

# 소셜빅데이터 수집 및 분석을 위한 아동청소년 학교폭력 온톨로지 개발

## Ontology Development of School Bullying for Social Big Data Collection and Analysis

한윤선\*, 김하영\*\*, 송주영\*\*\*, 송태민\*\*\*\*

서울대학교 사회복지학과\*, 한국청소년정책연구원\*\*, Pennsylvania State University-Schuylkill\*\*\*, 삼육대학교 보건관리학과\*\*\*\*

Yoonsun Han(yshan@snu.ac.kr)\*, Hayoung Kim(hayoungjoykim@gmail.com)\*\*,  
Juyoung Song(jxs6190@psu.edu)\*\*\*, Tae Min Song(tmsong@syu.ac.kr)\*\*\*\*

### 요약

소셜빅데이터는 아동청소년의 학교폭력 현상에 대한 풍부하고 다각적 시각을 제공할 수 있지만, 복잡하고 다양한 비정형 텍스트로 구성되어 있어 자료의 체계적인 수집과 활용이 어렵다. 소셜빅데이터의 수많은 정보 가운데 의미 있는 개념을 추출하고 자료를 효과적으로 수집하기 위해서는 연구주제에 상응하는 핵심용어를 명시하고, 해당 개념 간의 관계를 나타내주는 온톨로지의 역할이 중추적이다. 본 연구는 온톨로지의 개념을 정리하고, 7단계에 걸친 온톨로지 개발 과정을 구체적으로 설명한 후, 학교폭력 소셜빅데이터 수집 및 분석을 위한 온톨로지 구축에 적용하였다. 그 결과, 학교폭력의 대상, 원인, 유형, 장소, 지역, 대응방안 6가지 영역을 중심으로 최상위 계층인 대분류를 구성하고, 중분류 및 소분류 체계를 거쳐 총 2,400여 개의 핵심용어를 도출하였다. 본 연구의 의의는 온톨로지 수집 및 개발과정을 설명하고, 기존의 연구방법과는 다소 차별을 두는 소셜빅데이터를 활용한 연구모형을 학교폭력 연구에 제시하였다는 것이다. 소셜빅데이터 분석의 기초가 되는 온톨로지 개발 연구는 좁게는 학교폭력 대상자들에 대한 이해를 제고시킬 뿐 아니라, 거시적으로는 학교폭력이라는 사회현상을 바라보는 한국사회의 시각에 대한 정보를 제공하는 실천적 함의가 있다.

■ 중심어 : | 아동청소년 | 학교폭력 | 소셜빅데이터 | 온톨로지 |

### Abstract

Although social big data can provide a multi-faceted perspective on school bullying experiences among children and adolescents, the complexity and variety of unstructured text presents a challenge for systematic collection and analysis of the data. Development of an ontology, which identifies key terms and their intricate relationships, is crucial for extracting key concepts and effectively collecting data. The current study elaborated on the definition of an ontology, carefully described the 7 stage development process, and applied the ontology for collecting and analyzing school bullying social big data. As a result, approximately 2,400 key terms were extracted in top-, middle-, and lower-level categories, concerning domains of participants, causes, types, location, region, and intervention. The study contributes to the literature by explaining the ontology development process and proposing a novel alternative research model that uses social big data in school bullying research. Findings from this ontology study may provide a basis for social big data research. Practical implications of this study lie in not only helping to understand the experience of school bullying participants, but also in offering a macro perspective on school bullying as a social phenomenon.

■ keyword : | Child and Adolescent | School Bullying | Social Big Data | Ontology |

\* 이 논문은 2016년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임(NRF-2016S1A5A2A03925702).

접수일자 : 2019년 03월 25일

심사완료일 : 2019년 05월 20일

수정일자 : 2019년 05월 20일

교신저자 : 송주영, e-mail : jxs6190@psu.edu

## I. 서론

학교폭력 경험은 아동청소년들에게 비행, 자살, 사회적 부적응이라는 심각하고 장기적인 결과를 초래하고 있다[1][2]. 따라서 학교폭력을 아동청소년기의 불안정한 행동에 기인한 일반적·일시적 현상으로 여기는 등 문제의 심각성을 결코 간과해서는 안 된다[3]. 최근 학교폭력의 시작연령은 점차 낮아지고 있으며 학교폭력 방법들도 다양해지고 있다[4][5]. 특히 기술의 진보로 인해 SNS(Social Network Service)와 같은 매체를 통해 아동청소년들은 새로운 학교폭력 방법들을 생산하고 이를 빠르게 확산시킬 수 있게 되었다. 한국 고등학생의 스마트폰 보유율은 98.7%이고[6], 한국 청소년들이 온라인상에서 보내는 시간이 주 평균 16.9시간이라는 보고서 내용은[7] 한국 청소년들의 높은 디지털 기기 사용량을 입증한다. 한국 청소년 10명 중 2명은 온라인 폭력을 경험한 것으로 보고되어 실제로도 많은 아동청소년들이 온라인상에서 학교폭력 경험에 노출되어 있음을 보여준다[8]. 더욱이, 현대 아동청소년들의 학교폭력은 온라인 폭력과 전통적 폭력이 중첩적이고 순환적으로 행해지고 있다는 중요한 특징을 보이고 있다[9][10].

이처럼 높은 온라인 학교폭력 경험율과 변화하는 학교폭력 양상에도 불구하고, 대부분의 학교폭력 연구들은 여전히 기존의 설문지 등과 같은 전통적인 사회조사 방법론에 의존하고 있다. 하지만 전통적 사회조사 방법론은 한 번에 조사할 수 있는 변인의 개수가 한정적이며, 회상에 의한 기억에 의존하여 개념을 측정하고, 시간상으로 연속적인 데이터를 수집하는 데에 어려움을 지니고 있어 급격히 변화하는 학교폭력 양상을 신속하게 담지 못하는 한계를 지니고 있다.

소셜빅데이터는 방대한 양의 정보를 연속적으로 수집할 수 있다는 점에서 사회과학 연구의 새로운 대안으로 제시될 수 있다[11]. 온라인상의 글, 댓글, 뉴스 등을 포함하는 자료의 집합체인 소셜빅데이터는 장소, 방법, 원인 등과 같은 학교폭력의 복합적 차원들을 모두 포함하고 있고, 아동청소년들의 학교폭력 현상을 설명할 풍부하고 다각적 정보들을 제공하고 있다[12]. 소셜빅데이터 분석을 통해 학교폭력의 현상만을 기술하는 것뿐

아니라(텍스트 마이닝), 사회적 문제로 학교폭력을 대하는 일반인들의 반응 및 의견도 파악할 수 있다(오피니언 마이닝). 이러한 소셜빅데이터를 효율적으로 수집하고 아동청소년의 학교폭력 현상을 분석하는 데에 적절히 활용하기 위해서는 무엇보다도 학교폭력에 대한 체계적인 온톨로지 구축이 필수적이다[13][14].

이에 본 연구는 온톨로지의 개념을 이해하고, 소셜빅데이터를 통해 학교폭력 현상을 이해하는 데에 있어서 온톨로지의 효용성을 제시하고, 소셜빅데이터 수집에 활용될 학교폭력 온톨로지 개발과정 및 결과를 상세히 설명한다.

## II. 연구방법

### 1. 온톨로지의 개념

온톨로지(ontology)는 철학적 의미를 지닌 개념으로 '존재를 설명하는 특징들과 그 구조(nature and structure of reality)'라고 알려져 있다[15]. 최근 컴퓨터 과학 분야에서 온톨로지의 활용이 확산되고 있는 가운데, 온톨로지의 개념이 '서로 공유하는 개념적 이해에 대한 형식적이고 명시적인 구체화 과정의 결과물(formal, explicit specialization of a shared conceptualization)'로 재정의되어 사용되고 있다[16]. 즉, 온톨로지는 인간이 관심 갖는 주제에 대한 개념들을 정의하고, 그 개념들 간의 관계와 계층구조를 명시해줌으로써 컴퓨터가 스스로 정보처리를 할 수 있도록 돕는 역할을 한다[13][16]. 따라서 온톨로지를 인간의 생각과 정보기술의 기능을 연계시켜주는 용어체계를 명시하는 호환적 도구라고도 정의할 수 있다.

빅데이터 분석에 활용되는 온톨로지는 자료, 정보, 지식 세 가지 토대요소로 구성되어 있다[17]. 자료(data)는 우리에게 이미 주어진 것으로 정보 생성을 위한 기초를 제공하고, 정보(information)는 자료의 분석 및 탐색을 통해 생성되는 결과물을 가리키며, 지식(knowledge)은 정보를 분석함으로써 얻게 되는 현상에 대한 이해를 가리킨다[17]. 즉, 소셜빅데이터 상에 표현되는 자료(data)는 개별적인 정보(information)를 제공하는데, 이 때 온톨로지를 거쳐야만 정보 간의 관

계망이 형성되며 용어체계를 통해 의미를 부여받아 사회 현상을 해석하는 데 활용될 수 있는 지식(knowledge)이 생성된다고 할 수 있다. 이 때문에 학교폭력 문제와 같은 복잡한 사회현상을 분석하기 위한 소셜빅데이터 자료를 수집할 경우, 온톨로지의 개발과정은 현상 이해에 필요한 개념들을 정리할 뿐 아니라, 현상을 통찰력 있게 이해하기 위한 구조적인 틀을 제공한다고 할 수 있다. 이처럼 온톨로지는 특정 주제에 대한 데이터를 수집하고 분석하는 데 있어서 실용적인 틀로서의 역할을 한다[18].

특히 소셜빅데이터 수집을 하는 데 있어서 온톨로지의 구축은 필수적이다. 온라인상의 정보로 표현되는 소셜빅데이터는 양적으로 방대한 비정형 텍스트 데이터로서 자료의 체계적인 수집과 활용이 어렵기 때문이다. 따라서 수많은 정보 가운데 의미 있는 핵심용어를 추출하고 자료를 효과적으로 수집하기 위해서는 관심 분야의 핵심개념을 명시할 뿐 아니라, 해당 개념들 간의 관계를 나타내주는 온톨로지의 역할이 중요하다. 온톨로지는 체계적인 틀을 제공함으로써 웹크롤러(webcrawler)가 방대한 양의 비정형 데이터를 수집하고 처리할 수 있도록 구조적 시스템을 마련해 준다.

온톨로지 구축의 목적은 비구조화된 데이터로부터 정보의 체계적인 수집과 분석을 가능케 하기 위함이며, 온톨로지 구축 과정은 특정 주제와 관련하여 그동안 입증되어 온 이론들을 바탕으로 위계적인 체계를 정리하는 것으로 볼 수 있다. 이러한 온톨로지 구축단계 없이 소셜빅데이터 연구가 수행되는 경우, 주제와 관련된 일부 요인들을 간과할 수 있으며 비정형화된 단어들을 분류하기에 상대적으로 더 긴 시간이 소요될 가능성이 생긴다. 따라서, 아동청소년 학교폭력 문제와 같이 복잡한 사회현상을 분석하는 연구에서 효율적으로 소셜빅데이터를 활용하기 위해서는 온톨로지의 개발이 선행되어야 한다.

## 2. 온톨로지 개발

본 연구에서 소셜빅데이터 자료 수집 및 분석을 목적으로 개발하고자 하는 학교폭력 온톨로지의 방법론적 틀은 Noy와 McGuinness의 연구[19]<sup>1</sup>에 기초하며 이

를 번안, 수정 및 보완한 국내 선행연구[20-22]를 참고하였다. 구체적인 온톨로지 개발 과정에 대한 각 단계별 상세한 설명은 다음과 같다(그림 1).

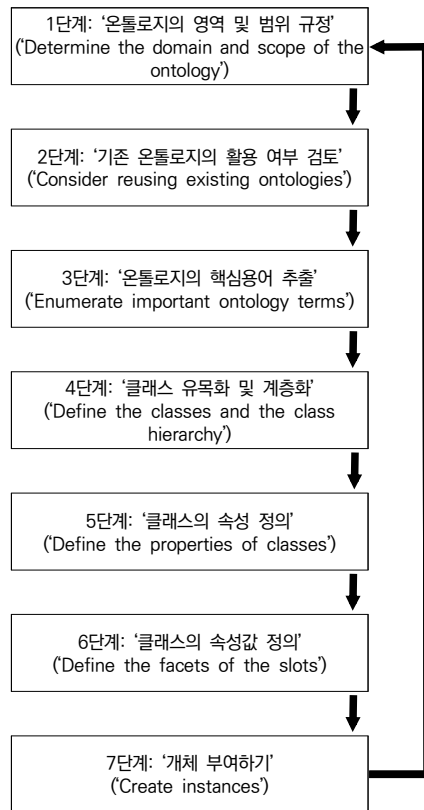


그림 1. 온톨로지 개발 과정[19-22]

### 2.1 1단계: '온톨로지의 영역 및 범위 규정'

첫 번째는 온톨로지의 영역 및 범위를 규정하는 단계이다. 이 단계에서 연구자는 여러 역량질문(competency question)을 통해 온톨로지의 영역 및 범위를 규정할 수 있다[19]. 역량질문은 연구자가 관심 갖는 주제나 핵심용어들의 개념적 정의, 원인(cause), 영향(effect) 요인 등에 대한 내용을 함축할 뿐 아니라,

1 Noy와 McGuinness의 연구[19]는 2019년 현재 Google

Scholar에 5,500 여회 인용되는 등 다양한 학문분야의 연구자들로부터 꾸준히 온톨로지 개발을 위한 핵심자료로 활용되어 왔다. 또한 Noy와 McGuinness[19]의 연구는 핵심내용만 간결하게 기술된 기술보고서(Technical Report) 형식의 논문으로 구성됨으로써 비전문가도 온톨로지 개발을 손쉽게 개발할 수 있도록 설명해준다는 장점을 갖고 있다.

연구결과가 누구를 위해, 어디에, 어떻게 활용될 것인지를 확인해주는 역할도 함으로써 수집되는 정보의 양과 내용을 일정 범위 내로 국한 시켜주는 역할을 한다[19]. 개발된 온톨로지를 기반으로 수집된 자료를 분석하여 역량질문에 대한 응답을 역으로 도출할 수 있는지 검증하는 등 평가과정을 거침으로써 온톨로지의 범위를 확장 또는 축소시킬 수 있다[19].

표 1. 학교폭력 영역과 세부 역량질문

영역	역량질문
대상	누가 학교폭력에 참여하는가? 가해자와 피해자 이외 다른 참여자는 있는가? 가해자와 피해자가 두명이 나누는가? 시간에 따라 학교폭력 역할자 분포가 달라지는가?
원인	학교폭력 가해와 피해 원인은 무엇인가? 신체적/심리적 요인 때문인가? 경제적 요인 때문인가? 사회관계적 요인 때문인가? 문화적 요인 때문인가? 제일 중요한 학교폭력 기제는 무엇인가? 시간에 따라 학교폭력 기제가 달라지는가?
유형	학교폭력의 방법은 무엇인가? 직접적 방법 또는 간접적 방법으로 일어나는가? 신체적, 언어적, 관계적 방법으로 나타나는가? 온라인/오프라인 영역에 따른 차이가 나타나는가? 학교폭력 방법들이 독립적인가 또는 복합적인가? 더 빈번히 발생하는 학교폭력 방법들이 무엇인가? 시간에 따라 학교폭력 방법들이 달라지는가?
장소 · 지역	학교폭력은 어디서 일어나는가? 학교교실이나 학교인근이 아닌 다른 장소에서도 일어나는가? 학교폭력은 어느 행정지역에서 일어나는가? 시간에 따라 학교폭력 발생 장소 및 지역이 상이한가?
대응 방안	학교폭력에 대한 정책적·제도적·사회적 대응방안은 무엇인가? 학교폭력 관련 서비스·프로그램은 무엇이며, 이를 운영하는 공공·민간 기관은 어디인가?

아동청소년의 학교폭력 경험을 확인하는 본 연구의 경우, 소셜빅데이터에 산재하여 있는 자료에서 학교폭력의 개념적 정의나 형태에 대한 이해를 높이고, 학교폭력의 원인이나 이로 인한 사회적 결과를 함축할 수 있는 정보를 추출할 수 있는 내용으로 역량질문을 구성하였다. 구체적으로 역량질문은 학교폭력의 대상자, 학교폭력의 형태 및 유형, 학교폭력의 원인 및 결과, 학교폭력 참여자들의 개인·가정·사회적 특징, 정책적·사회적 반응 및 대응방안과 관련된 내용을 포함한다. 구체적인 질문은 [표 1]에 제시하고 있다.

### 2.2 2단계: ‘기존 온톨로지의 활용 여부 검토’

두 번째는 기존 온톨로지의 활용 여부를 검토하는 단

계이다. 선행연구에서 개발한 온톨로지를 수정·보완하여 활용한다[19]. 연구주제 면에서 일치하는 온톨로지가 없는 경우 유사 분야의 온톨로지가 있는지 확인하고 이를 연구의 목적과 범위에 적합하게 수정하고 확장하여 활용한다[19].

아동청소년 학교폭력 소셜빅데이터 자료 수집을 위한 온톨로지 개발은 활발히 진행되지 않았으며, 기존 연구는 주로 우울[12], 비만[23], 상담[24], 치료[25], 예방접종[21] 등을 주제로 한 보건·의료 분야에 집중되어 있다. 예외적으로 최근 위기청소년 분류체계를 제시한 선행연구[26]가 있어 이 연구를 기반으로 하여 아동청소년 학교폭력 경험의 특수성을 반영할 수 있는 온톨로지 체계를 구축하였다.

### 2.3 3단계: ‘온톨로지의 핵심용어 추출’

세 번째는 온톨로지의 핵심용어를 추출하는 단계로, 이 단계에서는 연구자가 검증하고자 하는 핵심 주제와 연관된 모든 단어(명사, 동사 등)를 추출한다[19]. 이 단계에서는 단어들의 상하위 구조를 설정하지 않고 다양한 출처에서 최대한 풍부한 양의 용어들을 클래스와 속성 구별 없이 모두 확보한다[19].

본 연구에서는 대상, 원인, 유형, 장소, 지역, 대응방안 등의 영역에 대해 관련 분야 대학교수 3명과 대학원생 4명이 다양한 자료원을 검토하여 아동청소년 학교폭력 현상과 관련된 핵심용어를 추출하였다. 학교폭력 현상과 관련된 이론에 대한 전문서적이거나 논문과 같은 학술자료 뿐만 아니라, 학교폭력을 언급한 블로그, 게시판, 뉴스 등 소셜미디어에서 학교폭력 관련 용어들을 직접 검색하여 핵심용어를 추출하였다. 또한 관련 법령 및 제도와 각 부처에서 발행하는 보고서 등 행정 자료도 온톨로지의 핵심용어 추출에 활용하였다. 자료 수집을 위한 구체적인 근거와 방법은 다음과 같다.

#### 1) 학술자료 활용: 학교폭력 관련 이론

아동청소년의 학교폭력 현상을 설명하는 주요 이론을 선정한 후, 관련 저서나 학술논문 데이터베이스에서 이러한 이론을 활용한 자료를 검색하였다. 검색한 아동청소년 학교폭력 관련 학술자료에서 반복적으로 등장하는 단어들을 온톨로지의 핵심용어로 선정하였다. 학

술자료로부터 학교폭력의 원인 및 유형 관련 용어가 주로 추출되었으며, 온톨로지에 활용된 핵심용어의 구체적인 추출 근거 및 과정은 다음과 같다.

#### ① 학교폭력 원인

학교폭력의 원인을 설명하는 이론은 다양하다. 학교폭력 관련자들의 신체적·심리적 특징을 학교폭력 경험의 원인으로 설명하는 이론이 있는가 하면, 지역사회나 문화를 학교폭력의 주요 원인으로 꼽는 이론들도 있다.

우선, 생물학적·심리학적 요인을 학교폭력의 원인으로 꼽는 연구자(trait theorist)들은 피해자 및 가해자들의 신체적인 조건이나 심리적인 특성들을 강조한다. 신체적 미성숙, 장애, 비만 등의 특징을 지닌 아동청소년들은 학교폭력의 피해자 및 가해자가 될 가능성이 더 높다[27-29]. 또한, 학교폭력 피해자와 가해자들을 설명하는 심리적 요인으로 낮은 지능, 사회인지능력 부족, 민감성 부족, 공감 능력 부족, 자기통제력 부족, 충동성, 품행장애, 분노조절 장애 등을 들 수 있다[30-32].

긴장이론(strain theory)[33][34]은 일상에서의 긴장 요인들이 아동청소년들의 부정적 감정을 발달시켜 문제행동을 유발한다고 주장한다. 여기서 긴장요인이라 함은 목표를 이루지 못할 경우(예: 경제적 어려움, 사회적 지위 저하 등), 부정적 자극이 있거나 긍정적 자극이 결여된 경우(예: 학대적 부모, 학교 부적응, 부정적 교사관계, 또래 괴롭힘 등)를 가리킨다[35]. 일반 아동청소년도 위와 같은 긴장을 경험하게 될 때 분노, 우울, 불안 등의 부정적 감정을 느끼게 되는데, 주위 지원을 동원하는 등 적절한 대처방법을 모르거나 탄력성이 부족한 아동청소년들은 이러한 부정적 감정이 학교폭력 가해 행위와 같은 문제행동을 야기하는 요인이 될 수 있다.

사회통제이론(social control theory)[36][37]에 따르면 아동청소년의 문제행동 동기는 자연스러운 것이지만, 이러한 동기를 억누르고 통제할 수 있게 될 때 학교폭력 등 문제행동에 가담하지 않는다고 설명한다. 학교폭력을 감소시키는 역할을 하는 것으로 직접통제, 자기통제, 유대요인을 꼽을 수 있다. 직접통제는 체벌, 감독, 규제 등 직접 문제행동을 통제하는 역할을 하는 외적 기제요인인 반면, 자기통제는 스스로 문제행동을 통

제하는 내적 기제요인이다[35]. 자기통제는 선천적인 심리성향이 강하지만 직접통제이나 유대요인을 통해 후천적으로도 발달시킬 수 있다[37]. 유대요인으로 부모나 교사에 대한 애착(attachment), 사회적 성취를 위한 관여(commitment), 학교 활동 참여(involverment), 사회적 규범에 대한 신념(belief) 등을 들 수 있다[36].

생태체계이론(ecological systems theory)[38]은 미시체계(microsystem), 중간체계(mesosystem), 외체계(exosystem), 거시체계(macrosystem), 시간체계(chronosystem)가 학교폭력 행동에 영향을 줄 수 있다고 설명한다. 즉, 하나의 요인이 아닌, 서로 밀접하게 연결되어 상호영향을 미치는 여러 체계가 학교폭력의 기제로 꼽힌다. 부모·교사·또래와의 직접적 관계가 미치는 영향(미시체계)과 미시체계 간의 관계를 중시(중간체계)할 뿐 아니라, 미디어 및 지역사회 분위기(외체계)나 사회문화(거시체계)의 영향을 학교폭력의 원인으로 설명하고 있으며 전생애에 걸친 중단적 관점의 중요성(시간체계)을 강조하고 있다[39].

마지막으로, 문제행동이론(problem-behavior theory)[40]은 학교폭력의 원인을 성격체계(personality system)와 인지된 환경체계(perceived environmental system) 간의 상호교류가 행동체계(behavioral system)에 영향을 미친 결과라고 설명한다. 또한, 아동청소년이 한두 가지의 부적응 행동을 보일 경우, 다른 부적응 행동을 보일 가능성이 현저히 증가한다는 특성을 보인다[40]. 따라서 문제행동이론은 학교폭력 문제도 아동청소년의 여러 문제행동 중 하나로써, 음주, 흡연, 약물사용, 중독, 비행, 무분별한 성관계 등 아동청소년의 다양한 부적응 행동과 함께 고려되어야 한다는 입장을 취한다.

#### ② 학교폭력 유형

학교폭력의 유형들은 가해행위 방식에 따라 직접적 폭력과 간접적 폭력으로, 또는 가해 공간에 따라 오프라인과 온라인으로 나눌 수 있다. 먼저, 직접적 폭력과 간접적 폭력은 폭력의 종류나 결과에 따라 구분될 수 있다. 직접적 폭력은 폭행, 금품갈취, 감금 등과 같이 직접 신체적인 접촉이나 물리적으로 피해자를 괴롭히는

행위를 뜻하며, 간접적 폭력은 소문내기, 소외시키기, 비웃기 등과 같이 피해자가 정신적으로 고통을 느끼도록 하는 행위들을 뜻한다[41].

다음으로, 오프라인과 온라인은 폭력이 행해지는 공간에 따라 학교폭력 유형들을 구분한다. 오프라인 폭력에는 언어폭력, 신체폭력, 금품갈취 등 현실 공간에서 피해자와 가해자가 마주하고 있는 상황에서 행해질 수 있는 폭력 행위들을 의미하고, 온라인 폭력은 온라인채팅, SNS, 문자 등을 활용하여 피해자를 괴롭히거나, 또는 몰카, 개인정보 도용과 같이 디지털기기를 활용하여 피해자의 권리를 침해하는 행위들을 포함한다[42]. 특히, 현대사회에서는 온라인과 오프라인을 넘나들며 직·간접적 학교폭력 방법들이 복합적으로 행해지며 새로운 유형들도 빠르게 생성되고 있어 이에 대한 보다 정확한 파악 및 대처가 요구된다.

2) 정책·제도 자료 활용

학교폭력 문제에 대한 정책적·제도적 대응방안에서도 학교폭력 온톨로지 개발에 유용한 관련 핵심용어를 추출할 수 있다. 학교폭력과 관련하여 제정된 법이나 제도와 다양한 서비스나 프로그램을 운영·담당하는 행정기관에 대한 정보는 소셜빅데이터에 등장하는 학교폭력 문제를 대표하는 단어들을 제공할 수 있다. 구체적으로 청소년보호법, 청소년기본법, 청소년복지지원법, 아동청소년의 성보호에 관한 법률, 아동복지법 등 학교폭력과 관련된 다양한 법·제도적 대응방안을 확인할 필요가 있다. 또한 경찰, 검찰, 법무부, 여성가족부 등 학교폭력 문제를 다루는 행정기관과 이곳에서 제공하고 있는 서비스나 프로그램(청소년상담지원센터, 청소년비행예방센터, 1388헬프콜, 학교폭력신고전화 등)에 대한 탐색도 학교폭력 관련 핵심용어를 파악하는 데 도움을 줄 수 있다.

3) 장소 및 지역

학술자료나 정책·제도 자료 이외에도 아동청소년의 학교폭력 온톨로지 개발을 위한 핵심용어를 추출하기 위해 학교폭력이 이루어지거나 연관되는 장소 및 지역을 나타내는 자료를 참고하였다. 구체적으로 학교폭력이 이루어지는 공간적인 특성을 이해하기 위해 직접적

이거나(PC방, 교실, 화장실 등) 간접적인(카카오톡, 인스타그램, 페이스북 등) 가해가 이루어질 수 있는 학교폭력 장소와 학교폭력이 언급되는 시·군 단위의 행정지역에 대한 정보를 추출하기 위해 이와 관계된 핵심용어를 수집하였다. 위의 정보는 학교폭력의 위험성을 증가시키는 유해환경에 대한 이해를 높이고, 학교폭력이 핵심용어로서 이슈화되는 행정 지역에 대한 정보를 제공할 것이다. 그리하여 자료 분석 결과를 통해 위험요인을 제거하고, 적절한 대응방안을 제시할 수 있을 것이다.

4) 소셜미디어 검색

온톨로지의 핵심용어를 추출하기 위해 SNS(트위터), Blog, Café, Board, News Channel 등의 국내 오픈 사이트를 중심으로 학교폭력이라는 단어와 함께 문서에서 등장하는 단어들을 검색하였다. 국내 오픈사이트에서만 소셜빅데이터 수집이 가능하기 때문에 핵심용어 추출을 위한 검색도 상기 사이트를 중심으로 진행하였다. 각 채널의 검색브라우저에 검색 기간을 본 연구의 자료수집 기간인 2013년부터 2017년도까지로 설정하고, 검색단어를 “학교폭력 AND 청소년” 또는 “학교폭력 AND 중학생” 등으로 설정하여 검색을 시행하였다. 뉴스나 교육부 홈페이지 등과 같은 채널을 통해서도 학교폭력에 대한 표준어(폭력, 폭행, 금품갈취 등)를 파악하였다. 뿐만 아니라, 학술자료나 정책·행정 자료에 포함되지 않지만, 아동청소년들의 학교폭력 현상을 정확하게 이해하기 위해 필수적인 은어나 줄임말(‘고구마’(답답하다), ‘친삭’(친구삭제) 등), ‘카따’(카카오톡 왕따))을 네이버 카페나 트위터와 같은 공간에서는 찾아낼 수 있었다. 특히, 팬덤, 개인방송(인스타 라이브, 아프리카 방송 등)과 같이 학술자료에서 자주 다루어지지 않았으나 아동청소년들의 학교폭력 현상에 영향을 끼칠 수 있는 다양한 요인들에 대해서도 파악할 수 있었다.

2.4 4단계: ‘클래스 유목화 및 계층화’

4단계와 5단계는 온톨로지 개발 과정에 있어서 가장 핵심이 된다[19]. 4단계는 3단계에서 추출한 핵심용어를 클래스(또는 개념; class)에 따라 유목화하고 계층

화하는 과정으로서, 다음 세 가지 방법으로 가능하다 [19]. 첫째는 하향식(top-down) 방법으로, 온톨로지의 클래스를 형성하고 계층화하는 데 있어서 포괄적 개념에서 시작하여 세부적인 것으로 발전시킨다. 둘째는 상향식(bottom-up) 방법으로, 상세한 소분류 클래스에서 시작하여 일반적 대분류 클래스를 구축하는 것이다. 마지막으로 복합식(combination) 개발과정은 상향식과 하향식 클래스 개발 체계를 동시에 활용한다.

계층화의 특징 중 하나로 하위클래스의 용어는 상위클래스에 내포되며("is-a" relationship), 동일한 상위클래스에 포함된 하위클래스의 용어들은 유사한 속성을 공유한다[19]. 또한 하위클래스와 상위클래스 간에 추이적 성질(transitivity)이 성립되어 B가 A라는 상위클래스에 포함되고, C가 B라는 상위클래스에 포함되면, C도 A에 포함된다는 특징을 보인다[19].

본 연구에서는 학교폭력 관련 핵심용어들을 유목화하고 계층화하여 클래스를 구성하는 데 있어 하향식과 상향식을 함께 활용하는 복합식을 활용하였다. 학교폭력의 대상자, 영역 및 유형, 기제, 대응방안, 지역 등의 상위클래스를 설정하여 전체적인 용어체계를 구성함으로써 하위클래스를 이루는 체계를 도출하는 하향식 방법을 활용하는 한편, 동시에 신체적, 피해자심리, 가해자심리 등 하위개념을 설정한 후, 이 개념들이 공통으로 구성하는 학교폭력 원인이라는 상위개념을 도출하는 상향식 방법을 활용하였다.

### 2.5 5단계: '클래스의 속성 정의하기'

이 단계에서는 클래스의 속성을 정의하며, 클래스를 설정함으로써 핵심용어들의 유목화 및 계층화 과정을 거치면서 클래스 분류 체계를 구축한 후, 클래스들의 특징을 반영하는 속성(slots)을 파악한다[19]. 클래스의 고유한 특성을 나타내는 내적 속성(intrinsic property), 외부로부터 부여되는 외적 속성(extrinsic property), 구성(parts)을 나타내는 속성, 관계(relationship)를 나타내는 속성 등이 있다[19].

아동청소년 학교폭력 온톨로지의 경우, 클래스의 속성은 대부분 관계적이다. 대상자 영역의 용어들은 학교폭력에 직·간접적으로 연루된 대상들의 특징에 대한 용어(연령, 가족·도래·집단 등의 형태)로, 원인 영역은 다

양한 요인(신체·긴장·생태체계 요인 등)들과 학교폭력 간의 인과적인 관계를 설명하는 용어들로 관계적 속성을 지닌다고 할 수 있다. 반면 유형 영역에서의 학교폭력 가해 방법은 학교폭력 행위 자체의 고유한 특징에 대한 정보(직접·간접 폭력, 오프라인·온라인 폭력 등)를 제공하므로 내적 속성이라고 할 수 있으며, 학교폭력 발생 지역이나 가해 장소는 학교폭력의 본질 및 고유성과는 관계가 없으므로 외적 속성이라 할 수 있다.

### 2.6 6단계: '클래스의 속성값 정의하기'

여섯번째 단계에서는 클래스의 속성값을 정의한다. 사상수(cardinality)는 클래스 속성의 수를 칭하는데, 단 하나의 속성값을 갖는 경우는 단일형(single cardinality), 여러 속성값을 갖는 경우 다중형(multiple cardinality)이라 부른다[19]. 또한 클래스 내 핵심개념들의 속성에 따라서 상이한 속성값을 갖게 된다. 속성값의 유형으로 명명형(string), 연속적 숫자형(number), 예 또는 아니요를 묻는 이분형(boolean), 여러 옵션 중 하나를 선택하는 열거형(enumerated), 특정 클래스의 개체를 사용하는 개체형(instance) 등이 있다[19][22].

아동청소년 학교폭력 온톨로지의 경우 속성별 하나의 값만을 제시했으며(단일형 사상수), 클래스 별 다양한 속성값들이 활용되었다. 학교폭력 발생 지역을 나타내는 명명형(수도권: 서울, 강북, 강남, 인천 등), 관련 문서가 최초작성 문서인지 여부를 묻는 이분형('최초문서': 예, 아니요), 관련문서 작성 요일 정보가 담긴 열거형('요일': 월요일, 화요일 등)이 클래스의 속성값으로 활용되었다.

### 2.7 7단계: '개체 부여하기'

마지막 일곱번째 단계는 개체를 부여하는 단계이다. 이 단계에서는 클래스를 유형과 계층으로 분류하고, 클래스 별 속성과 속성값을 지정한 후, 마지막으로 클래스 별 속성에 대한 속성값에 개체(instances)를 부여하여 온톨로지를 완성한다[19]. 소셜빅데이터 자료수집을 위한 아동청소년 학교폭력 온톨로지 개발의 구체적인 결과는 아래에 제시하였다.

### III. 연구결과

학교폭력 온톨로지 개발 결과, 학교폭력의 대상, 원인, 유형, 장소, 지역, 대응방안의 6가지 영역으로 최상위 계층인 대분류가 구성되었다. 하위분류체계로서 대상, 원인, 유형, 장소는 대분류, 중분류, 소분류로 구성되었고, 지역과 대응방안은 대분류 및 중분류로 구성되었다. 각 대분류별 하위분류체계는 [표 2]와 같다.

표 2. 대분류 체계

대분류	정의	깊이 (개)	핵심용어 수 (건)	핵심용어(예시)
대상	학교폭력 관련 대상을 지칭하는 연령, 학년 또는 가족·사회·학교에서의 지위	3	385	8살, 초1, 여중생, 남고생, 선배 동생, 친구, 사촌 등
원인	학교폭력 피해 혹은 가해 경험에 영향을 끼치거나 원인이 될 수 있는 요인	3	1,035	외모, 사회성부족, 공격성, 가정불화, 비만, 호기심 등
유형	상대방을 괴롭히기 위해 온·오프라인에서 직·간접적인 형태로 행해지는 폭력행위	3	207	욕설, 구타, 따돌림, 악플, 해킹, 루머, 성희롱, 성폭행 등
장소	학교폭력의 발생 장소	3	141	교실, 노래방, PC방, 온라인 채팅방, SNS 등
지역	학교폭력의 발생 지역	2	42	서울, 경기, 도시, 시골 등
대응 방안	학교폭력의 사전 예방 및 사후 대처를 위한 대응방안	2	73	청소년 보호법, 청소년상담지원센터, 1388헬프콜 운영 등

#### 1. 학교폭력 대상

학교폭력 온톨로지의 첫번째 대분류인 대상자는 가해자, 피해자, 방관자, 일반인 등 학교폭력 관련 대상자 모두를 포함하는 개념으로서 연령과 형태의 두 가지 중분류로 구성되었다. 먼저, 연령 관련 정보를 통해 해당 문서가 아동청소년과 관련된 문서인지 아닌지를 파악할 수 있을 뿐만 아니라, 현재 온라인상에서 특정 연령대의 학교폭력 행동이 관심을 받고 있는지도 파악할 수 있다. 본 연구의 대상이 아동청소년이라는 점에 기반하여 연령을 다시 아동청소년 및 성인의 소분류로 구분하였고, 각 소분류에 해당하는 다양한 연령 표현 단어들(8살, 8세, 초1 등)을 정리하였다. 다음으로, 형태는 학교폭력에 직·간접적으로 관련된 대상(모, 친구, 선배 등)

이 무엇인지에 대한 정보를 제공한다. 형태는 아동청소년기에 중요한 관계에 해당하는 가족, 또래, 집단뿐만 아니라, 보편적으로 학생이나 기타 사람을 표현할 때 사용하는 단어들을 구분하여 5가지의 소분류체계로 이루어졌다[표 3].

표 3. 대상자 분류체계

대분류	중분류	소분류	핵심용어(예시)
대상자	연령	아동 청소년	8살, 8세, 초1, 초당1, 1학년, 여중생, 남고생 등
		성인	20대, 선생님, 아버지, 쌤, 과외선생님 등
	형태	가족	가족, 부모, 할머니, 형, 동생, 쌍둥이, 사촌 등
		또래	친구, 동네친구, 학원친구, 베프, 여사친 등
		집단	친구들, 선배들, 후배들, 동네애들, 노는애들 등
		학생	초등학생, 초당, 중고생, 여중생, 고베리 등
기타	남자, 어린이, 어린이, 아이, 여성, 남녀 등		

#### 2. 학교폭력 원인

아동청소년이 학교폭력의 가해자 또는 피해자가 되는 원인을 설명하는 단어들을 하나의 대분류로 구성하였으며, 이를 원인으로 명명하였다. 또한, 원인의 대분류 내에서는 특성요인, 긴장요인, 사회통제요인, 생태체계요인, 비행요인의 5가지 중분류로 분류하였다.

첫 번째, 특성요인은 신체적 요인, 피해자심리요인, 가해자심리요인으로 구성되었다. 학교폭력의 가해 또는 피해 경험을 유발시킬 수 있는 개인의 신체적 또는 심리적 특징을 나타내는 단어들 이에 포함되었다. 두 번째, 긴장요인은 가족긴장요인, 친구긴장요인, 학교긴장요인으로 구성되었다. 아동청소년기에 가족관계·친구관계·학교생활로부터 경험할 수 있는 다양한 긴장요인을 표현하는 단어들 이 해당 분류에 포함되었다. 세 번째, 사회통제요인은 직접통제요인, 낮은자기통제요인, 높은자기통제요인, 애착요인, 관여요인, 참여요인, 신념요인으로 구성되었다. 아동청소년의 학교폭력 행동을 통제할 수 있는 내적 또는 외적 통제요인들을 의미하는 단어들 이 포함되었다. 네 번째, 생태체계요인은 사회문화요인과 사회매체요인으로 구성되었으며 학교폭력 현상에 영향을 끼칠 수 있는 현대사회의 문화적 특징과 다양한 소셜 매체들이 해당 분류에 속하였다[표 4].



표 4. 원인 분류체계

대분류	중분류	소분류	핵심용어(예시)
원인	특성 요인	신체적	신체적미성숙, 장애, 비만, 동보, 열치, 못생김 등
		피해자 심리	사회성부족, 과잉보호, 공주병, 고구마 등
		가해자 심리	우월성과사, 욕심, 공격성, 분노, 극형, 복수 등
	간장 요인	가족 긴장	불화가정, 가정폭력, 이혼, 불륜, 별거, 가난 등
		친구 긴장	친구폭력, 친구 비밀유출, 비행친구, 절교 등
		학교 긴장	교사우관심, 꾸지람, 차별, 중간고사, 성적 등
	사회 통제 요인	직접 통제	교칙, 규범, 체벌, 감독, 학생부, 통금 등
		낮은 자기 통제	스릴, 재미, 호기심, 유혹, 순간 만족, 짜릿 등
		높은 자기 통제	통제, 절제, 인내, 참을성, 그만두다, 멈추다 등
		애착	애착, 애정, 대화, 소속감, 신뢰, 롤모델 등
		관여	전념, 열정, 몰입, 동기, 노력, 성공 등
		참여	참여, 수업, 학원, 동아리, 학교 밖 활동 등
	생태 체계 요인	사회 문화	핵가족화, 집단주의, 숭방망이처벌, 입시경쟁 등
		사회 매체	온라인 게임, 1인방송, 성인물, 일베, 불법 등
	비행 요인	-	흡연, 유흥업소출입, 성매매, 가출, 도둑질 등

3. 학교폭력 유형

학교폭력 유형 대분류는 온라인과 오프라인에서 직접적이거나 간접적인 형태로 행해지는 오늘날 아동청소년들의 학교폭력 방법들을 모두 포함한다. 유형 영역의 중분류로 먼저 직접폭력과 간접폭력으로 구분하였고, 이를 다시 각각 오프라인과 온라인 영역으로 구분하였다.

중분류에 따라 살펴보면, 첫 번째 직접폭력 유형에는 욕설, 성희롱과 같이 오프라인과 온라인에서 공통적으로 행해질 수 있는 방법들이 포함되며, 두 공간의 특성에 따라 차별적인 방법들이 각 소분류에 맞게 배치되었다. 예를 들면, 오프라인에는 구타, 집단폭행, 강간이 포함되었고, 온라인에는 악플과 동영상이 추가로 포함되었다. 두 번째, 간접폭력 유형에서도 이와 동일한 방식으로 단어들이 배치되었다. 소외, 따돌림, 루머와 같이 오프라인과 온라인에서 모두 행해질 수 있는 방법들이

두 소분류에 모두 포함되었다. 오프라인에는 빵셔틀 등이, 온라인에는 해킹, 몰카 등이 추가적으로 포함되었다 [표 5].

표 5. 유형 분류체계

대분류	중분류	소분류	핵심용어(예시)
유형	직접 폭력	오프라인	욕설, 성희롱, 구타, 집단폭행, 강간 등
		온라인	욕설, 성희롱, 악플, 동영상, 성적모욕 등
	간접 폭력	오프라인	소외, 따돌림, 루머, 장난, 뒷담화, 빵셔틀 등
		온라인	소외, 따돌림, 루머, 신상털기, 해킹, 몰카 등

4. 학교폭력 장소

학교폭력의 장소는 학교 내외, 지역사회, 사이버 공간 등과 같이 학교폭력의 발생 가능성을 지닌 모든 공간을 포함하고자 하였으며, 장소 영역의 중분류로 크게 오프라인과 온라인으로 분류하였다.

오프라인의 경우, 학교(내), 학교 밖 생활공간, 유흥시설로 분류하였다. 먼저, 학교 내 공간에는 교실, 복도와 같이 학교폭력 연구에서 꾸준히 중요하게 다루어져 왔던 장소들을, 학교 밖 생활공간은 놀이터, 공원과 같이 아동청소년 간의 학교폭력 행위가 이루어질 수 있는 장소들을 포괄한다. 마지막으로, 유흥시설은 PC방, 오락실과 같이 아동청소년들이 폭력적 매체를 접하거나 비행행동에 대해 접근 가능성이 큰 장소들을 포함한다. 온라인의 경우, SNS, 메신저, 기타로 소분류가 구성되었다. SNS에는 아동청소년들이 사회적 관계를 형성할 뿐만 아니라 공개적으로 상대방을 조롱하거나 비난할 수 있는 장소인 트위터, 인스타그램 등을, 메신저에는 카카오톡, 채팅방과 같이 일대일 대화 속 폭력뿐만 아니라 그룹채팅방에서의 따돌림이나 괴롭힘이 행해질 수 있는 장소들을 포함하였다. 온라인, 모바일과 같은 용어들도 학교폭력을 묘사하는 데에 사용될 수 있으므로 기타 소분류로 구분하였다[표 6].

표 6. 장소 분류체계

대분류	중분류	소분류	핵심용어(예시)
장소	오프라인	학교	교실, 복도, 매점, 체육관 등
		학교 밖 생활공간	학원, 독서실, 놀이터, 공원 등
		유흥시설	PC방, 오락실, 노래방 등
	온라인	SNS	트위터, 인스타그램, 페이스북 등
		메신저	카카오톡, 채팅, 메시지 등
		기타	온라인, 모바일, 컴퓨터 등

5. 학교폭력 지역

학교폭력의 지역은 크게 세 가지 중분류(수도권, 수도권 외, 지역특성)로 구성되었다. 지역의 하위분류체계는 소분류까지 세분화하지 않더라도 세 가지 중분류가 대다수의 핵심용어들을 포함하고 있었기에, 대분류 및 중분류의 하위분류체계로 구성되었다. 지역구분은 특정 지역에서 발생한 학교폭력 사건이 소셜빅데이터에서 어떻게 표출되는가를 나타내기 때문에 학교폭력의 지역적 집중도를 파악하는 데 도움을 줄 수 있다.

수도권과 수도권 외 중분류의 경우, 서울, 인천, 경북 등과 같이 수도권 내 시 또는 구 단위 지역 명칭들과 강원, 전주, 대구 등 수도권 외 지역의 명칭들을 모두 포함하였다. 또한, 지역명이 아닌 도시, 시골, 섬 등과 같이 지역의 특성을 하나의 중분류로 구성함으로써, 학교폭력에 영향을 끼치는 지역 특성에 대한 파악이 가능하도록 하였다[표 7].

표 7. 지역 분류체계

대분류	중분류	소분류	핵심용어(예시)
지역	수도권	-	서울, 경북, 경남, 인천 등
	수도권 외	-	강원, 대전, 전주, 대구, 제주 등
	지역특성	-	도시, 시골, 대도시 등

6. 학교폭력 대응방안

학교폭력 대응방안 영역은 학교폭력 문제의 사전예방 및 사후조치를 위하여 국가, 지자체, 또는 교육기관 등에서 행해지고 있는 여러 대응방안 관련 단어들을 모두 포함한다. 대응방안의 중분류는 세 가지 영역(법제도제정, 담당부처, 프로그램 운영)으로 구성하였다. 또한, 대응방안은 지역과 마찬가지로 소분류가 불필요하다고 여겨져, 대분류 및 중분류로 하위분류체계를 구성하였다.

먼저, 법·제도(청소년보호법, 학교전담경찰관 제도 등)에는 국가적 차원에서 학교폭력의 피해자를 보호하고 가해자를 처벌하기 위한 방침들을 포함하였다. 다음으로, 담당부처(청소년상담지원센터, 청소년비행예방센터 등)에는 학교나 지역사회에서 아동청소년들의 학교폭력 실태를 파악하고 개입 방안을 모색하는 다양한 기관들을 포함하였다. 마지막으로 프로그램은, 아동청소년

들에게 폭력 문제에 대한 심각성을 상기시키거나 학교폭력 상황에서 올바르게 대처할 수 있는 능력을 기를 수 있게 도움을 주는 다양한 예방교육 및 상담프로그램(1388헬프콜, 치유캠프 등)으로 구성되었다[표 8].

표 8. 대응방안 분류체계

대분류	중분류	소분류	핵심용어(예시)
대응방안	법제도 제정	-	청소년보호법, 소년법, 학원폭력방지 및 처리에 관한 지침, 학교전담경찰관 제도 등
	담당부처	-	청소년 상담지원센터, 청소년 비행예방센터, 학생안전통합시스템 등
	프로그램 운영	-	1388헬프콜 운영, 예방교육, 치유캠프운영 등

IV. 결론 및 고찰

본 연구는 Noy와 McGuinness[19]의 모형을 기본틀로 활용하여 한국 아동청소년의 학교폭력 소셜빅데이터 수집 및 분석을 위한 온톨로지를 개발하는 과정을 상세히 설명하고, 완성된 온톨로지 모형을 제시하였다. 본 연구에서 개발한 온톨로지는 학교폭력의 형태 및 유형, 학교폭력의 원인, 학교폭력 대상자들의 개인·가정·사회적 특징, 정책적·사회적 반응 및 대응방안 등의 내용을 포함하고 있다. 기존의 위기청소년 온톨로지[26]와 서적, 논문, 법·제도, 서비스·프로그램, 행정문서, 소셜미디어 등 다양한 자료원을 활용하여 학교폭력 분류체계를 구축하였다. 학교폭력의 핵심용어를 추출한 후, 용어들의 유목화와 계층화를 통해 클래스를 구성하였으며, 클래스의 속성, 속성값, 개체를 부여함으로써 학교폭력 현상을 설명하는 용어체계로 구성된 온톨로지를 구축·완성하였다.

이러한 온톨로지 구축의 실천적 함의는 최근 급변하는 학교폭력의 양상을 분석하고, 학교폭력에 대한 사회적 인식을 이해하는 데 있다. 학교폭력 가담자들이 저연령화되고 학교폭력 종류와 방법도 점점 다양해지는 가운데[4][5], 정보통신기술의 발전으로 온라인 매체를 통한 학교폭력 경험 또한 빠르게 변하고 있다. 온라인 상에서 실시간으로 새로운 학교폭력 가해 방법들이 생겨나고, 온라인 공간의 특징인 정보의 확산성, 지속성

등으로 인해 오늘날 아동청소년은 더욱 새로운 유형의 학교폭력 경험을 하고 있다. 이뿐 아니라, 전통적 오프라인 학교폭력과 온라인 학교폭력이 함께 이루어지고 있어 학교폭력은 더욱 복잡·다양한 양상을 띠고 있다 [10]. 또한, 디지털기 사용의 확산으로 학교폭력 관련 정보가 쉽고 빠르게 재생산되면서 특정 학교나 특정 지역에서 일어난 학교폭력 사건에 대한 정보의 공유가 용이해졌으며, 학교폭력 현상에 대한 개인의 의견도 SNS 등을 통해 쉽게 표출·공개하며 학교폭력에 대한 사회적 인식에 영향력을 미칠 수 있게 되었다.

이처럼 인터넷 공간은 학교폭력이 이루어지는 공간이면서, 학교폭력에 대한 사회적 시각에 대한 정보를 제공하기도 한다. 복잡·다양해지며 급변하고 있는 학교폭력 현상을 이해하고, 사회 문제에 대한 해결책을 찾기 위해서는 방대한 양의 정보를 제공할 뿐 아니라, 실시간 변화하는 정보에 민감한 소셜빅데이터 분석이 요구된다. 온톨로지 개발은 바로 이러한 소셜빅데이터 자료 수집 및 분석에 선행되어야 하는 단계이다. 따라서 아동청소년 학교폭력 온톨로지를 개발하고, 완성된 모형을 제시한 본 연구는 향후 실천가 및 연구자들에게 소셜빅데이터를 활용하여 아동청소년의 학교폭력 현상 뿐 아니라, 다양한 사회 문제에 대한 이해를 넓히는데 이바지할 것으로 기대한다.

본 연구는 다음 세 가지 제한점을 지닌다. 첫째, 온톨로지에 아동청소년 학교폭력 관련 모든 핵심용어를 전부 추출하여 클래스를 정하고 속성을 부여하기란 현실적으로 불가능하다. 본 연구에서 아동청소년 학교폭력의 대상, 원인, 유형, 장소, 지역, 대응방안의 6가지 영역과 관련된 핵심용어들을 이론, 법·제정, 서비스·프로그램에 대한 다양한 자료원에서 추출했을 뿐 아니라, 인터넷 자료를 직접 검색함으로써 아동청소년들이 흔히 사용하는 은어, 신조어도 파악하려는 노력을 취했다. 이처럼 다양한 자료원을 조사하여 핵심어를 최대한 포괄적으로 제시하려고 했음에도 불구하고, 현존하는 학교폭력 관련 용어들을 전부 제시하기란 어려울 것이다.

2 연구자들이 구축한 온톨로지를 웹크롤러에 적용했을 때, 학교폭력 관련 문서에서 반복적으로 등장하는 단어들(예: 학교폭력 가해 방법으로 활용되었던 바디워시, 담요 등). 연구자들은 추후 온톨로지 보완작업을 하면서 이렇게 새로 추출된 단어들을 최종 온톨로지에 포함시켰다.

둘째, 학교폭력이라는 사회현상을 용어들로 구성된 체계로 전환하는 데 있어서 기술적인 한계점이 있다. 인간이 개발한 온톨로지를 컴퓨터가 정확하게 인식하지 못하는 경우가 있기 때문이다. 아동청소년 학교폭력 현상과 관련된 용어들을 체계적으로 구성한 온톨로지는 추후 웹크롤러라는 온라인 데이터를 탐색하고 자료수집하는 컴퓨터 프로그램에 적용되어 소셜빅데이터를 수집하는 데에 사용된다. 하지만, 단일 글자로 이루어진 단어는 온라인 문서에서 추출하는 데에 있어 기술적인 어려움이 있어 대체 단어가 활용되거나 자료수집에서 누락 되었다. 예컨대, 본 연구에서는 가정간접요인 중 경제적인 어려움의 지표인 '빚'의 경우, '집안빚', '가정빚', '부모빚' 등으로, 학교폭력 발생 장소요인 중, '집'의 경우는 '친구집', '내집' 등으로 대체 사용하였다. 하지만, 학교폭력 가담자와 관련된 '패밀리(family)'의 약어인 '팸', 가족간접요인과 관련된 '돈', 신체적 요인과 연관이 있는 '옷' 등 대체가 어려운 단어들은 학교폭력을 이해하는 데에 중요한 용어임에도 온톨로지에 포함하지 못하였다.

셋째, 온톨로지 개발에 사용된 분류체계는 획일적이지 않으며 어떠한 참고자료를 기반으로 하는가에 따라 대분류별 하위분류체계가 다르게 설정될 수 있다. 예를 들어, 학교폭력 대응방안의 경우, 본 연구에서는 법제도, 담당부처, 프로그램 운영으로 중분류를 설정하였으나, 학교폭력 대응방안은 대응(지원)체계, 대응(지원)기관, 대응(지원)프로그램 등으로 분류가 가능하며, 소분류 역시 학교폭력 역할자 별 대응방안으로 구별이 가능할 수 있다.

이러한 제한점에도 불구하고, 본 연구는 온톨로지 수집 및 개발과정을 구체적으로 설명할 뿐 아니라, 최근 심각한 사회 문제인 아동청소년의 학교폭력을 연구 주제로 다루며, 기존의 연구방법과는 다소 차별을 두는 소셜빅데이터를 활용한 연구모형을 제시하였다는 데에 의미를 둔다. 온라인 공간은 학교폭력이 직·간접적으로 이루어지는 곳이면서 학교폭력 관련 정보를 상호공유하고 사회적으로 확산시키는 기능을 하는 곳이다. 이 때문에 소셜빅데이터 분석의 기초가 되는 온톨로지 개발 연구는 미시적으로는 학교폭력에 대상자들에 대한 이해를 제고시킬 뿐 아니라, 거시적으로는 학교폭력이

라는 사회현상을 바라보는 한국사회의 시각에 대한 정보를 제공하는 실천적 함의가 있다.

참 고 문 헌

[1] E. D. Barker, L. Arseneault, M. Brendgen, N. Fontaine, and B. Maughan, "Joint development of bullying and victimization in adolescence: Relations to delinquency and self-harm," *Journal of the American Academy of Child & Adolescent Psychiatry*, Vol.47, No.9, pp.1030-1038, 2008.

[2] D. Wolke and S. T. Lereya, "Long-term effects of bullying," *Archives of Disease in Childhood*, Vol.100, No.9, pp.879-885, 2015.

[3] F. Mishna, *Bullying: A guide to research, intervention, and prevention*, Oxford University Press, 2014.

[4] 신재한, 장기덕, 조용개, 학교폭력예방 및 학생의 이해, 양성원, 2017.

[5] 이경희, "학교폭력 문제 해결을 위한 패러다임의 전환," *윤리교육연구*, 제33권, pp.149-168, 2014.

[6] [http://m.kisdi.re.kr/mobile/repo/stat\\_view.m?key1=14242&selectPage=](http://m.kisdi.re.kr/mobile/repo/stat_view.m?key1=14242&selectPage=), 2018.01.10.

[7] [http://kostat.go.kr/portal/korea/kor\\_nw/2/1/index.board?bmode=read&aSeq=367381](http://kostat.go.kr/portal/korea/kor_nw/2/1/index.board?bmode=read&aSeq=367381), 2018.01.15.

[8] <http://www.kcc.go.kr/user.do?mode=view&page=A02060400&dc=K02060400&boardId=1030&cp=1&boardSeq=42403>, 2018.02.18.

[9] P. E. Jose, M. Kljakovic, E. Scheib, and O. Notter, "The joint development of traditional bullying and victimization with cyber bullying and victimization in adolescence," *Journal of Research on Adolescence*, Vol.22, No.2, pp.301-309, 2012.

[10] T. E. Waasdorp and C. P. Bradshaw, "The overlap between cyber bullying and traditional bullying," *Journal of Adolescent Health*, Vol.56, No.5, pp.483-488, 2015.

[11] I. Foster, R. Ghani, R. S. Jarmin, F. Kreuter,

and J. Lane, (Eds.), *Big data and social science: A practical guide to methods and tools*. CRC Press, 2016.

[12] H. Jung, H. A. Park, and T. M. Song, "Ontology-based approach to social data sentiment analysis: detection of adolescent depression signals," *Journal of Medical Internet Research*, Vol.19, No.7, pp.1-12, 2017.

[13] 박진수, 조원진, 노상규, "온톨로지 자동추출도구의 기능적 성능 평가를 위한 평가지표의 개발 및 적용," *지능정보연구*, 제14권, 제4호, pp.69-87, 2008.

[14] A. R. Kim, H. Park, and T. M. Song, "Development and evaluation of an obesity ontology for social big data analysis," *Healthcare Informatics Research*, Vol.23, No.3, pp.159-168, 2014.

[15] N. Guarino, D. Oberle, and S. Staab, *What is an ontology?*, In *Handbook on ontologies*, pp.1-17, Springer, 2009.

[16] R. Studer, V. R. Benjamins, and D. Fensel, "Knowledge engineering: Principles and methods," *Data & Knowledge Engineering*, Vol.25, No.1-2, pp.161-197, 1998.

[17] E. W. Kuiler, "From Big Data to Knowledge: An Ontological Approach to Big Data Analytics," *Review of Policy Research*, Vol.31, No.4, pp.311-318, 2014.

[18] R. Kishore and R. Sharman, "Computational ontologies and information systems I: foundations," *Communications of the Association for Information Systems*, Vol.14, No.1, pp.158-183, 2004.

[19] N. F. Noy and D. L. McGuinness, "Ontology development 101: A guide to creating your first ontology," *Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880*, pp.49-69, 2001.

[20] 배지영, 조미숙, "소셜빅데이터 분석을 위한 식품안전 온톨로지 구축," *보건정보통계학회지*, 제41권, 제3호, pp.278-289, 2016.

[21] 송태민, 이기호, 진달래, 천미경, 서동철, 박현애, *2016년 소셜 빅데이터 기반 보건복지이슈 동향 분석*,

- 한국보건사회연구원 정책보고 2016-02, 2016.
- [22] 정혜실, *청소년 우울 관련 소셜데이터 수집과 분석을 위한 온톨로지 개발 및 평가*, 서울대학교 대학원, 석사 학위 청구논문, 2016.
- [23] A. Shaban-Nejad, D. L. Buckeridge, and L. Dubé, "Cope: childhood obesity prevention [knowledge] enterprise," In Conference on Artificial Intelligence in Medicine in Europe, Springer, pp.225-229, 2011.
- [24] I. T. Afolabi, G. E. Olujinmi, and R. O. Nwokoye, "Ontology based decision support system for youth counseling," In The World Congress on Engineering and Computer Science, 2017.
- [25] J. P. Lousado, V. Rodrigues, N. Costa, R. Carvalho, C. Albuquerque, and V. Franco, "Medical ontology for treatment of clinical data from children and youth," *Atencion Primaria*, Vol.45, No.2, pp.223-227, 2013.
- [26] 송태민, 이연희, 정진욱, 이난희, 이기호, 진달래, 천미경, 진재현, 정우진, 김정선, 류시원, 박현애, 서동철, 이원상, 이충근, *빅데이터 분석기반의 위기청소년 예측 및 적시대응기술개발-소셜 빅데이터 분석을 통한 국가차원의 청소년 위기관리 체계 모형 개발*, 한국보건사회연구원 정책보고 2015-05, 2015.
- [27] L. J. Griffith, D. Wolke, A. S. Page, J. P. Horwood, and ALSPAC Study Team, "Obesity and bullying: Different effects for boys and girls," *Archives of Disease in Childhood*, Vol.91, No.2, pp.121-125, 2005.
- [28] J. Juvonen and S. Graham, "Bullying in schools: The power of bullies and the plight of victims," *Annual Review of Psychology*, Vol.65, pp.159-185, 2014.
- [29] R. Veenstra, S. Lindenberg, A. J. Oldehinkel, A. F. De Winter, F. C. Verhulst, and J. Ormel, "Bullying and victimization in elementary schools: a comparison of bullies, victims, bully/victims, and uninvolved preadolescents," *Developmental Psychology*, Vol.41, No.4, p.672-682, 2005.
- [30] D. L. Haynie, T. Nansel, P. Eitel, A. D. Crump, K. Saylor, K. Yu, and B. Simons-Morton, "Bullies, victims, and bully/victims: Distinct groups of at-risk youth," *The Journal of Early Adolescence*, Vol.21, No.1, pp.29-49, 2001.
- [31] T. R. Nansel, W. Craig, M. D. Overpeck, G. Saluja, and W. J. Ruan, "Cross-national consistency in the relationship between bullying behaviors and psychosocial adjustment," *Archives of Pediatrics & Adolescent Medicine*, Vol.158, No.8, pp.730-736, 2004.
- [32] J. A. Stein, R. L. Dukes, and J. I. Warren, "Adolescent male bullies, victims, and bully-victims: A comparison of psychosocial and behavioral characteristics," *Journal of Pediatric Psychology*, Vol.32, No.3, pp.273-282, 2007.
- [33] R. K. Merton, "Social structure and anomie," *American Sociological Review*, Vol.3, No.5, pp.672-682, 1938.
- [34] R. Agnew, "Foundation for a general strain theory of crime and delinquency," *Criminology*, Vol.30, No.1, pp.47-88, 1992.
- [35] R. Agnew and T. Brezin, *Juvenile delinquency: Causes and control (4th ed.)*, Oxford University Press, 2012.
- [36] T. Hirschi, *Criminology theory: Selected classic readings: A control theory of delinquency* (pp.289-305), Routledge, 1969.
- [37] M. R. Gottfredson and T. Hirschi, *A general theory of crime*, Stanford University Press, 1990.
- [38] U. Bronfenbrenner, Ecological models of human development. In T. Husen, & T. N. Postlethwaite (Eds.), *The international encyclopedia of education*(pp.1643-1647), (2nd ed.), Elsevier Sciences, 1994.
- [39] J. S. Hong and D. L. Espelage, "A review of research on bullying and peer victimization in school: An ecological system analysis," *Aggression and Violent Behavior*, Vol.17, No.4, pp.311-322, 2012.
- [40] J. E. Donovan and R. Jessor, "Structure of problem behavior in adolescence and young adulthood," *Journal of Consulting and Clinical*

Psychology, Vol.53, No.6, pp.890-904, 1985.

[41] K. Björkqvist, K. M. Lagerspetz, and A. Kaukiainen, "Do girls manipulate and boys fight? Developmental trends in regard to direct and indirect aggression," *Aggressive Behavior*, Vol.18, No.2, pp.117-127, 1992.

[42] J. J. Dooley, J. Pyżalski, and D. Cross, "Cyberbullying versus face-to-face bullying: A theoretical and conceptual review," *Zeitschrift für Psychologie/Journal of Psychology*, Vol.217, No.4, pp.182-188, 2009.

**저 자 소개**

**한 윤 선(Yoonsun Han)** 정회원



- 2012년 8월 : Univ. of Michigan (박사)
- 2018년 ~ 현재 : 서울대학교 사회 복지학과 조교수

〈관심분야〉 : 학교폭력, 청소년 부적응과 보호

**김 하 영(Hayoung Kim)** 정회원



- 2015년 8월 : 성균관대학교 공학사, 문학사
- 2018년 8월 : 성균관대학교 문학석사
- 2018년 ~ 현재 : 한국청소년정책연구원

〈관심분야〉 : 아동청소년, 진로발달, 연구방법론

**송 주 영(Juyoung Song)** 정회원



- 2012년 7월 : 미시간주립대학교 형사사법박사
- 2015년 ~ 현재 : 펜실베이니아주립대학교 범죄학과 조교수

〈관심분야〉 : Big Data, 머신러닝, 청소년범죄, 학교폭력

**송 태 민(Tae Min Song)** 정회원



- 2006년 2월 : 동국대학교 컴퓨터공학과(공학박사)
- 2017년 ~ 현재 : 삼육대학교 보건관리학과 교수

〈관심분야〉 : Big Data, 머신러닝, u-Health, 학교폭력