

소셜미디어를 통해 본 재난안전 분야 어휘 사용 양상 분석

A Study on the Analysis of Disaster Safety Lexicon Patterns in Social Media

김태영, 이정은, 오효정
전북대학교 기록관리학과

Tae-Young Kim(fnty127@hanmail.net), Jung-Eun Lee(lepina@naver.com),
Hyo-Jung Oh(ohj@jbnu.ac.kr)

요약

재난안전 분야 어휘의 표준화는 성공적인 재난안전사고 예방 및 대응을 위해 가장 기본적인 과정으로서 중요하다. 재난안전 분야의 어휘에 대한 이해 부족은 커뮤니케이션 및 정보공유의 부재로 이어지며, 이는 재난사고 발생 시 적절한 대응을 위한 의사소통에 문제가 될 수 있다. 현재 재난안전 유관기관별로 다양한 재난안전정보가 생산 및 관리되고 있으며, 정보공유를 위해 각 기관에서는 개별적으로 용어사전을 개발하여 활용하고 있다. 따라서 이용자에 따른 재난안전 분야의 어휘 사용 양상의 차이를 실제적으로 파악하는 것은 표준화를 위해 필수적이다. 이에 본 연구는 재난안전 분야의 어휘 사용 양상을 소셜미디어를 중심으로 분석하여 그 차이를 규명하였다. 구체적으로는 재난안전 분야에서 일반 이용자가 사용하는 어휘와 기존 재난안전 유관기관에서 활용하고 있는 어휘자원의 차이점을 비교·분석하였다. 이후 분석 결과를 기반으로 재난안전 분야 표준화 방안을 제안함으로써 용어사전 구축 방향성을 수립하였다.

■ 중심어 : | 재난안전 | 어휘자원 | 소셜미디어 | 용어사전 | 표준화 |

Abstract

Standardization of disaster safety lexicon is important as the most basic process for successful accident prevention and response. A lack of understanding of disaster safety lexicon leads lack of communication and information sharing, which can be a problem in communicating with appropriate responses in case of a disaster. Currently disaster and safety control agencies produce and manage heterogeneous information and they also develop and use word dictionaries individually. To solve this problem, identifying differences of disaster safety lexicon patterns by the user are essential for standardization. In this paper, we conducted lexicon patterns analysis based on social media and revealed the characteristics according to pattern types. At the result, we proposed the standardization and construction methods of disaster safety word dictionary.

■ keyword : | Disaster Safety | Lexicon Resource | Social Media | Word Dictionary | Standardization |

1. 서론

최근 빈번한 재난안전사고의 발생으로 인해 재난안

전 예방 및 대응 활동에 대한 관심이 급부상하고 있다. 이 중 다양한 재난안전 유관기관으로부터 생산·관리 되는 재난안전정보의 공유는 예방에서 복구에 이르기

* 본 연구는 2017년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(과제번호: NRF-2016M3D7A1912703).

접수일자 : 2017년 07월 07일

수정일자 : 2017년 09월 05일

심사완료일 : 2017년 09월 08일

교신저자 : 오효정, e-mail : ohj@jbnu.ac.kr

까지 재난관리단계의 전 과정을 효율적이고 신속하게 이루어지게끔 한다. 현재 재난안전 유관기관별로 2,500여 종의 재난안전정보가 생산 및 관리되고 있으며, 정보 공유를 위해 각 기관에서는 개별적으로 용어사전을 개발하여 활용하고 있다. 그러나 기관별로 각기 다른 용어와 정의를 사용함에 따라 발생하는 재난안전 용어의 다의성 및 의미적 모호성은 재난안전정보를 이용하려는 이용자 입장에서 큰 걸림돌이다[1]. 따라서 각 재난안전 유관기관에서 사용하고 있는 용어를 표준화하는 것은 재난안전정보 공유에 앞서 해결해야 하는 시급한 문제이다.

그 중에서도 표준화를 위해서는 전문용어와 일반용어의 괴리감을 파악하여 어휘를 순화하는 작업이 선행되어야 한다. 재난안전 분야의 어휘는 특정 재난 분야의 업무 실무자들이 사용하는 전문용어와 각 유관기관에서 중복적으로 활용되거나 일반인이 정보 검색 시 사용하는 일반용어가 혼재되어 있다. 과거 전문용어는 일반 이용자가 사용하지 않는 제한된 언어였으나, 최근 웹의 발전으로 일반 이용자도 전문 정보에 손쉽게 접근할 수 있게 되었다. 이에 전문용어는 더 이상 전문가만의 언어가 아니게 되었으며, 일반 이용자에게도 지식 전달력을 가질 수 있어야 한다. 이를 위해 전문용어는 소리를 들어 의미를 바로 파악할 수 있는 의미 환기성이 뛰어난 용어로 순화되어야 한다[2][3]. 실례로 해양수산부는 2015년부터 진행된 용어순화 프로젝트를 통해 일반 이용자가 이해하기 어려운 해양수산 분야의 용어 중 사용빈도가 높은 용어를 선정하여 일반인이 쉽게 이해할 수 있도록 순화어를 마련하고자 하였다[4].

이에 본 연구는 소셜미디어를 중심으로 재난안전 분야에서 일반 이용자가 사용하는 어휘와 기존 재난안전 유관기관에서 정의, 활용하고 있는 어휘자원 간의 차이점을 비교·분석하였다. 이후 분석 결과를 기반으로 재난안전 분야 어휘의 표준화 방안을 제안함으로써 용어사전 구축 방향성을 수립하였다.

II. 관련 연구

원활한 의사소통 및 업무의 효율성 향상을 위해 전문

용어사전을 구축하고자 하는 연구는 다양한 분야에서 진행되어져 왔다[5-7]. 그 중에서도 최중환 외(2011)는 국방과학기술분야의 전문용어 사전 구축에 대한 연구를 진행했으며, 사전 구축 프로세스 및 조직 구성, 표제어 정의, 전문용어 추출을 위한 대상 문서 선정, 전문용어 추출, 전문용어 후보군 생성, 위크벤치 등록, 사전 구축, 사전 검증 과정 등 전문 용어사전 구축을 위한 구체적인 사항을 제시하고 있다[7].

전문용어와 실제 이용자의 용어 사용 양상을 다룬 연구도 있다. 송영빈(2012)은 의학 분야에서 다섯 차체 이상 개정된 ‘의학용어집’ 과 실제 의학논문집에 사용된 전문용어를 네 가지 분류에 기초하여 비교·분석하였다. 이를 통해 전문용어가 현재 사용되는 양상을 파악함으로써, 의학용어집의 문제점을 밝히고자 하였다[3].

용어들 간의 상관관계 분석을 통하여 핵심용어를 추출하고자 하는 연구도 있다[8][9]. 박별나 외(2010)는 교과서에 사용된 과학 용어의 네트워크를 구축하여 네트워크 구조 및 연관 관계 분석을 위한 핵심용어를 추출하였다. 해당 연구는 복잡한 네트워크로부터 핵심 네트워크를 구축하여 유용한 정보를 도출할 수 있는 가능성을 제시한 연구이다[8]. 재난안전 분야에서도 신속한 재난 대응을 위해 기관 간 협업기능을 향상시키고자 재난안전 관련 핵심용어를 분석한 연구가 진행된 바 있다. 강성경 외(2016)는 재난현장에서 사용하는 용어들을 선정하고 텍스트 마이닝 및 소셜 네트워크 분석을 이용하여 핵심용어인지를 파악하였다. 구체적으로는 신문기사와 사설, 정책보고서 등의 정부 간행물을 검토한 후 재난관리단계 중 ‘재난 대응’에 해당되는 용어를 선정하였으며, 이 중 핵심용어를 용어 간 근접 중심성 계산을 통해 파악하였다[9].

가장 최근의 연구로는 재난안전 유관기관에서 관리하고 있는 어휘자원의 현황을 분석한 연구가 있으며, 본 연구의 선행연구이다. 각 기관에서 수집된 어휘자원의 중복도를 기반으로 네이버뉴스 검색 빈도를 통한 정보제공자 관점, 구글트랜드 검색 빈도를 통한 정보이용자 관점에서의 어휘자원 활용도를 분석하였다[1].

III. 연구내용 및 방법

본 연구는 소셜미디어에서 추출된 재난안전 관련 어휘와 재난안전 유관기관에서 사용하는 어휘와의 비교·분석을 통해 실제 사용 양상을 파악하였다. 이를 통해 재난안전 분야 용어 표준화 방향성 및 활용방안을 제시하고자 한다.

1. 재난안전 분야 어휘자원 수집

본 연구는 선행연구[1]를 통해 수집된 재난안전 분야 어휘자원[표 1]을 분석대상으로 중복도 분석을 수행하여 중복 횟수가 2회 이상인 어휘들을 추출하였다. 어휘 자원의 중복은 서로 다른 출처에서 동일한 표제어가 등장했다는 것을 의미하며, 1개의 표제어가 2개의 어휘자원 출처에서 등장했다면 중복 횟수는 2회이다. 수집된 어휘자원 총 114,782건 중 11,336건(약 9.9%)이 2회 이상 중복되었으며, 최대 13회까지 중복되었다. 이를 통해 수집된 어휘자원 중 보다 자주 활용되는 어휘를 추출할 수 있었으며, 추출된 어휘를 트위터(Twitter)로부터 추출된 어휘와 비교·분석함으로써 실제 사용 양상을 분석하였다.

표 1. 재난안전 분야 어휘자원 수집현황[1]

대상	용어 사전	국가 표준	단행본	법령	매뉴얼	합계
건수	33	4	3	300	48	388
어휘 수	109,278	138	2,725	2,053	588	114,782

2017.04.30. 기준

2. 재난안전 분야 트윗 수집 및 어휘 추출

실제 일반 사용자들이 활용하고 있는 어휘 양상을 파악하기 위해 본 연구에서는 다양한 연령의 이용자들이 자발적으로 작성하는 소셜미디어를 분석대상으로 선정하였다. 그 중에서도 특히 오픈 API를 통해 데이터 수집이 용이한 트위터를 대상으로 재난관련 트윗(Tweet)을 수집하였다. 재난관련 트윗 수집을 위해 자연 및 사회재난의 사건을 선정하였고, 사건명을 검색 키워드로 하였다. 재난 사건의 선정 기준은 국민안전처에서 2015년에 발행한 재난연감과 재해연보 통계자료이다

[10][11]. 통계자료로부터 최근 10년 간 발생한 ‘대형’ 자연 및 사회재난을 선정[표 2]하였다. ‘대형’ 재난 사건은 트위터 이용자의 관심도가 높아 분석 데이터 수집이 용이하기 때문에 선정하였다.

표 2. 트위터 어휘 추출 현황

재난 유형	재난사건	트윗수	어휘 빈도
자연 재난	태풍 에위니아	44	578
	태풍 나리	584	6,480
	태풍 곤파스	5,330	64,011
	태풍 무이파	2,625	28,641
	태풍 덴빈	6,550	74,563
	태풍 불라벤	22,593	275,359
	태풍 산바	1,945	26,411
	강풍 풍랑	941	17,806
	호우-서울	3,786	54,914
	호우-경기	2,246	37,870
	호우-전남	663	11,631
	호우-전북	533	10,107
	호우-충남	784	14,662
	호우-충북	271	5,217
	호우-경남	804	14,308
	호우-경북	536	10,589
	호우-강원	1,212	22,872
	호우-부산	1,266	15,503
	대설	대설-경북	456
대설-전남		102	1,908
대설-강원		1,312	22,073
사회 재난	고양종합터미널 지하공사장 화재사고	1,534	21,921
	담양대덕 펜션 화재사고	576	7,803
	남시어선 돌고래호 전북 사고	2,325	27,142
	경주 마우나오션리조트 체육관 붕괴사고	4,416	63,760
	중동 호흡기 증후군(메르스)	271,449	2,933,383
	상왕십리역 열차추돌사고	2,625	34,714
	신증 인플루엔자	31,728	286,122
	원양어선 제501오룡호 침몰사고	2,282	25,994
	의정부시 대봉그린아파트 화재	1,091	15,097
	장성 효사랑요양병원 화재사고	823	13,277
	화물선 청루15호(파나마)침몰	165	2,309
	용흥지구 산불피해	2,516	27,197
	합계	376,113	4,182,842 (어휘 수) 43,903 (중복제거)

수집된 총 376,113건의 트윗으로부터 4,182,842개의 어휘들이 추출, 각각의 트윗에는 평균 11개의 어휘로 구성되어 있음을 확인하였다. 추출된 어휘의 중복제거 결과 43,903개로 분석되었다. 어휘 추출을 위해 한글 형태소 분석기인 “RHINO” 코드를 활용[12], R[13]을 통해 트윗 내용을 형태소 분석하여 모든 명사형 어미를 추출하였다.

3. 어휘자원과 트위터 어휘 비교 분석

본 연구에서는 추출된 재난안전 분야 어휘자원에 나타난 어휘와 트위터에 나타난 어휘의 중복도를 비교·분석하였다(표 3). 2회 이상 중복된 재난안전 분야 어휘 총 11,336건 중 2,751건(약 24%)이 트윗 어휘와 매핑(Mapping)되었다. 3회 이상 중복된 재난안전 분야 어휘 총 3,254건 중 1,287건(약 39%)이 트윗 어휘와 매핑되었다. 분석 결과를 통해 재난안전 분야에서 재난안전 유관기관이 사용하는 어휘와 일반인이 사용하는 어휘에는 다소 차이가 있음을 확인할 수 있었다. 다음 절에는 이러한 차이에 따른 어휘별 특성을 파악하기 위해, 3회 이상 중복된 재난안전 분야 어휘와 트윗 어휘의 매핑 결과를 유형별로 자세하게 분석하였다.

표 3. 어휘 비교 분석 결과

중복 횟수	용어사전중복 표제어 수	트윗 어휘와 중복되는 표제어 수	중복 비율 (%)
2회 이상	11,336	2,751	24
3회 이상	3,254	1,287	39

IV. 재난안전 분야 어휘 사용 양상 분석

1. 유형별 어휘 사용 양상

비교 분석 결과를 토대로 다음과 같은 세 가지 기준을 갖고 유형별로 재난안전 분야 어휘 사용 양상을 분석하였다(그림 1).

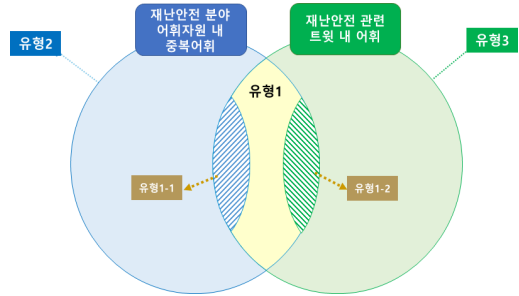


그림 1. 유형별 어휘 사용 양상 기준

- ① 유형1 : 재난안전 분야 중복어휘와 소셜미디어에 출현한 어휘가 일치하는 경우
- ② 유형2 : 재난안전 분야 중복어휘이나 소셜미디어에 출현하지 않은 경우
- ③ 유형3 : 소셜미디어에만 출현한 어휘인 경우

유형1의 경우에는 재난안전 분야 고빈도어와 소셜미디어 고빈도어로 구분하여 분석하였다. 그 이유는 각각의 출처에서 고빈도로 활용하는 중요 어휘의 특성을 파악하기 위해서이다.

선행 연구[1]를 통해서 재난안전 분야 어휘자원 중 중복 횟수가 높은 표제어들은 ‘일반용어’로 조사되었으며, 이 용어들은 대체로 정보제공자 및 정보이용자 측면에서 모두 유용한 정보이며 활용도가 높은 것으로 분석되었다. 반면 중복 횟수가 낮은 표제어들은 ‘전문용어’일 가능성이 높은 어휘로, 재난안전관련 업무 실무자들이 사용하면서도 분야별 전문성을 띄는 표제어인 것으로 파악되었다.

이에 본 연구는 소셜미디어를 통해 재난안전 분야에서 일반 이용자가 실제적으로 활용하는 어휘와 기존 어휘자원 간의 차이점을 비교·분석하고자 한다.

2. 유형1 분석 결과

2.1 재난안전 분야 고빈도어가 소셜미디어에서 얼마나 출현하는가(유형1-1)

재난안전 분야 중복어휘 중에서 트윗과 일치되는 어휘는 총 1,287건이며, 해당 어휘들은 소셜미디어 상에 평균 452건 등장한다. 재난안전 분야 고빈도어를 중심으로 소셜미디어에 나타나는 어휘들의 특징을 분석하

기 위해 다양한 재난을 대표하는 어휘로 표제어 중복도 순 상위 20개를 추출하였다[표 4]. 이를 통해 재난안전 분야에서 주로 쓰이는 업무 또는 전문용어가 일반적으로 얼마나 사용되는가를 확인할 수 있다.

표 4. 재난안전 분야 고빈도어의 소셜미디어 출현빈도

순위	표제어	중복도	소셜미디어 출현빈도
1	조류	13	1,201
2	방사선	13	43
3	생태계	13	23
4	복구	11	1,624
5	구조	11	1,502
6	지하수	11	11
7	세계보건기구	11	323
8	재해	11	523
9	전선	10	967
10	소독	10	819
11	잠복기	10	1,170
12	자외선	10	31
13	재난	9	3,459
14	적조	9	26
15	전도	9	80
16	방사능	9	232
17	선박	9	373
18	신뢰도	9	77
19	경계	8	368
20	환경	8	768
평균		10.2	681

재난안전 분야 고빈도어 상위 20개가 소셜미디어에 나타난 횟수는 평균 681건이다. 이는 전체 평균인 452건보다 높게 나타난 수치로, 상대적으로 중복 횟수가 높은 어휘는 일반 이용자에게 활용도가 높음을 확인할 수 있다. 보다 구체적으로 재난안전 분야 고빈도어의 소셜미디어 출현 현황 예시인 “조류”와 “방사선”을 살펴보면 [표 5]와 같다. “조류”의 경우에는 하나의 표제어에 대해 여러 개념들이 포함되어 다의적 용법을 가지는 ‘일반용어’에 해당된다. 반면 “방사선”은 하나의 표제어에 대해 하나의 개념을 지칭하는 ‘전문용어’에 해당된다.

이와 관련하여 정희찬 외(2017)는 재난안전 유관기관들이 특정 재난유형별로 전문화되어 있기 때문에 중복 그룹 내에서도 일반용어와 전문용어가 혼재되어 있다고 하였다[1]. 따라서 본 연구의 분석 결과는 이를 다시 보여주는 결과로 볼 수 있다. 이처럼 재난안전 분야 표준화 작업 시에는 일반용어와 전문용어가 혼재되는 현상이 발생하는 것에 유의해야 한다.

2.2 소셜미디어 고빈도어가 재난안전 분야 고빈도어와 얼마나 중복되는가(유형1~2)

재난안전 분야 중복어휘 중에서 트윗과 일치되는

표 5. 재난안전 분야 고빈도어의 소셜미디어 출현현황

표제어	중복도	정의	사용기관	소셜미디어	
				출현빈도	출현 내용
조류	13	-전기 에너지의 흐름 -원생생물계에 속하는 원시적인 진핵생물 -고등식물에 비해 비교적 간단한 조직과 형태를 갖는 하등식물군	산업통상자원부 국립수산과학원, 기상청, 부산광역시, 서울특별시, 한국수자원공사, 환경부	1,201	
		-조강의 철중용물을 일상적으로 통틀어 이르는 말	서울특별시		조류독감(마우나 리조트, 신중 인플루엔자, 메르스), 조류여왕(상왕십리), 조류인플루엔자(신중 인플루엔자, 메르스), 조류바이러스(신중 인플루엔자), 조류(메르스)
		-밀물과 썰물 때문에 일어나는 바닷물의 흐름 -주기적인 해수의 흐름	서울특별시, 해양수산부		조류예측기계(돌고래호), 조류(돌고래호, 강풍 풍랑)
방사선	13	-방사성원소의 붕괴에 따라 방사되는 고속도의 물질입자선. -물질 내에서 원자를 전리시키는 능력이 있는 입자선(알파입자, 베타입자 등) 또는 전자파(감마선, 엑스선 등)	국립수산과학원, 기상청, 산림청, 서울특별시, 식품의약품안전청, 원자력안전위원회, 한국수력원자력, 한국원자력안전기술원, 한국전력공사, 한국정보통신기술협회, 환경부	43	방사선(신중 인플루엔자), 방사선 음이온(신중 인플루엔자), 방사선 오염(신중 인플루엔자), 방사선 치료(메르스), 방사선실(메르스), 방사선사/방사선기사(메르스), 방사선 수처(서울 호우), 방사선 유출(경남 호우)

1,287건 중 트윗 고빈도어를 중심으로 상위 20개 어휘를 추출하였다[표 6]. 해당 어휘들이 소셜미디어 상에서 등장하는 횟수는 평균 10,803건으로 높으며, 일반 이용자들이 많이 사용하는 어휘들이 추출되었다. 이를 통해 재난안전과 관련되어 일반적으로 사용되는 어휘가 실제 업무에서는 어떻게 사용되는가를 확인할 수 있다.

유형1-2에서 추출된 어휘들을 살펴보면 중복도는 평균 5.15회로 대체로 낮다. 이로써 이용자들이 일반적으로 사용하는 어휘에 대한 재난안전 유관기관의 사용률은 낮다고 볼 수 있다. 이는 즉, 정보이용자와 정보제공자 간에 사용하는 어휘가 불일치하다는 것을 의미한다. 또한 격리, 확산, 대응, 발생, 예방, 사고 등 유형1-1보다 일반적인 어휘가 더 많이 나타난다. 특히 태풍, 감염, 호우, 바이러스, 주의보, 강풍 등을 제외하면 대부분이 일반용어로서, 이를 통해 소셜미디어 상에서는 주로 재난안전과 관련해서 일반용어가 의사소통에 많이 쓰이는 것을 확인할 수 있다.

유형1을 전체적으로 살펴봤을 때 재난안전 분야에서는 일반용어와 전문용어가 혼재되어 사용됨을 확인할 수 있으며, 방사선, 잠복기, 적조, 태풍, 감염 등 특정 재난상황 및 유형과 연계된 ‘사건 표제어’는 주로 일의성을 띄므로 전문용어일 가능성이 높다.

3. 유형2 분석 결과

재난안전 분야 중복어휘 중에서 트윗에 출현하지 않은 1,967건 중 표제어 중복도 상위 20개 어휘를 추출하였다[표 7]. 이를 통해 일반적으로는 사용되지 않지만 재난안전 분야에서 중요한 어휘들을 확인할 수 있다.

유형2의 어휘자원 중복어휘들이 소셜미디어에 나타나지 않은 이유는 해당 어휘들이 “메르스”와 같이 특정 ‘대형’ 재난사건과 관계가 없는 어휘들이기 때문이다. 따라서 해당 어휘들은 재난안전 분야에서 상시적으로 활용되는 ‘공통 표제어’로도 볼 수도 있다. 더불어 중복도 분석 결과 유형2에 해당하는 어휘들은 다른 유형 분석 결과에 비해 대체로 하나의 어휘에 하나의 개념을 나타내고 있다. 이는 ‘전문용어’의 특징 중 하나이다. 따라서 재난안전에서 상시적으로 활용되는 ‘공통 표제어’들은 ‘전문용어’일 가능성이 높다.

표 6. 소셜미디어 고빈도어의 재난안전 분야 출현빈도

순위	표제어	중복도	소셜미디어 출현빈도
1	태풍	7	55,964
2	감염	5	19,533
3	격리	5	16,711
4	호우	4	15,968
5	확산	4	15,006
6	대응	6	11,625
7	발생	4	10,775
8	바이러스	6	9,832
9	예방	5	9,039
10	사고	4	8,793
11	주의보	3	8,579
12	검사	7	4,213
13	경로	6	4,016
14	기상	5	3,941
15	비상	4	3,938
16	강풍	4	3,844
17	붕괴	5	3,840
18	사건	3	3,573
19	재난	9	3,459
20	경보	7	3,416
평균		5.15	10803.25

표 7. 소셜미디어에 출현하지 않은 재난안전 분야 중복 어휘

순위	표제어	중복도
1	비파괴검사	12
2	온실효과	11
3	질소산화물	11
4	공동구	10
5	산성비	10
6	온실가스	10
7	난류	10
8	슬러지	10
9	저기압	10
10	기상위성	9
11	대기오염	9
12	환경영향평가	9
13	관리지역	9
14	벤젠	9
15	방사선등위원소	9
16	오염물질	9
17	기후변화협약	9
18	열병합발전	9
19	방위각	9
20	부영양화	9
평균		9.65

결과적으로 유형1과 유형2를 분석한 결과 ‘사건 표제어’와 ‘공통 표제어’는 주로 전문용어에서 발견할 수 있었으며, 전문용어 중에서도 특정 재난상황 및 유형 관련 유무에 따라 어휘를 구분할 수 있다.

4. 유형3 분석 결과

유형3은 재난안전 분야 어휘자원에 등장하지 않고, 소셜미디어에만 출현한 어휘로서 자연 및 사회재난별로 출현 빈도를 중심으로 상위 10개씩 추출하였다[표 8]. 이를 통해 재난안전 유관기관의 사용 어휘에는 나타나지 않았지만 재난안전과 관련해서 일반적으로 활용되는 어휘를 확인할 수 있다.

표 8. 소셜미디어에만 출현한 어휘

자연재난			사회재난		
순위	등장 어휘	출현 빈도	순위	등장 어휘	출현 빈도
1	산간	2,765	1	확진	24,392
2	국지	1,064	2	독감	7,114
3	초속	861	3	감염자	5,459
4	비바람	652	4	탄저균	4,138
5	파도	320	5	감기	3,616
6	기상대	306	6	면역력	1,935
7	장대비	197	7	기침	1,931
8	피해액	183	8	폐렴	1,215
9	농경지	150	9	에볼라	820
10	내습	134	10	폭탄	697

유형3에 나타난 재난관련 기본 어휘들은 용어사전에 등재되지 않았으며, 주로 복합명사의 일부로서 사용됨을 확인할 수 있다. 예를 들어, 자연재난의 “국지”는 국지기상, 국지기후, 국지바람, 국지저기압 등으로, “초속”은 제동초속도, “기상대”는 중앙기상대로 사용되었다. 사회재난의 “확진”은 메르스 확진환자, 초기발생 확진 상황으로, “폐렴”은 과민성 폐렴, 화학성 폐렴, 방사선 폐렴으로 사용되었다. 즉, 유형3에 나타난 어휘들은 다른 용어를 재생산시킴으로써 높은 생산력을 가지는 재난안전 분야의 주요 기본 어휘이다. 결국 [표 8]의 어휘들은 비록 용어사전에 등재되지 않았으나 소셜미디어 상에서 출현빈도가 높으며, 생산성이 높은 어휘이므로 재난안전 용어사전의 주요 등재대상이 될 수 있다.

V. 특징별 표준화 방안

재난안전정보에 대한 국민들의 관심이 점점 높아지는 현 시점에서 재난안전 분야 어휘에 대한 사용 양상

을 분석하고, 이를 기반으로 제안된 표준화 방안을 적용하여 재난안전 용어사전을 구축하는 것은 매우 중요하다.

구체적으로 재난안전 분야에서 일반 이용자가 실제로 활용하는 어휘와 기존 어휘자원 간의 차이점을 통해 본 연구에서 도출한 특징은 다음과 같다. 첫째, 상대적으로 중복도가 높은 어휘는 일반용어로서 활용도가 높다(유형1-1). 둘째, 중복도가 높은 어휘 중에서도 재난 유형에 따라 일반용어와 전문용어가 혼재되어 있다(유형1-1). 셋째, 이용자들이 일반적으로 사용하는 어휘와 재난안전 유관기관에서 자주 활용하는 어휘 간에 간극이 있다(유형1-2). 넷째, 전문용어에서는 재난 관련 사전 표제어 및 공통 표제어를 추출할 수 있다(유형1-2, 유형2). 다섯째, 소셜미디어 상에서 출현빈도가 높으면서도 다른 용어를 재생산시키는 생산성 높은 어휘는 재난안전 분야에서 중요한 용어이다(유형3).

본 절에서는 상기한 유형별 어휘 특성에 따라 표준화 방안을 제시하고자 한다. 표준화 과정에서 중요한 점은 표준화할 대상을 선정하는 것이다. 본 연구에서 도출된 특징들을 통해 용어 정비 및 순화 대상, 표준화 우선순위 대상 선정이 가능하며, 선정된 재난안전 어휘들을 중심으로 용어사전을 구축해야 한다.

먼저 용어 정비 대상을 선정하기 위해서는 중복도가 높으면서도 소셜미디어 상에 자주 언급된 어휘를 추출해야 한다. 이현주와 조동성(2011)은 많은 분야에서 공통으로 활용되면서도, 빈도수가 높은 어휘를 중복용어 정비의 주요 지침으로 선정하였다[14]. 즉 중복도가 높아 여러 재난안전 유관기관에서 활용되면서도, 소셜미디어 상에 자주 언급된 유형1의 일반용어들은 정비의 주요 대상이면서도 용어사전 수록 대상이다. 특히, 유형 1-1에 출현한 어휘는 중복도가 높아 재난안전 분야와 관련성이 높으면서도, 정보이용자에게 자주 사용되는 어휘로서 대국민을 대상으로 하는 용어사전 구축 시에 주요 표준화 대상이 될 수 있다.

다음으로 용어 순화화를 위해서는 혼재된 재난안전 어휘들 중에서 전문용어를 추출할 수 있어야 한다. 본 연구는 유형1과 유형2에 대한 분석을 통해 전문용어를 확인할 수 있었다. 전문용어는 정보이용자가 실제 사용

하는 어휘들과 다소 차이가 있으므로 의미 환기성이 뛰어난 용어로 순화가 필요하다. 재난안전정보에는 재난 안전 유관기관에서 사용하는 어휘가 중점적으로 활용되는 반면에, 소셜미디어 상에서 정보이용자들은 자신의 의사를 표현하기 위해 재난안전 분야 전문용어와 일상용어를 조합하여 사용하는 경향을 보인다. 이는 곧 국민과 기관 간의 재난안전 분야 어휘 사용 양상이 불일치함을 의미한다. 이러한 어휘 사용 양상의 불일치는 잘못된 의사소통을 야기하고, 재난안전정보 공유에 큰 장애를 유발시킬 수 있다. 특히, 일반 이용자는 전문용어의 복잡성으로 인해 재난안전정보 이용에 방해 받아서는 안 되며, 정보에 대한 이용자의 접근이 용이하도록 정보제공자는 순화된 용어를 사용해야 한다.

또한 재난안전 분야 사건·공통표제어와 기본 어휘가 추출 가능함에 따라 표준화 대상 우선순위를 선정할 수 있다. 본 연구는 전문용어에서 사건 표제어 및 공통 표제어를 추출할 수 있음을 파악하였다. 정보이용자 측면에서 사건 표제어는 활용도가 높은 어휘로서[1], 대국민 서비스를 목적으로 용어 표준화 작업을 수행할 시에는 사건 표제어를 우선적으로 추출하여 표준화해야 한다.

소셜미디어 상에만 출현한 어휘(유형3)들은 생산성이 높아 용어사전에 등재될 필요가 있는 재난안전 분야의 기본 어휘이다. 재난안전 용어사전에는 61개 재난안전 유관기관에서 사용하는 모든 용어를 수록할 수 없으므로, 재난안전 분야 어휘의 기본적인 범위를 결정해야 한다. 이 때 높은 생산성을 가지면서도, 소셜미디어 상에서 활용성이 높은 어휘들은 용어사전의 주요 수록 대상이다. 그러나 소셜미디어 상에서만 출현한 어휘 중에서는 일반적으로 많이 쓰이는 어휘 즉, 일상어가 포함되어 있어, 용어사전에 등재하기 어려워 보이는 경우가 있다. 그러나 각 전문 분야별 용어의 정의는 해당 분야와 밀접한 관련이 있으며, 전문 분야에서 사용되어야만 용어로서의 지위를 가질 수 있다[14]. 즉, 재난안전 분야의 특수하고도 기본적인 개념을 가리킨다면 아무리 일반적으로 많이 쓰이는 어휘라 할지라도 재난안전 분야의 용어로 간주할 수 있다.

VI. 결론

재난안전 분야 어휘의 표준화는 성공적인 재난 예방 및 대응을 위해 가장 기본적인 과정으로서 중요하다. 재난안전 분야의 어휘에 대한 이해 부족은 곧 커뮤니케이션의 부재로 이어지며, 이는 크나큰 사고를 유발할 수 있다. 실제 어휘 사용 양상의 차이를 확인하는 것은 표준화를 위해 필수적인 부분이다.

본 연구는 미래창조과학부 및 국민안전처에서 수행하는 재난안전정보 표준화사업의 일환으로 진행되었으며, 소셜미디어를 중심으로 재난안전 분야 어휘 사용 양상을 분석함으로써, 재난안전 분야 어휘에 대한 표준화 방안 및 용어사전 구축 방향을 수립하는 데 그 목적이 있다. 이를 위해 재난안전 분야의 어휘자원 및 트윗으로부터 재난안전 관련 어휘를 수집하여 어휘 사용 양상을 유형별로 비교·분석하였다.

분석 결과 상대적으로 중복도가 높은 어휘는 일반용어로서 활용도가 높으며, 중복도가 높은 어휘 중에서도 일반용어와 전문용어가 혼재되어 있었다. 전문용어에서는 사건 표제어 및 공통 표제어를 추출할 수 있다. 또한 이용자들이 일반적으로 사용하는 어휘에 대한 재난안전 유관기관의 사용률은 낮으며, 소셜미디어 상에서 출현빈도가 높으면서도 생산성이 높은 어휘는 중요한 어휘임을 확인할 수 있었다. 이러한 특징들을 반영하여 재난안전 분야 어휘에 대한 표준화 작업을 진행해야 하며, 궁극적으로는 재난안전 용어사전을 구축해야 한다.

본 연구는 재난안전 분야의 어휘 사용 양상을 소셜미디어를 중심으로 분석하여 표준화 방안을 제안하였다는 데 의의가 있다. 향후에는 재난안전 분야 연구자와 현업 종사자 등 전문가 집단의 어휘자원에 대한 활용도를 분석하여 보다 더 구체적인 용어사전의 표준화 방안을 수립하고자 한다. 또한 분석 대상 미디어를 다양화하여 본 연구에서 도출한 결론을 검증하고자 한다.

참고 문헌

- [1] 정희찬, 김태영, 김용, 오효정, “재난안전정보 관리를 위한 어휘자원 현황분석 및 활용방안,” 정보

- 관리학회지, 제34권, 제2호, pp.137-158, 2017.
- [2] 권정현, *전문용어 순화정책을 위한 혼종어 연구*, 한성대학교 대학원, 박사학위논문, 2016.
- [3] 송영빈, “의학 논문을 통해 본 전문용어 사용 양상,” *한국사전학*, 제20권, pp.211-233, 2012.
- [4] <http://www.mof.go.kr/article/view.do?articleKey=13354&boardKey=17&menuKey=383¤tPageNo=1>
- [5] 김일태, “만화애니메이션 학술용어 표준화 및 용어사전 편찬 연구,” *만화애니메이션 연구*, 제10권, pp.17-31, 2006.
- [6] 김영애, “의도와 건축외피를 설명하는 주제 용어에 대한 기초연구,” *대한건축학회 논문집*, 제28권, 제11호, pp.39-48, 2006.
- [7] 최중환, 최석두, 김이걸, 박영옥, 정종희, 안희정, 정한민, 김평, “국방과학기술 전문용어 사전 구축을 위한 프로세스 표준화 및 활용 방안,” *한국콘텐츠학회논문지*, 제11권, 제8호, pp.247-259, 2011.
- [8] 박별나, 이윤경, 구자을, “고등학교 과학 및 생물 교과서 과학용어 네트워크 분석,” *한국콘텐츠학회논문지*, 제10권, 제5호, pp.427-435, 2010.
- [9] 강성경, 유환, 이영재, “텍스트마이닝과 소셜 네트워크 분석을 이용한 재난대응 용어분석,” *Information Systems Review*, 제18권, 제1호, pp.141-155, 2016.
- [10] 국민안전처, *2015 재난연감*, 국민안전처, 2016.
- [11] 국민안전처, *2015 재해연보*, 국민안전처, 2016.
- [12] <https://sourceforge.net/projects/koreanalyzer>
- [13] <https://www.r-project.org/>
- [14] 이현주, 조동성, “학술 전문용어 정비 및 표준화의 특징 및 과제,” *한국어 의미학*, 제35권, pp.245-283, 2011.

저 자 소 개

김 태 영(Tae-Young Kim)

정회원



- 2013년 2월 : 전북대학교 기록관리학과(문학석사)
- 2015년 9월 ~ 현재 : 전북대학교 기록관리학과 박사과정

<관심분야> : 이용자 행태, 빅데이터 분석, 온톨로지, 네트워크 분석, 시소러스

이 정 은(Jung-Eun Lee)

정회원

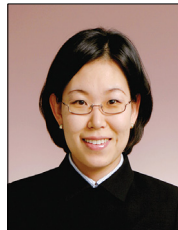


- 2017년 2월 : 전북대학교 기록관리학과(문학석사)
- 2017년 3월 ~ 현재 : 전북대학교 기록관리학과 박사과정

<관심분야> : 이용자 분석, 빅데이터 분석, 네트워크 분석

오 효 정(Hyo-Jung, Oh)

정회원



- 2000년 2월 : 충남대학교 컴퓨터과학과(이학석사)
- 2008년 2월 : 한국과학기술원 컴퓨터공학과(공학박사)
- 2000년 2월 ~ 2015년 5월 : 한국전자통신연구원 지식마이닝연구실 책임연구원

실 책임연구원

- 2015년 5월 ~ 현재 : 전북대학교 기록관리학과 조교수
- <관심분야> : 정보검색, 질의응답, 빅데이터정보처리, 소셜웹마이닝