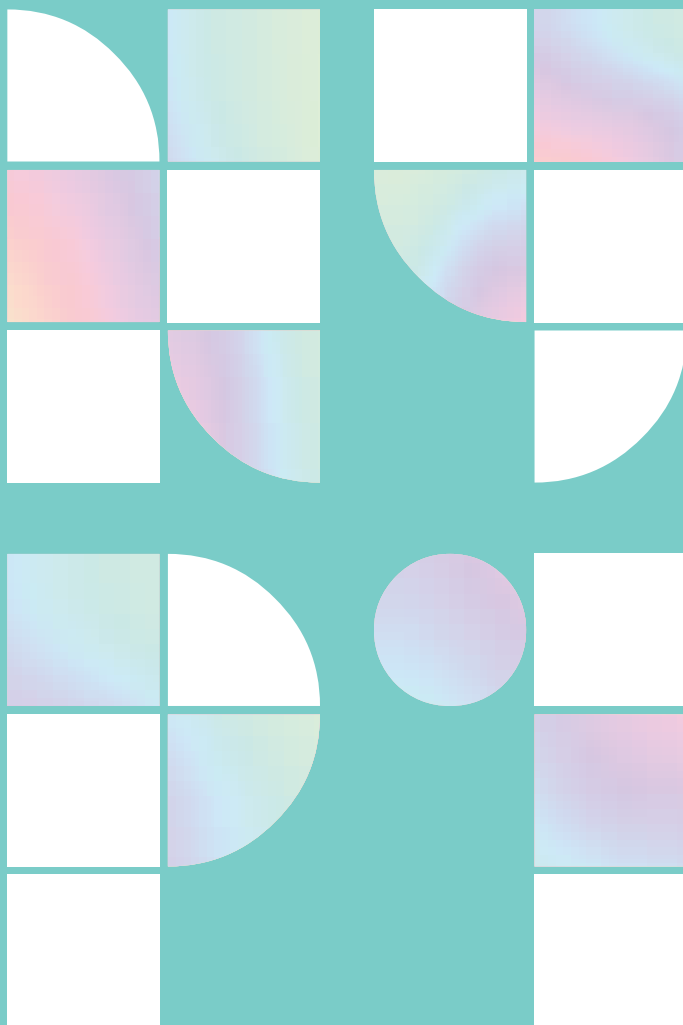


비정형 민원 데이터를 활용한 청소년 정책의제 도출

홍순구 · 임지원



연구보고 21-R12-1

비정형 민원 데이터를 활용한 청소년 정책의제 도출

저 자 홍순구, 임지원

연구진 연구책임자_홍순구(동아대학교 교수)
공동연구원_임지원(동아대학교 전임연구원)



한국청소년정책연구원
National Youth Policy Institute

보고서를 펴내며

연구를 끝마칠 즈음엔 포스트코로나 시대가 도래 할 것이라고 예측했던 것이 무색하게 우리는 지금 이 시점에도 코로나시대를 지나가고 있습니다. 이 재난이 언제 끝날지 우리를 어디로 데려갈지는 장담하기 어렵습니다. 여기저기에서 들리는 미래사회에 대한 다양한 예측을 이해하고 4차 산업 혁명, 뉴노멀 시대 등 다양한 변화에 적응해 나가기에도 벅찬데 우리는 코로나19 팬데믹이라는 재난을 맞아 복합적인 변화에 적응해야 하는 숙제를 받아들고 있습니다. 빠르게 앞으로 질주하는 미래라는 자동차 바퀴에 코로나19라는 변수가 어떠한 작용을 할지는 쉽게 알기 어렵습니다. 이러한 상황이 어른들에게도 혼란스러운데 청소년들에게는 어떠할까요? 마스크를 쓰는 삶이 당연하게 여겨지는 시대에 코로나19로 인해 화면 넘어 선생님과 친구들을 오랫동안 만나던 이들에게 학교는, 가족은, 친구는 어떤 의미일까요? 코로나19가 끝난 이후 이들의 삶을 위해 우리는 그리고 우리의 청소년정책은 어떻게 해야 할까요?

이 연구는 데이터의 의미와 가능성에 주목합니다. 비정형 민원 빅데이터, 현장 전문가의 목소리, 청소년의 생각과 욕구, 기존의 다양한 공공 빅데이터는 중요한 분석 대상입니다. 스마트거버넌스라는 새로운 정책 의제 도출 방법과 외국, 국제기구 사례에도 관심을 기울입니다. 자료 검토와 데이터 분석을 통해 현재와 미래에 대한 제안을 도출하고자 노력하였습니다. 이 과정에서 함께 해주신 모든 연구진, 집필진, 자문진들께 감사를 드립니다. 또한 연구결과가 포스트코로나 시대 청소년정책을 고민하는데 조금이나마 도움이 될 수 있기를 기대합니다.

2021년 12월
연구진 모두를 대표하여
연구책임자 서정아

연구요약

1. 연구목적

- 본 연구에서는 스마트 거버넌스 개념을 바탕으로 온라인에서 생성되고 있는 방대한 양의 비정형데이터를 분석하여 포스트코로나 시대 청소년과 관련된 정책의제를 도출하는 것을 목적으로 한다.
- 정책 이슈 도출을 위하여 토픽모델링이라는 대표적인 빅데이터 분석방법을 사용하여 기존 청소년정책 연구와의 차별성을 가지고 결과의 신뢰성과 객관성을 확보하고자 하였다.

2. 연구 방법

- 정부에서 운영하고 있는 국민신문고와 국민청원 사이트 내에 있는 총 9개 채널에서 2002년 1월부터 2021년 7월까지 ‘아이’, ‘아동’, ‘청소년’, ‘청년’, ‘학생’ 키워드를 내용에 포함하는 게시글 및 민원 총 265,744개의 데이터를 수집하였다.
- 수집된 데이터는 청소년과 관련한 여성가족부, 교육부 및 교육청, 지자체 및 기타 중앙행정기관 데이터로 재추출한 다음 코로나 이전과 이후로 구분하여 대표적인 비정형데이터 분석방법인 토픽모델링(Latent Dirichlet Allocation: LDA)을 적용하여 분석하였다.
- 토픽모델링 결과를 통하여 코로나 전후의 주요 중앙행정기관에서 청소년과 관련된 전반적인 토픽들을 파악하고 정책 이슈를 살펴보았다. 그리고 코로나 전후의 청소년관련 토픽들의 비중변화와 네트워크 분석을 통하여 청소년 정책 이슈의 흐름과 특징을 알아보았다.

3. 주요결과

1) LDA 분석결과

- 여성가족부의 경우 코로나 이전에는 8개의 토픽이 나타났으나, 코로나 이후에는 6개의 토픽들로 분류되었다. 코로나 이후 청소년관련 정책 이슈로는 청소년 및 가족대상 상담서비스 기관 및 지원 확대, 아동관련 기관의 성범죄·아동학대 정보 및 경력조회 시행, 맞벌이 가정에 대한 육아 지원서비스 확대, 성매매·자살 예방 의무교육 실시, 청소년대상 술·담배 판매자에 대한 처벌강화 및 청소년증 의무발급, 한부모가정 및 저소득층에 대한 양육비 지원 확대 등이 도출되었다.
- 교육부의 경우에는 코로나 이전에는 10개의 토픽으로 나타났으나, 코로나 이후에는 6개의 토픽들로 다소 감소하였다. 코로나 이후 등교연기 및 온라인 수업 문제, 아동폭력 예방 교육 및 보건교원 확대, 유해시설 인근 교육환경 개선 및 보상 요청, 코로나 방역 및 안전문제, 학교신설 및 교육투자 확대 요청, 대학수학능력 및 교육과정 개선 요청 등이 청소년관련 정책이슈로 도출되었다.
- 교육청은 코로나 이전에는 9개의 토픽이 나타났으나, 코로나 이후에는 학교 폭력문제 대응 및 피해대책, 돌봄 교실 운영 문제, 교사업무 및 수업평가 필요성, 과밀학급 해소를 위한 학교 신설 요청, 통학 안전 확보를 위한 학교 신설 요청, 유해시설 인근 학교 이전 및 신설 요청 등과 같은 7개의 이슈가 나타났다. 대부분이 학교 신설과 관련된 내용으로 청소년정책 이슈와는 관련성이 적었다.
- 지자체 및 기타 중앙행정기관 데이터를 LDA 분석한 결과, 코로나 이전에는 12개의 토픽이었으나 코로나 이후에는 8개의 토픽으로 다소 줄어들었다. 청소년과 관련된 주요 정책 이슈로는 코로나 방역 및 확진자 치료, 청년 및 청소년 교육사업 지원 및 지역 운영 확대, 어린이 안전사고 대비, 개학연기 및 보육문제, 아동학대 처벌 강화 및 피해자 보호, 대학입시제도 등이 도출되었다. 이와 함께 부동산 및 경제관련 이슈 등이 추가로 등장하였다.

2) 코로나 전후 토픽 비중 변화와 네트워크 변화

- 코로나 이후 지속적으로 높은 비중을 차지하던 대학입시제도, 저출산·결혼문제 등의 이슈는 감소한 반면 보육 서비스, 아동학대 및 폭력 문제, 성범죄 문제, 상담 서비스, 저소득층 양육비 지원 등의 이슈는 증가한 것을 확인할 수 있었다. 이와 함께 코로나로 인한 방역 및 확진자 치료, 온라인 수업 등의 이슈는 새롭게 등장하였다.
- 온라인 수업과 관련하여 저소득층의 컴퓨터 부족 문제, 수업 일수 부족 및 졸업 문제, 동영상 프로그램의 수준 및 강사 자질 문제 등의 이슈가 나타났는데, 특히 저학년의 경우 온라인 수업을 지양하는 것으로 확인되었다. 이는 상대적으로 보육이 중요한 저학년의 경우 등교를 하지 않을 시 생기는 돌봄 공백에 대한 학부모의 우려가 원인으로 판단된다. 코로나 이후 긴급보육 및 돌봄 교실 관련 토픽이 크게 증가한 것이 이를 뒷받침한다.
- 또한 코로나 이후 맞벌이 및 한부모 가정의 돌봄공백에 대한 우려와 보육서비스 요구 증가 이슈는 여성가족부, 교육부, 교육청, 기타 기관에서 모두 등장하는 것으로 나타나 그 중요성을 알 수 있다. 특히 저소득층에 대한 보육서비스는 경제상황 및 교육문제와 직결되어 있어, 다양한 형태의 한부모 가정에 대한 이해를 바탕으로 지원이 이루어져야 할 것이다.

4. 정책제언

- 중앙행정기관을 기준으로 분류하였지만, 기관별로 정책 이슈는 명확히 구분되지 않고 중복되어 나타나는 것을 확인할 수 있다. 이에 각 기관에서 두드러지게 나타난 청소년 정책의제를 파악하는 것이 중요할 것이다.

- 주요 정책이슈로 도출된 보육 서비스와 양육비 지원 확대, 가족·학생 상담 서비스 확대 요구, 학교폭력 및 아동학대 문제 등은 모두 코로나로 인해 개학이 연기되고 맞벌이 및 한부모 가정의 육아 부담이 증가함에 따라 파생된 이슈들로 판단된다. 이는 과거 일반적으로 국민의 소득수준과 의식 수준이 높아짐에 따라 사회서비스의 요구가 증가하게 되는 것과는 달리 코로나19로 인한 급격한 환경 변화에 기인한 것으로 정책적 접근도 달리해야 할 것이다.
- 정부 공식 온라인 플랫폼에서의 시민들의 게시글 및 민원만을 대상으로 빅데이터 분석을 실시하였으나 향후에는 SNS, 신문 기사 등 다양한 정보원으로부터 데이터를 수집할 수 있는 채널을 확보하고 전처리를 위한 사전 과정 마련 등 전반적인 빅데이터 관리 체계를 구축하여 지속적이고 체계적인 청소년 의제 도출을 위한 시도가 이루어져야 할 것이다.

비정형 민원
데이터를 활용한
청소년
정책의제 도출

연구보고 21-R12-1

I. 서론 1

1. 연구의 배경 3

2. 연구의 목적 4

II. 연구방법 5

1. 스마트 거버넌스: Co-creation과 지능화 기술의
 접목 7

2. 비정형데이터를 통한 청소년 정책의제 도출 20

3. 분석 대상 및 데이터 수집 23

III. 연구결과 29

1. 기초 분석 결과 31

2. 주요 중앙행정기관별 LDA 분석 결과 36

3. 코로나 전후 청소년 정책 이슈 변화 71

IV. 결론 및 정책 제언	79
1. 주요 연구 결과 및 의의	81
2. 향후 연구 과제	83
 참고문헌	 85
 부 록	 91
 국문초록	 105
 ABSTRACT	 107

표 목차

표 II - 1	데이터 수집 채널과 수집 내용	24
표 III - 1	수집 데이터 수	31
표 III - 2	채널 및 키워드 중복제거	33
표 III - 3	분석 데이터 현황	33
표 III - 4	코로나 이전 여성가족부 LDA 분석 결과	37
표 III - 5	코로나 이후 여성가족부 LDA 분석 결과	40
표 III - 6	여성가족부 코로나 이후 토픽 원문 예시	43
표 III - 7	코로나 이전 교육부 LDA 분석 결과	47
표 III - 8	코로나 이후 교육부 LDA 분석 결과	50
표 III - 9	교육부 코로나 이후 토픽 원문 예시	52
표 III -10	코로나 이전 교육청 LDA 분석 결과	56
표 III -11	코로나 이후 교육청 LDA 분석 결과	59
표 III -12	코로나 이후 교육청 주요 토픽 원문	62
표 III -13	코로나 이전 지자체 및 기타 중앙행정기관 LDA 분석 결과	64
표 III -14	코로나 이후 지자체 및 기타 중앙행정기관 LDA 분석 결과	68

그림 목차

그림 II- 1	스마트 거버넌스로의 진화	11
그림 II- 2	스마트 거버넌스를 적용한 혁신된 공공서비스 의사결정지원 프로세스	13
그림 II- 3	Text Preprocessing Step	15
그림 II- 4	워드 클라우드의 예시	18
그림 II- 5	pyLDAvis를 활용한 LDA 모형의 시각화 결과	19
그림 II- 6	네트워크 분석 시각화 예시	20
그림 II- 7	청소년 정책의제 도출을 위한 연구절차 및 방법	22
그림 II- 8	토픽 수 선정을 위한 Silhouette	28
그림 III- 1	검색키워드 시간별 추이변화	32
그림 III- 2	코로나 이전 전체데이터 LDA 결과 및 Top-30 단어	34
그림 III- 3	코로나 이후 전체데이터 LDA 결과 및 Top-30 단어	35
그림 III- 4	코로나 이전 여성가족부 토픽별 비중	38
그림 III- 5	코로나 이전 여성가족부 topic map	39
그림 III- 6	코로나 이후 여성가족부 토픽별 비중	41
그림 III- 7	코로나 이후 여성가족부 topic map	42
그림 III- 8	코로나 이전 교육부 토픽별 비중	48
그림 III- 9	코로나 이전 교육부 topic map	49
그림 III-10	코로나 이후 교육부 토픽별 비중	51
그림 III-11	코로나 이전 교육청 토픽별 비중	57
그림 III-12	코로나 이전 교육청 topic map	58
그림 III-13	코로나 이후 교육청 토픽별 비중	60

그림 Ⅲ-14	코로나 이후 교육청 topic map	61
그림 Ⅲ-15	코로나 이전 지자체 및 기타 중앙행정기관 토픽별 비중	65
그림 Ⅲ-16	코로나 이전 지자체 및 기타 중앙행정기관 topic map	66
그림 Ⅲ-17	코로나 이후 지자체 및 기타 중앙행정기관 토픽별 비중	69
그림 Ⅲ-18	코로나 이후 지자체 및 기타 중앙행정기관 topic map	70
그림 Ⅲ-19	여성가족부 코로나 전후 토픽의 비중 변화	72
그림 Ⅲ-20	여성가족부 코로나 이후 토픽 6 네트워크 분석 결과 ..	73
그림 Ⅲ-21	교육부 코로나 전후 토픽의 비중 변화	74
그림 Ⅲ-22	코로나이후 Topic -1 네트워크 분석 결과	75
그림 Ⅲ-23	교육청 코로나 전후 토픽의 비중 변화	76
그림 Ⅲ-24	교육청 코로나 이후 토픽 3 네트워크 분석 결과	77
그림 Ⅲ-25	지자체 및 중앙행정기관 코로나 전후 토픽의 비중 변화	78

○———— 제1장 서 론

- 1. 연구의 배경
- 2. 연구의 목적

1. 연구의 배경¹⁾

현재 우리는 인공지능과 빅데이터로 대변되는 4차 산업혁명 시대에 살고 있다. 애플이나 구글 등 세계 초일류 기업들은 사물인터넷(Internet of Things), 클라우드, 자율주행 자동차 등 우리의 삶을 편리하게 해주는 지능화 기술을 활용하여 새로운 비즈니스 기회를 창출하고 있다. 예를 들어 모바일 주문을 통해 새벽 배송을 신청하거나, 무인 점포에서 물건을 구매하면 자동으로 계산되어 청구서를 받게 된다. 투자 분야에서는 트레이더가 로봇 어드바이저로, 의료분야에서는 인간 의사가 인공지능 의사(IBM Watsons)와 같이 진료를 보는 등 우리의 삶이 급격하게 디지털로 전환되고 있는데, 이러한 현상을 디지털 전환(Digital Transformation)이라고 한다. 즉 디지털 전환은 인공지능, 빅데이터 등 지능화된 디지털 기술을 활용하여 기존 전통적 운영방식과 서비스를 급격히 개선(radical improvement)하는 파괴적 혁신으로 정의될 수 있다.

지능화된 정보기술이 사회 전반에 확산되면서 디지털 전환 등을 통한 새로운 산업구조의 변화 외에도 정부 운영이나 거버넌스에도 큰 영향을 미치고 있다. 일례로, 홍순구, 김나랑, 이태현 외(2017)는 인공지능 등 지능화된 신기술을 적용하여 시민과 이해관계자의 능동적 참여(co-creation)를 보장하는 ‘스마트 거버넌스’를

1) ‘제1장 서론’은 동아대학교 홍순구 교수가 작성하였다.

주장하고 있다. 그들은 스마트 거버넌스를 빅데이터와 인공지능과 같은 신기술을 접목하여 정부의 정책 수립 및 실행 과정 전반에 걸쳐 다양한 이해관계자들의 참여를 보장하고(co-creation) 기존 거버넌스의 한계점을 극복하는 새로운 정부 혁신체제로 정의하고 있다.

2. 연구의 목적

코로나19 팬데믹 이후 청소년을 대상으로 이루어지던 교육, 활동, 복지 등 모든 부문에서 전혀 없는 변화가 나타났으며, 이러한 급격한 환경 변화로 청소년의 삶에도 큰 혼란이 일어나고 있다. 이에 청소년 정책 및 서비스체계의 기존 운영 방향을 능동적이고 적극적으로 변화시켜 포스트코로나 시대를 대비할 필요성이 커지고 있다.

따라서 본 연구에서는 스마트 거버넌스 개념을 바탕으로 온라인에서 생성되고 있는 방대한 양의 비정형 데이터를 분석하여 포스트코로나 시대 청소년과 관련된 정책적 이슈를 도출하고자 한다. 특히 코로나19로 인해 발생한 청소년과 관련된 새로운 정책적 이슈는 어떠한 것이 있는지 살펴보고, 코로나19 이후로 청소년 정책 이슈의 변화를 분석하여 청소년 정책의제에 시사점을 줄 것이다.

이러한 스마트 거버넌스를 기반으로 한 접근은 온라인에서 실시간 대량으로 생성되는 시민들의 의견을 정책에 수렴 가능하게 함으로써 기존에 진행된 청소년 정책연구의 패러다임 전환과 함께 정책수혜자들의 실질적인 의견을 반영한 청소년 정책 수립을 가능하게 할 것이다.



제2장 연구방법

- 1. 스마트 거버넌스:
Co-creation과 지능화
기술의 접목
- 2. 비정형데이터를 통한
청소년 정책의제 도출
- 3. 분석 대상 및 데이터 수집

1. 스마트 거버넌스: Co-creation과 지능화 기술의 접목²⁾

최근 시민들은 모바일 플랫폼 기기와 스마트폰을 활용해 손쉽게 공공서비스를 제안하고 평가할 수 있게 되었다. 이러한 첨단 기술은 시간과 비용 문제로 불가능했던 시민과 정부 간의 의사소통을 통해 협업체계를 가능하게 하고 있다. 본 장에서는 지능화 기술을 소개하고, 지능화된 기술과 co-creation이 접목된 스마트 거버넌스를 설명한다.

1) 지능화 기술의 발전

지능화 기술이란 빅데이터, 사물인터넷(IoT), 인공지능 등과 같은 디지털 기술들을 말한다. 기업과 정부는 이런 기술들을 활용하여 수요자 맞춤형 서비스를 발굴하고 기존 서비스의 전달 과정을 차별화 시킬 수 있다.

(1) 인공지능(Artificial Intelligence, AI)

제4차 산업혁명과 함께 인공지능 시대가 본격적으로 오고 있다. 인공지능은 민간영역뿐만 아니라 버스, 택시, 전철 등 공공서비스 영역에도 깊숙이 침투하고

2) '제2장 연구방법'은 동아대학교 홍순구 교수와 임지원 전임연구원이 공동 작성하였다.

있다. 실제로, 세계 각국의 정부는 디지털 정부, 모바일 정부, e-government 등 ‘지능 정부(intelligent government)’를 추구하고 있다. 미국의 보스톤시 경우는 보스톤 311 Call 및 전용 App을 활용하여 민원 정보를 수집하고 분석하여 도로 파손, 낙서, 쓰레기 처리, 주차 문제 등을 해결하고 있다. 또한 남부 네바다주 보건당국은 식중독 예방을 위한 식당 위생 검사를 무작위 추출(random) 방식에서 인공지능을 통한 대상선정 방식으로 전환했다. 인공지능은 정부 운영의 효율성과 최적화된 맞춤형 서비스를 제공할 수 있어 많은 나라에서 이의 활용을 확대하고 있다(글로벌 과학기술정책정보 서비스, <https://now.k2base.re.kr>).

(2) 빅데이터 분석(Big Data Analysis)

지금까지 빅데이터는 주로 기업들의 경쟁력 향상 수단으로 많이 논의되어 왔으나 최근에는 공공서비스 부문에서의 활용 사례가 증가하고 있다. 예를 들어 한국의 전주시는 시간대별 대중교통 승·하차와 승객 환승 정보(교통카드 데이터), 유동 인구 정보(통신 데이터) 등을 분석하여 보다 정확하고 합리적인 대중교통 노선 조정과 신규 노선 결정에 활용하였다. 미국의 경우 미국지질조사소(United States Geological Survey: USGS)가 1900년 이래로 100년간의 지구지진 기록을 분석하여 향후 지진 발생률 등을 예측하였다. 이러한 분석을 통해 정부는 지진과 같은 자연 재난 사고뿐만 아니라 다양한 유형별 재난을 예방하고 훈련할 수 있는 재난서비스 매뉴얼을 시민들에게 제공하고 있다. 영국의 런던에서는 고질적인 교통혼잡을 줄이기 위하여 교통 데이터를 활용하여 혼잡 통행료를 부가하고 있다. 정부의 혼잡통행료 정책은 대중교통의 증가, 서비스 개선, 시내의 대기환경을 개선시켰다.

(3) 텍스트 마이닝 (Text Mining)

의견·생각·경험·관점 등을 서로 공유하기 위해서 사용되는 소셜 미디어(Social Media)는 공공서비스 분야에서도 활용되고 있다(Kowalski et al., 2019). 텍스트 마이닝은 소셜 미디어를 통한 다양하고 방대한 개인 의견이 포함된 텍스트 데이터를 분석하는 기법으로, 텍스트로 이루어진 빅데이터를 분석하는 데 다음과 같은 많은 장점이 있다. 첫째, 비정형 데이터의 분류와 감성분석 등의 기술을 활용해 수많은 글 속에 숨어있는 유의미한 정보와 지식을 추출할 수 있다. 둘째, 빅데이터 속에서 특정한 패턴이나 추세(trend)를 찾아내고 필요한 정보를 제시하여 효율적인 의사결정이 가능하다. 셋째, 폭넓고 다양한 데이터 원천(source)이 사회문제에 대한 더 나은 통찰력을 제공할 수 있다. 넷째, 공공부문에서의 텍스트 분석은 정책을 수립하고 집행하는데 있어서 증거 기반의 정책 추진을 가능하게 한다.

2) 디지털 전환(Digital Transformation)

최근 4차 산업혁명에 대한 관심과 함께 디지털 전환이 기업의 화두가 되고 있다. 디지털 전환은 기업이나 공공기관에서 사물 인터넷, 클라우드 컴퓨팅, 인공지능, 빅데이터 등 지능화 기술을 활용하여 기존의 조직 운영 방식과 서비스를 혁신하는 것이다. 혁신은 일반적으로 속도에 따라 점진적 진화(progressive evolution)와 급진적 혁명(radical revolution)으로 구분되는데, 디지털 전환은 ‘매우 빠른 변화’라는 점을 강조한다. 이러한 디지털 전환은 ICT 기술을 이용하여 기업의 일부 환경을 변화시키는 것이 아니라 제품 개발에서 생산, 마케팅 등 모든 분야에 적용하는 패러다임의 변화를 의미한다. 구글, 마이크로소프트, 아마존, 알리바바 등과 같은 사물 인터넷, 클라우드 컴퓨팅, 인공지능, 빅데이터 등의 기술을 활용하여 플랫폼을 구축 운영함으로써 소비자들에게 맞춤형 서비스를 제공하고 있다.

이러한 변화는 정부의 정책 또는 공공서비스 분야에서도 감지되고 있다. 예를 들어 뉴질랜드의 경우 임신 및 출산, 육아 등의 라이프 사이클(life cycle)에 맞추어 필요한 정보를 제공하는 온라인 맞춤형 서비스를 제공하고 있다. 우루과이 정부의 경우 국민이 요청했던 사항을 정부가 빠르게 요청자에게 정보를 제공하고 알림을 주는 시스템을 운영 중이다(글로벌 과학기술정책정보 서비스, <https://now.k2base.re.kr>).

3) 스마트 거버넌스

국가 정책의 수립과 운영은 Top-down 방식의 거버먼트(Government)에서 정부와 민간이 공동으로 참여하는 거버넌스로 변화였다. 거버먼트가 정부 중심의 국정운영이라면 거버넌스(Governance)는 일방적인 정부 주도에서 벗어나 다양한 이해관계자가 공동의 관심사에 대한 네트워크를 구축하여 문제를 해결하는 국정 운영 방식을 말한다(Pierre & Peters, 2000). 과거의 거버먼트 시스템은 획일화된 사회에 알맞은 방식으로 사회구성원 대부분이 유사한 문제를 마주하게 되는 사회에서 효과적으로 작동한다. 반면에 거버넌스는 국가행정에 있어서 분권화와 민영화, 시장화, 세계화 등 빠르게 변화하는 시대 상황에서 전통적인 통치 방식이 적절하지 않다는 인식에서 시작되었다. 이는 정부가 국민을 동반자적 관계로 보고 지속 가능한 사회적 발전이라는 공통된 목적을 위해 참여와 상호협력의 중요성을 강조하고 있다.

최근까지 사회적·정치적 문제 해결은 “협치”를 기반으로 하는 거버넌스에 의존하여 왔다. 정부 주도의 정책실행에서 벗어나 민간과의 협력을 중요하게 생각하는 협력적 거버넌스는 정부, 주민, 이해당사자의 협력을 통해 정책을 생성하고 실행하는데 어느 정도 성과를 내어왔다(홍순구 외, 2017). 하지만 기존의 거버넌스 구조에서는 물리적·시간적 제약으로 인해 이해관계자의 직접적 참여보다는 대리인을 통한 정책 수립과 실행이 이루어질 수밖에 없었다. 개인의 욕구가 다양해지고

의사 표명이 용이해진 현대 사회에서는 이해관계자의 참여가 제한되는 기존의 거버넌스 체계에서 벗어나 폭넓고 다양한 이해관계자와의 협력과 참여가 필요하다. 때마침, SNS와 인공지능 등 지능화된 정보통신 기술의 발전으로 이해관계자가 손쉽게 참여할 수 있는 여건이 마련됨에 따라 기존의 거버넌스에 지능화된 기술이 접목된 스마트 거버넌스(smart governance)로의 진화가 가능하게 되었다.

본 연구에서 언급하는 스마트 거버넌스(smart governance)는 홍순구 외(2017)에 의해 주장된 것으로 co-creation의 개념에 다양한 정보통신기술과 빅데이터와 인공지능과 같은 지능화 기술을 접목하여 정책 수립 및 실행과정에 이해관계자들과 정책 수혜자들의 참여를 보장하고 기존의 거버넌스가 가지고 있던 한계점을 극복하기 위한 새로운 개념이다.(<그림 II-1> 참조).



그림 II-1 스마트 거버넌스로의 진화

* 출처: 홍순구 외(2017). 신 지역혁신모델: 스마트 거버넌스. 유원복스.

따라서 스마트 거버넌스 모델의 핵심 가치는 빅데이터와 인공지능 등 지능화 기술을 적용하여 시민들의 실질적인 의견을 분석하여 보다 객관적이고 효율적인 정책이나 수요자 맞춤형 공공서비스를 제공하는데 있다.

4) 스마트 거버넌스 프로세스

최근 정부가 추진하는 공공서비스에 대한 시민들의 관심이 커지고 있다. 정부는 다양한 계층의 국민들을 대상으로 공공서비스를 제공하고 있다. 그러나 정책이 제대로 홍보되지 않아 이용하지 못하거나 실제 이용자의 요구사항에 적합 하지 않는 경우가 많다(Salge & Vera, 2012). 따라서 공공서비스가 성공하기 위해서는 정책을 신설하고 정비하는 것뿐만 아니라, 실제 이용자들의 관심도와 요구사항, 평가를 귀담아듣고 개선점을 파악하는 것이 매우 중요하다.

일반적인 공공서비스 분석은 정량적인 데이터로 나타나는 경제효과 분석, 통계치의 변화, 설문조사 결과를 바탕으로 이루어져 왔다. 그러나 사회 전체의 더 많은 이익 실현을 위해서는 시민이 공유하는 보편적 가치를 찾는 과정과 집단지성(collective intelligence)이 표출되는 과정이 선행되어야 한다. 즉, 시민들의 의사소통을 통한 집단지성의 표출이 정책 과정의 중심적인 역할이 되어야 한다. 의사소통과 집단지성의 발현은 절대다수의 시민 개개인으로 하여금 정부에 대한 발언권을 강화시킨다. 또한 의사소통 및 수렴과정이 확대될수록 공공서비스 정책의 효율성과 위기 대응력이 강화된다.

하지만 의사소통과정은 전적으로 인간의 언어에 의존할 수밖에 없기에 다양한 이해관계자의 의견을 수렴하고 정책 결정의 근거를 제시할 수 있는 기술이 요구된다. 즉, 스마트 거버넌스 구현을 위해서는 매우 많은 양의 인간 언어인 비정형 데이터를 처리하는 능력과 집단지성의 발현 과정을 모방한 기계 알고리즘이 필요하다. 또한 수치 데이터 분석을 통해 얻어지는 결과와 같은 수준의 값을 제시하기 위해서는 텍스트가 가진 비정형 데이터를 정형화할 수 있는 기술이 필요하다. 이를 위한 연구인 텍스트 마이닝과 머신러닝이 급속히 발전하고 있다.

본 연구에서 주장하는 스마트 거버넌스의 개념을 적용한 공공서비스 분석 모델은 아래 <그림 II-2>와 같다. 먼저 시민의 의견을 SNS, Twitter, 게시판, 언론 또는

민원 등을 통해 수집한다. 수집된 정보는 이슈 분석 또는 사용자 의견 분석을 위해 다양한 방법론이 적용된다. 분석 결과는 이용자 및 공공기관이 쉽게 보고 이해할 수 있는 시각화를 통해 제공되고, 공공기관은 결과물을 바탕으로 시민 의견을 반영하고 맞춤형 서비스 제공을 위한 정보로 활용할 수 있다.

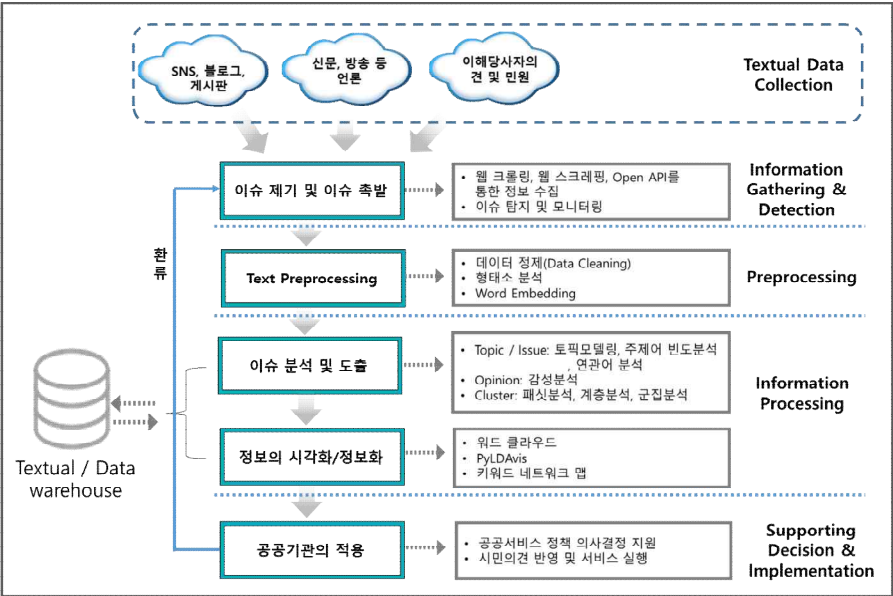


그림 II-2 스마트 거버넌스를 적용한 혁신된 공공서비스 의사결정지원 프로세스

(1) 시민 의견 데이터 수집

스마트 거버넌스를 적용한 공공서비스 의사결정 지원 시스템의 첫 번째 단계인 시민 의견 수집은 시간과 비용 절감을 위해 컴퓨터를 활용한 자동화된 방법을 사용하게 된다. 데이터의 수집 방법으로는 웹 크롤링(Web-crawling), 웹 스크래핑(Web-scraping), Open API(Application Programming Interface)를 주로 이용한다.

① 웹 크롤링(Web Crawling)

웹 크롤링은 웹 크롤러를 이용하여 인터넷에 공개된 정책 데이터를 자동으로 수집하는 방법이다. 웹 크롤러는 특정 사이트의 비정형 데이터를 수집할 때 사용된다.

② 웹 스크래핑(Web Scraping)

웹 스크래핑은 웹 크롤링과 유사한 의미로 사용되지만 웹 브라우저 화면 내에서 필요 정보만을 추출해내는 기술로 웹 크롤링이 모든 데이터를 수집하는 반면, 웹 스크래핑은 사용자가 필요한 정보만을 추출한다는 점에서 차이가 있다. 웹 크롤링과 웹 스크래핑은 오픈 소스(open source) 형태로 무료로 사용 가능하다.

③ Open API(Application Program Interface)

Open API란 데이터를 어디서나 쉽게 활용할 수 있도록 운영체제나 프로그래밍 언어가 제공하는 기능을 제어할 수 있도록 만든 인터페이스로, 운영체제와 응용 프로그램 간 통신에 사용된다.

(2) 텍스트 전처리

전처리 과정에서는 수집된 데이터의 형태소(Morpheme)를 분석하는 과정이 필요하다. 형태소 분석이란 품사, 어근, 접미사, 접두사 등 다양한 언어적 속성의 구조를 파악하는 것이다. 데이터 전처리 과정은 다음 <그림 II-3>과 같이 이루어진다. ① 데이터 정제: 결측값을 채우거나 이상치를 제거하는 과정, ② 데이터 통합: 다수의 정제된 데이터를 통합하는 작업, ③ 데이터 변환: 텍스트 마이닝의 효율을 높이기 위한 변환 작업, ④ 데이터 정리: 데이터 크기를 축소하지만 동일한 분석 결과를 만드는 작업

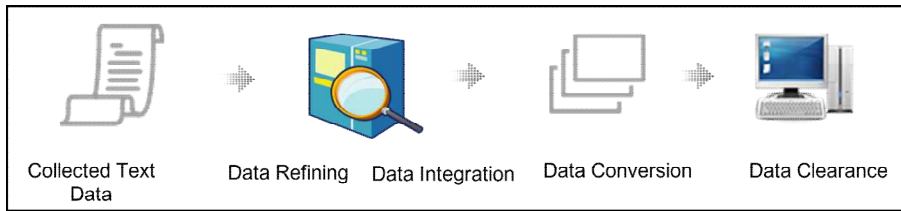


그림 II-3 Text Preprocessing Step

* 출처: Data Processing, <http://hanj.cs.illinois.edu/cs412/bk3/03.pdf>에서 2021년 10월 10일 인출

(3) 이슈 분석 및 도출

맞춤형 공공서비스 제공을 위하여 다양한 채널에 흩어져 있는 시민 의견을 수렴하여 주요 이슈를 도출하여야 한다. 이를 위하여 첫 번째 단계에서 수집된 공공서비스에 대한 시민 의견을 다음과 같은 다양한 텍스트 마이닝 기법으로 분석한다.

① 주제어 빈도분석(Keyword Frequency Analysis)

주제어 빈도분석(Keyword Frequency Analysis)은 특정 문서집단 내에서 빈번히 나타나는 주제어를 추출하고 이들이 언급되는 빈도에 따라 중요도를 찾아내는 방법이다. TF(단어 빈도, Term Frequency)는 특정한 단어가 문서 내에 얼마나 빈번히 나타나는지를 표시하는 값으로, 이 값이 높을수록 문서에서 차지하는 비중이 높다고 생각할 수 있다. 그러나 단어 자체가 문서집단 내에서 자주 사용되는 경우, 이것은 그 단어가 흔하게 등장한다는 것을 의미할 수도 있다. 예를 들어 영어에서 the는 빈번히 사용된다. 이것을 DF(문서 빈도, Document Frequency)라고 하며, 이 값의 역수를 IDF(역문서 빈도, Inverse Document Frequency)라고 한다. TF와 IDF를 곱하여 TF-IDF를 구한다.

② 연관어 분석(Association Keyword Analysis)

연관어 분석은 주제어 빈도 분석 결과를 바탕으로, 관심 주제어를 포함한 대상 문서에서 함께 언급된 주제어를 추출하여 국민들이 관심 주제어와 어떠한 주제를 연결시켜 생각하는지를 분석한다.

③ 토픽모델링(Topic Modeling)

토픽모델링은 텍스트 데이터에서 사용된 주제어들의 동시 사용 패턴을 바탕으로, 해당 텍스트들을 대표하는 특정 주제나 이슈, 주제 그룹들을 자동으로 추출하는 분석기법이다. 토픽모델링 기법에는 잠재의미 분석(Latent Semantic Analysis: LSA), 잠재 디리클레 할당(Latent Dirichlet Allocation: LDA)이 있다.

④ 네트워크 분석(Network Analysis)

네트워크 분석은 SNS 상에서 네트워크를 형성하고 있는 사용자 또는 문서 간의 상관관계를 분석하여 특정 이슈가 어떻게 나타나 타인에게 전파되는지를 보여주는 기법이다. 이 기법은 개인을 노드(node), 개인의 사회적 관계를 링크(link)로 하여 사회 네트워크를 구하고, 사회 네트워크의 연결 구조와 강도를 분석하여 어떤 메시지가 어떤 경로를 통해 전파되는지, 네트워크 상에서 입소문의 중심이나 허브(hub) 역할을 하는 influencer가 누구인지 등을 찾는 데 활용된다.

⑤ 분류분석(Classification Analysis)

텍스트 분류란 학습 데이터를 이용하여 분류모델을 만들고 텍스트를 입력하여, 어떤 종류의 범주(class)에 속하는지를 찾아내는 작업으로 주로 기계학습(Machine Learning) 기술을 사용한다. 이 기술에는 Naive Bayes, SVM (Support Vector Machine), Random Forest, Neural Network 등이 있다(Kim & Lee, 2018).

기계학습의 대표적인 방법인 딥러닝은 문서 분류 및 클러스터링, 문서 요약, 웹 마이닝 및 감성분석과 같은 다양한 텍스트 마이닝 문제에 대한 효과적인 해결책으로 부상하고 있다. 이에 딥러닝 기반의 CNN(Convolutional Neural Network)과 RNN(Recurrent Neural Network)을 이용한 연구가 늘어나고 있다.

⑥ 감성분석(Sentiment Analysis)

소셜 미디어는 SNS를 기반으로 급속히 성장하여 개인의 경험, 의견, 정보 등이 다양하게 생산되고 공유되고 있다. 소셜 미디어가 생산하는 대량의 텍스트로부터 사용자가 필요로 하는 정보를 신속히 찾아내고, 의미 있는 시사점을 도출하기 위해 텍스트 분석기술이 지속적으로 발전해오고 있다. 현재까지 텍스트 분석에 적용되는 기술들 중 오피니언 마이닝 기술을 활용한 대표적인 것으로 감성분석을 들 수 있다(Fang & Zhan, 2015). 감성분석은 사람들의 태도, 의견, 성향과 같은 데이터를 기반으로 특정 사안에 대한 긍정 혹은 부정을 파악하는 것으로, 문서의 주제가 무엇인지 찾아내는 것보다는 문서의 저자가 주제에 대해 가지고 있는 감정을 판별하는 것이다. 이러한 감성분석은 자연어처리(Natural Language Processing: NLP) 분야에서 개인의 감성, 감정 또는 의견을 추출하고 분석하는 연구로 가장 주목받고 있는 분석 방법이다. 특히 소셜 미디어 등과 같은 온라인으로 수집된 정형 또는 비정형 텍스트에 대하여 긍정과 부정의 선호도를 판별하는 데 유용하다. 감성분석에 대한 연구는 국내 및 해외에서 2000년 초반부터 지금까지 다양한 기법을 적용한 연구들이 활발하게 연구되고 있고, 연구 분야는 소셜 미디어, Twitter, 블로그 등을 적용한 영화 리뷰, 제품 리뷰, 여행 리뷰, 주가 관련 이슈 등이다.

(4) 결과를 위한 정보의 시각화 및 정보화

다양한 시민 의견을 분석한 결과는 공공서비스 정책실행자들이 분석 결과를 해석하기에 용이하도록 시각화하는 단계가 필요하다.

① 워드 클라우드(Word Cloud)

워드 클라우드는 데이터를 보여주는 다양한 방법 중 하나로 데이터를 구름 모양으로 보여줘서 붙여진 이름이다. 워드 클라우드는 그래프 형태로도 표현할 수 있으며 데이터가 가장 많은 것을 강조하기 위해 크게 표현할 수도 있으며 클라우드 상의 데이터의 크기를 넣어줄 수도 있다. 단어의 빈도수 데이터를 가지고 있을 때 이용할 수 있는 시각화 방법으로 많이 사용된다.



그림 II-4 워드 클라우드의 예시

② PyLDAvis

PyLDAvis는 토픽모델링에 이용되는 LDA 모델의 학습 결과를 시각화하는 Python 라이브러리다. 아래 <그림 II-5>는 아마존의 상품평에 대한 LDA 시각화 결과의 예이다.

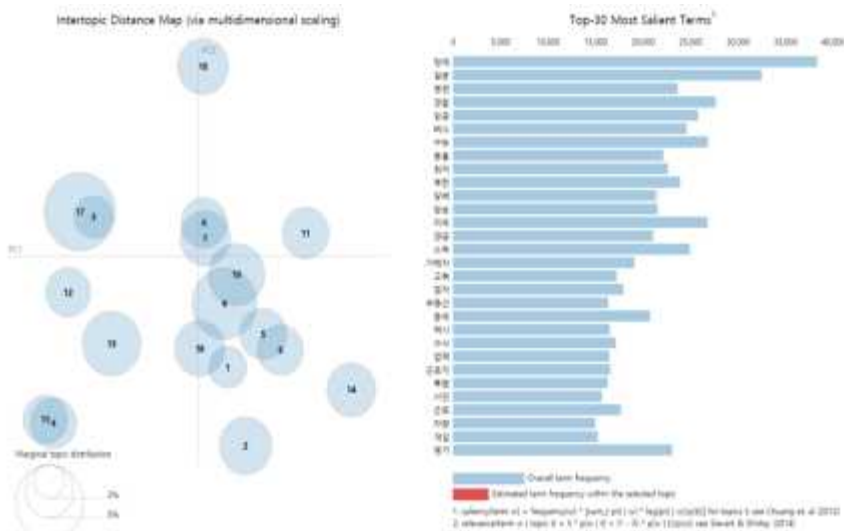


그림 11-5 pyLDAvis를 활용한 LDA 모형의 시각화 결과

③ 네트워크 분석 시각화

특정 이슈가 어떻게 발현되어 전파되는지 네트워크 분석을 통해 알 수 있다. 특히 분석 결과를 시각화하면 핵심어들의 빈도수뿐만 아니라 핵심어들 간의 관계를 한눈에 볼 수 있어 워드 클라우드보다 더 많은 정보를 얻을 수 있다.



그림 II-6 네트워크 분석 시각화 예시

2. 비정형데이터를 통한 청소년 정책의제 도출

1) 토픽모델링의 이해

빅데이터 분석 방법 중 토픽 모델링은 비정형데이터인 자연어를 특정한 목적에 유용한 정보를 추출하는데 적합한 분석 방법으로(Inzalkar & Sharma, 2015), 기존의 단어 빈도 분석에서 발생하는 희귀성 문제(Sparsity Problem), 다의성(Polysemy), 동의어(Synonymy), 의미 계층구조(Semantics Hierarchical Structure) 등의 문제점을 가지지 않는다. 즉 토픽모델링은 구조화되지 않은 방대한 문헌에서 일정한 패턴을 발견하여 주제를 찾아내는데 적합한 알고리즘이라 할 수 있다(Blei, 2012).

이러한 토픽 모델링은 일반적으로 두 가지 모형이 있는데, 첫 번째는 분절된 단어들에 벡터값을 부여하고 차원을 축소하여 근접한 단어들을 묶어내는 모형 (Latent Semantic Analysis: LSA)이고, 두 번째는 확률기법을 활용하여 문서의 토픽 분포를 추론하는 모형(Latent Dirichlet Allocation: LDA)이 있다(홍순구, 유승의, 안순재, 2019).

본 연구에서는 LDA 분석을 통하여 청소년 관련 토픽을 추출하고 각 토픽을 설명하는 상위 15개의 키워드를 보여줄 것이다. 이는 LDA가 문서를 구성하는 키워드의 확률분포를 구할 수 있으므로, 이를 기반으로 문서를 구성하는 토픽을 확률적으로 추정한 것이다. 이와 함께 LDA 분석 결과를 바탕으로 각 토픽을 구성하는 문서의 비중을 구하고, 코로나 전후 시기를 구분하여 토픽의 비중 변화 분석을 실시할 것이다. 코로나 이후 토픽 비중이 증가하였다면 중요한 정책 이슈라는 해석이 가능하고, 토픽의 비중이 감소했다면 비교적 덜 중요하거나 관심이 줄어든 정책 이슈라 볼 수 있다. 이처럼 토픽 모델링을 활용하여 주요 정책 이슈를 파악하고 변화를 확인할 수 있을 것이다.

2) 토픽모델링 분석을 통한 청소년 정책연구의 차별성

기존에는 청소년 관련 이슈를 도출하고 정책을 마련하기 위하여 문헌 조사 또는 전문가 인터뷰 및 자문회의 등과 같은 정성적 연구 방법이 주로 사용되어왔다. 그러나 이러한 방법은 대부분 연구자의 주관적 가치가 반영될 가능성이 크고, 소수의 전문가에 국한되어 이루어진다는 한계점을 내포하고 있다. 물론 대량의 정보에 접근할 수도 있겠지만, 정성적인 분석방법으로는 대량의 정보로부터 결과를 도출하는데 많은 시간과 비용이 소요될 뿐만 아니라 연구자의 주관적 가치가 반영될 가능성은 여전히 클 수밖에 없다.

따라서 이러한 정책 이슈 도출의 한계점을 보완하는데 텍스트 마이닝과 같은 정량적 연구 방법을 활용하여 보완하는 것이 필요하다. 이미 토픽모델링을 통한

비정형데이터의 분석은 심리학, 정치학, 서지분석 등 다양한 분야에서 활용되고 있는데, 2000년대 이후부터 본격적으로 이루어지고 있다(Blei, 2012). 청소년 정책연구에서도 빅데이터에 기반한 정책 수립과 미래 예견적 의사결정의 중요성이 점차 강조되면서, 빅데이터를 적용한 연구방법론으로 변화를 점차 시도하고 있다(김상겸, 장성용, 2016; 송태민, 2020; 홍성효, 2020).

본 연구에서는 토픽모델링을 활용하여 온라인에서 시민들이 남긴 대량의 비정형데이터를 통해 미래 청소년 정책의제를 도출 할 것이다. 시민들이 정책에 대한 다양한 의견을 제안하고, 토론하는 국민신문고와 국민청원 등을 대상으로 청소년 이슈를 도출하고자 하였다는 점이 기존 연구와 차별성을 가진다. 또한 이슈 도출을 위하여 토픽모델링이라는 계량적인 방법을 사용함으로써 신뢰성과 객관성을 확보하고자 하였다.

앞으로 이러한 연구의 차별성을 고려하여 지속적이고 누적적 분석을 통하여 청소년과 관련된 이슈를 분석하고 체계적이고 지속적으로 관리한다면 향후 정책 수립 및 대응 방안을 마련하는데 중요한 자산이 될 것이다.



그림 11-7 청소년 정책의제 도출을 위한 연구절차 및 방법

3. 분석 대상 및 데이터 수집

1) 분석 대상

본 연구에서는 대표적인 비정형 데이터인 온라인 게시글 및 민원을 연구 대상으로 하여 청소년 정책 이슈를 도출하기 위하여 토픽모델링 기법을 활용하였고, 그 결과에 대해 해석하였다.

현재 정부에서 운영하고 있는 국민신문고³⁾와 국민청원⁴⁾ 사이트는 시민들이 자유롭게 정책을 제안하고 정책을 토론하는 온라인 공간으로, 해당 사이트에 2002년 1월부터 2021년 7월까지 올라온 청소년 정책과 관련한 게시글 및 민원을 수집하여 기술통계분석, 토픽 모델링 분석, 네트워크 분석을 진행하였다.

연구의 분석 대상의 범위를 크게 대상적 범위, 시간적 범위로 다시 구분하여 정의한다면, 본 연구의 공간적·대상적 범위는 국민신문고와 국민청원 게시판의 제목과 내용이며, 연구의 시간적 범위는 2002년 1월 1일부터 2021년 7월 31일 까지의 기간이다.

2) 데이터수집 채널 및 정보

미래 청소년 정책의제 도출이라는 연구목적에 고려하여 ‘아이’, ‘아동’, ‘청소년’, ‘청년’, ‘학생’을 키워드로 설정하고 국민신문고와 국민청원 내에 있는 총 9개 채널에서 2002년부터 1월부터 2021년 7월까지의 데이터를 수집하였다.

3) 국민신문고(www.epeople.go.kr)는 정부에 대한 민원, 제안, 참여 등을 인터넷으로 간편하게 신청하고 처리하는 법정부 대표 온라인 소통 창구이다.

4) 국민청원(www.president.go.kr/petitions) 게시판은 청와대 홈페이지에서 운영하는 게시판으로 2017년 8월 17일 개설되었는데, 청원을 등록하고 30일 동안 20만 개 이상의 동의를 받으면 정부나 청와대 관계자들이 청원에 대한 답변을 제공하고 있다.

국민신문고는 대표적인 정부의 통합 온라인 소통창구로 정부에 대한 민원, 제안, 참여 등이 가능하다. 이러한 국민신문고에서 정보검색 및 수집이 가능한 공개 제안, 실시 제안, 민원질의응답, 생각 모음, 주제토론, 베스트 토론 6개의 채널에서 데이터를 수집하였다. 이와 함께 청와대에서 운영하고 있는 국민청원 게시판, 국민청원 분야별 게시판, 토론방에서 데이터를 수집하였다. 수집한 데이터의 정보는 <표 II-1>과 같다.

표 II-1 데이터 수집 채널과 수집 내용

	수집 채널	사이트	수집내용
1	공개 제안	국민신문고	글제목 평점 분야 처리기관 글본문(현황 및 문제점 개선방향 기대효과) 신청일 첨부파일여부
2	실시 제안	국민신문고	글제목 평점 분야 처리기관 글본문(현황 및 문제점 개선방안 기대효과) 신청일 첨부파일여부
3	민원질의응답 답변원문	국민신문고	사례명 답변의 글 본문 담당부서 관련법령(관련법령은 그 법령의 제목만 수집)
4	생각 모음	국민신문고	시작일 단계(진행중 등) 참여인원 제목 날짜 별점 좋아요 댓글개수 본문내용 참여기간 관련주제 태그 댓글 내용
5	주제토론	국민신문고	카테고리(대화·설문) 글쓴이 등록일 별 좋아요 댓글개수 본문내용 참여기간 관련주제 관련지역 태그 댓글작성자 댓글작성일 댓글내용 완료된 설문의 경우-결과
6	청년정책 제안	국민신문고	카테고리(대화·설문) 글쓴이 등록일 별 좋아요 댓글개수 본문내용 참여기간 관련주제 태그 관련지역 댓글작성자 댓글작성일 댓글내용
7	국민청원	청와대 국민청원	카테고리 청원시작 청원마감 청원인 진행단계 청원 내용 청원동의명수 동의내용
8	분야별 국민청원	청와대 국민청원	카테고리 청원 시작 청원 마감 청원인 진행단계 청원 내용 청원동의명수 동의내용
9	토론방	청와대 국민청원	제목 추천·비추천·작성일 작성자 본문내용 댓글(입력일 댓글 내용 추천·반대)

3) 분석 방법

국민신문고와 국민청원 등 9개 채널에서 수집한 데이터에 대해 기술통계분석, 토픽모델링 분석, 네트워크 분석을 진행하였다. 분석에는 LDA파이썬 라이브러리 중 LDA모델의 학습 결과를 시각화하는 PyLDAvis와 넷마이너(Netminer 4.0) 프로그램을 사용하였다.

기술통계 분석을 통해서 분석에 활용된 청소년 관련 데이터 분포 결과와 수집 키워드의 시기별 추이 변화를 파악하였다. 그리고 중앙행정기관별로 데이터를 구분하여 토픽모델링 분석을 실시하여 청소년과 관련된 전체적인 맥락을 파악하고, 청소년 정책 이슈를 도출하였다. 다음으로 코로나 전후의 토픽들의 비중 변화를 통하여 청소년 이슈 변화를 분석하였다.

(1) 데이터 전처리

국민신문고와 국민청원 홈페이지를 활용하여 텍스트를 수집하고 분석 데이터 준비를 위한 첫 번째 단계로 중복데이터를 삭제하였다. 그리고 추론한 데이터에서 이모티콘, 띄어쓰기, 초성 등 불용어를 제거하는 데이터 클렌징(cleansing) 작업을 하였다. 다음으로 어린이, 어린애, 어린아이 등과 같이 유사한 단어들을 찾아 통일하고 정제하였다. 이러한 형태는 다르나 의미가 같은 동의어 처리는 매우 중요한 단계이다(최수정, 2018). 그리고 문장의 의미에 영향을 미치지 않는 때, 시 등한 음절 단어들을 제거하는 등의 작업을 거쳐 명사 형태의 단어만을 추출하여 분석에 사용하였다.

이와 함께 텍스트 마이닝에서 중요정보를 추출하기 위해 TF-IDF(Term Frequency-Inverse)⁵⁾도 활용하였다(Spärck, 1972). 모든 문서에서 전반적으로

$$5) w_{i,j} = tf_{i,j} \times \log\left(\frac{N}{df_i}\right)$$

$tf_{i,j}$ =number of occurrences of i in j , df_i =number of documents containing i
 N =total number fo documents

자주 등장하는 단어에는 패널티를 주고, 단어 빈도 처리에서 많이 나타나지 않은 단어에 대해 가중치를 높게 부여하는 TF-IDF를 활용하여 토픽모델링에서 주제를 분류하는 데 큰 영향을 미치지 않거나 의미가 없는 단어를 제외하였다. 예를 들어 생각, 시간, 업무, 학교 등의 단어는 TF-IDF가 낮은 단어로 토픽을 분류하는데 큰 영향을 미치지 않으므로 제외하고 분석하였다.

(2) 토픽모델링 분석 방법

토픽모델링은 각 단어들이 통계적으로 특정 토픽에 포함될 확률을 파악하여 문서의 주제를 추정하는 기법으로, LDA는 의미적으로 일관성 있는 주제들을 추출하는 데 강점을 가지는 것으로 알려져 있다(Mimno & McCallum, 2012).

LDA(Latent Dirichlet Allocation)을 간단히 설명하면 문서들을 각 토픽의 분포로 표현하는 것이다. 양연희(2021)에서는 “이때 토픽은 각 키워드가 분포하는 특징으로 묶이게 된다. $p(\mathbf{z}|\mathbf{d}_i)$, $p(\mathbf{w}|\mathbf{z}_i, j)$ 를 각각 문서 i 에 대한 토픽 확률 밀도 함수와 문서 i 의 j 번째 단어에 할당된 토픽에 대한 단어 확률 밀도 함수라고 표기한다. 이러한 분포가 주어졌을 때 LDA는 다음 생성 프로세스를 통하여 새로운 문서를 생성한다. LDA는 문서 생성 프로세스를 모델링하도록 설계되지만, 이것은 메인 노드와 하위 노드 집합으로 이루어진 2-모드 네트워크를 생성 프로세스 모델링에 사용할 수 있다”고 설명하고 있다.

LDA 분석은 단어-문서 데이터에서 토픽이라는 새로운 레이어를 추출하는 분석이기 때문에, 연구자가 직접 토픽의 수, 알파(alpha), 베타(beta) 값, 반복 횟수(iteration)의 하이퍼파라미터(hyper-parameter) 값들을 지정해야 한다.

여기서 알파 값은 문서별 토픽의 할당(Distribution)에 직접적으로 영향을 주는 값으로, 알파 값이 크게 되면 한 문서에 여러 토픽이 비슷한 확률로 할당되게 된다. 이는 한 문서가 특정 토픽에 치우치지 않도록 한다. 다음으로 베타 값은

토픽별 단어의 할당에 영향을 주는 값으로, 베타 값이 크면 모든 단어가 비슷한 확률로 특정 토픽에 할당되게 된다. 즉 알파와 베타 값을 크게 하면 모든 문서, 모든 단어가 모든 토픽에 비슷한 확률로 할당되기 때문에, 작은 값을 쓸수록 소수의 토픽이 문서에 할당되고, 소수의 단어가 토픽에 할당된다.⁶⁾

본 연구에서는 알파 값은 50/토픽수, 베타 값은 0.01을 추천하는 연구를 토대로 해당 값을 기준으로 삼아 분석하였다(Griffiths & Steyvers, 2004). 그리고 최적 토픽 수의 산정⁷⁾에 있어서는 클러스터링 평가 방법인 Silhouette 방법을 활용하였다(Panichella et al., 2013). 이때 최적 토픽 수를 도출하기 위한 반복 횟수(iteration)로는 1000번을 적용하였다. 그러나 이러한 최적 토픽 수를 결정하는 방법만으로는 한계가 있어 최종 토픽 수를 결정하는 것은 분석자가 보기에 범주화가 잘 되었다고 판단하는 것이 중요하다(이수상, 2016). 따라서 본 연구에서는 선행 연구를 종합하여 Silhouette 방법을 통해 일차적으로 토픽을 분류하고, 다음으로 분류된 토픽 간에 의미적으로 주제가 잘 구분되고 범주화가 적절하게 이루어졌는지 분석자가 판단하여 최종 토픽 수를 결정하였다.

이 때, 토픽의 주제를 선정하는 데 결정적인 역할을 하는 핵심 키워드의 중복이 나타나지 않는 토픽의 수를 선정한다. 예를 들어, 청소년 정책에 있어서 주요 키워드이자 관련 단어인 ‘청소년’, ‘가정’, ‘정책’ 등의 단어와 같은 경우는 대부분의 토픽에서 핵심 키워드로 상위에 등장하지만, 여러 토픽에 동시에 나타남으로 토픽의 주제를 정할 때는 해당 키워드가 큰 영향을 주지 않았다.

따라서 토픽의 주제를 정함에 있어서는 이러한 키워드들의 중복 여부에 의미를 두기보다는 ‘육아’, ‘폭력’, ‘상담’, ‘아이돌보미’, ‘수능’ 등과 같이 주제의 범주화에 영향을 미치는 키워드에 의미를 두며 판단하였다.

6) 넷마이너(2020). 텍스트네트워크분석 교육교재와 넷마이너 4.0 프로그램 설명서를 참조하여 정리하였다.

7) 최적 토픽수는 선정에 있어서는 혼잡도(Perplexity)를 통해 선정하는 방법도 많이 활용되고 있으나, 혼잡도는 특정 확률모델이 실제로 관측되는 값을 얼마나 잘 예측하는지 평가할 때 사용하는 방법으로 텍스트 분석에서 해당 방법을 적용하는 것은 결과도출이 어렵다는 견해가 있다(이수상, 2016).

〈그림 II-8〉은 여성가족부 코로나 이전과 이후의 최적 토픽 수를 구하기 위해 토픽의 수, 알파(alpha), 베타(beta) 값 등의 하이퍼파라미터를 변경하면서 shihouette를 구한 표이다. shihouette 수치가 가장 높게 나타난 토픽의 수를 우선적으로 선택한 후 주제가 적정하게 범주화되었는지 판단하여 최종 토픽 수는 분석자가 결정하였다.

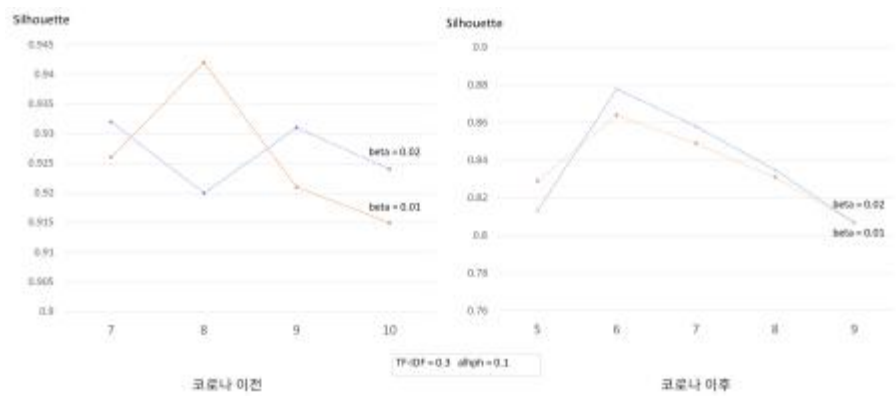


그림 II-8 토픽 수 선정을 위한 Silhouette



제3장 연구결과

- 1. 기초 분석 결과
- 2. 주요 중앙행정기관별
LDA 분석 결과
- 3. 코로나 전후 청소년 정책
이슈 변화

1. 기초 분석 결과⁸⁾

1) 데이터 수집 결과

국민신문고와 국민청원 내에 있는 총 9개 채널에서 ‘아이’, ‘아동’, ‘청소년’, ‘청년’, ‘학생’이라는 키워드를 제목과 내용에 포함하는 데이터를 2002년 1월부터 2021년 7월까지 수집한 결과 각각 45,346개와 288,025개의 데이터가 수집되었다.

표 III-1 수집 데이터 수

	아이	아동	청소년	청년	학생	데이터 수
제목	11,183	4,920	12,125	5,096	12,022	45,346
내용	95,169	17,174	29,504	24,547	101,631	288,025

8) ‘제3장 연구결과’는 동아대학교 임지원 전임연구원이 작성하였다.

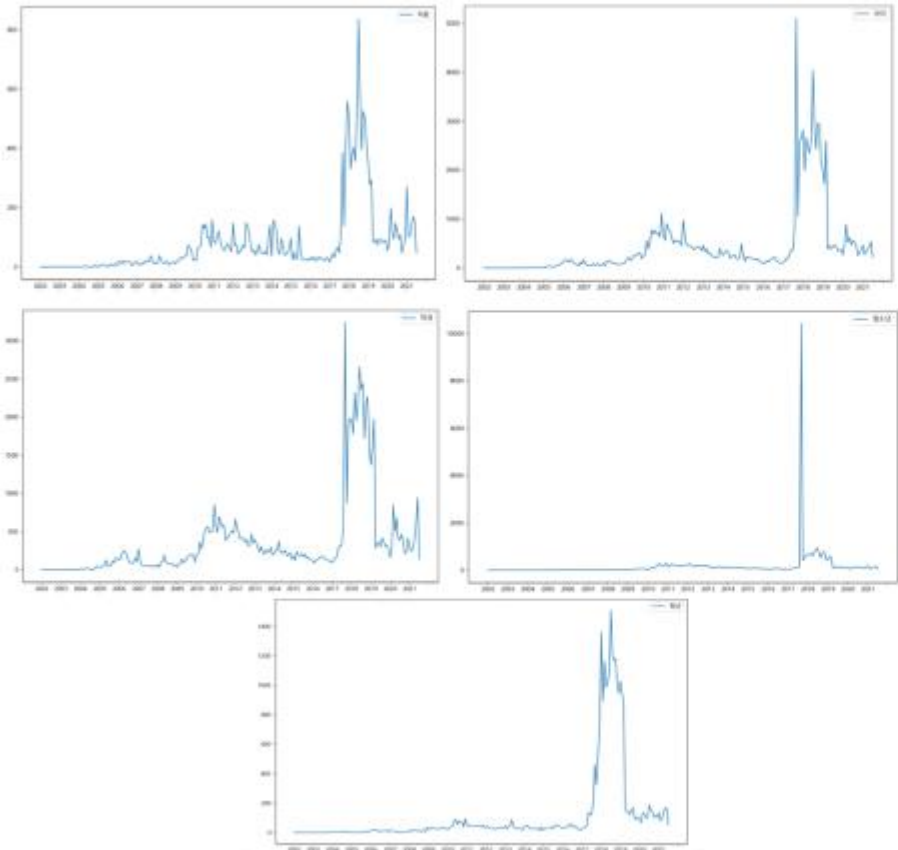


그림 Ⅲ-1 검색키워드 시간별 추이변화

수집된 데이터에서 다수의 키워드로 검색되어 중복되는 문서들을 제거하고, 동일한 사람이 여러 개 올린 문서는 최초 문서만 남기고 모두 제거하였다. 함께 데이터 전처리를 통해 불용어 등을 제거하였다. 키워드로 인한 중복 및 불용어를 제거한 데이터는 제목으로 수집된 경우는 42,616개, 내용으로 수집된 경우는 265,744개이다<표 Ⅲ-2 참조>

표 III-2 채널 및 키워드 중복제거

	아이	아동	청소년	청년	학생	불용어 제거 후
제목	6,896	4,497	11,147	4,763	9,963	42,616
내용	88,936	15,819	27,699	23,077	84,618	265,744

이를 다시 지자체 및 기타 중앙행정기관 등의 정보를 가지고 있는 데이터를 재추출하고 코로나가 발생한 2020년 1월 20일을 기준으로 코로나 이전과 이후로 분리하여 <표 III-3>과 같은 결과를 얻었다.

이때 수집된 데이터를 분석한 결과, 내용에 키워드를 포함한 데이터가 제목에 해당 키워드를 포함한 데이터를 90% 이상 포함하고 있어 내용을 기준으로 수집된 데이터를 분석 대상으로 삼았다.

즉 9개의 수집 채널에서 본문 내용에 ‘아이’, ‘아동’, ‘청소년’, ‘청년’, ‘학생’을 키워드로 포함하고 있는 데이터를 분석 대상으로 삼았으며, 최종 분석 대상 데이터 수는 코로나 이전 162,787개, 코로나 이후 31,461개이다.

표 III-3 분석 데이터 현황

기간	검색 범위	여성가족부	교육청	교육부	지자체 및 기타 중앙 행정기관	합
코로나 이전	제목	952	1,060	2,839	25,947	30,798
	내용	2,450	4,616	15,219	140,502	162,787
코로나 이후	제목	46	1,301	761	3,374	5,482
	내용	129	10,918	4,883	15,531	31,461

2) 전체 데이터 LDA 분석 결과

다음은 주요 중앙행정기관별로 데이터를 구분하여 분석하기에 앞서 전체 데이터를 대상으로 토픽의 범주화가 잘 되는지 등을 확인하기 위하여 1차적으로 분석 실시한 결과이다. LDA 파이썬 라이브러리 중 LDA모델의 학습 결과를 시각화하는 PyLDAvis를 활용하였다.

(1) 코로나 이전 LDA

지자체 및 중앙행정기관 정보를 포함하는 코로나 이전 162,787개 데이터를 빈도 분석한 결과, 장애, 일본, 병원, 경찰, 임금, 버스 등의 단어 순으로 상위에 나타났으며, 18개의 토픽으로 분류되는 것을 확인할 수 있었다.

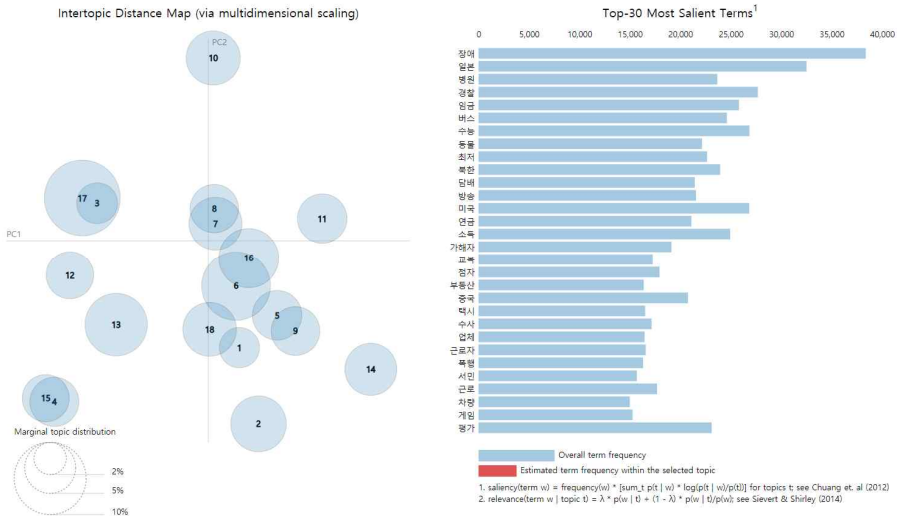


그림 III-2 코로나 이전 전체데이터 LDA 결과 및 Top-30 단어

(2) 코로나 이후 - topic 8개

지자체 및 중앙행정기관 정보를 포함하는 코로나 이후 31,461개 데이터를 빈도 분석한 결과, 대통령, 주택, 사건, 시험, 경찰 등의 단어가 상위에 나타났으며, 병원, 확진, 마스크, 감염 등 코로나와 관련된 다수의 단어들도 상위 30개 단어에 나타났다. 해당 데이터를 LDA 분석한 결과 8개의 토픽으로 코로나 이전에 비해 크게 감소한 것을 확인할 수 있다.

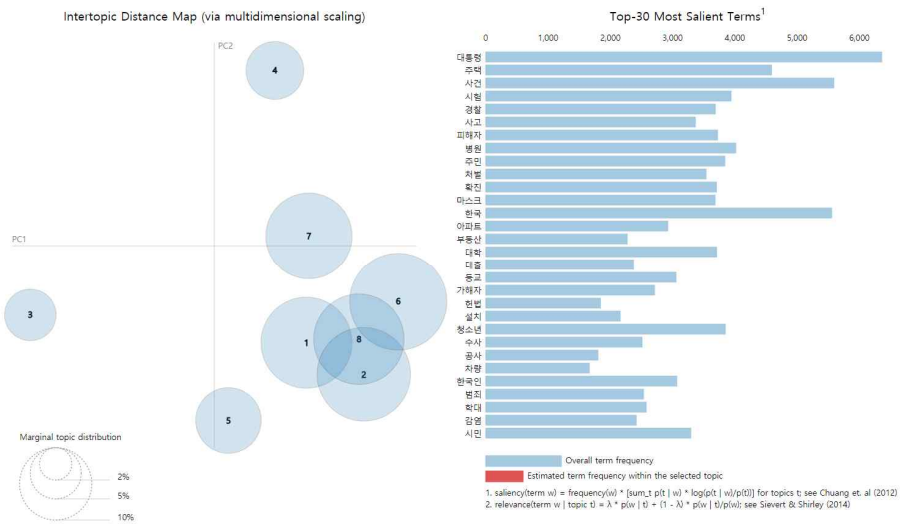


그림 III-3 코로나 이후 전체데이터 LDA 결과 및 Top-30 단어

키워드를 통하여 1차적으로 청소년과 관련된 데이터를 추출했다고 하더라도 적절한 토픽의 범주화 및 청소년 관련 정책의 도출에 용이한 결과가 아니라고 판단하여, 이를 다시 청소년과 유관한 중앙행정기관인 여성가족부, 교육부, 교육청, 기타 기관으로 분류하여 분석하는 것이 청소년 정책 이슈 도출에 적절하다 판단하였다. 이에 각 기관별로 LDA분석을 시도하여 청소년 정책 이슈를 도출하였다.

2. 주요 중앙행정기관별 LDA 분석 결과

1) 여성가족부

(1) 코로나 이전

코로나 이전 여성가족부 데이터를 토픽 모델링한 결과 8개의 정책 이슈가 도출되었다. Topic-1은 보호, 성매매, 처벌, 성범죄, 폭력 등의 단어의 연관 확률이 높은 것으로 나타나 ‘성매매 처벌·성범죄 피해자 보호’와 관련된 주제로 추론된다. Topic-2는 사람, 남성, 평등, 여자, 업무 등의 단어 등으로 구성되어 ‘양성 평등정책’과 관련된 주제로 판단된다. 그리고 Topic-3은 미성년자를 대상으로 한 ‘술·담배 판매자에 대한 처벌강화 및 청소년증 의무발급’에 관한 내용이 주제로 나타났으며, Topic-4는 ‘한부모가정 및 저소득층에 대한 양육비 지원 문제’에 대한 주제로 보인다. 다음으로 Topic-5는 학교, 복지, 사회, 센터, 상담 등이 연관 확률이 높은 단어로 나타나 ‘학교 및 지역사회 상담 기관 필요성 및 상담운영 지원’이 주제로 보인다. Topic-6은 ‘아이돌봄서비스 및 육아 지원 서비스’ 관련 주제로 돌보미, 선생, 육아, 지원, 신청 등의 단어가 나타났다. Topic-7은 최근 문제로 대두되고 있는 ‘스마트폰 및 인터넷게임 중독 방지를 위한 셧다운 및 인증제’와 관련된 내용으로 시간, 셧다운, 게임, 인증, 스마트폰 등이 연관성이 높은 단어로 도출되었다. 마지막으로 topic-8은 ‘결혼과 저출산’ 사회 문제에 대한 관심이 정책 이슈로 도출되었다<표 III-4 참조>.

표 III -4 코로나 이전 여성가족부 LDA 분석 결과

Topic	Topic-1	Topic-2	Topic-3	Topic-4	Topic-5	Topic-6	Topic-7	Topic-8
순번	성매매처벌· 성범죄 피해자 보호	양성평등 정책	미성년 대상 술· 담배 판매자 처벌 강화 및 청소년중 의무발급	한부모 가정 및 저소득층 양육비 지원	학교 및 지역사회 상담기관 필요성 및 상담운영 지원	아이돌봄 서비스 및 육아 지원 서비스 확대	스마트폰 및 인터넷게임 중독방지 셋다운 및 인증제	결혼· 저출산 사회문제
1	보호	사람	사람	생각	학교	업무	시간	문화
2	생각	여성부	처벌	해택	복지	시간	셋다운	부모
3	성매매	가족	발급	정책	사회	돌보미	게임	교육
4	사람	말	미성년자	양육비	지역	생각	정보	가정
5	사건	남자	신분증	교사	운영	말	번호	업무
6	업무	남성	청소년증	지급	문제	선생	인증	부부
7	처벌	사회	유해	소득	제공	육아	스마트폰	한국
8	성범죄	평등	판매	보육	센터	이용	가능	정부
9	법	정부	술	지원	상담	지원	생각	문제
10	범죄	여자	담배	부모	활동	신청	인터넷	결혼
11	성	정책	법	가정	교육	가정	중독	사회
12	여자	업무	출연	어린이집	시설	서비스	이용	생활
13	폭력	국민	생각	양육	기관	부모	부모	여자
14	사회	생각	성인	업무	필요	사람	사용	생각
15	피해자	공무원	보호	정부	지원	아이돌보미	주민	자녀

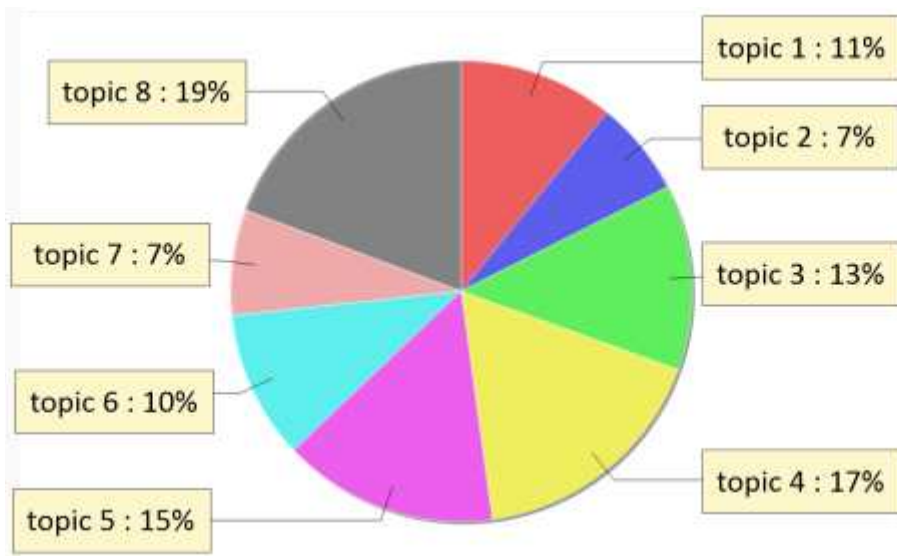


그림 III-4 코로나 이전 여성가족부 토픽별 비중

코로나 이전 여성가족부에서 도출된 토픽은 총 8개로, 각 토픽이 차지하는 비중은 다음과 같이 나타났다. 결혼·저출산과 같은 근본적인 여성가족부의 정책 이슈인 Topic-8이 19%로 가장 높은 비중을 차지하고 있는 것을 확인할 수 있으며, 다음으로 Topic-4인 한부모 가정 및 저소득층 양육비 지원에 관한 문제가 19%로 높은 비중을 차지함으로써 해당 이슈에 대한 사람들의 관심이 매우 높음을 알 수 있다(그림 III-4 참조). Topic-8을 제외하고 대부분 청소년과 관련성이 높은 정책적 이슈가 도출되었다고 할 수 있다.

다음은 LDA 분석을 통해 나타난 토픽들의 상위 단어 15개를 topic map으로 나타낸 것으로 사람, 가정, 생각, 부모, 지원, 교육, 사회, 여자 등의 단어가 여러 토픽에 동시에 등장하는 것으로 나타나 이러한 단어 이외에 해당 토픽에만 나타난 특징적인 단어를 참고하여 토픽 명을 지정하였다(그림 III-5 참조).

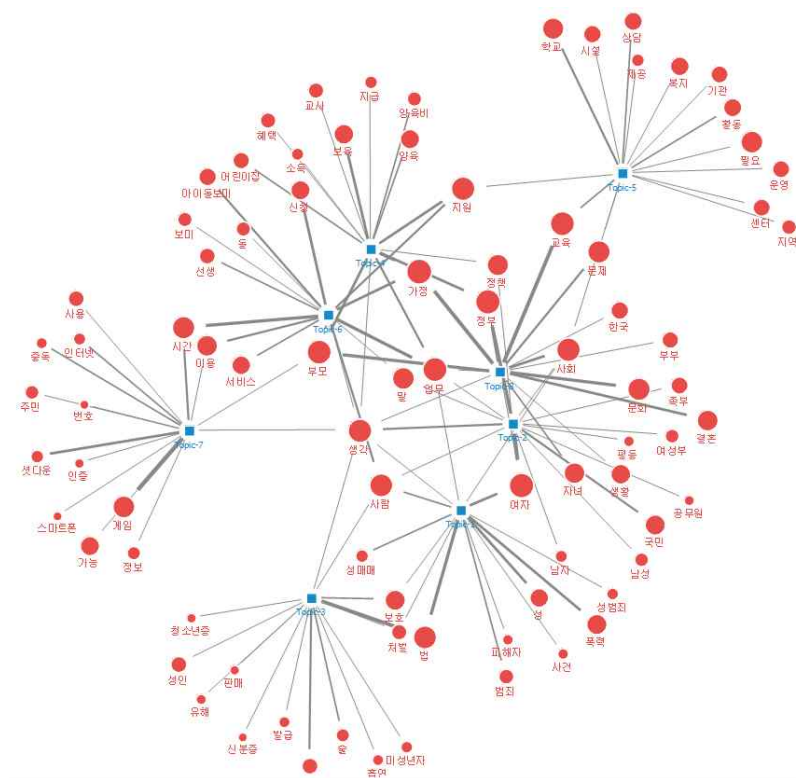


그림 III-5 코로나 이전 여성가족부 topic map

(2) 코로나 이후

코로나 이후 여성가족부 데이터를 토픽 모델링한 결과 6개의 정책 이슈가 도출되었다. Topic-1은 가족, 청소년, 지원, 상담, 기관 등의 단어의 연관 확률이 높은 것으로 나타나 ‘청소년·가족 상담 서비스’와 관련된 주제로 추론된다. Topic-2는 아동, 성범죄, 경력, 조회 등의 단어 등으로 구성되어 ‘아동 관련 기관의 성범죄자·아동학대범죄 정보 및 경력 조회 시행’과 관련된 주제로 판단된다. 그리고 Topic-3은 ‘코로나로 인한 맞벌이가정 육아 지원 서비스’에 관한 내용이 주제로

나타났으며, Topic-4는 ‘성매매·자살 예방 의무 교육’에 대한 주제로 보인다. 다음으로 Topic-5는 술, 구매, 청소년, 신분증 등이 연관 확률이 높은 단어로 나타나 ‘청소년 대상 술·담배 판매지도 및 벌금·청소년증 의무발급’이 주제로 보인다. 마지막으로 topic-6은 자녀, 양육비, 기준, 이혼, 양육 등 단어가 키워드로 나타나 ‘한부모가정 자녀 양육비’가 주제로 도출되었다<표 III-5 참조>.

표 III-5 코로나 이후 여성가족부 LDA 분석 결과

Topic	Topic-1	Topic-2	Topic-3	Topic-4	Topic-5	Topic-6
순번	청소년 가족 상담서비스	아동관련 기관의 성범죄자·아동학대 범죄 정보 및 경력 조치 시행	코로나로 워킹맘 육아지원	성매매· 자살 예방 의무 교육	청소년대상 술· 담배 판매 지도 및 벌금·청소년증 의무발급	한부모 가정 자녀양육비
1	가족	아동	아이	교육	술	부모
2	복지	관련	일	학생	과목	지원
3	청소년	성범죄	여성	학교	구매	가정
4	가정	경력	사람	사회	청소년	자녀
5	부	조치	나라	예방	발급	양육비
6	지원	법	정부	문제	신분증	기준
7	필요	학교	대한민국	성매매	생리대	아이
8	상담	성범죄자	육아	경우	시험	국가
9	서비스	여성	생각	자살	사용	이혼
10	신청	국민	사회	한국	담배	양육
11	제도	범죄	문제	방법	사항	혜택
12	기관	학대	필요	사실	학생	경우
13	가구	시행	어른	범죄	경우	가족
14	경제	정보	엄마	의무	벌금	지급
15	건강	기관	코로나	효과	지도	정책

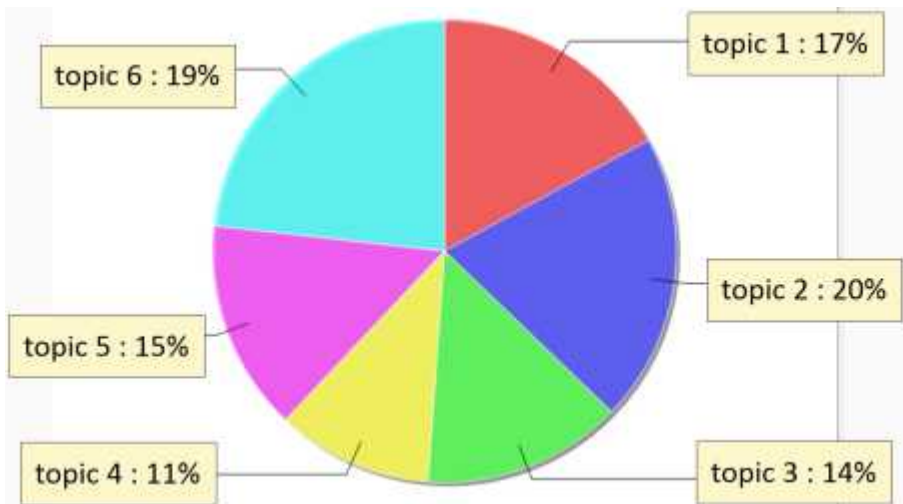


그림 III-6 코로나 이후 여성가족부 토픽별 비중

코로나 이후 여성가족부에서 청소년 정책과 관련된 토픽은 총 6개로, 각 토픽이 차지하는 비중은 <그림 III-6>과 같이 나타났다. 아동 관련 기관의 성범죄자·아동 학대 범죄 정보 및 경력 조회 시행을 주제로 한 Topic-2가 20%로 가장 높은 비중을 차지하고 있는 것을 확인할 수 있으며, 다음으로 Topic-6인 한부모가정의 자녀 양육비에 관한 문제도 19%로 높은 비중을 차지함을 알 수 있다.

표 III-6 여성가족부 코로나 이후 토픽 원문 예시

관련 토픽	원문
1	전문적인 청소년 미혼모(부) 가정(한부모 가정) 전담 공무원의 배치와 만족도 조사 실행: 변동이 잦고, 다양한 복지 서비스를 습득하는 과정에서 어려움이 존재하는 것은 분명하나, 복지서비스 부서를 보다 세분화하여 '한부모가정'과 '청소년한부모가정'의 전담 상담사를 배치하여 상담사의 개인 부담을 줄이고, 한부모가정에 대한 이해도가 높은 인력을 배치하여 그들에 대한 2차적인 피해를 방지하고, 다양한 복지 서비스 제공을 확대하였으면 좋겠습니다.
2	그러나, 시설·기관 또는 사업장을 특정하기 어려운 아동·청소년관련 사업*을 시행하는 자치단체의 장은 아동·청소년 관련기관 등에 포함되지 않아 관계 기관의 장(경찰서장)에게 취업자 등에 대하여 성범죄 등의 경력 조회 권한이 없음
2	개선방안은 대상자 1인이 아동 청소년이 있는 여러 곳을 방문할 경우 주무 관청에서 직접 채용하거나 계약을 한 대상자에 대하여 성범죄 및 아동학대 전력 조회를 허용하고 해당기간 동안 활동하는 모든 학교에 조회결과의 유효성을 일괄 적용하도록 관련 규정을 개선할 필요가 있음
3	- 청소년 한부모가정의 전담 공무원 배치로 인해 복지의 사각지대에 놓인 청소년 미혼모(부) 수 감소 - 청소년 한부모가정 아이돌봄서비스를 제공함으로써 청소년 미혼모(부)의 학업 중단 방지 - 다양한 가족의 형태를 이해하고 존중하는 사회적 인식 개선 한 아이의 생명을 책임지기 위해 노력하는 수 많은 어머니, 아버지에게 더 나은 삶을 제공할 수 있도록 위 제언들을 살펴봐 주시기를 간곡히 부탁드립니다.
4	- 학교에서 학생들을 대상으로 전문 상담 선생님과 상담을 실시한다. (집단 상담 형식으로 더 자주 실시한다) - 누구나 힘들면 상담을 받을 수 있도록 국가에서 적극적으로 지원한다. 또, 위 방법들이 현재 존재하지만 부실한 면이 있는 상담기관들과 자살 예방 교육의 효용성을 극대화할 것이라고 기대합니다
6	하지만 미혼모자 시설에 있지 않은 청소년 한부모가정은 정부 제공 아동 양육비 외에 경제 활동과 학업을 병행하기 어려운 상황으로 청소년 한부모 가정만을 위한 차별화 된 아이돌봄서비스가 제공되어야 합니다.
6	갑자기 예고도 없이 양육수당을 안주면, 그 돈으로 쌀도 사고, 아이 학교 교통비도 하고 책도 사고 했는데, 저소득 한부모가정 아이들은 고등학교 3학년 공부 하지 말고, 일해서 돈벌이란 말입니까. 그런데, 더욱 더 열심히 공부해야 하는 고등학교 3학년이 되었는데, 갑자기 여성가족부에서 지급되던 아동양육비 20만원을 주지 않았습니다.

관련 토픽	원문
5	이에 마스크 구매5부제에 따라 마스크를 구입할 때 청소년증을 사전에 발급받은 청소년은 발급하지 않은 청소년보다 상대적으로 쉽게 약국에서 본인의 신분을 입증할 수 있었음 (문제점) 청소년증을 신규로 발급할 경우 재발급과 달리 온라인 신청이 되지 않아, 본인이나 대리인이 직접 근처 주민센터를 방문해야 하므로 불편함이 있어, 청소년증의 신규 발급부터 비대면 방식을 검토할 필요가 있음
5	청소년증을 주민등록증처럼 온라인상의 본인인증 수단으로 사용할 수 있게 해주십시오
5	학교 밖 청소년은 청소년증을 발급받고 학교에 다니는 학생은 학생증을 발급받고 있다. 하지만 학교 밖 청소년들은 곱지 않은 사회시선으로 청소년증을 발급받지 않고(20%의 발급율) 있어 두가지 증을 하나로 통합하여 사용할 수 있는 청소년 통합증이 필요하다 1. 청소년증과 학생증을 합친 통합증을 고려하자 - 초기에 청소년증에 대한 홍보를 신분증명으로 접근하다 보니 학교 다니는 학생은 학생증, 학교 안다니는 학생은 청소년증이라는 이분법적인 관점으로 나뉘어 사회적 낙인이 찍힌 부분을 상기하자
4	중앙자살예방센터가 발행한 '2020 자살 통계 자료'에 따르면 2019년 한국인 자살자 수는 2018년과 비교해 0.9% 증가하였고, 한국의 자살률은 경제협력개발기구(OECD) 회원국 중 최고 수준입니다. 문제를 분석하는 과정에서 찾은 원인은 학업 스트레스, 주변과의 갈등, 상담기관과 자살 예방 교육의 부실 등이었기에 이를 토대로 모색한 아래의 4가지 해결책을 제안합니다. - 교내에 자살예방센터를 마련한다. - 자살을 생각하는 친구를 돕는 입장 1) 자살예방교육의 취지를 '자살을 생각하는 친구를 도와주자'로 하여 교육한다. 2) 그룹 프로젝트 형식으로 자살 예방 교육을 진행한다. 또, 위 방법들이 현재 존재하지만 부실한 면이 있는 상담기관들과 자살 예방 교육의 효용성을 극대화할 것이라고 기대합니다.
6	저소득 한부모가정, 특히 엄마 혼자 키우는 가정이 대한민국에서 얼마나 힘든건지 위정자들은 죽어도 알지 못할 것입니다. 기대효과: - 저소득 한부모가족의 경제적 지원으로 자녀 양육의 질을 높이고 교육 기회 박탈을 방지하여 일정 수준의 교육 이수를 통해 자녀 자립의 기초를 마련 - 복지사각지대 감소
6	기대효과: - 저소득 한부모가족의 경제적 지원으로 자녀 양육의 질을 높이고 교육 기회 박탈을 방지하여 일정 수준의 교육 이수를 통해 자녀 자립의 기초를 마련 - 복지사각지대 감소

관련 토픽	원문
6	아이들은 이미 청소년기에 접어들고 지원은 부족하고 다자녀임에도 특별히 지금은 도움 받는 것이 미미합니다. 실질적인 도움이 절실한데 지원은 적거나 없고 아동수당은 이미 성장해서 해당도 안 되고 번번이 정책이 만들어질 때마다 혜택에는 나이가 항상 걸립니다.
2	성범죄자의 생활권에서 주변에 아동이 있는 옆 집에만 알리는 것이 아니라, 성범죄자와 같은 블록내지 같은 단지의 주민 모두에게 주변에 성범죄자가 있다는 것을 알려 성범죄자 알림e를 통해 확인하고 조심할 수 있게 해야 한다. 현재 '성범죄자알림이' 앱라는 것이 있기는 하지만 전자발찌 GPS를 추적하는 실시간 알리미가 아닙니다. 조두순이 출소하기 전에 손 놓고 가만히 있지 말고 성범죄자가 일정거리 안으로 오면 경보메시지를 받을 수 있는 앱이 배포되어야 합니다.

여성가족부의 경우 코로나 이전에는 8개의 토픽이 나타났으나, 코로나 이후에는 6개의 토픽들로 분류되었다. 코로나 이후 6가지 주제들을 파악하고 토픽들의 원문들을 확인한 결과, 청소년관련 정책 이슈로는 청소년 및 가족대상 상담서비스 기관 및 지원 확대, 아동관련 기관의 성범죄·아동학대 정보 및 경력조회 시행, 맞벌이 가정에 대한 육아지원서비스 확대, 성매매·자살 예방 의무교육 실시, 청소년 대상 술·담배 판매자에 대한 처벌강화 및 청소년증 의무발급, 한부모가정 및 저소득층에 대한 양육비 지원 및 혜택 확대 등이 정책의제로 도출되었다.

2) 교육부

(1) 코로나 이전

코로나 이전 교육부 데이터를 토픽 모델링한 결과 10개의 정책 이슈가 도출되었다. Topic-1은 학교, 법 행정, 급식, 관리 등의 단어의 연관 확률이 높은 것으로 나타나

‘학교 행정 및 급식 관리’와 관련된 주제로 추론된다. Topic-2는 가정, 정부, 혜택, 등록금, 지원 등의 단어 등으로 구성되어 ‘사교육 및 등록금’과 관련된 주제로 판단된다. 그리고 Topic-3은 ‘전문상담 교사 현장근무 채용’에 관한 내용이 주제로 나타났으며, Topic-4는 ‘지역 특수장애 학교 운영 및 지원’에 대한 주제로 보인다. 다음으로 Topic-5는 수업, 영어, 수학, 교과서 등이 연관 확률이 높은 단어로 나타나 ‘교과서 중심 수업 강화’가 주제로 보인다. Topic-6은 ‘대한민국 미래 및 역사 교육’을 주제로 미래, 대한민국, 역사, 교육 등의 단어가 나타났다. Topic-7은 ‘학생건강 문제와 체육 및 운동시간 확대’와 관련된 내용으로 운동, 시간, 건강 등이 연관성이 높은 단어로 도출되었다. topic-8은 ‘학교폭력 상담 및 인성교육’에 관한 내용이 주제로 나타났고, Topic-9은 ‘대학입시제도·평가전형 관련’과 관련된 내용이 주제로 보인다. 마지막으로 Topic-10은 ‘다양한 체험 프로그램 운영’이 정책 이슈로 도출되었다<표 III-7 참조>.

표 III-7 코로나 이전 교육부 LDA 분석 결과

Topic	Topic-1	Topic-2	Topic-3	Topic-4	Topic-5	Topic-6	Topic-7	Topic-8	Topic-9	Topic-10
순번	학교 행정 및 급식관리	사교육 및 등록금	전문상담교사 현장근무 채용	지역 특수장애학교 운영 및 지원	교과서 중심 수업 강화	대한민국 미래 및 역사 교육	학생 건강문제	학교폭력 상담 및 인성교육	대학입시 제도· 평가전형	다양한 체험 프로그램 운영
1	학교	가정	업무	특수	생각	일본	운동	사회	졸업	도서관
2	법	정부	학교	장애	시간	미래	시간	말	고등학교	참여
3	발생	유치원	교사	학교	수업	세계	여자	문제	입학	학습
4	기관	방	교육	지역	과목	대학	생각	부모	생각	활용
5	정보	혜택	채용	교육	교육	업무	교복	교육	공부	효과
6	운영	돈	생각	법	영어	대한민국	체육	생각	입시	필요
7	확인	학교	상담	지방	필요	문화	학교	폭력	시험	체험
8	행정	부담	근무	운영	수학	생각	건강	선생	수능	운영
9	제안	등록금	현장	사회	학년	사람	문제	학교	대학	안전
10	교육청	업무	수업	수	학교	역사	남자	인성	제도	프로그램
11	급식	생각	강사	학급	문제	사회	효과	생활	문제	교육
12	관리	자녀	임용	지원	과정	한국	환경	교사	성적	시간
13	규정	학원	평가	교육부	교과서	교육	성	학부모	대학교	활동
14	업무	부모	교원	정책	공부	국민	사용	업무	평가	봉사
15	학부모	지원	전문	정부	학습	정부	개선	상담	전형	학교

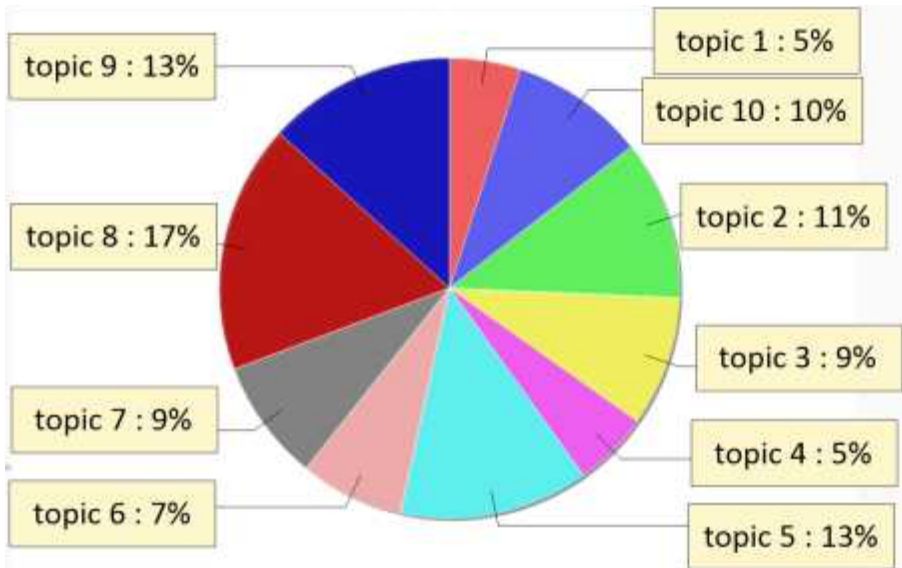


그림 Ⅲ-8 코로나 이전 교육부 토픽별 비중

코로나 이전 교육부에서 청소년 정책과 관련된 토픽은 총 10개로, 각 토픽이 차지하는 비중은 다음과 같았다. Topic-1 5%, Topic-2 11%, Topic-3 9%, Topic-4 5%, Topic-5 13%, Topic-6 7%, Topic-7 9%, Topic-8 17%, Topic-9 13%, Topic-10 10%를 각각 차지하는 것으로 나타났다.

즉 학교폭력과 인성교육에 대한 주제인 Topic-8이 17%로 가장 높은 비중을 차지하고 있는 것을 확인할 수 있으며, 다음으로 Topic-5인 교과서 중심 수업 강화가 높은 비중을 차지함으로써 해당 이슈에 대한 사람들의 관심이 매우 높음을 알 수 있다.

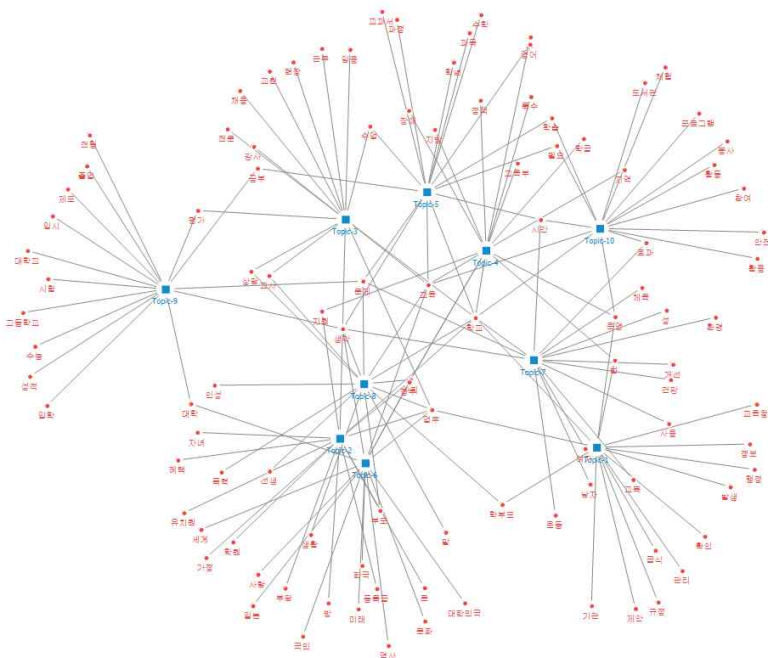


그림 III-9 코로나 이전 교육부 topic map

다음은 LDA 분석 결과 토픽별 상위 단어 15개를 topic map으로 나타낸 것이다. 코로나 이전 교육부에서 청소년과 관련된 토픽은 10개로 나타났는데 상담, 교사, 학부모, 교육, 학교, 공부 등의 단어가 여러 토픽에 동시에 등장하는 것을 살펴볼 수 있다. 이는 대부분 교육부의 역할과 관련된 내용이라 할 수 있다.

(2) 코로나 이후

코로나 이후 교육부 데이터를 토픽 모델링한 결과 6개의 정책 이슈가 도출되었다. Topic-1은 수업, 등교, 코로나, 온라인, 연기 등의 단어와 연관 확률이 높은 것으로 나타나 코로나로 인한 개학연기에 따른 ‘온라인 수업’과 관련된 주제로 추론된다. Topic-2는 ‘유해시설 인근 교육환경개선 및 보상’과 관련된 주제로 판단된다.

그리고 Topic-3은 아동, 폭력, 예방, 교원, 학대 등의 단어와 연관 확률이 높은 것으로 나타나 ‘아동 폭력 예방 교육’에 관한 내용이 주제로 판단되며, Topic-4는 코로나, 마스크, 급식, 안전, 방역 등이 주요 키워드로 나타나 ‘코로나 방역 및 안전 문제’에 대한 주제로 보인다. 다음으로 Topic-5는 ‘학교 신설 및 교육 투자’가 주제로 보이며, 마지막으로 topic-6은 시험, 과목, 평가, 수능, 학습 등의 단어가 상위에 나타나 ‘대학 수학능력’에 관한 관심이 정책 이슈로 도출되었다<표 III-8 참조>.

표 III-8 코로나 이후 교육부 LDA 분석 결과

Topic	Topic-1	Topic-2	Topic-3	Topic-4	Topic-5	Topic-6
순번	온라인 수업	유해시설 교육환경 개선 및 보상	아동 폭력 예방 교육	코로나 방역 및 안전 문제	학교 신설 및 교육 투자	대학 수학능력
1	수업	교육	학교	아이	설립	교육
2	학생	지원	교사	학교	교육부	학생
3	학교	학생	학생	코로나	축구	생각
4	개학	대학	교육	마스크	도시	사회
5	온라인	운영	업무	생각	요청	국민
6	등교	학교	아동	급식	심사	시험
7	코로나	소각장	폭력	가정	고덕	문제
8	상황	필요	법	업무	중앙	과정
9	학년	환경	예방	교실	투자	과목
10	시간	영통	경우	시간	매교	평가
11	학습	국가	활동	안전	고등학교	나라
12	생각	지역	보건	유치원	신설	필요
13	연기	가능	교원	부모	국제	수능
14	교사	장학금	보호	학원	부지	자신
15	진행	장애	행정	사용	초등학교	때
16	업무	사용	건강	방역	평택	공부
17	가능	과정	학대	문제	장현	내용
18	후	특수	피해자	상황	중학교	학습
19	경우	신청	관련	때	관련	사람

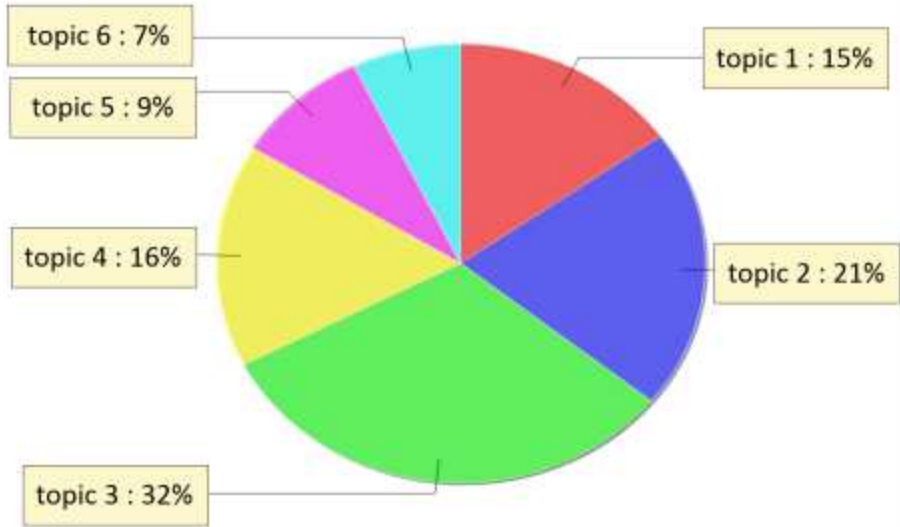


그림 III-10 코로나 이후 교육부 토픽별 비중

코로나 이후 교육부에서 청소년정책과 관련된 토픽은 총 6개로, 각 토픽이 차지하는 비중은 <그림 III-10>과 같다.

즉 아동 폭력 예방 교육에 대한 주제인 Topic-3이 32%로 가장 높은 비중을 차지하고 있는 것을 확인할 수 있으며, 다음으로 Topic-2인 교육환경개선 및 보상 문제도 21%로 높은 비중을 차지함으로써 해당 이슈에 대한 사람들의 관심이 매우 높음을 알 수 있다. 그러나 Topic-2의 경우 특정지역에 대한 요청으로 청소년 정책의제로써는 다소 부적절하다.

그 외 온라인수업과 방역문제 이슈는 코로나로 인해 새롭게 등장한 이슈로 각각 15%, Topic-4 16%를 각각 차지하는 것으로 나타났다.

표 III-9 교육부 코로나 이후 토픽 원문 예시

관련 토픽	원문
1	온라인 수업의 질 문제와 그로 인해 벌어진 교육격차, 그리고 온라인 교육 지원의 혜택을 받지 못하는 소외아동 문제 등 아직도 해결해야 할 문제가 많이 남아있습니다. 앞서 말했듯, 소외아동과 초등학교 저학년 학생 몇몇은 온라인 수업 방식에 어려움을 겪을 것입니다. 이 방과 후 프로그램은 기존의 프로그램처럼 국·영·수 등 주요 과목만 가르치는 것이 아니라 소프트웨어, 온라인 교육을 포함해야 합니다. 온라인 수업 방식을 잘 따라오고 적응할 수 있도록 기본적인 온라인 체계와 그 사용법 설명 및 체험을 중심으로 프로그램을 진행해야 합니다. 이러한 프로그램을 학교뿐만 아니라 지역아동센터나 문화센터에서도 적극적으로 시행하면 온라인 수업 방식과 스마트 기기로 인해 어려움을 겪는 소외아동들의 고충이 어느 정도 해결될 것입니다. 코로나 때문에 많은 사람들의 경제 활동에 어려움이 생겼고 현장 수업을 보내는 것도 불안하고 학교에서 하는 온라인 수업도 부실하다고 생각하는 추세이므로 방과후 수업도 수강 인원이나 부서가 많이 줄겠지만 그렇다고 개강이 되는 부서 수업의 질이 떨어지지는 않을 거라고 봅니다.
2	학습 분위기를 만드는 것은 학교의 역할이며 시험 위주보다 주도적 학교 생활의 주체가 바람직하다는 상황과 이에 대한 인식 고취가 필요합니다. 그러나 저녁 시간대, 7시에서 9시 사이에 학원을 굳이 다닐 필요가 없거나 가정 내 학습을 바람직하게 생각하는 자기 주도적 학습형 학생들에 대한 교육 서비스가 없어 폭력이나 쓸쓸함을 느낄 수 있는 여러 상황이 학교 주변에서 생기며 이는 저소득층 지역에서 특히 흔합니다. 담배, 술, 유흥, 이성교제가 저소득층 구역에서 특히 심하게 과시가 되는 이유입니다. 따라서 비대면 온라인 돌봄을 중학생 집단에 각 지역별로 실시하여 다수가 참가하여 자가 주도 학습 플랜을 짜고 일정 시간 온라인 돌봄 교사가 있는 온라인 공간에 있다면 안정감을 느끼고 학습이든 독서든 하기 쉬운 것입니다. 예를 들어 서울시 각 구청별로 실시하는 화상 영어처럼 10인 가입 온라인 방을 만들고 그 안에서 돌봄 교사와 소통하며 자기 주도 학습을 하고 이를 로그화 하는 것입니다. 이에 대해 교육부가 생기부 기록 등, 긍정적 권장 항목을 만든다면 자기 주도 학습 성향에 대해 더욱 강화하는 교육 효과를 낼 중학생 시기에 허전함을 느끼기 쉬운 저녁 시간에 일탈 방지든, 사교육 방지든 여러 긍정적 교육 목적으로 온라인 돌봄을 다양한 대학생들이나 휴직 교사 등의 고용 창출을 통해 실시할 수 있으며 서울은 서울 내 교육 가능자, 대구는 대구시 주민이 돌봄 등으로 지역 내 인프라 확보까지 추가할 수 있습니다. 온라인 돌봄은 이미 영국에서 유치원 생들을 상대로 실시되고 있으며 6살도 긍정적으로 수용하는 돌봄 형태입니다.
3	저는 학교의 주요한 기능 중 하나인 아동(학생) 보호와 행복증진을 위해서 교원양성기관(교육대학교, 사범대학)의 교육과정에 아동학대를 예방하거나, 발견 및 조치를 취할 수 있는 그런 내용을 배울 수 있는 과목이 추가되고, 교원자격시험검정을 위한 필수과목으로 지정되어야 한다고 생각합니다.

관련 토픽	원문
2	소일거리 없이 우두커니 집에서 생활하시거나 당장 조그마한 일자리라도 생계에 큰 도움이 되는 일정나이 이상의 노인(어르신)들을 이런 어린이 보호시설에 파견해서 보호 기관 측에서 아이들을 잘 케어하는 지 감시역할은 물론이고 아이들의 식사시간에 돌봐주고 간단한 일은 도움도 주는 어르신 일자리를 마련해 준다면 어떨까 생각해 봅니다 아이들은 대부분 할머니, 할아버지를 좋아하고 잘 따르며 노인들도 아이들을 좋아하는 특성을 잘 부합시키면 큰 효과를 거두리라 생각됩니다. 부모들은 아이들을 믿고 맡길 수 있고 노인들은 새로운 일자리 창출을 통해 생계에 보탬이 되고 서로 윈윈하는 효과가 있을 것이라고 생각합니다.
5	저는 시간이 오래 걸리긴 하겠지만 지금이라도 유치원에서 부터 고교과정까지 각 연령에 맞게 "가족" 또는 "사람"이라는 필수교과목을 만들어 가족이 어떻게 만들어 진건지(성인지 및 내가 가지고 있는 성기관의 자세한 이해와 사용 방법, 더불어 사랑과 책임감을 바탕으로 한 탄생이 그 결과물이란 교육, 부모가 되었을 때 부모의 역할, 그리고 다른 사람들과 소통하는 방법과 관계 맺는 방법) 가족 내에서의 구체적인 역할과 의무와 책임, 그리고 관계 맺음의 기본에 대한 제대로 된 교육, 스마트 기기에 대한 올바른 사용방법, 예절, 해서는 안 되는 것들에 대한 교육이 가치관이 정립되기 전, 규칙과 규범·질서 의식이 확립되는 유아기 시기부터 연결해 고교과정까지 국·영·수·과 보다도 더 주요과목으로 다루어져야 된다고 봅니다.
5	많은 시간을 할애하는 일은 아니고 우선 각 지역의 100명 이상 원아들이 다니는 큰 사립 유치원들을 시작으로 교육부 사이트에 등록된 시간이 맞는지 그 시간에 직접 등원하는 아이들의 수를 확인하고 방과후과정 하원시에도 아이들의 수를 확인하는 인원을 채용하여 각 유치원당 일주일에 2~3일씩 한 10군데 정도씩 지역별로 확인하면 많은 수의 부정수급을 막고 이미 이뤄진 부정 수급도 환수 할 수 있다고 생각합니다. 제가 제일 확실한 대책이라고 생각하는건 학부모인 것처럼 누군가가 유치원에 가서 직접적으로 교육시간과 등원 하원시간, 특성화 활동에 대한 모든 것을 관리 감독하는 것입니다.
5	소득구간 범위를 현실에 맞지 않아서 맞벌이 가정의 경우 네 자녀 이상(본인5자녀)임에도 아이들 배우처 교육(예:바른역사교실)등의 혜택이 전혀 없으며 세 자녀 이상 가정 대학 등록금의 경우도 혜택을 보지 못함 4자녀가구 이상(다다자녀)의 경우 소득구간 확대(현실적 반영) 및 지원 확대 4자녀 가구 이상의 경우는 소득에 대한 제한을 두지 말고 지원 확대 필요 특히 아이들 교육 배우처 우선지원 및 대학등록금 면제 등 학원비 지원 필요
5	단체복의 특성상 시즌 주문 상품이라서 주문이 없는 시기에 대량의 물품을 미리 만들어 놓았음에도 불구하고 각 유치원에서는 관에서 학부모님들께 단체복비를 못 받게 하니 (정보공시위반의 사유) 단체복을 못하겠다고 합니다.

관련 토픽	원문
5	물론 유치원 예산으로 설립자가 무상으로 단체복을 제공하는 경우도 있겠지만 대부분의 원에서는 설립자가 예산안에서 단체복 비용까지 부담하기에는 운영상의 어려움이 있는 것 같습니다. 일부 시도에서는 유치원 운영위에서 업체에 바로 주문하고 대금을 지급할 수 있다고 들었습니다. 운영위에서 결정하여 업체에 직접 체육복을 주문하고 업체에 개인적으로 송금하고 유치원에 단체 배송하는 것을 허용하여 주십시오. 위와 같은 방안으로 개선된다면 유치원은 적극적으로 단체복을 채택하는 것이 가능하고 학부모님들은 의복비 부담을 줄이며 단체복의 순기능을 더욱 극대화 할 수 있을 것입니다.
3	아동 학대 방지 교육을 포함한 바른 부모 되기 소양 교육 즉 '부모교육'이라는 교과목을 국·영·수에 버금갈 정도의 비중으로 고등학교 교과과정에 편성하여 교육하는 것이 절실하게 필요합니다. 고등학생부터는 한달 한번 일시적인 소양교육이 아닌 정식 교과목으로 '아동학대 방지'와 '바른 부모되기 교육'을 함께 가르치면 일석이조의 효과를 거둘 수 있다고 봅니다. 아동학대 방지 교육은 아동학대 근절을 위한 수동적인 교육으로 그칠 가능성이 높습니다. 그러기에 좀 더 적극적인 아동학대 방지 교육을 위해서라면 바른 부모되기 교육을 함께 가르치는 것이 요구된다고 생각합니다. 바른 부모되기 교육에서는 수동적인 아동학대 방지 교육을 비롯하여 부모로서의 기본 소양을 여러 학문적인 방식으로 접근하여 다차원적으로 교육하면 부모로서의 소양을 배양함 뿐 아니라, 사회 전반의 윤리의식을 드높임에도 기여한다고 생각합니다.
3	아동학대를 당해 신고 받고 보호 조치를 받던 아이도 전학을 가버리면 새 담임은 그전 아동학대 피해 상황을 알기 어렵습니다. 학폭 처리 내역을 생활기록부에 기록하는 것처럼 아동학대 피해 상황을 생활기록부 기본 학적 비교에 의무적으로 기록하여 학생이 어느 도시 어느 학교를 다니든지 졸업할 때까지 그 상황이 전달되어, 지속적으로 관심과 보호를 받을 수 있도록 해주세요. 아동학대 신고 후 처리 내역이 생활기록부 학적 비교에 기록되면 새 담임교사는 이를 자연스럽게 알게 되고 더욱 면밀히 아동을 관찰하여 추가 피해를 막을 수 있습니다.
3	아이들이 잘못된 행동을 했을때 교육차원이라는 명목하에 이루어지는 체벌, 학대는 절대 안 되고, 말 안듣고, 도벽, 거짓말에 대한 교육적 훈계나 방책을 위한 학부모 교육이 필요합니다. 학대를 받은 아동이 신고되어 분리된 경우 12세 이상의 아이들에게는 의사를 묻고 결정에 반영되어야 합니다. 아동센터에서 근무하는 경력자의 이야기에 따르면 아이들은 학대를 받던 부모라도 부모의 곁을 떠나는 것에 대하여 두려움과 슬픔을 가지고 나중에 성인이 되었을때 아동센터를 찾아와 원망의 말을 쏟아낸다고 합니다. 아이들을 부모가 분리시키되, 12세 이상의 아이들에게는 꼭 의사를 묻고 결정에 반영되어야 하며 센터차원의 학대방지시스템이 가동되어야 한다고 생각합니다.

교육부의 경우 코로나 이전에는 10개의 토픽으로 나타났으나, 코로나 이후에는 6개의 토픽들로 다소 감소하였다. 코로나 이후 6가지 토픽들을 파악하고, 토픽들의 원문 일부를 발췌한 내용을 살펴본 결과, 온라인수업 인프라 지원과 프로그램 및 강의 질 개선, 아동폭력 예방 교육 실시와 이를 위한 보건교원 확보, 방과후 온라인 돌봄확대 및 교육환경 개선 등이 청소년관련 정책 이슈로 도출되었다.

3) 교육청

(1) 코로나 이전

코로나 이전 교육부 데이터를 토픽 모델링한 결과 9개의 정책 이슈가 도출되었다. Topic-1은 학교, 폭력, 상담, 전문 등의 단어의 연관 확률이 높은 것으로 나타나 ‘학교폭력 전문 상담’과 관련된 주제로 나타났다. Topic-2는 학습, 건강, 운동 등의 단어 등으로 구성되어 ‘학습 건강과 운동’과 관련된 주제로 추론된다. 그리고 Topic-3은 ‘학교주변 차량 사고 방지와 통학 안전’에 관한 내용이 주제로 보이며, Topic-4는 ‘방학 중 학교 운영’에 대한 주제로 판단된다. 다음으로 Topic-5는 교사, 특수, 장애, 배치 등이 연관 확률이 높은 단어로 나타나 ‘특수교사 및 기간제 교사 배치’와 관련된 주제로 보인다. Topic-6은 ‘대학입시 고등영어 교육’을 주제로 교육, 영어, 고등학교, 대학 등의 단어가 나타났다. Topic-7은 ‘도서관, 급식 관리 지원’과 관련된 내용으로 도서관, 급식, 지원, 관리 등이 연관성이 높은 단어로 도출되었다. Topic-8은 활동, 문화, 체험, 참여, 봉사 등의 단어가 키워드로 ‘체험 및 봉사활동’이 정책 이슈로 도출되었다. 마지막으로 Topic-9는 치마, 바지, 옷, 불편, 체육복 등의 단어가 상위 키워드로 나타나 ‘여학생 교복치마 불편함 호소 및 졸업후 교복 나눔 활성화’에 관한 이슈를 보여준다(표 III-10 참조).

표 Ⅲ-10 코로나 이전 교육청 LDA 분석 결과

Topic	Topic-1	Topic-2	Topic-3	Topic-4	Topic-5	Topic-6	Topic-7	Topic-8	Topic-9	Topic-10
순번	학교폭력 전문상담	학습건강과 운동	학교주변 차량사고 방지와 통학안전	방학 중 학교운영	특수교사 및 기간제 교사 배치관련	대학입시 고등영어 교육	도서관, 급식관리 지원	체험 및 봉사활동	여학생 교복치마 불편함 호소 및 졸업후 교복 나눔 활성화	다양한 체험 프로그램 운영
1	학교	공간	학교	자녀	교사	교육	사용	활동	생각	도서관
2	선생	교실	안전	학교	교육	영어	도서관	정부	치마	참여
3	문제	효과	초등학교	업무	유치원	학습	급식	문화	고등학교	학습
4	폭력	시간	설치	시간	특수	수업	이용	학교	바지	활용
5	상담	학교	지역	수업	장애	고등학교	책	생각	옷	효과
6	교사	수업	시설	부모	지원	생각	지원	사회	출연	필요
7	생각	학습	차량	선생	학급	학교	업무	체험	불편	체험
8	교육	건강	교육청	학부모	유아	지역	활용	시간	여자	운영
9	업무	공부	사고	생각	교육청	정부	필요	사람	체육복	안전
10	말	운동	어린이	가정	학교	사회	개선	업무	졸업	프로그램
11	학부모	자율	곳	학년	배치	교육청	학부모	자원	졸업식	교육
12	전문	활동	주민	방학	강사	과정	관리	가정	착용	시간
13	사회	체육	가능	방	수	필요	운영	참여	교복	활동
14	필요	생각	발생	초등학교	기간제	대학	가정	교육	학교	봉사
15	답임	생활	통학	교실	과정	과목	학교	봉사	여학생	학교

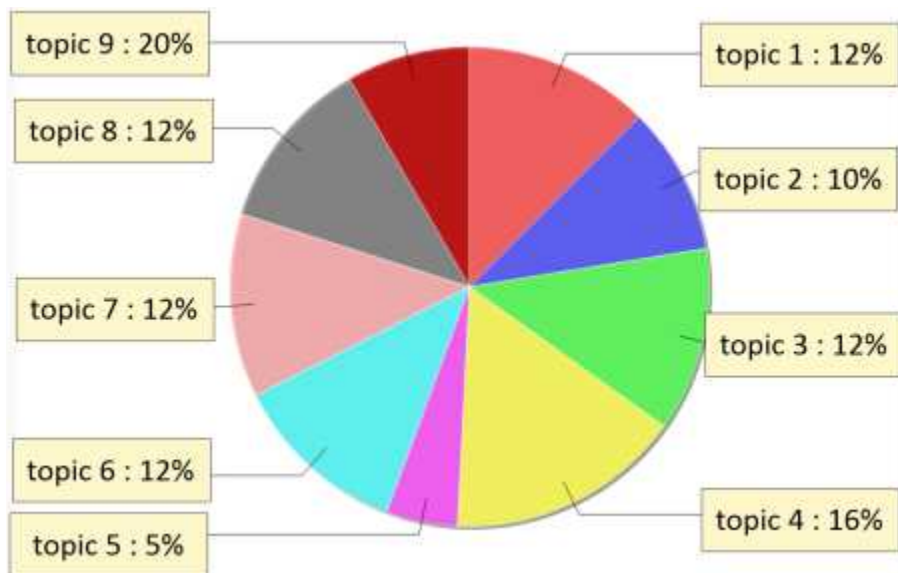


그림 III-11 코로나 이전 교육청 토픽별 비중

코로나 이전 교육청에서 청소년정책과 관련된 토픽은 총 9개로, 각 토픽이 차지하는 비중은 다음과 같다. Topic-1 12%, Topic-2 10%, Topic-3 12%, Topic-4 16%, Topic-5 5%, Topic-6 12%, Topic-7 12%, Topic-8 12%, Topic-9 20%를 각각 차지하는 것으로 각각 나타났다.

교복에 대한 주제인 Topic-9가 20%로 가장 높은 비중을 차지하고 있는 것을 확인할 수 있으며, 다음으로 Topic-4인 방학 중 학교 운영에 대한 주제가 높은 비중을 차지함으로써 교육현장에서 학생들의 생활과 밀접하게 관련되는 문제 및 실질적인 운영과 관련된 이슈가 나타나는 특징을 보이고 있다.

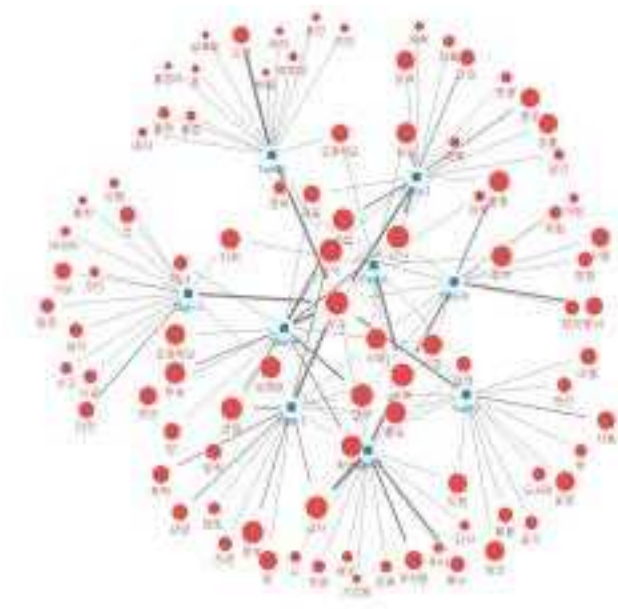


그림 III-12 코로나 이전 교육청 topic map

다음은 LDA분석 결과 토픽별 상위 단어 15개를 topic map으로 나타낸 것이다. 여러 토픽에 동시에 등장하는 단어를 제외하고 해당 토픽에만 나타난 특징적인 단어를 참고하여 토픽 명을 지정하였다.

(2) 코로나 이후

코로나 이후 교육청 데이터를 토픽 모델링한 결과 7개의 정책 이슈가 도출되었다. Topic 1은 학교, 폭력, 조치, 피해 등의 단어의 연관 확률이 높은 것으로 나타나 ‘학교폭력 조치 피해대책’과 관련된 주제로 추론된다. Topic-2는 요청, 고등학교, 초등학교, 신설 등의 단어 등으로 구성되어 ‘학교 신설 요청(특정지역)’과 관련된 주제로 판단된다. 그리고 Topic-3은 ‘코로나 돌봄 교실 운영’에 관한 내용이 주제로

나타났으며, Topic-4는 ‘과밀학교문제와 학교 신설’에 대한 주제로 보인다. 다음
 으로 Topic-5는 업무, 교사, 수업, 평가 등이 연관 확률이 높은 단어로 나타나
 ‘교사업무 및 수업평가 필요성’이 주제로 보인다. Topic-6은 ‘통학 안전문제 학교
 신설 요청’을 주제로 안전, 설치, 요청, 학교, 시설 등의 단어가 나타났다. 마지막으로
 Topic-7은 ‘유해시설 인근 학교 이전 요청’이 정책 이슈로 도출되었다(표 III-11
 참조).

표 III-11 코로나 이후 교육청 LDA 분석 결과

Topic	Topic-1	Topic-2	Topic-3	Topic-4	Topic-5	Topic-6	Topic-7
순번	학교폭력 조치	학교 신설	코로나 돌봄 교실	과밀 학교 문제	교사 수업평가	통학 안전 문제	학교 이전
1	학교	도시	지원	중학교	교육	교육청	설립
2	폭력	고덕	문의	신설	학교	안전	부지
3	배정	설립	교육	요청	업무	초등학교	지구
4	심의	국제	학교	초등학교	생각	설치	장현
5	중학교	요청	교실	배정	교사	교육	중학교
6	문의	축구	코로나	과밀	수업	요청	고등학교
7	위원회	고등학교	돌봄	요구	평가	학교	요청
8	조치	부지	운영	학급	교육청	개학	학교
9	전학	개교	방	학교	문제	연기	시흥
10	절차	매교	학원	교복	필요	부산	신설
11	가능	초등학교	수업	반대	지역	경기도	봉담
12	피해	안전	급식	문제	사용	통학	중
13	대책	육정	신청	아파트	시간	시설	소각장
14	교육청	신설	특수	철회	상황	장안	시흥시
15	요청	평택	방법	지구	사회	환경	목감

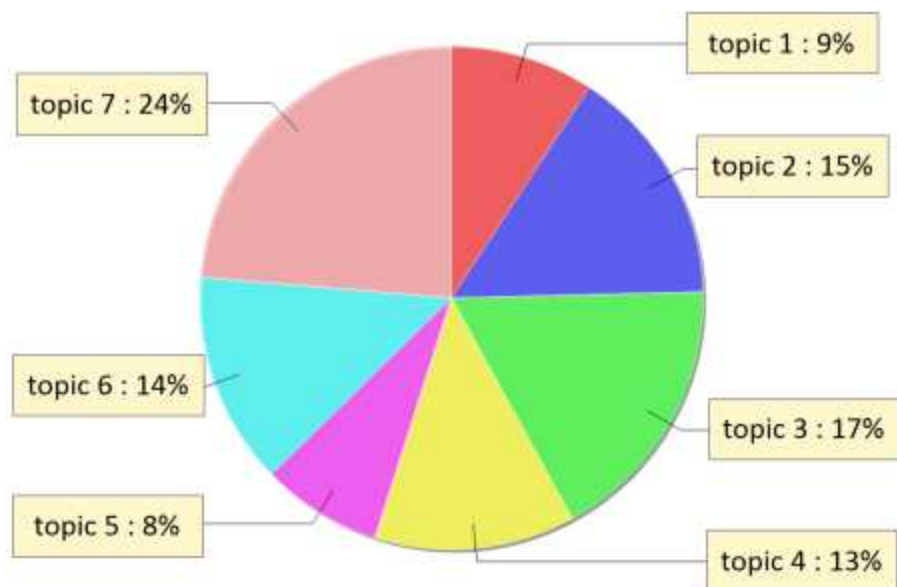


그림 Ⅲ-13 코로나 이후 교육청 토픽별 비중

코로나 이후 교육청의 청소년정책 관련 토픽은 총 7개로, 각 토픽이 차지하는 비중은 다음과 같다. Topic-2 15%, Topic-4 13%, Topic-6 14%, Topic-7 24%이며 4개의 토픽은 학교 이전 및 신설을 요청하는 내용으로 청소년 정책 이슈와는 관련성이 적었다.

주목할 점은 Topic-3인 코로나 돌봄 교실 운영에 대한 주제가 17%로 높은 비중을 차지함으로써 해당 이슈에 대한 사람들의 관심이 매우 높아진 것을 확인할 수 있었다.

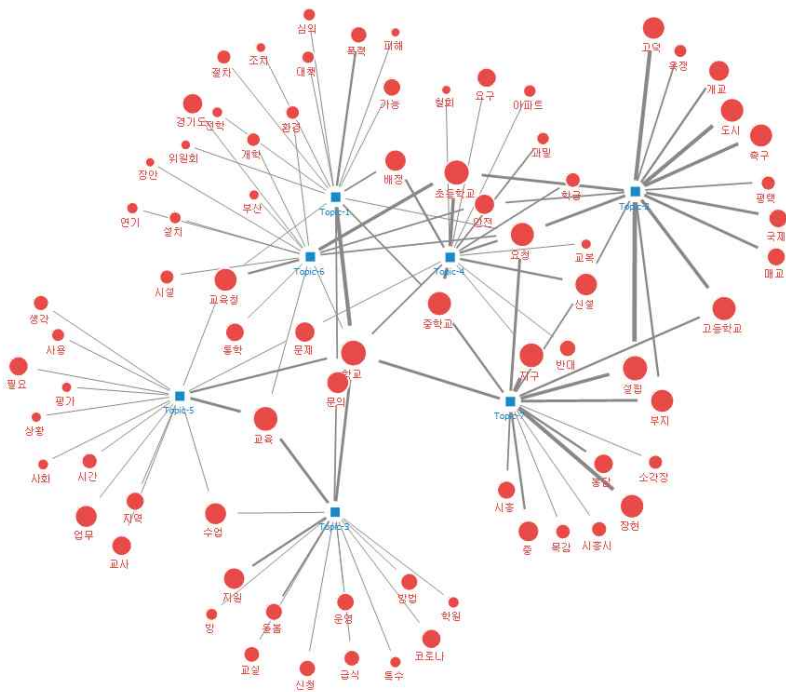


그림 III-14 코로나 이후 교육청 topic map

다음은 LDA 분석 결과 토픽별 상위 단어 15개를 topic map으로 나타낸 것이다. 코로나 이후 교육청에서 청소년과 관련된 토픽은 7개로 나타났으며, 학교, 안전, 교육, 수업, 배경, 설립 등의 단어가 여러 토픽에 동시에 등장하는 것을 살펴볼 수 있다.

표 Ⅲ-12 코로나 이후 교육청 주요 토픽 원문

관련 토픽	원문
3	돌봄교실은 모든 맞벌이 부부의 자녀들이 다 혜택을 받을 수 있는 수업이 아니더군요. 인원 수를 제한하는 것과 동시에 1학년 아이들을 우선적으로 선택하고 그 후 남은 자리를 2학년 아이들을 채우는 방식으로 하고 있습니다. 1학년 아이들을 다 채우고 남은 몇 안되는 자리.... 2학년 아이들은 "방과 후 돌봄교실" 과정을 이용 할 기회가 주어지기 힘든 상황입니다.
3	"방과 후 돌봄 교실"을 이용 할 수 없는 아이들은 부모들이 퇴근하고 오는 시간까지 돌봐줄 사람이 없어 돌봄을 위해서 사교육을 받아야 하는 상황입니다. 만약 경제적인 이유로 사교육을 받지 못 하는 아이들은 방과후 돌봄 수업 자리 부족으로 학교에서 보호 받지 못하는 상황입니다. 그리고 직장에서 일하고 있는 부모의 부재로 인해 큰 사고나 위험을 당할까 무섭습니다.
3	아동학대 경험이 있는 부모나 고위험군 가정에 대해서는 HELP 기능이 탑재된 원격교육용 스마트기기를 지원 아동이 가정 안에서 머물 수밖에 없는 환경 속에서 온라인 수업 시 접근성이 쉬운 신고와 도움을 요청할 수 있는 지원체계 마련
3	아이들의 교육은 학교에서 말고, 돌봄과 같은 보육은 지자체에서 맡게 하며 이에 대해 중앙정부의 제도적, 재정적 지원이 필요하다.
3	학생들의 등급이 조금이라도 늘어나면 온라인 수업으로만 이루어지는 것을 보완할 수 있고 요즘 급격히 심각하게 나타나고 있는 아동 방치 등의 문제를 조금이나마 줄일 수 있음을 꼭 긍정적으로 검토 바랍니다.
3	돌봄교실을 원하는 맞벌이 부부의 아이들이 모두 돌봄교실에서 수업을 받을 수 있도록 모든 초등학교에 인적, 경제적 지원을 부탁드립니다.

코로나 이후 교육청의 청소년관련 이슈의 특징은 돌봄 교실 확대 요청이 새롭게 등장한 것이다. 이는 온라인 교육으로 인한 돌봄공백 문제를 보여주는 결과라 할 수 있다. <표 Ⅲ-12>를 살펴보면 맞벌이 가정 학부모들의 온라인 수업에 대한 우려와 함께 돌봄 교실에 대한 요구가 증가하였음을 확인할 수 있다.

4) 지자체 및 기타 중앙행정기관

(1) 코로나 이전

코로나 이전 지자체 및 중앙행정기관 데이터를 토픽 모델링한 결과 12개의 정책 이슈가 도출되었다. Topic-1은 아동, 학대 범죄 등의 단어의 연관 확률이 높은 것으로 나타나 ‘아동 학대 및 범죄’와 관련된 주제로 추론된다. Topic-2는 아동, 복지, 프로그램, 청소년 등의 단어 등으로 구성되어 ‘아동복지 및 청소년 프로그램’과 관련된 주제로 판단된다. 그리고 Topic-3은 공부, 대학, 시험, 교육, 학년 등 단어가 상위에 나타나 ‘대학입시’에 관한 내용이 주제로 보인다. 다음으로 Topic-5는 ‘보육 지원’이 주제로 보육, 장애, 아동, 엄마, 유치원, 어린이집 등이, Topic-6은 ‘학생 흡연·음주·스마트폰 지도’를 주제로 흡연, 술, 스마트폰, 학생, 담배 등의 단어가 키워드로 나타났다. Topic-7은 세금, 혜택, 주택, 소득, 집 등의 연관 확률이 높은 단어로 나타나 ‘청년 정책’과 관련된 내용이 주제로 판단된다. Topic-8은 ‘경제 및 해외 경쟁력’을 주제로 해외, 정부, 국내, 경제, 세계 등의 단어가 나타났으며, Topic-9는 ‘소년법’에 관한 내용으로 소년, 피해자, 처벌, 법, 처벌 등의 단어가 나타났다. Topic-11은 ‘청년고용 및 실업정책’이 주제로 나타났으며, 마지막으로 Topic-12는 ‘어린이 안전’이 정책 이슈로 도출되었다<표 III-13 참조>.

표 III-13 코로나 이전 지자체 및 기타 중앙행정기관 LDA 분석 결과

Topic	Topic-1	Topic-2	Topic-3	Topic-4	Topic-5	Topic-6	Topic-7	Topic-8	Topic-9	Topic-10	Topic-11	Topic-12
순번	아동학대 및 범죄	이동복지 및 청소년 프로그램	대학입시	정책이슈 1	보육지원	학생출연· 음주·스마트폰 지도	청년 정책	경제 및 해외 경쟁력	소년법	정책이슈 2	청년고용 및 실업정책	어린이 안전
1	사건	지원	생각	경우	보육	출연	수당	국민	아이	정부	제도	설치
2	아이	사회	선생	정보	장애	술	세금	업무	부산	생각	근무	시간
3	처벌	센터	공부	사용	정부	경우	사람	해외	소년	문제	재용	집
4	말	교육	교사	제한	시간	스마트폰	집	청년	업무	교육	고용	때
5	아동	아동	학교	확인	아동	게임	정책	정부	생각	말	정책	어린이
6	폭력	지역	교육	관련	교사	생각	혜택	한국	사건	대통령	실업	선생
7	학대	복지	시간	법	가정	학교	주택	국내	피해자	사람	생각	아이
8	가해자	시설	대학	내용	업무	사용	가정	경제	가해자	국민	취업	말
9	업무	예산	활동	기관	생각	판매	지급	자리	폐지	사회	정부	생각
10	법	프로그램	시험	가능	엄마	때	업무	대통령	처벌	대한민국	지원	학교
11	사람	정부	수업	등록	지원	회장	정부	외국인	범죄	때	중소기업	곳
12	피해자	청소년	고등학교	해당	유치원	청소년	지원	세계	폭력	미래	청년	안전
13	경찰	사업	봉사	필요	아이	학생	돈	일본	청소년	여자	업무	업무
14	범죄	문화	학년	국민	부모	교복	소득	산업	법	인권	기업	사고
15	피해	운영	학생	발생	어린이집	담배	청년	중국	보호	자신	자리	차량

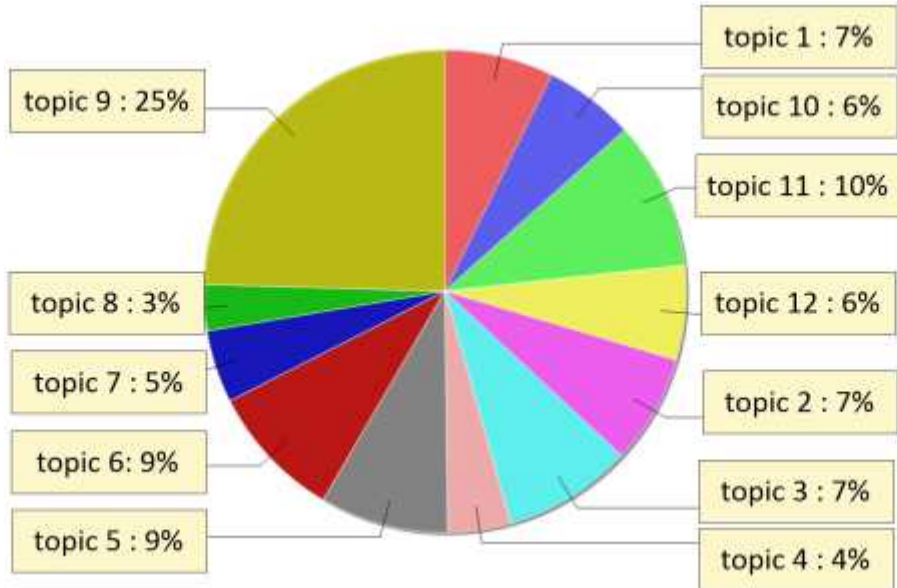


그림 Ⅲ-15 코로나 이전 지자체 및 기타 중앙행정기관 토픽별 비중

코로나 이전 지자체 및 기타 중앙행정기관에서 청소년정책과 관련된 토픽은 총 12개로, 각 토픽이 차지하는 비중은 <그림 Ⅲ-15>와 같이 나타났다. 지자체 및 기타 중앙행정기관을 모두 포함하는 데이터를 분석한 결과 다양한 이슈들이 도출되어, 기계적인 방법 만으로 즉 1차적인 방법으로 최적의 토픽 수를 산정하였다.

그 결과, 소년법에 대한 Topic-9가 25%로 가장 높은 비중을 차지하고 있는 것을 확인할 수 있어 청소년범죄와 소년법에 대한 관심이 매우 높음을 볼 수 있다. 다음으로 Topic-11인 청년고용 및 실업정책에 대한 주제가 높은 비중을 차지함으로써 코로나로 인해 청년 고용 및 실업문제가 심각해진 것을 확인할 수 있다.

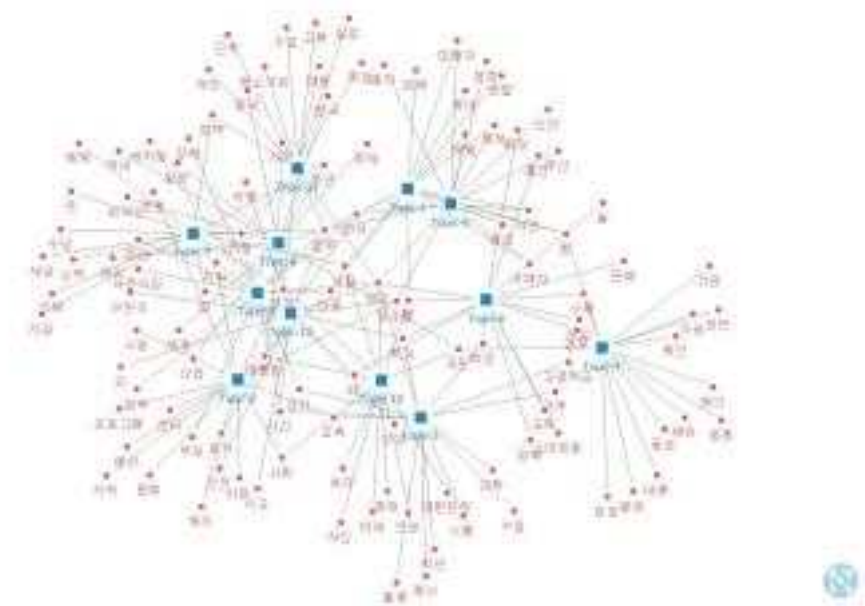


그림 III-16 코로나 이전 지자체 및 기타 중앙행정기관 topic map

다음은 LDA분석 결과를 topic map으로 나타낸 것이다. 코로나 이전 지자체 및 기타 중앙행정기관에서 청소년과 관련된 토픽은 아동학대범죄, 아동복지 및 청소년프로그램 운영, 학생 흡연·음주·스마트폰 사용 지도, 소년법 재정비 등이라고 할 수 있다. 대부분의 정책적 이슈에 대해 정책, 정부, 법, 사회, 학생, 사회 등이 등장함으로써 시민들이 정부와 사회의 역할을 중요하게 생각하고 있음을 알 수 있다.

(2) 코로나 이후

코로나 이후 지자체 및 중앙행정기관 데이터를 토픽 모델링한 결과 8개의 토픽으로 분류되었다. Topic-1은 코로나, 확진, 치료, 검사, 방역 등의 단어의 연관 확률이 높은 것으로 나타나 ‘코로나 방역 및 확진자 검사와 치료’와 관련된 주제로 추론된다. 그리고 Topic-3은 ‘청년 및 청소년 교육사업 지원 및 지역 운영 확대’에 관한 내용이 주제로 나타났으며, Topic-4는 ‘부동산 및 경제 관련’에 대한 주제로 보인다. 다음으로 Topic-5는 ‘어린이 안전사고’를 주제로 안전, 사고, 어린이, 시설, 설치, 차량 등의 단어가 나타났다. Topic-6은 ‘코로나로 인한 학교 개학 및 보육 문제’를 주제로 학교, 코로나, 개학, 보육, 어린이집 등의 단어가 키워드로 등장하였다. Topic-7는 ‘아동학대 처벌강화 및 피해자 보호’에 관한 내용이 주제로 나타났으며, 마지막으로 Topic-8은 ‘대학입시제도’가 정책 이슈로 도출되었다 <표 III-14 참조>.

표 III-14 코로나 이후 지자체 및 기타 중앙행정기관 LDA 분석 결과

Topic	Topic-1	Topic-2	Topic-3	Topic-4	Topic-5	Topic-6	Topic-7	Topic-8
순번	코로나 방역 및 확진자 검사와 치료	정책 이슈1	청년 및 청소년 교육사업 지원 및 지역운영 확대	부동산 및 경제 관련	어린이안전 사고 대비	코로나로 인한 학교개학 및 보육 문제	아동학대 처벌 강화 및 피해자 보호	대학입시제도
1	코로나	미국	청년	경제	안전	선생	사건	대학
2	마스크	생각	센터	지급	사고	학생	법	시험
3	사람	업무	활동	국민	주민	학교	아이	법
4	국민	사회	정보	정부	시설	수업	학대	업무
5	병원	세계	청소년	청년	설치	교사	아동	장애
6	업무	대통령	교육	집	생각	교육	업무	경우
7	확진	국민	지원	주변	법	아이	말	교육
8	생각	한국	참여	소득	지역	코로나	처벌	문제
9	상황	여자	국민	돈	차량	개학	경찰	기관
10	치료	나라	제안	업무	곳	어린이집	피해자	공무원
11	검사	법	사업	정책	환경	생각	폭력	사회
12	환자	사람	의견	대출	어린이	보육	청원	제도
13	아이	정부	정책	부동산	동물	상황	피해	국가
14	때	대한민국	지역	지원	도로	업무	가해자	근무
15	방역	국가	운영	사람	사용	시간	사람	학생

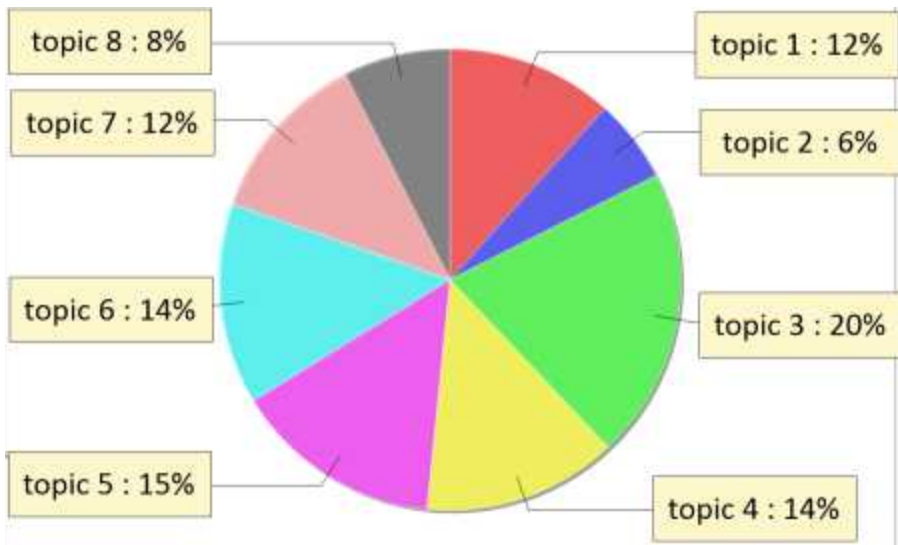


그림 III-17 코로나 이후 지자체 및 기타 중앙행정기관 토픽별 비중

코로나 이후 지자체 및 기타 중앙행정기관에서 청소년 정책과 관련된 토픽은 총 8개로, 각 토픽이 차지하는 비중은 <그림 III-17>과 같이 나타났다. 청년 및 청소년 교육사업 지원 및 지역 운영 확대에 대한 주제인 Topic-3이 20%로 가장 높은 비중을 차지하고 있는 것을 확인할 수 있으며, 다음으로 Topic-5인 어린이 안전사고 대비에 대한 주제가 15%로 높은 비중을 차지함으로써 해당 이슈에 대한 사람들의 관심이 매우 높음을 알 수 있다. 그 외 Topic-1은 12%, Topic-4 14%, Topic-6 14%, Topic-7 12% 등은 비슷한 수준으로 나타난 것을 확인할 수 있다.

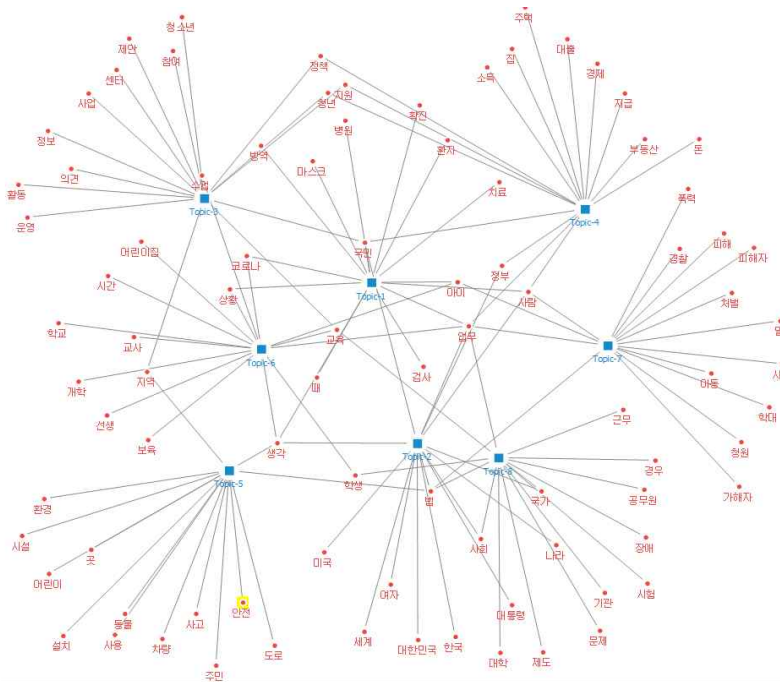


그림 III-18 코로나 이후 지자체 및 기타 중앙행정기관 topic map

다음은 LDA 분석 결과 토픽별 상위 단어 15개를 topic map으로 나타낸 것이다. 코로나 이후 지자체 및 기타 중앙행정기관에서 청소년과 관련된 토픽은 총 8개로 나타났으며, 코로나, 정부, 지원, 정책, 법, 교육 등의 단어가 여러 토픽에 동시에 등장하는 것을 살펴볼 수 있다. 따라서 대부분의 정책적 이슈에 대해 코로나가 중요한 영향을 끼치고 있다고 판단할 수 있다.

3. 코로나 전후 청소년 정책 이슈 변화

본 장에서는 코로나를 기점으로 청소년 정책 이슈 변화를 분석해 볼 것이다. 이를 위해 코로나가 국내에 처음 발생한 날로 보고된 2020년 1월 20일 기점으로 시기를 구분하였다. 총 분석 기간에서 코로나가 발생한 이후의 기간은 약 1년 6개월로 이전의 시기와 비교하기에는 짧지만, 포스트 코로나 청소년 정책 연구라는 관점에서 독립적인 두 시기로 구분하여 정책 이슈 변화를 살펴보았다.

우선 코로나 전후의 청소년 관련 토픽의 비중 변화를 살펴보았는데, 이러한 분석을 통해서 코로나 이후에 증가하거나 새롭게 등장한 청소년정책 이슈를 파악해 볼 수 있고, 역으로 코로나 이전에 비해 관심이 줄어드는 이슈를 알아낼 수도 있다. 그리고 추가적으로 특정 토픽으로 분류된 문서들을 대상으로 네트워크 분석을 실시하여 청소년 정책 이슈 변화의 특징을 알아보았다.

1) 여성가족부

코로나를 기점으로 한 기간 구분은 일수의 차이가 극명하므로 기간별 토픽 빈도 보다는 토픽의 비중 변화를 중심으로 분석하였다. 다음은 여성가족부의 코로나 이전과 이후에 도출된 토픽의 구성 비중 변화이다. 구체적으로 살펴보면, 코로나 이전 성매매 관련 토픽의 비중은 11%를 차지하고 있으나, 코로나 이후 아동성범죄 및 성매매 관련 토픽의 비중이 26%로 2배 이상 증가한 것을 확인할 수 있다. 이러한 토픽 구성 변화는 코로나 이후 아동 관련 성범죄 및 성매매 사건이 증가함에 따라 관련 정책에 대한 관심도 및 정책요구가 증가한 것으로 판단된다. 다음으로 한부모가정 및 저소득층에 대한 양육비 지원에 대한 토픽을 살펴보면 코로나 이전에 17%이었던 것이 약 2%p 증가하여 19%가 된 것을 확인할 수 있다. 이러한 증가는 코로나로 인하여 한부모가정 및 저소득층이 경제적 어려움을 겪게 된 결과로 추론

할 수 있다. 마찬가지로 상담 서비스의 확대 요구도 비슷한 수준으로 증가하였다. 학생들의 상담 서비스 확대에 대한 요구는 지속적으로 있어 왔는데, 코로나로 인해 학생들의 정상적인 학교생활이 어려워지면서 정신건강에 대한 우려가 증가한 것으로 보인다. 마지막으로 아이돌봄서비스를 포함한 육아 지원 서비스에 대한 토픽 비중도 코로나 이후 10%에서 크게 14%로 증가한 것을 확인할 수 있다. 즉 코로나로 인한 개학 연기 및 단축수업으로 보육 지원의 한계가 드러나면서 워킹맘에 대한 육아 지원 서비스 확대 필요성이 크게 증가하였다<그림 III-18 참조>.



그림 III-19 여성가족부 코로나 전후 토픽의 비중 변화

코로나 이후 비중이 증가한 정책 이슈 중에서, 양육비 지원에 관한 토픽(Topic-6) 문서를 대상으로 네트워크 분석을 실시하여 이슈의 특징을 살펴보았다. 양육비 지원 문제는 단순히 경제적 지원이 아닌 아이들의 보육 및 교육 문제와 직결되어 있음을 확인할 수 있으며, 특히 미혼모, 조손가정, 이혼으로 인한 한부모가정의 자녀 양육의 어려움이 예상된다. 따라서 다양한 형태의 한부모가정에 대한 이해와 함께 지원방식에 대한 고민이 필요할 것으로 보인다<그림 III-19 참조>.

코로나 이후 가정 보육이 증가하면서 아동학대 및 폭력 사건이 증가한 것이 원인으로 보여진다. 이에 학교에서의 폭력근절 교육 및 아동학대 가해자에 대한 처벌강화 등에 대한 요구가 크게 확대되었다<그림 III-21 참조>.

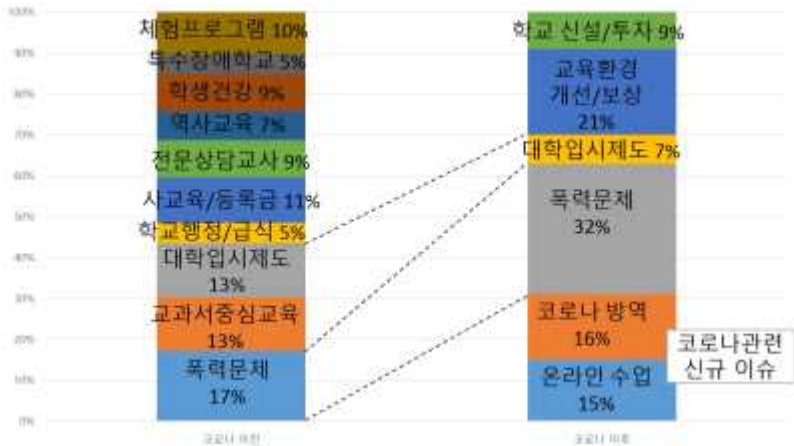


그림 III-21 교육부 코로나 전후 토픽의 비중 변화

새로운 이슈로 등장한 온라인 수업(topic-1)과 관련된 토픽 문서를 네트워크 분석한 결과, 개학 연기에 따른 온라인 수업에 관한 문제점을 살펴볼 수 있다. 저소득층의 컴퓨터 부족 문제와 수업 일수와 졸업 문제, 온라인 프로그램의 수준 및 질 문제 등이 나타나는 것을 확인할 수 있다. 주목할 점은 초등학교의 온라인 수업에 관한 것인데, 등교, 출석, 오프라인, 활동 등의 단어 링크가 많이 등장하는 것으로 보아 저학년의 온라인 수업을 지양하는 것을 확인할 수 있다. 이는 상대적으로 보육이 중요한 저학년의 경우 등교를 하지 않을 시 생기는 돌봄 공백의 문제와 연계된다고 추론할 수 있다. 즉 워킹맘들의 육아 지원 서비스에 대한 요구는 계속적으로 증가할 수밖에 없다. 반면에 고등학교는 평가, 수업, 강의, 학습 등의 단어 링크를 통하여 다른 결과를 보여주고 있다<그림 III-22 참조>.

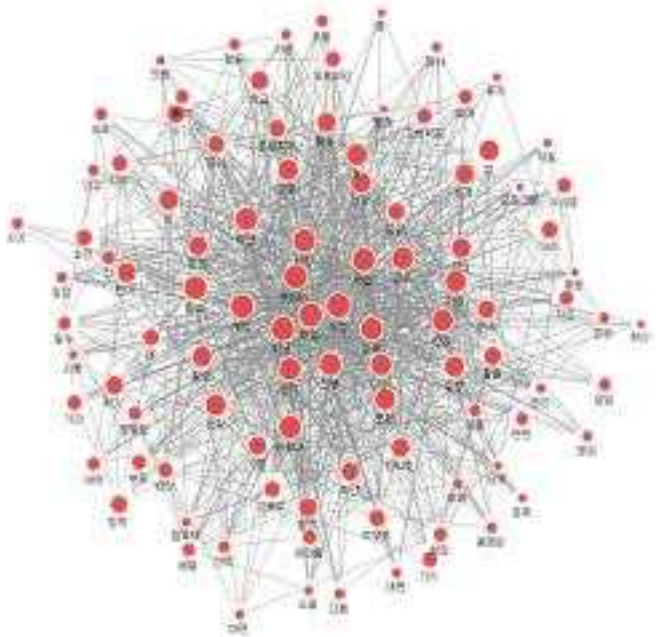


그림 Ⅲ-22 코로나이후 Topic -1 네트워크 분석 결과

3) 교육청

교육청은 앞서 두 기관과 달리 개인 및 지역사회의 문제해결을 위한 민원적 토픽이 주를 이루고 있음을 확인할 수 있다. 이는 교육청은 교육 관련 업무를 한다는 면에서 교육부와 유사하나 중앙행정기관이 아닌 각 지자체 관할이기 때문이다. 이러한 이유로 정책 이슈를 도출하는 데 다소 어려움이 있으나, 일반시민의 구체적인 요구를 기반으로 한다는 점에서 여전히 정책 이슈 도출에 시사하는 바가 있다.

코로나 이후 교육청의 청소년 관련 토픽들을 살펴보면, 근거만 다를 뿐이지 대부분 학교시설 및 이전을 요구하는 것이다. 근거로는 학교 주변 유해시설, 과밀 학급 문제, 통학 문제, 신도시 입지 등이 제시되고 있음을 확인할 수 있다. 이러한 토픽들은 코로나로 인한 이슈의 변화로 보기에 다소 무리가 있다.

그러나 코로나 이후 ‘돌봄 교실’ 토픽이 새롭게 도출된 것에 주목할 필요가 있다. 앞서 살펴본 여성가족부와 교육부의 코로나 이후 토픽에서도 아이돌봄 서비스를 포함한 육아 지원 서비스에 대한 요구가 크게 증가한 것과 마찬가지로 학교에서의 육아 지원 서비스 확대를 요구하는 사람들이 17%로 크게 증가한 것이다. 코로나 이후 학교의 돌봄 교실에 대한 필요성을 많은 사람들이 공감하고 있고, 학교의 교육적 역할 외에 방학 후 돌봄 교실에 대해 새롭게 정비할 필요성이 있는 것이다<그림 III-23>.



그림 III-23 교육청 코로나 전후 토픽의 비중 변화

추가적으로 실시한 돌봄 교실과 관련된 토픽 문서들의 네트워크분석 결과에서는 유치원과 초등학교의 등교 지연에 따른 긴급돌봄 신청에 관한 내용을 확인할 수

있다. 돌봄센터 및 돌봄 시간 강사에 대한 자격 및 프로그램의 수준에 대해서도 많은 관심을 가지고 있음을 볼 수 있다. 즉 코로나로 인해 돌봄 수요 증가와 함께 돌봄서비스에 대한 수요는 더욱 증가할 것으로 판단된다.



그림 III-24 교육청 코로나 이후 토픽 3 네트워크 분석 결과

4) 지자체 및 기타 중앙행정기관

다음은 지자체 및 중앙행정기관의 코로나 이전과 이후에 도출된 토픽의 구성 비중 변화이다. 결과를 살펴보면, 청소년과 관련된 다양한 토픽들이 도출된 것을 볼 수 있다. 그중 주목할 만한 결과는 아동학대 문제에 대한 사람들의 관심이 크게 증가한 것을 확인할 수 있다. 코로나 이전 아동학대 관련 토픽의 비중은 7%를 차지하고 있으나, 코로나 이후 아동학대 문제에 대한 토픽의 비중이 12%로

증가한 것이다. 이러한 토픽 구성 변화는 앞서 여성가족부와 교육부의 토픽 분석 결과에서도 확인할 수 있었는데 코로나 이후 돌봄 공백 및 가정 보육 증가 및 세대 내 우울감 증가 등이 원인으로 예상된다.

다음으로 코로나로 인한 방역 문제 및 확진자 치료에 관한 신규 이슈가 12%를 차지하는 것으로 나타났으며, 동시에 비슷한 원인에서 청소년의 안전에 대한 요구 및 관심도 6%에서 15%로 2배 이상 크게 증가한 것을 확인할 수 있다. 과거의 안전에 관한 이슈는 대부분 통학로에서의 안전사고 등에 관한 문제였다면, 최근의 안전에 관한 이슈는 더욱 확장되어 감염병 예방까지로 확대되는 경향을 보여준다. 코로나 이후 학교나 사회에서 청소년의 안전에 대한 새로운 개념을 정립하고 안심할 수 있는 안전 서비스가 제공되어야 할 것이다.



그림 III-25 지자체 및 중앙행정기관 코로나 전후 토픽의 비중 변화



제4장 결론 및 정책 제언

- 1. 주요 연구 결과 및 의의
- 2. 향후 연구 과제

1. 주요 연구 결과 및 의의⁹⁾

본 연구에서는 기존의 정성적인 연구 방식에서 벗어나, 스마트 거버넌스 개념을 바탕으로 한 빅데이터 분석 방법론을 통하여 청소년 정책 이슈를 파악하고자 하였다. 이를 위하여 2002년 1월부터 2021년 7월까지 청소년과 관련된 온라인 게시글 및 민원을 활용하여 코로나 이후 청소년 정책 이슈 변화와 특징을 분석하였다.

대표적인 청소년 관련 중앙행정기관을 기준으로 분석한 결과, 코로나 이후 청소년과 관련한 다양한 정책의제들이 도출되었다. 여성가족부의 경우에는 청소년 및 가족 대상 상담 서비스 지원, 성범죄·아동학대 경력조회, 육아 지원 서비스 확대 요청, 성매매·자살 예방 의무교육, 술담배 판매자에 대한 처벌강화, 한부모가정 및 저소득층에 대한 양육비 지원과 같은 총 6개의 이슈를 확인할 수 있었다. 다음으로 교육부에서는 6개의 토픽들이 나타났는데, 등교 연기 및 온라인 수업 문제, 아동 폭력 예방 교육 및 보건 교원 확대, 교육환경개선, 안전 문제, 학교 신설 및 교육 투자 확대, 대학입시제도와 같은 이슈들로 분류되었다. 그 외 지자체 및 기타 중앙행정기관에서는 부동산 및 경제 관련 이슈들이 추가로 등장하였기도 하였으나 여성가족부와 교육부와 유사한 이슈들이 나타나는 것으로 확인하였다. 중앙행정기관을 기준으로 분류하였지만, 정책 이슈는 명확히 구분되지 않고 중복되어 나타나므로, 각 기관에서 두드러지게 발생한 청소년 정책의제를 파악하는 것이 필요하다.

9) '제4장 결론 및 정책 제언'은 동아대학교 홍순구 교수와 임지원 전임연구원이 공동 작성하였다.

다음으로 코로나 전후의 토픽 비중 변화와 네트워크 분석을 통해서도 청소년정책 이슈의 흐름과 특징을 파악할 수 있었다. 코로나 이후 지속적으로 높은 비중을 차지하던 대학입시제도, 저출산·결혼문제 등의 이슈는 감소한 반면 보육 서비스, 아동학대 및 폭력 문제, 성범죄 문제, 상담 서비스 확대, 양육비 지원 등의 이슈는 증가한 것을 확인할 수 있었다. 그리고 코로나로 인한 방역 및 확진자 치료, 온라인 수업 등의 이슈는 새롭게 등장하였다. 이와 같은 결과는 코로나로 인해 개학이 연기되면서 맞벌이 및 한부모 가정의 육아 부담이 증가함에 따라 파생된 문제들로 보육 서비스와 양육비 지원 확대, 상담 서비스 등의 요구는 계속해서 증가할 것으로 보인다. 이는 과거 일반적으로 국민의 소득수준과 의식 수준이 높아짐에 따라 사회서비스의 요구가 증가하게 되는 것과는 달리 코로나로 인한 급격한 환경 변화에 기인한 것으로 정책적 접근도 달리해야 할 것이다.

이번 연구는 주관적인 해석 및 해결방안을 제시한 지금까지의 청소년 연구와는 달리, 빅데이터 분석을 통하여 청소년 관련 이슈를 도출하고 객관적인 근거를 제시하였다는 데에 의의가 있다. 선행연구를 통해서 제기한 바와 같이 기존의 소수의 전문가들을 중심으로 정성적인 연구를 통해 청소년 현황 및 정책의제를 도출하는 데에 그쳤던 한계를 극복하기 위하여, 최근 다양한 학문 분야에서 적극적으로 활용되기 시작한 ‘텍스트 마이닝’ 기법을 활용한 것이다. 온라인에서의 청소년 정책에 대한 시민들의 의견을 ‘토픽모델링’이라는 분석을 통하여 이슈를 도출하고 적용방안을 찾고자 하였다. 본 연구를 통하여 앞으로 다양한 분야에서 정책의제 도출 및 해결방안 마련에 적실성 있는 시사점을 제시할 것으로 기대한다.

2. 향후 연구 과제

본 연구에서는 비지도 학습을 이용하여 시민들의 민원 분석을 바탕으로 정책 과제를 도출하였다. 향후 청소년 정책 도출 및 평가 등을 위한 빅데이터 및 인공지능 활용 방안을 아래와 같이 제안한다. 첫째, 일부 정부 기관에서는 시민들과의 적극적 소통을 위한 도구로 빅데이터를 활용하기 위해 데이터의 수집 및 분석을 위한 거버넌스 체계 정립을 시도하고 있다. 본 보고서에서는 정부 공식 온라인 플랫폼에서의 시민들의 게시글 및 민원만을 대상으로 빅데이터 분석을 실시하였으나 향후에는 SNS, 신문 기사 등 다양한 정보원으로부터 데이터를 수집할 수 있는 채널 확보와 전처리를 위한 사전작업 등 전반적인 빅데이터 관리 체계를 구축할 필요가 있다. 둘째, 본 보고서에서는 비지도 학습인 LDA 기법으로 클러스터링을 통해 정책의제를 도출하였으나, 한 단계 더 나아가서 정책의 우선순위를 결정할 수 있도록 주요 키워드를 대상으로 추가 분석을 할 필요가 있다. 셋째, 실행된 청소년 정책 평가를 위해 감성평가 기법을 도입하여 정책수혜자나 이해관계자들의 직접적인 목소리를 분석해 볼 필요가 있다. 이를 위해서는 청소년 정책과 관련된 감성어 사전이나 지도학습을 위한 학습 데이터의 확보가 선행되어야 한다. 넷째, 업무의 효율성을 높이기 위해 빅 데이터와 인공지능 기법을 접목하여 청소년 정책의제 도출이나 평가에 대한 보고서 자동 생성 등을 실험적으로 시도해 볼 필요가 있다.



참고문헌

참고문헌

- 김상겸, 장성용(2016). 토픽모델링을 이용한 국내 산업공학의 연구동향 분석. 대한산업공학회 춘계공동학술대회 논문집, 3996-4018.
- 넷마이너(2020). 텍스트네트워크분석 교육교재. 사이람
- 송태민 (2020). 청소년 빅데이터 체계 구축 및 활용방안 연구_청소년 비만에 대한 비정형 빅데이터 연구 (연구보고 20-R08-1). 세종 :한국청소년정책연구원.
- 양연희 (2021). 토픽모델링을 활용한 공공갈등 유형 및 경향 분석. 지방행정연구, 35(2), 159-188.
- 이수상 (2016). 독후감 텍스트의 토픽모델링 적용에 관한 탐색적 연구. 한국도서관·정보학회지, 47(4), 1-18.
- 최수정 (2018). 빅데이터 감성분석을 통한 프로농구 발전 방안 모색. 숙명여자대학교 대학원 박사학위 청구논문.
- 홍성효 (2020). 청소년 빅데이터 체계 구축 및 활용방안 연구 데이터 분석 보고서 2_청소년 행복 결정요인에 대한 정형 빅데이터 연구 (연구번호 20-R08-2). 세종 : 한국청소년정책연구원.
- 홍순구, 김나람, 이태현, 한은정 (2017). 신 지역혁신모델 : 스마트 거버넌스(pp. 12-15). 유원북스.
- 홍순구, 유승의, 안순재 (2019). 도시재생정책에 대한 감성분석: 감천문화마을 방문객 리뷰를 중심으로. 한국데이터정보과학회지. 30(6), 1233-1244.

- Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4), 77-84.
- Fang, X., & Zhan, J. (2015). Sentiment analysis using product review data. *Journal of Big Data*, 2(1), 1-14.
- Griffiths, T. L. & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National academy of Sciences*, 101(1), 5228-5235.
- Inzalkar, S., & Sharma, J. (2015). A survey on text mining-techniques and application. *International Journal of Research In Science & Engineerin*, 24, 1-14.
- Kim, Y. & Lee, S. (2018). "Combinations of Text Preprocessing and Word Embedding Suitable for NErral Network Models for Document Classification." *Journal of KIISE*, 45(7), 690-700.
- Kowalski, R.M., Limber, S.P., & McCord, A. (2019). A developmental approach to cyberbullying: Prevalence and protective factors. *Aggression and Violent Behavior*, 45, 20-32
- Mimno, D. & McCallum, A. (2012). Topic models conditioned on arbitrary features with dirichlet-multinomial regression. *arXiv preprint arXiv 1*, 1-8.
- Panichella, A., Dit, B., Oliveto, R., Di Penta, M., Poshynanyk, D., & De Lucia, A. (2013). How to effectively use topic models for software engineering tasks? an approach based on genetic algorithms. 2013 35th International Conference on Software Engineering, 2013(25), 522-531.
- Pierre, J., & Peters, G. (2000). *Why the Concern with Governance Now?*, Politics and the State, St. Martin's Press, 50-69.
- Salge, T. O., & Vera, A. (2012). Benefiting from public sector

innovation: The moderating role of customer and learning orientation. Public Administration Review, 72(4), 550-559.

Spärck J. K. (1972). A Statistical Interpretation of Term Specificity and Its Application in Retrieval. Journal of Documentation, 28, 11-21.

국민신문고, www.epeople.go.kr, 2021년 8월 10일 인출.

국민청원, www.president.go.kr/petitions, 2021년 8월 10일 인출.

글로벌 과학기술정책정보 서비스, <https://now.k2base.re.kr/portal/main/main.do>, 2021년 10월 10일 인출.

넷마이너 4.0 프로그램 설명서.

http://www.cyram.net/reference_new/NetMiner_%EA%B3%B5%EC%8B%9D%EC%A0%9C%ED%92%88%EC%86%8C%EA%B0%9C%EC%84%9C_K.pdf 에서 2021년 10월 10일 인출.

Data Processing, <http://hanj.cs.illinois.edu/cs412/bk3/03.pdf>에서 2021년 10월 10일 인출.



부

록

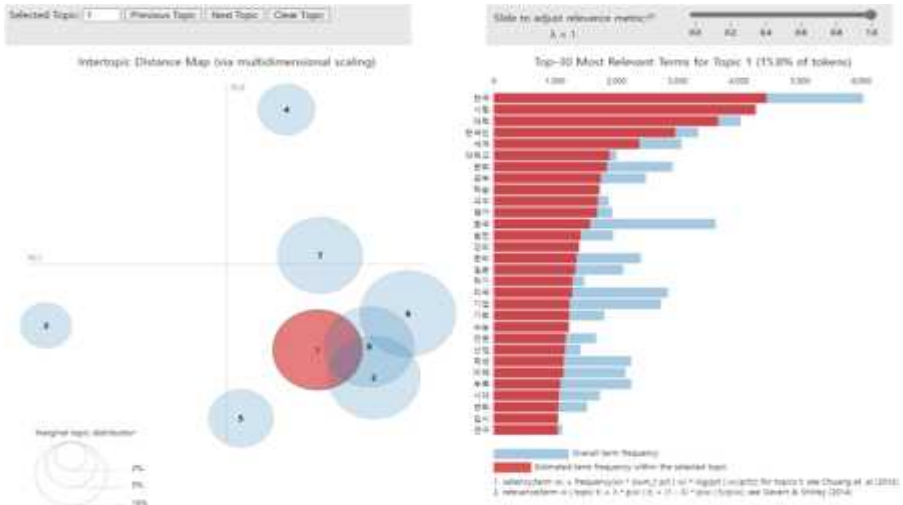
부 록

부록1. 코로나 이후 전체데이터 분석 결과

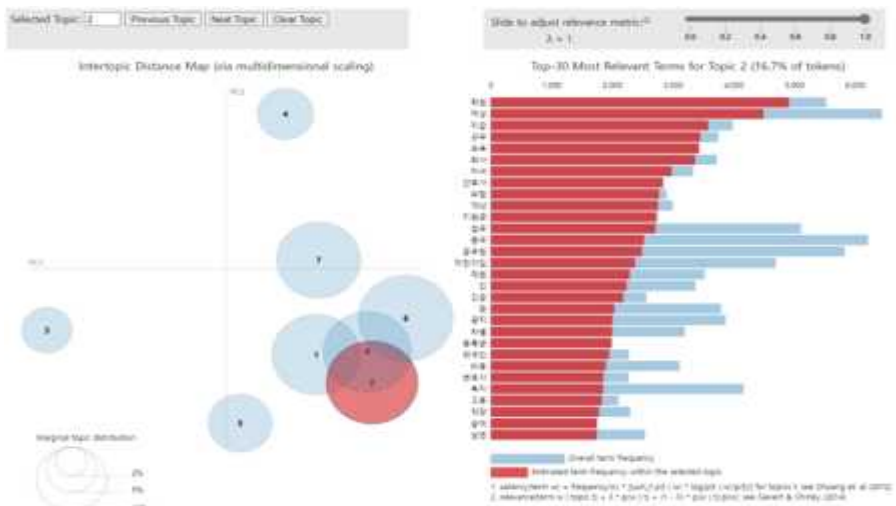
부록2. 코로나 이후 보건복지부 분석 결과

부록1. 코로나 이후 전체데이터 분석 결과

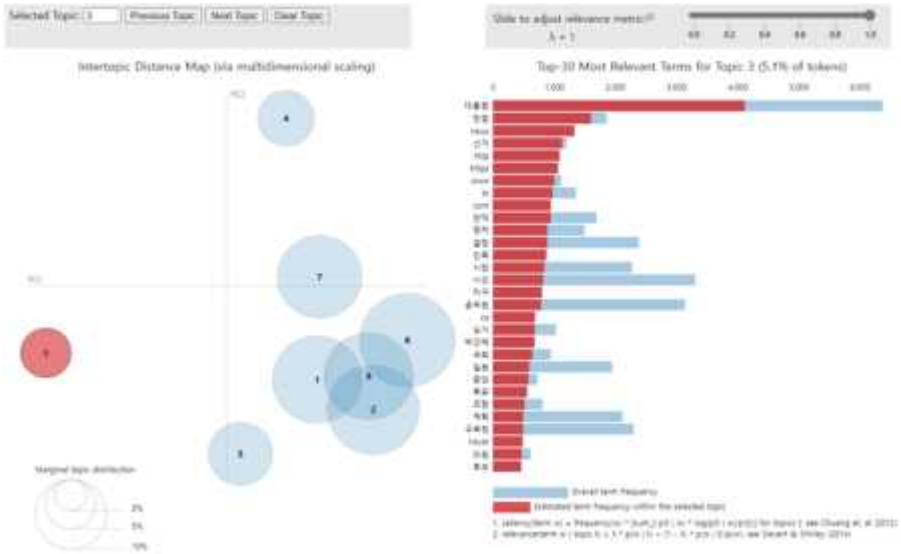
1) Topic 1



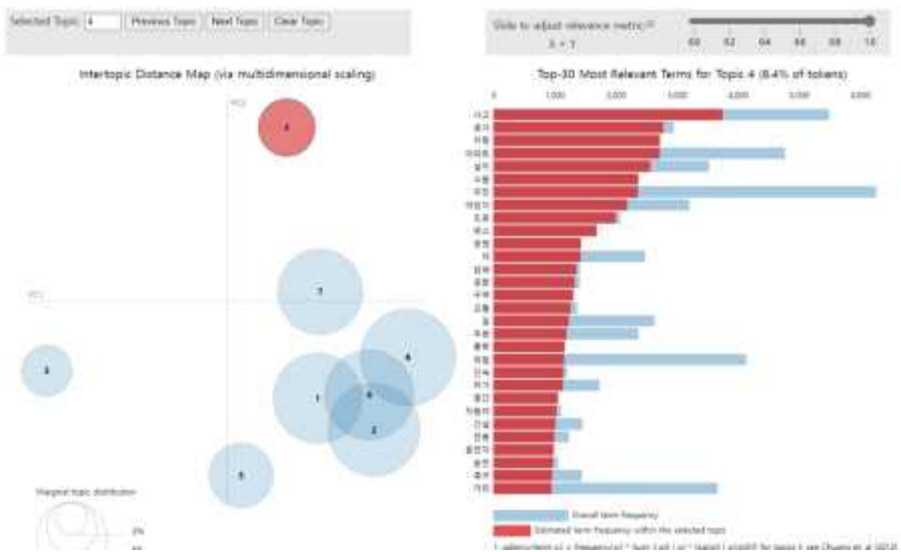
2) Topic 2



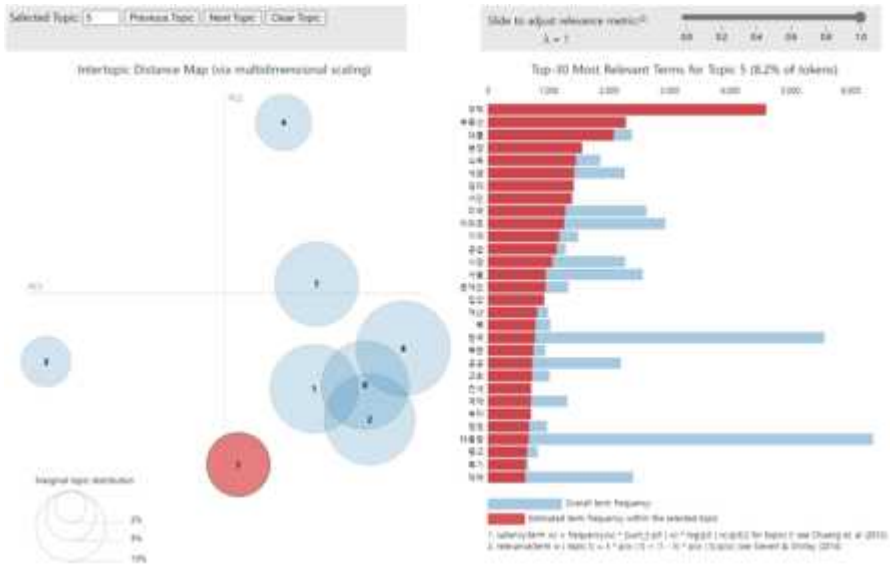
3) Topic 3



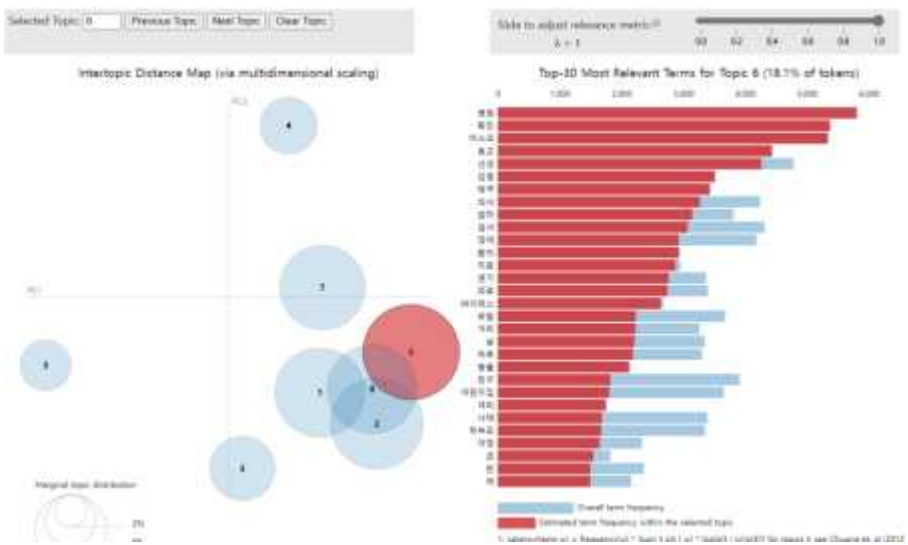
4) Topic 4



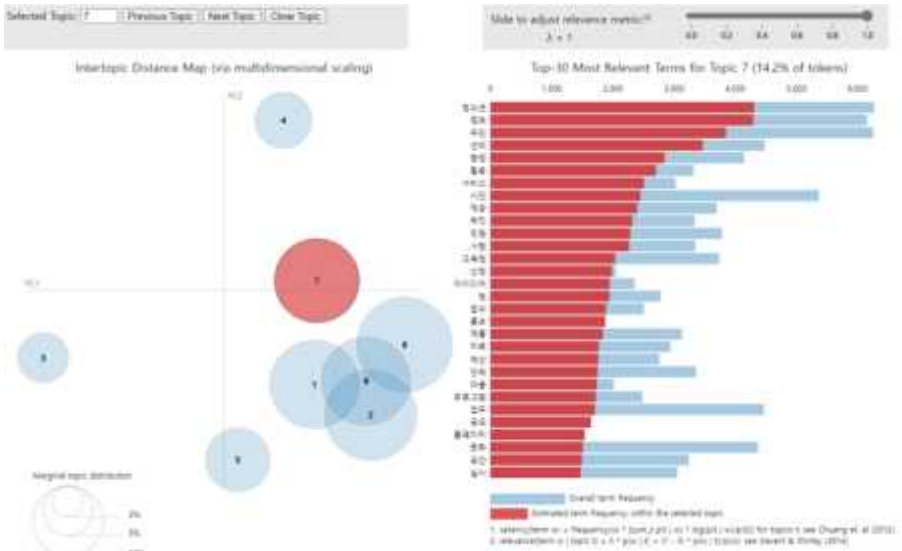
5) Topic 5



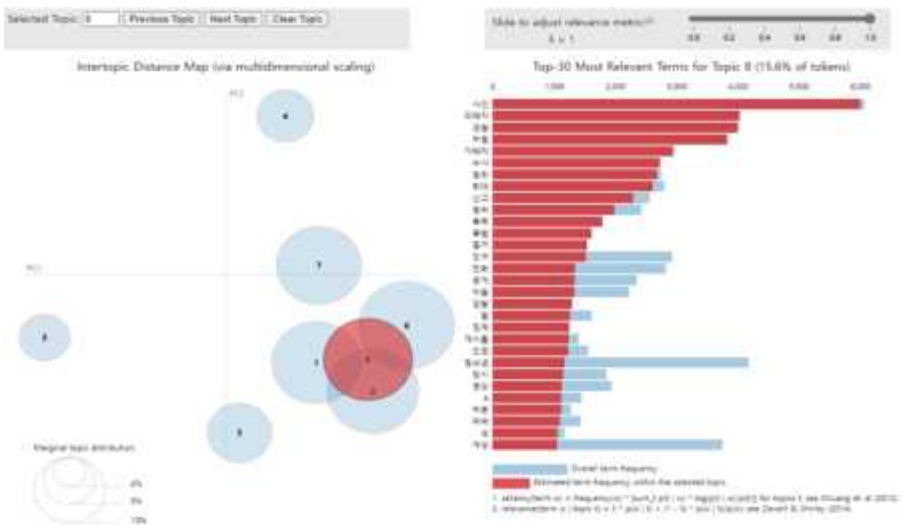
6) Topic 6



7) Topic 7

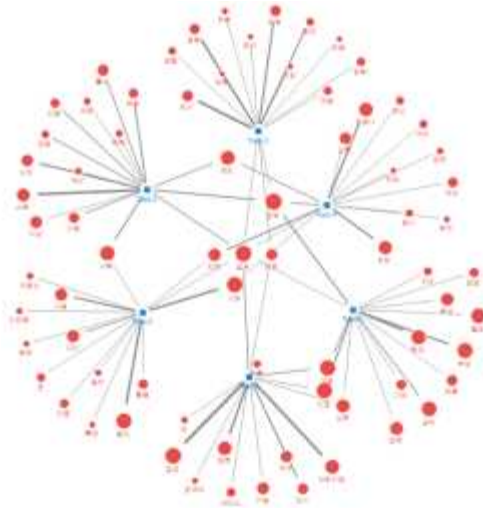


8) Topic 8



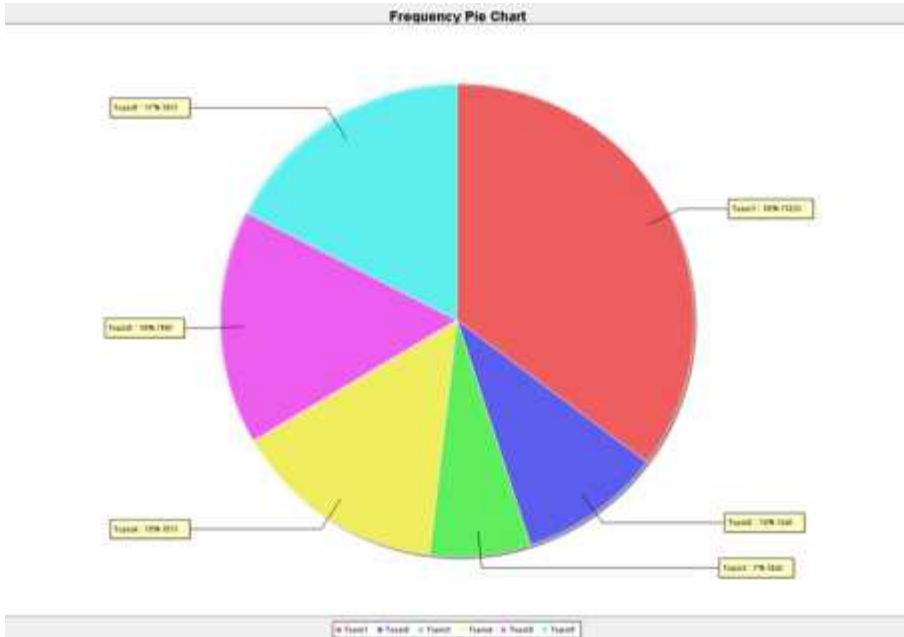
부록2. 코로나 이후 보건복지부 분석 결과

1) 코로나 이후 보건복지부 Topic-map



* 주: 분석대상 데이터 수(내용기준) : 코로나 이전 10,514개 · 코로나 이후 349개

2) 코로나 이후 보건복지부 Topic별 비중



3) 코로나 이후 보건복지부 최적 토픽 결정을 위한 실루엣 계수 값

Number	TFIDF	Topic	alpha	beta	Silhouette
1	0.25	6	0.1	0.01	0.864
2	0.25	6	0.1	0.02	0.878
3	0.25	6	0.2	0.01	0.883
4	0.25	6	0.2	0.02	0.85
5	0.25	8	0.1	0.01	0.849
6	0.25	8	0.1	0.02	0.858
7	0.25	8	0.2	0.01	0.856
8	0.25	8	0.2	0.02	0.825
9	0.25	10	0.1	0.01	0.831
10	0.25	10	0.1	0.02	0.835
11	0.25	10	0.2	0.01	0.875
12	0.25	10	0.2	0.02	0.838
13	0.35	6	0.1	0.01	0.829
14	0.35	6	0.1	0.02	0.813
15	0.35	6	0.2	0.01	0.813
16	0.35	6	0.2	0.02	0.817
17	0.35	8	0.1	0.01	0.807
18	0.35	8	0.1	0.02	0.807
19	0.35	8	0.2	0.01	0.817
20	0.35	8	0.2	0.02	0.823
21	0.35	10	0.1	0.01	0.795
22	0.35	10	0.1	0.02	0.776
23	0.35	10	0.2	0.01	0.781
24	0.35	10	0.2	0.02	0.793

4) 코로나 이후 보건복지부 LDA 분석결과 상위 15개 단어

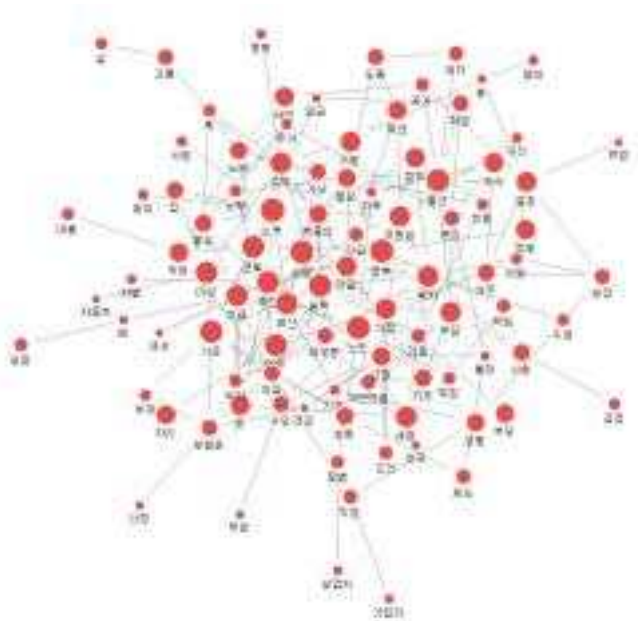
1	2	3	4	5	6
자녀학대부모 처벌 및 자녀 보호	고령화 및 저출산	건강보험 제도	코로나	아파트내 흡연문제	보육지원 및 교사업무개선
학대	소득	병원	건강	사회	장애
부모	사람	의사	지역	복지	보육
가정	노인	환자	코로나	시간	어린이집
정부	혜택	의료	국민	시설	지원
지원	정부	정부	정부	사람	교사
생각	기준	보험	상황	업무	개선
정책	출산	진료	정신	담배	업무
필요	국민	치료	사회	어르신	가능
문제	근로	제도	검사	아파트	서비스
교육	지급	의원	검진	곳	교육
도움	업무	전문	제안	지역	등록
보호	재산	지방	전국	흡연	제도
자녀	이상	업무	개인	설치	반
업무	세금	방법	업무	활동	장애인
기관	주택	회사	체크	이용	가정

5) 코로나 이후 보건복지부 트윅크 분석 결과

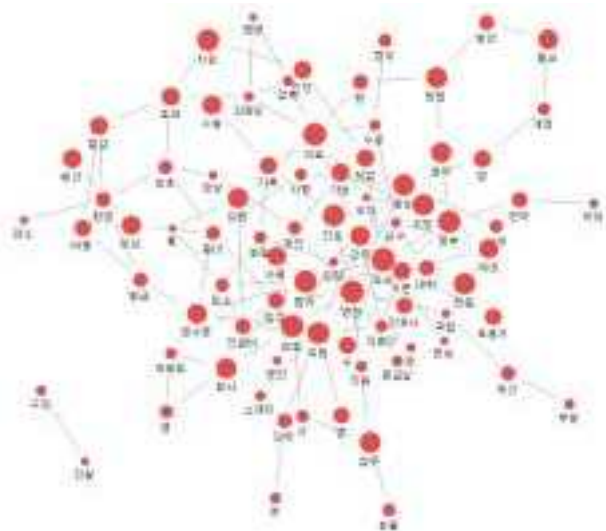
(1) Topic 1



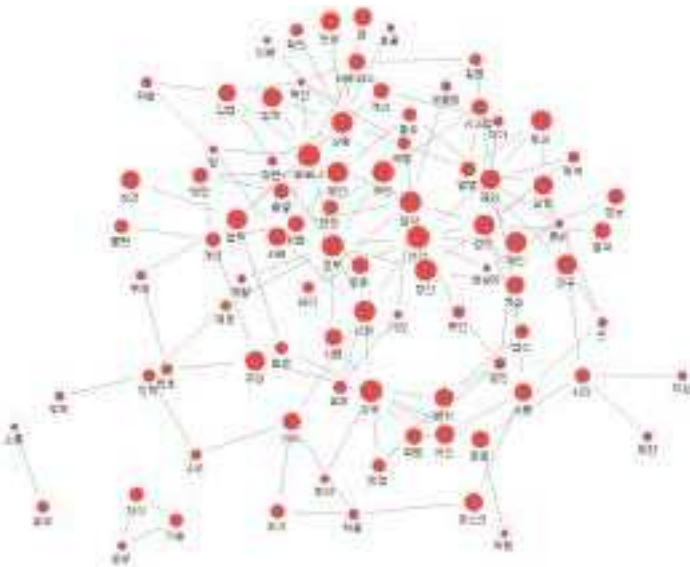
(2) Topic 2



(3) Topic 3



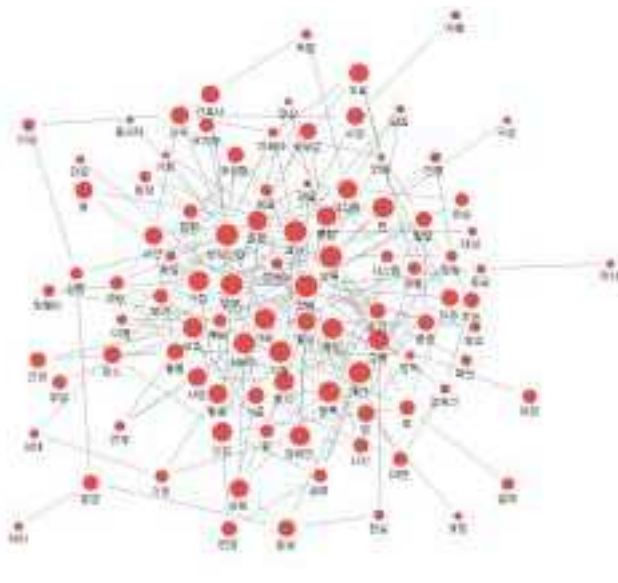
(4) Topic 4



(5) Topic 5



(6) Topic 6



국문초록

본 연구는 스마트 거버넌스 개념을 바탕으로 온라인에서 생성되고 있는 방대한 양의 비정형 민원데이터를 분석하여 포스트 코로나 시대 청소년과 관련된 정책의제를 도출하는 것을 목적으로 한다.

정부에서 운영하고 있는 국민신문고와 국민청원 사이트 내에 있는 총 9개 채널에서 2002년 1월부터 2021년 7월까지 ‘아이’, ‘아동’, ‘청소년’, ‘청년’, ‘학생’ 키워드를 내용에 포함하는 게시글 및 민원 총 265,744개의 데이터를 수집하였다. 수집된 데이터는 청소년과 관련한 여성가족부, 교육부 및 교육청, 지자체 및 기타 중앙행정기관 데이터로 재추출한 다음 코로나 이전과 이후로 구분하여 대표적인 비정형데이터 분석방법인 토픽모델링(Latent Dirichlet Allocation: LDA)을 적용하여 분석하였다.

분석 결과 여성가족부, 교육부, 교육청, 지자체 및 기타 중앙행정 기관의 토픽이 분류되었다. 주요 토픽으로는 가족대상 상담서비스 기관 및 지원 확대, 아동관련 기관의 성범죄·아동학대 정보 및 경력조회 시행, 맞벌이 가정에 대한 육아 지원서비스 확대, 성매매·자살 예방 의무교육 실시, 청소년대상 술·담배 판매자에 대한 처벌강화 및 청소년증 의무발급, 한 부모가정 및 저소득층에 대한 양육비 지원 확대, 아동폭력 예방 교육 확대 등이 있었다. 비정형 민원데이터 분석은 온라인에서 실시간으로 생성되고 있는 다양한 시민들의 의견을 정책에 수렴 가능하게 함으로써 보다 실질적인 청소년 정책의제 도출 및 수립을 가능하게 할 것이다.

ABSTRACT

A Study on Developing the Youth Policy Agendas Using Unstructured Civil Complaint Data

This study aims to draw up policy agendas related to the youth in the post-COVID-19 era by analyzing vast amounts of unstructured civil complaint data generated online based on the concept of smart governance.

From January 2002 to July 2021, a total of 265,744 posts and complaints were collected from a total of nine channels within the government-run e-People and National Petition site, including keywords such as “kids,” “children,” “youths,” “adolescence,” and “students.” The collected data were re-extracted for data from the Ministry of Gender Equality and Family, the Ministry of Education and the Office of Education, local governments, and other central administrative agencies related to adolescents. It was then analyzed, before and after the COVID-19 pandemic by applying Latent Dirichlet Allocation (LDA) as a representative unstructured data analysis method of topic modeling.

As a result of the analysis, topics from the Ministry of Gender Equality and Family, the Ministry of Education, the Office of Education, local governments, and other central administrative agencies were classified. Major topics included expanding counseling services and agencies for families, checking sex crime·child abuse information and background, increasing childcare support services for double-income families, implementing compulsory education to prevent prostitution·suicide, strengthening punishment for those who sell liquor or cigarette to the youth, mandatory issuing the youth card, expanding child support expenses for single-parent families and low-income brackets, increasing education to prevent the violence against children. The analysis of unstructured civil complaint data will enable the creation and establishment of more practical youth policy agendas by allowing the opinions of various citizens generated online in real time to be reflected in policies.

2021년 한국청소년정책연구원 간행물 안내

기관고유과제

- 21-R01 아동·청소년 대상 디지털 성범죄 현황 및 대응방안 연구 / 장근영·임지연
- 21-R02 지역사회 청소년 스포츠 활동 활성화 방안 연구 / 김정숙·정은주·정현우
- 21-R03 10대 청소년의 정신건강 실태조사 / 최정원·문호영·전진아·박용천
- 21-R04 청소년 정치참여 실태와 활성화방안 연구 / 모상현·최용환·남미자·정건희
- 21-R05 미래지향적 청소년 시설 및 공간 혁신방안 연구 / 김경준·강경균·변나향
- 21-R06 아동·청소년의 놀이 실태 및 성장지원 방안연구 / 김영한·이유진
- 21-R07 청년의 사회적 고립 실태 및 지원방안 연구 / 유민상·신동훈
- 21-R08 장애청소년 친화적 청소년정책 추진방안 연구 / 김지연·김승경·오욱찬·정소연
- 21-R09 청소년 분야 한국판 뉴딜 대응전략 연구 / 최창욱·좌동훈·성은모·남화성
- 21-R10 2021 아동·청소년 권리에 관한 국제협약 이행 연구 -한국 아동·청소년 인권실태: 총괄보고서 / 김영지·최홍일·유설희·이민희·김진호
- 21-R10-1 2021 아동·청소년 권리에 관한 국제협약 이행 연구 -한국 아동·청소년 인권실태: 심화분석보고서 / 유성상
- 21-R10-2 2021 아동·청소년 권리에 관한 국제협약 이행 연구 -한국 아동·청소년 인권실태: 총괄보고서 / 김영지·최홍일·유설희
- 21-R11 지역사회 네트워크를 활용한 청소년 성장지원정책 추진체계 구축 방안 연구Ⅲ / 황세영·이경상·송민경·조윤정
- 21-R12 포스트 코로나 시대의 청소년정책 대응 방안 연구 / 서정아·성윤숙·홍순구·권일남
- 21-R12-1 비정형 민원 데이터를 활용한 청소년 정책의제 도출 / 홍순구·임지원
- 21-R12-2 공공데이터를 활용한 청소년 정책의제 도출 / 홍성호
- 21-R13 2021 한국 아동·청소년 패널조사 / 하형석·황진구·김성은·이용해
- 21-R13-1 2021 한국 아동·청소년 패널조사: 데이터분석보고서 / 김성은·박찬호
- 21-R14 2021 다문화청소년 종단연구 / 양계민·장윤선·정윤미·한지형
- 21-R14-1 2021 다문화청소년 종단연구: 기초분석보고서(1기 패널) / 양계민·장윤선·정윤미·한지형
- 21-R14-2 2021 다문화청소년 종단연구: 기초분석보고서(2기 패널) / 양계민·장윤선·정윤미·한지형

협동연구과제

- 경제·인문사회연구회 협동연구총서 21-76-01 학교밖 청소년 지역사회 지원방안 연구Ⅳ:
질적 패널조사를 중심으로 / 김희진·서고운·김은정 (자체번호 21-R15)
- 경제·인문사회연구회 협동연구총서 21-76-02 학교밖 청소년 지역사회 지원방안 연구Ⅳ:
질적 종단자료 심층분석보고서 / 김희진·서고운·김미정·류정희·권지성 (자체번호 21-R15-1)
- 경제·인문사회연구회 협동연구총서 21-77-01 청소년의 건강권 보장을 위한 정책방안 연구Ⅲ:
후기청소년 / 임희진·황여정·유성렬 (자체번호 21-R16)
- 경제·인문사회연구회 협동연구총서 21-77-02 청소년의 건강권 보장을 위한 정책방안 연구Ⅲ:
후기청소년 - 해외사례 분석 / 임희진·박윤주·황여정 (자체번호 21-R16-1)
- 경제·인문사회연구회 협동연구총서 21-77-03 청소년의 건강권 보장을 위한 정책방안 연구Ⅲ:
후기청소년 - 기초분석보고서 / 임희진·황여정 (자체번호 21-R16-2)
- 경제·인문사회연구회 협동연구총서 21-78-01 청소년 미디어 이용 실태 및 대상별 정책대응방안
연구Ⅱ: 10대 청소년 / 배상률·이창호·김남두 (자체번호 21-R17)
- 경제·인문사회연구회 협동연구총서 21-78-02 청소년 미디어 이용 실태 및 대상별 정책대응방안
연구Ⅱ: 10대 청소년 - 청소년 미디어 교육 지원을 위한 부모 참여 중심 소셜리빙랩 운영
방안 / 정현선·조병영·권은선·김광희 (자체번호 21-R17-1)
- 경제·인문사회연구회 협동연구총서 21-78-03 청소년 미디어 이용 실태 및 대상별 정책대응방안
연구Ⅱ: 10대 청소년 - 기초분석보고서 / 배상률·이창호 (자체번호 21-R17-2)
- 경제·인문사회연구회 협동연구총서 21-79-01 청년 빈곤 실태와 자립안전망 체계 구축방안 연구Ⅰ/
김형주·연보라·정세정·변금선 (자체번호 21-R18)
- 경제·인문사회연구회 협동연구총서 21-80-01 청년 사회 첫 출발 실태 및 정책방안 연구Ⅰ:
일자리 / 김기현·배정희·김창환·성재민 (자체번호 21-R19)
- 경제·인문사회연구회 협동연구총서 21-80-02 2021년 청년 사회·경제 실태조사 - 기초분석보고서
/ 김기현·배정희 (자체번호 21-R19-1)

수시과제

- 21-R20 보호종료아동 및 자립준비청소년 자립지원 현안과 과제 / 김지연
- 21-R21 청소년 자살 원인 탐색 및 예방 대책 연구 / 서고운
- 21-R22 포스트 코로나 시대의 청소년활동 활성화 방안 연구 / 문호영
- 21-R23 청년정책연구 분야별 현황 분석 / 신동훈·김세현
- 21-R24 학교미디어교육센터 설치 유형별 지역 안착방안 / 배상률
- 21-R25 이주배경 아동·청소년 행정자료 연계 및 법제화방안 연구 / 양계민

연구개발직립금

21-R26 정책동향분석 / 이지연

수탁과제

- 21-R27 제4차 청소년보호종합대책 수립을 위한 기초연구 / 김승경·배정희
- 21-R28 청소년지도자 처우개선 및 청소년시설 운영기준 수립 연구용역 / 최용환·이희철·장혜윤
- 21-R29 청소년 자살예방 교육프로그램 효과성 분석 연구 / 김형주·강영배
- 21-R30 2021년 학교폭력 예방교육 컨설팅 매뉴얼 / 성윤숙·방인자·김유민·조아라
- 21-R31 2020 학교폭력 예방교육 어울림 프로그램 적용효과 분석 / 성윤숙·홍성호·김현수·장원빈
- 21-R32 한국 아동·청소년 삶의 질 지표 구축 방안 연구Ⅱ / 유민상
- 21-R33 2015년 개정교육과정에 따른 학교폭력 예방 어울림 교과 연계 프로그램
(중학교용-영어) / 성윤숙·임희진
- 21-R34 2015년 개정교육과정에 따른 학교폭력 예방 어울림 교과 연계 프로그램
(중학교용-기술·가정) / 성윤숙·서정아
- 21-R35 2015년 개정교육과정에 따른 학교폭력 예방 어울림 교과 연계 프로그램
(중학교용-체육) / 성윤숙·김영지
- 21-R36 2015년 개정교육과정에 따른 학교폭력 예방 어울림 교과 연계 프로그램
(고등학교용-영어) / 성윤숙·임희진
- 21-R37 2015년 개정교육과정에 따른 학교폭력 예방 어울림 교과 연계 프로그램
(고등학교용-기술·가정) / 성윤숙·서정아
- 21-R38 2015년 개정교육과정에 따른 학교폭력 예방 어울림 교과 연계 프로그램
(고등학교용-체육) / 성윤숙·김영지
- 21-R39 2020년 학부모용 학교폭력 예방교육 소식지 모음집 / 성윤숙·김옥수
- 21-R40 어울림 프로그램으로 열어나가는 학교폭력 예방교육 안내서(초등학교용) /
성윤숙·김정민·김효주·장원빈
- 21-R41 어울림 프로그램으로 열어나가는 학교폭력 예방교육 안내서(중학교용) / 성윤숙·김옥수
- 21-R42 어울림 프로그램으로 열어나가는 학교폭력 예방교육 안내서(고등학교용) / 성윤숙
- 21-R43 이주배경 아동·청소년 현황통계 생성을 위한 행정자료 연계방안 연구 / 양계민·장윤선
- 21-R44 2020년 청소년종합실태조사 / 김기현·문호영·황세영·유민상·이용해·김균희
- 21-R45 2021년 학교폭력 예방교육 어울림 학생서포터즈단 시범 운영 매뉴얼 /
성윤숙·장원빈·최동혁·김예원
- 21-R46 (초등용) 학업중단 위기 학생을 위한 비대면 꿈지락 / 황세영·송원일·배수인·유한별
- 21-R47 (중등용) 학업중단 위기 학생을 위한 비대면 꿈지락 / 황세영·송원일·배수인·유한별

- 21-R48 미디어 리터러시 교육 지원체계 구축방안 연구 / 배상률
- 21-R49 강원 2024 청소년 참여 기본계획 / 김정숙·김경준·정은주·이용해·김혁진
- 21-R50 보호대상아동 심리정서지원체계 통합구축방안 연구 / 김지연·이상정·정소연
- 21-R51 위기청소년 자립지원 표준화방안 및 평가지표 개발 연구 / 김지연·정경은·권해수
- 21-R52 인천 다문화교육 종합발전방안 연구 / 연보라·김정숙·신동훈·김재우
- 21-R53 지자체 및 중앙부처 청년정책 분야별 시행계획 평가방안 심화연구 /
배정희·장근영·이용해
- 21-R54 청소년상담사 자격제도 개선연구 / 최창욱·좌동훈
- 21-R55 청소년 참여활성화 및 권리의식 신장을 위한 제도적 기반 연구 / 좌동훈·김지수
- 21-R56 2021년 대학 성희롱·성폭력 전담기구 실태조사 / 정은주·문호영
- 21-R57 포스트코로나 시대의 성남시 청소년서비스 수요분석과 대안전략 수립 연구 /
최용환·곽창규·김민주
- 21-R58 2021 특수교육대상자 인권실태조사 / 김성은·김영지·김지연·강경균·서고운·
최홍일·임경원
- 21-R59 학교 성희롱·성폭력 사안처리 컨설팅 매뉴얼 / 정은주·임지연
- 21-R60 (2022년 개정판)어울림 프로그램으로 열어가는 학교폭력 예방교육 안내서
(초등학교용) / 성윤숙·유설희·장원빈
- 21-R61 (2022년 개정판)어울림 프로그램으로 열어가는 학교폭력 예방교육 안내서
(중학교용) / 성윤숙·유설희·김예원
- 21-R62 (2022년 개정판)어울림 프로그램으로 열어가는 학교폭력 예방교육 안내서
(고등학교용) / 성윤숙·유설희·신보람
- 21-R63 지속가능한 청소년국제교유사업 발전방안 연구 / 양계민·김민
- 21-R64 2021 경기도 다문화학생 직업교육 지원방안 / 양계민·장윤선·강희숙
- 21-R65 이주배경 청소년 실태조사 / 양계민·정윤미·장윤선
- 21-R66 학업중단 위기 청소년의 유형별·단계별 지원 방안 / 김영지·박하나·손진희
- 21-R67 학업중단 예방을 위한 지역자원 유형 조사 및 연계 방안 / 김성은·송원일·최홍일·배수인
- 21-R68 2021년 청소년방과후아카데미 효과만족도 조사 연구 / 서정아·조아미
- 21-R69 포스트코로나 시대의 한·중 청소년 생활실태 및 가치관 비교연구 / 배상률·장수·김수영
- 21-R70 닛트 등 비경제활동 청년층의 노동시장 유입을 위한 정책방안 연구 / 김기현·유민상·
배정희·신동훈
- 21-R71 청년정책 해외사례 및 시사점 / 김기현
- 21-R72 청년과 정의 연구 / 김기현
- 21-R73 이주배경 아동·청소년 지원 지역기관 연계 시범사업 모니터링, 평가 및 모델개발(2차) /
연보라·좌동훈·장윤선·김재우·오성배·장흔성
- 21-R74 2021 학교 밖 청소년 실태조사 / 김희진·임희진·김정숙·서고운
- 21-R75 다함께돌봄센터 운영 실무 개발 연구 / 황진구·좌동훈
- 21-R75-1 다함께돌봄센터 운영 실무 매뉴얼 / 황진구·좌동훈
- 21-R76 강진 청소년문화예술센터 건립 연구용역 / 김영한·이유진·선민철·이상선
- 21-R77 2021년 지역사회 청소년 성장지원 모델 운영사업 / 황세영·이경상·유민상·최홍일

세미나 및 워크숍 자료집

- 21-S01 2020 학교폭력 예방교육 어울림 프로그램 우수사례집 (1.28)
- 21-S02 2021년 학교폭력 예방교육 컨설팅단 워크숍(1.27)
- 21-S03 2020년 어울림 프로그램 운영 우수사례 시상 및 성과보고회(1.28)
- 21-S04 청년 연속기획포럼: 『2020 한국청년을 말한다』3차 포럼-한국청년, 무엇을 원하는가?(1.15)
- 21-S05 청소년정책 토론회 『코로나 팬데믹, 그리고 청소년』(2.26)
- 21-S06 2021년 학교폭력 예방교육 어울림 학생서포터즈단 발대식 및 사전교육 자료집(6.5)
- 21-S07 근거기반 이주배경청소년정책 추진방안모색을 위한 전문가 포럼(7.27)
- 21-S08 21세기 비전, 한국-몽골 청소년의 유라시아 이니셔티브(7.19)
- 21-S09 대안교실 프로그램(꿈지락) 이해 및 활용 직무연수 자료집(7.27)
- 21-S10 2021년 학교폭력 예방교육 컨설팅단 상반기 성과보고회(8.10)
- 21-S11 2021 학업중단 예방 및 대안교육 포럼: 미래 대안교육, 함께 길을 찾다-청소년에게 맞춤형 교육기회를(8.20)
- 21-S12 대안학교 및 대안교육특성화학교 역량강화 연수(8.18)
- 21-S13 2021년 학교폭력 예방교육 발전방안 포럼(8.18)
- 21-S14 2021 미디어교육 현장지원단 워크숍: 학교 미디어교육 활성화 제고 및 현장 적용 방안 수업의 실제와 사례 공 유(8.13)
- 21-S15 National Youth Policy Review: India(6.14)
- 21-S16 National Youth Policy Review: Malawi(6.21)
- 21-S17 제41회 청소년정책포럼: 미래 청소년 정책의 새로운 구상(11.5)
- 21-S18 National Youth Policy Review: Gambia(6.29)
- 21-S19 2021년 학교폭력 예방교육 어울림 학생서포터즈단 중간워크숍 자료집(8.21)
- 21-S20 제4회 한일진로교육포럼: 코로나 시대의 진로교육, '다양한 배움의 장'에서의 진로교육기회보장에 관한 한일비교 (9.25)
- 21-S21 미래를 대표하기: 핀란드 정부의 청소년 참여 정책(9.10)
- 21-S22 제12차 청년정책포럼: 청년의 사회적 고립 실태 및 대응방안(10.5)
- 21-S23 콜로키움 Mongolian Youth Policy Review 5차(10.1)
- 21-S24 콜로키움 Guatemalan Youth Policy Review 6차(10.8)
- 21-S25 Youth and Public Policy in Afghanistan(11.22)
- 21-S26 제42회 청소년정책포럼: 청소년 정치참여 현안과제 및 활성화 방안 모색(11.25)
- 21-S27 청소년정책 릴레이 토론회 자료집 '환경변화에 따른 청소년 위기상황 대응 및 보호지원 방향'(10.27)
- 21-S28 영국의 청년 고용 및 훈련 정책(11.23)
- 21-S29 청소년정책 릴레이 토론회 자료집 '뉴노멀 시대의 청소년 활동 환경 조성 방안'(11.3)

- 21-S30 청소년정책 릴레이 토론회 자료집 '뉴노멀 시대의 청소년 복지정책 방향'(11.10)
- 21-S31 직원역량 콜로키움: 실무자를 위한 예산 편성의 실제(11.22)
- 21-S32 2021 대안학교 관리자 및 담당자 역량 강화를 위한 랜선평 워크숍 자료집(11.25)
- 21-S33 청소년 정책 토론회: 청소년활동 프로그램 성인지 감수성 제고 방안(12.2)
- 21-S34 지역사회 네트워크를 활용한 청소년 성장지원 정책 추진체계 구축 시범사업Ⅲ “
시범사업기관 성과공유 워크숍 (11.24)
- 21-S35 미국 연방정부 및 주정부의 주요 청소년 정책 사례(12.3)
- 21-S36 제10회 한국아동청소년패널 학술대회(11.19)
- 21-S37 2021년 우수 대안교육 위탁교육기관 지원사업 연차평가 사전 워크숍(12.17)
- 21-S38 2021 학업중단예방 집중지원학교 및 대안교실 성과공유회(12.16)
- 21-S39 2021 대안교육 위탁교육기관 담당자 역량강화 연수(12.16)
- 21-S40 청소년 정책 토론회: 청소년 참여활성화 및 권리의식 신장을 위한 제도(12.15)

학 술 지

- 「한국청소년연구」 제32권 제1호(통권 제100호)
- 「한국청소년연구」 제32권 제2호(통권 제101호)
- 「한국청소년연구」 제32권 제3호(통권 제102호)
- 「한국청소년연구」 제32권 제4호(통권 제103호)

기타 발간물

- NYPI Bluenote 이슈 & 정책 130호 : 청소년활동의 사회적 가치 제고방안 연구
- NYPI Bluenote 이슈 & 정책 131호 : 청소년지도자 양성시스템 재구축 방안 연구Ⅱ
- NYPI Bluenote 이슈 & 정책 132호 : 국가 미래 인적자원으로서 재외동포청소년 성장과
지원방안 연구 Ⅲ
- NYPI Bluenote 이슈 & 정책 133호 : 지역사회 네트워크를 활용한 청소년 성장지원 정책
추진체계 구축방안 연구 Ⅱ
- NYPI Bluenote 이슈 & 정책 134호 : 학교 밖 청소년 지역사회 지원방안 연구Ⅲ :
질적 패널조사를 중심으로
- NYPI Bluenote 이슈 & 정책 135호 : 청소년 빅데이터 체계 구축 및 활용방안 연구
- NYPI Bluenote 이슈 & 정책 136호 : 청소년·청년의 플랫폼 노동 실태와 대응 방안 연구
- NYPI Bluenote 통계 53호 : 위기청소년 실태 파악을 위한 예비조사 결과

NYPI Bluenote 통계 54호 : 아동·청소년의 주거환경 인식조사
NYPI Bluenote 통계 55호 : Z세대 10대 청소년의 가치관 변화 연구
NYPI Bluenote 통계 56호 : 청년 사회·경제실태 및 정책방안 연구Ⅴ
NYPI Bluenote 통계 57호 : 위기청소년 건강권 실태
NYPI Bluenote 통계 58호 : 2020 다문화청소년 종단연구 2기 패널 주요 조사 결과
NYPI Bluenote 통계 59호 : 청소년 혐오표현 노출 실태 연구
NYPI Bluenote 통계 60호 : 청년 핵심정책 대상별 실태 및 지원방안 연구Ⅲ : 청년 이직자
NYPI Bluenote 통계 61호 : 10대 청소년 미디어 이용 실태
NYPI Bluenote 통계 62호 : 한국아동·청소년패널조사 2018 제3차년도 주요 조사 결과 및
데이터 분석·활용
NYPI Bluenote 통계 63호 : 아동·청소년인권실태조사

연구보고 21-R12-1

비정형 민원 데이터를 활용한 청소년 정책의제 도출

인 쇄 2021년 12월 23일

발 행 2021년 12월 31일

발행처 한국청소년정책연구원
세종특별자치시 시청대로 370

발행인 김 현 철

등 록 1993. 10. 23 제 21-500호

인쇄처 신진

사전 승인 없이 보고서 내용의 무단 전재·복제를 금함.
구독문의 : (044) 415-2125(학술정보관)

ISBN 979-11-5654-325-1
979-11-5654-323-7 (세트)

연구보고 21-R12-1

비정형 민원 데이터를 활용한 청소년 정책의제 도출

 **한국청소년정책연구원**
National Youth Policy Institute

30147 세종특별자치시 시청대로370 세종국책연구단지
사회정책동(D동) 한국청소년정책연구원 6/7층
Social Policy Building, Sejong National Research Complex, 370,
Sicheong-daero, Sejong-si, 30147, Korea
Tel. 82-44-415-2114 **Fax.** 82-44-415-2369



ISBN 979-11-5654-325-1
ISBN 979-11-5654-323-7(세트)