

워드 클라우드 기법을 이용한 국내 융복합 학술연구 트렌드 분석

김준환¹, 문형진^{2*}, 이항³

¹성결대학교 파이데이아학부 조교수, ²성결대학교 정보통신공학과 조교수, ³가천대학교 경제학과 교수

A Study on Trend Analysis in Convergence Research Applying Word Cloud in Korea

Joon-Hwan Kim¹, Hyung-Jin Mun^{2*}, Hang Lee³

¹Assistant Professor, Department of Paideia, Sungkyul University

²Assistant Professor, Department of Information & Communication Engineering, Sungkyul University

³Professor, Department of Economics, Gachon University

요약 융복합 트렌드는 4차 산업혁명의 핵심이며, 이런 기대와 가능성으로 인해 여러 분야에서 다양한 대응책이 모색되고 있다. 본 연구는 최근 10년 간 융복합 연구동향을 파악하기 위하여 정량적인 분석을 시행하였다. 구체적으로 주요 연구의 키워드를 추출하여, 워드 클라우드 기법을 적용하고 시각화하여 융복합에 대한 학술 연구동향을 파악하였다. 이를 위해 '디지털융복합연구'에 게재된 2012년-2020년간의 연구논문을 대상으로 조사하였다. 분석기간은 전반부 4년(2012년-2015년)과 후반부 4년(2016년-2019년) 두 기간으로 나눠서 비교분석하여 연구동향의 차이를 확인하였다. 추가적으로 코로나19 사태로 인한 최근 1년의 연구동향에 대한 변화를 보다 명확하게 파악하기 위해 2020년의 연구논문들을 대상으로 분석하였다. 본 연구의 결과는 융복합 분야의 핵심 주제로 연구동향을 파악하고 추후 연구를 위한 유용한 기초자료로 활용될 수 있다는 점에서 의의를 갖는다.

주제어 : 융복합 연구, 워드 클라우드 기법, 연구동향, 텍스트 마이닝

Abstract The convergence trend is the core of the 4th industrial revolution, and due to such expectations and possibilities, various countermeasures are being sought in diverse fields. This study conducted a quantitative analysis to identify the trend of convergence research over the past 10 years. Specifically, major research keywords were extracted, word cloud techniques were applied, and visualized to identify trends in academic research on convergence. To this end, research papers from 2012 to 2020 published in journal of digital convergence were investigated. The analysis period was divided into two periods: the former 4 years(2012-2015) and the latter 4 years(2016-2019) to confirm the difference in research trends. In addition, the research papers of 2020 were analyzed in order to more clearly understand the changes in the research trend of the last year due to the COVID-19. The results of this study are significant in that they can be used as useful basic data for future research and to understand research trends as keywords in the field of convergence.

Key Words : Convergence Research, Word Cloud Method, Research Trend, Text Mining

*Corresponding Author : Hyung-Jin Mun(jinmun@gmail.com)

Received January 17, 2021

Accepted February 20, 2021

Revised January 27, 2021

Published February 28, 2021

1. 서론

4차 산업혁명에는 제조공정에서부터 공급망, 사물, 인간, 서비스 간의 연결 및 디지털화를 의미하며 초연결, 초지능, 초융합 등을 그 특성으로 들고 있다. 특히 2개 이상의 기술을 융합하여 기존의 기술이 갖는 한계를 극복하고 고부가가치를 창출하려는 ‘융복합 기술’에 대한 관심이 높다. 이러한 융복합 기술은 각각의 본질은 그대로 유지하면서 서로 다른 기술들을 연결시켜 새로운 기술영역을 창출해내는 것을 의미하기 때문에 차세대 신산업의 창출은 물론이고 비즈니스 분야뿐만 아니라 ICT, 의료, 경영, 인문 및 콘텐츠 등 모든 산업 분야에 걸쳐 적용이 가능하다고 할 수 있다. 이는 21세기 융합의 시대로 기업과 사회가 필요로 하는 기업의 핵심 성장 전략으로써 주목 받고 있다. 또한 다양한 학문적 융합 및 연구결과를 토대로 디지털 뉴딜 시대에 대응하고 있다. 따라서 융복합 트렌드는 4차 산업혁명의 핵심이라고 할 수 있고 이러한 기대와 가능성으로 인해 각 부문에서 다양한 대응책이 모색되고 있다. 특히 융복합 연구는 새로운 패러다임을 제시하는 이론적 및 실무적 토대를 제공함으로써 연관되는 학문분야에 통합적이며 융합적인 연구들이 정립되고 있다.

이미 빠르게 발전하는 과학기술의 발전으로 여러 분야에서 융합연구를 통한 다양한 기술이 창출되고 있고 산학협력 등을 통한 학제 간 연구가 활성화되어 전문 영역에서부터 융합연구로 그 범위가 확장되고 있다. 따라서 융복합의 실질적인 성과는 각 영역 상호간에 융합의 요소가 되는 개별 기초학문의 토대가 지속가능하게 연계되어 구축되어져 있느냐에 달려 있다. 이상의 논의를 종합해보면, 융복합의 본질은 디지털 기술의 발전으로 각종 기술과 서비스가 하나로 협업하면서 주어진 문제를 해결하기 위하여 기존에 가지고 있던 지식들을 새로운 방식으로 융합하여 시대적 요구와 상황에 맞게 응용 및 적용하는 것이라고 할 수 있다[1].

최근의 융복합 연구동향을 살펴보면, 이론 중심의 실증적 분석보다 현상 주제의 확장, 그리고 이론 응용 접근법으로서의 새로운 현상을 발견, 강조하는 실증 및 현상 중심에 대한 관심 증대로 요약할 수 있다. 다음으로 연구 주제의 변화인데, 이러한 연구 동향에 맞추어 새로운 연구주제가 추진되거나 기존이론이 적용되면서 확장되고 있다. 그리고 기존의 접근법에 대한 대안으로 학문 간의 융합이 이루어지고 있는 것으로 보인다. 다시 말해, 이는 특정 학문 분야에 국한하여 연구가 개발되기 보다는 각

각의 산업현장에서 ICT 기술을 비롯한 새로운 기술의 산업화 및 신산업 발굴 등 사회적, 기술적 변화에 빠르고 적절하게 대응할 수 있는 기반을 둔 연구가 진행되고 있음을 보여주고 있다. 이처럼 최근 10년 동안 융복합 연구의 동향은 빠르게 변화하고 있다. 또한 융복합 연구는 아직은 초기에 머물고 있기 때문에 융복합 기술의 응용에 대한 성과는 아직 명확하게 평가하기는 어렵다. 그러나 융복합 연구의 중요성은 지속적으로 높아가고 있기 때문에 본 연구는 최근의 융복합 연구 트렌드를 반영한 최근 10년간의 연구동향을 파악하고자 하였다. 이를 위해서 키워드 출현 빈도에 대한 워드 클라우드 기법을 적용하고, 시각화하여 융복합에 대한 학술연구 동향을 파악하였다. 이를 바탕으로, 정량적으로 최근 10년 간 융복합 연구의 도출된 이슈들을 확인하였다.

2. 선행연구 고찰

2.1 워드 클라우드(Word Cloud)

융합연구의 필요성과 중요성이 부각됨에 따라, 관련 선행연구는 다양한 분야에서 접근하고 있으며, 실증적인 분석방법론을 제시하고 있다. 토픽모델링은 특정 분야에 대한 동향 분석에 주로 이용되고 있다. 또한 정성적 데이터는 분석방향에 따라 설계가 필요하다. 그러나 전문가 분석은 전문가의 지식이나 기술의 역량 및 접근 가능한 정보의 양과 수준에 따라 결과가 상이할 수 있으며, 인사이트를 찾아내려는 주관적인 판단이 개입된다. 따라서 분석에 많은 시간 및 고비용이 소요되는 한계를 지니고 있다. 그러므로 시간적으로 시급성을 요구하는 경우에는 대응할 수 없기 때문에 실시간으로 처리되는 토픽모델링이 직접적으로 활용될 수 있다. 특히 최근 학계에서 관심을 받고 있는 텍스트 데이터 분석기법 중 하나인 워드 클라우드(word cloud)는 텍스트 시각화 기법들 중 대표적인 방법이다. 빅데이터 분석에서 글 또는 데이터에서 언급된 핵심단어를 텍스트와 같은 비정형 데이터를 간결하게 원 그래프나 각 키워드(단어)가 출현한 빈도(각 단어의 빈도수)에 따라 크기의 형태로 시각화해준다. 키워드마다 중요성이 다르고, 키워드 사이에 연관관계 정보가 있다. 다양한 문맥에서 이런 정보를 키워드의 출현 빈도에 따라 추출하여 표시하는 가시화 기법이다. 워드클라우드에는 그 래프 내 키워드의 크기를 제시하므로 키워드의 노출정도를 파악하기 쉽도록 표현해준다. 또한 개론적으로 해당 핵심 키워드의 빈도를 최적화하여 표현하는 특징이 있다

[2]. 이는 방대한 데이터를 분석하여 의미 있는 결과를 비구조화된 텍스트 속에서 키워드를 통해 중요한 내용을 어떤 부분에 집중한지를 직관적으로 알 수 있게 신속하고 유용한 정보 또는 지식을 제공해주는 장점을 가지고 있다[3].

2.2 텍스트 마이닝을 활용한 연구동향

최근 다양한 분야에서 텍스트 마이닝을 통해 연구동향에 대한 분석이 활발히 이루어지고 있다. 예를 들어, 이재윤, 문주영, 김희정(2006)은 국내 문헌정보학 논문을 중심으로 문헌 클러스터링과 유사도 네트워크 분석을 통해 기록관리학 분야의 지적구조를 파악하였다[4]. 서은경(2010)은 정보학분야 학회지의 연구논문을 대상으로 표제어를 이용하여 기간별 주제 및 연구변화를 분석하였다[5]. 조재인(2011)은 7년간 문헌정보학분야의 주요 학술지에 게재된 논문 1,752건을 대상으로 단어 빈도분석과 네트워크 텍스트 분석을 수행하여 연구 경향을 분석하였다. 종전과의 변화 양상을 살펴본 결과, 다양한 주제의 개념에 대한 분포와 관계성이 발견되었다[6]. 박자현, 송민(2013)은 국내 문헌정보학 분야의 학술연구 동향을 살펴 보았다. 초록을 수집하여 LDA 기반의 토픽모델링 기법을 통해 문헌정보학자들이 공통적으로 관심을 갖는 연구 주제의 특성을 규명하였다[7]. 또한 한유진(2016)은 기업가정신 및 창업과 관련된 발표된 연구들에서 서지계량분석(bibliometric analysis)을 통해 저자 및 제목 분석을 하였다[8]. 손연정, 이수경(2019)은 국내 만성질환자를 대상으로 모바일 헬스 중재연구의 출현빈도와 시기별 키워드 변화를 확인하였다[9]. 이은영, 주경희, 이두희(2019)는 지난 10년 간 소비자학 연구동향을 주요 키워드를 대상으로 국내연구와 해외연구로 비교하고 소비문화 및 소비 트렌드에 대해 정량적·정성적 분석방법을 통해 파악하였다[3]. 한편 주우봉, 박한우(2019)는 블록체인 분야의 연구동향을 계량정보학적 네트워크 분석을 통해 연도별, 분야별, 논문저자 및 기관분석 등으로 살펴보았다[10]. 이남연, 김준환, 문형진(2019)는 국내 학술지에 게재된 감정도동관련 892편의 주제어를 텍스트 마이닝 및 네트워크 분석을 활용하여 연구동향을 확인하였다[11]. 문승리, 이원석(2020)은 젠트리피케이션 현상의 관점 차이를 워드 클라우드 기법을 통해 관광학과 비관광학의 연구로 비교하여 분석하였다[12]. 김민경(2020)은 대표적인 공유경제 숙박시설인 에어비앤비의 리스팅 타이틀을 워드 클라우드 분석으로 파악하였다[13]. 이외에도 텍스트 마이닝을 적용하여 연구동향을 살펴보는 연구

분야와 대상 자료가 최근 들어 광범위하게 활발해지는 경향이 나타나고 있다.

3. 연구방법 및 분석결과

3.1 연구방법

본 연구는 융복합 연구의 국내 대표저널인 '디지털융복합연구'를 선정하고 키워드를 중심으로 분석을 실시하였다. 그 이유는 기존 연구동향을 파악하는 연구들도 한 개의 대표 저널지를 집중적으로 분석하였으며[3], 해당 분야의 연구동향을 살펴보는 것이 보다 명확하고 시기별 트렌드 파악이 효율적이라 판단하였다. 워드 클라우드를 구성하기 위해서는 Python 라이브러리와 R studio 프로그램 등을 사용할 수 있다. '디지털융복합연구'의 전체 키워드 중에서 키워드 출현빈도 분석결과를 바탕으로 워드 클라우드(word cloud) 기법을 이용하였다. 이를 시각화하여 주요 키워드가 보다 명확하게 눈에 띄게 하여, 시기별 연구 동향에 대한 독자들의 이해를 보다 제고하고자 하였다. 본 연구는 앞서 언급한 바와 같이 Python에서 워드 클라우드 라이브러리를 활용하여 분석하였다(Python, version 3.8.3). Python에서 워드 클라우드를 만들어주는 'word cloud', 그림을 그려주는 'matplotlib.pyplot' 라이브러리를 사용하여 최종 워드 클라우드를 작성하였다.

3.2 분석결과

먼저 최근 10년간 '디지털융복합연구'의 연도별 논문수는 5,733건이며, Table 1과 같이 한국학술지인용색인(KCI)의 연도별 논문수와 피인용 횟수를 보여주고 있다.

Table 1. Number of Papers by Year

Year	2015	2016	2017	2018	2019	2020
Number of published papers	585	684	659	625	619	622
Number of citations	438	739	903	808	248	-
Self citation count	87	137	167	194	46	-

'디지털융복합연구'의 Topic Landscape로 분석한 결과와 같이 한국학술지인용색인(KCI) 전체년도의 논문 키워드 간의 관계를 살펴보았다. Figure 1에서 볼 수 있듯이, 맵 상에 나타난 고도는 키워드의 출현 빈도수가 높거

대한 공유뿐만 아니라 향후 연구로 다루어질 시의성 있는 중요한 키워드와 융합연구의 밑바탕이 되는 정보를 제공할 것이다. 또한 연구협업의 기회가 창출될 수 있다는 점에서 이론적 시사점을 제공할 것이다. 실무적으로는 주목할 만한 주요 융합 동인에 관한 자료를 확보하여 이를 분석하면 융합 증진을 위한 다양한 방안들을 도출할 수 있을 것이다.

본 연구에서는 융복합의 연구동향을 파악하고자 정량적인 분석으로 2012년부터 2020년간의 연구논문들의 주요 키워드를 대상으로 키워드 출현 빈도분석을 실시하였다. 이를 바탕으로 워드 클라우드(word cloud) 기법을 통해 연구동향에서 나타난 출현빈도와 관련성이 높은 키워드들을 시각적으로 제시하였다. 이를 위해, 최근 10년간 정책, 경영, 기술, 의과학, 엔터테인, 웰니스 등 융복합 분야의 국내 저널인 '디지털융복합연구'의 출간된 연구논문을 대상으로 2016년부터 4차 산업혁명에 대한 논의가 본격적으로 이루어지기 시작했기 때문에 2016년도 이전과 이후의 연구로 구분하고 최근 1년(2020년)의 연구동향에 대해 조사하였다.

연구결과, 전반 4년에는 우울, 스마트폰, 노인, 삶의 질, 대학생, 자아존중감, 직무만족, 자기효능감, 빅데이터, 스트레스, 조직몰입, 만족도 등 현상 영역(substantive domain) 등에서 발생하는 개인적인 인지 프로세스를 파악하고자 하는 연구들에 대한 이슈가 보다 많았다면, 후반 4년에는 간호대학생, 우울, 대학생, 자기효능감, 자아존중감, 매개효과, 간호사, 사용자 경험, 조절효과, 빅데이터, 삶의 질, 인공지능, 4차 산업혁명, 텍스트 마이닝, 블록체인, 사물인터넷 등 주로 사회와 기술 간의 융합 인프라 적용에 관한 가능성에 더 관심을 보이는 경향을 확인하였다. 한편 최근 코로나19로 인해 언택트(Un-tact, 비대면) 사회로 진입하면서 코로나19 전후의 연구주제가 변화될 수 있는지 확인해 볼 필요가 있다. 최근 1년(2020년)의 연구에서는 간호대학생, 대학생, 간호사, 자기효능감, 노인, 사용자 경험, 간호, 빅데이터, 인공지능, 우울, 스트레스, 불안, 삶의 만족도, 편의점 등이 주요 키워드로 떠오른 것으로 보아, '사회적 거리두기'로 인한 비대면 생활이 활성화되면서 건강과 소셜미디어 및 ICT 기술 등에 대한 관심이 급격히 높아져 산업 및 생활 현장에서 다양한 새로운 기술의 적용 가능성에 관심이 집중되었기 때문인 것으로 나타났다. 이러한 결과는 논문 주제에 따라 이론적 영역(theoretical domains)과 연구방법 영역(methodological)으로 나타나거나 사례 및 정책연구 등 다루는 구체적인 현상 이슈를 정량적 또는 행

동적 접근방법을 통해 제시한 것으로 보인다.

후속연구에서는 융복합 연구의 적용가능성에 대한 세부적인 주제와 영역을 세분화 시킬 뿐만 아니라 보다 구체적이고 체계적인 연구방법론을 적용하여 본 연구가 가지고 있는 한계점을 보완할 필요가 있다[14]. 또한 워드 클라우드는 편리하고 직관적인 장점이 있지만 단순히 빈도가 높다고 중요한 키워드로 해석하거나 출현 빈도별 크기의 키워드가 랜덤으로 분포되는 형태이기 때문에 의미추론이 어렵다는 한계도 있다. 따라서 추후연구에서는 키워드 네트워크 분석이나 word2vec 등을 활용하여 단순 출현 빈도수가 아니라 TF-IDF 기준으로 단어의 가치를 활용하거나 키워드 간의 중요도를 정량화하여[15], 반영된 더 정교한 연구들이 이루어지길 제안한다.

REFERENCES

- [1] J. Lind. (2005). Ubiquitous Convergence: market redefinitions generated by technological change and the Industry Life Cycle. *In DRUID Academy Winter 2005 Conference*. 27-29.
- [2] W. J. Lee (2020). A Study on Word Cloud Techniques for Analysis of Unstructured Text Data. *The Journal of Convergence on Culture Technology*, 6(4), 715-720.
- [3] E. Y. Lee, K. H. Chu & D. H. Lee. (2019). A Study on Recent Trend Analysis in Consumer Research Applying Word Cloud. *Journal of Product Research*, 37(1), 1-7.
DOI : 10.36345/kacst.2019.37.1.001
- [4] J. Y. Lee, J. Y. Moon & H. J. Kim. (2007). Examining the intellectual structure of records management & archival science in Korea with text mining. *Journal of the Korean Society for Library and Information Science*. 41(1), 345-372.
DOI : 10.4275/KSLIS.2007.41.1.345
- [5] E. G. Seo. (2010). Trends analysis on research articles in the Journal of Korean Society for Information Management. *Journal of the Korean Society for Information Management*. 27(4), 7-31.
DOI : 10.3743/KOSIM.2010.27.4.007
- [6] J. Cho. (2011). A study for research area of library and information science by network text analysis. *Journal of the Korean Society for Information Management*. 28(4), 65-83.
DOI : 10.3743/KOSIM.2011.28.4.065
- [7] J. H. Park & M. Song, (2013). A Study on the Research Trends in Library & Information Science in Korea using Topic Modeling. *Journal of the Korean Society for Information Management*, 30(1), 7-32.
DOI : 10.3743/KOSIM.2013.30.1.007

- [8] Y. J. Han. (2016). Research Trend of Entrepreneurship in Korea. *Journal of Digital Convergence*, 14(12), 121-131.
DOI : 10.14400/JDC.2016.14.12.121
- [9] Y. J. Son & S. K. Lee. (2019). Analysis of research trends on mobile health intervention for Korean patients with chronic disease using text mining. *Journal of Digital Convergence*, 17(4), 121-131.
DOI : 10.14400/JDC.2019.17.4.211
- [10] Y. P. Zhu & H. W. Park. (2019). Evaluating blockchain research trend using bibliometricsbased network analysis. *Journal of Digital Convergence*, 17(6), 219-227.
DOI : 10.14400/JDC.2019.17.6.219
- [11] N. Y. Lee, J. H. Kim & W. S. Lee. (2019). Exploration of Emotional Labor Research Trends in Korea through Keyword Network Analysis. *Journal of Convergence for Information Technology*, 9(3), 68-74.
DOI : 10.22156/CS4SMB.2019.9.3.068
- [12] M. S. Lee & W. S. Lee. (2020). A study on difference on gentrification perspectives: Comparison with the study of tourism and non-tourism. *Journal of Hospitality and Tourism*, 29(4), 229-243.
DOI : 10.24992/KJHT.2020.06.29.04.229
- [13] M. G. Kim. (2020). Examining popular key words listed for Airbnb properties using word cloud analysis. *Korean Journal of Hospitality and Tourism*, 29(4), 347-363.
DOI : 10.24992/KJHT.2020.06.29.04.347
- [14] B. E. Seo & C. M. Lee. (2014). Keyword Analyses of Diaspora Research and Its Trends in Korea. *Journal of Diaspora Studies*, 8(1), 43-69.
- [15] S. Y. Kwon. (2014). A Study on the Application to Network Analysis on the Importance of Author Keyword based on the Position of Keyword. *Journal of the Korean Society for Information Management*, 31(2), 121-142.
DOI : 10.3743/KOSIM.2014.31.2.121

김 준 환(Joon-Hwan Kim)

[정회원]



- 2003년 8월 : 서울시립대학교 경영학
과(경영학석사)
- 2010년 8월 : 서울시립대학교 경영학
과(경영학박사)
- 2014년 4월 ~ 현재 : 성결대학교 파
이데이아학부 조교수
- 관심분야 : 소비자 감정, B2B마케팅

· E-Mail : kjh715@sungkyul.ac.kr

문 형 진(Hyung-Jin Mun)

[정회원]



- 2008년 2월 : 충북대학교 전자계산학
과(이학박사)
- 2009년 3월 ~ 2012년 8월 : 중국 연
변과학기술대학교 컴퓨터전자통신학
부 조교수, 부교수
- 2017년 3월 ~ 현재 : 성결대학교 정
보통신공학부 조교수

· 관심분야 : 정보보안, 네트워크 보안, 빅데이터분석

· E-Mail : jinmun@gmail.com

이 항(Hang Lee)

[정회원]



- 1986년 8월 : 한국외국어대학교 무역
학과(경영학석사)
- 1992년 2월 : 국민대학교 무역학과(경
제학박사)
- 1993년 3월 ~ 현재 : 가천대학교 경
제학과 교수

· 관심분야 : 국제경영, 조직행동

· E-Mail : hlee@gachon.ac.kr