International Journal of Database Theory and Application
Vol.9, No.11 (2016), pp.245-256
http://dx.doi.org/10.14257/ijdta.2016.9.11.22

A Data Analysis Methodology for Measuring Practical Technology Impact Index by Analyzing Trends Data

Jae Hyuk Cho

Korea Institute of S&T Evaluation and Planning
cool@kistep.re.kr

Abstract

In the feasibility analysis of R&D program, the data used to analyze the impact/trends/level of technology derive mostly from patents and theses. However, there is limitation in reflecting the newest technology trends data based on patents and theses. That is because of the occurrence of a one or two year gap time before these patents (or theses) are actually published or granted. Therefore, not only are related patents and theses data collected but, the extensive trends data from public websites and social networks also need to be collected and analyzed. It takes a great deal of time, and manpower for these related feasibility analysis to happen successfully. To solve this issue, this analysis presents a methodology not only to rapidly and accurately collect data but, to efficiently analyze the newest technology trend flows. To analyze technology impact, phases of the data extraction, the application of measuring model and the determination of TIIB (Technology Impact Index based on Big Data) are processed. This theses proposes that the data analysis methodology used to find out the latest technology trends could also be useful for optimizing efficiency when analyzing. Moreover, the newly developed TIIB enables us to check the interest trends of the technology by reading the yearly changes.

Keywords: *TIIB (Technology Impact Index based on Big-data), Big Data, Feasibility Analysis, Panel Data Model, Quantitation Analysis*

1. Introduction

In the feasibility study of the government R&D program, researchers typically reference the data submitted by the government and or related public institutions or survey the related data to analyze the trend level and the status of technology [1]. Currently, the data to analyze the technology trends are mostly from patents and theses which provide a relatively objective measure of new technology. However, because there occurs the gap of 1~2 years before actual opening or publishing a patent or thesis, there is a limitation in ability to reflect the newest data. Therefore, the process to collect and analyze the extensive trend data provided in the Internet is required. It takes a great deal of time and manpower because of the limited ability of the related feasibility study. To solve this issue, this study presents a methodology not only to rapidly and accurately collect big data but, to efficiently analyze the newest technology trend flows.

2. Related Studies

Previous studies to provide efficient national research and developments by utilizing big data in Korea have been led by the Korea Institute of Science and Technology Information [3-4]. Actual research into big data platforms has been conducted, including research found in scientific and technological literature. This is in order to develop service models for big data analysis which are centric to the national R&D information, to support the establishment of science and technology strategy and to minimize possible risks from the large scaled investment and low success probability on such activities.

ISSN: 2005-4270 IJDTA

However, there have been no measures taken to scale the technology impact of web data like, the citation degree of patents and theses.

3. Methodology for Efficient Analysis of Feasibility Study Based on Big Data

There is a five step feasibility study involved in accessing the technology impact or promising nature of a core technology when utilizing big data (See Figure 1.) as shown below.

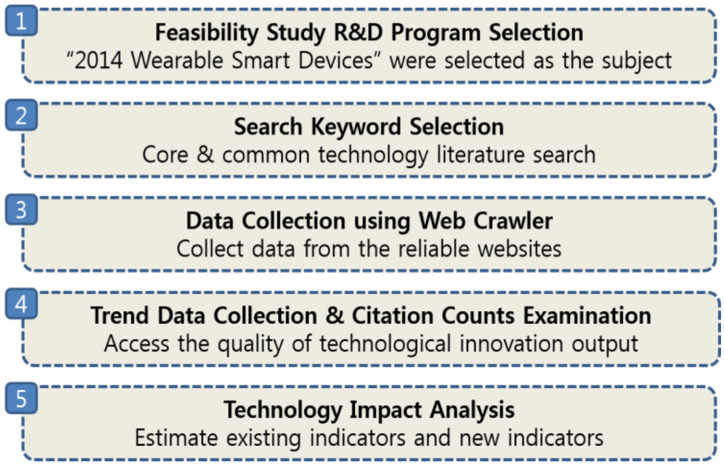

Figure 1. The Methodology for Efficient Analysis of Big Data Based on a Feasibility Study

(1) Feasibility Study R&D Program Selection

First, the R&D Program needs to be selected for the feasibility study based on big data analysis. In this study, "Wearable Smart Device" was selected as the subject for study—real datasets were used. This was partly in order to enhance the understanding of the methodology for feasibility study through a real dataset.

(2) Search Keyword Selection

In this study we examine the technology impact of the core technologies of the feasibility study cases versus other technologies in relevant fields. We discovered that the wearable smart device's core technology fields are "Wearable, Smart, Device, Material Part, Platform, Input, Output, Process, and Power" and the upper level technology area of the relevant technology is "Wearable". By comparing the technology impact of the core technology and of the upper level technological area of the relevant technology, we will find the technology impact of the core technology of the program which was subjected to a feasibility study.

o The keywords of the core technologies subjected for the feasibility study :

Wearable AND (Smart OR Device) AND (Part OR Platform OR Input OR Output OR Process OR Power)

o The keywords of the core technologies of the upper level area subjected for the feasibility study :

Wearable AND Device

(3) Data Collection using Web Crawler

Google Web Search API can search entire websites or it can search collections of predefined websites while Custom Search API can search over only a set of predefined websites. Custom Search API can have several rules (*e.g.* *.or.kr, *.re.kr *.go.kr, *.org *etc.*) so it can get a report on collective data resources—it returns up to 10 results per query: if you want to display more than 10 results, you can issue multiple requests. Whereas, Web Search API can apply only one rule so it cannot get such collective reports simultaneously. For these reasons, if you need to search over a collection of specific websites, you should use one website per one search and gather the related results manually in accordance with each specific rule. Moreover, it is highly impossible to get accurate search results if the keyword is broad and unobvious. A wide range of search keyword may generate very complex and far reaching search results and such results may not even include the entire search results. In order to increase the accuracy of Google API search results, it should be asked to draw up more specific and detailed keyword.

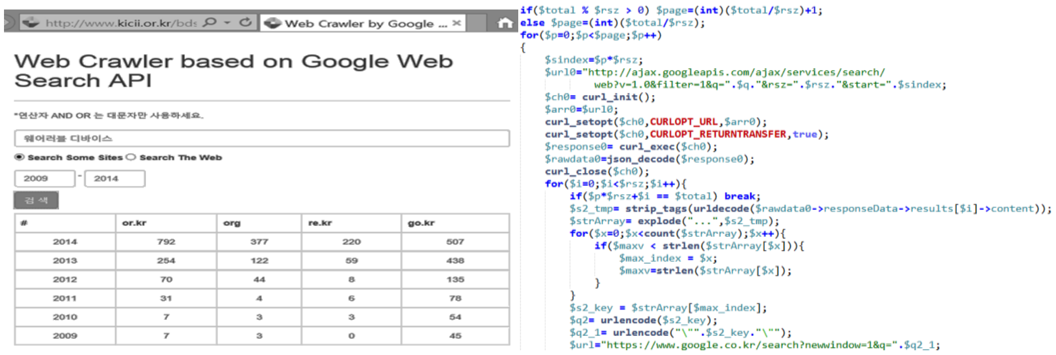

Figure 2. Web Crawler Based on Google Web Search API

It is expected that web resources generally are not appropriate to utilize for a feasibility study because the credibility of web resources in general has been lower than objective measurements such as theses and patents. However, we assume that government or related public institution’s web resources are enough to use as analysis data. Thus, we enhanced the credibility of data resources by collecting data from the government or related public institutions’ websites using web crawler. The web crawler we developed presents the number of keyword searches through Web Search API. Basically it collects search result using Web and Blog API offered by Google Web Search API. When you start the web crawler, the keyword and search period should be entered to configure the search period date — uses ‘date range’ attribute. The web crawler transforms the data after the user configures the period, because Google uses Gregorian time stamp to express time instead of using year, month and day. The web crawler takes the result data from web and blog in respects of search keywords and period of search time using Web Search API. Also one of the data results shows the number of keyword searches, ‘estimated Result Count’ attributes for each year. The problem is redundant data can be found calculating specific trends data using the web crawler. Because Google saves search history in a number of ways, it sometimes recognizes the same website as a different one and offers redundant data results. An in-depth examination for enhancing the credibility of the web crawler is needed. For this issue, the study presents how to correct duplicate data— those having the same URL and sub directory — concerning specific keyword.

(4) Trend Data Collection and Citation Counts Examination

Mostly, the number of theses and patents does not mean the level of quality but it stands for a fundamental indicator for the quantification which measures related technologies' productivity. The number of citation of theses and patents are used to access the quality of technological innovation output. On the other hand, the web data can be utilized as a quantitative data like theses and patents yet qualitative data. We believed that the number of citation of web data takes into account the occurrence of data scraped and linked to the other web sites but, the function of big data collection and analysis needs to be developed for the purpose of greater scalability and faster access or analytics.

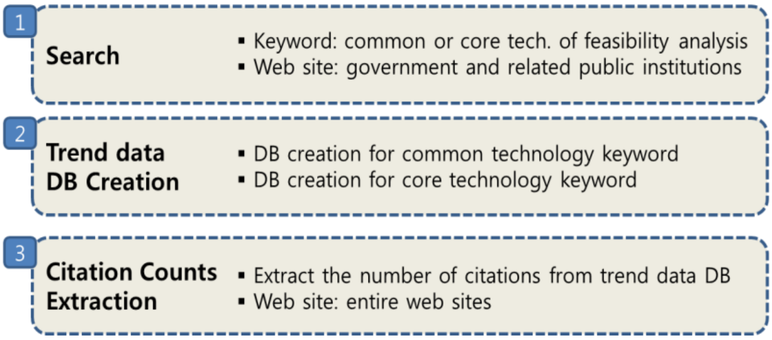

Figure 3. Extraction Process of the Number of Trend Data Citations

Google Web Search API offers the number of keyword searches that could be collected if necessary. However, the searched trend data does not reflect the number of scrap and link in the entire website. For this reason, this study creates a list of each year's trend data from the government or related public institutions' websites and stores them in the database for the qualitative judgment. Moreover, the web crawler presents the number of citations data (*e.g*. the number of the scraped and liked) by analyzing each trend data in the database intending entire web sites. By comparing the technology impact of the core technology and of the common technology of the relevant technology, we will find the number of citation and trend data concerning each technology.

(5) Results of Technology Impact Analysis

The study used real datasets such as the theses, patents and trend data in order to analyze the technology impact of the core technology of the program subjected to a feasibility study. In this study, the related thesis data needing to be analyzed was collected from RISS, NDSL, SCOPUS and Thomson and the relevant patent data was collected from NDSL, KIPRIS and patent web sites subscription based paid patent search websites. For the trend data, the number of keyword searches and citations was collected from official documents and social networks using the web crawler. The figures are estimated based on the number of citation of patents and technology trend data (see Figure 4, Figure5.) as shown below:

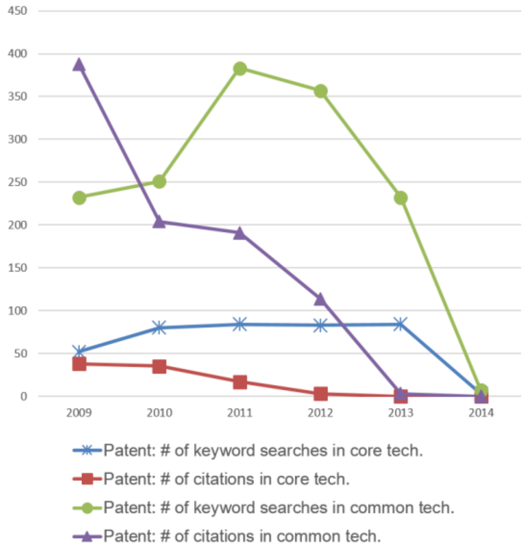

Figure 4-1 Patent Data

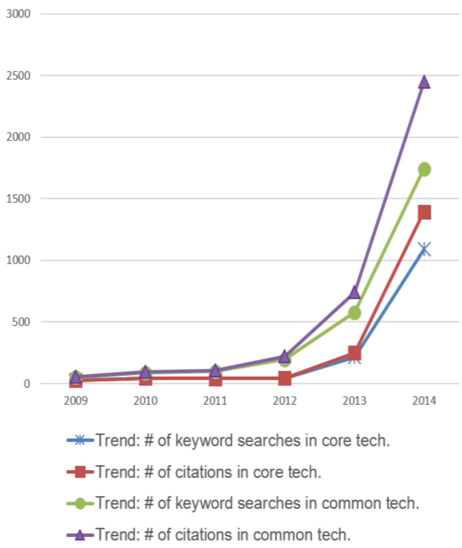

Figure 4-2 Trend Data

Figure 4. Changes in the Number of Keyword

Overall, as compared with the patents data, the trend data's total number of keyword searches and citations has steadily increased; it especially shows rapid growth in 2013. Therefore, we can infer that wearable device related technology has been a potential field since 2013. The rate of increase of patent data has grown only slightly, while the rate of increase of common technology trend data has increased dramatically—the rate of total keyword searches and citations in 2014 has increased about 4,533% and 5,362% in each case compared to 2009. Also, the rate of increase of core technology has grown dramatically—the rate of total keyword searches and citations in 2014 has increased about 3557% and 4371% in each case compared to 2009.

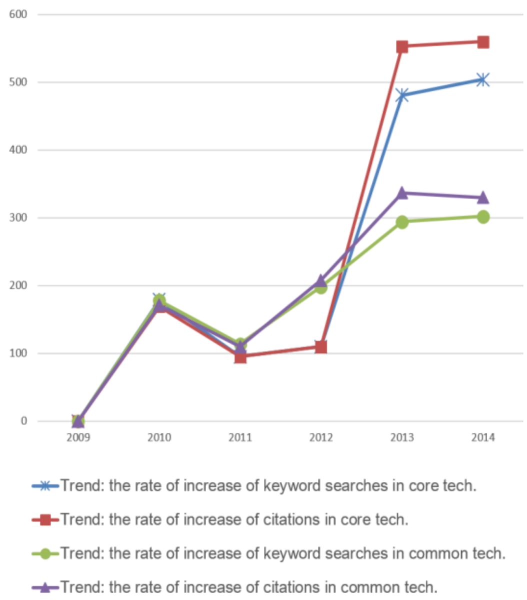

Figure 5. Common and the Core Technology Trends in Wearable Smart Device

In the case of the rate of increase of keyword searches and the citations, the core technology' rate of increase has been similar or lower than the common technologies was between 2009 and 2012. Since 2013, the common technology has remained higher than the core technology. Therefore, we've concluded that in the recent two years, the core technology is quickly emerging.

4. Extraction and Example Analysis of TIIB based on Trend Data

The application of big data technique is used as a means to adequately capture the independent variable data of measuring model in as timely a manner possible. In the traditional technique, data has been collected using statistical results, yet in the big data analysis technique the data could be collected in real time and it enables shortened processing time in between data preprocessing and data interpretation steps. It presents each step to discover the technology impact utilizing big data techniques and panel data models. To calculate Technology Impact Index based on Big Data (TIIB), phases of the data extraction, the application of measuring model and the determination of TIIB are processed. First, in the data extraction step, independent variable and dependent variables should be decided to apply to the panel data model. Second, based on the extracted data one of the appropriate panel data models should be selected. Through the selected panel data model, we can find out how much the independent variable influences the dependent variable. To execute TIIB we should draw up each variable and the weight of each variable and the weight is drawn from the panel data models.

1) Phase 1: Data Extraction

o Determining a dependent variable

In most applications of production functions, capital, labor or land are typically used as the independent variables. Recently, specific technology also needs to be regarded as the independent variable of the production function. We consider that these independent variables affect a dependent variable of the production function, especially market size. That is, through continuing R&D in technology, related market size will continue to improve and we can conclude that there is an association between specific technology and market size.

o Determining an independent variable

The extracted data with big data techniques are mainly used as the independent variable of production function to ensure up-to-date data. The limitation of the previous data collecting methodology was that the feasibility study only used public open data. Using big data techniques allows the ability to find the latest trends of technology and finally, it enhances data credibility.

In general, the factors that influence the market scale are determined to be quite huge. However, because this research does not purpose to identify all of factors affecting the market scale, it does not consider the other variables and instead considers only the impact of R&D on the market scale. Therefore, we utilize the data that may indicate interests in R&D on specific technology as the independent variables because the government R&D projects are closely related to the increase of the technology competitiveness. We collected the number of theses, patents, official documents and social networks and their citation number and we used it as the independent variables which could explain the dependent variables. Also, the data was split into two groups to utilize the panel model that minimizes the error arising from the data that are not considered in the independent and dependent variable and the relationship between dependent variables and independent variables.

o Determining a data period

The data period of time is an important consideration in the data extraction process. In order to perform a more accurate analysis than the panel model, the long-term period data could reduce error rates in the quantitative model. However, it is difficult to obtain data of more than ten years of R&D projects. Therefore, on average, the analysis utilizes five to ten years of data.

o Determining data groups

As stated above, it is difficult to collect long-term data. For this reason, splitting the data group will result in a more sophisticated analysis of panel model. In addition, the analysis based on the data of a different group could effectively control the errors when we analyze the cause and effect relationships (the impact on dependent variables) by changing the independent variables in the R&D projects. In this research, group 1 utilizes the number of searches as the independent variables of related technology and group 2 utilizes the number of citations as the independent variables of related technology. Furthermore these groups can be divided in more precise ways. For example, the number of searches can be subdivided into the number of searched in the core technology and the number of searched in the common technology. Group 2 can also be subdivided into two groups in the same way. After collecting independent variables using traditional and big data approaches, the relationship between the dependent and independent variables is identified by applying a quantitative model. The quantitative model utilizes panel models because there exist a variety of independent variables possible for each year. The following will be described in respects to the analysis method of the panel model.

2) Phase2: Analyzing a Quantitative Model

To find out how much the independent variables influence the dependent variables, the panel data analysis techniques is applied in order to derive the coefficients of the independent variables. The panel data analysis technique is also applied to take advantage of the modified nature of the panel model according to the business categories of the feasibility analysis. Basically panel data such as that found in Eq.1 could determine the influence of the independent variables.

$$y_{it} = \beta' X_{it} + u_i + \epsilon_{it} \quad \text{[Eq. 1]}$$

- i**:** independent variables.
- t: time variable
- β'**:** the coefficients of the independent variables.
- x_{1t}: theses number and its citation numbers
- x_{2t}: patents number and its citation numbers
- x_{3t}: official document number and its citation numbers
- x_{4t}: social network number and its citation numbers

In order to evaluate a model we usually take into account the following measures: ϵ_{it} represents random shock or idiosyncratic error. —the constant term reflects the number of errors that may occur in the analysis of the model.

u_i represents non-observed individual effects—this serves to control the effects of the other independent variables describing the market size. Coefficients in the model evaluate the average effect on the dependent variable y. This is according to the change in the independent variable x after controlling the individual-level differences. β' is drawn by applying the panel model to STATA program. Equation (1) is a general definition of the expression for the one group. It is applied to the other groups to analyze the model panel.

In order to apply the above model, we need to extract the data for independent variables. When the extraction is complete, it is necessary to distinguish whether all data is present or whether some data is partially not extracted. If the panel dataset is missing some observations during the experimental period, it is described as unbalanced panel. Therefore, different methods are applied depending on the extracted independent variables.

As described above, because of its ability to effectively reduce the number of errors that can occur in quantitative analysis, utilizing panel data is preferred over the use of simple cross-sectional data or time series data. Therefore, it is possible to take advantage of the modified model seen in Equation 2, in order to control the common factors of errors that may occur during specific periods of time.

$$y_{it} = \beta' X_{it} + \theta_t + \epsilon_{it} \quad \text{[Eq. 2]}$$

θ_t : represents common time effects. In order to control both individual effects and common time effects, Equation 3 can be applied for accessing market size.

$$y_{it} = \beta' X_{it} + u_i + \theta_t + \epsilon_{it} \quad \text{[Eq. 3]}$$

In this manner, when the panel material considers all of the individual effects and time effects a problem arises—depending on the sample size. There is no problem if the number of samples is very large. This is because usually the sample size of the R&D feasibility is not large. Panel data combines the size of the cross-sectional sample N and the size of time series sample T. In most micro panels, if N is large while T is small u_i could have a problem. In macro panels, if T is large while N is small, θ_t could have a problem. When analyzing the panel model of the R&D feasibility study we must closely analyze the individual effects rather than the time specific effects. Individual effects are the effects which are not observed, like culture, personality, abilities and so on. In labor pattern analysis, which is widely used in panel models, more than 30 years of data may be used, while in the case of R&D projects, a very small sample of the time is available. Depending on the relevance of unobserved intrinsic effects u_i and observed effect X_{it} it is assume to follow either the fixed effects or the random effects. In panel data, two-dimensional relationship or variation of independent variables are needs to be considered. Between-groups relationship appears when the independent variable, i is changed. Within-group variation is related to the change over the same independent variable i during the time variable, t. That means it is necessary to consider the effect caused by within-groups relationship in comparison to the between-groups relationship between independent variables—which represent market size—such as theses' numbers or citation numbers, patents' published numbers or citation numbers, public web site documents' exposed numbers or citation numbers and social networks' exposed numbers or citation numbers. Also, these independent variables variation depending on period of time should be considered. In other words, having the same relationship between the between-group and within-group relations are called random effects, and having the different relationship between them called the fixed effects. This process analyzes how the dependent variable could be influenced by independent variables with respects to the various groups (if the coefficient of the independent variable is positive: the relationship is positive. If the coefficient of the independent variables is negative: the relationship is negative) and how the dependent variables accurately describes the independent variables. To deal with these analyses, we should use appropriate quantitative model computer packages like E-views, STATA, and Gauss, which are mainly used as quantitative analysis tools for economic quantitative models. In this research we perform analysis using STATA, the most commonly used panel model analysis tool.

3) Phase 3: Measurement of TIIB using real-time data extraction techniques

Qualitative indicators including both current and previous interests are extracted from public web sites and social networks based on the Keyword searches of related science technology using a data extraction tool for real-time web data extraction. For derivation of quality indicators, first, we should find how much the citation and registration numbers of theses, patents, website documents, and social network influence technology impact index. That means each variable and the weight of each variable should be drawn up. The weight of each variable is drawn from the coefficient of each variable of panel model, which was measured in the second process. Second, after extracting the weight of each variable, equation 4 is used to measure the TIIB. The TIIB enables us to check interest in technology trends by reading the yearly changes from the reference year—set as 100.

$$TIIB = 100 + \alpha\frac{A_t - A_{t-1}}{A_{t-1}} + \beta\frac{B_t - B_{t-1}}{B_{t-1}} + \gamma\frac{C_t - C_{t-1}}{C_{t-1}} + \delta\frac{D_t - D_{t-1}}{D_{t-1}} \quad \text{[Eq. 4]}$$

α, β, γ and δ: the weight on the change of each variable

A: the theses number and its citation numbers

B: the number of patents and their citation numbers

C: official documents number and their citation numbers

D: the number of citations in social network and their numbers

The data of A, B, C and D should be acquired to calculate TIIB. To acquire the newest data and sufficient data group, big data technique is applied. α, β, γ and δ are drawn by applying the panel model to STATA program. The panel model can compensate of the shortcomings of time-series analysis caused by insufficient data according to the nature of R&D program. The data of A, B, C and D are used as independent variables to draw α, β, γ and δ from the panel model. And the market scale is used as the dependent variable influenced from such independent variables. Also, groups are set and analyzed to minimize the data error in the panel model.

4) Phase 4: comparison between CII and the Technology Impact Index based on Big-data (TIIB)

Economical and technological evaluations of the feasibility analysis are preceded based on the comparison between the Current Impact Index (CII) and the Technology Impact Index based on Big-data (TIIB). If the TIIB is greater than the CII under the technical basis, we must judge the technology as one having promising technical impact on future business. While if the CII is greater than the TIIB, we can determine that the technology has unpromising technical impact prospects on future business. If the TIIB continues to grow based on the reference year set as 100, we can conclude that the impact of the technology has increased continuously.

5) phase5: The quantification example utilizing big data techniques

For enhancing the understanding of CII and TIIB measurement for feasibility study, we analyzed the quantification example of “Wearable Smart Devices" in R&D projects. Generally, the low CII means low impact on technology. However, after analysis of TIIB to reflect the latest technology trends, even if the CII is low, the results determine it has influence on the related technology. In this paper, in order to deal with this problem, we focused on developing a new index which reflects the newest technology trends and shows technology impact. In order to extract the newly developed index the first step is extracting data. To do this, we should define the dependent variable and then extract independent variables. The dependent variable is the market size of wearable smart devices. The independent variables are citation and registration numbers of theses, patents, and public website documents. After collecting these variables, we can analyze the influence between dependent variable and independent variables using panel model. To execute TIIB we should draw up each variable and the weight of each variable from the panel data models. Panel model analysis is applied to measure the impact of core technology in the wearable market. The dependent variable, the market size of wearable technology from 2009 to 2017 is as follows (See Table 1.) The independent variables which were extracted above such as the citation and registration numbers of theses, patents, and website documents are used.

Table 1. Market Size: Wearable Technology from 2009 to 2014

Year	2009	2010	2011	2012	2013	2014
Market Value (in million U.S. dollars)	2008.3	2691.2	3606.1	4832.2	6475.2	8676.8

* Source: World Market for Wearable Technology, IMS Research 2012, August 2012 [2]

To accurately analyze panel models, we define four groups with respect to the search keywords and data types. For instance, Group1 and Group2 respectively present the number of search results for keyword1 and keyword2. Group3 and Group4 respectively present the number of citations for keyword1 and keyword2.The recent six year dataset is used. The dataset is from 2009 to 2014.

Table 2. Wearable Technology Dataset (Group 1~4)

Year	Group1				Group2			
	Market scale	Theses	Patents	Website documents	Market scale	Theses	Patents	Website documents
2009	2008.3	63	232	49	2008.3	1	52	24
2010	2691.2	65	251	87	2691.2	0	60	43
2011	3606.1	63	363	99	3606.1	0	64	41
2012	4832.2	59	357	196	4832.2	0	83	45

2013	6475.2	74	232	577	6475.2	2	84	216
2014	8676.8	85	7	1743	8676.8	4	3	1088

Year	**Group3**				**Group4**			
	Market scale	**Theses**	**Patents**	**Website documents**	**Market scale**	**Theses**	**Patents**	**Website documents**
2009	2008.3	80	386	56	2008.3	5	86	26
2010	2691.2	85	204	96	2691.2	0	35	44
2011	3606.1	71	191	106	3606.1	0	17	41
2012	4832.2	65	114	220	4832.2	0	3	45
2013	6475.2	70	3	741	6475.2	4	0	249
2014	8676.8	45	0	2488	8676.8	2	0	1394

We can reveal the affect the independent variable will have on the dependent variable by analyzing panel model.

For this analysis, it is important to align the data units of dependent and independent variables. In this study, the panel model evaluation is proceeded by comparing the proportions—where the denominator is the total of the recent six years, while the numerator is a year. Results of the analysis are shown in Table 3 which describes the wearable core technology's impact on market size according to the two scenarios.

Table 3. Market Size Impact of Wearable Core Technology

	Group 1&3	**Group 2&4**
Theses	0.0059	0.0819
Patents	-0.1851	-0.1801
Website documents	0.2528	0.1715

Scenario Group 1&3 presents the impact of common technology. Scenario Group 2&4 presents the impact of core technology. The reason to compare the different scenarios at the same time is because data can be missed when extracting the data using the core technology keywords in the data extraction phase. That is to say, for analyzing variables which can effect market size, the common technology keywords also need to be considered when extracting the data. In addition, the reason for analyzing two types of data groups at the same time is that a lot of errors can occur if we use only one group when analyzing the panel model. It was analyzed that the negative relationship is shown between the theses and patents data and the market size in the wearable common technology and between the theses data and the market size in the wearable core technology. However, in the wearable core technology, there exists positive relationship between the patents and website documents data and the market size. Thus the number of patents and website documents and their citation numbers have more possibility to have positive relationship than that of these numbers. In case of theses—which present more academic-like information, it does not have direct influence on market size. However, the publication of patents and the descriptions of technology development on website documents have positive influence on the market size. By utilizing the extracted data shown earlier in Table 2, Equation 5 is used to measure the TIIB. The weight of the change of each variable is drawn from the coefficient of each independent variable, which was measured in the panel model analysis.

$$TIIB = 100 + \alpha\frac{A_t - A_{t-1}}{A_{t-1}} + \beta\frac{B_t - B_{t-1}}{B_{t-1}} + \gamma\frac{C_t - C_{t-1}}{C_{t-1}} \quad \text{[Eq. 5]}$$

α, β, γ and δ: the weight on the change of each variable

A: the theses number and its citation numbers

B: the number of patents and their citation numbers

C: the number of official website documents and their citation numbers

Table 4. Annual Trends of the TIIB

	2009	2010	2011	2012	2013	2014
Group1&3	100	100.24	99.99	100.29	100.64	100.74
Group2&4	100	100.18	99.97	100.20	100.48	100.55

The Table 4 displays the annual trends of the TIIB from 2009 to 2014—the two scenarios (the search numbers and citation numbers), and dataset of the core and common keywords of wearable technology was used. The TIIB indexes were found to have little change during six years—2009 as the reference year set as 100. Overall, the interest trend of wearable technology has steadily increased in accordance with the newly developed TIIB indexes. In the case of common technology, the interest trend for the most recent six years has increased approximately 0.07% and in the case of core technology, it has increased approximately 0.02%. Whereas the CII index founded 0.4 implicates low technological value or economic value of wearable technology. Even though the CII index is low, we cannot conclude wearable technology is negatively related to market scale because of the limitation of patent data extraction—there is no choice but to use patent data which is missing a time gap of one or two years. This problem has been alleviated by utilizing the newly developed TIIB indexes and big data technology. Further exploration is necessary to develop an in-depth searching crawler and a way of resolving data inconsistencies and integrating multiple databases in order to deliver valuable and factual insight for actual research in feasibility study.

5. Conclusions

Currently, patents and theses are utilized to analyze the technology trend for the feasibility study. However, before the actual opening or publishing of a patent or thesis, a time gap of 1~2 years is required. So such material fails to reflect the newest data. Also, collecting qualitative data is time and resource consuming and the manpower for the feasibility studies are limited. In this study, the web crawler based on Google Web Search API is developed to enable the collection and analysis of the newest and most credible data from the Internet. This is achieved through use of big data technology and the newly developed TIIB indexes which reflect the technology impact of the newest trends proposed. By drawing the TIIB of recently emerging technology, the technology impact according to the correlation of the relevant market scale is analyzed.

Acknowledgements

This research was supported by the research program of KISTEP (Korea Institute of S&T Evaluation and Planning).

References

[1] "Korea Institute of S&T Evaluation and Planning (2011) A Study on Standard Guidelines for Preliminary Feasibility Study on R&D Program", http://mosf.go.kr/_upload/bbs/76/attach/20120221181135751.pdf , 1th edition, pp. 83-114.

[2] "A World Market for Wearable Technology", IMS Research, **(2012)**.
[3] W. K. Sung, "Research on analysis and application platform technologies for science and technology big data", Ministry of Education and Science Technology, **(2013)**, pp. 19-32.
[4] Y. K. Nam, "Development of Big data Analysis Model focused on National R&D Information", Korea Institute of Science and Technology Information, **(2013)**, pp. 81-86.

