

4차 산업혁명 관련 연구 논문 분석

조경원^{1*} · 우영운²

¹고신대학교 · ²동의대학교

Analysis of Research Papers Related to the Fourth Industrial Revolution

Kyoung Won Cho^{1*} · Young Woon Woo²

¹Kosin University · ²Dong-eui University

E-mail : kwcho@kosin.ac.kr / ywwoo@deu.ac.kr

요 약

본 연구에서는 "4차 산업"과 관련된 논문들의 세부 연구 주제를 파악하기 위하여 논문들을 분석하였다. 이를 위하여 2016년부터 2019년까지 한국학술지인용색인(KCI)에서 "4차 산업"이라는 키워드로 논문을 검색하여 총 685편의 논문을 수집하였으며, 논문 수집을 위해서는 Python 기반의 웹 스크래핑 프로그램을 사용하였다. 분석 결과, 인공지능, 빅데이터, 사물인터넷, 디지털, 네트워크 등이 상위 주요 기술들로 나타났으며, 산업, 정부, 교육 현장, 일자리 등 4차 산업과 관련한 다양한 분야에서 주요 기술들이 활용되고 있음을 확인할 수 있었다.

ABSTRACT

In this paper, we analyzed the papers related to the "4th Industry". In order to analyze the papers, total of 685 papers were collected by searching with the keyword "4th industry" in Korea Journal Index(KCI) from 2016 to 2019. We used Python-based web scraping program to collect papers. As a result of analysis, it was confirmed that artificial intelligence, big data, Internet of things(IoT), digital, network and so on have emerged as the major technologies, and it was confirmed that research has been utilizing the major technologies in various fields related to the 4th industry such as industry, government, education field, and job.

키워드

Fourth industry revolution, Artificial intelligence, Big data, KCI(Korea Citation Index)

1. 서 론

4차 산업혁명이란 18세기 후반 영국에서 시작된 증기기관의 사용과 기계화의 1차 산업혁명, 19세기 후반 증기기관에서 전기와 석유를 활용한 2차 산업혁명, 그리고 20세기 후반 컴퓨터와 인터넷, 인공위성을 이용한 정보혁명인 3차 산업혁명을 이어서 인공지능, 사물인터넷, 빅데이터, 모바일 등 첨단정보통신기술이 경제, 사회 전반에 융합되어 혁신적인 변화가 나타나는 차세대 산업혁명을 말한다[1].

국내에서는 '4차 산업'이란 용어가 사용된 발표

논문이 2017년부터 폭발적으로 증가하고 있으며, 연구 분야 역시 자연과학, 공학 뿐 아니라 인문, 사회과학과 예술체육 분야까지 광범위하게 분포하고 있다. 4차 산업혁명과 관련된 연구의 양적인 증가와 분야의 다양함은 4차 산업혁명에 대한 시대적 관심과 흐름을 반영하고 있다.

4차 산업혁명과 관련된 기존의 연구 분석은 4차 산업혁명과 관련 분야에서의 활용 및 분석이나 4차 산업혁명의 핵심기술에 대한 분석을 중심으로 연구되었다[4]. 본 연구는 빅데이터 분석에 주로 활용되는 텍스트 마이닝 기법[6]을 이용하여 "4차 산업"과 관련된 한글 대표 주제(토픽) 및 토픽별 빈도수 상위 단어, IDM(Intertopic Distance Map), 대표 토픽들의 연도별 변화 등을 분석하였다.

* speaker

II. 연구 방법

2.1 자료 수집

4차 산업혁명과 관련된 연구들을 수집하기 위해 한국학술지인용색인(KCI)[7]에서 논문 제목으로 “4차 산업”을 검색하였다. 분석 대상 논문은 인문학 134편, 사회과학 344, 자연과학 11편, 공학 43편, 의약학 3편, 농수해양 5편, 예술체육 47편, 복합학 98편으로 전체 685편이 검색되었다.

2.2 데이터 처리 및 분석

수집된 논문들을 대상으로 LDA 분석을 실시하여 토픽을 추출하고, 추출된 토픽들의 추이를 산출하였다. 이를 위해 먼저 토픽 수를 정하여야 한다. 본 연구에서는 perplexity를 이용하여 적절한 토픽 수를 결정하였다[9]. 그림 1의 그래프 변화에 나타나 듯이 perplexity 값의 변화가 최소화되는 지점이 9-10 구간임을 알 수 있다.

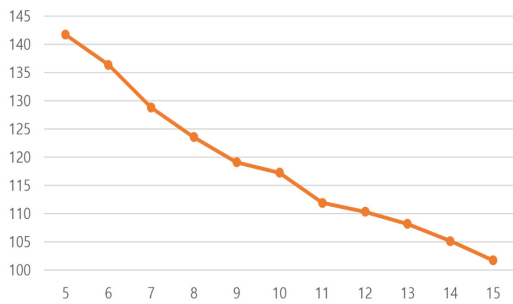


그림 1. 퍼플렉시티 변화

이상과 같은 분석 기법을 이용하여 9개의 토픽을 나타내는 빈도수 기준 상위 단어 20개씩을 추출하였다. 토픽별 구성 단어의 연관성을 중심으로 각 분야별 연구자가 토픽의 이름을 정하였다. 토픽 검토에 참여한 연구자는 경영공학, 공학, 교육학, 문헌정보학, 보건학 등의 전문가로 구성되었다.

이를 토대로 토픽의 이름을 정하였으며, 9개 토픽의 비중과 유사도를 확인하기 위해 IDM 분석을 실시하였다. 또한, 연도에 따른 논문의 대표 토픽들의 변화를 파악하기 위해 연도별 토픽 추이를 그래프로 분석하였다.

III. 연구 결과

3.1 키워드 빈도 분석

2015년부터 2019년까지의 “4차 산업” 관련 연구 685편의 논문에서 4차 산업 연구 동향을 분석하기 위해 전처리 과정을 거친 키워드를 이용하여 빈도를 분석하였다. 키워드 빈도수 상위 100개를 워드 클라우드(Word cloud)를 이용하여 시각화한 결과는

그림 2와 같다. 그림 2에서 알 수 있듯이 빈도수가 높은 단어일수록 그 값에 비례하는 크기로 표현됨을 알 수 있다.



그림 2. 상위 100개 단어에 의한 워드 클라우드

3.2 토픽 분석

“4차 산업” 관련 연구에서 중심이 되는 토픽과 관련된 핵심어를 추출하기 위하여 토픽 모델링을 실시한 결과는 표 1과 같다. 각 토픽은 LDA 기반 토픽 모델링 알고리즘을 사용하여 토픽들의 비중 순으로 제시하였다. 각 토픽을 나타내는 빈도수 기준 상위 20개 단어들을 각각 추출하였으며, 토픽별 상위 20개 단어들의 연관성을 중심으로 토픽의 이름을 정하였다.

3.3 IDM 분석

“4차 산업” 관련 연구의 토픽 모델링에서 추출된 9개 토픽에 대한 IDM 분석 결과는 그림 3과 같다. IDM은 토픽의 비중과 토픽 간 거리를 나타내는 그림으로 각 토픽이 다른 토픽과 어느 정도의 연관성을 지니는지와 토픽 간의 유사도를 파악할 수 있다.

그림 3을 분석해 보면 알 수 있듯이, 이 연구의 주제인 “4차 산업”과 관련된 토픽들의 분포를 보면 각 토픽들의 비중이 전체적으로 유사하게 나타나 편차가 크지 않음을 확인할 수 있다.

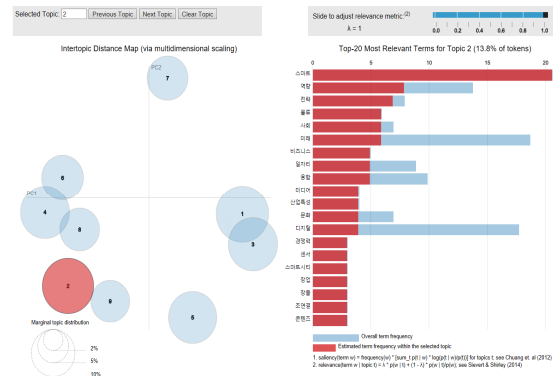


그림 3. IDM

표 1. 토픽 분석 결과

비중	토픽명	토픽별 상위 20개 단어
14.5	핵심주요 기술	인공지능, 빅데이터, 가상현실, 증강현실, 로봇, 사물인터넷, 블록체인, 기술, 저작권, 지식재산권, 컴퓨팅, 클라우드, 자동화, 기본소득, 정보보호, 프린팅, 시스템, 공유경제, 기술체계
13.8	산업화전략	스마트, 전략, 물류, 역량, 사회, 미래, 비즈니스, 융합, 미디어, 산업특성, 일자리, 문화, 경쟁력, 센서, 스마트시티, 창업, 창출, 초연결, 콘텐츠, 디지털
12.7	IOT	인공지능, 빅데이터, 산업, 데이터, 사물인터넷, 개인정보, 스포츠, 자율주행 자동차, 중국, 4차, 경제, 개인정보보호법, 소프트웨어, 인재상, 정보통신, 프라이버시, 관광, 교회교육, 변화, 보안
12.6	미래교육 방향	미래, 인간, 플랫폼, 교육, 교양교육, 교육과정, 대학, 모델, 인성교육, 교육혁신, 리더십, 핵심역량, 글쓰기, 노동의, 독일, 쓰기, 존엄성, 방향, 디지털화, 디지털
12.2	인문학과 윤리	인공지능, 인문학, 기독교, 상상력, 융합, 과학기술, 교회, 불평등, 소통, 알고리즘, 영성, 위원회, 코딩, 포스트휴먼, 윤리, 문화, 경험, 공감, 기계학습
9.2	인재상의 변화	창의성, 평생교육, 시대, 델파이, 융합교육, 디지털, 노동, 대학교육, 민간경비, 비서, 프로그램, 교육, 역량, 개발, 경영학, 공유재, 교과과정, 교수역량, 러닝
8.8	교육 패러다임 변화	분석, 빅데이터, 미래교육, 사회적, 교육패러다임, 세계경제포럼, 온라인, 유아교육, 차세대, 핵심, 활용, 부동산, 핵심역량, 기술, 고등교육, 권리, 기반, 네트워크, 논리
8.3	일자리 변화	서비스, 인식, 디자인, 사서, 의료, 진로교육, 도서관, 인성교육, 혁신, 일자리, 기본소득, 대량실업, 대학교육혁신, 복지, 윤리적, 이용자, 인간소외, 장애인복지관
7.9	정부정책	규제, 패러다임, 거버넌스, 정책, 및, 혁신, 정보통신기술, 미래이미지, 위험, 디자인, 시스템, 개방적, 건강, 기술변화, 네거티브, 대학무용교육, 도시, 미국, 융합적, 인적자원개발

3.4 고찰

앞의 토픽 모델링에 의한 IDM 분석과 트렌드 분석에 따라 알 수 있는 첫 번째 특징은 토픽의 비중이 큰 편차가 없다는 것이다. 9개 토픽의 비중이 7.9%~14.5%로 가장 낮은 토픽과 가장 높은 토픽 간의 차이가 크지 않은 것으로 나타났다. 즉, 4차 산업혁명과 관련한 연구들이 최근에 이루어지면서 관련 토픽들이 유사한 비중으로 연구되고 있

음을 알 수 있었다.

IV. 결 론

본 논문에서는 “4차 산업”의 제목으로 검색하여 추출된 KCI의 게재 논문들을 중심으로 데이터 분석을 실시하였다. “4차 산업” 주제로 2016년부터 2019년 현재까지 게재된 논문을 분석에 활용하였으며, 총 685건을 분석에 적용하였다. 분석 결과, 수집된 논문들이 발행된 전체 기간에 걸쳐 4차 산업혁명 시대의 정보 기술이 공통적으로 나타남을 확인할 수 있었으며, 인공지능, 빅데이터, 사물인터넷, 디지털, 네트워크 등의 상위순위 단어들이 4차 산업의 핵심 주요 기술들을 대표하는 것으로 나타났다.

References

[1] Telecommunication Technology Association. Dictionary of Telecommunication Terms [Internet]. Available: <http://terms.tta.or.kr/main.do>.

[2] J.-H. Ryu and Y.-Y. You, “The Fourth Industrial Revolution Core Technology Association Analysis Using Text Mining,” *Journal of Digital Convergence*, vol. 16, no. 8, pp. 129-136, Aug. 2018

[3] J.-Y. Cho and K. Y. Cho, “Topic Modeling on the Adolescent Problem Using Text Mining,” *Journal of the Korea Institute of Information and Communication Engineering*, vol. 22, no. 12, pp. 1589-1595, Dec. 2018.

[4] National Research Foundation of Korea. Korea Citation Index(KCI) [Internet]. Available: <https://www.kci.go.kr>.

[5] C. H. Nam, “An Illustrative Application of Topic Modeling Method to a Farmer’s Diary,” *Comparative Culture Study*, vol. 22, no. 1, pp. 89-135, Jan. 2016.