

텍스트 마이닝을 활용한 분야별 블록체인 연관 이슈 및 중요도 분석: 뉴스, 웹사이트, 학술분야를 중심으로

김 호 익¹ · 김 한 민² · 이 민 환^{3*}

¹성균관대학교 경영학과 석박통합과정, ²성균관대학교 경영학과 박사과정, ³성균관대학교 경영학과 초빙교수

An Analysis of Blockchain-related Issues and Importance Using Text Mining: Focused on News, Web Site, and Academic Field

Ho-Ik Kim¹ · Han-Min Kim² · Min-Hwan Lee^{3*}

¹Master & Doctoral's Course, ²Doctoral Course, ³Visiting Professor
Department of Business Administration, Sungkyunkwan University, Seoul 03063, Korea

[요 약]

블록체인은 암호화폐뿐만 아니라 현재 상호간의 믿음을 확보할 수 있는 기술로 떠오르고 있다. 블록체인에 대한 관심이 높아지면서 기존의 연구들은 블록체인 기술에 관련된 암호화폐에 대한 연구들을 주로 수행하였다. 하지만 암호화폐는 블록체인을 지속적으로 운영하기 위한 시스템적 요인으로 블록체인을 충분히 설명하기에는 한계가 존재한다. 한편, 다양한 분야에서 블록체인에 대한 논의와 질적 연구가 이루어짐에도 불구하고 블록체인의 관점에 대한 체계적인 분석은 거의 진행되지 않았다. 이러한 인식을 기반으로 본 연구는 분야별로 블록체인에 대한 연관 이슈와 중요도를 분석하고자 하였다. 텍스트 마이닝 기법을 활용한 본 연구는 뉴스, 웹사이트, 학술분야에서 바라보는 블록체인의 견해에 대한 해석을 제공한다. 블록체인에 대한 2,144개의 텍스트 데이터를 분석한 결과, 각 분야마다 공통적인 견해는 거의 존재하지 않지만 상반되는 관점으로 블록체인을 인식하고 있다는 사실을 발견하였다. 각 분야별 주요 블록체인 연관이슈는 뉴스에서 산업발전과 시장현황, 웹사이트에서 공공활용과 정책수립, 학술분야에서 연구개발과 기술적용이 주요 이슈인 것으로 나타났다. 끝으로 본 연구의 시사점과 한계점 및 향후 연구방향을 논의하였다.

[Abstract]

Blockchain is emerging as a technology that can secure mutual trust as well as cryptocurrency. As interest in Blockchain is increasing, previous studies have mainly focused on cryptocurrency related to Blockchain technology. However, there is a limit to explain Blockchain because cryptography is a systematic factor for continuously operating Blockchain. On the other hand, there is little systematic analysis of perspectives in Blockchain, despite discussions and qualitative researches on Blockchain in various fields. Based on this recognition, this study tried to analyze the blockchain-related issues and importance of Blockchain by field. This study, using text mining techniques, provides an interpretation of perspectives on Blockchain in news, web site, and academic field. As a result of analyzing 2,144 text data for Blockchain. As a results, it was found that Blockchain was perceived as conflicting issues, while there were few common issues in each field. The major blockchain-related issues in each field were found to be industrial development and market status in the news, public use and policy establishment in the website, and research and development and technical use in the academic field. Finally, the implications and limitations of this study and future research directions were discussed.

색인어 : 블록체인, 텍스트 마이닝, 뉴스, 웹사이트, 학술분야, 빅데이터

Key word : Blockchain, Text Mining, News, Web Site, Academic Field, Big Data

<http://dx.doi.org/10.9728/dcs.2020.21.8.1421>



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 18 July 2020; Revised 5 August 2020

Accepted 20 August 2020

*Corresponding Author; Min-Hwan Lee

Tel: +82-2-760-0453

E-mail: wlgid0727@skku.edu

I. 서론

블록체인에 대한 학문적 사회적 관심은 블록체인이 가지고 있는 기술적 특성과 암호화폐로부터 기인한다. 블록체인은 분산화된 원장을 네트워크 참여자에게 공유하는 기술로써 블록체인이 등장하기 이전에는 특정 네트워크의 참여자들이 일관화된 원장을 공유하고 서로를 신뢰하게 만드는 것에 대해서 큰 어려움을 겪고 있었다[1-3]. 블록체인이 사회적으로 관심을 받기 시작한 것은 블록체인 시스템에 존재하는 코인이 주목받으면서부터였다. 비트코인을 비롯한 많은 코인들의 가치가 급부상하면서 잠시나마 블록체인 기술에 대한 관심도 크게 늘어났다. 하지만 현재는 코인 가격의 거품이 사라지면서 블록체인 시스템을 연구하고 적용하고자 하는 일부 연구자들과 기업 등에서만 블록체인에 대해 관심을 가지고 있다. 블록체인에 대한 기존의 연구들은 블록체인 시스템의 일부분인 코인의 가격을 예측하거나 블록체인이 산업적으로 적용될 수 있는 분야에 대해서 연구를 진행하였다[4-13]. 반면에 다양한 분야에서 블록체인에 대해 각자의 관점을 가지고 논의가 이루어지고 있지만 아직까지 블록체인에 관련된 연관 이슈와 중요도에 대한 체계적인 연구는 크게 주목받고 있지 못한 상황이다. 블록체인은 뉴스, 소셜미디어, 커뮤니티 사이트, 학술분야 등 다양한 관점에서 언급되고 있다. 이러한 사실을 감안했을 때 각 분야의 관점에서 블록체인에 대한 연관 이슈의 평가와 공통적인 관점 및 차이점을 살펴보고 향후 블록체인이 나아가야 할 방향을 제시해 줄 필요가 있다. 다시 말해서 블록체인의 견해에 대한 체계적인 분석과 정리가 필요한 상황이라고 할 수 있다. 하지만 아직까지 이에 대한 연구는 거의 수행되지 않았다. 따라서 본 연구는 국내 블록체인 관련 텍스트 데이터를 수집하여 뉴스, 웹사이트, 학술분야에서 블록체인에 대한 연관 이슈와 중요도를 분석하고자 한다.

본 연구의 연구 문제는 다음과 같다. 첫 번째, 뉴스, 웹사이트, 학술분야에서 바라보는 블록체인 연관 이슈는 무엇인가? 두 번째, 각 분야마다 블록체인에 대한 공통적 연관 이슈는 무엇인가? 세 번째, 각 분야별로 블록체인에 대한 상반되는 연관 이슈는 무엇인가? 마지막으로, 각 분야별로 블록체인에 대해 보완되어야 할 진행 방향은 무엇인가? 본 연구는 서로 다른 세 분야에서 바라보는 블록체인에 대한 연관 이슈 및 중요도를 분석함으로써 블록체인에 대한 공통적인 관심사와 대비되는 관점을 제시할 수 있을 것이다. 또한, 본 연구의 분석을 토대로 향후 연구 방향 및 분야별 필요한 시사점을 제공할 수 있을 것이다.

II. 이론적 배경

2-1 블록체인

블록체인은 모든 참여자가 분산 네트워크를 통해 정보와 가치를 검증함으로써 임의적인 조작이 어렵도록 설계된 분산 시스템 기술이다[1]. 블록체인은 기존의 중앙 집중식 방식과 다르게 모든 참여자들이 조작이 거의 불가능한 원장을 공유할 수 있기 때문에 중개자 없이 거래의 신뢰를 확보할 수 있다는 장점을 가지고 있다[2]. 블록체인을 구성하는 블록은 블록 헤더와 바디로 구성된다. 블록의 헤더 부분은 블록 생성의 난이도, 논스값, 이전 블록의 해시값, 현재 블록의 생성 시간, 거래정보들을 해싱한 값이 포함되며, 블록 바디는 거래정보를 포함한다. 블록 헤더의 정보를 이용해 현재 블록의 고유한 해시값을 도출해낼 수 있으며, 이러한 해시값은 암호화 해시함수를 통과한 값이다. 해시함수는 입력값에 아주 미세한 변화가 발생할 경우 결과값이 완전히 다르게 도출되기 때문에 결과값을 보고 일정한 패턴이나 공식을 이용하여 입력값을 찾아내는 것은 거의 불가능하게 구성되어 있다[2]-[3]. 해시함수는 입력값을 알고 있을 경우 결과값을 계산하는 것은 매우 간단하지만 결과값만을 가지고 입력값을 찾아내는 것은 불가능하다. 입력값과 결과값 사이의 일정한 규칙이나 패턴이 존재하지 않기 때문이다. 결과값으로 입력값을 찾아낼 수 있는 방법은 임의의 입력값을 반복적으로 대입 계산하여 올바른 결과값이 나오는 입력값을 발견하는 방법뿐이다. 따라서 블록의 거래정보를 조작할 경우 결과값은 블록의 완전히 다른 해시값이 도출되고 원래의 블록 해시값으로 설정하기 위한 많은 노력이 필요로 한다. 새로운 블록을 생성하기 위해서는 이전 블록의 고유 해시값을 필요로 하기 때문에 블록 간의 체인 구조가 형성된다. 이러한 체인 구조로 인해 블록 내의 정보를 조작을 위해서는 블록체인 내에 등록되어 있는 모든 블록을 조작하여야 한다. 또한 새로운 블록이 지속적으로 생성되어 체인이 형성되고 있기 때문에 블록체인에 기록된 정보를 조작하는 것은 현실적으로 거의 불가능하다고 할 수 있다 [3].

하지만 블록체인은 보완되어야 할 부분이 존재한다. 그 중 첫 번째는 네트워크의 컴퓨터 연산능력이 51% 이상 장악당하면 거래정보가 조작당할 위험이 존재한다. 따라서 블록체인 네트워크를 구성할 때 컴퓨터 연산능력의 분산화가 선행되어야 한다. 두 번째, 현재 블록의 생성주기는 신속한 업무 프로세스를 지원하기에 한계가 존재한다. 블록체인에 정보가 등록되기 위해서는 정보를 담은 새로운 블록이 생성되고 블록체인 네트워크에 공식적으로 등록되어야 한다. 이러한 시스템 구조는 적시에 빠른 정보처리가 요구되는 작업에는 단점으로 작용할 수 있다.

2-2 블록체인 선행연구

해외 블록체인 관련 선행연구는 주로 블록체인의 시스템 요인인 코인 가격에 대해 연구를 수행하였다[3]-[10]. 코인 관련 대부분의 연구들은 주로 비트코인을 대상으로 가격예측 분석을 실시하였으며, 소셜미디어에서 언급되는 코인에 대한 의견이 코인 가격과 관련 있다는 사실을 발견하였다[3], [6]-[7].

또한, 원유 가격, S&P 500, 환율 등의 거시경제 변수가 코인 가격예측에 연관되어 있었다는 사실을 발견하였으며[4], 블록체인에서 발생하는 블록 생성 수, 블록 생성 난이도, 발행된 코인 양 등의 시스템 정보도 코인가격과 관련되어 있다는 사실을 발견하였다[5]. 블록체인 관련 연구가 대부분 코인에 대해서 이루어졌지만 산업에 적용될 수 있는 방안에 대한 연구도 일부 이루어졌다.

일부 연구에서는 사물 인터넷에 블록체인이 적용될 수 있는 방안에 대해 연구를 진행하였다. 사물인터넷과 블록체인을 연계하여 보안 관련 문제를 해결하고자 하는 논의가 이루어졌다[13]. 다른 일부 연구에서는 블록체인이 헬스케어 분야에 적용될 수 있는 예시를 제시하면서 블록체인을 활용한 의료데이터의 자유로운 권한 이동에 대해 설명하였다[11]. 또 다른 연구에서는 블록체인의 한계점과 대응방안을 발견하기 위해 개인을 대상으로 블록체인에 대한 설문 조사를 진행하였다[12].

국내 블록체인 관련 연구에서는 블록체인을 기반으로 IoT 기술을 구현할 때 발생할 수 있는 보안 문제에 대해 논의하거나[14], 핀테크의 보안 문제와 기술에 대해 블록체인의 역할과 적용방안을 제시하는 연구가 수행되었다[15]. 또한, 일부 연구에서는 블록체인이 부동산산업 분야에서 참여자들 간의 신뢰성을 향상시킬 수 있는 방안을 논의하거나 [16], 블록체인 기술을 실현하기 위해 극복해야 할 법적 과제에 대한 논의를 진행하였다[17]. 또 다른 연구에서는 블록체인 기술수용의도에 블록체인에 대한 성과 기대감과 사회적 영향력이 영향을 미친다는 사실을 발견하였다[18]. 국내 연구의 흐름을 요약하자면, IoT, 금융, 법률, 블록체인을 사용하는 이용자의 관점에서의 연구들이 이루어져 왔다는 것을 볼 수 있으며 대부분의 연구들이 블록체인 기술에 대한 논의 위주의 연구를 수행하였다는 것을 볼 수 있다.

반면에 아직까지 학술분야를 포함해서 뉴스 및 웹사이트 분야별로 블록체인에 대한 연관 이슈나 중요도를 체계적으로 분석한 연구는 거의 수행되지 않았다. 이러한 인식에 기반으로 본 연구는 텍스트 마이닝 기법을 활용하여 뉴스, 웹사이트, 학술분야에서 인식하는 블록체인 연관 이슈와 중요도를 분석하고자 하였다.

III. 연구방법

텍스트 마이닝은 대규모 텍스트 데이터로부터 의미 있는 정보와 가치를 추출하는 데이터 마이닝 기법이다[19]. 텍스트 마이닝은 전자상거래를 이용하는 고객의 리뷰나 특정 주제에 대한 소셜미디어 이용자의 의견을 분석하는 등 다양한 분야에서 활용되고 있다[6]. 텍스트 마이닝은 텍스트 데이터에서 불필요한 데이터를 제거 및 수정하는 전처리 작업을 통해 분석 가능한 정형화된 구조로 변환된다. 텍스트 데이터에서 추출한 정보를 시각화하고 텍스트 간에 관계를 파악하여 새로운 가치를 제시

한다는 점에서 텍스트 마이닝 기법은 기존의 정성적 분석의 한계를 넘어서는 체계적 분석방법이라 할 수 있다[20]-[21].

본 연구는 뉴스, 웹사이트, 학술분야에서 바라보는 블록체인에 대한 연관 이슈와 중요도를 살펴보기 위하여 텍스트(Textom)을 활용하여 텍스트 마이닝을 실시하였다. 텍스트를 활용하여 블록체인에 대한 키워드로 2017년 12월부터 2018년 12월까지 1년간의 데이터를 수집하였다. 네이버 검색 기준으로 뉴스기사, 일반 웹사이트, 학술분야 검색 분야가 포함되었다. 수집된 데이터는 뉴스, 웹사이트, 학술분야별로 데이터를 구분하였으며, 데이터 분석을 위해 수집한 데이터의 전처리를 수행하였다. 전처리 단계에서는 수집된 키워드의 교정, 통제 및 제거를 수행한다. 텍스트의 데이터 전처리 기능을 활용하여 전처리 과정에서 엑셀표 기능의 찾기 및 변경 기능을 활용하여 !@~& 등과 같은 특수문자를 제거하였으며, 명사를 제외한 형태의 텍스트는 모두 제외하였다. 또한 유의미한 의미를 가지지 않는 는, 을, 이다, 가 등과 같은 조사적 성격을 띠는 텍스트를 모두 제외하였다. 전처리 과정을 마친 정제데이터를 분야별로 나누는 후 텍스트의 분석기능을 활용하여 결과를 해석하기 위한 단어빈도, N-gram, TF-IDF(Term Frequency-Inverse Document Frequency, 단어빈도-역문서 빈도), 연결중심성을 활용하였다. 또한, N-gram 네트워크 분석을 수행하여 보다 직관적이고 구체적인 설명을 제공하고자 하였다. 최종적으로 텍스트 마이닝 분석에 사용된 정제된 데이터 수는 총 2,144개이며, 뉴스 592개, 웹사이트 810개, 학술분야 742개로 구성되었다.

IV. 분석결과

본 연구는 뉴스, 웹사이트, 학술분야에서 바라보는 블록체인에 대한 연관 이슈와 중요도를 확인하기 위해 텍스트를 이용하여 주요 단어에 대한 정제과정과 텍스트 마이닝을 실시하였다. 텍스트 마이닝을 통해 주요 단어들의 단어빈도(Frequency), N-gram, TF-IDF, 연결중심성(Degree Centrality)을 도출하여 블록체인에 대한 연관 이슈와 중요도를 살펴보았다. 단어빈도는 추출된 단어와 데이터 내 해당 단어의 빈도수를 나타내며, 빈도수가 높다는 것은 정제데이터 내에서 해당 단어가 등장하는 빈도가 높다는 것을 의미한다. N-gram은 n개 단어의 연쇄를 의미하며, 연결된 단어의 빈도가 높다는 것은 두 단어가 나란히 등장하는 빈도가 높다는 것을 의미한다. TF-IDF는 단어빈도(TF)와 IDF(문서빈도의 역수)를 곱한 값으로 단어가 특정 문서에서 얼마나 중요한지를 나타내고, 연결중심성은 연결된 단어가 얼마나 많이 연결되어 있는지를 나타낸다.

텍스트 마이닝의 10개 상위 단어를 기준으로 분석한 결과, 뉴스 분야의 단어빈도는 블록체인 (1,992), 빗썸 (473), 코인 (446) 등 순으로 나타났으며, N-gram에서는 블록체인시황 - 빗썸 (420), 코인 - 상승세 (322), 빗썸 - 코인 (254) 등 순으로 나타났다. TF-IDF에서는 암호화폐 (527.8956), 상승세 (415.2754),

플랫폼 (333.1536) 등 순으로 나타났다. 마지막으로 연결중심성을 기준으로서는 블록체인 (0.0686), 암호화폐 (0.0139), 플랫폼 (0.0074) 등 순으로 나타났다.

웹사이트 분야의 단어빈도는 블록체인 (7,806), 암호화폐 (1,562), 제로페이 (1,160) 등 순으로 나타났으며, N-gram의 경우는 제로페이 - 수수료 (580), 블록체인 - 암호화폐 (437), 이재갑 - 장관 (396) 등 순으로 나타났다. TF-IDF에서는 암호화폐 (1716.0323), 총리 (1312.0307), 이낙연 (1148.0268) 등 순으로 나타났으며, 연결중심성은 블록체인 (0.0985), 암호화폐 (0.0209), 기술 (0.0112) 등 순으로 나타났다.

학술분야에서 단어빈도는 블록체인 (1,030), 개발 (812), 시스템 (803) 등 순으로 나타났으며, N-gram은 시스템 - 개발 (85), 블록체인 - 기반 (66), 블록체인 - 기술 (62) 등 순으로 나타났다. TF-IDF는 개발 (1784.1463), 설계 (1150.3095), 시스템 (1113.1943) 등 순으로 나타났으며, 연결중심성은 개발 (0.0209), 블록체인 (0.0174), 시스템 (0.0137) 등 순으로 나타났다. 구체적인 분석결과는 아래의 표와 그림으로 정리하였다.

표 1. 텍스트마이닝 분석 결과: 단어빈도

Table 1. Text Mining Analysis Results: Frequency

Rank	News	Web Site	Academic Field
1	Blockchain (1,992)	Blockchain (7,806)	Blockchain (1,030)
2	Bithumb (473)	Cryptocurrency (1,562)	Development (812)
3	Coin(446)	Zeropay (1,160)	System (803)
4	Blockchain Market (420)	Platform (747)	Design (642)
5	Upward (378)	Technology (743)	Structure (633)
6	Cryptocurrency (328)	Commission (625)	Analysis (547)
7	Formation (217)	Exchange (593)	Block (448)
8	Plat from (207)	Gwangchul Go (592)	Base (414)
9	Technology (182)	Market (587)	Data (375)
10	Industry (171)	Prime Minister (528)	Configuration (374)

표 2. 텍스트마이닝 분석 결과: N-gram

Table 2. Text Mining Analysis Results: N-gram

Rank	News	Web Site	Academic Field
1	Blockchain Market - Bithumb (420)	Zeropay - Commission (580)	System - Development (85)
2	Coin - Upward (322)	Blockchain - Cryptocurrency (437)	Blockchain - Base (66)
3	Bithumb - Coin (254)	Jaekap Lee - Minister (396)	Blockchain - Technology (62)
4	Formation - Blockchain Market (194)	Nakyon Lee - Prime Minister (396)	Control - System (51)
5	Upward - Formation (174)	Market - Coin (297)	Performance - Comparison (35)
6	Bithumb - Lite (124)	Blockchain - Platform (280)	Dispersion - Restoration (33)

7	Lite - Coin (123)	Blockchain - Technology (273)	System - Design (32)
8	Bithumb - Bitcoin (88)	Cryptocurrency - Exchange (158)	System - Configuration (31)
9	Bitcoin - Indicator (88)	Blockchain - Expo (139)	State - Monitoring (30)
10	Blockchain - Technology (84)	Prime Minister - Speakings (132)	Application - Control (29)

표 3. 텍스트마이닝 분석 결과: TF-IDF

Table 3. Text Mining Analysis Results: TF-IDF

Rank	News	Web Site	Academic Field
1	Cryptocurrency (527.895635)	Cryptocurrency (1716.032394)	Development (1784.1463568)
2	Upward (415.275445)	Prime Minister (1312.030711)	Design (1150.309579)
3	Platform (333.153647)	Nakyon Lee (1148.026872)	System (1113.194371)
4	Bithumb (327.858616)	Minister (1093.727919)	Structure (877.524330)
5	Technology (326.100223)	Coin (1003.927100)	Analysis (758.303015)
6	Coin (309.143642)	Jaekap Lee (984.023033)	Blockchain (713.941595)
7	Industry (306.390869)	Market (944.7400549)	Control (694.072077)
8	Bitcoin (294.42809)	State Council (820.019194)	Base (666.307295)
9	Exchange (292.428306)	Statement (817.594459)	Comparison (645.033408)
10	Blockchain Market (291.121815)	Virtual Currency (701.714929)	Block (621.059873)

표 4. 텍스트마이닝 분석 결과: 연결중심성

Table 4. Text Mining Analysis Results: Degree Centrality

Rank	News	Web Site	Academic Field
1	Blockchain (0.06869868)	Blockchain (0.09852790)	Development (0.02098986)
2	Cryptocurrency (0.01395044)	Cryptocurrency (0.02090321)	Blockchain (0.01743380)
3	Platform (0.00744750)	Technology (0.01120013)	System (0.01375009)
4	Technology (0.00617597)	Platform (0.01053477)	Structure (0.01174410)
5	Exchange (0.00483179)	Virtual Currency (0.00873277)	Design (0.01150703)
6	Industry (0.00443217)	Transaction (0.00845554)	Analysis (0.01128820)
7	Investment (0.00406888)	Exchange (0.00643175)	Comparison (0.00837041)
8	Contents (0.00377824)	Company (0.00629314)	Block (0.00827923)
9	Transaction (0.00359659)	Investment (0.00615452)	Data (0.00780509)
10	Market (0.00348761)	Government (0.00615452)	Configuration (0.00758625)

분석결과 기준으로 뉴스, 웹사이트, 학술분야에서 바라보는 블록체인에 대한 연관 이슈와 중요도를 비교 분석해보도록 하겠다. 먼저 단어빈도에서는 블록체인이 모든 분야에서 공통적으로 나타났으며, 암호화폐, 플랫폼, 기술이라는 키워드가 뉴스와 웹사이트 분야에서 동일하게 높은 빈도를 갖는 것으로 나타났다. 뉴스는 블록체인이 적용된 암호화폐와 관련된 빙셋, 코인, 블록체인 시황과 같은 산업 내의 시장 관련 단어빈도가 주

를 이루었다. 웹사이트는 서울시 결제플랫폼이 제로페이와 같은 공공 결제 플랫폼, 고광철, 수수료, 총리와 같은 정책관련 키워드가 주로 나타났다. 이와 다르게 학술분야는 기술개발에 초점이 맞추어져 있었으며 개발, 시스템, 설계, 분석, 데이터 등의 필요한 기술적 자원과 기술적 개발 과정에 대한 단어빈도가 높게 나타났다.

다음으로 N-gram에서는 블록체인 키워드를 기반으로 단어 네트워크가 형성되어 있으나 가장 높은 빈도수는 뉴스에서 블록체인과 빗썸, 웹사이트에서 제로페이와 수수료, 학술분야에서 시스템과 개발로 다른 양상을 보였다. 뉴스는 블록체인과 빗썸, 코인과 상승세, 빗썸과 코인이 나란히 등장하는 빈도가 가장 높게 나타나 블록체인 기술기반의 암호화폐 기업인 빗썸의 코인 가격 및 시장 현황에 대한 관심이 주를 이루는 것으로 나타났다. 그 외에 블록체인 관련 기술, 산업, 플랫폼 등과 같은 블록체인 기술의 기술 발전을 위한 부분이 일부 네트워크를 구축하는 것으로 나타났다. 웹사이트는 제로페이와 수수료, 블록체인과 암호화폐, 이재갑과 장관, 이낙연과 총리가 나란히 등장하는 빈도가 높았는데, 이는 블록체인 기술이 활용한 암호화폐에 대한 관심과 블록체인 기반의 공공 결제 플랫폼인 제로페이가 플랫폼에 대한 관심이 수수료에 기반하여 작용했음을 확인할 수 있다. 또한 이재갑과 이낙연과 같은 블록체인 관련 정책에 목소리를 낸 정치인에 대한 관심 역시 크게 나타났다. 학술분야는 상대적으로 주요 관계를 파악하기보다는 시스템과 개발, 블록체인과 기반, 블록체인과 기술이 크게 나타나 블록체인 기술을 활용한 시스템을 개발 및 구현을 위한 다양한 기술과 효과에 관심을 보였다. 이에 대한 직관적인 이해를 돕고자 텍스트의 데이터 시각화 기능을 활용하여 N-gram 네트워크로 구체화하였다. 그림을 살펴보면 상위 10개 단어연결이 표현되어 있으며, 선의 연결이 굵고 진할수록 연결된 단어의 빈도가 높음을 의미한다.

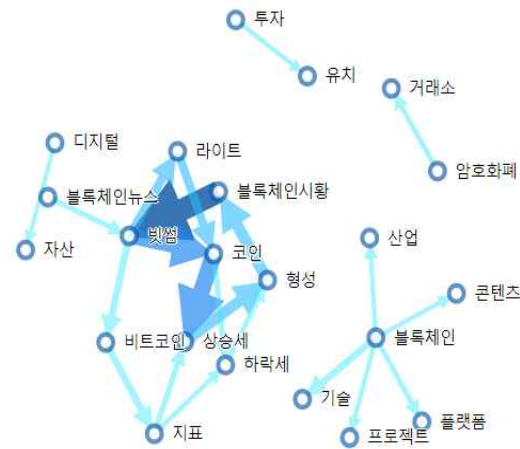


그림 1. 뉴스 N-gram 네트워크
Fig. 1. News N-gram Network

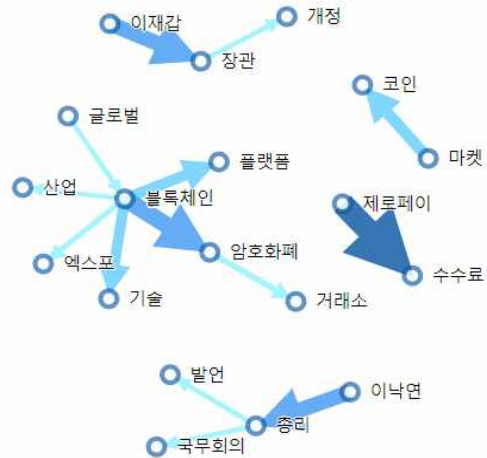


그림 2. 웹사이트 N-gram 네트워크
Fig. 2. Web Site N-gram Network

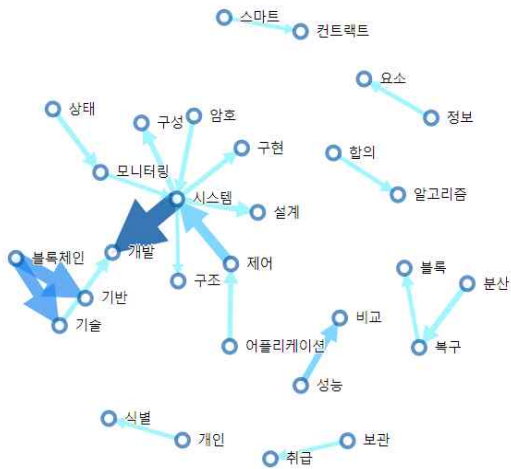


그림 3. 학술분야 N-gram 네트워크
Fig. 3. Academic Field N-gram Network

또한 TF-IDF 분석결과 뉴스와 웹사이트에서는 블록체인보다 이것이 구현된 암호화폐가 가장 중요한 단어로 여겨져 기술의 응용된 상품화에 집중되어 있지만, 학술분야에서는 블록체인 자체의 기술적 개발과 시스템을 구현하는 방안에 집중되어 뚜렷한 차이를 보였다. 뉴스에서는 암호화폐가 가장 중요했으며 관련 시장의 상승과 하락세에 대한 관심이 주를 이루었고, 웹사이트에서는 암호화폐가 가장 중요했으나, 암호화폐 시장에 대한 정치인의 발언 및 정책에 대한 관심이 주를 이루었다. 학술분야에서는 개발이 가장 중요하게 나타났으며, 이를 위한 설계, 시스템, 구조, 분석 등의 순으로 중요도를 보이며 블록체인에 대한 기반 기술 및 구현을 위한 관심이 주를 이루었다.

마지막으로 연결중심성 측면에서도 뉴스, 웹사이트, 학술분야에서 관련 단어 파생 정도에서 뚜렷한 차이를 보였다. 뉴스와 웹사이트 모두 블록체인이라는 핵심키워드가 많은 단어를 파

생하며 다른 단어보다 지배적으로 높은 중심성을 보였으나, 학술분야에서는 개발, 블록체인이라는 핵심키워드를 바탕으로 단어들이 파생되어 비교적 다양한 연구주제에 대한 관심이 나타나고 있음을 알 수 있었다. 구체적으로 뉴스와 웹사이트에서는 블록체인이 가장 높은 수치를 보였고, 암호화폐, 플랫폼, 기술의 순으로 나타나 암호화폐 측면의 기술과 플랫폼이 핵심 단어 역할을 수행하고 있었다. 즉, 블록체인 기술이 적용된 암호화폐 플랫폼과 시장에 대한 관심이 주를 이루나, 뉴스는 산업의 형성 및 거래 측면에서 웹사이트는 이에 대한 투자와 정부의 역할적 측면에서 주로 접근하는 것으로 보인다. 학술분야는 개발이 가장 높게 나타났으며, 블록체인, 시스템, 구조 등의 순으로 나타나 블록체인이라는 기술에 대한 연구개발 측면이 접근이 이루어지고 이를 위한 데이터, 설계, 시스템, 비교 등이 핵심 단어로서의 역할을 수행하고 있었다.

결과적으로 뉴스, 웹사이트, 학술분야에서 바라보는 블록체인에 대한 연관 이슈와 중요도는 서로 다르다는 것을 발견할 수 있었다. 뉴스에서는 암호화폐의 가격 등락에 따른 시세와 투자에 대한 관심이 지배적이었으며, 이러한 관심은 빗썸이라는 암호화폐거래소에 집중된 모습을 보였다. 웹사이트에서는 블록체인과 암호화폐 관련 정부 관련 인사의 의견 표현, 규제와 정책에 대한 관심이 지배적이었으며, 블록체인 기반의 결제플랫폼인 제로페이에 대한 관심도 함께 증가하는 것으로 나타났다. 제로페이에 대한 관심은 자연스럽게 결제플랫폼의 수수료와 연결되었다. 학술분야에서는 블록체인 관련 기술개발과 관련 시스템에 관련된 내용이 지배적이었으며, 블록체인 기술의 산업 적용방안과 시스템에 대한 의견이 지배적이었다. 요약하자면, 뉴스는 블록체인 산업의 시장형성과 동향 측면에서 접근하고, 웹사이트는 블록체인 기술의 사업화와 공공활용에 따른 정책 측면에서 접근하며, 학술분야는 블록체인 기술의 지속적인 발전을 위한 연구개발 측면에서 접근하는 것을 파악할 수 있었다. 이를 통해 뉴스, 웹사이트, 학술분야와 같은 서로 다른 영역에서 정보통신기술을 바라보는 연관 이슈와 중요도의 흐름을 살펴봄으로써 향후 블록체인 기술개발 및 발전 방향성 수립에 대한 역할 수행 측면에서 다양한 이해관계자의 의견 수렴 및 반영에 대한 필요성을 확인할 수 있었다.

V. 결 론

본 연구는 국내의 각 분야별 블록체인에 대한 연관 이슈와 중요도를 분석하고자 하였다. 뉴스, 웹사이트, 학술분야에서 언급되는 블록체인에 대한 텍스트 분석결과, 세 분야마다 블록체인에 대해 서로 상반되는 이미지를 가지고 있다는 사실을 발견하였다. 뉴스 분야에서 언급되는 블록체인은 주로 비트코인과 같은 암호화폐의 정보와 시세에 대한 관심이 지배적이었다. 암호화폐 가격의 등락과 이에 따른 투자에 관심이 상당히 높으며 관련 코인들을 거래할 수 있는 거래소에 대해서도 관심이 높은

것으로 나타났다. 특히 암호화폐 거래소인 빗썸에 대한 내용이 지배적이었다. 부가적으로는 블록체인을 활용한 플랫폼과 기술에 대한 내용도 언급되었다. 각종 웹사이트에서 언급되는 블록체인의 연관 이슈는 암호화폐 시황과 정부의 영향력에 대한 것들이었다. 특히 블록체인과 암호화폐에 관련한 정부 인사에 대한 정보가 빈번하게 등장하는 것을 볼 수 있었다. 정부의 암호화폐 규제로 인해 암호화폐 규제와 정책 실행에 관한 인물들의 관심도 높아지고 있다고 할 수 있다. 또한, 암호화폐에 대한 규제와는 다르게 블록체인을 이용한 산업 및 기술개발에 정부와 관련 인사들의 영향력이 반영된 결과라고 볼 수 있다. 추가적으로 블록체인을 기반으로 한 결제 플랫폼인 제로페이에 대한 의견도 등장하였으며 이러한 결제 플랫폼이 수수료와 관련되어 있다는 사실을 발견하였다. 제로페이는 블록체인 기반 결제 시스템으로써 수수료에 대한 부담이 거의 없기 때문에 이와 같은 결과가 나타난 것으로 볼 수 있다. 반면에 학술분야에서 바라보는 블록체인은 주로 블록체인의 기술개발과 관련 시스템에 대한 내용이 대부분이었다. 암호화폐나 정부 정책에 대한 의견보다는 블록체인을 산업에 적용할 수 있는 방안과 블록체인을 도입한 시스템이 현재 시스템에 비해 얼마나 큰 이점을 가져다줄 수 있는지에 대한 의견이 대부분이었다. 다시 말해서, 학술분야에서는 블록체인을 활용한 기술개발이 중요한 화두라고 할 수 있다.

본 연구는 블록체인이라는 정보통신기술의 도입과 적용 측면에서 다양한 분야의 연관 이슈 및 중요도를 도출함으로써 다음과 같은 시사점을 갖는다. 블록체인에 접근하는 다양한 분야 중 뉴스, 웹사이트, 학술분야의 텍스트마이닝을 통해 연관 이슈와 중요도를 비교 분석하고 공통점과 차이점을 발견했다는 점에서 의의가 있다. 구체적으로 뉴스는 기술기반의 산업형성 및 시장현황이 주요 이슈이며, 웹사이트는 기술기반의 사업화와 공공활용에 대한 정책수립과 여론 형성 및 반영이 주요 이슈이고, 학술분야는 기술 자체의 발전을 위한 다양한 연구개발과 기술적용에 대한 혜택이 주요 이슈인 것으로 나타났다. 각 분야에서 바라보고 인식하는 블록체인의 연관 이슈와 중요도를 살펴봄으로써 향후 기술기반의 산업이 형성되는 과정에서 사업화, 관련 정책 수립, 연구개발 측면에서 각각의 사회 구성원들의 주요 역할을 파악할 수 있었다. 이에 각각의 분야는 자신의 주요 역할 수행을 인지하고 서로 보완되어야 할 부분과 한계점을 인식한다면 블록체인의 지속적인 발전에 기여할 수 있을 것이다. 또한, 암호화폐 가격예측 분석에서도 소셜네트워크서비스(SNS) 사용자의 의견만을 고려하기보다 뉴스, 웹사이트, 학술분야의 의견도 적극 고려할 필요성이 있어 보인다. 기존 암호화폐의 가격예측에는 소셜 네트워크와 같은 대중의 심리가 반영되는데, 이러한 심리를 대변하듯 암호화폐 공포와 탐욕지수(Crypto Fear & Greed Index)[22]도 생겨났다. 그리고 이러한 심리는 유명인사의 가격전망이나 가치에도 큰 영향을 받기 때문에, 불안정하며 여러 요소가 중복적으로 적용됨에 따라 과대 또는 과소평가로 귀결될 수 있다. 따라서 암호화폐 가격예측 연구에서 서로 다른 관점을 가진 분야의 의견들을 고려하여 새로운

변수들을 발굴하고 가격 예측력을 높일 수 있을 것이다.

본 연구는 다음과 같은 한계점 및 향후 연구 방향을 제안한다. 첫 번째, 본 연구는 블록체인 연관 이슈와 중요도를 규명하기 위해 텍스트 데이터 범위를 국내의 주요 검색엔진인 네이버의 뉴스, 웹사이트, 학술분야로 한정하여 분석을 수행하였다. 따라서 결과의 일반화 측면에서 신중해야 할 것이며, 향후에는 페이스북, 인스타그램, 커뮤니티 등과 같은 다양한 데이터를 함께 수집하여 연구를 진행해야 할 것이다. 최근에는 소셜 네트워크에서 블록체인에 대한 논의가 활발하게 진행되고 있다. 따라서 향후 연구에서는 블록체인에 대한 소셜 네트워크 서비스 사용자의 의견을 고려해야 할 것이다. 이를 통해 일반인들의 연관 이슈와 중요도에 대한 분석 결과를 확보함으로써 보다 일반화 가능성이 높고 명확한 분석 결과를 확보할 수 있을 것이다. 두 번째, 본 연구는 블록체인에 대한 2018년에 해당하는 1년 동안의 데이터만을 활용하였다. 블록체인은 기술적인 한계도 존재하지만, 아직 발전 중이며 기술 발전과 서비스 다양화에 대한 논의가 활발하게 진행되고 있다. 블록체인은 향후 인공지능(AI)이나 사물인터넷(IoT)과 같은 4차 산업혁명을 이끄는 핵심 기술들과도 융복합될 가능성도 존재한다. 그럼에도 불구하고 2018년이 중요한 이유는 블록체인 기술의 세대를 구분하는 기준점이기 때문이다. 블록체인 기술은 1세대의 단순 지급 수단 기능에 대한 검증은 거친 후 2세대에는 다양한 거래 및 계약 기능이 추가되면서 활용범위를 넓혀나가고 있다[23]. 즉, 블록체인 기술을 사용한 가상통화가 본격적으로 출현하고 기능적 다양화로 활용범위를 넓혀나가기 시작한 기점이다[23]. 또한 한국에서 2001년 과학기술기본법 제정으로 기술영향평가가 의무화된 이래 총 19건의 기술을 평가하였는데, 2018년에는 블록체인 기술이 선정되었다[24]. 기술영향평가는 전문가 중심과 시민참여형을 절충한 방식으로 운영되므로 기술도입에 대한 다양한 분야의 이슈들이 본격적으로 제기된 해라고 볼 수 있다. 비록 가장 최신의 블록체인에 대한 연관 이슈와 중요도를 분석할 필요성도 있겠지만 본 연구는 혁신기술의 사회 전반 적용의 초기 단계 측면에서 블록체인에 대한 연관 이슈와 중요도를 분석하였다. 이는 새롭게 등장할 신기술이 사회 전반에 적용되는 단계에 겪는 다양한 연관 이슈를 중요도에 따라 어떻게 관리할 것인지에 대한 지침을 제공할 수 있을 것이다. 향후 연구에서는 연도별로 블록체인에 대한 연관 이슈와 중요도의 흐름을 체계적으로 분석하여 보다 장기적인 측면에서 접근할 필요가 있어 보인다. 세 번째, 본 연구는 국내의 블록체인에 대한 인식만을 텍스트 마이닝으로 분석하였다는 한계점을 가지고 있다. 국내에서 수행되는 블록체인 관련 사업이나 연구 동향을 파악하고 향후 방향성을 제시하였지만 국내 이외에도 해외에서 바라보는 블록체인의 관점과 동향은 국내와 차이가 존재할 수 있다. 따라서 향후 연구에서는 국내와 해외의 블록체인 연관 이슈 및 중요도를 분석하여 제시하는 연구가 필요할 것이다. 예를 들어서, 미국, 유럽 국가, 한국 등 국가별로 블록체인에 대한 연관 이슈와 중요도를 파악하고 비교 분석하는 연구를 진행할 수도 있을 것이다.

참고문헌

- [1] J. Abraham, D. Higdon, J. Nelson, and J. Ibarra, "Cryptocurrency Price Prediction Using Tweet Volumes and Sentiment Analysis," *SMU Data Science Review*, Vol. 1, No. 3, pp. 1-23, 2018.
- [2] A. M. Antonopoulos, *Mastering Bitcoin: Unlocking Digital Cryptocurrencies*, pp. 298, O'Reilly Media Inc, 2014.
- [3] A. M. Antonopoulos, G. Wood, *Mastering Ethereum: Building Smart Contracts and Apps*, pp. 200, O'Reilly Media, 2018.
- [4] P. Ciaian, M. Rajcaniova, and D. A. Kancs, "The Economics of Bitcoin Price Formation," *Applied Economics*, Vol. 48, No. 19, pp. 1799-1815, 2016.
DOI: <https://doi.org/10.1080/00036846.2015.1109038>
- [5] H. Jang and J. Lee, "An empirical Study on Modeling and Prediction of Bitcoin Prices with Bayesian Neural Networks based on Blockchain Information," *IEEE Access*, Vol. 6, pp. 5427-5437, 2018.
DOI: <https://doi.org/10.1109/ACCESS.2017.2779181>
- [6] Y. B. Kim, J. G. Kim, W. Kim, J. H. Im, T. H. Kim, S. J. Kang, and C. H. Kim, "Predicting Fluctuations in Cryptocurrency Transactions based on User Comments and Replies," Vol. 11, No. 8, *PLoS ONE*, 2016.
DOI: <https://doi.org/10.1371/journal.pone.0161197>
- [7] L. Kristoufek, "BitCoin Meets Google Trends and Wikipedia: Quantifying the Relationship between Phenomena of the Internet Era," *Scientific Reports*, Vol. 3 : 3415, 2013.
DOI: <https://doi.org/10.1038/srep03415>
- [8] D. C. Mallqui and R. A. Fernandes, "Predicting the Direction, Maximum, Minimum and Closing Prices of Daily Bitcoin Exchange Rate Using Machine Learning Techniques," *Applied Soft Computing*, Vol. 75, pp. 596-606, 2019.
DOI: <https://doi.org/10.1016/j.asoc.2018.11.038>
- [9] W. Mensi, K. H. Al-Yahyaee, and S. H. Kang, "Structural Breaks and Double Long Memory of Cryptocurrency Prices: A Comparative Analysis from Bitcoin and Ethereum," *Finance Research Letters*, Vol. 29, pp. 222-230, 2018.
DOI: <https://doi.org/10.1016/j.frl.2018.07.011>
- [10] O. Poyser, "Exploring the Dynamics of Bitcoin's Price: A Bayesian Structural Time Series Approach," *Eurasian Economic Review*, Vol. 9, No. 1, pp. 29-60, 2019.
DOI: <https://doi.org/10.1007/s40822-018-0108-2>
- [11] X. Yue, H. Wang, D. Jin, M. Li, and W. Jiang, "Healthcare

- Data Gateways: Found Healthcare Intelligence on Blockchain with Novel Privacy Risk Control,” *Journal of Medical Systems*, Vol. 40, No. 10(218), pp. 1-8, 2016.
DOI: <https://doi.org/10.1007/s10916-016-0574-6>
- [12] Z. Zheng, S. Xie, H. N. Dai, X. Chen, and H. Wang, “Blockchain Challenges and Opportunities: A survey,” *International Journal of Web and Grid Services*, Vol. 14, No. 4, pp. 352-375, 2018.
DOI: <https://doi.org/10.1504/IJWGS.2018.095647>
- [13] Z. Li, J. Kang, R. Yu, D. Ye, Q. Deng, and Y. Zhang, “Consortium Blockchain for Secure Energy Trading in Industrial Internet of Things,” *IEEE Transactions on Industrial Informatics*, Vol. 14, No. 8, pp. 3690-3700, 2018.
DOI: <https://doi.org/10.1109/TII.2017.2786307>
- [14] H. J. Park, H. E. Yang, and J. H. Jeon, “Blockchain-based IoT Security Vulnerability,” *Korea Information Processing Society Review*, Vol. 24, No. 3, pp. 13-21, 2017.
- [15] S. J. Park, “Blockchain Paradigm and Fintech Security,” *The Journal of The Korean Institute of Communication Sciences(Information and Communications Magazine)*, Vol. 34, No. 3, pp. 23-28, 2017.
- [16] S. Y. Oh and C. H. Lee, “Block Chain Application Technology to Improve Reliability of Real Estate Market,” *The Journal of Society for e-Business Studies*, Vol. 22, No. 1, pp. 51-64, 2017.
DOI: <https://doi.org/10.7838/jsebs.2017.22.1.051>
- [17] S. J. Seoung, “Legal Challenges for Activating Blockchain - Negative Regulation Discussion -,” *Business Law Review*, Vol. 31, No. 2, pp. 325-352, 2017.
DOI: <https://doi.org/10.24886/BLR.2017.06.31.2.325>
- [18] J. S. Kim and G. Y. Gim, “A Study on Factors Affecting the Intention to Accept Blockchain Technology,” *Journal of Information Technology Services*, Vol. 16, No. 2, pp. 1-20, 2017. DOI: <https://doi.org/10.9716/KITS.2017.16.2.001>
- [19] W. Fan, L. Wallace, S. Rich, and Z. Zhang, “Management Information System”, *Communications of the ACM*, Vol. 49 No. 9, pp. 76-82, 2006.
- [20] D. Paranyushkin, “Visualization of Text’s Polysingularity Using Network Analysis”, *Nodus Labs*, 2012. Available: <https://noduslabs.com/research/visualization-text-polysingularity-network-analysis>.
- [21] D. Paranyushkin, “Identifying the pathways for meaning circulation using text network analysis”, *Nodus Labs*, 2011. Available: <https://noduslabs.com/research/pathways-meaning-circulation-text-network-analysis>.
- [22] Alternative.me. Crypto Fear & Greed Index[Internet]. Available:<https://alternative.me/crypto/fear-and-greed-index>.
- [23] H. S. Cho, S. H. Kim, and S. M. Kim, Analysis on Blockchain Industry and Policy for Proliferation, Korea Institute for Industrial Economic & Trade, Sejong: Korea, Research Resort 2019-920, 2019.
- [24] Ministry of Science and ICT, Korea Institute of S&T Evaluation and Planning, The Future of Blockchain(2018 Technology Impact Evaluation Results Report), Sejong: Korea, Research Report, 2019.



김호익(Ho-ik Kim)

2014년 : 계명대학교 KAC International Business (학사)

2017년~현재 : 성균관대학교 경영대학 마케팅 (석박통합과정)

2018년(예정) : University of South Carolina International Business (Doctoral Degree)

※ 관심분야 : 국제 마케팅, 친환경 에너지 및 전기자동차(EV), IoT(Internet of Things), Fintech, Smart device



김한민(Han-Min Kim)

2017년 : 충남대학교 경영학과 경영정보시스템 (경영학석사)

2017년~현 재: 성균관대학교 경영대학 경영정보 (박사과정)

※ 관심분야 : 악성댓글, 빅데이터, 데이터마이닝, 블록체인



이민환(Minh-Hwan Lee)

2013년 : 성균관대학교 경영전문대학원 (경영학석사)

2017년 : 성균관대학교 일반대학원 (경영학박사-마케팅)

2017년~현 재: 성균관대학교 경영학과 초빙/겸임교수

2020년~현 재: 칼빈대학교 글로벌문화산업경영학과 대우/겸임교수

※ 관심분야 : 마케팅전략, 마케팅 커뮤니케이션, 디지털 마케팅, 플랫폼 비즈니스, 빅데이터