

한국청소년문화연구소
<청소년문화포럼>VOL.68

ISSN 1975-2733(Print)/ 2713-797X(Online)
<http://dx.doi.org/10.17854/ffyc.2021.10.68.59>

텍스트 마이닝을 통한 청소년 사이버불링 키워드 분석
Keyword Analysis of Adolescent Cyberbullying through Text
Mining

이 세 립 · 김 진 영 · 전 종 설

이화여자대학교 사회복지학과 박사수료 · 이화여자대학교 사회복지학과
박사수료 · 이화여자대학교 사회복지학과 교수

Lee, Serim · Kim, Jinyung · Chun, JongSerl

Ewha Womans University, Department of Social Welfare, PhD candidate ·

Ewha Womans University, Department of Social Welfare, PhD candidate ·

Ewha Womans University, Department of Social Welfare, Professor

텍스트 마이닝을 통한 청소년 사이버불링 키워드 분석

Keyword Analysis of Adolescent Cyberbullying through Text Mining

이 세 립*1) · 김 진 영**2) · 전 종 설***3)

Lee, Serim* · Kim, Jinyung** · Chun, JongSerl***

국 문 요 약

본 연구는 텍스트 마이닝을 통해 빠르게 진화하는 청소년 사이버불링의 주요 이슈를 파악하고자 했다. 최근 1년 동안의 뉴스, 블로그, 댓글 데이터를 ‘사이버불링’, ‘사이버폭력’ 과 같은 키워드를 사용해 수집하였다. 총 3,075건의 문서와 190,706개의 키워드가 추출되었다. 해당 데이터를 기반으로 핵심 키워드 및 빈도 수 분석, 토픽모델링, 워드클라우드 시각화, 연관어 분석, 감성어 분석, 기간별 시계열 분석을 진행하였다. 분석결과, 사이버불링과 관련된 상위 키워드로는 학교폭력, 범죄, 처벌 등이 추출되었으며, 성폭력, 학교폭력, 알페스, 대상 및 대응의 총 4개의 군집 및 토픽이 나타났다. 사이버불링 유형, 대상, 대응으로 구분하여 분석한 결과, 사이버불링 유형에서는 시각적 및 성폭력이 심각한 사회적 문제로 나타났으며, 지인능육, 딥페이크, 알페스 등과 같이 새로운 유형으로도 진화하고 있음을 확인했다. 또한, 사이버불링 유형 관련 용어가 혼용되고 있음을 확인했다. 대상으로는 청소년이 가장 핵심 키워드로 나타났으며, 대응에서는 법적 및 제도적 대응이 중점적으로 나타났다. 이에 사이버불링 예방 서비스 프로그램 제공의 중요성과 정신건강 관련 서비스의 제공이 보다 강화되어야 함을 제안했다.

주제어: 사이버불링, 텍스트 마이닝, 토픽 모델링, 연관어 분석, 감성어 분석, 시계열 분석

1) 이화여자대학교 사회복지학과 박사수료

2) 이화여자대학교 사회복지학과 박사수료

3) 이화여자대학교 사회복지학과 교수, 교신저자

I. 서론

정보통신발달과 인터넷 및 스마트폰 기기 등의 보급률이 높아지면서 최근 시간과 장소에 제한 없이 사이버 공간에서의 교류가 잦아지고 있고, 그로 인해 사이버불링 가해와 피해가 늘어나고 있는 추세이다(이창호, 신나민, 하은빈, 2014; 이승현, 강지현, 이원상, 2015). 그러나 충분한 온라인 교육과 윤리적 논의가 고려되지 않은 상태이기에 사이버불링 피해가 더욱 문제시 되고 있다. 사이버불링에 대한 개념적 정의가 아직 명확하지 않으므로, 다양한 학자들이 각기 다른 정의를 제공하고 있고(Chun et al., 2020), 한국에서도 사이버폭력과 사이버불링을 혼용해서 사용하는 것으로 확인된다. 예를 들어, 한국정보화진흥원(2019)이 발표한 사이버폭력 조사에서는 사이버폭력을 “사이버(인터넷, 휴대전화 등) 공간에서 언어, 영상 등을 통해 타인에게 피해 혹은 불쾌감을 주는 행위를 의미함(p.20)” 이라고 정의하면서 사이버폭력 아래 사이버 언어폭력, 명예훼손, 스토킹, 성폭력, 유출, 따돌림, 갈취, 강요와 같은 8가지 유형을 포함하였다. 여기서 사이버따돌림이란 “인터넷 대화방이나 스마트폰, 인스턴트 메시지 등에서 상대방을 따돌리는 행위(p.20)” 로 정의되고 있다. 즉, 사이버불링을 사이버폭력의 하위요인으로 간주한 것이다. 이와 달리 해외문헌에서는 사이버폭력(cyber violence)이 아닌 사이버불링(cyberbullying)이라는 용어를 보편적으로 사용하고, 이에 대한 개념을 정의하는 경우를 자주 살펴볼 수 있다. 그중에서도 일반적으로 쓰이는 사이버불링 정의는 다음과 같다: Patchin과 Hinduja(2011)은 사이버불링은 이메일이나 채팅룸, 소셜 네트워크 웹 사이트를 이용하여 친구들을 괴롭히는 행위라고 정의하였으며, Tokunaga(2010)은 사이버불링을 개인 또는 집단에 의해서 온라인 또는 디지털 미디어 매체를 사용하여 행해지는 모든 행위를 뜻하며, 이는 위협적이고 공격적인 메시지를 반복적으로 보내면서 타인에게 위협을 가하거나 불편감을 주기 위함이라고 정의하였다. Chun et al.(2020)의 경우, 보다 넓게 사이버불링을 전자 매체 또는 기기를 활용하여 행해지는 반복적이고도 고의적인 행동이자 타인을 협박, 괴롭힘, 창피를 주는 행위라고 일컬었다. 본 개념적 정의들을 종합적으로 고려해봤을 때, 본 연구는 사이버불링은 단순히 누군가를 온라인 상에서 따돌리거나 괴롭히는 행위만으로 보기보다는 Chun et al.(2020)의 정의에 따라 온라인 상에서 타인에게 행해지는 모든 “반복적이고도 고의적인 행동이자 타인을 협박, 괴롭힘, 창피를 주는 행위” 를 포함하고자 한다.

사이버불링의 피해는 국내·외 모두 증가하고 유형 또한 다양하게 진화하는 추세이며 그 영향과 심각성은 매우 높다. 전 세계 아동의 약 60%가 사이버 위협에 노출되어있으며, 그중에서도 사이버불링이 45%로 아동이 가장 많이 노출되는 사이버 위협인 것으로 나타났다

(DQ Institute, 2020). 국내와 미국, 호주, 캐나다 등의 국가에서 진행된 다수의 연구에서, 평균적으로 약 25%에서 30% 정도의 학생들이 사이버불링 피해를 당했다고 진술했다(Cénat et al., 2015; Chapin, 2016; Li, 2007; 이인태, 2012; Bilic, 2013; 이창호 외, 2014; Hemphil, Tollit, Kotevski & Heerde, 2015; Hinduja & Patchin, 2013). Li(2010)에 따르면 캐나다 중학생·고등학생들의 약 40%가 사이버불링 피해 경험이 있으며, Mishna et al.(2010)의 연구에서도 캐나다 중학생, 고등학생의 49.5%가 사이버불링 경험이 있다고 답했다. 한국의 경우, 국외보다는 상대적으로 낮은 사이버불링 피해 경험률을 보이는데, 그럼에도 불구하고 사이버불링은 증가 또는 높은 경험률을 유지하고 있으므로 각별한 주의가 필요하다. 예를 들어, 교육부(2021)가 발표한 학교폭력 실태조사에서 피해 유형별 비중이 언어폭력 33.6%, 집단따돌림 26.0%, 사이버폭력 12.3%로, 사이버폭력이 3번째로 높은 피해 유형인 것으로 밝혀졌다. 특히, 사이버폭력의 경우 2019년 1차 조사와 비교하여, 3.4% 증가한 것으로 확인되었다(교육부, 2021). 또한, 방송통신위원회와 한국지능정보사회진흥원(2020)에서 발표한 자료에 따르면, 청소년 중 사이버폭력 가해 경험률은 2018년 20.8%에서 2020년도에는 9.5%로 전반적으로 하락하는 추세를 보이나, 피해 경험률은 2018년 20.8%에서 2020년 19.7%로 큰 차이를 보이지 않는 것으로 나타난다. 피해유형별로는 언어폭력이 9.0%로 가장 높았고, 이어서 명예훼손(1.6%), 따돌림(1.0%) 순이었다(방송통신위원회, 한국지능정보사회진흥원, 2020).

최근 코로나-19 이후 온택트 시대가 도래하며 더 많은 사람들이 온라인 환경을 통해 소통하고 있는 상황이기 때문에 사이버불링은 더 심각한 사회적 문제로 떠오르고 있다. 특히, 아동·청소년의 경우 일반적으로 학교에서 전자 매체를 사용하지 못하다가 이제는 집에서 컴퓨터, 스마트폰, 또는 기타 스마트 기기로 수업을 듣는 상황이 되었으며, 친구들과도 온라인 상에서 소통하는 일이 증가했다. 이러한 교육 및 소통 환경의 변화는 학생들을 더 오랜 시간 온라인 환경에 노출 시키고, 소셜 네트워크 서비스와 미디어의 사용시간을 증가시키고 있다. 온라인 수업 전환 후 학생들의 미디어 사용실태 조사(전종설, 이해국, 2020)에 따르면 전체 응답자 중 인터넷·미디어 사용이 ‘매우 늘었다’ 27.1%, ‘조금 늘었다’, 38.4%로 65.5%의 청소년이 온라인 수업 전환으로 인해 인터넷·미디어 사용이 늘었다고 답했다. UNICEF(2020)는 이처럼 인터넷 사용시간이 길어질수록 아동·청소년의 사이버불링 피해 위험뿐만 아니라 온라인 성착취, 유해 콘텐츠에 노출될 가능성이 증가할 수 있다고 보고하고 있다. 실제로 위의 조사에 따르면 사이버폭력에 노출되는 청소년 역시 48.3%로 매우 높게 확인되었다(전종설, 이해국, 2020).

이처럼 사이버불링 피해는 점차 증가하고 있고 그 유형 또한 다양해지고 있다. 사이버불링 피해 유형은 개인정보 유출로 인한 피해, 온라인 매체를 통한 욕설, 따돌림, 원치 않는

성적 행위나 대화 강요, 음란물 노출 등이 있으며(이창호 외 2014; 이승현 외, 2015; Jones et al., 2012), 스마트폰 사용이 점차 보급됨에 따라 온라인 괴롭힘이 와이파이(WiFi) 셔틀이나 카카오톡 감옥 등과 같은 새로운 집단 따돌림의 형태로 진화하고 있다(Cetin, Yaman & Peker, 2011; 방송통신위원회, 한국인터넷진흥원, 2013; 최진오, 2015). 이뿐만 아니라 최근에는 지인능욕, 딥페이크, 알페스와 같은 새로운 유형의 사이버불링 역시 발생하고 있다. 딥페이크는 사람들이 말한 적 또는 한 적이 없는 것을 마치 한 것처럼 묘사하는 비디오로 주로 머신러닝이나 AI를 통해 만들어진다(Raymond, 2019). AI를 활용한 무료 소프트웨어가 보급됨에 따라, 딥페이크는 레드(Reddit), 트위터, 구글 등과 같은 다양한 온라인 플랫폼에서 만연하고 있는 추세이다(Raymond, 2019). 알페스는 ‘Real Person Slash(RPS)’의 약자로 왜곡된 성적 패러디물을 뜻한다(Wang, 2017). 알페스는 주로 배우, 정치인, 역사적 인물과 같은 실제 인물을 기반으로 형성되는 동성애적 판타지물을 동경하는 팬덤을 지칭한다(Wang, 2017). 이러한 알페스는 실존 인물을 성적으로 묘사한다는 점에서 문제시 될 수 있다(Wang, 2017). 지인능욕은 딥페이크와 비슷하나, 여성의 사진 또는 여성을 성적으로 합성한 사진을 신상정보와 함께 유포하여 성착취물로 만든다는 점에서 약간의 차이가 있다(안갑철, 2021). 이처럼 사이버불링은 빠르게, 다른 형태로 변화하고 있는 것으로 확인된다.

사이버공간에서 이뤄지는 사이버불링 피해는 기존의 전통적 학교폭력과 달리 시간과 장소를 불문하고 하루 24시간 내내 진행될 수 있어 피해자의 고통은 극단적이고 매우 심각할 수 있다(이성식, 전신현, 2011; 이주형, 안순철, 2013). 또한, 사이버불링은 단 한 번의 개인 정보 유포만으로도 기록이 영원히 남고 전파성이 높아 한 사람의 일생과 주변 사람들에게 막대한 부정적 피해를 유발하며 회복하는 것이 더욱 어렵다(조윤주, 정옥분, 2009). 이에 사이버불링 피해자들은 신체적 어려움을 호소하고(안화실, 박정도, 2013), 학교부적응, 약물복용, 무단결석(이창호 외, 2014), 등교거부나 학업성적 부진(Slonje, Smith & Frisen, 2008) 등의 일상생활 장애 및 비행 행동을 겪는다. 무엇보다도 사이버불링은 청소년 정신건강에 지대한 영향을 미치는데, 사회적 불안, 낮은 자아존중감 등을 유발(Hinduja & Patchin, 2007)하고, 자살 충동이나 우울증과 같은 정신학적인 특성과 연결되어(Bonanno & Hymel, 2013; Hinduja & Patchin, 2008; 안화실, 박정도 2013), 결국에는 정서적 황폐화를 경험하고(정여주, 김동일, 2012), 극단적으로는 스스로 목숨을 끊는 상황까지 이르기도 한다(Hinduja & Patchin, 2010; Ybarra, Mitchell, Wolak, & Finkelhor, 2006; 조윤오, 2013; 이창호 외, 2014).

사이버불링에 대한 사회적 관심이 높아지며 관련 연구들도 증가하고 있지만, 국내의 경우 사이버불링이 실제로 일어나고 있는 주 매체인 소셜 네트워크를 분석하여 사이버불링의 실태를 파악한 연구는 부족한 실정이다. 예를 들어, 국내 소셜 빅데이터를 활용하여 학교폭력의

트렌드를 파악하고자 한 Kim et al.(2019)의 연구에서는 성적폭력이 빠르게 증가하고 있고, 신체적 폭력과 사이버불링 역시 빠르게 성장하고 있다고 보고하였다. 해당 연구는 다양한 학교폭력 유형의 전반적인 트렌드를 살펴본다는 점에서 의의가 있으나, 사이버불링에 대한 트렌드를 파악하였기보다 학교폭력의 하위요인으로서 사이버불링을 연구하였다는 점에서 본 연구와 차이가 있다. 해외의 경우, Tahamtan과 Huang(2019)의 연구에서 소셜미디어인 트위터에서 데이터를 수집하고, 텍스트 마이닝(text mining) 기법을 통해 사이버불링에 대한 사람들의 인식을 분석했다. 분석결과, 사람, 아이들, 학생, 학교, 중지가 가장 중심 키워드로 나타났으며, 대처방안(예. 부모와 교사의 지도), 사람들의 관심을 끌었던 특정 사건(예. 미시간 사이버불링법), 사람들의 주요 우려(예. 학생들의 정신건강 문제)의 총 세 개의 토픽이 발견되었다. 본 연구는 다양한 플랫폼 중에서도 트위터에서만 정보를 수집하였다는 점에서 정보의 풀에 한계가 있었다. 이러한 선행연구를 토대로 본 연구는 텍스트 마이닝을 통해 사이버불링의 트렌드를 파악하고자 하며, 분석 결과를 대상자, 대처방안, 특정 사건, 우려 등과 같은 카테고리로 분류하여 이해하고자 한다. 또한, 본 연구는 선행연구들의 한계점을 극복하기 위하여, 네이버 뉴스뿐만 아니라 블로그에서도 자료를 수집하였으며, 다른 현상의 하위요인이 아닌 사이버불링이라는 현상 자체를 지정하여 탐색하였다는 점에서 의의가 있다.

기존의 사회 조사 연구는 정해진 변인들의 관계 파악에 있어서 유용하지만 실제 이용자들이 사이버상에서 언급한 방대한 정보를 빠르게 수집 및 분석하는 데에는 한계가 있을 수밖에 없다. 사이버불링은 인터넷 상에서 발생하는 것이기 때문에 인터넷을 기반으로 한 소셜미디어와 높은 연관성을 보인다. 또한, 소셜미디어 데이터를 활용한 사이버불링 연구는 더욱 유용하며(이지은, 전종설, 2021) 더욱 많은 양의 데이터를 활용하여 다양한 개인의 생각과 의견을 민첩하게 파악할 수 있다. 이와 같은 맥락에서 송태민, 송주영, 진달래(2014)는 사회복지 현장에서 소셜 빅데이터를 활용할 시 사이버불링의 예방 또는 개입과정에 대한 방대한 양의 데이터를 기반으로 수많은 인터넷 이용자의 목소리를 확인할 수 있기 때문에 기존의 제한된 문항과 표본추출을 통해 정보를 얻어내는 설문조사와 비교하여 보다 유용한 자료가 될 수 있다고 제안했다.

즉, 소셜 데이터는 개인의 심리적, 행동적 특성과 의견을 실시간으로 드러내고 사회적 트렌드를 그 어떤 데이터보다도 빠르고 민감하게 반영하기 때문에 기존의 분석에서 얻지 못했던 통찰력을 얻을 수 있다고 평가되는데(한국정보화진흥원, 2012), 무엇보다 사이버불링은 기술 플랫폼의 발전에 따라 매우 빠르게 진화 및 변모하고 있기 때문에 소셜 데이터를 활용하는 것이 요구된다. 따라서 본 연구에서는 소셜 미디어 텍스트 마이닝을 통해 청소년

사이버불링 키워드를 분석하여 사이버불링 관련 이슈를 파악하고 연구 및 실천현장에서의 함의점을 제시하고자 한다.

II. 연구 방법

1. 자료수집

본 연구에서는 청소년 사이버불링과 관련한 소셜 미디어(인터넷 기사, 댓글, 블로그)를 대상으로 텍스트마이닝 기법을 통해 핵심용어를 추출하였다. 인터넷 뉴스와 블로그 포스트는 사회적 의제설정 영향력을 지닌 대표 미디어이다(기재홍, 안승혁, 2020). 먼저, 기사를 발행하는 미디어 언론은 의제설정과 프레이밍을 통해 의제 형성에 영향을 미칠 뿐 아니라 가장 빠르게 사회의 이슈를 전달하는 역할을 한다(이동근, 2019). 또한, 인터넷 뉴스는 인터넷의 쌍방향성이라는 특성상 이용자가 능동적으로 정보를 찾고 수용하기 때문에 시민들의 여론 형성에 핵심적인 역할을 하고 있다(반현, 2003). 본 연구는 이러한 다양한 언론사의 뉴스 기사들을 종합적으로 수집하고 활용하기 위해 포털사이트 네이버 뉴스를 활용했다. 다음으로 블로그는 언론에 상응하는 사회적 의제 설정의 파급력을 가진 소셜미디어의 대표적인 매체로 특히 다른 소셜미디어 매체와 비교하여 더 심도 있고 긴 텍스트를 가지고 있기 때문에 텍스트 분석 대상으로 적합하다(안형준, 하영목, 2019). 이에 네이버 블로그 자료 역시 활용하였다.

뉴스, 블로그 문서 중 필수 포함 키워드 중 하나 이상과 각각의 검색 키워드를 포함하는 문서를 수집했으며 수집된 문서의 중복을 제거한 후 전체 문서에 대하여 빈도 상위 키워드를 조사했다. 검색 키워드는 ‘사이버불링’, ‘사이버폭력’, ‘온라인폭력’, ‘사이버학폭’, ‘사이버따돌림’, ‘온라인따돌림’, ‘사이버괴롭힘’, ‘온라인괴롭힘’ 중 한 가지와 더불어 ‘청소년’, ‘중학생’, ‘고등학생’ 중 한 가지의 키워드가 함께 검색되었다. 예를 들어 ‘사이버불링’ 그리고 ‘중학생’ 등의 조합으로 키워드가 검색되었다. 또한, 자료 수집 기간은 최근 1년의 자료(2020.3.9~2021.3.8)를 수집하였다. 소셜미디어 데이터는 프로그래머가 특정 기술을 사용하기 쉽게 만들어 주는 통신 기법(이지은, 전종철, 2021)인 Application Programming Interface(API)와 특정 문자 추출(parsing)을 통해 원하는 문자 정보를 얻는 것으로 웹상의 소스코드 분석을 통해 데이터를 수집하는 방식(이지은, 전종철, 2021; Mitchell, 2018)인 웹크롤링(web crawling)을 통해 수집하였다. 총 3,075건의 문서가 수집되었는데 채널

별로는 인터넷 뉴스 총 2,194건, 블로그가 881건이었다.

키워드 전처리의 경우, 제거 작업(불용어, 무관한 단어, 너무 일반적인 용어 등 제외), 교정 작업(단/복수, 약어, 띄어쓰기, 품사형태 변경 등과 같은 키워드 클렌징), 통제 작업(동의어, 유사어, 광의어, 협의어 등 통제)으로 나눌 수 있다(이수상, 2012). 그러나 일반적으로 소셜 빅데이터 분석에서는 연구자의 주관이 최대한 개입되지 않도록 하고 있으므로(이수상, 2012), 본 연구에서도 최소한으로 키워드 전처리를 진행하였다. 이에 먼저, 불용어를 삭제하고 사이버불링과 전혀 무관한 단어를 먼저 제거하였다. 또한, 키워드 정제 과정 역시 연구자의 주관이 최대한 개입되지 않는 선에서 단·복수 교정, 띄어쓰기, 품사형태 변경 등과 같은 교정 작업이 수행되므로(서선경, 정은경, 2013), 본 연구도 이에 따라 키워드 정제 과정을 수행하였다. 예를 들어, ‘사이버폭력’과 ‘사이버 폭력’은 모두 같은 개념인데도 불구하고 띄어쓰기로 인해 다른 키워드로 구분되기 때문에 사이버폭력으로 통일했다. 또한, 통제 작업의 경우에도 위에서 설명한 바와 같이 띄어쓰기, 품사 등의 형태로 인해 동의어, 유사어가 다르게 구분되는 것만을 통제할 것이 제안되기 때문에(서선경, 정은경, 2013) 최소한으로 진행하였다. 서선경과 정은경(2013)의 연구에서 제안하듯이 광의어, 협의어 등은 통제하지 않았는데, 이에 청소년과 청소년에 포함되는 고등학생, 중학생은 모두 다른 키워드로 보았다.

2. 자료분석

소셜 빅데이터의 분석은 사용자가 남긴 문서의 의미를 분석하는 것으로 주로 자연어 처리기술인 주제분석과 감성 분석 기술인 오피니언 마이닝(opinion mining) 등을 실시할 수 있다(이지은, 전종설, 2021). 즉, 텍스트 마이닝은 텍스트 형태의 비정형 데이터에 데이터 마이닝 기법을 적용한 것으로, 일반적으로 사용되는 문자로 저장된 데이터를 처리하는 다양한 관련 기술을 총칭한다(Miner et al., 2012). 결국, 텍스트 마이닝은 통계적 추론 기법을 활용하여 대규모 비정형 데이터인 텍스트에서 새로운 지식을 발견할 수 있게 해준다(강승지, 손유진, 2020).

본 연구는 FROTOMA(2021) 업체에서 구축한 Pro-TASTM(Trend Analysis System)을 활용하였다. Pro-TASTM는 FROTOMA 업체에서 Zava와 Python 언어를 이용하여 구축한 트렌드 동향분석 시스템이다. 해당 시스템은 정부의 경제, 사회, 문화, 교육 등 제반 정책 과제 또는 기업의 마케팅 등 주요 현안 이슈에 대한 온라인상의 여론을 분석한다(FROTOMA, 2021). 본 시스템은 언론 기사, 언론 기사의 댓글, 블로그상의 문서를 수집한 후 수집 문서의 정제를

통한 키워드 추출 및 키워드의 중요도 및 빈도와 키워드 간 연관성 분석이 이루어질 수 있도록 하며 워드 클라우드 인터페이스가 제공된다. 또한, 이슈 연관어 분석, 댓글 감정 분석, 댓글 상위 기사, 토픽 모델링 등 각종 통계 커스터마이징이 가능하다.

따라서 본 연구는 먼저 핵심 키워드와 빈도수를 분석하고 주요 키워드를 파악했다. 이어서 토픽모델링을 통해 군집, 토픽별 주요 키워드를 파악했다. 토픽모델링은 확률적 기법을 사용하여 여러 유형의 문서와 대량 데이터로부터 다양한 개념이나 토픽을 추출하여 숨겨진 의미 구조를 발견할 수 있도록 한다(Tsolmon & Lee, 2017). 토픽모델링을 통해 추출된 토픽 및 키워드에 따라 연구자가 토픽명을 부여하였다. 또한, Multidimensional Scaling을 통한 Intertopic Distance Map(IDM)을 생성하여 토픽들의 비중과 유사도를 파악하였다. 워드 클라우드를 통해 핵심 키워드를 시각화하고 핵심 키워드별 연관어 분석을 통해 연관어를 확인하였다. 이때, 블로그와 뉴스에 나타난 각각의 사이버불링 키워드를 비교하여 매체별 특성에 따라 나타난 사이버불링 이슈의 차이를 분석하였다. 사이버불링 현황 파악을 위해 댓글의 감성어 분석을 실시했다. 오피니언 마이닝이라고도 불리는 감성분석은 의견, 느낌, 기분을 표현하는 주관적인 내용의 문서에서 나타나는 감정이 긍정, 중립, 부정 중 어디에 해당하는지 정량적으로 예측하기 위해 이용하는 방법이다(Sarkar, 2019). 또한, 기간별 변화의 비교와 시계열 분석을 통해 최신 사이버불링 이슈의 변화 및 트렌드를 파악하고자 했다.

Ⅲ. 연구 결과

1. 키워드 빈도

분석결과, 총 190,706개의 키워드가 추출되었다. 핵심 키워드 및 빈도수 결과, 사이버불링 대상, 유형, 대응 방안과 관련된 핵심 키워드가 추출되었다. 뉴스와 블로그를 통합하여 사이버불링 상위 10위 키워드는 청소년(n=10,827), 학교폭력(n = 9387), 범죄(n = 9,006), 처벌(n = 8,334), 디지털(n = 8,294), 피해자(n = 8,104), 학생(n = 7,658), n번방(n = 7,053), 여성(n = 6,708), 성범죄(n = 6,552)였다. 뉴스와 블로그를 통합하여 최종 주요 100개의 키워드는 아래의 <표 1>과 같다.

<표 1> 최종 주요 100개 키워드

키워드	빈도	키워드	빈도	키워드	빈도	키워드	빈도	키워드	빈도
청소년	10827	성폭력	3421	미성년자	1702	삭제	1169	채팅	944
학교폭력	9378	사이버	3195	성적	1665	서비스	1161	성인	941
범죄	9006	변호사	3090	정책	1650	강훈	1143	선고	932
처벌	8334	예방	3080	교사	1559	폭행	1099	소지	917
디지털	8294	박사	3040	가족	1489	혐오	1067	수업	904
피해자	8104	촬영	2982	검찰	1447	알페스	1057	성폭행	866
학생	7658	불법	2767	청원	1361	선수	1036	교수	864
n번방	7053	제작	2725	참여	1355	인공지능	1020	검거	862
여성	6708	사이버폭력	2567	합성	1335	신체	1013	재판	856
성범죄	6552	조주빈	2345	남성	1320	형사	1013	방지	850
착취	6058	대응	2293	마돌림	1302	남자	1008	표현	845
학교	5683	답페이지	2191	기술	1267	자녀	994	경찰청	836
아동	5391	부모	2179	영상물	1262	사회적	989	공범	832
온라인	4395	징역	2026	전화	1251	차별	979	범행	826
폭력	4230	음란물	2021	안전	1234	중학생	969	여자	824
가해자	4088	친구	1982	얼굴	1229	법원	956	예방교육	821
코로나19	3804	신상	1920	양형	1222	기소	955	콘텐츠	800
텔레그램	3800	집단	1767	미디어	1215	가정	953	정신	778
유포	3768	협박	1760	게임	1210	의혹	946	교육부	776
아이	3611	인터넷	1730	인권	1209	심리	944	스타	770

2. 토픽모델링

토픽모델링 분석결과, 총 4개의 군집(cluster)과 토픽(topic)이 나타났다. 군집 수를 결정하기

위해 토픽 간의 차이를 반영하는 λ 값을 반복조정하며 최적의 토픽수를 산출하였다(강지연, 김수경, 노승국, 2019). 토픽들이 서로 겹치지 않고 독립적이고 토픽 간 경계가 명확해지도록 하였다(강지연 외, 2019). 결과적으로, 각 토픽 주제의 범주를 가장 적합하게 구성했다고 판단된 4개의 토픽이 산출되었으며 이때 λ 값은 0.8이었다. 또한, 토픽별로 빈도수가 높은 상위 30개 단어가 추출되었다. 4개의 토픽과 상위 30개 키워드는 다음 <표 2>와 같다. 군집 1, 2, 3은 사이버불링의 유형과 관련한 토픽이 나타났다. 군집 4는 사이버불링의 대상 및 대응에 대한 토픽이 나타났다.

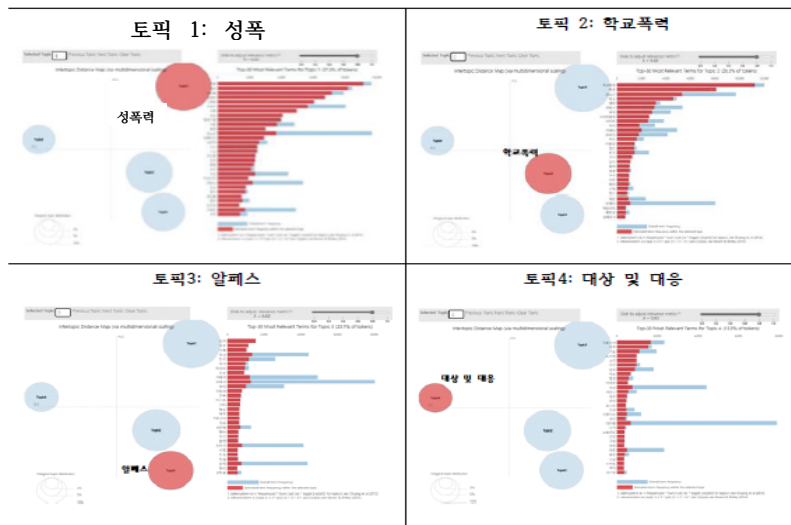
군집 1에서는 일명 N번방 사건으로 대표되는 성적 사이버폭력 사건과 관련된 키워드들이 나타나 “성폭력”으로 토픽명을 부여했다. 주요 키워드 예시로는 n번방, 딥페이크, 성폭력, 합성, 성범죄 등이 있다. 군집 2에서는 일명 학교폭력 폭로사건과 관련된 키워드들이 나타나 “학교폭력”으로 토픽명을 부여했다. 주요 키워드 예시로는 학교폭력, 따돌림, 학폭위, 교사, 학교 등이 나타났다. 군집 3에서는 알페스로 대표되는 유명한 대상 사이버폭력 관련 키워드들이 나타나 “알페스”로 토픽명을 부여했다. 주요 키워드 예시로는 여주, 선수, 스타, 배우, 성희롱 등이 있다. 군집 4는 “대상 및 대응”으로 토픽명을 부여했다. 주요 키워드 예시로는 정책, 서비스, 시스템, 인공지능, 보호처분 등을 들 수 있다.

<표 2> 토픽모델링을 통한 4개의 토픽과 상위 30개 키워드

군집	토픽	키워드
1	성폭력	처벌, 범죄, 디지털, 성범죄, n번방, 피해자, 착취, 유포, 텔레그램, 아동, 촬영, 청소년, 딥페이크, 성폭력, 박사, 신상, 음란물, 징역, 불법, 제작, 여성, 미성년자, 변호사, 합성, 형사, 영상물, 협박, 조주빈, 가해자, 성적
2	학교폭력	학교폭력, 학생, 청소년, 학교, 예방, 변호사, 폭력, 사이버폭력, 사이버, 아이, 가해자, 온라인, 부모, 따돌림, 집단, 친구, 교사, 심리, (학)폭위, 폭행, 자녀, 나무, 왕따, 신체, 정신, 대응, 피해자, 예방교육, 중학생, 실태조사
3	알페스	남자, 여주, 의혹, 여성, 친구, 여자, 미디어, 지수, 가해자, 피해자, 아이, 선수, 스타, 혐오, 배우, 커뮤니티, 가수, 성폭행, 멤버, 연간, 온라인, 시절, 진실, 차별, 폭력, 윤리, 성희롱
4	대상 및 대응	코로나-19, 정책, 기술, 소년법, 소년, 안전, 참여, 언택트, 환경, 의정부, 여성, 서비스, 투표, 유망, 포스트, 인권, 인공지능, 청년, 디지털, 자격, 보호처분, 안양, 구제, 변영, 대응, 통합, 시설, 스마트, 추이, 시스템

Multidimensional scaling 분석을 통한 IDM은 다음 <그림 1>과 같다. IDM은 토픽의 크기를 통해 토픽의 비중과 토픽 간 거리를 통해 토픽 간의 연관성 및 유사성을 파악할 수 있도록 하는 지도(map)이다(신명선, 조경원, 2019). 각 그림은 왼쪽에서 토픽모델링 분석에 따른 각 군집 지도를 보여주고, 오른쪽에서는 해당 군집의 상위 30개의 키워드를 보여준다. 아래

<그림 1>에서 확인할 수 있듯이 토픽의 크기가 가장 큰 것은 토픽 1(성폭력)이다. 즉, 성폭력이 가장 비중이 높은 토픽으로 나타났으며 학교폭력, 알페스, 대상 및 대응이 그 뒤를 이었다. 토픽 간 거리의 경우, 토픽 2(학교폭력)와 토픽 3(알페스)의 거리가 가장 가까워 비교적 연관성과 유사성이 높은 것으로 볼 수 있다. 반면 토픽 1(성폭력)과 토픽 3(알페스), 토픽 1(성폭력)과 토픽 4(대상 및 대응), 그리고 토픽 3(알페스)와 토픽4 (대상 및 대응) 간의 거리는 상대적으로 거리가 멀어 연관성과 유사성이 낮았다. 토픽 2(학교폭력)는 다른 토픽들에 비해 중앙에 위치하여 토픽1(성폭력), 토픽 4(대상 및 대응)와도 비교적 가까운 거리를 나타내었다.



<그림 1> Intertopic Distance Map(IDM)

3. 워드클라우드 및 연관어 분석

먼저 인터넷 뉴스와 블로그를 통합하여 전체적으로 상위 키워드 20개를 <그림 2>와 같이 워드 클라우드 형태로 시각화하여 제시하였다. 청소년이 가장 빈도가 높은 주요 키워드로 나타나 사이버불링이 청소년에게 두드러지는 문제임을 확인할 수 있었다. 또한, 사이버불링이 기존의 학교폭력이 사이버상으로 옮겨온 것인 만큼 학교폭력도 주요 키워드로 나타났다. 사이버불링은 단순히 사이버따돌림 수준을 넘어 “범죄”의 형태로 나타나고 있으며 이에 대한 “처벌”이 강조되고 있음을 유추해볼 수 있다. 그리고 최근 N번방 사건으로 더욱

화제가 된 “성범죄”와 “여성”에 대한 “착취”도 주요 이슈로써 상위 키워드로 워드 클라우드에 포함되었다.



〈그림 2〉 상위 키워드 20개 워드 클라우드_인터넷 뉴스 및 블로그 통합

보다 구체적으로 핵심 키워드를 파악하고자, 토픽모델링 분석결과를 활용했다. 군집 및 토픽 1(성적폭력), 2(학교폭력), 3(알페스)은 사이버불링의 유형과 관련된 것으로 볼 수 있고 군집 및 토픽 4(대상 및 대응)는 사이버불링의 대상과 대응과 관련된 것으로 볼 수 있다. 따라서 본 연구에서는 사이버불링의 (1)유형, (2)대상, (3)대응으로 나누어 관련 상위 키워드를 워드 클라우드로 나타냈다. 또한, 워드 클라우드를 인터넷 뉴스와 블로그 각각 제시하여 매체 간 차이를 한눈에 파악할 수 있도록 시각화하였다. 그리고 주요 키워드에 대해 연관어 분석결과를 제시했다.

먼저, 유형 관련 인터넷 뉴스와 블로그 통합, 인터넷 뉴스, 블로그 각각 매체별로 제시하여 상위 키워드를 비교할 수 있도록 다음 〈그림 3〉과 같이 시각화하였다. 인터넷 뉴스와 블로그를 통합하여 사이버불링 유형과 관련하여 가장 핵심 키워드는 “학교폭력”이었다. 유형 관련 상위 키워드를 매체별로 비교하여보면, 인터넷 뉴스는 음란물, 딥페이크, 포르노, 지인능욕과 같은 성폭력 유형을 집중적으로 보도했음을 추측해볼 수 있으며 이는 N번방 사건의 영향으로도 해석할 수 있다. 반면 블로그는 학교폭력 유형에 더욱 주목하며 따돌림, 왕따와 같은 키워드들이 함께 나타난 것을 확인할 수 있다.

통합	인터넷 뉴스	블로그
유포 범죄 성범죄 학교폭력 촬영 폭력 음란물 제작 딥페이크 사이버폭력 착취 성폭력 합성 협박	인터넷 뉴스 음란물 포르노 도박 언어폭력 디스 영상물 지인능욕 강간 사이버폭력 알페스 스토킹 딥페이크	블로그 학교폭력 유포 성폭력 합성 음란물 성범죄 사이버폭력 폭력 영상물 착취 촬영 지인능욕 사이버불링 따돌림 딥페이크 왕따 폭행 신상 협박 알페스

<그림 3> 유형 워드 클라우드_매체별 비교

다음으로, 대상 관련 인터넷 뉴스와 블로그 통합, 인터넷 뉴스, 블로그 매체별 상위 키워드는 <그림 4>와 같이 제시하여 비교할 수 있도록 하였다. 인터넷 뉴스와 블로그를 통합하여 가장 핵심 키워드는 청소년이었다. 대상 관련 상위 키워드를 매체별로 비교하여보면, 인터넷 뉴스는 조주빈, 강훈, 갓갓, 공범 등과 같은 N번방 사건의 대상에 집중되어 있는 것을 확인할 수 있다. 반면 블로그는 학생, 청소년, 피해자, 집단, 친구와 같은 학교폭력 폭로사건에 주목하고 있음을 알 수 있다.

통합	인터넷 뉴스	블로그
조주빈 가해자 피해자 미성년자 남성 박사 여성 청소년 친구 학생 아이 아동 집단	인터넷 뉴스 조주빈 공범 소년범 해린 강훈 수진 서신애 현진 갓갓 이다영 윤장현 금호 송명근 일진 육범소년 김근 홍국생명	블로그 미성년자 친구 아이 전달래 집단 여성 피해자 청소년 조주빈 학생 가해자 아동 가수

<그림 4> 대상 워드 클라우드_매체별 비교

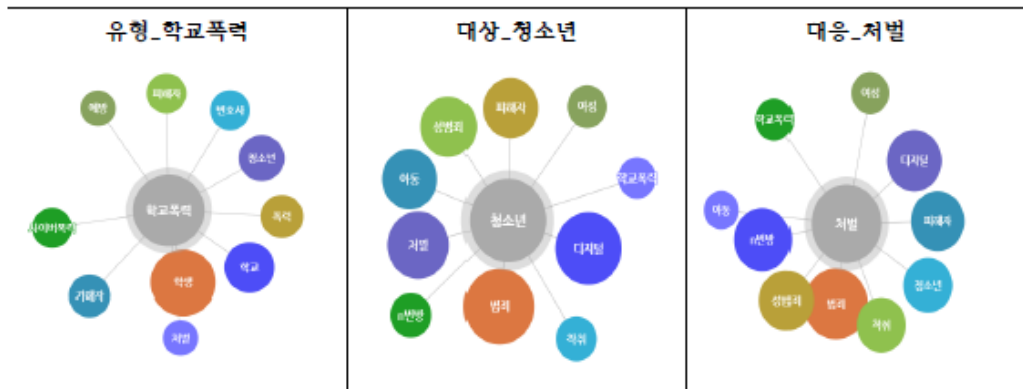
마지막으로 대응 관련 인터넷 뉴스와 블로그 통합, 인터넷 뉴스, 블로그 매체별 상위 키워드에 대한 워드 클라우드를 <그림 5>와 같이 제시하여 비교할 수 있도록 하였다. 인터넷 뉴스와 블로그를 통합하여 사이버불링 대응과 관련하여 가장 핵심 키워드는 “처벌”이었다. 대응 관련 상위 키워드를 매체별로 비교하여보면, 인터넷 뉴스와 블로그 모두 유사하게 검거, 처벌과 같은 법적 대응에 보다 중점을 두고 있는 것으로 나타났다. 두 매체 모두 서비스,

예방 교육 같은 대응에는 소홀한 것으로 볼 수 있다.



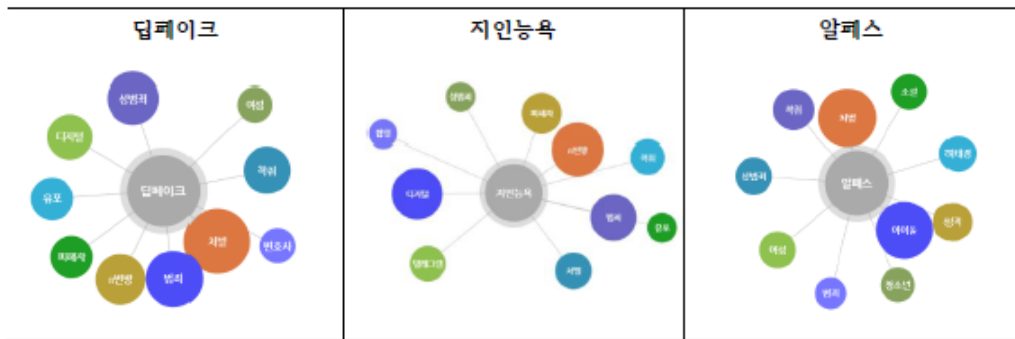
〈그림 5〉 대응 워드 클라우드_매체별 비교

유형, 대상, 대응의 각 범주에 대한 최상위 키워드에 대한 연관어 분석을 실시하였다. 즉, 유형 관련 키워드 중 최상위 키워드인 “학교폭력”, 대상 관련 키워드 중 최상위 키워드인 “청소년”, 대응 관련 키워드 중 최상위 키워드인 “처벌”에 대한 연관어 분석을 실시하여 〈그림 6〉에 제시하였다. 유형인 “학교폭력” 연관어 분석결과, 피해자, 예방, 변호사, 청소년, 폭력, 학교, 학생, 처벌, 가해자, 사이버폭력 키워드와 연관이 높은 것으로 나타났다. 대상인 “청소년” 연관어 분석결과, 피해자, 성범죄, 아동 차별, n번방, 범죄, 착취, 디지털, 학교폭력, 여성, 피해자 키워드와 연관이 높은 것으로 나타났다. 대응인 “처벌” 연관어 분석결과, 학교폭력, 여성, 디지털, 피해자, 청소년, 착취, 범죄, 성범죄, n번방, 아동, 학교폭력 키워드와 연관이 높은 것으로 나타났다.



〈그림 6〉 범주별 최상위 키워드 연관어 분석

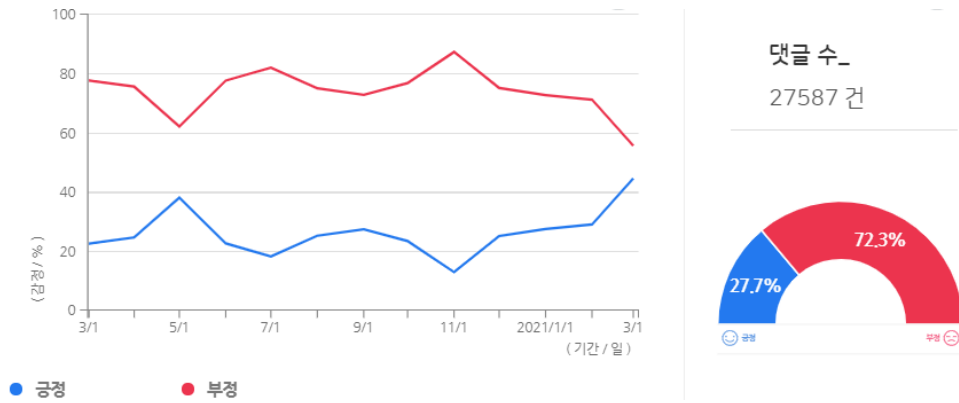
추가적으로, 최근들어 나타나고 있는 사이버불링의 새로운 유형인 “지인능욕”, “딥페이크”, “알페스”에 대한 깊이 있는 이해를 위하여 해당 단어를 중심으로 연관어 분석을 실시하여 <그림 7>에 나타냈다. “딥페이크”는 성범죄, 디지털, 유포, 피해자, n번방, 범죄, 처벌, 변호사, 착취, 여성과 연관성이 높은 것으로 나타났으며 “지인능욕”은 피해자, 성범죄, 합성, 디지털, 텔레그램, 처벌, 범죄, 유포, 착취, n번방과 연관성이 높은 것으로 나타났다. “알페스”는 차별, 착취, 성범죄, 여성, 범죄, 청소년, 아이들, 성적, 소설과 연관성이 높은 것으로 나타났다.



<그림 7> 최신 유형 사이버불링 연관어 분석

4. 감성어 분석

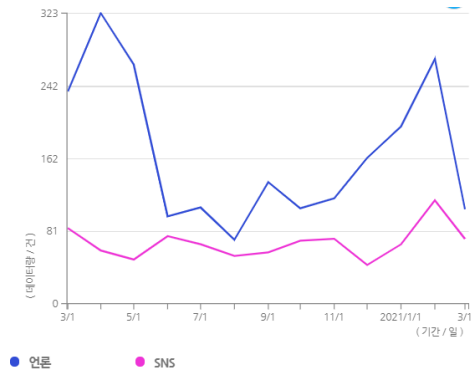
사이버불링 이슈 관련 인터넷 뉴스 댓글 총 27,587건의 감성어 분석 결과, 72.3%가 부정의 감정으로 나타났고 27.7%가 긍정의 감정으로 나타났다. 즉, 사이버불링은 사회적으로도 부정적 감정을 불러일으킨다고 볼 수 있다. 또한, 기간에 따른 감정 추이 변화를 살펴보았을 때 2020년 11월에 가장 부정적인 감정이 높은 것으로 나타났다. 반면 2021년 3월에는 부정적인 감정이 비교적 낮은 것을 확인할 수 있었다. 감성어 분석결과는 <그림 8>과 같다.



〈그림 8〉 사이버불링 이슈 댓글 감정 분석

5. 기간별 시계열 분석

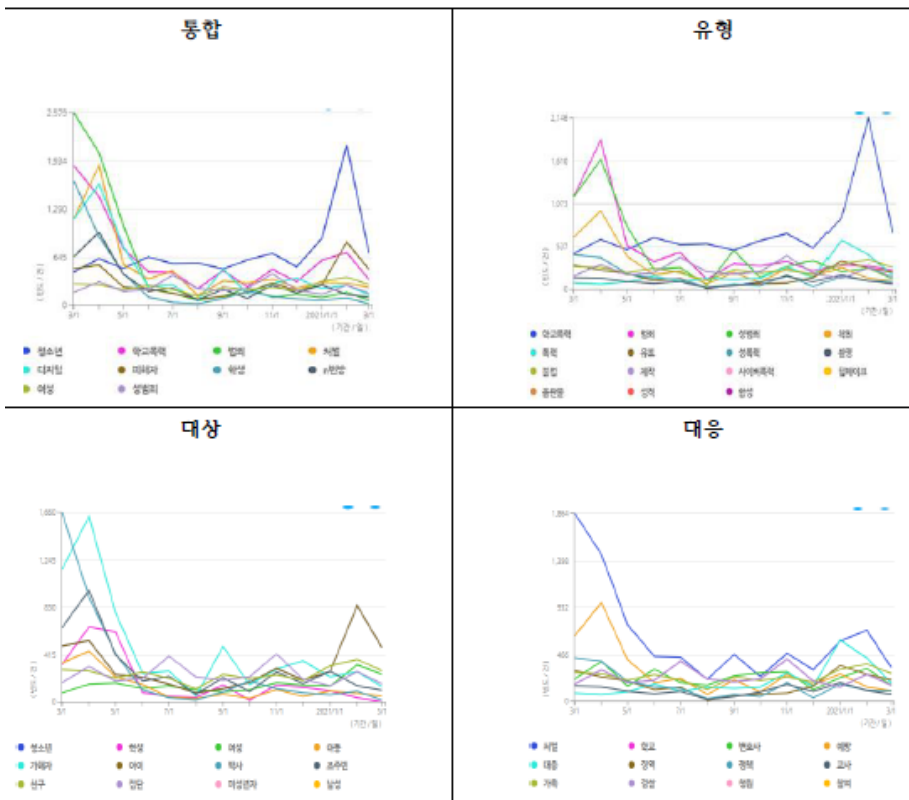
기간별 사이버불링 이슈 시계열 분석 결과는 〈그림 9〉에서 볼 수 있듯이, 언론의 경우 2020년 4월 323건으로 가장 건수가 많았고 2021년 2월 272건으로 두 번째로 건수가 많았다. 블로그의 경우 2021년 2월 115건으로 가장 높은 건수를 나타냈다.



〈그림 9〉 기간별 사이버불링 이슈 데이터량/건

모든 키워드를 통합하여 상위 키워드 10개의 기간별 빈도 변화를 분석한 결과, 〈그림 10〉의 통합에 나타난 것처럼 2020년 3월과 4월에는 키워드 범주의 빈도가 가장 높았고 처벌이 다음으로 높은 것으로 나타났다. 2021년 1월과 2월에는 청소년 키워드의 빈도가 가장 높았고 학교폭력과 피해자가 다음으로 높았다.

인터넷 뉴스와 블로그를 통합하여 사이버불링 유형과 관련된 키워드의 기간별 빈도/건 변화를 분석한 결과 <그림 10> 유형에 나타난 것처럼 2020년 3월부터 5월까지 범죄, 성범죄, 착취 순으로 키워드 빈도가 높았으며, 2021년 1월부터 3월까지 학교폭력, 폭력이 두드러지게 키워드 빈도가 높았다. 사이버불링 대상과 관련된 키워드의 기간별 빈도/건 변화를 분석한 결과 <그림 10> 대상에 나타난 것처럼 2020년 3월부터 5월까지 박사, 가해자, 조주빈 순으로 키워드 빈도가 높았으며 2021년 1월부터 3월까지는 아이, 친구가 높은 키워드 빈도를 나타냈다. 사이버불링 대응과 관련된 키워드의 기간별 빈도/건 변화를 분석한 결과 <그림 10> 대응에 나타난 것처럼 2020년 3월부터 5월까지 처벌, 예방, 정책 키워드 빈도가 높았으며 2021년 1월부터 3월까지 처벌, 대응, 가족 키워드가 높은 빈도를 나타냈다.



<그림 10> 기간별 사이버불링 관련 키워드 데이터 빈도/건

IV. 결론 및 함의

사이버불링의 피해는 국내·외 모두 증가하고 유형 또한 다양하고 급속하게 진화하는 추세이며 청소년 정신건강에 미치는 영향과 심각성은 매우 높다. 사이버공간에서 이뤄지는 사이버불링 피해는 시간과 장소를 불문하고 하루 24시간 내내 진행될 수 있어 피해자의 정신적 고통은 매우 심각할 수 있고 극단적으로는 자살 생각 및 행동으로도 이어지는 것이다. 사이버불링에 대한 사회적 관심이 높아지며 관련 연구들도 증가하고 있지만, 국내의 경우 사이버불링이 실제로 일어나고 있는 주 매체인 소셜 네트워크를 분석하여 사이버불링의 실태를 파악한 연구는 상대적으로 부족하였다. 따라서 본 연구에서는 소셜 미디어 빅데이터 분석 및 텍스트 마이닝을 통해 청소년 사이버불링 키워드를 분석 및 시각화하여 관련 이슈를 파악하고 연구 및 실천현장에서의 함의점을 제시하고자 했다는 점에서 의의가 있다.

분석결과, 인터넷 뉴스와 블로그를 통합하여 사이버불링 상위 5위 키워드는 청소년(n=10,827), 학교폭력(n=9,387), 범죄(n=9,006), 처벌(n=8,334), 디지털(n=8,294)이었다. 토픽모델링 결과, 성폭력, 학교폭력, 알페스, 대상 및 대응으로 총 4개의 군집 및 토픽이 나타났다. 이를 반영하여 본 연구에서는 사이버불링 유형, 대상, 대응으로 핵심 키워드를 구분하여 워드 클라우드 시각화 및 연관어, 감성어, 시계열 분석을 진행하였다.

먼저, 최근에는 사이버불링 유형 중에서도 시각적 및 성폭력이 심각한 사회적 문제로 나타났다으며, 이는 지인능욕, 딥페이크, 알페스 등과 같이 새로운 유형의 사이버불링으로 진화하고 있음을 확인할 수 있었다. 따라서 본 연구는 새롭게 제기되는 사이버불링 유형을 현실성 및 시의성 있게 파악하는 것의 중요성을 나타냈다. 지인 능욕은 일반인을 대상으로 영상을 조작하고 신상정보와 함께 유포하여 성착취물을 제작하는 범죄로 최근 새로운 디지털 성범죄 유형으로 추가되었다(김정호, 안재주, 양보성, 정주연, 우사이먼성일, 2020; 안갑철, 2021). 딥페이크는 인공지능 기술 및 딥러닝 기술을 이용해 악의적인 의도를 가지고 기존 인물의 얼굴 혹은 특정 부위를 영화의 컴퓨터 그래픽처럼 합성 및 조작된 음성, 영상, 이미지 편집물 등을 만들어 내는 방법을 말한다(김정호 외, 2020; 안갑철, 2021). 2018년 하반기 가해자들이 텔레그램 n번방과 박사방을 개설하고 운영하며 미성년자를 포함한 일반 여성을 대상으로 성착취 영상을 찍도록 협박하고, 해당 영상을 텔레그램 비밀대화방에서 판매한 잔인한 행각을 저지른 소위 ‘N번방’ 사건이 대표적이다(안갑철, 2021). 알페스(RPS)는 실존 인물 특히, 유명인을 대상으로 당사자의 동의 없이 창작물, 성적인 창작물을 만드는 것이다. 대표적인 예시로는 동성애 팬픽이 있다. 지금까지 진행된 다수의 사이버불링 연구들은 ‘성적/시각적 폭력’에 대한 내용을 내포하고 있었으나(정여주, 김한별, 전아영, 2016; Álvarez-García

et al., 2017; Lee et al., 2017), 조작된 사진이나 영상, 성적 창작물 등을 만드는 행위와 같이 구체적인 유형을 포착하진 못했었다. 이러한 측면에서 본 연구를 통해 발견된 지인능욕, 딥페이크, 알페스 등과 같은 새로운 사이버불링의 유형은 사이버불링을 연구하는 학자뿐만 아니라 사이버불링의 피해자에게 어떻게 개입할지를 고민해야 하는 실천가에게도 유용한 자료가 될 수 있으리라 생각한다.

본 연구는 사이버불링 유형 관련 용어가 혼용되고 있다는 것을 확인하였다. 예를 들어, 사이버불링 뿐만 아니라 사이버폭력, 디지털 범죄 및 범행 등 다양한 용어가 관련 키워드로 추출되었다. 사이버불링에 대한 사회적 관심과 더불어 학계의 관심이 커지고 있는 가운데 ‘사이버불링’이라는 용어가 쓰이기 시작한 이후 현재까지 사이버불링 외에도 사이버 따돌림, 사이버 일탈, 사이버폭력, 사이버 왕따 등 합의된 개념이 존재하지 않고 여러 용어가 혼용되고 있는 것이다(한희정, 정혜진, 2014). 미국에서는 사이버불링을 일반적으로 “이메일이나 채팅룸, 소셜 네트워킹 웹사이트를 이용하여 친구들을 괴롭히는 행위”(Patchin & Hinduja, 2011) 또는 “이메일, 핸드폰, 태블릿, 메시지, 인터넷 등의 전자 매체를 통해서 이루어지는 불링 행위”(Li, 2008)라고 정의한다. 그러나 한국에서는 사이버불링과 사이버폭력이라는 개념을 거의 구분하지 않으며 사이버불링이라는 개념도 다소 모호하여 사이버왕따, 사이버폭력, 사이버따돌림, 사이버괴롭힘, 온라인 폭력 등의 다양한 용어들로 혼용되어 사용되고 있다(이승현, 강지현, 이원상, 2015; 이창호 외, 2014). 형사정책연구원(2017)과 방송통신위원회·한국인터넷진흥원(2016)에 따르면, 사이버폭력과 사이버불링은 사이버 공간 안에서 타인을 공격하거나 피해를 줄 수 있다는 점에서 유사할 수 있으나, 사이버폭력은 사이버불링보다 더 광범위한 폭력의 유형을 다루고 있기에 구분이 되는 개념이라고 설명한다. 특히, 온라인 공간은 익명성을 전제로 하기때문에 사건 발생 후에도 오랜 기간 가해자, 피해자, 목격자가 제대로 파악조차 되지 않을 때가 많아 개념 정의 차원의 문제는 더욱 복잡해진다(정병삼, 2010). 따라서 향후 사이버불링 학문의 발전을 위해서라도 관련 용어들이 일관성 있게 통일될 필요가 있다고 본다.

학교폭력은 가장 핵심적인 사이버불링 유형으로 나타났다. 이는 사이버불링이 전통적인 학교폭력에서 사이버상으로 확대된 개념이라는 점을 고려하면 자연스러운 일이다. 즉, 사이버불링과 학교폭력은 매우 밀접하게 연계되어있다. 특히, 온택트 시대에는 대면 환경에 초점이 맞춰진 기존의 학교폭력에 대한 접근을 넘어서 비대면 환경에 적합한 학교폭력 예방이 이루어져야 할 것이며 사이버불링에 대한 고려가 함께 이루어져야 할 것이다. 예를 들어, 사이버불링 피해를 신고할 수 있는 앱이나 프로토콜을 개발하고, 관련 법안을 보다 학교와 긴밀하게 연계되도록 개정하여 사이버불링이 단순한 해프닝으로 끝나는 것이 아니라

적절한 징계와 처벌이 따를 수 있도록 해야 할 것이다.

사이버불링 대상과 관련하여서는 “청소년”, “미성년자”, “학생”, “자녀”, “아동”, “중학생”, “아이” 등이 빈도가 높은 주요 키워드로 나타났다. 전 세계 29개국에서 8-12세 어린이 38,000명을 대상으로 실시한 조사에서는 약 85%의 어린이들이 일부 형태의 소셜미디어를 사용하고 있으며, 절반 이상의 어린이들이 본인의 스마트폰을 사용하여 인터넷에 접속한다고 보고했다(DQ Institute, 2018). 한국의 경우, 청소년의 스마트폰 과의존 비율이 35.8%로 성인의 22.2%에 비해 약 1.6배 이상 높은 것으로 나타났다(과학기술정보통신부, 한국지능정보사회진흥원, 2020). 또한, 10대 전체 중 인터넷 이용자는 99%이며, 스마트폰 이용자는 97.6%로 청소년 대부분이 인터넷과 스마트폰을 이용하고 있는 것으로 확인된다(한국인터넷진흥원, 2014). 인터넷과 스마트폰의 일상화로 인해 청소년들의 사이버 공간에서의 활동 시간 및 영역이 계속해서 증가하고 있는 가운데 소셜미디어는 청소년들이 일상에서 보이는 생각이나 감정 등을 표현하고 이해할 수 있는 공간이다(이지은, 전종설, 2021). 이는 청소년들이 실시간으로 응답하고 소통 가능한 채널로 활용할 수 있다는 긍정적인 면을 보이기도 하지만, 인터넷 중독이나 사이버불링과 같은 부정적인 면도 동반하는 것으로 나타난다(이지은, 전종설, 2021).

한편, 이지은과 전종설(2021)은 청소년 대상이 아닌 신상이 파악되지 않는 온라인 콘텐츠 제공자를 대상으로 행해지는 사이버불링 역시 존재하므로, 사이버불링을 학교폭력 양상 중 하나로 보기보다는 더욱 확대되고 포괄적인 개념의 사이버불링 정의가 필요하다고 제안하였다. 본 연구와 비슷하게, 해당 연구에서도 “성인” 등이 주요한 키워드로 나타났다. 따라서 향후에는 청소년뿐만 아니라 다양한 대상 관련 사이버불링 현황을 구체적으로 파악하고 사이버불링 유형, 대상과 관련하여 명확한 개념 정의가 이루어질 필요가 있다.

본 연구에서 사이버불링 대응은 처벌 및 법적 대응 차원을 더 강조하는 것으로 나타났다. 이처럼 대응 키워드로 학교, 처벌, 검거, 변호사, 처벌, 징역과 같은 개념이 결과로 나온 것에 반해 현재 한국의 대응 방법은 오히려 예방적 차원에 더 중심을 두는 것으로 나타나 다소 상반된 양상으로 보여주고 있다. 한국에서는 「학교폭력 예방 및 대책 2020년도 시행계획」에 따라, 교육부, 방송통신위원회, 경찰청, 법무부, 여성가족부 등의 부처에서 사이버불링 예방 및 대처를 위한 다양한 프로그램을 통합적으로 진행하고 있다(관계부처 합동, 2020). 예를 들어, 사이버 어울림과 같이 학생, 교원, 학부모 대상 사이버불링 예방 및 대응 프로그램이 지원되고 있으며, 사이버불링을 조기에 감지할 수 있도록 사이버안심존, 실태조사, 스마트폰 사이버폭력 의심문자 스마트 알림 소프트웨어 등의 보급을 확산시키고 있다(관계부처 합동, 2020). 그러나 본 연구에서 강조된 결과가 법적 및 제도적 측면이 더 강한 만큼

사이버불링에 대한 처벌을 강화하거나 법적인 틀을 갖추는데 노력이 필요할 것으로 보인다. 특히, 학교가 사이버불링을 방어하는데 가장 효과적인 장소이고, 학교 밖에서 일어나는 사이버불링 역시 학교의 관할로 지정해야 한다는 의견이 대두되고 있으므로 학교와 연계된 법적 보호체계 마련과 처벌 규정이 강화되어야 할 필요가 있다(Chun et al., 2021; Erb, 2008; Lane, 2011). 예를 들어, 미국은 2013년을 기점으로 대다수의 주가 사이버불링 처벌 규정을 갖게 되었으며 학교와 긴밀한 연계를 갖추고 있다(조운오, 2015). 캐나다 정부 역시 사이버불링에 대한 강경 대응과 국가적 차원의 무관용 원칙을 강조하며 이에 대응하기 위한 법을 적극적으로 진행하여 공식적 처벌 근거를 마련하였다(조운오, 2015). 한국 역시 「학교폭력 예방 및 대책에 관한 법률(2004)」에 의거하여, ‘사이버 따돌림’을 학교폭력 안에 포괄하고, 본 법령하에 피해학생의 보호, 가해학생의 선도·교육 및 분쟁조정에 대한 내용을 설명하고 있다(국가법령정보센터, 2021). 비록 본 법률에서 사이버불링에 대한 정의(제2조)와 가해학생에 대한 조치(제17조)를 명시하고 있지만, 사이버불링의 유형이 빠르게 진화하고 있음에 따라 어디까지를 사이버불링으로 봐야하는지에 대한 문제가 있으며, 문제 해결에 있어 학교에게 어느 정도의 권한을 부여할지에 대한 고민 역시 필요하다.

앞서 한국은 제도 및 법적 규제보다는 예방적 차원의 접근방법이 더 중심이 되고 있다고 설명하였다. 그러나 예방적 차원의 개입방안에서도 사이버불링 피해로 인한 정신건강 서비스 개입은 비교적 부족한 것으로 평가할 수 있다. 본 연구에서 사이버불링 이슈 관련 인터넷 뉴스 댓글 총 27,587건의 감성어 분석결과, 72.3%가 부정의 감정으로 나타났고 27.7%가 긍정의 감정으로 나타났다. 즉, 사이버불링은 피해자 및 가해자 당사자들에게 정신적 피해를 가져올 뿐 아니라 사회적으로도 부정적 감정을 불러일으키는 것으로 확인된다. 이는 Lee와 Chun(2020)의 연구 결과와도 맥을 같이하는데 연구에 따르면 청소년 사이버불링 피해는 지인을 만나기 부끄러움, 사회적 공포, 다른 사람을 만나기 무서움, 밖에 나가고 싶지 않음, 세상을 부정적으로 바라봄, 관계에 회의적임, 복수에 대한 생각을 함, 모욕감을 느낌, 자존감이 낮아짐, 빼놓아짐, 소외감을 느낌, 무서움을 느낌, 우울감을 느낌, 성적 모욕감을 느낌, 트라우마를 느낌, 죄책감을 느낌, 자살생각을 함, 긴장감을 느낌 등의 정서적 피해를 불러일으킨다. 따라서, 사이버불링과 정신건강은 함께 연계되는 문제이기 때문에 이를 함께 통합적으로 고려하여 예방 및 개입 서비스를 제공하는 것이 필요하다.

기간별 시계열 분석 결과, 언론은 2020년 4월 건수가 가장 높았고 블로그는 2021년 2월 가장 높은 건수를 보였다. 이는 2020년 ‘N번방 사건’과 2021년 배구선수 관련 ‘학교폭력 폭로사건’의 영향으로 유추된다. 같은 맥락으로, 2020년 ‘N번방 사건’의 영향을 보여주듯, 2020년 3월부터 5월에는 키워드 범죄와 처벌의 빈도가 가장 높았고 “유형” 관련하여서는

범죄, 성범죄, 착취가, “대상” 과 관련하여서는 박사, 가해자, “대응” 에서는 처벌, 예방, 정책이 빈도가 높은 것으로 나타났다. 또한 ‘학교폭력 폭로사건’ 의 영향으로 2021년 1월부터 3월에는 키워드 청소년, 학교폭력, 피해자가 상위 키워드로 나타났고 “유형” 관련하여서는 학교폭력, 폭력이, “대상” 관련하여서는 아이, 친구가, “대응” 관련하여서는 처벌, 대응, 가족 키워드가 높은 빈도를 보였다. 이를 종합하였을 때 인터넷 뉴스와 블로그 모두 특정한 사건에만 집중하는 경향을 보인다는 것을 확인할 수 있었다. 그러나 특히 인터넷 뉴스의 경우, 특정 사건만 집중보도하는 것보다는 다양한 사건과 주제를 다룰 필요가 있을 것이다.

우리나라 교육부(2019)는 사이버폭력에 대처하는 7가지 자세에 대한 지침을 제공하고 있다. 그러나 청소년 개인적 대처만으로는 대응이 어렵기때문에 부모, 학교 등의 체계적 노력이 함께 요구된다. 특히 교육부에서는 청소년뿐만 아니라 학부모에게도 사이버폭력에 대해 규칙적으로 대화 나누기, 사이버폭력 신고 방법과 도움 요청 기관 알아두기, 자녀의 온라인 활동 관심을 갖고 지켜보기, 음란물, 폭력물 필터링 소프트웨어 설치, 그리고 자녀들에게 어른, 관계기관과 상의할 것을 강조하여 부모가 함께 참여하는 것의 중요성을 드러냈다. 부모 뿐 아니라 학교와의 연계도 매우 중요하다. 외국의 경우 다양한 사이버불링 예방 및 개입 프로그램들을 제공하는데, “총체적 학교 접근(Whole School Approach)” 이 대표적인 프로그램 개발 원칙이라고 할 수 있다. 호주의 “The Best of Coping” 프로그램은 크게 학교-중심의 전체적 개입전략을 중점적으로 사용하여 사이버불링 “대처” 기술을 강조한 프로그램을 개발 및 보급했다(Tangen & Campbel, 2010). 사이버불링을 예방하는데 긍정적인 효과를 거두었다고 평가되는 핀란드의 KiVa 프로그램은 총체적 학교 접근을 통해 안전하고 행복한 교육 환경을 만드는 것을 목표로 구성되었다(Kiva 사이트, 2020). 또한, 독일의 “Media Heroes” 프로그램 역시 총체적 학교 접근방식을 토대로 정규 교사가 정규 수업에서 사이버불링 예방 교육을 할 때 활용하기 위해 만들어진 프로그램이다(조운오, 2015). 따라서 청소년 사이버불링 대응을 위해서는 개인, 학교, 부모 등의 다양한 체계가 통합적으로 노력하여야 하며 법적 제도와 통합적 서비스를 함께 구축하는 노력이 요구된다.

본 연구의 한계점은 다음과 같다. 먼저, 데이터 자료가 연구 시작일 기준 최신 1년만을 수집하였기에 최신 트렌드를 반영할 수는 있을지언정 사이버불링에 대한 전반적인 트렌드를 파악하기는 어려울 것으로 사료된다. 따라서 향후에는 더 긴 기간의 데이터를 수집하여 트렌드의 변화를 살펴볼 수 있을 것이다. 또한, 본 연구는 인터넷 뉴스와 블로그에서만 자료를 수집했기 때문에 모든 소셜 네트워크 서비스의 데이터를 수집한 것은 아니기 때문에 모든 소셜네트워크를 대표하지 못할 수 있다는 한계가 있다. 향후 연구에서는 더 다양한 소셜네트워크 플랫폼에 대한 구체적인 비교분석이 요구된다. 그리고 소셜 네트워크의 특성상

사회적으로 이슈가 되는 사건들을 중점적으로 다루는 경향이 있을 수 있기 때문에 본 연구에 중점적으로 드러나지 않은 사이버불링의 주요한 비가시적 요소들이 더 있을 수 있다.

그럼에도 불구하고 본 연구는 국내의 기존 연구들과 달리 소셜 미디어 빅데이터 분석 및 텍스트 마이닝을 통해 현실성 및 시의성을 확보하였으며, 이를 통해 청소년 사이버불링 유형, 대상, 대응과 관련한 기간별, 매체별 키워드와 키워드 간 관계를 분석 및 시각화하여 사이버불링 관련 최신 이슈 및 여론을 파악하였다는 점에서 의의가 있다.

■ 참고 문헌

- 강승지, 손유진 (2020). 토픽 모델링 분석을 통한 유아 사교육 현상 연구: 국내 신문 기사와 블로그를 중심으로. *미래유아교육학회*, 27(1), 177-199.
- 강지연, 김수정, 노승국 (2019). 간호사의 직장 내 괴롭힘 관련 온라인 뉴스기사 댓글에 대한 토픽 모델링 분석. *기본간호학회지*, 49(6), 736-747.
- 과학기술정보통신부, 한국지능정보사회진흥원. (2020). 2020 스마트폰 과의존 실태조사. 세종: 과학기술정보통신부, 대구: 한국지능정보사회진흥원.
- 관계부처 합동 (2020). 학교폭력 예방 및 대책 2020년도 시행계획. 교육부, 방송통신위원회, 경찰청, 법무부, 여성가족부.
- 교육부 (2019). 여러분은 사이버폭력으로부터 안녕하신가요?. 세종: 교육부.
- 교육부 (2021). 2020년 학교폭력 실태조사 결과보고서. 세종: 교육부.
- 국가법령정보센터 (2021). 학교폭력예방 및 대책에 관한 법률. Retrieved from <https://law.go.kr/lsc.do?section=&menuId=1&subMenuId=15&tabMenuId=81&eventGubun=060101&query=%EC%82%AC%EC%9D%B4%EB%B2%84%ED%8F%AD%EB%A0%A5#undefined>
- 기재홍, 안승혁 (2020). 감성분석과 토픽모델링을 활용한 농촌태양광 관련 이슈 연구: 언론 기사와 블로그 포스트 비교. *디지털융복합연구*, 18(9), 17-27.
- 김아미 (2021). 청소년 사이버불링에 대응하기 위한 미디어 리터러시 교육 접근법: 해외의 논의 및 사례를 중심으로. *한국소통학회 세미나자료*, 29-42.
- 김정호, 안재주, 양보성, 정주연, 우사이머성일 (2020). 데이터 기반 딥페이크 탐지기법에 관한 최신 기술 동향 조사. *정보보호학회지*, 30(5), 79-92.
- 반현 (2003). 인터넷 뉴스 미디어의 신뢰도에 관한 실험 연구. *한국방송학보*, 17(2), 207-231.
- 방송통신위원회, 한국인터넷진흥원 (2016). 2015년 사이버폭력 실태조사 결과보고서. 방송통신위원회, 한국인터넷진흥원.
- 방송통신위원회, 한국지능정보사회진흥원 (2020). 2020년 사이버폭력 실태조사. 대구: 한국지능정보사회진흥원.
- 서선경, 정은경 (2013). 동시출현단어 분석 기반 오픈 액세스 분야 지적구조에 관한 연구. *한국비블리아학회지*, 24(1), 207-228.
- 송태민, 송주영, 진달래 (2014). 소셜 빅데이터를 활용한 인터넷 중독 위험예측 모형. *보건사회연구*, 34(3), 106-134.

- 신명선, 조경원 (2019). 텍스트마이닝을 활용한 한국어어치료학회지의 토픽 모델링 및 트렌드 분석: (2002~2018). 언어치료연구, 28(3), 81-91.
- 신소라 (2016). 폭력물시청과 폭력목격이 청소년의 사이버불링과 학교폭력에 미치는 영향: 자기통제력의 매개효과. 교정복지연구, 44, 143-162.
- 안갑철 (2021). 불법 성적 촬영물 구매행위 처벌규정에 관한 위헌성 검토. 인권과 정의, 495, 85-102.
- 안화실, 박정도 (2013). 사이버폭력에 관한 국내 연구물 분석, 한국소년정책학회, 23, 125-157.
- 안형준, 하영목 (2019). 경험의 유형과 블로그 글의 특성 간의 관계 분석. 한국정보기술학회논문지, 15(2), 131-140.
- 이동근 (2019). 일간신문과 블로그의 미디어 간 의제설정: 최순실 게이트 사건 보도 비교분석. 정치정보연구, 22(2), 53-90.
- 이성식, 전신현 (2011). 집단괴롭힘의 피해와 가해의 관계와 그 성별 차이. 피해자학연구, 19(1), 207-227.
- 이수상 (2012). 네트워크 분석 방법론, 부산대학교 사회과학연구원 연구총서. 서울: 논형.
- 이승우, 이정덕 (2016). 청소년의 사이버불링의 영향 요인 검증. 한국치안행정학회, 13(2), 41-66.
- 이승현, 강지현, 이원상 (2015). 청소년 사이버폭력의 유형분석 및 대응방안 연구. 형사정책연구원 연구총서, 1-481.
- 이인태 (2012). 초등학교의 사이버 불링 실태와 원인에 대한 조사연구. 청소년문화포럼, 32(1), 91-118.
- 이주형, 안순철 (2013). 학교폭력 예방을 위한 사이버불링(Cyberbullying)에 대한 연구: 미국의 법적 대응의 특징과 법제 도입을 위한 검토를 중심으로. 분쟁해결연구, 11(1), 133-157.
- 이지은, 전종설 (2021). 청소년 사이버불링 온톨로지 개발 연구. 청소년학연구, 28(4), 163-192.
- 이창호, 신나민, 하은빈 (2014). 청소년 사이버불링 실태 및 대응방안 연구. 한국청소년정책연구원 연구보고서, 1-268.
- 전종설, 이해국 (2020). 온라인 수업 전환 후 청소년들의 미디어 사용 실태. 코로나 19-온택트 시대, 청소년 미디어 중독 해법 포럼.
- 정병삼 (2010). 부모지도감독이 청소년의 비행 친구 접촉과 일탈행동에 미치는 중단적 억제효과, 청소년상담연구, 18(2), 149-165
- 정여주, 김한별, 전아영 (2016). 청소년 사이버폭력 피해 척도개발 및 타당화. 열린교육연구, 24(3), 95-116.

- 정여주, 김동일 (2012). 청소년의 사이버폭력 피해 경험과 정서조절. *상담학연구*, 13(2), 645-663.
- 조운오 (2013). 사이버불링 피해가 청소년의 비행에 미치는 영향-우울감의 매개효과. *한국 청소년학연구*, 20(10), 117-142.
- 조운오 (2015). 외국의 청소년 사이버불링 대응방안 및 개입 프로그램 고찰. *경찰학논총*, 10(3), 337-369.
- 조운오 (2017). 사이버불링 유형화 연구: 상담가 의견을 중심으로. *한국경찰학회보*, 19(5), 279-305.
- 조윤주, 정옥분 (2009). 청소년의 동조적 또래괴롭힘에 대한 매개된 조절모형 검증. 또래동조성의 작용을 중심으로. *인간발달연구*, 16(3), 91-111.
- 최진오 (2015). 사이버불링 피해경험이 사이버불링 가해행동에 미치는 영향: 불안의 매개효과와 스트레스 대처전략의 조절효과. *한국위기관리 논집*, 11(5), 195-214.
- 한국인터넷진흥원 (2014). 국내 SNS의 이용 현황과 주요 이슈 분석. *한국인터넷진흥원 INTER NET & SECURITY FOCUS*.
- 한국정보화진흥원 (2012). 소셜 분석으로 살펴본 청소년 자살예방정책의 시사점. *IT & Future Strategy*, 17.
- 한국정보화진흥원 (2019). 2019년 사이버폭력 실태조사. 대구: 한국정보화진흥원.
- 한국형사정책연구원 (2017). 청소년 사이버폭력의 유형분석 및 대응방안 연구. 서울: 한국형사정책연구원
- 한희정, 정혜진 (2014). 국내 사이버불링 연구 동향: 개념과 연구 경향성을 중심으로. *한국소통학보*, 24, 180-217.
- Mitchell, R. (2018). *Web scraping with Python: Collecting more data from the modern web*. Sebastopol, CA: O'Reilly Media.
- Álvarez-García, D., Núñez, J. C., Barreiro-Collazo, A., & García, T. (2017). Validation of the cybervictimization questionnaire (CYVIC) for adolescents. *Computers in Human Behavior*, 70, 270-281.
- Bilic, V. (2013). Violence among peers in the real and virtual world. *Pediatrics Today*, 9(1), 78-90.
- Bonanno, R. A., & Hymel, S. (2013). Cyber bullying and internalizing difficulties: Above and beyond the impact of traditional forms of bullying. *Journal of Youth and Adolescence*, 42(5), 685-697.
- Cénat, J. M., Blais, M., Hébert, M., Lavoie, F., & Guerrier, M. (2015). Correlates of bullying in Quebec high school students: The vulnerability of sexual-minority youth. *Journal of*

- Affective Disorders, 183, 315-321.
- Cetin, B., Yaman, E., & Peker, A. (2011). Cyber victim and bullying scale: A study of validity and reliability. *Computers & Education*, 57(4), 2261-2271.
- Chapin, J. (2016). Adolescents and cyber bullying: The precaution adoption process model. *Education and Information Technologies*, 21(4), 719-728.
- Chun, J., Lee, J., Kim, J., & Lee, S. (2020). An international systematic review of cyberbullying measurements. *Computers in Human Behavior*, 113, 106485.
- Chun, J., Lee, S., & Kim, J. (2021). Conceptualizing the protective factors of cyberbullying victimization in Korean adolescents. *School Mental Health*, 13(3), 473-486.
- DQ Institute (2018). The 2018 DQ impact report. Retrieved from https://www.dqinstitute.org/2018dq_impact_report/
- DQ Institute. (2020). 2020 Child online safety index. Retrieved from <https://www.dqinstitute.org/wp-content/uploads/2020/02/2020-COSI-Findingsand-Methodology-Report.pdf>
- Erb, T. D. (2008). A case for strengthening school district jurisdiction to punish off-campus incidents of cyberbullying. Retrieved August 15, 2021, from <https://heino.nline.org/HOL/LandingsPage?handle=hein.journals/arzj140&div=12>
- Hemphill, S. A., Tollit, M., Kotevski, A., & Heerde, J. A. (2015). Predictors of traditional and cyber-bullying victimization: A longitudinal study of Australian secondary school students. *Journal of Interpersonal Violence*, 30(15), 2567-2590.
- Hinduja, S., & Patchin, J. W. (2007). Offline consequences of online victimization: School violence and delinquency. *Journal of School Violence*, 6(3), 89-112.
- Hinduja, S., & Patchin, J. W. (2008). Cyberbullying: An exploratory analysis of factors related to offending and victimization. *Deviant Behavior*, 29(2), 129-156.
- Hinduja, S., & Patchin, J. W. (2010). Bullying, cyberbullying, and suicide. *Archives of Suicide Research*, 14(3), 206-221.
- Hinduja, S., & Patchin, J. W. (2013). Social influences on cyberbullying behaviors among middle and high school students. *Journal of Youth and Adolescence*, 42(5), 711-722.
- Jones, L. M., Mitchell, K. J., & Finkelhor, D. (2012). Trends in youth internet victimization: Findings from three youth internet safety surveys 2000-2010. *Journal of Adolescent Health*, 50(2), 179-186.

- Kim, H., Han, Y., Song, J., & Song, T. M. (2019). Application of social big data to identify trends of school bullying forms in South Korea. *International Journal of Environmental Research and Public Health*, 16, 2596.
- Kiva 사이트. (2020). Kiva program. Retrieved from <http://www.kivaprogram.net/program>
- Lane, D. K. (2011). Taking the lead on cyberbullying: Why schools can and should protect students online. *Iowa Law Review*, 96(5), 1791.
- Lee, J., Abell, N., & Holmes, J. L. (2015). Validation of measures of cyberbullying perpetration and victimization in emerging adulthood. *Research on Social Work Practice*, 27(4), 456-467.
- Lee, S., & Chun, J. (2020). Conceptualizing the impacts of cyberbullying victimization among Korean male adolescents. *Children and Youth Services Review*, 117, 105275.
- Li, Q. (2007). New bottle but old wine: A research of cyberbullying in schools. *Computers in Human Behavior*, 23(4), 1777-1791.
- Li, Q. (2010). Cyberbullying in high schools: A study of students' behaviors and beliefs about this new phenomenon. *Journal of Aggression, Maltreatment & Trauma*, 19(4), 372-392.
- Miner, G., Elder IV, J., Fast, A., Hill, T., Nisbet, R., & Delen, D. (2012). *Practical text mining and statistical analysis for non-structured text data applications*. Academic Press.
- Mishna, F., Cook, C., Gadalla, T., Daciuk, J., & Solomon, S. (2010). Cyber bullying behaviors among middle and high school students. *American Journal of Orthopsychiatry*, 80(3), 362-374.
- Patchin, J. W., & Hinduja, S. (2011). Traditional and nontraditional bullying among youth: A test of general strain theory. *Youth & Society*, 43(2), 727-751.
- Raymond, J. (2019). Fakebusters strike back: How to spot deep fakes, the manipulated videos that are the newest form of “fake news” to hit the internet. *Index on Censorship*, 48(1), 76-79.
- Sarkar, D. (2019). *Text analytics with Python: A practitioner's guide to natural language processing*. Berkeley, CA: Apress.
- Slonje, R., & Smith, P. K. (2008). Cyberbullying: Another main type of bullying?. *Scandinavian Journal of Psychology*, 49(2), 147-154.

- Tahamtan, I., & Huang, L. M. (2019). Cyberbullying in the digital age: Exploring people's opinions with text mining. *Proceedings of the Association for Information Science and Technology*, 56(1), 672-673.
- Tangen, D., & Campbel, M. (2010). Cyberbullying prevention: one primary school's approach. *Australian Journal of Guidance and Counseling*, 20(2), 225-234.
- Tokunaga, R. S. (2010). Following you home from school: A critical review and synthesis of research on cyberbullying victimization. *Computers in Human Behavior*, 26(3), 277-287.
- Tsolmon, B., & Lee, G. S. (2017). Topic model reflecting user behavior and time analysis for extracting disaster events from social data. *Information and Communications Magazine*, 34(6), 43-50.
- UNICEF. (2020). Children at increased risk of harm online during global COVID-19 pandemic. Retrieved from <https://www.unicef.org/press-releases/children-increased-risk-harm-online-during-global-covid-19-pandemic>
- Wang, Y. (2017). Resistance and compromise under power structures of sexuality: A case study on real person slash fans in China. [Master's dissertation, LUND University]. Google Scholar.
- Ybarra, M. L., Mitchell, K. J., Wolak, J., & Finkelhor, D. (2006). Examining characteristics and associated distress related to Internet harassment: Findings from the second youth internet safety survey. *Pediatrics*, 118(4), 1169-1177.

Keyword Analysis of Adolescent Cyberbullying through Text Mining

Lee, Serim*4) · Kim, Jinyung**5) · Chun, JongSerl***6)

Abstract

This study aimed to identify the major and rapid changing cyberbullying issues among adolescents through text mining. Online news, blog posts, and comments were selected with keywords as to ‘cyberbullying’ and ‘cyberviolence.’ We retrieved in total of 3,075 documents and 190,706 keywords. Using these keywords, frequency analysis, topic modelling, visualization of word cloud, related words analysis, sentiment analysis, time series analysis were conducted. In the keyword analysis, school violence, crime, punishment etc. were found as the major keywords. There were in total of four topics: sexual violence, school violence, real person slash, and subject and response. By cyberbullying types, ‘visual and sexual violence’ turned out as the most severe form of cyberbullying, and its form were changing rapidly including humiliation, deep-fake, and real person slash. Furthermore, words describing the types of cyberbullying were inconsistent, and ‘adolescent’ was found as the major target of cyberbullying. To counteract against cyberbullying, legal standards and policies were reported as important. Based on these findings, authors suggest for the provision of cyberbullying prevention programs and the reinforcement of mental health services related to cyberbullying.

Key Words: cyberbullying, text mining, topic modelling, association analysis, sentiment analysis, time series analysis

2021. 05. 29 투고, 2021. 08. 30 재심신청,
2021. 09. 15 심사완료, 2021. 09. 27 게재확정

4) Ewha Womans University, Department of Social Welfare, PhD candidate

5) Ewha Womans University, Department of Social Welfare, PhD candidate

6) Ewha Womans University, Department of Social Welfare, Professor, Corresponding author.