ICIC Express Letters
Volume **12**, Number **7**, July **2018**

ICIC International ©2018 ISSN 1881-803X
pp. **699–706**

A STUDY ON THE TECHNOLOGY CONVERGENCE OF ARTIFICIAL INTELLIGENCE THROUGH PATENT ANALYSIS

HoYeon Oh, Hong Joo Lee* and Tai-Woo Chang

Department of Industrial and Management Engineering
Kyonggi University
154-42, Gwanggyosan-ro, Yeongtong-gu, Suwon-si, Gyeonggi-do 16227, Korea
*Corresponding author: blue1024@kgu.ac.kr

Received January 2018; accepted April 2018

Abstract. *Since the patents contain original information on technology, it can investigate changes in technology and the convergence phenomenon with co-classification analysis of IPC (International Patent Classification) code indicating the technology group of the patents. In this research, we analyzed the changes of technological convergence for artificial intelligence (AI) technology, which was recently studied actively. We collected Korean patent documents using KIPRIS (Korea Intellectual Property Rights Information Service) and analyzed social networks of IPC codes of the patents. Additionally, the convergence types could be reviewed using network centrality and clique analysis. AI technology mainly focuses on data processing systems (G06F, G06Q). Although there were many technologies that utilized telecommunication services (H04) in the past, technologies related to computing, calculating and counting (G06) have been mainly used in recent years. We also could find that there was a change in the technology area where AI mainly focused.*

Keywords: Patent information analysis, Social network analysis (SNA), Artificial intelligence (AI)

1. **Introduction.** To develop new products and services, various technologies are integrated. Thus, grasping the state of technology convergence can help acquire the core technology and establish the strategy. One way to studying technological convergence is to use patent information. Recently, artificial intelligence (AI) technologies are developing quickly compared to the past, because the computing environment is also developing quickly. However, there is little research on the convergence of AI. Fujii and Managi studied the decomposition framework to clarify the determinants of AI technology [3]. Kim et al. only studied machine learning patents of AI [5]. Kim et al. analyzed the US class at category level [6]. Moreover, Bae and Shin conducted comparative studies for each country rather than Korea [18]. Therefore, it is necessary to review the AI convergence.

This research aims to investigate how the AI technology has changed with the International Patent Classification (IPC) code of patents in Korea. This research determines the convergence group that are chiefly shown by year, and analyzes its characteristics. We studied co-classification of IPC codes using social network analysis and collected patents using Korea Intellectual Property Rights Information Service (KIPRIS).

This research is structured as follows: in Section 2, we examine previous studies on patent information analysis and network analysis; Section 3 provides an overview of the data collection; Section 4 presents the analysis results; Section 5 shows the explanations of the conclusions and limitations.

DOI: 10.24507/icicel.12.07.699

2. Literature Review.

2.1. **Patent information analysis.** In patent, technology convergence studies are conducting using co-citation, co-classification and semantic analysis. Among them, we utilized co-classification. Co-classification grasps the relevance of pairs occurring in one document. For example, the IPC code is used in patents. The technologies belonging to one patent are related to each other [4]. The IPC code represents the technical area of the patents and it is an identifier of technology that is clustered into similar attributes. If different IPC codes are associated, it can be a convergence between heterogeneous technical areas [8]. Thus, the flow of technology can be estimated and analyzed using IPC codes [7]. Patents are categorized into classes with hierarchical structure according to functions and purposes. The IPC code has a classification system of "Section → Class → Sub Class → Main Group → Sub Group". The IPC code gives one main classification when a single technology is related to the invention. However, the information is divided into a main category that represents the invention and a sub category that includes the others, when there are two or more technical contents. Most patents match to many classes, and co-classification analysis utilizes class information that is categorized together [9]. Tijssen analyzed interdisciplinary convergence through co-classification analysis of the energy field in the Netherlands [15]. Suzuki et al. used the IPC code to divide the integration of technology in the same and heterogeneous fields for the R&D status of a company, and they were judged to be a technology convergence through co-classification analysis [11]. Curran and Leker studied the convergence technology by matching the IPC code with the industry [2]. We studied the technology convergence using IPC codes based on the premise [14] that the relationship between pairs is closer as the number of occurrences simultaneously increases.

2.2. **Network analysis.** Network analysis analyzes the level of union, the strength of relationship, and the direction of flow among documents, people, journals, groups, organizations, regions and countries. In addition, it can identify the flow of knowledge [1], and clarify the interrelationship or characteristics among them [17]. Network centrality indicates the power and influence of nodes in network areas [16]. Degree centrality, closeness centrality, betweenness centrality, and eigenvector centrality are mainly used as central indicators. Kim et al. showed changes in machine learning patents using keyword network analysis [5]. Kim et al. analyzed the US class of patents and papers, and studied a convergence map of AI technology [6]. Bae and Shin analyzed the social network with the IPC code using patents in each country, and showed the state of AI technology, the convergence pattern and cooperation pattern by country [18]. In this paper, to analyze about co-classification, we studied about theory of network analysis.

3. Research Process.

In this research, we investigated the flow of AI technology in Korea, and our target is limited to patent applications in Korea. The process of research is shown in Figure 1.

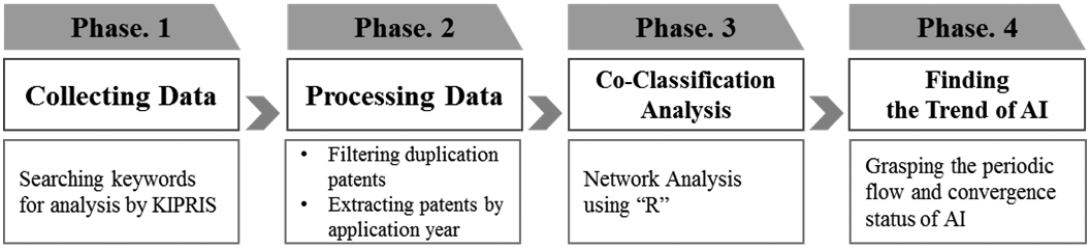


FIGURE 1. Research process

3.1. **Data collection and analysis.** We retrieved patent documents from KIPRIS using the keyword equation shown below. We added the keywords, machine learning and deep learning, in order that we concretely search the AI patents in Korea. We found 1,463 patent documents that were registered. There remain 1,192 pieces of data after filtering by years from 2005 to 2016.

$$Title = Artificial\ Intelligence + Machine\ Learning + Deep\ Learning$$

However, this study excluded the patents that only had one IPC code because there is not sufficient explanatory power in terms of technology convergence. Thus, 486 pieces of patent data with two or more IPC codes were analyzed using SNA. We also analyzed the subclass of IPC code for co-classification analysis because if the data is analyzed up to the group classification, the density of the network is too low to be tied to nodes. In this research, we used the statistical package R.

3.2. **Social network analysis.**

3.2.1. *Centrality analysis.* Centrality was considered from this research point of view. Degree centrality is connected with many other IPC codes, closeness centrality is close to other IPC codes in the whole network, betweenness centrality makes many connections between other IPC codes, and eigenvector centrality has great influence on the network. Generally, eigenvector centrality result rather than other centrality is more appropriate for identifying important nodes [13]. Therefore, we can determine the influential technologies in technology convergence as a result of eigenvector centrality. This research grasps the characteristics of AI via the meaning of each centrality.

3.2.2. *Clique analysis.* Social networks are not composed of one group. Several subgroups are composing them. A clique is a subgroup composed of three or more nodes [12]. All nodes in a clique are directly connected to each other, and none of the other nodes in the network have a direct connection relationship with the nodes in clique. This characteristic is used to find a community with strong connections in the network. Clique analysis can identify the set representing a specific technology [5]. In this research, we investigate the convergence of AI technology using the clique analysis.

4. **Result of Network Analysis.**

4.1. **Basic statistics.** In this research, we analyzed data from 12 years (2005-2016), and classified it into four periods (Q1, Q2, Q3, Q4) of three years each by applied year as shown in Table 1. Network analysis was conducted on the number of patents that included more than two IPC codes. The number of patent applications increased overall, but decreased during the period of Q4. Almost all registered patents in Q4 (2014-2016) were not included, due to the term that patents are approved and registered.

The characteristics of the network are density and transitivity. Density is the ratio of the number of nodes connected to the total number of possible connections between the nodes in the network. Transitivity can be shown as "a→b and b→c, then a→c", and the probability that a friend of a friend becomes my friend [12]. Density decreased over time. This means that the number of connected nodes increases and various IPC codes are

Table 1. Summary of analysis group

		Q1 (2005-2007)	Q2 (2008-2010)	Q3 (2011-2013)	Q4 (2014-2016)
Number of Patents	Total	172	317	385	322
	Networked	62	131	159	134
Network Characteristics	Density	11.8%	7.3%	6.1%	9.8%
	Transitivity	28.5%	30.7%	26.8%	37.0%

generated. Additionally, four periods were displayed similarly for transitivity especially Q4, which was high. This shows that various technologies related AI are more connected than before. Overall, various codes are displayed over time and the range of the network is expanding.

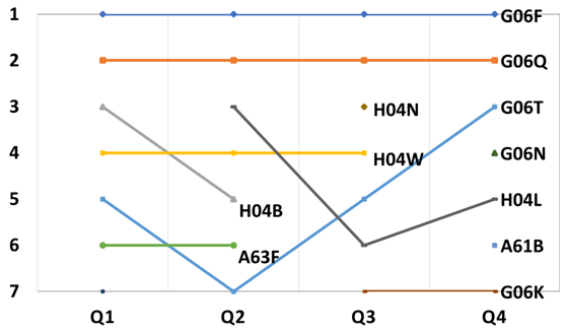
The frequency analysis in Table 2 was analyzed using a total of 1,192 pieces of data. Additionally, having a high frequency means higher utilization. Most of the patents include G06F (Electric Digital Data Processing) and G06Q (Data Processing Systems or Methods). These are the most important technologies in AI. Both codes relate to data processing systems and serve as a platform for technology. G06T (Image Data Processing) decreased for a while in Q2, but it increased thereafter. G06N (Computer Systems-based on Specific Computational Models) increased sharply in Q4. A63F (Card, Board or Roulette Games) only ranked in Q1 and Q2. H04W (Wireless Communication Network) appeared throughout, but it decreased to 10th in Q4.

TABLE 2. Result of frequency analysis

(a)

Rank	Q1	Q2	Q3	Q4
1	G06F	G06F	G06F	G06F
2	G06Q	G06Q	G06Q	G06Q
3	H04B	H04L	H04N	G06T
4	H04W	H04W	H04W	G06N
5	G06T	H04B	G06T	H04L
6	A63F	A63F	H04L	A61B
7	F24F	G06T	G06K	G06K

(b)



4.2. **Centrality analysis.** This research investigated the flow of IPC codes using centrality analysis and changes in the AI technology.

Table 3 shows the results of the degree centrality analysis, and it has many connections with other technologies. G06F has the largest number of connections. G06N and G06K (Recognition of Data) are higher in Q4, which means more connections with other technologies. H04N (Pictorial Communication) is connected in Q3, but not in most recent time period (Q4). G06T and H04W showed decreased connectivity.

Table 4 shows the closeness centrality results. A high closeness centrality means that it is easy to access for convergence with other technologies. In Q2 and Q3, the G06Q rank decreased. This means that it was difficult to connect via G06Q. Additionally, G06T recently decreased in Q4 in terms of accessibility. G08G (Traffic Control Systems) and

TABLE 3. Technological changes in the degree centrality rank by year

(a)

Rank	Q1	Q2	Q3	Q4
1	G06F	G06F	G06F	G06F
2	G06Q	G06Q	G06Q	G06Q
3	H04W	H04L	H04N	G06N
4	A63F	H04W	G06T	G06K
5	H04N	G06N	G01N	A61B
6	G06K	H04B	G08G	H04W
7	G06T	G06T	H04W	G06T

(b)

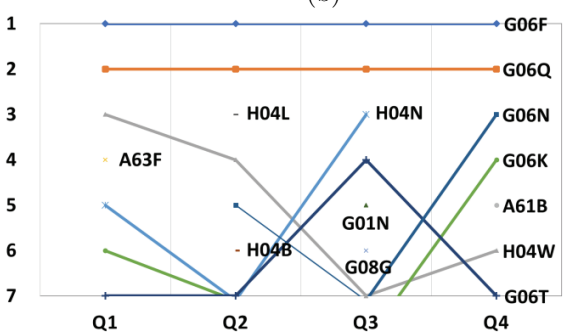


TABLE 4. Technological changes in the closeness centrality rank by year

(a)

Rank	Q1	Q2	Q3	Q4
1	G06Q	G06F	G06F	G06F
2	G06F	G06Q	H04N	G06Q
3	H04N	G06T	G08G	G06N
4	G06T	H04W	G06T	A61B
5	G06K	H04L	G06K	G06K
6	H04L	G01N	G06Q	G06T
7	A63F	H04Q	G01N	G01S

(b)

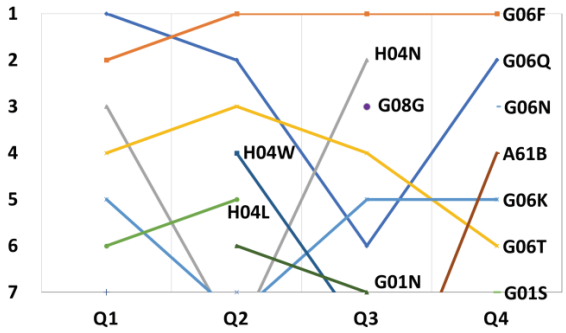


TABLE 5. Technological changes in the betweenness centrality rank by year

(a)

Rank	Q1	Q2	Q3	Q4
1	G06F	G06F	G06F	G06F
2	G06Q	G01N	G01N	G08B
3	H04N	H01L	H04N	A61B
4	G06K	G06T	B25J	G06Q
5	H01L	G06Q	G06T	G06N
6	G08B	A61K	G06K	H04N
7	B82Y	H04L	C12N	G01S

(b)

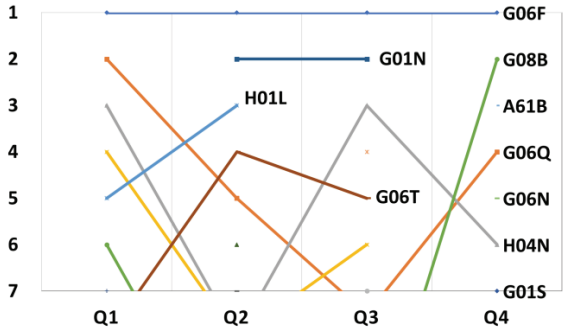
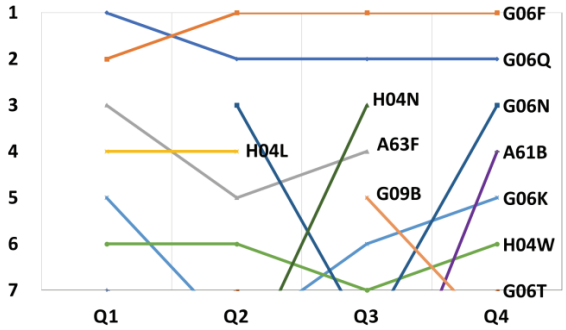


TABLE 6. Technological changes in the eigenvector centrality rank by year

(a)

Rank	Q1	Q2	Q3	Q4
1	G06Q	G06F	G06F	G06F
2	G06F	G06Q	G06Q	G06Q
3	A63F	G06N	H04N	G06N
4	H04L	H04L	A63F	A61B
5	G06K	A63F	G09B	G06K
6	H04W	H04W	G06K	H04W
7	H04B	G06T	H04W	G06T

(b)



H04N were new in Q3, and A61B (Diagnosis; Surgery; Identification) and G06N were new in Q4.

The higher the betweenness centrality is in Table 5, the more significant the interaction with other technologies is. Although the ranking of G06Q decreased, it increased in Q4, but it was not high compared with the others. G01N (Investigating or Analyzing Materials) also served as a mediator during Q2 and Q3. This technology is mainly used in medicines. G08B (Signaling or Calling Systems) and A61B increased sharply in Q4.

Table 6 shows the result of the eigenvector centrality analysis. It had a high impact on other technologies in the network. The most powerful technologies in Q4 are G06N and A61B excluding G06F and G06Q. H04W, G06T and G06K which are still emerging are consistently used in AI. Additionally, H04L (Transmission of Digital Information) until Q2 and A63F until Q3 had a high ripple power in AI.

The results are summarized as follows. G06F is the most important technology in the convergence of AI and other technologies. Additionally, G06Q is distributed in the upper part, but its relationship with various technologies is small. However, it has in the upper rank and it can be a major technology in AI. The mainly influential technologies have changed each time. The H04 (Electric Communication Technique) series was in the upper part in the past. However, now, it appeared in the lower part. Recently, the AI patents are networked around G06N (Computer Systems-based on Specific Computational Models). It means more influential in AI field than in the past. Also, A61 (Medical or Veterinary Science) series has more actively appeared.

4.3. **Clique analysis.** To understand the technology convergence of AI patents by year, a clique analysis of the network was performed. This research includes clique sizes of more than four. Because of the large size of the network, cliques with three or fewer nodes are less cohesive. Additionally, because G06F and G06Q appear often, these did not have a large influence on the results.

Table 7 shows the results of a clique analysis by year. The size of the maximum clique was 4 in Q1, 5 in each of Q2 and Q3, and 7 in Q4. For interpretation of the results, we examined the patents each year.

Table 7. Clique analysis results

	Clique Size	IPC code
Q1	4(2)	(A63F, G06F, G06Q, H04L) (G06F, G06K, G06Q, G06T)
Q2	4(8)	(G01B, G01N, G06F, H01L) (F24F, G06T, H02P, H04N) (B63B, G01C, G06Q, G08G) (G06F, H04H, H04N, H04W) (G06F, G06Q, H04N, H04W) (G06F, G06Q, G06T, H04N) (A63F, G06F, G06N, G06Q) (A61K, A61P, C07K, C12N)
	5(3)	(E02B, G06Q, G08B, G08C, H04Q) (G06F, G06Q, H04L, H04Q, H04W) (G06F, H04B, H04L, H04Q, H04W)
Q3	4(11)	(G01C, G06F, G06T, G08G) (B63H, G01C, G06F, G08G) (B63H, G01C, G05D, G08G) (B63B, G06F, G06Q, G08G) (B63B, B63J, G06Q, G08G) (B60C, G01M, G06F, G06N) (A61K, A61P, C07K, C12N) (G06F, H04B, H04L, H04W) (H04B, H04J, H04L, H04W) (G06K, G08B, G08G, H04N) (G06F, G06K, G06Q, G09B)
	5(2)	(G06F, G06K, G06T, G08G, H04N) (G06F, G06K, G06Q, G08G, H04N)
Q4	4(12)	(B64C, B64D, G06Q, H04N) (G01R, G06F, G06N, G06Q) (B65G, G06F, G06K, H04W) (B60R, G06T, G08G, H04N) (B60R, G06F, G06T, G08G) (B60R, G05D, G06T, G08G) (A61L, F21V, G08B, H05B) (G06F, H01H, H01R, H02J) (A61B, G06F, G06Q, G09B) (A61B, G06F, G06N, G06Q) (A61B, G02B, G06F, G06Q) (B25J, B64C, B64D, F03D)
	5(5)	(A61L, F21S, F21V, F21Y, H05B) (G01S, G06F, G06N, G06Q, G08G) (G01S, G06F, G06N, G06T, G08G) (A61B, A61H, A63B, G06F, G09B) (A61B, A61F, G01C, G06F, G06Q)
	6(4)	(G01C, G01S, G06F, G06K, G06T, H04W) (G06F, G06K, G06N, G06Q, H04L, H04W) (G01S, G06F, G06K, G06N, G06T, H04W) (G01S, G06F, G06K, G06N, G06Q, H04W)
	7(1)	(A61F, G01C, G01S, G06F, G06K, G06Q, H04W)

The areas strongly connected with G06F or G06Q are shown in the form of a game (Q1) → telecommunication (Q2, Q3) → medical (Q4). The H04 series mainly appeared in Q2 and Q3. The details are that this is a control system using a wireless network such as for a ship, waterway, crime prevention, or an unmanned system. In Q4, there are technologies from A61B, A61F (Invalid Operated Apparatus of Devices), A61H (Physical Therapy Apparatus), A63B (Training Equipment) and G09B (Educational or Demonstration Appliances). There are devices associated with a therapeutic robot and healthcare. In addition, G01S (Determining Distance of Velocity by Use of Radio Waves) in Q4 indicates a route search using wireless communication.

5. **Conclusion.** We investigated the trend of technology convergence through the analysis of patents related to AI. Summarizing the centrality analysis, the communication service of the H04 series had a high centrality in the past, but the arithmetic and logic operation of the G06 series were recently high. It has become technology changes from the control and transmission systems to the calculation model by the electric device were used. In the past, AI technology has been used in control systems, but now it is used more in knowledge processing systems. We also identified the group that had a strong cohesion using clique analysis. Over time, the technology area utilizing AI is changing (game to telecommunication to medicine). This study analyzed the convergent network of AI. We can show the influential technologies for each period. In Korea, it usually takes two years for the technology to be commercialized. Therefore, the applied technology can be said to be a technology to be spotlighted after about two years. The results showed that the technology is mainly used two years later. This can predict business trends in the future.

Limitations are that data collection target was limited to Korean patents, and that this analysis does not include the latest data because of length of the patent registration period. In addition, the interpretation of IPC codes was made through the patent documents for each year. Therefore, we need to compare actual events in each period to make a more valid interpretation.

Acknowledgements. This work was supported by the GRRC program of Gyeonggi province. [(GRRCKGU2017-B01), Research on Industrial Big-Data Analytics for Intelligent Manufacturing].

REFERENCES

[1] C. L. Borgman and J. Furner, *Scholarly Communication and Bibliometrics*, 2002.
[2] C. S. Curran and J. Leker, Patent indicators for monitoring convergence – Examples from NFF and ICT, *Technological Forecasting and Social Change*, vol.78, no.2, pp.256-273, 2011.
[3] H. Fujii and S. Managi, Trends and priority shifts in artificial intelligence technology invention: A global patent analysis, *Research Institute of Economy, Trade and Industry (RIETI)*, 2017.
[4] H. Grupp, Spillover effects and the science base of innovations reconsidered: An empirical approach, *Journal of Evolutionary Economics*, vol.6, no.2, pp.175-197, 1996.
[5] H. H. Kim, D. G. Kim and J. N. Jo, Patent data analysis using clique analysis in a keyword network, *Journal of the Korean Date of Information Science Society*, vol.27, no.5, pp.1273-1284, 2016.
[6] H. W. Kim, K. R. Noh, S. J. Ahn and O. J. Kwon, Technology convergence map creation and country profile analysis in the field of artificial intelligence, *The Journal of the Korea Institute of Electronic Communication Sciences*, vol.12, no.1, pp.139-145, 2017.
[7] J. H. Han, J. G. Na and C. B. Kim, Network analysis of technology convergence on decentralized energy by using patent information: Focused on Daegu city area, *Journal of the Society of Korea Industrial and Systems Engineering*, vol.39, no.3, pp.156-169, 2016.
[8] J. H. Kang, J. C. Kim, J. H. Lee, S. S. Park and D. S. Jang, A patent trend analysis for technological convergence of IoT and wearables, *Journal of Korean Institute of Intelligent Systems*, vol.25, no.3, pp.306-311, 2015.

[9] J. H. Yoon and Y. J. Geum, Analyzing dynamic patterns of technology convergence using patent co-classification analysis: A case of healthcare service, *The Korean Institute of Industrial Engineers Proceedings of Academic Conference*, vol.2016, no.11, pp.383-391, 2016.
[10] J. R. Lee, *A Study of Convergence Technology Classification System and Features by Using the Patent Data*, Master Thesis, Hanyang University, 2015.
[11] K. Suzuki, J. Sakata and J. Hosoya, Innovation position: A quantitative analysis to evaluate the efficiency of research and development on the basis of patent data, *Proc. of the 41st Annual Hawaii International Conference on System Sciences*, 2008.
[12] M. H. Huh, *Introduction to Social Network Analysis Using R*, JaYu Academy, 2014.
[13] M. J. Lee, Y. D. Kim and W. C. Jang, Patent citation network analysis, *The Korean Journal of Applied Statistics*, vol.29, no.4, pp.613-625, 2016.
[14] Q. He, Knowledge discovery through co-word analysis, *Library Trends*, vol.48, no.1, pp.133-159, 1999.
[15] R. J. Tijssen, A quantitative assessment of interdisciplinary structures in science and technology: Co-classification analysis of energy research, *Research Policy*, vol.21, no.1, pp.27-44, 1992.
[16] S. S. Lee, A preliminary study on the co-author network analysis of Korean library and information science research community, *Journal of Korea Library and Information Science Society*, vol.41, no.2, pp.297-315, 2010.
[17] S. Y. Kim, *A Comparative Study of Research Trend in the Field of Big Data Using Co-Word Analysis*, Master Thesis, Ewha University, 2016.
[18] Y. I. Bae and H. R. Shin, *A Study on Convergence and Cooperation Patterns of Artificial Intelligence Technology*, GRI, 2016.