

## LDA 토픽 모델링을 이용한 블록체인 학술연구 동향 분석

: 미국 · 중국 · 한국을 중심으로

조혜인<sup>1</sup> · 김진우<sup>2</sup> · 이봉규<sup>3\*</sup>

<sup>1</sup>연세대학교 정보대학원 석사과정

<sup>2</sup>연세대학교 정보대학원 석사과정

<sup>3</sup>연세대학교 정보대학원 교수

## A Study on Research Trends of Blockchain Using LDA Topic Modeling : Focusing on United States, China, and South Korea

Hye In Jo<sup>1</sup> · Jin Woo Kim<sup>2</sup> · Bong Gyou Lee<sup>3\*</sup>

1, 2, 3\*Graduate School of Information, Yonsei University, Seoul, South Korea

### [요약]

금융권을 시작으로 물류·유통, 나아가 정부의 공공·행정 서비스 등 다양한 분야에서 블록체인을 도입하고 있으며, 학계에서도 2018년을 기준으로 블록체인 연구가 급증하고 있는 추세이다. 본 연구에서는 문서의 주제 분류 토픽모델링 중 LDA 기법을 사용하여 블록체인 분야의 연구 키워드를 추출하여 미국, 중국, 한국의 학술 연구 동향을 분석하였다. 분석 결과, 미국은 ‘경제’, ‘금융’, 중국은 ‘기술’, ‘시스템’, 한국은 ‘정책’, ‘규제’ 등의 국가별 연구 동향의 차이를 확인하였다. 본 연구 결과를 통해 국가별 연구 동향을 파악하고, 향후 국내 연구자들의 블록체인 분야의 연구 방향 선정에 도움을 줄 수 있을 것으로 기대된다.

### [Abstract]

Starting from the financial field, various field such as distribution, logistics, governmental public service, and administrative service are introducing blockchain. Also, the research on blockchain rapidly increased in the academic field since 2018. This study used LDA to extract the research keywords in blockchain field and analyzed the research trends in the United States, China, and South Korea. The analysis result showed that the research trends differed for each country. While United States focused on ‘economy’ and ‘finance’ as keywords, China showed keywords such as ‘technology’ and ‘system’ and South Korean keywords were ‘policy’ and ‘regulation’. This study contribute in identifying the research trends in each country and helping the domestic researchers in determining the research direction in blockchain field.

색인어 : 블록체인, 잠재 디리클레 할당, 토픽 모델링, 연구 동향

Key word : Blockchain, LDA, Topic Modeling, Research Trend

<http://dx.doi.org/10.9728/dcs.2019.20.7.1453>



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 28 June 2019; Revised 10 July 2019

Accepted 25 July 2019

\*Corresponding Author; Bong Gyou Lee

Tel: +82-2-2123-6524

E-mail: [bglee@yonsei.ac.kr](mailto:bglee@yonsei.ac.kr)

## I. 서론

2016년 세계경제포럼(WEF)는 4차 산업혁명 시대를 이끌 핵심기술 중 하나로 블록체인을 선정했고 이것을 화두로 세계 각국의 블록체인 기술 경쟁력 확보를 위한 연구가 본격화되었다. 초기 블록체인은 금융권을 중심으로 발전하였으며, 점차 블록체인을 융합하여 기존의 비즈니스 프로세스를 혁신하고자 하는 기업들이 증가해왔다. 실제로 많은 글로벌 기업들이 블록체인 기술의 적용 및 상용화를 시도하였으며, 최근에는 금융뿐만 아니라 물류·유통, 나아가 정부의 공공·행정 서비스에서도 블록체인을 이용한 국가 차원의 시스템 개발과 관련된 연구가 활발하게 진행되고 있다[1]. 하지만 국내의 경우 블록체인과 암호 화폐에 대한 명확한 가이드라인 제시가 지체됨에 따라 다른 국가에 비해 연구와 정책적 지원이 뒤처지고 있는 실정이다. 더욱이 ‘가상화폐’, ‘비트코인’에 대한 일부 부정적인 시각에 의해 블록체인 기술 자체에 대한 산업과 연구에 간접적으로 부정적인 영향을 준다는 의견도 나오고 있다.

4차 산업혁명 기술에 대한 주요국의 정책지원, 정부규제 비교·조사 결과에 따르면, 정책지원에 있어 중국이 전 분야에서 앞서는 양상을 보였고, 미국 역시 블록체인, AI 등 6개 분야에서 한국보다 정책지원이 앞선 것으로 나타났다[2]. 미국의 경우 금융 시스템을 블록체인으로 대체하겠다는 범국가적 프로젝트인 ‘토큰노믹스(Tokenomics)’를 진행 중에 있다. 중국의 경우는 지난 2017년 ICO와 거래사이트 운영을 차단하며 암호화폐 거래를 무조건 금지하였다. 하지만 블록체인을 투자의 수단이 아닌 ‘기술’로 분류하여 국가적 차원에서 가이드라인을 제시하여 블록체인 산업 생태계 조성에 투자를 하고 있다. 또한 세계지적재산권기구(WIPO)에 따르면, 중국은 전 세계를 통틀어 블록체인 특허를 가장 많이 출원할 정도로 산업 전반적으로 우수한 인프라를 확보하고 있다[3]. 전 세계적으로 블록체인에 대한 관심이 높아지고 있는 가운데 블록체인 기술 활용을 위한 국제적 협업 및 학계 연구가 매년 증가하고 있는 추세이다. 블록체인 산업을 선도하기 위해 국가적 차원에서 투자가 진행되는 만큼 현재까지의 국가별 블록체인 산업 및 학계 연구 동향을 확인해볼 필요가 있을 것이다.

본 연구는 블록체인 산업을 선도하고 있는 미국, 중국과 국내의 연구의 동향을 비교한다. 연구 동향을 분석하기 위해 2019년 6월 기준의 Web of Science 데이터베이스에 누적된 SCIE급의 논문 초록과 DBpia에 게재된 논문 초록을 대상으로 LDA 토픽모델링과 워드클라우드 분석을 하였다. 본 논문의 구성은 서론에 이어서 2장에서는 이론적 배경에 대하여 기술하고, 3장에서 연구방법에 대해 설명하고자 한다. 4장에서는 LDA토픽모델링과 워드클라우드에 대한 결과 분석을 하며, 마지막 장에서 연구의 결론 및 한계점에 대해 서술하고 있다.

## II. 이론적 배경

### 2-1 토픽 모델링을 활용한 연구동향 분석

연구 동향에 대한 분석은 다양한 방법으로 발전해왔다. 초기 연구동향 분석은 시계열 데이터에 대한 추세분석이나 1차원 데이터에 대한 단순한 패턴 분석이었다[4]. 하지만 연구 자료의 내용적 측면까지는 고려하기 힘들다는 단점이 존재하였고, 이러한 한계점을 극복하기 위해 다양한 연구방법들이 개발되었다. 대량의 문서에서 토픽모델링은 내용적 측면을 고려할 수 있는 방법론 중 하나이다. 토픽모델링은 2012년에 D. M. Blei가 고안한 방법으로써 문서 안에 숨겨진 잠재적인 주제들을 찾아내기 위해 비정형 텍스트의 단어를 분석해서 각 주제들이 서로 어떤 관계가 있는지, 시간에 따라 어떻게 변화하는지 분석하는 방법론이다[5]. 특히 인터넷 리뷰, 신문 기사, 논문, 특허 데이터 분석에 있어 동향과약과 잠재 주제를 파악함에 있어 요긴하게 사용되며, 학계에서도 축적된 연구 자료를 바탕으로 토픽모델링을 사용하여 분야별 연구동향을 파악하고자 하는 연구가 진행되고 있다.

박주섭(2018)은 ‘A.I.’를 핵심어로 2002년부터 2016년까지의 미국 특허 13,618개의 초록 데이터를 분기별로 나누어 키워드 빈도 분석, 키워드 네트워크 분석, 키워드 연결 중심성 분석, 키워드 매개 중심성 분석을 하였으며, 시간이 지날수록 과학기술 응용 분야에 대한 중심성 수치가 높았고 기초 분야의 중요성이 상대적으로 감소하는 경향을 보였음을 확인하였다[6]. 김창식(2017)은 2002년부터 2016년까지 Web of Science 데이터베이스의 국내 저널 (APJIS, ISR, JIS)의 1,245편의 MIS 논문을 대상으로 토픽모델링 분석을 하였다[7]. 토픽모델링을 통해 국내 MIS 저널의 주요 연구 동향에 대해 파악하였으며, Hot&Cold 토픽 분석하여 정보보안, 개인정보, 고객만족, 소통혁신 분야는 떠오르는 주제, 시스템 구축, 웹사이트와 같은 분야는 하락 토픽인 것을 도출하였다. 김용환(2019)은 DBpia에서 ‘건강 관리’와 ‘헬스 케어’를 검색 키워드로 하여 4,561 건의 논문을 대상으로 토픽모델링을 하였다.[8] 도출한 연구 주제의 추세 분석을 통해 최근 가장 상승한 헬스케어 연구 주제는 ICT 기술과의 융합 토픽이며, 국내 헬스케어 연구들은 국가정책, 보건의료 및 체육 분야, 연구 디자인, IT기술, 생애주기, 환경, 여성과 관련된 내용들로 구성되는 것을 확인하였다. 윤지은(2018)은 2001년부터 2018년까지 Elsevier사의 Scopus에 게재된 2,690편의 논문을 대상으로 ‘스마트 헬스케어’ 분야의 주요 연구동향을 파악하고 향후 연구의 전개방향을 조망해 보기 위해 논문 초록을 추출하여 토픽모델링을 실시하였다[9]. 분석결과 “AI in healthcare”, “blockchain in healthcare”, “Smart health data”, “Mobile healthcare” 등 총 8개의 토픽 주제를 도출하였으며 이를 통해 AI in healthcare가 연구점유율이 가장 높음을 확인하였고, 에고네트워크 분석을 통해 주제별 핵심노드를 파악하였다.

블록체인 분야에서도 연구동향 파악을 시도하고자하는 연구가 몇 차례 진행되었다. Giungato(2017)는 2015년부터 2017년까지의 Scopus, Web of Science, Google Scholar의 데이터베이스에서 가상화폐 관련 문헌에 체계적 문헌 검토를 실시하였다[10]. 분석결과 가상화폐의 사회적 관점, 경제적 관점에서의 지속가능성을 확인하였으며, 문헌의 분석을 통해 블록체인 기술의 트렌드를 제시하였다. 조성환(2018)은 블록체인 기술의 활용 분야로 언급되는 ‘금융’, ‘물류’, ‘에너지’에 대한 7개월간의 49,093건의 뉴스 기사 데이터를 대상으로 텍스트 마이닝과 의미연결망 분석을 통해 블록체인+금융, 블록체인+물류, 블록체인+에너지 분야 각각의 주요 단어를 도출하였다[11]. 또한 CONCOR분석을 통해 분야 내 클러스터를 나누어 단어별 연관어와 관련 현상들을 확인하였다. 하지만 블록체인 학술데이터를 대상으로 연구동향 파악을 위해 토픽모델링을 실시한 연구는 부족한 실정이다.

### III. 연구방법

#### 3-1 분석 데이터

본 연구의 분석 데이터는 2019년 6월을 기준으로 Web of Science 데이터베이스에 누적된 SCIE급의 논문 초록과 DBpia에 게재된 논문 초록이다. 분석 대상 텍스트 수집을 위해 <표 1>과 같이 총 5가지의 검색 키워드 기준을 선정하였다. 논문을 검색할 때, 초록이 있는 논문 중 논문 제목과 키워드에 ‘블록체인(Blockchain)’이 들어간 경우는 가능한 모두 수집하였다. 또한, 블록체인 1.0, 블록체인 2.0, 블록체인 3.0 각각에 대한 대표적인 단어인 ‘비트코인(Bitcoin)’, ‘가상화폐, 암호화폐(Crypto Currency)’, ‘스마트 컨트랙트, 스마트 계약(Smart Contract)’, ‘이더리움(Ethereum)’이 들어간 단어의 논문도 블록체인과 관련이 있다고 보고, 논문의 제목과 키워드에 단어가 들어가 있는 경우 모두를 고려하여 데이터 분석에 필요한 논문의 제목과 초록을 연도별로 수집하였다. 국내논문의 경우 학술대회 논문은 분석 대상에서 제외하였으며, 최종적으로 국내 학술논문 125건, 미국 420건, 중국 334건을 수집하였다.

표 1. 검색 키워드

Table 1. Search Keyword

No	Keyword
1	Blockchain
2	Bitcoin
3	Crypto Currency
4	Smart Contract
5	Ethereum

#### 3-2 분석 방법

본 연구의 절차는 <표 1>의 검색 키워드를 대상으로 하는 논문 제목과 초록 데이터 수집, 데이터 필터링 및 전처리, LDA토픽 모델링과 워드클라우드, 결과 해석까지 총 4단계로 구성하였다. 연구동향 분석에 있어서 선행되어야 하는 활동은 분석의 범위를 명확히 하는 것으로 특히 전처리 과정에서는 수집된 데이터를 일관된 데이터로 만들기 위해서 데이터를 정제할 필요가 있다. 우선 수집한 텍스트 데이터의 필터링과 약어의 표준화 및 단어의 단수·복수형 통일, 띄어쓰기와 대·소문자 통일, 공백 제거 등의 처리를 하였다. 수많은 블록체인 주제의 논문에서 토픽을 도출하기 위해 대다수의 논문에서 언급되고 빈도가 가장 높은 'Blockchain'은 불용어로 처리하였다. 세 번째 단계로는 LDA 토픽모델링과 워드클라우드 분석이다. 워드클라우드는 토픽모델링의 결과치를 직관적으로 살펴 볼 수 있는 시각화의 방법 중 하나이다. 본 연구에서는 국가별 토픽 빈도를 기준으로 jupyter notebook(Python3)과 wordcloud 패키지를 이용하여 워드클라우드 분석을 실시하였다.

LDA(Latent Dirichlet Allocation)는 PLSA를 조건부 확률로 확장시킨 기법으로 잠재 주제의 확률적 분포에 대한 PLSA의 한계점을 보완한 모델이다. LDA모델은 무작위로 섞여있는 대량의 문서에서 단어들의 패턴을 추론하여 각 토픽의 특성을 도출하는데 용이하며, 텍스트 데이터의 의미구조를 파악하기에 적합한 방법 중 하나이다[12]. LDA토픽 모델링 단계를 통해 논문 초록에서 잠재되어 있는 주요 토픽(연구주제)을 파악하기 위해서는 도출할 토픽 수를 정해야 한다. 적절한 토픽 수에 대한 명확한 답은 없으며, 토픽 수의 결정은 도출된 토픽들의 해석가능성과 타당성 및 연구 질문에 따라 좌우되므로 토픽모델링을 통해 유의미한 결과를 도출하기 위해서는 해석상 가장 적절한 수로 결정해야 한다[13]. 따라서 본 연구는 의미 있는 결과 도출을 위해 jupyter notebook(Python3)과 gensim 패키지를 이용하여 LDA 분석을 실시하였으며, 최종적으로 1,000회의 샘플링 반복을 통해 6개의 토픽 수를 도출하였다. 본 연구의 전체 프로세스는 아래의 [그림 1]과 같다.

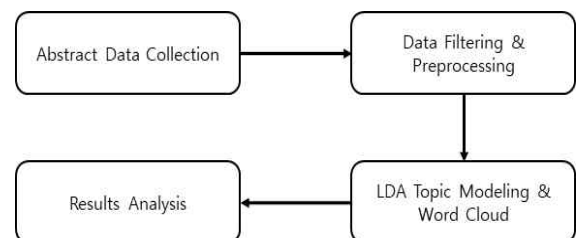


그림 1. 연구 프로세스

Figure 1. Research Process

IV. 분석 결과

4-1 분석 결과

[그림 2]는 연도별 미국, 중국, 한국 학술 논문 수의 변화를 그래프로 나타낸 것이다. [그림 2]를 보면, 2014년 이후 꾸준히 증가하는 추세를 나타내다 2018년에 급격히 연구의 수가 증가하는 것을 살펴볼 수 있다. 국가 별 추이를 세부적으로 보면 중국의 경우, 2017년까지는 미국에 비해 상대적으로 적은 수의 학술논문이 게재되었지만 2018년을 기점으로 미국의 157건과 근접한 수치인 156건의 논문이 게재되었으며 2019년 6월을 기준으로 133건의 논문을 게재함에 따라 미국의 84건을 넘어선 것을 확인할 수 있었다.

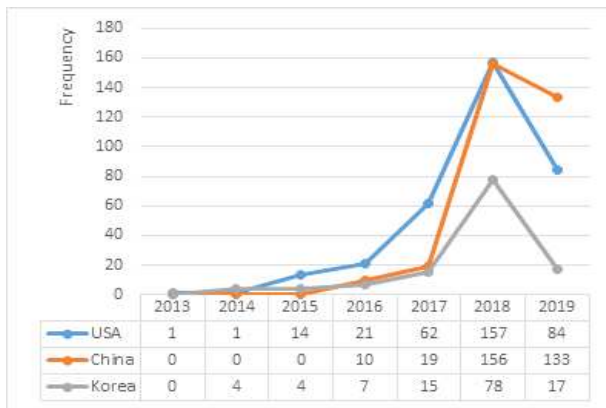


그림 2. 연도별 논문 수 추이  
Figure 2. Change in the Number of Papers by Year

4-2 미국 분석 결과



그림 3. 워드클라우드(미국)  
Figure 3. Word Cloud(USA)

Web of Science에서 블록체인을 주제로 미국에서 작성된 논문들의 초록들을 워드클라우드로 시각화한 결과는 [그림 3]과 같다. 분석결과 'bitcoin', 'techonology', 'system', 'transaction' 등의 순서로 토픽의 빈도가 높게 나타난 것을

확인할 수 있었다. 또한 'market', 'financial'과 같은 경제, 금융과 관련된 토픽이 상대적으로 크게 나타나는 것을 확인할 수 있었다.

표 2. 토픽모델링(미국)  
Table 2. Topic Modeling(USA)

	Topic1	Topic2	Topic3
word1	technology	technology	bitcoin
word2	security	financial	currency
word3	information	payment	market
word4	service	business	transaction
word5	attack	law	price
word6	transaction	economic	exchange
word7	privacy	market	volatility
word8	internet	contract	digital
word9	secure	bitcoin	bank
word10	industry	asset	evidence
	Topic4	Topic5	Topic6
word1	user	node	money
word2	digital	signature	monetary
word3	bitcoin	time	virtual
word4	governance	communication	tax
word5	cryptographic	transaction	speculative
word6	cloud	mining	bill
word7	protocol	delay	economy
word8	public	network	fraudulent
word9	trust	cost	policy
word10	anonymous	detection	criminal

Web of Science에서 블록체인을 주제로 미국에서 작성된 논문들의 초록들을 LDA토픽모델링 한 결과는 <표 2>와 같다. Topic1은 블록체인 기술의 특징 중 하나인 보안성과 관련된 단어들로서 'technology', 'security', 'transaction', 'privacy' 등이 나타났다. Topic2는 경제와 관련된 단어인 'financial', 'economic', 'payment', 'market', 'asset' 등이 나타났으며 'law'와 같은 토픽으로 보아 경제적 효과와 함께 법·규제에 대해서도 동시에 고려하며 학계 연구가 진행되는 것을 잠재적으로 유추할 수 있었다. Topic3은 비트코인과 암호화폐의 통화가치와 연관된 단어인 'bitcoin', 'currency', 'transaction', 'price', 'exchange', 'volatility' 등이 나타났다. Topic4는 Topic1과는 또 다른 블록체인 기술의 주요 특성을 나타내는 'digital', 'cryptographic', 'trust', 'anonymous' 단어들로 구성되었다. Topic5는 블록체인에 대하여 다른 토픽에 비해 심도 있는 단어들로 구성되었으며, 'node', 'signature', 'transaction', 'delay', 'mining' 등이 도출되었다. 마지막으로 Topic6은 Topic2와 같이 경제와 관련된 단어들이 등장한 것은 동일하지만, 'money', 'monetary', 'tax', 'bill' 외에도 'fraudulent', 'criminal'가 도출된 것을 통해 범죄와 사기 예방을 위해 블록체인 기술을 적용하고자 하는 연구가 되고 있는 것을 유추할 수 있었다.



표 4. 토픽모델링(한국)

Table 4. Topic Modeling(South Korea)

	Topic1	Topic2	Topic3
word1	information	virtual	doker
word2	iot	finance	network
word3	media	network	business
word4	system	invest	system
word5	industry	policy	spread
word6	trust	node	construct
word7	user	illegal	digital
word8	network	anonymity	industry
word9	privacy	exchange	centralized
word10	business	cryptocurrency	internet
	Topic4	Topic5	Topic6
word1	system	privacy	system
word2	distribution	market	smart_contract
word3	vote	trust	cryptocurrency
word4	network	asset	service
word5	control	algorithm	side effect
word6	contract	hash	problem
word7	enterprise	cryptocurrency	market
word8	infrastructure	digital	Satoshi
word9	exchange	problem	contract
word10	vehicle	vitalization	weak point

DBpia에서 블록체인을 주제로 국내에서 작성된 논문들의 초록들을 LDA토픽모델링 한 결과는 <표 4>와 같다. Topic1은 ‘information’, ‘industry’, ‘trust’, ‘privacy’ 등이 나타났으며, 블록체인의 대표적인 특성들이 도출되었다. Topic2는 가상화폐, 경제에 관련된 토픽들이 주를 이루었고, ‘virtual’, ‘finance’, ‘invest’, ‘exchange’, ‘cryptocurrency’가 도출되었다. 또한 ‘illegal’, ‘policy’가 같은 토픽으로 형성되었다. Topic3은 ‘doker’, ‘network’, ‘centralized’, ‘construct’ 등 블록체인의 기술적 특성을 나타내는 단어가 도출되었다. Topic4는 ‘vote’, ‘contract’, ‘vehicle’ 등 다른 토픽과 다소 다른 단어들이 등장하였으며, 블록체인을 다른 산업, 서비스와 융합하려는 학계 연구들이 같은 토픽으로 형성된 것으로 유추할 수 있었다. Topic5는 ‘privacy’, ‘market’, ‘trust’, ‘cryptocurrency’ 등 가상화폐와 연관된 토픽이 등장한 것을 확인할 수 있었다. 하지만 중국과 미국의 가상화폐 토픽에서 나타나지 않았던 ‘problem’이 나타난 것으로 보아 국내의 경우, 가상화폐의 부정적인 영향에 대하여 미국과 중국에 비해 다소 비중 있게 고려하고 있는 것으로 확인할 수 있었다. Topic6은 ‘Satoshi’가 나타난 것을 확인할 수 있었으며, 초록에 블록체인을 소개하는 토픽들이 일부 할당되어 있음을 알 수 있었다. 또한 Topic5와 동일하게 ‘problem’이 등장하였으며 ‘side effect’, ‘weak point’가 추가로 도출된 것을 확인하였다.

V. 결론 및 한계점

본 연구에서는 미국, 중국, 한국의 연구동향을 파악하기 위하여 학술논문 초록데이터를 대상으로 워드클라우드와 LDA 토픽모델링을 통해 연구주제에 대한 개괄적인 분석을 진행해 보았다. 전 세계적으로 블록체인에 대한 관심이 높아지고 있는 가운데 블록체인 기술 활용을 위한 국제적 협업 및 학계 연구가 매년 증가하고 있는 추세이며, 블록체인 산업을 선도하기 위해 국가적 차원에서 투자가 진행되는 만큼 현재까지의 국가별 블록체인 산업 및 학계 연구 동향을 확인해 보았다.

본 연구의 시사점은 다음과 같다. 워드클라우드와 LDA 토픽모델링을 통해 미국, 중국, 한국의 블록체인 관련 학술 연구 동향과 국가별 연구의 특성을 파악 할 수 있었다. 미국의 경우 한국과 중국에 비해 금융, 경제적 효과에 대한 토픽과 주제어들이 높게 나타났다. 실제로 ‘economy’, ‘financial’을 시작으로 비트코인의 화폐적 가치와 관련된 ‘volatility’가 등장하였으며, ‘tax’, ‘bill’ 등 공공서비스에 도입하려는 움직임을 잠재적으로 확인할 수 있는 토픽들이 다른 국가에 비해 높게 나타난 것을 확인할 수 있었다. 중국의 경우 미국, 한국에 비해 기술적 주제어들이 높게 나타났다. 워드클라우드 분석을 통해 ‘system’, ‘network’, ‘technology’, ‘user’, ‘scheme’ 등의 주제어들이 나타났다. 또한 범국가적 차원에서 블록체인 기술 자체에 대한 투자를 진행하고 있는 만큼 ‘energy’, ‘environment’, ‘IoT’의 토픽과 같이 블록체인 융합서비스에 많은 연구가 투자되고 있는 것을 예상 할 수 있었다. 국내의 경우 블록체인 기술에 대한 규제와 정책을 논의하고자 하는 연구의 빈도가 높게 나타난 것을 확인할 수 있었으며, 비트코인 투기 이슈와 함께 블록체인 기술에 대한 다소 부정적인 인식에 영향을 받았던 것으로 유추 할 수 있었다. 본 연구 결과를 통해 국내 연구자들의 블록체인 연구 방향 선정에 있어서 도움을 줄 수 있는 참고자료로 활용될 것으로 사료된다.

본 연구에서는 미국, 중국, 한국의 블록체인 연구동향을 분석하는 과정에 있어 최대한 객관적으로 분석하고자 하였으나, 현실적인 한계로 인해 다음과 같은 한계점을 갖고 있다. 분석에 사용된 데이터가 초록만을 대상으로 한 점이다. 향후 연구에서는 논문의 초록뿐만 아니라 블록체인 기술 관련 특허와 뉴스 기사 데이터까지도 분석데이터로 활용하여 보다 광범위한 연구를 진행할 필요가 있다. 또한 ‘암호화폐’, ‘블록체인 정책’, ‘블록체인 융합 서비스 제안’ 등의 세부적인 주제에 대하여 체계적 문헌검토 등의 연구방법론을 적용하여 보다 명확한 연구동향 파악을 함으로써 토픽모델링의 한계점을 보완할 필요가 있다.

## 감사의 글

본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 대학ICT연구센터지원사업의 연구결과로 수행되었음 (IITP-2019-2018-0-01799)

## 참 고 문 헌

- [1] J. Y. Lee, "Blockchain Technology Trends and Implications", *SCIENCE&TECHNOLOGY POLICY, STEPI*, Vol. 34, pp. 2-4, July 2017.
- [2] [Internet], Available:  
<https://www.yna.co.kr/view/AKR20190611058600003>
- [3] [Internet], Available:  
<http://www.donga.com/news/Issue/List/080912/article/all/20190530/95763252/1>
- [4] Kivikunnas, S., "Overview of process trend analysis methods and applications", *In ERUDIT Workshop on Applications in Pulp and Paper Industry*, pp. 395-408, 1998.
- [5] D. M. Blei., "Probabilistic topic models", *Communication of the ACM*, Vol. 55, No. 4, pp. 77-84, 2012.
- [6] J. S. Park., "Analysis of Trends in Science and Technology using Keyword Network Analysis", *Journal of the Korea Industrial Information Systems Research*, Vol. 23, No. 2, pp. 63-73, April, 2018.
- [7] C. S. Kim, "Investigation of Research Trends in Information Systems Domain Using Topic Modeling and Time Series Regression Analysis", *Journal of Digital Contents Society*, Vol. 18, No. 6, pp. 1143-1150, October, 2017.
- [8] Y. H. Kim, "Trend Analysis of Healthcare Research in Korea using Topic Modeling", *Journal of the Korean society for Wellness*, Vol. 14, No. 1, pp. 253-262, February, 2019.
- [9] J. E. Yoon, "Research Trend Analysis on Smart healthcare by using Topic Modeling and Ego Network Analysis", *Journal of Digital Contents Society*, Vol. 19, No. 5, pp. 981-993, May, 2018.
- [10] Giungato, P., "Current trends in sustainability of bitcoins and related blockchain technology", *Sustainability*, Vol. 9, No. 13, pp. 2214, 2017.
- [11] S. H. Cho, "A Study on Analysis of the Trend of Blockchain by Key Words Network Analysis", *Journal of Korea Institute of Information, Electronics, and Communication Technology*, Vol. 11, No. 5, pp. 550-555, October, 2018.
- [12] D. M. Blei. & A. Y. Ng & M. I. Jordan., "Latent Dirichlet Allocation", *Journal of Machine Learning Research*, Vol. 3, No. 4-5, pp. 993-1022, 2003.
- [13] C. H. Nahm., "An Illustrative Application of Topic Modeling Method to a Farmer's Diary", *Institute of Cross Cultural Studies*, Vol. 22, No. 1, pp. 89-135, 2016.



**조혜인(Hye In Jo)**

2018년 : 순천향대학교 컴퓨터공학과 (학사)

2018년~현 재: 연세대학교 정보대학원 ICT·콘텐츠트랙 (석사과정)  
※관심분야: 텍스트 마이닝, 데이터 아키텍처, 데이터 분석 등



**김진우(Jin Woo Kim)**

2017년 : 고려대학교 정보통계학과 (학사)

2017년~현 재: 연세대학교 정보대학원 ICT·콘텐츠트랙 (석사과정)  
※관심분야: 빅데이터 분석, 소셜 데이터, 미디어 정책 등



**이봉규(Bong Gyou Lee)**

1988년 : 연세대학교 상경대학 경제학과 (학사)  
1992년 : Cornell University, Dept. of CRP (MS)  
1994년 : Cornell University, Dept. of CRP (Ph.D)

1997년~2005년: 한성대학교 공과대학 정보전산학부 교수  
2016년~2017년: 연세대학교 정보대학원 원장  
2005년~현 재: 연세대학교 정보대학원 ICT·콘텐츠트랙 교수  
2018년~현 재: 연세대학교 학술정보원 원장(CIO, CPO)  
※관심분야: 디지털 트랜스포메이션 기술 및 전략 등