

교육과정평가연구
The Journal of Curriculum and Evaluation
2023, Vol. 26, No. 3, pp. 1~27
DOI: <https://doi.org/10.29221/jce.2023.26.3.1>

텍스트마이닝을 활용한 2022 개정 교육과정 정책에 대한 언론 보도의 쟁점 분석

서경숙 (이화여자대학교 교육학과 박사과정)*
김이현 (이화여자대학교 교육학과 박사과정)
이득기 (이화여자대학교 교육학과 석사수료)
이선복 (이화여자대학교 교육학과 교수)**

요약

본 연구는 2022 개정 교육과정 정책의 쟁점을 분석하기 위해 텍스트마이닝(text mining) 기법을 활용하여 뉴스 기사를 분석한 것이다. 언론매체의 관심도를 확인하기 위한 버즈 분석(buzz analysis)과 뉴스 기사의 키워드 빈도 분석 그리고 토픽모델링의 대표적인 기법인 잠재 디리클레 할당(Latent Dirichlet Allocation, LDA) 기법을 활용하여 정책의 쟁점을 분석하였다. 버즈 분석과 키워드 빈도 분석 결과, 첫째, 국내 주요 언론매체의 관심도는 정책 형성 단계의 초·중반에는 관련 보도가 거의 없다가 교과별 공청회 개최 전후로 급상승하였다. 둘째, 국내 주요 언론매체는 '고교학점제'와 '역사 교과서의 자유민주주의 용어' 논란에 집중한다는 점이 확인되었다. LDA 분석 결과, 2022 개정 교육과정 정책의 쟁점은 ① 고교학점제 전면 도입, ② 역사 교과에 '자유민주주의' 용어 논란, ③ 미래 지향적 교육과정과 디지털 환경, ④ 선택교과 확대와 국·영·수 시수 축소, ⑤ '국민참여소통채널'을 통한 국민 의견 접수, ⑥ 국가교육위원회 출범 늦어져 교육 정책 추진 차질 우려, ⑦ 초·중등 정보교육 시수 확대하고 코딩교육 필수, ⑧ 게임 섯다운제 폐지하고 시간선택제로 전환으로 집계되었다. 이 연구는 텍스트마이닝 분석 기법을 활용하여 정책 단계별로 2022 개정 교육과정 정책의 쟁점을 살펴보았으며, 빅데이터 분석의 유용성을 확인하였다는 데 의의가 있다.

주제어: 2022 개정 교육과정, 텍스트마이닝, 토픽모델링, LDA, 언론 보도, 뉴스 기사

* 제1저자, ksseo7575@gmail.com

** 교신저자, sunboklee@ewha.ac.kr

I. 서론

2015 교육과정 고시 이후 급속한 사회 변화 속에서 가장 두드러지는 것은 인공지능, 빅데이터, 사물 인터넷의 기술을 접목한 4차 산업혁명 시대로의 진입과 코로나19로 인한 지속가능발전에 대한 대응의 중요성이라 할 수 있다. 이는 미래 사회의 변화에 대응하기 위한 교육개혁의 요구를 가속하였으며, 국가 차원에서 이루어지는 교육과정 개정에서 고려할 중대한 일이라고 할 수 있다(교육부, 2022).

교육부는 이런 사회의 변화를 고려한 교육과정 개정을 논의하였으며, 2021년 11월 14일에 2022 개정 교육과정 총론 주요 사항을 발표하였다. 이 총론 주요 사항(교육부, 2021d)에서 제시하는 교육과정 개정의 배경은 디지털 대전환과 기후환경 변화, 인구구조 변화 등 예측할 수 없는 미래 사회 변화에 대응할 수 있는 교육혁신의 필요성, 학령인구 감소와 다양한 학습자 성향에 따른 맞춤형 교육에 대한 요구의 증가, 새로운 교육 환경에 적합한 역량 강화 교육, 그리고 현장에서 수용할 수 있는 교육과정 요구에 따른 미래 교육 비전과 수업 및 평가 개선을 강조하고 있으며, 이를 통해 교육과정 체제 전환의 필요성을 제시하고 있다.

한편, 교육부는 2022 개정 교육과정을 ‘국민과 함께하는 교육과정’으로 개발 추진하고자 하였다. 이는 교육과정 개발에 국민의 참여를 증대시켜 정책수요자들의 사회적 합의를 통한 최선의 대안을 설정해나가는 과정으로 의미가 있다(교육부, 2021a). 또한, 교육 주체인 국민과 전문가 등의 참여를 통해 현장 수용성이 높은 교육과정을 개발하고, 미래 교육에 대한 이해와 관심을 높여 교육 패러다임 전환을 이끌고자 하였다. 이를 위해 국가 차원에서 교육과정 전문가 중심으로 개발되던 초·중등학교 교육 과정에 교사, 학부모, 학생 그리고 범사회적 전문가 등 교육 주체들의 참여를 확대하고, 국가교육회의와 전국 시도교육감협의회 등 관련 기관과 협력하여 개발하도록 계획하였다. 이에 따라 교육부는 국민과 함께하는 교육과정이라는 개발 취지에 따라 ‘국민참여소통채널’을 통해서 총 7,860건의 의견을 모아 각 교과 정책 연구진에 전달하였다(교육부, 2021b).

또한, 2022 개정 교육과정에 대한 다수의 연구가 진행되어 왔다. 예컨대, 조용과 홍후조(2022)는 중학교 현장 운영 경험자의 입장에서 2022 개정 교육과정 총론 주요 사항을 내용 분석을 하였으며, 홍원표, 이광우, 임유나(2022)는 2022 개정 교육과정 총론 주요 사항에 담긴 내용을 중심으로 고등학교 현장 교원들의 인식을 설문 조사하여 비교 분석하였다. 김수영(2023)은 2015 개정 교육과정과 2022 개정 교육과정을 총론의 형식과 내용 측면에서 비교 분석하여 개선된 점을 확인하였다. 이 연구들은 주로 학교급별 현장 교사의 입장에서 쟁점을 검토하거나 2015 개정 교육과정과 비교하는 등 2022 개정 교육과정의 세부 내용을 다루는 연구를 하였다는 것에 의미가 있다. 그러나 2022 개정 교육과정 정책을 효과적이고 타당하게 집행하기 위해서는 관련 쟁점들을 파악하는 것이 중요하다고 할 수 있다.

본 연구에서 활용한 연구 대상인 언론매체의 뉴스 기사는 정책 의제 설정에 있어서 역할이 중요하다고 할 수 있는데, 이 때문에 정책 분석에서 중요한 자료로 활용된다(McCombs & Shaw, 1972). 특히, 정부 정책에 대한 특징을 파악할 때, 여론 추이 분석을 하게 되면 전수조사를 하는 것처럼 전반적인 내용을 확인할 수 있으므로 매우 유용하다(김용대, 조광현, 2013). 언론의 특성 중 더 많은 일반 대중이

관심을 두고 지켜보는 사회문제에 대하여 더 빈번하고 다양한 시각에서 정보를 생산 및 전달하는 특성으로 인하여, 집중된 언론의 보도를 통해 발전한 사회적 이슈는 정책 문제로 발전할 가능성이 높다. 이에 작은 이슈가 언론 보도에 의해 사회적 이슈로 발전하게 되면 일반대중들은 그 문제의 해결을 요구하게 되며, 정부가 해결책을 찾는 정부의제로 발전하는 경향이 있다고 설명되기도 한다(목진휴, 박기묵, 1993). 사회구성원들의 인식과 관련 내용들을 파악함에 있어서 뉴스 기사를 분석하는 것은 유용하다고 할 수 있는데, 뉴스 기사는 정책 관련 내용을 독자들에게 전달함으로써 해당 정책에 관심이 있는 이해관계자들에게 영향을 미친다. 이를 통해 여론이 형성되면 언론매체는 뉴스 기사들을 통해 관련된 이슈와 논의를 다루거나 사실을 통해 의견을 제기함으로써 정책 형성과 결정 단계에 영향을 미친다는 것을 알 수 있다(김병주, 김은아, 홍준영, 2006).

대규모 뉴스 기사인 텍스트 데이터로부터 의미 있는 정보와 지식을 추출하고 개념 간의 연결성을 시각화하여 의미를 도출하기 위하여 텍스트마이닝 기법이 유용하다고 할 수 있다(김선아 외, 2016). 이와 관련하여 교육 분야에서 텍스트마이닝을 활용한 연구를 살펴보면, 소셜미디어 데이터를 수집하여 교육 정보화 정책을 분석하고 관심 추이와 각 토픽에 대한 감성분석을 수행한 연구(김우주, 구찬동, 2015)와 언론 기사, 교육부 보도자료, 학술지 초록의 텍스트를 활용하여 자유학기제 정책을 분석한 연구(신안나, 2019)가 있다. 그뿐만 아니라 수능 절대평가 시안에 대한 언론 및 여론의 방향을 토픽모델링을 활용하여 분석하고 뉴스 기사의 댓글까지 분석하여 여론이 제도 개편에 '우려'와 '공정성'에 관한 불만이 많음을 도출한 연구도 있다(권순보, 유진은, 2018).

따라서 뉴스 기사는 정책 내용과 함께 문제점 및 해결 방안 등에 대한 다양한 관점을 종합적으로 평가하는 데 유용하므로 텍스트마이닝을 활용하여 국내 주요 언론매체의 보도 기사를 통해 나타난 2022 개정 교육과정 정책의 쟁점을 분석하고자 한다. 이를 위해, 국내 방송사를 포함한 주요 언론매체 54개에서 보도한 2022 개정 교육과정 정책에 대한 뉴스 기사 전체를 웹 크롤링 기법을 이용하여 수집하였으며 버즈 분석과 키워드 빈도 분석을 통해 해당 정책에 대한 언론매체의 관심도와 중요도를 분석한 후, 토픽모델링의 잠재 디리클레 할당(latent dirichlet allocation, LDA)을 활용하여 2022 개정 교육과정 정책의 주요 쟁점들을 정책의 형성과 결정 단계별로 비교 분석하였다.

이 연구의 연구 문제는 다음과 같다.

- 첫째, 2022 개정 교육과정 정책에 대해 언론매체의 관심도의 변화는 어떠한가?
- 둘째, 2022 개정 교육과정 정책에 대해 언론매체의 보도 기사에 나타난 키워드는 무엇인가?
- 셋째, 2022 개정 교육과정 정책에 대해 언론매체의 보도 기사에 나타난 주요 쟁점은 무엇인가?

II. 이론적 배경

1. 2022 개정 교육과정 정책의 도입과 전개 과정

2017년 문재인 정부 출범 이후 100대 국정과제의 하나로 고교학점제가 포함되어 2018년 8월 교육부는 「고교교육혁신 방향을 통한 고교학점제 추진 단계」를 발표하여 2021년까지 고교학점제 도입 기반을 마련하고 2024년까지 고교학점제 제도를 부분적으로 도입하고, 2025부터는 고교학점제를 본격적으로 시행하는 것을 목표로 삼았다(교육부, 2021c). 이는 2022 개정 교육과정의 출발로 여겨졌다. 2021년 2월에는 ‘포용과 성장의 고교교육 구현을 위한 고교학점제 종합 추진계획’을 발표하고, 이를 기반으로 같은 해 4월 ‘국민과 함께하는 미래형 교육과정’의 논의를 본격적으로 시작하여 2022 개정 교육과정 추진계획을 발표하였다.

2022 개정 교육과정의 주요 특징은 대국민 의견 수렴 체계와 협력적 관리체계(거버넌스)를 구축하여 정책 연구에 반영하는 것이다(교육부, 2021d). 교육과정 개정을 위해 교육계, 학부모, 미래학자, 생태환경, 공간 혁신 등 전문가로 구성된 ‘개정추진위원회’와 교육과정 개발 정책연구팀이 협력하여 주요 사항을 개발 및 조정하였다. 또한, 교육계, 공학 분야, 인문 사회 분야 등의 범사회적 전문가로 구성된 ‘정책자문위원회’를 통해 사회적 요구를 반영하고 개정안을 자문하였다(교육부, 2021d). 대국민의 의견 수렴을 위해 10만여 명의 설문조사를 실시하고, 학생과 학부모 공감 및 소통 콘서트를 개최하여 정책이해관계자들의 의견을 수렴하였다(교육부, 2021a). 교육 현장의 목소리를 듣기 위해 권역별 핵심 교원 연수, 교원단체 관계자 간담회, 현장 소통 포럼을 개최하고 정책연구팀 주관의 공청회를 통해 총론 주요 사항 관련 연구 결과를 공유하고 개정 방향을 논의하였다(교육부, 2021c). 이후 ‘교육과정 심의회’에 학생 특별위원회와 교육과정 특별위원회를 신설하여 다양한 의견을 수렴하였고, 2021년 11월에는 2022 개정 교육과정 총론의 주요 사항을 발표하였다. 이와 함께 정책 연구진은 2022년 9월에 교과별 공청회를 개최하고, ‘각론조정위원회’와 ‘개정추진위원회’ 등의 관련 위원회와 논의를 거쳐 총론 및 각론 행정예고를 진행하였다. 총론 주요 사항 발표 이후, 약 1년 동안 시안 개발 및 초·중·고등학교 교육과정 운영 방안 연구한 후, 국가교육위원회에서 심의와 의결을 거쳐 2022년 12월 22일에 2022 개정 교육과정을 확정·고시하였다.

정책의 단계는 여러 학자마다 조금씩 정의가 다르지만 대부분 정책 형성에서 정책 결정, 정책 집행 순으로 전개된다. 정부가 현재 당면하고 있는 사회문제를 정책의제(policy agenda)로 채택하여 대안을 탐색 및 개발하는 과정을 정책 형성으로 정의한다(김선정, 2014). 정책 형성 과정에서 나타난 여러 대안을 비교·평가하여 정책의제가 정책으로 확정되어 가는 과정은 정책 결정이다(이종수, 2009). 2022 개정 교육과정 정책은 현재 정책 형성, 정책 결정 과정을 거쳐 정책 집행 단계에 있으므로 정책 결정 단계까지만 구분하여 분석하였다. 이로써 학점제 도입 기반을 마련하기 시작한 2018년 8월 17일부터 2022 개정 교육과정 총론 주요 사항 시안 발표 전인 2021년 11월 23일까지를 정책 형성 단계, 총론 주요 사항 시안 발표일인 2021년 11월 24일부터 총론과 각론 확정 고시한 2022년 12월 22일까

지를 정책 결정 단계로 구분하였다. 정책의 단계별 주요사안은 <표 1>과 같다.

<표 1> 2022 개정 교육과정 정책의 단계 정의와 단계별 주요사안

정책 형성 단계 (2018.08.17. ~ 2021.11.23.)	
정의	현재 당면하고 있는 사회문제를 정부가 정책의제로 설정하여 대안 탐색 및 개발하는 단계
주요사안	<ul style="list-style-type: none"> • 2018.08.17. 「고교교육혁신 방향을 통한 고교학점제」 추진 발표 • 2021.02.17. 「고교학점제 종합추진계획」 발표 • 2021.04.20. 「국민과 함께하는 미래형 교육과정 개정 추진 계획」 발표 • 2021.04. '개정추진위원회, 정책자문위원회, 각론조정위원회' 구성 및 운영, '국가교육과정 정책자문위원회'를 통한 사회적 요구 반영 및 개정안 자문 • 2021.05.~09. 대국민 설문조사 및 '학생·학부모 공감 & 소통 콘서트'를 진행하여 권역별 학생·학부모 의견 수렴의 장 마련 • 2021.09.~10. '2022 개정 교육과정 총론 주요 사항 마련을 위한 공청회'를 통해 교육과정의 개정 방향에 관한 논의
정책 결정 단계 (2021.11.24. ~ 2022.12.23.)	
정의	제시된 여러 대안을 비교·분석하여 정책의제를 정책으로 확정해 가는 단계
주요사안	<ul style="list-style-type: none"> • 2021.11.24. 「2022 개정 교육과정 총론 주요 사항(시안)」 발표 • 2022.09.28. 교과별 공청회 개최 • 2022.11.09. 총론 및 각론 행정예고 • 2022.12.22. 총론 및 각론 최종 확정 고시

2. 2022 개정 교육과정 정책의 주요 내용

2022 개정 교육과정은 기후환경 변화, 코로나19, 학령인구 감소 등 디지털 전환에 따른 변화와 위기 상황에 대응하고 극복하는 힘을 키워주는 교육혁신의 필요성을 느끼며 '포용성과 창의성을 갖춘 주도적인 사람'을 비전으로 삼았다. 2022 개정 교육과정의 인간상은 2015 개정 교육과정의 큰 틀을 바탕으로 미래 사회가 요구하는 소양 및 역량을 반영하여 자주적인 사람이 아닌 자기 주도적인 사람으로 교체되었다. 핵심역량에는 2015 개정 교육과정의 6대 핵심역량인 '자기관리, 지식정보처리, 창의적 사고, 심미적 감성, 의사소통, 공동체'에서 의사소통을 빼고 '협력적 소통'이 추가되었다. 추진과제로는 첫째, 미래 대응을 위한 교육과정이다. 급변하는 미래 사회 변화에 대응하려면 디지털 기초소양을 강화하고 지속가능발전 과제에 대한 대응 능력 및 공동체적 가치를 함양하는 교육이 강화되어야 한다. 둘째, 학교 현장의 자율적인 혁신 지원이다. 학교 교육과정의 자율성을 확대하고 지역·학교 간 격차를 줄이는 책임교육을 구현한다. 셋째, 학습자 맞춤형 교육 강화이다. 넷째, 교육 환경 변화 대응 지원이 있다.

초·중·고 교육과정에서 공통점은 생태전환교육의 교육 목표를 교육적 인간상 '더불어 사는 사람'과 핵심역량을 연계하여 교육 목표를 수정하였고, 민주시민교육의 현행 목표를 수정 및 보완하여 관련 교육 내실화를 추진하였다. 또한, 디지털 AI 소양 함양 교육을 강화하기 위해 디지털 함양을 위한 과목 신설이 포함되었다. 개선 방안은 학교 급별로 차이가 나며 초등학교의 경우 한글 해득 교육을 위해 국어

교과 시수를 34시간 증설하였다. 기존의 안전한 생활 관련 요소를 더욱 체계적으로 재구조화하여 생활 중심 안전교육을 강조하였고 이에 따라 실외 활동과 관련된 시간이 증대되었다. 중학교는 3학년 2 학기에 진로연계학기를 추가하여 창의적 체험활동과의 중복을 최소화하였다. 학교스포츠클럽 활동 시간은 학교의 운영 상황을 고려하여 시간을 축소하였다. 고등학교는 고교학점제의 안착을 위해 수업 량을 16회로 전환하고 다양한 진로 선택과목을 재구조화했으며 융합 선택과목을 신설하였다. 필수이 수학점을 조정하고 자율 이수 학점 이수 범위를 확대하였다. 단, 균형적인 교과 학습을 위해 국·영·수 총 이수학점이 81점을 초과하지 않도록 규정하였고 이를 정리하면 다음 <표 2>와 같다.

<표 2> 학교 급별 추진과제

학교급	추진과제
초등학교	<ul style="list-style-type: none"> • 통합교과와 창의적 체험활동 구분 • 국어 교과 시수를 34시간 확대(한글 해득/익힘 학습 시간 확대) • 기존의 안전한 생활을 재구조화하여 체계적인 안전교육 실시 • 초등 저학년의 즐거운 생활 시간을 개선하여 실외 활동 기회 증대
중학교	<ul style="list-style-type: none"> • 자유 학기와 적용 학기 자율적 선택 • 진로연계학기를 추가하여 교과별 진로 단원을 신설 • 학교자율시간을 활용하여 진로 관련 선택과목 운영 • 학교스포츠클럽 102시간으로 축소
고등학교	<ul style="list-style-type: none"> • 고교학점제의 안정적인 도입을 위해 다양한 진로 선택과목을 재구조하고 융합 선택 과목 신설 • 필수 이수학점을 조정하고 자율 이수 학점 이수 범위 확대 • 다양한 교과의 균형 학습을 위해 국영수 총 81학점 초과하지 않도록 규정

3. 2022 개정 교육과정 정책 관련 선행연구

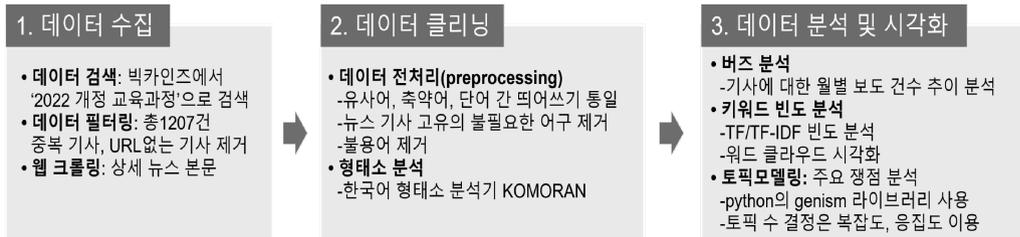
2022 개정 교육과정 정책 관련 선행연구를 정책 단계에 따라 분류하였다. 정책 형성 단계에서는 교과별로 새롭게 바뀌는 교육과정에 대비한 개선 방안을 다루는 연구가 많았다. 전용주, 김자미, 김한일(2021)은 전문가 FGI(focus group interview)를 통해 기존의 정보 교과의 부족한 시수, 통일성 있는 교재의 부재 등을 개선하고 미래 정보교육을 위한 내용 체계 안을 제안하였다. 문경호(2021)는 도덕 과목의 개편에 따라 온라인 도덕 수업의 확대를 예상하고 이에 맞춘 교재 개발의 필요성을 강조하였다. 모경환, 박형준(2021)은 일반사회 전공 교사들을 대상으로 ‘통합사회’ 과목의 운영 실태를 조사한 결과, 교사들이 ‘통합사회’의 의의에 대한 이해가 부족하고 교육과정 리터러시를 향상하는 것이 필요한 것으로 나타났다. 광영순, 신영준(2021)은 2015년 교육과정의 통합과학과 과학 탐구실험의 운영 실태를 점검하고 교사들을 대상으로 심층 면담을 통해 현장 전문가의 의견을 반영해 개선점을 도출하였다.

정책 결정 단계에서는 2022 개정 교육과정 총론 확정·고시를 앞두고 정책 형성 단계보다 2배 이상 더 많은 논문이 발표되었으며, 총론의 내용을 분석하고 교과별로 쟁점과 과제를 다룬 연구가 많았다. 특히 주로 현장에서 당면하는 문제를 반영하여 제안한 연구가 많았는데, 흥원표, 이광우, 임유나

(2022)는 2022 개정 교육과정 총론 주요 사항에 담긴 내용을 중심으로 고등학교 현장 교원들의 현장 교원들의 의견을 반영하여 총론을 분석한 결과, 선택과목의 양적 확대보다 교육과정 질적 개선, 고교 학점제 시행을 위해 학교 공간 및 시설 개선, 교원 수급이 필요하다고 제언하였다. 조용, 홍후조(2022)는 2022 개정 교육과정 총론을 중학교 현장 운영 경험자의 입장에서 총론 주요 사항의 내용을 분석하였다. 총론에서 체육형 클럽 활동을 ‘체육과’ 속에 통합하고 기술과 가정 과목은 교과를 분리하여 IT 융합 교육이 강화되어야 한다고 언급하였다. 끝으로, 백혜조(2022)는 2022 개정 교육과정의 초등학교 학교장 개설 선택과목을 검토하여 효과적인 운영방안을 제시했다. 선택과목이 교육 현장의 요구가 자세히 반영되어 있는지 재검토가 필요하며 교과 및 성취기준이 단위 학교와 교사의 자율에 맡기고, 재구성되어야 할 필요가 있다고 제언하였다.

III. 연구방법

이 연구는 2022 개정 교육과정 정책의 쟁점을 분석하고자 주요 언론매체의 보도 기사를 웹 크롤링 방법으로 수집하였으며, 수집된 자료는 데이터 전처리와 형태소 분석을 통해 분석하기 용이한 형태로 가공하여 버즈 분석, 키워드 빈도 분석 그리고 토픽모델링을 활용한 쟁점 분석을 실시하였다. 모든 과정은 파이썬(python)을 이용하였다. 연구 절차는 [그림 1]과 같다.



[그림 1] 연구 분석 절차

1. 데이터 수집

이 연구에서는 한국언론진흥재단이 1990년부터 운영하는 공익적 성격의 뉴스 빅데이터 시스템 빅카인즈(www.bigkinds.or.kr)에 등록된 중앙지 11개, 경제지 8개, 지방지 28개, 방송사 5개, 전문지 2개에서 2022 개정 교육과정 정책 관련 뉴스 기사를 수집하였다. 언론매체는 <표 3>과 같이 총 54개이다.

〈표 3〉 분석 대상 언론매체

종류	언론사
전국일간지	경향신문, 국민일보, 내일신문, 동아일보, 문화일보, 서울신문, 세계일보, 조선일보, 중앙일보, 한겨레, 한국일보
경제일간지	매일경제, 머니투데이, 서울경제, 아시아경제, 아주경제, 파이낸셜뉴스, 한국경제, 헤럴드경제
지역일간지	강원도민일보, 강원일보, 경기일보, 경남도민일보, 경남신문, 경상일보, 경인일보, 광주매일신문, 광주일보, 국제신문, 대구일보, 대전일보, 매일신문, 무등일보, 부산일보, 영남일보, 울산매일, 전남일보, 전북도민일보, 전북일보, 제민일보, 중도일보, 중부매일, 중부일보, 충북일보, 충청일보, 충청투데이, 한라일보
방송사	KBS, MBC, OBS, SBS, YTN
전문지	디지털타임스, 전자신문

자료수집을 위한 검색 키워드는 ‘2022 개정 교육과정’을 기본으로 하고, 검색어의 띄어쓰기와 단어 조합도 함께 고려하여 ‘2022개정교육과정’, ‘2022 교육과정’, ‘2022 교육과정 개정’, ‘새 교육과정’을 추가 검색하였다. 검색 기간은 학습자 특성 및 진로와 적성에 맞는 맞춤형 교육을 위한 교육체계라 할 수 있는 고교학점제의 도입 기반을 마련한 ‘고교교육혁신 방향을 통한 고교학점제’ 추진 단계 발표를 한 2018년 8월 17일부터 ‘2022 개정 교육과정 총론 및 각론’ 최종 확정 보도한 2022년 12월 22일까지로 설정하였다(교육부, 2021d).

본 연구에서는 1차로 빅카인즈에서 제공하는 메타데이터 파일을 다운받아 중복하여 수집된 기사, 광고와 홍보성 기사, 표나 그림으로만 채워져 분석할 텍스트 내용이 존재하지 않는 기사는 제외하여 데이터를 정리하였다. 빅카인즈에서 제공하는 데이터는 뉴스 기사 전문이 아닌 일부만 제공하므로 뉴스 기사 텍스트 전문을 수집하기 위하여 언론매체별로 사이트에 접속하여 웹 크롤링(web crawling) 기법을 이용하여 텍스트를 수집하였다. 여기서 웹 크롤링은 웹 크롤러를 이용해 인터넷에 공개된 정보를 자동으로 긁어 수집하는 방법으로, 빅카인즈에서 엑셀로 제공하는 메타데이터의 칼럼의 내용 중 기사 URL을 이용하여 웹 문서를 수집하였다. 웹 크롤링을 통해 수집된 분석 자료는 총 1,249건으로 URL이 없거나, 1차에서 중복 수집으로 삭제하였어도 웹 크롤링 후에도 중복된 뉴스 기사를 발견하여 삭제하였다. 그리고 URL은 있지만 페이지가 존재하지 않는 뉴스 기사도 삭제하였다. 따라서 최종 분석 자료는 총 1,207건이며, 수집 자료를 정책 단계별, 매체 종류별로 나누어 살펴보면 〈표 4〉와 같다.

〈표 4〉 수집 자료의 정책 단계별, 매체 종류별 건수

구분	일간지(건)	전문지(건)	방송사(건)	계(건)
정책 형성 단계	302	28	12	342
정책 결정 단계	783	35	47	865
총 계				1,207

2. 데이터 클리닝

수집한 뉴스 기사 텍스트 데이터를 한국어 자연어 처리(koNLP) 기법을 활용해 분석이 용이한 형태로 변환하였다. 먼저 수집된 텍스트 데이터 중 분석에 불필요한 문장을 삭제하는 작업과 유사어지만 다르게 표현된 단어들을 정규화(normalization)하는 작업을 반복하였다. 예컨대, 뉴스 기사의 하단에는 광고성 문구나 해당 매체의 공식 홈페이지 주소, 관련 뉴스 기사를 작성한 기자의 이름이나 이메일 주소, 저작권과 관련된 문구(예: ‘무단전재 및 재배포 금지’) 등의 불필요한 문장이 포함되기도 한다. 그리하여 파이썬의 내장된 re 라이브러리를 사용하여 연구자가 직접 정규표현식(regular expression)을 활용한 코드를 작성하여 자동화된 방식으로 불필요한 문장을 삭제하였다. 불필요한 문장은 또한 텍스트에 포함된 각종 특수기호(예: ‘◇’), 문장부호, 의존명사 등 분석 시 의미를 부여할 필요가 없는 불용어(stop word)를 제거하였다. 연구자가 별도로 구성하는 불용어는 모든 뉴스 기사에서 차별적인 의미를 갖지 않는 단어인 ‘2022 개정 교육과정’, 시제 표현(예: ‘어제’, ‘지난해’), 지역명(예: ‘서울시’), 학교명(예: ‘서울대’) 등으로 구성하여 직접 파이썬 코드를 작성하여 일괄 삭제하였다. 그리고 유사한 의미지만 축약어로 다양하게 표현되는 용어들(예: ‘수학능력시험’, ‘대학수학능력시험’)을 동일한 단어(예: ‘수능’)로, 여러 가지 언어로 표기된 용어(예: ‘AI 교육’)를 하나의 언어(예: ‘인공지능 교육’)로 통일하였다. 완전히 동일하지 않아도 의미상 연결되는 단어들(예: ‘교사’, ‘교원’, ‘선생님’, ‘교장’, ‘교감’)은 해석의 용이성을 확보하기 위해 통일(예: ‘교원’)하였다. 이렇게 유사어, 동의어, 축약어의 용어를 통일하고, 띄어쓰기도 통일하는 전처리 작업을 하였다.

이 연구에서는 수집한 뉴스 기사 텍스트 데이터에 대한 형태소 분석을 하기 위해 한국어 형태소 분석기를 사용하였다. 형태소 분석은 문장을 의미의 최소 단위인 형태소로 바꾸는 작업으로, KAIST의 SWRC(semantic web research center)에서 개발한 ‘한나눔’ 형태소 분석기와 Shineware에서 개발한 ‘KOMORAN’ 형태소 분석기, 서울대학교 IDS(intelligent data systems)에서 개발한 ‘꼬꼬마’ 형태소 분석기 등이 활용되고 있다. 이 중 ‘KOMORAN’ 형태소 분석기를 사용하였다. ‘KOMORAN’ 형태소 분석기는 공백이 포함된 고유명사(예: 영화 제목, 음식점명, 노래 제목, 전문 용어 등)를 더 정확하게 분석할 수 있고, 분석 속도가 빠르다. 그러나 띄어쓰기 없는 문장 분석에는 취약하지만, 언론매체의 보도 기사들은 형식에 맞춰진 텍스트이므로 선택의 문제점은 없다. 그리고 복합명사 사용자 용어 목록을 구성하여 연구자가 원하는 형태의 명사로 추출할 수 있는 큰 장점이 있다.

3. 데이터 분석 및 시각화

이 연구에서는 주어진 연구 문제를 해결하기 위해 텍스트마이닝 기법을 활용하였다. 텍스트마이닝(text mining)은 비정형 텍스트 데이터를 분석하는 기술로서, 특성과 패턴을 분석하여 유용한 지식을 발견하는 이론과 방법론이다(Sharkar, 2016). 이러한 텍스트마이닝은 주로 언론 보도 기사 분석에 활용되며, 텍스트의 문맥에 따른 쟁점 파악과 텍스트 간 연계 분석이 가능한 장점이 있다. 이 연구에서는 텍스트마이닝의 대표적 기법인 버즈 분석과 키워드 빈도 분석, 토픽모델링을 사용하였다(유예림, 2017; 손승덕, 2023).

가. 버즈 분석

버즈 분석은 일종의 문서량이나 검색량 분석이라 할 수 있으며(송태민, 송주영, 2016), 특정 주제를 포함한 문서의 등장 빈도 또는 검색 빈도수를 확인하여 해당 주제에 대한 국민의 관심과 반응을 파악하는 분석 방법이다. 그리고 문서의 등장 빈도를 시계열적으로 확인하며 해당 주제에 대한 관심도 변화를 살펴보는 추이 분석(trend analysis)이 가능하다(유예림, 2017). 이 연구에서 버즈(buzz)는 언론매체를 통해 보도된 2022 개정 교육과정 정책 관련 보도 기사의 건수를 의미한다. 이 연구에서는 2022 개정 교육과정 정책에 대한 언론매체의 관심도를 분석하기 위해 버즈 분석을 실시하였다. 따라서 본 연구에서 버즈 분석은 보도 기사의 건수에 대한 추이 분석을 실시하여 2022 개정 교육과정 정책에 대한 주요 언론매체의 관심도가 정책의 형성과 결정 단계에 따라 어떻게 변하는가를 탐색하였다.

나. 키워드 빈도 분석

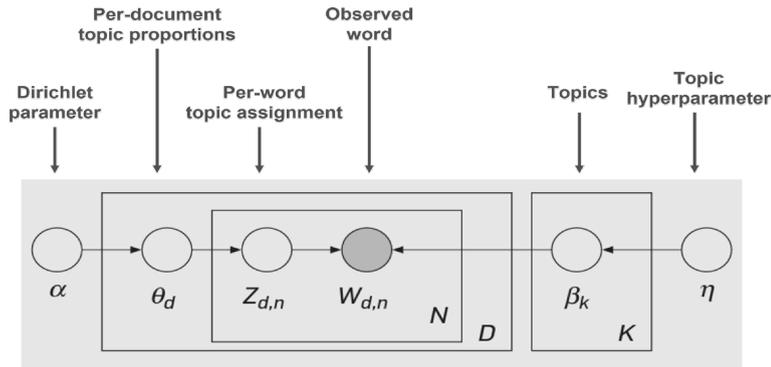
키워드 빈도 분석(keyword frequency analysis)은 특정 문서 내에서 자주 언급되는 키워드들을 추출하고 이들이 언급되는 빈도에 따라 순위를 부여하는 분석 방법이다(황명화, 2014). 키워드 빈도 분석은 버즈 분석의 분석 단위를 텍스트 문서의 키워드로 바꾸면 가능하다(손승덕, 2023). 키워드의 빈도는 특정 단어가 문서에서 얼마나 자주 등장하는지를 나타내는 '단어 빈도(term frequency, TF)'로 결정할 수 있다. 그러나 TF 값만을 가지고 문서에서의 중요성을 표현하기엔 부당한 측면이 있다. 이는 TF 값이 지나치게 큰 단어가 오히려 불용어 수준의 단어일 수 있기 때문이다. 이를 해결하기 위해 '역문서 빈도(inverse document frequency, IDF)'가 도입되었다. IDF 값은 문서 집합에 포함되어 있는 문서 수를 특정 단어가 나타난 문서의 수(document frequency, DF)로 나눈 것이다. 이는 상대적으로 많은 문서에 출현한 단어의 IDF 값은 작게 되고, 반대로 한쪽으로 편중하여 나타난 단어는 IDF 값이 커진다. 따라서 IDF 값이 작은 단어는 상투적인 단어일 가능성이 크고, 반대로 IDF 값이 큰 단어는 문서 내에서 주요한 단어일 가능성이 크다는 것을 의미한다(Robertson, 2004; 이성직, 김한준, 2009). TF-IDF는 단순한 빈도 처리가 아닌 단어의 출현 확률을 기준으로 출현 빈도를 재가공한 것이다(임화진, 2014). 즉, TF-IDF는 특정 문서 내에서 단어 빈도가 높을수록, 그리고 전체 문서 중 그 단어를 포함한 문서가 적을수록 그 값이 커진다(Salton & Buckley, 1988; Robertson, 2004; 이성직, 김한준, 2009; 김지은, 2017).

이 연구에서는 2022 개정 교육과정 정책과 관련하여 보도 기사에서 어떤 내용에 주목하는지를 탐색하기 위해 키워드 빈도 분석을 실행하였다. 이때, 키워드는 언론매체에서 뉴스 기사를 보도할 때 자주 등장하는 단어들을 의미한다. 빈도 분석은 TF와 TF-IDF를 모두 활용하여 수행되었으며, 그 결과를 비교하였다. 또한 뉴스 기사에서 2022 개정 교육과정 정책의 형성과 결정 단계에 따라 키워드가 어떻게 변화하는지를 살펴보기 위해 키워드 빈도 분석에 대한 추이 분석을 실시하였다. 이를 시각화하기 위해 워드 클라우드(word cloud)를 활용하였다. 워드 클라우드는 문서에서 언급된 단어의 빈도에 비례하여 단어의 크기를 나타내는 방법으로, 어떤 단어가 자주 등장하였는지를 빠르게 파악할 수 있다.

다. 토픽모델링

2022 개정 교육과정 정책이 언론매체에 담긴 쟁점을 파악하기 위하여 토픽모델링 기법을 실시하였다. 토픽모델링(topic modeling)은 텍스트 데이터 내 단어들의 빈도를 통계적으로 분석하여 전체 데이터를 관통하는 잠재적 주제, 즉 토픽(topic)들을 자동으로 추출하여 분류한다는 점에서 쟁점 분석 시 유용하다(황명화, 2014; 유예림, 백순근, 2016). 특히 이 연구에서는 다양한 토픽모델링 기법 중에서 잠재 디리클레 할당(LDA) 기법을 활용하였다.

LDA는 비지도 머신러닝 기법으로, 문서 집합 내의 각각의 문서들 또한 여러 개의 토픽을 포함하고 있으며, 각각의 토픽에 따라 단어가 분포한다(Blei & Jordan, 2003). 이렇게 한 개의 문서당 한 개의 토픽만 대응되는 것이 아니라, 여러 가지 토픽들을 포함할 수 있다는 측면에서 실제 텍스트의 본질을 잘 반영한다는 장점을 갖고 있다(김정수, 이석준, 2016). LDA는 데이터로부터 잠재 변인을 유추하고 복잡한 데이터의 차원을 축소해 데이터를 효과적으로 이해하고자 한다는 점에서 탐색적 요인 분석과 유사하다(Griffiths & Steyvers, 2004). 그리고 LDA는 잠재 변인인 토픽에 따른 단어 분포(β_k), 문서별 토픽 분포(θ_d), 토픽 할당($Z_{d,n}$)을 예측한다(황은희, 장지현, 양현경, 2019). Blei(2012)는 이러한 과정을 [그림 2]와 같이 시각화하였다.



[그림 2] LDA 모델의 시각화(Blei, 2012)

LDA는 사전 정보로 토픽의 개수를 지정해 주어야 한다. ‘토픽의 수’를 추출하기 위한 토픽 모형 평가 방법으로 복잡도(perplexity)와 응집도(coherence) 지수를 이용하였다. 복잡도는 낮을수록 모델 적합도가 적절하고(Blei, 2003), 응집도는 높을수록 모델 적합도가 적절하다(Aletras & Stevenson, 2013). 그러나 최종적인 토픽의 수는 해석의 가능성을 고려하여 종합적으로 결정하였다(신안나, 2019).

이 연구에서 LDA는 파이썬의 gensim 라이브러리를 활용하였다. 이 연구에서 gensim 라이브러리를 선택한 이유는 분석 속도가 빨라 대규모 문서 집합 처리에 유용하고(김우주, 구찬동, 2015; 유예림, 2017), 다른 분석 도구에서 활용하고 있는 추정 방법을 모두 포함하고 있기 때문이다. 그리고 LDA 모

델을 pyLDAvis 라이브러리를 사용하여 시각화하였다.

pyLDAvis는 잠재 디리클레 할당을 사용하여 추정된 토픽의 웹 기반 대화형 시각화인 LDAvis를 기반으로 만들어졌다(Siever & Shirley, 2014). 그리고 사용자 주제 모델의 주제를 해석하는 데 도움이 되는 대화형 주제 모델 시각화를 위한 파이썬 라이브러리이다.

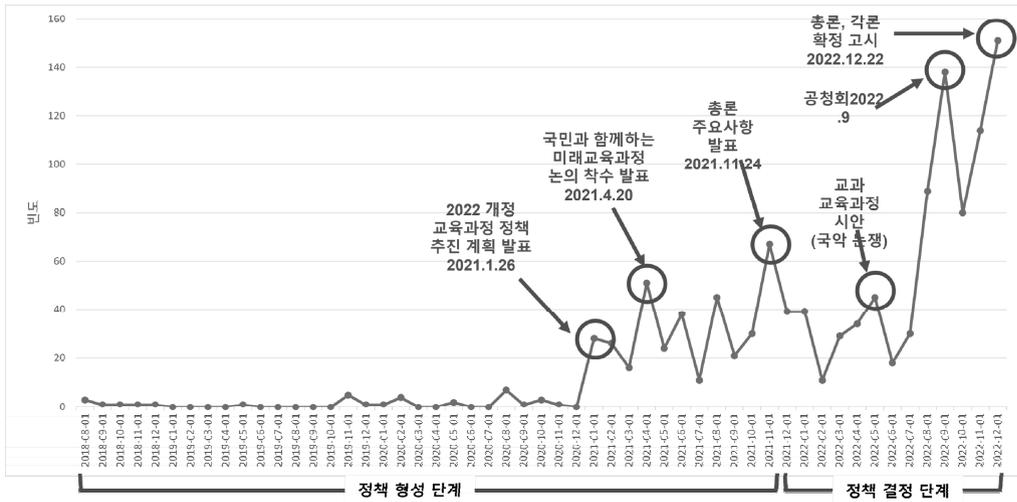
마지막으로 토픽 분석 후 산출된 토픽 명은 각 토픽을 구성하는 단어와 이론적 배경, 2022 개정 교육과정 정책의 맥락 등을 종합적으로 고려하여 연구자가 일차적으로 설정한 후, 도출한 쟁점의 타당성에 대한 전문가 4명의 의견 수렴을 거쳐 최종 토픽 명을 도출하였다. 전문가 4인은 교육학 분야(교육행정, 교육빅데이터, 교육상담) 석·박사과정생으로 교육 분야 경력이 5년~20년 이내로 구성하였다. 도출한 토픽 명에 대해 5점 척도(① 전혀 타당하지 않다, ② 타당하지 않다, ③ 보통이다, ④ 타당하다, ⑤ 매우 타당하다)로 타당도를 평가받았으며, 결과 4점 미만인 토픽 명에 대해서는 타당성이 부족하다 판단하여 전문가의 의견을 적절히 반영한 후 최종 토픽 명을 선정하였다.

IV. 연구 결과

1. 언론매체의 관심도 분석을 위한 버즈 분석 결과

2022 개정 교육과정 정책에 대한 언론매체의 관심도를 분석하기 위해 2018년 8월 17일부터 2022년 12월 22일까지 보도된 기사들을 대상으로 버즈 분석을 실시하였다. [그림 3]은 2022 개정 교육과정 정책에 대한 언론매체의 관심도를 분석하기 위해 보도된 기사 건수를 월별로 나타낸 것으로, 이를 토대로 2022 개정 교육과정 정책에 대한 언론매체의 관심도를 정책의 형성·결정 단계로 나누어 분석하면 다음과 같다.

정책 형성 단계를 살펴보면, 고교교육 혁신 방향을 통해 새 교육과정에 고교학점제를 부분 도입한다고 언급한 2018년 8월 17일부터 교육부의 「2021년 업무계획」을 발표한 2021년 1월 26일 전까지는 언론매체에서 거의 보도되지 않은 것을 확인할 수 있다. 이처럼 정책 형성 단계의 초·중반에는 거의 주목 받지 못하던 2022 개정 교육과정 정책은 2021년 1월, 교육부의 「2021년 업무계획」 추진과제 중 ‘2022 개정 교육과정 본격 시행 준비 계획’을 발표한 이후 보도 건수가 늘어난 것을 볼 수 있다. 또한 2021년 4월 20일 「2022 개정 교육과정 추진계획」을 발표, 2021년 11월 24일 「2022 개정 교육과정 총론 주요 사항」을 발표 전후까지 꾸준히 언론매체의 보도 건수가 증가하고 있는 것을 볼 수 있다.



[그림 3] 언론 보도 건수 월별 추이(총 1,207건 기준)

또한 정책 결정 단계를 살펴보면, 「2022 개정 교육과정 총론 주요 사항」 발표를 전후로 보도 건수가 급증하였으나, 정책 결정 단계의 초반엔 주요 언론매체의 보도 건수가 주춤하다가 2022년 4월 공개한 「2022 개정 음악과 교육과정」 시안에서 국악 관련 기준이 현재 초·중고 국악 비중 30~40%에서 대폭 축소된 것에 관해 국악계는 문제를 제기하였고, 이에 교육부에서 음악 교과서에서 국악 내용은 삭제되거나 축소되지 않는다는 보도자료를 발표하며 언론의 관심이 증가하였다. 또한 2022년 8월 교육부에서는 「디지털 인재 양성 종합 방안」을 발표하며 코딩교육 필수화를 언급하였고, 「2022 개정 교육과정 교과별 공청회」 개최를 통하여 현장 교원, 유관기관 관계자, 학부모 등이 참여하여 2022 개정 교육과정 시안 논의 발표와 국민의 의견을 수렴한다는 계획을 언급하면서 관련 보도 건수가 상승하였다.

선행연구에 의하면, 2015 개정 교육과정 정책에 대한 국내 주요 언론매체의 관심도는 정책의 형성 단계에는 거의 보도되지 않았으며, 정책의 결정 단계의 초·중반에도 주요 언론매체의 관심도가 낮았으나 공청회가 진행되기 시작하고 총론 및 각론 확정 고시 발표할 전후에는 보도량이 급증하였다(유예림, 백순근, 2016). 반면에 2022 개정 교육과정 정책은 정책의 형성 단계에 「2022 개정 교육과정 추진 계획」 발표 이후부터 보도 건수는 꾸준히 상승하였으며, 「국민과 함께하는 교육과정」 개발 취지에 맞추어 ‘국민참여소통채널’을 통해 국민의 의견을 수렴하고, 교과별 교육과정 공청회 개최하면서 2022 개정 교육과정 정책의 보도 건수는 이전보다 2배 이상 증가하였다. 이러한 언론의 관심도는 2022 개정 교육과정 총론과 교과 교육과정이 확정·고시된 2022년 12월에 정점을 이룬 것을 버즈 분석을 통해 확인할 수 있다.

2. 키워드 빈도 분석 결과

가. 분석 기간 전체

2018년 8월 17일부터 2022년 12월 22일까지 국내 주요 언론매체를 통해 보도된 기사에서 높은 빈도로 출현한 TF 기준의 상위 빈출 단어와 자주 출현하는 상투어의 중요도를 감소시키기 위해 TF-IDF를 적용한 상위 빈출 단어는 <표 5>와 같이 각 20개씩 정리하였다.

분석 기간 전체를 기준으로 가장 출현 빈도가 높은 단어는 ‘고교학점제’였고, 그다음으로는 ‘수업’, ‘위원회’, ‘미래’, ‘사회’, ‘선택’ 순이다.

<표 5> 언론매체의 보도 기사 상위 빈출 20개 단어: 전체

순서	단어	TF	순위	단어	TF-IDF
1	고교학점제	2,053	1	정책	115.6
2	수업	1,987	2	고등	111.4
3	위원회	1,872	3	자유	111.3
4	미래	1,834	4	민주	109.8
5	사회	1,831	5	개발	107.7
6	선택	1,828	6	설명	106.8
7	교사	1,695	7	강화	105.8
8	민주	1,581	8	수업	105.0
9	의견	1,566	9	논의	105.0
10	교과	1,566	10	전환	104.7
11	자유	1,562	11	우려	104.4
12	정책	1,478	12	의견	104.2
13	지원	1,456	13	미래	103.4
14	국민	1,280	14	계획	102.8
15	고등	1,220	15	시행	102.6
16	학점	1,206	16	교과	102.4
17	학습	1,180	17	확정	102.3
18	평가	1,170	18	위원회	101.6
19	도입	1,094	19	필요	101.4
20	정보	1,083	20	학습	101.2

이 결과는 국내 주요 언론매체는 2022 교육과정 개정에 따른 가장 큰 변화인 고교학점제 도입에 관심이 가장 많음을 알 수 있으며, 2021년 국가교육위원회 설치 및 운영에 관한 법률이 제정되고, 2022년 9월 국가교육위원회가 정식 출범되었기에 언론매체에서도 큰 관심을 가졌음을 알 수 있다.

한편, 분석 기간 전체를 기준으로 가장 중요하게 고려된 단어는 ‘정책’, ‘고등’, ‘자유’, ‘민주’ 순이다.

〈표 6〉 언론매체의 보도 기사 상위 빈출 20개 단어: 정책 형성과 정책 결정

순서	정책 형성		정책 결정	
	키워드	TF	키워드	TF
1	고교학점제	1192	위원회	1700
2	미래	944	민주	1473
3	수업	834	자유	1446
4	교사	761	사회	1266
5	선택	740	의견	1245
6	학점	628	수업	1153
7	도입	603	교과	1127
8	사회	565	정책	1102
9	계획	516	선택	1088
10	지원	495	지원	961
11	평가	459	국민	938
12	고등	450	교사	934
13	대입	447	미래	890
14	교과	439	고교학점제	861
15	학습	438	역사	823
16	교원	398	고등	770
17	필요	386	디지털	746
18	정책	376	학습	742
19	정보	373	평가	711
20	고교	371	정보	710

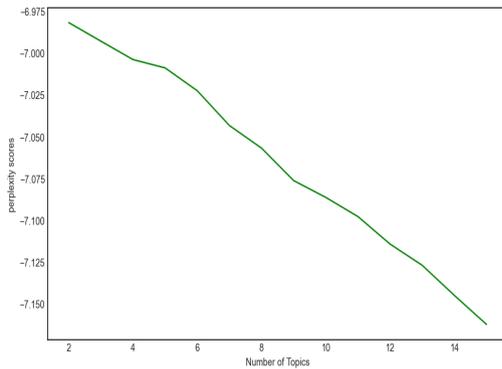


〈그림 5〉 TF 상위 빈출 50위 단어 워드 클라우드(좌:형성, 우:결정)

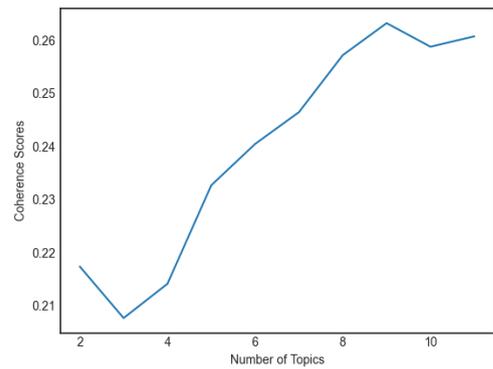
3. 토픽모델링 분석을 활용한 쟁점 분석 결과

키워드 빈도 분석은 정책 단계에 따라 키워드가 출현하고 사라지는 변화를 효과적으로 보여주지만, 단어의 존재 여부만으로는 표현되지 않는 쟁점을 드러내는 데에는 한계가 있다(박중희, 박은정, 조동준, 2015). 이러한 한계를 극복할 수 있는 분석법인 토픽모델링은 단어들의 배치를 통해 숨겨진 토픽의 수와 내용을 추정하기 때문에 쟁점 분석에 유용하다(유예림, 2017).

LDA 토픽모델링 분석과 복잡도(그림 6), 응집도 지수(그림 7)를 참고하여 토픽 수를 결정한다. 토픽 수는 복잡도 지수가 낮고 응집도 지수가 높은 가장 적절한 토픽의 수를 최종 토픽 수로 도출한다. 그리고 일차적으로 토픽 명을 연구자가 도출한 후, 전문가 4인의 타당도 검토를 거쳐 분석 기간 전체와 정책 단계별로 최종 토픽 명을 선정하였다.



[그림 6] 복잡도 지수 그래프(전체)



[그림 7] 응집도 지수 그래프(전체)

가. 분석 기간 전체

2022 개정 교육과정 정책 형성과 정책 결정 단계를 모두 포함하는 분석 기간 전체에서 토픽을 도출하기 위해 먼저 적정 토픽 수를 구한다. 분석 기간 전체에서의 토픽은 복잡도, 응집도 지수를 참고하여 총 8개로 도출되었다. 복잡도와 응집도를 살펴 가장 적절해 보이는 토픽 수가 9개였지만, 실제 토픽을 추출하였을 때 두 토픽이 중복되는 단어들이 많아 토픽 수를 8로 정하였다.

〈표 7〉은 토픽별로 각 토픽에 나타난 확률이 높은 단어 상위 10개와 그 외 단어와 토픽 명을 정리한 표이다.

〈표 7〉 도출된 주요 쟁점(분석 기간 전체)

구분	쟁점	키워드				
		1	2	3	4	5
		6	7	8	9	10
토픽1	18.2% 고교학점제 전면 도입	고교학점제	도입	선택	학점	진로
		대입	평가	고등	지원	계획
토픽2	16.8% 역사 교과에 '자유' 강조 방향 확정 논란	위원회	민주	자유	역사	의견
		심의	논란	국민	의결	용어
토픽3	14.1% 미래 지향적 교육과정과 디지털 환경	미래	교사	지원	사업	그린스마트
		혁신	정책	변화	디지털	환경
토픽4	13.1% 선택교과 확대와 국영수 시수 축소	선택	국악	수학	고등	국어
		고교학점제	기초	단위	도입	공통
토픽5	11.5% 국민참여소통채널을 통한 정보교육 초중등 연계 필요 국민 의견 접수	의견	정보	국민	교과	참여
		디지털	공청회	수렴	정책	소통
토픽6	10.6% 국가교육위원회 출범 늦어져 교육정책 추진 차질 우려	정책	위원회	자사고	국회	대입
		고교학점제	평가	문제	논란	출범
토픽7	8.3% 초중등 정보교육 시수 확대하고 코딩교육 필수화	디지털	정보	평가	지원	교과서
		교사	교원	문해력	문제	양성
토픽8	7.5% 게임 섯다운제 폐지하고 시간선택제로 전환	게임	청소년	지원	선택	온라인
		섯다운	보호	위원회	변화	환경

토픽1의 경우 '고교학점제', '도입', '선택', '학점', '대입'의 단어가 포함되어 있으며, 단어들의 의미와 맥락을 고려할 때 고교학점제 종합 추진계획 발표 내용과 관련되므로 '고교학점제 전면 도입'이라고 토픽 명을 정하였다. 나머지 토픽도 같은 방식으로 명칭을 부여하였으며, 그 결과 2022 개정 교육과정 정책의 쟁점은 “고교학점제 전면 도입”, “역사 교과에 '자유' 강조 방향 확정 논란”, “미래 교육을 위한 학교 디지털 전환 수립 고민 시작”, “사회, 과학, 역사 교과의 선택지 확대와 국·영·수 시수 축소”, “국민 참여소통채널을 통한 정보교육 초중등 연계 필요 국민 의견 접수”, “국가교육위원회 출범 늦어져 교육 정책 추진 차질 우려”, “초중등 정보교육 시수 확대하고 코딩교육 필수화”, “게임 섯다운제도 폐지하고 시간선택제로 전환”의 8가지로 도출되었다.

나. 정책 단계별

1) 정책 형성 단계

2022 개정 교육과정 정책 형성 단계에서의 토픽은 <표 8>과 같이 ‘고교학점제와 성취평가제’, ‘미래형 수능 및 대입제도 변화 논의’, ‘국민 의견 온라인 설문조사 수렴’, ‘디지털 리터러시 소양 강화’, ‘청소년 게임 섯다운제 폐지’ 5개로 도출되었다. 먼저 토픽1은 2021년 2월 10일 발표한 고교학점제 종합 추진계획 발표 내용 중 학점제형 교육과정의 내신 평가제도인 성취평가제 확대이며, 토픽2는 교육부 2021년 업무계획을 발표하며 고교학점제 등 새 교육과정을 반영한 미래형 수능 및 대입 방향을 수립하기 위한 논의 착수 쟁점과 연결된다. 토픽3은 국가교육회의의 2022 개정 교육과정을 위한 국민 의견을 수렴하기 위하여 온라인 설문조사를 실시한 활동에 대한 쟁점이라 할 수 있다. 토픽4는 코로나19로 인해 빠르게 가속화된 디지털 대전환 시기 미래 교육을 위해 디지털 소양 강화, 토픽5는 청소년의 게임을 할 자유를 침해한다는 이유로 비판이 제기되었던 게임 섯다운제의 폐지와 게임 과몰입 예방제도가 청소년 자기 결정권과 행복추구권 등을 존중하는 방식으로 전환되길 기대하는 쟁점이라 할 수 있다.

<표 8> 도출된 주요 쟁점(정책 형성 단계)

구분	쟁점	키워드					
		1	2	3	4	5	
		6	7	8	9	10	
토픽1	26.9%	고교학점제와 성취평가제	학점	고교	진로	교과	대입
			단위	평가	성취평가제	전환	공통
토픽2	23.8%	수능 및 대입제도 논의	수능	대입	평가	학습	정책
			정시	논의	입시	특수	고교
토픽3	20.4%	국민 의견 설문조사 수렴	국민	의견	학습	참여	설문
			전문가	수렴	진행	학부모	온라인
토픽4	18.0%	디지털 리터러시 소양 강화	교과	미디어	시민	교과서	학습
			디지털	연계	환경	리터러시	정보
토픽5	10.9%	청소년 게임 섯다운 폐지	게임	청소년	교과서	대입	이용
			섯다운	연계	교과	자유	평가

2) 정책 결정 단계

2022 개정 교육과정 정책의 결정 단계에서 토픽은 총 7개로 도출되었으며 <표 9>와 같다. 정책 결정 단계는 분석 기간을 전제로 설정하였을 때의 유사한 결과가 나타났지만, 한정된 범위 내에서의 토픽 분석한 결과 ‘국악 교육 비중 축소 반발’, ‘민주시민교육’, ‘늦은 국가교육위원회 출범’의 토픽을 도출하였다.

〈표 9〉 도출된 주요 쟁점(정책 결정 단계)

구분	쟁점	키워드				
		1	2	3	4	5
		6	7	8	9	10
토픽1	27.8% 교과서 자유민주주의 용어 논란	위원회	민주	자유	심의	의견
		국민	위원	정책	용어	교과서
토픽2	20.4% 미래형 평가-대입 제도 개편	선택	평가	수능	교과	진로
		수업	대입	지원	미래	고교학점제
토픽3	12.5% 국악 교육 축소 반발과 선택 과목 확대	국악	디지털	수업	지원	미래
		학습	강화	선택	기준	교과
토픽4	11.9% 교과별 공청회 전 국민 의견 수렴	지원	고교학점제	의견	선택	교과
		수업	위원회	공청회	국민	강화
토픽5	11.7% 민주시민교육 및 노동 교육	교사	노동	민주	경제	시민
		수학	선택	수업	환경	고교학점제
토픽6	11.2% 디지털 교육 강화	디지털	수업	미래	지원	강화
		정보	교과	교사	진로	역량
토픽7	4.5% 늦은 국가교육위원회 출범	위원회	교과서	역사	추천	수립
		대한민국	출범	기술	행정 예고	해설

V. 논의 및 결론

이 연구는 2018년 8월부터 2022년 12월까지 주요 언론매체에서 보도한 2022 개정 교육과정 정책 관련 뉴스 기사에서 분석 기간 전체와 정책 단계별로 어떠한 쟁점들이 나타났는지를 키워드 빈도 분석과 토픽모델링 분석을 통해서 살펴보았다. 이를 위해 우선, 국내 주요 17개 일간지와 3개 방송사에서 보도한 2022 개정 교육과정 정책 관련 뉴스 기사를 웹 크롤링 기법을 이용하여 자체적으로 수집하였다. 그리고 최종적으로 사진, 이미지로만 이루어진 기사, 중복된 기사 등을 제외한 기사 총 1,207건을 수집하였다. 이후 데이터 클리닝과 형태소 분석, 자료 정제 과정을 거쳐 토픽모델링 분석이 가능한 문서-단어 행렬 형태로 데이터를 가공하였다.

연구의 구체적인 세 가지 연구 문제의 분석 결과는 다음과 같다.

본 연구의 첫 번째 연구 문제는 2022 개정 교육과정 정책에 대한 언론매체의 관심도의 변화를 파악하는 것이었다. 언론매체의 관심도를 분석하기 위해 버즈 분석한 결과, [그림 3]의 언론 보도 건수 월별 추이를 보면, 정책 형성 단계의 초·중반에는 관련 보도가 거의 없다가 2022 개정 교육과정 추진계획을

발표한 2021년 2월 시점부터 상승하였다. 이후 언론매체의 관심도는 계속 증가 추세였으며, 2022 개정 교육과정 교과별 공청회를 개최한 2022년 9월 전후로 2배 이상의 상승을 보이다가 교육과정 총론 및 교과 교육과정 확정 고시된 2022년 12월에 정점을 이루었다. 이러한 분석 결과는 2015 개정 교육과정 정책에 대한 언론의 관심도를 연구한 선행연구(김소영, 강신원, 2015; 유예림, 백순근, 2016)에서 정책의 형성, 결정 단계에서 관련 보도가 거의 없어 언론매체는 해당 정책에 소극적 태도를 보였다는 연구 결과에 비교해, 2022 개정 교육과정은 교육과정 추진 확정 이후 꾸준히 언론매체의 관심도가 증가하고 특히 교과별 공청회 개최 전후로 더 급증한 것으로 보아 언론매체의 관심도는 이전 교육과정보다 관심도가 증가함을 나타낸다. 이는 언론매체의 보도 건수의 증가로 2022 개정 교육과정 정책에 대한 평가와 피드백이 자연스럽게 이루어질 수 있어 차후 정책의 개선과 수정에 필요한 정보들을 얻을 수 있으므로 교육정책의 향상에 도움이 됨을 시사하고 있다.

본 연구의 두 번째 연구 문제는 2022 개정 교육과정 정책에 대한 언론매체의 보도 기사에 나타난 키워드를 파악하는 것이었다. 키워드 빈도 분석한 결과, <표 5>와 <표 6>에 제시된 언론매체의 보도 기사 상위 빈출 결과에 제시된 바와 같이 '고교학점제'가 가장 많이 등장하였고 그다음으로는 '수업', '위원회', '미래'와 같은 키워드가 많이 언급되었다. 이 단어들은 일반적인 단어일 수도 있지만 '온라인 수업', '원격수업' 등의 단어의 빈도가 높았기에 '수업' 단어가 많았을 것으로 보인다. 빈도 분석이 아닌 문장 내에서 설명력이 높은 키워드를 부각하는 TF-IDF를 활용하여 재분석한 결과, '정책', '고등', '자유', '민주'가 최상위 빈도를 보였다. 이러한 분석 결과는 국내 주요 언론매체는 2015 개정 교육과정 정책의 TF-IDF의 빈도 분석한 선행연구(유예림, 백순근, 2016)의 결과와 비슷하게 국가·사회적으로 주목받는 '역사 교과서의 자유민주주의 용어' 논란에 대한 쟁점에 더 많은 기사화와 방송을 할애한 것을 확인할 수 있었다. 이는 교육부가 제시한 2022 개정 교육과정 정책의 의도가 언론매체를 통해 일반 대중들에게 제대로 전달되지 않음을 시사하고 있다. 그리고 2022 개정 교육과정 정책의 형성 단계에서 키워드는 '고교학점제', '미래', '수업', '교사', '선택', '학점', '도입' 순으로 높게 나타났지만, 정책의 결정 단계에서는 '위원회', '국민', '자유', '역사', '디지털', '평가' 순으로 나타났다. 이러한 분석 결과는 2022 개정 교육과정 정책의 형성 단계의 기간에 교육부에서 2021년 2월 「고교학점제 종합 추진 계획」 발표와 2021년 8월 「2025년 고교학점제 전면 적용을 위한 단계적 이해 계획(안)(2022-2024)」 발표 등 고교학점제 운영 체제로의 전환을 위한 이행 계획 수립 준비를 하였는데, 이를 언론매체에서는 중요한 이슈로 다루었음을 볼 수 있다. 정책의 결정 단계에서는 예상보다 늦은 2022년 5월에 국가교육위원회 출범된 이슈와 사회와 역사 교과서의 자유 용어에 대한 이슈를 언론매체는 중요하게 다루었음을 볼 수 있다.

본 연구의 세 번째 연구 문제는 2022 개정 교육과정 정책에 대해 언론매체의 보도 기사에 나타난 주요 쟁점을 파악하는 것이었다. 토픽모델링 기법을 활용하여 분석한 결과, <표 7>의 분석 기간 전체를 기준으로 토픽 분석한 결과에 제시된 바와 같이 ① 고교학점제 전면 도입, ② 역사 교과에 '자유민주주의' 용어 논란, ③ 미래 지향적 교육과정과 디지털 환경, ④ 선택교과 확대와 국·영·수 시수 축소, ⑤ '국민참여소통채널'을 통한 국민 의견 접수, ⑥ 국가교육위원회 출범 늦어져 교육정책 추진 차질 우려, ⑦ 초·중등 정보교육 시수 확대하고 코딩교육 필수, ⑧ 게임 섯다운제 폐지하고 시간선택제로 전환의 8가지로 나타났다. 그리고 정책의 단계별로 분석한 결과, 정책 형성 단계에서는 <표 8>의 정책 형성 단계의 토픽 분석 결과에 제시된 바와 같이 ① 고교학점제와 성취평가제, ② 수능 및 대입제도 논의, ③ 국민 의견 설

문조사 수렴, ④ 디지털 리터러시 소양 강화, ⑤ 청소년 게임 셧다운 폐지 등 주로 고교학점제 정책이나 대입제도 등과 같이 교육부에서 제시한 정책이 쟁점으로 나타났으며, 10년간 이루어졌던 셧다운제도 폐지는 게임 과몰입 예방제도가 청소년 자기 결정권과 행복추구권 등을 존중하는 방식으로 전환되길 기대하는 쟁점이라 할 수 있다. 정책 결정 단계에서는 <표 9>의 정책 결정 단계의 토픽 분석 결과에 제시된 바와 같이 ① 교과서 자유민주주의 용어 논란, ② 미래형 평가-대입 제도 개편, ③ 국악 교육 축소 반발과 선택과목 확대, ④ 교과별 공청회 전 국민 의견 수렴, ⑤ 민주시민교육 및 노동 교육, ⑥ 디지털 교육 강화, ⑦ 늦은 국가교육위원회 출범 등 교육부에서 제시하는 2022 개정 교육과정 주요 사항의 내용과 국민 의견 수렴 활동과 같은 새 교육과정에서 추구하는 가치에 관한 쟁점이 55.2% 도출되었고, 사회적으로 주목받는 '역사 교과서 자유민주주의 용어 논란', '국악 교육 축소 반발', '국가교육위원회 늦은 출범으로 정책 추진 차질 우려'와 같은 쟁점은 44.8%로 도출되어 주요 언론매체를 통해 2022 개정 교육과정 정책의 취지나 의도가 일반 대중들에게 잘 전달되었음을 시사하고 있다. 이처럼 정책 단계에 따라 2022 개정 교육과정의 쟁점이 다르다는 것은 언론매체의 관심사가 정책의 단계에 따라 변화하고, 언론의 역할이 정책에 대한 대중의 이해도 증진에 중요한 역할을 함을 확인할 수 있다.

본 연구는 텍스트마이닝 분석 기법을 활용하여 정책의 형성과 결정 단계별로 2022 개정 교육과정의 언론매체의 관심도 분석, 키워드 빈도 분석 그리고 토픽모델링 기법인 LDA를 활용한 토픽 분석을 하여 해당 정책을 살펴보았으며, 교육과정의 쟁점 분석을 위한 뉴스 기사와 같은 빅데이터 분석의 유용성을 확인하였다는 데 의의가 있다. 이 연구에서 활용한 텍스트마이닝 기법은 다양한 교육정책 연구에서 적용될 수 있을 것이다.

본 연구의 한계점과 이를 바탕으로 제안할 수 있는 후속 연구는 다음과 같다.

첫째, 정책적 쟁점은 정책 과정에 따른 변화의 특성이 있어(유예림, 2017) 정책의 형성과 결정 그리고 집행 단계에 따라 쟁점 추이 분석이 필요하지만, 본 연구는 정책의 형성과 결정 단계만 살펴보았기에 정책 집행 단계에도 언론매체의 관심도와 쟁점을 확인할 수 없는 한계가 있다. 그리하여 정책의 집행 단계의 언론매체의 관심도와 토픽 분석을 통한 쟁점을 도출하고, 좀 더 면밀한 시사점을 도출하기 위해 이해 당사자인 초중등 교육 현장의 교사와 학생, 학부모, 교육과정 분야의 전문가, 교육 당국 관계자들과의 심층 면담을 통한 질적 연구를 함께할 필요가 있다.

둘째, 이 연구는 2022 개정 교육과정 정책에 기반한 키워드와 토픽 분석에 집중되어 도출되었으며로 국민의 반응에 대한 이해가 다소 부족하여 이에 언론 기사 댓글을 통해 일반 국민의 2022 개정 교육과정에 대한 인식이 텍스트 데이터로 더 활발하게 나타날 가능성이 있으므로, 향후 기사 댓글을 분석에 활용한다면 다양한 쟁점들이 도출될 수 있을 것이다. 그리고 감성분석(sentiment analysis) 기법을 활용하여 교육 정책에 대한 긍정 또는 부정 인식을 효과적으로 분석할 필요가 있다.

셋째, 이 연구는 정책 단계별 토픽을 도출하는 잠재 디리클레 할당(LDA) 기법을 적용하였다. 토픽의 변화를 확인하기 위한 LDA의 확장 모형인 구조적 토픽 모형(structural topic model, STM)과 토픽 간의 위계 구조를 상징하는 위계적 잠재 디리클레 할당(hierarchical latent dirichlet allocation, hLDA) 등 모형에 따라 주안점이 상이하므로, 다른 토픽 분석 기법을 적용하여 결과를 비교 분석하여 분석 기법들을 더욱 정교하게 발전시킬 필요가 있다.

참고문헌

- 교육부(2021a). 2022 개정 교육과정, 학생·학부모와 함께 이야기하다. 자료출처(보도일: 2021.08.26.): <https://www.korea.kr/briefing/pressReleaseView.do?newsId=156468044>
- 교육부(2021b). 국민참여 소통채널 주요의견 공개 브리핑. 자료출처(보도일: 2021.09.19.): <https://www.moe.go.kr/boardCnts/viewRenew.do?m=060302&s=moe&page=1&boardID=412&boardSeq=92631&lev=0&opType=N>
- 교육부(2021c). 2022 개정 교육과정 총론 주요 사항에 대한 공청회 개최. 자료출처(보도일: 2021.10.21.): <https://www.moe.go.kr/boardCnts/viewRenew.do?boardID=294&lev=0&boardSeq=89363>
- 교육부(2021d). '2022 개정 교육과정' 총론 주요 사항 발표. 자료출처(보도일:2021.11.24.): <https://www.moe.go.kr/boardCnts/viewRenew.do?boardID=294&lev=0&boardSeq=89671>
- 교육부(2022). 2022 개정 교육과정 총론 시안 개발 연구
- 곽영순, 신영준(2021). 2022 개정 교육과정에 대비한 과학과 통합과학 및 과학탐구실험 교육과정 개선 방안 탐색. **과학교육연구지**, 45(2), 143-155.
- 권순보, 유진은(2018). 텍스트 마이닝 기법을 통한 수능 절대평가. **열린교육연구**, 26(2), 57-79.
- 김경근(2008). 한국 사회의 대안교육 수요 결정요인. **한국교육학연구**, 14(1), 45-69.
- 김규하, 박철웅(2015). 토픽 모형 및 사회연결망 분석을 이용한 한국데이터정보과학회지 영문 초록 분석. **한국데이터정보과학회지**, 26(1), 151-159.
- 김병주, 김은아, 홍준영(2006). 개정 사립학교법에 대한 신문의 보도태도 분석. **교육행정학연구**, 24(4), 363-387.
- 김선아, 박진희, 이현정, 정유진(2016). 텍스트마이닝 기법을 활용한 다문화 미술교육 연구 동향 분석 연구. **다문화교육연구**, 9(2), 203-227.
- 김선정(2014). **행정학의 이해**. 서울: 북큐브네트웍스
- 김수영(2023). 2015 개정 교육과정과 2022 개정 교육과정 문서의 비교 분석. **교육과정개발연구**, 1(2), 131-144.
- 김우주, 구찬동(2015). **교육정보화 정책에서 빅데이터 활용 방안 탐색 연구: 소셜미디어 데이터 활용을 중심으로**. 한국교육학술정보원.
- 김용대, 조광현(2013). 빅데이터와 통계학. **한국데이터정보과학회지**, 24(5), 959-974.
- 김정수, 이석준(2016). 중국 주식상장폐지제도와 개인투자자 반응:개인투자자의 유형과 토픽 분석. **부산대학교 중국 연구소**, 19(1), 123-152.
- 김정숙(2006). 여자 대학생의 직업인식과 직업선택 과정. 고려대학교 박사학위논문.

- 김지은(2017). 빅데이터를 활용한 대학구조개혁 평가의 키워드 및 토픽 분석. 서울대학교 박사학위논문.
- 모경환, 박형준(2021). 고등학교 '통합사회' 운영 실태와 2022 개정 교육과정의 개발 방향. **시민교육연구**, 53(4), 55-93.
- 목진휴, 박기묵(1993). An Issue Attention Cycle of Air Pollution Issue. **정책분석평가학회보**, 3(1), 199-211.
- 문경호(2021). 고교학점제 도입과 2022 교육과정 개정에 따른 도덕과의 변화와 과제. **도덕윤리과교육**, 4(73), 61-79.
- 박종희, 박은정, 조동준(2015). 북한 신년사(1946-2015)에 대한 자동화된 텍스트 분석. **한국정치학회보**, 49(2), 27-61.
- 백혜조(2022). 2022 개정 교육과정의 초등학교 선택과목(활동)에 대한 비판적 고찰. **한국교육학연구**, 28(4), 149-171.
- 손승덕(2023). 텍스트 마이닝을 활용한 특성화고등학교 교육과정 운영에 대한 언론 보도 경향 분석. 서강대학교 석사학위논문.
- 송태민, 송주영(2016). **R을 활용한 소셜 빅데이터 연구방법론**. 서울: 한나래 출판사.
- 신안나(2019). 빅데이터를 활용한 자유학기제 정책의 키워드 및 토픽 분석. 서울대학교 박사학위논문.
- 오옥환(2005). **교사전문성**. 서울: 교육과학사.
- 유예림, 백순근(2016). 자동화된 텍스트 분석을 활용한 2015 개정 교육과정 정책에 대한 언론 보도의 쟁점 분석. **교육과정평가연구**, 19(3), 127-156.
- 유예림(2017). 빅데이터 분석기법을 활용한 2015 개정 교육과정 정책에 대한 언론 보도 분석. 서울대학교 박사학위논문.
- 윤옥한(2022). 2022 개정 교육과정 의미 분석. **국제문화기술진흥원**, 8(5), 59-69.
- 이성직, 김한준(2009). TF-IDF의 변형을 이용한 전자뉴스에서의 키워드 추출 기법. **한국전자거래학회지**, 14(4), 59-73.
- 이종수(2009). **행정학 사전**. 서울: 대영문화사.
- 이종재(2003). **교육행정시스템 혁신의 방향**. 한국교육개발원.
- 임화진(2014). **빅데이터를 이용한 충남도 정책 키워드 분석**. 충남발전연구원.
- 전용주, 김자미, 김한일(2021). 2022 개정 교육과정을 대비한 정보과 과목 구성 및 내용 체계 제안. **컴퓨터교육학회**, 24(6), 1-15.
- 조용, 홍후조(2022). 2022 개정 중학교 교육과정 총론 개정안의 쟁점과 과제에 관한 시론적 탐색. **학습자중심교과교육연구**, 22(23), 131-146.

- 최예림, 신서경(2022). 테크놀로지 활용 교육정책의 현황과 과제: 토픽모델링을 중심으로. **교육정보미디어연구**, 28(3), 681-706.
- 홍원표, 이광우, 임유나(2022). 2022 개정 고등학교 교육과정의 남은 쟁점과 과제: 교육과정 담당 교원들의 의견을 중심으로. **교육과정연구**, 40(1), 157-183.
- 황명화(2014). **여론 모니터링을 위한 비정형 빅데이터 시공간분석 방법론 연구**. 국토연구원.
- 황은희, 장지현, 양현경(2019). 텍스트마이닝 기법을 활용한 자유학기제 연구동향 분석. **교육문화연구**, 25(3), 299-318.
- Aletras, N., & Stevenson, M. (2013). Evaluating topic coherence using distributional semantics. *Proceedings of the 10th International Conference on Computational Semantics(IWCS 2013)-Long Papers*, 13-22.
<https://www.aclweb.org/anthology/W13-0102.pdf>
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3, 993-1022. <https://doi.org/10.5555/944919.944937>
- Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4), 77-84. <https://doi.org/10.1145/2133806.2133826>
- Brookhart, S. M., & Freeman, D. (1992). Characters of entering teacher candidates. *Review of Educational Research*, 62(1), 37-60. <https://doi.org/10.3102/00346543062001037>
- Ding, W., & Chen, C. (2014). Dynamic topic detection and tracking: A comparison of HDP, C-word, and cocitation methods. *Journal of the Association for Information Science and Technology*, 65(10), 2084-2097. <https://doi.org/10.1002/asi.23134>
- Gold, N. C. (1981). *Meta-evaluation of selected bilingual education projects*. University of Massachusetts. https://scholarworks.umass.edu/dissertations_1/3667
- Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National Academy of Sciences*, 101(1), 5228-5235. <https://doi.org/10.1073/pnas.0307752101>
- McCombs, M., & Shaw, D. L. (1972). The Agenda-Setting function of mass media. *Public Opinion Quarterly*, 36(2), 176-187. <https://doi.org/10.1086/267990>
- Robertson, S. (2004). Understanding inverse document frequency: on theoretical arguments for IDF. *Journal of Documentation*, 60(5), 503-520.
<https://doi.org/10.1108/00220410410560582>
- Salton, G., & Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information Processing & Management*, 24(5), 513-523.

[https://doi.org/10.1016/0306-4573\(88\)90021-0](https://doi.org/10.1016/0306-4573(88)90021-0)

Sarkar, D. (2016). Text Analytics with Python. In *Apress eBooks*.

<https://doi.org/10.1007/978-1-4842-2388-8>

Sievert, C., & Shirley, K. E. (2014). LDAvis: A method for visualizing and interpreting topics. In *Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces*(pp. 63-70). <https://doi.org/10.3115/v1/W14-3110>

Teh, Y. W., Jordan, M. I., Beal, M. J., & Blei, D. M. (2012). Hierarchical dirichlet processes. *Journal of the American Statistical Association*, *101*(476), 1566-1581. <https://doi.org/10.1198/016214506000000302>

· 논문접수 : 2023.07.05. / 수정본접수 : 2023.07.28. / 게재승인 : 2023.08.09.

ABSTRACT

An Analysis of Issues in Media Reports on the 2022 Revised Curriculum Policy Using Text Mining

Seo, Kyungsook

Doctoral Candidate, Department of Education, Ewha Womans University

Kim, Yihyun

Doctoral Candidate, Department of Education, Ewha Womans University

Lee, Deuk Gi

Master's Completion, Department of Education, Ewha Womans University

Lee, Sunbok

Professor, Department of Education, Ewha Womans University

This study analyzed news articles using a text mining technique to analyze the issues of the 2022 revised curriculum policy. Buzz analysis to identify media attention and keyword frequency analysis of news articles and Latent Dirichlet Allocation (LDA), a classic technique in topic modeling, was used to analyze policy issues. The results of the buzz analysis and keyword frequency analysis revealed the following: Firstly, major domestic media's interest in the policy was almost non-existent during the early and middle stages of policy formulation, but it surged before and after subject-specific public hearings., and Secondly, major media outlets in Korea focused on the controversy over the 'High School Credit System' and the 'Liberal Democracy Terms in History Textbooks'. The LDA analysis identified the key points of the revised 2022 education curriculum policy as follows: ① Full introduction of the High School Credit System, ② Controversy over the term 'Liberal Democracy' in history subjects, ③ Future-oriented curriculum and digital environment, ④ Expansion of elective subjects and reduction of national, English, and entrance exams, ⑤ Receipt of public opinions through the 'Communication Channel with Public Participation', ⑥ Concern over disruptions in education policy due to delayed launch of the National Education Committee, ⑦ Expansion of elementary and middle school information education hours and mandatory coding education, ⑧ Abolition of the game shutdown system and conversion to a part-time system. This research utilized text mining analysis techniques to confirm the issues of the revised 2022 education curriculum policy at different policy stages and demonstrated the usefulness of big data analysis.

Key Words: 2022 revised curriculum, text mining, topic modeling, LDA, news articles