JDCS
디지털콘텐츠학회논문지
Journal of Digital Contents Society
Vol. 24, No. 10, pp. 2355-2362, Oct. 2023

Check for updates

# 국가 사이버안보에 관한 사회적 인식 및 함의: 빅데이터 분석기법을 중심으로

김 동 훈[1] · 김 법 헌[2*]
[1]조선대학교 군사학과 박사과정, 국립전파연구원 주무관
[2*]조선대학교 군사학과 교수

# Societal Perception and Implications of National Cybersecurity: Focusing on Big Data Analysis Techniques

**Dong-Hun Kim[1] · Beob-Heon Kim[2*]**

[1]Ph.D. Candidate, Department of Military Science, Chosun University, Gwangju 61452, Korea
[2*]Professor, Department of Military Science, Chosun University, Gwangju 61452, Korea

**[요 약]**

최근 러시아-우크라이나 전쟁에서 사이버안보는 새로운 게임체인저로 대두되었고 국가 기밀 정보 탈취, 사회 기반 시설파괴, 범죄 자금 확보, 가상화폐 탈취 등 다양한 사이버 활동들이 국가안보를 위협하고 있다. 이에 본 연구는 사이버안보에 관한 사회 전반의 논의와 토픽을 언론보도 분석을 통해 사회적 인식과 함의를 제시하는 데 그 목적이 있다. 연구 방법으로는 최근 3년의 언론보도 총 1,211건을 수집하고 다양한 빅데이터 분석기법을 활용하였다. 연구 결과 최근 한미 양국의 동맹 범위를 사이버공간으로까지 확장한 선언과 사이버안보 협력체계 구축, 사이버안보 관련 기관, 위협 국가 등이 특징적 키워드로 도출되었다. 연구 결과를 바탕으로 국가안보의 핵심인 사이버안보 역량 확보를 위한 인재 육성과 제도, 사이버안보법의 조속한 제정, 공동 대응을 위한 민관 협력 강화 등이 요구되었다.

**[Abstract]**

In the recent Russia-Ukraine war, cybersecurity has emerged as a pivotal game-changer. Threats to national security have manifested in various cyber activities, including the theft of national classified information, destruction of societal infrastructure, procurement of crime funds, and theft of virtual currencies. Through an analysis of relevant media reports, this study clarifies the societal perception and implications surrounding cybersecurity. Big data analysis techniques were employed to scrutinize 1,211 media articles from the past three years. The findings revealed notable keywords such as the recent declaration of expanding the alliance scope between the U.S. and South Korea into cyberspace, the establishment of a cybersecurity cooperation system, relevant cybersecurity agencies, and threat nations. Based on these findings, the need for nurturing talent with cybersecurity capabilities at the core of national security, establishing legal frameworks such as cybersecurity laws promptly, and strengthening public-private collaborations is underscored.

**색인어 :** 사이버안보, 언론보도, 사회연결망 분석, 텍스트마이닝, LDA 토픽모델링
**Keyword :** Cybersecurity, Media Reporting, Social Network Analysis, Text Mining, LDA Topic Modeling

**http://dx.doi.org/10.9728/dcs.2023.24.10.2355**



**Received** 09 August 2023; **Revised** 21 August 2023
**Accepted** 06 September 2023

***Corresponding Author; Beob-Heon Kim**
**Tel:** +※ 개인정보 표시제한
**E-mail:** kbh@chosun.ac.kr

## Ⅰ. 서 론

초연결과 초고속을 특징으로 하는 디지털 정보통신기술의 발전으로 현대 사회는 PC 기반의 인터넷 환경과 스마트 기기를 기반으로 한 모바일 네트워크를 통해 모든 일상의 생활 관련된 정보와 기기가 연결된 편리한 사이버 환경 속에서 살아가고 있다. 일상의 삶뿐만 아니라 다양한 사회 인프라, 금융, 산업 등을 포함한 다양한 국가 기반 시설 역시 초연결을 기반으로 운영되고 있다. 이러한 사이버안보의 중요성이 사회적으로 크게 대두된 사건은 지난 2011년 발생하였다. 북한의 소행으로 밝혀진 사이버테러 공격으로 인해 농협 전산망에 있는 대규모 자료가 손상되어 수일 동안 일부 서비스 이용이 마비되며 전 국가적인 혼란을 초래하였다. 당시 정보 보안 분야 전반에 대한 다양한 문제점이 제기되며 이를 보완하는 조치와 노력이 이루어졌고 이러한 사이버안보의 중요성과 위기관리 및 대응 체계 구축의 필요성이 대두되기 시작하였다. 그리고 사이버테러 등에 의한 혼란뿐 아니라 네트워크 기반 시설에 대한 물리적 취약 요소로 인해 사회적 혼란이 발생한 경우가 있었다[1]. 또 2018년 11월 발생한 서울 서대문구 KT 아현지사 통신구 화재 사고로 서울 한강 이북 서부 지역의 통신망이 일시적으로 마비되었고 인근 대학병원의 통신장애 원인이 되어 응급실이 폐쇄되는 등의 문제가 발생하였다. 지난 2022년 10월에는 경기도 성남시 SK 판교 데이터센터 화재로 인해 전 국민이 사용하는 카카오의 대부분 서비스가 중단되며 전 국민적 혼란이 벌어졌다. 그리고 최근 러시아-우크라이나 전쟁을 통해 국가적 차원의 사이버안보의 중요성이 새롭게 대두되었으며 사이버안보는 전쟁의 새로운 게임체인저로 떠오르기도 하였다[2]. 이처럼 네트워크를 기반으로 한 현대 사회는 관련 사이버공격 및 테러뿐만 아니라 네트워크 기반 시설에 대한 물리적 손상 등으로 제 기능을 하지 못할 때, 다양한 사회적 혼란이 발생할 가능성이 존재하며 이러한 부분까지도 포괄적으로 위협을 방지하는 개념의 사이버안보의 중요성이 대두되고 있다. 현대 사회의 급속한 디지털화는 사회 전반을 디지털 전환으로 바꾸어 놓았으며 정치, 경제, 사회, 문화를 포함한 모든 영역에서 사이버 의존성이 높아졌다. 본 연구는 이러한 사회적 상황에 비추어 국가가 지켜야 할 가치의 중심으로 대두된 사이버공간을 보호하는 핵심 개념인 국가적 차원의 사이버안보와 관련된 부분이 언론보도에서 어떠한 내용으로 다루어지고 있으며 어떠한 토픽과 이슈가 있는지를 살펴보고자 한다. 이를 위해 텍스트마이닝 연구방법론을 활용한 최근 3년(2020년 7월 1일 ~ 2023년 6월 30일)의 언론보도 빅데이터 분석을 통해 '사이버안보'와 관련된 언론보도에 등장하는 중점 키워드(Keyword)를 도출하고, 의미연결망 분석을 활용하여 키워드의 상관관계를 살펴본 후, 주요 토픽과 이슈를 분석하여 사이버안보와 관련된 시사점을 도출하고자 한다. 이를 위해 본 연구의 연구 문제를 다음과 같이 설정하여 연구를 진행하고자 한다.

첫째, 언론보도 빅데이터에 나타난 '사이버안보'와 관련된 주요 키워드 특성과 의미연결망의 구조적 특성은 어떠한가?

둘째, 언론보도 빅데이터에 나타난 '사이버안보' 관련 주요 토픽과 이슈는 무엇인가?

셋째, 분석 결과를 통한 '사이버안보' 대응 관련 정책적 시사점은 무엇인가?

이러한 연구 문제 해결을 통해 본 연구는 사이버안보에 관한 사회 전반의 논의 이슈와 토픽을 언론보도를 통해 확인한다는 점에 학술적 의의가 있으며 분석 도출된 연구의 결과는 사이버안보 관련 학술 분야 및 현장 정책 진단과 개선을 위한 시사점을 제공할 수 있을 것으로 기대한다.

본 논문은 서론으로부터 시작하여 이론적 논의, 데이터 분석, 사회연결망 분석 및 LDA 토픽모델링, 논의, 결론까지의 총 6장으로 구성하였다.

## Ⅱ. 이론적 논의

### 2-1 사이버안보의 개념

사이버안보(cybersecurity)는 사이버와 안보의 합성어로 2000년대 들어 다양한 사이버안보 위협이 증대되며 사용되기 시작하였으며 사이버 보호 대책과 관련하여 통용되어 사용되는 단어이다. 이는 국가 차원의 사이버 위협이 다수 발생하자 정부 차원에서 사이버안보 대응 체계를 고민하고 국가 사이버 안전 전략 회의를 통해 종합적인 대책을 마련하는 과정에서 생긴 개념이라고 할 수 있다. 사이버보안 개념과의 차이는 사이버보안은 사이버 공간상의 정보 유출이나 침해 방지를 위한 모든 종류의 보호 대책과 기기, 통신, 사이버공간의 보호를 뜻하며 사이버안보는 국가안보 차원에서 사이버안보 위협 및 예방, 대응에 관한 정책에 활용하는 개념으로 구분하여 이해할 수 있다[3]. 즉, 사이버안보는 사이버상의 위협이나 공격으로부터 안전을 보장하는 것으로 사이버테러를 방지하거나 대응하는 방법 등 모든 개념을 포괄한다고 할 수 있다.

### 2-2 선행연구

사이버안보를 주제로 한 선행연구들은 학술적으로 주로 사이버안보 정책[4], 사이버안보 관련 법[5] 관련 연구, 사이버안보 대응 체계 및 전략에 관한 연구[6] 등이 주를 이루었다. 빅데이터 분석기법인 텍스트마이닝을 활용한 사이버안보 관련 특징적 연구로는 김두환·박호정[7]이 텍스트마이닝 기법을 활용한 연구를 통해 육군이 안고 있는 군사보안과 관련된 정책 방향 연구의 방안을 제시하고자 하였으나 '군사보안'에 관한 제한된 주제와 범위의 연구를 진행하였다. 또 조원선[8]이 국가 사이버안보 담론과 안보화 이론에 관하여 관련 언론보도 분석을 통해 연구하고자 하였다. 그러나 주요 기사의 내용을 해석하고 포털 등의 관련 검색 빈도 정도만 제시하고 텍스트마이닝 분석이 심도 있게 이루어지지는 않았다. 언론보도도

동향 분석을 통한 사회적 논의 및 이슈를 도출하고자 하는 연구는 다양한 분야에 적용되었다. 박해선[9]은 1998년부터 2022년까지의 아동학대를 다룬 언론보도를 수집하여 토픽모델링을 통해 시기별 토픽을 도출하고 그 결과를 토대로 아동보호와 아동학대 예방을 위한 대안을 논의하였다. 민기연·주란[10]은 2010년부터 2021년도까지 보도된 자폐성장애 관련 언론보도를 토픽모델링기법을 통해 분석하여 자폐성장애에 대한 언론과 사회 인식 변화를 확인하고 시사점에 관해 논하였다.

선행연구를 통해 살펴본 것과 같이 사이버보안과 관련된 연구는 2010년대 이후로 활발하게 이루어지고 있지만 대부분 사이버안보 정책, 관련 법제, 대응 체계 및 전략에 초점을 맞춘 연구가 대부분이었고 언론보도 등에 나타난 '사이버안보'의 특성을 빅데이터를 활용한 텍스트마이닝 기법을 활용하여 텍스트 자체에 관하여 심도 있게 분석 접근한 연구는 다소 미진하였다. 이에 본 연구는 기존 연구와 차별화된 텍스트마이닝 기법 등을 활용하여 언론보도를 분석하여 사이버안보에 관한 보도 동향을 파악하고 사회적 이슈와 토픽을 확인하여 시사점을 도출하는 연구를 진행하고자 한다. 이를 위해 지난 3년간의 '사이버안보' 관련 언론보도 전체를 수집하여 빅데이터 기반의 텍스트마이닝 기법과 사회연결망 분석 방법을 활용하여 실증적이고 객관적인 연구 방법을 통해 중점 키워드 및 핵심 토픽을 도출하고자 하였다.

## Ⅲ. 데이터 분석

### 3-1 데이터 수집

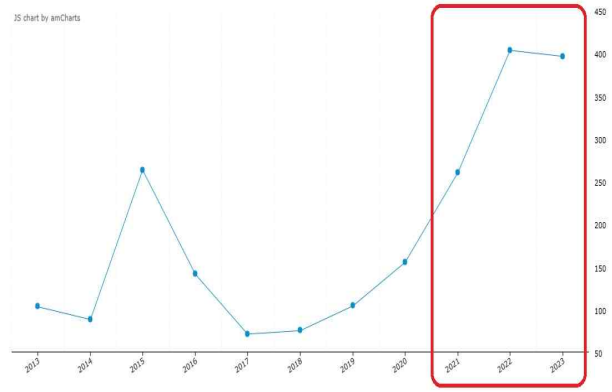



**그림 1.** '사이버안보' 언론보도 보도량 변화(최근 10년)
**Fig. 1.** 'Changes in the amount of media coverage of 'CyberSecurity' (last 10 years)

본 연구는 최근 3년에 해당하는 2020년 7월 1일부터 2023년 6월 30일까지를 분석 대상 기간으로 설정하였다. 한국언론진흥재단의 뉴스 분석 시스템인 '빅카인즈(BIG KINDS)' 시스템을 통해 54개 언론사의 관련 언론보도를 수집하고 정제 및 전처리 과정을 거쳐 부적절한 데이터를 정제하였다. 연구를 위해 수집한 언론보도는 총 1,211건이었으며,연구 분석에는 해당 기간 관련 보도 전체에서 제목과 키워드, 본문 전체의 텍스트를 분류하여 분석 데이터로 활용하여 형태소 분석기를 통해 363,340개의 명사를 추출하였다. 사이버안보를 키워드로 포함하는 보도량 변화는 그림 1과 같이 나타났다. 이를 통해 언론보도 보도량이 최근 3년 증가하였음을 확인할 수 있다.

### 3-2 분석 절차 및 방법

본 연구는 최근 사이버안보에 관한 사회적 논의와 이슈를 분석하기 위해 빅카인즈 시스템을 활용하여 국내 언론보도를 수집하였다. 언론보도를 통해 관련 핵심 논의와 토픽을 살펴보기 위해 수집 데이터에서 정제 후 추출한 명사를 대상으로 텍스트마이닝 분석과 사회연결망 분석을 수행하였다. 텍스트마이닝 분석은 TF-IDF(Term Frequency-Inverse Document) 값을 기준으로 1-mode matrix 데이터를 생성하여 사회연결망 분석 통해 키워드에 나타난 의미와 상관관계를 파악하였다. 또 LDA(Latent Dirichlet allocation) 토픽모델링 기법을 사용하여 언론보도에 나타난 토픽을 추출하여 주요 토픽 주제와 이슈를 확인하고자 하였다. 사회연결망 분석은 UCINET 프로그램을 사용하여 연결망의 네트워크 구조적 속성을 확인하고, 중심성 분석, 의미연결망 도식, CONCOR (CONvergence of iteration CORrealtion) 군집 분석을 수행하였다.

TF-IDF는 자연어 처리에서 단어 중요도를 평가하는 수치이다. TF-DF는 단어가 나타난 빈도를 의미하며 문서에서 높은 빈도로 노출된 단어일수록 가중치가 높다고 생각할 수 있다. 하지만 언어 체계에 조사와 관사 등 출현 빈도가 높으나 의미가 크지 않은 단어들이 높은 빈도로 추출될 수 있다. 따라서 이를 보완하기 위해 IDF 즉 역 문서빈도 개념을 적용하여 특정 단어가 출현한 문서의 수에 역수를 취하여 단순 단어 빈도와 비교하여 유의미한 결과를 얻어 낼 수 있는 방식으로 텍스트마이닝 분석 방법으로 널리 사용되고 있다[11]. LDA 토픽모델링은 비정형화된 대량의 텍스트 데이터에서 유의미한 주제, 즉 토픽을 추출할 수 있는 확률적 모델 알고리즘으로 유사한 의미의 키워드를 연결하는 방법으로 토픽을 추론하고 분석하는 기법이다[12]. 언론보도 동향을 파악하고 핵심 토픽을 분류하는 방법으로 다양한 분야의 연구에 활용되고 있다.

본 연구는 사이버보안과 관련하여 수집된 언론보도 빅데이터에서 텍스트 키워드 TF-IDF 값 상위 50개의 키워드를 도출하여 사이버보안에 관한 핵심적 키워드를 확인하고 Matrix 분석을 통해 사회연결망을 활용하여 키워드 사이의 상관관계와 특징을 확인하고자 한다. 그리고 LDA 토픽모델링 기법을 활용하여 사이버보안 관련 핵심 논의와 주제(토픽)를 추출하는 등 관련 사회 전반의 인식을 언론보도 분석을 통해 확인하고 그에 따른 시사점을 도출하고자 한다.

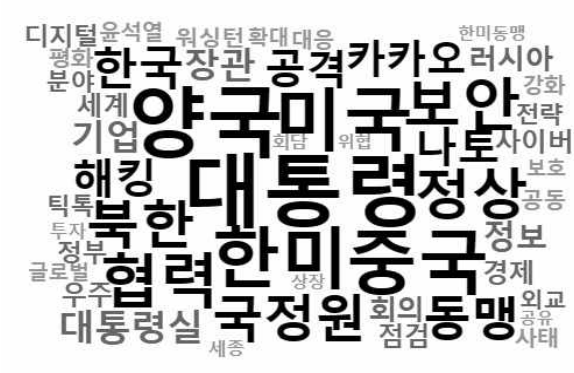
**표 1.** '사이버안보' TF-IDF, 빈도 분석
**Table 1.** 'CyberSecurity' TF-IDF, frequency analysis

| Ranking | Keyword | TF-IDF | Keyword | frequency |
|---|---|---|---|---|
| 1 | 대통령 | 3317.27 | 미국 | 4324 |
| 2 | 양국 | 3196.68 | 대통령 | 4215 |
| 3 | 중국 | 2915.89 | 협력 | 3754 |
| 4 | 미국 | 2842.12 | 국가 | 3344 |
| 5 | 한미 | 2802.45 | 대응 | 2501 |
| 6 | 협력 | 2408.57 | 양국 | 2272 |
| 7 | 보안 | 2399.63 | 중국 | 2271 |
| 8 | 정상 | 2299.13 | 정보 | 2221 |
| 9 | 북한 | 2217.15 | 강화 | 2212 |
| 10 | 국정원 | 2103.18 | 정부 | 2204 |
| 11 | 동맹 | 1983.81 | 북한 | 2154 |
| 12 | 한국 | 1923.57 | 보안 | 2100 |
| 13 | 나토 | 1796.41 | 한국 | 2078 |
| 14 | 해킹 | 1783.11 | 정상 | 1882 |
| 15 | 공격 | 1745.82 | 분야 | 1875 |
| 16 | 카카오 | 1712.02 | 한미 | 1821 |
| 17 | 대통령실 | 1581.41 | 기업 | 1794 |
| 18 | 장관 | 1562.07 | 위협 | 1663 |
| 19 | 기업 | 1551.21 | 공격 | 1562 |
| 20 | 정보 | 1459.84 | 해킹 | 1456 |
| 21 | 러시아 | 1330.80 | 회의 | 1452 |
| 22 | 회의 | 1307.74 | 사이버안보 | 1373 |
| 23 | 세계 | 1290.52 | 국정원 | 1330 |
| 24 | 경제 | 1288.62 | 동맹 | 1317 |
| 25 | 사이버 | 1278.95 | 전략 | 1218 |
| 26 | 점검 | 1269.59 | 대통령실 | 1183 |
| 27 | 우주 | 1267.46 | 경제 | 1172 |
| 28 | 정부 | 1253.11 | 세계 | 1059 |
| 29 | 디지털 | 1235.78 | 확대 | 1040 |
| 30 | 분야 | 1225.98 | 윤석열 | 980 |
| 31 | 틱톡 | 1221.37 | 상황 | 930 |
| 32 | 전략 | 1197.91 | 러시아 | 924 |
| 33 | 외교 | 1191.42 | 논의 | 917 |
| 34 | 워싱턴 | 1190.21 | 국제 | 905 |
| 35 | 공동 | 1185.85 | 글로벌 | 893 |
| 36 | 평화 | 1171.33 | 강조 | 890 |
| 37 | 사태 | 1166.66 | 외교 | 889 |
| 38 | 대응 | 1148.84 | 장관 | 887 |
| 39 | 강화 | 1139.16 | 공유 | 880 |
| 40 | 윤석열 | 1136.93 | 포함 | 867 |
| 41 | 확대 | 1129.67 | 공동 | 867 |
| 42 | 글로벌 | 1115.54 | 핵심 | 850 |
| 43 | 보호 | 1101.60 | 기관 | 845 |
| 44 | 투자 | 1069.77 | 정책 | 844 |
| 45 | 세종 | 1065.38 | 디지털 | 838 |
| 46 | 회담 | 1039.88 | 중요 | 836 |
| 47 | 위협 | 1035.01 | 개최 | 826 |
| 48 | 상장 | 1033.85 | 보호 | 787 |
| 49 | 공유 | 1033.50 | 국민 | 786 |
| 50 | 한미동맹 | 1024.57 | 점검 | 781 |

*Korean text data was analyzed, so English was not used.

### 3-3 데이터 분석 결과 (TF-IDF, 빈도 분석)

빅카인즈 시스템으로 수집한 언론보도 데이터로 텍스트마이닝을 수행하였으며 TF-IDF 값과 빈도 상위 50개의 키워드를 위 표 1과 같이 도출하였다. 도출된 키워드의 상관관계 및 네트워크상 구조적 특성 등을 파악하기 위해 UCINET을 활용하여 의미연결망과 CONCOR 분석 등 사회연결망 분석을 수행하였으며 LDA 토픽모델링 분석을 통해 핵심 이슈와 토픽을 도출하였다. TF-IDF 값 상위 50개의 키워드를 기반으로 도식한 워드클라우드는 아래 그림 2와 같이 도식되었다.



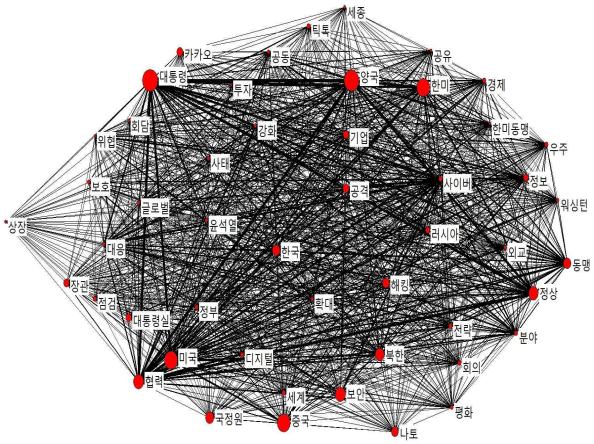

*Korean text data was analyzed,so English was not used.

**그림 2.** '사이버안보' 언론보도 워드클라우드
**Fig. 2.** 'CyberSecurity' media report WordCloud

TF-IDF 값은 '대통령' 키워드가 3,317.27로 가장 높게 나타났으며 '양국'이 3,196.68, '중국'이 2,915.89, '미국'이 2,842.12, '한미'가 2,802.45, '협력'이 2408.27, '보안'이 2,399.63, '정상'이 2299.13, '북한'이 2217.15, '국정원'이 2103.18의 값으로 나타났다. 빈도는 '미국'이 4,324회, '대통령'이 4,215회, '협력'이 3,754회, '국가'가 3,344회, '대응'이 2,501회, '양국'이 2,272회, '중국'이 2,271회, '정보'가 2,221회, '강화'가 2,212회, '정부'가 2,204회 노출된 것으로 나타났다.

## Ⅳ. 사회연결망 분석 및 LDA 토픽모델링

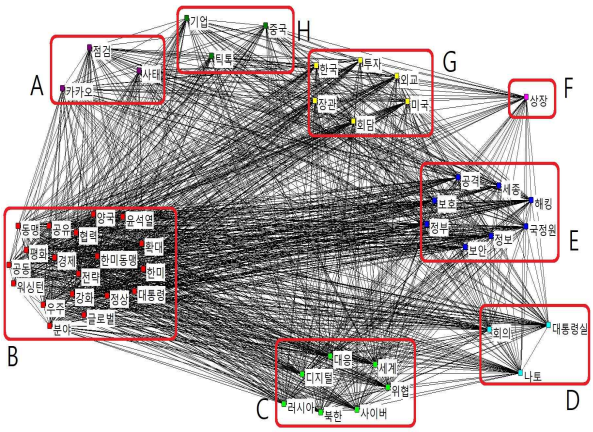
1-Mode Matrix를 기반으로 도식한 의미연결망은 아래 그림 3과 같이 도식되었다. 네트워크의 구조적 속성은 총 50개의 노드와 2,420개의 연결선으로 나타났으며 평균 연결 강도는 48.400, 밀도는 0.988, 평균 연결 거리는 1.012, 지름은 2임을 확인할 수 있었다. 노드 간 연결 강도는 '협력'과 '대통령'이 53,719로 가장 강한 것으로 나타났으며 '양국'과 '협력'이 52,536, '대통령'과 '양국'이 51,401, '미국'과 '대통령'이 42,356, '정상'과 '협력'이 41,464 '양국'과 '정상'이 40,339로 높은 연결 강도를 보였다.




*Korean text data was analyzed,so English was not used.

**그림 3.** '사이버안보' 언론보도 의미연결망
**Fig. 3.** 'CyberSecurity' media report semantic network

CONCOR 군집 분석을 통해 상위 50개의 키워드를 구조적 등위성 측정 의한 상관관계에 따라 유사한 속성을 가진 키워드 8개의 군집으로 분류하였으며 군집 A는 '카카오', '점검', '사태'의 키워드가 군집을 형성하였다. 군집 B는 '동맹', '평화', '공동', '워싱턴', '우주', '분야', '공유', '경제', '협력', '전략', '강화', '글로벌', '양국', '한미동맹', '정상', '윤석열', '확대', '한미', '대통령'로 가장 많은 19개의 키워드가 하나의 군집을 형성하는 것을 살펴볼 수 있었다.




*Korean text data was analyzed,so English was not used.

**그림 4.** '사이버안보' 언론보도 CONCOR
**Fig. 4.** 'CyberSecurity' media report CONCOR

군집 C는 '러시아', '디지털', '북한', '대응', '세계', '위협', '사이버'의 키워드가 군집을 이루었고 군집 D는 '회의', '나토', '대통령실', 군집 E는 '정부', '보호', '공격', '보안', '정보', '세종', '국정원', '해킹'의 키워드가 군집을 형성하였다. 군집 F는 '상장', 군집 G는 '한국', '장관', '투자', '회담', '외교', '미국', 군집 H는 '기업', '틱톡', '중국'의 키워드가 군집을 형성하였다.

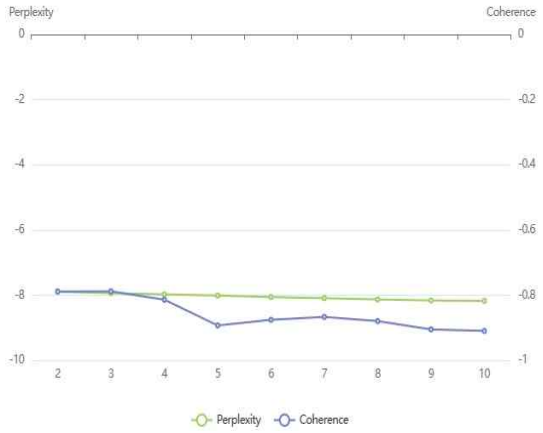
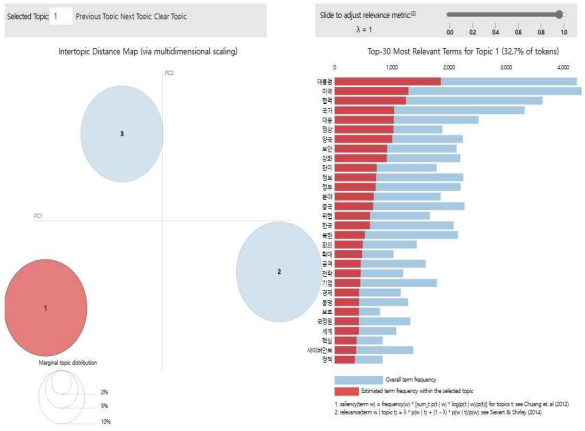



**그림 5.** '사이버안보' 언론보도 토픽 혼잡도 및 응집도
**Fig. 5.** 'CyberSecurity' media report topic perplexity and coherence

LDA 토픽모델링 분석의 토픽 수를 정하기 위한 혼잡도와 응집도 분석 결과는 위 그림 5와 같이 나타났다.

최적의 토픽 수 설정과 관련된 연구[13]에 근거하여 혼잡도와 응집도를 고려한 최적의 토픽 수를 3개로 정하여 토픽모델링 분석을 수행하였다. LDAvis를 활용하여 3개로 분류된 토픽의 2차원 위치를 확인하였다. 아래 그림 6을 통해 살펴볼 수 있는 것과 같이 분류된 3개의 토픽이 각각 다른 사사분면에 위치하였기 때문에 토픽의 분류 결과를 성공적으로 평가할 수 있다.




*Korean text data was analyzed,so English was not used.

**그림 6.** '사이버안보' 언론보도 LDAvis 시각화
**Fig. 6.** Visualization of 'CyberSecurity' media reports LDAvis

토픽 1은 '대통령', '미국', '협력', '국가', '대응' 등의 핵심어를 포함하고 있으며 토픽 2는 '해킹', '정보', '북한', '중국', '한국', '정부'의 핵심어를 포함하였다. 토픽 3은 '강화', '분야', '정보', '기업', '동맹'의 핵심어를 포함하고 있다.

## V. 논의

본 연구는 언론보도 빅데이터를 기반으로 텍스트마이닝과 사회연결망 분석기법을 활용하여 사이버안보와 관련된 사회 전반의 논의와 토픽을 언론보도를 통해 확인하고 이를 통한 시사점을 도출하고자 하였다. 언론보도 보도량이 급증한 2020년부터 최근 3년간의 사이버안보 관련 언론보도를 통해 관련 사회적 논의가 어떠한 핵심적 키워드를 중심으로 이루어져 있는지 살펴보고자 TF-IDF 분석과 의미연결망 분석을 수행하였으며 CONCOR 군집 분석과 LDA 토픽모델링 분석을 수행하였다. 이러한 연구를 통해 중점적으로 다루어진 핵심 키워드와 상관관계에 의한 의미연결망 군집을 분석하고 핵심 토픽을 추출하였다. 연구 결과에 의하면 '대통령', '양국', '중국', '미국', '한미', '협력', '보안', '정상', '북한', '국정원'이 TF-IDF 값 상위 순위의 핵심 키워드로 나타났다. 이러한 결과는 사이버안보와 관련하여 중국, 미국, 북한과 관련이 매우 크며 특히 한미간의 협력, 국가 정상 간의 협력 등이 매우 중요하다는 점을 유추해 볼 수 있는 결과였다. 특히 중국이나 북한과 같은 키워드는 사이버안보를 위협하는 대상으로서 해커의 활동이 많은 중국과 주요 사이버안보 위협에 관한 이슈를 일으켰던 북한의 해킹 공격 등을 반영하는 결과라고 볼 수 있다.

의미연결망 분석 결과에서도 연결 강도 상위 키워드로 '협력'과 '대통령', '양국'과 '협력', '대통령'과 '양국', '미국'과 '대통령', '정상'과 '협력', '양국'과 '정상'이 높은 강도를 나타낸 것을 통해 사이버안보 확립에 있어 미국과의 협력에 관한 보도가 중점적으로 이루어졌음을 확인할 수 있었다.

CONCOR 군집 분석 결과를 통해서는 '카카오', '점검', '사태'의 키워드가 군집 A를 이루며 카카오톡 서비스 장애로 인한 혼란이 언론보도에서 비중 있게 다루어졌음을 유추해볼 수 있었다. 또 한미동맹과 협력에 관한 다수의 키워드가 군집을 형성한 군집 B를 확인할 수 있었고 '러시아', '북한'이 포함된 세계의 위협에 관한 키워드가 군집 C를 이루었다. 사이버안보의 주체가 되는 '정부', '국정원', '보호' 등의 키워드를 포함한 군집 E를 확인할 수 있었다. 이러한 군집 분석 결과를 통해 카카오 서비스 장애로 인한 점검과 사태가 사이버안보와 관련된 하나의 이슈임을 확인할 수 있었고 한미간의 사이버안보 강화를 위한 동맹, 협력에 관한 부분과 북한과 러시아가 위협 관련 하나의 군집을 이루는 것도 살펴볼 수 있다. '중국'의 키워드는 '틱톡'과 '기업' 키워드와 함께 하나의 군집을 형성하였는데 이는 중국산 앱에 대한 미국의 대대적 퇴출과 미국 정부 행정기관의 틱톡 금지 조치가 전 세계로 확대되고 있는 상황에 관한 내용이 반영된 군집 분류 결과라고 볼 수 있다.

LDA 토픽모델링 분석 결과 3개의 토픽 중 첫 번째 토픽은 사이버안보 위협에 대한 대응을 위해 미국과의 협력의 중요성과 한미 사이버 안보협력에 관한 내용을 보여주는 토픽이

추출되었다. 두 번째 토픽은 북한과 중국의 해커들에 의한 사이버안보 위협과 정부의 대응이 하나의 핵심 토픽으로 도출되어 사이버안보 측면에서의 위협의 대상을 북한 또는 중국으로 인식하고 있음을 나타내는 결과이다. 세 번째 토픽은 사이버안보 위협에 대응하기 위한 기업 차원의 대응 등 민관 협력 사이버안보 강화 체계 구축에 관한 이슈를 보여주는 결과라고 할 수 있다.

지금까지의 연구 결과를 종합하여 보았을 때 최근 3년간 언론보도에서는 사이버안보에 관하여 미국과의 협력과 동맹관계 측면의 기사를 중점적으로 다루었다는 점을 확인할 수 있었다. 이는 최근 2023년 관련 언론 보도량의 급증과도 관련이 있다. 2023년 4월 미국 워싱턴에서 열린 한미 정상회담에서 양국의 동맹 범위를 사이버공간으로까지 확장하기로 한 선언을 하였기 때문이다. 사이버공격을 받아 사이버안보에 위협을 받을 때 공동 대응하는 상호 방위 조약을 사이버안보에 적용할 논의가 시작되었다고 볼 수 있으며 이는 다양한 사이버 적대 세력뿐만 아니라 특히 북한과 대치하고 있는 상황에서 사이버안보 측면의 큰 변화라고 할 수 있다. 또한 사이버안보의 중요성과 위상이 국가 정책 및 전략적 우선순위로 설정할 만큼 중요한 부분으로 인식되고 있음을 보여주는 결과이다. 사이버안보 관련 기관을 살펴보면 '국정원', '대통령실', '정부'가 상위 순위의 키워드로 도출되었다는 것은 사이버안보와 관련된 대응 체계의 변화를 반영해 주는 결과이다. 국정원이 사이버안보 민·관·군 협력 대응을 위해 만든 '사이버안보협력센터'가 언론의 관심을 받았다는 점을 유추해 볼 수 있다. 또 '중국', '미국', '북한', '나토' '러시아', '글로벌'과 같은 키워드는 TF-IDF 값 상위 50개 이내로 도출되어 사이버안보에 관한 언론보도에서 중점적으로 다루어지는 협력 대상 국가 또는 위협의 대상을 확인할 수 있었다. 그리고 카카오 서비스 마비 사태에 관한 특징적 군집의 형성이 도출된 것을 통해 최근 3년간 사이버안보 문제와 관련하여 발생한 문제가 사회에 미치는 영향을 유추해 볼 수 있다. 이는 우리 사회의 삶 전반에 사이버 영역이 미치는 영향력이 매우 크다는 점을 유추할 수 있는 연구 결과이다. 사이버상의 문제가 발생하였을 때 전 사회적으로 미치는 충격과 영향, 불편이 매우 크게 나타난다는 점이 언론보도를 통해 투영된 것이라고 볼 수 있다.

## Ⅵ. 결 론

지금까지의 연구 결과를 통해 사이버안보와 관련된 시사점을 다음과 같이 제시하고자 한다. 첫째, 한미동맹의 사이버공간 확장에 부합하는 사이버안보 역량과 체계 구축을 위해 노력해야 한다. 이를 위해서는 가장 먼저 사이버안보 전문가를 양성하기 위한 인적 투자가 적극적으로 이루어져야 할 것이다. 현 정부가 추진하고 있는 사이버보안 인재 10만 양성 계획이 결실을 거두기 위해서는 인력 양성을 위한 체계적인 시스템 마련과 투자가 뒷받침되어야 하며 단순한 인력의 숫자 충족이 아닌 사이버보안 인재 수준 고도화와 질적 향상을 위한 노력이 필요하다. 또 한미 상호방위조약의 사이버공간 확대가 실질적으로 적용될 수 있는 발동 조건에 대해 지속 고민하여 정상 간의 약속이 실제 실천으로 이어질 수 있는 준비 과정이 필요할 것이다.

둘째, 사이버안보 분야의 민관 협력 강화가 필요하다. 연구 결과 상위 순위로 도출된 기관 측면의 핵심 키워드는 '국정원', '대통령실', '기업'으로 나타났다. 이는 언론보도에서 높은 비중을 차지하는 각 키워드에 해당하는 기능이 증대되는 사이버 위기에 효과적으로 대응하기 위해 범정부 차원에서 민관이 협력체계를 구축해 나가야 할 필요성을 보여주는 결과라고 할 수 있다. 또한 최근 개소한 국정원의 사이버안보협력센터를 중심으로 다양한 과학기술정보통신부, 국방부 등 다양한 유관기관, 민간 보안기업이 긴밀히 협력하여 사이버위협 대응 체계를 공고히 해 나가야 할 것이다.

셋째, 사이버안보의 중요성에 대한 국민적 공감대 형성을 통해 사이버안보법을 제정하여 사이버위협에 대한 효율적이고 체계적 대응을 위한 법체계 마련과 전문성을 갖춘 조직 구축이 요구된다. 현행 사이버안보 위협 대응은 국가적 대응 체계 구축이 미비한 상황으로 각 부처의 역할이 분리되어 있어 위기 상황에 일원화된 종합적 대응에 한계를 가지고 있다. 그러나 지난 2006년 법 제도 마련의 논의가 처음 시작되었으나, 무려 17여 년이 지난 지금까지도 관련 법안이 제정되지 못하고 있다. 최근 사이버안보법 제정의 필요성에 대한 공감대 및 여론의 조성은 과거에 비해 높은 수준이라고 할 수 있으나, 언론보도 등을 통해 사이버안보와 관련된 적극적 정책 홍보 활동을 강화해 나간다면 사회적 인식 측면에서의 조속한 문제 해결에 도움이 될 것이다.

위와 같은 제언을 중심으로 사이버안보에 관한 발전적 방향성을 지향한다면 고도화되는 각종 국가 사이버안보 위협 상황에서 국가안보와 국익을 지키기 위한 효과적 대응에 도움이 될 것이라고 예상한다.

본 연구는 최근 3년간의 언론보도 데이터를 대상으로 실증적인 분석을 수행하여 사이버안보에 관한 최신 논의와 토픽 등을 도출하였다. 또 그에 따른 시사점을 도출하고 관련 정책에 관한 발전적 제언을 제시하였다는 점에서 학술적 의의가 있을 것이다. 그러나 제한된 분석 기간 설정으로 언론보도 동향의 변화나 추이 변화까지는 살펴보지 못하였다는 데 연구의 한계점을 가진다. 향후 사이버보안에 관한 다양한 연구 및 빅데이터 분석기법을 활용한 사회적 인식, 언론보도 동향 변화 등에 관한 후속 연구가 추가로 이루어지기를 기대하며 본 연구가 사이버안보 관련 정책 수립과 대응 체계 발전을 위한 기초자료로 활용되기를 기대한다.

## 참고문헌


[1] D. H. Kim and B. H. Kim, “A Study on the Social Awareness Using Big Data Analysis Regarding Cyber Terror,” *Korean Journal of Convergence Science (KJCS)*, Vol. 11, No. 2, pp. 147-158, February 2022. http://dx.doi.org/10.24826/KSCS.11.2.9

[2] Y. S. Lee and K. D. Jung, “Russian vs Ukraine Cyber War Lessons and Implications,” *The Quarterly Journal of Defense Policy Studies,* Vol. 38, No. 3, pp. 37-39, October 2023. http://dx.doi.org/10.22883/jdps.2022.38.3.002

[3] Y. H. Cho and H. H. Kim, “A Study on The Legal System for Professional Training to Strengthen Cybersecurity - Focusing on The Enactment of The「CyberSecurity Act」-,” *Kookmin Law Review*, Vol. 36, No 1, pp. 195-229, June 2023. http://dx.doi.org/10.17251/legal.2023.36.1.195

[4] B. W. Kim, “Cybersecurity Policy for Hyper-connected Industrial Society,” *Hannam Journal of Law & Technology*, Vol. 22, No. 3, pp. 85-122, October 2016. http://dx.doi.org/10.32430/ilst.2016.22.3.85

[5] J. K. Kim, “Coping with Legal Issues on Cyber-Security Threat,” *Kyungpook National University Law Journal (KNU Law Journal),* No. 58, pp. 145-177, May 2017. http://dx.doi.org/10.17248/knulaw..58.201705.145

[6] J. J. Park and S. H. Lee, “Korea’s Security·Strategic Capabilities to Counter Cyber Attack and Future Challenges,” *The Journal of Political Science & Communication,* Vol. 20, No. 3, pp. 79-114, October 2017. http://dx.doi.org/10.15617/psc.2017.10.31.3.79

[7] D. H. Kim and H. J. Park, “Military Security Policy Research Using Big Data and Text Mining,” *Journal of Convergence Security,* Vol. 19, No. 4, pp. 23-34, October 2019. http://dx.doi.org/10.33778/kcsa.2019.19.4.023

[8] O. S. Cho, “Cyber Security Discourses and Securitization Theory: On the Analysis of Korean Cyber Security Issues,” *The Quarterly Journal of Defense Policy Studies,* Vol. 33, No. 2, pp. 145-177, July 2017. http://dx.doi.org/10.22883/jdps.2017.33.2.006

[9] H. S. Park, “Comparison of Text Mining by Timeline for Media Coverage of Child Abuse ‑ Focusing on Topic Modeling -,” *Korean Journal of Family Social Work,* Vol 70, No 2, pp. 133-175, July 2023. http://dx.doi.org/10.16975/kjfsw.2023.70.2.133

[10] K. Y. Min and J. Ran, “Big Data Analysis on the Trend of Media Coverage for Autism Spectrum Disorders,” *The Journal of the Korean Association on Developmental Disabilities (KADD),* Vol. 27, No. 2, pp. 215-234, July 2023. http://dx.doi.org/10.34262/kadd.2023.27.2.12

[11] Y. E. Lee and J. H. Chang, “Child Abuse Analysis and Keyword Extraction through Unstructured Data Collection and TF-IDF,” *Korean Criminal Psychology Review (KCPR)*, Vol. 18, No. 4, pp. 171-182, December 2022. http://dx.doi.org/10.25277/KCPR.2022.18.4.171

[12] D. M. Blei, A. Y. Ng, and M. I. Jordan, “Latent Dirichlet Allocation,” *Journal of Machine Learning Research,* Vol. 3, pp. 993-1022, January 2003.

[13] D. Y. Lee and H. S. Yi, “Exploring Methods for Determining the Appropriate Number of Topics in LDA: Focusing on Perplexity and Harmonic Mean Method,” *Journal of Educational Evaluation*, Vol. 34, No 1, pp. 1-30, March 2021. http://dx.doi.org/10.31158/JEEV.2021.34.1.1



**김동훈(Dong-Hun Kim)**

2009년 : 홍익대학교 경영학과 (경영학사)

2015년 : 아주대학교 경영대학원 (경영학석사)

2023년 : 조선대학교 군사학과 (군사학 박사수료- 안보정책 전공)

2009년～2017년: 국방부 육군 예비역 소령

2018년～현 재: 과학기술정보통신부 국립전파연구원 주무관

※관심분야 : 국방정책, 빅데이터 분석, 사회연결망 분석, 사회적 인식 등

**김법헌(Beob-Heon Kim)**

1984년 : 조선대학교 무역학과 (경영학사)

1994년 : 전남대학교 행정학과 (행정학석사)

2017년 : 한남대학교 행정학과 (행정학박사-정책학 전공)

1984년～2016년: 국방부 육군 예비역 준장

2016년～현 재: 조선대학교 군사학과 교수

※관심분야 : 정책학, 사회과학 분야 교육, 군사통합, 전쟁론, 북핵 등