

소셜 네트워크 분석을 이용한 4차 산업혁명 기술 분야의 연구 동향 분석

The Analysis of Research Trends in Technology to the Fourth Industrial Revolution using SNA

김홍광* · 안종욱**
Kim, Hong-Gwang · Ahn, Jong-Wook

Abstract

The fourth industrial revolution technology focused on the fusion of infrastructure and various advanced technologies related city. Therefore, technical cooperation in various fields of research is essential. In order to activating the fourth industrial revolution technologies, it is necessary to research the state of technology in various fields. Consequently, this paper aims to analysis of domestic and foreign research trends on technology to the fourth industrial revolution using SNA and text mining for web site. We collected text, date data of research paper and report in web site for five years, that is, from January 1st in 2014 to December 31st in 2018. Next, we have deduced the major keywords in public data through analyzing the morphemes. Then we have analyzed the core and related keyword lists through an SNA. In Korea, the focus is on R&D and legal/institutional solution in relation to the fourth industrial revolution technology. On the other hand, in the case of foreign, there was focus on practical technologies for urban services in detail aspects.

Keywords: SNA, Text Mining, IoT, Big Data, AI

1. 서 론

우리나라는 ‘4차 산업혁명 위원회의 설치 및 운영에 관한 규정’에 따라 2017년 10월에 4차 산업혁명 위원회를 출범하고, 사람 중심의 혁신 성장을 목표로 하여 4차 산업혁명 관련 정책을 추진하고 있다. 이 정책은 다양한 산업 창출과 주력 산업 육성, 사회 문제 해결에

따른 국민의 삶의 질 제고, 새로운 일자리 창출 등을 토대로 사회 안전망 강화, 지능화 기술, 데이터 및 네트워크 확보를 기본 방향으로 설정하였다.

최근에는 4차 산업혁명을 대표하는 핵심 기술인 사물인터넷, 빅데이터, 인공지능 등이 비약적으로 발전하기 시작하면서 상호 연결된 기술과 다양한 플랫폼을 토대로 사물-인간을 연결하는 패러다임이 새롭게

* 건국대학교 행정문제연구소 초빙연구원 (first author: woosim99999@hanmail.net)

** 안양대학교 도시정보공학과 교수 Department of Urban Information Engineering, Anyang University (corresponding author: ajw0603@anyang.ac.kr)

창출되고 있다.

이러한 4차 산업혁명 기술은 도로, 건축물, 교량 등 도시를 구성하는 각종 물리적 인프라에 정보 통신 기반의 다양한 첨단 기술을 융합하는 것에 기반을 두고 있기 때문에 여러 연구 분야 간의 기술 협력이 반드시 필요하다. 따라서 4차 산업혁명 기술들의 실질적인 활성화를 위해서는 어떠한 대상에 특정 기술이 적용 또는 연구되고 있는가에 대한 현황과 연구 분야에 대한 고찰이 필요하다(Lim et al. 2014).

이러한 배경을 바탕으로 본 연구는 텍스트 마이닝(Text mining) 및 소셜 네트워크 분석(Social network analysis) 기법을 이용하여 사물인터넷, 빅데이터, AI를 대상으로 국내·외 연구 현황 및 동향을 정량적으로 분석하고, 국내·외 기술 동향의 차이점에 대한 시사점을 제시하고자 한다. 본 연구는 4차 산업혁명 관련 기술의 연구 동향에 대한 이해를 도모함으로써, 연구자들에게 학술적 관점에서의 통찰 기회를 제공할 수 있을 것으로 사료된다.

2. 이론적 고찰

2.1. 텍스트 마이닝

소셜 마이닝은 텍스트 마이닝, 오피니언 마이닝, 소셜 네트워크 분석, 클러스터 분석 등으로 구분된다(Jeong et al. 2016). 본 연구에서는 웹 상에 존재하는 텍스트들을 수집 및 분석하기 위하여 텍스트 마이닝 기법을 활용하였으며, 해당 기법은 비정형 데이터(텍스트 등) 위주의 문서에 다양한 분석 기법을 적용함으로써 데이터에서 지식을 발견 및 도출해내는 기술이다(Jeong et al. 2016). 텍스트 마이닝은 대표적인 정량적 기법으로, 다양한 분야에의 연구 및 기술 동향을 파악하거나 연구 기회의 발견 등에 활용되어 왔다. 일반적인 데이터 마이닝 기법과의 차이점은 데이터 소스가 문서 집합으로 구성되고, 소스의 형태도 정형화

된 데이터베이스에서 찾기 보다는 비정형의 텍스트 데이터를 대상으로 한다는 점이다.

2.2. 동시 출현 단어 분석

동시 출현 단어 분석은 텍스트 마이닝 기법 기반의 단어들 간의 연관성을 파악하고 이를 통해 수집된 문서들의 주요한 이슈를 도출하는 것에 활용된다(Jeong et al. 2016). 단어의 동시 출현은 말뭉치 내에서 두 단어가 같이 나타난 수를 말한다. 주로 Pointwise mutual information (PMI) 방법이 동시 출현 단어 분석법으로 사용된다(Church and Hanks 1990). PMI는 두 단어가 발생할 확률과 두 단어의 결합 확률, 즉 두 단어가 동시에 발생할 확률로 구성된다. 동시 출현 단어 분석은 주로 연구 동향의 파악, 집단 특징/특성 분석과 관련된 연구에 활용되고 있다. 또한 동시 출현 단어 분석법을 사용해 얻어진 단어 쌍들을 시각화하기 위하여 Pathfinder networks (PFNET) 알고리즘을 활용하였다. 이 알고리즘은 특정 객체들 간의 관계를 네트워크화 할 때 덜 중요한 객체들을 제거하고 더 중요한 객체들을 유지시키기 위한 알고리즘으로 복잡한 관계를 주요 네트워크로 단순화하는 방법이다(Schvaneveldt 1990). 대부분의 동시 출현 단어는 행렬 기반의 네트워크로 표현될 수 있으므로, 단어 간의 주요한 관계를 시각화 표현하기 위하여 PFNET이 사용될 수 있다.

2.3. 소셜 네트워크 분석

소셜 네트워크 분석은 사람·조직·사물·사건·개념 등 분석 대상들 간의 관계를 네트워크로 구성하여 네트워크 구조를 계량적으로 분석하는 기법으로, 네트워크를 구성하는 정보들을 과학적으로 분석하여 사회학, 경영학, 문헌 정보학 등 다양한 분야에서 활용하는 복합적 사고의 과학이자 고도의 학제적 연구 영역이

다(Jeong et al. 2016). 특히 텍스트 마이닝에서의 네트워크 분석은 광범위한 텍스트를 토대로 문맥의 의미를 텍스트와의 조합을 통해 파악하는데 활용되며, 텍스트의 시각화를 통해 주요 개념과 관계를 직관적으로 파악 가능하다. 네트워크를 활용한 분석 방법에는 크게 하위 네트워크 분석, 구조적 등위성 분석, 중심성 분석 등이 있으며, 중심성 분석은 크게 연결, 매개, 근접 부문으로 구분된다. 중심성은 영향력이라는 개념으로 해석되기도 하며 일반적으로 네트워크 분석에서 가장 많이 사용되는 분석 단위 또는 분석 지표 가운데 하나이다(Jeong et al. 2016). 연결 중심성은 하나의 노드에 얼마나 많은 노드 간 링크가 연결되어 있는가를 나타내며, 네트워크 상에서 허브 역할을 담당한다. 매개 중심성은 하나의 노드가 네트워크 상에서 기타 노드와의 위치하는 정도를 측정하며, 연결 정도와는 상관 없이 네트워크/그룹 간의 정보 유통이나 소통 역할 등을 담당한다. 근접 중심성은 각 노드들에서 다른 모든 노드들에 대한 거리치 총합의 역수이며, 네트워크 전역에서 일반적인 영향력을 갖는 노드로 정보, 권력, 접근이 쉬운 노드를 찾을 수 있다(Kim et al. 2015).

3. 선행연구 검토

3.1. 선행연구 검토

텍스트 마이닝 분석 기법의 특성을 이용하여 계량 서지학 분석(Bibliometric analysis)에 적용하고 특정 분야의 연구 동향을 파악한 연구들이 다수 존재한다. 이는 서지 정보 검색 방법이 개선되고 텍스트 마이닝 기법이 발전하면서 논문이나 보고서 등에 언급된 키워드를 추출하는 것이 용이하게 변화하였으며, 추출된 키워드 간의 관계를 분석하여 관련 연구 분야에 대한 동향 등의 패턴을 더욱 정밀하게 파악하는 것이 가능하기 때문이다. Patra and Misha(2006)는 생물 정

보학 분야의 연구 동향을 계량서지학 분석 기법을 사용하여 파악하였다. Heo and Song(2013)은 의료 정보학 분야의 저널 중 저자 동시 인용 분석(Author co-citation analysis)과 동시 출현 단어 분석을 통해 의료 정보학 연구의 학문적 성향을 분석하였으며, Cho and Kim(2011)은 학술 분야별 중심 키워드 및 분야 간 융합을 위한 연계 키워드를 도출하여 학술 분야별 복잡한 연관 관계를 분석하였다.

본 연구와 관련 있는 정보 통신 기술을 기반으로 한 융·복합 관련 연구 분야에 관해서도 텍스트 마이닝 기법 적용 연구들이 존재한다. Oh(2015)는 교통 체계가 지능형 교통시스템, 텔레매틱스와 같이 IoT 기술로 주목받으면서 이와 관련된 연구 동향을 정량적으로 분석하기 위해 텍스트 마이닝 기법을 활용하였다. 여기서는 'ICT'와 '교통' 등 두 가지 키워드가 동시에 출현하는 국내에서의 논문 초록을 추출하고 그 결과에 토픽 모델링 기법을 적용하여 연구 토픽의 시계열 분석을 통해 분야 간의 융합을 위한 필요 연구들을 제시하였다. Kim et al.(2015)은 키워드 네트워크 분석을 이용하여 융·복합 정보인 공간정보의 글로벌 연구 동향을 분석하였다. 이 연구에서는 'Web of Science' 검색 엔진을 이용해 공간정보와 관련된 논문 초록의 주요 키워드를 도출하고 이들 간의 연결, 근접, 매개 중심성을 분석하여 미래의 유망한 공간정보 융합 기술 항목을 제시하였다.

3.2. 연구의 방법

텍스트 마이닝 및 소셜 네트워크 분석 기법은 다양한 분야의 연구 동향을 정량적으로 파악하고 근 미래의 발전 방향을 예측하며, 이를 통한 후속 연구를 도출하는 취지에서 매우 폭넓게 활용되고 있다. 본 연구에서는 여러 가지 도구 중 빅데이터 분석 프로그램으로서 활용도가 높은 R을 이용하였다. R은 통계와 그래픽 시각화를 위한 프로그래밍 언어로서, 빅데이터 이슈

및 Comprehensive R archive network (CRAN)을 통한 패키지 배포 및 공유를 토대로 여러 분야에서의 사용이 증가하고 있다. R 패키지를 통한 Web과 SNS 데이터 연동과 Word Cloud를 통한 시각화 기법은 데이터를 분석하고 표현하기에 최적의 도구로서 인식되는 추세다. 다음으로 소셜 네트워크 분석을 수행하기 위해서 NodeXL 1.01을 사용하였으며, NodeXL은 Microsoft Excel에 키워드 네트워크 분석 기능을 추가한 도구이다. 소셜 네트워크 분석에 필요한 주요 함수와 분석 결과의 시각화를 위한 각종 차트 기능을 지원한다.

본 연구의 분석 절차는 Figure 1과 같다. 텍스트 마이닝을 위해 ‘사물인터넷/IoT’, ‘빅데이터/Big data’, ‘인공지능/AI’ 등의 용어 및 유사 용어를 포함하여 검색어를 설정하였다. 이후 ‘네이버 학술정보’에서 국내·외 논문을 R 기반의 크롤링(Crawling) 기법을 이용하여 수집 및 저장하였다(2014년~2018년). 국외 논문 수집과 관련하여 google scholar 등 외국 저널 사이트를 포함하지 않은 점은 연구 한계점으로 남는다. 국내의 경우 검색어로 ‘사물인터넷’을 적용했을 때 558편의 논문이 검색되었으며, ‘빅데이터’는 6,434편, ‘인공지능’은 3,039편이 검색되었다. 국외의 경우 검색어로 ‘IoT’를 적용했을 때 86,439편이 검색되었으며, ‘Big data’는 69,010편, ‘AI’는 86,439편이 검색되었다. 이를 종합하면 국내 논문 10,031편, 국외 논문 164,804편을 연구 대상으로 선정하였다.

다음으로 연구 대상이 되는 논문의 제목, 발간일 등의 항목을 기준으로 텍스트를 추출하였다. 이를 토대로 국내·외 논문의 연도별 발간 추이를 살펴보았다. 이후 제목에 대한 형태소 분석을 통해 제목 내 단어들을 우선 추출하여, 비정상 및 품사 중 조사 등의 단어들을 제거하거나 유사 단어를 대표 단어로 치환하는 등의 작업을 수행하였다. 또한 유사 단어(‘사물인터넷’, ‘IoT’ 등)를 통해 수집된 논문의 경우에는 논문 제

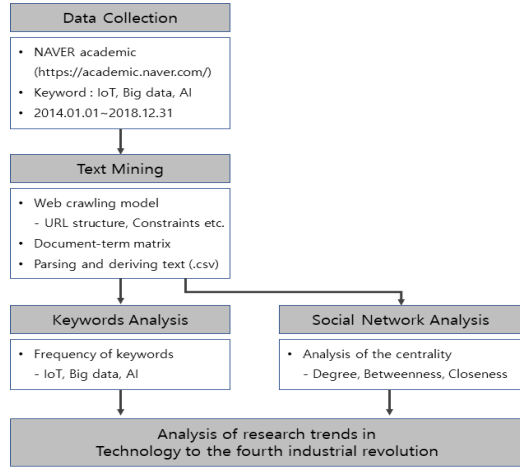


Figure 1. Flow chart of the analysis

목간의 중복성 검사를 실시하여 정제하였으며, 해당 결과를 토대로 키워드 분석을 통해 국내·외 4차 산업혁명 기술별 출현 빈도가 높은 키워드를 살펴보았다. 마지막으로 기술별 연관도 분석을 위해 소셜 네트워크 분석을 수행하여 국내·외 4차 산업혁명 기술 연구 동향에 대한 시사점을 제시하였다.

4. 분석 결과

4.1. 기술 관련 논문 발간 추이

4차 산업혁명 관련 기술에 대한 국내·외 논문의 발간 추이를 살펴보면 Figure 2와 Figure 3과 같다.

국내에서는 ‘IoT’ 관련 논문이 2016년도를 기점으로 하여 점진적으로 증가하고 있음을 알 수 있다. ‘빅데이터’ 관련 논문 또한 2014년도부터 꾸준히 증가하고 있음을 알 수 있으며, IoT 및 AI 논문에 비해 큰 비중을 차지하고 있다. ‘AI’ 관련 논문 또한 증가 추세를 보이며, 다른 기술에 비해 가파른 증가를 보이고 있다.

소셜 네트워크 분석을 이용한 4차 산업혁명 기술 분야의 연구 동향 분석

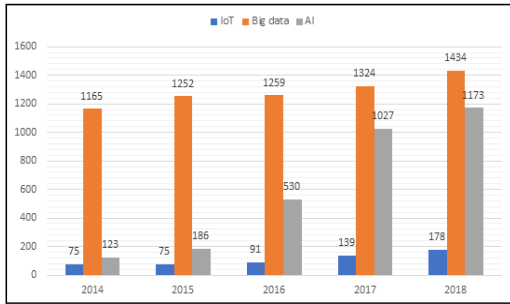


Figure 2. Annual trends of domestic

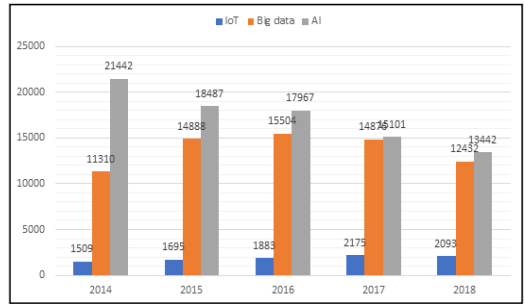


Figure 3. Annual trends of foreign

Table 1. Main keywords and appearance rates

	Domestic		Foreign	
	Keywords	Rate (%)	Keywords	Rate (%)
IoT	도시(city)	0.04	데이터(data)	0.05
	건설(building)	0.04	어플리케이션(application)	0.05
	분석(analysis)	0.03	스마트시티(smart city)	0.04
	산업(industry)	0.03	시스템(system)	0.03
	데이터(data)	0.03	분석(analysis)	0.03
	기술(technology)	0.03	지속 가능한(sustainable)	0.03
	서비스(service)	0.03	교통(traffic)	0.02
	동향(trends)	0.02	환경(environment)	0.02
	사례(case)	0.02	모니터링(monitring)	0.02
	개발(development)	0.02	모바일(moblie)	0.02
Big data	분석(analysis)	0.12	분석(analysis)	0.05
	이용(use)	0.12	연구(research)	0.04
	연구(research)	0.08	클라우드(cloud)	0.03
	데이터(data)	0.08	챌린지(challenge)	0.03
	개발(development)	0.02	학습(learning)	0.03
	소셜(social)	0.02	건강(health)	0.02
	기술(technology)	0.02	관리(management)	0.02
	보안(security)	0.02	스마트시티(smart city)	0.02
	시스템(system)	0.02	마이닝(mining)	0.02
	서비스(service)	0.02	소셜(social)	0.02
AI	연구(research)	0.07	사물인터넷(IoT)	0.10
	기술(technology)	0.05	지능형(intelligence)	0.04
	산업(industry)	0.03	교육(education)	0.03
	지능형(intelligence)	0.06	시간(time)	0.03
	이용(use)	0.03	학습(learning)	0.03
	개발(development)	0.03	인간(human)	0.03
	로봇(robot)	0.03	게임(game)	0.03
	법(law)	0.03	데이터(data)	0.03
	교육(education)	0.02	시스템(system)	0.03
	인간(human)	0.02	연구(research)	0.02

국외에서는 'IoT' 관련 논문이 점진적으로 증가하다가 2017년도에 최고점을 통과하였으며, 2018년도에는 적정 수준을 유지하고 있다. '빅데이터' 관련 논문은 2014년도부터 증가 추세를 보이다가 2016년도에 최고점에 도달하였으며, 이후 감소하는 추세를 보이고 있다. 'AI' 관련 논문은 2014년도에 최고점에 도달하였다가 지속적으로 감소하는 추세를 보이고 있다.

국내·외 모두 'IoT' 기술과 관련해서 지속적인 연구가 진행되고 있음을 알 수 있으나, '빅데이터' 및 'AI' 기술의 경우 국내에서는 지속적으로 기술에 대한 이슈가 발생하고 이와 관련된 연구가 활발해지고 있음을 알 수 있었다. 반면에 국외에서는 특정 시점까지 연구가 활발하게 진행되었다가, 최근 전반적으로 관련 연구들이 감소하고 있음을 확인할 수 있었다.

4.2. 키워드 빈도 분석 결과

4차 산업혁명 기술별 연구 논문에서 주요 키워드의 출현 비율을 살펴보면 Table 1과 같다. 키워드 출현 비율은 수집한 논문의 제목 중 키워드가 1번 이상 출현하는 비율을 의미하며, 도출된 전체 키워드 중 상위 100개 키워드를 대상으로 비율을 설정하였다.

국내의 경우 'IoT' 관련 가장 빈번하게 출현하는 키워드는 도시(0.04%)로 나타났다. 이 외에 건설, 분석, 산업 등의 용어가 상위를 차지하고 있다. IoT와 관련하여 주로 데이터, 서비스, 플랫폼, 정책 등의 연구 이슈가 존재하는 것을 알 수 있다. 국외에서는 'IoT' 관련 가장 빈번하게 출현하는 키워드는 데이터(0.05%)로 나타났다. 이 외에 어플리케이션, 스마트시티, 시스템 등의 용어가 상위를 차지하고 있다. 국외에서는 주로 교통, 환경, 모니터링 등 서비스 차원에서의 IoT 연구 이슈가 많은 것으로 나타났다.

'빅데이터' 관련 국내에서 가장 빈번하게 출현하는 키워드는 분석(0.12%)으로 나타났다. 이 외에 활용, 연구, 데이터 등의 용어가 상위를 차지하고 있다. 빅데

이터의 경우 소셜, 개인 등 사회적인 이슈와 환경, 교통 등 서비스 차원에서의 이슈로 구분될 수 있다. 국외의 경우 '빅데이터' 관련 가장 빈번하게 출현하는 키워드는 분석(0.05%)으로 나타났다. 이 외에 연구, 클라우드, 도전 등의 용어가 상위를 차지하고 있다. 빅데이터의 경우 개인 건강 등을 중점으로 연구가 활발하게 진행됨을 알 수 있으며, 다양한 모바일 서비스에 대한 이슈 나타나는 것을 알 수 있다.

국내 'AI' 관련 가장 빈번하게 출현하는 키워드는 연구(0.07%)로 나타났다. 이 외에 기술, 산업, 지능형 등의 용어가 상위를 차지하고 있다. AI와 관련하여 로봇, 교육/학습, 인간, 자율주행 등의 연구가 많이 수행된 것으로 사료된다. 'AI' 관련 국외에서 가장 빈번하게 출현하는 키워드는 IoT(0.1%)로 나타났다. 이 외에 지능형, 교육, 인간 등의 용어가 상위를 차지하고 있다. 국내와 마찬가지로 국외에서도 AI 학습에 대한 이슈가 중점적으로 다루어지고 있음을 알 수 있다.

4.3. 국내·외 소셜 네트워크 분석 결과

키워드 간의 연관성을 분석하기 위해서 국내와 국외로 구분하여 사물인터넷, 빅데이터, 인공지능 기술에서 나타난 키워드를 통합한 후, 논문의 초록들을 대상으로 동시 출현 빈도 기반의 연관성 분석을 수행하였다. 분석 결과, 높은 연결 중심성 수치를 보이는 대표 키워드 상위 10종을 핵심 키워드로 정의하였다.

국내의 경우 데이터, 기술, 플랫폼, 서비스, 모델, 스마트시티, 법·제도, 시스템, AI, 정책 등으로 도출되었다(Figure 4 참조).

국외는 ICT, 데이터, 모바일, 시스템, 스마트시티, 서비스, 어플리케이션, 건강, 환경, 에너지 등으로 나타났다(Figure 5 참조).

국내·외 모두 데이터, 스마트시티, 플랫폼 등에서 공통적인 이슈가 존재하고 있음을 알 수 있었다. 한편, 국내는 법·제도 및 정책, 기술 등 거시적인 관점에서

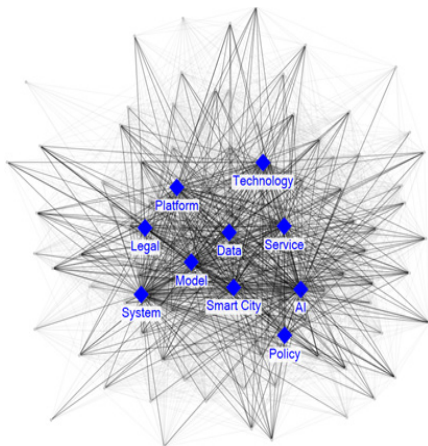


Figure 4. SNA result of domestic

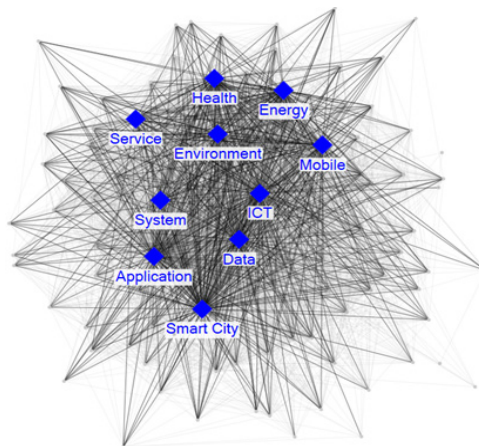


Figure 5. SNA result of foreign

Table 2. Words associated with main keywords

Domestic				Foreign			
Field	Related	Field	Related	Field	Related	Field	Related
Data	City	Smart City	Platform	ICT	Transportation	Service	Software
	Mining		Energy		Traffic		Solution
	Issue		Link		Sensing		Smart City
	Privacy		Fusion		Security		Transportation
	Social		Data		Sustainable		Society
Technology	Company	Legal	Change	Data	Legal	Application	Vehicle
	Robot		Privacy		World		Traffic
	Model		Industry		Vehicle		Sustainable
	Industry		Cooperation		Traffic		Storage
	Case		Use		Social		Platform
Platform	Innovation	System	Real time	Mobile	Vehicle	Health	Support
	Activation		Algorithm		Transportation		Monitoring
	Cooperation		Link		Service		Management
	Status		Fusion		Reality		Social
	Korean type		Integration		Sensing		Privacy
Service	Intelligent	AI	Auto driving	System	Transportation	Environment	Monitoring
	Real time		Cloud		Vehicle		Service
	Disaster		Learning		Traffic		IoT
	Energy		Disaster		World		Mining
	Medical		Platform		Real time		Intelligent
Model	Legal	Policy	Area	Smart City	Transportation	Energy	Sustainable
	Case		Evaluation		Real time		Solution
	Business		Platform		Storage		Sensing
	Industry		Cooperation		Solution		IoT
	Service		Activation		Sustainable		Monitoring

의 이슈가 주를 이루었으나, 국외의 경우 건강, 환경 분야 등의 세부적인 이슈들이 중점적으로 도출되었다.

핵심 키워드별 연관성이 높은 키워드를 도출하기 위해서, 핵심 키워드를 기준으로 매개 중심성이 높은 키워드 목록을 살펴보았다(Table 2 참조). 국내의 경우 '데이터' 관점에서는 도시, 마이닝, 이슈, 개인(정보), 소셜 등의 연관 키워드가 출현하였으며, 이를 통해 소셜 정보 활용 및 정보 보안에 대한 이슈가 주를 이루고 있음을 알 수 있다. '플랫폼' 관점에서는 공공·민간 협력 기반의 한국형 플랫폼의 개발 및 활성화에 대한 내용이 주를 이루고 있었으며, '서비스'에서는 재난, 에너지, 의료 분야를 대상으로 기존 서비스의 지능화 및 실시간성 확보 등에 대한 주요 이슈가 도출되었다. 추가적으로 '법·제도'에서는 개인정보 보호, 부처간 협력, 4차 산업혁명 기술의 활용 활성화, 규제 완화 등 제도적 차원에서 변화 요구가 나타났다. 국외의 경우 핵심 키워드 간의 공통적으로 교통, 차량 관련 시스템, 서비스, 데이터 등에 대한 이슈가 도출되었으며, 그중 '실시간, 지속 가능한 데이터'에 대한 연구가 중요하게 진행되고 있음을 알 수 있었다. 또한 건강, 환경, 에너지 등 특정 분야들을 대상으로 IoT 및 빅데이터, 모니터링 등의 세부적인 연구들이 활발하게 진행되고 있음을 알 수 있었다.

5. 결 론

본 연구에서는 4차 산업혁명과 관련한 사물인터넷, 빅데이터, 인공지능 기술들을 검색어로 설정하여 국내외의 논문을 수집하고, 해당 논문들의 키워드와 연관성에 대한 분석을 수행하였다. 이를 통해 국내·외 간의 4차 산업혁명 기술 관련 연구 동향에 대한 차이점을 확인하고자 하였다.

우선 논문 비중 및 발간 추이를 보면 국내의 경우 사물인터넷, 빅데이터, 인공지능 모두 점진적으로 증가하는 추세로 나타났다. 특히 다른 기술에 비해 빅데이

터와 관련된 연구가 주를 이루고 있음을 알 수 있었으며, 인공지능 관련 연구는 가파른 상승 곡선을 보였다. 국외의 경우 해당 기술들과 관련하여 전반적으로 적정 수준을 유지하거나 점차 감소하는 추세를 보였다. IoT만이 약간 감소 및 적정 수준을 유지하였으며, 빅데이터와 인공지능의 경우 이미 충분한 연구 활성화가 진행된 상태임을 알 수 있었다.

키워드 분석을 위하여 4차 산업혁명 기술별 국내·외 빈도 분석을 수행하였으며, 그 결과 국내에서는 주로 데이터, 정책, 플랫폼, 서비스 등 큰 단위 개념의 키워드들의 비중이 높게 나타났다. 반면 국외는 서비스를 중심으로 국내에 비해 더 세부적인 단위의 키워드들이 중점적으로 나타났다.

키워드 기반의 소셜 네트워크 분석 결과를 살펴보면, 국내·외 모두 중심 역할을 하는 키워드로써 공통적으로 데이터, 스마트시티, 플랫폼 등이 도출되었다. 그중 국내에서는 주로 거시적 차원에서의 키워드들이 도출되었으며, 국외에서는 국내에 비해 상대적으로 세부적인 키워드들이 도출되었다.

키워드 간 연관성 결과를 살펴보면, 우리나라는 4차 산업혁명 기술 관련 법·제도, 시스템 등 큰 틀에서의 구조적 변화를 도모하고 있으며, 국외의 경우 실시간·지속 가능성을 강조한 실질적인 기술의 활용 및 적용 측면에서의 연구들이 활발하게 진행되고 있었다.

이러한 점을 미루어 보아, 국내는 4차 산업혁명 기술 관련 연구 개발 및 법·제도적 완화 등에 대한 초점을 두고 있다고 유추할 수 있다. 반면, 국외는 세부적인 서비스 측면에서의 도시에 대한 실질적 적용을 중요하게 여기고 있음을 알 수 있다. 그러나 본 연구는 논문의 제목과 키워드를 중심으로 분석하였다는 한계점을 지니므로 단어의 유사어 및 관련 보고서 등을 대상으로 추가적인 연구를 수반하여, 국내·외 연구 동향 간의 구체적인 차이점을 확인하는 과정이 필요할 것으로 사료된다.

참고문헌

References

- Cho I, Kim N. 2011. Recommending core and connecting keywords of research area using social network and data mining techniques. *Journal of intelligence and information system*. 17(1):127-138.
- Church KW, Hanks P. 1990. Word association norms, mutual information, and lexicography. *Computational Linguistics*. 16(1):22-29.
- Heo GE, Song M. 2013. Examining the intellectual structure of a medical informatics journal with author co-citation analysis and co-word analysis. *Journal of the Korean society for information management*. 30(2):207-225.
- Lim SY, Lim YM, Lee JY. 2014. Study on the trends of U-City and Smart City researches using text mining technology. *Journal of the Korean society for geo-spatial information science*. 22(3):87-97.
- Jeong DW, Kim BS, Shin DB. 2016. Analysing spatial pattern of urban fire accidents based on causal factors using text mining [Master's thesis]. University of Anyang.
- Kim BS, Jeong MW, Jeon SE, Shin DB. 2015. Global research trends on geospatial information by keyword network analysis. *Journal of Korea spatial information society*. 23(1):69-77.
- Oh JS. 2015. Identifying research opportunities in the convergence of transportation and ICT using text mining techniques. *Journal of transport research*. 22(4):93-110.
- Patra SK, Mishra S. 2006. Bibliometric study of bioinformatics literature. *Scientometrics*. 67(3):477-489.
- Schvaneveldt RW. 1990. Pathfinder associative networks: studies in knowledge organization. Ablex.

2019년 05월 02일 원고접수(Received)
 2019년 05월 17일 1차심사(1st Reviewed)
 2019년 06월 03일 2차심사(2st Reviewed)
 2019년 06월 18일 게재확정(Accepted)

초 록

본 연구에서는 국내·외 4차 산업혁명 관련 기술 분야의 연구 동향을 분석하기 위해서 웹 기반의 텍스트 마이닝 및 소셜 네트워크 분석 기법을 이용하였다. 이를 위해 2014년 1월 1일부터 2018년 12월 31일까지 국내·외 4차 산업혁명 관련 기술에 대한 연구 논문 및 보고서의 제목 텍스트와 낱자를 대상으로 하여 텍스트 마이닝을 수행하였다. 이후 개념적인 차원에서의 키워드 간 연관성을 분석하기 위해서 형태소 분석을 통한 대표 키워드를 도출하였다. 이후 사회 연결망 분석을 활용하여 핵심 키워드 및 연관 키워드 등을 도출하였다. 그 결과, 우리나라에서는 4차 산업혁명 기술 관련 연구 개발 및 법·제도적 완화 등에 대한 초점을 두고 있다고 유추할 수 있다. 반면, 국외는 단위 서비스 형태로의 접근을 통해 도시에 대한 실질적 적용 기술에 초점을 두고 있음을 파악할 수 있었다.

주요어 : 소셜 네트워크 분석, 텍스트 마이닝, 사물인터넷, 빅데이터, 인공지능