

# 뉴스 데이터를 활용한 재난에 대한 사회적 관심 측정방법 개발

신은혜<sup>1</sup> · 김도우<sup>2</sup> · 정재학<sup>3</sup> · 장성록<sup>4\*</sup>

## Development of a Method for Measuring Social Interest Index on Disaster Using News Data

Eun Hye Shin<sup>1</sup> · Do Woo Kim<sup>2</sup> · Jae Hak Chung<sup>3</sup> · Seong Rok Chang<sup>4\*</sup>

### <sup>1</sup>Corresponding Author

Seong Rok Chang

Tel : +82-51-629-6468

E-mail : srchang@pknu.ac.kr

Received : June 16, 2023

Revised : July 29, 2023

Accepted : September 1, 2023

**Abstract** : Social interest in disasters is a significant factor in shaping disaster management policies, enabling the enhancement of disaster safety management and prevention activities according to the specific needs of society. However, in the past, there were limits to measuring which disasters society was particularly interested in. Hence, in this study, a method of measuring social interest using news data was devised. Specifically, we classified news reports into natural and social disasters, creating a comprehensive Social Interest Index (SII) on disasters covering from 2011 to 2021. Additionally, we quantitatively compared the SII with budgets allocated to disaster-related efforts. Our primary findings are as follows: First, our methodology not only distinguishes natural disasters from social disasters but also identifies emerging areas of societal concern. Second, in recent years (2014-2021), social disasters gained more attention than natural disasters. Third, the disaster safety budget accounted for approximately 3.5% of Korea's total budget, closely paralleling the SII we measured. However, exceptions were noted in cases such as heavy snow, cold waves, and heat waves, where the SII remained high, but the disaster safety budget was relatively low, indicating potential outliers. The findings of this research are projected to contribute to the improvement of national disaster management policies by providing a quantitative measure of social interest in disaster, enabling more informed and effective policy decisions.

**Key Words** : classification of disaster news, measuring social interest

Copyright©2023 by The Korean Society of Safety All right reserved.

## 1. 서론

재난에 대한 사회적 관심은 재난관리정책의 중요한 요소이다<sup>1)</sup>. 사회적 관심이 높을 때 정부는 이를 재난에 대한 예산 투자 확대 등 재난안전관리 강화의 기회로 활용할 수 있으며, 사회적 관심이 낮을 때에는 안전불감증으로 인한 사고가 발생하지 않도록 예방활동을 강화할 수 있다.

그 동안 재난에 대한 사회적 관심은 주로 대중에 대한 설문조사를 통해 측정되어져 왔다<sup>2)</sup>. 대표적으로 통

계청에서 실시하는 “한국의 사회조사<sup>3)</sup>”와 행정안전부의 “국민 안전의식 조사<sup>4)</sup>”가 있다. 하지만 설문조사 방법은 장기간 동안 설문 대상 집단을 동일하게 유지하기 힘들므로 수집된 자료의 균질성에 한계가 있다. 뿐만 아니라 설문조사에 필요한 비용적 측면 때문에 설문조사를 일부 집단에 한정하여 수행할 수 밖에 없다. 따라서 설문조사 결과의 대표성에도 한계가 존재한다. 마지막으로 재난에 얼마나 관심을 갖고 있는지에 대한 설문은 설문을 수행하는 시기에 민감하게 반응할 수 있다. 예를 들어 설문조사가 대형재난이 발생한 직후

<sup>1</sup>국립재난안전연구원 사회재난연구센터 책임연구원 (Center for Societal Disaster Research, National Disaster Management Research Institute)

<sup>2</sup>국립재난안전연구원 사회재난연구센터 기상연구관 (Center for Societal Disaster Research, National Disaster Management Research Institute)

<sup>3</sup>국립재난안전연구원 사회재난연구센터 사회재난연구센터장 (Center for Societal Disaster Research, National Disaster Management Research Institute)

<sup>4</sup>부경대학교 안전공학과 교수 (Department of Safety Engineering, Pukyong National University)

에 수행된다면 재난에 대한 관심이 과도하게 높다고 추정될 수 있다.

이러한 한계점들 때문에 연구자들은 뉴스 데이터를 활용하여 재난에 대한 사회적 관심을 측정하는 시도를 하였다. 2007년 미국에서 처음으로 재난에 대한 사회적 관심을 신문 지면 상 보도 횟수와 지면의 양을 통해 수치적으로 측정할 수 있음이 제시되었다<sup>5)</sup>. 국내에서도 신문 기사의 수, 기사의 면적, 기사의 글자 수로 계량화하여 이슈에 대한 사회적 관심도를 측정할 수 있음이 제시되었다<sup>6)</sup>. 김철민과 최충익<sup>7)</sup>에서는 시간에 따른 특정 재난사고(18개)에 대한 뉴스 건수 감소비율을 계산하여 한국에서 대형재난이 발생하면 그 이후 사회적 관심이 얼마나 오래 지속되는지를 측정하였고, 유사한 방법으로 유호선 등<sup>8)</sup>은 뉴스 건수 감소 패턴에 따라 재난 시 사회적 관심이 지속되는 유형을 2가지(돌발형, 주기형)로 구분하였다. 그 밖에도 태풍, 홍수, 호우, 폭염 등 특정 자연재난 유형의 피해와 관련하여 어떤 키워드가 사회적 관심을 많이 받았는지를 뉴스 자료를 이용하여 분석한 연구들이 있었다<sup>9-10)</sup>.

하지만 앞서 언급된 국내 연구들은 대부분 특정 재난 유형 혹은 특정 재난 사건에 제한되어 분석이 이루어졌으며, 다양한 재난유형을 포괄하는 부분에 대한 사회적 관심을 측정한 연구는 발견되지 않는다. 이는 다음의 2가지 어려움이 있기 때문이라 사료된다. 첫 번째 어려움은 어떤 뉴스를 재난 관련 뉴스라고 정의할 것인가이다. 홍수, 태풍, 가뭄과 같은 자연재난의 경우, 재난현상을 나타내는 단일 단어로 해당 뉴스가 재난 관련 뉴스인지를 대부분 판별할 수 있다. 하지만 사회재난의 경우, 해당 단어가 반드시 재난과 관련되어 한정적으로 사용되는 것이 아니기 때문에 단일 단어만으로는 관련 뉴스를 판별하기 어렵다. 예를 들어, 붕괴의 경우 해당 단어로 뉴스를 탐색할 경우 ‘의료체계 붕괴법’, ‘상권 붕괴’, ‘코스피 붕괴’ 등과 같이 재난과 관련 없는 뉴스도 함께 탐색 된다. 이 때문에 다양한 재난에 대한 뉴스를 종합적으로 분석하려고 시도한 연구들도 대부분 자연재난 분야에 제한되어 이루어졌다<sup>11-12)</sup>. 두 번째 어려움은 뉴스 데이터 전수를 획득하는 것이다. 대부분의 뉴스 기반 연구들은 뉴스를 온라인상에서 크롤링하여 수집하였다<sup>13-14)</sup>. 따라서 이들 뉴스 데이터는 전수 데이터가 아닌 일종의 샘플링 데이터에 해당된다. 즉, 특정 기간동안 생산된 전체 뉴스에서 재난 관련 뉴스가 얼마나 큰 비율을 차지하는지를 산출할 수 없다.

본 연구에서는 자연재난 뿐만 아니라 사회재난 뉴스를 정의하고 찾아낼 수 있는 방법을 고안하여 첫 번째 어려움을 극복하고자 하였다. 또한, 일부 뉴스에 대한

크롤링 방식의 획득이 아닌 언론사에서 생산되는 뉴스 전수 데이터를 확보함으로써 두 번째 어려움을 해결하고자 하였다. 이를 바탕으로 재난에 대한 사회적 관심도를 측정하는 방법을 고안하고, 사회적 관심도를 국가 전체 예산 중 재난 관련 예산과 비교하여 시사점을 얻고자 한다.

## 2. 연구방법

### 2.1 뉴스 데이터

본 연구에서는 한국언론진흥재단 30종 언론사 뉴스 데이터를 활용하였으며, 일간지 6종, 방송사 4종, 경제지 3종, 지역지 15종, 전문지 2종으로 구성되어 있다. 분석 대상은 2011년 3월 1일 ~ 2021년 12월 31일까지, 총 18,714,583건의 뉴스기사이다.

Table 1. List of target media in this study

Type	Media
Central media	Kookmin Ilbo, Naeil Newspaper, Seoul Newspaper, Segye Ilbo, Hankyoreh, Hankook Ilbo
Broadcast media	KBS, MBC, SBS, YTN
Economic media	Money Today, Seoul Economic, Korea Economic
Local media	Gangwon Ilbo, Gyeonggi Ilbo, Gyeongnam Provincial Daily, Kyongbuk Ilbo, Kwangju Ilbo, Daejeon Ilbo, Maeil Newspaper, Busan Ilbo, Ulsan Maeil, Incheon Ilbo, Jeonnam Ilbo, Jeonbuk Provincial Daily, Chungbuk Ilbo, Chungcheong Today, Halla Ilbo
Specialized media	Electronic Newspaper, Hwankyong Ilbo

### 2.2 재난안전예산 데이터

행정안전부는 2014년 세월호 사고 이후 재난 및 안전과 관련한 국가 예산을 체계적으로 관리하고자 「재난 및 안전관리 기본법」 제10조의2(재난 및 안전관리 사업예산의 사전협의 등)를 신설하여 재난안전예산 사전협의 제도를 운영하고 있다<sup>15)</sup>. 2021년 재난안전예산 사전협의 대상은 풍수해, 화재·폭발 등 43개의 재난 유형으로, 자연재난 9개, 사회재난 및 안전사고 27개, 공통 8개로 분류된다<sup>16)</sup>.

### 2.3 재난 유형

미국의 연방재난관리청(Federal Emergency Management Agency, FEMA)은 재난을 통상적으로 사망, 상해, 재산 피해를 가져오고 일상적인 절차나 정부의 자원으로는 관리할 수 없는 심각하고 규모가 큰 사건으로 정의하고 있다. 우리나라는 「재난 및 안전관리 기본법」을 통해

Table 2. Category by disaster type

Type	Disaster
Natural Disaster	Typhoon, Flood, Strong Wind, Wind&Wave, Tidal Wave, Heavy Snowfall, Cold Wave, Lightning, Drought, Heat Wave, Earthquake, Sandy Dust, Hypertrophied of Algae(Green Tide, Red Tide), Ebb&Flow, Landslide, Volcano, Collision of a Natural Space Object
Social Disaster	Fire, Explosion, Collapse, Traffic Accidents (Aviation&Marine Accidents, etc.), Radioactive&Chemical Accidents, Environmental Pollution(Water Pollution, Marine Pollution, Soil Pollution), Paralyzation of the National Infrastructure, Fine Dust, Infectious Disease, Contagious Animal Disease

재난이란 국민의 생명·신체·재산과 국가에 피해를 주거나 줄 수 있는 것으로 정의하고 자연재난, 사회재난으로 분류하고 있다. 자연재난은 태풍, 홍수, 호우, 강풍, 풍랑, 해일, 대설, 한파, 낙뢰, 가뭄, 폭염, 지진, 황사, 조류, 조수, 화산활동, 소행성·유성체 등 자연우주물체의 추락·충돌 그 밖에 이에 준하는 자연현상으로 인하여 발생하는 재해로 규정하고 있으며, 사회재난은 화재·붕괴·폭발·교통사고·화생방사고·환경오염사고 등으로 인하여 발생하는 일정 규모 이상의 피해와 국가핵심기반의 마비, 감염병, 가축전염병, 미세먼지 등으로 규정하고 있다. 본 연구에서는 한국의 「재난 및 안전관리 기본법」 상의 재난 유형을 준용하여 Table 2와 같이 자연재난 17종, 사회재난 10종으로 분류하여 분석하였다.

2.4 재난에 대한 사회적 관심지수 정의 및 측정방법

사회적 관심은 특정 사건이나 이슈에 대한 대중의 관심을 말한다. 특정 이슈에 관한 사회적 관심도는 언론 보도량으로 측정될 수 있으며, 이는 특정 주제에 관한 언론기사의 수로 계량화된다<sup>6)</sup>. 특정 이슈에 관한 언론 보도량은 대중의 관심과 밀접한 상관관계가 있기 때문에<sup>17)</sup>, 많은 선행연구에서도 사회적 관심을 대변하는 지표로 언론기사의 보도량을 주로 활용하고 있다<sup>6,18)</sup>. 따라서 본 연구에서는 재난이라는 특정 이슈에 대한 사회적 관심지수를 해당 기간동안의 전체 뉴스 중 특정 재난의 뉴스 건수의 비율로 정의하였다.

재난에 대한 사회적 관심지수는 n일이라는 기간 동안 생산된 총 뉴스 건수 중 재난 관련 뉴스 건수의 비율을 식1의 방법으로 측정하였다. 즉, 재난뉴스의 비율은 특정 기간동안 생산되는 뉴스량 대비 재난 관련 뉴스량의 백분율로 표현된다. 재난 관련 뉴스 건수는 2.3에서 제시한 자연재난과 사회재난 유형별 뉴스 건수 합이며, 이를 같은 기간 생산된 총 뉴스 건수 대비 백분율로 표현한 것이 재난에 대한 사회적 관심지수이다.

$$\text{사회적 관심지수} = \frac{\sum_{i=1}^n i\text{번째 날짜에 생산된 재난관련 뉴스 건수}}{\sum_{i=1}^n i\text{번째 날짜에 생산된 뉴스 총 건수}} \times 100 \quad (1)$$

3. 연구결과

3.1 재난뉴스의 분류체계 개발 및 식별

전체 뉴스에서 재난과 관련있는 단어가 뉴스 제목이나 본문에 포함될 경우 재난뉴스로 정의할 수 있으며, 자연재난과 사회재난을 식별할 수 있는 분류체계는 Table 3과 같다.

자연재난 관련 뉴스의 경우 비교적 쉽게 식별할 수 있다. 태풍, 홍수, 대설, 지진과 같은 단어들은 단어 자체가 재난현상을 의미한다. 따라서 뉴스 제목이나 본문 내 자연재난의 유형을 지칭하는 키워드의 유무에 따라 판별할 수 있다.

사회재난의 경우, 재난 유형에 따라 다양한 키워드 패턴을 활용해야 관련 뉴스를 판별할 수 있다.

첫 번째, ‘미세먼지’, ‘화재’, ‘폭발’은 그 자체로 재난 현상을 의미하므로, 자연재난과 같이 단일 단어로 관련 뉴스를 가려낼 수 있다.

두 번째는 ‘감염병’, ‘가축전염병’, ‘화생방사고’, ‘환경오염사고’ 유형이다. 이들 유형은 해당 재난을 일으키는 구체적인 단어(코로나19, 오미크론 등)와 확산을 표현하는 어휘(전염, 유출 등)가 조합되어 표현된다. 재난을 일으키는 구체적인 단어는 재난유형별 해당 법령에서 정의하는 용어로 정하였다. 감염병은 「감염병의 예방 및 관리에 관한 법률」, 가축감염병은 「가축전염병 예방법」의 법정 용어를 사용하였다. 화생방사고는 유해화학물질사고와 방사능사고를 포함하는 사고로, 유해화학물질은 「산업안전보건법」, 방사능은 「원자력시설 등의 방호 및 방사능 방재 대책법」에서 정의하는 물질을 주요 단어로 사용하였다. 환경오염사고는 수질오염, 해양오염, 토양오염을 포함하는 사고로, 수질오염물질은 「물환경보전법」, 해양오염물질은 「해양환경관리법」, 토양오염물질은 「토양환경보전법」에서 규정하는 오염물질을 주요 단어로 사용하였다.

세 번째는 ‘붕괴’, ‘교통사고’, ‘국가핵심기반 마비’ 유형이다. 이들 유형에는 해당 재난으로 피해를 입는 대상에 대한 단어(교량, 건물 등)와 피해 형태를 표현하는 어휘(붕괴, 화재 등)가 조합되어 표현된다. 특히 붕괴의 대상은 수많은 시설에 대한 단어가 될 수 있다. 본 연구에서는 270만여건의 뉴스 본문을 Word2Vec을 통해 학습시켜 구축한 워드벡터를 활용하여 건물, 교량

등의 유사어 1,138개를 사용하였다<sup>19)</sup>. 교통사고는 자동차사고, 지하철사고, 고속철도사고, 항공교통사고, 선박사고를 포함하였고, 국가핵심기반 마비는 에너지마비, 통신마비, 교통마비, 금융마비, 의료마비를 포함하였다.

Table 3. Search keywords by disaster type

Type	Disaster	Search keywords
Natural Disaster	Typhoon	Typhoon, hurricane, cyclone, etc.
	Flood	flood, heavy rain, etc.
	Strong Wind	Strong wind, tornado, whirlwind, etc.
	Wind&Wave	wind and wave, storm, etc.
	Tidal Wave	tidal wave, tsunami, etc.
	Heavy Snowfall	heavy snowfall, a lot of snow, etc.
	Cold Wave	cold wave, cold snap, cold spell, etc.
	Lightning	lightning, thunderbolt, thunderclap, etc.
	Drought	drought, dry season
	Heat Wave	heat wave, hyperthermia, etc.
	Earthquake	earthquake
	Sandy Dust	Sandy Dust, yellow sand, etc.
	Hypertrophied of Algae (Green · Red Tide)	green tide, water bloom, red tide, harmful algal bloom, etc.
	Ebb&Flow	ebb and flow
	Landslide	landslide, landslip, etc.
Volcano	volcano, volcanic ash, etc.	
Collision of a Natural Space Object	Collision of a Natural Space Object, flying saucers, etc.	
Social Disaster	Fire	fire, flame
	Explosion	explosion, explode
	Collapse	collapse, sink whole, destruction, etc. apartment, housing, school, market, etc.
	Traffic Accidents (Aviation&Marine Accidents, etc.)	car, traffic, subway, train, plane, etc. crash, accident, death, injure, etc.
	Radioactive&Chemical Accidents	radioactive, nuclear accident, chemical, noxious liquid, etc. contagion, infection, diagnosis, spill, spread, etc.
	Environmental Pollution (Water · Marine · Soil Pollution)	water pollution, marine pollution, soil pollution, etc. contagion, infection, diagnosis, spill, spread, etc.
	Paralyzation of the National Infrastructure	energy, network, traffic, medical, etc. paralysis, block, disturbance, etc.
	Fine Dust	fine dust
	Infectious Disease	Infectious Disease, MERS, cholera, etc. contagion, infection, diagnosis, spread, etc.
	Contagious Animal Disease	animal disease, avian influenza, etc. contagion, infection, diagnosis, spread, etc.

### 3.2 재난에 대한 사회적 관심지수 산출

2011년 3월부터 2021년 12월까지 생산된 전체 뉴스 건수는 18,714,583건이며, 이 중 재난 관련 뉴스는 539,755건으로 지수는 2.9%의 값을 가진다(Fig. 1).

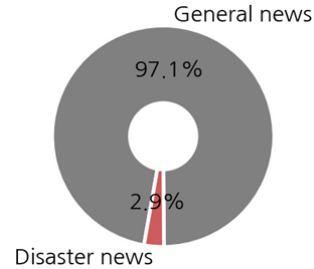


Fig. 1. Social interest index of total news(%).

자연재난과 사회재난을 나누어 살펴보면 자연재난의 관심지수는 0.99%, 사회재난의 관심지수는 1.89%로 사회재난에 대한 관심지수가 더 높게 나타났다.

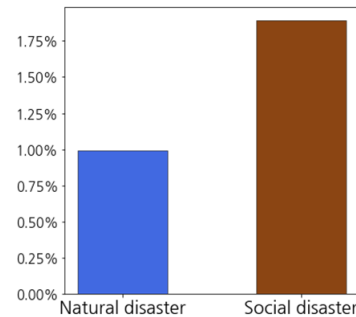


Fig. 2. Social interest index of both natural and social disaster(%).

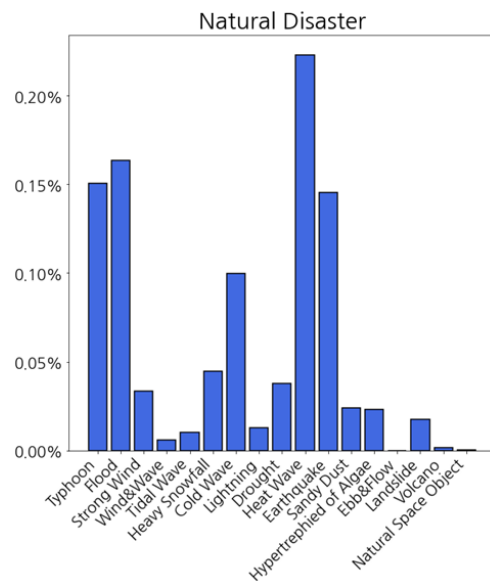


Fig. 3. Social interest index of by type of natural disaster(%).

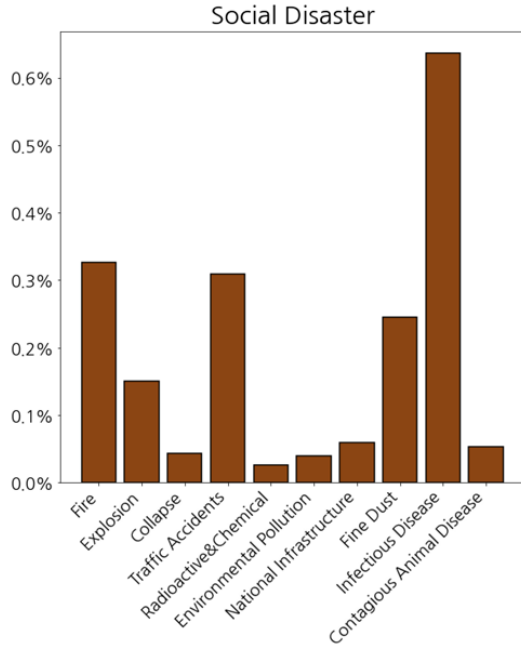


Fig. 4. Social interest index of by type of social disaster(%).

자연재난에 대한 관심 지수를 개별 재난유형별로 분리해서 살펴보면, 폭염이 0.22%로 가장 높게 나타났으며, 그 다음으로 홍수(0.16%), 태풍(0.15%), 지진(0.15%), 한파(0.10%) 순으로 높았다. 대설(0.04%), 가뭄(0.04%), 강풍(0.03%), 황사(0.02%), 조류(0.02%), 산사태(0.02%)는 0.01%~0.05% 사이의 값을 가졌으며, 풍랑, 해일, 낙뢰, 조수, 화산, 우주물체 추락은 0.01% 이하로 지수가 매우 낮았다(Fig 3).

사회재난에 대한 관심 지수를 개별 재난유형별로 분리해서 살펴보면, 감염병이 0.64%로 가장 높았으며, 그 다음으로 화재(0.33%), 교통사고(0.31%), 미세먼지(0.25%), 폭발(0.15%) 순으로 높았다. 그 밖에 국가핵심기반 마비(0.06%), 붕괴(0.04%), 가축전염병(0.05%), 환경오염(0.04%), 화재방사고(0.03%)는 상대적으로 매우 낮았다(Fig 4).

자연재난과 사회재난을 합친 전체 재난에 대한 사회적 관심지수를 연도별로 살펴보면 분석기간 초기 2011~2013년에는 약 0.16~1.7% 수준이었다. 2014년에

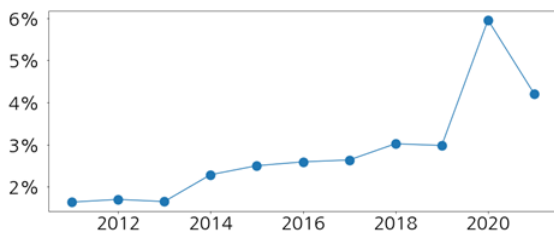


Fig. 5. Social interest index of total disaster by year(%).

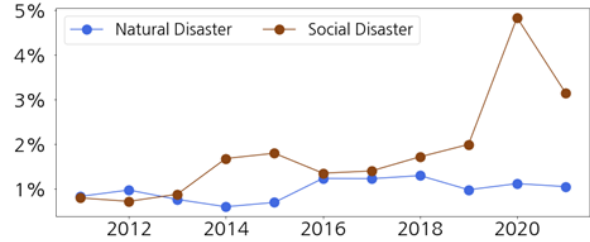


Fig. 6. Social interest index of both natural and social disaster by year(%).

2.29% 수준으로 높아졌으며, 이후 2019년까지 2.98%까지 지속적으로 상승하였다. 2020년에 들어 5.96%까지 급증하였고 2021년에 4.20%로 다소 감소하였다(Fig 5).

자연재난과 사회재난을 분리해서 살펴보면(Fig 6) 2011~2013년에는 자연재난과 사회재난에 대한 지수가 모두 0.8% 수준으로 서로 유사하였으나, 이후에는 모든 연도에서 사회재난에 대한 관심 지수가 자연재난 보다 더 높게 나타난다. 앞서 살펴본 2014년부터 2020년까지 전체 재난에 대한 관심 지수의 지속적인 상승은 2014~2015년은 사회재난, 2016~2018년은 자연재난, 2019~2021년은 사회재난에서 기여가 크음을 확인할 수 있다. 2014~2015년의 사회재난 이슈를 연도별로 살펴보면 2014년에는 4·16 세월호 참사로 인해 교통사고(선박사고)가 높게 나타났고, 2015년에는 메르스 발생으로 인해 감염병의 관심 지수가 높게 나타났다. 2016~2018년의 자연재난 이슈를 연도별로 살펴보면 2016년과 2018년에는 온열질환으로 인한 사망자가 81명, 162명<sup>20)</sup> 발생하면서 폭염 이슈가 높게 나타났고, 2016년 경주지진, 2017년 포항지진 발생으로 지진에 대한 관심 지수도 높게 나타났다. 또한 2017년은 7월 충북지역에 250 mm 이상의 기록적인 폭우로 인해 피해가 발생하면서 홍수 관련 관심 지수가 높게 나타났다. 2019~2022년의 사회재난 이슈를 연도별로 살펴보면 2019년에는 미세먼지, 2020년 이후로는 코로나19로 인한 감염병이 크게 기여한 것으로 확인된다.

### 3.3 재난에 대한 사회적 관심지수와 재난안전예산의 비교

「재난 및 안전관리 기본법」 제10조의2(재난 및 안전관리 사업예산의 사전협의 등)에 의해 ‘재난안전예산 사전협의’ 제도가 시행되고 있다. 이 제도는 재난 및 안전관리 사업 예산의 체계적 관리를 위해 2015년부터 시행되고 있다. 행정안전부는 재난안전예산의 효율적 운용과 투자 방향성 제시를 위해, 3가지 지표(피해통계, 사회적 이슈강도, 전문가 향후 위험 전망)를 종합하여 43개 재난 및 안전사고 유형 중 중점투자유형을 선정한다<sup>15)</sup>. 본 연구에서 측정한 사회적 관심지수는 3가지 지표 중 사회적 이슈강도로서 투자 우선순위에

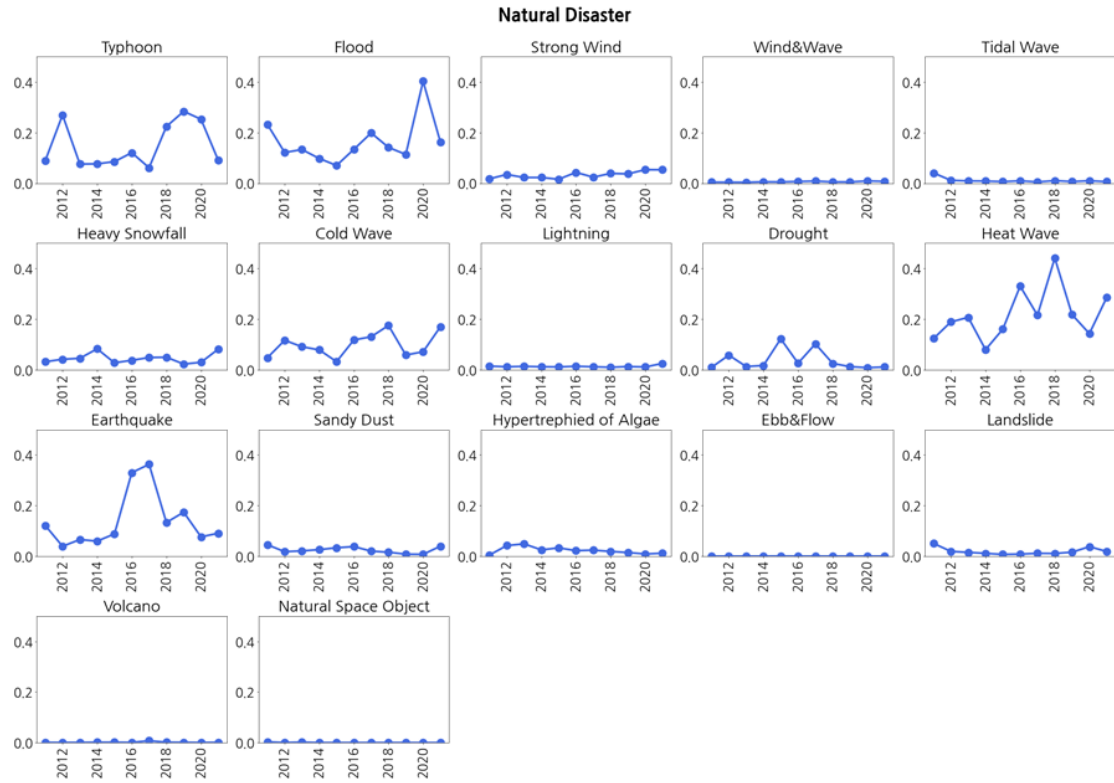


Fig. 7. Social interest index by type of natural disaster by year.

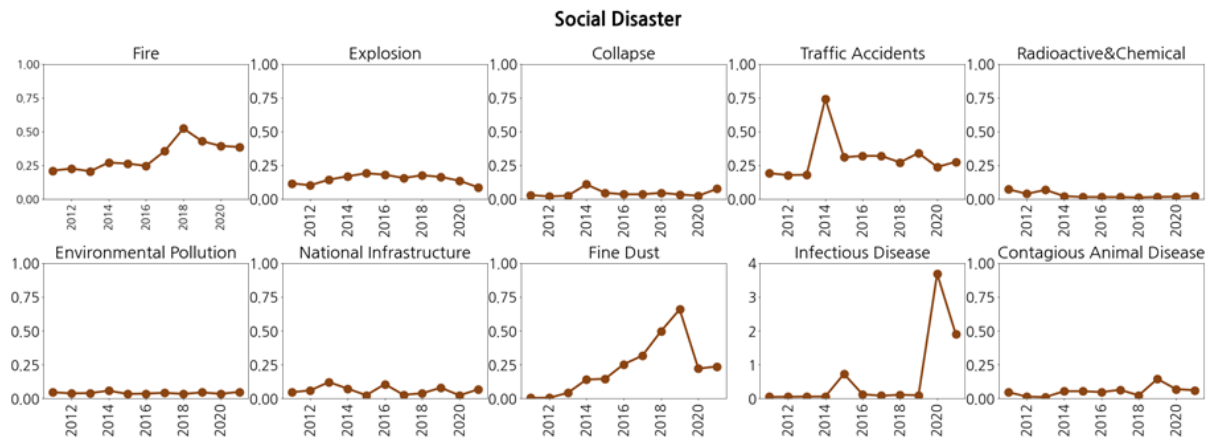


Fig. 8. Social interest index by type of social disaster by year.

반영된다. 따라서 재난에 대한 사회적 관심지수와 재난안전예산을 비교함으로써, 사회적 관심은 높지만 재난안전예산 편성은 미흡한 재난유형을 발견하는 등 재난관리 체계 개선을 위한 시사점도 발견할 수 있을 것이라 판단된다.

2019년부터 2021년까지 3년간 대한민국 중앙정부의 전체 예산과 재난 관련 예산을 Fig. 9에 제시하였다. 3년간 연평균 대한민국 전체 예산은 513.3조원이며, 이 중 재난과 직접적으로 관련 있는 예산은 연평균 18조

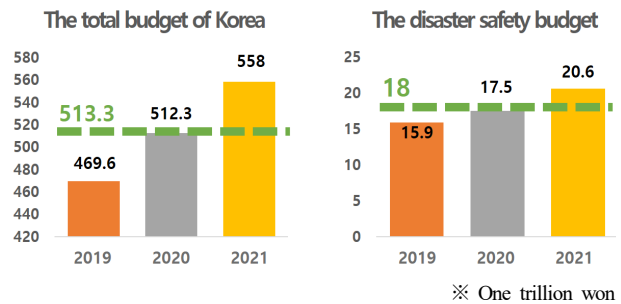


Fig. 9. National budget from 2019 to 2021.

원으로 3.5% 수준이다. 이는 2011년부터 2022년까지의 재난에 대한 관심 지수 2.9%와 규모적으로 상당히 유사하다. 다만 사회적 관심지수는 과거에 비해 급격하게 증가하였지만, 정부 예산은 그 정도의 급격한 증가는 나타나지 않았다. 그러나 정부 예산도 시간에 따라 지속적으로 증가하고 있음이 확인된다.

재난안전예산에서 분류한 재난유형 중 사회적 관심 지수 산출 시 분류한 재난유형과 비교하여 분석이 가능한 재난은 가뭄, 대설·한파, 산사태, 조류, 지진, 폭염, 풍수해(태풍, 홍수), 황사, 가축전염병, 감염병, 교통사고, 미세먼지, 방사능, 수질오염, 유해화학물질 사고, 해양오염, 화재·폭발, 국가기반마비로 총 18개이다.

18개의 재난유형에 대해 재난안전예산과 사회적 관심지수를 비교한 산점도는 Fig 10과 같다. x축은 재난안전예산이며, y축은 사회적 관심지수로 두 축 모두 로그스케일로 표현하였다. 대부분의 재난유형은 두 변수가 선형관계식에 가깝게 위치하였다. 그러나 대설·한파와 폭염의 경우 사회적 관심지수는 높으나 상대적으로 재난안전예산은 적게 나타났다.

18개의 재난에 대한 사회적 관심지수 대비 재난안전 예산의 비율을 시각화한 상자그림(Box Plot)은 Fig. 11과 같다. 상자그림에서 음영 상자의 아랫부분은 1사

분위수(Q1), 윗부분은 3사분위수(Q3)를 의미하며, 상자로부터 뻗어 나온 수평선은 양 극면으로부터 3사분위수(Q3)와 1사분위수(Q1)의 차이(Q3-Q1)의 1.5배의 범위를 의미하는 펜스이다<sup>21)</sup>. 위·아래 펜스 너머에 있는 값들은 이상값으로 간주한다. 본 연구에서 1사분위수(Q1)의 값은 0.49이며, 3사분위수(Q3)의 값은 0.69이다. 이를 통해 산출한 펜스의 범위는 0.38~0.83으로 이 값을 벗어난 관측치는 이상값으로 간주된다. 따라서 본 연구에서는 대설·한파와 폭염이 이상값으로 나타났다.

산점도와 박스플롯의 결과를 바탕으로 가능한 원인을 유추해보면 다음과 같다. 한파와 폭염의 경우 다른 재난에 비해 시설물 피해의 규모가 작을 수 있다. 또한 한파와 폭염은 2018년에 「재난 및 안전관리 기본법」의 자연재난으로 포함되면서, 재난 유형 중 가장 늦게 명시되었다. 따라서 사회적 관심도는 높지만 상대적으로 재난안전예산이 적게 나타난 것은 다른 재난에 비해 재산피해의 규모가 작고, 국가에서 재난으로 관리를 시작한지 얼마 되지 않아 예산 편성의 규모가 작을 수 있다는 것이 또 다른 원인 중 하나가 될 수 있다고 생각된다.

#### 4. 결론

본 연구에서는 뉴스데이터를 활용하여 재난안전에 대한 사회적 관심을 측정하기 위해 다음과 같이 연구를 진행하였다. 먼저, 자연재난과 사회재난을 식별할 수 있는 분류체계를 개발하였으며, 이 분류체계를 통해 재난 유형별 사회적 관심지수를 측정하였다. 또한 재난에 대한 사회적 관심지수와 재난안전 예산을 비교하였다. 이러한 결과를 통해 도출한 결론은 다음과 같다.

첫째, 자연재난의 경우 태풍, 홍수, 대설 등 재난 현상을 의미하는 단어만으로도 뉴스를 식별할 수 있었다. 그러나 사회재난의 경우 3가지 방법으로 식별할 수 있었다. 먼저 미세먼지, 화재, 폭발은 자연재난과 마찬가지로 재난 현상을 의미하는 단어만으로도 뉴스를 식별할 수 있었으며, 감염병과 가축전염병, 화재방사고, 환경오염사고는 재난의 구체적인 명사와 확산을 표현하는 어휘를 조합하여 탐색할 수 있었다. 붕괴, 교통사고, 국가핵심기반 마비는 피해를 입는 대상과 피해 형태를 표현하는 어휘를 조합하여 표현된다. 이러한 분류체계를 통해 뉴스데이터에서 재난뉴스를 식별할 수 있다.

둘째, 재난유형별 분류체계를 통해 사회적 관심지수를 산출한 결과, 최근에는 자연재난보다 사회재난에 대한 관심지수가 더 높게 나타났으며, 연도별로 살펴 보았을 때 2020년도까지 지속적으로 증가하고 있다.

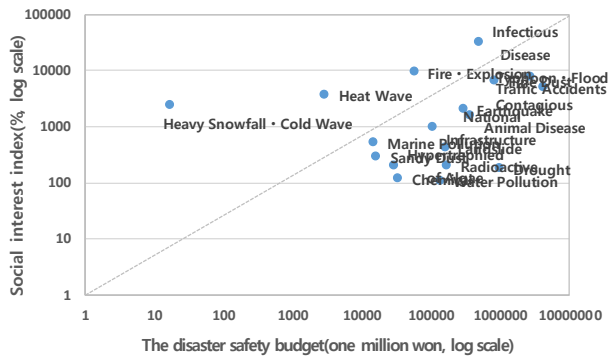


Fig. 10. A Scatter plot of social interest index and disaster safety budget by disaster type.

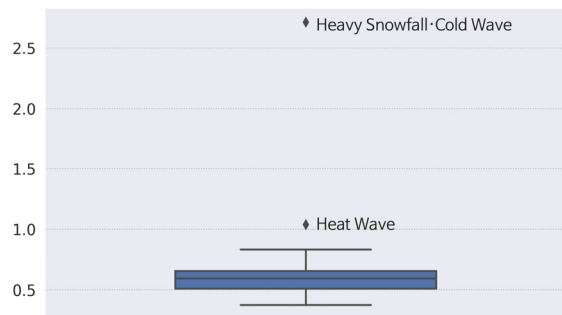


Fig. 11. A boxplot of social interest index and disaster safety budget by disaster type.

이는 2011~2013년까지는 자연재난과 사회재난에 대한 관심지수가 유사하였으나, 그 이후로 4·16 세월호참사, 메르스, 미세먼지, 코로나19 등으로 인해 사회재난의 관심지수가 높게 나타난 것을 알 수 있었다.

셋째, 재난안전예산을 한국의 전체예산과 비교한 결과 약 3.5%로, 이는 본 연구에서 측정한 사회적 관심지수와 매우 유사한 수치로 나타났다. 그러나 대설·한파와 폭염의 경우 사회적 관심지수는 높으나 상대적으로 해당 재난의 예산이 적어 이상값으로 나타났다. 이는 「재난 및 안전관리 기본법」의 재난 유형 중 폭염과 한파가 가장 늦게 명시되어 다른 재난에 비해 예산 규모가 작을 수 있다는 것이 한가지 원인으로 생각된다.

본 연구에서 개발한 재난유형별 분류체계는 뉴스데이터에서 자연재난, 사회재난 각각의 재난유형에 대해 식별할 수 있으며, 이를 통해 사회적 관심도를 측정하여 국가의 재난관리 정책에 반영할 수 있을 것으로 기대된다.

본 연구에서는 온라인상에서 크롤링을 통해 수집할 경우 일정 샘플 데이터로 분석하는 한계점을 극복하고자 한국언론진흥재단의 언론사 뉴스 전수 데이터를 받아 분석을 진행하였다. 그러나 예산상의 한계로 30종 언론사에 대해서만 분석하였으며, 한국언론진흥재단에서 제공하는 언론사는 총 104개(2023년 2월 기준), 문화체육관광부에서 제공하는 정기간행물 등록 현황에서는 일반일간신문의 언론사가 총 602개로 본 연구에서 분석했던 30종 언론사만으로는 전체 뉴스를 대변하기에는 한계가 존재한다. 그러나 본 연구에서 활용한 30종 언론사의 경우, 언론사의 보도 성향에 따라 편향된 결과를 도출하지 않도록 언론진흥재단에서 제공하는 일간지를 모두 확보하였으며, 지역지 또한 16개 시도에 대해 구독수가 높은 언론사를 선정하여 대표성을 확보하고자 하였다.

언론사 유형별로 전체 뉴스 중 재난 관련 뉴스 보도량을 비교해보았을 때, 일간지, 경제지, 지역지, 전문지는 평균 2.4% 정도였으며 방송사는 평균 8%의 비율로 재난뉴스를 보도하였다. 특히 방송사 중 KBS는 9.9%의 비율로 재난뉴스를 보도하였는데, KBS는 「방송통신발전 기본법」 제40조의2(재난방송 등의 주관방송사)에 의해 재난방송 등의 주관 방송사이다. 이처럼 재난에 대한 보도가 일간지, 방송사, 지역지 등 언론사 유형별 보도 특성에 따라 달라질 수 있는 점을 고려하여 추후 연구가 진행될 필요가 있다.

재난안전예산 사전협의 제도(재난 및 안전관리 기본법 제10조의2)와 관련하여 행정안전부는 43개 재난 및 안전사고 유형별 투자우선순위를 매년 정하고 있다.

재난유형별 투자우선순위는 총 3가지 지표(피해통계, 사회적 이슈강도, 전문가 향후 위험 전망)를 종합하여 정해지며, 본 연구에서 개발한 재난유형별 사회적 관심지수가 사회적 이슈강도로서 투자우선순위에 반영된다.

현재는 재난유형에 대한 사회적 관심지수가 언론 데이터만을 활용하여 산출되지만, 지속적인 연구를 통해 국민이 직접 생산하는 SNS 데이터 등 더 다양한 사회적 주체의 관심도가 반영된다면 향후 재난안전예산 편성 시 국민이 체감할 수 있는 국민 중심의 중점투자 유형 선정과 재난안전예산의 효율적 운용의 향상을 기대할 수 있을것이라 판단된다

**Acknowledgement:** The study financially supported by the ‘National Disaster Management Research Institute’, Ministry of the Interior and Safety(NDMI-Main-2023-04-02).

## References

- 1) D. Nohrstedt, “When do Disasters Spark Transformative Policy Change and Why?”, *Policy & Politics*, Vol. 50, No. 3, pp. 425-441, 2022.
- 2) C. S. Gwak and K. S. Moon, “Influence of Risk Perception of Fine Particles on Protection Behaviors: Focused on Mediating Effects of Risk Attitude”, *J. Korean Soc. Saf.*, Vol. 36, No. 6, pp. 48-55, 2021.
- 3) Y. S. Kim and D. W. Kim “The Study on Development of Crime Safety Indicators: Focused on Application of Social Indicators in Korea”, *Korean Criminological Review*, Vol. 24, No. 4, pp. 295-328, 2013.
- 4) J. Y. Lee, C. S. Shin, M. G. Suh and D. Y. Park, “A Study on the Public Safety Awareness for the Development of Social Safety Indicators”, 2005.
- 5) M. J. Robinson, “The News Interest Index, 1986-2007: Two Decades of American News Preferences”, *Pew Research Center for the People & The Press*, 2007.
- 6) S. D. Yoo and J. S. Lee, “Media Interest and Policy Decision of Child Welfare Issue: An Empirical Study on the Downs’ Issue-Attention Cycle”, *Journal of the Government Administration*, Vol. 13, pp. 29-60, 2017.
- 7) C. M. Kim and C. I. Choi, “An Exploration of Issue Attention Cycle of Great Disasters”, *Seoul Studies*, Vol. 16, No. 4, pp. 147-162, 2015.
- 8) H. S. Yoo, H. J. Kim and H. J. Oh, “Analysis of Issue Life Cycle Types of Disaster Events”, *Journal of Korean*



- Institute of Information Technology, Vol. 16, No. 3, pp. 125-134, 2018.
- 9) W. Y. Cho and D. G. Lee, “An Exploratory Research based on Keyword Network Analysis about Natural Disaster Issue of Busan: Focusing on the News Articles of KOOKJE Newspapers of Busan Media(2016~2018)”, Vol. 3, No. 2, pp. 201-228, 2018.
  - 10) J. E. Kim, “An Analysis of Heat Wave Disaster Issue through Social Media Big Data”, The Korea Policy Journal, Vol. 22, No. 1, pp. 50-60, 2022.
  - 11) T. Nugent, F. Petroni, N. Raman, L. Carstens and J. L. Leidner, “A Comparison of Classification Models for Natural Disaster and Critical Event Detection from News“ In 2017 IEEE International Conference on Big Data (Big Data), IEEE, pp. 3750-3759, 2017.
  - 12) Téllez Valero, A., Montes y Gómez, M., and Villaseñor Pineda, L., “Using Machine Learning for Extracting Information from Natural Disaster News Reports”, Computación y Sistemas, Vol. 13, No. 1, pp. 33-44, 2009.
  - 13) Y. Y. Na, J. H. Park and J. M. Chol, “Analysis of the Characteristics of a Complex Disaster Using Bigdata”, The Geographical Journal of Korea, Vol. 53, No. 4, pp. 375-384, 2019.
  - 14) D. S. Hur, J. E. Yoo, S. J. Jeung and S. K. Jung, “A Study on Urban Flood Damage Topic Model through News Article Text Mining and Latent Dirichlet Allocation”, Journal of the Korean Data & Information Science Society, Vol. 34, No. 2, pp. 315-330, 2023.
  - 15) H. U. Cho and S. Y. Choi, “Investment Priorities of the Budgets for Disaster and Safety Management Projects Based on Euclidean Distance with AHP”, Journal of The Korean Society of Hazard Mitigation, Vol. 21, No. 3, pp. 23-37, 2021.
  - 16) Ministry of the Interior and Safety(MOIS), “Report of 2021 Prior Consultation for Disaster and Safety Management Budget”, 2020.
  - 17) K. M. Park, S. C. Kim, “Type of Life Cycle of Regulatory Reform Issue and Performance of Regulatory Reform: From Kim Dae-Jung to Park Geun-hye Administration”, The Korean Governance Review, Vol. 27, No. 2, pp. 83-117, 2020.
  - 18) T. J. Lee, “A Comparative Assessment of the Social Attention to Public Issues Amplified by the Press and Digital Discourse”, Korean Public Administration Review, Vol. 50, No. 2, pp. 215-241, 2016.
  - 19) D. W. Kim, Y. S. Kim, B. G. Bae and E. H. Shin, “Development of Knowledge base and its Application Technologies for Disaster Management( I)”, National Disaster Management Research Institute, 2022.
  - 20) W. I. Cho, “The Government Changes the Heatwave Death Toll Standards, It’s Four Times More than before”, NEWSTAPA, 2021.
  - 21) I. S. Min and P. S. Choi, “STATA Introduction and Practical Applications”, JIPHIL Media, pp. 204-2012, 2021.