

키워드 네트워크 분석을 활용한 과학기술동향 분석

(Analysis of Trends in Science and Technology using Keyword Network Analysis)

박 주 섭¹⁾, 김 나 랑^{2)*}, 한 은 정³⁾

(Ju Seop Park, Na Rang Kim, and Eun Jung Han)

요 약 학계나 연구소에서는 연구동향이나 과학기술동향을 파악하고 예측하기 위해 전문가들의 판단에 의존하는 정성적인 방법을 주로 활용하여 왔다. 이 기법은 많은 시간과 비용이 드는 단점이 있기에 본 논문에서는 키워드 네트워크 분석을 활용하여 과학기술 동향을 예측하였다. 이를 위해 미국 특허 중 AI(Artificial Intelligence) 특허 초록 13,618개를 대상으로 키워드 네트워크 분석을 활용하여 분석 1기(2002.1.1. ~ 2006.12.31.), 분석 2기(2007.1.1. ~ 2011.12.31.), 분석 3기(2012.1.1. ~ 2016.12.31.)로 구분하여 분석하였다. 빈도 분석 결과, 분석 1기에서 3기로 시간이 경과할수록 AI 응용 분야의 방법에 관련된 핵심어들이 부각되었다. 키워드 네트워크 분석에서도 시간이 경과함에 따라 응용 분야의 방법에 관련된 핵심어와 다른 핵심어 간의 연계성이 높아졌다. 또한 분석 전체 기간 중 상승 및 하락 추세를 보인 연계 핵심어를 분석하면 응용 분야의 방법과 관리에 대한 연계성은 강화되는 반면에 기초 분야의 연계성은 약화되었다. 키워드 연결 중심성 분석에서도 시간이 경과할수록 응용 분야에 대한 중심성 수치가 높았다. 키워드 매개 중심성 분석에서 분석 3기는 응용 분야의 방법론관련 핵심어가 가장 높은 매개 수치를 보였다. 이는 앞으로 응용 분야의 방법들이 AI 분야의 강력한 중개자 역할을 할 것으로 예상된다. 본 논문에서 제시한 기법은 지역혁신과 관련된 과제 발굴이나 사회문제 이슈의 시각화 등 지역혁신 분야에 활용되어 질 수 있을 것이다.

핵심주제어 : 키워드 네트워크 분석, 과학기술동향, 인공지능, 키워드 연결 중심성 분석, 키워드 매개 중심성 분석

Abstract Academia and research institutes mainly use qualitative methods that rely on expert judgments to understand and predict research trends and science and technology trends. Since such a technique has the disadvantage of requiring much time and money, in this study, science and technology trends were predicted using keyword network analysis. To that end, 13,618 AI (Artificial Intelligence) patent abstracts were analyzed using keyword network analysis in three separate lots based on the period of the submission of each abstract: analysis period 1

* Corresponding Author : whitecoral@hanmail.net

+ 이 논문은 2015년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임(NRF-2015S1A3A2046781)

Manuscript received April 3, 2018 / revised April 16, 2018 / accepted April 20, 2018

1) 동아대학교 경영문제연구소, 제1저자

2) 동아대학교 경영정보학과, 교신저자

3) 동아대학교 공동가치창출혁신연구소 연구원

(January 1, 2002 – December 31, 2006), analysis period 2 (January 1, 2007 – December 31, 2011), and analysis period 3 (January 1, 2012 – December 31, 2016). According to the results of frequency analyses, keywords related to methods in the field of AI application appeared more frequently as time passed from analysis period 1 to analysis period 3. In keyword network analyses, the connectivity between keywords related to methods in the field of AI application and other keywords increased over time. In addition, when the connected keywords that showed increasing or decreasing trends during the entire analysis period were analyzed, it could be seen that the connectivity to methods and management in the field of AI application was strengthened while the connectivity to the field of basic science and technology was weakened. According to analysis of keyword connection centrality, the centrality value of the field of AI application increased over time. According to analysis of keyword mediation centrality during analysis period 3, keywords related to methodologies in the field of AI application showed the highest mediation value. Therefore, it is expected that methods in the field of AI application will play the role of powerful intermediaries in AI hereafter. The technique presented in this paper can be employed in the excavation of tasks related to regional innovation or in fields such as social issue visualization.

Key Words : Keyword Network Analysis, Science and Technology Trends, AI (Artificial Intelligence), Keyword Degree Centricity Analysis, Keyword Betweenness Centricity Analysis

1. 서 론

과학기술 동향을 파악하고 미래 기술을 예측하는 방법으로 델파이 기법이 주로 활용되어왔다. 그러나 이 기법은 절차의 복잡성, 과도한 시간과 비용, 소수 전문가에 대한 지나친 의존성 등 많은 단점이 있다[1]. 본 연구에서는 특히 데이터를 이용한 키워드 네트워크 분석을 통해 정성적 기술예측 방법의 단점을 극복하고자 하였다. 최근 과학기술 동향 및 예측 분석시 대표적 정량적 방법인 특허분석을 많이 활용하고 있다. 특허 문서에는 출원 및 공개일자, 발명자, 특허 정보 및 초록, 인용정보 등 개발 기술에 대한 많은 정보가 포함되어 있다. 특허 데이터 분석을 통한 기술예측 방법 중에서 텍스트마이닝을 활용한 SNA(Social Network Analysis) 방법론이 있다. 이 방법론에는 관계(Citation)를 이용한 방법, 국제 특허 분류(International Patent Classification, IPC)를 이용한 방법, 핵심어를 이용한 방법으로 나눌 수 있다[2]. 본 논문에서는 특허초록의 중복성을 배제하기 위해 핵심어를

이용한 방법을 사용한다.

네트워크는 인간이나 사물이 만들어 내는 다양한 유형의 시스템을 구조적으로 표현하는 방식의 하나이다. 즉, 사물이나 핵심어는 노드로 표현하고, 노드 간 연결은 링크로 표현하여 각종 시스템을 모델링하는 방법이다. 이렇게 모델링된 네트워크는 과학적인 방법으로 다양한 특성을 분석하고 이해할 수 있다[3]. 키워드 네트워크 분석은 문서 텍스트내 단어들의 동시출현 관계를 분석하여 의미를 추론하는 기법이다.

본 연구의 목적은 키워드 네트워크 분석을 활용하여 과학기술 동향을 예측하는 것이다. 이를 위해 과학기술 동향 및 예측 분야에서 많이 활용되는 미국 AI(Artificial Intelligence) 특허 초록 데이터를 이용하여 실증분석을 실시한다.

본 논문의 구성은 다음과 같다. 2장에서는 네트워크 개요를 설명하고 네트워크 분석을 활용한 선행 연구를 요약하며, 3장에서는 AI 핵심어 간 네트워크 분석 및 시사점을 제시한다. 마지막 4장에서는 결론을 제시한다.

2. 이론적 배경

정보통신기술의 발전은 데이터의 저장, 처리 능력을 크게 향상시키고 있다. 많은 데이터를 가공하여 필요한 정보를 전달해주는 다양한 기술들이 발전하고 있다. 특히 기술동향 및 예측을 위해선 특허정보를 대상으로 텍스트마이닝 기법을 활용한 연구가 있다. 또한 추세나 이슈 파악을 위해선 웹문서를 검색하고 연관성 네트워크 분석을 통한 이슈를 도출하는 연구도 있다. 핵심어를 대상으로 단순 빈도수 파악에 의한 해당 분야의 특징이나 주제를 파악하는 방법은 연구가 단순하고 의미있는 시사점을 내는 데도 한계가 있다. 이러한 점을 극복하기 위해 네트워크 분석 연구가 활발히 진행되고 있다. 네트워크 분석은 개인과 집단들 간의 관계를 노드와 링크로 나타내며, 그 위상구조, 확산 및 진화 과정을 분석하는 방법론이다[4]. 키워드 네트워크 분석은 동일 문서 내에서 핵심어 간의 동시 출현 빈도수를 통해 어느 시기에 어떤 핵심어들 간의 동시 출현 빈도를 파악하여, 시기별로 어떤 핵심어들 간의 연결 정도를 알아냄으로써 이슈가 되는 핵심어를 분석할 수 있다.

네트워크 연구의 핵심은 연구 대상을 네트워크로 표현하고 분석하는 작업인 네트워크 분석이며, 네트워크 분석 방법을 사용하는 분야는 사회과학, 자연과학, 인문과학 등 다양하다. 네트워크 분석 구조는 연결구조, 응집구조, 중심구조, 역할구조의 4가지로 분류된다. 중심성은 소셜 네트워크에서 개인이 가지는 권력과 영향력이며, 네트워크의 분석에서 가장 많이 사용되는 지표이다[3]. 소셜 네트워크에서 노드의 중심성은 전체 네트워크에서 노드가 중심에 차지하는 정도라 할 수 있다. 즉 어느 노드가 가장 중심에 있는가를 나타낸다. 중심성 분석을 통해 네트워크에서 중요한 역할을 하는 핵심 노드가 누구인지를 알 수 있다. 중심성의 종류에는 연결정도 중심성, 매개 중심성, 근접 중심성, 위세 중심성이 있다.

네트워크 분석을 바탕으로 선행연구들을 살펴보면 학술분야, 기업·기술분야, 사회분야로 세 가지로 분류될 수 있다.

첫째로, 학술분야에서 보면 Kho[5]는 최근 10년 동안 기술경영분야의 해외저널에 게재된 논문의 핵심어를 분석하여 기술경영연구의 동향을 파악하고 융합 및 유망기술 예측하였고, Park[6]은 교과서에 포함된 생물학 내용에 대해 학교급 및 생물학의 핵심 주제별로 내용 분석과 개념 관계망 분석하였고, Kim et. al.[7]은 기계 학습 분야의 특허를 수집하여 키워드 네트워크를 구축하고 네트워크 구조 분석, 중요 키워드 분석 및 클릭 분석하였다.

둘째로, 기업·기술분야에서 보면 Heo et. al.[8]은 네트워크 분석을 통해서 연구 분야간에 어떤 유기적 관계를 형성하면서 융합연구가 이루어지는지를 분석하였고, Kim[9]은 사회연결망분석을 활용한 연구동향 분석 방법론을 통하여 핵심어 빈도 분석, R&D 성과 연구의 시기별 핵심어 분석, 연도별 네트워크 분석 실시하였고, Lee[10]는 기술의 동향을 파악하고자 미국 특허청에 등록된 특허를 중심으로 특허 인용 네트워크 분석하였다.

셋째로, 사회분야에서 보면 Jeong et. al.[11]은 네트워크기법 및 텍스트마이닝 기법을 활용하여 건설분야를 대상으로 미래예측에 필요한 핵심어를 도출하였고, Noh et. al.[12]는 온라인 뉴스 기사를 대상으로 텍스트 마이닝 방법들을 사용하여 규제역으로 인한 경제적, 환경적, 그리고 정책적 파급효과를 분석하였다.

3. 연구방법론

3.1 분석의 틀

과학기술동향 예측을 위해 빈도분석으로 시기별 핵심기술을 살펴봤다. 키워드 네트워크 분석을 통해 시기별 핵심어들의 동시 출현 빈도를 파악하여, 특정한 키워들 간의 연결 강도를 알아냄으로써 이슈가 되는 토픽들을 분석할 수 있다. 키워드 연결 중심성 분석은 핵심어들 간의 연결 중심성 수치를 통해 네트워크 내에서 중심에 있는 핵심어가 어떤 핵심어인지와 중심어에 근접한 핵심어들을 파악함으로써 시기별 핵심어

를 분석할 수 있다. 키워드 매개 중심성 분석은 핵심어들 간의 매개 중심성 수치를 통해 네트워크를 구축하는 데 중개자 역할을 하는 핵심어들

을 파악함으로써 시기별로 새로운 기술의 매개 역할을 하는 핵심어들을 분석할 수 있다. 분석의 틀을 도식화하면 Fig. 1과 같다.

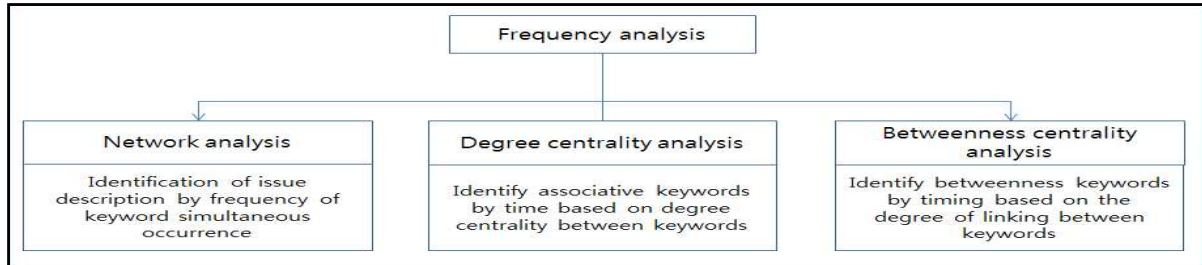


Fig. 1 Analysis Framework

3.2 분석 개요

과학기술 핵심어 간 연관성 분석을 위해 미국 특허 AI(Artificial Intelligence) 초록을 분석대상으로 주제어인 ‘Artificial Intelligence’를 핵심어로 2002년 이후 미국 특허 DB에서 13,618개의 특허문서를 연단위로 추출하였다. 수집된 특허문서 중 초록들을 대상으로 전처리 작업을 수행하였다. 빈도분석을 통해 1차로 단어들을 검색한 후 관련문헌 검토 및 AI 전문가 3명이 브레인스토밍을 통해 핵심어를 선정하였다. 2002년에서 2016년까지 초록은 동향분석 용이성을 위해 5년

단위의 3개 그룹으로 나누었다. 즉 분석 1기(2002.1.1. ~ 2006.12.31.) 1956개 초록, 분석 2기(2007.1.1. ~ 2011.12.31.) 3791개 초록, 분석 3기(2012.1.1. ~ 2016.12.31.) 7871개 초록이다. 이를 대상으로 키워드 빈도 분석, 키워드 네트워크 분석, 키워드 연결 중심성 분석, 키워드 매개 중심성 분석 등을 실시하였다.

3.3 키워드 빈도 분석

초록을 분석 1기, 분석 2기, 분석 3기로 구분하여 빈도 분석한 결과는 Table 1과 같다.

Table 1 Frequency Analysis

Div.	Period 1 (2002.1.1 ~ 2006.12.31)			Period 2 (2007.1.1 ~ 2011.12.31)			Period 3 (2012.1.1 ~ 2016.12.31)		
	Keyword	Freq.	Rat. (%)	Keyword	Freq.	Rat. (%)	Keyword	Freq.	Rat. (%)
1	system	2678	22	system	4523	23	system	7134	17
2	information	1448	12	method	2383	12	device	5374	13
3	method	1373	11	information	2316	12	information	4905	12
4	network	830	7	device	1301	7	method	4082	10
5	invention	706	6	network	1229	6	network	3355	8
6	device	652	5	invention	1118	6	image	2326	6
7	computer	644	5	computer	974	5	communication	1821	4
8	image	631	5	image	859	4	time	1789	4
9	control	578	5	time	810	4	computer	1719	4
10	time	522	4	data	801	4	control	1487	4
11	data	452	4	control	770	4	location	1484	4
12	model	438	4	component	733	4	invention	1413	3
13	database	371	3	model	684	3	data	1396	3
14	signal	328	3	database	589	3	application	1371	3
15	communication	302	3	memory	566	3	service	1255	3

분석결과, 전 기간을 보면 system, information, method 등이 상위에 랭크되었다. 분석 3기에는 6위에 image, 7위에 communication, 8위에 time 등 응용 분야의 방법론에 대한 관심이 점차로 많아진다는 것을 알 수 있다.

3.4 키워드 네트워크 분석

주요 핵심어를 대상으로 분석 1기 핵심어 간 네트워크 관계를 보면 Fig. 2와 같다. 노드의 개수는 16개이고 빈도가 높을수록 노드가 크다. 노드 간의 연관성을 나타내는 링크는 강도(weight)가 높을수록 동시출현 빈도가 높다는 의미이며 해당 핵심어가 함께 많이 이슈화된다는 것을 의미한다. Table 2를 보면 분석 1기에서는 system-information, system-method의 연계가 높게 나타났다.

Table 2 Period 1 Weighting High Order

	Node 1	Node 2	Weight
1	system	information	2,669
2	system	method	1,879
3	network	information	1,653
4	system	control	1,194
5	system	network	1,193
6	method	information	801
7	system	communication	668
8	system	database	649
9	network	control	620
10	system	environment	580
11	information	control	575
12	system	service	512
13	information	database	509
14	system	engine	487
15	system	software	480
16	system	managemant	478
17	network	method	457
18	information	engine	423
19	service	information	421
20	system	machine	397

분석 2기 핵심어 간 네트워크 관계를 보면 Fig. 3과 같다. Table 3을 보면 분석 2기에서도 분석 1기와 마찬가지로 system-information, system-method의 연계가 높게 나타났다.

분석 3기 핵심어 간 네트워크 관계를 보면 Fig. 4와 같다. Table 4를 보면 분석 3기에서는 system-method, system-information의 연계가 높게

나타났다. 특히 system-method와의 연계성이 분석 1,2기와 비교해서 가장 높게 나타났다. 이는 시간이 경과할수록 핵심어 연계측면에서 살펴보면 기초 분야에서 응용 분야로 확산되는 경향을 보여준다.

Table 3 Period 2 Weighting High Order

	Node 1	Node 2	Weight
1	system	information	1,590
2	system	method	1,447
3	system	control	903
4	system	network	846
5	information	method	746
6	information	network	617
7	system	service	517
8	system	communication	501
9	system	database	492
10	system	environment	443
11	system	software	409
12	system	performance	407
13	method	network	403
14	system	application	357
15	information	database	324
16	system	managemant	309
17	network	control	283
18	system	engine	272
19	information	control	233
20	method	control	228

Table 4 Period 3 Weighting High Order

	Node 1	Node 2	Weight
1	system	method	1568
2	system	information	1546
3	system	network	858
4	information	method	759
5	system	control	689
6	information	network	563
7	system	communication	473
8	system	service	416
9	system	environment	400
10	method	network	378
11	system	database	370
12	system	management	340
13	environment	information	340
14	system	application	308
15	information	management	267
16	management	information	267
17	information	service	261
18	information	database	260
19	system	software	257
20	network	communication	255

Analysis of Trends in Science and Technology using Keyword Network Analysis

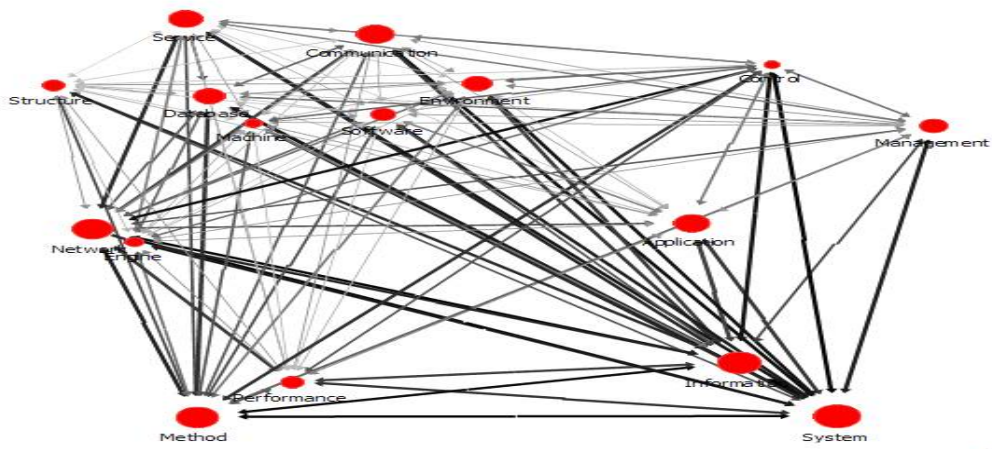


Fig. 2 Period 1 Network Analysis

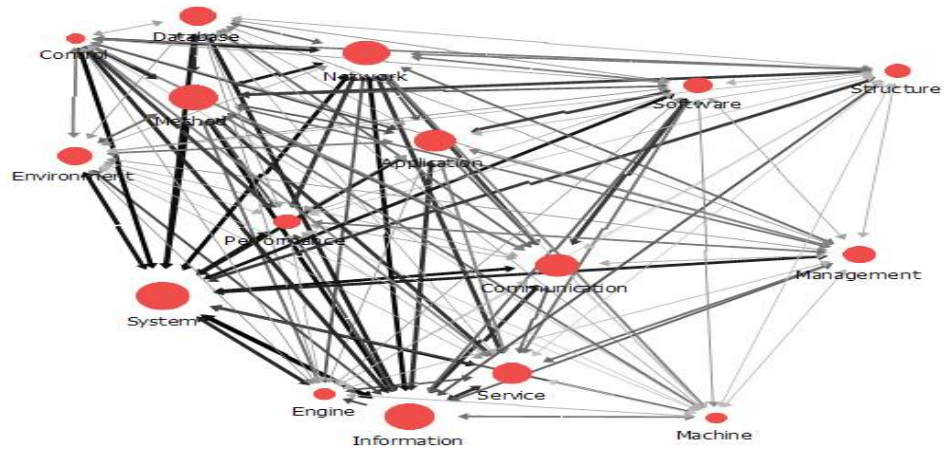


Fig. 3 Period 2 Network Analysis

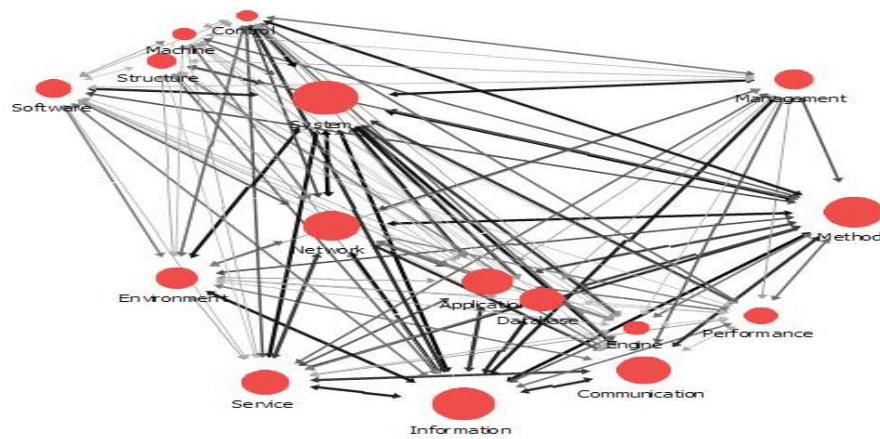


Fig. 4 Period 3 Network Analysis

분석 전체 기간 중에 상승추세를 보이는 연계 핵심어는 Table 5, 하락추세를 보이는 연계 핵심어는 Table 6에 있다. Table 5를 보면 method-information, system-network, management, application 등 응용 분야의 방법이나 관리에 연계성이 강화되는 것을 알 수 있다. Table 6을 보면 system-database, information-database, network-control, information-control 등 기초 분야의 연계성이 약화되는 것을 알 수 있다.

Table 5 Linked Keywords Rising over the Entire Period

Node 1	Node 2	1 Period	2 Period	3 Period
system	network	5위	4위	3위
method	information	6위	5위	4위
system	management	16위	16위	12위
system	application	28위	14위	14위
information	environment	30위	34위	13위

Table 6 Linked Keywords Falling over the Entire Period

Node 1	Node 2	1 Period	2 Period	3 Period
system	database	8위	9위	11위
network	control	9위	17위	33위
information	control	11위	19위	23위
information	database	13위	15위	18위
system	engine	14위	18위	28위

3.5 키워드 연결 중심성 분석

주요 핵심어를 대상으로 연결정도 중심성(degree centrality)을 분석 1기, 분석 2기, 분석 3기로 구분하여 분석하였다. 연결 중심성(degree centrality) 분석은 네트워크의 핵심어들이 얼마나 많은 연결을 가지고 있는가를 측정한다. 연결 중심성이 높다는 것은 연결된 핵심어의 수가 많다는 의미이며, 연결 중심성 분석을 통해 네트워크 내에서 중요한 핵심어가 무엇인지, 중심에 있는 핵심어에 어느 정도 접근하고 있는지를 알 수 있다(이수상, 2012). Table 7을 보면 연결 중심성 수치가 1~5위까지는 비슷하며, 6위부터는 다소 차이가 있다. AI 기초 기술을 의미하는 database 핵심어의 연결 중심성 순위는 분석 1기 6위, 분석 2기 8위, 분석 3기 11위로 하향 추세를 나타낸다. 이는 database 핵심어의 연결 중심성이 상대적으로 낮아진다는 것으로서 database같은 기초 기술이 AI 핵심 분야에서 점점 멀어짐을 의미한다. 반면에 AI 응용 및 융합 분야에서 많이 쓰이는 communication 핵심어는 분석 1기 9위, 분석 2기 7위, 분석 3기 6위로 상향 추세를 나타낸다. 이는 communication 핵심어의 연결 중심성이 점진적으로 높아진다는 것으로서 응용 및 융합 기술이 AI 응용 핵심 분야로 점차 근접함을 의미한다.

Table 7 Degree Centrality Value

Ranking	Period 1 (2002.1.1 ~ 2006.12.31)	Value	Period 2 (2007.1.1 ~ 2011.12.31)	Value	Period 3 (2012.1.1 ~ 2016.12.31)	Value
1	system	801	system	587	system	530
2	information	596	information	329	information	338
3	network	400	method	279	method	301
4	method	354	network	219	network	208
5	control	247	control	154	control	129
6	database	155	service	108	communication	118
7	service	137	communication	107	service	118
8	engine	129	database	100	environment	90
9	communication	126	application	99	management	87
10	management	106	software	87	application	87
11	performance	96	performance	83	database	76
12	environment	96	environment	70	software	55
13	application	87	engine	63	performance	46
14	software	85	management	63	engine	43
15	machine	73	structure	46	structure	41
16	structure	70	machine	30	machine	32

3.6 키워드 매개 중심성 분석

주요 핵심어를 대상으로 매개 중심성 (betweenness centrality)을 분석 1기, 분석 2기, 분석 3기로 구분하여 분석하였다. 매개 중심성 (betweenness centrality)은 한 핵심어가 다른 핵심어와 네트워크를 구축하는 데 있어서 중개자(매개) 역할을 얼마나 수행하느냐를 측정하는 개념으로 한 핵심어가 네트워크내의 다른 핵심어들의 정보흐름을 통제하는 데 큰 영향력을 가질 수 있다(이수상, 2012). AI 응용 분야의 방법

을 나타내는 method 핵심어의 매개 중심성 순위는 분석 1기 3위, 분석 2기 3위, 분석 3기 1위를 나타낸다(Table 8). 이는 앞으로 응용 분야의 방법론이 AI 분야의 강력한 중개자 역할을 할 것으로 예상된다. 또한 분석 3기의 매개 수치를 살펴보면 environment, performance 핵심어가 중개자 역할을 할 수 있게 된 반면에 database는 더 이상 중개자 역할을 할 수 없다는 것을 알 수 있다. 이는 앞으로 기초 분야는 AI 분야의 중개자 역할을 하기가 힘들 것으로 예상된다.

Table 8 Betweenness Centrality Value

Ranking	Period 1 (2002.1.1 ~ 2006.12.31.)	Value	Period 2 (2007.1.1 ~ 2011.12.31)	Value	Period 3 (2012.1.1 ~ 2016.12.31)	Value
1	system	0.1836	system	0.4577	method	0.1465
2	information	0.1328	information	0.1292	system	0.1465
3	method	0.1219	method	0.0569	information	0.1053
4	network	0.0947	network	0.0307	network	0.0821
5	control	0.0398	application	0.0161	control	0.0136
6	database	0.0142	communication	0.0095	service	0.0105
7	management	0.0015	service	0.0023	communication	0.0080
8	engine	0.0015	control	0.0019	application	0.0057
9	communication	0	management	0	environment	0.0038
10	application	0	database	0	performance	0.0013
11	service	0	environment	0	database	0
12	environment	0	software	0	management	0
13	software	0	performance	0	software	0
14	performance	0	structure	0	structure	0
15	structure	0	engine	0	engine	0
16	machine	0	machine	0	machine	0

3.7 시기별 AI(Artificial Intelligence)기술동향 변화의 특징 및 시사점

본 연구에서는 2002년 이후 미국 AI 특허초록을 대상으로 동향분석 용이성을 위해 분석 1기, 분석 2기, 분석 3기로 구분하여 분석하였다. 앞에서 논의된 AI기술동향 변화의 특징을 종합하면 시기에 따라 AI기술 주요 핵심어가 변화하고 있는 것으로 나타났다. AI기술 주요 핵심어 간 네트워크 관계를 보면 분석 1기와 2기에서는 system-information의 연계가 가장 높게 나타났

고, 분석 3기에서는 system-method의 연계가 가장 높게 나타났다. 이는 핵심어 연계측면에서 살펴보면 시간이 경과할수록 기초 분야에서 응용 분야로 확산되는 경향을 보여준다. AI기술 주요 핵심어 간 연결 중심성을 보면 1~5위까지의 연결 중심성 수치가 비슷하게 나타나며 6위부터는 다소 차이가 있다. AI 기초 기술을 의미하는 database 핵심어의 연결 중심성 순위는 분석 1기 6위, 분석 2기 8위, 분석 3기 11위로 하향 추세를 나타냈다. 이는 database 핵심어의 연결 중심성이 상대적으로 낮아진다는 것으로써

database같은 기초 기술이 AI 핵심 분야에서 점점 멀어짐을 의미한다. AI기술 주요 핵심어 간 매개 중심성을 보면 분석 1기와 2기는 system 핵심어가, 분석 3기는 method 핵심어가 가장 높은 매개 수치가 나타났다. 이는 앞으로 응용 분야의 방법론이 AI 분야의 강력한 중개자 역할을 할 것으로 예상된다.

본 연구에서의 시사점은 다음과 같다. 첫째, AI기술에 있어서 기초 분야의 중요성이 상대적으로 감소하는 경향을 보이고 있다. AI기술 동향을 분석한 결과, 핵심어 간 네트워크 연결성을 보면 system-database, information-database, network-control, information-control 등 기초 분야의 연계성이 약화되는 것을 알 수 있다.

둘째, AI기술에 있어서 응용 분야의 중요성이 상대적으로 증가하는 경향을 보이고 있다. AI기술 동향을 분석한 결과, 핵심어 간 네트워크 연결성을 보면 method-information, system-network, management, application 등 응용 분야의 방법이나 관리에 연계성이 강화되는 것을 알 수 있다.

셋째, AI기술 분야의 새로운 매개(중개자)의 화두는 AI 응용기술의 방법론이다. 키워드 빈도 분석 결과, 전 기간(1기에서 3기) 1위~5위까지는 비슷하지만, 분석 3기에는 6위에 image, 7위에 communication, 8위에 time 등 응용 분야의 방법론에 대한 관심이 점차로 많아진다는 것을 알 수 있다.

4. 결론

본 연구는 키워드 네트워크 분석을 통해 과학 기술동향 파악을 위한 분석 방법론을 제시하고 특히 데이터를 대상으로 실증분석을 실시하였다. 핵심어 간 연관성 분석을 위해 미국 AI(Artificial Intelligence) 특허 문서를 대상으로 2002년부터 2016년까지 13,618개의 특허 초록을 수집하여 토큰화, 불용어삭제 작업 등 전처리 과정을 수행하였다. 초록은 동향분석 용이성을 위해 분석 1기(2002.1.1. ~ 2006.12.31.), 분석 2기

(2007.1.1. ~ 2011.12.31.), 분석 3기(2012.1.1. ~ 2016.12.31.)로 기간별 구분하여 분석하였다. 빈도 분석 결과, 분석 1기에서 3기로 갈수록 방법론에 관련된 핵심어들(system, information, method 등)이 부각됨을 알 수 있다. 키워드 네트워크 분석에서 분석 1기와 2기는 시스템(system)-정보(information), 시스템(system)-방법(method) 순으로 높은 연계성이 나타났으며, 분석 3기는 시스템(system)-방법(method), 시스템(system)-정보(information) 순으로 높은 연계성이 나타났다. 특히 방법(method) 연계 핵심어는 시스템(system)을 제외하면 분석 1기 방법(method)-네트워크(network)(17위), 분석 2기 방법(method)-정보(information)(5위), 분석 3기 방법(method)-정보(information)(4위)로 방법(method)과 다른 핵심어 간의 연계성이 시간이 지날수록 높아짐을 알 수 있다. 또한 분석 전체 기간중에 상승 추세를 보인 연계 핵심어는 방법(method)-정보(information), 시스템(system)-네트워크(network), 시스템(system)-관리(management), 시스템(system)-응용(application) 등으로 응용 분야의 방법과 관리에 대한 연계성이 강화되는 반면, 하락 추세를 보인 연계핵심어는 시스템(system)-데이터베이스(database), 정보(information)-데이터베이스(database), 네트워크(network)-제어(control), 네트워크(network)-정보(information) 등으로 기초 분야의 연계성이 약화되는 것을 알 수 있다. 키워드 연결 중심성 분석에서도 시간이 경과할수록 응용 분야에 대한 중심성 수치가 높았다. 키워드 매개 중심성 분석에서 분석 3기는 방법(method)이 가장 높은 매개 수치를 보였다. 이는 방법론이 새로운 분야에 대한 중개자 역할이 강화됨을 의미한다.

본 연구를 통해 과학기술 정책 수립을 위한 빅데이터 분석 방법을 제시하였고, 실제 행정기관에서 자료로 활용될 수 있는 실무적 기여도가 있다. 또한 AI 산업 기술에 대한 초기 학술 연구로서 의의가 있다. 특히 이런 기법은 핵심어 동시 출현 관계의 유사성에 따라 네트워크를 구성하여 분석하는 것으로 지역혁신과 관련된 과제 발굴이나 사회문제 이슈의 시각화 등 지역혁신 분야에 활용이 가능할 것으로 생각된다. 향

후 연구 과제로는 첫째, 텍스트 전처리 과정 중 어미 제거 및 어간 추출하는 스테밍 작업처리를 위한 효율적인 방법론이 개발된다면 작업 시간을 단축하는 효과가 발생할 것이다. 둘째, 기술 예측 분야에서는 텔파이 기법이 주로 활용되어 왔다. 텔파이 기법과 네트워크 분석을 유기적으로 적용할 수 있는 방법이 활용된다면 예측 정확성을 높일 수 있을 것이다.

References

- [1] Suh, C. K., Kim, E. J. and Lee, Y. S. "The Development of Delphi Support Decision System on the Web," *The Journal of Information Systems*, Vol. 10, No. 1, pp. 5-25, 2001.
- [2] Kim, J. H., "A Study on Apple's Technology Innovation Strategy Using Patent Analysis: Focusing on iPod, iPhone, and iPad," Graduate School of Technology, Korea University, 2015.
- [3] Lee, S. S., "Network Analysis Methods," *Nonhyeong*, pp. 5-25, 2012.
- [4] http://www.cyram.com/product/product_00_01.jsp, 2015.
- [5] Kho, J. C., "A Study on Research Trend in Management of Technology using Keywords Network Analysis," *Management of Technology*, Sungkyunkwan University, 2014.
- [6] Park, H. Y., "Analysis of the Biological Themes and Their Articulations in Science and Biology Textbooks Developed under the 2009 Revised Science Curriculum Using Concept Networks," Graduate School Seoul National University, 2016.
- [7] Kim, H. H. Kim, D. G. and Jo, J. N., "Patent Data Analysis using Clique Analysis in a Keyword Network," *Journal of the Korean Data & Information Science Society*, Vol. 27, No. 5, pp. 1273-1284, 2016.
- [8] Heo, J. E. and Yang, C. H., "Applying Network Analysis in Convergent Research Relationships: The Case of High-Tech Convergence Technology Development Program," *Journal of Technology Innovation*, Vol. 16, No. 4, pp. 883-912, 2013.
- [9] Kim, M. M., "Analysis of Research Trend on R&D Performance using Social Network Analysis," Graduate School Hanyang University, 2015.
- [10] Lee, M. J., "Patent citation network analysis", Graduate School Seoul National University, 2016.
- [11] Jeong K. H. and Jeong, C. W., "Future Prediction Method using Text Mining and Network Analysis," Korea Institute of Science & Technology Evaluation and Planning, 2010.
- [12] Noh, B. J. Xu, Z. S. Lee, J. U. Park, D. H. and Chung, Y. H. "Keyword Network Based Repercussion Effect Analysis of Foot-and-Mouth Disease Using Online News," *Journal of the Korean Institute of Information Technology*, Vol. 14, No. 9, pp. 143-152, 2016.



박 주 섭 (Ju Seop Park)

- 울산대학교 전산학과 공학사
- 부산대학교 전산학과 이학석사
- 동아대학교 경영정보학과 경영학박사
- 동아대학교 경영문제연구소 연구원
- 관심분야 : 빅데이터, 텍스트마이닝, 기술예측, 지역혁신



김 나 랑 (Na Rang Kim)

- 부산대학교 문헌정보학과 문헌정보학사
- 동아대학교 경영정보학과 경영학석사
- 동아대학교 경영정보학과 경영학박사
- 동아대학교 경영정보학과 연구교수
- 관심분야 : 정보시스템 성과, 네트워크 분석, 이노베이션, 공동가치창출



한 은 정 (Eun Jung Han)

- 순천향대학교 물리학과 이학사
- 숙명여자대학교 경영학석사
- 숙명여자대학교 경영학박사
- 동아대학교 공동가치창출혁신연구소 연구원
- 관심분야 : 혁신네트워크, 산업생태계, 플랫폼 비즈니스, 비즈니스모델