

텍스트 마이닝을 활용한 ‘경제안보’ 담론 분석

오정주¹, 임예람², 천혜수², 박원형³

¹성신여자대학교 연구산학협력단 연구원

²성신여자대학교 융합보안공학과 석사과정, ³교수

{winteroot, 220246056, 220246058, whpark}@sungshin.ac.kr

‘Economic Security’ Discourse Analysis Using Text Mining

Jungjoo Oh¹, Yeram Lim², Hyesu Cheon², Wonhyung Park³

¹R&DB Foundation, Sungshin Women’s University

^{2,3}Dept. Convergence Security Engineering, Sungshin Women’s University

요 약

미·중 기술 패권 경쟁이 심화되면서 경제안보는 국가안보의 핵심 요소로 부상하였다. 주요국들은 각국이 도입한 경제안보 개념에 따라 입법과 정책을 추진하고 있다. 그러나 우리나라에서 경제안보 개념은 아직까지 불분명한 상황이다. 이에 본 연구는 국내 뉴스 빅데이터를 통해 경제안보 관련 담론을 파악하여 한국식 경제안보 개념화를 위한 토대를 만드는 것을 목적으로 하였다. 빅카인즈를 통해 경제안보 관련 뉴스 기사를 수집하고 텍스트 마이닝을 활용하여 분석하였다. TF-IDF 분석과 LDA 토픽 모델링이 분석에 활용되었다. 그 결과 세 개의 주요 토픽이 도출되었고, 경제안보의 이중 구조를 확인할 수 있었다. 본 연구는 향후 한국식 경제안보를 개념화하고 그에 대한 전략을 마련하기 위한 기초자료로 활용할 수 있을 것으로 기대한다.

1. 서론

2020년대 미·중 기술 패권 경쟁이 본격화되면서 ‘경제안보’는 핵심적인 국가안보의 구성요소로 부상하였다[1]. 각 국가들은 국가 목표에 따라 경제안보에 대한 개념화하고, 그에 따른 입법과 정책을 추진하고 있다. 미국은 기술격차 유지를 통한 압도적 우위 유지, EU는 전략적 자율성과 위협 회피, 일본은 ‘전략적 자율성’과 ‘전략적 불가결성’ 확보를 통한 경제성장이라는 국가목표를 중심으로 경제안보를 개념화하고 있다[2][3].

우리나라 정부는 국가안보 전략 기초 중 하나로 경제안보를 제시하며 글로벌 경제안보 대응체제를 확립하겠다고 밝혔다. 구체적으로는 전략적 파트너들과의 협력 네트워크 구축, 핵심 공급망 위기 대응 능력 확보, 핵심·신흥기술 보호 및 협력 강화 등이 있다[4]. 그러나 아직까지 경제안보의 개념은 불분명하여 우리나라의 전략적 목표에 맞게 경제안보 개념을 정립하는 것이 필요하다는 의견이 제시되고 있다[2][3].

우리나라의 실정에 맞게 경제안보 개념을 정립하기 위해서는 경제안보의 동향을 파악하고 관련된 논

의를 탐색할 필요가 있다. 이에, 본 연구는 국내 뉴스 기사를 바탕으로 경제안보와 관련하여 형성된 담론을 살펴보고자 한다. 이를 위하여 경제안보와 관련된 뉴스 기사를 수집 후 핵심 키워드를 추출하여 TF-IDF 분석과 LDA 토픽 모델링을 시행한다. 본 연구는 경제안보와 관련하여 형성된 담론을 살펴봄으로써, 향후 한국식 경제안보의 개념화를 위한 기초연구로 활용될 수 있을 것으로 보인다.

2. 이론적 배경

최근 미·중 패권경쟁의 심화, 코로나19, 러시아-우크라이나 전쟁 등 국제정세가 크게 변화함에 따라, 경제안보의 개념은 크게 확대되었다[5]. 이러한 변화에 따라 ‘경제안보’에 대한 개념화가 시도되었지만, 아직까지 명확히 정리되지 않는 상황이다[1][2].

최윤정(2024)은 주요국의 경제안보 논의를 바탕으로, 경제안보 개념을 4가지 범주로 구분하였다. 첨단 기술과 같이 안보적 중요성을 갖는 경제에 대한 안보 관점의 관리, 국가안보 달성을 위한 경제적 통치, 강력한 경제력 기반의 안보 역량 확보, 안보력 기반의 경제력 확보 등이다[3]. 곽성일(2023)의 연구

에서는 물리적 자원을 최대한 확보하고 자국의 이익을 극대화될 때 경제안보가 달성될 수 있다고 정의하며, 주요국의 경제안보 조치를 바탕으로 ‘양면성’과 ‘상호연동성’이라는 특성을 도출하였다[6].

백우열(2022)과 안병수(2023)에 의하면, 경제안보는 이중적인 구조로 영역이 확대되고 있다. 상층구조에서는 첨단기술 경쟁 및 기술안보를 중심으로, 하층구조에서는 기존의 경제안보 문헌에서 주로 다루은 영역(자원과 물자의 안정적인 공급)이 중심이 된다[1][5]. 즉, 경제안보는 ‘첨단기술의 개발과 보호’ 그리고 ‘공급망 보안’의 두 축을 중심으로 논의가 확장되고 있다. 이처럼 확장된 경제안보 이슈에 대응하기 위하여 주요국은 대내적으로는 강력한 산업정책을 통해 첨단기술을 전략적으로 관리하는 한편, 대외적으로는 전략적인 동맹 관계를 맺음으로써 국제 안보협력체계를 공고히 하고 있다[1][2].

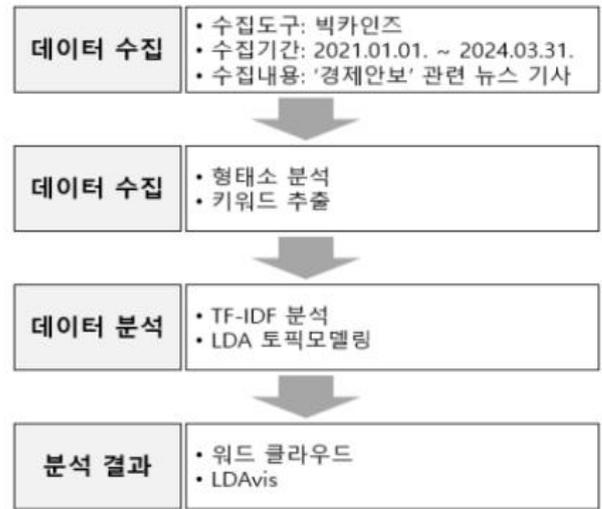
우리나라의 경우, 2023년 「국가첨단전략산업 경쟁력 강화 및 보호에 관한 특별조치법」과 「국가전략기술 육성에 관한 특별법」, 「산업기술의 유출방지 및 보호에 관한 법률」 등의 제정을 통해 경제안보상 중요한 기술들을 육성 및 보호하고 있다. 한편 「소재·부품·장비산업 경쟁력 강화 및 공급망 안정화를 위한 특별조치법」과 「경제안보를 위한 공급망 안정화 지원 기본법」, 「국가자원안보 특별법」 등 일명 ‘공급망 3법’을 제정하여 공급망 보안 이슈에 대응하고 있다. 대외적으로는 인도-태평양 경제 프레임워크(Indo-Pacific Economic Framework; IPEF)와 역내포괄적경제동반자협정(Regional Comprehensive Economic Partnership; RCEP) 등의 경제 동맹에 참여함으로써, 미·중 기술 패권 경쟁으로 인하여 우리나라에 발생하는 리스크를 완화하고자 노력하고 있다.

3. 연구방법

3.1. 연구절차

본 연구는 국내 뉴스 기사를 통해 경제안보와 관련된 담론을 분석하는 것을 목표로 한다. 이를 위하여 한국언론진흥재단에서 제공하는 뉴스 빅데이터 서비스인 빅카인즈(BIGKinds)를 통해 ‘경제안보’를 키워드로 하는 뉴스 기사를 수집하였다[7]. 수집된 기사의 키워드들은 텍스트마이닝 기법을 적용하여 분석하였다. 먼저, TF-IDF 분석을 통해 핵심 키워드를 도출하였고, 워드클라우드를 통해 도출 결과를 시각화하였다. 그 다음 LDA 토픽 모델링을 통하여

경제안보와 관련된 주요 토픽들을 도출한 뒤 (그림 1)과 같이 LDAvis로 시각화하였다.



(그림 1) 연구절차

3.2. 데이터 수집과 분석

본 연구에서 사용되는 데이터의 수집에는 빅카인즈 플랫폼이 이용되었다. 빅카인즈를 통해 2021년 1월 1일부터 2024년 3월 31일까지의 국내 뉴스 기사를 수집하였다. 중복된 기사를 제외하고 총 12,935건의 기사가 수집되었다. 빅카인즈를 통해 추출된 뉴스 기사의 키워드는 정제과정을 거친 뒤 텍스트롬(TEXTOM)을 통해 분석을 진행하였다. 텍스트롬은 대용량 데이터 수집 및 저장, 처리 및 정제, 분석 등의 서비스를 제공하는 플랫폼이다[8].

본 연구에서는 텍스트롬을 통해 TF-IDF 분석과 LDA 토픽모델링을 수행하였다. LDA 토픽모델링(Latent Dirichlet Allocation)은 텍스트 내에서 단어와 주제의 결합 확률 분포를 계산하는 비지도학습 확률적 모델링 방법으로, 주제 내에서 가장 빈번하게 사용되는 단어가 주제의 내용을 대표한다고 가정한다[9]. LDA 토픽모델링은 빅데이터를 통해 대규모 사회현상을 연구할 수 있도록 해준다는 장점을 가지고 있다[10].

4. 분석결과

4.1. ‘경제안보’ 관련 핵심 키워드

TF-IDF 분석은 빈도에 중요도에 따른 가중치를 적용한 기법으로, 단어의 중요도를 판단하기 어렵다는 단어빈도 분석의 한계를 보완한다[11]. TF-IDF 분석 결과 ‘중국’, ‘대통령’, ‘반도체’, ‘정상’, ‘한미’ 등의 키워드가 가장 높은 출현 빈도를 차지하였다.

TF-IDF 분석을 통해 도출된 ‘경제안보’ 관련 핵심 키워드 상위 50개는 <표 1>과 같다.

<표 1> TF-IDF 핵심 키워드 추출 결과

No.	단어	TF-IDF	No.	단어	TF-IDF
1	중국	27560.709	26	회장	12695.236
2	대통령	25546.511	27	공급망	12689.796
3	반도체	23709.222	28	정부	12490.421
4	정상	21121.877	29	기시다	12172.447
5	한미	20795.299	30	지원	12128.551
6	일본	20369.092	31	전략	12115.568
7	미국	20241.075	32	대통령실	11427.521
8	양국	18759.682	33	정책	11277.999
9	협력	17639.015	34	방문	11119.646
10	총리	17383.726	35	정상회담	11083.633
11	기업	16713.223	36	IPEF	10598.624
12	북한	16584.590	37	국가	10551.200
13	장관	15944.838	38	강화	10503.320
14	외교	15493.132	39	글로벌	10479.536
15	한국	15468.344	40	세계	10426.530
16	산업	14310.101	41	분야	10317.252
17	동맹	14165.373	42	우크라이나	10286.429
18	회담	13702.403	43	생산	10211.184
19	투자	13697.348	44	시장	10163.776
20	한일	13446.911	45	대응	10052.970
21	수출	13388.557	46	러시아	10040.447
22	바이든	13280.094	47	EU	9786.244
23	후보	12935.391	48	공동	9726.580
24	국민	12784.844	49	공장	9723.665
25	회의	12780.270	50	논의	9709.311

이상 TF-IDF 값이 큰 핵심 키워드 상위 50개를 워드 클라우드로 시각화한 결과는 (그림 2)와 같다.



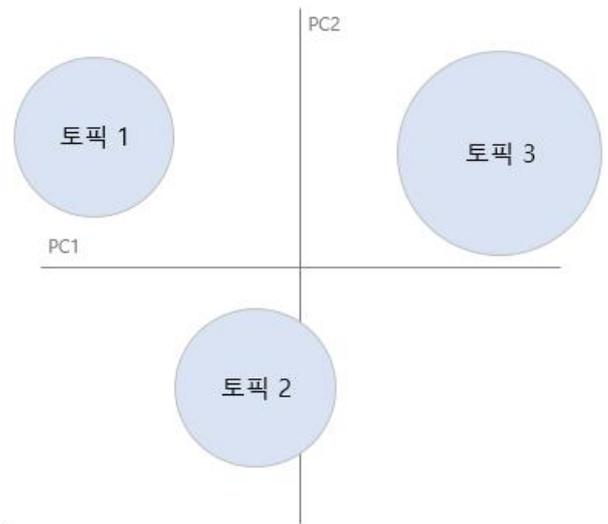
(그림 2) 워드 클라우드 시각화 결과

4.2. LDA 토픽 모델링

LDA 토픽 모델링을 진행한 결과, 혼잡도와 일관성 점수를 고려하여 토픽 수는 3개가 가장 적합할 것으로 판단하였다. 토픽 1의 주제는 ‘글로벌 반도체 산업 블록화’로 명명하였다. 토픽 1의 점유율은 30.5%이며, ‘중국’, ‘미국’, ‘반도체’, ‘경제’, ‘정부’ 등의 키워드가 도출되었다. 토픽 2의 주제는 ‘공급망 안정화’로, 29.4%의 점유율을 차지하였다. 핵심 키워드로는 ‘정부’, ‘경제’, ‘공급망’, ‘안보’, ‘요소수’ 등의 키워드들이 있다. 토픽 3은 ‘한·미·일 안보협력’으로 명명하였다. 가장 높은 점유율인 40.1%를 차지하며, 핵심 키워드로는 ‘대통령’, ‘미국’, ‘안보’, ‘정상’, ‘경제’ 등이 도출되었다. 위의 분석을 토픽모델링 시각화 도구인 LDAvis로 시각화한 결과는 (그림 3)과 같다.

<표 2> LDA 토픽모델링 결과

토픽	주제	주요 키워드 (상위 10개)
토픽 1	글로벌 반도체산업 블록화	중국, 미국, 반도체, 경제, 정부, 기업, 일본, 한국, 안보, 산업
토픽 2	공급망 안정화	정부, 경제, 공급망, 안보, 요소수, 장관, 대통령, 산업, 협력, 전략
토픽 3	한·미·일 안보협력	대통령, 미국, 안보, 정상, 경제, 협력, 한국, 한미, 일본, 외교



(그림 3) LDA 토픽 모델링 시각화 결과

5. 결론

본 연구에서는 ‘경제안보’ 키워드가 포함된 국내의 언론 기사를 수집한 뒤 텍스트 마이닝 기법을 활용하여 빅데이터 분석을 진행하였다. TF-IDF 분석 결과 국가와 관련 키워드(중국, 일본, 미국 등)와 정부 관련 키워드(대통령, 총리, 바이든 등)가 핵심 키워드로 도출되었다. 이는 경제안보와 관련된 논의가 주로 국가적 차원에서 진행되고 있으며, 정부 주도로 경제안보 정책이 추진되고 있음을 시사한다.

LDA 토픽 모델링을 진행한 결과, 세 개의 토픽(글로벌 반도체산업 블록화, 공급망 안정화, 한·미·일 안보협력)이 도출되었다. 반도체 산업의 블록화는 선행연구에서 언급된 경제안보의 상층구조와 관련된 것으로 해석된다. 반도체 기술은 기술 패권 시대의 핵심이 되는 첨단기술이다. 블록화를 통해 블록 내부의 국가들은 연구개발 협력을 함으로써 반도체 산업을 육성하지만, 블록 외부에 있는 국가에 대해서는 차별적으로 대우함으로써 육성한 반도체 산업을 보호할 수 있다. 이에 토픽 1은 경제안보의 상층구조와 관련된 것으로 사료된다.

2021년 중국에서 요소수의 수출을 제한하자 우리나라에서 요소수 대란이 발생했을 정도로 요소수는 산업계에서 필수적인 원료이다. 2021년 요소수 대란은 국민경제에 큰 파장을 일으켰으며, 이후 공급망 안정화는 경제안보 시대 중요한 핵심 키워드로 부상하였다. 우리나라에서는 제2의 요소수 사태를 막고자 공급망 3법 제정, 글로벌 경제 동맹 참여 등을 통해 대응하고 있다. 따라서 토픽 2는 경제안보의 하층구조와 관련된 것으로 보인다.

마지막으로 토픽 3은 위의 두 토픽에 대응하기 위한 국가들의 전략이라고 해석된다. 각 국가들은 첨단기술의 개발 및 보호를 통한 우위 확보 그리고 공급망 안정화를 위하여 안보협력 체계를 구축하는 것이다. 특히 우리나라는 미국, 일본과 긴밀한 협력체계를 구축하여 왔기에, 한·미·일 안보협력이라는 토픽이 도출된 것으로 판단된다.

본 연구는 국내 뉴스 기사를 통해 경제안보에 대한 논의를 살펴보았다. 주요국들은 국가 목표에 따라 경제안보를 개념화하고 관련 정책을 시행하고 있다. 경제안보를 위한 입법과 정책을 전략적으로 시행하기 위해서는 경제안보에 대한 명확한 개념 정의가 필요할 것이나, 우리나라에서는 아직 경제안보의 개념이 불분명한 상태이다. 본 연구는 빅데이터 분석을 통해 선행연구에서 언급된 경제안보의 이중구

조를 확인하였다. 이는 추후 한국식 경제안보의 개념화를 위한 하나의 접근방식이 될 수 있을 것으로 기대된다. 한편, 본 연구에서는 우리나라의 국가적 목표에 부합하는 경제안보의 개념화까지는 이르지 못하였다. 추후 연구에서는 경제안보의 이중구조에 따라 경제안보를 개념화하고, 우리나라의 경제안보 전략의 방향성을 모색할 필요가 있다.

6. Acknowledgement

본 논문은 2024년도 산업통상자원부 및 한국산업기술진흥원의 산업혁신인재성장지원사업 (RS-2024-00415520)과 과학기술정보통신부 및 정보통신기획평가원의 ICT혁신인재4.0 사업의 연구결과로 수행되었음 (No. IITP-2022-RS-2022-00156310)

참고문헌

- [1] 백우열, “경제안보 개념의 확장: 2020년대 안보 맥락에서”, 국제정치논총, Vol. 62, No. 4, pp. 325-364, 2022.
- [2] 차정미 외, “세계 경제안보의 실제와 한국 경제안보에의 제언 - 경제안보 전략과 경제안보 거버넌스-”, 국회 미래의제, No. 23-14, pp. 1-27, 2023.
- [3] 최윤정, ““경제안보”, 진영을 넘어 국익으로”, 세종정책브리프, No. 2023-20, pp. 1-34, 2024.
- [4] 국가안보실, “윤석열 정부의 국가안보전략”, 2023.
- [5] 안병수, “경제안보 관점의 FDI 심사제도 개선에 관한 연구”, 무역상무연구, Vol. 97, pp. 37-68, 2023.
- [6] 광성일, “한국의 경제안보 추진 방향”, 계간 외교, No. 145, pp. 54-65, 2023.
- [7] 빅카인즈, <https://www.bigkinds.or.kr/>
- [8] 텍스트툼, <https://www.textom.co.kr/home/sub/overview.php?pnm=1>
- [9] Asmussen, C. B., Møller, C., “Smart literature review: a practical topic modelling approach to exploratory literature review”. Journal of Big Data, Vol. 6, No. 93, 2019.
- [10] Mohr, J. W., Bogdanov, P., “Introduction—Topic models: What they are and why they matter”. Poetics, Vol. 41, No. 6, pp. 545-569, 2013.
- [11] 김수현 외. “거시경제 분석을 위한 텍스트 마이닝”, 한국경제의 분석, Vol. 26, No. 1, pp. 1-85, 2019.