J. Inf. Technol. Appl. Manag. 27(1): 59~73, February 2020
https://doi.org/10.21219/jitam.2020.27.1.059

ISSN 1598-6284 (Print)
ISSN 2508-1209 (Online)

국가 융합 R&D 특성 분석에 관한 연구: 텍스트분석을 중심으로

유기철[*] · 이태희[**] · 최상현[***] · 이정환[****]

Feature Analyze and Research of National Convergence R&D: With Focus on the Text Mining

KiCheol Yoo[*] · TaeHee Lee[**] · SangHyun Choi[***] · JungHwan Lee[****]

Abstract

There is a growing interest in convergence. National R & D is also providing various policies and institutional support to promote convergence research. Convergence research, however, does not clearly specify its characteristics at the academic and government levels. This research proceeds with the process of collecting, refining, analyzing, modeling, verifying and visualizing national R & D data through the National Science and Technology Information Service (NTIS). The method is to derive the convergence research characteristics and to derive through text mining, focusing on the unstructured information of national R & D project data. The study confirmed that there was a difference in perception between the definition of converged research and the research site. In order to improve this, the research suggested that convergence among research subjects, collaboration among research topics reflecting various backgrounds and characteristics of researchers, and analysis of characteristics of convergence research using information were suggested in the process of establishing convergence policy.

Keywords : National R & D, NTIS(National Science &Technology Information Service), Convergence Research, Unity Research, Text Ming

Received : 2019. 12. 15. Revised : 2020. 02. 12. Final Acceptance : 2020. 02. 15.

* Fist Author, Masters Course, Chungbuk National University, Department of Management Information System, e-mail : ryugami07@chungbuk.ac.kr
** Co-Author, Undergraduate Course, Chungbuk National University, Department of Management Information System, e-mail : taehee8264@hanmail.net
*** Co-Author, Professor, Chungbuk National University, Department of Management Information System, e-mail : chois@chungbuk.ac.kr
**** Corresponding Author, Assistant Professor, Chungbuk National University, Department of Management Information System, 1, Chungdae-ro, Seowon-gu, Cheongju-si, Chungcheongbuk-do, 28644 Korea, Tel : +82-2-3290-2811, e-mail : junghwan@chungbuk.ac.kr

1. 서 론

전 세계 산업과 과학은 기술 및 학문간 융합이 확대되는 방향으로 발전하고 있다[Jung and Lee, 2014]. 이에 따라 동종 기술 영역의 한계를 극복하고, 이종 기술의 장점과 효용성을 융합한 새로운 연구 패러다임이 요구되는 가운데 글로벌 선진국들은 국가 차원에서 융합연구 발전 방안을 마련하고 있다[KIST, 2018]. 우리나라 역시 1990년대 이후부터 디지털기술을 바탕으로 정보통신(IT), 바이오(BT), 나노(NT) 등 신기술간 융합에 관심을 가지기 시작했다. 이후 2008년 국가융합기술 발전 기본계획(2009~2013)을 통해 제도적 발판을 마련하였고 융합을 주요 정책 방향으로 R&D 기반을 확대하고 있다[NSTC, 2008]. 그 규모는 2.56조 원으로 전체 R&D 투자의 13.2% 수준이다[KISTEP, 2017]. 최근에는 초연결 지능화, 정밀의료, 스마트시티, 스마트농축수산, 스마트공장, 신재생에너지, 자율주행차, 고기능무인기, 지능형로봇, 국민생활문제해결 등 10대 융합연구개발(R&D) 과제를 선정 투자 계획을 발표하였다[Ministry of Information and Technology, 2018].

이처럼 융합에 대한 관심은 증가하고 있지만 그 동안 융합연구에 대한 체계적 분석은 미흡한 실정이다. 본 연구는 이 같은 상황을 감안하여 방대한 국가 R&D 데이터의 수집, 정제, 분석, 검증, 모델링, 시각화를 통해 융합연구의 특성을 살펴보고 더 나아가 과제목표, 과제 내용에 대한 텍스트 분석을 통해 융합 R&D 특성을 고찰해 보았다. 다시 말해 단순히 융합 R&D에 대한 기술통계 중심의 분석 결과를 제시하는 수준이 아니라 주요 키워드 선정부터 직관성 있는 해석까지 융합연구에 대한 전략적, 정책적 시사점을 제공하고자 한다. 이 과정에서 본 연구가 기존 국가 R&D 데이터를 분석하는 연구와 차별화 되는 점은 다음과 같다.

먼저 국가 R&D 전체를 대상으로 융합 R&D를 구분하여 특성을 분석한다. 현재 국가과학기술지식정보서비스(NTIS)[1]에서 구분하는 융합 R&D는 국가과학기술표준분류에서 연구책임자가 지정한 대분류 2개 이상에 속하는 과제로 정의되고 있다[KISTEP, 2016]. 그 결과 기존 융합 R&D 특성을 분석한 연구들은 대부분 다양한 기술 분야 가운데 한 개의 기술을 선정하거나 특정 R&D 사업을 대상으로 기술 통계학적 정보(기술분류, 개발단계, 연구비, 수행 주체 등)를 제공하는 수준에서 분석이 이루어졌다. 본 연구는 2017년 R&D 과제(연간 6만 개, 20조 원 규모)를 분석 대상으로 보다 직관적인 정보 제공이 가능한 데이터 시각화 측면에서 초점을 맞추었다. 다음은 국가 R&D 관련 연구 데이터 정보 활용 범위를 정형적 부분에서 비정형적인 부분까지 확장한 것이다. 지금까지 융합 R&D와 관련된 연구는 인력 양성, 조직관점의 문제점 개선, 연구 과정에서 시행착오를 줄이는 방법 등 설문 조사를 활용한 결과 도출 방식이 많아 기술적 발전 방향과 연구 현장 정보를 있는 제대로 확인할 수 없었다[Lee, 2014]. 이를 개선하기 위해 본 연구는 융합 R&D와 관련한 텍스트 정보(사업목적, 연구내용, 과제요약서 등)를 통해 맥락적(Contextual) 의미를 추출하여 융합 R&D 특성을 파악해 보고 그 변화 동향을 분석 하였다. 특히 이 과정에서 2017년, 2018년 2년 R&D 과제를 비교 분석함으로 융합 연구 변화 흐름을 파악하였다.

2. 기존 연구

2.1 분석 보고서

최근 데이터 중심의 국가경쟁력 강화에 대한 관심이 높아지면서 정책수립과정에서도 데이터를 활용한 의사결정 필요성이 제기되고 있다. 구체적으로 국가과학기술지식정보서비스(NTIS)나 연구 현장 설문조사 결과 등의 정량적, 정성적 자료를 활용하여 국가 융합기술 R&D 조사, 정책 및 이슈사항 도출, 투자 및 성과 분석 등의 연구가 확대 되었다. 이외에도 최근 관심이 증가하는 SW 분야 R&D를 중심으로 융합에 대한 인식 차이를 확인하는 연구도 있다. 하지만 기존 조사분석 보고서에는 '융합'의 필요성을 당위적이고 규범적으로 받아들이거나, 연구 특성에 따라 융합에 대한 기준 적용이 상이한 경우가 많다는 것을 확인할 수 있다. 이는 융합연구에 대한 정의가 2가지로 구분되는 것과 관련되어 있다. 융합 연구는 첫 번째 서로 다른 연구 주체간 협동·공동 연구를 의미한다. 이 경우는 주로 출연연, 기업,

1) NTIS(National Science & Technology Information Service): R&D 사업, 과제, 인력, 성과 등 국가연구개발사업에 대한 정보를 한 곳에서 서비스하는 국가과학기술지식정보 포털.

대학 등 서로 다른 목적을 갖는 기관들이 협력하여 공동으로 연구하는 경우를 의미한다. 두 번째는 다양한 학제 및 이종 기술 간의 결합을 통해 확보되는 혁신기술로 서로 다른 분야의 기술이나 지식을 결합하는 연구로 정의하고 있다[Lee et al., 2012]. 이러한 구분 가운데 융합연구와 관련된 기존 연구보고서들은 융합의 동태적 변화 흐름을 확인하고 거시적 특성을 조망하여 그 안에서 새로운 가치 창출 가능성을 모색하는 본연의 융합연구는 제한적이다. 실제로 '2017년도 융합연구연감(2018)', '국가 융합 기술 R&D 조사 분석(2017)', '국가연구개발정보를 활용한 사업화 성과 연계구조 분석(2017)' 등은 융합연구의 정성적, 정량적 성과나 경제적 효과를 제시하는 정도이다. 다음 <Table 1>은 관련 연구 보고서를 요약한 것이다.

해외 주요국 역시 여러 가지 기술 중심의 융합연구를 진행하고 있다. 구체적으로 미국은 주요 다부처 공동국책사업과 연방 기관 또는 소규모 기관 간 협동 프로그램 중심의 목표지향적 연구에 중점을 두며, 정책과 사업의 높은 연관성이 있는 편이다[Wagnera et al., 2011]. 다시 말해 목표로 삼는 기술(예: 무인자동차, 스텔스 전투기 등)을 개발하는 것에 우선순위를 두고 있으며, 연구 과정에서 발생한 한계나 기술적 어려움 또한 새로운 연구 결과로 인정한다. 일부는 개발 후 사업화까지 연계되어 연구자가 기술 개발 이외의 성과 압박으로부터 자유롭다는 평가가 있다. EU는 기초과학이나 경제성 보다 건강, 식량안보, 환경오염, 신재생 에너지 등 사회문제 해결을 위한 중장기 대형 프로젝트 중심의 융합연구를 추진하는 경우도 있다[Nordmann, 2004]. 이처럼 융합연구가 특정한 목표 혹은 문제 해결 중심으로 중장기적이고 연속적인 연구 형태로 주제 선정부터 사업화 전 단계에서 프로세스화 되어 있는 특성으로 우리나라의 정부 주도 융합연구의 한정된 분야의 연구와는 다른 면이 있다[Kim et al., 2018].

<Table 1> Major Reports about Convergence Research

Report	Main Contents
Analysis and Implications of Software Convergence Research and Development(2019)	- Sampling 2,005 projects during all national R&D project(54,827) in 2016 - Analysing SW projects convergence by using Delphi method with 30 experts.
Convergence Research 2017 (2018)	- Introducing related policy, issues and trend about convergence, - Giving implications to vitalize convergence research with quantitative and qualitative assessment from survey results.
National R&D of Convergence Technology Research (2017)	- Classifying convergence R&D projects by type at the task level, research areas etc. - Presenting the current status of convergence technology R&D centering on quantitative indicators - Uutilizing the evidence data of government policy establishment such as basic development plan and annual implementation plan of convergence technology.
Analysis of Commercialization Performance Linkage Structure Using National R&D Information (2017)	- Designing methods to analyze the economic effects of government R&D investment by using information on the commercialization performance of national R&D projects. - Presenting through network visualization and linkage probability tables.
Convergence R&D Progress Analysis and Activation Plan (2015)	- Analyzing the current status and performance of converged R&D projects - Analyzing the recognition status of the definition and policy of converged R&D. - Identifing policy and strategic implications for future convergence R&D.
Promotional Strategy for Convergence Technology R&D Vitalization (2012)	- Many definitions and foreign example of convergence technology
The ERA of International R&D Investment(2019)	The findings support the role of EU policy instruments.

2.2 연구 논문

융합과 관련한 여러 연구 논문을 살펴보면 다음과 같이 특성을 가지고 있다

첫 번째는 특정 기술 또는 산업, 사업 중심의 융합 R&D 연구이다. 여기서는 플랙서블 디스플레이, 나노바이오와 같은 기술이나 제조업, 첨단융합 기술개발사업, 인문사회기반 연구 같은 사업, 그리고 정부출연연구소와 같은 연구주체의 융합 연구에 대한 기술적 통계 결과를 중심으로 연구 분야 간 유기적 관계, 협력적 연계구조, 연구효율성 등을 분석하였다. 두 번째는 융합 연구와 관련한 사람, 운영의 문제를 고민하는 연구이다. 구체적으로 융합연구 입문 과정에서 인재의 역량을 강화시키는 방안이나, 융합연구자의 시행착오를 극복하는 방법, 조직 행위론적 관점에서 융합 연구팀 내 조직과 리더의 역할 등을 중심으로 문제 해결책을 마련하는 연구가 있었다.

세 번째는 융합 생태계 구축을 통해 효과적인 연구 환경을 조성하는 것과 관련된 것이다. 여기서는 창의적이고 도전적인 연구 체계 수립, 국가별 주제와 성과 차이, 정부 부처 간 중복을 피하기 위한 연계맵 도출, 성과 영향 요인 규명, 소셜 미디어에서 관련 동향 정보 수집 등을 통해 융합 연구 활성화를 추진하였다. 다음 〈Table 2〉는 관련 연구 논문을 요약한 것이다.

〈Table 2〉 Major Research Area of Convergence Research

Area	Author	Main Contents
Convergence with technology and Industry	Joh et al.〔2018〕	- Comparison and analysis of the results of convergence on the institutes conducting convergence research projects.
	Yang and Heo〔2017〕	- Analysis of collaborative linkage structure and attributes among research fields - Identify the structural properties of interdisciplinary research networks
	Hwang〔2017〕	- Investigate national standard policies in key manufacturing technology sectors where convergence is active.
	Jung et al.〔2015〕	- Deriving technical ripple effects using patent data and analyzing technology convergence for nanobio convergence technology, a convergence field
	Lee〔2014〕	- R&D technology and specialized design convergence research technology trend and case study of flexible display device.
	Heo and Yang〔2013〕	- Analyze the formation of organic relations between research fields through network analysis and analysis of convergence research.
Convergence Research Reasercher and management	Kim et al.〔2018〕	- Researchers participated in the KRF Convergence research project summarize the difficulties & ways of overcoming interdisciplinary research.
	Baek〔2016〕	- Research on convergence research teams from the perspective of organizational behavior that can help practically manage the team.
	Oh et al.〔2012〕	- An in-depth analysis of whether researchers in the field of science and technology carry out the process of entering convergence research.
	Blanco et al.〔2018〕	- Comparred R&D expenditure between convergence research and divergence research by EU13, EU15, EU28
Environment of Convergence Research	Roh et al.〔2019〕	- Analyze convergence research trends by bigdata from posts and titles posted on social media using text mining techniques.
	Roh et al.〔2018〕	- Each department has a separate project to address the overlap and efficiency issues between projects.
	Minin et al.〔2017〕	- Explore significant difference European, USA and Chinese R&D Investment
	Joh et al.〔2015〕	- Identify linkages in convergence research projects and derive linkage maps in terms of policy, technology and departments.
	Park et al.〔2013〕	- Identify strategies to activate convergence research in basic research projects - Establish a support system for creative and challenging convergence research.

3. 연구방법

본 연구는 일반적인 기술통계 분석 결과를 제시하는 것이 아니라 국가 R&D 데이터[2]를 수집하여 직관적 해석[3]이 가능하도록 표현하고 텍스트마이닝을 중심으로 융합연구 특성을 분석하였다[Berinato, 2019]. 분석 데이터는 국가과학기술지식정보서비스(NTIS)의 중심으로 국가융합기술 R&D 조사분석 보고서에서 정의된 융합연구 기준[4]에 따라 국가과학기술표준 분류표 대분류 기준이 없는 과제(6,928개)와 1개 대분류에 속하는 과제(43,446개), 2개에 속하는 과제(8,486개), 3개에 속하는 과제(2,063개)를 0, 1, 2, 3으로 구분하여 단일연구과제(0,1)와 융합연구과제(2, 3)로 정의하였다.

다음으로 텍스트 분석은 '연구내용', '과제분류', '연구수행주체' 등 필수정보가 없는 과제를 제외한 2017년 50,930개, 2018년 53,601개 과제를 분석 대상으로 하였다. 구체적으로 텍스트 분석 과정에서는 문장 형태로 이루어진 연구내용을 단어 형태소로 분리하고 명사를 추출[5]하였는데 이 과정에서 파이썬에서 제공하는 KONLPY패키지를 활용하였다. 그리고 불용어 처리[6]를 위해 TF-IDF 알고리즘 활용하는데 이는 KONLPY패키지 적용에서 완벽한 명사 추출이 되지 않기 때문에 불용어 목록 구축 및 처리를 위한 과정으로 진행하였다. 이 알고리즘은 특정 단어가 해당 문서 내에서 얼마나 중요한지 나타내는 통계적 수치로 특정 문서 내에서 단어 빈도가 높을수록, 전체 문서 중 그 단어를 포함한 문서가 적을수록 TF-IDF 값은 높아진다. 이를 통해 흔하게 나타나는 단어를 걸러내는 전처리 과정을 거쳐 텍스트 분석을 진행하였다[Choi and Yoo, 2019]. 〈Figure 1〉은 각 단어와 문서들 간의 빈도수를 행렬형태로 표현하여 선형대수학 형태로 연산할 수 있는 방법을 나타낸 것이다.

$$W_{x,y} = tf_{x,y} \times \log(\frac{N}{df_x})$$

$tf_{x,y}$: 문서 또는 문장 y에서 출현한 단어 x의 빈도
df_x: 단어 x를 포함한 문서/문장의 개수
N : 모든 문서/문장의 개수

	word 1	word 2	word 3	…	word n
document 1					
document 2					
document 3	The values of the matrix are the respective weight values.				matrix of n * m
⋮					
document m					

〈Figure 1〉 Matrix from Each Documents and Words

4. 분석결과

4.1 융합 R&D 특성

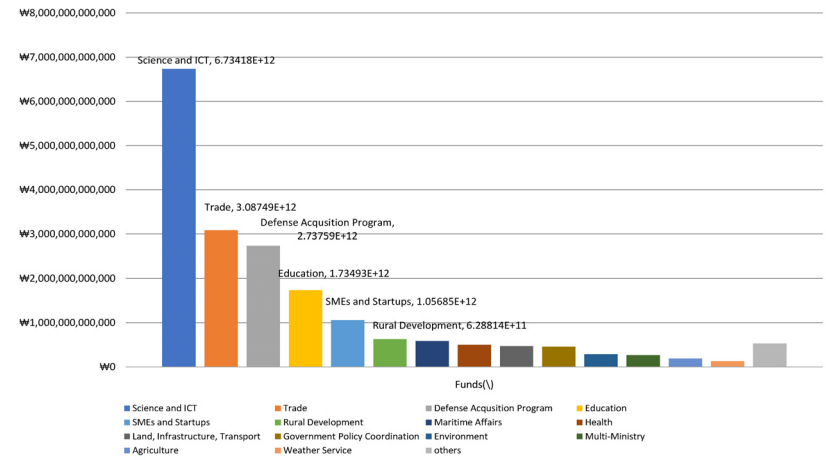
2017년 국가 R&D 과제는 전체 61,280개, 19조 3926억 원 규모이다. 사업부처별 연구비 비중은 과학기술정보통신부 34.7%(6.7조 원), 산업통상자원부 15.8%(3.1조 원), 방위사업청 14.1%(2.7조 원) 순으로 나타났으며 3개 기관이 전체의 64.7%를 차지하고 있다(〈Figure 2〉 참고).

연구수행 주체별 연구비 규모 보면 출연연 40.7%(7.8조 원), 대학 15.8%(4.4조 원), 중소기업은 14.1%(3.1조 원) 순으로 나타났으며 R&D 과제 개수의 경우 대학 53.26%(32,636개), 중소기업 23.67%(14,502개), 출연연 10.89%(6,671개)로 프로젝트 평균적 규모 차이를 확인할 수 있다(〈Figure 3〉 참고). 연구수행 주체별로 단일과제와 융합과제를 수행한 비율을 보면 모두 출연연구소가 1위, 대학이 2위, 중소기업이 3위로 확인되었다. 특히 대학은 단일과제에서는 18.1%를 차지하지만, 융합과제에서는 34.6%로 융합연구의 비중이 높게 나타났다(〈Figure 4〉 참고).

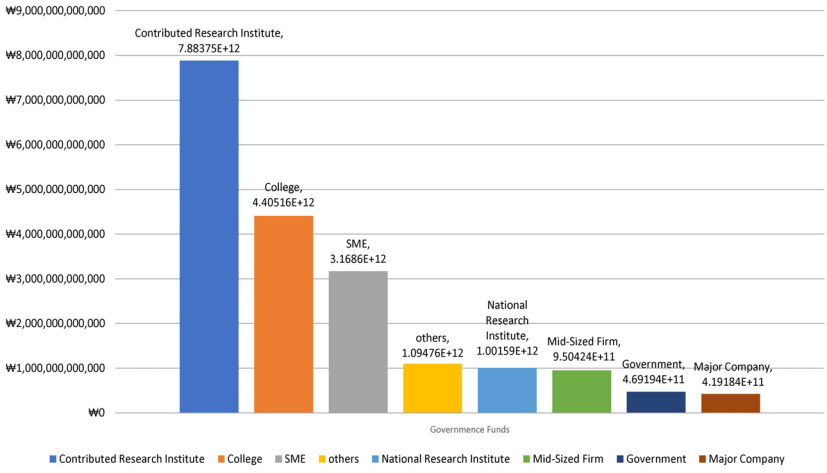
2) 2017년 기준 61,280개, 2018년 63,697개 과제.
3) 데이터시각화라는 개념으로 vRik, Qlikview, 태블로(Tableau), 스팟파이어(Spotfire)와 같은 다양한 도구(tool)를 활용.
4) 국가과학기술표준분류표 대분류 기준으로 서로 다른 2개 이상의 과학기술표준분류를 동시에 포함하는 다학제 연구. 서로 다른 2개를 포함하는 과제를 단일융합과제, 서로 다른 3개 분류를 포함하는 과제를 복합융합과제로 구분.
5) 명사추출 과정 예시

Before extract nouns (명사 추출 전 연구내용)	After extract nouns (명사 추출 후 연구내용)
or vessel collision, collision with harbor facillty and...(선박 충돌이나, 항구시설과의 충돌 등으로 …)	Vessel, Collision, Harbor facility, Inside..(선박, 충돌, 항구시설, 충돌・파공, 내부, 오염물질 …)

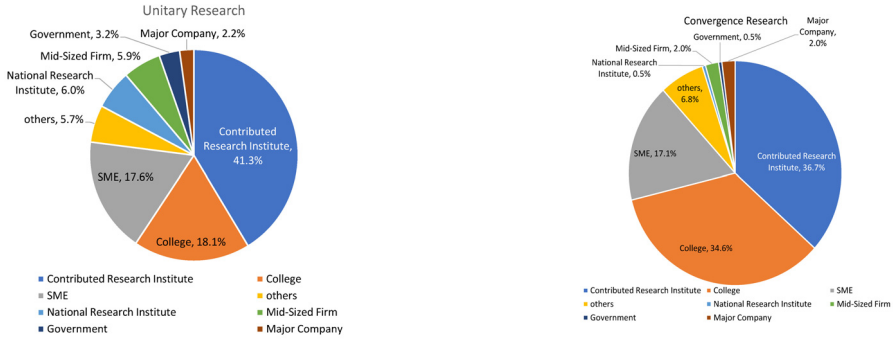
6) 본 연구에서의 불용어는 ① '입니', '합니' 와 같이 동사가 제대로 제외되지 않은 단어, ② 관사, 전치사, 조사 등 의미 없는 단어, ③ 오타, ④ 사업과제정보의 특성상 빈발하게 사용되는 단어 (ex) 개발, 분석, 기술, 이용 등)를 고려.



〈Figure 2〉 Government Research Funds of Each Ministry



〈Figure 3〉 Number of Type to use Government Research Funds

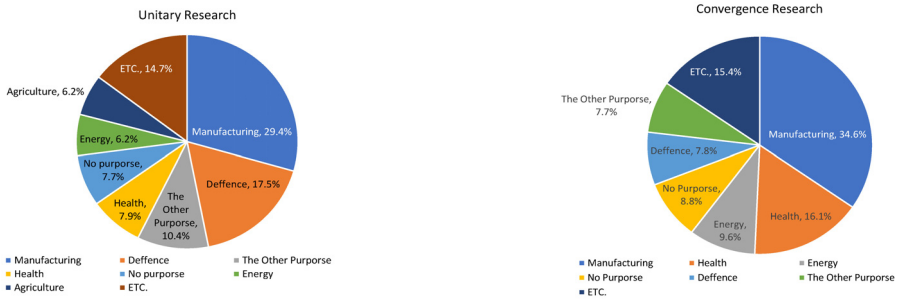


Unity Research Convergence Research

〈Figure 4〉 Rate Difference of Research Group

〈Figure 5〉는 단일연구와 융합연구별로 연구분야를 표현한 것으로 단일연구와 융합연구 모두 산업생산 및 기술(각각 29.4%, 34.6%)이 가장 높은 비율을 차지하며 다음으로 단일과제는 국방(17.5%)이, 융합과제는 건강(16.1%) 순으로 나타났다.

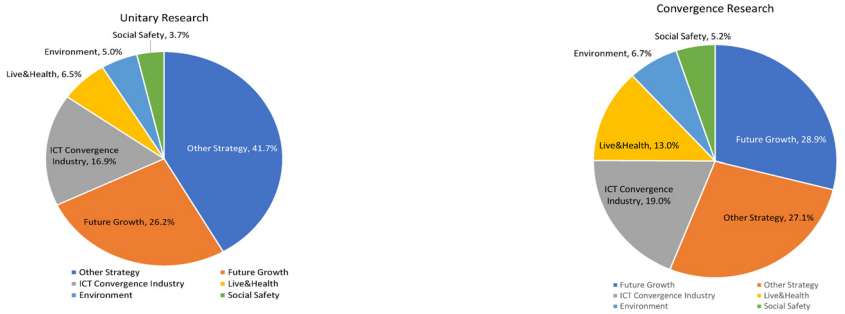
〈Figure 6〉는 6T 기술 관점에서 이루어지는 R&D 과제수를 표현한 것으로 단일연구는 IT(정보기술)가 18.6%로 가장 큰 비중을 차지하며, 융합연구는 BT(생명공학기술)가 27%로 가장 큰 비중을 차지한다. ET(환경기술)는 융합, 단일연구 모두 높은 비중을 차지하고 있다. 국가전략기술로 보면, 단일연구와 융합연구 모두 미래성장동력 확충에 관심이 높고, 융합연구는 건강장수시대 구현이 13%로 단일연구의 2배 수준이다. 하지만 단일연구와 융합연구가 아래와 같은 분석결과 차이가 있음에도 불구하고 그 원인을 유추할 수 있는 인과관계를 일반화하기는 어려운 면이 있다.



Unity Research Convergence Research

〈Figure 5〉 Rate Difference of Major Research Area

Unity Research Convergence Research



Unity Research Convergence Research

〈Figure 6〉 Comparison between 6T(Upper)Technology and National Strategic Technology(Under)

4.2 텍스트 마이닝을 통한 융합 R&D 특성 분석

연구과제 정보가 있는 과제(2017년 50,930개, 2018년 53,601개) 가운데 특별한 의미는 없지만 자주 나타나는 단어[7]는 걸러내고 연구수행주체를 중심으로 텍스트분석을 진행하였다. 먼저 연구수행주체 기준으로 국공립연구소, 출연연구소, 대학, 대기업, 중견기업, 중소기업 등 범주로 분류하고 각 범주별 누적빈도수를 가지고 워드클라우드[8]를 구성 하였다(〈Figure 7〉 참고). 분석결과 국공립연구소, 대학, 출연연구소는 주로 농림식품, 화학 등의 연구와 밀접한 단어가 높은 빈도를 갖는 것을 확인할 수 있는데 특히 국공립연구소의 거의 대부분은 농림수산식품과 관련된 것을 확인할 수 있다.

〈Figure 7〉 Keyword Coloud of Each Research Group

반면 기업은 소재, 생산품, 기술과 관련된 단어들이 빈도가 높게 나타나는데 특히 대기업은 플랜트, 복합소재, 나노카본, 자율주행 등 고급기술과 관련된 단어의 빈도가 높고, 중견기업, 중소기업으로 갈수록 필름, 카메라, 모터, 엔진, 배터리, 스마트폰 등 규모가 작고, 구체적인 단어들이 많아지는 것을 알 수 있다. 이는 중소기업은 우리나라의 산업 구조 상황에서 그 규모에 맞는 국가적 정책 지원이 추진되며 대기업은 소재-부품-완제품으로 이어지는 제조업 구조상의 특성이 반영된 것으로 볼 수 있다. 아울러 국공립연구소, 출연연구소, 대학 등은 부족한 기초학문과 핵심 산업 이외의 연구를 지원하는 결과가 반영되고 있음을 의미한다.

다음으로 국가 R&D 키워드 특성을 분석해 볼 때 2년에 걸쳐 한쪽에서만 등장하는 단어는 27~28개, 동시에 등장한 단어가 72~73개인 것을 확인할 수 있으며〈Table 3〉, 100순위 기준으로 단일연구와 융합연구에 출현하는 세부 단어는 〈Table 4〉과 〈Table 5〉와 같다.

실제 2017년을 기준으로 각 범주별 Top 20개 단어를 비교 했을 때〈Table 6〉, 14개(70%) 단어가 중복되는 것을 볼 수 있는데, 과제분류상의 기준으로는 융합연구와 단일연구가 구분 되어 있으나, 실제로 과제 내용으로 볼 때 단일연구와 융합연구 간 크게 차이가 없는 것을 알 수 있다. 이 결과는 국가 R&D 과제가 2년 이상의 연속된 과제로 진행되는 경우가 많고 1년 시간 간격을 고려했을 때 시점의 차이가 크지 않기 때문에 나타난 결과로 파악된다.

〈Table 3〉 Words Difference about Project Type

Category(분류)	Words(단어개수)	
	2017	2018
The words number of only in unity research (단일연구에서만 등장한 단어)	28	27
The words number of only in convergence research (융합연구에서만 등장한 단어)	28	27
The words number in two category (동시에 등장한 단어)	72	73

7) 대표적 불용어: 개발(96,811개), 분석(52,979개), 기술(34,532개), 이용(30,678개), 설계(28,613개), 시스템(23,985개) 등.

8) 단어가 클수록 빈도가 많다는 의미.

〈Table 4〉 Words Frequency Comparison Between Unity Research and Convergence Research

Category	2017			2018		
	word	Unity research	Convergence	word	Unity research	Convergence
The words in both category	Structure(구조체)	756	464	2 dimensions(2차원)	1121	455
	flora(균주)	1312	405	Structure(구조체)	835	423
	graphene(그래핀)	1203	710	flora(균주)	1518	452
	nanoparticle(나노입자)	1478	939	graphene(그래핀)	1087	451
	tolerance(내성)	1238	474	machine learning(기계학습)	728	256
	aging(노화)	675	344	nanoparticle(나노입자)	1596	817
	porosity(다공성)	679	371	tolerance(내성)	1263	469
	metabolism(대사)	1040	442	aging(노화)	890	306
	metabolome(대사체)	798	329	porosity(다공성)	673	331
	colorectal cancer(대장암)	783	311	metabolism(대사)	1023	512
	mutant(돌연변이)	903	408	metabolome(대사체)	935	300
	display(디스플레이)	700	214	colorectal cancer(대장암)	961	234
	digital(디지털)	955	279	mutant(돌연변이)	981	357
	deep learning(딥러닝)	882	461	display(디스플레이)	717	269
	laser(레이저)	1410	479	digital(디지털)	1145	300
	robot(로봇)	1522	398	deep learning(딥러닝)	1552	558
	immune(면역)	1082	296	laser(레이저)	1710	518
	wireless(무선)	1248	358	robot(로봇)	1796	436
	microbe(미생물)	1892	775	immune(면역)	1153	321
	mitochondria(미토콘드리아)	778	239	wireless(무선)	1357	328
	virus(바이러스)	1811	530	microbe(미생물)	2247	677
	bio-marker(바이오마커)	1274	547	pine dust(미세먼지)	912	242
	thin membrane(박막)	1508	830	mitochondria(미토콘드리아)	711	338
	semiconductor(반도체)	997	495	virus(바이러스)	1955	568
	radiation(방사선)	935	694	bio-marker(바이오마커)	1377	587
	battery(배터리)	914	291	thin membrane(박막	1655	785
	vaccine(백신)	781	244	semiconductor(반도체)	1158	465
	complex(복합체)	1174	537	radiation(방사선)	988	562
	nonlinear(비선형)	698	260	battery(배터리)	1292	232
	oxide(산화물)	640	379	security(보안)	1166	225
	fiber(섬유)	705	315	complex(복합체)	1388	492
	surgery(수술)	886	234	nonlinear(비선형)	693	272
	acceptor(수용체)	1144	517	oxide(산화물)	784	314
	stress(스트레스)	1534	532	fiber(섬유)	838	258
	simulator(시뮬레이터)	634	217	surgery(수술)	1033	234
	plants(식물)	966	299	acceptor(수용체)	1210	415
	food(식품)	1026	322	stress(스트레스)	1578	554
	nerve cell(신경세포)	789	447	plants(식물)	1139	254
	silicon(실리콘)	779	276	food(식품)	1022	337
	antenna(안테나)	969	295	nerve cell(신경세포)	908	425
	cancer cell(암세포)	1027	524	silicon(실리콘)	845	222
	restrainer(억제제)	669	253	antenna(안테나)	959	276
	derivatives(유도체)	870	238	cancer cell(암세포)	1117	506
	breast cancer(유방암)	748	263	restrainer(억제제)	699	259
	genes(유전자들)	691	218	breast cancer(유방암)	707	313
	dielectric(유전체)	2176	717	dielectric(유전체)	2652	688
	clinical test(임상시험)	1602	279	AI(인공지능)	993	310
	resilience(저항성)	1581	420	resilience(저항성)	1553	293
	transcriptome(전사체)	724	266	transcriptome(전사체)	850	245
	inosculation(접합)	749	244	inosculation(접합)	922	224
	cyst(종양)	1287	469	cyst(종양)	1378	564
	stem cell(줄기세포)	1715	785	stem cell(줄기세포)	1749	637
	car(차량)	1517	386	car(차량)	1713	415
	natural product(천연물)	715	300	natural product(천연물)	699	285
	ultrasonic wave(초음파)	1123	335	ultrasonic wave(초음파)	1300	341
	extract(추출물)	1510	379	extract(추출물)	1627	270
	camera(카메라)	1544	317	camera(카메라)	1783	325
	concrete(콘크리트)	1013	225	concrete(콘크리트)	1193	218
	solar battery(태양전지)	640	667	cloud(클라우드)	1245	228
	peptide(펩타이드)	937	422	solar battery(태양전지)	743	555
	printing(프린팅)	790	352	peptide(펩타이드)	1022	431
	plasma(플라즈마)	1103	483	printing(프린팅)	1103	335
	skin(피부)	942	269	plasma(플라즈마)	1310	333
	film(필름)	940	269	skin(피부)	1078	269
	hybrid(하이브리드)	1539	646	film(필름)	1090	241
	anticancer(항암제)	1040	355	hybrid(하이브리드)	1656	686
	antigen(항원)	706	241	anticancer(항암제)	1044	381
	antibody(항체)	1426	440	antigen(항원)	727	269
	vein(혈관)	630	422	antibody(항체)	1488	509
	fluorescence(형광)	736	303	vein(혈관)	762	436
	compound(화합물)	1597	557	fluorescence(형광)	756	338
	candidate substance(후보물질)	1166	364	compound(화합물)	1793	539
				candidate substance(후보물질)	1184	311

〈Table 5〉 Words Frequency Comparison by Each Research Type and Year

Category	2017			2018		
	word	Unity research	Convergence	word	Unity research	Convergence
The words only in Unity research	object(객체)	637		drone(드론)	804	
	drone(드론)	674		lens(렌즈)	686	
	memory(메모리)	820		memory(메모리)	965	
	immune cell(면역세포)	639		motor(모터)	1315	
	motor(모터)	1001		fermentation(발효)	795	
	fermentation(발효)	779		vaccine(백신)	749	
	pathogenicity(병원성)	646		vector(벡터)	681	
	security(보안)	1201		powder(분말)	1100	
	powder(분말)	911		block(블록)	787	
	block(블록)	720		growth(생육)	952	
	growth(생육)	830		ceramic(세라믹)	910	
	rat(생쥐)	620		export(수출)	761	
	germ(세균)	655		smart phone(스마트폰)	933	
	ceramic(세라믹)	772		aluminium(알루미늄)	686	
	export(수출)	747		engine(엔진)	1290	
	smart phone(스마트폰)	955		derivatives(유도체)	789	
	engine(엔진)	1122		gens(유전자들)	724	
	genetic resource(유전자원)	1067		genetic resource(유전자원)	1161	
	crop(작물)	791		voice(음성)	742	
	cultivation(재배)	758		crop(작물)	963	
	seed(종자)	634		cultivation(재배)	932	
	cohort(코호트)	748		seed(종자)	724	
	cloud(클라우드)	1203		cohort(코호트)	866	
	soil(토양)	1098		sunlight(태양광)	843	
	alloy(합금)	1012		soil(토양)	1221	
	antibiotics(항생제)	635		alloy(합금)	1105	
	transformation(형질전환)	690		cosmetics(화장품)	847	
	cosmetics(화장품)	727				
The words only in Convergence research	machine learning(기계학습)		239	CO2		246
	nanostructure(나노구조)		299	convergence research(공동연구)		218
	nanostructure(나노구조체)		261	nanostructure(나노구조)		245
	nanofiber(나노섬유)		216	nanostructure(나노구조체)		249
	doping(도핑)		286	macrophage(대식세포)		258
	ligand(리간드)		238	igand(리간드)		260
	separation membrane(분리막)		222	rat(생쥐)		256
	hydrogen(수소)		333	germ(세균)		237
	synapse(시냅스)		277	hydrogen(수소)		255
	nerve(신경)		349	nerve(신경)		347
	quantum dot(양자점)		294	quantum(양자)		245
	wearable(웨어러블)		304	quantum dot(양자점)		229
	imaging(이미징)		375	wearable(웨어러블)		248
	CO2(이산화탄소)		264	imaging(이미징)		318
	AI(인공지능)		277	CO2(이산화탄소)		254
	phosphorylation(인산화)		242	phosphorylation(인산화)		212
	recombination(재조합)		274	recombination(재조합)		308
	inhibitor(저해제)		213	low energy(저전력)		214
	precursor(전구체)		219	electrochemical(전기화학적)		311
	electrochemical(전기화학적)		360	conductivity(전도성)		320
	conductivity(전도성)		340	battery(전지)		232
	battery(전지)		299	dementia(치매)		258
	supporter(지지체)		260	Perovskite(페로브스카이트)		330
	sunlight(태양광)		216	lung cancer(폐암)		211
	Perovskite(페로브스카이트)		375	hydrogel(하이드로겔)		224
	hydrogel(하이드로겔)		242	antibiotics(항생제)		232
	anti cancer(항암)		232	adsorption(흡착)		333
	adsorption(흡착)		292			

〈Table 6〉 Top 20 Words Comparison Between Research Type

Frequency rank	Unity research	Convergence research
1	dielectric(유전체)	nanoparticle(나노입자)
2	microorganism(미생물)	hin membrane(박막)
3	virus(바이러스)	stem cell(줄기세포)
4	stem cell(줄기세포)	microorganism(미생물)
5	solar battery(태양전지)	dielectric(유전체)
6	compound(화합물)	graphene(그래핀)
7	resilience(저항성)	radiation(방사선)
8	tolerance(내성)	olar battery(태양전지)
9	hybrid(하이브리드)	hybrid(하이브리드)
10	stress(스트레스)	compound(화합물)
11	robot(로봇)	bio marker(바이오마커)
12	car(차량)	complex(복합체)
13	extract(추출물)	stress(스트레스)
14	thin membrane(박막)	virus(바이러스)
15	nanoparticle(나노입자)	cancer cell(암세포)
16	antibody(항체)	acceptor(수용체)
17	laser(레이저)	semiconductor(반도체)
18	flora(균주)	plasma(플라즈마)
19	cancer cell(암세포)	laser(레이저)
20	bio marker(바이오마커)	tolerance(내성)

5. 결 론

5.1 연구 요약

본 연구는 국가과학기술지식정보서비스(NTIS)를 통해 수집한 연간 20조 원 규모의 국가연구개발사업 정보를 활용하여 융합연구 특성을 시각화해 보고 텍스트 마이닝을 통해 구체적 특성을 확인해 보았다. 본 연구가 보다 객관적이며 목적 중심적인 융합정책 및 연구 기획에 활용될 수 있기를 기대하면서 분석 과정에서 확인한 사항은 다음과 같다.

먼저 융합연구와 관련된 다양한 분석 결과를 볼 때 융합연구만의 특정한 패턴이나 특성을 뚜렷하게 구분하기에 제한적인 것을 볼 때 현재 융합연구와 관련한 정책 현실과 연구현장의 시각 차이가 존재하는 것을 확인할 수 있었다. 이는 우리나라 R&D가 정부 중심의 융합연구 기준에서 구분하는 연구와 실제 연구현장에서 이루어지고 있는 융합 연구와 차이가 있기 때문이다. 다시 말해 연구현장에서는 융합이라는 가치 창출 중심의 연구가 진행되는 경우에도 단일연구로 인식되는 경우가 있고 연구자의 주관에 따라 단일 기술 분류 중심의 연구임에도 불구하고 융합연구로 행정적 목적에 따라 구분되는 경우도 있을 수 있다는 것이다. 그럼에도 불구하고 지금 R&D 현장 및 정책 실행 과정에서는 연구비규모, 연구개발 단계, 공동/위탁연구 형태 등 여러 가지 융합연구 현황분석 결과가 활용되고 있는데, 본 연구의 결과로는 타 분야와 시너지를 만드는 융합연구나 여러 기관이 함께 수행하는 융합형 연구는 '융합'으로 제대로 파악되지 않는다는 것이다. 따라서 융합의 물리적 특성과 융합정책 환경의 불일치 극복을 위해 기술의 융합 이외 연구주체 간 융합이나 연구자의 다양한 배경과 특성을 반영한 연구주제간 협업이 이루어 질 수 있도록 보다 심도 있는 융합에 대한 이해를 제안한다. 궁극적으로는 이를 반영하여 융합연구 정의를 재정립하는 것에 대한 논의가 필요한 시점이다.

4차 산업혁명 시대 융합현상은 점점 확대가 예상되고 그 자체로도 의미를 갖는 바람직한 활동일 수 있다. 그러나 일부는 그 자체가 목적이 되어 융합을 위한 융합연구를 추진하거나 융합의 목적과 본질이 명확하지

않은 연구가 융합연구로 인식되어 연구자의 혼선, 그리고 그 과정에서 정부 재원의 합리적 배분에 영향을 준다면 국가 정책 차원에서 진지하게 개선 방향을 고민해야 하는 부분이다. 예를 들어 결과론적 관점에서 성과 중심(논문, 특허, 기술료, 사업화, 인력양성, 연수지원)의 융합연구인지, 연구추진 과정에서의 연구주체간의 융합인지, 지금과 같은 기술융합 관점인지 연구정보 입력단계부터 명확하게 목표로 제시하는 것도 방안이 될 수 있다.

무엇보다 이 모든 과정이 연구자에 의해 이루어지는 만큼 정부 주도의 국가 R&D 프로세스를 연구자와 현장 중심의 프로세스로 변화하는 것이나 당장의 프로세스 변화가 어렵다면 보다 중장기적인 프로젝트 진행과 구체적인 목표를 제시하는 것 또한 논의를 해 볼 수 있을 것이다.

5.2 연구 의의 및 향후 연구

데이터를 활용하고, 분석하는 능력이 국가경쟁력을 좌우하는 '데이터 주도의 혁신 시대'가 도래 하였다〔OECD ,2015〕. '데이터 주도'라 함은 데이터 분석이 단순한 사실, 인과관계의 증명 등을 넘어 우리의 의사결정의 방향성을 제시하고, 주관적인 경험이 아닌 객관적인 데이터를 통해 보다 효율적인 가치창출이 이루어질 수 있음을 의미한다.

본 연구는 최근 R&D 가 요소기술 및 개별 산업 중심에서 기술 및 학문 간 융합되는 흐름 가운데 동종 기술 영역의 한계를 극복하고, 이종 기술의 장점과 효용성을 융합(Convergence)하는 패러다임에 주목하고 그 실체를 규명해 보고자 하였다. 그 결과는 기존에 이루어졌던 정성적 연구나 보고서와 비교해 보면, 정성적 연구 분석과 데이터 기반의 정량적 분석 사이에 차이가 있음을 확인하였다.

이는 선진국과 같이 융합 R&D를 실행함에 있어 정부가 명확한 목표를 설정하는 것과 기술 개발자체를 목적으로 하는 것에 차이가 있음을 보여준다. 우리나라의 국가 R&D 특히 융합연구 분야는 모호한 목적에서 벗어나 보다 명확한 목표를 제시해 줄 필요가 있음을 확인할 수 있다.

또한 우리가 사용한 국가 R&D와 같이 방대한 연구 정보가 제대로 활용하면 복잡한 현상의 분석과 성과 예측에서 유용한 정보로 사용될 수 있지만 그렇지 않으면 의미 없는 평범한 대상에 불과하기 때문에 과학적이고 체계적 분석이 필요하다는 것을 시사해 주고 있다.

향후에는 본 연구가 가진 융합 R&D 국가연구개발정보 활용 이외에 공개가 제한되어 있는 정보(기술료, 인력양성 등)를 활용한 연구나 융합연구 선진국과의 융합연구 주제 및 결과, 단위 과제별 비용 및 인력 규모 비교 등을 통해 우리나라 융합 R&D 보완점을 보다 다양한 시각에서 찾아보는 시도 또한 필요하겠다.

References

〔1〕 Berinato, S., The Harvard Business Review Good Charts Collection: Tips, Tools, and Exercises for Creating Powerful Data Visualizations, Harvard Business Review Press; Workbook edition, February 19, 2019.

〔2〕 Blanco, F. A., Delgado, F. J., and Presno, M. J., "R&D Expenditure in the EU: Convergence or Divergence?", Governance and Economics research Network, GEN Working Paper B 2018 - 4, June 2018.

〔3〕 Cho, Y. R., Yang, I., Suh, Y. Y., and Jeon, J. H., "Analysis of Current Situation and Relationship among National R&D Projects for Technology Convergence", *Journal of the Korean Institute of Industrial Engineers*, Vol. 41, No. 3, 2015, pp. 305-323.

〔4〕 Choi, W. S. and Yoo, K. C., "Create List of Stopwords and Typing Error by TF-IDF Weight Value", *Korean Society for Big data Service*, BIGDAS 2019, August 2019.

〔5〕 Damioli, G., Vertesy, D., Castellani, D., "The ERA of International R&D Investments", Jrc Technical Reports, Innova Measure 3 D3.1, 2019.

〔6〕 Heo, J. G. and Yang, C. H., "Applying Network Analysis in Convergent Re-

search Relationships: The Case of High-Tech Convergence Technology Development Program", *Journal of Korea Technology Innovation Society*, Vol. 16, No. 4, 2013, pp. 883-912.

〔7〕 Heo, J. G. and Yang, C. H., "Research Networking in Convergence Relations: A Network Analytic Approach to Interdisciplinary Cooperation", *The Korea Contents Society*, Vol. 17, No. 12, 2017, pp. 49-63.

〔8〕 Hwang, G. S., "Standardization Policy for Newly Developing Area of Technology Convergence within the Manufacturing R&D", Korea Association for Public Management, Vol. 31, No. 3, 2017, pp. 263-289.

〔9〕 Joh, J. H., Kang, B. S., Choi, H. C., and Kim, Y. J., "The Study on Convergence Research Efficiency of Government-Funded Research Institute", *Korean Association for Policy Science*, Vol. 22, No. 2, 2018, pp. 1-24.

〔10〕 Jung, E. D., Ko, S. G., and Seo, W. C., "RLA Study on Analysis of Technology Fusion by Identifying Technological Knowledge Spillovers: The Case of Nanobiotechnology", *The Journal of Intellectural Property*, Vol. 10, No. 4, 2015, pp. 255-290.

〔11〕 Jung, S. H. and Kim, S. G., "Network Analysis of National Convergence R&D Project", *The Korean Institute of Industrial Engineers*, Vol. 2014, No. 11, 2014, pp. 415-423.

〔12〕 Jung, W. J. and Lee, S. Y., "R&D Performance Analysis on Convergence Technologies Using Patent Citation: Comparison of IT/ET Convergence with Others", *Journal of Information Technology Applications & Management*, Vol. 21, No. 4, 2014, pp. 65-96.

〔13〕 Kim, H. J., Kim, E. J., and Lee, S. Y., " Suggestion of Successful Convergent Research Method to Overcome Trial and Error of Interdisciplinary Convergent Researcher", *Culture and Convergence*, Vol. 40, No. 1, 2018, pp. 183-214.

〔14〕 KIST, Research and Analysis of National R&D Project 2017, 2018.

〔15〕 KISTEP, 2015 National R&D Project Result Analysis Report, 2016.

〔16〕 KISTEP, 2016 National R&D Research and Analysis Report, 2017.

〔17〕 Lee, J. I., "A study on the Flexible Display R&D Technology Design Convergence", *Korea Science & Art Forum*, Vol. 18, 2014, pp. 519-529.

〔18〕 Lee, J. M., Om, K. Y., Choi, M. S., Song, C. H., and Kim, K. Y., "Scientists and Engineers in Convergence Technologies", *Technological and Economic Development of Econom*, Vol. 20, No. 3, 2014, pp. 434-456.

〔19〕 Lee, S. G., Juh, H. G., Moon, H. N., Lee, B. Y., and Kim, G. S., "Promotional Strategy for Convergence Technology R&D Vitalization", *PD Issue Report*, KETI, Vol. 12-5, 2012.

〔20〕 Minin, A. D., Quan, X. I., Zhang, Z., "A comparison of international R&D strategies of Chinese companies in Europe and the USA", *International Journal of Technology Management*, Vol. 74, No. 1/2/3/4, 2017.

〔21〕 Ministry of Information and Technology, The 3rd Convergence R&D Vitalization Basic Plan, 2018.

〔22〕 Noh, Y. H., Lee, K. H., and Jeong, D. K., "A Study of the Factor of Humanities and Social Science based Convergence Research Performance", *The Korea Contents Society*, Vol. 18, No. 6, 2018, pp. 667-678.

[23] Noh, Y. H., Lee, K. H., Jeong, D. K., and Kim, T. Y., "Trend Analysis of Convergence Research based on Social Big Data", *The Korea Contents Society*, Vol. 19, No. 2, 2019, pp. 135-146.

[24] Nordmann, A. R., "Converging Technologies-Shaping the Future of European Societies", *EC Report*, 2004.

[25] NSTC, Basic Plan for National Convergence Technology R&D('09-'13): Misistry of Education, 2008.

[26] OECD, "Data-Driven Innovation: Big Data for Growth and Well-Being", 2015.

[27] Oh, H. S., Bae, H. J., and Kim, D. Y., "Interdisciplinary Researchers: How Did They Cross the Boundaries and Do Interdisciplinary Research?", *Asian Journal of Education*, Vol. 13, No. 4, 2012, pp. 297-335.

[28] Paik, Y. J., "Overcoming Barriers of Multidisciplinary Research Teams: From the Perspective of Organizational Behavior", *Korea Journal of Business Administration*, Vol. 29, No. 2, 2016, pp. 237-264,

[29] Park, K., Shin, S. K., Song, Y. D., Lee, S. J., Lee, E., and Cha, E. J., "Developing a Convergence Research Support Framework to Promote Creative and Transformative Basic Research", *The Korea Contents Society*, Vol. 13, No. 12, 2013, pp. 221-234.

[30] Wagnera, C. S., Roessner, J. D., Julie, K. B., Kevin, T. Boyack, K. W., Keyton, J., Rafols, I., and Börner, K., "Approaches to understanding and measuring interdisciplinary scientific research (IDR): A review of the literature", *Journal of Informetrics*, Vol. 5, No. 1, 2011, pp. 14-26.

Author Profile

유 기 철

현재 충북대학교 경영학 석사과정으로 재학 중이며, 주요 관심 분야는 데이터시각화, 데이터 마이닝, R&D 전략 등이다.

이 태 희

현재 충북대학교 경영정보학과 재학 중이며 주요 관심 분야는 스마트팩토리, 계량경영, 프로그래밍 등이다.

최 상 현

현재 충북대학교 경영정보학과 교수로 재직 중이다. 한국과학기술원(KAIST)에서 경영정보공학 박사학위를 취득하였으며, 주요 관심 분야는 데이터 마이닝, 스마트팩토리, 프로세스 최적화 등이다.

이 정 환

현재 충북대학교 경영정보학과 교수로 재직 중이다. 한국과학기술원에서 기술경영공학 박사학위를 취득하였으며, 주요 관심 분야는 모바일 비즈니스, 데이터사이언스, 기술혁신 등이다.