

## 국내 인공지능분야 연구동향 분석 -토픽모델링과 의미연결망분석을 중심으로-

황서이<sup>1</sup> · 김문기<sup>2</sup>

<sup>1</sup>중앙대학교 인문콘텐츠연구소

<sup>2</sup>한양대학교 경영학과

# An Analysis of Artificial Intelligence(A.I.)\_related Studies' Trends in Korea Focused on Topic Modeling and Semantic Network Analysis

SeoI Hwang<sup>1</sup> · Mun-Ki Kim<sup>2\*</sup>

<sup>1</sup>Humanities Research Institute, Chung-Ang University, Seoul 06974, Korea

<sup>2</sup>Department of Business Administration, Han-Yang University, Seoul 04763, Korea

### [요 약]

본 연구는 국내 인공지능분야 학술논문의 연구동향을 파악하고, 향후 연구방향을 모색하기 위해 토픽모델링과 의미연결망분석을 활용하여 연구하였다. 연구 범위는 인공지능 분야에 대해 다룬 1985년부터 2018년까지 총 1,691편의 학술논문이며 서명, 주제어, 초록을 대상으로 연구토픽과 토픽 간 분석을 진행했다. 분석 프로그램은 오픈 소프트웨어인 R을 활용하였다. 본 연구의 주요한 시사점은 첫째, 인공지능분야 연구에서 인공지능과 인간, 인공지능과 기술이라는 주요 토픽을 도출했다는 것이며, 둘째, 현재까지는 인공지능 기술 자체에 초점을 맞춘 연구들이 시도되고 있으나 향후 인공지능과 인간 및 인공지능 인문학 등 인공지능에 큰 영향을 받는 분야에 대한 연구 동향과 심도 있는 논의가 필요하다는 것이다.

### [Abstract]

The main purpose of this study was to use Topic Modeling and Semantic Network Analysis to examine research trend of Artificial Intelligence-related papers in Korea. The scope of the study was to analyze the relationship between research topics and topics using the themes, key words and abstracts of a total of 1,691 academic papers from 1985 to 2018 on the field of artificial intelligence. The analysis program used open software R. Based on the analysis results, the implications and limitations of the study and suggestions for future research were discussed.

**색인어** : 인공지능, 토픽모델링, 의미연결망분석, 빅데이터, 연구동향, 융합연구

**Key word** : Artificial Intelligence, A.I., Topic Modeling, Semantic Network Analysis, Big Data, Research Trend, Convergent Research

<http://dx.doi.org/10.9728/dcs.2019.20.9.1847>



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 11 August 2019; Revised 10 September 2019

Accepted 20 September 2019

\*Corresponding Author; Mun-Ki Kim

Tel: [REDACTED]

E-mail: seiojoy282@gmail.com

## I. 서론

4차 산업혁명 영역 중 하나인 인공지능은 인간의 인지, 학습, 추론, 이해 등 높은 수준의 정보처리 기술을 모방한 기술로서 인공지능을 활용한 산업 영역의 확장과 잠재적 가치 창출을 위한 핵심 분야로 주목받고 있다[1]. 특히 글로벌 IT기업 중심으로 인공지능 기술개발 및 투자, 인수 합병이 활발하게 이루어지고 있는데 비카리우스(Vicarious), 킨드레드(Kindred), 누멘타(Numenta) 등 인공지능 기술 스타트업의 경우 회사의 핵심이나 사업 내용, 매출 등의 실적이 뚜렷하지 않더라도 인공지능 기술에 대한 가능성과 전문 인력 확보를 위해 대기업으로부터 많은 투자를 유치하고 있다. 또한 구글 CEO인 선다 피차이 역시 “인공지능은 모든 것이 될 수 있고, 어디에나 있을 수 있다(AI is everything, and everywhere)”고 언급했는데 이는 인공지능이 사회 산업 등 전(全) 분야에 걸쳐 파괴적 혁신을 일으키며 타 산업과의 융합을 통해 새로운 가치를 창출할 수 있다는 것을 의미한다

국내 인공지능 기술의 경우 2022년까지 기술 최강국과 비교해 85% 이상의 기술 수준을 확보를 목표로 하며, 이를 통한 글로벌 인공지능융합 산업 선진국이 될 것이라는 비전을 제시하고 있다[1]. 이에 국내 IT기업 및 연구소를 중심으로 인공지능 기술 개발 및 연구가 진행 중이며 2016년 알파고와 이세돌 9단의 바둑 대국으로 인공지능에 대한 국민적 관심이 높아지기도 하였다. 그러나 해외 기술 선진국과 비교했을 때 아직까지 그 수준이나 규모가 미약한 상황이라고 볼 수 있다[2]. 따라서 국내 인공지능 산업의 육성과 장기적 가치 창출을 위해서 기술 연구의 방향이나 투자 분야에 대한 불확실성을 줄일 필요가 있으며 보다 체계적이고 구체적인 로드맵이 필요한 시점이다[3].

이처럼 빠르게 변화하며 적응해나가는 산업의 흐름을 반영 하듯 학문 영역에서도 다양한 연구가 진행되고 있다. 인공지능을 키워드로 한 국내 학술논문이 2015년 88건에서 2016년 474건, 2017년 885건, 2018년 1,279건으로 짧은 기간 동안 많은 연구가 진행되었으며, 현 시점이 인공지능 기술의 초기 단계임을 고려했을 때 향후에는 더 양질의 연구가 진행될 것이라는 예상을 할 수 있다.

이에 본 연구는 국내 인공지능 연구 동향을 살펴봄으로써 인공지능과 관련한 국내 연구가 어느 정도까지, 어떤 방향으로 진행되었는지를 살펴보고자 한다. 특히 그 간의 인공지능 동향 연구가 인공지능 기술 동향, 기술 특허 동향 등 ‘기술’을 고려한 연구가 주를 이루었다면 본 연구는 기술 영역을 포함한 전(全) 분야의 국내 학술 데이터를 수집해 빅데이터 분석 방법으로 살펴봄으로써 인공지능의 연구 토픽들을 살펴보고 통찰을 시도해 보려고 한다. 이러한 시도는 인공지능이라는 분야가 산업의 기술적 융합 뿐 아니라 인문, 사회, 예술, 문화 등 사회 모든 영역과도 융합될 수 있다는 점에서 반드시 필요하다고 할 수 있으며, 현재까지 인공지능과 관련한 국내 연구 수준을 조망하는데 유용한 지침이 될 수 있을 것이다.

## II. 이론적 배경

### 2-1 인공지능

인공지능(A.I.; Artificial Intelligence)이란 인간의 인지, 학습, 추론, 이해 등 높은 수준의 정보처리 기술을 컴퓨터로 구현한 기술이라고 할 수 있다[1]. 사실 이러한 인공지능의 개념은 근래에 발생한 것이 아니며, 1956년 다트머스 회의에서 최초로 다루어진 후 지금까지 몇 번의 진화와 쇠퇴 과정을 거쳤다. 이는 산업이나 실생활에 적용하기에 양질의 데이터가 충분하지 않았으며 데이터를 처리하기 위한 기술도 뒷받침되지 못했기 때문이다[4]. 그러나 2010년을 전후로 클라우드 컴퓨팅, 빅데이터, 네트워크의 발전, 딥러닝 알고리즘의 발전 등 IT분야 전반의 기술이 혁신적으로 발전하면서 현대 사회는 지능화라는 새로운 시대 패러다임을 형성하고 있다.

인공지능은 스스로 사고하는 정도에 따라 약한 인공지능과 강한 인공지능으로 구분할 수 있다고 주장하였다[5]. ‘스스로 사고한다’ 관점은 컴퓨터가 인간과 유사한 형태의 인지능력을 구현한다고 해서 실제 ‘인간과 같은 지능’을 가졌다고 할 수 있는가에 대한 논의로 출발한 것이지만, 현실적으로 인공지능이 인간처럼 포괄적이고 복합적인 영역을 다루지 못할 것이라는 논의가 일반적이므로[5],[6], 인공지능 자체의 발전과 적용에 따른 기대 성과를 기준으로 약한 인공지능과 강한 인공지능을 구분할 수 있을 것이다. 따라서 약한 인공지능이란 자의식이 없는 상태의 인공지능을 말하며 특정 기능에 특화된 형태로 인간의 한계를 극복하기 위해 활용되는 개념이라고 할 수 있으며, 강한 인공지능은 자유로운 사고가 가능하며 지능과 감정을 가져 인간의 지혜 수준의 합리적 능력을 발휘하는 개념이라고 할 수 있다[7].

이처럼 인공지능의 기술 발전을 토대로 다양한 산업 영역에서 이를 적용하기 위한 시도가 이루어지고 있으며, 4차 산업혁명을 이끌어갈 주요 기술 중 하나로 주목받고 있다[8]. 대표적으로 금융(투자, 트레이딩, 신용평가), 의료(헬스케어, 처방/치료), 자동차(자율주행, 교통 서비스), 제조업(공정 최적화, 스마트 팩토리), 미디어(콘텐츠, 광고), 농업(기상데이터, 농장관리), 에너지(에너지관리), 통신(통신자원 배분), 유통(옴니채널) 등의 영역에서 활발히 논의되고 있다[2],[4].

### 2-2 토픽모델링과 의미연결망분석

#### 1) 토픽모델링

토픽모델링(Topic Modeling)은 대량의 문헌이나 문서와 같은 텍스트에 내재되어 있는 단어를 근간으로 토픽을 추출하고, 토픽 간의 관계를 분석하여 어떠한 변화를 보이고 있는지 파악할 수 있는 통계 기법이다[9].

특히 빅데이터 분석방법이 각광받고 있는 현 시점에서 토픽 모델링은 빅데이터를 분석하는 일종의 기계학습(Machine

Learning)이다. 즉, 문서나 문헌 텍스트의 토픽과 토픽별 주제어의 생성 확률에 기초한 분석기법[10]이기 때문에 텍스트가 잠재적으로 가지고 있는 토픽의 개수와 그 토픽을 대표할 수 있는 주제어의 개수를 연구자가 임의로 지정하여 분석한다[11]. 따라서 주제어를 지정할 때는 여러 번의 사전 테스트를 통해 적절한 토픽의 개수를 설정하는 것이 중요하다.

## 2) 의미연결망분석

의미연결망분석(SNA; Semantic Network Analysis)은 의미연결망분석 뿐 아니라 동시출현네트워크분석, 언어네트워크분석, 키워드네트워크분석, 키워드연결망분석 등으로 다양하게 불려지고 있다. 이는 외국어를 한국어로 번역하는 과정에서 다양한 학자들의 의견이 반영되었기 때문이다. 본 연구에서는 ‘의미연결망분석’이라는 용어로 사용하였으며, 이는 문서의 의미구조와 맥락을 통해 변화 양상을 고찰한다는 점에 의미가 있다[12].

의미연결망분석은 텍스트로 구성된 문서나 문헌에서 자동적으로 단어를 추출하고, 단어들 간의 동시출현 연결 관계를 근간으로 문헌이나 문서의 텍스트에 내재된 의미를 분석하는 기법이다[13].

의미연결망분석은 문서나 문헌의 텍스트에 잠재되어 있는 의미구조와 맥락을 고찰하는데 매우 유용하기 때문에 다양한 분야에서 변화 양상, 트렌드 분석 등을 시도할 때 많이 활용되고 있다. 또한 문헌이나 문서의 텍스트에서 추출한 텍스트의 패턴을 분석하여 시각화 할 수 있다.

이러한 장점으로 인해 토픽모델링과 의미연결망분석은 연구의 추세, 동향, 경향, 추이 등의 트렌드 분석 연구에 매우 유용한 방법론이다[12],[14],[15],[16].

의미연결망분석은 문서 내 단어들 간의 관계를 근간으로 하기 때문에 관계의 정도를 측정하는 것이 중요한데, 그 척도는 밀도(density)와 중심성(centrality)이다. 밀도는 단어들 간의 밀집정도를 뜻하며, 중심성은 연결중심성, 매개중심성, 근접중심성, 아이겐벡터 중심성의 네 가지로 구분하며 이를 통해 문서나 문헌에서 추출된 단어가 어떤 역할을 하는지 분석한다. 네 가지 중심성이 모두 높은 수준으로 나타났다면 이는 문서나 문헌에서 의미구조나 내용을 형성하는 매우 주요한 역할을 한다는 것이며, [16]을 참조하면 중심성의 개념을 파악할 수 있을 것이다.

본 연구는 토픽모델링을 통해 추출한 토픽과 간의 연결성을 분석하기 위해 의미연결망 구조를 나타낼 것이며, 이러한 과정은 지식의 흐름을 고찰하는데 매우 유용하다. 의미연결망 구조는 토픽모델링에서 추출된 연구토픽을 구성하는 주제어들을 노드(node)로 설정하고, 노드의 가중치를 에지(edge)로 하여 의미연결망분석을 시도하고자 한다.

## 2-3 빅데이터 분석을 활용하여 인공지능분야 연구동향

국내 인공지능에 대한 관심은 2016년은 알파고와 이세돌과의 바둑 대전이 큰 영향을 미친 것으로 보인다. 이 시점에 인공지능 관련 연구가 급격히 증가했다.

인공지능 연구 동향을 다룬 국내 연구들은 특허 데이터를 분석한 연구[17],[18]나 인공지능 기술 관련 연구 활동[3],[19] 등 인공지능 기술을 중심으로 한 동향 분석이 주로 진행되었다. 특히나 기술 개발 관련 데이터는 출원, 등록날짜, 기술 초록, 기술에 대한 상세 정보, 인용 등 다양한 정보를 가지고 있기 때문에 기술의 현재 수준을 잘 살펴볼 수 있고[20] 기술 동향이나 관련 시장 동향을 살펴보기에 매우 유용한 접근이라 할 수 있다. 이와 관련한 연구는 다음과 같다.

특허 데이터를 활용한 토픽 모델링으로 인공지능 기술 연구 동향을 살펴보았는데, 인공지능 관련 기술 중 인간을 대신해 효율적인 자동화 작업이 가능한 기술과 이를 위한 하드웨어 최적화 기술 연구 주제가 활발히 논의되고 있다고 주장했다[21]. 그리고 해외의 인공지능 동향 분석을 통해 미국, 중국, 일본 등 주요 해외 기술 선진국들이 자국 정부조직이나 정책 등을 고려해 인공지능 분야에 다양한 접근을 모색하고 있으며, 범정부 거버넌스나 개발거점 플랫폼 구축, 인재 확보에 노력을 기울이고 있다는 점을 짚고 이에 대한 국가 차원의 정책을 제안하였다[22].

국내 기술을 통한 인공지능 특허 출원이 1990년대부터 시작되었으며 2013년 이후 출원량이 급격히 증가했지만, 인공지능 주요 기술 분야에서의 출원은 아직 미흡한 실정이라고 언급했다[17],[22]. 따라서 전문가 언어/시각 지능의 연구는 아직 초기단계 수준이고 예산 규모도 상대적으로 낮기에 보다 국가적 차원에서 지원이 필요하다는 점을 강조했다. 또한 기업의 투자, 특허출원 등 실무 적용을 위한 기술 보다는 국책연구기관, 대학교 등 이론적 연구 분야에서의 특허출원이 활발하며, 삼성전자의 경우 행동인식 기술과 음성처리 기술, 감정이해 기술을 중점적으로 연구개발하고 있다는 것을 밝혔다.

이 외에도 정명석 등(2017)[18]은 SCI(E)에 등재된 국내 연구자의 논문데이터를 활용해 키워드 네트워크 분석으로 인공지능 연구 동향을 분석했다. 그 결과 국내 연구자의 논문이 양적인 측면에서는 매년 10%정도 성장했지만 전체 논문에서 차지하는 비율은 감소하여 연구 활성화가 더딘 상황이며, 연구 초기에는 예측이나 패턴 인식과 같은 알고리즘 위주의 연구가 주로 수행되었고 최근에는 SVM모델과 같이 실용적인 연구가 진행되고 있다는 점을 밝혔다.

또한 SCI(E)의 논문들 중 ‘인공지능(Artificial Intelligence)’ 키워드가 포함된 논문데이터를 통해 LDA 토픽모델링을 사용하는 등 인공지능 연구 동향을 살펴보기 위한 여러 시도가 진행되었다[3].

위와 같은 논의들을 살펴봤을 때 그간의 인공지능 동향 연구는 주로 인공지능 기술 자체에 대한 연구나 해외 저널을 중심으로 이루어졌다는 것을 알 수 있다. 그러나 위 연구들이 기술

분야에만 국한되거나, SCI(E)에 게재된 논문들만을 대상으로 진행하였기에 국내 각 학계에서 진행되는 인공지능 연구 동향의 다양성을 파악하기는 어렵다. 물론 인공지능이 산업의 다양한 영역과 융합되는 만큼 각 분야의 연구자에 의해 보다 구체적인 전문적인 의의를 도출할 수 있겠지만, 다학제적 성격을 가지고 있는 인공지능을 거시적 관점에서 살펴봄으로써 연구자들에게 인공지능 분야의 사회구조 지표 제시 및 연구 동향 파악 등 폭 넓은 통찰을 제공할 필요가 있다.

본 연구는 인공지능 키워드를 중심으로 국내 학술 데이터를 수집하여 다양한 빅데이터 분석방법을 활용해 인공지능분야 국내 연구 동향과 향후 연구방향을 조망해 보고자 한다.

### III. 연구설계

#### 3-1 분석대상

본 연구는 토픽모델링과 언어네트워크 분석기법을 활용해 국내 인공지능분야의 연구동향을 분석하고자 한다. 이 연구는 총체적인 인공지능분야의 연구동향을 파악하고자 학술연구정보서비스 홈페이지(RISS)에서 상세검색을 실시하였다. 상세검색 시, 제목, 주제어, 초록에 검색어 ‘인공지능’과 ‘A.I.’를 입력하였고, 국내학술지논문, 한국어, KCI우수등재, KCI 등재지를 설정하여 1,723편의 논문을 수집하였다. 1,723편의 논문 중 검색어로 인한 중복 논문, 오류로 인해 중복된 논문을 제거하고 1,619편의 학술논문의 서명, 주제어, 초록을 분석 대상으로 삼았다.

분석대상은 인공지능과 관련한 1985년부터 2018년까지의 1,619편 학술논문의 서명, 초록, 주제어이며, 인공지능분야 논문 편수는 [그림 1]와 같다. [그림 1]에서 나타났듯이 인공지능 관련 연구들은 2016년도부터 급격하게 늘어나기 시작하였고, 이 시기는 인공지능 알고리즘과 이세돌의 바둑 대전이 있었고, 많은 사람들이 인공지능에 대한 관심이 급격히 높아진 시점이다.

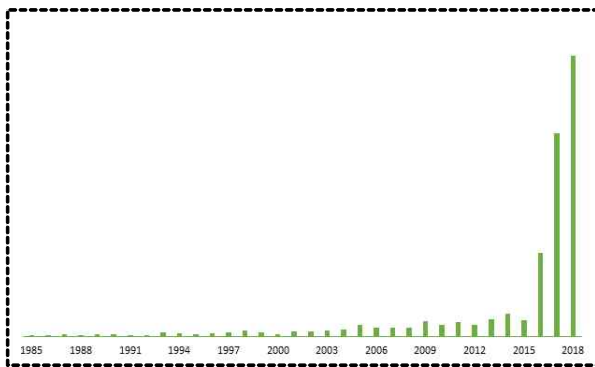


그림 1. 인공지능분야 논문 수(편)(RISS, 1985-2018)  
Fig. 1. Artificial Intelligence Paper Output (RISS, 1985-2018)

#### 3-2 연구문제

본 연구를 통해 인공지능분야 관련 연구들이 어떠한 변화양상을 보이고 있는지를 파악하기 위해, 학술논문을 대상으로 단어와 토픽을 추출하였고, 그에 따른 관계성을 살펴보고자 한다.

세부적으로 키워드 분석을 통해 단어를 추출하여 빈도분석을 실시하고, 토픽모델링을 활용해 토픽을 추출한 후, 의미연결망분석을 이용해 토픽들 간의 관계를 살펴보고자 한다. 이런 단계에 걸쳐 인공지능분야의 변화 양상을 거시적 측면뿐만 아니라 미시적 측면에서 고찰할 수 있어 인공지능분야 연구 흐름을 정리하고 앞으로의 추세를 예측할 수 있다. 이와 같은 연구목적 을 위해 이 연구는 다음과 같은 연구문제를 설정하였다.

연구문제1. 인공지능분야 학술논문에 나타난 주요단어들은 무엇인가?

연구문제2. 인공지능분야 학술논문에 나타난 연구토픽들은 무엇이며, 가장 많은 연구토픽이 발견된 학문영역은 무엇인가?

연구문제3. 인공지능분야 학술논문에 나타난 주요 연구토픽 간의 관계는 어떠한가?

### IV. 분석결과

#### 4-1 키워드 분석결과

키워드 분석으로 바이너리 빈도분석과 워드클라우드분석으로 빈도분석을 시각화 할 수 있다. 인공지능분야 학술논문 1,619편의 빈도분석을 실시한 결과를 [표 1]에 정리하였다. 바이너리 빈도분석은 무의미한 단어가 상위에 랭킹되는 것을 방지하기 위하여 문서를 기준으로 빈도를 분석하는 방법이다.

바이너리 키워드 분석결과, 1,619편의 인공지능분야 학술논문에서는 ‘인공지능’, ‘기술’ 등이 1,000편 이상의 논문에서 나타났다으며, 최상위 빈도를 차지하고 있음을 확인할 수 있다.

키워드 분석기법인 빈도분석을 시도하여 인공지능분야에 가장 많이 출현한 상위 20개 주요단어를 살펴보았다. 이어 워드클라우드 분석기법으로 상위 20개의 주요단어를 시각화하였다. 워드클라우드 분석결과는 [그림 2]와 같다.

인공지능분야 연구의 워드클라우드 분석결과, ‘인공지능’을 중심으로 ‘기술’, ‘인간’, ‘지식’ 이 굵고 큰 글씨로 중앙에 위치하고 있음을 볼 수 있다.

키워드 분석결과, 인공지능과 직·간접적으로 관련된 학술논문에 대한 전반적인 흐름은 ‘인공지능과 인간’, ‘인공지능 기술’ 등에 대한 지식을 탐구하는 연구들이 주를 이루었다.

표 1. 인공지능분야 문서빈도

Table 1. Frequency Analysis of Article Data

word	Frequency	word	Frequency	word	Frequency	word	Frequency
artificial intelligence	1,619	system	721	robot	522	industry	475
technology	1,347	data	637	bigdata	494	automated	466
human	845	information	580	realization	485	algorithm	438
community	816	behavior	577	future	480	prediction	433
industry 4.0	782	recognition	574	learning	479	education	427

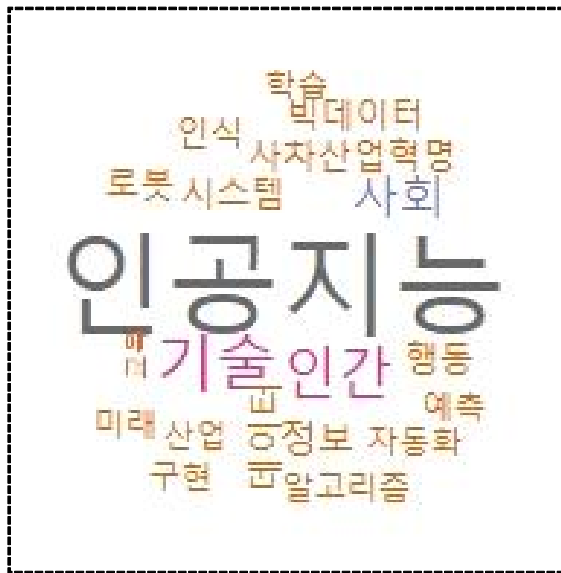


그림 2. 인공지능분야 워드클라우드분석  
Fig. 2. Word Cloud Result of Article Data

4-2. 토픽모델링 분석결과

토픽모델링은 토픽모델링을 활용해 두 가지 결과를 확인할 수 있다. 첫째는 연구토픽들을 확인할 수 있으며, 토픽비중을 통해 어떤 토픽의 중요한지를 파악할 수 있다.

토픽분석 시, 먼저 토픽 개수를 지정해야 한다. 이어서 텍스트에 포함된 모든 단어들이 해당 토픽에 맞게 자동적으로 확률이 계산되어 값이 큰 순서대로 배열된다. 해당 토픽의 주제어들은 기계학습을 통해 자동적으로 배열되는 것이며, 확률이 큰 순서대로 나타난다[12].

본 연구는 1985년부터 2018년까지 발간된 인공지능분야 학술논문 1,619편에서 추출된 9개의 토픽과 그 토픽에 포함된 주제어 15개를 [표 2]에 제시하였다.

토픽명은 연구자가 [17], [18], [19], [21]의 분석결과를 고려

표 2. 인공지능분야 학술지에서 추출된 9개 토픽과 각 토픽에 포함된 15개 주제어 & 토픽비중

Table 2. Topic Modeling Result of Article Data

no	Subject	proportion	Topic Word
1	autonomous car	0.066	artificial intelligence, automated, responsibility, liberalisation, automatic driving, robot, science of law, human, motoring, driver, criminal, rights, duties, autonomous car, compensation for damages
2	Humanities	0.271	human, artificial intelligence, techniques, robot, society, machine, philosophy, education, intelligent, industry4.0, medicine, post human, emotion, arts, mind
3	education	0.040	education, learning, artificial, intelligence, design, industry4.0, society, future, professor, knowledge, software, convergence, creativity, computer, job
4	Natural language	0.019	language, translation, system, classification, artificial intelligence, machine learning, pattern, ontology, korean, automated, natural language, semantic, construction, knowledge, machine learning
5	copyright	0.015	artificial intelligence, creation, protection, copyright, literary work, copyright law, IPRs, invention, property, system, rights, ICT, computer, knowledge, music
6	recognition	0.028	space, design, service, UX, artificial intelligence, recognition, techniques, phonetics, chatter robot, speaker, sensibility, smart, interface, augmented, city
7	ethics	0.024	artificial intelligence, morality, behavior, robot, human, robot ethics, posthumanism, ethics, responsibility, philosophy, society, education, autonomous, rights, transhumanism
8	general	0.340	artificial intelligence, techniques, industry4.0, data, bigdata, industry, information, system, medical, ICT, IoT, human, smart, service, city
9	Deep Learning Platforms	0.197	system, artificial intelligence, data, cognitive, classification, image, techniques, automated, deep learning, artificial neural network, neural network, machine learning, model, prediction, knowledge

하고, 토픽 내 주제어 간의 연관성을 파악하여 직접 명명하였다.

인공지능분야 학술논문에 나타난 토픽은 총 9개이며, topic1은 ‘자율주행’, topic2는 ‘인공지능과 인문학’, topic3은 ‘인공지능 교육’, topic4는 ‘자연언어생성’, topic5는 ‘인공지능과 저작권’, topic6은 ‘인공지능과 인식’, topic7은 ‘인공지능과 윤리’, topic8은 ‘인공지능분야 일반’, topic9는 ‘딥러닝 플랫폼’으로 토픽명을 부여하였다.

다음으로 각 문서 별로 토픽비중(topic proportion)을 분석하였다. 토픽비중은 연구의 중요도를 나타내는 정도로 볼 수 있다. 총 토픽비중의 합은 1이며, 각 토픽이 어느 정도의 비율로

중요한지를 분석하는 방법이다. 이를 통해 어떤 연구토픽이 높은 비중을 차지하고 있는지를 확인할 수 있다. 인공지능분야 학술논문에서 추출된 연구토픽들의 비중은 [표 4]에 제시하였다.

토픽비중 분석결과, 인공지능분야 학술논문은 ‘인공지능분야 일반’, ‘인공지능과 인문학’, ‘딥러닝 플랫폼’, ‘자율주행’,

‘인공지능 교육’, ‘인공지능과 인식’, ‘인공지능과 윤리’, ‘자연언어생성’, ‘인공지능과 저작권’순으로 연구 비중이 나타났다.

### 4-3. 의미연결망 분석결과

의미연결망분석은 연구토픽과 주제어 간의 관계를 고찰하는데 그 목적이 있으며 토픽분석을 통해 추출된 연구토픽과 주제어를 노드로 설정하여 연구토픽 간의 관계를 분석하였다.

#### 1) 인공지능분야 주제어 중심성 분석

인공지능분야 학술논문에 나타난 의미연결망 구조를 확인하기 위하여 중심성 값을 [표 3]에 제시하였다. ‘인공지능’, ‘지

표 3. 인공지능분야 주제어 및 중심성

Table 3. Key words and Centrality Result of Article Data

no.	Keyword	degree centrality	Keyword	closeness centrality
1	artificial intelligence	182	artificial intelligence	0.011
2	knowledge	94	knowledge	0.007
3	techniques	90	techniques	0.007
4	human	90	human	0.007
5	rights	74	rights	0.007
6	automated	72	automated	0.007
7	industry4.0	72	industry4.0	0.007
8	society	70	society	0.007
9	education	70	education	0.007
10	system/robot	68/68	system/robot	0.007/0.007

no.	Keyword	eigenvector centrality	Keyword	betweenness centrality
1	artificial intelligence	0.350	artificial intelligence	3309.783
2	techniques	0.210	knowledge	549.733
3	human	0.207	human	383.144
4	knowledge	0.193	techniques	379.534
5	industry4.0	0.178	rights	284.881
6	society	0.168	artificial neural network	270.277
7	education	0.168	industry4.0	195.060
8	robot	0.159	automated	164.539
9	automated	0.158	society	151.285
10	system	0.158	education	151.285

식’, ‘기술’, ‘인간’, ‘권리’, ‘자동화’, ‘사차산업혁명’, ‘사회’, ‘교육’, ‘시스템’, ‘로봇’ 연결중심성, 근접중심성, 아이겐벡터 중심성, 매개중심성 이 네 개의 모든 중심성에서 상위를 차지하였다.

연결중심성, 근접중심성, 아이겐벡터중심성, 매개중심성 네 개의 중심성이 모두 상위에 있다는 것은 문헌이나 문서에서 의미구조나 내용, 맥락을 형성하는데 매우 중요한 역할을 한다는 것이다.

그러므로 앞서 언급한 단어들은 인공지능분야 학술논문에서 의미구조와 내용, 맥락을 형성하는데 주요한 역할을 하는 단어 및 주제어라고 할 수 있다. 즉, 학술논문의 토픽과 이슈를 구성하는데 매우 중요한 역할을 한다.

#### 2) 인공지능분야 의미연결망의 구조 및 하위군집 분석

인공지능분야 학술논문에는 92개의 단어가 주요단어로 출현하였고, 밀도는 0.207이고, 의미연결망 구조와 군집화한 결과는 [그림 3]과 같다.

인공지능분야 학술논문은 의미연결망 구조의 중앙에 노드의 크기가 가장 크고 명암이 뚜렷한 ‘인공지능’이 존재한다. 이는 인공지능분야 학술논문에서 인공지능은 핵심이며, 그 주위로 ‘인간’, ‘지식’, ‘기술’ 등이 존재하며 연결 관계를 보였다.

인공지능분야 학술논문은 크게 4개의 군집으로 형성되어 있었다. <군집1>의 총 노드 수는 21개이고, 자연언어 생성과 딥러닝 플랫폼과 관련한 주제어들이 하나의 군집을 형성하였다. <군집2>의 총 노드 수는 9개로 나타났으며, 저작권에 관련한 주제어들이 하위 군집을 이루고 있었다. <군집3>은 인문학, 윤리 그리고 자율주행과 관련한 주제어들이 하나의 군집을 형성하였고, 총 노드의 수는 31개였다. <군집4>는 총 노드 수 38개이고, 교육, 디자인, 인공지능분야 일반들의 주제어들이 하위 군집을 구성하였다. 이를 통해 <군집1>은 ‘인공지능 기술 관련 군집’, <군집2>는 ‘지적 재산 관련 연구 군집’으로 특징지을 수 있다. <군집3>은 ‘인문학 관련 연구 군집’으로 볼 수 있으며, <군집4>는 ‘인공지능 응용산업 관련 연구 군집’으로 볼 수 있다.

### V. 결 론

본 연구는 인공지능분야 학술적 연구동향을 고찰하고자 토픽모델링과 의미연결망분석을 시도하였다. 연구에는 학술연구정보서비스인 RISS를 통해 1985년부터 2018년까지 인공지능분야 학술논문을 수집하였고, 1,619편 논문의 서명, 초록, 주제어를 분석대상으로 삼아 연구토픽을 도출하였고, 가장 비중 있게 다루어진 토픽을 살펴보고, 토픽 간의 관계성을 파악하였다. 분석에는 오픈 소프트웨어 프로그램인 R을 사용하였다.

분석결과는 다음과 같다. 첫째, 키워드 분석결과, 인공지능분야 연구는 인공지능, 기술이 최상위 주요 단어로 나타났으며,

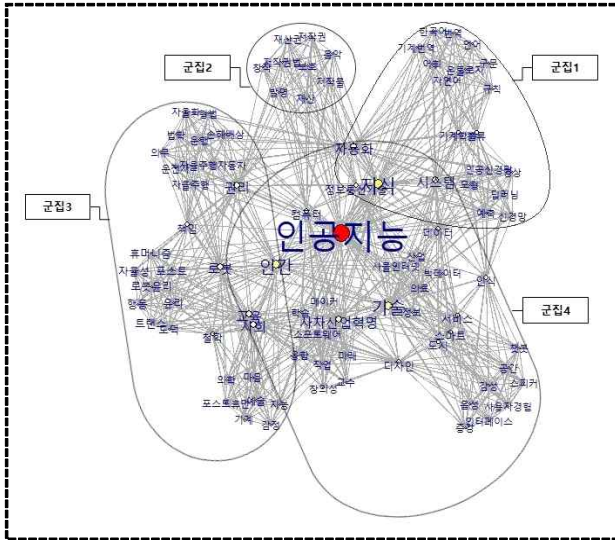


그림 3. 인공지능분야 네트워크 구조 및 하위군집 분석  
 Fig. 3. As a result of Semantic Network Structure and Cluster Analysis

인간, 사회, 4차 산업혁명, 시스템, 데이터, 정보, 행동, 인식, 로봇, 빅데이터, 구현, 미래, 학습, 산업, 자동화, 알고리즘, 예측, 교육 등이 그 뒤를 따르고 있었다.

둘째, 의미연결망에 따른 군집분석과 토픽모델링 분석 결과를 고려해 군집에 따른 대주제와 토픽모델에 따른 소주제를 범주화하고 그에 따른 핵심 키워드를 분류하였다. 세부 내용은 다음 [표 4]와 같다.

본 연구는 다음과 같은 시사점과 제언을 도출하였다. 첫째, 인공지능 키워드를 전(全) 분야에 걸쳐 분석했다는 것이며, 기존 연구에서 많이 다루어 왔던 인공지능 기술, 인공지능 응용산업 외에 철학/인문 분야와 지적재산 분야의 중요한 키워드를 도출했다는 것이다.

둘째, 인공지능분야는 인공지능과 인간, 인공지능과 기술 영역이 주요 주제로 나타났다. 인공지능 분야에서 인간에 대한 문제가 매우 주요하게 나타났음에도 불구하고 현재까지는 인공지능 기술에 대한 논의가 매우 활발히 진행되고 있다. 인공지능 기술의 결과들이 인간에 미칠 문제에 대해서도 다양한 논의가 필요하다는 것이다.

셋째, 자율주행차가 인공지능 인간과 기술에서 주요한 단어로 나타났다. 이는 자율주행차가 인간의 생활, 안전과 직결되는 영역이기에 기술과 인문학적 측면을 필연적으로 반영할 수밖에 없으며, 그 결과 책임이라는 키워드와도 중요하게 연결된 것이라 예상 할 수 있다. 이후 완전자율주행 자동차에 대한 기술력이 논의될수록 인문학적 관점에서 많은 논의들이 이루어 질 것으로 보인다.

넷째, 인공지능이 만든 창작물에 의해 인공지능과 저작권 문제에 대해 고민이 필요하다. 인공지능분야에서 지적재산 영역이 중요한 것은 인공지능의 발전이 인간의 고유 활동에 대한 기존 질서를 완전히 바꿀 수도 있기 때문이다. 문학, 예술, 영화뿐

표 4. 분석결과 종합  
 Table 4. Result of Article Data

ultimate theme	sub-themes	key words
Humanities	humanities	philosophy, responsibility, rights, duties, humanism, society, arts, etc
	ethics	ethics, morality, robot ethics, philosophy, rights, transhumanism, responsibility, etc
	education	education, learning, future, knowledge, professor, convergence, job, creativity, etc
application industry	general	industry4.0, future, job, creativity, data, industry, information, city, medical, etc
	recognition	space, design, service, UX, phonetics, interface, etc
technology	autonomous car	automated, automatic driving, autonomous car, criminal, compensation for damages, responsibility etc
	Natural language	deep learning, neural, translation, korean, language, construction, pattern, natural language, machine learning etc
	platform	data, recognition, classification, image, automated, machine learning, model, prediction, knowledge etc
intellectual property	copyright	copyright, property rights, invention, creation creativity, etc

만 아니라 신문기사, 명함 제작까지 점차적으로 인공지능이 직접 만들 수 있는 분야들이 많아지고 있다. 이러한 상황에서 인공지능이 만든 창작물들의 지적재산권에 대한 문제, 그 가치에 대한 평가를 어떻게 인정할 것인지에 대해 심도 깊은 논의가 이루어져야 할 것으로 보인다.

마지막으로 인문학적 인공지능과 미래사회를 위한 인문학적 역할을 고민하여 인문학과 과학 기술이 융합된 4차 산업혁명 시대에 필요한 새로운 관점들에 대한 논의도 이루어져야 한다.

연구의 한계점 및 향후 연구를 위한 과제는 다음과 같다. 인공지능분야는 2016년부터 급격한 성장세를 보이고 있다. 본 연구에서는 학계에 나타난 인공지능분야의 변화 양상을 파악하기 위해 학술논문을 대상으로 풍부한 연구결과를 얻기 위해 연구보고서, 신문기사, 페이스북, 트위터, 블로그 등과 같이 소셜 데이터 분석이 요구된다.

본 연구는 인공지능분야 연구동향을 파악하기 위해 1985년부터 2018년까지 데이터를 분석하였다. 인공지능분야의 진화 과정을 시기별로 파악하기 위해서는 시기를 구분하여 체계적으로 연구동향을 고찰한다면 관련한 연구들이 양적, 질적으로 더욱 풍부해 질 것이다.

## 감사의 글

본 연구는 2017년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2017S1A6A3A01078538)

## 참고문헌

- [1] O. Y. Han, J. H. Kim, "A Study on Artificial Intelligence Trend in Industry 4.0," *Korea Society for Internet Information*, Vol. 18, No. 1, pp. 19-26, Jun 2017.
- [2] Goodinfo, Prospective Market Trend, *Technology and Case Analysis of AI Industry*, May 2016.
- [3] M. S. Chung, J. Y. Lee, "Systemic Analysis of Research Activities and Trends Related to Artificial Intelligence(A.I.) Technology Based on Latent Dirichlet Allocation (LDA) Model," *Journal of the Korea industrial information systems society*, Vol. 23, No. 3, pp. 87-95, May 2018.
- [4] J. P. Kim, H. Na, "Artificial Intelligence to be fully developed," *KT management economic research institute*, Digieco, pp. 1-19, March 2016.
- [5] Searle, John. R, "Minds, brains, and programs," *Behavioral and Brain Sciences*, Vol. 3, No. 3, pp. 417-457, Aug 1980.
- [6] J. S. Kim, "The Problem of Distinction Between 'weak AI' and 'strong AI'," *Philosophical Studies*, Vol. 117, pp. 111-137, May 2017.
- [7] J. S. Park, "Response Strategy on Artificial Intelligence by Major Countries and Proposal for Political Development of Korea," *Dankook Law Review*, Vol. 41, No. 3, pp. 35-73, Aug 2017.
- [8] J. H. Kim. (2017). A Study on Strategic Response to Future Social Change in the Age of the Fourth Industrial Revolution, Korea Institute of Science & Technology Evaluation and Planning, 『KISTEPT InI』 Vol. 15, pp. 45-58, Aug 2017.
- [9] B. I. Kang, M. Song, and W. S. Jho, "A Study on Opinion Mining of Newspaper Texts based on Topic Modeling," *Journal of the Korean Library and Information Science Society*, Vol. 47, No. 4, pp. 315-334, Dec 2013.
- [10] D. Blei, A. Ng, and M. Jordan (2003), "Latent Dirichlet allocation," *Journal of Machine Learning Research*, No. 3, pp. 993-1022, June 2003.
- [11] J. Y. An, K. B. An, and M. Song, "Text Mining Driven Content Analysis of Ebola on News Media and Scientific Publications," *Journal of the Korean Library and Information Science Society*, Vol.50, No. 2, pp. 289-307, May 2015.
- [12] S. I. Hwang, D. Y. Hwang, "A Study on the Research Trends in Arts Management in Korea using Topic Modeling and Semantic Network Analysis," *Journal of Arts Management and Policy*, Vol. 47, pp. 5-29, Aug 2018.
- [13] S. S. Lee, "A Content Analysis of Journal Articles Using the Language Network Analysis Methods," *Korea Society for Information Management*, Vol. 31, No. 4, pp. 49-68, June 2014.
- [14] D. Y. Hwang, G. E. Hwang, "Examining of Semantic Map of Humanities Contents through Semantic Network Analysis: Published Articles in the Human Contents Association Published for the Past 13 Years," *Korea Humanities Content Society*, Vol. 43, pp. 229-255, Dec 2016.
- [15] S. I. Hwang, Y. W. Park, "An Analysis of Arts Management\_related Studies' Trend in Korea using Topic Modeling and Semantic Network Analysis," *Journal of Arts Management and Policy*, Vol. 43, pp. 5-37, May 2019.
- [16] G. E. Hwang, S. J. Moon, "Analysis of Big Data by Regimes of Image Contents Field," *Journal of Digital Contents Society*, Vol. 18, No. 5, pp. 911-921, May 2017.
- [17] S. M. Rho, "Artificial Intelligence technology R&D Trend by Patent Analysis," *Journal of Digital Contents Society*, Vol. 18, No. 2, pp. 423-428, Feb 2017.
- [18] M. S. Chung, S. H. Park, B. H. Chae, and J. Y. Lee, "Analysis of major research trends in artificial intelligence through analysis of thesis data," *Journal of Digital Convergence*, Vol. 15, No. 5, pp. 225-233, May 2017.
- [19] J. S. Chong, D. S. Kim, H. J. Lee, and J. W. Kim, "A Study on the Development Trend of Artificial Intelligence Using Text Mining Technique: Focused on Open Source Software Projects on Github," *Journal of Intelligence and Information Systems*, Vol. 25, No. 1, pp. 1-19, Mar 2019.
- [20] D. Hunt, L. D. Nguyen, and M. Rodgers, *Patent Searching Tools & Techniques*, Wiley, 2007.
- [21] M. S. Chung, S. H. Jeong, and J. Y. Lee, "Analysis of major research trends in artificial intelligence based on domestic/international patent data," *Journal of Digital Convergence*, Vol. 16, No. 6, pp. 187-195, June 2018.
- [22] B. Y. Kim, "Trend Analysis and National Policy for Artificial Intelligence," *Information policy*, Vol. 23, No. 1, pp. 74-93, Mar 2016.





**황서이(Seol Hwang)**

2014년 : 중앙대학교 대학원 (광고학석사)

2019년 : 중앙대학교 대학원 (예술학박사)

2019년~현 재: 중앙대학교 인문콘텐츠연구소 HK연구교수

※ 관심분야 : 인공지능인문학(AIH; Artificial Intelligence Humanities), 빅데이터(Big Data),  
트렌드분석(research trend), 융합연구(convergent research) 등



**김문기(Mun-Ki Kim)**

2014년 : 중앙대학교 대학원 (광고학석사)

2018년~현 재: 한양대학교 대학원 경영학 박사과정

※ 관심분야 : 빅데이터(Big Data), 하이테크 마케팅(high-tech marketing), 벤처경영(venture management),  
B2B 마케팅(Business to Business) 등