eval.* **2007**, *12*, 1–8.
18. Rowe, G.; Wright, G. The Delphi technique as a forecasting tool: Issues and analysis. *Int. J. Forecast.* **1999**, *15*, 353–375.
19. Hilbert, M.; Miles, I.; Othmer, J. Foresight tools for participative policy-making in inter-governmental processes in developing countries: Lessons learned from the eLAC Policy Priorities Delphi. *Technol. Forecast. Soc. Chang.* **2009**, *76*, 880–896.
20. Helmer, O.; Rescher, N. On the Epistemology of the Inexact Sciences. *Manag. Sci.* **1959**, *6*, 25–52.
21. Klee, A.J. The utilization of expert opinion in decision-making. *Aiche J.* **1972**, *18*, 1107–1115.
22. Oh, K.S. *Forecasting through Hierarchical Delphi*; Ohio State University: Columbus, OH, USA, 1974.
23. Pill, J. The Delphi method: Substance, context, a critique and an annotated bibliography. *Soc. Econ. Plan. Sci.* **1971**, *5*, 57–71.
24. Witkin, B.R.; Altschuld, J.W. *Planning and Conducting Needs Assessment: A Practical Guide*, 1st ed.; Sage Publications: Thousand Oaks, CA, USA, 1995.
25. Brooks, K.W. Delphi Technique: Expanding Applications. *North Cent. Assoc. Q.* **1979**, *53*, 377–385.
26. Custer, R.L.; Scarcella, J.A.; Stewart, B.R. The Modified Delphi Technique—A Rotational Modification. *J. Vocat. Techn. Educ.* **1999**, *15*, 50–58.
27. Cyphert, F.R.; Gant, W.L. The Delphi Technique: A Case Study. *Phi Delta Kappan* **1971**, *52*, 272–273.
28. Ludwig, B. Predicting the Future: Have you considered using the Delphi Methodology? *J. Ext.* **1997**, *35*, 1–4.
29. Marmur, A.; Park, S.K.; Mulholland, J.A.; Tolbert, P.E.; Russell, A.G. Source apportionment of PM_{2.5} in the southeastern United States using receptor and emissions-based models: Conceptual differences and implications for time-series health studies. *Atmos. Environ.* **2006**, *40*, 2533–2551.
30. Park, S.K.; Marmur, A.; Kim, S.B.; Tian, D.; Hu, Y.T.; McMurry, P.H.; Russell, A.G. Evaluation of fine particle number concentrations in CMAQ. *Aerosol Sci. Technol.* **2006**, *40*, 985–996.
31. Park, S.K.; Marmur, A.; Russell, A.G. Environmental Risk Assessment: Comparison of Receptor and Air Quality Models for Source Apportionment. *Hum. Ecol. Risk Assess.* **2013**, *19*, 1385–1403.

32. Park, S.K.; Russell, A.G. Regional adjustment of emission strengths via four dimensional data assimilation. *Asia Pac. J. Atmos. Sci.* **2013**, *49*, 361–374.
33. Park, S.K.; Cobb, C.E.; Wade, K.; Mulholland, J.; Hu, Y.T.; Russell, A.G. Uncertainty in air quality model evaluation for particulate matter due to spatial variations in pollutant concentrations. *Atmos. Environ.* **2006**, *40*, 563–573.
34. Cheong, T.S.; Felix, M.L.A.; Jeong, S.M. Development of GIS-based disaster assessment system to reduce flood risks in urbanized creeks. *Desalin. Water Treat.* **2014**, *52*, 2817–2825.
35. Mahendra, R.S.; Mohanty, P.C.; Bisoyi, H.; Kumar, T.S.; Nayak, S. Assessment and management of coastal multi-hazard vulnerability along the Cuddalore-Villupuram, east coast of India using geospatial techniques. *Ocean Coast. Manag.* **2011**, *54*, 302–311.
36. Van Haaren, R.; Fthenakis, V. GIS-based wind farm site selection using spatial multi-criteria analysis (SMCA): Evaluating the case for New York State. *Renew. Sustain. Energy Rev.* **2011**, *15*, 3332–3340.
37. Cho, H.W.; Kim, S.B.; Jeong, M.K.; Park, Y.; Miller, N.G.; Ziegler, T.R.; Jones, D.P. Discovery of metabolite features for the modelling and analysis of high-resolution NMR spectra. *Int. J. Data Min. Bioinform.* **2008**, *2*, 176–192.
38. Cho, H.W.; Kim, S.B.; Jeong, M.K.; Park, Y.; Ziegler, T.R.; Jones, D.P. Genetic algorithm-based feature selection in high-resolution NMR spectra. *Expert Syst. Appl.* **2008**, *35*, 967–975.
39. Kim, S.; Tsui, K.; Borodovsky, M. Multiple hypothesis testing in large-scale contingency tables: Inferring pair-wise amino acid patterns in beta-sheets. *Int. J. Bioinform. Res. Appl.* **2006**, *2*, 193–217.
40. Kim, S.B.; Lee, J.W.; Kim, S.Y.; Lee, D.W. Dental Informatics to Characterize Patients with Dentofacial Deformities. *PLoS One* **2013**, doi:10.1371/journal.pone.0067862.
41. Kim, S.B.; Temiyasathit, C.; Chen, V.C.P.; Park, S.K.; Sattler, M.; Russell, A.G. Characterization of spatially homogeneous regions based on temporal patterns of fine particulate matter in the continental United States. *J. Air Waste Manag. Assoc.* **2008**, *58*, 965–975.
42. Kim, S.B.; Wang, Z.; Oraintara, S.; Temiyasathit, C.; Wongsawat, Y. Feature selection and classification of high-resolution NMR spectra in the complex wavelet transform domain. *Chemom. Intell. Lab. Syst.* **2008**, *90*, 161–168.
43. Mei, Y.J.; Kim, S.B.; Tsui, K.L. Linear-mixed effects models for feature selection in high-dimensional NMR spectra. *Expert Syst. Appl.* **2009**, *36*, 4703–4708.
44. Yang, D.H.; Kang, J.H.; Park, Y.B.; Park, Y.J.; Oh, H.S.; Kim, S.B. Association Rule Mining and Network Analysis in Oriental Medicine. *PLoS One* **2013**, doi:10.1371/journal.pone.0059241.
45. Voros, S.; Elashoff, M.R.; Wingrove, J.A.; Budoff, M.J.; Thomas, G.S.; Rosenberg, S. A peripheral blood gene expression score is associated with atherosclerotic Plaque Burden and Stenosis by cardiovascular CT-angiography Results from the PREDICT and COMPASS studies. *Atherosclerosis* **2014**, *233*, 284–290.
46. Paoletti, M.G.; Pimentel, D. Genetic engineering in agriculture and the environment. *Bioscience* **1996**, *46*, 665–673.
47. Ko, E.-N. A Multimedia Replicated Architecture for Transportation Safety Service. *J. Korea Multimed. Soc.* **2014**, *17*, 226–231.
48. Lee, G.B.; Ko, M.J.; Ku, T.J. Analysis of energy efficiency in PCB manufacturing process. *Int. J. Precis. Eng. Manuf.* **2012**, *13*, 1215–1220.

49. Lee, J.-E.; Shin, M.-S. An Exploratory Study on the Feasibility of M2M Services. *J. Korea Soc. IT Serv.* **2012**, *11*, 125–139.
50. Tao, F.; Cheng, Y.; Xu, L.D.; Zhang, L.; Li, B.H. CCIoT-CMfg: Cloud Computing and Internet of Things-Based Cloud Manufacturing Service System. *IEEE Trans. Ind. Inform.* **2014**, *10*, 1435–1442.
51. Fan, Z.; Haines, R.J.; Kulkarni, P. M2M Communications for e-health and Smart Grid: An Industry and Standard Perspective. *IEEE Wirel. Commun.* **2014**, *21*, 62–69.
52. Lopez, G.; Moura, P.; Moreno, J.I.; Camacho, J.M. Multi-Faceted Assessment of a Wireless Communications Infrastructure for the Green Neighborhoods of the Smart Grid. *Energies* **2014**, *7*, 3453–3483.
53. Dijkstra, L.J.; Yakushev, A.V.; Duijn, P.A.C.; Boukhanovsky, A.V.; Sloot, P.M.A. Inference of the Russian drug community from one of the largest social networks in the Russian Federation. *Qual. Quant.* **2014**, *48*, 2739–2755.
54. Ginsberg, J.; Mohebbi, M.H.; Patel, R.S.; Brammer, L.; Smolinski, M.S.; Brilliant, L. Detecting influenza epidemics using search engine query data. *Nature* **2009**, *457*, 1012–1014.
55. Choi, S.; Kim, H.; Yoon, J.; Kim, K.; Lee, J.Y. An SAO-based text-mining approach for technology roadmapping using patent information. *R&D Manag.* **2013**, *43*, 52–74.
56. Seol, H.; Lee, S.; Kim, C. Identifying new business areas using patent information: A DEA and text mining approach. *Expert Syst. Appl.* **2011**, *38*, 2933–2941.
57. Terada, M.; Nagata, T.; Kobayashi, M. Population estimation technology for mobile spatial statistics. *NTT DOCOMO Techn. J.* **2013**, *14*, 10–15.
58. Delbecq, A.L.; van de Ven, A.H.; Gustafson, D.H. *Group Techniques for Program Planning: A Guide to Nominal Group and Delphi Processes*; Scott, Foresman, and Co.: Glenview, IL, USA, 1975.