

소셜 빅데이터 기반 융합연구 동향 분석

Trend Analysis of Convergence Research based on Social Big Data

노영희*, 김태연**, 정대근***, 이광희****

건국대학교 문헌정보학과*, 건국대학교 지식콘텐츠연구소**, 전남대학교 문헌정보학과***,
한국연구재단 인재양성실****

Younghee Noh(irs4u@kku.ac.kr)*, Taeyoun Kim(ktykdioh@kku.ac.kr)**,
Dae-Keun Jeong(jdk1319@jnu.ac.kr)***, Kwang Hee Lee(thomas@nrf.re.kr)****

요약

본 연구는 4차 산업혁명과 함께 학제간 융합연구의 중요성이 부각되는 시점에서 소셜미디어 빅데이터 분석을 통하여 학술적 연구를 넘어 융합연구 전반에 대한 동향을 분석하고자 하였다. 이를 위해 텍스트마이닝 기법을 활용하여 소셜미디어에서 융합연구와 관련하여 2009년 1월부터 2018년 9월까지 약 10년간 게시된 글과 제목 등 약 150,000건을 수집하였으며, 이를 바탕으로 기간별로 워드클라우드와 네트워크 분석을 실시하였다. 분석결과, 각 기간별로 활발하게 진행된 연구분야는 2009년과 2010년에는 친환경, 2011년과 2012년에는 스마트, 2013년과 2014년에는 정보통신, 2015년과 2016년에는 로봇, 2017년과 2018년에는 인공지능이다. 또한 약 10년간 지속적으로 수행되고 있는 연구분야는 문화, 디자인, 화학, 나노, 바이오, 로봇, IT, 정보통신이다. 본 연구에서는 기간별 융합연구 동향을 파악하므로써, 융합연구를 기획하고 있는 연구자들에게 연구방향을 설정하는데 있어 도움이 될 수 있다.

■ 중심어 : | 소셜미디어 | 빅데이터 | 융합연구 | 워드클라우드 | 네트워크 분석 |

Abstract

This study was designed to analyze trends in the entire convergence research beyond academic research through social media big data analysis at a time when interdisciplinary convergence research is emphasized along with the fourth industrial revolution. For this purpose, about 150,000 cases of texts and titles were acquired for about 10 years from January 2009 to September 2018 in connection with the convergence research in social media, and word cloud and network analysis were conducted. As a results, the research fields that were actively conducted for each period were eco-tech in 2009 and 2010, smart technology in 2011 and 2012, information and communication in 2013 and 2014, robots in 2015 and 2016, and artificial intelligence in 2017 and 2018. Also, the research areas that have been consistently conducted for about 10 years are culture, design, chemistry, nanotechnology, biotechnology, robot, IT, and information and communication. Since this study identifies trends in convergence research over time, it can be helpful to researchers who are planning convergence research direction by understanding the trends of convergence research.

■ keyword : | Social Media | Big Data | Convergence Research | Word Cloud | Network Analysis |

I. 서론

과학기술이 발달함에 따라 점점 복잡한 문제를 해결하고 단일학문 분야의 한계를 극복하기 위해 학제간 융합연구의 필요성이 부각되고 있다. 이에 따라 교육과학기술부는 과학기술분야와 인문사회분야의 융합과 소통을 강조하면서 융합연구의 기반을 마련하기 위해 지속적으로 노력하고 있다. 이를 바탕으로 2009년 한국연구재단은 학제간 융합연구지원사업을 추진하고 여러 학문분야들이 공동으로 연구할 수 있는 융합연구 기반을 구축하고 있다[1][2].

융합연구는 다학제적 융합연구, 학제간 융합연구, 초학제적 융합연구, 이렇게 크게 세 가지로 분류할 수 있다[3]. 다학제적 융합연구는 연구문제를 해결하기 위해서 두 개 이상의 분야가 협력하지만 각자의 분야에 해당하는 연구문제를 해결하는 연구활동을 말한다[4]. 학제간 융합연구는 연구문제를 해결하는 과정에서 서로 다른 분야의 지식을 결합하여 새로운 접근을 생성하고, 이를 기반으로 연구문제를 해결하는 연구활동을 말한다. 초학제적 융합연구는 다학제적 융합연구와 학제간 융합연구보다는 통합적으로 현실의 문제를 해결하고 사회적 지식을 생성하는 연구활동을 말한다.

현재 다양한 분야에서 수행되고 있는 융합연구의 동향을 파악하기 위해서는 각 정부부처 및 출연연구기관, 그리고 대학에서 진행한 융합연구 사업계획서를 모두 수집하고 다각적으로 분석해야 한다. 하지만 대중에게 공개되는 사업계획서는 한정되어 있고 연구과제마다 사업성격과 특성이 모두 다르기 때문에 직접적으로 융합연구의 동향을 파악하는 것은 매우 어렵다. 또한 사업계획서와 논문과 같이 학술적 측면에서 동향 분석을 수행하게 되면 주로 정부출연금의 지원을 받아 수행되고 있는 연구과제에 대한 동향을 파악하게 된다. 하지만 대학 혹은 기업연구소에서 자체적으로 수행하고 있는 융합연구 관련 동향을 파악하기 어렵다.

이에 본 연구에서는 연구수행주체에 관계없이 융합연구의 동향을 파악하고자 소셜미디어 상에 노출되어 있는 융합연구 관련 데이터를 수집하여 융합연구 동향을 간접적으로 파악하고자 하였다. 먼저, 웹크롤링을 통

하여 2009년 1월 부터 2018년 9월까지 네이버 뉴스 및 신문기사와 블로그에 게시된 글과 제목을 수집하였다. 취득된 데이터를 기간별로 분류하였고, 이 후에 자연어 처리를 통하여 분석을 수행하는데 있어 불필요한 조사와 동사 등을 제거하고 빈도수가 높은 핵심키워드를 추출하여 워드클라우드 방법으로 시각화 하였다. 또한 기간별로 동시에 나타나거나 특정기간에만 국한되어 나타나는 핵심키워드를 파악하기 위하여 네트워크 분석을 수행하였다. 여기에서는 각 기간별로 빈도수가 높은 순으로 200개의 핵심키워드를 선정하여 이를 기반으로 네트워크 분석을 수행하였다. 그 결과 기간별로 소셜미디어에서 나타나는 융합연구 분야와 이와 관련된 정부정책을 파악할 수 있었다.

본 연구는 소셜미디어 빅데이터를 이용하여 기간별 융합연구 동향을 파악하는 연구라는 점에서 의의가 있으며, 융합연구에 관심있는 연구자들이 연구목표 및 방향을 결정하는데 도움이 될 수 있다.

II. 선행연구

1. 소셜 빅데이터를 활용한 동향 분석

최근 데이터마이닝 기법의 발전으로 소셜미디어 상에 존재하는 데이터를 이용한 동향분석, 요소간 관계분석, 미래 예측 등에 대한 연구가 활발하게 진행되고 있다. 다양한 종류의 소셜미디어 중에서 트위터와 페이스북에서 취득된 데이터를 이용하여 동향분석을 수행하는 연구가 진행되었다. 이오준 등[5]은 트위터 데이터를 수집하고 분석하는 시스템을 구축하였으며, 이를 기반으로 영화<변호인>에 대한 트위터 데이터를 수집하고 흥행요인을 분석하였다. 트위터에서 빈도수가 높은 명사를 추출하여 이를 영화 관객의 관심을 나타내는 것으로 간주하였고, 이로부터 영화흥행에 영향을 미친 주요 요인들을 제시하였다. 오효정 등[6]은 트위터 문서에서 GPS 정보가 포함된 데이터를 추출하였다. GPS 데이터와 내용을 기반으로 주소, 나이, 성별 정보를 추출하였으며, 이를 기반으로 트위터 사용자들의 선호지역과 이동패턴 등을 파악하였다. 이수진과 전유내[7]는 트위터

에 작성된 관광 관련 내용을 수집하였고, 이를 기반으로 연관키워드 분석과 네트워크 분석을 수행하였다. 이로부터 경기도 지역 관광이 모험과 힐링의 경향을 나타내는 것을 확인하였고 경기도 지역의 관광 발전 방향을 제시하였다. Joo and Teng[8]은 페이스북에서 취득된 소셜 미디어 데이터를 활용하여 페이스북이 사회구성원들에게 미치는 영향을 분석하였다.

트위터 이 외에 주요 포털 사이트의 뉴스 및 신문기사 혹은 기타 홈페이지에서 취득된 데이터를 이용하여 동향분석을 수행하는 연구가 진행되었다. 김대경과 이현수[9]는 웹크롤링을 통하여 뉴스기사와 블로그에서 전국장애인체육대회에 관한 글을 수집하였으며, 전치리를 통하여 데이터를 정형화 시키고 핵심 키워드 등에 대한 연관성 분석을 수행하였다. 그 결과를 이용하여 네트워크 분석을 수행함으로써 키워드들 간에 관련성이 높은 요소들을 추출하였다. 서운채와 김형중[10]은 기업에 대한 정보를 제공하는 소셜 미디어인 잡플래닛에 작성된 기업의 평판 데이터를 수집하였다. 이를 이용하여 기업의 내부평판요인과 직무만족도 사이의 상관관계를 분석함으로써 분석 대상에 해당하는 기업에 대한 문제점을 지적하고 기업의 경영방식을 개선할 수 있는 방향을 제시하였다. 오효정 등[11]은 뉴스기사 데이터를 수집하여 사회적 이슈 리스크를 학습시키고 이를 기반으로 리스크 유형을 자동으로 분류하는 모델을 생성하였다. 이로부터 미래의 위험요소를 감지하고 빠르게 대응할 수 있음을 시사하였다. 박성건 등[12]은 웹크롤링을 통하여 리우올림픽 축구에 대한 웹 뉴스 및 뉴스 댓글을 수집하였으며, 전치리를 통하여 축구와 관련된 신조어와 인물의 별칭 등을 따로 저장하였다. 이를 기반으로 한국축구 대표팀의 주요 선수에 대한 연관단어를 추출하고 네트워크 분석을 수행하였다. 이로부터 올림픽과 같이 대중들이 관심을 보이는 스포츠 행사에 대하여 관심을 높일 수 있는 방법을 제안하였다. 나경식과 이지수[13]는 20년간의 신문기사 데이터를 취득하여 기간별 도서관과 관련된 연관이 분석을 수행하였다. 이로부터 도서관에 대한 인식이 공부의 장소에서 문화를 위한 공간으로 변경되었음을 확인하였다.

2. 융합연구 동향 분석

김종범[14]은 정부에서 추진하고 있는 융합연구의 필요성과 현황, 그리고 활성화 방안을 분석하여, 항공우주 기술에 융합연구를 적용할 수 있는 방안을 살펴보았다. 이태희와 정영철[15]은 의료정보학 분야 온라인 저널 JMIR(Journal of Medical Internet Research)에 2011년부터 2013년까지 게재된 논문을 대상으로 정보기술 융합연구 동향을 파악하였다. 3년간 게재된 논문의 절반 이상을 차지하는 4가지 주제를 선택하여 통계적으로 분석함으로써 의료분야에 적용된 정보 전략기술 동향을 파악하였다.

데이터마이닝 기법을 이용하여 한국연구재단 홈페이지에 공개된 연구과제 정보를 취득하고 이로부터 융합연구에 대한 동향을 분석하는 연구가 진행되고 있다. 최하영 등[16]은 한국연구재단이 공개한 융합연구관련 연구제안서를 데이터마이닝 기법을 이용하여 수집하고 분석함으로써 국내 융합연구과제에 대한 동향분석을 수행하였다. 연구제안서 내에 있는 키워드를 추출하여 융합연구과제의 세부주제를 도출하고 주제들 간의 연관규칙 분석을 적용하여 주제들 간의 과급효과를 파악하고 네트워크 분석을 수행하였다. 장재혁과 윤장혁[17]은 한국연구재단에서 수행된 학제간융합연구사업의 연구제안서 내용을 분석하여 인문사회기반 융합연구과제의 동향을 분석하였다. 또한 연구제안서의 내용을 기반으로 토픽모델링 기법을 적용하여 연구주제를 유추하고 연관성 분석을 수행하였다. 연관성 분석결과를 이용하여 연도별로 네트워크 분석을 수행하였고 근접중심성과 매개중심성 지수를 계산하여 매년 인문사회기반 융합연구과제의 관심이 높은 주제와 중개자 역할을 수행하는 주제를 파악하였다.

그밖에도 데이터마이닝 기법을 활용하지 않고 사례 분석을 통하여 융합연구의 동향을 분석하는 연구가 진행되고 있다. 이재익[18]은 국내에서 연구하고 있는 플렉서블 디스플레이에 대한 융합연구 사례를 고찰함으로써 디스플레이에 대한 융합연구 동향을 분석하였다. 이를 통하여 미래 디스플레이 소비자에게 미칠 수 있는 영향을 분석하는 것이 중요함을 파악하였고, 수익창출에 있어 중요한 요소들을 고려하여 사업화 전략을 구상

하는 것이 플렉서블 디스플레이 융합연구의 핵심문제를 확인하였다. 김정운 등[19]은 미국 국립과학재단의 “환경 속의 생명복합성 프로그램”과 한국연구재단의 학제간 융합연구지원사업을 비교·분석함으로써 두 나라의 융합연구에 대한 동향을 분석하였다. 이를 바탕으로 국내 융합연구에 대한 발전방향과 전문인력 양성에 대한 필요성을 제시하였다.

선행연구의 경우 융합연구에 대한 동향의 대부분은 연구제안서 및 논문 등 학술적 측면에서의 동향을 분석하고 있어 본 연구에서는 뉴스, 기사, 블로그 등 소셜미디어를 통해 융합연구에 대한 전반적인 동향을 분석하고자 하였다.

III. 연구방법

1. 연구질문

본 연구는 소셜 데이터 분석을 통하여 소셜미디어에 나타나고 있는 융합연구에 대한 경향 및 관심사를 파악하고자 한다. 이를 위해 다음과 같은 2가지 연구문제를 설정하였다.

RQ 1: 융합연구 핵심 키워드는 기간에 따라 어떻게 변화하는가?

RQ 2: 융합연구 핵심 키워드의 네트워크는 어떤 구조를 보이는가?

2. 분석 대상 및 데이터 수집

본 연구에서는 2009년 1월부터 2018년 9월까지 약 10년 간 네이버 블로그, 뉴스 및 신문기사에서 “융합연구”를 키워드로 검색하여 결과를 추출하였다. 글 제목, 본문, 날짜, URL(Uniform Resource Locator) 정보를 수집하였다. 블로그, 뉴스 및 신문기사 모두 웹페이지 구조가 다르기 때문에 파이썬 Selenium 라이브러리를 사용하여 원격으로 웹 브라우저를 제어하는 방식으로 데이터를 취득하였다. 파이썬 Selenium 라이브러리는 웹 페이지에 나타나는 데이터 중에서 취득하고자 하는 요소의 xpath를 이용하여 데이터를 취득할 수 있도록 돕는다. 또한 클릭, 새로고침, 스크롤 등의 기능을 포함하

고 있기 때문에 웹페이지를 자유롭게 이동하면서 원하는 데이터를 자동으로 취득할 수 있다.

네이버 검색엔진에서 1회 검색으로 웹페이지에 나타나는 글의 개수가 최대 1,000개이므로 1개월 단위로 검색하여 나타나는 결과를 추출하였다. 또한 잦은 웹페이지 요청에 의한 아이피차단과 인터넷 속도 등에 대한 문제 해결을 위하여 다음 웹페이지를 요청하는데 있어 충분한 시간간격을 두었다. “융합연구”로 검색했을 때 가장 많이 나타나는 결과는 핵융합과 관련된 글이다. 이는 본 연구의 목적과 관련이 없으므로 검색을 수행할 때 핵융합과 관련된 내용이 제외되도록 설정하였다. 추출된 데이터는 네이버 뉴스 및 신문기사 48,298건, 블로그 105,038건 이다.

3. 데이터 전처리 및 분석

본 연구에서 분석을 위해 사용하는 데이터는 글제목과 본문이고, 취득된 데이터는 작성된 기간별로 5개의 파일로 나누어 저장하였다. 이를 파이썬 KoNLPy 라이브러리를 이용하여 글제목과 본문에 대한 형태소 분석을 수행하였고, 문장에 포함된 어미, 조제어, 구두점, 동사, 특수문자를 제외하였다. 이렇게 생성된 키워드들을 빈도수 별로 정리하였고, 상위 200개의 핵심키워드를 추출하여 워드클라우드 방법으로 시각화하였다.

IV. 분석결과

1. 융합연구 소셜 데이터 경향 분석

시간에 따른 키워드 변화 경향을 파악하기 위하여 분석 기간을 5기(2009~2010년: 1기, 2011~2012년: 2기, 2013~2014년: 3기, 2015년~2016년: 4기, 2017년~2018년: 5기)로 구분하여 분석하였다. 뉴스 및 신문기사에서 “융합연구”를 키워드로 검색한 결과를 이용하여 워드클라우드를 제작하였다[그림 1].

1기에서 뚜렷하게 나타나는 키워드는 융합기술, 기업, 산업, 융합, 미래, 한국, 분야, 연구, 기술 개발, 바이오 등이다. 키워드 ‘한국’이 많이 나타나는 이유는 정부 출연연구소 기관명에 ‘한국’이 들어가서 나타나는 현상

으로 보인다. 즉, 1기에서는 융합연구와 관련하여 정부 출연연구소에서 활발하게 수행하였음을 알 수 있다. 또한 바이오 분야에서 융합연구를 활발하게 수행했다는 것을 확인할 수 있다.

2기 워드클라우드에서 뚜렷하게 나타나는 키워드는 융합 기술, 융합 연구, 융합과학 기술, 기술 대학원, 교수, 서울대 융합과학 등이다. 2011년 6월 안철수 박사의 서울대 융합과학기술대학원 교수 임용에 대한 내용과 2012년 9월 안철수 서울대 융합과학기술대학장 대선출마에 대한 내용이 뉴스 및 신문기사에 많이 포함되어 있어서 나타나는 현상이다. 2기 워드클라우드에서는 융합연구에 대한 경향을 파악하기 어렵다.

3기 워드클라우드에서 뚜렷하게 나타나는 키워드는 융합연구, 융합 기술, 연구 센터, 기술 연구원, 교수, 산학 융합, 시험 연구원, 나노 융합, 화학 융합 등이다. 나노 분야와 화학 분야에 대한 융합연구가 매우 활발하게 진행되고 있음을 알 수 있다. 또한 지역별로 산학융합지구 조성 사업이 진행되고 있음을 추측 할 수 있다.

4기 워드클라우드에서 뚜렷하게 나타나는 키워드는 융합 연구, 연구 센터, 융합 기술, 기술 연구원, 융합 연구원, 시험 연구원, 화학 융합, 산학 융합, 로봇 융합 등이다. 3기에 이어서 화학 분야에 대한 융합이 활발하게 진행되고 있으며, 산학융합지구 조성 사업 또한 지속적으로 진행되고 있음을 알 수 있다. 4기에서 새롭게 등장한 키워드가 로봇이다. 4차 산업혁명의 핵심산업으로 로봇이 주목 받으면서 타 분야와 융합된 로봇의 실용화 연구가 진행되기 시작했다는 것을 확인할 수 있다.

5기 워드클라우드에서 뚜렷하게 나타나는 키워드는

융합 연구, 연구 센터, 융합 기술, 기술 연구원, 한국 화학, 융합 시험, 인공 지능, 나노 융합, 로봇 융합, 바이오 융합이다. 5기에서 새롭게 등장한 단어가 인공 지능이다. 4차 산업혁명에서 가장 중요한 키워드인 인공지능이 바이오로봇과 국방 등 타 분야와 융합하기 시작했다는 것을 확인할 수 있다. 뉴스 및 신문기사 워드클라우드에서는 가장 빈번하게 발생하는 단어들이 크게 나타나기 때문에 중요 키워드를 한 눈에 확인할 수 있지만, 본 연구에서 파악하고자 하는 연구동향을 파악하기 어렵다. 따라서 뉴스 및 신문기사에서 가장 빈번하게 나타나는 연구분야 관련 키워드 순위는 [표 1]과 같다.

표 1. 기간별 뉴스 및 신문기사에서 나타나는 연구분야 키워드 순위

순위	1기	2기	3기	4기	5기
1	바이오	화학	나노	화학	화학
2	녹색성장	나노	화학	로봇	인공지능
3	로봇	스마트	정보통신	나노	나노
4	문화	문화	바이오	정보통신	로봇
5	스마트	정보통신	로봇	생명	바이오
6	에너지	에너지	디자인	의료	스마트
7	환경	바이오	스마트	문화	에너지
8	디자인	의료	환경	빅데이터	해양
9	생명	로봇	소재	원자력	의료
10	의료	디자인	의료	에너지	지질

[표 1]을 살펴보면 1기에서 빈번하게 나타나는 연구 분야 관련 핵심 키워드는 바이오, 녹색성장, 로봇, 문화, 스마트 등이다. 워드클라우드에서와 마찬가지로 1기에 바이오 분야에 대한 융합연구가 가장 활발하게 진행되



그림 1. 뉴스 및 신문기사 워드클라우드

었다는 것을 파악할 수 있다. 또한 1기에 정부에서 녹색 성장 정책을 강조함에 따라 친환경과 관련된 융합연구가 많이 수행되었음을 파악할 수 있다[21].

2기에서 빈번하게 나타나는 연구분야 관련 핵심 키워드는 화학, 나노, 스마트, 문화, 정보통신 등이다. 화학과 나노 분야에서 융합연구가 2기에 본격적으로 진행되었다는 것을 확인할 수 있다. 스마트 기술은 IT기술과 타 분야와의 융합을 의미한다. 1기에 이어서 2기까지 스마트 기술에 대한 융합연구가 이루어지고 있음을 확인할 수 있다.

3기에서 빈번하게 나타나는 연구분야 관련 핵심 키워드는 나노, 화학, 정보통신, 바이오, 로봇 등이다. 2기에 이어서 3기에도 나노와 화학 분야에 대한 융합연구가 활발하게 진행되고 있음을 확인할 수 있다. 3기에는 로봇 분야에 대한 융합연구가 다시 부각되고 있다.

4기에서 빈번하게 나타나는 연구분야 관련 핵심 키워드는 화학, 로봇, 나노, 정보통신, 생명 등이다. 워드클라우드에서와 마찬가지로 4기에 들어서 4차 산업혁명이 부각되면서 로봇 분야에 대한 융합연구가 활발하게 진행되고 있음을 알 수 있다.

5기에서 빈번하게 나타나는 연구분야 관련 핵심 키워드는 화학, 인공지능, 나노, 로봇, 바이오 등이다. 5기에 들어서 인공지능과 관련된 융합연구가 활발해지기 시작했다는 것을 알 수 있다. 화학과 나노는 2기부터 5기까지 지속적으로 나타나고 있다. 국내 융합연구와 관련해서 가장 중요한 분야가 화학과 나노임을 확인할 수 있다.

뉴스와 신문기사가 객관적인 사실을 기반으로 작성

되는 것이라면, 블로그는 개인의 주관적인 생각을 공유하거나 특정 기관에서 홍보하는 공간이다. 하지만 융합연구는 대중들에게 관심있는 키워드가 아니기 때문에 본 연구에서는 블로그에 나타나는 융합연구와 관련한 내용을 특정 기관에서 대중들에게 홍보의 용도로 사용하였다고 간주하였다. 또한 대중들과 상호작용을 수행하면서 대중들의 참여를 유도하는 블로그 뉴스 내용도 포함하고 있기 때문에 뉴스 및 신문기사와는 다른 관점에서 융합연구와 관련된 동향파악이 가능하다[20]. 네이버 블로그에서 “융합연구”를 키워드로 검색한 결과를 이용하여 워드클라우드를 제작하였다(그림 2).

1기 워드클라우드에서 뚜렷하게 나타나는 키워드는 연구, 융합, 융합기술, 한국, 기술 융합, 융합 연구, 기술 개발, 교육, 교수, 기업, 문화, 정보통신, 녹색성장 등이다. 뉴스 및 신문기사에서와 동일하게 블로그에서도 키워드 ‘한국’이 들어간 정부출연연구소와 공공기관에서 진행하고 있는 융합연구에 대한 홍보를 많이 하고 있는 것으로 추측한다. 키워드 ‘기업’은 기업과 공동으로 융합연구를 수행하고 있다는 것으로 보인다. 키워드 ‘문화’는 과학과 문화의 융합연구에 대한 내용과 융합연구를 지향하는 연구문화에 대한 홍보가 블로그에서 이루어지고 있음을 확인할 수 있다.

2기 워드클라우드에서 뚜렷하게 나타나는 키워드는 연구, 융합, 융합기술, 융합연구, 기술융합, 한국, 교수, 교육, 문화, 디자인 등이다. 1기와 비교해서 융합 디자인에 대한 홍보가 증가하고 있는 것으로 보인다. 하지만 전체적으로 1기와 비교했을 때 빈도 순위만 조금 바뀌었을 뿐 키워드가 매우 유사하게 나타난다. 따라서



그림 2. 블로그 워드클라우드

블로그에 나타나는 내용이 1기와 매우 유사한 것으로 보인다.

3기 워드클라우드에서 뚜렷하게 나타나는 키워드는 연구, 융합연구, 융합기술, 교육, 기술융합, 융합, 연구센터, 교수, 미래, 문화, 나노융합, 정보통신 등이다. 키워드 ‘융합연구센터’는 대학 내에 있는 연구기관 중에서 기관명에 ‘융합연구센터’가 들어있는 연구기관의 융합연구관련 홍보가 많이 이루어지고 있음을 나타낸다.

4기 워드클라우드에서 뚜렷하게 나타나는 키워드는 연구, 융합연구, 융합기술, 기술융합, 연구센터, 교수, 한국, 교육, 대학, 기술개발, 스마트 등이다. 4기에서 IT와 타분야가 융합한 스마트 기술에 대한 홍보가 본격적으로 이루어지고 있음을 추측할 수 있다.

5기 워드클라우드에서 뚜렷하게 나타나는 키워드는 융합연구, 연구, 융합기술, 연구센터, 기술융합, 교수, 창의 융합, 산업혁명, 교육, 인공지능 등이다. 5기에 들어와서 4차 산업혁명과 관련하여 인공지능 융합연구에 대한 사례 홍보가 많이 이루어지고 있음을 추측할 수 있다. 1기부터 5기까지 융합교육에 대하여 홍보하는 내용이 블로그에 많이 게시되어 있을 것으로 보인다. 또한 키워드 ‘교수’가 모든 기간에서 빈도가 매우 높게 나타나는 핵심키워드인 것으로 보아 정부출연연구소가 아닌 대학에서 융합연구를 한 사례를 많이 홍보하고 있는 것으로 보인다.

네이버 블로그 검색결과에서 융합연구와 관련된 연구분야를 빈도가 높은 순으로 도시하였다[표 2].

표 2. 기간별 블로그에서 나타나는 연구분야 순위

순위	1기	2기	3기	4기	5기
1	문화	문화	문화	스마트	인공지능
2	정보통신	디자인	디자인	나노	스마트
3	녹색성장	나노	나노	문화	나노
4	로봇	환경	정보통신	로봇	바이오
5	에너지	스마트	환경	정보통신	정보통신
6	정보통신	에너지	로봇	에너지	수학
7	환경	인문학	에너지	의료	에너지
8	디자인	의료	바이오	빅데이터	화학
9	바이오	생명	인문학	화학	빅데이터
10	의학	정보통신	화학	생명	문화

뉴스 및 신문기사에서와는 다르게 1기부터 3기까지 문화에 대한 융합연구가 많이 홍보되고 있음을 확인할 수 있다. 또한 2기와 3기에는 융합 디자인에 대한 사례도 많이 홍보 되었을 것으로 추측한다. 뉴스 및 신문기사와 다르게 화학 분야에 대한 융합연구 사례가 많이 홍보되지 않고 있음을 확인할 수 있다. 4기에 들어와서 스마트 기술의 융합연구 사례 홍보가 급격하게 늘어났고, 5기에 들어와서 인공지능에 대한 융합연구 사례 홍보가 급격하게 증가했을 것으로 보인다. 나노와 관련된 융합연구 내용은 뉴스 및 신문기사와 마찬가지로 2기부터 5기까지 많이 홍보되고 있음을 알 수 있다.

2. 융합연구 소셜 데이터 네트워크 분석

워드클라우드에서는 빈도가 높은 키워드가 크게 나타나기 때문에 연구분야처럼 상대적으로 빈도가 적은 키워드들을 찾아내는 것은 매우 어렵다. 또한 각 기간 별로 빈도가 높은 키워드를 찾는 것만으로는 기간 전체에 대한 전반적인 경향을 파악하는 것도 쉽지 않다. 따라서 본 연구에서는 네트워크 분석을 통하여 융합연구에 대한 전반적인 경향을 파악하고자 하였다. 뉴스 및 신문기사의 워드클라우드를 제작하는데 사용된 데이터를 이용하여 1기부터 5기에 대한 네트워크 분석을 수행하였다. [그림 3]에서 노드 1기부터 5기까지, 이 5개 노드의 안 쪽에 위치할수록 국내 융합연구에서 일반적인 경향을 나타낸다. 연구분야와 관련된 단어는 스마트, 로봇, 나노, 화학, 의료가 여기에 속한다. 키워드 ‘바이오’는 1기, 3기, 4기와 연결되어 있으나 2기에 ‘바이오 센터’ 키워드가 있는 것으로 보아 1기부터 4기까지 생명 공학과 관련된 연구가 활발하게 진행되었을 것으로 추측한다. 정부정책과 관련해서는 ‘인재 양성’, ‘산업 육성’, ‘융합 산업’, ‘교육’ 등이 있다. 1기부터 5기까지 융합연구와 관련하여 인재와 산업 등을 육성하고자 하는 정부의 의지를 확인할 수 있다. 2기부터 5기까지는 키워드 ‘융합 교육’이 존재한다. 인재 육성 정책과 관련하여 융합교육을 지속적으로 수행하고 있음을 확인할 수 있다.

반면에 5개 노드의 바깥 쪽에 위치할 경우 특정기간에만 활발하게 진행된 융합연구 분야라고 볼 수 있다.

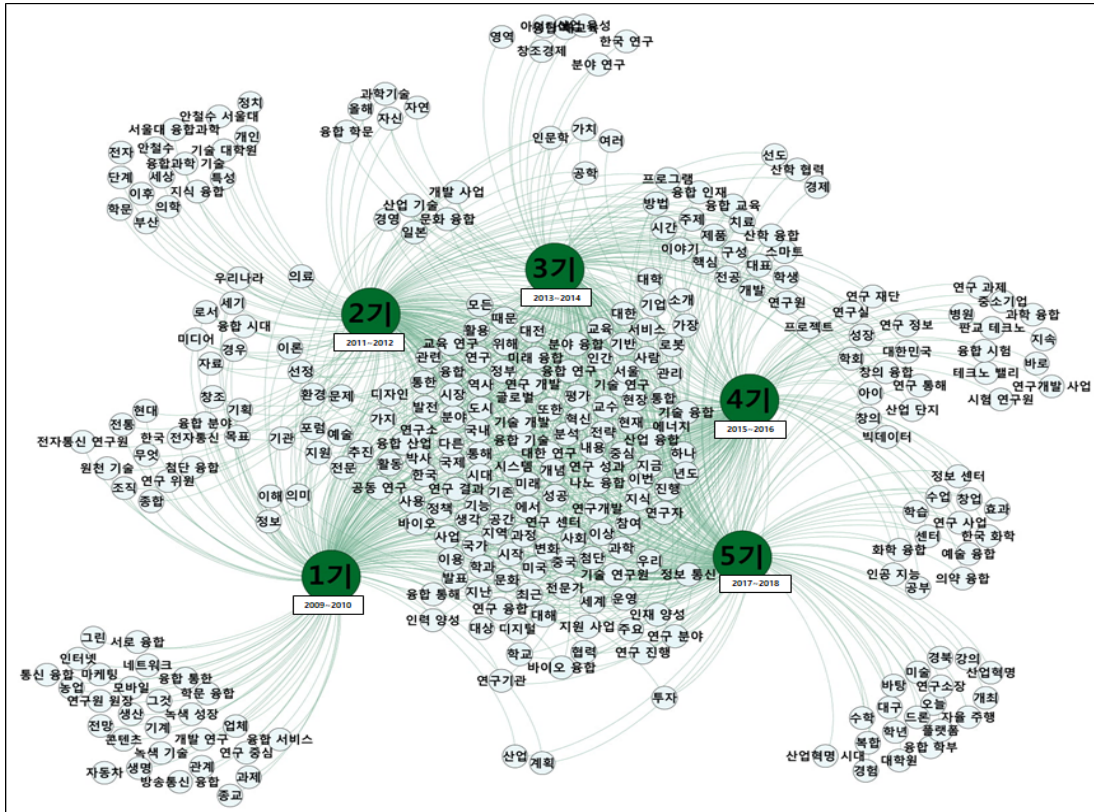


그림 4. 블로그 네트워크 분석

기와 2기에 나타나는 융합연구와 관련된 연구분야 키워드는 ‘의료’이다. 의료분야는 신문 및 뉴스기사에서는 부각되지 않은 연구분야로 1기와 2기에 이와 관련된 융합연구 내용이 많이 홍보되고 있음을 알 수 있다. 3기와 4기에 융합연구와 관련하여 나타나는 키워드는 ‘스마트’이다. IT 기술과 융합하는 연구가 3기와 4기에 많이 이루어 졌음을 알 수 있다. 4기와 5기에 나타나는 키워드는 ‘빅데이터’, ‘화학 융합’, ‘예술 융합’, ‘인공 지능’, ‘의약 융합’이다. 뉴스 및 신문기사에서 드러나지 않은 빅데이터, 예술, 의약학 분야의 융합연구와 관련된 내용이 홍보되고 있음을 알 수 있다. 뉴스 및 신문기사에서는 화학 분야가 1기부터 5기까지 전체적으로 나타나지만 블로그에서는 4기와 5기에서만 나타난다. 5기에서만 나타나는 키워드는 ‘드론’, ‘자율 주행’, ‘산업혁명’이다. 4차 산업혁명과 관련하여 자율주행 자동차와 드론에

대한 융합연구가 블로그에 홍보되고 있음을 추측할 수 있다.

정부정책과 관련해서는 1기부터 5기까지 ‘인재 양성’과 ‘인력 양성’이 나타나며, 1기에는 ‘녹색 기술’, 3기에서는 ‘창조경제’, 4기와 5기에서는 ‘창의 융합’, ‘중소기업’이 나타나고 있다. 블로그에서도 융합연구와 관련하여 인재양성에 대한 내용이 많이 홍보되고 있다. 4기와 5기에 나타나는 키워드 ‘창의 융합’은 창의융합형 인재 양성 정부정책에 대한 홍보가 4기와 5기에 활발하게 이루어 졌음을 알 수 있다.

V. 논의

본 연구에서는 네이버 뉴스 및 신문기사와 블로그에

서 키워드 ‘융합연구’로 검색한 결과를 바탕으로 워드클라우드 분석 및 네트워크 분석을 수행하였다.

첫 번째 연구문제로 융합연구에 관한 핵심 키워드가 시간에 따라 어떻게 변화하는지 파악하였다. 기간별 키워드 순위를 보았을 때, 각 기간별로 뉴스 및 신문기사에서 10위 안에 들어가는 연구분야를 살펴보면, 1기는 바이오, 친환경, 로봇, 문화, 스마트이고, 2기는 화학, 나노, 스마트, 문화, 정보통신이다. 3기는 나노, 화학, 정보통신, 바이오, 로봇이고, 4기는 화학, 로봇, 나노, 정보통신, 바이오이다. 5기는 화학, 인공지능, 나노, 로봇, 바이오이다. 블로그의 경우, 1기는 문화, 정보통신, 친환경, 로봇, 에너지이고, 2기는 문화, 나노, 환경, 스마트이다. 3기는 문화, 디자인, 나노, 정보통신, 환경이고, 4기는 스마트, 나노, 문화, 로봇, 정보통신이다. 5기는 인공지능, 스마트, 나노, 바이오, 정보통신이다. 뉴스 및 신문기사에서는 이공계열 위주의 융합연구 동향을 보여주고 있지만 블로그에서는 문화와 디자인과 같은 인문학과 관련된 융합연구 동향이 많이 나타나고 있다. 하지만 5기에서는 4차 산업혁명의 영향으로 이공계와 관련된 융합연구 동향을 보여주고 있다.

블로그와 뉴스 및 신문기사에 공통으로 10위 안에 들어가는 연구분야를 살펴보면, 1기는 친환경, 로봇, 문화이고, 2기는 스마트와 문화이다. 3기는 나노와 정보통신이고 4기는 로봇, 나노, 정보통신이다. 5기는 인공지능, 나노, 바이오이다. 이 분야들은 각 기간 동안 활발하게 수행된 연구분야라고 볼 수 있다. 1기에서는 정부의 녹색성장 정책으로 인하여 친환경 기술에 대한 융합연구가 활발하게 진행되었을 것으로 추측할 수 있다. 2기에서 공통으로 나타난 키워드 ‘스마트’는 IT기술이 발전하면서 다양한 분야에 접목되는 융합연구가 활발하게 진행될 시기임을 확인할 수 있다. 1기부터 4기까지 등장하지 않았던 인공지능 분야가 5기에서 등장하는 것은 4차 산업혁명 시대를 대표하는 핵심기술인 인공지능과 관련된 융합연구가 활발하게 진행되고 있음을 추측할 수 있다. 또한 소셜미디어 상에서 과학기술 중심의 융합연구에 대한 보도가 이루어지고 있으며, 인문사회기반의 융합연구는 실질적으로 나타나지 않음을 알 수 있다. 이는 최근 융합연구에 있어 인문사회과학의

중요성이 높아지고 있음을 감안할 때 더욱 홍보가 필요한 부분이라 할 수 있다.

두 번째 연구문제로 융합연구 핵심키워드 네트워크는 어떤 구조를 보이는지 파악하였다. 블로그와 뉴스 및 신문기사 모두 핵심키워드의 절반이상이 적어도 2개 기간과 연결되어 있으며, 모든 기간과 연결된 핵심키워드가 전체의 20% 이상이다. 이는 융합연구 핵심키워드 네트워크 구조가 여러 기간과 고르게 연계되어 있는 분산형 구조라는 것을 의미한다. 따라서 시각화된 네트워크 구조를 보면 전기간에 걸쳐 전반적으로 수행된 연구분야 혹은 특정기간에만 수행된 연구분야를 쉽게 파악할 수 있다. 또한 네트워크 내에 존재하는 요소들 중에 융합연구와 관련된 정부정책 키워드들이 존재하기 때문에 이와 관련된 동향 파악이 가능하다.

예를들면, 인재양성, 산업육성, 그리고 융합교육은 1기부터 5기까지 5개 기간과 연결되어있다. 이 키워드들은 10년간 전반적으로 수행된 정부정책임을 알 수 있다. 또한 블로그와 뉴스 및 신문기사 모두 1기와 2기에 키워드 ‘융합 시대’가 연결되어있다. 이는 1기와 2기에 국내에서 본격적으로 융합연구를 수행했다고 볼 수 있다. 또한 블로그에서는 5기에만 키워드 ‘융합 학부’가 나타난다. 이는 융합연구가 활발하게 진행되어 5기에 대학에서 융합 교육을 수행할 수 있는 기반이 마련되었다고 볼 수 있다.

VI. 결론 및 제언

본 연구에서는 2009년 1월부터 2018년 9월까지 약 10년 동안 수행된 융합연구 동향을 파악하고자 하였다. 2년 단위로 기수를 잡았을 때, 1기의 대표적인 융합연구 분야는 친환경이고, 2기는 스마트, 3기는 정보통신, 4기는 로봇, 5기는 인공지능이다. 1기, 2기, 5기는 정부정책으로 인하여 나타나는 동향으로 볼 수 있다. 이는 융합연구를 수행하는데 있어 특정분야를 육성하고자 하는 정부의 정책이 많이 반영되고 있음을 나타낸다.

이와는 대조적으로, 융합연구 키워드에 대한 네트워크 분석 결과가 분산형 구조라는 것은 10년 동안 지속

적으로 수행되고 있는 융합연구 분야와 연구기관들이 존재하며, 융합연구와 관련된 정부정책에 있어서도 일관성 있게 진행된 부분이 있다고 볼 수 있다.

하지만 정부정책에 따라 주로 과학기술에 집중적으로 투자하고 있는 융합연구 현황을 볼 때, 미래 융합연구의 발전을 위하여 인문사회분야와 과학기술분야의 융합연구에 많은 투자가 필요하다. 또한 소셜미디어 상에서 나타나는 연구기관 관련 키워드로 보았을 때, 정부출연연구소와 같은 특정 기관에서 수행된 융합연구 위주로 많이 나타나고 있다. 이는 융합연구와 관련하여 연구비가 특정 연구기관에 편중되어 있음을 시사하고 있으며, 융합연구 주제의 다양성과 많은 연구자들에게 융합연구의 기회를 부여하기 위해서 다양한 연구기관에 연구비를 지원해야 한다.

본 연구에서는 블로그와 뉴스 및 신문기사에서 나타나는 융합연구 관련 데이터를 수집하였다. 블로그에서는 문화와 디자인과 같이 예체능 계열에 대한 융합연구를 많이 홍보하고 있으며, 뉴스에서는 화학, IT 같은 공학 계열에 대한 융합연구 성과를 많이 보여주고 있다. 이러한 점은 워드클라우드와 네트워크 분석결과를 해석하여 동향분석을 수행하는데 있어 어려움을 준다. 또한 키워드 ‘생명’과 ‘바이오’는 같은 연구분야를 나타내는 핵심키워드인데, 이와 같은 연구분야 키워드에 대한 동의어 처리를 하지 않음으로써 연구분야 키워드 순위를 정확하게 파악하기 어려웠다. 그리고 소셜 미디어에서 나타나는 검색결과가 항상 융합연구에 대한 내용을 나타내는 것이 아니기 때문에 간접적으로 동향을 파악했다는 점이 본 연구의 한계라고 볼 수 있다.

따라서 추후 연구에서는 검색키워드를 ‘융합연구’가 아닌 ‘인문사회 융합연구’ 혹은 ‘이공계 융합연구’와 같이 세부분야를 명확하게 구분하고, 이음동의어에 대한 전처리를 수행함으로써 융합연구에 대한 동향을 보다 세부적으로 분석하고자 한다.

참 고 문 헌

[1] 경정운, 김영우, 한승환, “융합연구의 새로운 발전

방향 모색: NSF의 환경 속의 생명복합성과 NRF의 학제간 융합연구 지원사업 비교,” 한국정책과 학회보, 제15권, 제1호, pp.151-177, 2011.

[2] 노영희, 이광희, 정대근, “인문사회기반 융합연구 성과 영향요인 연구,” 한국콘텐츠학회논문지, 제18권, 제6호, pp.667-678, 2018.

[3] 오현석, 배형준, 김도연, “과학기술분야 융합연구자의 융합연구 입문과 과정에 관한 연구,” 아시아 교육연구, 제13권, 제4호, pp.297-335, 2012.

[4] D. H. Kim, “Analysis of Perception Difference in R&D Planning and Evaluation Following Characteristics of Convergence Researcher: Focused on Humanities, Social Sciences, and Science and Technology Majors and Areas,” The Journal of Transdisciplinary, Vol.2, No.1, pp.5-28, 2018.

[5] 이오준, 박승보, 정다울, 유은순, “소셜 빅데이터를 이용한 영화 흥행 요인 분석,” 한국콘텐츠학회논문지, 제14권, 제10호, pp.527-538, 2014.

[6] 오효정, 윤보현, 최남현, 유철중, 김용, “소셜 빅데이터 내용 분석 기반 사용자 그룹별 선호지역 및 이동패턴 시각화,” 한국정보기술학회논문지, 제12권, 제12호, pp.195-203, 2014.

[7] 이수진, 전유나, “소셜미디어 빅데이터 분석을 통한 관광 인식 조사: 경기도 지역을 중심으로,” GRI 연구논총, 제18권, 제1호, pp.83-109, 2016.

[8] T. M. Joo and C. E. Teng, “Impacts of Social Media (Facebook) on Human Communication and Relationships: A View on Behavioral Change and Social Unity,” International Journal of Knowledge Content Development & Technology, Vol.7, No.4, pp.27-50, 2017.

[9] 김대경, 이현수, “소셜 빅데이터 텍스트 마이닝을 활용한 전국장애인체육대회 분석 연구,” 한국체육학회지, 제55권, 제6호, pp.801-810, 2016.

[10] 서운채, 김형중, “직무만족도에 영향을 미치는 내부평판 요인에 관한 연구: 기업정보 제공 소셜 미디어 빅데이터를 중심으로,” 한국디지털콘텐츠

학회 논문지, 제17권, 제4호, pp.295-305, 2016.

- [11] 오효정, 안승권, 김용, “소셜 빅데이터 기반 사회적 이슈 리스크 유형 분류,” 한국콘텐츠학회논문지, 제16권, 제8호, pp.1-9, 2016.
- [12] 박성진, 이수원, 황영찬, “소셜 빅데이터 기반 2016리우올림픽 축구 관련 이슈 및 인물에 대한 연관단어 분석,” 한국체육학회지, 제56권, 제2호, pp.303-320, 2017.
- [13] 나경식, 이지수, “신문 빅데이터를 바탕으로 본 국내 정보화의 경향과 도서관의 역할,” 한국콘텐츠학회논문지, 제18권, 제9호, pp.14-33, 2018.
- [14] 김종범, “융합연구사업과 항공우주기술,” 항공우주산업기술동향, 제12권, 제2호, pp.3-9, 2014.
- [15] 이태희, 정영철, *의료분야에서의 정보기술 융합 연구 동향과 시사점*, 보건복지포럼, 2014.
- [16] 최하영, 장재혁, 윤장혁, “NRF 연구제안서 분석을 통한 한국의 융합 R&D 동향 분석,” 대한산업공학회 추계학술대회 논문집, pp.986-989, 2016.
- [17] 장재혁, 윤장혁, “한국의 인문사회과학기술 융합 연구의 동향분석,” 한국경영과학회 학술대회논문집, pp.2326-2332, 2017.
- [18] 이재익, “플렉서블 디스플레이 R&D기술 디자인 융합연구,” 한국과학예술포럼, 제18권, pp.519-529, 2014.
- [19] 김정운, 김영우, 한승환, “융합연구의 새로운 발전방향 모색: NSF의 환경 속의 생명복합성과 NRF의 학제간 융합연구 지원사업 비교,” 한국정책과학학회보, 제15권, 제1호, pp.151-177, 2011.
- [20] 손주연, 이장재, 김시정, “녹색성장 정책의 변화: 정책네트워크 분석을 중심으로,” 기술혁신학회지, 제18권, 제3호, pp.516-538, 2015.
- [21] 김은미, 김예란, 임영호, 김경모, “블로그 뉴스의 속성과 이용자의 반응,” 방송통신연구, 제74호, pp.9-41, 2011.

저 자 소 개

노 영 희(Younghee Noh)

정회원



- 2000년 2월 : 연세대학교 문학박사
- 2009년 6월 ~ 현재 : 건국대학교 문헌정보학과 교수
- 2011년 3월 : 지식콘텐츠연구소 소장
- 2017년 9월 ~ 현재 : 융합연구

총괄센터 센터장

<관심분야> : 융합연구성과, 지식콘텐츠, 정보시스템

김 태 연(Taeyoun Kim)

정회원

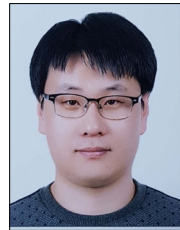


- 2018년 2월 : 과학기술연합대학원대학교 석유자원공학 박사
- 2018년 10월 ~ 현재 : 지식콘텐츠연구소 전임연구원

<관심분야> : 융합연구, 국제공동연구, 빅데이터

정 대 근(Dae Keun Jeong)

정회원



- 2014년 2월 : 전남대학교 문헌정보학 박사
- 2014년 9월 ~ 현재 : 전남대학교 문헌정보학과 강사
- 2018년 2월 ~ 현재 : 융합연구 총괄센터 부센터장

<관심분야> : 융합연구, 도서관경영, 청소년독서

이 광 희(Kwang Hee Lee)

정회원



- 2000년 2월 : 서울대학교 행정학 박사
- 2009년 6월 ~ 현재 : 한국연구재단 인재양성실 팀장
- 2006년 4월 : 한국학술진흥재단 학술정책단 연구위원

▪ 2004년 10월 : 행정자치부 전자정부/행정개혁본부 전문위원(가급)

<관심분야> : R&D정책, 전자정부, 도시 및 지역정책