

탄소중립 기술의 미래신호 탐색연구: 국내 뉴스 기사 텍스트데이터를 중심으로

정지송¹, 노승국^{2*}

¹경찰대학 치안대학원 범죄학과, ²경찰대학 치안대학원 데이터사이언스 전공 조교수

Detecting Weak Signals for Carbon Neutrality Technology using Text Mining of Web News

Jisong Jeong¹, Seungkook Roh^{2*}

¹Dept. of Criminology, Graduate School of Police Studies, Korean National Police University

²Assistant Professor. Data Science Major, Graduate School of Police Studies, Korean National Police University

요약 우리나라는 기후변화 위기에 대응하기 위해 2050 탄소중립을 선언하였으며, 이를 위해 다양한 감축 계획 및 입법화 과정을 진행 중이다. 탄소중립의 실현은 산업기술 전반에서의 근본적 변화를 필요로 하기 때문에 이를 위한 구체적 대응체계 마련이 매우 중요하다. 본고는 탄소중립 관련 산업기술 확보 경쟁에서 선제적으로 대비하기 위하여 글로벌 탄소중립 기술분야의 현황과 발전 트렌드를 파악하고자 한다. 이를 위해, 탄소중립 관련 온라인 뉴스기사 데이터를 웹 크롤링하여 수집하였고, 미래신호분석방법론과 인공지능망 딥러닝 기술인 Word2Vec 알고리즘을 적용하여 탄소중립 기술 트렌드를 분석 및 예측하였다. 분석결과, 탄소 과배출 업종인 철강업 및 석유화학 분야의 기술고도화가 요구되고 있었으며, 전기차 분야에의 투자 타당성 확보와 기술 고급화가 추세인 것으로 드러났다. 이에 대한 정부의 적극적인 지원과 글로벌한 기술협력/인프라 조성이 밀반침되어야 할 것으로 보인다. 그 외에도 탄소중립 관련 인력양성이 시급한 것으로 나타났으며, 기업에서 필요한 탄소중립 인력을 양성할 수 있도록 간접지원정책 마련의 필요성을 확인할 수 있었다.

키워드 : 산업기술, 탄소중립, 미래신호예측, 약신호, Word2Vec

Abstract Carbon neutrality is the concept of reducing greenhouse gases emitted by human activities and making actual emissions zero through removal of remaining gases. It is also called "Net-Zero" and "carbon zero". Korea has declared a "2050 Carbon Neutrality policy" to cope with the climate change crisis. Various carbon reduction legislative processes are underway. Since carbon neutrality requires changes in industrial technology, it is important to prepare a system for carbon zero. This paper aims to understand the status and trends of global carbon neutrality technology. Therefore, ROK's web platform "www.naver.com." was selected as the data collection scope. Korean online articles related to carbon neutrality were collected. Carbon neutrality technology trends were analyzed by future signal methodology and Word2Vec algorithm which is a neural network deep learning technology. As a result, technology advancement in the steel and petrochemical sectors, which are carbon over-release industries, was required. Investment feasibility in the electric vehicle sector and technology advancement were on the rise. It seems that the government's support for carbon neutrality and the creation of global technology infrastructure should be supported. In addition, it is urgent to cultivate human resources, and possible to confirm the need to prepare support policies for carbon neutrality.

Key Words : Industrial technology, Carbon neutrality, Future sign, Weak signal, Word2Vec

*Corresponding Author : Seungkook Roh(skroh@police.ac.kr)

Received February 7, 2023

Accepted May 20, 2023

Revised March 13, 2023

Published May 28, 2023

1. 서론

탄소중립은 미래 유망 산업기술 중 현재 가장 큰 이슈가 되고 있는 주제이다. 탄소중립이란 인간활동에 의해 배출되는 온실가스를 최대한 감축하고, 남은 온실가스는 흡수 및 제거를 통해 실질적 배출량을 0으로 만든다는 개념이다. 획기적으로 배출량을 줄이고 남은 탄소와 흡수되는 탄소량을 같게 하여, 순 탄소배출량이 0이 되게 하는 것으로, '넷 제로(Net-Zero)', '탄소제로'라고도 한다. 국제사회는 기후변화에 대응하기 위해 온실가스 배출 감축 방안을 지속적으로 논의해왔으며, 2020년 유럽연합을 선두로 2021년 미국, 중국 등 주요국들이 '탄소중립' 달성 목표 시점을 설정하고 법제화를 추진중에 있다. 이러한 상황을 보았을 때, 국가 단위의 탄소중립은 향후 기후변화 대응과 지구 평균온도 상승 억제를 위한 중요한 추세가 되었음을 알 수 있다.

국내에서도 탄소중립을 위한 다양한 시도가 이뤄지고 있다. 우리나라는 2020년 10월 제시한 탄소중립이라는 도전적인 목표하에 2021년 9월에는 국내 탄소중립 관련법이 제정되었다. 2050년 탄소중립 달성을 위해 '기후위기 대응을 위한 탄소중립/녹색성장기본법(탄소중립기본법)'이 제정되었고, 2022년 3월에는 탄소중립기본법이 본격 시행되었다. 2050년 탄소중립 목표는 기본법에도 명시되어 있는 만큼 그 이행 전략은 국가 수준의 중장기 계획이 될 것이며, 향후 경제·에너지·사회 등 많은 분야에서 크나큰 변화를 야기할 것이다. 반면, 시행된 탄소중립 정책을 살펴보면 구체적이고 타당한 이행방안이 담겨있지 않아 실효성에 의문이 제기된다는 비판이 다수 존재한다.

탄소중립 목표 달성은 기존 사회 전반에서의 대전환을 요구하기 때문에 이에 대한 선제적인 점검과 현 정책의 검토 병행이 필수적인 것이다. 그에 따라 현재까지 각 분야별 감축 목표와 방식에 대해서 다양한 논의가 이뤄져왔으나, 그 세부 사항에 대한 검토는 미비한 상황이다. 이러한 점으로 미루어보았을 때 2050년까지 탄소중립 달성을 위한 사회기술적 기반마련은 매우 중요한 과제로 보여진다. 효과적인 탄소중립 정책지원을 위해 주요 부문에서 점검해야 하는 현안과 수행해야 하는 이행방안 및 기술을 제시하고, 정부 측면에서 고려해야 할 방안에 대해 모색해야 할 것이다. 따라서, 본 연구는 국가 간 치열한 산업기술 확보 경쟁시장에서 선제적으로 대비하기 위하여 미

래예측방법론을 활용해 글로벌 탄소중립 산업 및 기술 분야의 핵심 트렌드를 체계적으로 파악하고자 한다.

2. 선행연구 분석

2.1 미래신호예측 분석방법론에 대한 이론적 논의

과학기술정책연구원(2022)에 의하면, 미래연구란 미래에 대비한 의사결정이나 전략실행을 목표로 하여, 다수의 참여자가 중장기적 시각에서 과거 및 현재의 경험과 자료를 기초로 구체적 대안을 수립하는 체계적인 과정이라 정의한다[1]. 거시적 측면으로는 사회, 경제 및 기술적 발전과 수요에 대해서 구조화된 전망을 제공하며, 미시적으로는 정책 결정자에게 있어 정보의 폭과 분석의 깊이를 확대시키는 역할을 수행한다. 현대와 같이 빠른 속도로 변화하는 시대에, 올바른 의사결정 및 선제적 대비를 위한 미래연구의 가치가 증가하고 있다.

미래예측 방법론은 가장 먼저 1907년 Column Gilfillan에 의해 본격적으로 연구되었으며, 1944년 Flechtheim이 '미래학(Futurology)'이라는 용어를 널리 사용하며 주목을 받았다. 이후 1960년대 Herman Kahn에 의해 미래예측 연구방법론이 다듬어졌으며, 이를 기반으로 미래학이 하나의 학문으로 확립되었다. Theodore J. Gordon과 제롬 글렌 등의 미래학자들은 유엔 대학(United Nations University)의 밀레니엄 프로젝트(The Millenium Project)에서 체계화된 미래연구방법론(Futures Research Methodology)을 발표하는 등 미래예측을 위한 방법론을 지속적으로 개발하고 있다[2-5].

밀레니엄 프로젝트에서 발표된 미래 연구방법론은 기준과 목적에 따라 그 방법이 다양하게 분류된다. 정성적 기법으로는 델파이, 환경 스캐닝, 델파이, 천재적 예측 등이 있으며 정량적 기법으로는 교차영향분석, 추세영향분석, 의사결정모델링 등이 있다. 미래연구방법들은 어떠한 목적으로 사용하는가에 따라라도 분류될 수 있다. 즉 판단을 모을 때, 시계열과 다른 양적 방법 예측하기, 사건과 유행 그리고 움직임 사이의 관계에 대한 이해, 불확실성 하의 행동을 정할 때, 그럴듯한 대안 미래 제시, 미래가 개선되고 있는지에 대한 이해관계 성립, 변화와 가정 추적, 시스템의 안정성 확인의 목적에 따라 적절한 방법들이 구분되어 사용된다[6, 7]. Glenn과 Aaltonen의 선행 연구를 토대로 속성과 목적성에 근거한 주요 미래연구방법론의 분류를 정리하면 Table 1과 같다[8, 9]. 본 연구는

Table 1. Purpose and attributes of future research methodology

| Purpose | Methodology | Attribute | | | |
|---|---------------------------------------|--------------|-------------|-----------|-------------|
| | | Quantitative | Qualitative | Normative | Exploratory |
| Collection of Judgments | Genius Forecasting | | ○ | ○ | ○ |
| | Delphi | | ○ | ○ | ○ |
| | Futures Wheel | | ○ | ○ | ○ |
| Non-Time Series Quantitative Forecasting | Econometrics and Statistical Modeling | ○ | | | ○ |
| | Trend impact analysis | ○ | | | ○ |
| | Structural Analysis | ○ | ○ | | ○ |
| | Regression Analysis | ○ | | | ○ |
| The Relationship between Event, Trend, and Movements | Agent modeling | | ○ | | ○ |
| | Trend impact analysis | ○ | | | ○ |
| | Cross-impact analysis | ○ | | | ○ |
| | Decision Tree | ○ | | | ○ |
| | Futures Wheel | | ○ | ○ | ○ |
| | Simulation Gaming | | ○ | | ○ |
| | Multiple Perspective | | ○ | ○ | ○ |
| | Causal Layered Analysis | | ○ | | ○ |
| | Field anomaly relaxation | | ○ | | ○ |
| Action Decisions Under Uncertainty | Decision Tree | ○ | | | ○ |
| | Road Mapping | ○ | ○ | ○ | ○ |
| | Technology sequence analysis | | ○ | ○ | |
| | Genius forecasting, | | ○ | ○ | ○ |
| Plausible Suggestion of future alternatives | Scenarios | ○ | ○ | ○ | ○ |
| | Futures Wheel | | ○ | ○ | ○ |
| | Simulation Gaming | | ○ | | ○ |
| | Agent Modling | | ○ | | ○ |
| Establishing a stake in whether the future is improving | State of the Future Index | ○ | ○ | ○ | ○ |
| Tracking Changes and Assumptions | Environmental Scanning | | ○ | | ○ |
| | Text mining | | ○ | ○ | ○ |
| System Stability | System Modeling | ○ | | | ○ |

탄소중립 관련 기술의 미래 변화를 예측하고자 하였기에, 미래연구 방법론 중 '변화와 가중추적' 목적의 환경 스캐닝, 텍스트마이닝 분석방법을 연구방법으로 고려하였다. 이 두 가지 방법론 중, 보다 넓은 범주의 방법론인 텍스트 마이닝을 최종 분석방법으로 결정하였다.

2.2 텍스트마이닝에 의한 미래신호 분석연구

빅데이터 분석방법은 분석하고자 하는 데이터의 형태에 따라 그 분석방법이 나누어진다. 고정된 필드에 수치로 저장된 데이터인 정형 데이터의 경우, 명목척도, 서열척도 등의 개념을 적용해 회귀분석과 같은 통계적 방법을 통해 분석이 수행되는 반면, 고정된 필드에 저장되어 있지 않은 동영상, 이미지, 텍스트데이터와 같은 비정형데이터의 경우에는 세부적으로 분석방법이 나뉜다.

영상이나 이미지 데이터의 경우 영상인식을 위한 지도 학습기술, 비지도 학습기술, 강화학습 등의 분석기술이 사용되게 된다[10, 11]. 음성 및 텍스트데이터의 경우는 확보된 텍스트를 기반으로 분석이 수행되게 된다. 텍스트

는 가장 기본적이고 광범위한 비중을 차지하는 비정형데이터이고, 이러한 텍스트를 컴퓨터 공학, 통계학, 사회학 등의 학문적 지식을 이용하여 특정 목적에 맞게 유의미한 정보를 추출하는 분석 및 처리 과정을 텍스트마이닝(Text mining)이라고 한다.

소셜 미디어의 확산으로 온라인상에 남긴 정치·경제·문화에 대한 의견들이 그 시대의 감성과 정서를 파악할 수 있는 원천으로 등장함에 따라, 많은 국가와 기업에서는 SNS를 통하여 생산되는 소셜 빅데이터를 분석 및 활용하여 사회적 문제를 해결하고 미래를 예측하기 위해서 노력한다[12]. 산업기술 분야의 미래예측 연구에서도 텍스트마이닝에 의한 미래신호 분석방법론이 다수 활용되고 있다. 그 중, Roh(2020)는 미래신호예측 방법론 적용을 통해 원자력 산업관련 UAE 원전수출과 탈원전 정책 시행이슈가 약신호에서 강신호로 변화했다는 것을 확인하였다[13]. 이를 통해 산업기술 분야에의 미래신호 분석 방법 적용이 타당함을 증명하였다.

3. 데이터 수집 및 분석방법

3.1 데이터 수집

본 연구는 비정형데이터 분석방법을 통해 국가 간 치열한 산업기술 확보 경쟁 시장에서 선제적으로 대비하기 위한 탄소중립 관련 산업 및 기술 분야 핵심 트렌드를 파악하고자 한다. 이를 위해 네이버 뉴스에서 '탄소중립' 키워드가 포함된 뉴스 기사의 제목, 언론사, url, 날짜, 기사 전문, 댓글, 대댓글 등의 정보를 모두 수집하였다. 데이터 수집은 파이썬(Python) 코드를 통해 이루어졌으며, '탄소중립' 키워드를 포함하는 기사 81,186건의 데이터를 대상으로 분석을 진행하였다. 데이터 수집 기간은 2019년 1월 1일부터 수집 당시인 2021년 12월 31일까지, 3년 동안 보도된 자료에 한하였다. 수집된 탄소중립 관련 데이터를 바탕으로 형태소 분석(Stemming), 불용어(Stopwords) 처리 등의 데이터 전처리과정을 거쳐 미래신호 분석, Word2Vec분석을 수행하였다(Fig. 1.). 데이터 분석은 2019년부터 2021년까지의 잠재신호, 약신호, 강신호들에 속한 단어들의 변화를 추적하였다. 이러한 분석과정을 바탕으로 본 연구는 탄소중립 분야의 산업기술 트렌드에 대한 방향을 제시하고 산업기술 경쟁력을 확보하기 위한 선제적 대비의 근거를 마련하고자 한다.

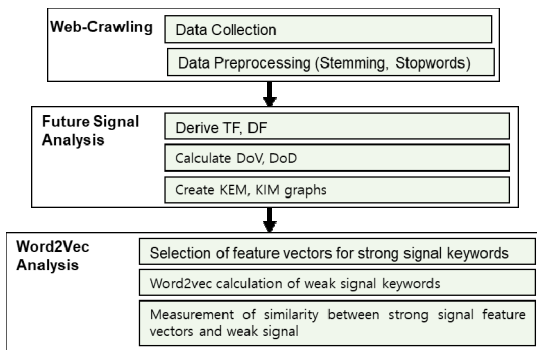


Fig. 1. Research procedure

3.2 데이터 분석방법

온라인 채널에서 생산되는 소셜 빅데이터는 정보로서의 높은 가치를 가지고 있지만, 미래를 탐지하고 예측하기 위해서는 이를 보다 효과적으로 수집 및 분석하기 위한 분석틀이 필요하다. 특히, 온라인 속에서 표현된 빅데이터는 '비정형데이터'로 이들 가운데 의미있는 키워드를 추출하고 자료를 효과적으로 수집하기 위해서는 산업 분

야의 메가 프로젝트 발굴을 위한 핵심 개념을 추출하고 해당 개념들 간의 관계를 나타내는 온톨로지가 필요하다. 온톨로지는 '해당 개념을 명시적으로 정의하며, 컴퓨터가 처리할 수 있는 형태로 표현하는 용어의 논리적인 집합이면서 개념 간 관계를 명시한 사전의 역할'을 의미한다. 즉, 온톨로지는 관심주제의 공유된 개념을 형식화하고 표현하기 위한, 컴퓨터가 해석 가능한 지식모델이다. 따라서 온톨로지가 있어야만 거대한 비정형 빅데이터를 분류하여 처리하고, 이를 기존 연구방법들을 통해 다양한 분석을 시도할 수 있다. 본 연구에서의 소셜 빅데이터 주제 분석(수집 및 분류)은 온톨로지¹⁾의 키워드를 수집하여 분류하는 Top-down 방법과 해당 토픽을 웹 크롤링 후 범용사전이나 사용자 사전으로 범주화시키는 Bottom-up 방법을 병행하여 사용하였다.

3.2.1 미래신호예측 분석방법론

과거 미래연구방법으로는 델파이, 시나리오 기법 등 전문가 견해를 집약하는 방법이 다수 이용되었으나, 방대한 비정형데이터 생성 및 급격한 기술발달로 현재는 빅데이터를 활용한 미래신호, 즉, 약신호 예측방법론이 주목 받고 있다[12, 14-16]. 미래신호 중 가장 핵심적인 약신호(Weak signal)는 임박한 충격적 사건에 대한 미약한 초기 징후라고 정의된다[17]. 약신호는 처음에는 매우 모호하고 감지하기 어렵다가 시간이 경과함에 따라 더 강력하고 명확해지는 신호이며, 향후에는 정치, 경제, 기술 및 사회적 사고를 동반한다. 즉, 약신호는 '미래에 가능한 변화의 징후'로, 시간이 흐르면서 강신호(Strong signal)로, 다시 강신호는 트렌드(Trend)나 메가트렌드(Mega Trend)로 발전이 가능하다[18, 19].

Yoon(2012)은 Hiltunen(2008)의 미래신호 모델을 기반으로 하여 텍스트마이닝 분석을 통한 미래신호 예측분

1) 탄소중립 산업분야의 메가 트렌드 발굴을 위한 탄소중립 산업기술 특성을 분석하고 미래신호를 예측하기 위한 분석틀로서 산업통상자원부 및 KIAT(한국산업기술진흥원)의 내부 전문가 회의를 통해 온톨로지를 개발하였다. 2021년 12월 21일에 산업기술 트렌드 분석을 위한 미래신호예측 분석 방법론과 온톨로지 구축에 관한 1차 회의가 진행되었다. 분석 방법론에 대한 타당성을 확보하고자 델파이 방법론을 통해 2022년 산업부의 핵심 정책 방향에 맞춰 기술 키워드를 도출하였다. 2021년 1월 4일부터 한달간은 '탄소중립' 키워드를 통해 실시한 미래신호예측 분석 파일럿 테스트 결과를 공유 및 향후 추진 방향에 대한 협의를 진행하였다. 결과의 정책화와 미래 메가트렌드에 적합한 전문적 시그널(키워드) 도출 유도를 위해 키워드별 서브 키워드를 다음과 같이, 탄소중립(규제&정책, 기술, 인재, 인프라, 정책, 지역), 으로 분류하였다.

석 방법을 고안하였다. 해당 분석방법은 특정 단어가 한 문서 내 얼마나 자주 등장하는지를 나타내는 단어빈도(TF, Term Frequency)를 미래신호의 가시성으로, 특정 단어가 나타나는 문서의 수인 문서빈도(DF, Document Frequency)를 미래신호가 얼마나 확산되고 있는지를 나타내는 사건의 수로 가정하였다[18, 20]. 이러한 가정을 바탕으로 시간에 따른 단어 및 문서빈도 증가율을 기준으로 미래신호를 분류하기 위해 가시성(DoV, Degree of Visibility)와 확산성(DoD, Degree of Diffusion) 값을 도출해야 한다.

가. 가시성(DoV, Degree of Visibility)

시간 가중치를 고려하여 시간 흐름에 따른 특정 단어의 가시성 변화정도를 파악하기 위한 값이다. 가시성을 도출하기 위한 식은 아래 식 (1)과 같으며, 도출된 DoV값을 바탕으로 X축에는 평균 단어빈도(average TF), y축에는 DoV 평균증가율로 설정하여 KEM(Keyword Emergence Map)을 도출할 수 있다.

$$DoV_{ij} = \left(\frac{TF_{ij}}{NN_j} \right) \times 1 - tw \times (n - j) \quad (1)$$

TF: 단어빈도 (Term Frequency)
 NN_j: 전체문서 수 (문장 수)
 tw: 시간가중치 (0.05)
 n: 전체 시간 구간
 j: 시점

나. 확산성(DoD, Degree of Diffusion)

시간 가중치를 고려하여 시간 흐름에 따른 특정 단어의 확산정도를 파악하기 위한 값이다. 확산성을 도출하기 위한 식은 아래 식 (2)과 같으며, 도출된 DoD값을 바탕으로 X축에는 평균 문서빈도(average DF), y축에는 DoD

평균 증가율로 설정하여 KIM(Keyword Issue Map)을 도출할 수 있다.

$$DoD_{ij} = \left(\frac{DF_{ij}}{NN_j} \right) \times 1 - tw \times (n - j) \quad (2)$$

DF: 문서빈도 (Document Frequency)
 NN_j: 전체 문서 수 (문장 수)
 tw: 시간가중치 (0.05)
 n: 전체 시간 구간
 j: 시점

본 연구는 Yoon(2012)과 Song(2016)이 적용한 시간 가중치 0.05를 적용하여 DoV와 DoD를 산출하였다. 산출된 DoV, DoD값을 바탕으로 키워드 포트폴리오인 KEM, KIM을 작성하였다[14-16, 20]. KEM은 x축과 y축으로 이뤄진 평면을 나타내는 것으로, x축에는 평균TF를 나타내고, y축에는 DoV의 평균증가율을 나타낸다. 그리고 각 값의 중앙값을 기준으로 사분면을 구분한다. 1사분면에 속하는 키워드는 평균단어빈도가 높으면서 DoV 평균증가율이 높은 경우로 강신호라 일컫는다. 2사분면에 해당하는 영역의 키워드는 평균단어빈도가 낮으면서 DoV 평균증가율이 높은 약신호가 된다. 3사분면에 위치한 키워드는 아직 두드러지지 않은 잠재신호로 구분할 수 있으며 4사분면 키워드는 단어출현빈도수는 높으나 DoV 평균증가율이 높지 않은 키워드들로 강하지만 출현증가율이 낮은 신호라 일컫는다. 4사분면에 위치한 키워드는 다수의 사람들이 이미 익숙하게 알고 있지만 현시점에서 가시성 정도가 둔화된 것이라 볼 수 있다. KIM은 x축에 평균 DF를 나타내고 y축에는 DoD 평균증가율을 나타내고 각각의 중앙값으로 사분면을 나눈 그래프이다. KIM 또한 KEM과 마찬가지로 1사분면에 위치한 키워드는 평

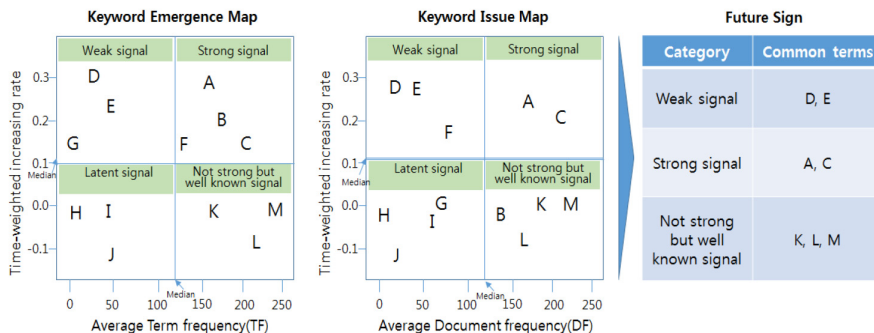


Fig. 2. Future signal classification by KEM and KIM portfolios

군문서빈도와 DoD 평균증가율이 높은 강신호를 의미하며, 2사분면은 약신호, 3사분면은 잠재신호, 4사분면은 강하지만 출현증가율이 낮은 신호라 해석한다. KEM과 KIM 공통적으로 확인된 강신호, 약신호, 잠재신호, 강하지만 출현증가율이 낮은 신호 키워드들을 미래신호로 최종 도출하게 된다.

3.2.2 Word2Vec을 통한 어휘간 유사도 분석

자연어 처리를 위해서는 사람의 언어를 기계 모델이 이해할 수 있는 숫자 형태(벡터)로 변환시켜야 한다. 이러한 과정을 단어 임베딩(Word Embedding)이라 부르며, 대표적인 분석방법으로는 TF-IDF, PMI(Pair-wise Mutual Information), Word2Vec 등이 있다. TF-IDF와 PMI 분석방법은 두 개체 간의 연관 정도를 단순히 단어가 문장에서의 출현하는 횟수만을 계산하여 단어들의 관계를 파악한다[21, 22]. 이러한 계산 방법들은 단어들 이 문장에서의 맥락 또는 문맥상 의미를 바탕으로 한 단어들 간 관계를 전혀 고려하지 않은 계산 방법이기 때문에 정확도의 한계가 존재한다[23]. 본 논문에서는 이 같은 한계를 가진 유사도 계산 방법들을 보완하기 위하여 단어의 의미와 문장에서의 맥락을 고려하여 단어를 벡터로 표현하는 Word2Vec 분석방법을 활용하였다. Word2Vec이란 특정 단어가 등장할 확률을 주변 어휘를 통해 추측하는 것으로, 비슷한 위치에서 등장하는 단어들은 비슷한 의미를 가진다고 가정하고 유사한 벡터값을 준다. 해당 분석방법은 단어의 의미를 벡터 형태로 표현하는 계량기법으로 의미적으로 유사한 단어들은 거리상 근접한 곳에 위치하게 된다. 이러한 분석 특성을 이용하여 단어 Word2vec에 기반 변의 가중치를 계산한다. 각 단어에 대한 Word2vec의 벡터화 결과로 나타내어진 단어 벡터로부터 코사인 유사도 계산 식 (3)을 이용하여 단어들 간 유사도를 계산한다.

$$\text{Word2vec}(u, v) = \frac{\sum_{i=1}^n (u_i v_i)}{\sqrt{\sum_{i=1}^n (u_i)^2} \sqrt{\sum_{i=1}^n (v_i)^2}} \quad (3)$$

u_i : 단어 벡터 u 의 i 번째 요소

v_i : 단어 벡터 v 의 i 번째 요소

(코사인 유사도를 편의상 Word2vec으로 표기)

Word2Vec 결과에 따라 단어 간의 공간벡터의 코사인 유사도가 산출되고, 이를 통해 실질적으로 연구자의 관심 주제와 연관성 높은 단어를 추출할 수 있다[24-26]. 위 식 (3)에 따라 단어 간 유사도가 1에 가까울수록 특정 토픽과 연관성이 높은 단어라고 해석할 수 있다.

Word2vec은 결국 벡터의 계산과 이를 통한 등장 확률 계산이 그 핵심인 알고리즘이라고 할 수 있다. 즉, 학습 데이터를 ‘탄소중립’과 같이 특정 키워드를 주제로 한 기사로 한정할 경우에는 벡터가 나타내는 값은 의미로서의 지표가 아니라 기사 안에서의 어휘의 쓰임새나 경향 등으로 해석할 수 있을 것이다. 따라서 본 연구에서는 도출된 벡터값이 곧 강신호가 될 확률이라 가정하였고, 해당 과정의 구현을 위해 아래와 같은 과정을 진행하였다.

가. 강신호 키워드의 feature vector 선정

각 주제별로 추출한 기사들을 학습시켜 해당 기사에 등장한 어휘들을 벡터로 변환하였다. 히든 레이어의 크기는 100으로 설정하였으며 해당 학습 결과 계산된 강신호의 feature vector를 산출하였다. feature vector은 여러 벡터들을 대표할 수 있는 벡터값으로 강신호에 해당하는 어휘들이 어떤 벡터를 지니는지를 나타낸다고 할 수 있다 [27].

나. 약신호 키워드의 word2vec 계산

다음으로 각 주제별 약신호에 해당하는 벡터값을 계산하였다. 약신호에 해당하는 키워드는 서브 키워드와 관계 없이 주제를 기준으로 나누어 진행하였다.

다. 강신호의 feature vector과 약신호 간의 유사도 측정

마지막으로 강신호의 feature vector과 약신호 간 벡터의 유사도를 측정하여 해당 어휘가 강신호의 영역으로 진입할 확률을 계산하였다.

본 연구에서는 이처럼 텍스트 분석 시 출현 빈도수뿐만 아니라 단어 간의 연관성 분석까지 가능한 word2vec을 미래신호 분석방법과 함께 추가 분석함에 따라 탄소중립 관련 특정 어휘의 강신호, 즉 메가트렌드로의 변화 가능성도 함께 예측 및 제한하였다.

4. 데이터 분석결과

2019.1.1.부터 2021.12.31.까지 보도된 ‘탄소중립’ 관

련 네이버 뉴스 기사 텍스트데이터 81,186건을 수집하였다. 뉴스기사 데이터에서 사용된 키워드의 수는 총 1,759건이었으며, 분석단위 기간은 2019, 2020, 2021년 1년 단위로 설정하였다. 각 기간별 강신호, 약신호, 잠재신호, 강하지만 출현증가율이 낮은 신호 영역에 속한 1,759개 키워드의 변화를 추적하였다.

4.1 미래신호 분석

4.1.1 탄소중립 관련 키워드의 발생증가율

미래신호 연구방법론에 따라 탄소중립 관련 키워드 1,759개의 DoV 증가율과 평균단어빈도를 산출하였다. 아래 Table 2는 1,759개의 키워드 중, 평균 단어출현빈도 최상위 15개의 단어를 나열한 것이다. 분석 결과, 탄소중립 관련 키워드의 DoV 평균증가율 중앙값은 0.112로 나타나 관련 키워드들의 노출 정도는 평균적으로 증가하고 있는 것으로 나타났다. 단어출현빈도 상위 15개의 키워드만 고려해보았을 때, '기업', '산업', '사업', '기술', '수소', '정부', '경제', '발전' 등의 키워드가 평균 단어빈도의 중앙값보다 높으면서 DoV 평균증가율 역시 중앙값보다 높게 나타났다. 이는 해당 단어가 시간이 지남에 따라 신호가 강해지는 강신호였음을 의미한다. '탄소', '환경', '에너지', '위해', '기후', '배출', '가스' 등의 단어는 높은 단어빈도를 보이고 있으나, DoV 평균증가율은 중앙값보다 낮게 나타나 시간이 지남에 따라 신호가 약해지는 것으로 나타났다. 즉, 강하지만 증가율이 낮은 신호

라고 볼 수 있다.

Table 3은 평균 단어출현빈도 최하위 15개 키워드 나타난 표이다. '하이닉스', '생물 다양성', '재사용', '권소 사업' 등의 키워드는 단어출현빈도는 낮지만, DoV 평균증가율이 가장 높은 약신호 단어들이었다. '대기오염', '휘발유', '침체', '소외', '도달', '고속도로' 는 단어출현빈도가 낮고 DoV 평균증가율 또한 낮게 나타나는 잠재신호라는 것을 알 수 있다. 정량적 분석 관점에서 약신호를 떠는 주제는 발생빈도가 낮을 가능성이 다분하지만 발생빈도 증가의 변동폭은 높다. 반대로 발생빈도가 낮고 발생증가율이 높은 키워드를 가진 주제는 중요도가 높고 외부에 자주 노출되는 것으로 간주되어 강신호가 될 수 있다는 점에서 큰 의미를 갖는다.

탄소중립 관련 상위 문서출현빈도 키워드 15개의 DoD 증가율과 평균단어빈도를 산출한 결과는 Table 4와 같다. 분석결과, 탄소중립 관련 키워드의 DoD 평균증가율 중앙값은 .085로 나타나 키워드 관련 주제들의 확산 정도는 평균적으로 증가하고 있는 것으로 나타났다. 키워드별 확산율을 살펴보면 '정부', '산업', '추진', '기업', '사업', '정책', '기술', '지원' 등은 문서빈도가 높고 문서증가율 정도도 높아 강신호가 될 수 있는 가능성을 보이고 있다. '탄소', '위해', '환경', '에너지', '계획', '경제' 등의 단어는 문서빈도는 높으나 확산증가율이 높지 않은 단어들로, 강하지만 확산 정도는 낮은 잘 알려진 주제들로 간주된다.

Table 2. Average term frequency and DoV time weighted increasing rate of carbon neutrality – the top 15 keywords in term frequency

(n=3, tw=0.05)

| Keyword | DoV | | | DoV Time weighted increasing rate | Average Term Frequency |
|-------------|--------|-------|-------|-----------------------------------|------------------------|
| | 2019 | 2020 | 2021 | | |
| carbon | 3.715 | 3.818 | 3.637 | -0.007 | 102,538 |
| Environment | 2.032 | 2.054 | 2.017 | -0.002 | 56,659 |
| energy | 1.558 | 2.310 | 1.982 | 0.083 | 56,495 |
| Enterprise | 0.687 | 1.306 | 1.871 | 0.397 | 50,297 |
| industry | 0.847 | 1.631 | 1.716 | 0.265 | 47,595 |
| business | 0.685 | 1.365 | 1.678 | 0.348 | 45,756 |
| technology | 0.754 | 1.118 | 1.578 | 0.279 | 42,531 |
| Hydrogen | 0.701 | 0.978 | 1.560 | 0.305 | 41,625 |
| government | 0.822 | 1.984 | 1.320 | 0.171 | 39,015 |
| Harm | 1.270 | 1.383 | 1.364 | 0.024 | 38,259 |
| economy | 0.896 | 1.879 | 1.266 | 0.122 | 37,367 |
| development | 0.926 | 1.405 | 1.287 | 0.116 | 36,359 |
| climate | 2.324 | 1.742 | 1.199 | -0.198 | 35,630 |
| support | 0.399 | 1.028 | 1.156 | 0.426 | 31,784 |
| plan | 0.795 | 1.148 | 0.995 | 0.078 | 28,322 |
| ... | ... | ... | ... | ... | ... |
| | Median | | | 0.112 | 1,649 |

Table 3. Average term frequency and DoV time weighted increasing rate of carbon neutrality – The lower 15 keywords in term frequency (n=3, tw=0.05)

| Keyword | DoV | | | DoV Time weighted increasing rate | Average Term Frequency |
|---------------|-------|-------|-------|-----------------------------------|------------------------|
| | 2019 | 2020 | 2021 | | |
| Air pollution | 0.084 | 0.025 | 0.017 | -0.419 | 507 |
| gasoline | 0.025 | 0.026 | 0.017 | -0.125 | 507 |
| stagnation | 0.015 | 0.028 | 0.017 | 0.052 | 514 |
| Isolation | 0.013 | 0.031 | 0.017 | 0.078 | 516 |
| arrival | 0.048 | 0.016 | 0.018 | -0.273 | 519 |
| Hynix | 0.001 | 0.012 | 0.020 | 1.452 | 523 |
| biodiversity | 0.009 | 0.010 | 0.020 | 0.286 | 524 |
| highway | 0.024 | 0.010 | 0.020 | -0.060 | 528 |
| recycle | 0.005 | 0.028 | 0.018 | 0.492 | 530 |
| consortium | 0.003 | 0.005 | 0.021 | 0.985 | 530 |
| Justice Party | 0.027 | 0.024 | 0.018 | -0.119 | 532 |
| North America | 0.008 | 0.013 | 0.020 | 0.354 | 533 |
| know-how | 0.003 | 0.007 | 0.021 | 0.981 | 534 |
| collection | 0.011 | 0.017 | 0.019 | 0.220 | 534 |
| damage | 0.036 | 0.034 | 0.017 | -0.223 | 535 |
| ... | ... | ... | ... | ... | ... |
| Median | | | | 0.112 | 1,649 |

Table 4. Average document frequency and DoD time weighted increasing rate of carbon neutrality – The top 15 keywords in document frequency (n=3, tw=0.05)

| Keyword | DoD | | | DoD Time weighted increasing rate | Average Document Frequency |
|-------------|-------|-------|-------|-----------------------------------|----------------------------|
| | 2019 | 2020 | 2021 | | |
| carbon | 0.819 | 0.858 | 0.920 | 0.040 | 25,566 |
| harm | 0.560 | 0.551 | 0.605 | 0.026 | 16,760 |
| Environment | 0.600 | 0.545 | 0.571 | -0.016 | 15,934 |
| government | 0.300 | 0.554 | 0.459 | 0.153 | 13,138 |
| industry | 0.290 | 0.443 | 0.467 | 0.171 | 12,958 |
| energy | 0.396 | 0.476 | 0.460 | 0.051 | 12,937 |
| plan | 0.383 | 0.438 | 0.460 | 0.063 | 12,801 |
| economy | 0.373 | 0.552 | 0.445 | 0.060 | 12,792 |
| enforcement | 0.175 | 0.471 | 0.450 | 0.369 | 12,612 |
| Enterprise | 0.300 | 0.369 | 0.448 | 0.143 | 12,267 |
| business | 0.199 | 0.358 | 0.431 | 0.293 | 11,786 |
| Policy | 0.257 | 0.442 | 0.397 | 0.156 | 11,230 |
| technology | 0.273 | 0.334 | 0.402 | 0.138 | 11,023 |
| support | 0.226 | 0.362 | 0.395 | 0.204 | 10,912 |
| climate | 0.444 | 0.446 | 0.361 | -0.067 | 10,411 |
| ... | ... | ... | ... | ... | ... |
| Median | | | | 0.085 | 1,046 |

Table 5는 탄소중립 관련 문서출현빈도 하위 15개 키워드를 나타낸 표이다. ‘제네시스’, ‘기본소득’, ‘대구시’, ‘한수원’, ‘주유소’, ‘창원시’, ‘한울’, ‘연료비’ 등의 단어들은 관련 키워드들이 등장하는 문서빈도는 낮지만, 상대적으로 확산 증가율은 높아 앞으로 더 강해지거나 약해질 수 있는 변동 가능성을 보인다.

4.1.1 탄소중립 관련 키워드 포트폴리오

미래신호예측분석을 위하여 앞서 도출한 1,759개의 키워드별 발생증가율 수치를 바탕으로 키워드 포트폴리

오 맵인 KEM과 KIM을 작성하였다(Fig. 3, 4). Fig. 5는 KEM과 KIM에서 공통적으로 나타나는 미래신호를 표시한 그래프이다. KEM과 KIM에서 공통적으로 나타나는 강신호(1사분면)에는 뉴딜, 생태계, 신재생에너지, 과학기술, 경쟁력 등 91개의 키워드가 출현하였다. 약신호(2사분면)에는 철강업, 석유화학, 이차전지와 같은 산업기술분야 키워드들과 롯데, 하이닉스, 삼성과 같은 국내 산업기술 관련 대기업이 도출되었다. 또한, 테슬라와 제네시스와 같은 전기차 개발에 박차를 가하고 있는 기업 키워드 외에도 총 424개의 키워드들이 도출되었다. 잠재신

Table 5. Average document frequency and DoD time weighted increasing rate of carbon neutrality – The lower 15 keywords in document frequency (n=3, tw=0.05)

| Keyword | DoD | | | DoD Time weighted increasing rate | Average Document Frequency |
|--------------|-------|-------|-------|-----------------------------------|----------------------------|
| | 2019 | 2020 | 2021 | | |
| Audi | 0.051 | 0.006 | 0.005 | -0.551 | 144 |
| Volvo | 0.015 | 0.010 | 0.005 | -0.292 | 163 |
| Genesis | 0.003 | 0.001 | 0.007 | 0.373 | 173 |
| basic income | 0.001 | 0.009 | 0.006 | 0.674 | 183 |
| Hungary | 0.064 | 0.003 | 0.007 | -0.530 | 191 |
| Benz | 0.047 | 0.008 | 0.006 | -0.486 | 192 |
| Volkswagen | 0.090 | 0.009 | 0.006 | -0.587 | 206 |
| Daegu | 0.001 | 0.006 | 0.008 | 0.795 | 208 |
| KHNP | 0.004 | 0.002 | 0.008 | 0.273 | 211 |
| gas station | 0.005 | 0.006 | 0.008 | 0.136 | 213 |
| mixed fire | 0.008 | 0.002 | 0.008 | 0.013 | 215 |
| Changwon | 0.004 | 0.006 | 0.009 | 0.284 | 229 |
| Hanul | 0.001 | 0.010 | 0.008 | 0.835 | 235 |
| fuel cost | 0.005 | 0.012 | 0.008 | 0.143 | 238 |
| Spain | 0.092 | 0.006 | 0.008 | -0.558 | 239 |
| ... | ... | ... | ... | ... | ... |
| Median | | | | 0.085 | 1,046 |

호(3사분면)에는 684개, 강하지만 출현증가율이 낮은 신호(4사분면)에는 285개의 키워드가 위치하고 있다.

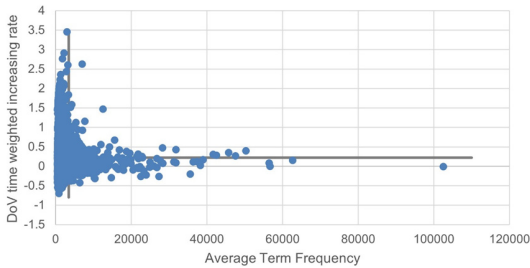


Fig. 3. 2019–2021 carbon neutrality KEM portfolios

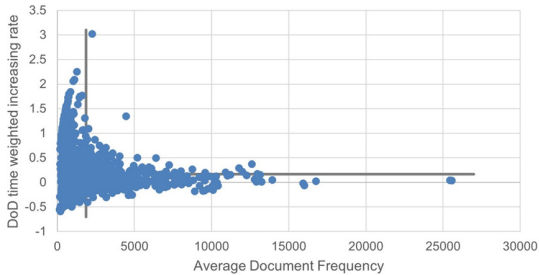


Fig. 4. 2019–2021 carbon neutrality KIM portfolios

특히 약신호인 2사분면에 출현한 키워드는 시간이 지남에 따라 강신호로 발전할 수 있기 때문에 이에 대한 대응책을 적극적으로 마련해야 한다. 이러한 근거로 2사분면에 출현한 키워드를 중심으로 탄소중립 관련 현황과 대응책을 유추해보자면 다음과 같다. 먼저, 탄소 과배출 업종의

위기 돌파를 위한 기술개발 필요성이 강조되고 있음을 확인할 수 있다. 그에 따라 철강업 및 석유화학 분야 기술의 고도화 외에도 기술 이노베이션을 꾀해야 할 것이며, 하이닉스, 삼성과 같은 반도체 회사의 탄소저감을 위한 기술개발 지원이 요구될 것으로 보인다. 또한, 탄소저감을 위한 혁신적인 기술개발이 필요함을 보여준다. 전기차 분야의 투자 타당성 및 고급화 기술을 확보해야 할 것이며, 삼성 등 이차전지 대기업에 대한 분야 지원이 요구될 것이다. 해당 기술분야의 선두인 중국과의 기술협력을 검토해야 할 것이며, 베트남 등 해외공장 투자 활성화도 고려해야 할 것이다. 무엇보다도 공기기업들의 자체적 기술 개발을 위한 정부의 적극적 지원이 밑받침되어야 할 것이다.

| | Weak signal (424 keywords) | Strong signal (91 keywords) |
|-------------------------------|---|---|
| Time weighted increasing rate | Steel industry Lotte Kakao Hynix advancement Xi jingping reability Petrochemicals enterprise High potential venture Tesla Genesis | Rechargeable batteries Samsung innovation Vietnam added value Public enterprise ... Demand Supply job finance researcher ... |
| | Latent signal (684 keywords) start-up company KEPCO Civic organization Public transportation Hydrogen car Waste Digitization Ammonia Natural gas Thermal power plant Safety subsidy | New Deal Ecosystem Renewable energy Petroleum Wind power Activation Democratic Party Demand Supply job finance researcher ... |
| | Not strong but well known signal (285 keywords) Recycling Solar Eco-friendly Hyundai Material Effort Battery Plastic Nuclear power | competitiveness Science and technology Moon Jae In Semiconductor discovery hydrogen Capability Small and medium enterprise chemistry recovery KHNP ... |
| | Average term frequency | Average document frequency |

Fig. 5. Future signals related to carbon neutrality in 2019–2021

4.2 Word2Vec을 통한 약신호의 강신호 진입 가능성 예측

각 keyword별로 약신호 중 어떤 keyword가 강신호로 갈 것인가에 대한 확률값 계산을 위해 machine learning 기법 중 Word2Vec 방법론을 활용하기로 하였다. Word2Vec은 일반적인 단어 임베딩 모델(embedding model)과 같이 언어학의 분산원리를 따르며, 이는 비슷한 의미의 단어는 비슷한 문장에서 나타난다는 개념으로 단어를 공간 벡터에 임베딩한다. 그 결과에 따라 단어 간의 공간벡터의 코사인 유사도가 산출되고, 이를 통해 실질적으로 연구자의 관심 주제와 연관성 높은 단어를 추출할 수 있다[28].

본 연구에서의 Word2Vec 분석은 탄소중립 관련 약신호에 해당하는 키워드 424개 중 강신호로 넘어갈 확률이 높은 어휘를 선정하고자 진행하였으며 이를 위해 다음과 같은 과정을 진행하였다. 먼저 424건의 약신호 키워드들을 각 토픽별로 분류하였다. 키워드들의 토픽 주제를 분류는 델파이 분석법을 적용하여 산업부 내 온톨로지를 기반으로 하였다. 토픽은 '기술', '인재', '인프라', '정책', '지역'으로 결정되었으며 해당 토픽에 맞춰 키워드들을 분류하였다. 각 토픽별로 강신호에 해당하는 어휘들을 선정하고 해당 키워드들의 벡터값을 수렴한 뒤 강신호에 해당하는 feature 벡터를 산출하였다. feature 벡터와 유사한 값을 지닌 어휘는 입력한 텍스트 안에서 비슷한 역할을 하며 텍스트 안에서 유사할 기능을 수행한다는 것을 시사한다[24]. 따라서 feature 벡터의 값과 유사한 값을 지니는 약신호 어휘는 강신호의 영역으로 진입할 가능성이 높다고 추론할 수 있다. 아래 Table 6은 위에서 언급한 각 토픽별 서브키워드와 해당 서브키워드에 해당하는 약신호 단어들 중, 강신호의 feature 벡터와 유사도가 0.5 이상인 단어만 표시한 것이다. 토픽별 Word2Vec 연관성 분석결과, 5개의 토픽 중 '기술', '인재', '인프라' 부문에서만 연관성이 높은 키워드들이 도출되었다. 해당 분석을 통해 탄소중립 관련 Word2vec 분석 결과에 따라 각 부문별 약신호(Weak Signal) 중 강신호(Strong Signal)로 갈 수 있는 키워드들을 찾아내었다.

Table 6. Word2vec-based carbon-neutrality sub-keyword Weak signal

| division | Carbon Neutrality Weak Signalz |
|----------------|---------------------------------------|
| technology | Raw Material (0.54), patent(0.58) |
| talent | Human Resources (0.76) |
| infrastructure | hub (0.60), research institute (0.57) |

4.2.1 기술 부문

먼저 '기술' 토픽에 해당하는 키워드들은 '시멘트', '제철', '중공업', '선점', '광물' 등이 있었으며, 이 중 Word2Vec 연관도가 높게 도출된 것은 '원자재', '특허'였다. '원자재'의 토픽 연관성 수치는 0.54이고 '특허'는 0.58로, 해당 두 키워드가 1과 가장 가까운 값을 가진 키워드로 도출되었다. '원자재'와 '특허' 키워드가 '기술' 토픽과 가장 연관성이 높은 단어라 볼 수 있다. 이는 곧 탄소중립을 위한 글로벌 구조 개편에 의한 원자재 시장의 동향이 빠르게 일어나는 가운데, 고철과 같이 탄소배출량이 적은 원자재를 중심으로 이 분야의 기술확대가 일어날 것이라고 볼 수 있을 것이다. 또한, 전기차 관련 이차전지 등 기술 확보에 국가 간 사력을 다할 것으로 보이며, 해당 분야 기술특허 확보를 위해 정부 차원의 지원이 필요할 것으로 보인다.

Table 7. Articles on carbon neutrality weak signal_technology

| No. | Keyword | Date | Article |
|-----|--------------|-------------|--|
| 1 | Raw Material | 2021.12.24. | '친환경 원자재값 상승'에 재활용 뜨는데...규제 지원 '제자리걸음' |
| 2 | Raw Material | 2021.12.21. | [2021결산-철강] 역대급 실적에도...탄소중립 리스크에 '웃음 독!' |
| 3 | Raw Material | 2021.12.23. | 탄소중립에 사활 건 포스코-현대제철, 라이벌 타협종 다 손잡는다 |
| 4 | patent | 2021.12.29. | 페로브스카이트 태양전지 특허출원 급증...상용화 연구개발 활발 |
| 5 | patent | 2021.12.28. | '탄소중립 본격화...전기차 수소차 내년까지 50만대 보급(종합) |
| 6 | patent | 2021.12.23. | 대한상의, 중소 중견기업에 탄소중립 R&D 특허전략 지원한다 |

4.2.2 인재 부문

'인재' 토픽에 해당하는 연관성 높은 단어로는 '인력'이 도출되었다. 이는 곧 탄소중립에 관한 인력이 부족하며 이에 대한 양성이 시급한 상황임을 함의한다. 기관별 인력 교류, 기업과의 산학협동을 증가시키는 방향으로 가

Table 8. Articles on carbon neutrality weak signal_talent

| No. | Keyword | Date | Article |
|-----|-----------------|-------------|--|
| 1 | Human Resources | 2021.12.27. | 文대통령 "인재는 기업의 확실한 투자처"... 삼성 현대차 직접 언급 |
| 2 | Human Resources | 2021.12.22. | '탄소중립 대응' 2030년까지 에너지 기술인력 2만명 육성 |
| 3 | Human Resources | 2021.12.03. | 과기부 내년 예산 18.6조원... '디지털 뉴딜 2.0' 인재 양성 탄소 중립 중점 투자 |

야할 것이며, 과학기술부에서도 탄소중립 관련 에너지 기술인력 양성에 집중 투자하고 있는 것으로 보여진다. 뿐만 아니라 산업부에서는 대학, 연구소보다 기업에서 필요한 탄소중립 인력을 양성할 수 있도록 간접지원정책을 강구해야 할 것으로 보인다.

4.2.3 인프라 부문

'인프라' 부문에서는 '허브'와 '연구기관'의 토픽 연관성 수치가 각각 0.60, 0.57로 높게 도출되었다. 탄소중립과 깊게 관련된 중화학공업 등의 '탄소중립화 허브'가 확산되는 것을 물론 연구기관 간 협업 중요성이 날로 증가할 것으로 예측된다. 정부는 중화학 공업단지가 모여있는 여수시, 포항시 등 허브 도시를 중심으로 연구기관을 묶어서 그 지역 내 탄소 발생 문제를 자체해결할 수 있도록 그 방안을 모색해야 할 것이다.

Table 9. Articles on carbon neutrality weak signal_infrastructure

| No. | Keyword | Date | Article |
|-----|--------------------|-------------|--|
| 1 | hub | 2021.12.29. | 산업부, 내년 탄소중립 실현 본격화...핵심기술 개발 속도 |
| 2 | hub | 2021.12.26. | [탄소와의 전쟁] 탄소판 '아나버다', CCUS 기술로 실현하다 |
| 3 | hub | 2021.12.22. | 여수시, 13개 기관·기업 간 협약 체결...탄소중립 실현·수소산업 육성 |
| 4 | research institute | 2021.12.28. | 지구를 구하는 길, 넷제로...주요국 준비 상황은[탄소중립, 시대적 과제] |
| 5 | research institute | 2021.12.27. | 탄소중립 이끌 '한국화학연구원 여수분원' 유치 |
| 6 | research institute | 2021.12.21. | 롯데케미칼·삼성엔지니어링 등 컨소시엄, 탄소 중립 위한 암모니아 분해 수소 기술개발 |

5. 결론 및 시사점

탄소중립 관련 기술개선 방안 모색을 목표로 시행된 본 연구는 탄소중립 구현을 위한 구체적 이행방안 수립 근거를 제시할 수 있었다. 탄소중립 관련 온라인 뉴스기사 비정형 텍스트데이터를 바탕으로 미래신호분석방법과 Word2Vec 알고리즘을 사용하여 탄소중립 기술의 핵심 트렌드를 분석하였다. 미래신호 분석 결과, 미래신호 그래프 중 약신호 영역에 도출된 어휘들로 유추해보았을 때 탄소저감 기술개발과 정부의 적극적인 지원이 요구되고 있음을 알 수 있었다. 특히나 탄소배출량이 적은 원자재 관련 기술과 전기차 부문의 기술혁신이 필요함을 확인할 수 있었으며, Word2Vec 분석에 따르면, '원자재', '특

허', '인력', '허브', '연구기관' 등의 약신호 키워드가 강신호 영역으로 진입할 가능성이 높은 것으로 나타났다. 이는 활발한 기술개발연구를 위해 연구기관 간 인프라 허브 구축과 관련 인력양성 지원정책도 함께 강구해야 함을 내포하고 있다.

본 논문에서는 텍스트마이닝을 통해 탄소중립 관련 미래유망 기술을 파악해보았고, 더 나아가, 미래신호 분석 결과를 기반으로 Word2Vec 도출을 수행함으로써 객관적 방법의 분석을 수행할 수 있었다. 또한, 앞으로 탄소중립 실현과정에서 요구될 기술적인 이슈에 대한 현황검토와 발전 방향 예측을 시도했다는 것에 실무적 차원의 의의를 가진다. 이러한 예측은 향후 탄소중립 관련 산업기술 역량 증진과 글로벌 선진기술 선점과 같은 선제적 대비에 도움이 될 것으로 보인다. 뿐만 아니라, 산업기술 혁신역량 제고를 위한 미래산업기술 트렌드 및 동향을 파악할 수 있을 것이며, 산업기술 혁신을 위한 방안 수립 등에 필요한 기초분석 자료로 활용될 수 있을 것이다.

그러나 본 연구는 다음과 같은 한계점을 가지고 있다. 먼저, 국외 데이터를 고려하지 않고 국내 데이터만을 대상으로 분석되었다는 한계가 있다. 타 국가의 웹 데이터도 함께 고려하여 비교 분석하였다면 우리나라의 기술발전 정도를 객관적으로 파악하여 심도있는 분석을 진행할 수 있었을 것이다. 두 번째로, 본 연구에서는 탄소중립 관련 특정 토픽을 기준으로 키워드를 필터링한 것이 아니라, 81,186건의 뉴스 데이터 내에서 사용된 1,759개의 키워드 전체를 대상으로 연구를 진행한 후 약신호로 검출된 키워드에 한하여 세부 분류했다는 한계가 있다. 탄소중립 관련 기술이나 인재, 인프라, 정책, 지역 등 알고자 하는 세부 토픽을 먼저 선정한 후에 키워드 분류 및 분석을 진행했다면, 보다 구체적인 결과해석과 타당성 확보 또한 가능했을 것이다. 향후 탄소중립에 국한된 것이 아닌 여러 산업기술들을 대상으로 세부 온톨로지를 적용하여 분석한다면 좀 더 신뢰도 높은 결과를 도출할 수 있을 것이며, 국내 산업기술 보완과 역량 증진에 큰 도움이 될 것으로 기대한다.

REFERENCES

- [1] Science & Technology Policy Institute, (2022. 01.01.). what is the future. Retrieved from <https://stepi.re.kr>.

- [2] Flechtheim, Ossip. (1971). "Is Futurology the Answer to the Challenge of the Future?" In *Mankind 2000*, edited by Robert Jungk and Johann Galtung, 264-269. London: Allen & Unwin.
- [3] Aligica, P. D. (2004). The challenge of the future and the institutionalization of interdisciplinarity: notes on Herman Kahn's legacy. *Futures*, 36(1), 67-83.
DOI : 10.1016/S0016-3287(03)00136-8
- [4] Gordon, T. J., & Helmer, O. (1964). Report on a long-range forecasting study. RAND CORP SANTA MONICA CA SANTA MONICA United States.
- [5] International Future Society. (2014). *Strategic Future Forecasting Methodology Bible*. Seoul: Dunam Publishing Co.
- [6] Bae, Y. J., (2016). Foresight on future fashion in Korea applying futures research methodology. Seoul National University.
- [7] Park, S. C., (2013). *The Understanding of Futurology*, Seoul: Chungmook Publishing Co.
- [8] Glenn, J. C. (1994). *Introduction to the futures research methodology series*. Washington, DC: United Nations University.(Part of Glenn 1994a).
- [9] Aaltonen, M., & Irene Sanders, T. (2006). Identifying systems' new initial conditions as influence points for the future. *Foresight*, 8(3), 28-35. DOI : 10.1108/14636680610668054
- [10] Bae, Y. S., Park, J. Y., & Park, G. (2014). Object recognition and visual intelligence platform technology trends based on image big data. *Korea Information Processing Society Review*, 21(3), 59-67.
- [11] Kim, S.Y., & Jeong, Y.J. (2017). Learning machine learning for the first time: from basics to modeling, practical examples, and problem solving. Hanbit Media.
- [12] Park, C. K. (2017). A study on how to utilize big data for the future prospects of the electric power industry. *Korea Energy Economics Institute*.
- [13] Roh, S.K., & Choi, J. Y. (2020). Exploring Signals for a Nuclear Future Using Social Big Data. *Sustainability*, 12(14), 5563.
- [14] Song, T.M. & Song, J.Y. (2016). Future Signals of Health and Welfare Policies and Issues using Social Big Data. *Journal of Health Informatics and Statistics*, 41(4), 417-427.
DOI : 10.21032/jhis.2016.41.4.417
- [15] Song, T.M. & Song, J.Y. (2016). Using Social Big Data Predictive Future Signal: With Special Reference to the Major Policy Issues of Health and Welfare. *Health and Welfare Policy Forum*, 41(4), 417-427.
- [16] Song, T.M. & Song, J.Y. (2017). *Social Big Data and Future Prediction with Machine Learning*. Hannarae.
- [17] Ansoff, H.I., (1975). Managing strategic surprise by response to weak signals, *Californian Management Review* 18(2), 1975, pp.21-33.
- [18] Hiltunen, E. (2008). The future sign and its three dimensions. *Futures*, 40(3), 247-260.
- [19] Jeong, D.Y., (2018). STEPI Future Research Glossary. *FUTURE HORIZON*, (35), 30-31.
- [20] Yoon, J., Detecting weak signals for long-term business opportunities using text mining of Web news, *Journal Expert Systems with Applications* 39(16), 2012, pp.12543-12550.
DOI : 10.7472/JKSII.2015.16.2.41
- [23] Chang, J.Y.. (2013). A Study on Research Trends of Graph-Based Text Representations for Text Mining. *The Journal of The Institute of Internet, Broadcasting and Communication*, 13(5).
- [22] Oh, Y.J. & Chae, S.H.. (2015). Movie Rating Inference by Construction of Movie Sentiment Sentence using Movie comments and ratings. *Journal of Internet Computing and Services*, 16(2), 41-48.
- [23] Xue, B., Fu, C., and Shaobin, Z. (2014). A study on sentiment computing and classification of sinaweibo with word2vec, In *Big Data, BigData Congress, IEEE International Congress on*, pp. 358-363, IEEE.
- [24] Kang, H.S. & Yang, J.H. (2019). Analyzing Semantic Relations of Word Vectors trained by The Word2vec Model. *Journal of KIISE*, 46(10), 1088-1093. DOI : 10.5626/jok.2019.46.10.1088
- [25] Kang, H.S. & Yang, J.H. (2019). Optimization of Word2vec Models for Korean Word Embeddings. *Journal of Digital Contents Society*, 20(4),

- 825-833. DOI : 10.9728/dcs.2019.20.4.825
- [26] Kang, H.S. & Yang, J.H. (2019). On Characteristics of Word Embeddings by the Word2vec Model. *Proceedings of the Korea Information Processing Society Conference*, 26(1), 263-266.
- [27] Lee, J.S., Shin, J.C. & Ock, C.Y. (2016). An Word vector(USenseVector) constructed Korean Lexical Semantic Map(UWordMap) and Korean Dictionary. *Proceedings of the Korean Information Science Society Conference*, 750-752.
- [28] Kang, H.S. & Yang, J.H. (2020). Performance Comparison of Word2vec and fastText Embedding Models. *Journal of Digital Contents Society*, 21 (7), 1335-1343. DOI : 10.9728/dcs.2020. 21.7.1335
- [29] Kim, Y.H. & Han, C.K. (2021). Social Conflict Issues and Searching for Future Signal since COVID-19: Focusing on the Keywords in the News Articles. *Korean Journal of Social Welfare Studies*, 52(1), 5-37. DOI : 10.16999/KASWS.2021.5 2.1.5
- [30] Park, E. Y. (2019). Do it! Jump to Python. easyspublishing.
- [31] Park, J.M. & Song, T.M. (2020). Identifying Strong and Emerging Signals Related to Personal Debt using Social Big Data, *Korean Journal of Social Welfare*, 72(1), 35-61. DOI : 10.20970/kasw.2020.72.1.002
- [32] Park, C.K., Kim, H.J. (2015). A study on the development direction of the new energy industry through the Internet of Things - Exploring future signals using text mining. *Korea Energy Economics Institute*.
- [33] Baek, C.W., & Roh, S.K. (2020). Analysis of Portal News Articles and User Opinion Data on Sexual Crimes: Focused on Nth Room Sex Crime Case. *The Journal of Police Policies*, 34(3), 41-72. DOI : 10.35147/KNPSI.2020.34.3.41
- [34] Lim, H.W., Lee, S.Y., Choi, H.S. & Jeong, E.H. (2022). Policy Proposals to Achieve the New Government's Carbon Neutrality Goal, *Korea Environment Institute Environment Forum*, 1-25.
- [35] Jeong, J. S., & Kim, H. D. (2021). Analysis of media trends related to spent nuclear fuel treatment technology using text mining techniques. *Journal of Intelligence and Information Systems*, 27(2), 33-54. DOI : 10.13088/ JIIS.2021.27.2.033
- [36] Choi, J.M. & Oh, H.Y. (2021). Analysis of whether the feeling of relative deprivation is shown in the comments of the Luxury Howl YouTube video - Focusing on modern sentiment analysis using TF-IDF, Word2vec, LDA and LSTM -. *Journal of the Korea Institute Of Information and Communication Engineering*, 25(3), 355-360. DOI : 10.6109/JKIICE.2021.25.3.355
- [37] Heo, C., & Ohn, S.Y. (2017). A Novel Method for Constructing Sentiment Dictionaries using Word2vec and Label Propagation. *The Journal of Korean Institute of Next Generation Computing*, 13(2), 93-101.
- [38] Vargas-Calderón, V., & Camargo, J. E. (2019). Characterization of citizens using word2vec and latent topic analysis in a large set of tweets. *Cities*, 92, 187-196. DOI : 10.1016/j.cities.2019.03.019

정지송(Ji song Jeong)

[정회원]



- 2017년 8월 : 동국대학교 원자력공학과(공학사)
- 2021년 2월 : 과학기술연합대학원대학교 한국원자력연구원 양자에너지화학공학과(공학석사)
- 2023년 2월: 경찰대학 범죄학과(범죄학석사)

- 관심분야: 빅데이터, 자연어처리
- E-Mail : jisong9410@gmail.com

노승국(Seungkook Roh)

[정회원]



- 2001년 2월 : 고려대학교 화학과(이학사)
- 2003년 2월 : 포항공과대학교 전자및컴퓨터공학부(공학석사)
- 2014년 8월 : 한국과학기술원 문화기술대학원(공학박사)

- 2020년 1월~현재 : 경찰대학 치안대학원 데이터사이언스 전공 조교수
- 관심분야 : 범죄통계분석, 수사
- E-Mail : skroh@police.ac.kr