

Examination on the Trend of Researches on Absorptive Capacity using Text Mining Technique

Eun-MiPark¹, Joung-Hae Seo^{*2}

^{1,*2}*School of Business Administration, Kyungpook National University, 80 Daehakro, Bukgu, Daegu 41566, South Korea*
Issack38317@naver.com¹, knight6407@naver.com^{*2}

Article Info

Volume 81

Page Number: 448 - 457

Publication Issue:

November-December 2019

Abstract

Background/Objectives: This paper is investigated the trend of researches on absorptive capacity through search of Scopus Database by using text mining technique for academic theses related to absorptive capacity.

Methods/Statistical analysis: This study used Text Mining Technique, discovering hidden meanings in big data such as extraction of information from large volume of documents or finding linkages, classification, clustering and summarization by using the method of NLP.

Findings: As a result of analysis on the keywords related to absorptive capacity, most frequently used words were 'capacity' followed by 'absorptive', 'knowledge', 'firm', 'innovation', 'study', 'performance', 'research', 'results', 'model', 'effect', 'development' and 'technology' in descending order. Absorptive capacity is the very core of current convergence era, many studies are expected to be conducted related to absorptive capacity in the digital transformation era.

Improvements/Applications: The trend of researches on absorptive capacity and related studies in the leading academic thesis database, Scopus, and drew out and analyzed 20 keywords. In the future, related researches will be continuously conducted as well.

Keywords: Big Data, Absorptive Capacity, Text mining, Trend, Scopus, R.

Article History

Article Received: 3 January 2019

Revised: 25 March 2019

Accepted: 28 July 2019

Publication: 22 November 2019

1. Introduction

In the wake of the global financial crisis, Korean economy has also been declining, which is manifested in various domestic indicators of consumption, investment and employment. Accordingly, companies have been making a variety of efforts to fight against deteriorating management.

Companies are using technological innovation as a way to survive in an uncertain and unstable business environment, which increases the importance of technological innovation day by day[1,2]. Diverse attempts are being made by companies to enhance their performances of technological innovations[3]. Companies need to use the

capabilities they own and to effectively utilize external knowledge. In order to effectively utilize external knowledge internally, companies need absorptive capacity.

In current era of the 4th Industrial Revolution in which technology is rapidly advancing, the companies require high level of absorptive capacity to utilize new external knowledge within the companies since they cannot develop and acquire all of the necessary technologies by themselves.

The more absorptive and innovative capacity a company possesses which are difficult to reproduce or imitate, the higher the innovative performance of the company and, the higher the absorptive capacity and learning mechanism of the company, the more likely the company is to pursue innovation. In the long run, such a company will secure a competitive advantage in the market.

In addition, a company which carries out technological innovations uses both internal and external knowledge and ideas to better their technology than before and such a company utilizes various forms of internal and external resources in diverse ways regardless of whether internal or external to develop new products and services[4].

As a result, there is ever-growing interest in absorptive capacity in various academic fields as well. This study aims to extract and analyze prior studies related to absorptive capacity from Scopus DB. That is, this study attempts to investigate the trend of research on absorptive capacity by conducting frequency analysis and network analysis on major words centering on the titles, keywords and abstracts of the theses.

The structure of this paper is organized as follows; chapter 2 enunciates theoretical background, chapter 3 deals with big data analysis methodology, chapter 4 presents the results of the study and chapter 5 suggests conclusions and implications of the study.

2. Theoretical Background

2.1. Bigdata

Recently, big data has emerged as an essential keyword in our society. Big data can be defined as information technology to extract valuable information and actively responds to or predicts changes based on knowledge generated by utilizing and analyzing large amounts of data. In addition, big data is enormous set of data which cannot be managed by existing management and analysis system, and is defined as a concept which includes all the technologies and tools (collection, storage, retrieval, sharing, analysis and visualization) related to large-scale data[5,6].

Considering such definition, while big data is enormous amount of data, it is also used as a concept covering analysis and utilization of data beyond mere quantitative meaning. Big data consists of '5 V' components, which are Volume, Variety, Velocity, Veracity and Value. Beyond calling the data over a certain volume as 'big data', it requires relative interpretation on the data to acquire desired 'big values'[5].

2.2. Text mining

Most of the materials produced in our real lives including internet materials, e-mails, theses of various areas, newspaper and magazine articles and reports on opinion polls, are in the form of text. Text mining refers to the technique of discovering hidden meanings in big data such as extraction of information from large volume of documents or finding linkages, classification, clustering and summarization by using the method of natural language processing on unstructured text data written in human languages[5].

The text mining (treatment process) is a process of discovering knowledge from various kinds of text data. In terms of knowledge discovery, the purpose of text mining is to extract high-dimensional and meaningful information or knowledge necessary for decision making by

processing unstructured, structured or semi-structured data. In general, the text mining treatment process is conducted in the order of preparation stage, pre-processing stage and knowledge extraction stage[5,6].

Preparation stage establishes and converts the data of the various text documents entered into ones appropriate for the scope of the problem. Pre-processing stage transforms the texts which are appropriately organized for the scope of the problem in the preparation stage into standardized modes of expression. Knowledge extraction stage discovers knowledge such as meaningful patterns or relationships from the standardized data which were transformed to fit the problems.

2.3.Absorptive Capacity

The ability to digest and imitate new

information and knowledge acquired from external sources varies from company to company. Calling the ability ‘Absorptive Capability’, Cohen and Levinthal (1990) defined it as an organizational ability to recognize, assimilate, transform and exploit external knowledge. Prior knowledge possessed by a company forms absorptive capacity, which in turn develops existing knowledge to enhance understanding of new things and positively promotes creation of new ideas and development of new products[7].

Knowledge produced this way in turn acts as absorptive capacity. Zahra and George (2002) argued that companies with high level of absorptive capacity can create excellent performance by taking advantage of the ‘first mover advantage’, quick response to customers, ‘lock-in effect’ or by avoiding ‘competency trap’[8,9].

Table 1: Absorptive Capacity Research

Researchers	Content
Cohen & Levinthal(1990) [7]	Factors affecting the formation of absorptive capacity (classified into the levels of individual, organization and corporation)
Mowery & Oxley(1995) [10]	Using absorptive capacity as a control variable from the perspective of technology transfer route and national innovation system
Lane & Lubatkin(1998) [11]	Using absorptive capacity as a prediction value of organizational learning in terms of affiliation
Zahra & George(2002) [8]	Classifying absorptive capacity into acquisition, digestion, transformation and utilization by specifying existing concepts
Lane et al.(2006) [12]	Classifying absorptive capacity into sequential processes of exploratory learning, transformative learning and exploitative learning
Lichtenthaler(2009) [13]	Analyzing the relationship between exploratory learning, change learning, utilization learning and performance and using absorptive capacity as a mediating variable

In other words, absorptive capacity enables companies to effectively utilize knowledge acquired from external sources and to promote better ways of doing things and generating profits by modifying practices within the organizations[14].

3. Research Methods and Procedure

Procedure of big data analysis is illustrated in Figure 1. First, study defined the problem(s) and collected data. Then, the data was refined and analyzed to visualize them. This study was conducted in the research procedure used in the prior studies.

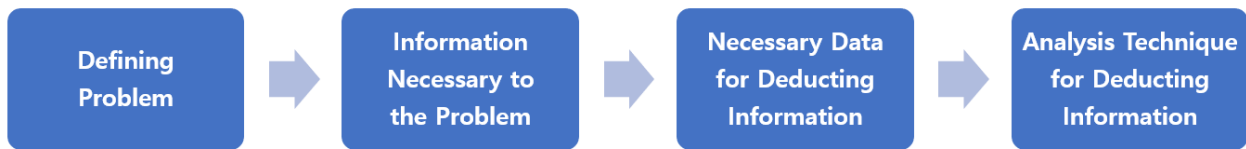


Figure 1. Research Procedure

3.1. Defining Problem

This study conducted its search for absorptive capacity within the limit of abstracts and keywords and downloaded the final data[5].

3.2. Information Necessary to the Problem

At this era when the interest in absorptive capacity increases, this study aims to identify trend of research on absorptive capacity.

3.3. Necessary Data for Deducting Information

In order to collect related data, this study collected data from the theses provided by Scopus homepage, a leading academic theses database website, by using the keyword ‘absorptive capacity’.

3.4. Analysis Technique for Deducting Information

The analysis was performed by using R, Tagxedo, etc. and the data needed for analysis were collected from www.scopus.com [5].

4. Data Collection and Analysis

4.1. Data Collecting and Preprocessing

Frequencies of the words were calculated for the collected data by using natural language processing technique[15].First, NLP package was used to analyze the collected data. Frequencies of key words were calculated by repeating data purification process. As for article types, there were 299 open access publications and 1,543 others.

Table 2: Number of Absorptive Capacity Paper

Years	Number of paper
2019	338
2018	445
2017	369
2016	363
2015	327

As shown in Table 2 and Figure 2, there were a total of 1,834 theses from 2015 through September 2019, and studies on absorptive capacity turned out to be on the increase.

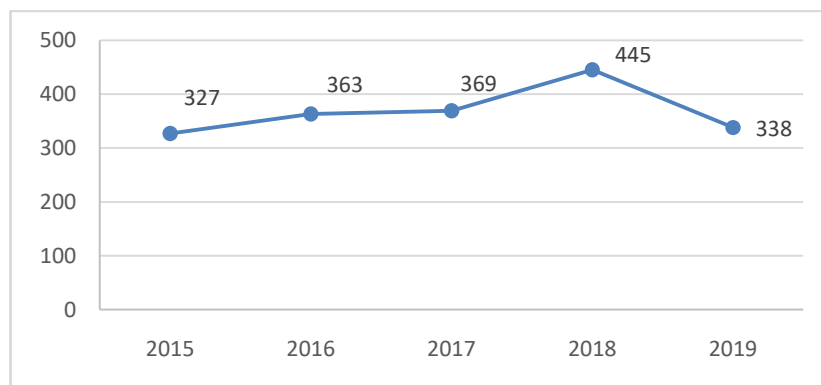


Figure 2. Trend of Absorptive Capacity Paper

4.2. Data Analysis

As of September 30, 2019, a total of 1,834 documents were retrieved. In the process of drawing out 1,834 keywords,

the search was performed by using the words ‘absorptive capacity’ and the search was limited to the keyword ‘absorptive capacity’ in Table 3.

Table 3: Absorptive Capacity Paper in Subject, Type, Keyword

Subject Area		Document Type		Keyword	
Business, Management and Accounting	949	Article	1,486	Absorptive Capacity	867
Social Sciences	347	Conference Paper	188	Innovation	300
Economics, Econometrics and Finance	307	Review	60	Knowledge Management	145
Engineering	284	Book Chapter	55	Article	97
Computer Science	230	Conference Review	8	Human	85
Decision Sciences	209				
Environmental Science	149				
Agricultural and Biological Sciences	119				
Medicine	99				
Energy	87				

As for the subject areas, studies related to ‘absorptive capacity’ are being conducted most frequently in Business

followed by Management and Accounting, Social Sciences, Economics, Econometrics and Finance and Engineering in

descending order.

As for document types, researches on absorptive capacity are being most frequently conducted in the form of articles followed by conference papers and reviews. For the keyword, researches on absorptive capacity are being most frequently conducted in the keyword ‘absorptive capacity’ followed by innovation and knowledge management.

4.3. Keyword Analysis

The first stage in keyword analysis is to extract the same nouns as the dictionary vocabulary from the raw data. Nouns were extracted based on the abstracts from the collected paper data. The results of extraction of the top 20 keywords from the collected data are presented in Table 4.

Table 4: Absorptive Capacity Keyword Analysis

	rev	Freq
1	capacity	3274
2	absorptive	3098
3	knowledge	2836
4	firm	2314
5	innovation	2276
6	study	1561
7	performance	1400
8	research	1183
9	results	891
10	model	764
11	effect	741
12	development	716
13	technology	690
14	relationship	686
15	data	656
16	external	623
17	organizational	605
18	analysis	591
19	new	587
20	role	579

Top 1st through 20th keywords were drawn out to investigate the trend of change by using keyword extraction algorithm. Top 20 keywords were drawn out except the words such as survey, noun and article which are not related to this study. The numbers of the derived keywords were 3,274 for the word ‘capacity’, 3,098 for ‘absorptive’, 2,836 for ‘knowledge’, 1,324 for ‘firm’ and 2,276 for ‘innovation’.

What is important in absorptive

capacity is ‘capacity’ and ‘absorptive’ and the results of this study also demonstrated high level of importance for ‘capacity’ and ‘absorptive’. In addition, as absorptive capacity is the sum of knowledge, ‘knowledge’ ranked the 3rd place. And because the companies are making innovations by using their absorptive capacity, the word ‘company’ and ‘innovation’ ranked the 4th and 5th place. As learning is required (6th) for companies to digest external knowledge into their own and generate good performance,

learning ranked the 7th place.

Since companies need various ‘efforts’ and ‘researches’ to be able to accommodate and utilize absorptive capacity within organization, ‘efforts’ and ‘researches’ ranked the 8th and 9th place respectively followed by ‘results’, ‘model,

‘development’, ‘technology’, ‘relation,’ ‘data’ and ‘external’ in descending order.

The results of the derived top 20 data are presented in Figure 3 by using Tagxedo. In the case of text mining, large and bold texts indicate high level of significance.



Figure 3. Keyword Analysis (Deriving 20 Nouns)

4.4. Explore Text Data

Pre-processing procedure is as follows;

1. Create DTM
2. Delete Hangul Words
3. Delete Sparse Terms
4. Delete Low TF-IDF col and row

TF-IDF indicates weighted value used in information search and text mining. That is, it means frequency of a specific word appearing in the document.

The expression to be established is as follows;

$$\frac{\text{a term count in a document}}{\text{total term count in a document}} \times \log \frac{\text{total Doc count}}{N \text{ doc (the term appears)}}$$

As shown in Figure4, This is the visualized result of the words mentioned

over 100 times in the word cloud.



Figure 4. Absorptive Capacity Paper Word Cloud

The hierarchical structure of the data expressed in tile from is called a 'Treemap'. As shown in Figure 5, it has hierarchical attribute and hierarchies have the advantage that they can be expressed in colors. In addition, words are visualized

in the size of rectangular surface area calculated according to their frequencies of appearance and, and hierarchical structure of the words can be expressed together, if any in Figure 5.

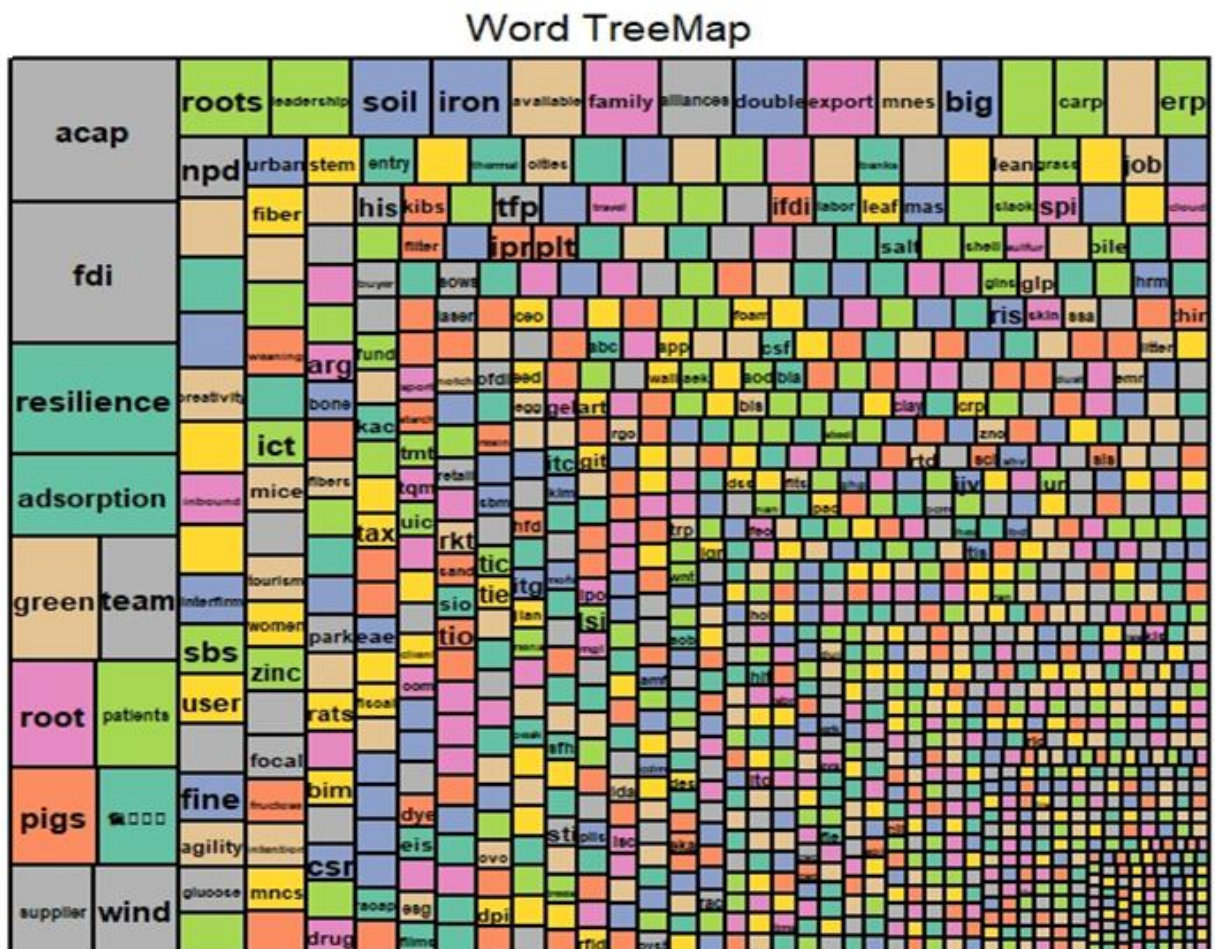


Figure 5. Absorptive Capacity Paper Word Tree Map**5. Conclusion**

This study investigated the trend of researches on absorptive capacity through search of Scopus DB by using text mining technique for academic theses related to absorptive capacity. This study collected and analyzed a total of 1,834 academic data and drew out following conclusions;

First, as a result of analysis, it turned out that researches of absorptive capacity have been continuously increasing since 2015 and are being conducted now as well in various subject areas. Since absorptive capacity is the very core of current convergence era, many studies are expected to be conducted related to absorptive capacity in the digital transformation era.

Second, as a result of analysis on the keywords related to absorptive capacity, most frequently used words were 'capacity' followed by 'absorptive', 'knowledge', 'firm' and 'innovation' in descending order. As mentioned earlier, since absorptive capacity is the ability to convert external knowledge into internal knowledge, it is expected to become even more important at this time of technological convergence and, in the future, related researches will be continuously conducted as well.

This study has significant implication in that, under the domestic situation which lacks empirical researches on absorptive capacity, it investigated the trend of researches on absorptive capacity and related studies in the leading academic thesis database, Scopus, and drew out and analyzed 20 keywords.

Limitation of this study is that it did not utilize various databases for academic theses. It did not consider such foreign sources as EBSCO Host DB, electronic journals as Science Direct and Springer, riss of Korea, e-article, etc. Future researches are expected to obtain data from more diverse databases for academic theses and to use various methods in addition to text mining technique used in

this study.

6. References

1. Park ST, Jung JR, Liu C. A study on policy measure for knowledge-based management in ICT companies: focused on appropriability mechanisms. *Information Technology and Management*. 2019:1-3.
2. Park ST, Li G, Hong JC. A study on smart factory-based ambient intelligence context-aware intrusion detection system using machine learning. *Journal of Ambient Intelligence and Humanized Computing*. 2018 Aug:1-8.
3. Park ST, Oh MR. An empirical study on the influential factors affecting continuous usage of mobile cloud service. *Cluster Computing*. 2019 Jan 16;22(1):1873-87.
4. Chesbrough HW. *Open innovation: The new imperative for creating and profiting from technology*. Harvard Business Press; 2003.
5. Ko MH, Park EM, Park,ST. Mining the Open Science: Themes and Trends 10 Years of Open Science. *Journal of Advanced Research in Dynamical and Control Systems*.2018;10(1):263-70.
6. Park ST, Lee SW, Kang TG. A study on the trend of cloud service and security through text mining technique.2018;7(2.33):127-32.
7. Cohen WM, Levinthal DA. Absorptive capacity: A new perspective on learning and innovation. *Administrative science quarterly*. 1990 Mar 1;35(1):128-52.
8. Zahra SA, George G. The net-enabled business innovation cycle and the evolution of dynamic capabilities. *Information Systems Research*. 2002 Jun;13(2):147-50.
9. Tsai W. Knowledge transfer in intraorganizational networks: Effects of network position and absorptive capacity on business unit innovation and performance. *Academy of management journal*. 2001 Oct 1;44(5):996-1004.
10. Mowery DC, Oxley JE. Inward technology transfer and competitiveness: the role of national innovation systems. *Cambridge journal of economics*. 1995 Feb 1;19(1):67-93.
11. Lane PJ, Lubatkin M. Relative absorptive capacity and interorganizational learning.

- Strategic management journal. 1998
May;19(5):461-77.
12. Lane PJ, Koka BR, Pathak S. The reification of absorptive capacity: A critical review and rejuvenation of the construct. Academy of management review. 2006 Oct 1;31(4):833-63.
 13. Lichtenthaler U, Lichtenthaler E. A capability-based framework for open innovation: Complementing absorptive capacity. Journal of management studies. 2009 Dec;46(8):1315-38.
 14. Park EM, Park ST. The effectiveness of absorptive capacity formation mechanism on innovation performance by industry. Indian Journal of Science and Technology. 2015;8(21):1-9.
 15. Kim DY, Park ST, Ko MH. A Study on the Analysis of IT-related Occupational Cluster using Big Data. IAENG International Journal of Computer Science. 2018 Feb 1;45(1):7-11.