

재난안전정보 관리를 위한 어휘자원 현황분석 및 활용방안*

A Study on the Utilization Plan of Lexical Resources for Disaster and Safety Information Management Based on Current Status Analysis

정힘찬 (Him-Chan Jeong)**, 김태영 (Tae-Young Kim)***
김 옹 (Yong Kim)****, 오효정 (Hyo-Jung Oh)*****

초 록

재난은 국민의 생명·신체·재산에 직접적인 영향을 미치는 사건으로, 재난 발생 시 신속하고 효과적인 대응을 위해서는 관련 정보들을 효율적으로 공유, 활용하는 협조 과정이 무엇보다도 중요하다. 현재 재난안전 유관기관별로 다양한 재난안전정보가 생산 및 관리되고 있지만, 각 기관별로 개별적인 용어와 의미를 정의하여 활용하고 있다. 이는 재난안전정보를 검색하고 접근하려는 실무자 입장에서 큰 걸림돌이며, 기관별 정보 활용도를 저해시키는 요인 중에 하나이다. 이러한 문제점을 해결하기 위해 재난안전정보의 통합적 관리를 위한 어휘자원의 표준화 작업의 선행 연구로, 본 연구에서는 재난안전 유관기관에서 관리하고 있는 어휘자원의 현황분석을 수행하였다. 또한 수집된 어휘자원을 대상으로 정보제공자 및 이용자 관점에서의 활용도 분석을 통해 어휘 그룹별 특성을 파악하고 이에 기반해 재난안전정보 관리를 위한 활용방안을 제안하였다.

ABSTRACT

Disaster has a direct influence on the lives of the people, the body, and the property. For effective and rapid disaster responses, coordination process based on sharing and utilizing disaster information is the essential requirement. Disaster and safety control agencies produce and manage heterogeneous information. They also develop and use word dictionaries individually. This is a major obstacle to retrieve and access disaster and safety information in terms of practitioners. To solve this problem, standardization of lexical resources related disaster and safety is essentially required. In this paper, we conducted current status analysis about lexical resources in disaster and safety domain. Consequently, we identified the characteristics according to lexical groups. And then we proposed the utilization plan of lexical resources for disaster and safety information management.

키워드: 재난안전정보, 정보관리, 어휘자원, 용어사전, 표준화
disaster and safety information, information management, lexical resource, word dictionary, standardization

* 이 논문은 2017년도 전북대학교 연구기반 조성비 지원에 의하여 연구되었음.

이 논문은 2016년도 정부(미래창조과학부)의 재원으로 한국연구재단-재난안전플랫폼기술개발사업의 지원을 받아 수행된 연구임(과제번호: NRF-2016M3D7A1912703).

** 전북대학교 일반대학원 기록관리학과 석사과정(hc-1106@naver.com) (제1저자)

*** 전북대학교 일반대학원 기록관리학과 박사과정(fnty127@hanmail.net) (공동저자)

**** 전북대학교 문헌정보학과 교수, 문화융복합이카이빙연구소장(yk9118@jbnu.ac.kr) (공동저자)

***** 전북대학교 기록관리학과 조교수, 의료정보융합 연구원(ohj@jbnu.ac.kr) (교신저자)

■ 논문접수일자: 2017년 6월 2일 ■ 최초심사일자: 2017년 6월 9일 ■ 게재확정일자: 2017년 6월 15일

■ 정보관리학회지, 34(2), 137-158, 2017. [http://dx.doi.org/10.3743/KOSIM.2017.34.2.137]

1. 서론

1.1 연구배경 및 필요성

재난은 국민의 생명·신체·재산에 직접적인 영향을 미치는 사건이다. 최근 빈번한 자연 및 사회재난의 발생으로 인하여 국민들은 다양한 신체적·정신적 피해를 받고 있다. 재난 발생 시 신속하고 효과적인 대응을 위해서는 관련 정보들을 효율적으로 공유, 활용하는 협조 과정이 무엇보다도 중요하다. 현재 재난안전 유관기관별로 2,500여 종의 다양한 재난안전 정보가 생산 및 관리되고 있다(국가민방위재난안전교육원, 2017). 뿐만 아니라 각 기관에서는 개별적으로 용어사전을 개발하여 정보 공유에 활용하고 있다. 그러나 같은 업무기능을 수행하는 재난안전 관련 기관 내에서도 동일한 용어를 다양한 의미로, 혹은 같은 의미를 다양한 용어로 사용하고 있어 기관별 실무 담당자 간

의 이해도를 저해시키며, 그 결과 기관별 정보 활용도도 감소시키는 요인이 된다. 이에 대한 구체적인 예시로 <표 1>을 살펴보면, 동일 표제어 '냉해' 내에서 각기 다른 의미의 정의문들이 기술되었음을 알 수 있다. 반면 다른 용어사전에서는 '냉해'와 같은 의미의 정의문에 해당하는 표제어로 '한상', '냉수장애'가 쓰이고 있다. 이처럼 재난안전 용어의 의미적 모호성은 재난 안전정보를 검색하고 활용하려는 실무자 또는 일반 이용자 입장에서 큰 걸림돌이 된다.

국가는 재난의 예방과 피해경감을 위하여 재난에 대한 대응·복구 계획을 신속히 수립·시행해야 한다(국민안전처, 2016). 재난 대응 및 복구 계획을 수립하기 위해서는 해당 분야에 대한 다양한 정보들이 필요하다. 위에서 언급했다시피 재난안전과 관련된 정보들은 다양한 유관기관으로부터 생산된다. 따라서 재난안전 분야는 기관 간의 정보 공유가 필수적이며 효율적인 정보 공유를 위해서는 표준화된 기술

<표 1> '냉해' 관련 표제어 및 정의문 예시

표제어	정의문	출처
냉해	수온이 낮은 물을 농업용수로 이용함에 따라 곡물의 생육에 장애가 발생하는 것을 냉수해라고 말한다. 눈 녹은 물이 직접 흘러드는 산간지 등에서 자주 발생하며, 작물의 생육시기에 따라 냉수해의 정도도 다르게 나타난다.	방재사전
냉해	5~9월의 여름철에 농작물의 성장 기간 중 작물의 생육에 지장을 초래할 정도의 <u>저온이 오래 지속되어</u> 농작물의 작황이 나빠지는 재해를 말한다.	재난안전 용어사전
냉해	찬물에 의한 장애로, 작물 생육시기에 따라 냉수해 정도가 다르게 나타난다.	재난안전 용어사전
냉해	여름 작물이 생육기간 중에 <u>냉온 장애</u> 에 의하여 생육이 저해되고 수량의 감소나 품질의 저하를 가져오는 기상재해를 말한다.	기상청
냉해	열대 또는 난대식물이 0℃ 이상의 저온에 의하여 받는 피해로서 0℃ 이상의 세포동결에 의하여 나타나는 동해(凍害)와 구별된다.	산림청
한상	0℃ 이상의 온도에서 일어나는 냉해와 0℃ 이하의 온도에서 얼음결빙에 의해 일어나는 동해를 모두 한해(寒害) 혹은 한상(寒傷)이라고 한다.	산림청
냉수장애	눈에 담수된 물의 온도가 낮아서 생기는 벼의 생육장애.	수자원공사

(description) 체계가 기반이 되어야 한다. 이러한 체계는 표준화된 용어를 기초로 개발된 분류 체계, 메타데이터, 온톨로지를 통해 마련될 수 있다. 그러나 현재 국내 재난안전 분야의 경우, 각 유관기관마다 독자적으로 용어사전을 개발, 관리함에 따라 기관 간의 정보 공유 및 의사소통이 원활하지 않은 실정이다.

또한 재난안전에 대한 공통의 이해를 바탕으로 성공적인 업무 지시와 커뮤니케이션을 수행하기 위해서는 재난안전정보 공유를 위한 플랫폼 개발이 필수적이다. 표준화된 용어를 활용하여 재난안전정보 공유 플랫폼을 개발하고, 효율적으로 재난안전정보를 관리하면 해당 정보에 대한 정확한 표현을 국민 또는 기관에게 전달할 수 있게 된다. 용어의 표준화는 효율적인 정보 공유 및 활용의 기본 토대이다. 따라서 본 연구는 재난안전정보의 통합적 관리를 위한 어휘자원의 표준화에 앞서 재난안전 유관기관에서 수집한 어휘자원을 대상으로 재난안전 분야 어휘자원이 지닌 특징과 활용도를 분석하였으며, 효율적인 정보 공유와 활용을 위해 몇 가지 활용방안을 제시하였다.

1.2 연구목적 및 방법

본 연구는 재난안전 분야의 표준화된 용어사전 구축을 위해 기존에 관련 유관기관에서 관리하고 있는 어휘자원에 대한 현황을 파악하고, 수집된 어휘자원이 활용도 측면에서 유효한가를 분석하여 재난안전정보 관리를 위한 활용방안을 제시하고자 한다. 이를 위한 세부 연구 방법은 다음과 같다.

첫째, 재난안전 용어와 관련된 문헌조사 및

현황조사를 수행하였다. 문헌조사를 통해서 재난안전정보 및 재난 관련 법령·표준 현황에 대하여 전반적으로 검토하였다. 더불어 재난안전 분야의 어휘자원에 대한 현황조사를 수행하였으며, 그 중에서도 재난안전 유관기관의 용어사전 보유 현황을 중점적으로 조사하였다. 조사 결과 총 61개 중 40개의 재난안전 유관기관에서 100개의 용어사전을 보유하고 있음을 최종적으로 확인하였다.

둘째, 현황조사 내용을 토대로 실제 재난안전 유관기관을 대상으로 현장방문조사 및 전화 연락을 수행하여 용어사전을 수집하였다. 또한 분석 결과의 타당성을 위해 용어사전 외에 추가적으로 재난안전 관련 문헌으로부터 표제어(entry)를 수작업으로 수집하였다. 본 연구에서는 총 388건의 용어사전 및 관련 문헌으로부터 수집된 114,782건의 어휘자원을 디지털화하여 기초 분석에 활용하였다.

셋째, 수집된 어휘자원에 대한 활용도 분석 수행을 통해 어휘 그룹별 특성을 파악하고, 이에 기반해 재난안전정보 관리를 위한 활용방안을 제안하였다. 특히, 정보제공자 및 정보이용자별로 어휘자원 활용도를 분석하고자 네이버 뉴스 및 구글트렌드 검색을 활용하였다.

1.3 선행연구

재난안전 분야에 대한 관심이 집중됨에 따라 재난관리체계 개발 및 재난안전정보 활용 등 재난피해 경감을 위한 다양한 연구가 현재 진행되고 있다. 이에 본 연구에서는 재난안전 분야에서의 용어 관련 연구들을 구체적으로 살펴 보았으며, 더불어 통합적인 재난안전정보 관리

를 위한 용어의 역할에 대하여 선행연구를 통해 고찰해 보았다.

우선 재난안전 분야에서의 용어 관련 선행연구들은 구체적으로 다음과 같다. 국립재난안전연구원(2013)은 재난안전관리 모니터링의 효율성을 향상시키기 위하여 재난안전 용어 간 관계모형을 구축하고자 하였다. 이에 재난 관련 비정형 빅데이터에서 주로 나타난 용어를 추출한 후, 분류체계를 기반으로 재정리하였다. 구축된 재난안전 용어 간 관계모형은 연관어 네트워크 연결에 활용될 수 있다. 해당 연구의 관계모형은 비정형 빅데이터의 모니터링을 위한 것으로 전체 재난안전 분야의 적용에는 한계가 있다.

강성경과 이영재(2015)는 재난현장 통합대응 표준체계 마련의 일환으로 각 기관별 법령 및 표준, 매뉴얼 일부를 검토하여 표준용어 권고안을 제안하였다. 원활한 재난 통합대응을 위해서는 대응인력 간에 표준화된 공통용어를 사용해야 함을 주장하였다.

Mayner와 Arbon(2015)은 용어사전을 포함한 광범위한 출처로부터 파생된 재난안전 관련 용어들의 표준화를 강조하였다. 이에 110개의 용어사전으로부터 재난안전 관련 용어들을 수집하고, 텍스트 분석기인 Leximancer 소프트웨어를 사용하여 재난안전 주요 용어들을 도출하였다.

강성경, 유환, 이영재(2016)는 재난안전 분야의 원활한 업무 협업을 위해서는 의사소통이 매우 중요함을 피력하였다. 의사소통은 용어를 기반으로 이루어지기 때문에 해당 연구에서는 재난안전 분야의 용어 표준화를 강조하였다. 용어 표준화 기준 마련의 일환으로 재난관리단계 중 재난대응에 대한 핵심 용어를 텍스트 마

이닝 및 소셜 네트워크 분석을 통해 파악하는 연구를 진행하였다.

특정 분야의 다양하고 이질적인 정보 관리를 위한 용어의 역할에 대해 고찰할 수 있는 선행연구는 다음과 같다. 왕순주(2008)는 인간 사회의 모든 활동의 기초는 의사소통이며, 의사소통의 도구는 인간이 사용하는 언어이기 때문에 용어의 표준화가 중요함을 피력하였다. 정보관리 프로그램의 효율적 운영을 위해서는 용어체계의 확립이 중요하며, 이와 관련된 보건 의료 분야의 표준화 활동에 대한 기초 사항들을 조사하였다.

신윤주와 김정은(2009)은 건강정보 소비자 용어 시스템을 구축하기 위한 기초 연구로서, 건강정보 소비자 및 의료전문가 간 사용용어의 차이를 규명하고 소비자 용어의 특징을 분석하였다. 특히, 소비자와 의료전문가의 건강 관련 용어 차이는 효율적인 건강정보검색을 가로막는 장애물이기 때문에, 해당 연구에서는 건강정보시스템을 구성하는데 용어 간 차이 해소는 필수적인 단계라 강조하였다.

서진순, 김상균, 김안나, 장현철(2015)은 혼재된 용어는 의사소통 및 정보 공유에 혼선을 일으키기 때문에 용어를 수집하고 체계적으로 정리하는 등의 표준화가 필요하다고 하였다. 특히나 한의학 용어는 오랜 기간 동안 생성 및 소멸을 반복하면서 여러 용어들이 혼재되어 있으므로, 한의학용어를 일반 사용자가 쉽게 검색할 수 있는 서비스 및 시스템 개발과 표준화된 용어를 통한 분류체계가 필요함을 밝혔다.

재난안전에 대한 이용자의 관심과 요구가 급증한 반면 재난안전 분야에서의 용어 연구는 아직 미비한 상태이다. 선행 연구 대부분은 다양

한 출처에서 생산되는 재난안전 용어의 표준화 필요성을 강조하고 있으며, 이를 위한 사전작업으로 주요 용어들을 도출하였다. 또한 전반적으로 재난안전정보를 관리하고 획일화된 의사소통 체계 마련을 위해서는 용어의 표준화가 필요함을 인지하고 있었다. 대부분의 연구가 일부 범위에 한정하여 재난안전 용어를 추출하거나 주요 용어를 파악하는데 그친 반면, 본 연구에서는 재난안전 분야 전 범위의 어휘자원을 추출하고 이에 대한 실질적인 활용성을 분석하고자 한다. 특히 정보를 체계적으로 관리하고 공유하기 위한 표준화 작업의 선행 연구로써 기본 체계에 용어 사전을 활용하는 방안을 제시한다는 점에 본 연구의 의의가 있다.

2. 재난안전 분야 어휘자원 현황분석

2.1 재난안전 분야 어휘자원 수집절차

본 연구에서는 재난안전 분야 어휘자원을 4단계의 절차를 거쳐 수집하였으며, 단계적으로 수집된 어휘자원은 재난안전 분야에서의 활용도 분석에 사용되었다. 세부 수집 절차는 <표 2>와 같다.

<표 2> 재난안전 분야 어휘자원 수집절차

단계	내용
1단계	재난안전 분야 어휘자원 온·오프라인 현황조사
2단계	재난안전 분야 관련 문헌조사
3단계	재난안전 분야 용어사전 수집 및 표제어(entry) 식별
4단계	재난안전 분야 어휘자원 디지털화

수집 1단계는 재난안전 분야 용어사전에 대한 온·오프라인 현황조사이다. 본 연구에서는 총 61개 기관(국민안전처 및 재난안전 유관기관)을 대상으로 재난안전 분야 용어사전에 대한 온라인 현황조사를 수행하였다. 재난안전 유관기관의 선정 기준은 「재난 및 안전관리기본법(이하, 재난안전관리법)」 시행령의 “재난 및 사고유형별 재난관리주관기관”과 국가재난관리정보시스템(NDMS)의 정보제공기관 목록이다(소방방재청, 2005). 재난안전 유관기관 홈페이지 및 국립중앙도서관을 전수 조사하여 용어사전 여부를 확인하는 것으로 온라인 현황조사가 수행되었다. 1차 온라인 현황조사 결과를 바탕으로 용어사전과 관련된 기관에 대한 현장 방문조사 및 전화연락 등 오프라인 조사를 수행하였다. 이를 통해 총 40개의 재난안전 유관기관에서 100개의 용어사전을 보유하고 있음을 최종적으로 확인하였다.

수집 2단계는 재난안전 분야 관련 문헌조사이다. 본 연구에서는 재난안전 분야의 어휘자원을 수집하기 위하여 관련 표준·법령·매뉴얼·단행본 등 문헌 현황을 조사하였다. 그 결과 사회안전 ISO 국가표준(4건), 재난 및 사고 유형 관련 법령(300건), 위기관리 표준매뉴얼 및 위기대응 실무매뉴얼(48건), 재난안전 관련 단행본(3건) 등이 어휘자원 수집 대상으로 선정되었다. 어휘자원 수집 대상의 선정 기준은 온라인 조사 및 현장방문조사 결과를 따르고 있다. 총 61개의 기관(국민안전처 및 재난안전 유관기관)을 대상으로 현장방문조사 및 전화연락을 수행하였다.

현장방문조사 결과, 국민안전처는 국제표준 ISO/TC 223(사회안전)에서 정의한 용어를 따

르는 것으로 파악되었다. 따라서 이에 대응하는 사회안전 관련 국가표준 4건의 문헌을 용어 수집 대상으로 선정하였다. 또한 국민안전처 및 재난안전 유관기관은 법령에서 정의한 용어를 따르므로 선정된 유관기관의 소관 법령 총 3,690건(2017년 4월 20일 기준, 중복 제거) 중 재난 및 사고유형과 관련된 법령 300건을 용어 수집 대상으로 선정하였다. 더불어 본 연구에서는 국내 법률에 근거한 총 5,302건의 매뉴얼 중 수집 가능한 48건의 위기관리 표준매뉴얼 및 위기대응 실무매뉴얼과 최신의 재난안전정보를 종합적으로 포함하며 공신력 있는 기관에서 출판된 단행본 3건을 용어수집 대상으로 선정하였다.

수집 3단계는 재난안전 분야 용어사전 수집 및 표제어(entry) 식별이다. 본 연구에서는 앞서 파악한 용어사전 현황조사 결과를 토대로 국민안전처의 공식 협조를 통해 재난안전 유관 기관으로부터 해당 사전을 실제적으로 수집하였다. 또한 재난안전 분야 어휘자원 수집 대상으로 선정된 관련 문헌으로부터 대상 어휘 후보를 추출, 용어사전의 표제어(entry)로 정의문을 식별하였다.

수집 4단계는 재난안전 분야 어휘자원의 디지털화이다. 확보된 용어사전 중에는 전자 파일 형태로 작성된 것 뿐 아니라 인쇄본 형태의 문헌 기록자료로 보관된 경우도 있었다.

이를 위해 본 연구에서는 수작업 입력 작업을 병행하였으며 입력된 자료를 정제(cleansing)하고 정렬(alinement)하는 등의 전처리(text preprocessing) 과정을 통해 최종 DB화하였다. 디지털화된 어휘자원은 중복도 및 활용도 분석에 활용되었다.

2.2 재난안전 분야 어휘자원 수집현황

수집 및 디지털화된 총 33개 용어사전(〈표 4〉 참조)에 포함된 표제어 109,278건은 본 연구의 주요 어휘자원으로 활용되었다. 현황조사 결과 나타난 총 100개의 용어사전 중 보유 기관으로부터 사용 협의를 완료한 33개 용어사전을 본 연구에 활용하였다. 이와 함께 수집된 관련 문헌으로부터 식별된 표제어(entry) 및 정의문의 수집현황은 〈표 3〉과 같다.

〈표 4〉는 재난안전 분야 용어사전 수집현황을 표제어 수 중심으로 정리한 것이며, 중복 표제어 수 데이터도 포함되어 있다. 중복 표제어 수는 중복도 분석에 나타난 각 용어사전별 표제어 개수이다. 총 용어사전 표제어 109,278건 중 중복 표제어는 26,601건으로, 이를 통해 용어사전 중 약 24%의 어휘들이 재난안전 분야에서 주로 활용됨을 확인할 수 있다. 또한 표를 통해 표제어 수 상위 5개와 중복 표제어 수 상위 5개(전력관련 용어사전 정보(2016), 국가기

〈표 3〉 재난안전 분야 어휘자원 수집현황

대상	용어사전	국가표준	단행본	법령	매뉴얼	합계
건 수	33	4	3	300	48	388
어휘 수	109,278	138	2,725	2,053	588	114,782

* 2017.4.30. 기준

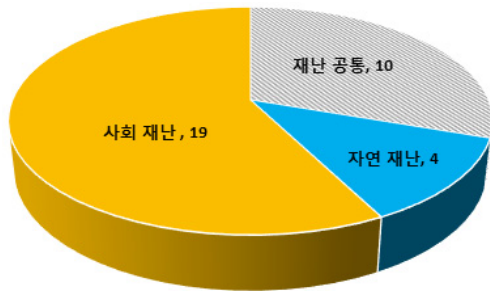
〈표 4〉 재난안전 분야 용어사전 수집현황

	사전명	기관명	유형	표제어 수	중복 표제어 수
1	TTA정보통신 용어사전	한국정보통신기술협회	웹사전	25,103	1,362(5%)
2	국가기후데이터센터 대기과학 용어집 및 용어사전	기상청	웹사전	21,651	4,344(20%)
3	서울시 행정데이터 표준용어사전 정보	서울특별시	웹사전	13,664	3,621(27%)
4	전력관련 용어사전 정보(2016)	한국전력공사	웹사전	10,103	4,968(49%)
5	산림입업용어사전	산림청	웹사전	9,937	2,005(20%)
6	항공관련 용어(2013년 기준)	국토교통부	웹사전	8,460	1,267(15%)
7	전력 관련 용어 정보(2016)	한국전력공사	웹사전	4,365	3,583(82%)
8	안전보건용어사전	안전보건공단	웹사전	3,599	920(26%)
9	국가공간정보포털 공간정보용어사전	국토교통부	웹사전	1,758	563(32%)
10	원자력 용어	한국수력원자력	웹사전	1,331	272(20%)
11	물용어사전	한국수자원공사	인쇄본	1,236	569(46%)
12	해양교육포털 해양용어사전	해양수산부	웹사전	1,164	501(43%)
13	철도 용어해설집	한국철도공사	웹사전	1,135	161(14%)
14	한중일 수산용어집 데이터베이스	국립수산과학원	웹사전	839	330(39%)
15	토지이용 관련 쉬운 용어사전(2015)	국토교통부	웹사전	597	249(42%)
16	발전설비 검사업무 용어집(2013)	한국전기안전공사	웹사전	563	156(28%)
17	환경성질환 용어집(2011)	환경부	웹사전	558	131(23%)
18	케미스토리 환경용어사전	환경부	웹사전	436	197(45%)
19	연안포털 용어사전	해양수산부	웹사전	423	312(74%)
20	서울특별시 대기환경정보 용어사전	서울특별시	웹사전	390	230(59%)
21	보건환경연구원 용어설명상식	부산광역시	웹사전	279	148(53%)
22	2013 우리 본부 전문 용어집 ver.2	농림축산검역본부	인쇄본	264	73(28%)
23	물환경정보시스템 용어사전	환경부	웹사전	255	136(53%)
24	산지정보시스템 용어해설집	산림청	웹사전	247	93(38%)
25	2015 알기쉬운 도시계획 용어집	서울특별시	웹사전	190	113(59%)
26	원자력안전용어	한국원자력 안전기술원	웹사전	181	77(43%)
27	항공용어 해설('16)	한국공항공사	웹사전	152	50(33%)
28	화학물질안전관리정보시스템 화학관련용어	화학물질안전원	웹사전	93	47(51%)
29	지진용어해설	부산광역시	웹사전	87	29(33%)
30	메르스 백서 용어 정의	질병관리본부	웹사전	69	7(10%)
31	화학물질안전관리정보시스템 해양관련용어	화학물질안전원	웹사전	62	32(52%)
32	Air Korea 용어사전	한국환경공단	웹사전	59	47(80%)
33	저녹스버너 관련 용어 정의	한국환경공단	웹사전	28	8(29%)
합계				109,278	26,601(24%)

후데이터센터 대기과학 용어집 및 용어사전, 서울시 행정데이터 표준용어사전 정보, 전력 관련 용어 정보(2016), 산림입업용어사전)가 유사함

을 확인할 수 있다. 중복도 상위에 위치한 용어 사전은 상대적으로 활용도가 높을 것으로 예측 된다.

또한 본 연구는 재난안전 분야 용어사전에 대한 재난유형별 분포 현황을 살펴보았다. 재난유형은 크게 공통, 자연재난, 사회재난으로 구분하였으며, 이에 대한 세분류 기준은 재난안전관리법에서 정의한 재난 유형 기준에 따른다. 예를 들어 TTA정보통신사전의 경우에는 ‘사회’재난으로 구분하였다. 재난안전관리법에서는 에너지·통신·교통·금융·의료·수도 등 국가기반체계의 마비를 사회재난의 유형 중 하나로 정의하였으며, 여기에서 TTA정보통신사전이 “통신”이란 분야에 포함된다고 보았다. 분석 결과 총 33개의 용어사전 중 10개가 재난공통, 4개가 자연재난, 19개가 사회재난에 해당하는 용어사전임을 확인하였다. 전체 용어사전 중 약 30%만이 재난공통에 해당되었으며, 이는 곧 용어사전이 재난유형별로 전문화되어 있다는 것을 의미한다(〈그림 1〉 참조).



〈그림 1〉 재난유형별 용어사전 현황

2.3 재난안전 분야 어휘자원 중복도 분석

효율적인 재난안전정보의 공유 및 활용을 위해서는 어휘자원의 표준화가 필수적이다. 그러나 〈표 4〉에서 나타나듯이 현재 재난안전 유관 기관에서 활용하는 어휘자원들은 기관별로 상

이하다. 표준화를 위해서는 현재 각 기관별로 얼마나 다양한 어휘자원을 사용하며, 어떠한 의미로 정의되어 있는지를 확인할 필요가 있다. 이에 따라 본 연구에서는 재난안전 분야 어휘자원의 특성을 파악하기 위하여 디지털화된 어휘자원을 대상으로 중복도 분석을 실시하였다.

어휘자원의 중복은 서로 다른 출처에서 동일한 표제어가 등장한 것을 의미한다. 예를 들어 1개의 표제어가 2개의 어휘자원 출처(sources)에서 등장했다면 중복 횟수는 2회로, 본 연구에서 나타난 최대 중복 횟수는 13회이다. 중복도 분석은 동형의어를 고려하지 않고 표제어의 일치도를 기반으로 수행하였으며, 이를 통해 하나의 표제어가 기관별로 다양한 의미로 사용됨을 확인할 수 있었다. 본 연구에서는 해당 용어가 얼마나 자주 활용되는지를 분석하는 것으로, 전체 총합을 기준으로 삼았다. 본 연구에서는 중복 횟수별로 어휘자원을 그룹핑하여 그 차이를 파악하고 세부 특성을 분석하고자 한다. 이러한 차이를 검증하기 위해서, 실제 중복이 발생한 11,336건의 표제어를 중복 횟수에 따라 총 4개의 그룹으로 구분하였다. 〈표 5〉에서는 중복 횟수에 따른 그룹별 표제어 수와 비율을 살펴볼 수 있다. 그룹A, 그룹B, 그룹C, 그룹D는 각각 10회 이상, 7-9회, 4-6회, 2-3회 중복이 발생한 표제어들의 집합이다.

중복도 그룹별 어휘자원의 특성 도출을 위해 본 연구에서는 그룹별로 표제어 20개씩을 표본 표제어로 선정하였다. 20개의 표본 표제어를 추출한 이유는 중복 횟수가 높은 A그룹의 표제어가 총 28개이므로, 20개 내외로 추출된 표본 표제어가 대표성이 있다고 보았기 때문이다. 이 때 추출 대상이 되는 표제어는 재난안전 분

〈표 5〉 중복 횟수 그룹별 표제어 수/비율

구분	중복 횟수	표제어 수		비율(%)
그룹A	13회	5	28	0.25
	12회	2		
	11회	8		
	10회	13		
그룹B	9회	27	151	1.33
	8회	44		
	7회	80		
그룹C	6회	152	1,188	10.48
	5회	335		
	4회	701		
그룹D	3회	1,887	9,969	87.94
	2회	8,082		
합계		11,336		100

야와 관련성이 높은 표제어로, 그 중에서 동음이의어 측면에서 다른 분야에서 일반적으로 통용되는 표제어와 '땀', '독'과 같이 한 글자인 표제어는 추출 대상에서 제외하였다. 본 연구는 어휘자원의 활용도 분석을 위해 네이버 뉴스 및 구글트렌드 검색을 활용하였다. 다른 분야에서 일반적으로 통용되거나 한 글자인 표제어는 검색 결과에 재난안전 분야와 관계없는 정보들이 다수 포함될 것으로 보고, 제외하였다.

그룹A는 10회 이상 중복이 발생한 표제어들이다. 해당 그룹의 표제어는 총 28개밖에 되지 않기 때문에 대부분이 표본 표제어로 추출되었다. 추출된 표제어에는 '복구', '재해', '생태계' 등이 있다. 그룹B는 7-9회 중복이 발생한 표제어들이며 '지진', '태풍', '오염물질' 등이, 그룹C는 4-6회 중복이 발생한 표제어들이며 '가뭄', '폭우', '홍수' 등이 추출되었다. 그룹D는 2-3회 중복이 발생한 표제어들로 '조류인플루엔자', '감염병', '진앙' 등이 추출되었다. 이상으로 4개의 그룹에서 각각 20개씩 총 80개의 표본 표제

어를 선정하였다. 선정된 표본 표제어는 〈표 6〉에 상세히 기술되어 있다.

3. 재난안전 분야 어휘자원 활용도 분석

재난안전정보를 관리하는 1차적인 목적은 효율적인 공유 및 활용이다. 하지만 수집된 어휘자원들은 각 기관에서 사용되는 전문분야별 업무용어로서 전문용어와 일반용어가 혼재되어 있다. 국민과 기관 또는 각 기관 간의 효율적인 재난안전정보 공유 및 활용을 위해서는 혼재된 어휘자원들의 특징을 분석하고, 적절한 표준화 대상을 선정하여 표준화 작업을 수행해야 한다. 특히, 재난안전정보의 사용주체가 실제로 활용하는 재난안전 분야 어휘자원을 확인하고 그 특징을 파악할 필요가 있다. 이에 본 연구는 앞서 추출한 표본 표제어를 대상으로 정보제공자 및 정보이용자 관점에서의 활용도를 분석하였다.

〈표 6〉 그룹별 전체/최근1년 검색결과 건수

구분	용어+중복횟수	전체 검색결과 건수	최근1년 검색결과 건수	구분	용어+중복횟수	전체 검색결과 건수	최근1년 검색결과 건수
그룹 A	복구11	543,715	73,969	그룹 C	가뭇5	219,989	31,171
	재해11	513,087	88,671		폭우6	201,654	23,531
	생태계13	468,292	97,861		홍수4	194,289	26,315
	조류13	290,199	70,246		산불6	188,262	34,538
	소독10	194,028	39,904		구제역4	172,798	23,936
	온실가스10	187,631	27,178		폭설4	172,713	13,801
	동상13	183,868	30,773		집중호우4	151,078	17,846
	자외선10	157,066	25,495		산업재해4	64,219	10,069
	방사선13	132,642	17,883		해일4	56,130	6,252
	세계보건기구11	121,765	15,686		유해화학물질4	47,047	10,023
	지하수11	87,010	12,677		기상재해4	44,098	7,141
	온실효과11	57,723	8,643		해외재난4	34,086	7,087
	저기압10	48,143	6,713		우박5	21,233	2,581
	잠복기10	38,776	4,582		방화구조4	16,387	2,376
	질소산화물11	27,335	7,671		냉해5	10,404	641
	접지13	17,912	2,579		지진파4	7,203	1,337
	다이옥신10	12,593	1,332		비점오염원4	5,231	1,132
	난류10	8,101	1,027		내화구조6	4,345	608
	산성비10	4,247	338		홍수조절용량4	1,676	174
공동구10	3,329	533	광화학스모그4	712	66		
그룹 B	지진7	507,665	109,813	그룹 D	조류인플루엔자2	125,801	45,758
	태풍7	479,094	76,472		감염병3	99,803	27,510
	오염물질9	158,829	31,117		진앙3	35,677	9,183
	위협도7	158,774	25,559		가축전염병3	34,198	6,788
	방사능9	109,212	9,565		호우주의보3	33,170	3,104
	지구온난화8	107,147	12,919		방사능오염3	29,177	2,180
	대기오염9	104,563	24,220		내진설계3	28,316	14,061
	유해성7	56,581	14,820		도시홍수2	25,186	3,407
	방사성물질8	42,042	3,325		구난3	20,327	1,975
	방화벽7	39,397	3,053		특별재난지역2	19,594	5,254
	기후변화협약9	37,460	4,753		겨울가뭇2	17,883	1,954
	스모그7	27,991	6,019		재난지원금2	10,389	2,477
	방사성폐기물8	26,136	4,841		갈수기3	9,092	808
	위험물8	24,769	4,919		방화지구3	8,303	864
	환경호르몬7	22,394	2,881		설해3	7,340	1,099
	적조9	21,690	2,413		계획홍수위3	1,732	157
	엘니뇨7	13,041	1,825		동상해3	1,389	161
	부영양화9	4,321	388		사보타주3	1,178	147
	잔류성유기오염물질7	1,261	266		재해복구보조금3	1,066	108
	먼지회오리8	1,016	179		떠돌이태풍3	106	11

재난안전 분야 어휘자원의 사용주체별 활용도 분석을 위해 본 연구는 '네이버뉴스'¹⁾ 검색과 '구글트렌드'²⁾ 검색을 활용하였다. '네이버뉴스'와 '구글트렌드' 검색은 사용주체 측면에서 서로 다른 특징을 지닌다. 네이버뉴스 검색은 뉴스 기사를 작성하는 기자, 즉 정보를 생성하는 주체들이 사용한 어휘 양상을 대변하는 것으로, 정보제공자 측면에서 주로 어떠한 어휘자원이 재난안전 분야에서 활용되는가를 확인할 수 있다. 반면 구글트렌드 검색은 일반 이용자들이 입력한 키워드의 추이를 보여주는 것으로, 정보이용자 측면에서 일반 이용자들이 주로 어떠한 어휘자원에 관심을 가지는가를 유추할 수 있다.

보다 구체적인 분석방법을 설명하면, 네이버뉴스 검색을 통해서도 표본 표제어들을 검색했을 때 검색결과 건수가 어떠한지, 표제어 그룹간에 특별한 차이가 있는지를 분석하였다. 다음으로 각 그룹별로 네이버뉴스 검색결과 건수가 최상위인 표제어들을 대상으로 구글트렌드와의 비교검색을 실시하였다. 이 때 비교검색 결과는 해당 표제어에 대한 일반 이용자들의 관심도 차이를 나타낸다. 또한 구글트렌드를 통해 각 그룹별 최상위 표제어의 상대적인 검색량 변화를 시간의 흐름에 따라 가시화함으로써, 각 그룹별 표제어들이 지닌 특성을 분석하였다. 나아가 어휘 그룹별 특성 분석 결과를 토대로 재난안전 분야 어휘자원의 활용방안을 제시하고자 한다.

3.1 정보제공자 관점에서의 활용도 분석

네이버뉴스는 뉴스기사를 대상으로 키워드

검색기능을 제공하고 있다. 이를 통해 뉴스기사의 '전체 검색결과 건수'와 '최근1년 검색결과 건수'를 확인할 수 있다. 이 때 '검색결과 건수' 지표는 정보제공자 측면에서의 어휘자원 활용도를 의미한다. 정보제공자가 활용하는 어휘자원에 대해 식별하는 것은 중요하다. 그 이유는 정보제공자가 활용하는 어휘자원들이 재난안전 분야에서 중요한 용어이면서도 정보이용자의 재난안전정보 수집 시에 활용되는 중요한 접근점이기 때문이다. 이에 정보제공자의 어휘자원 활용도 분석을 위해 본 연구는 표본 표제어들을 중복 횟수에 따라 그룹화하고, 각 그룹별로 네이버뉴스 검색을 실시하였다. <표 6>은 각 그룹별 뉴스기사의 전체 및 최근 1년 검색결과 건수를 나타낸 표이다.

먼저 그룹A 표제어들 중 전체 검색결과 건수가 가장 높은 표제어는 '북구'로, 총 11개의 용어사전에 다양하게 정의되었으며, 뉴스 도메인에서는 약 54만 건이 검색되었다. 그 다음으로 '재해', '생태계', '조류'에 대해서 약 30만 건 이상의 뉴스기사가 검색되었다. 반면에, 그룹A에서 전체 검색결과 건수가 낮은 표제어는 '공동구', '산성비', '난류'이며, 약 3천에서 8천 건의 뉴스기사가 검색되었다. 또한 상위 표제어 중에서도 최근 1년 검색결과 건수를 통해 '생태계' 표제어가 최근 들어 많이 활용되었음을 확인할 수 있다.

다음으로는 그룹B 표제어들의 검색결과를 살펴보았다. 전체 및 최근 1년 검색결과 건수가 가장 높게 나타난 표제어는 '지진'으로 총 7개의 용어사전에 정의되었으며, 뉴스 도메인에서는 약 50만 건이 검색되었다. 이와 함께 '태풍'도 약

1) <http://news.naver.com/>

2) <https://trends.google.com/trends/>

48만 건의 검색결과 건수를 보여, 그룹A의 상위 표제어인 ‘복구’, ‘재해’, ‘생태계’의 검색결과와 큰 차이가 없음을 확인할 수 있다. 그룹B에서 검색결과 건수가 낮은 표제어는 ‘먼지회오리’, ‘잔류성유기오염물질’, ‘부영양화’이며, 약 1천 건에서 4천 건의 뉴스기사가 검색되었다.

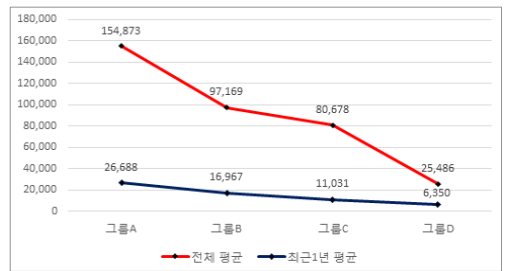
그룹C의 표제어들을 살펴보면, 전체 검색결과 건수가 가장 높은 표제어는 ‘가뭄’으로, 5개의 용어사전에 정의되었으며, 뉴스 도메인에서는 약 21만 건이 검색되었다. 앞서 살펴본 그룹A와 B의 최상위 표제어와 비교했을 때, 확실히 검색결과 건수가 많이 하락된 것을 확인할 수 있다. 그 다음으로 검색결과 건수가 높은 표제어들은 ‘폭우’, ‘홍수’, ‘산불’, ‘구제역’, ‘폭설’, ‘집중호우’이며, 검색결과 건수는 대체로 약 15만에서 20만 건으로 비슷한 수치를 보인다. 이 중에서도 ‘산불’은 최근 1년 검색결과 건수가 가장 높으므로, 최근 들어 많이 활용되는 표제어이다.

마지막으로 그룹D의 검색결과를 살펴보면 전체 및 최근 1년 검색결과 건수가 가장 높은 표제어는 ‘조류인플루엔자’로, 2개의 용어사전에 정의되었으며, 뉴스 도메인에서는 약 12만 건이 검색되었다. 마찬가지로 그룹A와 B와는 검색결과 건수에서 확연한 차이를 보이고 있다. 또한, 그룹D에서는 ‘조류인플루엔자’와 ‘감염병’을 제외한 모든 표제어의 전체 검색결과 건수가 약 4만 건 미만이다.

표본 표제어들을 대상으로 네이버뉴스 키워드 검색을 통한 각 그룹별 검색결과 건수를 종합 분석한 결과, 다음과 같은 두 가지 특징이 드러난다.

첫째, 중복 횟수가 높은 표제어는 정보제공자의 입장에서 활용도가 높으며, 재난안전 분

야에서 일반적으로 활용되는 어휘자원일 가능성이 크다. 분석 결과를 살펴보면, 고빈도 중복 그룹A에서 저빈도 그룹 D로 갈수록 검색결과 건수는 감소하고 있다. 각 그룹별 검색결과 건수가 최상위인 표제어들을 살펴보면 그룹A는 54만, B는 50만, C는 21만, D는 12만이다. 그룹간 간격에는 차이가 있지만 전체적인 검색결과 수치는 감소하고 있다. 보다 구체적으로 그룹별 검색결과 건수 차이를 확인하기 위해서, 그룹별 전체 및 최근 1년 검색결과 건수의 평균을 그래프로 표현하였다(〈그림 2〉 참조).



〈그림 2〉 그룹별 전체/최근 1년 검색결과 건수의 평균

〈그림 2〉의 그래프 결과를 보면 그룹A의 전체 검색결과 건수 평균은 그룹D에 비해 약 13만 건이 높으며, 최근 1년 검색결과 건수 평균도 그룹D에 비해 약 2만 건이 높다. 이를 통해 중복 횟수가 높은 표제어일수록 정보제공자 입장에서 활용도가 높음을 확인할 수 있다. 중복 횟수는 해당 표제어를 다양한 출처에서 정의하고 있다는 의미이며, 검색결과 건수는 실제로 해당 표제어가 얼마나 자주 활용되는가를 의미하는 지표이다. 따라서 다양한 출처에서 다뤄지면서도, 활용도가 높은 표제어는 비교적 일반적인 용어라고 할 수 있다. 즉, 네이버뉴스 검색

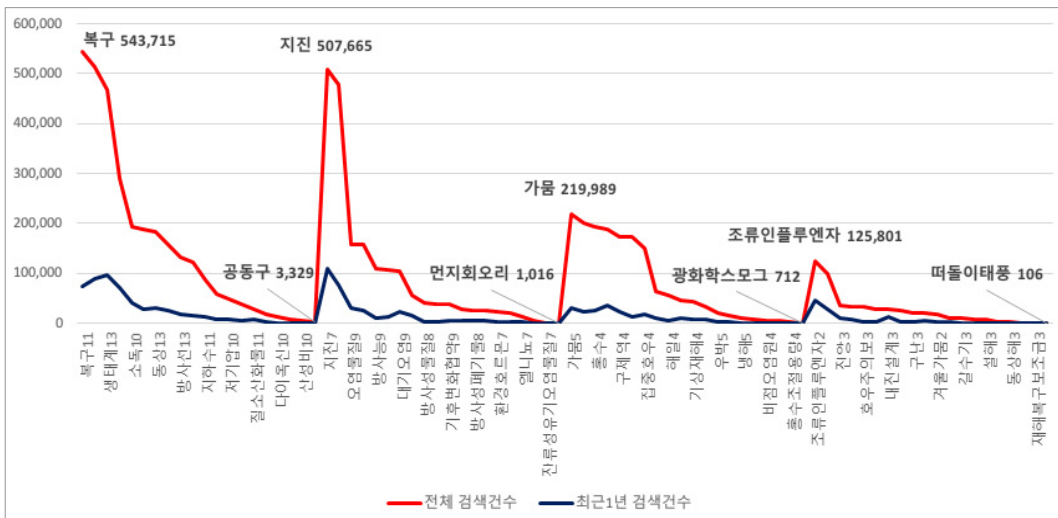
색 결과를 비추어 볼 때 그룹A에 포함된 표제어들은 ‘일반용어’일 가능성이 크며 반대로 그룹D에 포함된 표제어들은 ‘전문용어’일 가능성이 크다고 할 수 있다.

둘째, 각 그룹별로 검색결과 건수가 많은 표제어는 그룹 내 다른 표제어에 비해 재난안전 분야에서 많이 활용되는 일반용어일 가능성이 크다. <표 6>을 참조하여 각 그룹별 최상위와 최하위 표제어 간의 검색결과 건수 차이를 살펴 보았다. 그룹A의 최상위 표제어와 최하위 표제어는 54만 건과 3천 건, 그룹B는 50만 건과 1천 건, 그룹C는 21만 건과 7백 건, 마지막으로 그룹D는 12만 건과 1백 건이 검색되었다. 각 그룹별 최상위 표제어와 최하위 표제어 간 전체 검색결과 건수에는 큰 차이가 있으며, 이는 각 그룹별로 유사한 패턴을 보인다(<그림 3> 참조).

보다 구체적인 차이점을 분석하기 위해 <그림 3>과 같이 각 그룹별로 전체 검색결과 건수 상위 5개와 하위 5개의 표제어들을 비교하였다.

결과적으로 재난안전 분야에서 활용도가 높은 표제어들은 대체로 상위권에 분포되었음을 확인할 수 있다. 즉, 각 그룹의 상위 5개 표제어들은 대체로 재난안전관리법 및 NDMS 분류체계(소방방재청, 2005)에 등장하거나, 국립재난안전연구원(2015)에서 수행한 ‘재난안전교육체계 정립을 위한 재난유형 분류 및 표준화(이하, 재난유형 분류 및 표준화 연구)’의 분류체계 개발에 활용된 표제어들이다. 반면에, 각 그룹의 하위 5개 표제어는 재난안전관리법 및 NDMS 분류체계, 재난유형 분류 및 표준화 연구의 분류체계에서 전혀 활용되지 않는다.

법령 및 분류체계에 등장한 표제어들을 살펴 보면, 그룹A의 ‘복구’는 재난안전관리법에서 재난관리를 위한 예방, 대비, 대응, 복구 활동 중 하나로 명시되어 있다. ‘재해’는 재난안전관리법에 등장하는 표제어임과 동시에 ‘재난’과 함께 혼용되어 사용되는 표제어이다. ‘조류’는 재난안전관리법에서 ‘조류(藻類) 대발생’으로 활용되



<그림 3> 그룹별 네이버뉴스 검색결과 건수 비교

고 있다. 그룹B의 '지진'과 '태풍'은 재난안전관리법의 자연재난에 명시되어 있다. '방사능'은 재난유형 분류 및 표준화 연구에서 '방사능 재난'으로 활용되고 있다. 그룹C의 '가뭄'과 '홍수'도 재난안전관리법의 자연재난에 포함되며, '산불'은 재난유형 분류 및 표준화 연구의 분류체계에서 '화재 안전'의 세부 재난유형이다. 또한 '폭우'와 유사한 개념으로 재난안전관리법 상에서는 '호우'를 명시하고 있다. 마지막으로 그룹D의 '감염병'과 '가축전염병'은 재난안전관리법의 사회재난에 포함되며, '호우주의보'는 재난안전관리법 상의 '호우'로 볼 수 있다.

한편, 법령 및 분류체계에서 활용되는 표제어들이 각 그룹의 상위권에 분포되어 있음을 확인하였다. 이는 동일 그룹 내에서도 일반용어와 전문용어가 구분되며, 그룹 내 상위권 표제어가 일반용어일 가능성이 크다는 것을 시사한다. 이때, 재난안전 분야의 일반용어가 각 그룹별로 발생하는 이유는 재난유형별 용어사전 분포를 통해 유추할 수 있다. 앞서 설명한 바와 같이 본 연구에서는 재난안전 유관기관 33개의 용어사전과 관련 문헌에서 어휘자원을 수집했다. 이때 재난안전 유관기관들은 기상, 수자원, 원자력, 정보통신, 산림, 교통, 전기, 보건, 환경, 농림축산 등 각각의 재난 유형 분야에 전문기관인 경우가 대부분이다. <그림 1>에서도 나타나듯이 '재난공통' 영역에 해당하는 용어사전보다는 특정 유형의 '재난전문' 용어사전이 70%로 다수를 차지한다. 이에 따라 재난 영역별로 표제어들 간에 간격이 발생하기 때문에 각 영역별로도 일반용어와 전문용어가 구분되는 것으로 해석된다.

결론적으로 중복도 분석 및 네이버뉴스 검색

을 통해서 정보제공자 관점에서 주로 활용되는 재난안전 분야 일반용어 및 전문용어를 식별할 수 있음을 확인하였다. 이를 통해 재난안전 분야 용어 표준화 작업의 목적에 따라 일반용어 및 전문용어별로 표준화 작업의 우선순위를 정할 수 있다.

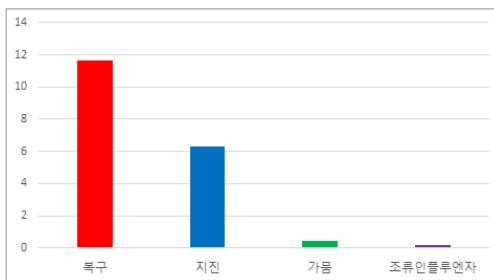
3.2 정보이용자 관점에서의 활용도 분석

구글트렌드에서는 정보이용자들이 검색에 활용한 키워드 및 검색빈도 등의 데이터를 분석하여, 시간별·지역별 관심도, 연관검색어, 연관주제 리스트 등의 정보를 제공하고 있다. 구글트렌드는 실제 일반 이용자들이 정보검색의 과정에서 입력한 질의어를 대상으로 하는 것으로, 정보이용자 측면에서의 어휘자원 활용도를 파악할 수 있다. 이는 앞서 네이버뉴스가 정보제공자 관점에서의 활용도를 함의하는 특성과 대비된다. 본 연구에서는 구글트렌드가 제공하는 다양한 기능 중에서도 '시간의 흐름에 따른 관심도 변화' 기능을 활용하였으며, 해당 기능은 여러 표제어들을 동시 검색함으로써 시간 흐름에 따른 표제어 관심도 추이를 한 눈에 비교할 수 있게 한다.

중복 횟수에 따른 재난안전 분야 어휘자원이 지닌 특성과 활용도를 파악하기 위해, 앞서 실시한 각 그룹별 네이버뉴스 검색결과 건수가 최상위인 표제어들을 대상으로 구글트렌드 비교검색을 실시하였다. 이 때, 비교검색을 위한 표제어 선정은 정보제공자와 정보이용자 간의 어휘자원 활용양상에 대한 차이를 보다 가시적으로 비교하고자, 네이버뉴스 검색에 사용된 표제어 중 각 그룹에서 검색결과 건수가 최상위인

표제어들을 대표어로 선정하였다. 이를 통해 다음과 같은 특징이 도출되었다.

첫째, 정보이용자는 재난안전 분야의 일반용어에 대하여 관심이 높은 것으로 분석되었다. <그림 4>를 확인하면 관심도는 복구가 가장 높으며 지진, 가뭄, 조류인플루엔자 순으로 관심이 점차 감소하는 것을 확인할 수 있다. <그림 4>와 <그림 2>는 유사한 패턴을 가진다. <그림 2>에서도 그룹A부터 D까지 중복 횟수가 줄어들수록 뉴스 검색결과 건수도 감소하고 있다.



〈그림 4〉 그룹별 최상위 표제어의 관심도 평균

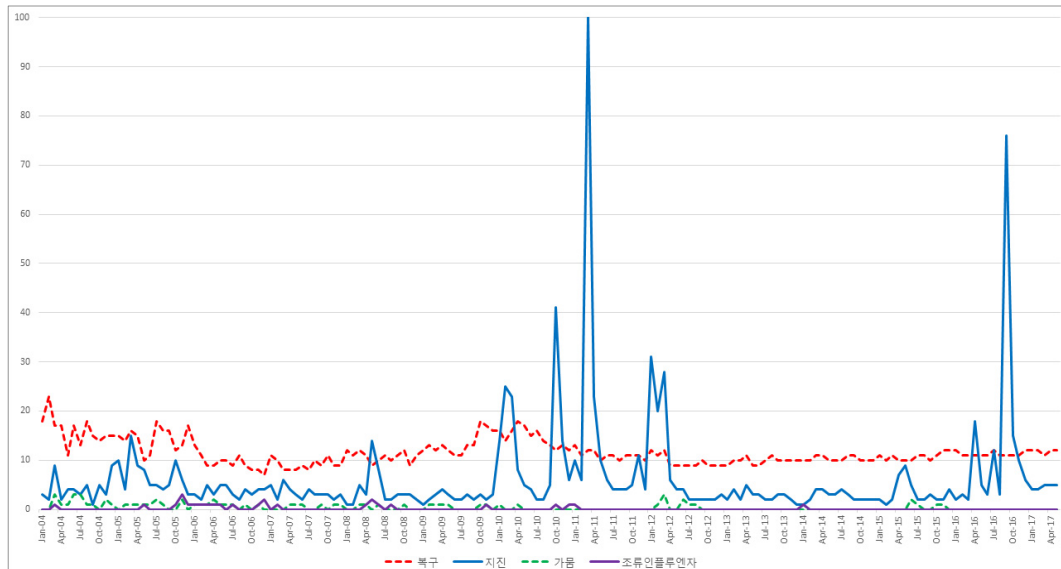
앞서 정보제공자 측면에서 뉴스검색 결과 건수를 분석할 결과, 중복 횟수가 높은 그룹A는 재난안전 분야에서 활용되는 일반용어, 그룹D는 전문용어일 가능성이 크다고 하였다. 따라서 정보제공자가 활용하는 어휘자원과 정보이용자가 관심을 가지는 어휘자원은 어느 정도 비례 관계에 있다고 할 수 있다. 즉, 정보이용자는 재난안전 분야의 일반용어에 대한 관심이 비교적 더 높다.

둘째, 특정 재난 발생여부에 따라 정보이용자가 관심을 갖는 표제어를 파악할 수 있었다. 분석 결과 특정 재난이 발생했을 경우, 해당 재난과 관련된 표제어 검색 빈도가 증가함을 확

인할 수 있었다. 반면, 재난안전 분야에서 상시적으로 활용되는 표제어의 검색 빈도는 일정한 수준을 유지한다. 이에 대한 구체적인 분석 내용은 아래와 같다.

본 연구에서는 정보이용자들이 재난안전과 관련해 상시적으로 활용하는 표제어(이하 ‘공통 표제어’)와 특정 재난 상황과 관련된 표제어(이하 ‘사건 표제어’) 간 시간의 흐름에 따른 관심도 변화를 비교하였다. 비교검색은 <그림 4>와 동일하게 각 그룹별 네이버뉴스 검색결과 건수가 최상위인 ‘복구’, ‘지진’, ‘가뭄’, ‘조류인플루엔자’ 4개 표제어를 대상으로 실시하였다. 비교검색 결과는 <그림 5>에 나타나 있으며, 2004년 1월부터 2017년 5월까지의 관심도 변화 추이를 보여주고 있다. 그래프에 나타난 수치는 단순한 검색빈도 결과가 아니라, 특정 기간 동안 검색된 표제어의 상대적인 관심도를 나타낸다. 비교검색 결과 100은 최고 관심도를, 100을 기준으로 50은 관심도가 절반 수준임을 의미한다. 예를 들어, <그림 5>에서는 2011년 3월(Mar-11)에 ‘지진’ 표제어에 대한 관심도가 100으로 가장 높다. 이는 곧 검색된 4개의 표제어 중 지진의 관심도가 가장 높으며, <표 5>는 2011년 3월의 ‘지진’ 표제어의 관심도를 중심으로 나머지 표제어들의 관심도가 표현된 결과이다.

<그림 5>에서는 ‘복구’ 표제어에 대한 관심도 변화폭이 크지 않음을 확인할 수 있다. ‘복구’에 대한 일반 이용자들의 관심이 일정한 이유는 해당 표제어가 재난안전 관리단계에서 상시적으로 활용되기 때문이다. 더불어 <그림 4>에서 관심도 평균이 가장 높은 표제어임을 고려해볼 때, ‘복구’는 정보이용자들이 꾸준히 검색하는 고빈도어이다. 반면, ‘지진’은 실제 해당 재난 상황이



〈그림 5〉 그룹별 최상위 표제어의 시간흐름에 따른 관심도 변화 비교

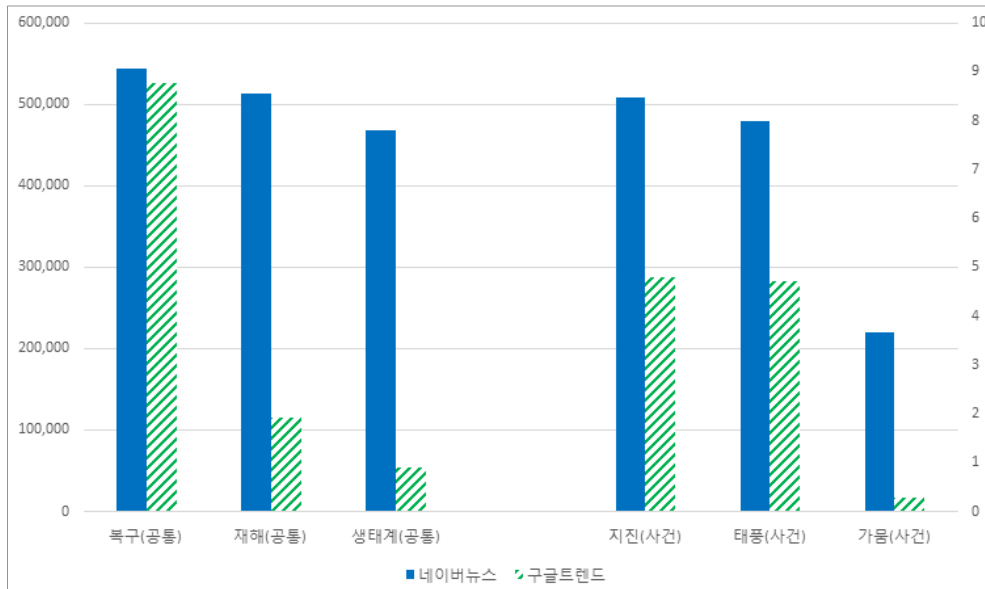
발생했을 때, 정보이용자의 관심도가 크게 향상되었으며, 검색 빈도도 함께 증가했다.

〈그림 4〉에서는 ‘지진’의 관심도 평균이 ‘복구’와 비교했을 때 낮게 나타났지만, 〈그림 5〉에서는 특정 기간 동안 ‘지진’에 대한 관심도가 ‘복구’보다 높게 나타났다. 특히, 2011년 3월에 지진에 대한 일반 이용자들의 관심도(관심도 100)가 큰 폭으로 증가했음을 확인할 수 있다. 이 때 실제로 ‘동일본 대지진’이라는 재난 상황이 발생하여 일본에 큰 피해를 입혔으며, 이와 동반하여 정보이용자들의 검색 빈도도 증가했다. 뿐만 아니라 불규칙적으로 그래프가 치솟은 다른 구간을 살펴보면, 2010년 10월 수마트라 대지진, 2016년 9월 경주 지진 등 특정 재난 상황이 발생했을 때 일반 이용자의 관심도가 증가함을 확인할 수 있다.

분석 결과를 정리해보면 ‘복구’와 같이 정보 이용자들이 상시적으로 사용하는 재난안전 관

련 공통 표제어에 대한 관심도는 시간의 흐름에 따라 비교적 일정한 패턴을 보이며 검색빈도가 꾸준히 높았다. 반면 ‘지진’과 같이 실제 재난 상황과 관련된 사건 표제어는 공통 표제어에 비해 평균 검색빈도는 낮을 수 있지만, 재난 상황 발생 시에는 정보이용자들의 관심도가 급격히 증가한다.

셋째, 정보제공자 및 정보이용자가 자주 사용하는 어휘자원 간에 차이를 발견할 수 있었다. 실제로 어휘자원의 사용주체에 따라 활용되는 표제어에 차이가 있는지를 파악하기 위하여, 표본 표제어들을 대상으로 네이버뉴스 검색결과 건수와 구글트렌드 관심도 평균을 비교 분석하였다. 특히 공통 표제어와 사건 표제어를 구분하여, 어휘자원 사용주체와 이들 표제어 간의 상관관계를 분석하였다. 〈그림 6〉은 〈표 6〉에서 그룹에 관계없이 네이버뉴스 검색결과 건수가 높은 공통 표제어 3개(‘복구’, ‘재



〈그림 6〉 공통표제어와 사건표제어의 네이버뉴스 검색결과 건수와 구글트렌드 관심도 평균 비교

해’, ‘생태계’)와 사건 표제어 3개(‘지진’, ‘태풍’, ‘가뭄’)를 선정, 총 6개 표제어를 대상으로 네이버뉴스 검색결과 건수와 구글트렌드 관심도 평균을 비교한 그래프이다.

〈그림 6〉을 살펴보면 공통 표제어 ‘복구’는 재난안전 분야에서 일반적이고 상시적으로 사용되는 표제어로서, 전체 6개 표제어 중 네이버뉴스와 구글트렌드 검색 결과에서 모두 높은 수치를 보이고 있다. 한편, 공통 표제어 ‘재해’, ‘생태계’와 사건 표제어 ‘지진’, ‘태풍’은 네이버뉴스 검색결과 건수가 서로 유사하다. 이는 정보제공자 측면에서 위 4개의 표제어들이 비슷한 활용도를 가짐을 시사한다.

정보이용자 측면에서 활용도가 높은 표제어를 분석하기 위해 구글트렌드 검색결과 평균을 살펴보았다. 비교 값에 대한 타당성을 위해 네이버뉴스 검색결과가 유사한 재해, 생태계, 지

진, 태풍 4건의 표제어를 대상으로 관심도의 평균을 구했다. 그 결과 공통 표제어 관심도 평균(1.40)이 사건 표제어 관심도 평균(4.75)보다 낮게 나타났다. 이는 정보이용자 측면에서는 공통 표제어보다 사건 표제어인 ‘지진’, ‘태풍’의 활용도가 높다는 것을 의미한다. 따라서 정보제공자와 정보이용자가 실제로 활용하는 재난안전관련 어휘자원 간에는 다소 차이가 존재하며, 재난안전 분야 용어 표준화 작업을 위한 우선순위 선정 시에는 이러한 차이점을 반영해야 한다.

이상의 어휘 활용도 분석을 통해 재난안전 분야 어휘자원의 특징을 파악하였다. 결론적으로 중복도 그룹 또는 용어 유형별 차이에 따라 다음과 같은 특성이 도출되었다. 먼저, 재난안전 분야 어휘자원 중 중복 횟수가 높은 표제어들은 ‘일반용어’이며, 대체로 정보제공자 및 정

보이용자 측면에서 모두 유용한 정보로서 활용도가 높은 것으로 분석되었다. 반면 중복도 횡수가 낮은 표제어들은 '전문용어'일 가능성이 높은 어휘로, 재난안전 관련 업무 시 실무자들이 사용하면서도 분야별 전문성을 띄는 표제어인 것으로 파악되었다. 즉, 재난안전 유관기관에서 사용하는 용어사전 간의 중복도는 재난안전 분야에서의 활용도로 해석할 수 있다. 용어사전 간에 중복해서 발생하는 용어는 여러 기관에서 사용하는 용어로서 재난안전 분야에서 일반적으로 활용되는 단어라고 유추할 수 있다. 반대로 특정 기관의 용어사전에만 나타나는 용어라면 해당 기관에서 다루는 특정 재난 유형에만 국한된 용어라고 유추할 수 있다. 다만, 중복도가 동일한 표제어 그룹에도 일반용어와 전문용어가 혼재되어 나타났는데, 이는 재난안전 유관기관들이 특정 재난유형별로 전문화되어 있기 때문에 나타난 현상이다. 마찬가지로 중복도가 낮은 그룹에도 재난안전 일반용어가 포함되어 있음을 인지해야 한다. 이에 대한 세부적인 분석은 향후 연구로 진행할 계획이다.

또한 특정 재난 발생여부에 따라 정보이용자가 관심을 가지는 표제어에는 차이가 있다. 그중 '공통 표제어'는 재난안전에서 상시적으로 활용되는 표제어로, 특정 재난 상황 발생 여부와는 무관하게 일정한 패턴으로 꾸준히 활용되는 특징을 가진다. 반면, '사건 표제어'는 특정 재난 상황과 관련된 표제어로 재난 상황 발생 시 정보이용자의 관심도와 검색량이 급격히 증가하는 특징을 가진다.

본 연구에서 수행한 어휘자원의 중복도 분석은 차후 재난안전 분야 용어 표준화 대상 선정 시 중요한 기준이 될 수 있다. 다음은 분석 결과

를 토대로 재난안전 분야 어휘자원을 어떻게 활용할 수 있을지 구체적인 활용방안을 제안하고자 한다.

4. 재난안전정보 관리를 위한 어휘자원 활용방안

급변하는 환경 속에 재난이 사람에게 미치는 영향이 갈수록 커지는 만큼 재난 관련 정보에 대한 국민들의 관심은 점점 높아지고 있다. 재난 상황 발생 시 여러 기관으로부터 생산된 다양한 정보들을 공유하고 국민들에게 효과적으로 전달하기 위해서는 표준화된 재난안전 분야 용어사전이 무엇보다도 필요하다. 본 연구에서는 재난안전 분야 용어 표준화를 위한 기초자료로서, 재난안전 분야의 어휘자원을 조사 및 수집하여 현황을 분석하고 활용방안을 제시하고자 한다.

첫째, 재난안전 분야 어휘자원은 업무활용성이 높은 '전문용어'와 많이 활용되는 '일반용어'로 구분할 수 있다. 이는 각각의 어휘자원을 사용할 대상을 고려하여 표준화 작업을 수행해야 함을 시사한다. 전문용어의 주 사용주체는 해당 재난안전 업무를 수행하는 기관, 업무담당자 또는 전문가이다. 정보 공유 및 활용이 빈번하게 발생하는 재난안전정보의 특성상 다양한 기관 및 실무자들이 어휘자원에 접근하게 된다. 따라서 전문용어 표준화 작업은 의미적인 모호성을 제거하고 용어를 통일하는 과정에 집중해야 한다. 반면 일반용어의 경우에는 재난안전정보에 접근하고자 하는 모든 이용자가 서비스 대상이며, 수집된 어휘자원들 중 중복도 및 활용

도가 높은 표제어들을 중심으로 표준화 작업을 수행해야 한다. 뿐만 아니라 전문용어 중에서도 한자어나 외래어로 구성되어 해석하기 어려운 용어나 전문성이 너무 높아 일반 이용자의 접근성이 낮은 어휘자원에 대해서 '일반용어화' 하는 작업도 요구된다. 또한 앞서 분석한 결과를 통해서 정보제공자 및 정보이용자가 주로 활용하는 용어에 차이가 있음을 파악하였다. 이러한 문제를 해결하기 위해서는 정보제공자가 정보이용자 관점에서 자주 활용되는 용어를 사용하도록 유도해야 한다. 실제로 해양수산부는 2017년 1월 18일 「해양수산 분야 전문용어 표준화」 고시를 시행하여, 법령 제·개정, 공문서, 보도자료, 해양수산 현장에서 활용되던 전문용어에 대해 일반용어로의 순화를 촉구하였다. 이처럼 재난안전 분야의 어휘자원은 궁극적으로 대국민 서비스를 목적으로 표준화되어야 한다.

둘째, 재난안전정보에 대한 접근성 및 검색의 정확성을 향상시키기 위하여 재난안전 분야 어휘자원을 활용한 시소러스를 구축할 수 있다. 재난안전정보 관리의 궁극적인 목적은 재난안전정보의 공유 및 활용에 있다. 특히 국민이 보다 쉽게 재난안전정보에 접근할 수 있도록 접근점을 제공하는 것은 재난안전정보 관리를 위한 핵심적인 기능이다. 재난안전 분야 시소러스 구축을 위해서는 어휘자원의 활용성과 재난 유형별 전문성 두 가지 측면을 고려해야 한다. 본 연구에서는 중복도 및 활용도 분석을 통해, 재난안전 분야 어휘자원 중 활용성이 높은 표제어들을 확인하였다. 이 때 활용성이 높은 표제어는 시소러스 구축을 위한 중심어가 될 수 있으며, 중심어로부터 상위어, 하위어, 관련어

등의 용어 관계를 설정해야 한다. 또한 본 연구에서는 정보이용자 관점에서의 활용도 분석을 통해, '지진'과 '태풍'과 같은 실제 재난 상황이 발생하였을 경우, '사건 표제어'에 대해서 이용자 관심도가 크게 향상하는 것을 확인하였다. 따라서 정보이용자들은 재난 유형별로 시소러스에 접근할 가능성이 크며, 시소러스 구축 시 재난 유형별 전문성을 고려해야 한다.

셋째, 재난안전 유관기관에서 생산되는 재난안전정보의 체계적 관리를 목적으로 개발되는 분류체계에 재난안전 분야 어휘자원을 활용할 수 있다. 다수의 재난안전 유관기관 및 연계 시스템으로부터 산출되는 다양한 종류의 재난안전정보를 통합 관리하는 과정에서 상호호환성의 문제가 발생할 수 있다. 어휘자원의 표준화는 이러한 문제를 해결하기 위한 가장 기초적인 작업이라고 할 수 있다. 본 연구에서는 각각의 재난안전 유관기관별로 가지는 전문적인 특성으로 인해, 재난 유형별 중복이 적게 발생하는 현상을 살펴보았다. 또한 네이버뉴스 검색결과 건수 분석을 통해 각 그룹별로 비교적 활용도가 높은 재난관련 '일반용어'가 존재함을 확인한 바 있다. 따라서 재난안전 분야의 표준화된 분류체계 개발을 위해서는 우선 법령이나 표준 및 재난안전 분야에서 수행된 다양한 연구들을 참고하여 기본적인 프레임워크를 설계해야 한다. 이후 활용도가 높은 어휘자원을 선정하여 분류체계 개발에 활용해야 한다. 개발된 분류체계는 재난안전정보의 통합적이고 체계적인 관리에 이바지할 수 있다. 나아가 정보제공자 및 정보이용자 측면 모두에서 활용성이 높은 어휘자원을 분류체계 정의에 사용하는 것은 재난안전정보에 대한 접근성 향상에 기여할 것이다.

5. 결 론

본 연구는 미래창조과학부 및 국민안전처에서 수행하는 재난안전정보 표준화사업의 일환으로 진행되었으며, 재난안전정보의 통합적 관리를 위한 어휘자원의 표준화에 앞서 기본적인 현황 분석을 통해 어휘자원의 활용방안을 찾는 데 그 목적이 있다. 이를 위해 재난안전 분야 어휘자원에 대한 현황을 온라인으로 분석하였으며, 실제로 현장방문조사 및 전화연락을 통해 수집된 용어사전을 대상으로 중복도 및 활용도 분석을 수행하였다. 분석 방법으로는 정보제공자 및 정보이용자별로 어휘자원 활용도를 분석하고자 네이버뉴스 및 구글트렌드 검색을 활용하였다.

분석 결과 중복 횟수가 높은 표제어는 정보제공자 측면에서 활용도가 높으며, 재난안전 분야에서 일반적으로 활용되는 어휘자원임을 파악할 수 있었다. 이는 각각의 어휘자원을 사용할 때를 고려하여 표준화 작업을 수행할 수 있음을 시사하며, 궁극적으로는 대국민 서비스를 목적으로 재난안전 분야 어휘자원들이 표준화되어야 한다. 또한 정보이용자는 재난안전 분야의 일반용어에 대해 관심도가 높았으며, 재난 발생 여부에 따른 정보이용자의 주요 관심 표제어를

파악할 수 있었다. 특히 정보이용자는 재난사건과 관련된 표제어 즉, 재난 유형 표제어를 활용하여 재난안전정보에 접근할 가능성이 높다. 따라서 정보제공자가 사건 표제어와 같이 정보이용자 관점에서 자주 활용되는 어휘자원을 사용하도록 촉구해야 한다.

본 연구는 재난안전 분야 전 범위의 어휘자원을 수집하여 분석한 결과를 바탕으로 어휘자원의 활용방안을 제안하였다는 데 의의가 있다. 특히 재난안전정보의 효율적인 공유 및 활용을 위한 용어 표준화 측면에서 활용할 수 있는 방안을 제안하였으며, 궁극적으로는 대국민 서비스를 위하여 표준화되어야 한다. 향후 연구로 재난안전 분야 어휘자원의 사용 양상을 소셜미디어를 활용하여 분석하고자 한다. 이로써 재난안전 분야에서 일반 이용자가 활용하는 용어와 기존 어휘자원 간의 차이점을 비교·분석하고, 재난안전 분야 용어사전의 표준화 방안을 구체적으로 모색하고자 한다. 또한 재난안전 분야 연구자와 현업 종사자 등 전문가 집단의 어휘자원 활용도 분석을 실시하여, 용어사전의 표준화를 도모하고자 한다. 이를 통해 공유와 활용이 빈번하게 발생하는 재난안전정보에 대하여, 실무상에서 생길 수 있는 의미적 모호성을 제거할 수 있을 것이다.

참 고 문 헌

- 강성경, 유환, 이영재 (2016). 텍스트 마이닝과 소셜 네트워크 분석을 이용한 재난대응 용어분석. 정보시스템연구, 18(1), 141-155.
- 강성경, 이영재 (2015). 한국형 재난현장 통합대응 표준체계 개발연구. 한국방재안전학회 논문집, 8(2), 1-10.

- 국립재난안전연구원 (2013). 비정형 빅데이터 수집 및 분석모델 개발. 울산: 국립재난안전연구원.
- 국립재난안전연구원 (2015). 재난안전교육체계 정립을 위한 재난유형 분류 및 표준화. 울산: 국립재난안전연구원.
- 국민안전처 (2016). 국가안전관리기본계획(2015~2019). 세종: 국민안전처.
- 서진순, 김상균, 김안나, 장현철 (2015). 한의학 표준용어집 서비스와 향후 발전 방향. 한국콘텐츠학회, 15(11), 571-581. <https://doi.org/10.5392/jkca.2015.15.11.571>
- 소방방재청 (2005). 국가안전관리 2단계 BPR/ISP 사업완료보고서. 서울: 소방방재청.
- 신윤주, 김정은 (2009). 건강정보 소비자 용어 시스템 구축을 위한 기반연구. 대한의료정보학회지, 15(1), 31-40. <https://doi.org/10.4258/jksmi.2009.15.1.31>
- 왕순주 (2008). 보건의료 분야의 표준화 활동: 용어표준화를 통한 국가보건의료정보 구축의 예. 한국재난관리표준학회지, 1(1), 11-16.
- 한국방재학회 (2015). 재난안전 용어사전. 서울: 씨아이알.
- 한국방재협회 방재사전편찬위원회 (2010). 방재사전. 서울: R&D프레스.
- Mayner, L., & Arbon, P. (2015). Defining disaster: The need for harmonisation of terminology. *Australasian Journal of Disaster & Trauma Studies*, 19, 21-25.

[참고웹사이트]

- 구글트렌드. Retrieved from <https://trends.google.com/trends/>
- 국가민방위재난안전교육원. Retrieved from <http://www.ndti.go.kr/>
- 네이버뉴스. Retrieved from <http://news.naver.com/>

[참고 법령 및 규정]

- 「재난 및 안전관리기본법」
- 「재난 및 안전관리기본법 시행령」
- 「해양수산 분야 전문용어 표준화 고시」

• 국문 참고문헌에 대한 영문 표기
(English translation of references written in Korean)

- Kang, Seong Kyung, & Lee, Young Jai (2015). A study on developing Korean standardized integrated response system for disaster site. *Korean Society of Disaster & Security*, 8(2), 1-10.
- Kang, Seong Kyung, Yu, Hwan, & Lee, Young Jai (2016). Analyzing disaster response terminologies

- by text mining and social network analysis. *Information Systems Review*, 18(1), 141-155.
- Korea Disaster Prevention Association (2010). *Dictionary of disaster terminology*. Seoul: R&D Press.
- Korean Society of Hazard Mitigation (2015). *The glossary for disaster and safety*. Seoul: CIR.
- Ministry of Public Safety and Security (2016). *National safety management plan(2015~2019)*. Sejong: Ministry of Public Safety and Security.
- National Disaster Management Research Institute (2013). *Development of collection and analysis models for unstructured bigdata of disaster-related*. Ulsan: National Disaster Management Research Institute.
- National Disaster Management Research Institute (2015). *Standard classification of disaster for establishing customized disaster safety education system*. Ulsan: National Disaster Management Research Institute.
- National Emergency Management Agency (2005). *The final report on national safety management 2-level BPR/ISP*. Seoul: National Emergency Management Agency.
- Seo, Jinsoon, Kim, Sang-Kyun, Kim, Anna, & Jang, Hyunchul (2015). A service of Korean medical standard terminology and directions for future development. *The Journal of the Korea Contents Association*, 15(11), 571-581. <https://doi.org/10.5392/jkca.2015.15.11.571>
- Shin, Yoonju, & Kim, Jeongeun (2009). A study for building a system of consumer vocabulary for health information. *Healthcare Informatics Research*, 15(1), 31-40. <https://doi.org/10.4258/jksmi.2009.15.1.31>
- Wang, Soon-Joo (2008). Standardization activities on public health: The example of the implementation of national public health information by the terminology standardization. *Korean Society of Disaster Management Standard*, 1(1), 11-16.