



## 저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

교육학박사학위논문

빅데이터 분석 기법을 활용한  
2015 개정 교육과정 정책에 대한  
언론보도 분석

2017년 2월

서울대학교 대학원

교육학과 교육학전공

유 예 립



# 빅데이터 분석 기법을 활용한 2015 개정 교육과정 정책에 대한 언론보도 분석

지도교수 백 순 근

이 논문을 교육학박사 학위논문으로 제출함  
2016년 11월

서울대학교 대학원  
교육학과 교육학전공  
유 예 림

유예림의 박사 학위논문을 인준함  
2017년 1월

위 원 장	_____	(인)
부 위 원 장	_____	(인)
위 원	_____	(인)
위 원	_____	(인)
위 원	_____	(인)



## 국 문 초 록

각종 정보가 실시간으로 생성되고 유통되는 ‘빅데이터’(big data) 시대에서는 빅데이터 분석 기법을 활용하여 주요 교육정책에 대한 언론보도 내용을 분석하는 것이 중요하다. 이 연구는 우리나라 초·중등 교육정책의 핵심 중 하나인 2015 개정 교육과정 정책에 대한 언론보도 내용을 빅데이터 분석 기법을 활용하여 분석하는 것을 목적으로 한다. 이 연구의 연구 문제는 다음과 같다.

첫째, 2015 개정 교육과정 정책에 대한 언론 매체의 관심도는 정책의 형성·결정·집행 단계에 따라 변화하는가?

둘째, 2015 개정 교육과정 정책에 대한 뉴스 기사의 주제어는 무엇이며, 이는 정책의 형성·결정·집행 단계에 따라 변화하는가?

셋째, 2015 개정 교육과정 정책의 주요 쟁점은 무엇이며, 이는 정책의 형성·결정·집행 단계에 따라 변화하는가?

이러한 연구 문제를 해결하기 위해 우선, 국내 여론 영향력 점유율이 가장 높은 네이버 뉴스 홈페이지를 통해 보도된 2015 개정 교육과정 정책 관련 뉴스 기사 총 1,598건을 프로그래밍 언어인 파이썬(Python) 기반의 웹 스크래핑(web-scraping) 기법을 이용하여 자체적으로 수집하였다. 그리고 분석 목적에 부합하지 않는 기사를 제외한 총 1,512건의 최종 분석 데이터에 대해 전처리(preprocessing)와 형태소 분석(morphological analysis)을 실시하였다. 이후 빅데이터 기반의 자동화된 텍스트 분석(automated text analysis) 기법인 버즈(buzz) 분석과 주제어(keyword) 빈도 분석, 토픽 모델링(topic modeling)의 대표적 기법인 잠재 디리클레 할당(Latent Dirichlet Allocation: LDA)을 적용하여 분석하였다. 이 연구의 주요 결과는 다음과 같다.

첫째, 2015 개정 교육과정 정책에 대한 국내 언론 매체의 관심도는 정책의 형성·결정·집행 단계에 따라 변화하였다. 즉, 2015 개정 교육과정 정책의 형성 단계에서는 관련 보도가 거의 이뤄지지 않았고, 해당 정책이 결정되고 집행되기 직전과 국가적으로 역사교과서 국정화 논란이 극대화된 시점에서만 관련 내용을 집중적으로 보도하는 양상이 나타났다.

둘째, 국내 언론 매체의 뉴스 기사에 등장한 2015 개정 교육과정 정책의 주제어는 전체 분석 기간 동안 단순 단어 빈도(Term Frequency: TF)를 기준으로 분석한 경우 ‘교과서’, ‘교육부’, ‘역사’, ‘학생’, ‘국정’, ‘교과’, ‘고등학교’, ‘교원’, ‘초등학교’, ‘중학교’ 순으로 최상위 빈도를 나타냈다. 반면, 모든 뉴스 기사에서 자주 출현하는 상투어의 중요도를 감소시키기 위해 단어 빈도-역문서 빈도(TF-Inverse Document Frequency: TF-IDF)를 기준으로 분석한 경우에는 ‘역사’, ‘국정’, ‘한자’, ‘소프트웨어’, ‘수학’, ‘정부’, ‘과학’, ‘수업’, ‘수능’, ‘교육청’ 순으로 빈도가 높았다.

셋째, 국내 언론 매체의 뉴스 기사에 등장한 2015 개정 교육과정 정책의 주제어는 정책의 형성·결정·집행 단계에 따라 변화하는 것으로 나타났다. 즉, 2015 개정 교육정책의 형성 단계에서는 주제어의 빈도(TF-IDF 기준)가 ‘소프트웨어’, ‘수능’, ‘과학’, ‘기술’, ‘이수단위’ 순으로 높게 나타난 반면, 정책의 결정 단계에서는 ‘한자’, ‘국정’, ‘브리핑’, ‘정부’, ‘수학’이, 정책의 집행 단계에서는 ‘역사’, ‘국정’, ‘소프트웨어’, ‘교육청’, ‘한글’이 최상위 빈도를 나타냈다.

넷째, 국내 언론 매체를 통해 나타난 2015 개정 교육과정 정책의 주요 쟁점은 전체 분석 기간 동안 ① 고교 공통 과목 운영, ② 역사교과서 국정화 고시, ③ 현장교원 의견 수렴, ④ 초등 교과서 한자 병기, ⑤ 수능·대입 제도 개선, ⑥ 역사교육 세부 내용, ⑦ 소프트웨어 교육 강화, ⑧ 역사교과서 국정화 관련 정치적 대립, ⑨ 역사교과서 국정화 찬반 논란, ⑩ 역사교과서 집필진 구성, ⑪ 교수·학습 방법 개선, ⑫ 체험·인성교육 강화, ⑬ 교육과정 개정의 필요성, ⑭ 안전교육 강화, ⑮ 교육과정 지원체제 구축, ⑯ 자유학기제 운영(이상 문서수 기준 정렬)의 16개로 나타났다.

다섯 째, 국내 언론 매체의 뉴스 기사에서 나타난 2015 개정 교육과정 정책의 주요 쟁점은 정책의 형성·결정·집행 단계에 따라 변화하는 것으로 나타났다. 즉, 2015 개정 교육과정 정책의 형성 단계에서는 ‘과학교육 축소 우려’, ‘고교 공통 과목 도입’, ‘수능·대입제도 개선’, ‘소프트웨어 교육 확대’ 등의 5개 주요 쟁점이 나타난 반면, 정책의 결정 단계에서는 ‘고교 공통 과목 운영’, ‘초등 교과서 한자 병기’, ‘역사교과서 국정화 찬반 논란’ 등의 11개 주요 쟁점이, 정책의 집행 단계에서는 ‘교원의 전문성 개발’, ‘공교육 정상화’, ‘역사교과서 국정화 고시’ 등의 9개 쟁점이 나타났다.

이 연구는 빅데이터 분석 기법을 활용하여 주요 교육정책에 대한 분석을 새롭게 시도했다는 점에서 의의를 찾을 수 있다. 이 연구에서 사용한 빅데이터 분석 기법은 새로운 교육정책 평가 연구의 가능성을 제시하고 있는 만큼, 후속 연구를 통해 다양한 교육정책들로 분석 대상을 확대하고 관련 분석 기법을 더욱 정교하게 발전시킬 필요가 있다.

**\* 주요어 :** 빅데이터 분석, 2015 개정 교육과정, 언론보도 분석, 토픽 모델링, 잠재 디리클레 할당, 교육정책 평가

**\* 학 번 :** 2008-30394





## < 목 차 >

I. 서론 .....	1
1. 연구의 필요성 및 목적 .....	1
2. 연구 문제 .....	4
II. 이론적 배경 .....	5
1. 교육정책 평가를 위한 언론보도 분석 .....	5
가. 교육정책 평가의 개념 .....	5
나. 언론보도 분석을 통한 교육정책 평가 연구 .....	7
2. 2015 개정 교육과정 정책 .....	11
가. 2015 개정 교육과정 정책의 도입과 전개 과정 .....	11
나. 2015 개정 교육과정 정책의 주요 내용 .....	21
다. 2015 개정 교육과정 정책 관련 선행연구 분석 .....	47
3. 빅데이터 분석 기법 .....	54
가. 빅데이터의 개념과 특성 .....	54
나. 빅데이터의 분석 절차와 방법 .....	57
다. 빅데이터 분석 기법을 활용한 정책평가 .....	81
III. 연구 방법 .....	90
1. 분석 데이터 .....	90
가. 데이터 수집 .....	90
나. 데이터 클리닝 .....	98

2. 분석 방법 .....	103
가. 버즈 분석 .....	103
나. 주제어 빈도 분석 .....	103
다. 토픽 모델링 .....	105
<b>IV. 연구 결과 .....</b>	<b>110</b>
1. 2015 개정 교육과정 정책에 대한 언론 매체의 관심도 분석 .....	110
가. 정책 형성 단계 .....	111
나. 정책 결정 단계 .....	111
다. 정책 집행 단계 .....	112
2. 2015 개정 교육과정 정책 관련 언론보도의 주제어 분석 .....	114
가. 분석 기간 전체 .....	114
나. 정책 단계별 .....	116
3. 2015 개정 교육과정 정책의 주요 쟁점 분석 .....	119
가. 분석 기간 전체 .....	119
나. 정책 단계별 .....	128
<b>V. 요약 및 논의 .....</b>	<b>137</b>
1. 요약 .....	137
2. 논의 .....	140
참고문헌 .....	149
부    록 .....	171
Abstract .....	198

## < 표 목 차 >

<표 II-1> 언론보도 분석을 활용한 교육정책 평가 관련 선행연구 .....	9
<표 II-2> 2015 개정 교육과정이 추구하는 인간상과 핵심역량 .....	22
<표 II-3> 2015 개정 교육과정의 인간상과 학교 급별 교육 목표(온정덕, 2015 수정) .....	24
<표 II-4> 초등학교 교육과정 편제와 시간배당(교육부, 2015d) .....	26
<표 II-5> 중학교 교육과정 편제와 시간배당(교육부, 2015d) .....	28
<표 II-6> 고등학교 교육과정 편제와 시간배당(교육부, 2015d 수정) .....	31
<표 II-7> 고등학교 ‘통합사회’ 내용 체계(교육부, 2015f) .....	40
<표 II-8> 고등학교 ‘통합과학’ 내용 체계(교육부, 2015g) .....	42
<표 II-9> 2009 개정 교육과정과 2015 개정 교육과정의 비교 .....	45
<표 II-10> 언론보도 분석을 활용한 2015 개정 교육과정 관련 선행연구 .....	52
<표 II-11> 토픽 모델링 분석 도구 비교 .....	73
<표 II-12> 빅데이터를 활용한 교육정책 평가 관련 선행연구 .....	86
<표 III-1> 데이터 수집 단계별 뉴스 기사의 사례수 및 제외 기준 .....	97
<표 III-2> 분석 데이터의 정책 단계 및 매체 종류별 빈도(건수) .....	97
<표 IV-1> 뉴스 기사 내 상위 빈출 20개 단어(TF 기준): 전체 .....	114
<표 IV-2> 뉴스 기사 내 상위 빈출 20개 단어(TF-IDF 기준): 전체 .....	115
<표 IV-3> 뉴스 기사 내 상위 빈출 20개 단어(TF-IDF 기준): 정책 단계별 .....	117
<표 IV-4> LDA 분석 결과: 전체 .....	121
<표 IV-5> LDA 분석 결과: 정책 형성 단계 .....	129
<표 IV-6> LDA 분석 결과: 정책 결정 단계 .....	132
<표 IV-7> LDA 분석 결과: 정책 집행 단계 .....	135
<표 V-1> 2015 개정 교육과정 정책의 주요 쟁점 비교·분석 .....	142

## < 그 립 목 차 >

[그림 II-1] 2015 개정 교육과정 정책의 단계별 주요 추진 사안 .....	20
[그림 II-2] NCS 기반 전문 교과Ⅱ(군)의 구조(교육부, 2015a) .....	44
[그림 II-3] LDA 기법의 기본 가정(Blei, 2012) .....	66
[그림 II-4] LDA에 대한 그래프 모형(Blei, 2003) .....	67
[그림 II-5] 상위 모수 분포에 따른 LDA의 그래프 모형(김은희, 표신지, 김문철, 2012) .....	68
[그림 II-6] pyLDAvis를 활용한 LDA 모형의 시각화 결과 예시(1) .....	74
[그림 II-7] pyLDAvis를 활용한 LDA 모형의 시각화 결과 예시(2) .....	76
[그림 II-8] 텍스트 빅데이터의 분석 절차 및 방법 .....	80
[그림 II-9] 정책 단계별 빅데이터의 활용(송영조, 2012 수정) .....	85
[그림 III-1] 네이버 뉴스 홈페이지의 뉴스 기사 검색창 화면 .....	91
[그림 III-2] 1차 URL을 이용한 310개 웹문서의 HTML 정보 수집 code .....	92
[그림 III-3] 뉴스 기사의 ‘보도일자’, ‘언론사명’, ‘2차 URL’ 관련 HTML 정보 .....	93
[그림 III-4] 네이버 뉴스 홈페이지의 뉴스 기사 검색 결과 화면 .....	94
[그림 III-5] 뉴스 기사의 ‘본문’ 관련 HTML 정보 .....	95
[그림 III-6] 1, 2차 URL을 이용한 웹 스크래핑 code .....	96
[그림 III-7] 불용어 제거를 위한 Python code 예시 .....	99
[그림 III-8] 단어 정규화를 위한 Python code 예시 .....	100
[그림 III-9] 분리된 복합명사를 재조합하기 위한 Python code .....	102
[그림 III-10] 주제어 빈도 분석을 위한 R 프로그램 code .....	104
[그림 III-11] LDA 분석을 위한 Python의 주요 code .....	107
[그림 III-12] 토픽 응집성을 도표화하기 위한 Python code .....	108
[그림 III-13] 이 연구의 분석 절차 .....	109
[그림 IV-1] 뉴스 기사 보도 건수 월별 추이(총 1,512건 기준) .....	110
[그림 IV-2] 뉴스 기사의 고빈도 30개 주제어에 대한 워드 클라우드 .....	116
[그림 IV-3] TF-IDF에 따른 정책 단계별 워드 클라우드 .....	118
[그림 IV-4] LDA 모형의 토픽 수에 따른 복잡도: 전체 .....	119

[그림 IV-5] LDA 모형의 토픽 수에 따른 토픽 응집성: 전체 .....	120
[그림 IV-6] 토픽 13 관련 뉴스 기사의 내용 및 문서-토픽 확률값 예시 .....	122
[그림 IV-7] pyLDAvis를 활용한 LDA 모형의 시각화 결과(1) .....	125
[그림 IV-8] pyLDAvis를 활용한 LDA 모형의 시각화 결과(2) .....	126
[그림 IV-9] LDA 모형의 시각화 결과: 토픽 9와 토픽10 .....	127
[그림 IV-10] 토픽 수에 따른 LDA 모형의 복잡도: 정책 형성 단계 .....	128
[그림 IV-11] 토픽 수에 따른 LDA 모형의 토픽 응집성: 정책 형성 단계 .....	128
[그림 IV-12] 토픽 수에 따른 LDA 모형의 복잡도: 정책 결정 단계 .....	130
[그림 IV-13] 토픽 수에 따른 LDA 모형의 토픽 응집성: 정책 결정 단계 .....	131
[그림 IV-14] 토픽 수에 따른 LDA 모형의 복잡도: 정책 집행 단계 .....	134
[그림 IV-15] 토픽 수에 따른 LDA 모형의 토픽 응집성: 정책 집행 단계 .....	134
[그림 V-1] 네이버 통합 검색 빈도와 뉴스 기사 보도 건수의 추이 비교 .....	141

# I. 서론

## 1. 연구의 필요성 및 목적

21세기 지식정보화 사회가 도래하고 ‘알파고’(AlphaGo)로 대변되는 인공지능시대가 개막함에 따라 교육현장 역시 많은 변화를 요청받고 있다. 다양한 매체를 통해 언제 어디서나 필요한 지식과 정보를 손쉽게 획득할 수 있는 ‘빅데이터’(big data) 시대에서 지식을 단순히 습득하고 인출하는 방식의 교육 체제는 더 이상 큰 힘을 발휘할 수 없기 때문이다. 또한 과거에는 정책 의사결정자들이 관련 쟁점을 연구하고 적절한 대응 체제를 구축하는 데 충분한 시간이 있었다. 그러나 각종 정보가 실시간으로 생성·유통되는 제4차 산업혁명 시대에서 파괴적 혁신이 일으키는 변화에 빠르고 효율적으로 대처하기 위해서는 빅데이터 분석 기법과 같은 새로운 방법론(Schwab, 2016)을 활용하여 교육정책을 분석할 필요가 있다.

빅데이터는 ‘새로운 처리·분석 방법을 활용해 저비용으로 가치를 창출할 수 있는 대규모 데이터’로, 최근 각종 포털(portal)과 사회관계망서비스(Social Network Service: SNS)가 활성화되면서 텍스트(text) 빅데이터가 폭증하였다. 텍스트 빅데이터는 정책 수요자를 비롯한 다양한 이해당사자들의 관점과 의견을 적극 반영하고 있다. 때문에 정책 문제 식별 및 정책 형성과정 추적, 정책 결정·집행에 대한 실시간 모니터링, 정책 수용도 및 실효성 평가와 같은 정책평가 연구의 타당성을 확보하는 데 기여할 수 있다(이승룡 외, 2013; 이윤식, 2011). 즉, 텍스트 빅데이터는 다양한 이해당사자들의 의견을 폭 넓게 반영한 의사 결정과 즉각적 정책 대응에 기여할 수 있다는 점에서 정책평가에 유용하게 활용될 수 있다.

다양한 유형의 텍스트 빅데이터 중에서도 언론 매체의 뉴스 기사를 대규모로 수집한 ‘뉴스 빅데이터’는 정책평가에서 중요한 역할을 담당한다. 뉴스 기사는 정책 내용뿐 아니라 정책 과정에서 발생하는 문제점과 해결

방안 등을 다양한 관점에서 보도하므로, 정책에 대한 다양한 이해당사자들의 관심도, 관련 쟁점과 여론 등을 종합적으로 평가할 때 유용하기 때문이다. 특히 최근에는 뉴스 기사가 디지털 형태로 생산·유통되는 것이 보편화되어 자동화된 방식으로 뉴스 빅데이터를 구축하기 용이하고(박대민, 2013), 뉴스 빅데이터의 주요 공급원인 포털은 주요 언론사뿐 아니라 정부 및 정당의 보도자료, 정책 소비자가 직접 참여하는 대안 언론의 뉴스 기사를 모두 공급하기 때문에 다양성을 확보할 수 있다. 또한 뉴스 기사는 댓글이나 SNS 메시지 등의 다른 텍스트에 비해 형식을 갖춰 작성되었기 때문에 텍스트 분석에 유리하다.

아울러 최근 사회과학 분야에서는 텍스트 빅데이터에 대한 분석 기법으로 컴퓨터를 이용한 ‘자동화된 텍스트 내용 분석’(automated text content analysis)이 주목받고 있다(Hopkins & King, 2010; Iliev Dehghani, & Sagi, 2015; Lewis, Zamith, & Hermida, 2013). 이 기법은 텍스트 빅데이터를 일관적, 체계적으로 분석할 수 있을 뿐 아니라 데이터의 수집·가공·분석 과정이 알고리즘(algorithm) 형태로 공개되기 때문에 연구의 투명성과 재현가능성(reproducibility)을 확보(박종희, 박은정, 조동준, 2015)한다는 점에서 기존의 내용 분석 방법을 보완할 수 있다.

이에 따라 국내 교육 분야에서도 자동화된 텍스트 내용 분석을 활용하는 연구가 등장하고 있다. 예컨대, ‘텍스트 네트워크 분석’(text network analysis)<sup>1)</sup>과 ‘토픽 모델링’(topic modeling)을 활용해 교과서의 핵심 개념(고인규, 2016; 황홍섭, 박지수, 2016)이나 특정 분야의 연구 동향 및 수요(권기석, 박진수, 구찬동, 2013; 권현범, 2016), 정책의 쟁점 및 관련 여론을 분석(김우주, 구찬동, 2015; 김정원, 김지수, 최유림, 2014; 박선주,

---

1) 기존에도 이를 활용해 특정 분야의 연구 동향(강정배, 2015; 배상훈 외, 2013)이나 관련 인식(신세인 외, 2015), 정책 네트워크 및 쟁점을 분석한 연구(김수철, 이환철, 2014; 김정원 외, 2012; 류민영, 2013; 박수정, 2011; 이기종, 2015 등)는 다수 존재하였다. 그러나 이 연구들은 주로 연구자들의 직접 코딩(coding)에 의해 분석 가능한 비교적 소규모의 뉴스 기사, 공문서, 학술 자료를 대상으로 내용 분석을 실시했다는 점에서 빅데이터 분석과는 다르다.



2015; 최인봉, 이채희, 이은경, 2015)한 연구, ‘기계학습’(machine learning)을 활용해 대규모 평가의 서답형 문항 자동채점 프로그램을 개발한 연구(노은희, 성경희, 임은영, 2015)가 이에 해당한다.

이들 선행연구는 그동안 교육 분야에서 잘 활용되지 않았던 빅데이터 분석 기법을 활용해 기존 이론에서 간과되었던 새로운 변수나 변수 간 관계, 정책 여론 등을 탐색함으로써 교육 연구의 지평을 넓혔다는 점에서 의의가 있다(유진은, 2016). 그러나 대다수의 선행연구에서는 상용 서비스를 이용해 분석하였기 때문에 데이터의 수집, 가공, 분석 방법이 투명하게 공개되지 않았고, 교육학 분야의 특수성이 고려되지 않았으며, 전국적 여론 추이 분석에 머무른 경향이 있다(황명화, 2014). 또한 정보 접근과 분석 방법의 어려움으로 여전히 관련 연구는 부족한 실정이다(송영조, 2013; 송태민 외, 2013; 유예림, 백순근, 2016; 이승룡 외, 2013).

따라서 이 연구에서는 빅데이터 분석 기법을 활용하여 2015 개정 교육과정 정책에 대한 언론보도 내용을 분석하고자 하였다. 이 연구에서 2015 개정 교육과정 정책<sup>2)</sup>을 분석 대상으로 선정한 이유는 다음과 같다. 첫째, 이 정책은 창의·융합교육을 위해 교과별 교수·학습·평가 방법과 편제, 성취기준 등을 조정할 뿐 아니라, 이와 연계해 교과서, 대학 입시, 교원양성·연수 등의 제반 교육정책을 종합적으로 개선한다(교육부, 2015a)는 점에서 중요한 의미를 갖는다. 둘째, 정책 과정에서 교육과정 개발에 대한 무관심, 이기주의 등은 합리적 의사 결정을 방해하기도 하는데(백승훈, 남상준, 2009), 특히 2015 개정 교육과정 정책은 교과 통합 및 신설을 통해 교육 내용의 재구조화와 학습량 적정화를 시도하고 있어 이러한 문제가 더욱 발생하기 쉽다. 따라서 2015 개정 교육과정 정책을 타당하고 효과적으로 집행하기 위해서는 이해당사자들의 관심도와 관련 쟁점들을 즉시적·지속적으로 파악하는 정책평가 연구가 반드시 필요하다.

---

2) 국가 교육과정은 사회적 수준에서 개발·결정되고, 국가·공공기관의 자원이 투입돼 운영의 전 과정이 정책평가의 대상이 된다는 점에서 ‘교육정책’으로 간주될 수 있다(김용, 2005).

## 2. 연구 문제

이 연구는 빅데이터 분석 기법을 활용하여 초·중등 교육정책의 핵심 중 하나인 2015 개정 교육과정 정책을 과정 평가(process evaluation)의 측면에서 평가하기 위한 것이다. 이를 위해 교육정책의 공급자, 수요자, 전문가, 교육당국의 다양한 입장이 반영되어 있는 국내 인터넷 포털의 뉴스 기사를 활용하여 2015 개정 교육과정 정책에 대한 언론 매체의 관심도와 뉴스 기사의 주제어를 파악하고, 언론 보도를 통해 나타난 2015 개정 교육과정 정책의 주요 쟁점을 분석하고자 하였다. 이 연구에서 다루고자 하는 연구 문제들을 구체적으로 진술하면 다음과 같다.

첫째, 2015 개정 교육과정 정책에 대한 언론 매체의 관심도는 정책의 형성·결정·집행 단계에 따라 변화하는가?

둘째, 2015 개정 교육과정 정책에 대한 뉴스 기사의 주제어는 무엇이며, 이는 정책의 형성·결정·집행 단계에 따라 변화하는가?

셋째, 2015 개정 교육과정 정책의 주요 쟁점은 무엇이며, 이는 정책의 형성·결정·집행 단계에 따라 변화하는가?

## Ⅱ. 이론적 배경

### 1. 교육정책 평가를 위한 언론보도 분석

교육 분야의 정보기술 동향에 대한 연간 보고서를 발행하는 미국의 뉴미디어컨소시엄(The New Media Consortium: NMC)은 향후 5년 간 파급력이 클 것으로 예상되는 동향으로 ‘데이터 기반 학습 및 평가의 증가’를 지목하였다(Johnson et al., 2014). 이 절에서는 데이터 기반의 정책평가 접근으로서 언론보도 분석을 활용한 교육정책 평가에 관해 정리하였다.

#### 가. 교육정책 평가의 개념

교육정책 평가(教育政策 評價, educational policy evaluation 또는 policy evaluation in education)는 교육평가의 하위 영역 중 하나로, ① 일반적인 ‘정책평가’의 개념을 교육의 영역으로 확장하거나, ② 교육의 특수성을 고려한 ‘교육평가’의 개념에 교육정책을 적용함으로써 정의할 수 있다(유예림, 2008). 우선 ‘정책평가’(policy evaluation)는 “정책의 내용, 집행 및 그 영향을 추정하거나 평정하기 위해 체계적 연구방법론을 응용하는 것으로, 특정 정책의 과정이나 결과를 이해하고 그 값어치를 판단하는 사회적 과정”으로 정의할 수 있다(노화준, 2015: 4). 흔히 정책평가는 정책 집행이 종료된 이후 효과성을 분석하는 것으로 인식되고 있다. 그러나 넓은 의미<sup>3)</sup>에서 정책평가는 정책에 대한 사후 평가뿐 아니라 정책 수요 분석, 최적화 분석, 집행과정 평가 등 정책의 전(全) 과정에

---

3) 좁은 의미에서는 ‘정책평가’를 경험적 자료를 통해 정책 집행의 효과와 적합성 등을 검토하는 사후적·회고적 분석으로, ‘정책분석’은 정책 결정에 필요한 정보를 산출하기 위한 사전적·조망적 분석으로 구분하기도 한다(노화준, 2015).

걸쳐 이뤄지는 활동이라고 할 수 있다(이종수, 2009). 예컨대, Anderson(2015)은 정책평가를 ‘정책 내용, 집행, 효과 등을 추정·사정·평가하는 것’으로 정의하였고, 정정길 외(2004)는 ‘모든 정책 과정에 대한 가치판단적 논의’라고 하였다. 김신복(1983) 역시 정책평가는 광의적으로 정책을 집행하는 도중에 원래의 의도나 지침대로 추진되고 있는지를 점검하는 과정 평가(process evaluation)까지 포함한다고 하였다.

한편 교육평가(educational evaluation)는 ‘교육 관련 특성이나 현상에 대해 가치를 따져 값을 매기는 전문적 활동이며, 동시에 그 값을 더 높이기 위한 개선·발전 방안을 탐색하는 체계적 활동’으로 정의할 수 있다(백순근, 2006). 전통적으로 ‘교육평가’는 ① 교육목표의 달성도를 밝히는 과정(Tyler, 1949), ② 교육 관련 정보를 수집·분석해 교육적 의사결정에 도움을 주는 활동(황정규, 2005; Cronbach, 1982; Stufflebeam, 1968), ③ 교육 관련 대상의 장점이나 가치를 판단하는 활동(박도순, 홍후조, 2006; 백순근, 2002; House, 1980; Scriven, 1973)이라는 세 가지 관점으로 개념화되었다(채선희, 1997). 그리고 ‘교육정책’은 정치적 과정을 통해 이뤄지는 중요한 의사결정으로, 교육 목적을 실현하기 위한 총체적 지침(윤정일 외, 2015)이라고 할 수 있다. 이상의 내용을 종합하면, ‘교육정책 평가’는 교육정책의 최종 목표 달성도뿐 아니라 매 단계 의사결정에 도움을 주기 위해 관련 여론과 쟁점 등의 정보를 수집·분석하며, 이를 통해 해당 정책의 장점이나 한계, 가치 등을 판단하는 활동이라고 할 수 있다.

요컨대, 교육정책 평가는 ‘교육정책을 개선·발전시키기 위해 정책의 형성·결정·집행 과정이나 그 결과 및 파급효과 등 교육정책의 전 과정을 체계적으로 평가하는 활동’이라고 정의할 수 있다. 궁극적으로 교육정책 평가는 교육 수요자의 요구와 필요에 부응하고, 교육정책의 다양화·전문화·특성화를 유도하는 동시에, 질 높은 교육을 실천하려는 데 기본 취지가 있다(한경수 외, 1998). 특히 정책 관련 이해당사자들의 이해가 상충하여 정책의 흐름이 올바르게 흐르지 못할 때, 이들을 설득하고 올바른 방향으로 정책을 이끌고자 할 때 교육정책 평가는 필수적으로 요청된다.

## 나. 언론보도 분석을 통한 교육정책 평가 연구

인터넷과 SNS 등의 새로운 정보 기술이 도래하고 정책 환경이 다원화됨에 따라 대중의 정치사회 참여 방식이 다양해지면서 정책의 형성·결정·집행 단계에서 여러 가지 쟁점들이 발생하고 있다. 이러한 쟁점들은 관련 분야에서 개혁을 단행하는 하나의 자극제가 되기도 하지만(Zhou & Moy, 2007), 쟁점이 부정적 혹은 감정적으로 치우쳐 확산되다 보면 이해당사자 간 소모적 갈등을 야기하여 정책을 효과적으로 실행하는 데 어려움을 겪기도 한다(문빛, 김찬석, 이철한, 2013). 따라서 타당한 교육정책 수립 및 효과적 집행을 목적으로 하는 교육정책 평가 연구에서 특정 정책에 대한 쟁점 분석은 중요한 의미를 갖는다.

또한 정책 쟁점은 언론 매체와 여론 공간이 상호작용하면서 형성·발전되고, 다양한 사회·정치적 맥락과 결부되어 확대 재생산되는 경향이 있다(왕재선, 김선희, 2013). 즉, 정책 쟁점은 정책 과정(policy process)에 따라 동태적(dynamic) 특성을 갖고 있기 때문에 정책의 형성·결정·집행 단계에 따른 쟁점 추이 분석이 필수적이다. 쟁점에 대한 추이 분석은 쟁점에 대한 즉각적 대응뿐 아니라 쟁점이 위기로 발전할 가능성을 차단하는 사전 예방적 접근(김찬석, 2012)으로 활용될 수 있기 때문이다. 그러나 지금까지 수행된 대부분의 쟁점 분석 연구는 쟁점을 특정 시점에 한정하여 제한적으로 도출하거나, 정책이 형성·결정된 이후 회고적으로 분석하는 경우가 많았다. 따라서 쟁점의 변화 과정을 즉시적으로 파악하고 이에 대해 선제 처방을 내림으로써 쟁점 연구의 실효성을 높일 필요가 있다(문빛, 김찬석, 이철한, 2013).

정책 관련 쟁점을 분석할 때 대표적으로 활용되는 분석 데이터 중 하나는 언론 매체의 뉴스 기사다. 언론 매체는 정보 전달과 비판의 기능을 통해 정책 과정 전반에 영향을 주기 때문이다(윤정일 외, 2015). 즉, 언론 매체의 뉴스 기사는 정책의 주요 내용을 국민에게 전달함으로써 해당 정책

에 대한 이들의 관심도와 이해도에 영향을 미친다. 그리고 이를 바탕으로 정책 여론이 형성되면, 언론 매체는 뉴스 기사를 통해 관련 쟁점과 논의들을 심층보도하거나 사실 및 칼럼 등을 통해 의견을 개진함으로써 정책의 형성과 결정 단계에 영향을 미친다(유예림, 백순근, 2016). 윤정일 외(2015)에 따르면, 언론 매체는 정부와 국민과의 매개 역할을 통해 정책의 수정 및 정교화 등 교육정책 형성 단계에 강력한 영향을 미치고, 비판 기능을 통해 정책 대안을 제시하는 등 정책 결정 단계에 영향을 미친다.

이러한 언론 매체의 역할에 주목하여 그동안 교육 분야에서도 뉴스 기사에 대한 텍스트 분석 연구(강정배, 2015; 김병주, 김태완, 김은아, 2006; 김수철, 이환철, 2014; 류민영, 2013; 문성호, 윤동엽, 2011; 박수정, 2011; 배상훈 외, 2013; 신세인 외, 2015; 양길석, 2010; 이기중, 2015 등)가 다수 수행되었다(선행연구들의 분석 데이터, 내용 및 방법은 <표 II-1> 참고). 예컨대, 김병주, 김태완, 김은아(2006)는 국가의 교육정책 결정 과정에서 언론 매체의 올바른 방향 설정과 비판 기능의 필요성을 제기하며 교원평가제에 대한 국내 주요 6대 일간지의 보도태도를 평가 주체, 평가 태도, 평가 내용, 평가 방법, 평가 결과라는 준거에 따라 분석하였다. 분석 결과, 언론 매체는 교원평가제에 대해 단순 사실 보도 및 그 중요성을 홍보하는 수준에 그치고 있었으며, 향후 언론 매체는 교육정책 관련 기사수를 단순히 늘리기보다는 심층보도를 통해 제도에 대한 올바른 시각과 대안 제시 등을 보도할 필요가 있음을 강조하였다. 국내 3개 언론사의 학교 폭력 관련 신문기사의 형식과 내용을 분석한 문성호와 윤동엽(2011)의 연구 역시 각 언론 매체가 정보 제공과 문제적 접근에 치중함으로써 교육정책에 대한 구체적 대안 제시나 심층 분석은 결여되었다고 지적하였다.

또한 양길석(2010)은 쟁점 분석을 ‘문제를 보는 시각과 문제 해결 방안에 대한 대립되는 입장을 살펴보는 것’으로 정의하고, 이를 위해 신문 사설을 활용하였다. 즉, 국내 2개 일간지에서 보도한 수능 관련 사설을 중심으로 수능의 주제 영역별·연도별 경향과 주요 문제 및 논지를 분석하고 관련 쟁점을 도출하였다. 한편 박수정(2011)은 국내 4개 일간지에 보도된

교육감 선거 관련 사실에 대해 네트워크 텍스트 분석(network text analysis)을 실시함으로써 해당 보도에서 나타난 주제어(keyword)를 통해 관련 정책에 대한 언론 매체의 보도 경향을 분석하였다.

<표 II-1> 언론보도 분석을 활용한 교육정책 평가 관련 선행연구

연구자 (연도)	분석 데이터	연구 내용	분석 방법
김병주, 김태완, 김은아 (2006)	국내 6개 일간지 기사 637건	<ul style="list-style-type: none"> <li>· 교원평가제 관련 신문의 보도 태도 분석</li> <li>- 기사 유형별: 뉴스, 사설, 칼럼, 기획/연재, 인터뷰, 통계/설문조사</li> <li>- 쟁점별: 평가 주체, 의도, 내용, 방법, 결과 활용</li> <li>- 논조별: 긍정, 중립, 부정</li> </ul>	<ul style="list-style-type: none"> <li>· 기사유형별 빈도 분석</li> <li>· 연구자의 직접 코딩에 의한 내용 분석</li> </ul>
양길석 (2010)	국내 2개 일간지 사설 93건	<ul style="list-style-type: none"> <li>· 수능 관련 사실에 나타난 주요 관심 문제와 논제, 쟁점 분석</li> <li>- 주제영역별: 수능 성격 및 체제, 출제·시행관리, 난이도, 점수체제</li> <li>- 매체별: 보수, 진보언론</li> <li>- 연도별: 1993년~2009년</li> </ul>	<ul style="list-style-type: none"> <li>· 주제영역별, 연도별 빈도분석</li> <li>· 연구자의 직접 코딩에 의한 내용 분석</li> </ul>
문성호, 윤동엽 (2011)	국내 3개 일간지 기사 212건	<ul style="list-style-type: none"> <li>· 학교폭력 관련 기사의 형식 및 내용 분석</li> <li>- 기사 형식별: 유형(스트레이트, 통계·가십, 해설, 의견), 형태(단순사실, 기자 의견, 정보원 의견, 혼합), 작성자(기자, 전문가, 독자, 미상), 의도(정보제공, 문제제기, 고발, 발굴)</li> <li>- 단순사실 기사의 서술방식별: 초점(특정 사건, 정책), 출처(개인·학교, 지역, 국가), 정보경향(심각, 악화, 비판, 극복), 정보원(교사, 학생, 전문가, 부모)</li> </ul>	<ul style="list-style-type: none"> <li>· 기사형식별, 빈도 분석</li> <li>· 연구자의 직접 코딩에 의한 내용 분석</li> </ul>

연구자 (연도)	분석 데이터	연구 내용	분석 방법
박수정 (2011)	국내 4개 일간지 사설 64건	<ul style="list-style-type: none"> <li>· 교육감 선거 관련 사설 제목 및 내용에 나타난 특징 분석</li> <li>- 시점별: 선거 전, 선거 후</li> <li>- 매체별: 보수, 진보언론</li> </ul>	<ul style="list-style-type: none"> <li>· 주제어 빈도 분석</li> <li>· KrKwic, Netminer를 활용한 네트워크 텍스트 분석</li> </ul>
류민영 (2013)	국내 8개 일간지 기사 2,000건	<ul style="list-style-type: none"> <li>· 교원평가 정책 네트워크 특성에 따른 정책 과정 분석</li> <li>- 정책 네트워크 차원별: 구성원, 통합성, 자원, 권력</li> <li>- 연도별: 2000년~2012년</li> </ul>	<ul style="list-style-type: none"> <li>· 연도별 빈도분석</li> <li>· 연구자의 직접 코딩 및 Ucinet을 활용한 사회 연결망 분석</li> </ul>
김수철, 이환철 (2014)	네이버 를 통해 보도된 기사 120건	<ul style="list-style-type: none"> <li>· 스토리텔링 수학 관련 기사의 외형적 특성 및 보도프레임 분석</li> <li>- 기사 유형별: 일간지, 뉴스, 잡지, 연재기사</li> <li>- 취재 대상별: 학생·학부모, 교수·교사, 교육부·교육청, 출판사·학원, 수학 관련 내용</li> <li>- 프레임별: 이해, 교육방법, 논조(긍정/부정/혼합), 강조, 상업광고</li> </ul>	<ul style="list-style-type: none"> <li>· 기사유형별, 취재 대상별 빈도 분석</li> <li>· 연구자의 직접 코딩에 의한 내용 분석</li> </ul>

언론보도 분석을 활용한 교육정책 평가 관련 선행연구에서는 주요 일간지 기사들을 대상으로 연구자들의 직접적 코딩에 근거하여 내용 분석(content analysis)을 실시한 경우가 많았다. 그러나 최근에는 검색포털, SNS 등과 같은 다양한 의사소통망이 발달함에 따라, 이를 통해 생성·유통된 정책 쟁점과 관련 의견들을 텍스트 데이터 형태로 추가 확보하여 정책 여론을 전반적으로 조망할 필요가 있다. 그런데 이러한 정보량의 증가는 정보 처리에 부담을 주어 생략, 처리 실수, 지연, 여과, 대충 훑어보기 등의 정보 처리 문제를 발생시킨다(박병량, 주철안, 2012). 따라서 대규모의 뉴스 기사(또는 뉴스 빅데이터)를 활용하여 언론보도를 통해 나타난 정책 쟁점을 분석하기 위해서는 자동화된 텍스트 내용 분석(automated text content analysis)과 같은 새로운 분석 기법의 도입이 요청된다(이와 관련된 자세한 사항은 ‘3. 빅데이터 분석 기법’ 참고).



## 2. 2015 개정 교육과정 정책

### 가. 2015 개정 교육과정 정책의 도입과 전개 과정

이 절에서는 2015 개정 교육과정 정책의 도입 및 전개 과정을 체계적으로 개관하기 위해 이를 일반적인 정책 과정인 ① 정책 형성 단계, ② 정책 결정 단계, ③ 정책 집행 단계의 세 단계로 구분하여 정리하였다.

#### 1) 정책 형성 단계

‘정책 형성’(policy formation) 단계는 특정 정책을 결정하기에 앞서 논의 대상이 될 수 있는 정책 의제(policy agenda)를 채택하고, 무엇을 정책 문제의 핵심으로 할 것인가 하는 목표를 설정한 후, 그 목표에 따라 주어진 문제를 해결하기 위한 대안들을 탐색·개발하여 각각의 효과성과 실현 가능성 등을 비교·분석하는 단계이다(이종수, 2009). 이에 비추어 볼 때, 2015 개정 교육과정 정책의 형성 단계는 국가와 사회가 현행 2009 개정 교육과정의 문제점과 한계를 인식하고, 이를 바탕으로 새로운 국가교육과정 정책 마련의 필요성을 역설하며 그에 대한 청사진을 모색하는 교육과정 개정의 추진 배경에 해당된다고 할 수 있다.

우리나라의 국가교육과정은 다양한 국가·사회적 요구, 과학기술과 학문의 발전, 교육 여건과 환경의 변화 등에 효과적으로 대응하기 위해 지속적으로 개정되어 왔다(김경자 외, 2015). 현행 2009 개정 교육과정 역시 미래역량 중심의 교육에 대한 시대적 요청을 반영하여 학습자의 자율성과 창의성을 신장하기 위해 추진되었는데(교육과학기술부, 2009), 2009 개정 교육과정 체제에서는 여전히 학습량이 많아 수업 개선이 어려운 점, 교과 간 연계성과 시수 부족으로 융합이나 통합이 어려운 점 등의 문제가 지속적으로 제기되어 왔다(김경자 외, 2015).

개정 작업이 진행 중이던 2008년 11월, 국가과학기술위원회 운영위원회는 창조적 융합기술 전문 인력 양성을 골자로 한 「국가융합기술발전 기본계획('09~'13)」을 발표함으로써 ‘창의·융합형 인재 양성’을 요청하였다. 그러나 이러한 요구는 발표의 시점 상 2009년 12월 고시된 2009 개정 교육과정에 즉각적으로 반영되지는 못하였다. 교육과학기술부는 2010년 12월, 융합인재교육(STEAM)이라는 교수·학습 접근을 통해 부분적으로 대응하였으나(이경진, 김정자, 2012), 충분한 효과를 거두지 못하였고 인재상이 달라졌다는 점에서 전체 교육과정의 변화가 불가피해졌다(김정자 외, 2015). 즉, 2009 개정 교육과정 체제에서는 여전히 학습량이 많아 수업을 개선하기 어렵고, 교과 간 연계성과 시수 부족으로 융합 및 통합교육을 실시하기 힘들다는 문제가 제기됨에 따라(김정자 외, 2015) 새로운 국가교육과정의 마련을 통해 현행 교육과정 실행의 문제점을 보완할 필요가 있다는 주장이 나타나기 시작하였다.

이러한 국가교육과정 개정에 대한 논의는 2013년 8월 27일 발표된 「학생·학부모 부담 완화와 학교교육 정상화를 위한 대입전형 간소화 및 대입제도 발전방안(시안)」을 통해 재점화되었다고 할 수 있다. 2017학년도 대학수학능력시험(이하 ‘수능’) 체제를 검토하기 위해 마련된 이 시안에서는 ‘문·이과 완전 융합안’이 처음 언급되었는데, 이는 “창의적, 융·복합적 인재 양성을 위해 모든 학생이 균형 있는 학습을 할 수 있도록 문·이과로 구분되어 있는 수능시험 체제를 개선하는 방안”(교육부, 대입제도 발전방안 연구위원회, 2013: 18)으로, 모든 학생이 공통으로 학습해야 하는 과목을 평가하기 위해서는 고등학교 교육과정의 개정이 필요하다고 진술되었다. 비록 2013년 10월 24일 발표된 2017학년도 대입제도 확정안에서는 “2009 개정 교육과정 내에서의 운영 가능성, 제도의 안정성, 학생·학부모 부담 경감 측면”(교육부, 2013: 2)에서 현행 방식을 유지하는 것으로 결정되었지만, 융합형 인재 육성의 필요성 및 의견 수렴 기간에 나타난 문·이과 융합에 대한 폭넓은 지지를 고려”(교육부, 2013: 2)하여 2013년 말부터 본격적인 교육과정 개정이 착수되었다.

우선, 2013년 11월부터 약 4개월 간 한국교육과정평가원에서는 ‘국가 교육과정 총론 개선을 위한 기초 연구’(박순경 외, 2014)가 진행되었다. 2014년 1월에는 교육부가 ‘국가교육과정 정책자문위원회’를 구성·운영하고, ‘교육과정 심의위원회’ 위원 등이 참석한 ‘문·이과 통합형 교육과정 개발에 관한 전문가 대토론회’와 한국교육과정학회가 주최한 ‘국가교육과정 포럼’ 등을 통해 개정 방향에 관한 각계의 의견을 수렴하기 시작하였다. 이를 바탕으로 2월 13일 교육부 업무보고를 통해 발표된 「문·이과 통합형 교육과정 기본방향(안)」에서는 “모든 학생이 최소한의 인문·사회·과학 교과를 학습하고 창의·융합형 인재로 자라도록 지원”(교육부, 2014a: 3)한다는 원론적 수준의 방향이 제시되었다(황규호 외, 2014a).

특히 교육부는 2월 26일 ‘국가교육과정 개정연구위원회’를 구성하여 ‘문·이과 통합형 교육과정 총론 시안 개발 연구(총괄)’를 비롯한 총 11개의 관련 정책연구 과제를 수행하고 관련 의견 수렴을 진행하도록 하였다(김경자 외, 2015). 이때 교육부와 국가교육과정 개정연구위원회는 개정 교육과정의 비전을 “미래사회가 요구하는 창의·융합 인재 양성과 학습경험의 질 개선을 통한 행복한 학습의 구현으로 결정하였는데, 전자는 사회적 요구차원에서 문·이과 통합형 교육과정을 지향하는 것이고, 후자는 그동안 추진된 교육과정 개정의 연속선상에서 학습자의 행복교육을 실현하기 위한 것”(김경자, 2015: 8)이라고 할 수 있다.

이후에는 정책연구를 통해 설계된 개정 교육과정의 요소와 예상되는 문제점 및 개선 방안 등을 탐색하기 위해 국가 차원의 각종 의견 수렴이 수시로 이어졌다(김경자 외, 2015). 즉, 2014년 6월부터 추가적인 ‘국가교육과정 포럼’을 통해 현장 교원 및 학계를 중심으로 의견 수렴이 수차례 이루어졌으며, 8월부터는 ‘국가교육과정개정 자문위원회’와 시·도별 전문직 및 핵심교원을 대상으로 의견 수렴이 이루어졌다. 또한 8월과 9월 중에는 이러한 과정을 통해 마련된 개정 교육과정 총론 시안을 최종적으로 검토하기 위해 두 차례에 걸쳐 공청회가 진행되었고, 학교급별 교육 전문가 등으로 구성된 교육과정 심의위원회의 검토 역시 이루어졌다.

## 2) 정책 결정 단계

‘정책 결정’(policy-making) 단계는 정책 형성 단계를 통해 제시된 여러 가지 대안들을 탐색하고, 다양한 이해당사들 간의 논의를 바탕으로 이를 비교·평가함으로써 정책을 확정해 가는 단계다(이종수, 2009). 이에 비추어 볼 때, 2015 개정 교육과정 정책의 결정 단계는 정책 형성 단계를 통해 모색된 교육과정 총론 주요 사항을 구체화하고, 이에 근거하여 마련된 교과별 교육과정 시안을 다양한 이해당사자들이 함께 검토한 후 최종안을 채택하는 단계라고 할 수 있다. 2015 개정 교육과정 정책의 결정 단계를 추진된 사안을 중심으로 정리하면 다음과 같다.

2014년 9월 24일, 교육부는 ‘2015 문·이과 통합형 교육과정’이라는 이름으로 교육과정 총론 주요사항을 발표하였다. 즉, “2009 개정 교육과정이 추구한 ‘창의적 인재 양성’의 기본 정신을 유지하되, 현장 적용 과정에서 제기된 문제점을 개선하고, 교육환경 변화에 적극적으로 대응하기 위해 미래 사회가 요구하는 창의·융합형 인재를 양성”(교육부, 2014b: 5)하기 위한 새로운 국가교육과정을 마련한다는 것이다. 총론 주요사항에서는 창의·융합형 인재를 “인문학적 상상력, 과학기술 창조력을 갖추고 바른 인성을 겸비하여 새로운 지식을 창조하고 다양한 지식을 융합하여 새로운 가치를 창출할 수 있는 사람”(교육부, 2014c: 2)으로 정의하였다. 또한 ① 초등학교는 누리과정과의 연계를 강화하고, ② 중학교는 자유학기제의 취지가 전체 교육과정에 반영될 수 있도록 학습내용을 적정화하며, ③ 고등학교는 모든 학생이 필수적으로 배워야 할 ‘공통 과목’을 구성해 기초소양을 함양하되 내용과 수준은 적정화할 것을 강조하였다.

이후에는 총론 주요사항을 바탕으로 2015 개정 교육과정 총론 및 24개 교과별 개정 교육과정 시안 개발 연구가 진행되었다. 이때 교과별 개정 교육과정 시안은 ① 교과목표 및 성취기준의 제시를 위한 1차 시안과, 이를 바탕으로 하되 ② 교수·학습 및 평가 방법 개발에 주안을 둔 2차 시안으로 구분된다(박철웅 외, 2015). 1차 시안은 2014년 12월부터 2015

년 4월까지 수차례의 연구진 합동 워크숍과 내용 조정 협의회, 내용 타당성 검토 및 개선안 협의회, 내용 구성안에 대한 공개 토론회 등을 거쳐 마련되었다(구정화 외, 2015). 특히 3월에는 인문·사회, 과학기술, 체육·예술의 3개 분과별로 각계 인사와 교육과정 전문가, 현장 교원 등이 참여하는 ‘국가교육과정 각론조정위원회’가 구성되었다(교육부, 2015a). 국가교육과정 각론조정위원회에서는 총론과 교과 교육과정 간의 연계성과 일관성을 확보하고, 교과별 내용 중복을 해소하며, 교과 간 이해관계를 조정하기도 하였다(김경자 외, 2015). 2차 시안은 5월부터 8월까지 개발되었으며, 개발 과정은 1차 시안과 유사하다. 2월부터 11월까지는 교사와 학부모를 대상으로 개정 교육과정의 현장 적합성과 실행 관련 의견을 묻는 설문조사가 진행되기도 하였다.

이와 같은 과정을 거쳐 개발된 2015 개정 교육과정 총론 및 교과 교육과정 시안을 최종적으로 검토하기 위해 2015년 8월 6일과 9월 4일, 두 차례에 걸쳐 공청회가 진행되었다. 또한 2015년 9월에는 총 21회에 걸친 국가교육과정 심의위원회의 검토를 받아 2015년 9월 23일, ‘2015 개정 교육과정’이라는 명칭으로 개정 교육과정의 총론과 교과 교육과정이 확정·고시되었다(교육부 고시 제2015-74호, 관련 내용은 ‘나. 2015 개정 교육과정 정책의 주요 내용’ 참고).

### 3) 정책 집행 단계

‘정책 집행’(policy implementation) 단계는 정책 수단들을 마련하고 동원함으로써 정책을 실질적으로 집행하는 과정으로, 일련의 행정적, 정치적 결정과 활동으로 이뤄진다(이종수, 2009). 이에 비추어 볼 때, 2015 개정 교육과정 정책의 집행 단계는 정책 결정 단계를 통해 확정된 개정 교육과정의 총론과 교과 교육과정을 실제 교육 현장에서 효과적으로 실천(practice)하기 위해 ① 새로운 교과서와 교사용 지도서(이하 ‘교과용도서’), 교육과정 해설서 등의 관련 기반을 준비하고, ② 대입제도 및 교원

양성·연수 등의 지원 체제를 개편하여 ③ 제반 정책을 모두 실행하는 단계라고 할 수 있다.<sup>4)</sup> 2015 개정 교육과정 정책의 집행 단계를 추진된 사안을 중심으로 정리하면 다음과 같다.

우선, 2015 개정 교육과정이 교육현장에서 실행되기 위해서는 교과용 도서와 교육과정 해설서가 새로 마련되어야 한다. 이와 관련하여 2015년 7월, 교육부는 국정도서 ‘실험본’의 명칭을 ‘현장검토본’으로 변경하고, 검정도서의 집필 기간을 최소 1년 이상으로 늘리는 것을 골자로 한 ‘교과용도서 개발 체제 개선 방안’을 발표하였다(교육부, 2015e). 또한 10월에는 「초등학교 교과용도서 국·검·인정 구분」(교육부 고시 제2015-76호)이, 11월에는 10월에 행정예고된 「중·고등학교 교과용도서 국·검·인정 구분」(교육부 고시 제2015-78호)이 확정·고시되었다. 후자의 경우 중학교 『역사』와 고등학교 『한국사』 교과서 발행체제를 현행 검정에서 국정으로 전환하는 내용이 포함되어 있다. 한편 2016년 6월에는 「초·중등학교 디지털교과서 국·검정 구분(안)」이 행정예고되었고, 8월에는 해당안이 확정·고시되었다(교육부 고시 제2016-98호). 이에 따라 2015 개정 교육과정에서는 초등학교 3학년~중학교 3학년의 사회(역사/도덕 포함), 과학(중학교 기술·가정/정보 포함), 영어 교과와 고등학교 영어(검정 5종)교과가 디지털교과서 형태로 개발·보급될 예정이다(교육부, 2016c).

2015 개정 교육과정의 교과용도서는 2017/2018/2019년 대상 도서로 구분되어 있으며, 이는 집필진 공모, 개발, 심사, 선정 등의 과정을 거쳐 최종 확정된다. 예컨대, 2015년 12월 고시된 「초·중등학교 교육과정 일부 개정」(교육부 고시 제2015-80호)에 따라 2017년 대상 국정 도서로 예정되었던 중학교 『역사』와 고등학교 『한국사』 교과서의 경우 2015년 9

---

4) 교육과정 분야에서 교육과정의 ‘집행’ 또는 ‘실행’(implementation)은 “교육과정의 실제적 사용 혹은 교육과정이 실제로 이루어지는 것”(Fullan & Pomfret, 1977; 소경희, 2003, p. 131에서 재인용)을 의미한다. 그러나 이 연구의 분석 대상은 ‘교육과정’이 아닌 ‘교육과정 정책’이기 때문에, 이 연구에서 ‘2015 개정 교육과정 정책의 집행 단계’라는 용어는 2015 개정 교육과정의 실행기와 실행준비기를 모두 포함하는 의미로 사용되었다.

월부터 편찬기관 및 집필자 공모와 심의위원 모집, 집필 등이 이루어져 2016년 12월, 현장 검토본을 공개하였다. 그리고 현장 교원과 국민이 제출한 의견, 국사편찬심의회의 심의 내용을 바탕으로 2017년 1월에는 최종 완성본이 마련되었으며(교육부, 2016e), 2017학년도에는 해당 교과서를 희망 연구학교에 한해 사용함으로써 국·검정 도서를 혼용할 예정이다<sup>5)</sup>(교육부, 2017b). 또한 2017년 대상 검·인정 도서(초등학교 3, 4학년, 중학교 1학년, 고등학교 검정도서)의 경우 2016년 1월 예정자 등록을 마쳤고, 해당 심사본이 접수된 12월까지 집필·개발되었다. 개발된 교과서들은 이후 검·인정 심사를 거쳐 2017년 9월 최종 선정되어 2018년 3월에 실제 적용될 예정이다(주형미 외, 2015; 한국교육과정평가원, 2016).

‘교육과정 해설서’는 교사들에게 교육과정 개정의 방향과 중점, 이에 따른 학교급별 개정의 주요 변화 내용을 자세하고 친절하게 안내하는 실질적 활용 자료다(정영근 외, 2010). 이를 위해 교육부와 한국교육과정평가원은 2015년 9월부터 12월까지 전문가 협의회 및 실무진 심층면담 등을 통해 2015 개정 ‘교육과정 총론 해설서’(온정덕 외, 2015; 한혜정 외, 2015), ‘창의적 체험활동 교육과정 해설서’(유제순 외, 2015), ‘특수교육 교육과정 총론 해설서’(박은혜 외, 2015)를 개발·배포하였다. 아울러 학교에서 평가활동이 원활하게 이루어질 수 있도록 2015 개정 교육과정에 따른 교과별 성취기준, 평가기준, 성취수준 및 그에 따른 예시 평가도구를 교과별로 개발하여 보급하였다(이미경 외, 2016).

또한 2015 개정 교육과정을 반영하여 2017년, 2018년부터는 새로운 학교생활기록부(이하 ‘학생부’) 제도가 시행될 예정이다. 이는 학생들이 진로와 적성에 따라 선택한 과목을 이수하면서 쏟은 열정과 노력, 이를 통

---

5) 이에 따라 2017년 1월, 2015 개정 교육과정에 따라 개발된 국·검정도서를 혼용하고, 검정도서의 개발기간을 1년 6개월에서 1년으로 단축하기 위한 「교과용 도서에 관한 규정」 일부 개정령(안)이 입법 예고되었다(교육부, 2017a). 그러나 1월 20일, 국회 교육문화체육관광위원회는 역사교과서의 국정 발행 금지를 골자로 한 ‘국정교과서 금지법’(역사 교과용도서의 다양성 보장에 대한 특별법)을 의결·상정하는 등 2017년 1월 30일 현재 관련 논란이 계속되고 있다.

해 성장한 역량 등이 학생부 종합전형 등에 반영될 수 있도록 하기 위함이다(교육부, 2017c). 이에 따라 2016년 4월에는 「학교생활기록 작성 및 관리지침」 훈령 일부가 개정되었다. 즉, 2016년부터는 학생부에 자유학기제의 ‘자유학기 활동’(진로탐색활동, 주제선택활동, 예술·체육활동, 동아리활동) 이수 상황 기재란과, ‘자격증 및 인증 취득상황’의 ‘국가직무능력표준(National Competency Standards: NCS)으로 구성된 전문 교과Ⅱ의 실무과목 이수상황’란을 신설하였다. 아울러 수행평가를 포함한 과정 중심의 평가를 강화하여 초·중학교는 과목 특성 및 수업활동과 연계해 필요한 경우, 고등학교는 전문 교과의 실기과목과 보통 교과의 체육 및 예술 교과(군) 중 실기 중심인 경우 수행평가만으로 평가를 실시할 수 있도록 하였다(교육부, 2016a). 또한 2016년 11월에는 ‘학생부 기재 개선 방안’을 통해 항목별 입력 주체를 명확히 규정하고, 서술형 평가 항목을 중심으로 표준 가이드라인과 기재 예시를 제시하였다(교육부, 2016d).

2015 개정 교육과정이 처음 적용되는 2021학년도 수능<sup>6)</sup>(2020년 11월 시행) 개편안은 ‘대학입학 사전 예고제’에 따라 해당 시험 응시자들이 고등학교에 입학하기 이전인 2017년에 확정·발표될 계획이다. 2015 개정 교육과정 연구에서는 수능 개편의 기본 방향을 ① 고등학교 수업의 정상화를 가능하게 하고, ② 문·이과 통합의 취지에 부합하며, ③ 수험생의 부담을 과도하게 늘리지 않는 수능체제로 규정(황규호 외, 2015)하였으나, 구체적 방안은 2017년 1월 현재 논의 중이다. 참고로, 2013년 8월에 발표된 ‘문·이과 완전 통합안’에서는 문·이과 구분 없이 모든 학생이 이수하는 ‘공통 과목’을 중심으로 수능을 출제하는 방안이 제시되었다(교육부, 대입제도 발전방안 연구위원회, 2013).

아울러 교육부는 2015 개정 교육과정 정책이 교육현장에 제대로 안착

---

6) 다만, 2020학년도 수능의 경우 ‘한국사’ 과목의 문항을 2009 개정 교육과정과 2015 개정 교육과정의 공통 성취기준 범위 내에서 출제할 계획이다. 이는 해당 시험 응시자들이 한국사를 배우기 시작하는 2017년의 경우, 한국사 교과서가 국·검정이 혼용된 형태로 운영되기 때문에, 어느 교과서로 공부한 학생이든 대학 입시에서 불이익을 받지 않도록 조치하기 위함이다.(교육부, 2017b).



될 수 있도록 2016년부터 관련 교원 연수를 지속적으로 추진하고 있다. 교원 연수는 교원 참여형, 현장 밀착형으로 진행되며, 시·도교육청 권역별 교육과정 총론 핵심교원(교육과정 부장, 수석교사, 교감 등) 연수, 신설 과목 교원 및 전문직 연수, 교과별 선도교원 연수, 시·도교육청 현장 교원 연수, 시·도(시·군)교육청 전문직 연수 등의 형태로 구성되어 있다(교육부, 2016b). 고등학교를 예시로 들면, 2016년의 경우 국어, 수학, 영어, 한국사, 사회, 과학의 6개 교과에 핵심교원 및 선도교원 대상의 연수가 이루어졌다(교육부, 2017c). 아울러 2016년 5월부터는 교사의 수업역량 강화를 위해 교과 연구회 등 기존의 자생적 교사공동체, 신규교사공동체 등 교사학습 공동체를 발굴하여 지원하고 있다.

이와 함께 2015 개정 교과 교육과정에 적절히 대응할 수 있는 예비 교사를 양성하기 위해 교원양성기관의 교육과정 역시 개편할 예정이다. 교육부(2016f)에 따르면, 교직 과목과 교과교육 과목에 다양한 수업 실현, 과정 중심 평가방법, 진로교육, 융합교육, 창의적 체험활동 등의 내용을 확대·신설하고, 예비 교사들이 자유학기제를 접할 수 있는 기회를 확대하기 위해 ‘자유학기제 수업지원단’을 구성·운영할 계획이다. 아울러 교원 자격 표시 과목을 변경하고, 관련 자격별 기본 이수과목 개선을 추진한다. 또한 교원임용시험에 평가 역량 등을 반영하고, 심층면접 시 교육환경 변화 및 교육정책 이해도 평가를 강화할 예정이다(교육부, 2017c).

이러한 단계를 거쳐 2015 개정 교육과정 정책은 2017년부터 초·중등교육 현장에 연차적으로 시행된다.<sup>7)</sup> 즉, 2017년 3월에는 해당 시점에 교과용도서가 개발된 초등학교 1~2학년에, 2018년 3월에는 초등학교 1~4학년과 중·고등학교 1학년에, 2019년 3월에는 초등학교 전 학년과 중·고등학교 1~2학년에, 2020년 3월에는 초·중·고 전 학년에 적용될 예정이다. 지금까지 논의한 2015 개정 교육과정 정책의 도입과 전개 과정을 정책의 형성·결정·집행 단계별로 정리하면 [그림 II-1]과 같다.

7) 단, 교과서 개발과 상관없이 진행되는 중학교 자유학기제의 경우 2016년부터 전면 시행되었다(교육부, 2015a).

### 1. 정책 형성 단계 (2013. 8. 27~2014. 9. 23)

- **현행 정책의 문제점을 인식·진단하며, 새로운 정책 의제를 설정하는 단계**

- 2013. 8. 27. 「대입전형 간소화 및 대입제도 발전방안(시안)」, ‘문·이과 융합안’ 첫 언급
- 2013. 10. 24. 「2017학년도 대입 확정안」, 융·복합형 교육과정 개편 언급
- 2014. 2. 13. 「문·이과 통합형 교육과정 기본방향(안)」 발표
- 2014. 2. 26. 국가교육과정개정연구위원회, 총론 시안 개발 시작
- 2014. 6. 30. ~ 7. 24. 교육과정 재구조화, 개발지침 공개 토론회
- 2014. 9. 12. 총론 주요사항 공청회



### 2. 정책 결정 단계 (2014. 9. 24~2015. 9. 23)

- **정책 형성 단계를 통해 제시된 여러 가지 대안들을 선정·검토하여 정책을 확정해 가는 단계**

- 2014. 9. 24. 총론 주요사항 확정
- 2014. 12. 2015 개정 교육과정 총론 시안(최종안) 및 교과별 교육과정 개발 연구 시작
- 2015. 4. 15. ~ 5. 7. 각론 시안 공개 토론회
- 2015. 7. 30. ~ 8. 24. 총론 및 각론 1차, 한자교육 관련 공청회
- 2015. 8. 31. ~ 9. 4. 총론 및 각론 시안 2차 공청회
- 2015. 9. 교육과정심의회위원회, 총론 및 각론 시안 검토
- 2015. 9. 23. 2015 개정 교육과정 총론 및 각론 확정·고시



### 3. 정책 집행 단계 (2015. 9. 24~ )

- **정책 결정 단계를 통해 확정된 정책을 집행하기 위해 관련 기반을 준비하고 실제 실행하는 단계**

- 2015. 10. 12. 「중·고등학교 교과용도서 국·검·인정 구분(안)」 행정예고
- 2015. 10. 21. 초등학교 교과용도서 구분 고시
- 2015. 11. 3. 중·고등학교 교과용도서 국·검·인정 구분 확정·고시
- 2015. 12. 1. 초·중등학교 교육과정 일부개정 고시
- 2016. 4. 28. 「학교생활기록 작성 및 관리지침」 훈령 일부 개정
- 2016. 8. 29. 초·중등학교 디지털교과서 국·검정 구분 고시
- 2015. 9. ~ 2017. 8. 교과용도서 공모, 개발 및 심사, 선정(2017년 대상 기준)
- 2017년에 2021학년도 수능 개편안 발표(예정)

[그림 II-1] 2015 개정 교육과정 정책의 단계별 주요 추진 사안

## 나. 2015 개정 교육과정 정책의 주요 내용

2015 개정 교육과정 정책의 주요 내용으로는 국가 교육과정 문서를 통해 발표된 ① 초·중·고등학교 교육과정 개정 내용과 ② 학교 교육과정 편성·운영 지침, 그리고 해당 교육과정의 실행을 지원하기 위한 ③ 교과서, 대입, 교원 등 관련 제반 정책의 변경 내용이 포함된다. 이 중 세 번째 내용에 대해서는 앞선 ‘가. 2015 개정 교육과정 정책의 도입과 전개 과정’에서 기술하였다. 따라서 이 절에서는 2015 개정 교육과정의 주요 내용과 편성·운영 지침에 대해 교육부(2015b, 2015c, 2015d)가 발표한 ‘총론’과 ‘교과’ 주요 개정 내용을 중심으로 재구성하였다.

### 1) 총론 주요 개정 내용

2015 개정 교육과정은 「초·중등교육법」 제23조 제2항에 의거해 고시된 것으로, ‘총론’은 초·중등학교에서 편성·운영해야 할 학교 교육과정의 공통적, 일반적 기준을 제시한 것이다. 총론에서는 2015 개정 교육과정의 성격을 “① 국가수준의 공통성과 지역, 학교, 개인 수준의 다양성을 동시에 추구하고, ② 학습자의 자율성과 창의성을 신장하기 위한 학생 중심의 교육과정이며, ③ 학교와 교육청, 지역사회, 교원, 학생, 학부모가 함께 실현해 가는 교육과정일 뿐 아니라, ④ 학교교육 체제를 교육과정 중심으로 구현하고 ⑤ 학교교육의 질적 수준을 관리하고 개선하기 위한 교육과정”(교육부, 2015d)이라고 기술하였다.

우선, 2015 개정 교육과정 총론에서는 교육과정이 추구하는 교육이념과 인간상을 구현함에 있어 “교과교육을 포함한 학교 교육 전 과정을 통해 중점적으로 기르고자 하는 핵심역량”을 제시하였다(김경자 외, 2015: 277). 즉, 2015 개정 교육과정에서는 현행 2009 개정 교육과정이 추구하는 인간상을 기초로 하여, 창의·융합형 인재를 구현하기 위한 인간상을

‘자주적인 사람’, ‘창의적인 사람’, ‘교양 있는 사람’, ‘더불어 사는 사람’의 네 가지로 설정하였다. 그리고 이들이 갖추어야 할 핵심역량으로는 자기 관리 역량, 지식정보처리 역량, 창의적 사고 역량, 심미적 감성 역량, 의사소통 역량, 공동체 역량의 여섯 가지가 제시되었으며, 각 교과에서는 이들 역량과 연계하여 교과에 맞는 역량을 제시함으로써 교과 특성에 적합한 교육과정을 운영하도록 하였다. 2015 개정 교육과정이 추구하는 인간상과 핵심역량의 구성요소별 정의는 <표 II-2>와 같다.

<표 II-2> 2015 개정 교육과정이 추구하는 인간상과 핵심역량

구분	구성요소	정의
인간상	자주적인 사람	전인적 성장을 바탕으로 자아정체성을 확립하고 자신의 진로와 삶을 개척하는 사람
	창의적인 사람	기초 능력의 바탕 위에 다양한 발상과 도전으로 새로운 것을 창출하는 사람
	교양 있는 사람	문화적 소양과 다원적 가치에 대한 이해를 바탕으로 인류 문화를 향유하고 발전시키는 사람
	더불어 사는 사람	공동체 의식을 가지고 세계와 소통하는 민주 시민으로서 배려와 나눔을 실천하는 사람
핵심역량	자기 관리 역량	자아정체성과 자신감을 가지고 자신의 삶과 진로에 필요한 기초 능력과 자질을 갖추어 자기주도적으로 살아갈 수 있는 능력
	지식정보처리 역량	문제를 합리적으로 해결하기 위하여 다양한 영역의 지식과 정보를 처리하고 활용할 수 있는 능력
	창의적 사고 역량	폭넓은 기초 지식을 바탕으로 다양한 전문 분야의 지식, 기술, 경험을 융합적으로 활용하여 새로운 것을 창출하는 능력
	심미적 감성 역량	인간에 대한 공감적 이해와 문화적 감수성을 바탕으로 삶의 의미와 가치를 발견하고 향유할 수 있는 능력
	의사소통 역량	다양한 상황에서 자신의 생각과 감정을 효과적으로 표현하고 다른 사람의 의견을 경청하며 존중하는 능력
	공동체 역량	지역·국가·세계 공동체의 구성원에게 요구되는 가치와 태도를 가지고 공동체 발전에 적극적으로 참여하는 능력

\* 이 표는 교육부(2015d)의 내용을 재구성한 것임.

또한 총론에서는 2015 개정 교육과정에서 특히 추구하고자 하는 변화의 내용과 방향을 ‘교육과정 구성의 중점’이라는 항목을 통해 제시하였다(김경자 외, 2015). 즉, 창의·융합형 인재를 양성하기 위해 2015 개정 교육과정이 중점을 두고 있는 6개 항목을 제시하였는데, 이는 크게 세 가지 측면으로 정리할 수 있다. 첫째, 인문·사회·과학기술에 관한 기초 소양 교육을 강화한다. 이를 위해 고등학교에서는 문·이과 구분 없이 모든 학생이 배우는 공통 과목(『국어』, 『수학』, 『영어』, 『한국사』, 『통합사회』, 『통합과학』, 『과학탐구실험』)을 도입 및 신설한다. 또한 심화선택 『정보』 과목의 내용을 소프트웨어 중심으로 개편하고 일반선택 과목으로 전환하며, 예술 교과군의 일반 선택에는 『연극』 과목을 신설해 체험 중심의 인성교육을 강화한다. 아울러 각 학교 급에서는 심폐소생술 교육을 강화하고, 체육, 기술·가정, 과학, 보건 등 관련 교과(목)에 ‘안전’ 단원을 신설하여 체계적 안전교육을 실시한다.

둘째, 학생들의 꿈과 끼를 키울 수 있는 교육과정을 마련한다. 이를 위해 단위학교의 교육과정 편성·운영의 자율권을 확대하여 학생의 진로와 적성을 고려한 다양한 선택 과목을 개설할 수 있도록 하고, 2016년부터 중학교 한 학기를 ‘자유학기’로 운영할 수 있는 근거를 마련하였다. 셋째, 미래 사회가 요구하는 핵심역량을 함양할 수 있는 교육과정을 마련한다. 이를 위해 교과별로 반드시 배워야 할 핵심 개념을 중심으로 학습내용을 정선·감축하고, 교육 목표, 교육 내용, 교수·학습 및 평가의 일관성을 강화한다. 또한 교수·학습 및 평가 방법을 학생 참여 및 과정 중심으로 개선해 학생들이 진정한 배움의 즐거움을 느낄 수 있도록 한다. 아울러 특성화 고등학교(이하 ‘특성화고’)와 산업수요 맞춤형 고등학교(이하 ‘산업수요 맞춤형고’)에서는 NCS를 활용해 산업사회가 필요로 하는 기초 역량과 직무 능력을 함양한다.

참고로, 2015개정 교육과정의 교육 목표는 추구해야 할 인간상을 바탕으로 학교급별로 나누어 제시되었다. 2015 개정 교육과정에서 추구하는 인간상과 학교 급별 교육 목표를 함께 나타내면 <표 II-3>과 같다.

<표 II-3> 2015 개정 교육과정의 인간상과 학교 급별 교육 목표(온정덕, 2015 수정)

인간 상	교육 목표		
	초등학교	중학교	고등학교
	초등학교 교육은 학생의 일상생활과 학습에 필요한 기본 습관 및 기초 능력을 기르고 바른 인성을 함양하는 데에 중점을 둔다.	중학교 교육은 초등학교 교육의 성과를 바탕으로, 학생의 일상생활과 학습에 필요한 기본 능력을 기르고 바른 인성 및 민주 시민의 자질을 함양하는 데에 중점을 둔다.	고등학교 교육은 중학교 교육의 성과를 바탕으로, 학생의 적성과 소질에 맞게 진로를 개척하며 세계와 소통하는 민주 시민으로서의 자질을 함양하는 데에 중점을 둔다.
자주 적인 사람	자신의 소중함을 알고 건강한 생활 습관을 기르며, 풍부한 학습 경험을 통해 자신의 꿈을 키운다.	심신의 조화로운 발달을 바탕으로 자아존중감을 기르고, 다양한 지식과 경험을 통해 적극적으로 삶의 방향과 진로를 탐색한다.	성숙한 자아의식과 바른 품성을 갖추고, 자신의 진로에 맞는 지식과 기능을 익히며 평생학습의 기본 능력을 기른다.
창의 적인 사람	학습과 생활에서 문제를 발견하고 해결하는 기초 능력을 기르고, 이를 새롭게 경험할 수 있는 상상력을 키운다.	학습과 생활에 필요한 기본 능력 및 문제 해결력을 바탕으로, 도전 정신과 창의적 사고력을 기른다.	다양한 분야의 지식과 경험을 융합하여 창의적으로 문제를 해결하고, 새로운 상황에 능동적으로 대처하는 능력을 기른다.
교양 있는 사람	다양한 문화 활동을 즐기고 자연과 생활 속에서 아름다움과 행복을 느낄 수 있는 심성을 기른다.	자신을 둘러싼 세계에서 경험한 내용을 토대로 우리나라와 세계의 다양한 문화를 이해하고 공감하는 태도를 기른다.	인문·사회·과학기술 소양과 다양한 문화에 대한 이해를 바탕으로 새로운 문화 창출에 기여 할 수 있는 자질과 태도를 기른다.
더불어 사는 사람	규칙과 질서를 지키고 협동정신을 바탕으로 서로 돕고 배려하는 태도를 기른다.	공동체 의식을 바탕으로 타인을 존중하고 서로 소통하는 민주 시민의 자질과 태도를 기른다.	국가 공동체에 대한 책임감을 바탕으로 배려와 나눔을 실천하며 세계와 소통하는 민주 시민으로서의 자질과 태도를 기른다.

아울러 총론에서는 ‘학교 급별 교육과정 편성·운영의 기준’을 제시하였다. 우선, 전 학교 급 공통 사항은 다음의 9가지로, 2015 개정 교육과정은 ① ‘공통 교육과정’(초등학교 1학년~중학교 3학년)과 ‘선택 중심 교육과정’(고등학교 전 학년)으로 편성·운영하고, ② 학년 간 상호 연계와 협력을 통해 학교 교육과정을 유연하게 편성·운영할 수 있도록 ‘학년군’을 설정한다. ③ 공통 교육과정의 교과는 교육 목적상의 근접성, 학문 탐구 대상 또는 방법상의 인접성, 생활양식에서의 연관성 등을 고려하여 ‘교과군’으로 재분류하고, ④ 선택 중심 교육과정에서는 학생들의 기초 영역 학습을 강화하며, 진로와 적성에 맞는 학습이 가능하도록 4개 교과 영역으로 구분해 교과(군)별 필수이수단위를 제시한다. ⑤ 고등학교 교과는 보통 교과와 전문 교과로 구분하며, 학생들의 기초 소양 함양과 기본 학력을 보장하고자 보통 교과에 ‘공통 과목’을 개설하여 모든 학생이 이수하도록 한다. ⑥ 학습 부담을 적정화하고 의미 있는 학습 활동이 이루어질 수 있도록 학기당 이수 교과목 수를 조정하여 ‘집중이수’를 실시할 수 있다. ⑦ ‘창의적 체험활동’은 학생의 소질과 잠재력을 계발하고 공동체 의식을 기르는 데에 중점을 두며, ⑧ ‘범교과 학습’ 주제는 안전·건강 교육, 인성교육, 진로 교육, 민주 시민 교육, 인권 교육, 다문화 교육, 통일 교육, 독도 교육, 경제·금융 교육, 환경·지속가능발전 교육의 총 10가지로, 교과와 창의적 체험활동 등 교육 활동 전반에 걸쳐 통합적으로 다루도록 하고, 지역사회 및 가정과 연계하여 지도한다. ⑨ 학교는 필요에 따라 ‘계기 교육’을 실시할 수 있으며, 이 경우 계기 교육 지침에 따른다. 이를 학교 급별로 나누어 정리하면 다음과 같다.

#### 가) 초등학교

2015 개정 초등학교 교육과정은 2009 개정 교육과정과 마찬가지로 ‘교과(군)’과 ‘창의적 체험활동’으로 편성한다. 즉, 교과(군)에는 국어, 사회/도덕, 수학, 과학/실과, 체육, 예술(음악/미술), 영어가 포함되며(1, 2학년

은 국어, 수학, 바른 생활, 슬기로운 생활, 즐거운 생활), 창의적 체험활동은 자율 활동, 동아리 활동, 봉사 활동, 진로 활동으로 구성된다(1, 2학년은 안전한 생활 포함). 단, 창의적 체험활동의 경우 하위 영역 구분으로 인해 단위 학교의 교육과정 편성 운영 자율권을 충분히 발휘할 수 없다는 지적이 제기됨에 따라(온정덕, 2015), ‘학교는 창의적 체험활동의 영역을 학생들의 발달 수준, 학교의 여건 등을 고려하여 학년(군)별로 선택적으로 편성·운영할 수 있다.’는 지침이 신설되었다. 초등학교 교육과정의 편제와 시간 배당 기준을 도식화하면 <표 II-4>와 같다.

<표 II-4> 초등학교 교육과정 편제와 시간배당(교육부, 2015d)

구 분		1~2학년	3~4학년	5~6학년
교 과  (군)	국어	국어 448	408	408
	사회/도덕		272	272
	수학	수학 256	272	272
	과학/실과	바른 생활 128	204	340
	체육	슬기로운 생활 192	204	204
	예술(음악/미술)	즐거운 생활 384	272	272
	영어		136	204
	소계	1,408	1,768	1,972
창의적 체험활동		336 안전한 생활 (64)	204	204
총 수업 시간 수		1,744	1,972	2,176

- ① 학년군 및 교과(군)별 시간 배당은 연간 34주를 기준으로 한 2년간의 수업 시수를 뜻함.  
 ② 학년군별 총 수업 시간 수는 최소 수업 시수를 나타낸 것임.  
 ③ 실과의 수업 시간은 5~6학년 과학/실과의 수업 시수에만 포함된 것임.

2009 개정 교육과정과 비교해 보았을 때, 2015 개정 초등학교 교육과정에서 달라진 점은 1~2학년(군)에 한글교육을 강조하는 등 유아 교육과정인 누리과정과 연계를 강화한 점이다. 또한 초등학교 수업 시수의 국제 비교 연구 결과와 맞벌이 부부의 증가에 따른 돌봄 기능의 강화에 근거하여(온정덕, 2015) 1~2학년(군)의 수업시수를 주당 1시간(40분) 늘



리되, 학생들의 학습 부담이 추가되지 않도록 창의적 체험 활동 시간으로 활용하여 체험 중심의 『안전한 생활』을 편성·운영하도록 하였다. 그 결과, 창의적 체험활동은 기존 272시간에서 336시간으로 변경되었으며, 안전한 생활에서는 생활안전, 교통안전, 신변안전, 재난안전의 4개 영역에 대한 내용을 다루고 지식보다는 체험 중심 학습으로 자연스럽게 안전한 생활습관과 의식을 습득하게 할 예정이다. 3~6학년의 경우에는 관련 교과와 단원의 신설을 통해 안전교육을 실시할 예정이다.

한편 소프트웨어 교육 활성화의 필요성이 대두됨에 따라, 기존에 ICT 활용 중심으로 편성되어 있던 초등학교 5~6학년 실과 단원의 내용을 소프트웨어 기초 소양 중심으로 재편하였다. 이 단위에서는 놀이 중심의 알고리즘 체험과 교육용 도구를 활용한 프로그래밍 체험을 통해 기술적 문제해결능력을 개발할 수 있도록 하였으며, 정보윤리의 내용을 포함해 정보기술을 올바르게 사용할 수 있도록 하였다(김경자 외, 2015).

#### 나) 중학교

2015 개정 중학교 교육과정은 교과군 및 학년군 설정, 교과(군)별 20% 범위 내 시수 증감 허용, 교과 집중 이수 및 학기당 이수 교과목 수 축소, 선택 과목 및 진로 교육 활성화 등 2009 개정 교육과정의 기본 체제를 유지하고 있다(김경자 외, 2015). 즉, 중학교 교육과정 역시 초등학교와 마찬가지로 ‘교과(군)’와 ‘창의적 체험활동’으로 편성하며, 교과(군)에는 국어, 사회(역사 포함)/도덕, 수학, 과학/기술·가정/정보, 체육, 예술(음악/미술), 영어, 선택 교과가 포함된다. ‘선택 교과’는 한문, 환경, 생활 외국어(독일어, 프랑스어, 스페인어, 중국어, 일본어, 러시아어, 아랍어, 베트남어), 보건, 진로와 직업 등의 과목으로 한다. 이때 선택 교과목인 ‘환경’은 기존의 ‘환경과 녹색성장’에서 명칭이 변경된 것이다. 중학교 교육과정의 편제와 시간 배당 기준을 도식화하면 <표 II-5>와 같다.

<표 II-5> 중학교 교육과정 편제와 시간배당(교육부, 2015d)

구 분		1~3학년
교 과 (군)	국어	442
	사회(역사 포함)/도덕	510
	수학	374
	<b>과학/기술·가정/정보</b>	<b>680</b>
	체육	272
	예술(음악/미술)	272
	영어	340
	선택	170
	소계	3,060
창의적 체험활동		306
총 수업 시간 수		3,366

- ① 1시간 수업은 45분을 원칙으로 하되, 기후 및 계절, 학생의 발달 정도, 학습 내용의 성격, 학교 실정 등을 고려하여 탄력적으로 편성·운영할 수 있음.  
 ② 학년군 및 교과(군)별 시간 배당은 연간 34주를 기준으로 3년간의 기준 수업 시수임.  
 ③ 총 수업 시간 수는 3년간의 최소 수업 시수를 나타낸 것임.

2009 개정 교육과정과 비교해 보았을 때, 2015 개정 중학교 교육과정에서 크게 달라진 점은 한 학기를 ‘자유학기’로 운영할 수 있는 근거를 마련한 점이다. 즉, 자유학기제 운영을 통해 학생들이 중간·기말고사에 대한 부담에서 벗어나 체험 중심의 교과 활동과 진로 탐색 활동을 적극 수행하도록 하였으며, 이러한 활동은 창의적 체험활동의 관련 영역과 연계하여 운영할 것을 강조하였다(김경자 외, 2015). 또한 학생들이 소프트웨어에 대한 기초 소양을 충실히 갖추 수 있도록 『정보』 교과를 필수 과목으로 지정하였다. 이에 따라 2009 개정 교육과정의 ‘과학/기술·가정’ 교과군은 ‘과학/기술·가정/정보’로 명칭이 변경되었고, 정보 교과는 해당 교과군에서 34시간을 기준으로 편성·운영할 수 있도록 하였기 때문에 ‘선택 교과’ 수업 시간은 204에서 170으로 줄어들고, 대신 과학/기술·가정/정보 교과군의 기준 수업 시간은 646에서 680으로 늘어나게 되었다.

## 다) 고등학교

2015 개정 고등학교 교육과정은 2009 개정 교육과정 때와 마찬가지로 ‘교과(군)’과 ‘창의적 체험활동’으로 편성되었다. 교과(군)은 보통 교과와 전문 교과로 구분되는데, ‘보통 교과’는 기초, 탐구, 체육·예술, 생활·교양 영역으로 구성되며, 국어, 수학, 영어, 한국사, 사회(역사/도덕 포함), 과학, 체육, 예술, 기술·가정/제2외국어/한문/교양 교과가 포함된다. 이때 ‘기초 영역’(국어, 수학, 영어, 한국사)의 이수단위는 2009 개정 교육과정 때와 마찬가지로 교과 총 이수단위의 50%를 넘을 수 없도록 하였는데, 2015 개정 교육과정에서는 기초 교과 영역에 『한국사』(필수 이수 단위 6단위)가 포함되므로 교과 총 이수 단위의 50%는 기존의 90단위가 아닌 84단위를 의미한다. 따라서 국어, 영어, 수학 비중을 적정화하여 다양한 교과에 대한 균형 있는 학습을 권장하는 이 규정은 2015 개정 교육과정에서 더욱 강화되었다고 할 수 있다(온정덕, 2015). 또한 2009 개정 교육과정에서는 이 규정을 자율형 사립 고등학교에 대해 권장 사항으로 두었지만, 2015 개정 교육과정에서는 이를 의무 사항으로 제시하였다. 참고로, 『한국사』 과목은 2009 개정 교육과정에서는 사회 교과(군)에 포함되어 있었으나, 2015 개정 교육과정에서는 모든 학생이 필수로 이수하도록 기초 영역에 포함하였다(김경자 외, 2015).

보통 교과는 ‘공통 과목’과 ‘선택 과목’으로도 구분되는데, 이 부분이 2015 개정 교육과정에서 가장 크게 달라진 점이라고 할 수 있다. 즉, 모든 학생들은 ‘공통 과목’을 통해 기초소양을 함양한 후, 학생 각자의 적성과 진로에 따라 맞춤형으로 교육받을 수 있도록 하는 ‘선택 과목’이 개설된 것이다. 선택 과목은 교과별 주요 학습 영역을 일반적 수준에서 다루는 ‘일반 선택 과목’과 교과 융합학습, 진로 안내학습, 교과별 심화학습, 실생활 체험학습 등이 가능한 ‘진로 선택 과목’으로 구성하였다(교육부, 2014c). 이때 일반 선택 과목은 학생이 원할 경우 학교가 가능한 이수 기회를 제공하도록 하는 지침이, 진로 선택 과목은 학생의 진로에 따

른 선택권을 확대하기 위해 3개 이상 이수하도록 하는 지침이 신설되었다. 아울러 진로에 따라 다양한 과목을 조합한 맞춤형 과정 이수가 가능하도록 과목 선택에 대한 진로 지도를 강화하는 지침도 신설되었다.

전문 교과는 전문 교과Ⅰ과 전문 교과Ⅱ로 구분하며, ‘전문 교과Ⅰ’은 과학, 체육, 예술, 외국어, 국제 계열에 관한 과목으로 한다. 즉, 2009 개정 교육과정에서 보통 교과의 심화 과목으로 배치되어 있던 특수 목적 고등학교 대상 과목을 보통 교과에서 분리해 전문 교과Ⅰ로 제시하고, 일반 고등학교(이하 ‘일반고’)와 자율 고등학교(이하 ‘자율고’)에서도 학생의 희망과 학교 여건에 따라 과목을 편성·운영할 수 있게 하였다(김경자 외, 2015). 그리고 특성화고와 산업수요 맞춤형고의 교육과정은 전문 교과를 공통 과목, 기초과목, 실무과목으로 개편하여 NCS와의 연계성을 강화하였는데 ‘전문 교과Ⅱ’가 바로 이에 해당한다. 이때 NCS는 산업현장에서 직무를 수행하기 위해 요구되는 지식·기술·소양 등의 내용을 국가가 산업부문별, 수준별로 체계화한 것으로, 산업계의 교육훈련 요구를 반영한다고 간주할 수 있다(김경자 외, 2015). 즉, 전문 교과Ⅱ는 NCS를 활용하여 산업사회에서 요구하는 기초 역량과 직무 능력을 함양하기 위한 것으로, 전문 교과Ⅱ의 실무 과목은 성취평가제와 연계하여 내용 요소를 구성하는 ‘능력단위’ 기준으로 평가할 수 있다.

고등학교 교육과정의 편제와 시간 배당 기준은 <표 Ⅱ-6>과 같으며, 세부 내용은 고등학교의 유형에 따라 다르다. 예컨대, 기초영역 중 국어, 수학, 영어 교과의 필수 이수단위의 경우 일반고와 특목고는 각 10단위인 데 반해, 특성화고와 산업수요 맞춤형고의 경우 세 교과 통합 24단위로 운영된다. 참고로 ‘자율 편성 단위’는 2009 개정 교육과정에서 ‘학교 자율 과정’으로 사용되었는데, 학교 자율 과정은 특정 과정을 개설하는 것이 아니라는 점에서 국가 수준에서 정한 ‘필수 이수 단위’와 대비되는 용어로 변경하였다(김경자 외, 2015).

<표 II-6> 고등학교 교육과정 편제와 시간배당(교육부, 2015d 수정)

교과 영역	교과(군)	공동 과목 (단위)	필수이수단위		자율편성단위	
			일반고, 특목고	특성화고, 산업수요 맞춤형고	일반고, 특목고	특성화고, 산업수요 맞춤형고
교과 (군)	기초	국어	국어(8)	10	학생의 적성과 진로를 고려 하여 편성	학생의 적성· 진로와 산업계 수요를 고려 하여 편성
		수학	수학(8)	10		
		영어	영어(8)	10		
		한국사	한국사(6)	6		
	탐구	사회 (역사/도덕 포함)	통합사회(8)	10		
		과학	통합과학(8) 과학탐구실험(2)	12		
	체육· 예술	체육		10		
		예술		10		
	생활· 교양	기술·가정/ 제2외국어/ 한문/교양		16		
	소계			94	66	86
	전문 교과 II	17개 교과(군) 등	—	86	—	28
창의적 체험활동			24(408시간)			
총 이수 단위			204			

- ① 1단위(1시간)는 50분을 기준으로 하여 17회를 이수하는 수업량임.
- ② 공동 과목은 2단위 범위 내에서 감하여 편성·운영할 수 있음. 단, 한국사는 6단위 이상 이수하되 2개 학기 이상 편성하도록 함.
- ③ 과학탐구실험은 이수 단위 증감 없이 편성·운영하는 것을 원칙으로 하되, 과학 계열, 체육 계열, 예술 계열 고등학교의 경우 학교 실정에 따라 탄력적으로 운영할 수 있음.
- ④ 필수 이수 단위의 단위 수는 해당 교과(군)의 ‘최소 이수 단위’로 공동 과목 단위 수를 포함함. 특수 목적 고등학교와 자율형 사립 고등학교의 경우 예술 교과(군)는 5단위 이상, 생활·교양 영역은 12단위 이상 이수할 것을 권장함.
- ⑤ 창의적 체험활동의 단위는 최소이수단위이며 ( )안의 숫자는 시간 수로 환산한 것임.
- ⑥ 총 이수 단위 수는 고등학교 3년간 이수해야 할 ‘최소 이수 단위’를 의미함.

한편, 2015 개정 교육과정 총론에서는 ‘학교 교육과정 편성·운영’ 지침을 신설하였다. 이는 기본 사항, 교수·학습, 평가, 모든 학생을 위한 교육 기회의 제공의 네 가지 항목으로 구성되었는데, 이 중에서도 ‘교수·학습’과 ‘평가’ 부분은 2015 개정 교육과정의 기본 방향에 맞게 추가된 내용이 많다. 우선 ‘교수·학습’ 부분에서는 “단편적 지식의 암기를 지양하고 핵심 개념과 일반화된 지식의 심층적 이해에 중점을” 두며, “융합적 사고를 기를 수 있도록 교과 내, 교과 간 내용 연계성을 고려하여 지도한다”(교육부, 2015d: 31)는 항목을 신설하였다. 그리고 학생이 “능동적으로 수업에 참여하고 자신의 생각을 표현하는 기회를 가질 수 있도록 토의·토론 학습을 활성화”하며, “자신의 학습 과정과 학습 전략을 점검하고 개선하며 자기주도적으로 학습할 수 있도록 지도”(교육부, 2015d: 31)할 뿐 아니라 “교사와 학생 간, 학생과 학생 간 상호 신뢰와 협력이 가능한 교수·학습 환경을 제공한다”(교육부, 2015d: 32)는 내용 역시 추가되었다.

‘평가’ 부분에서는 “학교는 학생에게 평가 결과에 대한 적절한 정보 제공과 추수 지도를 통해 학생이 자신의 학습을 지속적으로 성찰하고 개선할 수 있도록 지도한다”(교육부, 2015d: 33)는 항목을 추가함으로써 평가 목적을 확장시켰다. 아울러 “학습의 결과뿐만 아니라 학습의 과정을 평가하여 모든 학생이 교육 목표에 성공적으로 도달할 수 있도록 한다”, “학교는 학생의 인지적 능력과 정의적 능력에 대한 평가가 균형 있게 이루어질 수 있도록 한다”(교육부, 2015d: 33)는 항목을 신설하여 과정 중심의 평가, 전인적 평가의 중요성을 강조하였다.

## 2) 교과 주요 개정 내용

2015 개정 교육과정의 교과 교육과정은 총 24개 교과에 해당하는 방대한 내용을 다루고 있다. 여기에서는 교육부(2015b, 2015c)에서 제시한 학교 급별 국어, 수학, 영어, 사회, 과학과 및 전문 교과Ⅱ 교육과정으로 범위를 한정하여 관련 내용을 재구성하였다.

## 가) 초등학교

### (1) 국어

2015 개정 초등학교 국어과 교육과정에서는 ‘한글 교육’을 강화하였다. 즉, 2009 개정 교육과정에서 총 27차시였던 한글 교육은 1학년 1학기에 최소 45차시 이상이 되도록 편성하였으며, 모든 학생이 기초학습능력 갖출 수 있도록 2학년까지 꾸준히 학습하도록 하였다. 또한 ‘연극 교육’ 역시 강화되어 기존에 4차시, ‘이야기를 희곡으로 바꾸기’로 다루어졌던 연극 교육을 대단원으로 구성하고 체험 중심의 연극 활동을 수행하도록 하였다. 또한 초등학교 국어를 비롯한 전 학교 급 모든 과목의 ‘4. 교수학습 및 평가의 방향’에 1학기 1권 통합적 독서 활동을 위한 수업 시간 확보하고 도서 준비를 안내함으로써 ‘독서 교육’을 강화하였다. 또한 예상 수업 시간 내에 배울 수 있는 적정 학습량을 제시하도록 하였다.

한편 ‘학습량 감축’ 차원에서 2009 개정 교육과정에서 97개였던 성취기준은 82개로 축소되었다. 이 과정에서 핵심 내용에 변화가 생겼는데, ‘면담’, ‘다양한 관점’ 등의 내용은 중학교로 이동했으며, ‘절의 연결’, ‘문학 언어와 일상언어의 차이’, ‘갈래 변환’ 등의 내용은 삭제되었고, ‘글자’, ‘낱말’, ‘문장 관찰’ 등의 기초 내용이 신설되었다.

### (2) 수학

2015 개정 초등학교 수학과 교육과정에서는 핵심개념과 원리를 중심으로 학습 내용을 재구성하는 데 개정의 초점을 맞추었다. 즉, 1~2학년(군)의 경우에는 누리과정과의 연계성을 확보하기 위해 ‘물체의 위치와 방향’, ‘규칙 찾기’ 활동을 강화하였다. 또한 3~4학년(군)에서는 학습자의 발달 단계를 고려하여 ‘자연수의 혼합계산’, ‘규칙과 대응’, ‘수의 범위’, ‘어림하기’의 내용이 5~6학년(군)으로 이동되었고, 대신 5~6학년(군)의

‘무게 단위 톤(t)’은 3~4학년(군)으로 이동하되, ‘아르(a)/헥타르(ha) 단위’의 내용은 실생활 활용이 낮다는 현장교사의 의견을 반영하여 삭제되었다. 그리고 ‘원주율 근삿값 다양화’라는 내용은 신설되었다. 한편 ‘분수·소수 혼합계산’, ‘정비례와 반비례’는 중학교로 통합·이동되었다.

### (3) 영어

2015 개정 초등학교 영어과 교육과정에서도 핵심 성취기준을 선별하고 학습 내용을 재구성하는 데 개정의 초점을 맞추었다. 예컨대, ‘과거, 미래 시제’의 경우 단순 문법 지식을 경감하기 위해 삭제되었다. 또한 학습자의 발달단계를 고려하여 ‘영어로 노래하기’, ‘영어로 게임하기’ 등에 해당하는 성취기준은 삭제하고, 5~6학년(군)의 성취기준인 “자신이나 가족 등에 관해 짧고 간단하게 쓴다”는 중학교 1학년으로 이동하였다. 단, 이와 같이 특정 내용이 삭제·이동되었으나 수업 활동에 해당하는 성취기준은 2009 개정 교육과정의 14개와 동일하기 때문에 영어과에서 학습량이 감축된 것은 아니다. 아울러 실생활 속 영어회화 능력을 강화하는 데 중점을 두기 위해 ‘듣기’와 ‘말하기’ 영역의 학습 요소를 확대하였다.

### (4) 사회

2015 개정 초등학교 사회과 교육과정에서는 2009 개정 교육과정과 마찬가지로 ‘지리’, ‘역사’, ‘일반사회’ 영역으로 구성하되, 주제 중심의 통합 접근을 강화하기 위해 통합 대단원을 구성하였다. 예컨대, ‘여성’, ‘신분제도’, ‘인권’과 같은 내용을 지리, 일반사회 영역과 통합하는 방식이다. 또한 2009 개정 교육과정의 ‘사람이 모이는 곳’은 삭제하고 이를 ‘이동과 소통하기’로 변형하였으며, ‘시대마다 다른 삶의 모습’은 신설, ‘영토 교육’과 ‘통일 교육’은 강화하는 방향으로 교육과정 내용을 재구성하였다.



## (5) 과학

2015 개정 초등학교 과학과 교육과정에서도 핵심 내용을 재구조화하는 방식으로 개정이 이루어졌다. 예컨대, 2009 개정 교육과정의 ‘동물의 생활’과 ‘식물의 생활’은 각각 동일 학년에서 학습하도록 조정·통합되었으며 ‘액체의 부피 측정’의 경우 수학 교과에서 활용하는 것으로 변경한 후 삭제하였다. 그리고 학습자의 발달단계를 고려하여 ‘기체에 대한 입자적 설명’은 중학교 교육과정으로 이동되었고, 통합 단원으로는 ‘물의 여행’과 ‘에너지와 생활’이 신설되었다. 또한 안전교육 강화의 측면에서 2009 개정 교육과정의 ‘지진 발생시 대처 방법’ 부분은 안전 관련 내용으로 보완하였다. 아울러 초등학교 과학 학습 단계를 실험 수행을 위한 기능숙달 과정으로 간주하고 이러한 측면을 보강하였다.

### 나) 중학교

#### (1) 국어

2015 개정 중학교 국어과 교육과정에서는 초등학교와 마찬가지로 ‘독서교육’과 ‘연극교육’을 강화하였다. 우선, 2009 개정 교육과정에서 교과서의 짧은 글을 읽는 방식으로 이뤄졌던 독서교육을 보완하기 위해 읽고, 생각을 나누고, 쓰는 통합적 독서활동에 관한 성취기준, 학습요소, 교수·학습 방법 및 유의사항을 교육과정 문서에 제시하도록 하였다. 그리고 중학교에서는 연극 교육을 위한 소단원이 새롭게 구성되었고, 이 역시 체험 중심으로 운영될 예정이다. 또한 2009 개정 교육과정에서 55개였던 성취기준이 51개로 축소되었다. 이 과정에서 ‘음운의 변동’, ‘문법요소’, ‘협상’ 등의 내용은 고등학교로 이동했으며, ‘로마자/외래어 표기법’, ‘전통적인 말하기 문화 비교’, ‘독자의 정체성’, ‘작가의 태도’ 등은 삭제되고 ‘말하기 불안 대처’, ‘고전 재해석’ 등은 신설되었다.

## (2) 수학

2015 개정 중학교 수학과 교육과정은 수학에 대한 흥미와 자신감을 가질 수 있도록 핵심개념과 원리를 중심으로 학습내용을 재구성하는 데 개정의 초점을 맞추었다. 1학년의 경우에는 실생활 맥락의 통계교육을 강화하기 위해 ‘공학적 도구 이용’을 신설하였고, ‘도수분포표 평균’은 지나친 계산을 유발한다는 점에서 삭제되었다. 2학년의 경우에는 단순 문제풀이를 유발하는 ‘등식의 변형’은 삭제하고, ‘기하 교육’은 교수·학습 방법 다양화 측면에서 강화하였다. 또한 ‘곱셈공식’은 연계성을 강화하기 위해 3학년으로 이동·통합된 대신, 3학년에서는 국제적 추세를 고려하여 ‘피타고라스 정리’를 2학년으로 이동시켰다. 3학년에서는 빅데이터 시대상을 반영하기 위한 목적으로 ‘산점도와 상관계수’가 추가되었으며, ‘이차함수의 최대, 최소’는 고등학교 1학년으로 이동·통합되었다.

## (3) 영어

2015 개정 중학교 영어과 교육과정은 학습량을 감축하기보다는 하위 성취기준을 통합함으로써 핵심 성취기준을 선별하는 방식으로 개정되었다. 예컨대, ‘차이점 말하기’, ‘호불호의 이유 말하기’, ‘줄거리 말하기’ 등 타 성취기준에 포함되거나 중복되는 성취기준은 ‘세부 정보 파악 및 중심 내용 이해’ 등에 통합시켰다. 또한 2009 개정 교육과정에서는 초등학교만 권장 어휘를 제시하였는데, 2015 개정 영어과 교육과정에서는 중학교에서도 학습 어휘를 제시하였다.

## (4) 사회

2015 개정 중학교 사회과 교육과정에서는 2009 개정 교육과정과 마찬가지로 ‘일반사회’와 ‘지리’ 영역으로 구성하되, 내용 중복을 해소하고 학

습의 질을 제고하기 위해 단위 수를 기존 28개에서 24개로 조정하였다. 예컨대, ‘산지 지형’, ‘해안지형’, ‘카르스트지형’, ‘화산지형’ 등 지형에 대한 학습은 ‘산지 지형’과 ‘해안지형’으로 내용을 적정화하였고, 국제 경제 단원에서 다루었던 ‘환율’, ‘국제수지’ 등에서 ‘국제 수지’ 부분은 삭제하였다. 또한 ‘툰드라 지역의 생활’은 ‘기후 환경 극복’으로 변경하였고, ‘남북 분단과 통일’은 초·중·고등학교 간 연계성을 확보하기 위해 지리 영역으로 이동시켰다. 한편 금융 교육 등에 대한 국가·사회적 요구를 반영하기 위해 ‘금융’, ‘기업가정신’ 등의 내용은 추가하였다.

#### (5) 과학

2015 개정 중학교 과학과 교육과정은 초등학교와 마찬가지로 핵심 내용을 재구조화하는 방식으로 개정이 이루어졌다. 예컨대, 2009 개정 교육과정의 ‘빛과 파동에서 상의 작도’는 단순 지식 중심으로 구성되어 있다는 문제제기 하에 관련 내용이 삭제되었다. 그리고 학습자의 발달단계를 고려하여 ‘산·염기, 산화 반응’은 고등학교로, ‘생명의 진화’와 ‘염색체와 유전의 관계’, ‘지진과 이용 지구 층상구조 파악’은 각각 고등학교 『생명과학 I』과 『지구과학 II』로 이동하였다. 또한 안전교육 강화의 측면에서 ‘재해재난과 안전’과 같은 통합 단원이 신설되었고, ‘과학과 나의 미래’, ‘과학과 현대문명’ 등과 같은 통합 단위 역시 신설되어 과학에 대한 참여와 평생학습 능력 등이 교육과정에 다각도로 반영되었다.

#### 다) 고등학교

2015 개정 고등학교 교육과정의 가장 큰 변화는 모든 학생들이 인문·사회·과학기술에 대한 균형 잡힌 기초소양을 함양할 수 있도록 ‘공통 과목’을 도입한 것이다. 공통 과목에는 국어, 수학, 영어, 사회, 과학(이상 8단위), 한국사(6단위)가 포함되며, 이때 사회와 과학은 『통합사회』와

『통합과학』으로 개발되었고, 실험·실습·탐구 중심의 과학교육을 위해 『과학탐구실험』(2단위) 과목도 신설되었다. 관련 주요 내용을 교과별로 살펴보면 다음과 같다.

### (1) 국어

2015 개정 고등학교 국어과 교육과정은 초등학교, 중학교와 마찬가지로 핵심 내용을 선별하고, 인문학적 소양을 갖추 수 있도록 수업 시간에 한 학기 한 권의 책을 읽고, 생각을 나누며, 쓰는 통합적 독서 활동을 강화하였다. 이에 따라 모든 선택 과목(일반 선택: 『화법과 작문』, 『독서』, 『언어와 매체』, 『문학』 / 진로 선택: 『실용 국어』, 『심화 국어』, 『고전 읽기』)의 ‘4. 교수학습 및 평가의 방향’에 1학기 1권 통합적 독서 활동을 위한 수업 시간을 확보하고 도서 준비를 안내하였다. 또한 교과에서 다룰 핵심 내용을 선별하여 국어과 성취기준을 26개로 줄이고, ‘반모음 첨가’, ‘고대-중세-근대의 음운 변화’ 등의 내용은 삭제하였다. 반면 ‘문제해결을 위한 독서/작문’ 등과 같은 내용은 신설하였다.

### (2) 수학

2015 개정 고등학교 수학과 교육과정의 경우, 학생들이 수학에 흥미와 자신감을 잃지 않도록 학생 발달 단계와 국제적 기준(global standards)을 고려하여 핵심 개념과 원리를 중심으로 학습 내용을 재구성하였다. 예컨대 『수학 I』과 『수학 II』의 ‘연립일차방정식(미지수 3개)’은 단순 문제풀이를 유발한다는 측면에서 삭제하였고, 현장 교사 의견에 따라 좀 더 강조해야 할 내용으로 판단된 ‘사인법칙과 코사인법칙’은 추가되었다. 또한 학생의 진로와 적성에 따른 맞춤형 교육과 수월성 교육이 이루어질 수 있도록 진로 선택 과목으로 『실용 수학』, 『경제 수학』, 『수학과제 탐구』, 『심화 수학 I, II』를 신설함으로써 선택 과목을 재구조화하였다.

아울러, 활동과 탐구 중심으로 교수·학습 방법을 개선하고자 하였다. 예컨대, 『확률과 통계』에서는 ‘표본조사의 의미’를 추가함으로써 미디어 등의 활용도를 높이하고자 하였다. 또한 ‘평가 방법 및 유의 사항’을 신설하여 교육과정을 벗어난 내용을 평가하지 않도록 안내(예: ‘집합의 개념이나 집합의 포함관계는 개념을 이해하는 수준에서 간단히 평가한다.’) 하고 학생들의 학습 부담을 실질적으로 경감시키고자 하였다.

### (3) 영어

2015 개정 영어과 교육과정의 경우, 초·중학교에서는 ‘듣기’와 ‘말하기’에 중점을 둔 것과 달리 고등학교에서 ‘읽기’, ‘쓰기’ 학습을 강조하는 등 언어발달 단계와 학생발달 단계를 고려하여 의사소통 중심 교육을 강화하였다. 예컨대, 초등학교 31%, 중학교 26%였던 ‘듣기’ 비율을 고등학교에서는 24%로 감소시키고, 초·중학교 때 각각 18%였던 ‘쓰기’ 비율을 고등학교에서는 28.5%로 증가시켰다. 또한 고등학교에서는 학습 내용을 정선하기 위해 ‘논리적 관계’를 ‘인과관계 파악’에 포함·통합시켰다.

아울러 기본적으로 학습해야 할 어휘 수는 3천 개 수준으로 유지하되, 어휘 목록과 언어 형식을 학교급별로 구분하여 제시하였다. 즉, 전체 학습량은 유지하되 학습 부담은 경감하기 위해 일반 선택(『영어 회화』, 『영어 I』, 『영어 독해와 작문』, 『영어 II』), 진로 선택(『실용 영어』, 『영어권 문화』, 『진로 영어』, 『영미 문학 읽기』) 및 전문 교과 I(『심화 영어 회화 I·II』, 『심화 영어 I·II』, 『심화 영어 독해 I·II』, 『심화 영어 작문 I·II』)별로 학습 어휘를 구분하여 제시하였으며, 언어 형식 역시 고등학교 권장 형식을 별도로 제시하였다.

### (4) 사회

2015 개정 고등학교 사회과 교육과정은 2009 개정 교육과정에서의 ‘지

리’, ‘역사,’ ‘일반사회’의 구분은 유지하되, 핵심 개념을 중심으로 초·중·고 연계를 강화하고자 하였다. 특히 그 동안 학생들이 학교에 개설된 사회과 과목을 수능 위주로 선택 이수함에 따라, 지식을 편중적으로 학습한다는 비판이 지속적으로 제기되어 왔다. 이에 2015 개정 고등학교 사회과 교육과정에서는 공통 과목으로 『통합사회』를 신설하였다. 이 과목은 “인간을 둘러싼 자연과 사회 현상에 대해 시간적, 공간적, 사회적, 윤리적 관점을 적용하여 사회 현상을 종합적으로 이해하도록 한 것”(교육부, 2015c)으로, 학생들이 삶 속에서 중요하게 다루어야 하는 9개의 핵심 개념을 선정하여 사회 현상의 특징, 사회 문제의 발생 원인과 해결 방안, 자연과 인간 삶의 조화, 사회적 갈등 해결방안 등을 모색하도록 하고 있다. 2015 개정 고등학교 교육과정 문서에 제시된 『통합사회』 과목의 내용 체계는 <표 II-7>과 같다.

<표 II-7> 고등학교 ‘통합사회’ 내용 체계(교육부, 2015f)

영역	핵심 개념	내용 요소	기능
삶의 이해와 환경	행복	<ul style="list-style-type: none"> <li>• 통합적 관점</li> <li>• 행복의 조건</li> </ul>	파악하기 설명하기
	자연환경	<ul style="list-style-type: none"> <li>• 자연환경과 인간 생활</li> <li>• 자연관</li> <li>• 환경 문제</li> </ul>	조사하기 비교하기 분석하기
	생활공간	<ul style="list-style-type: none"> <li>• 도시화</li> <li>• 산업화</li> <li>• 정보화</li> </ul>	제안하기 적용하기 추론하기
인간과 공동체	인권	<ul style="list-style-type: none"> <li>• 시민 혁명</li> <li>• 인권 보장</li> <li>• 인권 문제</li> </ul>	분류하기 예측하기 탐구하기
	시장	<ul style="list-style-type: none"> <li>• 합리적 선택</li> <li>• 국제 분업</li> <li>• 금융 설계</li> </ul>	평가하기 비판하기
	정의	<ul style="list-style-type: none"> <li>• 정의의 의미</li> <li>• 정의관</li> <li>• 사회 및 공간 불평등</li> </ul>	종합하기 판단하기 성찰하기

영역	핵심 개념	내용 요소	기능
사회 변화와 공존	문화	<ul style="list-style-type: none"> <li>• 문화권</li> <li>• 문화 변동</li> <li>• 다문화 사회</li> </ul>	표현하기
	세계화	<ul style="list-style-type: none"> <li>• 세계화</li> <li>• 국제사회 행위 주체</li> <li>• 평화</li> </ul>	
	지속가능한 삶	<ul style="list-style-type: none"> <li>• 인구 문제</li> <li>• 지속가능한 발전</li> <li>• 미래 삶의 방향</li> </ul>	

또한 협력학습, 프로젝트 수업 등 학생 활동 중심의 수업을 통해 문제 해결력, 의사결정능력 등을 함양할 수 있도록 탐구활동의 예시를 제시하고 학습량을 적정화하였다. 예컨대, 『경제』의 ‘경제주체의 역할과 의사결정’, 『사회·문화』의 ‘일상생활과 사회 제도’는 삭제하였다. 또한 『법과 정치』는 학생들의 선호도를 회복하기 위해 과목명을 『정치와 법』으로 변경하고, ‘국제 관계’ 부분에 ‘한반도의 국제 질서’ 내용을 추가하였다. 그리고 학생들의 다양한 진로를 고려하여 『여행 지리』, 『사회문제 탐구』와 같은 진로 선택 과목을 신설하였다.

#### (5) 과학

2015 개정 과학과 교육과정의 경우 ‘모두를 위한 과학(Science for all) 교육’을 목표로, 고등학교 1학년까지는 주변 자연 현상에 대한 궁금증을 과학 개념과 연결시켜 삶의 즐거움을 경험할 수 있도록 재미있고 쉽게 구성하였다. 이를 위해 고등학교에서는 공통 과목으로 『통합과학』 과목을 신설하였다. 이 과목은 자연현상에 대한 네 가지 핵심 영역, 즉 ‘물질과 규칙성’, ‘시스템과 상호작용’, ‘변화와 다양성’, ‘환경과 에너지’를 중심으로, 분과 학문적 지식을 넘어 다양한 형태의 통합을 통해 융·복합적 사고력을 신장할 수 있도록 구성되었다(<표 II-8> 참조).

<표 II-8> 고등학교 ‘통합과학’ 내용 체계(교육부, 2015g)

영역	핵심 개념	내용 요소	기능
물질과 규칙성	물질의 규칙성과 결합	<ul style="list-style-type: none"> <li>우주 초기의 원소(생성)</li> <li>태양계에서 원소 생성</li> <li>지구의 고체 물질 형성</li> <li>금속과 비금속</li> <li>최외각 전자</li> <li>이온 결합</li> <li>공유 결합</li> </ul>	<ul style="list-style-type: none"> <li>문제 인식</li> <li>탐구 설계와 수행</li> <li>자료의 수집·분석 및 해석</li> <li>수학적 사고와 컴퓨터 활용</li> <li>모형의 개발과 사용</li> <li>증거에 기초한 토론과 논증</li> <li>결론 도출 및 평가</li> <li>의사소통</li> </ul>
	자연의 구성 물질	<ul style="list-style-type: none"> <li>지각과 생명체 구성 물질의 규칙성</li> <li>생명체 주요 구성 물질</li> <li>신소재의 활용</li> <li>전자기적 성질</li> </ul>	
시스템과 상호작용	역학적 시스템	<ul style="list-style-type: none"> <li>중력</li> <li>자유 낙하</li> <li>운동량</li> <li>충격량</li> </ul>	
	지구 시스템	<ul style="list-style-type: none"> <li>지구 시스템의 에너지와 물질 순환</li> <li>기권과 수권의 상호 작용</li> </ul>	
	생명 시스템	<ul style="list-style-type: none"> <li>세포막의 기능</li> <li>세포 소기관</li> <li>물질대사, 효소</li> <li>유전자(DNA)와 단백질</li> </ul>	
변화와 다양성	화학 변화	<ul style="list-style-type: none"> <li>산화와 환원</li> <li>산성과 염기성</li> <li>중화 반응</li> </ul>	
	생물다양성과 유지	<ul style="list-style-type: none"> <li>지질 시대</li> <li>화석, 대멸종</li> <li>진화와 생물다양성</li> </ul>	
환경과 에너지	생태계와 환경	<ul style="list-style-type: none"> <li>생태계 구성요소와 환경</li> <li>생태계 평형</li> <li>지구 온난화와 지구 환경 변화</li> <li>에너지 전환과 보존</li> <li>열효율</li> </ul>	
	발전과 신재생 에너지	<ul style="list-style-type: none"> <li>발전기</li> <li>전기 에너지</li> <li>전력 수송</li> <li>태양 에너지</li> <li>핵발전</li> <li>태양광 발전</li> <li>신재생 에너지</li> </ul>	



즉, 『통합과학』은 여러 분야에서 활용되고 있는 기초 개념을 융합하고, 동료 간 협업적 탐구활동을 통해 새로운 가치를 창출하고 역량을 습득할 수 있는 대주제 학습 접근을 취하였다. 아울러 학교 밖 현장 체험을 통한 실생활 연계, 전통적인 과학-기술-사회 연계 학습인 STS 학습 등으로 구성된다. 또한 문·이과 공통으로 『과학탐구실험』 과목을 신설함으로써 탐구 활동과 체험 중심의 학습을 강화하였다.

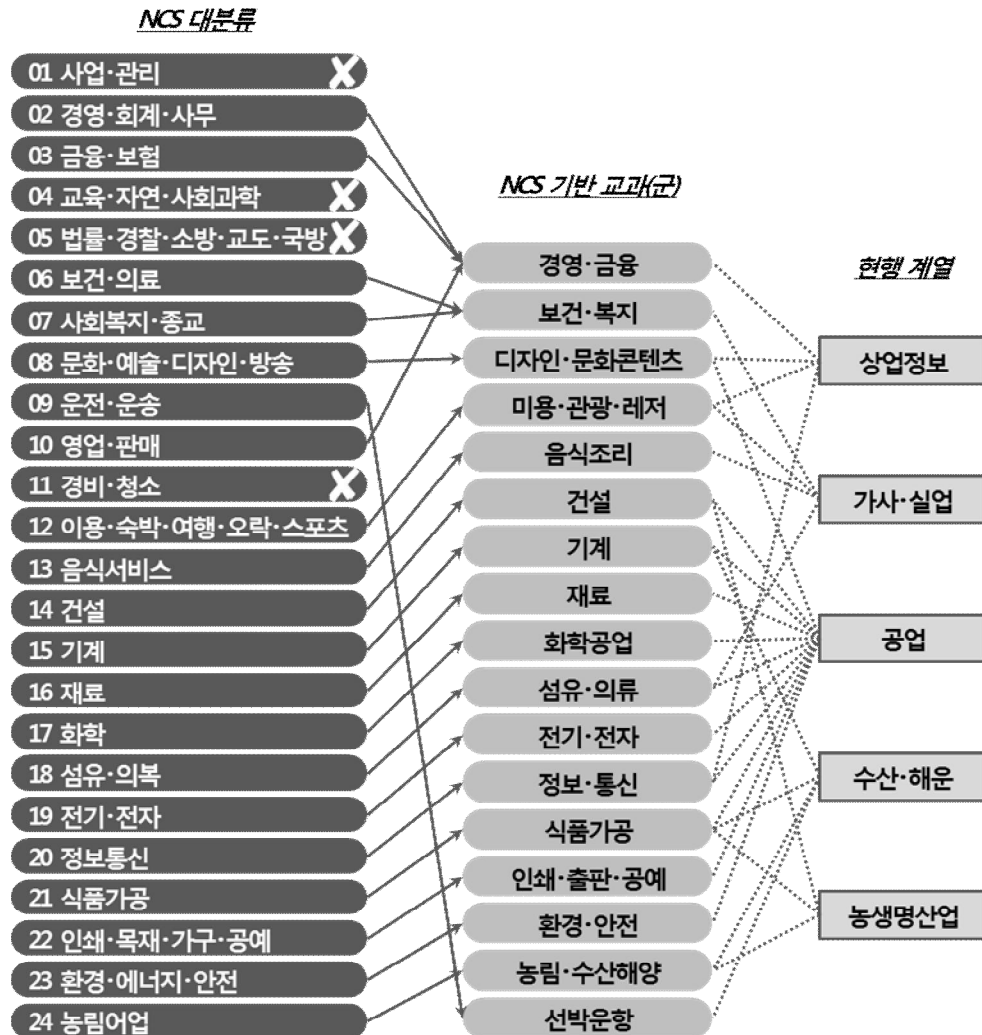
한편 고등학교 2학년부터는 자신의 진로를 고려하여 진로 선택 과목 및 심화 과목을 이수할 수 있도록 과목을 유기적으로 구성하였다. 즉, 진로에 따라 『물리학Ⅱ』, 『화학Ⅱ』, 『생명과학Ⅱ』, 『지구과학Ⅱ』를 선택 이수하고, 『고급 물리학』, 『고급 화학』, 『고급 생물학』, 『고급 지구과학』 등의 전문 교과Ⅰ 과목을 선택할 수 있게 함으로써 수월성 교육이 가능하도록 하였다.

## (6) 전문 교과Ⅱ

2015 개정 교육과정의 고등학교 전문 교과Ⅱ는 산업 현장에서 요구하는 직무를 중심으로 직업교육체제를 구축하여 ‘할 줄 아는 교육’으로 개선하고자 하였다. 즉, 전문 교과Ⅱ는 NCS 대분류와 교육과정 간의 연계를 강화하기 위해 한국고용직업분류(KECO)에 따른 인력구조 특성에 대한 분석을 바탕으로, 현행 5개 계열(상업정보, 가사·실업, 공업, 수산·해운, 농생명산업)에서 17개 교과(군)로 개편되었다. 이때 교과 체제는 ‘전문 공통 과목’, ‘기초 과목’, ‘실무 과목’으로 구성되어 있으며, NCS 기반 전문 교과Ⅱ(군)의 구조는 [그림 Ⅱ-2]와 같다.

또한 특성화고 및 마이스터고 졸업생이 수행할 수 있는 직무를 명확히 설정하기 위해 NCS 중분류를 활용하여 기준 학과를 설정하고, 이에 해당하는 인력양성 유형(예: 고객 상담원)도 설정하였다. 예컨대, ‘음식·조리’ 교과(군)의 경우 전문 공통 과목은 『성공적인 직업생활』이고, 기초 과목은 『식품과 영양』과 『급식 관리』, 실무 과목은 『한국 조리』, 『서

양 조리』, 『중식 조리』, 『일식 조리』, 『소믈리에』, 『바리스타』, 『바텐더』이며, 기준 학과는 ‘조리·식음료과’로 설정된다.



[그림 II-2] NCS 기반 전문 교과Ⅱ(군)의 구조(교육부, 2015a)

이상의 내용을 바탕으로 2009 개정 교육과정과 2015 개정 교육과정을 비교하면 <표 II-9>과 같다.

<표 II-9> 2009 개정 교육과정과 2015 개정 교육과정의 비교

구분			주요 내용	
			2009 개정	2015 개정
교육과정 개정 방향			<ul style="list-style-type: none"> <li>· 창의적 인재 양성</li> <li>· 전인적 성장을 위한 창의적 체험활동 강화</li> <li>· 국민공통교육과정 조정, 학교교육과정 편성·운영의 자율성 강화</li> <li>· 교육과정 개편을 통한 수능 제도 개혁 유도</li> </ul>	<ul style="list-style-type: none"> <li>· 창의융합형 인재 양성</li> <li>· 모든 학생이 인문·사회·과학 기술에 대한 기초 소양 함양</li> <li>· 학습량 적정화, 교수·학습 및 평가 방법 개선을 통한 핵심역량 함양</li> <li>· 수능·대입제도 연계, 교원 연수 등 교육 전반 개선</li> </ul>
총론	공통 사항	핵심역량 반영	· 명시적 규정 없이 일부 교육과정 개발에서 고려	<ul style="list-style-type: none"> <li>· 총론 ‘추구하는 인간상’ 부문에 6개 핵심역량 제시</li> <li>· 교과별 역량 제시, 역량 함양 위한 성취기준 개발</li> </ul>
		인문학적 소양 함양	· 예술고 심화선택 ‘연극’ 개설	<ul style="list-style-type: none"> <li>· ‘연극’ 일반선택으로 개설</li> <li>· 독서교육 강화</li> </ul>
		소프트웨어 교육 강화	· 심화선택 ‘정보’	· ‘정보’ 과목을 일반선택으로 전환, SW 중심 개편
		안전교육 강화	· 교과, 창의적 체험활동에 안전 내용 포함	<ul style="list-style-type: none"> <li>· ‘안전한생활’ 신설</li> <li>· 관련 교과에 안전 단원 신설</li> </ul>
		범교과학습 주제 개선	· 39개의 범교과 학습 주제 제시	· 10개의 범교과학습 주제로 범주화
		NCS 직업 교육과정 연계	<신설>	· ‘교육과정 구성의 중점’ 등에 반영
	초등학교	초1,2 수업시수 증배	<개선>	<ul style="list-style-type: none"> <li>· 주당 1시간 증배, ‘안전한 생활’ 신설(64시간)</li> <li>- 창의적 체험활동에서 체험 중심 교육으로 실시</li> </ul>
		한글교육 강화	· 초1~2학년군: 27차시	· 초등학교 교육과정과 누리과정의 연계 강화 위해 한글교육 61차시로 확대

구분			주요 내용	
			2009 개정	2015 개정
중 학 교	자유학기제 편제 방안	<신설>	· ‘교육과정 편성·운영의 중점’에 자유학기제 교육과정 운영 지침 제시	
	고 등 학 교	공통 과목 신설 및 이수단위	· 공통 과목 없이 전 학 년 선택 과목으로 구 성	· 공통 과목 및 선택 과목(일 반선택/진로선택)으로 구성 · 진로선택 및 전문 교과 통한 맞춤형 교육, 수월성 교육
		특목고 과목	· 보통교과 심화과목으 로 편성	· 보통교과에서 분리하여 전 문 교과로 제시
		국·수·영 비중 적정화	· 교과 총 이수단위의 50%를 초과할 수 없 음(국·수·영 90단위)	· 기초 교과 이수단위 제한 규 정(50%) 유지(국·수·영·한국 사 84단위)
		특성화고 교육과정	· 특성화고 전문 교과로 제시	· 총론과 NCS 교과의 연계 - NCS 기초 과목, 실무 과목
교과교육과정 개정 방향		<개선>	· 총론과 교과교육과정 간의 유기적 연계 강화	
		<개선>	· 교과교육과정 개정 기본방 향 제시 - 핵심개념 중심의 학습량 적정화 - 핵심역량 반영 - 학생참여중심 교수·학습 - 과정중심 평가 확대	
지 원 체 제	교과서	<개선>	· 흥미롭고 재미있는 질 높은 교과서 개발	
	대입 제도 및 교원	<개선>	· 교육과정에 부합하는 수능 및 대입 제도 도입 검토 · 교원양성기관 질 제고, 연수 확대	

\* 이 표는 교육부(2015a), 교육부(2017c)의 내용을 재구성한 것임.

## 다. 2015 개정 교육과정 정책 관련 선행연구 분석

이 절에서는 2015 개정 교육과정 정책 관련 선행연구들을 전반적으로 개관하기 위해 앞서 구분한 정책 과정(policy process)에 따라 관련 내용을 분석하였다. 그리고 언론보도 분석을 활용해 2015 개정 교육과정 정책의 쟁점을 도출한 선행연구들을 별도로 분석함으로써 이 연구에 대한 시사점을 얻고자 하였다.

### 1) 정책 과정에 따른 선행연구 분석

2015 개정 교육과정 정책 관련 선행연구의 내용은 정책 과정, 즉 정책의 형성, 결정, 집행 단계에 따라 변화하는 양상을 보였다. 선행연구의 내용을 정책 단계별로 나누어 살펴보면 다음과 같다.

#### 가) 정책 형성 단계

정책 형성 단계에서는 2017학년도 대입제도 확정안을 통해 발표된 ‘문·이과 통합안’을 바탕으로 문·이과 통합형 교육과정의 개발 방향을 검토한 연구(김경자, 2014; 한혜정, 2014; 허경철, 2014; 홍후조, 2014; 홍후조, 임유나, 2014; 황규호, 2014a)가 많았다. 이 연구들은 주로 국가교육과정 개정연구위원회에 참여한 연구진에 의해 수행되어 최종 고시된 2015 개정 교육과정의 방향과 일치하는 경우가 많았다(유예림, 백순근, 2016). 일례로, 황규호(2014a)는 문·이과 통합형 교육과정이 학생 개개인의 꿈과 끼를 키워주는 ‘다양하고 유연한 개별화 교육과정’임과 동시에 원리 이해와 탐구력 함양을 중시하는 ‘참된 학습경험 중심 교과 교육과정’이 될 수 있는 방향으로 개발되어야 한다고 하였는데, 이러한 방향 설정은 다양한 선행연구들에서 대체로 일치하는 경향을 보였다.

다만 고등학교에서의 공통 과목과 선택 과목의 편제 방향에 대해서는 학자들마다 이견이 있었다. 예컨대, 허경철(2014)은 문·이과가 통합된 단일 과정에서 공통 과목과 선택 과목의 종류와 비중이 제시되어야 한다고 하였다. 반면, 홍후조(2014)는 학생의 다양한 선택권이 보장될 수 있도록 고등학교 교육을 진로집중과정 방식으로 다양화하고, 이공계열의 융합인재교육(STEAM: Science·Technology·Engineering·Arts·Mathematics)뿐 아니라 인문·사회계열 교과를 아우르는 총론(LASS: Language·Arts·Society·Science)을 만들어 계열 내 통합교육을 유도해야 한다고 제안하였다.

또한 선행연구에서는 새로운 교육과정을 성공적으로 운영하기 위해 학교 현장에서의 교수·학습 및 평가활동, 수능·대입제도, 교원 양성·연수 등 제반 정책의 개선 방안을 함께 논의할 필요가 있다고 언급하였다. 예컨대, 한혜정(2014)은 문·이과 통합형 교육과정을 반영한 수능이 공통 과목 중심으로 출제될 경우 고등학교 교수·학습이 획일화될 수 있고, 선택 과목을 포함할 경우 학습 부담이 증가될 수 있으므로 이를 종합적으로 고려해 개선 방안을 마련해야 하며, 장기적으로는 선다형 중심의 평가체제를 재검토할 필요가 있다고 하였다. 한편, 교원 연수와 관련하여 홍후조(2014)는 교사들이 교과를 문과형, 이과형, 예술형, 체육형으로 전환하고 재구성할 수 있는 능력을 배양할 수 있도록 지원해야 한다고 하였다.

#### 나) 정책 결정 단계

정책 결정 단계에서는 2014년 9월에 발표된 문·이과 통합형 교육과정 총론 주요사항 또는 이후 개발된 2015 개정 교육과정 시안이 전반적(소경희, 2014; 이광우, 2014; 황규호, 2014b) 또는 교과별 교육과정 개발 및 편성·운영에 주는 시사점을 논의한 연구(구정화, 2015; 김병수, 2015; 소경희, 2015; 은지용 외, 2015; 조대현, 2014; 한혜정, 백경선, 2014 등)가 주를 이루었다. 예컨대, 소경희(2015)는 개정 교육과정 총론의 주요 특징을 ① 학습에 대한 수사학적 강조, ② 강한 교수법적 방향 제시, ③ 국

가·사회적 요구에 대한 민감한 대응으로 정리한 후, 교과별로 특수 역량에 대한 조작적 의미와 고유한 실행적 지식을 추가적으로 탐색·규명할 필요가 있다고 하였다. 이와 관련하여 김병수(2015)는 통합형 국어과 교육과정의 경우, 핵심역량을 기반으로 영역의 통합을 추구하고, 교사 친화적이되 학생 이해에 기반이 되어야 한다고 제안하였다.

이외에도 2015 개정 교육과정 정책에 대한 이해당사자들의 인식을 분석하여 향후 발전 방향을 도출하거나(강현석, 이지은, 전호재, 2015; 박소영, 이정렬, 강현석, 2015; 박채형, 2015; 조상연, 2015), 해당 정책의 도입에 따른 교원양성기관의 대처 방안을 제안한 연구(이원희, 이지은, 2014)도 수행되었다. 예컨대, 강현석, 이지은, 전호재(2015)는 2015 개정 교육과정의 쟁점을 ① 과목 신설, ② 범교과 학습, ③ 교육과정 적정화 및 운영 방식의 문제로 설정하고, 이에 대한 이해당사자별 요구 사항을 분석하기 위해 델파이조사와 설문조사를 실시하였다. 그 결과, 총론과 각론 개발의 연계 강화를 위해 내용 체계 제시방식에서 요소별 구분을 명확히 하고, 교과별 핵심역량과 특성에 부합하는 교수·학습 및 평가 방법을 모색할 것을 강조하였다. 또한 이원희와 이지은(2014)은 초등학교 교원양성 대학들이 다양한 수업 내용과 방법을 모색하고, 통합 수업의 가능성을 검토하며, 현직 교사와의 멘토링제를 도입할 것 등을 제안하였다.

#### 다) 정책 집행 단계

정책 집행 단계에서는 2015년 9월에 확정·고시된 2015 개정 교육과정 문서에서 제시된 핵심역량과 추구하는 인간상 설정의 타당성을 분석하고(강이화, 2015; 박선화, 2016; 이원희, 2015; 임유나, 2016; 한혜정, 김영은, 이주연, 2016; 홍은숙, 2015), 교과별 핵심개념 및 내용, 성취기준 등을 분석함으로써 교과서 집필(방정숙, 이지영, 서은미, 2016; 양정혜, 2015; 유병건, 김자미, 이원규, 2016; 유병열, 2016; 유종열, 2016) 및 교수·학습 및 평가 방법의 개선 방안(옥일남, 2016; 임은진, 2015; 장은주, 2016; 홍

원표, 2016)을 논의한 연구가 주를 이루었다. 예컨대, 강이화(2015)는 언어 네트워크 분석을 통해 총론에 제시된 핵심역량이 초등학교 통합교과에 타당하게 반영되었는지 확인하였다. 분석 결과, 의사소통 역량의 경우 타 역량에 비해 중요하게 사용되지 못하였으며, 이를 통해 교과별 교과역량의 타당한 선정과 역량 간 균형의 문제를 재검토할 필요가 있다고 하였다. 한편, 장은주(2016)는 국어과 교육과정의 중학교 언어기능 영역 성취기준에 대한 적합도를 현직 국어교사에게 평가하게 한 결과 대체로 적합한 것으로 나타났다. 다만, 신설된 성취기준의 경우 내용 타당성이 낮게 평가된 만큼 이들 성취기준의 교육적 타당성에 대한 합의, 관련 교수·학습 전략 및 사례에 대한 정보가 보충될 필요가 있음을 지적하였다. 또한 임은진(2016)은 핵심역량을 타당하게 평가하기 위해 과제 중심의 분석적 루브릭을 적극 활용할 것을 제안하기도 하였다.

또한 교원양성 및 연수 측면에서 개선 방안을 논한 연구(김갑수, 2016; 김남희, 심규철, 2015; 김이경 외, 2016; 윤지현, 강성주, 2016; 이성숙, 2016)도 수행되었다. 예컨대, 김이경 외(2016)는 설문조사를 통해 2015 개정 교육과정에 대한 교원들의 관심도와 새로운 내용을 현직 연수를 통해 함양하려는 의지가 강하다는 것을 밝히며, 향후 교원 연수의 내용을 좀더 구체적으로 마련하고, 교원들의 참여를 독려하기 위해 학교 업무 감면, 연수 비용 지원 등의 유인책을 마련해야 한다고 제안하였다. 다만 김갑수(2016)는 소프트웨어 교육에 대한 초등 교원의 이해도는 부족한 편이므로 관련 연수를 적극적으로 시행할 필요가 있음을 강조하였다.

2015 개정 교육과정 도입에 따른 수능 체제 개편의 방향과 쟁점에 관해 구체적으로 논의한 연구(박경호 외, 2016; 이용상 외, 2015)도 일부 수행되었다. 예컨대, 이용상 외(2015)는 현 수능 체제 하에서 공통과목만을 수능 응시과목으로 지정할 경우 고등학교 교육과정의 파행 운영, 문·이과 계열 간 유불리 등의 문제가 발생할 수 있기 때문에, 공통과목 이외에 선택과목을 어떻게 반영할 것인가에 대한 부분이 중요한 쟁점이 되고 있다고 하였다. 또한 박경호 외(2016)는 선행연구 분석과 포커스 그



를 인터뷰를 통해 문·이과 통합형, 교과별 성취수준 평가, 영향력 축소라는 기본원칙을 도출한 후 이에 기반하여 수능 개선 방안을 세 가지로 제시하였다. 2019학년도 수능과 유사하게 마련한 1안과 달리, 2안의 경우 2015 개정 교육과정에 따른 고교 공통 과목만을 반영한 것이며, 3안의 경우 수능을 이원화하여 수능 I에서는 공통 과목을, 수능 II에서는 2~3개 교과(군) 중 최대 3개의 선택 과목을 응시하는 형태이다.

끝으로, 정책평가 연구의 맥락에서 2015 개정 교육과정 정책의 결정 단계에 대해 비판적으로 고찰한 연구(김응현, 2015; 박선미, 2016; 박희경; 2016; 양정실, 2016)도 수행되었다. 예컨대, 박선미(2016)는 2015 개정 중학교 지리교육과정의 경우 교육과정 개발에 누가 참여하느냐보다 의사결정구조가 어떻게 조직되어 있는가의 문제가 결과에 큰 영향을 미친다고 언급하며, 관련 학계가 상호 협력적 거버넌스를 구축할 필요가 있음을 강조하였다.

## 2) 언론보도 내용을 활용한 쟁점 분석 관련 선행연구 고찰

언론보도 내용을 활용해 2015 개정 교육과정 정책의 쟁점을 분석한 선행연구는 다음과 같다(<표 II-10> 참조). 우선, 김소영과 강신원(2015)은 관련 온라인 뉴스의 보도 내용을 분석하기 위해 개정 방향, 총론 내용, 지원체제, 기타 전반적 내용을 분석기준으로 설정하였다. 그리고 이를 바탕으로 내용 분석(content analysis)을 실시한 결과, ‘국가 교육과정 개정에 대한 수요자의 무관심’, ‘교과서 국정화 관련 갈등’, ‘찾은 개정주기로 인한 혼란’, ‘수능 체제에 대한 동반 개편 요구’, ‘국가 정책에 대한 언론의 소극적 태도’의 총 5개 쟁점을 도출하였다. 또한 이윤복과 강현석(2015)은 주요 언론사와 교육단체의 기사 내용 및 보도자료에 대한 내용 분석을 통해 관련 쟁점을 도출하였다. 이 연구에서는 분석 기준을 교육과정 총론, 각론, 개정 방식, 초등학교, 중학교, 고등학교로 설정하고, ‘문·이과 통합형의 취지 반영’, ‘학습 부담과 내용량 감축’, ‘총론과 각론의

연계’, ‘개정 방식과 절차’ 등의 총 11개 쟁점을 도출하였다.

한편, 강지영과 소경희(2016)는 국내 주요 10대 일간지를 통해 보도된 기사들을 대상으로 교육부와 교육과정 개발진을 중심으로 하는 ‘융합형 인재 양성 교육과정’ 담론, 과학계 인사를 중심으로 형성된 ‘이과 말살 교육과정’ 담론, 그리고 시민사회단체와 교원단체를 중심으로 형성된 ‘현장과 유리된 교육과정’ 담론을 2015 개정 교육과정 총론의 주요 쟁점으로 분석하였다. 그리고 유예림과 백순근(2016)은 국내 주요 20개 언론사를 통해 보도된 뉴스 기사들을 대상으로 자동화된 텍스트 내용 분석 방법인 토픽 모델링을 활용하여 ‘역사 교과서 국정화’, ‘소프트웨어 교육’, ‘수업 및 평가 방법 개선’, ‘역사교육 내용’, ‘수능 및 대입제도 개선’, ‘수학교육 내용’, ‘교과서 한자 병기’, ‘안전 교육’의 8개 쟁점을 도출하였다.

<표 II-10> 언론보도 분석을 활용한 2015 개정 교육과정 관련 선행연구

연구자 (연도)	분석 데이터 및 수집 기간	분석 방법	분석 결과(쟁점)
강지영, 소경희 (2016)	<ul style="list-style-type: none"> <li>· 국내 주요 10대 일간지의 온라인 뉴스 총 94건</li> <li>· 2013. 10. 24. ~ 2015. 9. 24.</li> </ul>	<ul style="list-style-type: none"> <li>· Fairclough의 비판적 담론분석</li> <li>· 키워드: 2015 개정 교육과정, 문·이과 통합 교육과정, 통합과학, 통합사회, 교육과정 개정</li> </ul>	<ul style="list-style-type: none"> <li>· 융합형 인재 양성 교육과정</li> <li>· 이과 말살 교육과정</li> <li>· 현장과 유리된 교육과정</li> </ul>
김소영, 강신원 (2015)	<ul style="list-style-type: none"> <li>· 국내 검색포털 온라인 뉴스 총 103건</li> <li>· 2014. 9. 24. ~ 2014. 10. 23.</li> </ul>	<ul style="list-style-type: none"> <li>· 내용분석               <ul style="list-style-type: none"> <li>- 키워드 코딩</li> <li>- 연구자, 검토자 간 주제 분류 일치도 검토</li> </ul> </li> <li>· 키워드: 2015 개정 교육과정</li> </ul>	<ul style="list-style-type: none"> <li>· 국가 교육과정 개정에 대한 수요자의 무관심</li> <li>· 교과서의 국정화 논란</li> <li>· 잦은 개정으로 인한 혼란</li> <li>· 국가 교육과정 개정과 수능의 동반 개편 요구</li> <li>· 국가 정책에 대한 언론의 소극적 태도</li> </ul>

연구자 (연도)	분석 데이터 및 수집 기간	분석 방법	분석 결과(쟁점)
유예림, 백순근 (2016)	<ul style="list-style-type: none"> <li>· 국내 주요 20대 언론사의 온라인 뉴스 총 367건</li> <li>· 2013. 8. 1. ~ 2016. 5. 30.</li> </ul>	<ul style="list-style-type: none"> <li>· 자동화된 텍스트 내용 분석               <ul style="list-style-type: none"> <li>- 토픽 모델링</li> </ul> </li> <li>· 키워드: 문이과 통합 교육과정, 2015 개정 교육과정</li> </ul>	<ul style="list-style-type: none"> <li>· 역사 교과서 국정화</li> <li>· 소프트웨어 교육</li> <li>· 수업 및 평가 방법 개선</li> <li>· 역사교육 내용</li> <li>· 수능 및 대입제도 개선</li> <li>· 수학교육 내용</li> <li>· 교과서 한자 병기</li> <li>· 안전 교육</li> </ul>
이윤복, 강현석 (2015)	<ul style="list-style-type: none"> <li>· 국내 주요 10대 일간지의 인터넷 뉴스 기사(분석 건수 미보고)</li> <li>· 2013. 10. 1. ~ 2015. 9. 30.</li> </ul>	<ul style="list-style-type: none"> <li>· 내용분석               <ul style="list-style-type: none"> <li>- 키워드 코딩 및 분류</li> </ul> </li> <li>· 키워드: 2015 개정 교육과정</li> </ul>	<ul style="list-style-type: none"> <li>· 문·이과 통합형의 취지반영</li> <li>· 역량중심 교육과정</li> <li>· 학습부담과 내용량 감축</li> <li>· 총론과 각론의 연계</li> <li>· 개정 방식과 절차</li> <li>· 초: 교과서 한자병기, 안전교육</li> <li>· 중: 자유학기제, 소프트웨어 교육</li> <li>· 고: 수능과의 연계, 진로선택 과목</li> </ul>

이 연구들은 대부분 전문 연구자들의 질적 연구방법에 근거하여 2015 개정 교육과정 정책의 쟁점을 체계적으로 도출했다는 점에서 의미가 있다. 그러나 연구 수행의 시점 상 정책의 형성, 결정, 집행 단계의 뉴스 기사를 모두 포괄하지 못한 점, 검색포털 등을 통해 공급되는 인터넷 뉴스의 파급력과 다양성을 고려하지 않고 주요 일간지의 보도 기사로 분석대상을 제한한 점 등은 보완될 필요가 있다. 또한 연구 방법론 측면에서 유예림과 백순근(2016)의 연구를 제외하고는 모든 선행연구에서 전통적인 내용 분석을 활용하였는데, 이 경우 전문 연구자들은 분석 기준에 대한 일련의 주제어들(keywords)을 사전에 마련해야 한다(Lewis, Zamith, & Hermida, 2013). 그러나 자동화된 내용 분석 기법의 경우 주제어를 자동으로 탐색·제공함으로써 연구의 객관성과 재현 가능성을 제고할 수 있기 때문에 쟁점 분석시 적용하는 것을 고려해 볼 필요가 있다.

### 3. 빅데이터 분석 기법

#### 가. 빅데이터의 개념과 특성

빅데이터(big data)란 맥킨지 보고서에서 언급된 바와 같이 “일반적인 데이터베이스 프로그램이 수집·저장·관리·분석하기 어려운 대규모 데이터”(Manyika et al., 2011)로 간단히 정의할 수 있다. 그러나 가트너 보고서에서는 빅데이터를 “더 나은 의사결정, 시사점 발견 및 프로세스의 최적화를 위해 새로운 형태의 정보처리가 필요한 대용량, 초고속 및 다양성의 특성을 가진 정보 자산”(Beyer & Laney, 2012)이라고 정의함으로써 단순히 규모 차원이 아닌 활용 측면에도 초점을 맞추었다. 국제데이터협회(The International Data Corporation: IDC) 역시 빅데이터를 “초고속 수집·발굴·분석을 통해 대규모 데이터로부터 저비용으로 가치를 추출할 수 있도록 고안된 새로운 세대의 기술이자 아키텍처”(Vesset et al., 2012)라고 정의하면서 빅데이터의 가치 창출 측면을 부각시켰다. 이를 종합하면, 빅데이터는 ‘새로운 처리·분석 방법을 활용해 저비용으로 가치를 창출할 수 있는 대규모 데이터’를 뜻한다고 할 수 있다.

빅데이터는 빅(big)이라는 수식어로 인해 양적 측면이 부각되는 경향이 있지만, 실제로는 질적 의미가 빅데이터의 본질에 더욱 가깝다고 할 수 있다. 이러한 경향으로 최근에는 1세대 빅데이터와 차별화하자는 의미에서 빅데이터2.0이 회자되고 있는데, 단순한 데이터의 양적 증가가 빅데이터 1.0이라면 초고속망 인프라와 데이터 분석력에 기반해 정밀하게 여과되어 최적화된 양질의 데이터는 ‘빅 데이터 2.0’이라고 할 수 있다(신동희, 2015). 이에 따라 빅데이터의 특성은 주로 3V(Volume, Variety, Velocity)에 2V(Value, Veracity)나 1C(Complexity)의 특성을 추가하는 방식으로 설명된다(송태민, 송주영, 2015). 즉, 빅데이터는 복잡하고(Complexity) 방대한 양(Volume)의 데이터로 다양한(Variety) 정보의 결

함이 가능하며, 데이터의 축적 속도(Velocity)가 빠를 뿐 아니라, 이를 통해 신뢰롭고(Veracity) 새로운 가치(Value)를 창출할 수 있다는 것이다. 따라서 한신잡(2015) 역시 빅데이터에 대해 논할 때에는 데이터의 양적 증가에만 주목할 것이 아니라, 데이터의 원천과 생산방식, 구성과 사용방식이 지금까지와는 근본적으로 다른 새로운 형태와 내용이라는 질적 변화의 측면에도 주목할 필요가 있다고 강조하였다. 즉, 빅데이터는 인위적 개입에 의해 별도로 만들어지는 것이 아니라, 생활의 모든 측면에서 ‘자연스럽게’ 만들어진, 있는 그대로를 반영하는 데이터라는 점에서 기존의 사회과학 데이터와 차별화된다는 것이다.

그 결과, 빅데이터는 지금까지 분석에 많이 활용된 정형화된 데이터뿐 아니라 다양한 형태의 비정형, 반정형 데이터까지 포괄한다. 조성우(2012), 김대중과 윤서연(2013)의 논의를 중심으로 각각에 대해 설명하면 다음과 같다. 첫째, 정형(structured) 데이터는 정보를 저장하는 항목(필드)이 고정되어 있는 데이터로, 관계형 데이터베이스나 스프레드시트 등을 활용해 수행하는 과정에서 축적되는 정보이기도 하다. 예컨대, 각 단위학교의 물적·인적 인프라 관련 정보는 학교의 학생수, 교원수, 교실수 등과 같은 고정된 필드에 저장되어 있다. 둘째, 비정형(unstructured) 데이터는 고정된 필드에 저장되어 있지 않은 데이터로, 검색포털과 SNS 등을 통해 생산된다. 여기에는 텍스트뿐 아니라, 이미지, 동영상, 음성 등도 포함되며, 스마트폰의 보편화와 SNS의 활성화로 인해 다른 형태의 데이터보다 훨씬 빠르게 급증하고 있다. 셋째, 반정형(semi-structured) 데이터는 고정된 필드가 일부 있는 데이터로 메타데이터(속성 정보)를 포함한다. 예컨대, 사진 자체는 비정형 데이터지만 사진의 메타데이터에 포함된 촬영시간, 위치 등은 정형 데이터이기 때문에 반정형 데이터로 간주될 수 있다. 워드프로세서, 이메일, 프리젠테이션, PDF와 같은 문서들 역시 메타데이터가 포함된 반정형 데이터로 볼 수 있다.

한편, 김용대와 조광현(2013)은 빅데이터의 통계학적 특징을 ① 관측자료, ② 전수자료, ③ 고차원적 자료로 정리하고 각 특징이 통계학적 방법

론과 어떻게 연결되는지를 고찰하였다. 첫째, 빅데이터는 기본적으로 관측자료로, 분석 목적과는 별개의 이유로 모아진 것이기 때문에 이를 활용하여 결론을 내릴 때에는 주의를 기울여야 한다. 특히 관측자료를 이용하여 인과관계를 분석하려면 관련된 교란 효과(confounding effect)를 모두 제거해야 하는데 관련 변수의 탐색과 적용이 쉽지 않다. 때문에 한신갑(2015)은 빅데이터를 활용한 사회과학 연구의 틀이 인과관계에서 상관관계로 방향이 전환될 수 있다고 하였다.

둘째, 빅데이터는 관심 모집단을 잘 정의하면 전수자료로서의 역할을 담당할 수 있으므로, 통계적 추론보다는 탐색적 자료 분석과 예측이 더 주효하다. 단, 예측을 위해서는 다양한 종류의 통계학적 가정이 필요하며, 그 중 가장 중요한 것은 현재의 자료가 미래를 잘 설명한다는 가정이다. 따라서 빅데이터가 수집된 데이터의 모집단을 벗어나는 경우까지 확장을 하여 결론을 내려서는 안되며, 빅데이터가 생성되는 모집단이 시시각각 변한다는 점을 잘 고려하여 분석해야 한다. 예컨대, 트위터 데이터를 사용할 경우 트위터 사용자라는 특정한 범주에 관한 것일 뿐 전체 인구를 대상으로 한 것이 아니므로 체계적 대표성이 보장되지 않음에 유의해야 한다(한신갑, 2012). 또한 데이터를 너무 자세하게 분석하면 과적합(overfitting) 문제로 인해 예측 성능이 저하될 수 있다.

셋째, 다양한 빅데이터가 결합될 경우 데이터는 필연적으로 차원이 높아지는데, 이때 연구자가 분석의 정확도를 일정 수준으로 유지하되 다수의 변수를 사용하고자 할 경우 데이터의 수 역시 지수적으로 증가해야 한다. 예컨대, 100개의 변수를 사용하고자 할 경우 필요한 데이터수는  $2^{100}$ 개 정도다. 따라서 고차원적 데이터를 활용할 경우 고차원 모형을 저차원 모형의 결합으로 나타내는 통계모형(예: functional ANOVA decomposition)이나 차원 축소방법(예: 주성분 분석)을 적절하게 활용하여야 한다.

## 나. 빅데이터의 분석 절차와 방법

빅데이터 분석 방법론이란 상대적으로 가치가 낮은 대량의 데이터로부터 통찰력 있는 고급 정보를 얻는 과정으로(이만재, 2012), 여기에는 데이터 분석 자체 뿐 아니라 데이터를 수집하고 클리닝하는 단계 모두를 포함한다. 따라서 이 절에서는 빅데이터의 분석 절차와 방법을 ① 데이터 수집, ② 데이터 클리닝, ③ 데이터 분석 및 시각화의 세 단계로 구분하고, 선행연구(김대중, 윤서연, 2013; 김용대, 조광현, 2013; 김정선, 권은주, 송태민, 2014; 박대현, 송동현, 2014; 서동민, 정한민, 2013; 신수정, 2014; 이미숙, 이창훈, 김지연, 2014; 진설아 외, 2013; 황명화, 2014; Liu, 2012; Maynard & Funk, 2012; Medhat, Hassan, & Korashy, 2014) 내용을 종합하여 단계별로 기술하였다.

### 1) 데이터 수집

빅데이터를 수집할 때에는 대규모의 데이터에 접근해야 하기 때문에 컴퓨터를 활용한 자동화된 방법을 활용하게 된다. 빅데이터의 수집 방법으로는 크게 ① 공개 API(Open Application Programming Interface: Open API)를 이용하는 방법과 ② 웹 크롤링(web-crawling) 또는 ③ 웹 스크래핑(web-scraping) 방법으로 구분할 수 있다.<sup>8)</sup>

우선, ‘공개 API’란 데이터를 어디서나 쉽게 활용할 수 있도록 운영체제나 프로그래밍 언어가 제공하는 기능을 제어할 수 있도록 만든 인터페이스로, 운영체제와 응용 프로그램 간 통신에 사용되는 언어나 메시지

---

8) 이 밖에도 카산드라(cassandra, 분산 시스템에서 방대한 분량의 데이터를 처리할 수 있도록 설계된 오픈소스 데이터베이스 관리시스템), 로그 수집기(log collector, 조직 내부에 존재하는 웹 서버의 로그, 클릭 로그, 데이터베이스 로그 데이터 등을 수집하는 것), RSS(Really Simple Syndication or Rich Site Summary, 데이터의 생산·공유·참여 환경인 웹 2.0을 구현하는 기술로, 필요한 데이터를 프로그래밍을 통해 수집하는 것)와 같은 방법이 있다.

형태로 구성되어 있다. 공개 API를 이용하여 데이터를 수집할 때에는 주로 검색포털이나 공공 기관의 데이터베이스와 연계되며, 이 경우 트래픽, 시스템 과부하, 데이터 소유권 등의 문제로 인해 해당 기관과 사전 협의가 필요하다. 따라서 개인은 데이터를 제공하는 기관에 인증키를 요청하고, 해당 기관의 관리자가 요청 정보를 확인한 후 API 인증 키와 적용 소스를 보내주면, 이를 해당기관의 서비스 제공 방식에 따라 접근하면 된다. 예컨대, 정형 데이터 중 지난 10여 년 간 시·도별 교원 1인당 학생 수 관련 정보를 수집하고자 하는 경우, 통계청에서 공개한 API 규약을 따르면 되며, 네이버 뉴스 검색 결과와 같은 비정형 데이터의 경우에도 해당 검색포털에서 공개한 검색 API를 활용하면 수집 가능하다.

또한 웹 크롤러를 이용하여 인터넷에 공개된 정보를 자동으로 수집하는 방법을 ‘웹 크롤링’이라고 한다. 웹 크롤러는 특정 사이트의 웹 문서를 자동으로 수집하는 기술로, 수많은 웹 문서를 자동으로 돌아다닌다는 점에서 웹 로봇(robot)이나 봇(bot), 스파이더(spider)라고 불리기도 한다. 웹 크롤러는 주로 비정형 데이터를 수집할 때 사용되며 야후나 페이스북 등의 민간 기업이 개발한 후 오픈소스로 공개한 사례가 대부분이다. 웹 크롤링은 웹 크롤러가 수집할 웹 문서 주소(URL) 목록인 시드(Seed)를 기준으로 웹 문서를 수집하고, 수집된 웹 문서에 포함된 URL을 다음 시드로 활용해 그 다음 웹 문서를 수집하는 방식을 따른다.

한편 웹 스크래핑은 일반적으로 웹 크롤링과 유사한 의미로 사용되는데, 엄밀하게는 두 개념이 구분된다. 즉, ‘웹 스크래핑’은 웹 브라우저 화면 내에서 사용자가 저장한 필요 정보만을 추출해 내는 기술(서명구, 박규석, 2004)로, 웹 크롤링이 기본적으로 모든 데이터를 수집한다면, 웹 스크래핑은 사용자가 원하는 특정 정보만 추출한다는 점에서 차이가 있다. 웹 크롤링과 웹 스크래핑을 오픈 소스(open source) 형태로 무료 지원하는 도구로는 R 프로그램의 httr 패키지(Hadley, 2016), 프로그래밍 언어인 파이썬(Python)의 ‘beautifulsoup’(Richardson, 2016) 라이브러리와 ‘Scrapy’(http://scrapy.org) 라이브러리가 대표적이다.



## 2) 데이터 클리닝

수집한 데이터를 분석하기 용이한 형태로 정제하는 것은 모든 데이터에서 요구되는 작업이지만, 비교적 처리가 용이한 정형 데이터보다 비정형 데이터에서 이 작업은 더욱 필수적이고 또한 어렵다. 특히 자동화된 방식으로 수집된 텍스트 빅데이터의 경우 그 규모가 엄청나기 때문에 자동화된 데이터 클리닝(data cleaning) 작업이 필수적이다. 이러한 작업은 그동안 공학 분야에서 주로 활용되었던 ‘텍스트 마이닝’(text mining) 기법을 통해 실시되며, 최근 인문·사회과학 분야에서는 ‘자동화된 텍스트 분석’(automated text analysis)으로 명명되고 있다. 이 기법은 자연어 처리(Natural Language Processing: NLP)를 활용해 비정형 데이터를 정형화하는 것으로, 데이터 클리닝 단계에서는 텍스트 데이터에 대한 전처리(preprocessing), 형태소 분석(morphological analysis)을 수행한다.

우선, ‘전처리’ 단계에서는 수집된 텍스트 데이터 중 동일·유사어지만 다르게 표현된 단어들을 통일하는 ① 정규화(normalization) 작업과, ② 분석시 불필요한 단어 및 어구를 삭제하는 작업을 수행한다. 우선, ‘정규화’ 작업에서는 띄어쓰기가 다양하게 표현된 용어(예: ‘개정 교육과정’, ‘개정교육과정’)나 축약어(예: ‘수능’, ‘대수능’), 여러 가지 언어로 표기된 용어(예: ‘SW 교육’, ‘初’) 등을 동일한 용어로 변환하는 과정이 이루어진다. 또한 유사어(예: ‘고사’와 ‘시험’)를 하나의 단어로 통일하고, 의미가 다소 다를지라도 분석 목적을 고려할 때 동일한 단어로 처리해야 하는 단어 역시 통일할 필요가 있다. 또한 마침표, 쉼표, 괄호 등의 문장 부호와 각종 기호로 표현된 특수 문자(예: ‘▷’), 의존명사 및 기타 분석시 의미를 부여할 필요가 없는 불용어(stopword)들을 제거해야 한다.

‘형태소 분석’은 문장을 의미의 최소 단위인 형태소로 바꾸는 작업이다. 형태소는 그 의미와 기능에 따라 어휘 형태소(명사, 동사, 형용사, 부사 등)와 문법 형태소(조사, 어미 등)로 분류된다. 특정 텍스트로부터 주요 내용을 파악하고자 할 때에는 대부분 어휘 형태소가 활용되며, 품사

부착(part-of-speech tagging)을 통해 특정 품사를 선별, 선택할 수 있다. 예컨대, 한국어 형태소 분석기인 KOMORAN을 사용하여 “교육의 질은 교사의 질을 능가하지 못한다”라는 문장에 대해 형태소 분석을 실시하면, “교육/NNG+의/JKG 질/NNG+은/JX 교사/NNG+의/JKG 질/NNG+을/JKO 능가/NNG+하/XSV+지/EC 못하/VX+ㄴ다/EC”와 같이 결과가 도출된다. 이러한 결과를 활용하면 특정 품사의 단어로 한정하여 텍스트 분석을 실시할 수 있으며, 예컨대 명사로만 텍스트 분석을 실시할 때에는 일반명사(NNG)로 품사가 부착된 ‘교육’, ‘질’, ‘교사’, ‘질’, ‘능가’와 같은 용어들을 선택적으로 활용할 수 있다.

참고로, 국내에 오픈 소스의 형태로 공개된 한국어 형태소 분석기로는 KOMORAN 이외에도 서울대학교 IDS(Intelligent Data Systems)에서 개발한 ‘꼬꼬마 형태소 분석기’, KAIST의 SWRC(Semantic Web Research Center)에서 개발한 ‘한나눔 형태소 분석기’, 은전한닢 프로젝트에 따라 일본어 형태소 분석기 엔진에 세종 말뭉치를 학습시킨 ‘mecab-ko’ 형태소 분석기가 있다. 형태소 분석기는 분석 텍스트의 내용에 따라 그 성능이 좌우될 수 있으므로, 최종적으로 사용할 형태소 분석기를 선택할 때에는 여러 가지 형태소 분석기를 활용해 분석한 후 그 결과를 비교·대조할 필요가 있다.

### 3) 데이터 분석 및 시각화

빅데이터 분석에는 통계학과 컴퓨터 공학 분야의 지식이 고루 활용되고 있으며, 텍스트 빅데이터의 경우 언어학 분야의 이론도 중요하게 반영되고 있다. 이로 인해 연구자들은 각 학문 분야의 특성을 반영하여 서로 다른 분류 체계에 의해 빅데이터 분석 방법을 다양하게 논하고 있다. 이 절에서는 데이터 분석의 대상을 텍스트 데이터로 한정하고, 최근 자동화된 텍스트 분석에서 많이 활용되고 있는 기법들을 개관하였다.

## 가) 텍스트 마이닝

텍스트 마이닝(text mining)은 텍스트 형태로 이루어진 비정형 데이터들을 자연어 처리 방식을 이용하여 정보를 추출하거나 연계성을 파악하는 기법이다. 이 기법은 웹 문서에서 특정 주제어(keyword)와 매칭되는 단어를 찾아 수를 부여하는 인덱싱(indexing) 검색 기법에서 발전되어 왔으며, 점차 특정 주제어나 문맥(context)을 기반으로 데이터의 숨은 의미를 탐색하는 데 활용되고 있다. 텍스트 마이닝이 두드러지는 분야는 뉴스 기사 분석으로, 이를 활용하면 텍스트의 문맥에 따라 쟁점을 파악하고 텍스트 간 연계를 분석할 수 있다는 장점이 있다. 텍스트 마이닝의 대표적 기법은 다음과 같다.

### (1) 버즈 분석

‘버즈 분석’(buzz analysis)은 주로 경영학 분야에서 사용되는 용어(송태민, 송주영, 2016)로, 일종의 문서량 또는 검색량 분석이라고 할 수 있다. 버즈 분석은 관심 주제어를 포함한 문서의 발현 빈도 또는 검색 빈도 총량을 확인할 수 있기 때문에 이를 통해 해당 주제에 대한 국민의 관심도를 분석하는 데 주로 활용된다. 그리고 특정 기간 동안 해당 문서의 발현 빈도를 시계열적으로 추적하면 해당 주제에 대한 국민의 관심도 변화를 살펴보는 추이 분석(trend analysis)이 가능하고, 특정 사건 전후로 발현된 문서 빈도를 비교하면, 사건에 대한 반응도를 확인하는 이슈 반응 분석(issue response analysis)을 수행할 수 있다.

### (2) 주제어 빈도 분석

‘주제어 빈도 분석’(keyword frequency analysis)은 특정 문서 집단 내에서 자주 언급되는 주제어를 추출하고 이들이 언급되는 빈도에 따라 중

요도를 분석하는 방법(황명화, 2014)이다. 앞서 언급한 버즈 분석의 분석 단위를 텍스트 데이터의 주제어로 바꾸면 주제어 빈도 분석을 수행할 수 있다. 이때 주제어의 빈도는 특정 단어가 문서에서 얼마나 자주 등장하는지를 나타내는 ‘단어 빈도’(Term Frequency: TF)로 단순히 결정할 수 있다. 특정 단어를  $t_i$ , 총  $k$ 개의 단어로 구성된 특정 문서를  $d_j$ 라고 하고 할 때, TF를 수식으로 표현하면 다음과 같다(이성직, 김한준, 2009).

$$TF_{ij} = \frac{n_{i,j}}{\sum_k n_{k,j}} \quad (\text{단, } n \text{은 출현 횟수})$$

직관적으로는 TF 값이 큰 단어일수록 중요도가 높다고 판단할 수 있다. 그러나 TF 값이 큰 단어는 모든 문서에서 자주 등장하는, 즉 ‘문서 빈도’(Document Frequency: DF) 값이 큰 단어일 수 있다. 이처럼 모든 문서에서 자주 출현하는 상투어를 걸러내기 위해 ‘단어 빈도-역문서 빈도’(Term Frequency-Inverse Document Frequency: TF-IDF)가 도입되었는데, 이는 단순한 빈도 처리가 아닌 단어의 출현 확률을 기준으로 출현 빈도를 재가공(임화진, 2014)한 것이다. 즉, TF-IDF는 특정 문서에서의 단어 중요도를 나타내는 TF에 전체 문서와 관계된 단어의 중요도인 IDF를 곱한 값으로, 특정 문서 내에서 단어 빈도가 높을수록, 그리고 전체 문서들 중 그 단어를 포함한 문서가 적을수록 그 값이 커진다. IDF를 문서 집합에 포함되어 있는 전체 문서의 빈도,  $|d_j|_{t_j \in d_j}$ 를 특정 단어  $t_j$ 가 출현하는 문서의 빈도라고 할 때 TF-IDF를 수식으로 나타내면 다음과 같다(이성직, 김한준, 2009).

$$\begin{aligned} TF-IDF_{ij} &= TF_{ij} \times IDF_i \\ &= \frac{n_{i,j}}{\sum_k n_{k,j}} \times \log \frac{|D|}{|d_j|_{t_j \in d_j}} \end{aligned}$$

한편 ‘연관어 분석’(association keyword analysis)은 주제어 빈도 분석을 발전시킨 것으로, 관심 주제어를 포함한 대상 문서에서 함께 언급된 주제어를 추출하여 국민들이 관심 주제어와 어떠한 주제를 연결시켜 생각하는지를 분석한다. 예컨대, 하나의 뉴스 기사에서 동시 출현한 용어의 쌍을 추출하고 전체 문서집합에서 주제어의 쌍별 발생 빈도와 연결 관계를 분석하면 언론 매체 상의 주요 관심 이슈와 그 연계성의 변화를 추적할 수 있다. 또한 주제어를 특정 주제별로 분류한 후 연관어 분석을 실시하면 주요 쟁점 간의 관련성을 구조적으로 파악하기에 용이하다.

### (3) 토픽 모델링

‘토픽 모델링’(topic modeling)은 텍스트 데이터에서 사용된 주제어들의 동시 사용 패턴을 바탕으로, 해당 텍스트들을 대표하는 특정 주제나 이슈, 주제 그룹들을 자동으로 추출하는 분석 기법이다. 이 기법은 개별 문서가 다수의 주제, 즉 토픽(topic)을 다룰 수 있다고 가정한다. 즉, 수집된 문서 데이터는 이러한 토픽들의 확률적 혼합체로 간주되며, 각 토픽은 추출된 주제어들로 나타낼 수 있다는 것이다.

#### ① 토픽 모델링의 발전 과정

토픽 모델링의 시초는 잠재 의미 분석(Latent Semantic Analysis: LSA)이다. 이 기법은 대량의 텍스트 문서에서 발생하는 단어들 간의 연관 관계를 분석함으로써 잠재적 의미 구조를 도출하기 위해 고안되었으며, 이때 연관 관계는 단어들의 동시 출현성(co-occurrence)을 통해 파악한다(Deerwester et al., 1990). LSA 기법은 토픽을 추출하기 위해 행렬 인수분해 기법인 특이값 분해(Singular Value Decomposition: SVD)를 활용하여 텍스트 문서 집합을 내용의 유사도에 따라 여러 개의 소집단으로 분류한다.

이를 수식을 통해 살펴보면, LSA 기법은 SVD를 이용해 단어-문서 행렬(term-document matrix)  $A$  ( $m \times n$  직사각 행렬)를 다음과 같은 3개의 독립적 행렬로 분해한다. 이때  $U$ 와  $V$ 는 각각 행렬  $A$ 의 좌·우 특이벡터로 구성된  $m \times r$ ,  $r \times n$  직교(orthogonal) 행렬로 각각 용어 및 문서와 관련되어 있다. 그리고  $\Sigma$ 는 특이값들로 구성된  $r \times r$  대각(diagonal) 행렬로, 분해된 행렬의 계수(rank)  $r$ 에 따라 단어-문서 행렬  $A$ 의 차원을 축소해 토픽을 추출한다(Berry, Dumais, & O'Brien, 1995).

$$A = U\Sigma V^T$$

즉, LSA 기법을 통해 단어-문서 행렬의 차원을 축소하면 원래의 차원에서는 관찰되지 않았던 단어들이 같은 차원에서 관찰되며, 이러한 단어들은 서로 높은 상관을 보인다(김성근, 조혁준, 강주영, 2016). 이때 축소된 각 차원은 하나의 토픽을 의미하며, 이는 최소 2개 이상의 동시출현 빈도가 높은 주제어 집합이 추출된 것이다(이수연, 김현정, 정수연, 2016). 이 기법은 분석시 특별한 제약 조건이 없고, 직교 행렬을 기반으로 하기 때문에 계산이 편리하다는 장점이 있다(Steyvers & Griffiths, 2007). 그러나 계산량이 많고, 분석 결과를 현실 상황에 적용하여 의미를 도출하기 어렵다는 한계가 있다(Blei, Ng, & Jordan, 2003).

이에 따라 Hofmann(1999)은 LSA를 확률적 모델링 기법으로 확장한 확률론적 LSA(Probabilistic LSA: pLSA) 기법을 제안하였다. 이 기법은 데이터가 주제(thematic structure)와 같은 은닉 변수(hidden variable)를 포함하는 생성적 확률 과정(generative probabilistic process)을 통해 발생된 관찰 값이라고 가정한다. 즉, pLSA 기법은 사후 추론(posterior inference)을 이용해 은닉 구조를 추론하며, 추정된 모형에 새로운 데이터를 학습시킨다. 이 기법은 쉽게 확장될 수 있고 다른 복잡한 모형과 함께 사용할 수 있다는 장점을 가지나, 과적합(overfitting)되는 경향이 있어 학습(training)시 유의해야 한다.

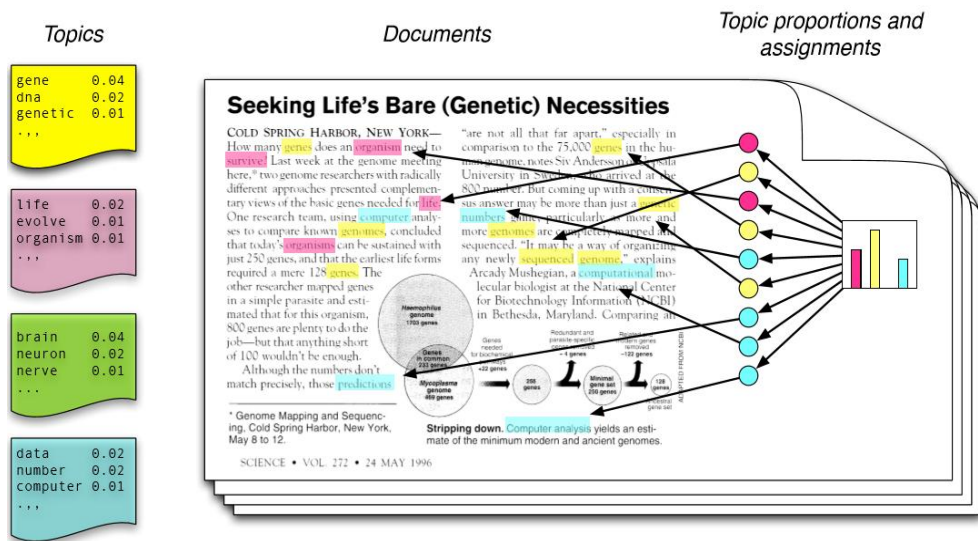
잠재 디리클레 할당(Latent Dirichlet Allocation: LDA)(Blei, Ng, & Jordan, 2003)은 이러한 한계를 극복하기 위해 등장한 기법이다. LDA 기법은 텍스트 문서들을 저차원으로 축소한다는 점에서 LSA 기법과 유사하고 단어 및 문서 벡터가 확률 분포를 갖고 있다고 가정한다는 점에서 pLSA와 유사하다. 그러나 LDA 기법은 LSA 기법 등 다른 토픽 모델링에 비해 결과 해석이 용이하고(Blei, 2012), 상위 모수(hyper parameter) 값을 임의의 변수로 평활화(smoothing)함으로써 과적합 문제를 해결하기 때문에 방대한 비정형 데이터의 차원을 축소해 여러 토픽들을 도출하는 데 장점이 있다(Griffiths & Steyvers, 2004).

참고로, 토픽 모델링은 감독/지도 분류(supervised classification)와 무감독/비지도 분류(unsupervised classification)로 구분할 수 있다. ‘감독 분류’의 경우 연구자가 미리 문서의 토픽과 관련 단어 및 문장 구조들을 선정한 다음 그것에 따라 문서들을 분류하는 방식이라면, ‘무감독 분류’는 특정 문서들에 포함된 단어들을 그 맥락과 의미를 바탕으로 유사도에 따라 몇 개의 그룹으로 분류한 후 포괄적 주제어를 부여하는 방식을 택한다. 토픽 모델링에서 가장 일반적으로 많이 활용되고 있는 LDA 기법은 무감독 분류에 속하며, 이하에서는 LDA 기법에 대해 고찰하였다.

## ② LDA 기법의 개요

LDA 기법은 데이터로부터 잠재 변수를 유추하고 복잡한 데이터의 차원을 축소해 데이터를 효과적으로 이해하고자 한다는 점에서 탐색적 요인 분석(exploratory factor analysis)과 유사하다(Steyvers & Griffiths, 2007). 그리고 기존의 군집분석 기법과 비교해 볼 때, 다대다 대응이 가능하다는 차이점을 갖고 있다. 즉, 1개의 문서당 1개의 주제만 대응되는 것이 아니라, 여러 가지 주제들을 포함할 수 있다는 측면에서 실제 텍스트의 본질을 잘 반영한다는 장점을 갖고 있다(김정수, 이석준, 2016).

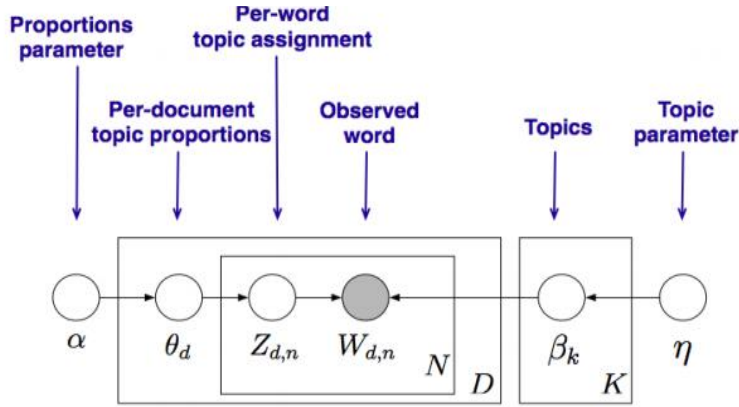
LDA는 문서들이 알려지지 않은 여러 개의 토픽들로 구성되어 있고, 각 토픽은 다양한 단어들로 표현된다고 전제한다([그림 II-3]의 맨 왼쪽 부분 참고). 또한 각 문서마다 토픽들의 구성 비율이 다르며, 이 비율을 결정하는 확률분포인 디리클레 분포가 존재한다고 가정한다. 이러한 기본 전제 하에 LDA 기법에서는 각 문서가 다음의 과정을 통해 생성된다고 가정한다. 우선, ([그림 II-3]의 맨 오른쪽 히스토그램에서 보이는 것처럼) 여러 토픽들에 대한 하나의 분포를 무작위로(randomly) 선택한다. 그리고 해당 분포로부터 (색칠된 원들로 표시된) 할당될 수 있는 토픽을 선택한 후, 그에 대응되는 어휘 분포로부터 하나의 단어를 선택한다.



[그림 II-3] LDA 기법의 기본 가정(Blei, 2012)

즉, LDA는 문서 내 토픽 구성 비율을 결정짓는 베이시언(bayesian) 확률모형의 사후 분포(posterior distribution)를 추정하며, 그 과정에서 전체 텍스트 데이터를 주도하는 토픽들과 그 토픽들을 구성하는 주요 단어 집합들을 찾아낸다(김용대, 조광현, 2013). Blei(2003, 2012)는 이러한 과정을 [그림 II-4]와 같이 그래프 모형(graphical model)으로 나타냈다.



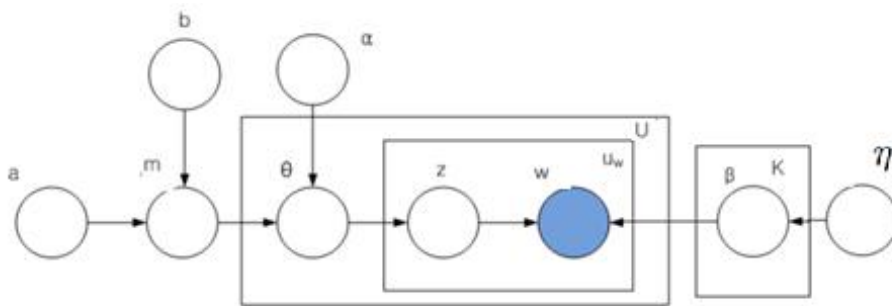


[그림 II-4] LDA에 대한 그래프 모형(Blei, 2003)

그림에서 각 원은 확률 변수이며, 화살표의 방향은 확률 모형 내에서 상위 변수가 하위 변수를 결정함을 의미한다. 회색 원으로 표시된  $W_{d,n}$ 은 관찰된 데이터로  $d$ 는 문서,  $n$ 은 단어를 지칭한다. 그리고 LDA에서는 문서들이  $K$ 개의 토픽들의 집합으로부터 생성된다고 가정한다. 만약 문서에  $K$ 개의 토픽과  $V$ 개의 중복되지 않은 단어가 있다고 가정하면, 토픽의 사전분포 모수  $\alpha$ 는 양의 값을 갖는  $K$ -벡터이고 토픽의 단어 분포 확률을 나타내는 모수인  $\beta_k$  역시  $K$ -벡터이다.  $\beta_k$  값은 문서 전체에 대한 디리클레 분포를 통해 할당되는데, 상수  $\eta$ 는  $\beta_k$ 의 사전분포를 결정하는 모수다. 또한  $\theta_d$ 는 각 문서  $d$ 의 토픽 구성 비율을 표현하는  $d$ -벡터 모수이고,  $Z_{d,n}$ 는 문서  $d$ 에 등장하는 각 단어  $n$ 에 토픽을 할당하는 모수로 둘 다 잠재 변수이다. LDA에서는  $Z_{d,n}|\theta_d$ 가 다항 분포(multinomial)를 따른다는 정보를 갖고 있기 때문에, 공액 사전분포(conjugate prior)로 디리클레 분포를 사용하게 되면 사후확률의 분포는 다항 분포가 된다(김규하, 박철용, 2015). 즉, 즉, 문서  $d$ 에 있는 각 단어  $n$ 을 생성하기 위해 먼저  $\theta_d$ 에서 토픽 인덱스(index)  $Z_{d,n}$ 을  $Z_{d,n} \sim \text{Mult}(\theta_d)$ 와 같이 생성한다.

상위 모수(hyper parameter)라고 불리우는  $\alpha$ 와  $\eta$ 는 데이터에 의해 추정되는 것이 아니라 연구자의 사전 지식에 기반해 설정되기 때문에, 토픽 모델링의 분류 성능을 높이기 위해서는 해당 모수를 신중하게 설정해

야 한다. 이와 관련하여 Wallach, Mimno & McCallum(2009)은 깃스 표집(Gibbs sampling) 방법을 이용해 ‘비대칭’(asymmetric) 상위 모수를 LDA 모형에 최적화시키는 방법을 제안하였다. 이때 비대칭이란 토픽마다 해당 모수의 값을 다르게 부여하는 것을 허용한다는 의미이다. 연구 결과, LDA에서는  $\alpha$ 를 비대칭 모수,  $\eta$ 는 대칭 모수로 설정하였을 때 복잡도(perplexity)를 기준으로 토픽의 분류 성능이 가장 좋다는 것이 확인되었다. [그림 III-5]는 기본적인 LDA 모형에서 상위 모수  $\alpha$ 와  $\eta$ 를 각각 비대칭, 대칭으로 설정하였을 때 [그림 II-4]에서 제시한 Blei(2003)의 그래프 모형이 어떻게 변화하는지를 나타낸 것이다.



[그림 II-5] 상위 모수 분포에 따른 LDA의 그래프 모형  
(김은희, 표신지, 김문철, 2012)

### ③ LDA 기법에서의 사후 분포 추정 방법

LDA 모형은 베이지안 확률 모형이므로 사후 분포를 추정하게 되는데, 사후 분포를 직접 다루는 것은 어려운 작업이므로 주어진 데이터로부터 반복적으로 사후 확률의 근사치를 추정하게 된다. 이때 추정 방법으로는 ML(maximum likelihood), MAP(maximum approximate posterior), VEM(variational expectation-maximization), 깃스 표집(Gibbs sampling) 등이 있는데, LDA에서는 ‘깃스 표집’과 ‘VEM’ 방법이 주로 사용된다.

우선, ‘깁스 표집’은 두 개 이상의 변수들로부터 연쇄적으로 표집을 시행함으로써 사후 분포를 추정하는 방법으로, 주어진 모수에 의해 데이터를 분류하고 분류된 데이터 속에서 모수를 또 다시 추정하는 경험적 접근을 취한다. 깁스 표집은 불편(unbiased) 추정치를 제공하는 등 정확도가 높고 추론 과정을 이해하기 쉬운 편이다(Asuncion et al., 2009). 그러나 Markov chain monte carlo(MCMC) 기법에 근거하기 때문에 계산이 복잡하고 변수 간 의존성이 커서 대규모의 텍스트 데이터를 다룰 때에는 비효율적일 수 있다. 이러한 계산의 복잡성을 개선하기 위해 Griffiths와 Steyvers(2004)는 붕괴된 깁스 표집(Collapsed Gibbs Sampling) 방법을 제안하였다. 이 기법은 수렴될 때까지 반복적으로 추정해야 하는 문서-토픽 확률과 토픽-단어 확률 모수를 적분해 제거함으로써 변수 간 의존성을 줄이고 속도를 개선하는 장점을 갖는다(김용대, 정구환, 2016).

한편, Blei, Ng, & Jordan(2003)이 제안한 VEM은 모수화된(parameterized) 특정 분포를 미리 상정한 후 사후 분포에 가장 가까운 멤버를 찾는 결정론적 접근(deterministic)을 취한다. 즉, 상위 모수를 이미 안다고 전제하고 LDA 모형의 로그-우도(log-likelihood)를 최대화하는  $\theta_d$ 와  $\beta_k$ 를 계산한 다음, 이 값을 이용해 다시 상위 모수 값을 추정한다. 이 추정 방법은 속도 처리가 빠르기 때문에 최근 빅데이터 분석 환경에서 많이 활용되기 시작하였으나 최적화(optimization) 문제가 발생하므로 정확도는 다소 떨어지는 편이다.

#### ④ LDA 기법에서의 모형 평가 방법

LDA 모형은 모수 통계 모형(parametric model)이기 때문에 사전 정보로 토픽의 개수를 지정해 주어야 한다. 그러나 대규모 분석 데이터의 경우 연구자 본인만의 역량으로 토픽의 개수에 대한 정보를 파악하기는 쉽지 않다. 이에 따라 최근에는 토픽 모델링의 기반이 되는 확률모형을 최적화하여 적절한 수의 토픽을 비모수적(non-parametric)으로 추정하는

계층적 디리클레 과정(Hierarchical Dirichlet Process: HDP)(Teh et al., 2012; Wang, Paisley, & Blei, 2010)이 도입되고 있으나 분석 결과를 일반화하는 데 다소 어려움이 있어 사용시 주의해야 한다. LDA 모형에서 토픽의 개수를 추정할 때에는 토픽의 개수가 상이한 토픽 모형들을 비교·평가하게 되는데, 모형의 유용성을 파악하기 위해 사용되는 지표로는 ‘복잡도’(perplexity)가 대표적이다.

우선, ‘복잡도’는 각 단어 우도(likelihood)의 기하 평균에 대한 역수와 동일한 값으로, 그 값이 작을수록 해당 모형의 일반화 가능성이 높은 것을 의미한다(Blei et al., 2003). LDA에서 복잡도를 산출하는 가장 일반적인 방법은 분석 데이터를 임의로 훈련용 데이터(training data)와 시험용 데이터(test data)로 나눈 후, 훈련용 데이터를 통해 도출한 LDA 모형을 시험용 데이터에 적용하였을 때의 분류 성능을 수치화하는 것이다.

단어들의 집합을  $W$ 라고 할 때 복잡도는 다음의 수식으로 표현할 수 있으며,  $H(W)$ 는 언어학에서 언어 모형을 평가할 때 주로 사용되는 교차 엔트로피(cross entropy) 값인  $-\frac{1}{N}\log_2 P(w_1w_2...w_N)$ 이다(Jurafsky & Martin, 2008).

$$\text{Perplexity}(W) = 2^{H(W)} = \sqrt[N]{\prod_{i=1}^N \frac{1}{P(w_i|w_1...w_{i-1})}}$$

이때 복잡도가 클수록 사전에 훈련된 토픽 모형이 시험용 데이터에서 토픽별 단어를 잘못 추출(misrepresentation)한다는 것을 뜻하므로, 복잡도는 작을수록 바람직하다고 할 수 있다(Chen, & Wang, 2007).

그런데 최근 연구에 따르면, 텍스트 데이터에 대해 사람이 판단한 토픽과 복잡도 계산을 통해 결정된 토픽 간의 상관(correlation)이 통계적으로 유의하지 않거나 오히려 부적(negative)인 경우가 종종 나타난다(Chang et al., 2009; Jacobi, van Atteveldt, & Welbers, 2016; Newman,

et al., 2010; Röder, Both, & Hinneburg, 2015). 이에 따라 사람이 텍스트를 해석하는 방식을 반영한 새로운 토픽 모델링 평가 방법이 개발되어야 한다는 의견이 제기되고 있으며, 그 결과 개발된 개념이 바로 ‘토픽 응집성’(topic coherence)이다.

토픽 응집성은 토픽의 해석 가능성을 추출된 토픽을 구성하고 있는 단어들로부터 찾는 기법으로, LDA를 통해 추정된 토픽-단어 확률분포( $\phi$ )로부터 토픽 내 상위 N개의 단어를 추출하는 방식이 이에 해당한다 (O’Callaghan et al., 2015). 대표적인 토픽 응집성 지수로는 Newman et al. (2010)의 Pointwise Mutual Information(PMI)이 있는데, 이는 단어들의 동시 출현 빈도(co-occurrence frequency)를 외부 데이터가 아닌 LDA를 실시한 데이터에서 계산함으로써 토픽 응집성을 도출한다. 즉, 토픽 모형으로부터 추출된 특정 토픽의 상위 n개 단어 집합을  $T = \{w_1, w_2, \dots, w_n\}$  라고 할 때, PMI는 단어 집합 T에서 단어들 간 짝 유사도(pairwise similarity)의 평균이 높을수록 토픽의 응집성이 높다고 가정한다. 대칭적(symmetric) 단어 유사성 측도  $Sim(w_i, w_j)$ 가 주어졌을 때 PMI를 수식으로 표현하면 다음과 같다.

$$Coherence_{sim}(T) = \frac{\sum_{\substack{1 \leq i \leq n-1 \\ i+1 \leq j \leq n}} Sim(w_i, w_j)}{\binom{n}{2}} \quad \text{where } w_i, w_j \in T.$$

Röder, Both, & Hinneburg(2015)은 수많은 토픽 응집성 지수 중 ‘C<sub>v</sub>’의 성능이 가장 우수하다고 평가하였다. 이 지수는 데이터 마이닝의 교차 타당화(cross-validation) 방법 중 하나인 슬라이딩 윈도우(sliding window)<sup>9)</sup>를 활용하는데, C<sub>v</sub>의 경우 하나의 윈도우가 110개의 단어로 구

9) 텍스트의 특정 패턴을 검색하기 위해 이동하는 영역(rectangle)의 단위를 슬라이드(slide)라고 부르며, 슬라이드의 크기를 변경해 가면서 패턴을 파악하는 방법을 슬라이딩 윈도우(sliding window)라고 부른다.

성되어 있으며, 여러 윈도우들을 미끄러지듯 이동하면서 연속적으로 훈련과 시험을 반복한다. 즉, 하나의 윈도우가 열릴 때마다 그 곳에 포함된 단어들의 동시 출현 빈도를 계산해 고빈도 단어들의 NPMI(Normalized PMI)를 산출하고, 이 과정을 윈도우가 종료될 때까지 반복한다. 그리고 단계별 고빈도 단어의 벡터 집합과 분석 데이터 내 모든 단어 벡터 간의 코사인(cosine) 유사도를 계산한 후 이를 산술 평균하면  $C_v$ 가 산출된다.  $C_v$  지수가 높을수록 토픽 각각의 응집성이 높다는 것을 의미하므로 토픽 모델링에서는 이 값이 가장 큰 토픽 모형의 토픽 개수를 채택한다.

그러나 이러한 지표에만 근거하여 토픽 모형을 평가하고 토픽의 개수를 추정하는 것은 다소 편의적인 접근이라는 지적도 있다. Ding & Chen(2014)에 따르면, 토픽 모델링을 활용한 선행연구에서는 대체로 복잡도와 같은 지수를 활용하지만 ① 여러 가지 경우의 수를 고려하여 토픽 분석을 실시한 후 용어 분류의 정확성이 높은 결과를 선택하거나 (Griffiths & Steyvers, 2004), ② 비교적 많은 수의 토픽을 설정하여 LDA를 실시한 후 유사한 토픽들을 결합하여 최종 결과를 생성하는 경우(Song, 2010; Yu, 2014)도 많았다.

## ⑤ 토픽 모델링의 분석 및 시각화 도구

토픽 모델링을 실시할 수 있는 오픈 소스 형태의 분석 도구로는 R 프로그램의 ‘topicmodels’(Bettityna & Kurt, 2015)와 ‘lda’(Chang, 2015) 패키지, Python의 ‘gensim’(Rehurek & Sojka, 2010) 라이브러리, Java 기반의 MALLET(McCallum, 2002) 등이 있다. 이 분석 도구들은 토픽 모델링에 활용되는 사후 분포 추정 방법이나 상위 모수의 기본 값이 서로 다르기 때문에 연구의 목적에 부합하는 분석 도구를 이용해 토픽 모델링을 수행해야 한다. <표 II-11>은 토픽 모델링을 실시할 수 있는 여러 가지 분석 도구들의 특징을 비교하여 정리한 것이다.

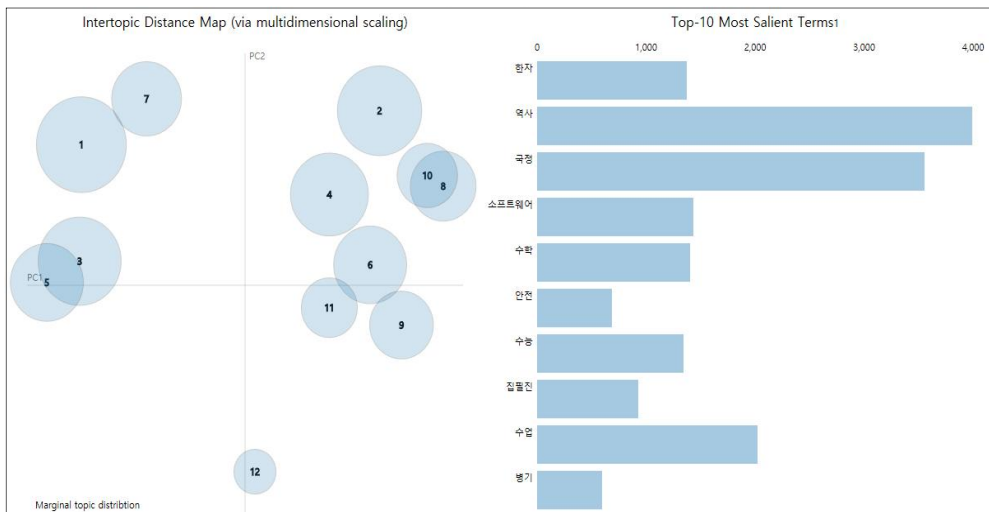
<표 II-11> 토픽 모델링 분석 도구 비교

구분		R		Python	Java
		topicmodels	lda	gensim	Mallet
사후 확률의 근사치 추정방법		· collapsed Gibbs sampling · VEM	collapsed Gibbs sampling	· (on-line) VEM · collapsed Gibbs sampling	collapsed Gibbs sampling
상위 모수 (기본값)	$\alpha$	50/k	—	1/k, 대칭적	50/k, 비대칭적
	$\eta$	0.1	—	1/k, 대칭적	0.01
기타 분석 가능 모형		CTM(correlated topic modeling)		HDP	hLDA (Hierarchical LDA)

이 중에서 R의 topicmodels와 Python의 gensim이 토픽 모델링의 대표적 분석 도구로 활용되고 있다. 우선, topicmodels는 VEM 알고리즘을 활용한 LDA와 CTM 기법, 붕괴된 깁스 표집을 이용한 LDA 기법을 지원하며, 문서-단어 행렬(document-term matrix)를 투입하여 분석을 실시한다. 반면, gensim의 경우에는 일반 텍스트 문서들(plain text documents)의 말뭉치(corpus)를 투입시키며, LSI, LDA, HDP 등의 다양한 알고리즘이 포함되어 확장 가능성이 높다. 그리고 일단 토픽이 발견되면, 일반 텍스트 문서들은 토픽의 유사성에 기반하여 분류될 수 있다(Liu et al., 2016). 참고로, 기존의 gensim에서는 VEM 방법만을 지원하였으나, 최근에는 Java 기반의 MALLET을 Python에서 활용할 수 있도록 래핑(wrapping)한 API를 추가하여 좀 더 다양한 분석이 가능해졌다.

끝으로, 토픽 모델링 결과를 시각화하기 위한 도구로는 HTML 형식으로 시각화 결과를 제공하여 토픽들 간의 관계를 동적(dynamic)으로 확인할 수 있는 ‘LDAvis’(Sievert & Shirley, 2014) 라이브러리가 있다. 그 중에서도 pyLDAvis는 LDAvis를 Python에서 사용할 수 있도록 만든 것

으로, 사용자가 LDA 분석 결과 산출된 토픽-단어 분포 행렬, 문서-토픽 분포 확률, 문서당 길이, 분석시 사용된 단어와 빈도수(term frequency), 시각화하고 싶은 토픽의 개수와 토픽별 상위 단어수를 입력하면 두 가지 그래프를 산출한다. [그림 II-6]은 토픽의 개수를 12개, 상위 단어수는 10개로 설정하였을 때, pyLDAvis를 이용하여 특정 LDA 모형을 시각화한 결과의 첫 화면이다.



[그림 II-6] pyLDAvis를 활용한 LDA 모형의 시각화 결과 예시(1)

우선, 오른쪽 그림은 전체 분석 데이터에서 가장 중요하다고 판단된 상위 10개 단어의 빈도를 막대그래프로 나타낸 것이다. 이때 각 단어의 중요도는 단어의 돌출도(saliency)에 의해 결정되는데, 이 값은 각 단어

의 TF에 각 단어의 특수도(distinctiveness)  $\sum_T P(T|w) \log \frac{P(T|w)}{P(T)}$  를 곱한 값으로 이때  $P(T|w)$ 는 토픽 T에 단어 w가 관찰될 우도(likelihood),  $P(T)$ 는 토픽 T에 무선적으로 선택된 단어 w'가 관찰될 우도다(Chuang, Manning, & Heer, 2012). 이 데이터에서는 ‘역사’의 단어 빈도수가 가장 높지만 중요도는 ‘한자’가 가장 높다.

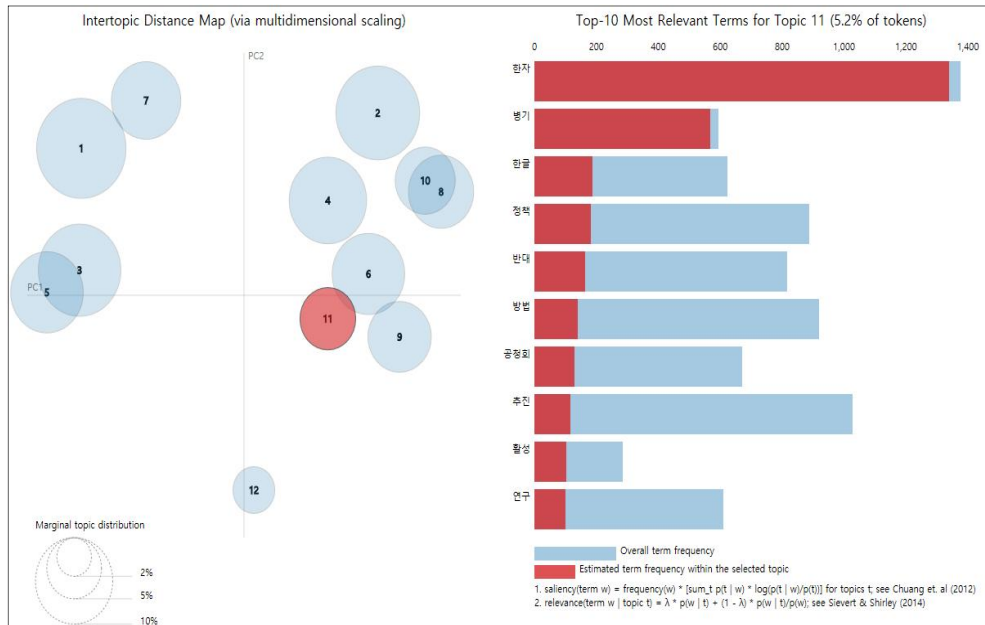


왼쪽의 그래프는 토픽 간 거리 지도(intertopic distance map)로, 토픽  $k$ 개의 차원을 2차원( $PC_1$ ,  $PC_2$ )으로 축약한 것이다. 이 그래프는 분석 결과 도출된 토픽을 원의 형태로 보여주는데, 원의 면적은 분석 데이터 내에서 토픽의 상대적 중요도(relative prevalence)를 의미한다(Sievert & Shirley, 2014). 또한 각 토픽의 위치는 다차원 척도법(multidimensional scaling)<sup>10)</sup>에 따라 유사성 거리에 의해 표현된 것으로, 이를 통해 토픽 간의 관계나 유사성 혹은 토픽 간 차이(inter-topic differences)를 파악하는 데 유용하다. 이때 유사성 거리로는 Jensen-Shannon divergence<sup>11)</sup>가 기본적으로 사용되며, 토픽 간 거리 집합(set)을 척도화(scaling)할 때에는 주성분 분석(principle components analysis)이 기본적으로 적용되나 다른 알고리즘을 사용할 수도 있다(Sievert & Shirley, 2014). [그림 II-6]에서는 토픽 3과 토픽 5, 토픽 8과 10이 유사한 것을 시각적으로 확인할 수 있다.

이후 [그림 II-6]의 왼쪽 그래프에서 원하는 토픽을 선택하여 컴퓨터 마우스의 커서를 위치시키면 [그림 II-7]처럼 선택한 토픽이 붉은 색(검정색)으로 변하고 오른쪽 그래프는 해당 토픽을 설명하는 단어로 바뀐다. [그림 II-7]에서는 11번 토픽이 선택되었으며, 오른쪽 그림의 붉은 색 막대는 이 토픽에서 추정된 TF다. 붉은 색과 푸른 색(회색) 막대그래프를 비교하면 각 단어가 해당 토픽과 얼마나 잘 관련되어 있는지를 파악할 수 있다.

10) 다차원 척도법은 데이터로부터 도출된 잠재적 차원(latent dimension)으로 이루어진 공간에 주어진 데이터를 시각적으로 표현하기 위한 기법이다(박광배, 2000).

11) Jensen-Shannon divergence는 분포 두 쌍의 유사성을 측정하기 위한 척도로, 두 분포가 동일할 때(identical distributions)에만 0이고 다를 경우 무한대로 접근한다. 이 값은 각 분포의 쿨백 라이블러 다이버전스(Kullback-Leibler divergence: KL divergence)의 평균 값을 도출함으로써 계산할 수 있다(Hall, Jurafsky, & Mannung, 2008). 참고로, KL divergence는 최근린(nearest neighbor) 정보, 즉, 가장 가까운 데이터들까지의 거리를 직접 이용하는 연속적 정보를 이용하여 도출한다(노영균, 김기웅, 2014).



[그림 II-7] pyLDAvis를 활용한 LDA 모형의 시각화 결과 예시(2)

## 나) 오피니언 마이닝

오피니언 마이닝(opinion mining)은 텍스트 마이닝에서 발전된 분석 기법으로, 주로 블로그/카페나 SNS의 여론 향방을 추적하는 데 사용된다. 텍스트 마이닝이 특정 단어와 문맥의 연관성을 분석하는 데 비해, 오피니언 마이닝은 문맥과 연계된 감정분석(sentiment analysis)을 활용하여 특정 텍스트의 어조와 감정을 파악해 내는 특징이 있기 때문이다. 학자들에 따라서는 감정분석과 오피니언 마이닝을 구분하기도 하는데, Tsytarau & Palpanas(2012)에 따르면 오피니언 마이닝은 개체에 대한 사람들의 의견을 추출하여 분석하나, 감정 분석은 텍스트에 표현된 감정을 규정한 후 그것을 분석한다. 그러나 중요한 것은 오피니언 마이닝과 감정분석 모두 의견을 찾고, 그들이 표현한 감정들을 규정한 후 그것들의 극성을 분류하는 것을 목표로 한다는 것이다(Medhat, et al., 2014).

오피니언 마이닝은 분석 대상인 주제어와 함께 나오는 긍정, 부정 등 감정적 언어의 빈도수를 추적하여 여론의 향방을 추적한다. 이때 대상에 대한 선호도는 긍정, 부정, 중립(neutral) 또는 다양한 범주의 감정 분류(예: 기쁨, 감동, 신뢰, 감사, 슬픔, 두려움, 싫어함, 실망) 및 일정 범위의 감정 점수(예: 10점 척도)를 통해 판별된다. 이에 따라 감정 분석은 일반적으로 세 단계의 분석 절차를 거친다(신수정, 2014). 첫 번째는 데이터 수집 단계이고, 두 번째는 수집된 정보 중 사용자의 주관이 드러난 부분만 남기고 나머지는 걸러주는 ‘주관성 탐지’(subjectivity detection) 단계, 세 번째는 추출한 감정 데이터를 긍정, 부정으로 분류하는 ‘극성 탐지’(polarity detection) 단계다. 컴퓨터는 텍스트 안에 있는 긍정, 부정적 단어들을 탐지하여 이를 정량화한 후 통계적 기법을 적용한다. 예컨대, 문서에서 각 단어가 나타나는 빈도나 긍/부정과 같은 속성에 따라 점수나 가중치를 부여한 후, 각 단어가 나타내는 점수의 총합이나 평균을 구해 전체 텍스트가 과연 긍정적인지 혹은 부정적인지를 알아내는 것이다.

감정분석은 문서(document) 단위의 극성 분석, 문장(sentence) 단위의 극성 분석, 속성(aspect) 단위의 극성 분석으로 분류할 수 있다. 우선 ‘문서’ 단위의 극성 분석은 전체 문서를 기본 단위로 간주하고 해당 문서를 긍정 또는 부정적 문서로 분류하는 것이다. 그리고 문장이라는 것은 결국 짧은 문서일 뿐이므로 ‘문장’ 단위의 극성 분석은 사실 문서 단위의 분류와 근본적으로 차이가 없다. 이 두 분석의 첫 번째 단계는 그 문장이 주관적인지, 객관적인지를 규명하는 것이다. 그러나 단지 텍스트를 문서나 문장 수준으로 나누어 기계적으로 긍정어와 부정어의 빈도를 확인하는 것만으로는 텍스트 전체의 의미를 파악하는 데 불충분할 수 있다. 때문에 ‘속성’ 단위의 극성 분석이 필요하며, 이를 위해서는 평가 대상이 어떠한 속성으로 구성되어 있는지 확인하고, 각 속성에 따른 감정어를 찾은 후 그 감정어가 긍정과 부정 중 어디에 해당하는지를 분류한다. 최근에는 여러 문장 사이의 관계와 문맥을 고려해 문장들의 의미관계를 분석하는 담화 분석(discourse analysis)의 필요성도 제기되고 있다.

#### 다) 사회 네트워크 분석

사회 네트워크 분석(social network analysis)은 SNS 상에서 네트워크를 형성하고 있는 사용자 또는 문서 간의 상관관계를 분석하여 특정 이슈가 어떻게 발현되어 전파되는지를 보여주는 기법이다. 이 기법은 개인을 노드(node), 개인의 사회적 관계를 링크(link)로 간주하여 사회 네트워크를 구하고, 사회 네트워크의 연결 구조와 강도를 분석하여 어떤 메시지가 어떤 경로를 통해 전파되는지, 네트워크상에서 입소문의 중심이나 허브(Hub) 역할을 하는 영향 요소(influencer)는 누구인지 등을 찾는 데 주로 활용된다. 때문에 네트워크 구조가 비교적 뚜렷한 SNS를 주로 분석하며, 트위터의 경우 리트윗(Retweet: RT) 관계를 통해 심층 분석을 수행할 수 있다. 네트워크 분석의 대표적 통계 모형으로는 그래프 모형(graphical model)이 있는데, 이는 변수(노드)들 간의 조건부 독립성을 그대로 표현하는 방법으로 공분산 행렬의 역행렬을 추정하는 문제로 바꿀 수 있다. 고차원 데이터에서 네트워크를 구성할 때에는 공분산 행렬의 역행렬을 벌점화(penalization) 방법을 이용해 추정할 수도 있다.

#### 라) 군집 분석

군집분석(cluster analysis)은 유사한 관측치들을 동일한 집단으로 묶는 통계적 방법이다. 즉, 변화가 많은 대상 집단을 일정한 군집으로 나누어 특성을 분석하고, 각 데이터 간의 거리를 분석하여 비슷한 유형에 대한 성향을 파악하거나 타 집단과의 차이점을 관측하기 위한 목적으로 활용되는 기법이다. 예컨대, 교육이라는 유사한 관심사를 가진 그룹의 특성을 파악하기 위해 군집 분석이 사용될 수 있다. 고전적인 군집분석 방법으로는 거리와 같은 유사성을 기반으로 관측치를 묶는 k-평균 군집분석, 계층적 군집분석 등이 있다. 최근에는 잠재 구조(latent structure)를 이용한 군집분석, 군집분석에서의 변수 선택 등이 주로 연구되고 있다.

## 마) 분류 분석

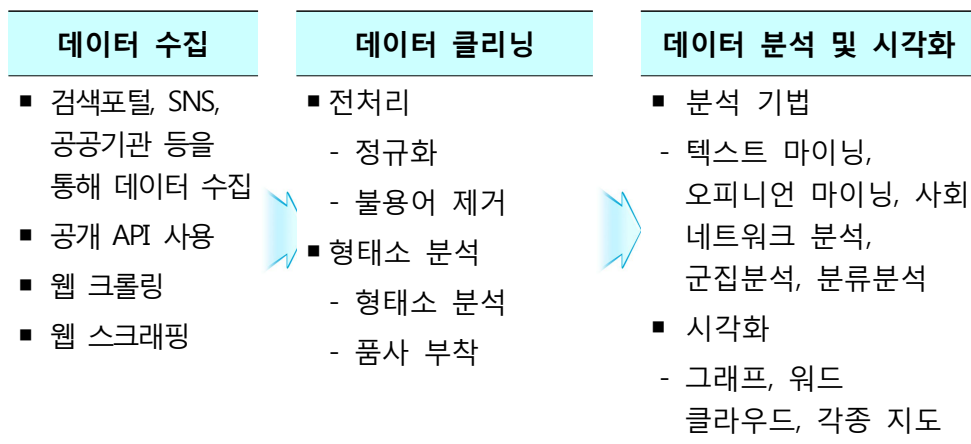
분류 분석은 설명변수에 기반하여 반응변수를 구분하는 통계 기법이다. 대표적 분류방법으로는 로지스틱 회귀분석(logistic regression), 선형 판별 분석(linear discriminant analysis), 서포트 벡터 머신(support vector machine: SVM), 나이브 베이즈 분류기(naive Bayes classifier) 등이 있다. 고차원 데이터인 경우에는 많은 경우 기존의 분류 방법을 적용하기 불가능하며, 변수 선택 또는 정규화 작업이 필요하다. ‘선형 판별 분석’이나 ‘나이브 베이즈 분류기’의 경우 독립성 가정 하에 추정된 공분산행렬을 이용하는 방법이 제안되었으며, 나이브 베이즈 분류기는 앞서 언급한 오피니언 마이닝에서도 자주 활용된다. 그리고 ‘로지스틱 회귀분석’의 경우 LASSO(Least Absolute Shrinkage and Selection Operator) 벌점함수를 이용해 벌점화된 로그 우도비(log-likelihood)를 최대화시키는 방법을 사용한다. ‘SVM’은 커널(kernel) 기법을 이용해 차원이 매우 큰 경우에도 쉽게 분류모형을 구축할 수 있다는 장점이 있다.

한편, 빅데이터 분석에서는 분석 기법뿐 아니라 그에 대한 표현 기술 또는 시각화(Visualization) 방법도 매우 중요한 요소이다. 데이터를 어떻게 표현하느냐에 따라 새로운 트렌드나 패턴을 찾아낼 수도, 찾아내지 못할 수도 있기 때문이다. 또한 분석 결과에 대한 적절한 시각화는 타인과 의사소통할 때 효과적이다. 따라서 데이터 시각화는 정보를 분명하고 효과적으로 전달하는 것을 목적으로 한다.

데이터 시각화에는 간단한 그래프에서부터 3차원 표현까지 가능하며, 분석자의 창의성에 따라 다양하게 제시될 수 있다. 또한 시계열 데이터, 통계적 분포, 계층 구조, 네트워크, 지도, 각종 애니메이션 등이 시각화에 주로 활용된다. 최근 손쉽게 많이 활용되는 시각화 방법으로는 워드 클라우드(word cloud) 분석이 있다. 워드 클라우드는 문서 내에서 단어들이 언급된 수에 비례하여 단어의 크기로 표현한 것으로, 단어의 언급 정

도를 빠르게 파악함으로써 어떤 단어가 중요한 의미를 갖는지를 탐색할 수 있다. 앞서 언급한 pyLDAvis 역시 토픽 모델링의 결과를 시각화하기 위한 도구로, 고차원의 데이터를 저차원의 평면을 통해 나타냄으로써 토픽에 대한 해석을 용이하게 하는 역할을 수행한다.

지금까지 기술한 빅데이터의 분석 절차와 방법을 텍스트 빅데이터를 중심으로 도식화하면 [그림 II-8]과 같다.



[그림 II-8] 텍스트 빅데이터의 분석 절차 및 방법

## 다. 빅데이터 분석 기법을 활용한 정책평가

### 1) 빅데이터를 활용한 정책평가 연구의 필요성

최근에는 소셜 미디어(social media)의 진화로 정책 의제 설정의 패러다임이 대중매체나 정치적 영향력을 가진 이들이 주도하는 것에서 일반 시민과 시민집단들로 전이(송영조, 2012)됨에 따라, 빅데이터가 정책 평가의 새로운 증거로 주목 받고 있다. 이에 따라 세계 각국에서는 각종 데이터 마이닝(data mining)과 데이터 시각화(data visualization) 등 빅데이터 분석 방법에 대한 연구가 활발히 이뤄지고 있다(최인봉, 박도영, 이은경, 2014; Manyika et al., 2011). 우리나라에서도 다양한 영역에서 빅데이터의 활용을 촉구하는 ‘정부3.0 추진 계획’(관계부처 합동, 2014)이 발표되었다.

‘증거기반 실천’(evidence-based practice)은 각종 분야에서 정책을 평가할 때 적극 활용되고 있다. 이는 정책을 개발하고 시행할 때 연구를 통해 최상의 증거를 제시함으로써 이해당사자들이 증거에 기반을 두고 의사결정할 수 있도록 돕는 것(윤광석, 정혜경, 김양분, 2011; Davies, 1999)으로, 증거는 신뢰롭고, 시기적절해야 하며, 이해하기 쉬워야 한다(OECD, 2012). 지금까지 정책평가 연구에서 활용된 ‘증거’는 설문조사 결과가 대부분이었는데, 설문조사법은 연구자가 원하는 데이터를 체계적으로 수집할 수 있고, 표집이 적절하게 이뤄진다면 모집단에 대한 통계적 추론이 가능하다는 장점이 있다. 그러나 이 방법은 데이터 수집에 많은 시간과 비용이 소모되고, 의식적으로 표현된 의견만 파악할 수 있으며, 인터넷과 스마트폰을 통해 정보가 실시간으로 생산·공유·확산되는 현 시대에서는 조사가 이뤄지는 동안 정책 여론이 바뀔 수 있는 한계가 있다(김대중, 윤서연, 2013; 송태민, 2014; 이미숙, 이창훈, 김지연, 2014).

이러한 정책평가 측면에서 빅데이터가 갖는 가치를 김기환(2013), 김대중과 윤서연(2013)의 논의를 바탕으로 정리하면 다음과 같다. 첫째, 국민

의 목소리를 신속하고 정확하게 반영함으로써 정책 집행의 투명성과 의사결정의 시의 적절성을 실현할 수 있다. 미디어나 SNS를 통해 표현되는 비정형 텍스트 빅데이터는 특정 주제나 서비스에 대한 감정의 흔적이기 때문에 이를 분석하면 현재의 의견 수렴 방법으로는 제대로 감지되지 않는 국민의 다양한 요구와 정서 등을 빠르고 정확하게 파악할 수 있기 때문이다. 이 경우 정책 모니터링 및 정책에 대한 선제적 대응에 기여함으로써 정책에 대한 만족도 역시 제고할 수 있다. 둘째, 보다 정확하게 미래를 예측함으로써 정책 집행의 효율성과 효과성을 증진시킬 수 있다. 빅데이터는 현실에서 일어나는 현상을 지속적으로 기록한 흔적이므로, 다양한 출처의 빅데이터를 함께 분석할 경우 정책의 파급 효과를 광범위하게 파악할 수 있다. 따라서 기존의 데이터로는 파악할 수 없었던 숨겨진 패턴이나 상관관계를 찾아내거나 신뢰성 있는 예측을 도출할 수 있어 문제해결에 가장 우수한 정책 대안을 효율적으로 선택할 수 있다.

다음에서는 송영조(2012), 김대중과 윤서연(2013), 이정아(2015)의 논의를 참고하여, 정책의 형성, 결정, 집행, 평가 단계별로 빅데이터를 어떻게 활용할 수 있는지 고찰하였다.

#### 가) 정책 형성 단계

‘정책 형성’ 단계는 현재 시행되고 있는 정책의 문제가 무엇인지 인식하고 진단하며, 새로운 정책 의제를 설정하는 단계이다. 과거에는 특정 정책에 대한 문제점이 정치권 인사의 문제 제기 또는 미디어의 보도에 의해 이뤄지는 경우가 많아 영향력이 큰 특정 개인이나 집단의 이익을 대변하거나 수동적인 특성을 띠 수밖에 없었다. 이러한 의견 개진의 비대칭성과 의견 수렴의 수동성을 해소하기 위한 방법으로 빅데이터를 활용한 정책평가를 고려해 볼 수 있다. 즉, 빅데이터를 활용하여 현재 시행 중인 정책의 쟁점이나 그에 대한 국민의 정서를 실시간으로 분석한다면, 전통적 설문조사를 활용할 경우 발생하는 회고적 데이터 수집의 단점을



극복하고, 시의적절한 정책 의제를 설정하는 데 기여할 수 있을 것이다. 아울러 빅데이터를 활용하여 정책의 세부 분야별로 정책 수요를 분석한다면, 정책분야별 우선순위를 설정하는 데에도 기여할 수 있을 것이다.

#### 나) 정책 결정 단계

‘정책 결정’ 단계는 정책 형성 단계를 통해 제시된 여러 가지 대안들을 선정·검토하여 정책을 확정해 가는 단계이다. 지금까지 이러한 정책 결정 과정에는 제한된 수의 전문가, 이해집단, 정부인사, 정당 등이 참여하여 영향력을 미쳤는데, 이 경우 정책에 대한 실제 수요자들의 의견이나 지난 정책 효과에 대한 과학적 분석 결과는 간과되는 경향이 있다. 이에 비해 오랜 시간 동안 누적된 빅데이터에는 어떠한 정책을 시행했을 때 나타나는 다양한 측면의 효과들과 개인의 의견 등이 포함되어 있어 정책 결정시 활용도가 높을 것으로 예상된다. 즉, 빅데이터를 활용해 정책평가를 실시한다면, 다양한 여론을 실시간 모니터링하여 정책 결정에 반영할 수 있다. 또한 이를 통해 정책에 대한 국민의 참여와 공감대를 형성하고 시행착오를 최소화하여 잘못된 정책으로 인한 사회적 비용을 감소시킬 수 있을 것이다.

다만, 구체적인 정책 대안을 마련하고 결정하는 것은 주로 전문가 집단에 의해 이뤄지기 때문에, 정책 결정 단계의 경우 다른 단계에 비해서는 빅데이터의 활용이 제한적일 수 있다. 또한 정책 의사결정자들이 빅데이터를 통해 획득한 증거를 정책 수립에 반영하도록 유도하기 위해서는 증거의 정당성이라고 할 수 있는 빅데이터의 품질, 정확성, 타당성, 신뢰성, 당면 문제와의 관련성, 실현가능성 등을 확보하고, 서로 다른 형태를 지닌 무수한 종류의 증거들을 어떻게 체계화하여 축적할 것인지에 대해서도 숙고할 필요가 있다.

#### 다) 정책 집행 단계

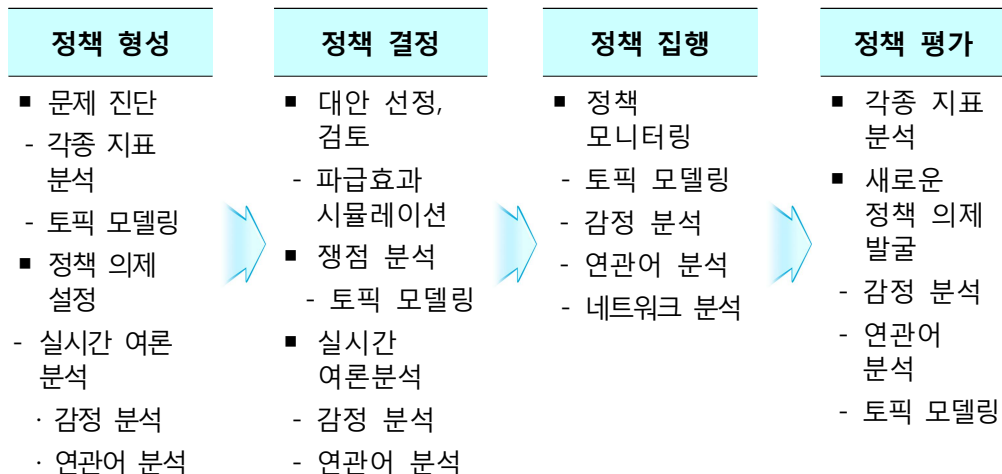
‘정책 집행’ 단계는 정책의 형성, 결정 단계를 거쳐 확정된 정책을 집행하기 위해 관련 기반을 준비하고 실제 실행하는 단계이다. 정책을 집행한 이후에는 정책에 대한 반응 결과가 지속적으로 축적되지만, 기존의 방법으로는 이것을 분석하고 그 결과를 알게 되는 데까지 상당한 시간이 소요된다. 이때 빅데이터를 활용하면 정책에 대한 이해당사자들의 의견이나 정책 발표에 의해 관심이 집중되는 데이터를 실시간으로 조회할 수 있어 지속적인 정책 모니터링 시스템을 구축할 수 있다.

또한 정책 집행 단계에 대한 실시간 모니터링 데이터에는 정책이 의도한 영향 이외에도 다양한 간접적 파급 효과 및 여론 추이가 포함되어 있어 집행시 문제점을 발견할 경우 즉각적으로 보완 조치할 수 있을 것이다. 예컨대, 토픽분석을 활용할 경우 특정 정책에 대해 국민들이 어떠한 부분을 문제로 인식하고 있는지를 파악할 수 있으며, 이를 오피니언 마이닝과 결합하여 분석할 경우 정책에 대한 긍정 또는 부정적 인식과 관련된 이슈가 무엇인지, 왜 그러한 반응을 보였는지에 대한 원인도 살펴볼 수 있을 것이다.

#### 라) 정책 평가 단계

‘정책 평가’는 정책이 집행되는 중에도 상시적으로 이루어지는 활동이나, 이 단계에서는 정책 집행이 완료된 후 실시되는 총괄평가로서의 정책 평가를 의미한다. 빅데이터 형태로 누적된 정책 반응을 이용하면 시행된 정책의 인과 관계를 파악할 수 있고, 정책에 대한 다각적 평가가 가능하다. 예컨대, 정책 완료 후 SNS 빅데이터를 활용하여 관련 정책의 쟁점과 그에 대한 만족도를 분석한다면, 이후 새로운 정책의제를 도출함으로써 다음 단계의 정책 개선을 위한 의사 결정에 이용할 수 있다.

[그림 II-9]는 이상의 논의를 바탕으로 각 단계별 정책과정에서 빅데이터가 어떻게 활용될 수 있는지를 나타낸 것이다.



[그림 II-9] 정책 단계별 빅데이터의 활용(송영조, 2012 수정)

## 2) 빅데이터를 활용한 정책평가 관련 선행연구 분석

정부 3.0 기조의 확산으로 인해 빅데이터를 활용한 정책평가 연구 사례가 조금씩 등장하고 있다. 그러나 사회과학 분야에서 빅데이터를 활용하여 새로운 분석의 틀을 짠 경우는 아직 보기 어렵다(한신갑, 2015). 즉, 기존 연구 틀에 빅데이터를 수용하여 기존의 연구를 확장·심화한 것이 최근의 추세인데, 특히 이러한 추세는 그동안 데이터의 집적과 전산처리가 상대적으로 어려웠던 텍스트를 대상으로 한 연구들에서 두드러진다.

또한 관련 선행연구의 경우 현재까지는 대부분 상용서비스를 이용한 전국적 여론 트렌드 분석에 머무르고 있다. 특히 이들 상용서비스들은 대체로 데이터의 수집, 가공, 분석 과정을 투명하게 공개하지 않고, 키워드 빈도 분석, 감정 분석, 연관어 분석 등으로 제한된 범위의 분석 기법을 제공한다는 한계점이 있다(황명화, 2014).

이는 국내 교육 분야에서도 마찬가지로, 빅데이터를 활용한 연구는 아직 시작 단계다. 이에 따라 향후 다양한 교육 빅데이터의 발굴 및 분석을 통해 최적의 교육 방법과 교육정책을 마련해야 한다는 점이 강조되고 있다(나일주, 임철일, 조영환, 2015; 조용상, 2013; 최제영, 2012). 이는 교육 분야에서도 마찬가지로, 지금까지 수행된 빅데이터를 활용한 정책평가 선행연구들로는 권정은과 정지선(2012), 권기석, 박진수, 구찬동(2014), 김우주와 구찬동(2015), 김정미와 윤미영(2012), 김정원, 김지수, 최유림(2014), 김지은과 백순근(2016), 박선주(2015), 최인봉, 이채희, 이은경(2015), 홍영란 외(2016)의 연구가 있으며, 이를 통해 주로 오피니언 마이닝(혹은 감정분석)과 텍스트 네트워크 분석, 토픽 모델링이 교육정책을 평가하는 데 활용되었음을 확인할 수 있다(해당 연구들의 분석 데이터, 연구 내용, 연구 방법은 <표 II-12> 참고).

<표 II-12> 빅데이터를 활용한 교육정책 평가 관련 선행연구

연구자 (연도)	분석 데이터	연구 내용	분석 방법
권정은, 정지선 (2012)	온라인 뉴스, 블로그, 카페, 트위터, 페이스북	· 청소년이 자살과 관련하여 언급한 내용 분석	· 버즈 분석 · 연관어 분석 · 감정 분석 · 영향력자 분석
김정미, 윤미영 (2012)	트위터, 페이스북, 블로그, 카페, 유튜브	· 영유아 보육 정책에 대한 여론 분석 - 영유아 보육료 - 양육수당 - 보육교사 처우개선	· 대화점유율 및 미디어 유형 빈도 분석 · 토픽 모델링 · 연관어 분석 · 감정 분석
권기석, 박진수, 구찬동 (2014)	교육학 분야 논문의 서지데이터, 온라인 뉴스, 블로그, 웹문서	· 교육 분야의 정책 아젠다 및 수요 분석	· 주제어 빈도 분석 · 텍스트 네트워크 분석

연구자 (연도)	분석 데이터	연구 내용	분석 방법
김정원, 김지수, 최유림 (2014)	남·북한교사용 정기간행물, 남·북한 교사 대상 집담회 전사 데이터	<ul style="list-style-type: none"> <li>· 남·북한 교사에 대한 공적 요구 분석</li> <li>· 남·북한 교사가 느끼는 주관적 교사 역할 분석</li> </ul>	<ul style="list-style-type: none"> <li>· 토픽 모델링</li> <li>· 텍스트 네트워크 분석 (의미망 분석)</li> </ul>
최인봉, 이채희, 이은경 (2015)	온라인 뉴스, 블로그, 트위터 데이터	<ul style="list-style-type: none"> <li>· 학업성취도, 학업성취도 평가, 국가수준 학업성취도 평가와 관련 여론 현황 및 이슈 분석</li> </ul>	<ul style="list-style-type: none"> <li>· 이슈 트렌드 분석</li> <li>· 정책에 대한 긍·부정어 주제어 분석</li> <li>· 세부 주제별 연관어 분석</li> </ul>
박선주 (2015)	소프트웨어 교육 관련 뉴스데이터 (네이버 뉴스)	<ul style="list-style-type: none"> <li>· 소프트웨어 교육에 대한 대중의 관심 방향 분석</li> </ul>	<ul style="list-style-type: none"> <li>· 주제어 빈도 분석</li> <li>· 연관성 분석</li> <li>· 토픽 모델링</li> </ul>
김우주, 구찬동 (2015)	소프트웨어 교육, 스마트 교육 관련 블로그, 트위터 데이터	<ul style="list-style-type: none"> <li>· 교육정보화 주요 정책(디지털교과서, SW 교육 등)에 대한 여론 분석</li> </ul>	<ul style="list-style-type: none"> <li>· 주제어 빈도 분석</li> <li>· 토픽 모델링</li> <li>· 감정 분석</li> <li>· 텍스트 네트워크 분석</li> </ul>
김지은, 백순근 (2016)	언론보도 기사 자료, 교육부 보도자료	<ul style="list-style-type: none"> <li>· 대학구조개혁 평가 정책에 대한 쟁점 분석</li> </ul>	<ul style="list-style-type: none"> <li>· 주제어 빈도 분석</li> <li>· 토픽 모델링</li> </ul>
홍영란 외(2016)	검색포털, 트위터, 언론보도 기사, 전문학술자료	<ul style="list-style-type: none"> <li>· 고등교육 및 평생교육 정책의 주요 쟁점 분석</li> </ul>	<ul style="list-style-type: none"> <li>· 텍스트 네트워크 분석</li> <li>· 화제어 추이 분석</li> </ul>

예컨대, 권정은과 정지선(2012)은 효과적인 청소년 자살 예방 정책의 수립과 추진의 방향성을 제시하기 위해 온라인 뉴스, 블로그/카페/게시판, SNS와 같은 소셜 데이터를 분석하였다. 이를 위해 우선, ‘자살’을 키워드로 하여 온라인 채널에서 발생한 텍스트 문서들을 수집하고, 이 중

작성자가 청소년임을 추정할 수 있는 내용(예: 나이, 학년, 성적, 담임과 같이 학교생활을 의미하는 단어)이 포함된 데이터를 별도로 추출하여 용어 사전을 작성하였다. 이후 매체별 자살 관련 언급 빈도수와 해당 키워드가 가장 많이 언급되는 요일과 시간, 언급 형태(대화형, 전파형, 독백형, 정보링크형), 확산속도를 분석하고, 연관어 분석을 통해 자살의 주요 원인을 탐색하고자 하였다. 또한 감정분석을 활용해 자살에 대해 느끼는 감정의 종류와 주제어를 도출하였다. 아울러 자살에 대한 문서를 가장 많이 작성한 사람을 추출하고 이들의 영향력과 성향을 파악함으로써 이들을 활용한 자살예방 정책이 필요함을 역설하였다.

김정미와 윤미영(2012)은 트위터, 페이스북, 블로그, 유튜브 등에서 언급된 영유아 보육정책을 영유아 보육료, 양육수당, 보육서비스, 장애아/다문화 지원, 보육교사 처우 개선의 다섯 가지 범주로 구분하고, 범주별 키워드를 활용해 수집한 유효한 멘션의 양(Voice of Customer: VOC)이 전체 대비 50% 이하인 보육서비스, 장애아/다문화 지원은 제외한 후 국민 여론을 분석하였다. 분석 결과 영유아 보육정책에 대한 여론을 주도하는 미디어는 트위터, 블로그 순으로 나타났으며, 해당 정책에 대한 긍정 비율은 1%, 부정 비율은 35%로 여론은 좋지 않았다. 즉, 국민들은 영유아 보육료 정책에 대한 불신과 불만을 표현하였으며, 보육료 지급에 대한 개선 의견을 제시하고 있는 것으로 나타났다.

또한 권기석, 박진수, 구찬동(2014)은 교육학 분야 논문의 서지데이터, 언론기사, 연구보고서, 한국교육학술정보원의 교육정보통계시스템(EDSS) 데이터를 기반으로 사회 연결망 분석과 의미망 분석을 활용하여 우리나라 교육정책의 미래 수요를 분석하였다. 분석 결과, 교육 연구의 문헌 데이터에서는 사교육(비), 학업성취(도), 고등교육재정, 책무성, 학교폭력, 학자금, 국제화 등의 아젠다가 도출되었고, 언론 빅데이터 분석에서는 교육과정, 평생교육, 사교육비, 학교폭력 등의 영역이 도출되었다.

김정원, 김지수, 최유림(2014)은 남·북한 교사에 대한 공적 요구를 도출하기 위해 1990년대 이전과 2000년대 이후 발행된 남북한 교사용 정기

간행물의 기사 제목을 토픽 모델링을 활용하여 분석하였다. 분석 결과, 교사용 정기간행물에서 강조된 주제들은 남·북한 교육법에 나타난 교사에 대한 공적 요구와 밀접한 관련을 보였다. 그리고 남·북한 교사가 자신의 역할을 어떻게 규정하는지를 확인하기 위해 관련 집담회 데이터에 대한 의미망 분석을 실시한 결과, 남한 교사들은 학생들의 행복 증진에 필요한 전문적 의견을 개진할 수 있는 교육전문가로서의 역할을 지향하였다. 반면, 북한 교사들은 당 정책의 전달자가 아닌, 지식 교육과 교양을 담당했던 과거의 역할이 회복되어야 한다고 판단하였다.

한편 최인봉, 이채희, 이은경(2015)은 소셜 빅데이터(뉴스, 블로그, 트위터)를 기반으로 트렌드 분석, 연관어 분석, 감정분석을 활용해 학업성취도 평가와 관련된 이슈와 여론을 분석하였다. 분석 결과, 2012년에는 학업성취도 평가가 학교 서열화와 경쟁 교육을 부추긴다는 비판 여론이, 2013년에는 초등학교 학업성취도 평가 폐지와 중·고등학교의 기초학력 미달 학생 비율의 증가에 대한 논쟁이, 2014년에는 학업성취도 평가 시행과 관련된 좌파 교육감 정책이 포퓰리즘이라는 비판 여론이 형성되었다. 그리고 학업성취도에 대한 긍정적 의견으로는 학생들의 기초학력 정착, 성취수준 파악을 통한 맞춤형 지원, 일관성 있는 평가 등이, 부정적 의견으로는 평가에 대한 과도한 스트레스 및 부정행위, 학교 간 경쟁으로 인한 부작용 발생 등이 나타났다.

끝으로, 김우주와 구찬동(2015)은 블로그, 트위터 등 소셜 미디어에서도 출되는 텍스트 빅데이터를 활용하여 우리나라 교육정보화 정책에 대한 여론을 분석하였다. 이를 위해 ‘소프트웨어 교육’, ‘스마트 교육’ 관련 키워드 총 28개를 이용하여 데이터를 수집하고, 토픽 모델링을 이용해 관심 빈도와 추이를 분석하였다. 또한 수집된 빅데이터 문서 각각에 대해 감정점수를 계산한 후 극성을 분류하여 토픽별 감정 트렌드 분석을 실시하였다. 분석 결과, ‘컴퓨팅 사고력’, ‘소프트웨어 의무 교육’, ‘스크래치 프로그램, 네이버 소프트웨어’와 같은 토픽이 추출되었으며, 이와 관련된 감정은 중립 수준이나 점차 긍정적으로 변화하고 있음을 확인하였다.

### Ⅲ. 연구 방법

#### 1. 분석 데이터

##### 가. 데이터 수집

이 연구에서는 네이버의 뉴스 홈페이지(<http://news.naver.com>)(이하 ‘네이버 뉴스’)를 활용해 2015 개정 교육과정 정책 관련 뉴스 기사를 수집하였다. 신문, 텔레비전 등의 다양한 언론 매체 중에서 ‘인터넷 포털’을 선택한 것은 최근 해당 매체의 영향력이 상승한 점을 반영한 결과다. 현 시점에서 전통적 매체 영향력을 고려해 주요 언론 매체로 분석 데이터를 제한하는 것은 정책에 대한 국내 여론을 전체적으로 조망하고 텍스트의 사회 맥락이나 담론 속 논증 구조를 드러내는 데 한계가 있을 수 있다(박대민, 2013). 또한 수많은 국내 인터넷뉴스 서비스 포털 중에서 ‘네이버’를 선택한 이유는 국내에서 해당 매체의 여론 영향력이 가장 크기 때문이다. 2013년부터 2015년까지 여론집중도조사위원회의 조사 결과, 국내 여론영향력 점유율은 네이버에서 가장 높게 나타났다(윤영철 외, 2015).

네이버 뉴스에서 제공하는 인터넷 뉴스 기사는 비정형 텍스트 데이터 형식이다. 웹 문서 내 텍스트 데이터를 자동화된 방법으로 수집하기 위해서는 공개 API나 웹 크롤링 또는 웹 스크래핑 방법을 사용할 수 있다. 그런데 네이버 뉴스에서 제공하는 공개 API는 기사 본문 내용 검색을 지원하지 않기 때문에 이 연구에서는 사용할 수 없다. 또한 이 연구에서는 뉴스 기사가 탑재된 웹 문서 내 모든 정보가 필요한 것이 아니라, 뉴스 작성 날짜, 언론사명, 기사 제목 및 본문 내용, 관련 URL 등과 같이 뉴스 기사와 직접적으로 관련된 특정 정보만을 선택적으로 수집하면 되기 때문에 웹 크롤링이 아닌 ‘웹 스크래핑’ 기법을 활용하였다.



웹 스크래핑은 웹 문서 내 HTML(Hypertext Markup Language)의 특정 패턴만을 분석하여 추출하는 기법이다. 따라서 웹 스크래핑 기법을 이용해 네이버 뉴스의 뉴스 기사들을 수집하기 위해서는 수집할 웹 문서의 주소(URL)를 알아야 한다. 해당 URL은 네이버 뉴스 홈페이지에서 ‘검색어’와 ‘검색 기간’을 입력하면 컴퓨터 화면을 통해 바로 확인할 수 있다([그림 III-1] 참고). 이 연구에서 활용한 검색어는 ‘문·이과 통합 교육과정’과 ‘2015 개정 교육과정’으로, 2015 개정 교육과정 정책 과정(policy process)에서 언급된 명칭들을 모두 고려하였다<sup>12)</sup>. 이때 문·이과 ‘통합형’ 교육과정이 아닌 문·이과 ‘통합’ 교육과정을 검색어로 사용한 것은 좀 더 유연하게 데이터를 수집하기 위함이다. 또한 검색 기간은 이론적 배경에 근거하여 문·이과 융합안이 최초 언급된 ‘2013년 8월 27일’부터 이 연구의 분석 시점인 ‘2016년 9월 30일’까지로 설정하였다.



[그림 III-1] 네이버 뉴스 홈페이지의 뉴스 기사 검색창 화면

12) 참고로, 네이버에서 이 검색 키워드를 사용하면 자동적으로 띄어쓰기에 따른 단어 조합을 모두 고려하여 검색 결과를 반환한다. 예컨대, ‘2015 개정 교육과정’을 검색 키워드로 할 경우 ‘2015 개정교육과정’, ‘2015개정교육과정’, ‘2015 역사과 개정 교육과정’과 같은 단어도 모두 검색된다.

네이버 뉴스에서 2015 개정 교육과정 정책 관련 뉴스를 검색한 결과, 총 222개 언론 매체의 뉴스 기사 3,371건이 검색되었다. 네이버 뉴스에서는 ‘제목과 본문’ 보기로 검색할 경우 하나의 페이지(화면)에 10개의 뉴스 기사가 나타나므로 이 경우 총 338페이지에 해당하는 웹 문서 URL (이하 ‘1차 URL’, 프로그래밍 코드에서는 `base_url`)이 얻어진다. 즉, 1차 URL은 총 338개로, 검색어가 “문이과 통합 교육과정”인 경우 28개, “2015 개정 교육과정”인 경우 310개다. 이때 동일 검색어에 대한 1차 URL은 앞부분이 모두 동일하고 마지막 부분의 페이지 숫자만 페이지에 따라 변경되는 것을 확인할 수 있었다. 따라서 이 연구에서는 네이버 뉴스의 338개 웹 문서에 대한 HTML 정보를 수집할 때, 이러한 특성을 고려하여 Python 프로그래밍 코드(code)를 작성하였다([그림 III-2] 참고).

```
def parse_naver_news_page(uls):
    result = list()
    for item in uls:
        result.append(parseItem(item))
    return result

base_url = 'http://news.naver.com/main/search/search.nhn?query=%26quot%3B2015+%B0%B3%C1%A4+%B1%B3%C0%B0%B0%FA%C1%A4%26quot%3B&st=news.all&q_enc=EUC-KR&r_enc=UTF-8&r_format=xml&rp=none&sm=all.basic&ic=all&so=datetime.dsc&detail=0&pd=1&start=1&display=10&page=' # 1차 URL

naverNews = []
Num_pages = 310 # 검색어가 '2015 개정 교육과정'인 경우의 웹문서 수

for page_num in range(1, Num_pages + 1):
    res = None
    url = base_url + str(page_num)
    soup = BeautifulSoup(requests.get(url).text, 'html5lib')
    res = parse_naver_news_page(soup('ul', 'srch_lst'))
    naverNews.extend(res)
```

[그림 III-2] 1차 URL을 이용한 310개 웹문서의 HTML 정보 수집 code  
(검색어: 2015 개정 교육과정)

이와 같이 1차 URL을 기반으로 추출한 338개 웹 문서의 HTML 정보를 이용하면, 해당 웹문서에 포함된 3,371건의 개별 뉴스 기사에 대한

‘보도일자’, ‘언론사명’, ‘제목’, ‘언론사 URL’(해당 매체 홈페이지 상의 URL)과 ‘네이버 뉴스 URL’(네이버 뉴스에서 별도로 편집한 자체적 URL)(이하 ‘2차 URL’) 정보를 추출할 수 있다. 즉, 1차 URL을 바탕으로 웹 스크래핑을 실시해 수합한 HTML 문서의 문법적 표시, 즉 태그(tag)를 이용하여 개별 뉴스 기사들의 ‘보도일자’, ‘언론사명’, ‘제목’, ‘언론사 URL’, ‘네이버 뉴스 URL’ 관련 정보를 추출하는 것이다.



[그림 III-3] 뉴스 기사의 ‘보도일자’, ‘언론사명’, ‘2차 URL’ 관련 HTML 정보

예컨대, [그림 III-3]의 뉴스 기사는 ‘2016년 9월 29일’ ‘뉴스’라는 언론 매체에서 보도되었는데, 관련 정보를 추출하기 위한 HTML 정보를 살펴보면 보도일자는 ‘span’이라는 태그에 class 속성명이 ‘time’인 부분에, 언론사명은 ‘span’이라는 태그에 class 속성명이 ‘press’인 부분에 나타나는 것을 확인할 수 있다.

그러나 1차 URL을 통해 확인할 수 있는 웹 문서 화면에는 뉴스 기사  
의 본문이 일부분만 표시되기 때문에([그림 III-4] 참고) ‘본문’ 전체를 추  
출하기 위해서는 별도의 완결된 웹 문서 형태로 작성되어 있는 뉴스 기  
사별 2차 URL도 알아야 한다. 네이버 뉴스의 경우 2차 URL로 ‘언론사  
URL’과 ‘네이버 뉴스 URL’을 모두 제공하는데, 전자의 경우 ‘a’라는 하  
이퍼링크(hyperlink) 태그에 class 속성명이 ‘tit’인 ‘href’ 속성 부분에,  
네이버 뉴스 URL은 ‘a’라는 하이퍼링크 태그에 class 속성명이  
‘go\_naver’인 ‘href’ 속성 부분에 나타나 있다([그림 III-3] 참고).



[그림 III-4] 네이버 뉴스 홈페이지의 뉴스 기사 검색 결과 화면

그런데 네이버의 경우 뉴스 기사를 ‘뉴스검색제휴’와 ‘뉴스콘텐츠제휴’  
서비스 방식으로 나누어 공급한다. 전자는 기사콘텐츠에 대한 별도의 금  
전적 대가 없이 검색 결과에서 아웃링크(out-link) 방식으로, 후자는 금전  
적 대가에 기반해 인링크(in-link) 방식으로 네이버에 뉴스를 제공하는  
것이다(네이버 홈페이지, 2016). 그 결과, ‘뉴스검색제휴’를 통해 서비스되  
는 뉴스 기사(예: [그림 III-4]의 첫 번째, 두 번째 기사)는 ‘언론사 URL’  
정보만 존재하고 ‘네이버 뉴스 URL’은 존재하지 않는다. 이때 언론사  
URL의 경우 언론사 홈페이지마다 HTML 구조와 패턴이 모두 다르기

때문에 동일한 프로그래밍 코드로 웹 스크래핑을 구현하는 것이 불가능하다. 따라서 자동화된 텍스트 분석 방법을 활용하는 이 연구에서는 ‘네이버 뉴스 URL’을 제공하는 ‘뉴스콘텐츠제휴’ 기사(예: [그림 III-4]의 세 번째 기사)<sup>13)</sup> 총 1,598건으로 분석 대상을 한정하였다.

개별 뉴스 기사들의 본문 텍스트 전체를 추출하기 위해 ‘네이버 뉴스 URL’을 2차 URL 주소로 갖는 웹 문서의 HTML 구조를 살펴보면, 기사 본문은 ‘div’라는 태그에 id 속성명이 ‘articleBodyContents’인 부분에 나타나는 것을 확인할 수 있다([그림 III-5] 참고).



[그림 III-5] 뉴스 기사의 ‘본문’ 관련 HTML 정보

13) 각 언론 매체의 영향력을 가늠할 수 있는 한국ABC협회(2015)와 여론집중도 조사위원회(윤영철 외, 2015)의 발표 결과를 참고할 때, 이 뉴스 기사들은 대다수 사용자의 이용 점유율이 높은 매체에서 보도되었다. 즉, 해당 뉴스 기사들이 여론에 대한 대표성을 갖는다는 점에서 이러한 데이터의 제한은 무리가 없는 것으로 판단하였다.

이상의 논의를 종합하여, 이 연구에서는 Python의 beautifulsoup4 라이브러리를 기반으로 뉴스 기사를 웹 스크래핑하였다. 웹 스크래핑을 위한 오픈 소스 패키지는 R 프로그램에서도 제공하고 있으나, 이 연구에서 R이 아닌 Python을 활용한 이유는 웹 스크래핑 처리 속도가 Python이 더 우수하기 때문이다. [그림 III-6]은 전술한 것처럼 네이버 뉴스 홈페이지에서 확인할 수 있는 웹문서의 URL과 HTML 패턴 정보(태그)들을 활용하여 뉴스 기사를 수집하기 위해 연구자가 직접 작성한 Python 코드의 내용이다. [그림 III-6]에서 1차 URL은 base\_url, 뉴스 기사의 보도일자는 time(or date), 언론사명은 press, 2차 URL은 link\_naver와 link\_original, 제목은 title, 본문은 content에 해당한다.

```
def parseItem(uls):
    time = uls.find("span", "time").text      # 보도일자 추출
    title = uls.find("div", "ct").a.text      # 제목 추출
    press = uls.find("span", "press").text    # 언론사명 추출
    link_naver = uls.find("a", "go_naver")    # 네이버 뉴스 URL 추출
    link_original = uls.find("a", "tit").get("href") # 언론사 URL 추출
    date = "NA"
    content = "NA"
    if link_naver is not None:
        link_naver = link_naver.get("href")
        target = BeautifulSoup(requests.get(link_naver).text, 'html5lib')
        try:
            content = " ".join(target.find('div', {'id': 'articleBodyContents'} ).strings) # 본문 추출
        except:
            return None
    return {
        "title" : title,
        "date" : time,
        "press" : press,
        "content" : content,
        "link" : link_naver,
        "link_o" : link_original
    }
```

[그림 III-6] 1, 2차 URL을 이용한 웹 스크래핑 code

이러한 과정을 거쳐 분석 대상 데이터로 수집된 뉴스 기사는 64개 매

체의 총 1,598건으로 나타났다. 이 중 ‘문이과 통합 교육과정’과 ‘2015 개정 교육과정’이 모두 포함된 뉴스 기사 6건은 중복되어 수집된 것이기 때문에 ‘언론사 URL’을 활용하여 동일한 것은 한 건만 남겨두고 삭제하였다. 또한 수집된 뉴스 기사 중 분석 목적에 부합하지 않는 기사 80건(사교육 업체, TV 프로그램 등에 대한 광고·홍보성 기사 76건, 교육부 보도자료의 <표>로만 채워진 기사 4건)은 기사 제목에 나타난 사교육 업체명 등의 특정 문구를 활용하여 별도의 코딩(coding)을 통해 제외, 최종 분석 데이터는 61개 매체의 총 1,512건의 뉴스 기사다. 최종 분석 데이터를 수집하기까지 단계별 뉴스 기사의 사례수 및 제외 기준을 정리하면 <표 III-1>과 같다.

<표 III-1> 데이터 수집 단계별 뉴스 기사의 사례수 및 제외 기준

구분	사례수(건)	제외 기준
1단계	3,371	· 자동화된 텍스트 수집 방법의 적용 용이성 확보 → 네이버 ‘뉴스콘텐츠제휴’ 기사로 한정
2단계	1,598	· 분석 내용 및 방법의 타당성 확보 → 중복 수집 기사, 광고·홍보성 기사, <표>로만 채워져
3단계	1,512	분석할 텍스트 내용이 존재하지 않는 기사 삭제

또한 최종 분석 데이터를 정책 단계 및 매체 종류별로 살펴보면 <표 III-2>와 같으며 이때 ‘기타’에는 교육부와 청와대의 ‘정책브리핑’, 더불어민주당의 보도자료가 포함된다(언론 매체별 빈도수는 <부록 1> 참고).

<표 III-2> 분석 데이터의 정책 단계 및 매체 종류별 빈도(건수)

구분	신문·잡지	TV	인터넷	기타	계
정책 형성 단계	30	6	13	2	67
정책 결정 단계	308	58	152	23	620
정책 집행 단계	421	91	194	43	825
계	759	155	359	68	1,512

## 나. 데이터 클리닝

데이터 클리닝(data cleaning) 단계는 수집한 텍스트 데이터를 자연어 처리(NLP) 기법을 활용해 추후 분석이 용이한 형태로 변환하는 과정이다. 이 연구에서 데이터 클리닝은 텍스트 데이터에 대한 각종 ① 전처리(preprocessing) 작업과 ② 형태소 분석의 두 단계로 구성되었다.

### 1) 데이터 전처리

‘전처리’ 단계에서는 수집된 텍스트 데이터 중 ① 분석에 불필요한 단어 및 어구를 삭제하는 작업과 ② 동일·유사어지만 다르게 표현된 단어들을 통일하는 정규화(normalization)하는 작업을 매우 반복적으로 실시하였다. 이 연구에서는 웹 스크래핑 기법을 통해 분석에 필요한 정보만을 추출하고자 해당 웹 문서의 HTML 패턴을 최대한 정교하게 적용하였다. 그럼에도 불구하고 뉴스 기사의 속성이나 보도 매체의 특성상 기사 본문에는 분석시 불필요한 부분이 포함될 수밖에 없었다.

예컨대, 뉴스 기사의 본문 하단에는 해당 매체명이 포함된 광고성 문구나 공식 홈페이지 및 SNS 주소(예: “SNS에서 만나는 00일보 [페이스북 링크] [트위터 링크]”), 관련 뉴스 기사 링크 등의 불필요한 부분이 포함되기도 한다. 또한 뉴스 기사의 본문에는 해당 기사를 작성한 기자의 이름이나 직위, 장소(예: “서울=000 기자”), 이메일 주소 등이 표기되며, 신문·잡지 기사의 경우 [기획], <전문보도>와 같은 문구, TV 뉴스의 경우 “000 기자가 보도합니다”와 같은 불필요한 문장이 포함되기도 한다. 아울러 수집된 데이터가 인터넷 포털 뉴스인 만큼 “이미지를 크게 보려면 여기를 클릭하세요”와 같은 하이퍼링크 관련 문구도 포함된다. 이에 이 연구에서는 Python에 내장된 ‘re’ 패키지를 기반으로 연구자가 직접 정규식(regular expression)을 활용하는 코드를 작성하여 관련 내용들을 자동화된 방식으로 삭제하였다(주요 코드 예시는 [그림 III-7] 참고).





이후에는 동일한 의미지만 축약어로 다양하게 표현되는 용어들(예: ‘대학수학능력시험’, ‘수능시험’, ‘대수능’)을 동일한 단어(예: ‘수능’)로, 여러 가지 언어로 표기된 용어(예: ‘SW 교육’) 역시 하나의 언어(예: ‘소프트웨어 교육’)로 통일하였다. 아울러 완전히 동일하진 않아도 의미상 연결되는 단어들은 해석의 용이성을 확보하기 위해 통일하였다. 예컨대, ‘교사’, ‘교감’, ‘교장’은 모두 ‘교원’으로, ‘부모’, ‘엄마’, ‘아빠’는 모두 ‘학부모’로, ‘공부’는 ‘학습’으로 용어를 통일하였다. 또한 같은 단어이지만 띄어쓰기가 다르게 표현된 경우(예: ‘교과 중점 학교’, ‘교과중점 학교’) 컴퓨터는 각각을 서로 다른 단어로 인식할 수 있기 때문에 띄어쓰기도 통일(예: ‘교과중점학교’)하였다. 이러한 정규화 작업 역시 [그림 III-8]과 같이 연구자가 직접 Python 코드를 작성하여 전처리 작업을 수행하였다.

#### # 용어 통일① 유사어

```
content = re.sub('[0-9]+단위', '이수단위', content)
content = re.sub(" 교사대 ", " 교원양성기관 ", content)
content = re.sub("선생님", "교원", content)
content = re.sub("학?교장", "교원", content)
content = re.sub("교감", "교원", content)
content = re.sub("부장 ?교사", "교원", content)
content = re.sub("교사", "교원", content)
content = re.sub("교?과목", "교과", content)
content = re.sub(" 평가 ?역량", " 평가 전문성", content)
content = re.sub(" 평가전문성", " 평가 전문성", content)
```

#### # 용어 통일② 동의어/축약어

```
content = re.sub('융합 ?인재 ?교육', "STEAM", content)
content = re.sub('스팀 ?교육', "STEAM", content)
content = re.sub("SW", "소프트웨어", content)
content = re.sub("초 ?중 ?고(등학교)?", "초등학교 중학교 고등학교", content)
content = re.sub("고1", "고등학교 1학년", content)
content = re.sub(" 수학능력시험", " 수능", content)
content = re.sub(" 대학수학능력 시험", " 수능", content)
content = re.sub("대학수학능력시험", "수능", content)
content = re.sub("대입수학능력시험", "수능", content)
content = re.sub("수능시험", "수능", content)
content = re.sub("한문 ?병기", "한자 병기", content)
content = re.sub("한자 ?병용", "한자 병기", content)
content = re.sub("공교육 ?정상화 ?(촉진)?법", "공교육정상화법", content)
content = re.sub("수학 ?포기자", "수포자", content)
content = re.sub(" 외국 ", " 해외 ", content)
content = re.sub("자기 주도적으로 학습", "자기주도학습", content)
```

[그림 III-8] 단어 정규화를 위한 Python code 예시

## 2) 형태소 분석

이 연구에서는 분석 데이터에 대한 형태소 분석을 실시하기 위해 java 기반의 한국어 형태소 분석기인 ‘KOMORAN 2.0’(ver 2.4)의 Python 버전([https://github.com/shineware/komorany\\_2.0](https://github.com/shineware/komorany_2.0))을 활용하였다. 여러 가지 한국어 형태소 분석기 중 KOMORAN을 선택한 이유는 이 연구의 분석 데이터에 대해서는 해당 형태소 분석기의 성능이 가장 좋았기 때문이다. 즉, 한나눔 분석기는 조사가 제대로 분리되지 못하는 경우가 많았고, 꼬꼬마와 mecab-ko 분석기는 명사가 중복으로 추출되는 경우가 많았다.

KOMORAN은 입력 문장을 구성하고 있는 모든 형태소를 찾는 단계(Parsing)와 분석된 형태소를 조합해 최적의 품사열을 찾아내는 품사 부착(POS tagging) 부분으로 구성되어 있는데, 이때 품사 부착은 비교적 간단한 기계학습 알고리즘인 HMM(Hidden Markov Model)을 통해 수행된다. HMM 방법은 연속된 단어 패턴을 인식하는 데 무난한 성능을 보이며, 별도의 학습(training)이나 추후 사용자 사전(user dictionary), 기 분석 사전 등을 소스 코드 수정 없이도 기존 성능을 최대한 유지한다는 점에서 유용하다(신준수, 2014). 또한 KOMORAN의 경우 타 형태소 분석기와 달리 공백을 포함시켜 단일 형태소로 분석할 수 있기 때문에 앞서 실시한 전처리 작업을 단순화할 수 있다. 예컨대, 타 형태소 분석기에서 “바람과 함께 사라지다”라는 영화 제목을 하나의 고유명사로 추출하려면 띄어쓰기를 모두 붙인 후 “바람과함께사라지다”와 같은 형태로 사용자 사전에 추가해야 하는데, KOMORAN에서는 띄어쓰기를 하지 않아도 사용자 사전에 추가된 형식 그대로 고유명사를 추출한다.

이 연구에서는 해당 형태소 분석기에 내장된 기 분석 사전을 바탕으로 형태소 분석 결과 고유명사(NNP)와 일반명사(NNG)에 해당하는 ‘명사’만 선택하여 분석하였다. 다만, KOMORAN은 명사를 최대한 분리하도록 학습되어 복합명사까지도 분리(예: ‘핵심역량’→ ‘핵심’, ‘역량’)하는 경향이 있었다. 이러한 문제를 해결하기 위해 우선, 2015 개정 교육과정 문서(교

육부, 2015d)와 교육학용어사전(서울대학교 교육연구소, 2011), 교육평가용어사전(한국교육평가학회, 2004) 등을 참고하여 953개의 단어들을 연구자가 직접 사용자 사전에 추가한 후 형태소 분석을 실시하였다(사용자 사전에 포함된 전체 단어 목록은 <부록 3> 참조). 이와 같이 사용자 사전을 활용하면 복합명사 이외에도 연구자가 원하는 형태로 명사를 추출할 수 있어 분석시 유용하다.

그러나 사용자 사전을 추가한 후에도 여전히 분리되는 복합명사가 발생하였다. 따라서 이 연구에서는 복합명사의 형태로 추출되기 원하는 사용자 사전의 단어들의 경우 형태소 분석 후 분리된 명사를 다시 붙여줄 수 있는 별도의 프로그래밍 코드를 연구자가 직접 작성하여 함께 분석함으로써 이러한 문제점을 해결하였다(관련 코드는 [그림 III-9] 참고).

```
def combind(words, countDict):
    resultWords = list()
    resultCount = copy.deepcopy(countDict)

    def combindFind(idx, s):
        if idx >= len(words):
            return -987654321
        if s == words[idx]:
            return 1
        if s.startswith(words[idx]):
            return combindFind(idx + 1, s[len(words[idx]):]) + 1
        return -987654321

    index = 0
    while index < len(words):
        count = 0
        for u in userDict:
            count = max(count, combindFind(index, u))
        if count < 1:
            count = 1
        jCombindWord = "".join(words[index:index + count])
        index += count
        resultWords.append(jCombindWord)
        resultCount[jCombindWord] = resultCount.setdefault(jCombindWord, 0) + 1

    return resultWords, resultCount
```

[그림 III-9] 분리된 복합명사를 재조합하기 위한 Python code

## 2. 분석 방법

이 연구에서는 설정된 연구 문제를 해결하기 위해 빅데이터 기반의 자동화된 텍스트 분석 기법인 버즈 분석과 주제어 빈도 분석, 토픽 모델링을 사용하였다.

### 가. 버즈 분석

이 연구에서는 2015 개정 교육과정 정책에 대한 이해당사자들의 ‘관심도’를 분석하기 위해 버즈(buzz) 분석을 실시하였다. 이 연구에서 ‘버즈’는 네이버 뉴스 홈페이지를 통해 보도된 2015 개정 교육과정 정책 관련 뉴스 기사의 건수를 의미하며, 데이터 수집 단계를 통해 전체 분석 기간 동안 총 버즈는 총 1,512건이라는 것이 이미 확인되었다.

따라서 이 연구에서 버즈 분석은 2015 개정 교육과정 정책에 대한 언론의 관심도가 정책의 형성·결정·집행 단계에 따라 어떻게 변화하는가를 탐색하기 위해 활용되었다. 즉, 분석 데이터인 총 1,512건의 뉴스 기사를 최초 보도 일자를 기준으로 월별 합산하여 시간에 따른 빈도로 나타내고, 정책의 형성·결정·집행 단계에 따라 그 값이 어떻게 변화하는지 추이(를) 분석(trend analysis)하였다. 아울러 특정 사건 전후로 발현된 뉴스 기사의 월별 빈도를 비교함으로써 사건에 대한 반응도를 확인하는 이슈 반응 분석(issue response analysis)도 함께 실시하였다.

### 나. 주제어 빈도 분석

이 연구에서는 뉴스 기사에서 2015 개정 교육과정 정책의 내용 중 어떠한 부분에 주목했는지를 일차적으로 탐색하기 위해 주제어 빈도 분석(keyword frequency analysis)을 실시하였다. 우선, 이 연구에서 주제어

는 언론 매체에서 뉴스 기사를 보도할 때 등장한 주요 단어들을 의미하며, 뉴스 기사별 주제어들은 앞서 언급한 데이터 수집 및 클리닝 단계를 통해 csv 형식의 텍스트 파일로 저장되었다. 또한 뉴스 기사의 주제어들이 2015 개정 교육과정 정책의 형성·결정·집행 단계에 따라 어떻게 변화하는지를 살펴보기 위해 주제어 빈도에 대한 추이 분석을 실시하였다.

주제어 빈도 분석은 해당 텍스트 파일(keywords.csv)과 R 프로그램의 tm 패키지(Ingo & Kurt, 2015)를 활용하여 TF와 TF-IDF를 모두 산출한 후 그 결과를 비교하는 방식으로 이루어졌다. TF-IDF 값을 산출할 때에는 문서 크기에 따른 편중 현상을 방지하기 위해(이성직, 김한준, 2009) 문서의 단어 출현 빈도를 모든 단어의 총 출현 횟수로 나눈 정규화된(normalized) TF를 사용하였다. 그리고 분석 결과를 시각화하기 위해 R 프로그램의 wordcloud 패키지(Ian, 2014)를 활용하여 워드 클라우드(word cloud)를 작성하였다. 이때 Python 대신 R 프로그램을 사용한 이유는 R에서 제공하는 관련 패키지의 사용자 편의성이 더욱 우수하다고 판단하였기 때문이다(관련 코드는 [그림 III-10] 참고).

```
## 주제어 담긴 csv 파일 읽기 ##
txt_noun <- read.csv('keywords.csv', header=F, col.names=c("v1","v2"))
txt <- unlist(subset(txt_noun, select="v2"))
corpus <- Corpus(VectorSource(txt))

## Document Term Matrix 형성 ##
library(tm)
uniTokenizer <- function(x) unlist(strsplit(as.character(x), "[[:space:]]+"))
dtm <- DocumentTermMatrix(corpus, control= list(tokenize = uniTokenizer))
dtm_tfidf <- weightTfidf(dtm) # TF-IDF 산출: 정규화된 TF 이용

## 워드 클라우드 작성##
TermFreq <- colSums(as.matrix(dtm_tfidf)) # 단어별 빈도 계산
wordcloud(names(TermFreq), TermFreq,
  scale = c(8, .5), max.words = 30, # 최대 30개 단어만 표현
  random.order = 'False', # 빈도 높은 단어를 중앙으로 위치
  rot.per = 0, # 단어 회전 안함(모두 가로로 표현)
  colors = brewer.pal(7, 'Greys')) # 단어 색상 설정 흑백
```

[그림 III-10] 주제어 빈도 분석을 위한 R 프로그램 code

## 다. 토픽 모델링

주제어 빈도 분석은 정책 단계에 따라 어떠한 주제어가 출현하고 사라졌는지에 대한 변화를 효과적으로 보여주지만, 단어의 존재 유무만으로는 표현되지 않는 쟁점을 드러내는 데에는 한계가 있다(박종희, 박은정, 조동준, 2015). 이때 ‘토픽 모델링’(topic modeling)은 텍스트 데이터 내 단어들의 빈도를 통계적으로 분석하여 전체 데이터를 관통하는 잠재적 주제, 즉 토픽(topic)들을 자동으로 추출하는 분류한다는 점에서 쟁점 분석시 유용하다. 따라서 이 연구에서는 2015 개정 교육과정 정책의 주요 쟁점을 추출하기 위해 토픽 모델링을 실시하였다. 그리고 정책의 형성·결정·집행 단계에 따라 해당 정책의 쟁점이 어떻게 변화하는지를 파악하기 위해 토픽 트렌드 분석(topic trend analysis)을 실시하였다.

특히 이 연구에서는 다양한 토픽 모델링 기법 중에서도 Blei, Ng, & Jordan(2003)이 발표한 잠재 디리클레 할당(Latent Dirichlet Allocation: LDA) 기법을 활용하였다. 이는 LDA 기법이 다른 토픽 모델링 기법에 비해 결과 해석이 용이하고(Blei, 2012), 과적합(overfitting) 문제를 해결하기 때문에 방대한 비정형 데이터로부터 여러 가지 토픽(쟁점)들을 도출하는 데 유리하기 때문이다(Griffiths & Steyvers, 2004). 최근 발표되고 있는 다양한 토픽 모델링 관련 확장 모형 역시 LDA 기법에 기반하고 있기 때문에 이 분석 기법은 향후 연구의 확장성 측면에서도 유용하다.

LDA 기법에서는 주제어 빈도 분석과 마찬가지로 TF와 TF-IDF를 선택적으로 활용할 수 있는데, TF-IDF의 경우 일반적 단어를 분석할 때에는 매우 유용하나 고유명사나 새로운 개념을 분석할 때에는 중요도가 과소 및 과대평가될 수 있다(임화진, 2014). 따라서 이 연구에서는 TF를 이용하여 LDA를 수행하였다. 또한 전체 분석 데이터에서 5회 미만 등장한 단어와 거의 모든 문서에서 언급되는 최상위 고빈도 단어 중 내용 분석시 큰 의미를 갖지 않는 7개 단어(교과서, 교육부, 학생, 교과, 초등학교, 중학교, 고등학교)는 분석시 제외하였다.

또한 LDA 분석 시에는 ① 사후 분포의 추정 방법, 토픽 모형 생성시 표준화(normalization)를 위한 ② 반복수행 횟수(iteration), 상위 모수인 ③  $\alpha$ 와  $\eta$ 값, ④ 토픽의 개수를 사전에 설정해야 한다. 우선, ‘사후분포 추정 방법’으로는 선행연구에서 정확성이 높고 분석 속도가 개선되었다고 보고되는 붕괴된 깁스 표집(collapsed Gibbs sampling) 방식을 사용하였다. 그리고 토픽 모형 생성의 ‘반복수행 횟수’는 분석 시간과 반비례 관계에 있기 때문에, 분석의 정확성과 효율성 모두를 충족하기 위해서는 적당한 수를 설정해 주어야 한다. 선행연구들에서 보통 1,000에서 2,000 사이의 값을 권고(안주영, 안규빈, 송민, 2016)하는 점을 참고하여, 이 연구에서는 반복횟수를 1,000으로 설정하였다.<sup>14)</sup>

‘상위 모수’인  $\alpha$ 와  $\eta$ 값은 선행연구에서 가장 우수한 성능을 보인 비대칭적  $\alpha$ -대칭적  $\eta$  조합을 선택했다. 이때  $\alpha$ 의 초기 값과  $\eta$ 값은 Steyvers & Griffiths(2007)에 근거하여 대다수의 선행연구에서 채택하고 있는 ‘50/토픽수’와 0.01 값으로 각각 지정하였다. 아울러 ‘토픽의 개수’를 추출하기 위한 토픽 모형 평가 방법으로는 ‘복잡도’(perplexity)와 토픽 응집성(topic coherence) 지수 중 하나인 ‘ $C_v$ ’ 지수를 상호 보완적으로 활용하였다. 그리고 토픽의 해석 가능성을 고려하여, 추출된 토픽 내 주요어들이 조합을 통해 의미 있는 쟁점을 만들어 내는지를 확인하였다.

이상의 LDA 분석을 실시하기 위해 이 연구에서는 우분투(Ubuntu) 운영체제를 기반으로 Python의 ‘gensim’ 라이브러리를 활용하였다. 이때 윈도우즈(Windows) 대신 우분투 운영체제를 사용한 이유는 해당 운영체제가 리눅스(Linux)를 기반으로 하고 있어 텍스트 마이닝 관련 라이브러리 설치 및 분석 등의 측면에서 좀 더 효율적이기 때문이다. 또한 이

---

14) 실제 분석 단계에서는 반복횟수를 1,000, 1,500, 2,000의 세 가지 경우로 나누어 각각 실시한 후 그 결과를 비교하였다. 분석 결과, 반복횟수가 늘어날수록 LDA 분석에 소요되는 시간은 크게 증가하였으나 분석 결과 추출되는 토픽의 개수와 토픽별 단어의 구성, 복잡도와 토픽 응집성 지수에는 의미 있는 차이가 나타나지 않았다. 따라서 이 연구에서는 LDA 분석의 정확성과 효율성을 모두 고려하여 최종적으로 반복횟수를 1,000으로 지정하였다.



연구에서 gensim 라이브러리를 선택한 이유는 분석 속도가 빨라 대규모 문서집합 처리에 유용(김우주, 구찬동, 2015)하고, 다른 분석도구에서 활용하고 있는 추정 방법을 모두 포함하고 있기 때문이다.

아울러 gensim에서는 다양한 토픽 모델링 기법 관련 API를 제공하고 있는데, 이 연구에서는 gensim에서 Java 기반의 MALLET을 활용할 수 있도록 래핑(wrapping)한 ‘wrappers.LdaMallet’ API를 사용하였다. 해당 API는 LDA 분석을 실시할 때 이 연구에서 설정한 사후 분포의 추정 방법, 토픽 모형 생성의 반복수행 횟수,  $\alpha$ 와  $\eta$  값 모두를 기본 값으로 설정하고 있기 때문이다. 또한 토픽 모델링 결과를 시각화하기 위한 분석 도구로는 Python의 ‘pyLDavis’(Sievert & Shirley, 2014) 라이브러리를 사용하였다(관련 코드는 [그림 III-11], [그림 III-12] 참고).

```
# 각 단어에 숫자 속성 부여(Encode tokens to integers)
from gensim import corpora
dictionary = corpora.Dictionary(texts) # texts=전처리가 끝난 뉴스 기사의 본문
dictionary.save_as_text('dict.txt')

# TF 계산
from gensim import models
tf = [dictionary.doc2bow(text) for text in texts]

# LDA 분석
import numpy as np; np.random.seed(2016)
for ntopics in range(1, 16):
    lda = models.wrappers.LdaMallet('/mallet-2.0.7/bin/mallet',
                                     tf, num_topics=ntopics, id2word=dictionary)
    print(lda.print_topics(num_topics=ntopics, num_words= 10))
    lda.save('mallet_lda.pkl')

# LDA 분석 결과 시각화
import pyLDavis

## 이하에서 topic_term_dists, doc_topic_dists, doc_lengths, vocab, term_frequency 도출 과정은 생략

ldavis = pyLDavis.prepare(topic_term_dists, doc_topic_dists,
                           doc_lengths, vocab, term_frequency, R=10)
pyLDavis.save_html(ldavis, 'lda_mallet_vis.html')
```

[그림 III-11] LDA 분석을 위한 Python의 주요 code

```

import matplotlib.pyplot as plt

def evaluate_graph(dictionary, corpus, texts, limit):
    c_v = []
    lm_list = []
    for num_topics in range(1, limit):
        lm = models.wrappers.LdaMallet('/mallet-2.0.7/bin/
            mallet', num_topics=num_topics, id2word=dictionary)
        lm_list.append(lm)
        cm = models.CoherenceModel(model=lm, texts=texts,
            dictionary=dictionary, coherence='c_v')
        c_v.append(cm.get_coherence())

    # 토픽 개수에 따른 토픽 응집성(c_v) 도표화
    x = range(1, limit)
    plt.plot(x, c_v)
    plt.xlabel("num_topics")
    plt.ylabel("Coherence score")
    plt.legend(("c_v"), loc='best')
    plt.show()
    plt.close()
    return lm_list, c_v

lm_list, c_v = evaluate_graph(dictionary, tf, texts, 21)

```

[그림 III-12] 토픽 응집성을 도표화하기 위한 Python code

끝으로, 분석 결과 산출된 토픽(쟁점)들에 대한 명칭은 각 토픽을 구성하고 있는 단어와 이론적 배경, 2015 개정 교육과정 정책의 맥락 등을 종합적으로 고려해 연구자가 일차적으로 설정하였다. 그리고 해당 토픽명의 타당도를 확보하기 위해 교육학 전공 박사 및 박사과정 대학원생 총 13명을 대상으로 전문가 협의회를 개최하여 타당도를 검증하였다. 타당도는 Likert 방식의 5점 척도(① 전혀 타당하지 않다, ② 타당하지 않다, ③ 보통이다, ④ 타당하다, ⑤ 매우 타당하다)를 통해 평정하였으며, 평정 결과 4점 미만인 토픽명에 대해서는 논의를 거쳐 수정하였다(토픽명 결정을 위한 전문가 협의회 데이터 및 결과는 <부록 4> 참고).

이상의 연구 방법을 종합하면, 이 연구의 분석 절차는 [그림 III-13]과 같이 크게 세 단계로 구분할 수 있다.

## 1. 데이터 수집

### • 데이터 검색: 총 3,371건의 뉴스 기사

- 네이버 뉴스 홈페이지(<http://news.naver.com>)의 상세 검색 이용  
(검색어: 2015 개정 교육과정, 문이과 통합 교육과정) → 웹문서의 1차 URL 도출

### • 웹 스크래핑 (web-scraping) : 총 1,598건의 뉴스 기사

- 1차 URL, 웹문서 내 HTML 패턴, 파이썬(python)의 BeautifulSoup4 활용  
→ [뉴스콘텐츠제휴기사] 보도일자, 제목, 본문, 2차 URL(네이버뉴스/언론사) 추출

### • 데이터 필터링 (filtering) : 총 1,512건의 뉴스 기사

- 기사별 언론사 URL 이용 → 중복 기사 6건 제거
- 기사 제목 이용 → 분석 목적에 부합하지 않는 기사(예: 광고·홍보성 기사) 80건 제거

## 2. 데이터 클리닝

### • 데이터 전처리(preprocessing)

- 단어 정규화(regularization): 유사어, 축약어, 단어 간 띄어쓰기 통일
- 뉴스 기사 고유의 불필요 어구 제거: 기자명, 언론사명, 이메일, 홈페이지 주소 등
- 불용어(stopwords) 제거: 특수기호, 문장부호, 학교명, 검색어 등 총 472개
- 파이썬의 re 기반으로 연구자가 직접 정규식(regular expression)을 활용해 코드 작성

### • 형태소 분석

- 한국어 형태소 분석기 KOMORAN 2.0의 파이썬 버전 활용 → 명사(고유/일반) 추출
- 과도하게 분리된 복합명사들은 연구자가 작성한 사용자 사전 및 코드 활용해 재조합

## 3. 데이터 분석 및 시각화

### • 버즈(buzz) 분석: 언론 매체의 관심도 분석

- 기사에 대한 월별 보도 건수 추이 분석

### • 주제어(keyword) 빈도 분석

- R 프로그램의 tm 이용 → 빈도(TF/TF-IDF) 분석
- R 프로그램의 wordcloud 이용 → 단어 빈도 시각화

### • 토픽 모델링(topic modeling): 주요 쟁점 분석

- python의 gensim 라이브러리 중 Mallet을 래핑(wrapping)한 LDA 분석 API 사용
- 추정 방법: Collapsed Gibbs Sampling( $\alpha=50/k$ ,  $\eta=0.01$ )
- 평가 방법: 복잡도(perplexity), 토픽 응집도(topic coherence) 이용
- python의 pyLDAvis 이용 → LDA 분석 결과 시각화(토픽 간 거리 지도)



### • 전문가 협의회를 통한 최종 토픽(쟁점)명 결정

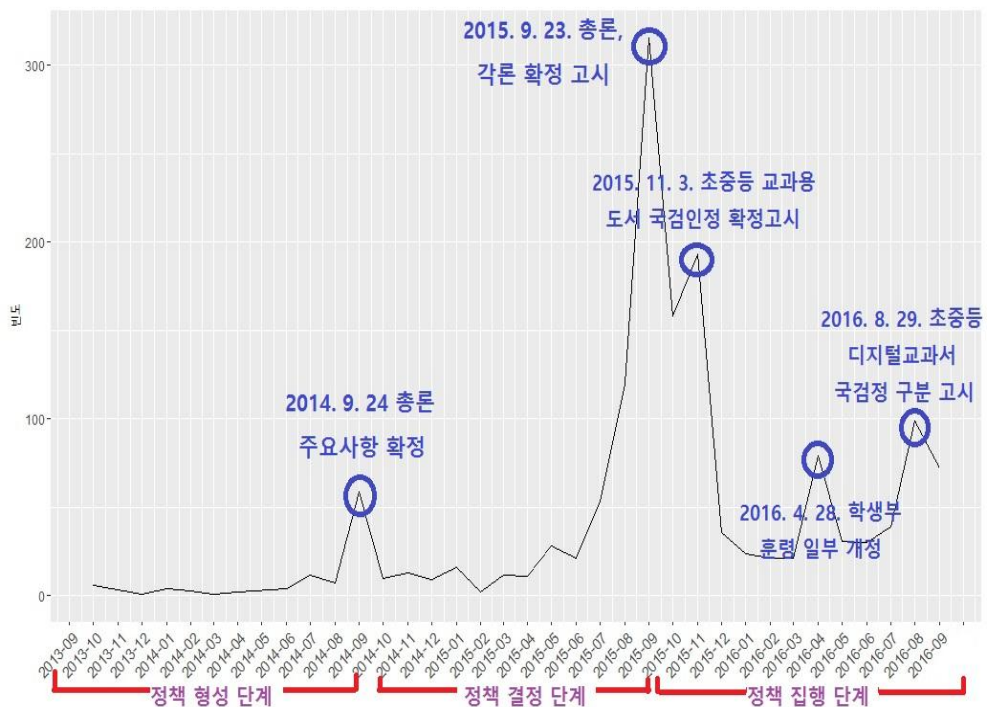
- 교육학 전공 박사 및 박사과정 대학원생 총 11명의 토픽명 타당도 평정
- 평정 결과 4점 미만(5점 만점 기준)인 토픽명은 논의 후 수정

[그림 III-13] 이 연구의 분석 절차

## IV. 연구 결과

### 1. 2015 개정 교육과정 정책에 대한 언론 매체의 관심도 분석

2015 개정 교육과정 정책에 대한 국내 언론 매체의 관심도를 분석하기 위해 2013년 8월부터 2016년 9월까지 보도된 관련 뉴스 기사들을 대상으로 버즈 분석을 실시하였다. [그림 IV-1]은 버즈 분석 결과를 월별로 나타낸 것으로, 이를 토대로 2015 개정 교육과정 정책에 대한 언론 매체의 관심도를 정책의 형성·결정·집행<sup>15)</sup>단계로 나누어 분석하면 다음과 같다.



[그림 IV-1] 뉴스 기사 보도 건수 월별 추이(총 1,512건 기준)

15) 이 연구에서는 연구의 시점상 2015 개정 교육과정 정책의 집행 단계로 2015 개정 교육과정의 ‘실행 준비기’만을 다루었다.

## 가. 정책 형성 단계

‘정책 형성’ 단계의 뉴스 기사 보도 건수를 살펴보면, 2015 개정 교육과정 정책은 문·이과 융합안이 처음 언급된 2013년 8월 27일부터 총론이 개발되고 있는 중에도 거의 보도되지 않은 것으로 나타났다. 이처럼 정책 형성 단계에서 거의 주목을 받지 못하던 2015 개정 교육과정 정책은 2014년 7월, 관련 보도 건수가 소폭 증가하였는데, 이는 2014년 7월 21일, 한국과학기술단체총연합회가 과학교육 축소를 반대하며 ‘교육과정 개편안 재논의 촉구’에 대한 긴급 기자회견을 열었기 때문이다.

2009 개정 교육과정에서 필수이수단위가 15단위였던 과학 교과가 2015 개정 교육과정(당시 문·이과 통합형 교육과정)에서는 12단위 혹은 10단위로 축소된다는 것이 비판의 주요 근거였는데, 이에 따라 8월 5일, 정책연구진에서는 과학 교과뿐 아니라 국·수·영·사와 같은 타 주요 교과도 모두 같은 수준으로 필수이수단위를 조정할 예정이라고 발표하였다. 그러나 과학계의 반발은 그치지 않았고, 한국과학기술단체총연합회는 2014년 8월 25일에 20개 과학기술 단체와 함께 공동성명을 내는 등 다시 한번 강하게 비판하였다. 그러나 버즈 분석 결과, 이러한 “‘이과 말살 교육과정’ 담론”(강지영, 소경희, 2015)의 경우 과학계의 반발 수준에 비해서는 사회적으로 크게 주목받지 못한 것으로 나타났다. 오히려 해당 시점에는 교육과정 재구조화 및 개발지침 공개 토론회를 통해 알려진 소프트웨어 교육 강화가 새롭게 관심을 받았다.

## 나. 정책 결정 단계

정책 형성 단계에서 일부 교과를 제외하고는 이해당사자들의 관심을 끌지 못하던 2015 개정 교육과정 정책은 ‘정책 결정’ 단계가 시작된 2014년 9월 24일, 총론 주요사항이 발표되자 관련 뉴스가 일제히 보도되었다.

당시에는 2015 개정 교육과정에서 신설되는 문·이과 통합교육, 안전교육과 같은 주제가 주로 소개되었는데, 여론의 큰 반향 없이 단발성 보도로 그쳤다. 즉, 총론 주요사항이 발표되었음에도 불구하고 그에 대한 이해당사자들의 관심은 크지 않아 관련 보도 건수는 다시 정책 형성 단계 수준으로 하락하였다. 2015년 4월과 5월에는 교과 교육과정 시안 공개 토론회도 수시로 열렸지만 이에 대한 관심은 미미한 수준이었다.

2015 개정 교육과정 정책에 대한 언론 매체의 관심이 다시 상승한 것은 총론과 교과 교육과정에 대한 공청회가 진행되기 시작한 2015년 7월부터다. 또한 2015년 8월 13일에는 ‘한글문화연대’ 등이 주관한 초등 교과서 한자 병기 관련 반대 집회가 열리면서 한자 병기 문제를 비롯한 2015 개정 교육과정 정책이 큰 주목을 받았다. 이러한 대중의 관심도는 총론과 교과 교육과정이 확정·고시된 2015년 9월 23일을 전후로 정점을 이룬 것을 버즈 분석을 통해 확인할 수 있다.

## 다. 정책 집행 단계

정책 집행 단계부터는 2015 개정 교육과정 정책 관련 뉴스 보도 건수가 소폭 줄었지만 2015년 10월 12일, 역사교과서 국정화 전환 내용이 포함된 「중·고등학교 교과용도서 국·검·인정 구분」(안)이 행정 예고되고 11월 3일에 해당 안이 확정·고시되기까지 ‘역사교과서 국정화 반대’ 관련 논쟁이 전국적으로 확산되어 다시 언론 매체의 큰 주목을 받았다. 그러나 역사교과서 국정화 사안이 최종 확정됨에 따라 관련 논란은 점차 소강 상태에 이르렀고 2015 개정 교육과정 정책에 대한 보도 기사 건수는 해당 정책이 집행 단계임에도 불구하고 급격히 감소하였다.

2015 개정 교육과정 정책에 대한 뉴스 보도는 2016년 4월에 소폭 증가하였다. 이 시기에는 「학교생활기록 작성 및 관리 지침」 훈령이 일부 개정되었음이 교육부를 통해 발표되었는데, 해당 시점에 보도량이 증가한 것은 이 사안 때문이 아니었다. 2016년 4월 20일, 서울특별시교육청에서

2016년 2학기부터 ‘개방형 선택 교육과정’을 운영하는 일반계 고등학교를 대상으로 시범학교를 지정해 ‘개방-연합형 종합캠퍼스 교육과정’을 도입한다고 발표한 것이 언론의 주목을 받았다. 이는 문·이과 과정 없이 보통교과 중심, 예체능교과 중심, 자유교양 중심, 직업기능 중심 등 개인별로 맞춤형 교과 과정을 운영한다(서울특별시교육청, 2016a)는 점에서 2015 개정 교육과정 정책과 ‘다툼은 풀’ 정책인데, 뉴스 보도에서는 이를 통해 일반계 고등학교 교육의 질적 향상에 대한 기대감을 전하였다.

이후 논의가 다시 소강 상태를 보인 2015 개정 교육과정에 대한 관심은 2016년 8월 29일 ‘초·중등 디지털 교과서 국·검정 구분 고시’가 발표된 시점 전후로 다시 상승하였다. 그런데, 실제 이 시점에서 디지털 교과서보다 더 언론매체로부터 관심을 받은 사안은 8월 23일 ‘사교육걱정없는세상’에서 발표한 초등학교 1, 2학년 수학교과서 현장 검토본 검토 내용과 8월 30일 서울특별시교육청이 발표한 초등학교 1, 2학년 대상의 ‘안정(안정과 성장) 맞춤 교육과정’이다. ‘사교육걱정없는세상’이 초등학교 1, 2학년 수학 교육 내용이 선행학습을 하지 않으면 이해하기 어렵다는 문제를 제기한 데 이어 서울특별시교육청(2016b)에서는 초등학교 1, 2학년 학생의 선행학습 부담을 덜어주기 위해 초등학교 1학년 공교육 과정 내에서 기초 한글 교육과 수학 교육을 책임지도록 한다는 정책을 발표한 것이다.

요컨대, 정책 집행 단계에서는 역사교과서 국정화 논란 이후 2015 개정 교육과정 정책에 대한 논의가 충분히 이뤄지지 못하였는데, 서울특별시교육청에서 발표한 교육정책들과 ‘사교육걱정없는세상’에서 발표한 수학교육 적정화 문제, 디지털 교과서의 도입 등은 공교육 정상화와 내실화 측면에서 언론 매체와 정책 이해당사자들의 관심을 받은 것으로 분석된다. 이를 통해 전국 시·도교육청 중 서울특별시교육청의 교육 방향과 지침이 언론 매체 및 교육 전반에 미치는 영향력을 추론할 수 있으며, ‘사교육걱정없는세상’과 같은 시민 단체들이 교육정책의 과정에서 담당하는 역할이 증대되고 있음을 확인할 수 있다.

## 2. 2015 개정 교육과정 정책 관련 언론보도의 주제어 분석

### 가. 분석 기간 전체

2013년 8월부터 2016년 9월까지 언론 매체를 통해 보도된 뉴스 기사에서 높은 빈도로 출현한 상위 주제어 20개를 TF를 기준으로 정리하면 <표 IV-1>과 같다. 즉, 분석 기간 전체를 기준으로 가장 많이 등장한 단어는 ‘교과서’였고, 그 다음으로는 ‘교육부’, ‘역사’, ‘학생’, ‘국정’, ‘교과’ 순으로 높았다. 이를 통해 국내 언론 매체는 2015 개정 교육과정에서 사용할 ‘교과서’에 가장 많은 관심을 보였고, 이해당사자 중에서는 ‘학생’을, 학교급으로는 ‘고등학교’, 교과 중에서는 ‘역사’ 교과를 가장 많이 언급한 것으로 나타났다.

<표 IV-1> 뉴스 기사 내 상위 빈출 20개 단어(TF 기준): 전체

순위	주제어	빈도(TF)	순위	주제어	빈도(TF)
1	교과서	8,474	11	정부	2,231
2	교육부	4,480	12	수업	2,064
3	역사	4,135	13	내용	2,059
4	학생	4,133	14	사회	1,990
5	국정	3,583	15	한국사	1,743
6	교과	3,370	16	한자	1,450
7	고등학교	3,190	17	수학	1,429
8	교원	2,971	18	소프트웨어	1,415
9	초등학교	2,881	19	개발	1,369
10	중학교	2,557	20	수능	1,350

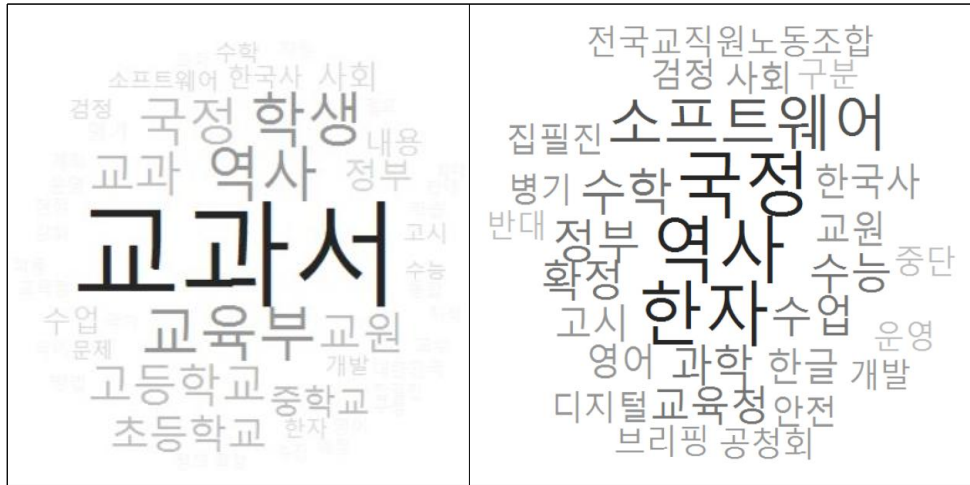


한편 모든 뉴스 기사에서 자주 출현하는 상투어의 중요도를 감소시키기 위해 TF-IDF를 기준으로 상위 빈출 단어를 나타내면 <표 IV-2>와 같다. 즉, 분석 기간 전체를 기준으로 가장 중요하게 고려된 단어는 ‘역사’였고, 그 다음으로는 ‘국정’, ‘한자’, ‘소프트웨어’, ‘수학’ 순으로 높았다. 이를 통해 국내 언론 매체는 2015 개정 교육과정 정책에서 강조하는 문·이과 융합교육의 측면 보다 역사교과서 국정화와 같은 사회적 논란을 더욱 중요하게 다루었던 것으로 분석된다(유예림, 백순근, 2016). 또한 2015 개정 교육과정의 정책 형성 및 결정 단계에서 끊임없이 문제로 제기된 초등 교과서 한자 병기, 소프트웨어 교육, 수학교육 내용 적정화 역시 언론 매체를 통해 중요하게 거론되어 왔음을 간접적으로 확인할 수 있다.

<표 IV-2> 뉴스 기사 내 상위 빈출 20개 단어(TF-IDF 기준): 전체

순위	주제어	빈도(TF-IDF)	순위	주제어	빈도(TF-IDF)
1	역사	33.66	11	확정	19.25
2	국정	33.32	12	고시	18.38
3	한자	32.58	13	한글	18.06
4	소프트웨어	26.84	14	교원	17.42
5	수학	22.31	15	영어	16.40
6	정부	21.50	16	한국사	16.35
7	과학	20.64	17	사회	15.85
8	수업	20.63	18	검정	15.45
9	수능	20.46	19	집필진	15.44
10	교육청	19.48	20	디지털	15.37

[그림 IV-2]는 전체 주제어 중 가장 많이 언급된 상위 30개 단어들로 워드 클라우드를 나타낸 것이다. 이때 왼쪽 그림은 TF를 기준으로, 오른쪽 그림은 TF-IDF를 기준으로 나타낸 것이다.



[그림 IV-2] 뉴스 기사의 고빈도 30개 주제어에 대한 워드 클라우드  
(좌: TF, 우: TF-IDF)

## 나. 정책 단계별

국내 언론 매체를 통해 보도된 뉴스 기사에서 가장 많이 출현한 주제어를 정책 단계별로 비교하면 <표 IV-3>, [그림 IV-3]과 같다. 단, 이때에는 모든 뉴스 기사에서 등장하는 상투어의 중요도를 감소시키기 위해 TF-IDF를 사용하여 분석하였으며, 워드 클라우드는 정책 단계별 주제어 중 고빈도 상위 30개 단어를 이용하여 나타냈다.

우선, ‘정책 형성’ 단계에서 가장 많이 등장한 단어는 ‘소프트웨어’였으며, 그 다음으로는 ‘수능’, ‘과학’, ‘기술’, ‘이수단위’ 순으로 높았다. ‘정책 결정’ 단계에서는 정책 형성 단계와 달리 초등 교과서 한자 병기와 역사 교과서에 관련된 주제어(예: ‘한자’, ‘병기’, ‘국정’, ‘역사’)가 중요하게 언급되었다. ‘정책 집행’ 단계에서 가장 많이 등장한 단어는 ‘역사’였고, 그 다음으로는 ‘국정’, ‘소프트웨어’, ‘교육청’, ‘한글’ 순이었다. 정책 집행 단계에서는 정책 결정 단계와 마찬가지로 역사교과서 국정화 관련 주제어

가 주요하게 언급되었으나 ‘집필진’이 등장하는 등 보도 내용이 구체화되었다. 또한 새로운 교수·학습 방법을 모색하는 ‘디지털’과 ‘수업’, 교육과정 운영을 위한 지원체제 구축과 관련된 ‘교육청’, ‘교육지원청’, ‘지원’과 같은 단어들이 새롭게 등장하였다.

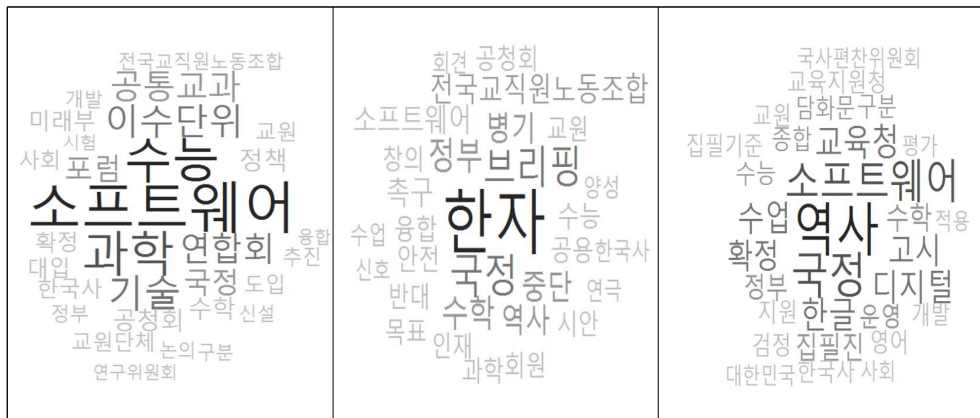
<표 IV-3> 뉴스 기사 내 상위 빈출 20개 단어(TF-IDF 기준): 정책 단계별

순위	정책 형성 단계		정책 결정 단계		정책 집행 단계	
	주제어	빈도	주제어	빈도	주제어	빈도
1	<u>소프트웨어</u>	2.34	한자	20.99	역사	21.52
2	<u>수능</u>	2.04	<u>국정</u>	13.81	<u>국정</u>	18.05
3	과학	1.96	브리핑	12.96	<u>소프트웨어</u>	17.47
4	기술	1.55	정부	11.18	교육청	13.61
5	이수단위	1.49	<u>수학</u>	10.24	한글	13.39
6	연합회	1.38	중단	9.81	디지털	13.36
7	공통교과	1.34	병기	9.75	수업	13.19
8	포럼	1.25	역사	9.57	고시	13.17
9	<u>국정</u>	1.15	전교조	9.41	확정	12.67
10	정책	0.98	<u>소프트웨어</u>	8.95	집필진	11.44
11	공청회	0.97	촉구	8.86	<u>수학</u>	10.72
12	<u>수학</u>	0.96	공용	8.48	정부	10.25
13	미래부	0.94	융합	8.43	운영	9.69
14	확정	0.90	안전	8.37	종합	9.58
15	대입	0.88	<u>수능</u>	8.29	<u>수능</u>	9.53
16	한국사	0.87	반대	8.23	구분	9.36
17	도입	0.84	교원	8.12	답화문	9.33
18	사회	0.83	창의	7.88	교육지원청	9.17
19	교원	0.83	인재	7.83	지원	9.16
20	교원단체	0.82	과학	7.76	영어	9.13

\* 밑줄 친 단어는 정책의 형성·결정·집행 단계에서 모두 출현한 주제어임.

한편 정책의 형성·결정·집행 단계에서 공통적으로 나타난 단어는 ‘소프트웨어’, ‘수능’, ‘국정’, ‘수학’이었다. 우선, ‘소프트웨어’의 경우 정책 형성 단계에서는 가장 많이 언급된 최상위 빈도의 주제어였으나 정책 결정 단계에서는 10위로 하락하였으며, 정책 집행 단계에서는 관련 교원 연수와 선도학교 운영이 활성화됨에 따라 3위로 상승하였다. 그리고 ‘수능’의 경우 2015 개정 교육과정 정책의 도입 배경과 맞닿아 있는 만큼 정책 형성 단계에서 2위로 주요하게 보도되었으나, 정책의 결정·집행 단계에서는 모두 15위로 하락하였다.

이와 반대로 ‘국정’의 경우 정책 형성 단계에서는 9위로 나타났으나, 정책의 결정·집행 단계에서는 모두 2위로 주제어의 빈도가 상승하는 등 언론 매체는 역사교과서 국정화 관련 주제어를 중요하게 보도한 것으로 나타났다. ‘수학’의 경우에는 정책 결정 단계에서 상대적으로 가장 큰 주목을 받았으며, 정책의 형성·집행 단계에서도 빈도수는 소폭 하락하였으나 여전히 주요하게 다뤄지고 있었다.

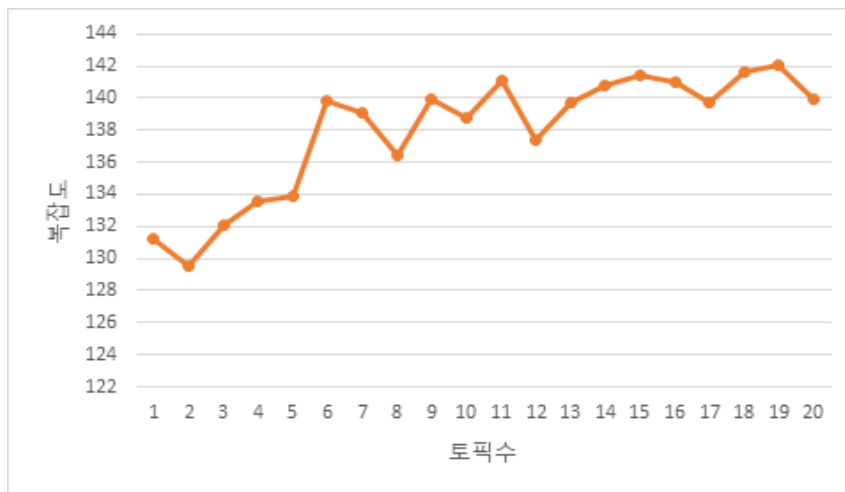


[그림 IV-3] TF-IDF에 따른 정책 단계별 워드 클라우드  
(좌: 정책 형성 단계, 중: 정책 결정 단계, 우: 정책 집행 단계)

### 3. 2015 개정 교육과정 정책의 주요 쟁점 분석

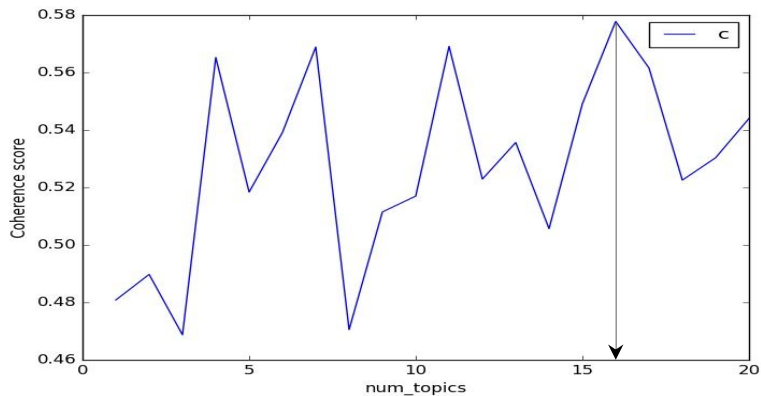
#### 가. 분석 기간 전체

2013년 8월부터 2016년 9월까지 국내 언론 매체를 통해 분석된 2015 개정 교육과정 정책의 쟁점은 총 16개이다. 우선, 토픽의 개수를 추정하기 위해 토픽 모델링의 평가 방법으로 자주 활용되는 복잡도(perplexity)와 토픽 응집성(topic coherence) 지수를 모두 산출하여 보완적으로 사용하고자 하였다. 우선, 분석 전체 기간에 대한 LDA 분석에서 토픽 개수에 따른 복잡도는 [그림 IV-4]와 같다. 복잡도는 작을수록 바람직한데, 토픽을 최대 20개로 설정해 계산한 결과 토픽이 2개일 때 가장 작고 그 이후로는 계속 커지면서 수렴되지 않았다. 이는 분석 데이터의 규모가 커 훈련할 문장이 많아지면, 단어들이 이루는 확률 분포 범위가 넓어져 그 자체로 높은 복잡도를 갖게 되기 때문에 오인식을 일으킨 결과일 가능성이 높다(김현호, 왕지현, 김상훈, 2015).



[그림 IV-4] LDA 모형의 토픽 수에 따른 복잡도: 전체

이에 따라 토픽 모형의 새로운 평가 방법으로 사용되고 있는 토픽 응집성 지수를 함께 산출하였는데, 분석 결과 토픽이 16개일 때 토픽 응집성이 가장 높은 것으로 나타났다([그림 IV-5] 참조). 토픽 응집성이 높다는 것은 각 토픽이 하나의 분명한 요인을 이룬다는 것을 의미하므로, 이처럼 토픽 응집성이 가장 높은 토픽 모형의 토픽 수를 최종 토픽 수로 설정하는 것이 바람직하다. 따라서 이 연구에서도 분석 기간 전체에 대한 토픽의 개수를 16개로 설정하고, 해당 결과가 해석 가능성 측면에서 합리적인지를 살펴보았다.



[그림 IV-5] LDA 모형의 토픽 수에 따른 토픽 응집성: 전체

<표 IV-4>는 분석 기간 전체에 대한 16개의 토픽과 각 토픽별 주요어<sup>16)</sup>, 해당 토픽이 가장 대표적으로 나타난 문서(뉴스 기사)의 수를 나타낸 것이다. 이때 각 문서를 대표하는 토픽은 LDA 분석 결과 도출된 문서별 토픽에 따른 확률값(가중치)을 이용해 지정할 수 있다. 예컨대, [그림 IV-6]과 같은 뉴스 기사(문서)의 경우 분석 데이터 상에서 309번 문

16) LDA 분석 결과 생성된 토픽에 어떤 단어가 할당될 확률은 디리클레 확률 분포를 따른다고 가정되므로 반복 시행에 따라 토픽을 구성하는 주요 단어들은 조금씩 달라질 수 있다. 이에 이 연구에서는 분석 데이터에 대해 LDA 분석을 반복적으로 실시해 보았는데, 그 결과 토픽을 구성하는 주요어들의 확률값(가중치)만 조금씩 변화하였을 뿐 주요어 구성 자체에는 큰 변화가 없었다.

<표 IV-4> LDA 분석 결과: 전체

구분	1	2	3	4	5	6	7	8
	역사교과서 국정화 관련 정치적 대립	역사교과서 국정화 찬반 논란	역사교육 세부 내용	교육과정 운영 지원 체제 구축	고교 공통 과목 운영	교육과정 개정의 필요성	역사교과서 집필진 구성	역사교과서 국정화 고시
1	야권	역사	대한민국	대학	수학	사회	집필진	국정
2	정부	국정	북한	지원	과학	문제	개발	검정
3	국회의원	한국사	정부	추진	사회	융합	국사편찬위	한국사
4	대통령	국민	수립	확대	영어	창의	구성	고시
5	국회	반대	근현대사	개혁	국어	필요	역사	정부
6	박근혜	대한민국	서술	사회	내용	학습	집필	확정
7	여권	사회	일본	확산	학습	내용	집필기준	논란
8	정치	국가	내용	재정	신설	인재	공개	인정
9	정권	교원	역사	직업	공통교과	지식	검토	수정
10	비판	정부	주체사상	강화	수업	해결	확정	전환
문수	87	86	101	60	168	65	85	143
구분	9	10	11	12	13	14	15	16
	수능·대입 제도 개선	교수·학습 방법 개선	초등 교과서 한자 병기	현장교원 의견 수렴	자유학기제 운영	소프트웨어 교육 강화	안전교육 강화	인성·체험 교육 강화
1	수능	교원	한자	공청회	운영	소프트웨어	안전	활동
2	평가	수업	병기	의견	교육청	과학	강화	예술
3	대입	참여	한글	고시	계획	기술	예산	연극
4	시험	개선	학부모	현장	자유학기제	개발	교육청	체육
5	영어	연수	정책	총론	지원	정보	신설	체험
6	제도	방법	방법	논의	확대	디지털	재난	인성
7	출제	환경	반대	국가	마련	미래부	조직	강화
8	학부모	현장	국어	연구	실시	활용	대응	음악
9	부담	교실	연구	지적	협력	필수	업무	교육지원청
10	반영	토론	학습	전교조	시행	진행	정책	문화
문수	103	85	118	139	41	100	63	68

서이고, 토픽에 따른 확률 값을 살펴보면 토픽 13에서 가장 높은 값 (0.6212)을 갖는 것을 확인할 수 있다. 따라서 이 뉴스 기사의 경우 토픽 13에 해당되는 문서로 결정하고, 이를 토픽 13에 대한 문서수 계산시 포함시켰다.

## 내년 중학교 자유학기제 전면 시행(뉴시스, 2015-08-06)

내년부터 중학교에서 한 학기 동안 시험 없이 진로탐색 활동을 하는 ‘**자유학기제**’가 전면 **시행**된다. (중략) ‘교과통합 진로교육’을 활성화해 학습 동기를 부여하고, 현장직업체험 등 다양한 진로 체험활동을 학교별로 2회 이상 **실시**한다. 진로활동 기록 사항을 상급학교에 연계해 진로 상담·지도를 체계적으로 관리할 수 있도록 했다. 주제선택 활동과 동아리 활동은 학습동기 유발과 자율적 문제해결력 함양을 위해 학생 주도의 프로그램을 **운영**한다. (중략)

교육당국은 내년 전면 시행 시 필요한 체험처 4만7300개와 체험 프로그램 9만4600개를 확충한다는 목표다. (중략) 기업의 각종 시설을 개방하고, 산업현장 체험 프로그램 운영도 **확대**한다. (중략) 또 원격영상 진로 멘토링과 진로체험 버스 방방곡곡 프로그램, 찾아가는 진로 체험단을 확대 운영해 놓여준 학생의 체험활동을 **지원**한다. (중략)

교육부는 이달 중으로 자유학기제의 법적 근거를 마련하기 위한 ‘초·중등교육법 시행령’을 개정하고, ‘**2015 개정 교육과정**’에서 자유학기제 편성·운영방안을 제시할 예정이다. 시안에 대한 의견수렴 절차를 거쳐 오는 10월 자유학기제 **시행계획**을 확정할 방침이다.

문서 No.	토픽1	토픽2	토픽3	토픽4	토픽5	토픽6	토픽7	토픽8
309	0.0178	0.0178	0.0221	0.0307	0.0178	0.0652	0.0135	0.0264
	<b>토픽9</b>	<b>토픽10</b>	<b>토픽11</b>	<b>토픽12</b>	<b>토픽13</b>	<b>토픽14</b>	<b>토픽15</b>	<b>토픽16</b>
	0.0221	0.0307	0.0135	0.0221	<b>0.6212</b>	0.0135	0.0393	0.0264

[그림 IV-6] 토픽 13 관련 뉴스 기사의 내용 및 문서-토픽 확률값 예시

또한 LDA 분석 결과 도출된 토픽의 명칭(이하 ‘토픽명’)은 연구자가 직접 부여해야 한다. 예컨대 토픽 13의 경우 ‘운영’, ‘교육청’, ‘계획’, ‘자유학기제’, ‘지원’ 등의 단어로 구성되어 있는데, 단어들의 의미와 맥락을 고려할 때 2015 개정 교육과정 정책의 내용 중 자유학기제와 관련되므로



토픽명을 “자유학기제 운영”으로 설정하였다. 그리고 교육학 전공 박사 및 박사과정 대학원생 총 13명을 대상으로 전문가 협의회를 개최하여, 이러한 토픽명의 타당도를 검증하였다.

나머지 토픽에 대해서도 같은 방식으로 명칭을 부여한 결과, 2015 개정 교육과정 정책의 쟁점은 전체 분석 기간 동안 ① 역사교과서 국정화 관련 정치적 대립, ② 역사교과서 국정화 찬반 논란, ③ 역사교육 세부 내용, ④ 교육과정 운영 지원체제 구축, ⑤ 고교 공통 과목 운영, ⑥ 교육과정 개정의 필요성, ⑦ 역사교과서 집필진 구성, ⑧ 역사교과서 국정화 고시, ⑨ 수능·대입제도 개선, ⑩ 교수·학습 방법 개선, ⑪ 초등 교과서 한자 병기, ⑫ 현장교원 의견 수렴, ⑬ 자유학기제 운영, ⑭ 소프트웨어 교육 강화, ⑮ 안전교육 강화, ⑯ 체험·인성교육 강화의 16개로 도출되었다 (토픽별 관련 뉴스 기사의 예시는 <부록 4> 참고).

그런데 도출된 토픽들을 살펴보면 의미상 연결되는 토픽들이 있음을 추론할 수 있다. 특히 ‘역사교과서 국정화’와 관련해서는 4개의 쟁점이 포함되어 있는데, 일견 연관되어 있는 것으로 판단되는 이들 쟁점이 실제로 유사성을 갖는지, 또는 그 밖의 다른 쟁점들 간에 연구자가 파악하지 못한 관련성은 없는지 등 토픽들 간의 구조적 관계에 대한 추가적 고찰이 필요하다. 이를 위해 이 연구에서는 토픽들을 ‘시각화’함으로써 토픽들 간 상호 연관성 여부를 확인하고자 하였다.

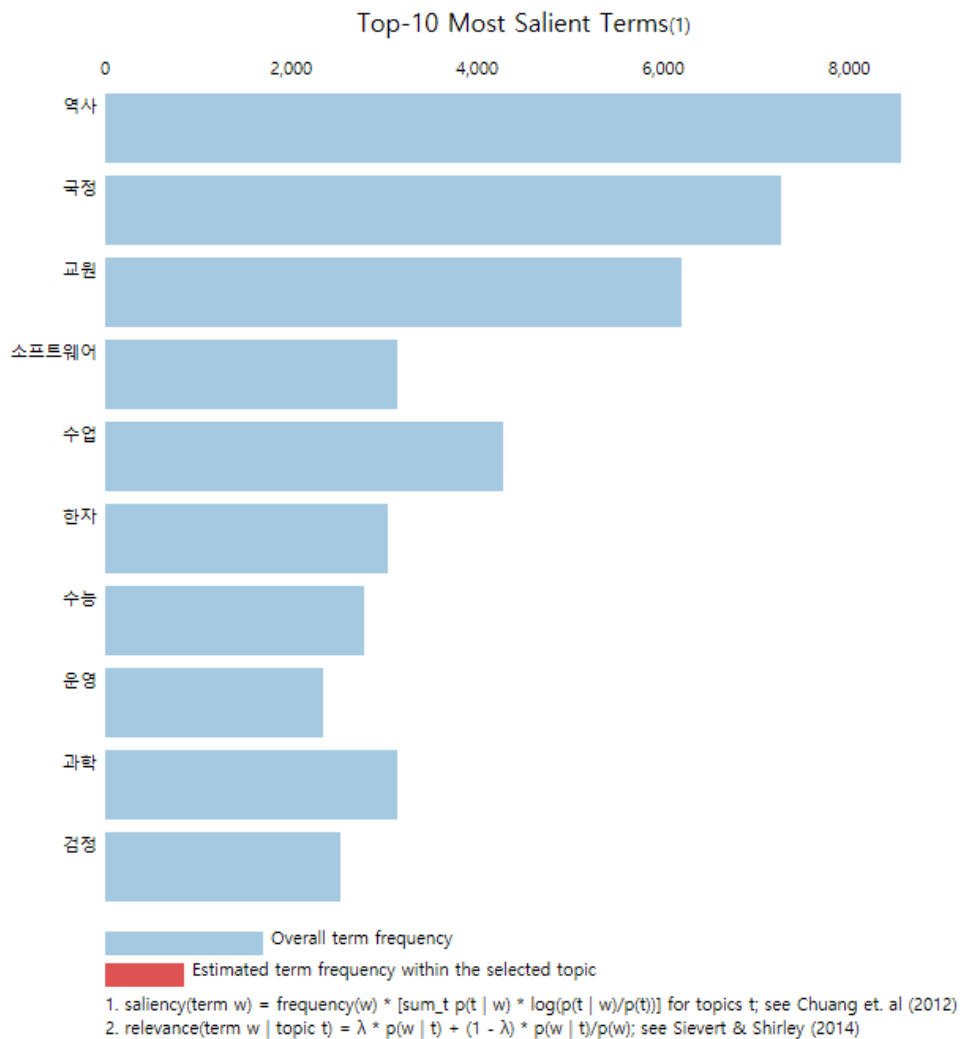
[그림 IV-7]과 [그림 IV-8]은 Python의 ‘pyLDAvis’ 라이브러리를 활용하여 2015 개정 교육과정 정책 관련 토픽들의 관계를 시각화한 첫 화면이다. 우선, 첫 화면의 오른쪽 그림인 [그림 IV-7]에서 막대 그래프는 전체 분석 데이터에서 가장 중요한 단어 10개를 제시한 것으로, 워드 클라우드와 유사한 기능을 갖는다고 볼 수 있다. 즉, 이 연구에서는 ‘역사’, ‘국정’, ‘교원’, ‘소프트웨어’, ‘수업’, ‘한자’, ‘수능’, ‘운영’, ‘과학’, ‘검정’이 가장 중요한 10개 단어로 분석되었으며, 이를 통해 전체 분석 기간 동안 “역사교과서 국정화”, “소프트웨어 교육”, “초등 교과서 한자 병기”, “수능” 관련 쟁점이 형성되었음을 재확인할 수 있다.

시각화 결과 첫 화면의 왼쪽 그림인 [그림 IV-8]의 사사분면 그래프는 LDA 분석 결과 도출된 16개 토픽의 구조적 관계를 나타낸 것으로, 원에 표기되어 있는 숫자는 <표 IV-4>에 제시된 토픽들의 번호와 동일하다. 이를 통해 토픽2(“역사교과서 국정화 찬반 논란”)와 토픽8(“역사교과서 국정화 고시”), 토픽 5(“고교 공통 과목 운영”)와 토픽 14(“소프트웨어 교육 강화”), 토픽 7(“역사교과서 집필진 구성”)과 토픽 12(“현장교원 의견 수렴”), 토픽 9(“수능·대입제도 개선”)와 토픽 10(“교수·학습 방법 개선”)이 밀접한 관련을 맺고 있음을 확인할 수 있다. 예컨대, ‘토픽2’와 ‘토픽8’의 경우 둘 다 주요어로 ‘국정’과 ‘한국사’를 포함한다는 점에서 유사성의 근거를 찾을 수 있다. 다만 이때 같은 단어라도 각 토픽별로 단어의 중요도는 서로 다를 수 있고, 기타 다른 단어들의 조합을 통해 두 토픽의 의미를 재탐색할 수도 있다. 이와 관련해서는 ‘토픽 9’와 ‘토픽 10’에 대한 비교·분석을 통해 설명하고자 한다.

[그림 IV-9]는 토픽 9와 토픽 10에 대한 시각화 결과를 각각 나타낸 것으로 각 그림의 왼쪽을 보면 해당 토픽 번호가 기재된 원이 붉은색(검은색)으로 표시된 것을 확인할 수 있다. 그리고 그림의 오른쪽에는 각 토픽을 구성하고 있는 상위 주요어 10개가 나열되어 있는데, 파란색(회색) 막대그래프는 분석데이터 전체에서의 빈도, 붉은색(검은색)의 경우 해당 토픽 내에서의 빈도를 막대의 길이로 나타낸 것이다. 이때 전체 분석 데이터 중 특정 토픽에서 출현한 빈도의 비율이 높을수록 토픽에 대한 해당 단어의 기여도가 크다고 할 수 있다. 즉, 전체 빈도와 특정 토픽 내 빈도수가 크게 차이나지 않는 단어들은 해당 토픽에 집중적으로 출현한 것이므로 토픽 해석시 유용하다.

예컨대 ‘토픽 9’의 경우 ‘수능’, ‘대입’, ‘시험’, ‘출제’가, ‘토픽 10’의 경우 ‘연수’, ‘환경’, ‘교실’, ‘토론’이 이러한 단어에 해당된다. 즉, 토픽 9와 토픽 10은 각각 “수능·대입제도 개선”, “교수·학습 방법 개선” 관련 쟁점으로 별도의 분야를 다루고 있는 것처럼 보이나, 토픽 9가 넓은 의미에서는 평가 방법과 관련 있고 평가는 교수·학습과 연계되어 이루어지므로

두 토픽 간 연관성이 탐지된 것으로 보인다. 이처럼 토픽들 간의 관계를 시각화하면 그 유사성과 차별성을 함께 검토할 수 있기 때문에, 시각화 작업은 2015 개정 교육과정 정책의 쟁점을 다각도로 해석할 수 있는 기회를 제공한다는 점에서 중요하다고 볼 수 있다.

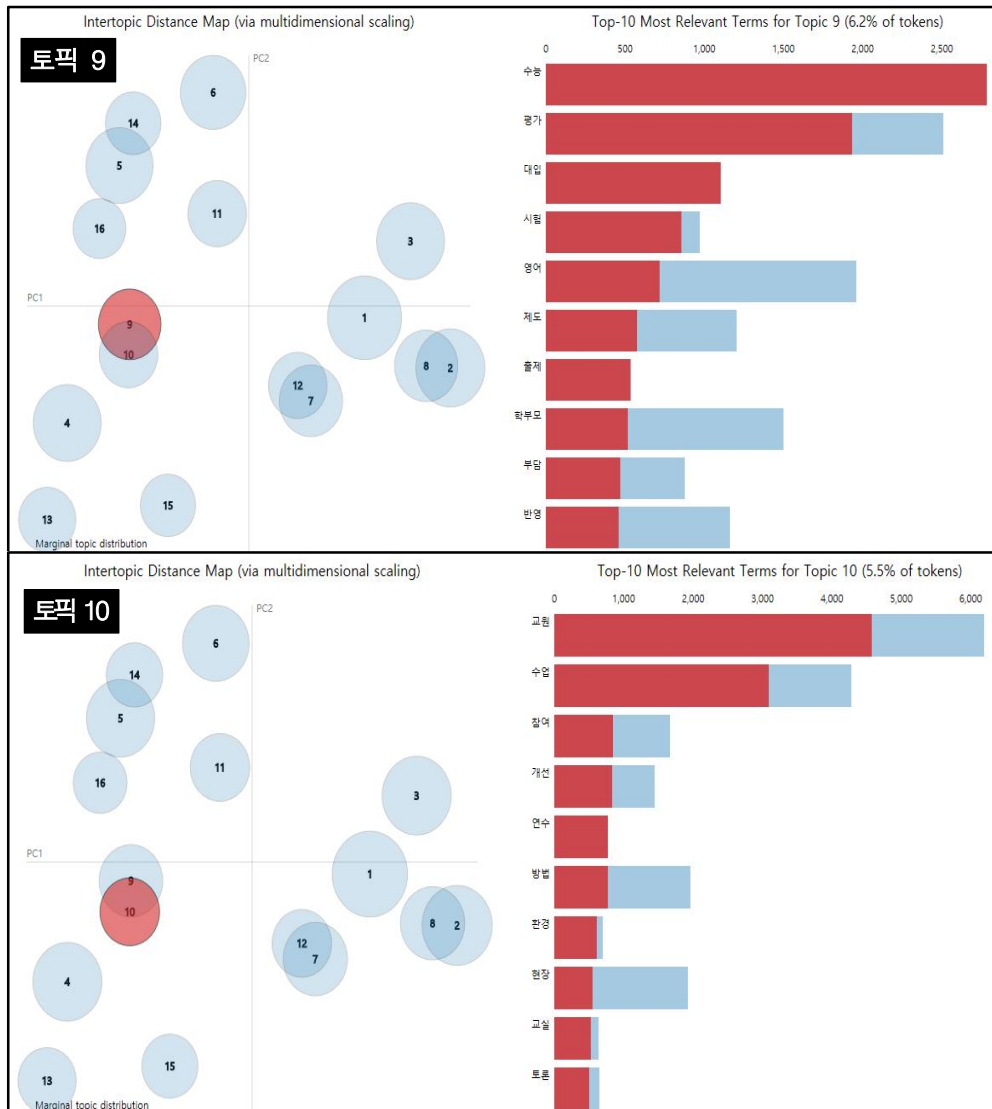


[그림 IV-7] pyLDavis를 활용한 LDA 모형의 시각화 결과(1)



[그림 IV-8] pyLDavis를 활용한 LDA 모형의 시각화 결과(2)

1. 역사교과서 국정화 관련 정치적 대립, 2. 역사교과서 국정화 찬반 논란, 3. 역사교육 세부 내용, 4. 교육과정 운영 지원체제 구축, 5. 고교 공통 과목 운영, 6. 교육과정 개정의 필요성, 7. 역사교과서 집필진 구성, 8. 역사교과서 국정화 고시, 9. 수능·대입제도 개선, 10. 교수·학습 방법 개선, 11. 초등 교과서 한자 병기, 12. 현장교원 의견 수렴, 13. 자유학기제 운영, 14. 소프트웨어 교육 강화, 15. 안전교육 강화, 16. 인성·체험교육 강화



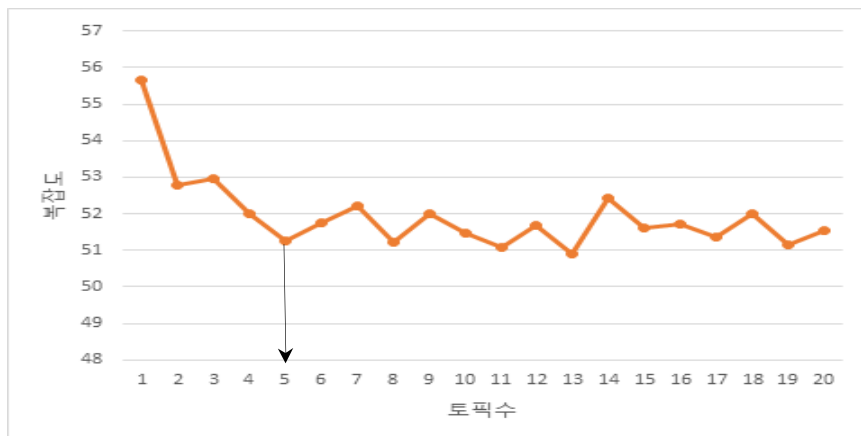
[그림 IV-9] LDA 모형의 시각화 결과: 토픽 9와 토픽10

1. 역사교과서 국정화 관련 정치적 대립, 2. 역사교과서 국정화 찬반 논란, 3. 역사교육 세부 내용, 4. 교육과정 운영 지원체제 구축, 5. 고교 공통 과목 운영, 6. 교육과정 개정의 필요성, 7. 역사교과서 집필진 구성, 8. 역사교과서 국정화 고시, 9. 수능·대입제도 개선, 10. 교수·학습 방법 개선, 11. 초등 교과서 한자 병기, 12. 현장교원 의견 수렴, 13. 자유학기제 운영, 14. 소프트웨어 교육 강화, 15. 안전교육 강화, 16. 인성·체험교육 강화

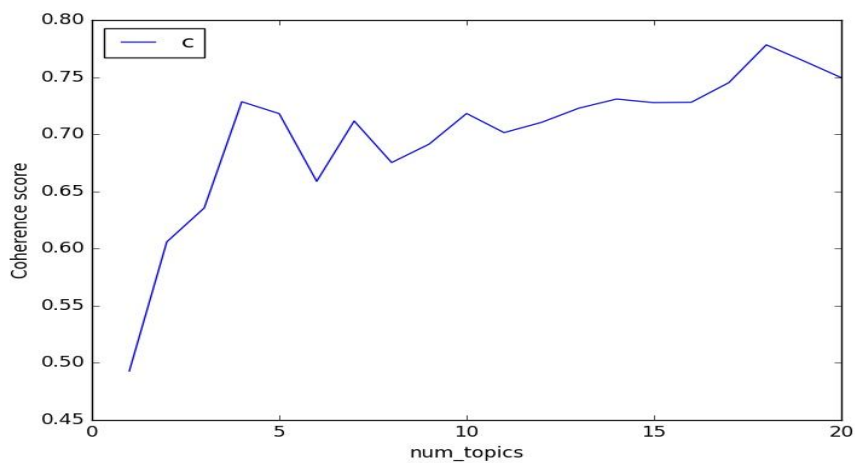
## 나. 정책 단계별

### 1) 정책 형성 단계

2015 개정 교육과정 정책의 형성 단계에서도 복잡도와 토픽 응집성을 고려하여 토픽의 개수를 추출하였다([그림 IV-10], [그림 IV-11] 참조).



[그림 IV-10] 토픽 수에 따른 LDA 모형의 복잡도: 정책 형성 단계



[그림 IV-11] 토픽 수에 따른 LDA 모형의 토픽 응집성: 정책 형성 단계

정책 형성 단계의 경우 뉴스 기사가 67건으로 전체 분석 데이터인 1,512건에 비해 적은 편이다. 때문에 이 단계에서는 토픽 수가 증가함에 따라 복잡도가 낮아지며 특정 값으로 수렴되는 결과를 보였다. 즉, [그림 IV-10]을 살펴보면, 복잡도는 토픽수가 최초 5개일 때의 값으로 수렴되므로 LDA 모형의 복잡도를 기준으로 토픽의 개수를 설정하면 5개라고 할 수 있다. 다만, 정책 형성 단계에서 토픽 응집성 지수는 토픽 수가 커짐에 따라 계속 증가하는 양상을 보여 의미를 찾을 수 없었다.

LDA 분석 결과, 2015 개정 교육과정 정책의 형성 단계에서는 ① 과학 교육 축소 우려, ② 고교 공통 과목 도입, ③ 수능·대입제도 개선, ④ 소프트웨어 교육 확대, ⑤ 역사교과서 국정화의 총 5개 토픽이 도출되었다(<표 IV-5> 참조). 문서수를 기준으로 보면 ‘과학 교육 축소 우려’와 ‘고교 공통 과목 도입’ 관련 쟁점이 가장 많이 다루어졌으나, 정책 형성 단계의 총 보도 건수가 작고 쟁점 간 문서수의 차이가 크지 않기 때문에 이러한 특정 쟁점이 해당 단계에서 두드러졌다고 보기는 다소 어렵다.

<표 IV-5> LDA 분석 결과: 정책 형성 단계

구분	1	2	3	4	5
	과학교육 축소 우려	고교 공통 과목 도입	수능·대입제도 개선	소프트웨어 교육 확대	역사교과서 국정화
1	과학	사회	수능	소프트웨어	교원
2	기술	이수단위	도입	정책	문제
3	수학	과학	대입	미래부	검정
4	학계	필수	정부	추진	국정
5	논의	영어	평가	확대	내용
6	의견	수업	확정	정보	수정
7	연합회	국어	융합	개발	연구
8	요구	공통교과	개발	지원	전환
9	국가	내용	선택	인재	현장
10	포럼	신설	한국사	강화	고시
문서수	16	16	12	12	11

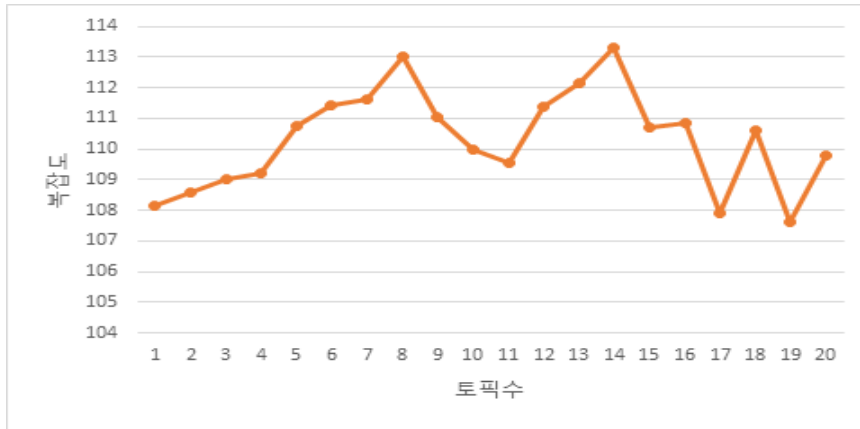
‘토픽1’의 경우 국가교육과정 개정연구위원회가 2014년 5월 발표한 개정 초안의 “과학 교과” 필수 이수단위가 2009 개정 교육과정보다 축소된 것에 대해 과학계가 심하게 반발한 이른바 ‘이과 말살 교육과정’ 담론(강지영, 소경희, 2016)과 연결된다. 또한 ‘토픽 3’는 2015 개정 교육과정 정책이 「2017학년도 대입 확정안」을 통해 논의되기 시작한 만큼, 개정 교육과정이 “수능·대입제도 개선”과 연계되어야 함을 강조하는 쟁점이라고 할 수 있다.

한편 교육부는 2014년 7월 23일 「SW교육 관련 교육과정 개편 방안」을 발표하고, 9월 초 미래부와 함께 과학교육 강화와 소프트웨어 교육 확대 방안을 논의함에 따라 토픽 4인 “소프트웨어 교육 확대”에 대한 논의가 함께 이루어졌다. 이외에도 2015 개정 교육과정 총론 주요사항이 발표되기 이전, 공청회를 통해 공개된 “고교 공통 과목(의) 도입” 내용을 소개하거나 “역사교과서 국정화” 가능성에 대해 언급하고 우려하는 뉴스 기사들이 나타나기 시작, 토픽 2와 토픽5를 형성하였다.

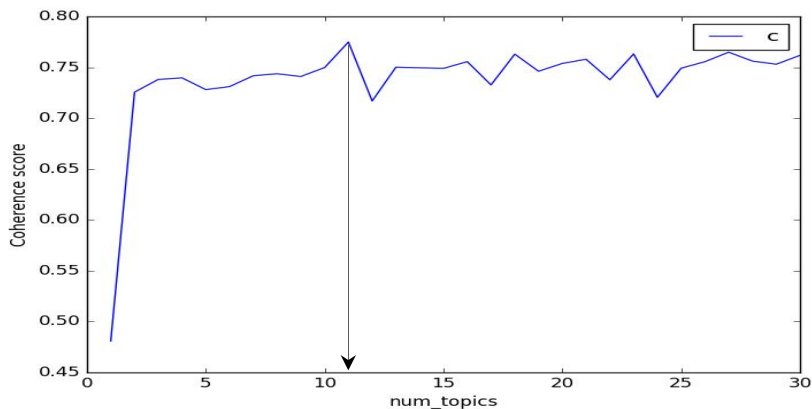
## 2) 정책 결정 단계

2015 개정 교육과정 정책의 ‘결정’ 단계에서는 총 11개의 토픽이 추출되었으며(<표 IV-6> 참조), 이 역시 [그림 IV-12], [그림 IV-13]과 같은 복잡도와 토픽 응집성 지수를 근거로 산출하였다. 정책 형성 단계에서는 복잡도가 토픽 수를 결정하는 근거 자료로 활용되었는데, 정책 결정 단계에서 분석 데이터의 수가 620건으로 증가하자 다시 수렴 실패의 문제가 발생하였다. 대신 분석 데이터의 수가 적었던 정책 형성 단계에서는 제대로 기능하지 못했던 토픽 응집성 지수가 정책 결정 단계에서는 충분한 사례수가 확보됨에 따라 일정 수렴 값 범위 내에서 등락을 반복하였다. 토픽 응집성 지수에 따르면 정책 결정 단계의 토픽은 총 11개인데, 이러한 토픽 분류가 유의미한지를 LDA 분석 결과를 통해 확인하였다.





[그림 IV-12] 토픽 수에 따른 LDA 모형의 복잡도: 정책 결정 단계



[그림 IV-13] 토픽 수에 따른 LDA 모형의 토픽 응집성: 정책 결정 단계

분석 결과, 2015 개정 교육과정 정책의 결정 단계에서는 ① 고교 공통 과목 운영, ② 수학교육 내용 적정화, ③ 초등 교과서 한자 병기, ④ 교육과정 개정의 필요성, ⑤ 역사교과서 국정화, ⑥ 현장교원 의견 수렴, ⑦ 수능·대입제도 개선, ⑧ 역사교육 세부 내용, ⑨ 안전교육 강화, ⑩ 역사교과서 국정화 관련 정치적 대립, ⑪ 소프트웨어 교육 강화의 총 11개 토픽이 추출되는 등 정책 형성 단계보다 더욱 다양한 쟁점이 나타났다. 이는 정책 형성 단계에서 문·이과 융합교육이라는 취지 정도만이 강

조되었던 반면, 정책 결정 단계에서는 총론 주요사항이 확정·발표됨에 따라 교과별로 역점을 두어야 할 방안들이 구체적으로 제시되고 그에 대한 논의가 여러 과정을 거쳐 이어졌기 때문이다(유예림, 백순근, 2016).

<표 IV-6> LDA 분석 결과: 정책 결정 단계

구분	1	2	3	4	5	6
	고교 공통 과목 운영	수학교육 내용 적정화	초등 교과서 한자 병기	교육과정 개정의 필요성	역사교과서 국정화 찬반 논란	현장교원 의견 수렴
1	사회	수학	한자	창의	국정	교원
2	과학	내용	병기	융합	역사	현장
3	신설	학습	한글	사회	한국사	의견
4	국어	학습량	방법	과학	검정	고시
5	공통교과	시안	반대	인재	반대	공청회
6	진로	학습부담	정책	기술	개발	적용
7	강화	사항	공청회	필요	정부	국가
8	이수단위	국어	제시	개발	인정	총론
9	선택교과	사고	활성	변화	검토	논의
10	적용	성취기준	연구	환경	학계	추진
문서수	86	72	71	67	66	57

구분	7	8	9	10	11
	수능·대입 제도 개선	역사교육 세부 내용	안전교육 강화	역사교과서 국정화 관련 정치적 대립	소프트웨어 교육 강화
1	수능	역사	안전	정부	수업
2	평가	대한민국	계획	야권	소프트웨어
3	대입	국가	지원	국회의원	시안
4	영어	일본	추진	박근혜	운영
5	문제	근현대사	대학	국회	체육
6	시험	정부	재난	대통령	활동
7	제도	비중	개혁	국감	자유학기제
8	대학	내용	확대	규정	정보
9	도입	축소	현장	여권	교원
10	공교육	수립	정부	진행	체험
문서수	50	39	38	37	37

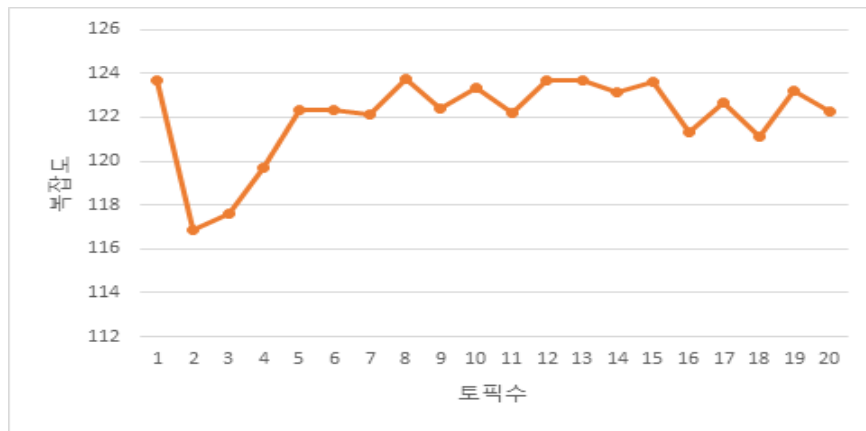
2015 개정 교육과정 정책의 결정 단계에서 가장 쟁점이 된 토픽은 “역사교과서 국정화 찬반 논란”(토픽 5)이었다. 역사교과서 국정화 문제는 국가적 논쟁으로 확대되어 학계와 일반 대중뿐 아니라 여·야의 “역사교과서 국정화 관련 정치적 대립”(토픽 10)이 연일 보도되었다. 또한 이러한 논란의 중심에는 “역사교육 세부 내용”(토픽 8)으로서 ‘대한민국 정부 수립’을 어떻게 서술할 것인지, ‘근현대사 비중’은 어느 정도가 적절한지, ‘일본’과 관련해 독도나 위안부 문제는 얼마나 다룰 것인지 등이 문제로 제기되었다. 그리고 정책 형성 단계와 마찬가지로 향후 도입될 “고교 공통 과목 운영”(토픽 1) 사항을 소개하는 기사가 많이 보도되었다.

또한 각종 시민단체를 중심으로 “수학교육 내용 적정화”(토픽 2)와 “초등 교과서 한자 병기”(토픽 3) 관련 비판도 많았다. 두 쟁점 모두 사교육 부담과 연관되어 논의가 진행되었고, 초등 교과서 한자 병기의 경우 공청회를 통해 나타난 비판 의견이 뉴스 기사를 통해 다수 보도된 이후 해당 방안이 유보되기도 하였다. 특히 수학교육 내용 적정화와 관련해서는 시민단체인 ‘사교육걱정없는세상’이 우리나라 학생들의 수학 학습량이 과다하여 학습부담을 가중시킨다는 비판을 지속적으로 제기해 새로운 담론을 형성한 것이다.

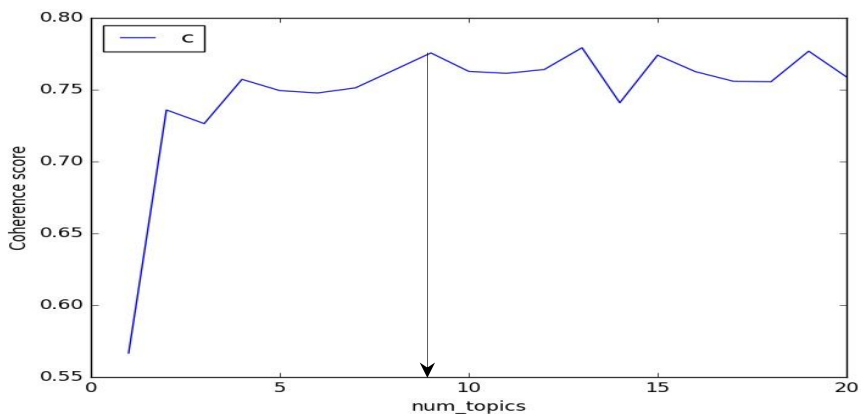
이러한 논란 속에 2015 개정 교육과정의 총론과 교과 교육과정을 최종 확정하기 전 “현장교원 의견 수렴”(토픽 6)을 위한 공청회도 끊임없이 개최되었다. 특히 “소프트웨어 교육 강화”(토픽 11), “안전교육 강화”(토픽 9) 부분은 최근 우리나라가 처한 시대적 상황에 비추어 주목을 받았다. 한편으로는 교육과정과 함께 다루어져야 할 “수능·대입제도 개선”(토픽 7)에 대한 논의가 제대로 이뤄지지 못하고 있음을 비판하거나, 창의·융합형 인재 양성이라는 국가적, 사회적 요구에 부응하기 위해 2015 “개정” 교육과정 정책의 도입 “필요성”(토픽 4)을 강조하는 전문가의 사설이나 칼럼 보도가 수차례 등장하기도 하였다.

### 3) 정책 집행 단계

2015 개정 교육과정 정책의 ‘집행’ 단계에서의 토픽 개수 역시 복잡도와 토픽 응집성 지수, 토픽의 해석 가능성을 근거로 산출되었다([그림 IV-14], [그림 IV-15] 참조). 분석 결과, 정책 결정 단계와 마찬가지로 복잡도는 제대로 수렴되지 않는 반면, 토픽 응집성 지수는 토픽이 9개일 때 최댓값을 보여 이를 기준으로 토픽의 개수와 토픽별 주요어가 타당하게 구성되어 있는지를 살펴보았다.



[그림 IV-14] 토픽 수에 따른 LDA 모형의 복잡도: 정책 집행 단계



[그림 IV-15] 토픽 수에 따른 LDA 모형의 토픽 응집성: 정책 집행 단계

2015 개정 교육과정 정책의 집행 단계에서는 ① 교원의 전문성 개발, ② 공교육 정상화, ③ 역사교과서 국정화 고시, ④ 역사교육 세부 내용, ⑤ 역사교과서 집필진 구성, ⑥ 교실 수업·평가 개선, ⑦ 교육과정 운영 지원체제 구축, ⑧ 체험·인성교육 강화, ⑨ 역사교과서 국정화 찬반 논란의 총 9개 토픽이 추출되었다(<표 IV-7> 참조).

<표 IV-7> LDA 분석 결과: 정책 집행 단계

구분	1	2	3	4	5	6	7	8	9
	교원의 전문성 개발	공교육 정상화	역사교과서 국정화 고시	역사교육 세부 내용	역사교과서 집필진 구성	교실 수업·평가 개선	교육과정 운영지원 체제 구축	체험·인성 교육 강화	역사교과서 국정화 찬반 논란
1	교원	적용	역사	정부	집필진	수능	지원	운영	국정
2	수업	학부모	검정	대한민국	국사편찬위	사회	강화	계획	역사
3	소프트웨어	교원	국정	북한	역사	문제	예산	교육청	정부
4	연수	평가	고시	수립	집필기준	과학	정책	확대	한국사
5	참여	한글	확정	내용	집필	학습	사회	자유학기제	반대
6	창의	수학	한국사	주체사상	구성	영어	안전	예술	야권
7	교실	정책	발행	서술	개발	평가	교육청	지원	대통령
8	진행	수업	사회	기술	공개	디지털	조직	수업	박근혜
9	미래	공교육	교원	문제	확정	필요	대학	활동	국회
10	방법	지적	국민	총리	검토	내용	환경	연극	정치
문서 수	157	117	110	92	86	80	79	77	27

정책 집행 단계에서는 정책 결정 단계와 마찬가지로 역사교과서 국정화 관련 사항이 가장 큰 쟁점으로 기능하였다. 특히 2015년 10월 12일 관련 행정예고가 실시되고 11월 3일에 초·중등 교과용도서 국·검·인정 구분이 “확정·고시”(토픽 3)됨에 따라 더욱 쟁점화되었다. 이와 함께 “역사교육 세부 내용”(토픽 4), “역사교과서 국정화 찬반 논란”(토픽 9)이

지속적으로 논의되었다. 또한 정책 집행 단계는 향후 정책 실행을 준비하는 단계인 만큼 교원 연수를 통한 “교원의 전문성 개발”(토픽 1), “역사교과서 집필진 구성”(토픽 5), “교실 수업·평가 개선”(토픽 6), 각종 “교육과정 운영 지원체제 구축”(토픽 7)이 필요하다는 논의가 새롭게 등장하였다. 특히 국정 역사교과서 집필진을 투명하게 공개하라는 요구가 많았는데 이는 수용되지 않아 논란이 지속되었다. 그리고 교원의 전문성 개발과 관련해서는 미래 창의·융합교육을 실천하기 위한 학생 참여 중심의 교실 수업 방법이나 신설·강화되는 소프트웨어 교육과 관련된 교원연수가 많이 요구된다는 내용이 다수 보도되었다.

아울러 ‘사교육걱정없는세상’은 우리나라 초등학교 1, 2학년 학생들의 한글 및 수학 선행학습 실태를 보고하며, “공교육”이 선행학습의 부담을 덜어줄 수 있도록 “정상화”(토픽 2)되어야 한다고 주장하였다. 그리고 비슷한 시기에 서울특별시교육청이 이러한 문제를 해소하기 위한 정책을 발표하여 관련 쟁점이 형성되었다. 또한 창의·융합형 인재 양성을 강조하는 2015 개정 교육과정 정책은 우리나라 창의성 교육이 융합, 도전, 상상력 등의 인성적 특성도 함께 강조(소경희, 허예지, 이종아, 2016)하는 것과 마찬가지로, ‘자유학기제’를 중심으로 연극을 비롯한 각종 “체험·인성교육(을) 강화”(토픽 8)할 수 있도록 지원해야 함을 강조하여 또 하나의 쟁점을 형성하였다.

한편, 변화하는 교육 환경에 부응하기 위해 새롭게 도입된 “디지털 교과서” 경우 2016년 8월 29일 관련 국·검정 구분 고시가 발표되었다. 그러나 해당 정책의 발표 시점이 이 연구의 데이터 수집 종료 시점인 2016년 9월 30일과 시간적으로 큰 차이가 없기 때문에 이 연구에서는 쟁점으로 드러나지 않았다.

## V. 요약 및 논의

### 1. 요약

각종 정보가 실시간으로 생성되고 유통되는 ‘빅데이터’(big data) 시대에서는 빅데이터 분석 기법을 활용하여 주요 교육정책에 대한 언론보도 내용을 분석하는 것이 중요하다. 이는 교육정책에 대한 신속한 의사 결정과 즉각적 대응에 기여할 수 있기 때문이다. 따라서 이 연구는 우리나라 초·중등 교육정책의 핵심 중 하나인 2015 개정 교육과정 정책에 대한 언론보도 내용을 빅데이터 분석 기법을 활용하여 분석하는 것을 목적으로 한다. 이 연구에서 알아보고자 하는 주요 연구 문제는 다음과 같다.

첫째, ‘2015 개정 교육과정’ 정책에 대한 언론 매체의 관심도는 정책의 형성·결정·집행 단계에 따라 변화하는가?

둘째, ‘2015 개정 교육과정’ 정책에 대한 뉴스 기사의 주제어는 무엇이며, 이는 정책의 형성·결정·집행 단계에 따라 변화하는가?

셋째, ‘2015 개정 교육과정’ 정책의 주요 쟁점은 무엇이며, 이는 정책의 형성·결정·집행 단계에 따라 변화하는가?

이러한 연구 문제를 해결하기 위해 우선, 국내 최대 포털이자 매체합산 여론영향력 점유율이 1위(윤영철 외, 2015)인 ‘네이버’ 뉴스 홈페이지(<http://news.naver.com>)를 활용하여 신문·잡지, TV, 인터넷 뉴스, 기타 청와대, 교육부의 정책브리핑 및 정당 보도자료가 포함된 2015 개정 교육과정 정책 관련 뉴스 기사 총 1,598건을 프로그래밍 언어인 파이썬(Python) 기반의 웹 스크래핑(web-scraping)을 통해 자체적으로 수집하였다. 그리고 분석 목적에 부합하지 않는 사교육 업체 및 TV 프로그램 홍보 기사 등을 제외한 총 1,512건의 뉴스 기사를 대상으로, 전처리

(preprocessing) 과정과 한국어 형태소 분석기 KOMORAN 2.0을 이용한 형태소 분석(morphological analysis)을 거쳐 자동화된 텍스트 분석이 가능한 형태로 데이터를 클리닝하였다. 이 연구의 주요 분석 과정 및 결과는 다음과 같다.

첫째, 버즈(buzz) 분석을 실시한 결과, 2015 개정 교육과정 정책에 대한 국내 언론 매체의 관심도는 정책의 형성·결정·집행 단계에 따라 변화하였다. 즉, 2015 개정 교육과정 정책의 형성 단계에서는 관련 보도가 거의 이뤄지지 않았고, 해당 정책의 내용(2015 개정 교육과정의 총론 주요 사항, 총론, 교과 교육과정)과 국가적으로 역사교과서 국정화 논란이 극대화된 시점에서만 관련 내용을 집중적으로 보도하는 양상이 나타났다.

둘째, 국내 언론 매체의 뉴스 기사에 등장한 2015 개정 교육과정 정책의 주제어(keyword)는 단순 단어 빈도(Term Frequency: TF)를 기준으로 분석한 경우 ‘교과서’, ‘교육부’, ‘역사’, ‘학생’, ‘국정’, ‘교과’, ‘고등학교’, ‘교원’, ‘초등학교’, ‘중학교’와 같은 일반적인 단어가 가장 많이 언급되었다. 반면, 대부분의 뉴스 기사에 포함되어 있는 단어들의 중요도를 약화시키고 설명력이 높은 또 다른 주제어를 부각시키기 위해 단어 빈도-역문서 빈도(TF-Inverse Document Frequency: TF-IDF)를 활용하여 분석을 실시한 경우에는 ‘역사’, ‘국정’, ‘한자’, ‘소프트웨어’, ‘수학’, ‘정부’, ‘과학’, ‘수업’, ‘수능’, ‘교육청’ 순으로 빈도가 높았다.

셋째, 국내 언론 매체의 뉴스 기사에 등장한 2015 개정 교육과정 정책의 주제어는 정책의 형성·결정·집행 단계에 따라 변화하는 것으로 나타났다. 즉, 2015 개정 교육과정 정책의 형성 단계에서는 주제어의 빈도(TF-IDF 기준)가 ‘소프트웨어’, ‘수능’, ‘과학’, ‘기술’, ‘이수단위’ 순으로 높게 나타난 반면, 정책의 결정 단계에서는 ‘한자’, ‘국정’, ‘브리핑’, ‘정부’, ‘수학’이, 정책의 집행 단계에서는 ‘역사’, ‘국정’, ‘소프트웨어’, ‘교육청’, ‘한글’이 최상위 빈도를 나타냈다.

넷째, 국내 언론 매체를 통해 나타난 2015 개정 교육과정 정책의 쟁점을 잠재 디리클레 할당(Latent Dirichlet Allocation: LDA) 기법을 활용



하여 도출한 결과 ① 고교 공통 과목 운영, ② 역사교과서 국정화 고시, ③ 현장교원 의견 수렴, ④ 초등 교과서 한자 병기, ⑤ 수능·대입제도 개선, ⑥ 역사교육 세부 내용, ⑦ 소프트웨어 교육 강화, ⑧ 역사교과서 국정화 관련 정치적 대립, ⑨ 역사교과서 국정화 찬반 논란, ⑩ 역사교과서 집필진 구성, ⑪ 교수·학습 방법 개선, ⑫ 체험·인성교육 강화, ⑬ 교육과정 개정의 필요성, ⑭ 안전교육 강화, ⑮ 교육과정 지원체제 구축, ⑯ 자유학기제 운영(이상 문서수 기준 정렬)의 16개로 나타났다.

다섯 째, 국내 언론 매체의 뉴스 기사에서 나타난 2015 개정 교육과정 정책의 주요 쟁점은 정책의 형성·결정·집행 단계에 따라 변화하는 것으로 나타났다. 우선, 2015 개정 교육과정 정책의 ‘형성’ 단계에서는 ‘과학 교육 축소 우려’, 창의·융합형 인재를 양성·평가하기 위한 ‘고교 공통 과목 도입’ 및 ‘수능·대입제도 개선’, ‘소프트웨어 교육 확대’ 등의 5개 주요 쟁점이 나타났다. 그리고 정책의 ‘결정’ 단계에서는 역사교과서 국정화와 관련된 쟁점이 가장 부각되었으며, 이는 ‘역사교과서 국정화 찬반 논란’, ‘역사교육 세부 내용’, ‘역사교과서 국정화 관련 정치적 대립’의 하위 쟁점으로 나타났다. 아울러 2015 개정 교육과정에 새롭게 도입되는 ‘고교 공통 과목(의) 운영’, ‘수학교육 내용 적정화’, ‘초등 교과서 한자 병기’ 등의 11개 주요 쟁점이 나타났다.

정책의 ‘집행’ 단계에서도 ‘역사교과서 국정화 고시’와 같은 역사교과서 국정화 관련 논란이 여전히 쟁점으로 기능하였다. 그리고 변화가 많은 2015 개정 교육과정 정책을 현장에 안착시키기 위한 노력으로 교원연수를 강조하는 등 ‘교원의 전문성 개발’, ‘공교육 정상화’ 등을 비롯한 총 9개 주요 쟁점이 나타났다. 또한 정책 실행을 준비하는 정책 집행 단계의 특성상 ‘역사교과서 집필진 구성’, ‘교실 수업·평가 개선’, ‘교육과정 운영 지원체제 구축’ 쟁점도 새롭게 등장하였다.

## 2. 논의

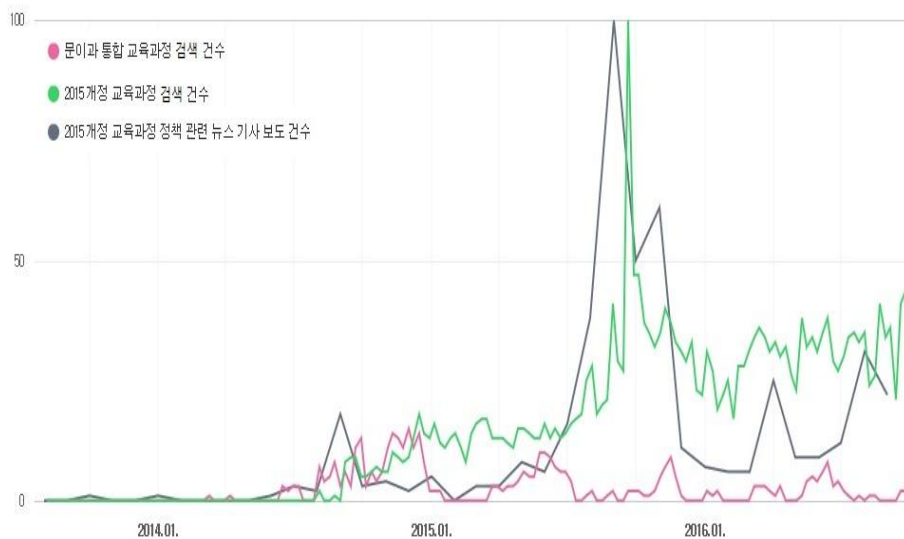
이 연구에서는 빅데이터 분석 기법을 활용하여 우리나라 초·중등 교육 정책의 핵심 중 하나인 2015 개정 교육과정 정책 관련 언론 보도 내용을 분석하였다. 그 결과 몇 가지 논의할 사항이 발견되었다.

첫째, 이 연구에서는 2015 개정 교육과정 정책에 대한 언론 매체의 관심도가 정책의 형성·결정·집행 단계에 따라 변화하는 것으로 나타났다. 즉, 정책 형성 단계에서는 관심도 수준이 낮았으나, 정책이 결정·집행되기 직전에는 관심도가 급증하였으며, 집행 단계에서는 다시 하락하였다. 이와 같은 언론 매체의 관심도 변화가 일반 대중을 비롯한 정책 환경과 어떠한 관련성을 갖는지를 탐색한다면, 정책에 대한 과정 평가(process evaluation) 측면에서 언론보도 분석의 유용성을 확인할 수 있을 것이다.

이를 위해 네이버의 데이터랩(<http://datalab.naver.com>)에서 제공하는 ‘데이터 융합분석’ 서비스를 활용하여 일반 대중들이 네이버 홈페이지의 ‘통합검색’에서 2015 개정 교육과정 정책을 얼마나 검색했는지를 조사해 보았다. 이때 검색어로는 이 연구의 분석 데이터인 뉴스 기사 수집 과정과 마찬가지로 ‘문이과 통합 교육과정’과 ‘2015 개정 교육과정’을 사용하였다. 이러한 과정을 거쳐 산출된 네이버 통합검색 빈도는 검색 포털 내 특정 검색어에 대한 검색 횟수를 주간으로 합산해 기간 내 최대 검색량을 100으로 설정하고, 이 값을 기준으로 다른 시기의 검색 횟수를 상대적으로 산출한 지표다.

[그림 V-1]을 살펴보면, 2015 개정 교육과정 정책에 대한 일반 대중의 관심도 변화 양상은 언론 매체의 관심도 양상과 거의 유사한 것을 확인할 수 있다. 즉, 언론 매체와 일반 대중 모두 정책의 형성 및 결정 초반 단계에서 2015 개정 교육과정 정책에 대한 관심이 상대적으로 낮았다. 이는 2015 개정 교육과정 정책에 대한 일반 대중의 관심도를 제고하고 대중 스스로의 자발적 참여를 유도하기 위한 방안 마련이 필요함을 시사

한다. 즉, 그동안 교육당국의 정책 홍보 활동이나 공청회 등을 통한 의견 수렴은 정책 내용이 일정 수준 이상 결정·집행된 이후에 활발히 이루어져 왔다. 그러나 정책 내용이 타당하게 수립되기 위해서는 정책 형성 단계에서부터 특정 분야의 전문가 집단뿐 아니라 일반 대중도 함께 정책 도입의 필요성 및 발전 방안 등을 모색할 수 있는 장(場)을 마련하고 이에 대중 역시 적극 참여하려는 태도를 갖는 것이 바람직할 것이다.



[그림 V-1] 네이버 통합 검색 빈도와 뉴스 기사 보도 건수의 추이 비교

또한 정책의 형성·결정 단계에서는 대중의 관심도가 언론 매체의 집중적 보도 직후 급증하는 양상을 보였다. 이는 대중의 교육정책에 대한 관심과 이해도를 증진시키는 데 언론 매체의 역할 역시 중요함을 시사한다. 그리고 언론 매체의 관심도와 일반 대중의 관심도 변화 양상이 정확하게 일치하지 않는다는 것은 다양한 이해당사자들의 정책에 대한 관심도를 분석하고자 할 때 언론 매체의 뉴스 기사뿐 아니라, 다양한 이해당사자들의 의견이 직접적으로 표출된 텍스트 데이터(예컨대, SNS 게시물, 공청회 토론자료 등)가 추가적으로 확보되어야 함을 시사하기도 한다.

둘째, 이 연구를 통해 국내 언론 매체는 2015 개정 교육과정 정책에서 강조하는 학생 중심의 융합교육, 핵심역량 함양과 같은 주제보다 ‘역사교과서 국정화’, ‘초등 교과서 한자 병기’ 논란과 같은 쟁점에 더욱 초점을 맞춘 것을 확인할 수 있었다. 언론 매체를 비롯한 이해당사자가 일부 쟁점에만 지나치게 민감하게 반응할 경우 합리적 결정이 저해될 수 있으므로(윤정일 외, 2015) 이와 관련된 논의는 교육정책 연구에서 중요하다고 볼 수 있다. 따라서 관련 논의를 구체화하기 위해 교육부, 교육과정 분야 전문가의 선행연구, 이 연구를 통해 도출된 2015 개정 교육과정 정책의 주요 쟁점을 비교하면 <표 V-1>과 같다.

<표 V-1> 2015 개정 교육과정 정책의 주요 쟁점 비교·분석

교육부	선행연구	이 논문의 분석 결과
<ul style="list-style-type: none"> <li>· 인문·사회·과학기술 기초 소양 함양</li> <li>· 학습 내용 구조화, 학습량 적정화를 통한 학습의 질 개선</li> <li>· 학생 참여형 수업 활성화를 통한 자기주도적 학습 능력 및 학습의 즐거움 배양</li> <li>· 교과외 교육 목표, 교육 내용, 교수·학습 및 평가의 일관성 강화</li> <li>· 산업사회가 필요로 하는 기초 역량과 직무 능력 함양</li> </ul>	<ul style="list-style-type: none"> <li>· 융합형 인재 양성</li> <li>· 역량중심 교육과정</li> <li>· 학습부담과 내용량 감축</li> <li>· 총론과 각론의 연계</li> <li>· <u>같은 개정으로 인한 혼란</u></li> <li>· <u>개정 방식과 절차</u></li> <li>· <u>언론의 소극적 태도</u></li> <li>· <u>수요자의 무관심</u></li> <li>· <u>학교급별 신설 내용</u> <ul style="list-style-type: none"> <li>- 초: 교과서 한자 병기, 안전교육</li> <li>- 중: 자유학기제, 소프트웨어 교육</li> <li>- 고: 수능과의 연계, 진로선택 과목</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>· 고교 공통 과목 운영</li> <li>· <u>역사교과서 국정화</u> <ul style="list-style-type: none"> <li>- 고시, 역사교육 세부 내용, 정치적 대립, 찬반 논란, 집필진 구성</li> </ul> </li> <li>· <u>현장교원 의견 수렴</u></li> <li>· 초등 교과서 한자 병기</li> <li>· 수능·대입제도 개선</li> <li>· 소프트웨어 교육 강화</li> <li>· 교수·학습 방법 개선</li> <li>· 체험·인성교육 강화</li> <li>· 교육과정 개정의 필요성</li> <li>· 안전교육 강화</li> <li>· <u>교육과정 지원체제 구축</u></li> <li>· 자유학기제 운영</li> </ul>

우선, 교육부(2015a)는 해당 정책을 통해 학생들에게 중점적으로 길러주고자 하는 핵심역량을 설정하고 문·이과 공통 과목을 신설하며, 연극과 소프트웨어 교육 등 인문·사회·과학기술에 대한 기초 소양 교육을 강

화하고자 하였다. 또한 학습 내용을 적정화하며, 학생 중심의 교수·학습 및 평가 방법을 제시하고자 하였다. 이러한 강조점은 교육과정 분야 전문가가 수행한 관련 선행연구들(강지영, 소경희, 2016; 김소영, 강신원, 2015; 손민호, 2016; 이윤복, 강현석, 2015 등)에서도 동일하게 나타났다. 다만, 잦은 개정으로 인한 혼란, 관련 정책에 대한 언론과 수요자의 소극적 태도, 개정 방식과 절차의 타당성 등 향후 개선이 필요한 사항과 학교급별 신설 내용이 전문가들의 선행연구에서 추가적으로 검토되었다.

그에 반해, 이 연구의 분석 데이터인 2015 개정 교육과정 정책 관련 언론보도에서는 새로운 교육과정이 추구하는 교육 이념과 방향보다는 사회적 논란이 된 역사교과서 국정화 문제가 가장 큰 쟁점으로 논의되었다. 2015 개정 교육과정이 정책 형성 단계에서 문·이과 통합형 교육과정으로 언급되었던 만큼, 언론 매체에서도 ‘고교 공통 과목 운영’과 같은 문·이과 통합교육 관련 쟁점을 큰 비중으로 다루긴 하였으나, 초·중등교육 전 과정에서의 운영 방안을 구체적으로 논의하기보다는 신설 과목에 대한 소개에 그친 경우가 많았다. 이는 교육부가 제시한 2015 개정 교육과정 정책의 취지나 의도, 관련 분야 전문가가 제안하는 교육개혁의 방향 등이 국내 언론 매체를 통해 일반 대중들에게 제대로 전달되지 못하였음을 시사한다. 따라서 2015 개정 교육과정 정책에 대한 이해도와 운영의 충실도(fidelity of implementation)를 증진시키기 위해서는 교육당국의 정책 홍보 활동 강화가 요청된다고 할 수 있다.

셋째, 2015 개정 교육과정 정책의 쟁점은 분석 기간 전체를 기준으로 총 16개가 도출되었으며, 정책의 형성·결정·집행 단계에 따라 각각 5개, 11개, 9개로 변화하였다. 우선, 쟁점의 내용 측면에서 2015 개정 교육과정 정책의 쟁점이 다양하다는 것은 해당 정책을 교육현장에 제대로 정착시키기 위해 국민적 합의와 이에 바탕을 둔 체계적 준비가 필요함을 시사한다(유예림, 백순근, 2016). 또한 쟁점의 개수 측면에서 해당 정책의 쟁점이 달라진다는 것은 정책 단계 구분에 따른 정책평가 연구가 필요함을 시사한다. 아울러 이 연구에서 각 정책 단계별 분석 기간이 약 1년으

로 거의 동일함에도 불구하고 정책 결정 단계의 쟁점 개수가 정책의 형성·집행 단계보다 많다는 것은 정책 결정 단계에서의 사회적 합의가 특히 중요함을 시사한다.

넷째, 이 연구를 통해 도출된 2015 개정 교육과정 정책의 쟁점이 논의된 배경을 고려할 때, 교육정책의 쟁점을 제시하는 새로운 주체로 관련 시민단체가 부상한 점을 확인할 수 있었다. 이는 특히 정책 결정 단계의 ‘초등 교과서 한자 병기’와 ‘수학교육 내용 적정화’, 정책 집행 단계의 ‘공교육 정상화’에서 두드러졌다. 즉, 초등 교과서 한자 병기와 관련해서는 이를 반대하는 민족문제연구소 등 전국 46개 교육시민단체가 ‘초등교과서 한자병기 반대 국민운동본부’를 발족하였고(뉴시스, 2015), 수학교육 내용 적정화와 공교육 정상화와 관련해서는 시민사회단체인 ‘사교육걱정없는세상’을 중심으로 적극적인 의견 개진이 이루어졌다.

이러한 정책 결정 구조의 다원화는 관련 선행연구(강지영, 소경희, 2016; 장수빈, 2016)에서도 유사하게 확인되었다. 이진영(2014) 역시 정책 여론 형성의 주요 참여자가 정책 형성 단계에서는 정부와 오피니언 리더인 반면, 정책 결정·집행 단계에서는 언론과 정책 수요자도 활발하게 참여하는 등 정책 과정에 따라 정책 여론의 형성이 하향식에서 상향식으로 변화한다고 주장하였다. 따라서 향후에는 빅데이터 분석 기법을 활용하여 해당 정책에 대한 다양한 이해당사자들의 여론을 실시간으로 분석해 이들의 의견을 다각도로 수렴함과 동시에, 이에 기반하여 정책의 향후 과제를 지속적으로 모색할 필요가 있다.

다섯째, 2015 개정 교육과정 정책의 쟁점으로 도출된 ‘수능·대입제도 개선’과 ‘(수학)교육 내용 적정화’에 대한 구체적 방안 마련 및 검토가 시급하다. 우선, 이 연구의 분석 결과 ‘수능·대입제도 개선’ 쟁점에 대한 논의는 정책 결정·집행 단계에서 적극적으로 이뤄지지 않았다. 2015 개정 교육과정 총론의 ‘교육과정 구성의 중점’ 부분에서도 제시되었듯이 “교과의 교육목표, 교육 내용, 교수·학습 및 평가의 일관성을 강화”(교육부, 2015d: 3)하는 것이 중요함에도 불구하고 정책 집행 단계인 2017년 1월

현재까지도 수능·대입제도 개선과 관련된 구체적 방안이 발표되지 않았다. 학교 현장에서 교육과정을 적절하게 운영하기 위해서는 대입 전형 방안을 적절한 시기에 확정하여 발표하는 것이 중요하다(홍원표, 2016).

또한 ‘학습량 적정화’ 관련 쟁점은 언론 보도 내용에서 수학 교과에 한정하여 논의가 이뤄지고 있었는데, 수학 이외의 교과에서도 관련 논의가 핵심역량과 함께 종합적으로 이뤄질 필요가 있다. 사실 학습량 적정화 문제는 교육과정이 개발될 때마다 지속적으로 지적되어 온 문제이나, 교과 전문가들 사이에서도 적정한 학습량에 대한 기준 설정, 핵심 개념이나 지식에 대한 합의 등이 이루기 매우 어려워 실질적인 개선 방안이 도출되지 못하였다(이화진 외, 2016). 또한 학습량을 감축하기 위해 새롭게 도입된 교육과정 구성 방식은 학습량을 오히려 증가시키는 부작용을 초래할 수도 있으므로(손민호, 2016) 학습량을 적정화할 수 있는 실효성 있는 방안을 적극 모색하여야 할 것이다.

이 연구의 의의는 다음과 같다.

첫째, 이 연구는 교육 분야에서 빅데이터 분석 기법을 활용한 새로운 정책평가 연구의 가능성을 제시했다는 점에서 의의를 찾을 수 있다. 특히 토픽 모델링의 경우, 개별 연구자의 통찰력에 의존하는 전통적인 질적 연구방법과 달리 토픽 추출 방법이 명료하여 타 연구자에 의한 재생가능성이 높고, 자동화된 분석 방법을 활용하기 때문에 코더(coder)들 간 일치도 등의 신뢰도 검증 측면에서 유리하다(남춘호, 2016). 따라서 향후에는 이 연구에서 활용한 분석 기법을 자유학기제, 혁신학교, 대입 정책, 고등교육 지원 특성화 사업(예: ACE, PRIME) 등 다양한 교육정책을 대상으로 확대할 필요가 있다. 이는 기존 이론이나 연구 방법에서 간과되었던 쟁점이나 쟁점 간의 관계, 정책 여론 및 대응 방안 등을 즉시적으로 탐색함으로써 장기적인 교육 발전에 기여할 수 있을 것이다.

둘째, 이 연구는 ‘교육학’ 연구자의 관점에서 빅데이터 분석 기법을 활용하였기 때문에, 텍스트 빅데이터 분석 기법을 교육 분야에 적용할 때

고려해야 하는 사항들을 발견하고 제공했다는 점에서 의의가 있다. 따라서 이 연구에서 사용한 빅데이터 분석 기법을 후속 연구를 통해 더욱 정교하게 발전시키고, 다양한 기법을 적용함으로써 방법론적 가능성을 좀 더 모색할 필요가 있다. 예컨대, 이 연구에서는 연구자가 토픽 내 주요어들의 조합이 갖는 의미를 고려하여 토픽명을 추출하고 이를 타당화하기 위해 전문가 협의회를 개최하였는데, 향후에는 언어 네트워크 분석(semantic network analysis) 기법을 활용해 주요어들 간의 관계를 파악하는 방식으로 추가적인 타당화 작업을 수행할 수도 있을 것이다.

셋째, 이 연구는 2015 개정 교육과정 정책에 대한 언론 매체의 관심도와 이를 통해 형성된 주요 쟁점을 분석함으로써 해당 정책에 대한 이해의 폭을 넓혔다는 점에서 의의가 있다. 정책 추진 과정에서 발생하는 쟁점이 부정적 혹은 감정적으로 치우쳐 확산되다 보면 이해당사자 간 소모적 갈등을 야기해 정책을 효과적으로 실행하기 어렵다(문빛, 김찬석, 이철한, 2013)는 점에서 정책에 대한 쟁점 분석 연구는 매우 중요하다. 다만, 이러한 분석 결과가 2015 개정 교육과정 시행 및 향후 교육과정 개정 작업에 주는 시사점을 면밀히 도출하기 위해서는 관련 이해당사자인 초·중등 교육현장의 교사와 학생 및 학부모, 교육과정 분야의 전문가, 교육당국 관계자들과의 심층 면담 등을 통한 질적 연구를 병행할 필요가 있을 것이다.

한편, 이 연구의 제한점 및 제언은 다음과 같다.

첫째, 이 연구에서는 2015 개정 교육과정 정책의 관심도와 그에 따라 형성된 쟁점을 언론 매체를 통해 보도된 뉴스 기사를 활용하여 분석하였다. 뉴스 기사는 그 특성상 여러 이해당사자들의 관점이 혼재되어 있어 특정 정책에 대한 전반적인 쟁점과 여론을 파악하는 데 용이하다. 그러나 한편으로 일반적인 뉴스 기사의 경우 대체로 정보 전달의 성격을 갖고 있기 때문에 지배적인 쟁점을 찾기 어려울 수도 있다. 따라서 향후에는 일반 보도기사와 특정 이해당사자가 작성한 사설 및 칼럼 기사에서



나타나는 쟁점 간에 차이가 있는지를 구분하여 확인할 필요가 있다. 예컨대, 토픽 모델링 중 디리클레 다항 회귀(Dirichlet-multinomial regression: DMR)를 활용하면 뉴스 기사를 통해 추가적으로 확보할 수 있는 정보(예: 매체 특성, 기사의 종류 등)가 2015 개정 교육과정 정책에 대한 쟁점에 어떠한 영향을 미치는지 등을 추가적으로 탐색할 수 있다.

둘째, 국내 언론 매체의 뉴스 기사에서 나타난 2015 개정 교육과정 정책의 쟁점은 상호 연관된 구조적 관계를 갖고 있었다. 분석 결과 도출된 토픽들 간의 관계를 파악하기 위해 ‘pyLDAvis’ 라이브러리를 활용하여 토픽 구조를 시각화한 결과, 역사교과서와 관련된 하위 쟁점인 ‘역사교과서 국정화 찬반 논란’과 ‘역사교과서 국정화 고시’, ‘역사교과서 집필진 구성’과 ‘현장 교원 의견 수렴’이 서로 유사한 토픽으로 구분되었다. 또한 ‘수능·대입제도 개선’과 ‘교수·학습 방법 개선’ 역시 서로 유사한 토픽으로 시각화되었다. 이는 향후 토픽 간의 상관이나 위계적 구조를 반영하여 분석할 수 있는 상관 토픽 모형(Correlated Topic Models: CTM), 위계적 잠재 디리클레 할당(Hierarchical Latent Dirichlet Allocation: hLDA) 등을 적용하여 이 연구 결과를 확장할 필요가 있음을 시사한다.

셋째, 이 연구에서 분석 대상으로 설정한 2015 개정 교육과정 정책은 연구의 시점상 본격적인 시행 이전 단계이므로 관련 이해당사자들의 의견이 개진된 텍스트(예: SNS, 카페 게시글 등)가 많지 않아 분석에 어려움이 있었다. 따라서 2015 개정 교육과정 정책이 교육현장에서 시행된 이후 각 이해당사자 집단의 입장을 대표하는 텍스트(예: 공급자인 교사의 경우 관련 교원단체의 보도자료 및 토론회 자료, 수요자인 학생의 경우 온라인 커뮤니티에서의 질문·토론 내용, 전문가의 경우 학술지, 교육당국의 경우 보도자료 등의 공문서)를 추가적으로 확보·분석하여 이해당사자들 간 쟁점에 차이가 있는지 등을 좀 더 모색할 필요가 있다. 나아가 관련 데이터가 증가한 시점에는 감정 분석(sentiment analysis)을 활용하여 이해당사자들의 정책 선호도를 분석·평가한다면 빅데이터 분석 기법을 활용한 교육정책 평가 연구에 좀 더 기여할 수 있을 것이다.



## 참 고 문 헌

- 강이화(2015). 2015 개정 교육과정 교과역량의 타당성 분석: 통합교과 교육과정을 중심으로. **학습자중심교과교육연구**, 15(12), 815-835.
- 강정배(2015). 언어 네트워크 분석(Semantic Network Analysis)을 활용한 특수교육 분야 연구동향 분석, 1997년~2014년. **특수교육저널: 이론과 실천**, 16(1), 281-306.
- 강지영, 소경희(2016). 2015 개정 교육과정 총론에 대한 미디어 담론 분석. **교육과정연구**, 34(3), 1-27.
- 강현석, 이지은, 전호재(2015). 2015 문·이과 통합형 교육과정 개정의 요구 조사에 나타난 교육과정적 의미 탐색. **교육문화연구**, 21(5), 5-38.
- 고인규(2016). 빅데이터 분석 방법에 기반 한 실과 농업 단원의 핵심개념 연구. **학습자중심교과교육연구**, 16(4), 431-450.
- 관계부처 합동(2014). 2014년도 정부3.0 추진계획.
- 교육과학기술부(2009). 초·중등학교 교육과정 총론. 교육과학기술부 고시 제2009-41호 [별책 1] (2009. 12. 23).
- 교육부(2013). 2017학년도 대입제도 확정. 보도자료(2013. 10. 25).
- 교육부(2014a). 2014년 교육부 업무보고. 보도자료(2014. 2. 13).
- 교육부(2014b). 2015 문·이과 통합형 교육과정 총론 주요 사항(시안).
- 교육부(2014c). 2015 문·이과 통합형 교육과정의 총론 주요사항 발표. 보도자료(2014. 9. 24).
- 교육부(2015a). 2015 개정 교육과정 총론 및 각론 확정·발표. 보도자료 (2015. 9. 23).
- 교육부(2015b). 배움을 즐기는 행복교육, 2015 개정 교육과정이 함께 하겠습니다. 자료 출처(검색일 2015. 11. 7.): <http://blog.naver.com/moeblog/220490016283>.
- 교육부(2015c). 2015 개정 교육과정 질의·응답자료. 자료 출처(검색일 2015. 11. 7.): <http://blog.naver.com/moeblog/220489069781>.

교육부(2015d). 초·중등학교 교육과정 총론. 교육부 고시 제2015-80호 [별책 1] (2015. 9. 23).

교육부(2015e). 교육부, 「교과용도서 개발 체제 개선 방안」 발표. 보도자료(2015. 7. 30).

교육부(2015f). 사회과 교육과정. 교육부 고시 제2015-74호 [별책 7] (2015. 9. 23).

교육부(2015g). 과학과 교육과정. 교육부 고시 제2015-74호 [별책 9] (2015. 9. 23).

교육부(2016a). 「학교생활기록 작성 및 관리지침」 훈령 일부 개정. 보도자료(2016. 4. 5).

교육부(2016b). 2015 개정 교육과정 생생(生生) 교원 연수 추진: 교육부, 2015 개정 교육과정 및 학생참여형 수업개선을 위한 교원역량 강화 추진. 보도자료(2016. 6. 9).

교육부(2016c). 2015 개정 교육과정에 따른 초·중등학교 디지털교과서 국·검정 구분 고시. 보도자료(2016. 8. 29).

교육부(2016d). 교육부, 「학교생활기록부 기재 개선 방안」 발표. 보도자료(2016. 11. 24).

교육부(2016e). 올바른 역사교과서 현장 검토본 의견 수렴 결과 중간 발표. 보도자료(2016. 12. 5).

교육부(2016f). 자유학기제와 2015개정교육과정 정착을 위한 교원 양성과정 개선계획 수립. 보도자료(2016. 8. 18).

교육부(2017a). 「교과용도서에 관한 규정」 일부개정령안 입법예고. 보도자료(2017. 1. 3).

교육부(2017b). 2015 개정 교육과정에 따른 역사교육 연구학교 운영계획 발표. 보도자료(2017. 1. 10).

교육부(2017c). 배움을 즐기는 행복교육, 2015 개정 교육과정이 함께 하겠습니다: 고등학교용. 자료 출처(검색일 2017. 1. 20.): <http://ncic.go.kr/mobile.brd.ntc.view.do?articleNo=10000069>.

교육부, 대입제도 발전방안 연구위원회(2013). 학생·학부모 부담 완화와 학

- 교교육 정상화를 위한 대입전형 간소화 및 대입제도 발전방안(시안).  
구정화(2015). 문·이과 통합형 교육과정의 “통합사회” 과목 개발에 대한  
성찰적 연구. **시민교육연구**, 47(3), 1-29.
- 구정화, 박철웅, 박병기, 모경환, 조철기, 박병석, 최정순, 김재준, 김남준,  
윤영진, 박주현(2015). **2015 개정 교과 교육과정 시안 개발 연구 I:  
통합사회 교육과정**. 한국교육과정평가원, 연구보고 CRC 2015-5.
- 권기석, 박진수, 구찬동(2014). **교육 (빅)데이터를 이용한 아젠다 개발:  
사회연결망 분석을 중심으로**. 한국교육학술정보원, 연구보고 KR  
2014-10.
- 권정은, 정지선(2012). **소셜 분석으로 살펴본 청소년 자살예방정책의  
시사점**. 한국정보화진흥원, IT & Future Strategy 제17호.
- 권현범(2016). 언어 네트워크 분석방법을 이용한 교육의 의미와 연구 영  
역 탐색. **교육혁신연구**, 26(1), 125-148.
- 김갑수(2016). 2015 개정 교육과정의 소프트웨어 교육에 대한 초등 교사  
들의 인식 분석. **한국정보교육학회 논문지**, 20(1), 47-56.
- 김경자(2014). 창의·융합 인재 양성을 위한 통합 교육과정 접근. **2014년  
한국교육과정학회 3월 월례학술대회 자료집**(pp. 43-52). 서울:  
한국교육과정학회.
- 김경자(2015). 2015 교육과정 개정 정책 방향. **교육광장**, 55, 8-11.
- 김경자, 박상훈, 백남진, 송호현, 온정덕, 이승미, 한혜정, 허병훈, 홍은숙  
(2015). **2015 개정 교육과정 총론 시안[최종안] 개발 연구**. 교육  
부, 국가교육과정개정연구위원회.
- 김규하, 박철웅(2015). 토픽 모형 및 사회연결망 분석을 이용한 한국데이  
터정보과학회지 영문초록 분석. **한국데이터정보과학회지**, 26(1),  
151-159.
- 김기환(2013). 공공부문 빅데이터의 활용성과 위험성. **정책분석평가학회  
보**, 23(2), 1-27.
- 김남희, 심규철(2015). 2015 개정 과학과 교육과정 “통합과학”과 사범대  
학 예비 과학 교사 교육 내용의 분석을 통한 예비 과학 교사 교

- 육에 대한 시사점. **한국과학교육학회지**, 35(6), 1039-1048.
- 김대중, 윤서연(2013). 빅데이터를 활용한 국토정책반응 모니터링 및 정책수요 예측방안. 국토연구원 연구보고, 국토연 2013-41.
- 김병수(2015). 통합형 국어과 교육과정의 과제와 전망. **학습자중심교과교육연구**, 15(8), 247-265.
- 김병주, 김태완, 김은아(2006). 교원평가제에 대한 신문의 보도태도 분석. **한국교원교육연구**, 23(1), 349-371.
- 김성근, 조혁준, 강주영(2016). 학술연구에서의 텍스트 마이닝 활용 현황 및 주요분석기법. **정보화연구**, 13(2), 317-329.
- 김소영, 강신원(2015). 뉴스 기사를 통해 본 2015 문·이과 통합형 교육과정의 쟁점과 과제. **교육혁신연구**, 25(2), 91-112.
- 김수철, 이환철(2014). ‘스토리텔링 수학’ 관련 언론 보도 내용 분석. **수학교육논문집**, 28(2), 179-193.
- 김신복(1983). 정책개발의 과정과 방법. 유훈 외, **정책학**(pp.11-13), 서울: 법문사.
- 김양분, 유한구, 남궁지영(2004). 교육 만족도 조사 도구 개발 연구. 한국교육개발원 현안연구, OR 2004-5.
- 김용(2005). 교육정책과 교육과정, **교육과정연구**, 23(1), 283-303.
- 김용대, 조광현(2013). 빅데이터와 통계학. **한국데이터정보과학회지**, 24(5), 959-974.
- 김용대, 정구환(2016). 토픽모형을 이용한 빅데이터 기반 마이크로 세그멘테이션 방법론 연구. **통계연구, 특별호**, 17-34.
- 김우주, 구찬동(2015). 교육정보화 정책에서 빅데이터 활용 방안 탐색 연구: 소셜미디어 데이터 활용을 중심으로. 한국교육학술정보원, 연구보고 KR 2015-4.
- 김은희, 표신지, 김문철(2012). LDA 기반 은닉 토픽 추론을 이용한 TV 프로그램 자동 추천. **방송공학회논문지**, 17(2), 270-283.
- 김응현(2015). 문·이과 통합형 사회과 교육과정 개정을 둘러싼 사회과교육의 이상과 현실의 괴리에 관한 연구: 사회과 통합교육과정에 관

- 한 갈등을 중심으로. **사회과교육연구**, 22(4), 63-78.
- 김이경, 안지윤, 김정, 김지혜, 유완(2016). 2015 개정 교육과정에 대한 교원의 관심이 연수 참여의지에 미치는 영향 분석. **교육과정평가연구**, 19(1), 23-47.
- 김정미, 윤미영(2012). **영유아 보육정책에 대한 국민정서분석: 빅데이터 분석 따라하기-사례①**. 한국정보화진흥원.
- 김정수, 이석준(20016). 취업준비생 토픽 분석을 통한 취업난 원인의 재탐색. **경영과 정보연구**, 35(1), 85-116.
- 김정원, 김기수, 정미경, 홍인기(2012). **미래형 교사교육체제 구안 연구**. 한국교육개발원, 연구보고 RR 2012-03.
- 김정원, 김지수, 최유림(2014). **남북한 교사 역할 비교 분석 연구**. 한국교육개발원, 연구보고 RR 2014-08.
- 김지은, 백순근(2016). 텍스트 빅데이터 분석기법을 활용한 대학구조개혁평가의 쟁점 분석. **아시아교육연구**, 17(3), 409-436.
- 김찬석(2012). 공공갈등 관리의 성공 요인과 커뮤니케이션 메시지 특성. **홍보학연구**, 15(4), 5-31.
- 김현호, 왕지현, 김상훈(2014). 토픽 클러스터링 기반 언어모델링을 위한 문장 분류. **2014 한국정보과학회 제41회 동계학술발표회 논문집**(pp. 465-467). 한국정보과학회.
- 나일주, 임철일, 조영환(2015). **학습분석 모델 및 확장 방안 연구**. 서울특별시교육청 위탁연구 보고서.
- 남춘호 (2016). 일기자료 연구에서 토픽모델링 기법의 활용가능성 검토. **비교문화연구**, 22(1), 89-135.
- 노영균, 김기웅(2014). 거리 학습과 최근린 방법을 이용한 빅데이터 분석. **정보과학회지**, 32(7), 26-32.
- 노은희, 성경희, 임은영(2015). 한국어 문장 수준 서답형 문항 자동채점 적용 가능성 탐색. **교육평가연구**, 28(2), 523-551.
- 노화준(2015). **정책평가론**(제5판). 파주: 법문사.
- 류민영(2013). 교원능력개발평가 정책네트워크의 특성을 통한 정책과정

- 분석. **교육행정학연구**, 31(2), 227-255.
- 문빛, 김찬석, 이철한 (2013). 전략적 쟁점관리를 위한 정책 쟁점의 역동성: 정책 쟁점별 매체 및 프레이밍 변화과정에 관한 연구. **한국언론학보**, 57(6), 121-148.
- 문성호, 윤동엽(2011). 학교폭력에 관한 뉴스프레임 연구. **미래청소년학회지**, 8(1), 87-109.
- 박경호, 길혜지, 김주아, 박병영, 최유리(2016). 2015 개정 교육과정에 따른 대학수학능력시험 개선방향 탐색 연구. 한국교육개발원, 이슈페이퍼 IP 2016-02.
- 박광배(2000). **다차원척도법**. 서울: 교육과학사.
- 박대민(2013). 뉴스 기사의 빅데이터 분석 방법으로서 뉴스정보원연결망 분석. **한국언론학보**, 57(6), 234-262.
- 박대현, 송동현(2014). 비정형 데이터 활성화의 정치, 경제, 문화적 함의. 한국인터넷진흥원, Internet & Security Focus, 2014년 2월호.
- 박도순, 홍후조(2006). **교육과정과 교육평가**. 서울: 교육과학사.
- 박병량, 주철안(2012). **교육행정 및 교육경영**. 서울: 학지사.
- 박선미(2016). 2015 개정 중학교 사회과교육과정 개발 과정의 의사결정 구조에 대한 비판적 고찰. **한국지리환경교육학회지**, 24(1), 33-45.
- 박선주(2015). R을 활용한 SW교육 텍스트데이터 토픽분석. **정보교육학회논문지**, 19(4), 517-524.
- 박선화(2016). 2015 개정 수학과 교육과정의 수학 교과 역량의 의미. 토의·토론식 수학 수업 개선 방안 모색. 2016 한국수학교육학회 춘계학술대회 발표 자료집(pp. 15-22). 부산: 한국수학교육학회.
- 박성태(2011). 사회 갈등 공공이슈에 한 언론의 보도태도연구: 정권교체기 보수와 진보언론의 교육정책 관련 보도태도 분석. **한국공공관리학보**, 25(3), 97-118.
- 박소영, 이정렬, 강현석(2015). 2015 문·이과 통합형 교육과정 개정에 관한 전문가 인식 연구. **수산해양교육연구**, 76, 1172-1183.
- 박수정(2011). 교육감 선거 관련 신문사설에 대한 네트워크 텍스트 분석.



- 교육정치학연구, 18(2), 183-203.
- 박은혜, 이대식, 김정연, 이경선(2015). 2015 개정 창의적 체험활동 교육과정 해설서. 교육부.
- 박종희, 박은정, 조동준(2015). 북한 신년사(1946-2015)에 한 자동화된 텍스트 분석. 한국정치학회보, 49(2), 27-61.
- 박채형(2015). 2015 개정 교육과정 개발 방향의 현실성 분석: 초등학교를 중심으로. 학습자중심교과교육연구, 15(4), 335-354.
- 박철웅, 박병기, 모경환, 조철기, 박병석, 최정순, 김재준, 이영호, 김남준, 이철훈, 박주현(2015). 2015 개정 교과 교육과정 시안 개발 연구Ⅱ: 통합사회 교육과정. 한국교육과정평가원, 연구보고 CRC 2015-25-6.
- 박희경(2016). 교과 교육과정 개발을 위한 속의 과정의 개선 과제: 2015 개정 교육과정 사례를 중심으로. 교육과정연구, 34(3), 185-208.
- 방정숙, 이지영, 서은미(2016). 문제 해결에 관한 초등학교 수학과 교육과정 및 교과용도서 분석. 수학교육학연구, 26(3), 583-605.
- 배상훈, 장환영, 오세희, 장창성, 이태희(2013). 우리나라 중등단계 직업교육 관련 연구동향 분석: 역대 정부별 분석. 직업교육연구, 32(2), 45-71.
- 백순근(2002). 시·도교육청평가의 문제점 및 개선 방안. 아시아교육연구, 3(2), 135-151.
- 백순근(2006). 학교정책평가의 필요성 및 지향점. 학교경영, 19(9), 46-51.
- 백승훈, 남상준(2009). 사회과 교육과정 결정과정의 정책네트워크 분석: 2007년 개정 초등 사회과교육과정을 중심으로. 사회과교육연구, 16(1), 49-65.
- 서동민, 정한민(2013). 빅데이터 분석 서비스를 위한 지능형 웹 크롤러. 한국콘텐츠학회논문지, 13(12), 575-584.
- 서명구, 박규석(2004). 효율적인 웹 스크래핑을 위한 확장 WIDL에 관한 연구. 한국인터넷정보학회 학술발표대회 논문집, 5(2), 241-244.
- 서울대학교 교육연구소(2011). 교육학용어사전. 서울: 하우동설.
- 서울특별시교육청(2016a). 서울 일반고에 ‘개방-연합형 종합캠퍼스 교육과정’ 도입. 보도자료(2016. 4. 20).

- 서울특별시교육청(2016b). 서울시교육청, 초 1~2 안성맞춤 학년제 도입. 보도자료(2016. 8. 30).
- 소경희(2003). 국가수준에서 개발된 교육과정의 실행 양상에 대한 이해: 초등학교 국어과 사례를 중심으로. **교육과정연구**, 21(1), 129-153.
- 소경희(2014). 2015 개정 초·중학교 교육과정 개정(안)의 기본 방향 및 남겨진 과제. **2015 문·이과 통합형 교육과정 개정(안)의 기본 방향에 대한 소통과 논의**. 2014년도 한국교육과정학회 추계 학술대회 자료집(pp. 23-44). 서울: 한국교육과정학회.
- 소경희(2015). 2015 개정 교육과정 총론 개정안이 남긴 과제: 각론 개발의 쟁점 탐색. **교육과정연구**, 33(1), 195-214.
- 손민호(2016). 2015개정교육과정이 남긴 문제들. **교육비평**, 37, 37-47.
- 송영조(2012). 빅데이터 시대! SNS의 진화와 공공정책. 한국정보화진흥원, IT & Future Strategy 제13호.
- 송영조(2013). 정책도구로서의 빅데이터 역할에 대한 탐색적 연구. 2013 한국정책분석평가학회 추계 학술세미나 자료집. **한국 과학기술정책의 조망: 성과평가와 성과관리**(pp. 235-261). 서울: 한국정책분석평가학회.
- 송태민(2014). 소셜 빅데이터를 활용한 보건복지정책 동향분석. **보건복지포럼**, 213, 101-113.
- 송태민, 송주영(2015). **빅데이터 연구 한 권으로 끝내기**. 서울: 한나래출판사.
- 송태민, 송주영(2016). **R을 활용한 소셜 빅데이터 연구방법론**. 서울: 한나래출판사.
- 송태민, 송주영, 안지영, 진달래(2013). 소셜 빅 데이터를 활용한 자살검색 요인 다변량 분석. **보건교육·건강증진학회지**, 30(3), 59-73.
- 신동희(2015). 빅데이터와 사용자. **대학교육**, 188, 10-16.
- 신세인, 이준기, 하민수, 이태경, 정영희(2015). 과학영재 중학생들과 일반 중학생들의 과학과 관련된 직업에 대한 인식 비교: 언어 네트워크 분석법 중심으로. **영재교육연구**, 25(5), 673-696.
- 신수정(2014). **글에서 감정을 읽다, 감성 분석의 이해**. IDG Tech Report.

- 신준수(2014). **KOMORAN 2.0 한국어 형태소 분석기**. 자료 출처(검색  
일 2016. 8. 1.): <http://shineware.tistory.com/attachment/cfile30.uf@2369B250547217C6290AF3.pdf>
- 안주영, 안규빈, 송민(2016). 텍스트 마이닝을 이용한 매체별 에볼라 주제  
분석. **한국문헌정보학회지**, 50(2), 289-307.
- 양길석(2010). 대학수학능력시험의 변천과정과 쟁점 분석. **교육평가연구**,  
23(4), 765-791.
- 양정실(2016). 국어과 교육과정 정책 결정 양상과 합리성 제고 방안. **국  
어교육학연구**, 51(2), 55-79.
- 양정혜(2015). 2015 개정 교육과정에서 초등 가정생활영역의 내용요소  
분석. **한국실과교육학회지**, 28(4), 327-341.
- 옥일남(2016). 2015 개정 교육과정기 ‘통합사회’ 개발 과정과 쟁점. **교육  
과정평가연구**, 19(2), 49-75.
- 온정덕(2015). **2015 개정 교육과정의 방향과 주요 개정 내용**. 한국교육  
개발원 이슈페이퍼 07, CP 2015-01-7.
- 온정덕, 김정자, 박희경, 홍은숙, 황규호(2015). **2015 개정 교육과정 총  
론 해설서 (초등학교) 개발 연구**. 교육부.
- 왕재선, 김선희(2013). 정책이슈 확산의 다이내믹스: 무상급식 논쟁 사례  
를 중심으로. **한국정책학회보**, 22(1), 389-422.
- 유병건, 김자미, 이원규(2016). 2015 개정 교육과정의 컴퓨팅 시스템 단원  
의 집필에 대한 시사점. **한국컴퓨터교육학회 논문지**, 19(2), 31-40.
- 유병열(2016). 초등 도덕과 새 교과용 도서 개발의 방향과 과제. **한국초등  
교육**, 27(1), 329-348.
- 유예림(2008). **중등교사 임용시험 정책 평가 연구: PUPA 모형을 중심으  
로**. 서울대학교 대학원 석사학위 논문.
- 유예림, 백순근(2016). 자동화된 텍스트 분석을 활용한 2015 개정 교육과  
정 정책에 대한 언론 보도의 쟁점 분석. **교육과정평가연구**,  
19(3), 127-156.
- 유제순, 고은정, 서창호, 김진숙, 김영훈, 박지만, 김승욱, 박영일(2015).

- 2015 개정 창의적 체험활동 교육과정 해설서 개발 연구. 교육부.
- 유종열(2016). 2015 개정 사회과 교육과정에서의 인구교육 내용 분석. **학습자중심교과교육연구**, 16(3), 759-783.
- 유진은(2016). 데이터 마이닝 기법을 통한 교육 패널데이터 분석: 벌점회 귀모형과 KYPS 자료. **아시아교육연구**, 17(3), 1-19.
- 윤광석, 정혜경, 김양분(2011). 교육정책 및 프로그램의 효과성에 관한 증거기반 수립 모형: 미국의 사례를 중심으로. 한국교육개발원 포지션 페이퍼, 현안보고 OR 2011-03-20.
- 윤영철, 윤석민, 박상인, 배진아, 변종석, 심미선, 양승찬, 이민규, 이재진, 정준희, 김유정(2015). **여론집중도조사 보고서**. 서울: 제2기 여론집중도조사위원회.
- 윤정일, 송기창, 조동섭, 김병주(2015). **교육행정학 원론**(제6판). 서울: 학지사.
- 윤지현, 강성주(2016). 2015 개정 교육과정에서 통합과학과 과학탐구실험 교과에 관해 고등학교 과학 교사들이 기대하는 부분과 우려하는 부분에 대한 분석. **학습자중심교과교육연구**, 16(5), 515-546.
- 은지용, 이간용, 최병택, 한춘희(2015). 2015 개정 초등 사회과 교육과정 시안의 기본 방향과 주요 특징 및 변화. **사회과교육**, 54(4), 65-83.
- 이경진, 김경자(2012). 통합교육과정 접근으로서의 '융합인재교육'(STEAM)의 의미와 실천 가능성 탐색. **초등교육연구**, 25(3), 55-81.
- 이광우(2014). 문·이과 통합형 교육과정을 위한 교과 교육과정 개정(안)의 설계 방향. 2015 문·이과 통합형 교육과정 개정(안)의 기본 방향에 대한 소통과 논의. 2014년도 한국교육과정학회 추계 학술대회 자료집(pp. 91-122). 서울: 한국교육과정학회.
- 이기종(2015). 대학 구조 개혁 평가의 배경, 쟁점, 대안. **교육평가연구**, 28(3), 933-954.
- 이만재(2012). 빅 데이터 어널리틱스와 공공 데이터 활용. **정보과학회지**, 30(6), 33-39.
- 이미경, 정영근, 권점례, 이근호, 김희경, 이주연, 이명애, 가은아, 김현수,

- 박은아, 박진동, 김현정, 진의남, 김기철, 이경언, 양윤정, 주형미, 백경선, 김경훈, 장호성, 이근님, 한혜정, 서민철(2016). **2015 개정 교육과정에 따른 초·중학교 교과 평가기준 개발 연구(총론)**. 한국교육과정평가원, 연구보고 CRC 2016-2-1.
- 이미숙, 이창훈, 김지연(2014). **빅데이터를 활용한 환경분야 정책수요 분석**. 한국환경정책·평가연구원, 연구보고서 2014-01.
- 이성숙(2016). 2015개정 실과교육과정에 기반한 예비교사 교육 방안. **한국실과교육학회지**, 29(3), 115-132.
- 이성직, 김한준(2009). TF-IDF 의 변형을 이용한 전자뉴스에서의 키워드 추출 기법. **한국전자거래학회지**, 14(4), 59-73.
- 이수연, 김현정, 정수연(2016). 텍스트마이닝 기반 토픽 분석을 통한 온라인 성차별성의 이해. **사이버커뮤니케이션학보**, 33(3), 159-199.
- 이승룡, 최문정, 민경미, 이정재, 이승규, 최한림, 김연화(2013). **2012년도 기술영향평가: 빅데이터 분석 기술과 활용**. 한국과학기술기획평가원, 연구보고 2013-070.
- 이용상, 김현미, 신진아, 정혜경, 박상욱, 권태현, 변태진(2015). **수능 체제 개편의 쟁점 및 현황**. 한국교육과정평가원, 연구자료 ORM 2015-50-15.
- 이원희(2015). 2015 개정 교육과정의 역량 개념 도입과 해결 과제: 행동주의와 인지론의 접근. **통합교육과정연구**, 9(4), 113-134.
- 이원희, 이지은(2014). 2015 개정 교육과정에 따른 교원양성 대학의 대처 방안. **교육학논총**, 35(2), 79-95.
- 이윤복, 강현석(2015). 2015 개정 교육과정 내용 분석 및 그 시사점 탐구. **중등교육연구**, 63(4), 435-464.
- 이윤식(2011). 우리나라에서 지식정보화에 따른 정책평가연구의 변화 방향에 관한 탐색적 연구. **정책분석평가학회보**, 21(4), 37-59.
- 이정아(2015). **데이터 증거기반의 과학적 정책 수립 방안**. 한국정보화진흥원 IT & Future Strategy 보고서.
- 이종수(2009). **행정학 사전**. 서울: 대영문화사.

- 이진영(2014). 정책여론 형성 과정 연구. **미디어, 젠더 & 문화**, 29(2), 147-186.
- 이화진, 조용기, 이창훈, 이미경, 이명애, 오상철, 조지민, 박소영, 이용상, 김진구, 양윤정, 정영근, 권점례, 박진용, 주형미, 김수진, 노은희, 박인용, 김태은, 홍선주, 이미숙, 구자옥, 박영수, 정은주, 조보경, 김혜숙(2016). **미래사회 변화 및 학령기 인구 감소 대비 학교교육 진단과 교육방향 탐색**. 한국교육과정평가원, 연구보고 RRO 2016-4.
- 임유나(2016). 2015 개정 교육과정의 핵심역량 제시 방식 검토와 스코틀랜드 ‘Curriculum for Excellence’ 분석을 통한 시사점 탐색. **학습자중심교과교육연구**, 16(5), 143-167.
- 임은진(2015). 핵심역량 기반 교육과정에서의 사회과 평가에 대한 연구. **사회과교육**, 54(4), 143-155.
- 임화진(2014). **빅데이터를 이용한 충남도 정책 키워드 분석**. 충남발전연구원, 전략연구 2014-15.
- 장수빈(2016). 교육과정 속의, 틀 짓기, 미디어의 관계: 2015 국가 교육과정 수학과 개정 사례를 중심으로. **교육과정연구**, 34(1), 29-49.
- 장은주(2016). 2015 개정 국어과 교육과정 학교 언어기능영역 성취기준 적합성 분석. **교육과정평가연구**, 19(2), 27-48.
- 정영근, 박순경, 이광우, 박제윤, 백인화, 나태순, 진동섭, 안종선, 류병구(2010). **고등학교 교육과정 총론 해설 개발 연구**. 한국교육과정평가원, 연구보고 CRC 2010-15.
- 정정길, 성규탁, 이장, 이윤식(2004). **정책평가: 이론과 적용**. 서울: 법영사.
- 조대현(2014). 2015 교육과정 개정안의 문제점 및 대안 연구: 총론과 음악과 내용을 중심으로. **예술교육연구**, 12(4), 207-223.
- 조상연(2015). 2015 개정 교과 교육과정 시안에 대한 비판적 검토. **초등교육연구**, 28(3), 199-227.
- 조성우(2012). **Big Data** 시대의 기술. KT경제경영연구소 Theme Article.
- 조용상(2013). **표준화 이슈리포트: 학습 분석 기술 활용 가능성 및 전망**. 한국교육학술정보원, 연구자료 RM 2013-15.

- 주형미, 윤현진, 이경언, 한혜정, 윤지훈(2015). 2015 개정 교육과정에 따른 교과서 검정 심사 운영 방안(I). 한국교육과정평가원, 연구보고 RRC 2015-9.
- 진설아, 허고은, 정유경, 송민(2013). 트위터 데이터를 이용한 네트워크 기반 토픽 변화 추적 연구. **정보관리학회지**, 30(1), 285-302.
- 채선희(1997). 교육평가의 내재적 가치와 그 실현방안. **교육학연구**, 35(4), 66-80.
- 최인봉, 박도영, 이은경(2014). 국가단위 평가의 수요자 중심 정보 활용 서비스 시스템 구축 및 운영: 시스템 개선 및 고도화 방안 수립. 한국교육과정평가원, 연구보고 RRE 2014-10.
- 최인봉, 이채희, 이은경(2015). 소셜 빅데이터를 활용한 학업성취도 평가 이슈 분석. 한국교육과정평가원 이슈페이퍼, 연구자료 ORM 2015-50-10.
- 최제영(2012). 스마트교육 환경에서의 빅데이터 동향. 2012 KERIS 이슈리포트, 연구자료 RM 2012-19.
- 한경수, 정태범, 유현숙, 서정화, 김신복(1998). 교육(행정)기관 평가체제의 진단. **교육행정학연구**, 16(1), 1-7.
- 한국교육과정평가원(2016). 2015 개정 교육과정에 따른 교과 교육과정 적용을 위한 2017년 교과용도서 검정 신청 안내. 한국교육과정평가원, 기타자료 OM 2016-2.
- 한국교육평가학회(2004). **교육평가용어사전**. 서울: 학지사.
- 한국ABC협회(2015). ABC협회, 2015년 일간신문 발행·유통부수 발표. 보도자료(2015. 11. 27).
- 한신갑(2012). 혼합식 조사와 웹패널의 (열은) 빛과 (짙은) 그늘. **조사연구**, 13(3), 1-31.
- 한신갑(2015). 빅데이터와 사회과학하기. **한국사회학**, 49(2), 161-192.
- 한혜정(2014). 문·이과 통합형 교육과정 개정의 의의와 쟁점. **교육과정연구**, 32(3), 45-69.
- 한혜정, 김영은, 이주연(2016). 교육 목적으로서 ‘일반 능력’ 설정에 대한

- 논의 고찰: 2015 개정 교육과정의 핵심역량과 교과 역량의 관계에 주는 함의. **교육과정연구**, 34(2), 1-18.
- 한혜정, 김영은, 이주연, 박상훈, 김광하, 김기탁, 김용진, 박지만, 진동섭, 최동선, 한충희, 황유진(2015). **2015 개정 교육과정 총론 해설서 (중·고등학교) 개발 연구**. 한국교육과정평가원, 연구보고 CRC 2015-28.
- 한혜정, 백경선(2014). 일반 고등학교 직업교육 지원을 위한 교육과정 편성·운영 방안과 2015 국가 교육과정 개정을 위한 정책 제언. **교육과정연구**, 32(4), 137-161.
- 허경철(2014). 문·이과 통합형 교육과정 개발의 가능 범위 탐색. **2014년 한국교육과정학회 3월 월례학술대회 자료집**(pp. 25-40). 서울: 한국교육과정학회.
- 홍영란, 최정윤, 서정인, 나민주, 장덕호, 현영섭, 김우철, 김현진(2016). **교육개혁의 전망과 과제(Ⅰ): 고등·평생교육 영역**. 한국교육개발원, 연구보고 RR 2016-29.
- 홍원표(2016). 2015 개정 고등학교 교육과정의 적용 방안과 후속 지원에 대한 탐색적 연구. **교육과정연구**, 34(2), 69-94.
- 홍은숙(2015). 국가 교육과정의 교육적 인간상 진술에 대한 비판적 고찰. **교육과정연구**, 33(4), 227-250.
- 홍후조(2014). 고교 진로별 교육과정 속에서 문·이과 통합적 교양교육의 구현 방안. **2014년 한국교육과정학회 3월 월례학술대회 자료집**(pp. 55-97). 서울: 한국교육과정학회.
- 홍후조, 임유나(2014). 고교 문·이과 통합형 교육과정 개발 및 발전 방향에 관한 연구. **교육과정연구**, 32(2), 67-99.
- 황규호(2014a). 문·이과 통합형 교육과정 개발 방향의 검토. **2014년 한국교육과정학회 3월 월례학술대회 자료집**(pp. 1-21). 서울: 한국교육과정학회.
- 황규호(2014b). 2015 문·이과 통합형 교육과정을 위한 고등학교 교육과정 개정(안)의 기본 방향. **2015 문·이과 통합형 교육과정 개정(안)의 기본 방향에 대한 소통과 논의**. 2014년도 한국교육과정학회 추계



- 학술대회 자료집(pp. 59-78). 서울: 한국교육과정학회.
- 황규호, 강태중, 박화식, 백경선, 이영호, 이현주, 임유원, 정광희, 진동섭, 한혜정, 홍은숙(2015). **문·이과 통합형 교육과정 구성 방향 연구**. 교육부.
- 황명화(2014). **여론 모니터링을 위한 비정형 빅데이터 시공간분석 방법론 연구**. 국토연구원, 국토연 2014-10.
- 황정규(2005). **학교학습과 교육평가**. 서울: 교육과학사.
- 황홍섭, 박지수(2016). 지도 토픽 빅데이터 분석과 2009 개정 초등사회과 교과서 지도 활용의 비교분석. **사회과교육**, 55(1), 71-89.
- Anderson. J. E. (2015). *Public policymaking: An Introduction* (8th ed.). Stamford, CT: Cengage Learning.
- Asuncion, A., Welling, M., Smyth, P., & Teh, Y. W. (2009). On smoothing and inference for topic models. *In Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence* (pp. 27-34). AUAI Press.
- Berry, M. W., Dumais, S. T., & O'Brien, G. W. (1995). Using linear algebra for intelligent information retrieval. *SIAM review*, 37(4), 573-595.
- Bettityna, G., & Kurt, H. (2015). Package 'topicmodels' (Topic Models). Retrieved December 1, 2015, from <https://cran.r-project.org/web/packages/topicmodels/topicmodels.pdf>.
- Beyer, M. A., & Laney, D. (2012). *The importance of 'big data': a definition*. Stamford, CT: Gartner.
- Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4), 77-84.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, 3, 993-1022.
- Chang, J. (2015). lda: Collapsed Gibbs Sampling Methods for Topic Models. R package version 1.4.2. Retrieved December 1, 2015,

- from <https://CRAN.R-project.org/package=lda>.
- Chang, J., Gerrish, S., Wang, C., Boyd-Graber, J. L., & Blei, D. M. (2009). Reading tea leaves: How humans interpret topic models. *In Advances in neural information processing systems* (pp. 288–296). Retrieved August 30, 2016, [http://machinelearning.wustl.edu/mlpapers/paper\\_files/NIPS2009\\_0125.pdf](http://machinelearning.wustl.edu/mlpapers/paper_files/NIPS2009_0125.pdf)
- Chen, K. Y. M., & Wang, Y. (2007). Latent Dirichlet Allocation. Retrieved August 3, 2016, <https://pdfs.semanticscholar.org/4574/d77fff19e093782178595a8988a7f3aa1969.pdf>.
- Chuang, J., Manning, C. D., & Heer, J. (2012). Termite: Visualization techniques for assessing textual topic models. *In Proceedings of the International Working Conference on Advanced Visual Interfaces* (pp. 74–77). ACM.
- Cronbach, L. J. (1982). *Designing evaluators of educational and social programs*. San Francisco: Jossey-Bass.
- Davies, P. (1999). What is evidence-based education?. *British journal of educational studies*, 47(2), 108–121.
- 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.
- Fullan, M., & Pomfret, A. (1977). Research on curriculum and instruction implementation. *Review of educational research*, 47(2), 335–397.
- Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National academy of Sciences*, 101(suppl 1), 5228–5235.
- Hadley, W. (2016). httr: Tools for Working with URLs and HTTP. R package version 1.1.0. <https://CRAN.R-project.org/package=httr>.

- Hall, D., Jurafsky, D., & Manning, C. D. (2008). Studying the history of ideas using topic models. In *Proceedings of the conference on empirical methods in natural language processing* (pp. 363–371). Association for Computational Linguistics. Retrieved January 10, 2017, from [http://delivery.acm.org/10.1145/1620000/1613763/p363-hall.pdf?ip=223.39.190.253&id=1613763&acc=OPEN&key=4D4702B0C3E38B35%2E4D4702B0C3E38B35%2E4D4702B0C3E38B35%2E6D218144511F3437&CFID=893500059&CFTOKEN=89437629&\\_acm\\_=1485443729\\_d82910a02633f53335d9869f83048b06](http://delivery.acm.org/10.1145/1620000/1613763/p363-hall.pdf?ip=223.39.190.253&id=1613763&acc=OPEN&key=4D4702B0C3E38B35%2E4D4702B0C3E38B35%2E4D4702B0C3E38B35%2E6D218144511F3437&CFID=893500059&CFTOKEN=89437629&_acm_=1485443729_d82910a02633f53335d9869f83048b06).
- Hofmann, T. (1999). Probabilistic latent semantic analysis. In *Proceedings of the Fifteenth conference on Uncertainty in artificial intelligence* (pp. 289–296). Morgan Kaufmann Publishers Inc..
- Hopkins, D. J., & King, G. (2010). A method of automated nonparametric content analysis for social science. *American Journal of Political Science*, 54(1), 229–247.
- House, E. R. (1980). *Evaluation with validity*. Beverly Hills, CA: Sage.
- Ian, F. (2014). wordcloud: Word Clouds. R package version 2.5. <https://CRAN.R-project.org/package=wordcloud>.
- Iliev, R., Dehghani, M., & Sagi, E. (2015). Automated text analysis in psychology: methods, applications, and future developments. *Language and Cognition*, 7(2), 265–290.
- Ingo, F., & Kurt, H. (2015). tm: Text Mining Package. R package version 0.6-2. <https://CRAN.R-project.org/package=tm>.
- Jacobi, C., van Atteveldt, W., & Welbers, K. (2016). Quantitative analysis of large amounts of journalistic texts using topic modelling. *Digital Journalism*, 4(1), 89–106.
- Johnson, L., Adams Becker, S., Estrada, V., & Freeman, A. (2014). *NMC Horizon Report: 2014 Higher Education Edition*. Austin, Texas: The New Media Consortium.

- Jurafsky, D., & Martin, J. H. (2008). *Speech and language processing* (2nd ed.). Prentice Hall.
- Kim, Y. M. (2009). Issue Publics in the New Information Environment, Selectivity, Domain Specificity, and Extremity. *Communication Research*, 36(2), 254-284.
- Lewis, S. C., Zamith, R., & Hermida, A. (2013). Content analysis in an era of big data: A hybrid approach to computational and manual methods. *Journal of Broadcasting & Electronic Media*, 57(1), 34-52.
- Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies*, 5(1), 1-167.
- Liu, L., Tang, L., Dong, W., Yao, S., & Zhou, W. (2016). An overview of topic modeling and its current applications in bioinformatics. *SpringerPlus*, 5(1), 1608.
- Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. H. (2011). *Big data: The next frontier for innovation, competition. and productivity*. McKinsey Global Institute.
- Maynard, D., & Funk, A. (2012). Automatic detection of political opinions in tweets. *In The semantic web: ESWC 2011 workshops* (pp. 88-99). Springer Berlin Heidelberg.
- McCallum, A. K. (2002). MALLET: Machine Learning for Language Toolkit. (URL: <http://mallet.cs.umass.edu>)
- Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 5(4), 1093-1113.
- Newman, D., Lau, J. H., Grieser, K., & Baldwin, T. (2010). Automatic evaluation of topic coherence. *In Human Language Technologies: The 2010 Annual Conference of the North*

- American Chapter of the Association for Computational Linguistics* (pp. 100–108). Association for Computational Linguistics.
- O’Callaghan, D., Greene, D., Carthy, J., & Cunningham, P. (2015). An analysis of the coherence of descriptors in topic modeling. *Expert Systems with Applications*, 42(13), 5645–5657.
- OECD(2012). *Better Governance for Inclusive Growth*. The session briefs for the OECD Global Forum on Public Governance. Retrieved December 1, 2015, from [http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=GOV/PGC/GF\(2012\)1&docLanguage=En](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=GOV/PGC/GF(2012)1&docLanguage=En).
- Ponweiser, M. (2012). Latent Dirichlet allocation in R. Diploma thesis, Vienna University of Economics and Business.
- Rehurek, R., & Sojka, P. (2010). Software framework for topic modelling with large corpora. *In Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks* (pp. 46–50). University of Malta, Retrieved September 7, 2016, from [https://radimrehurek.com/gensim/lrec2010\\_final.pdf](https://radimrehurek.com/gensim/lrec2010_final.pdf).
- Röder, M., Both, A., & Hinneburg, A. (2015). Exploring the space of topic coherence measures. *In Proceedings of the eighth ACM international conference on Web search and data mining* (pp. 399–408). Retrieved October 1, 2016, from [http://svn.aksw.org/papers/2015/WSDM\\_Topic\\_Evaluation/public.pdf](http://svn.aksw.org/papers/2015/WSDM_Topic_Evaluation/public.pdf)
- Scriven, M. (1974). Standards for evaluation of educational programs and products. In G. D. Borich (Ed.). *Evaluating educational programs and products*. NJ: Educational Technology.
- Schwab, K. (2017). 제4차 산업혁명[*The fourth industrial revolution*]. (송경진 역). 서울. 새로운현재. (원전은 2016에 출판)
- 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). Association for Computational Linguistics, Retrieved September 4, 2016, from <http://nlp.stanford.edu/events/illvi2014/papers/sievert-illvi2014.pdf>.
- Song, Z. (2010). Research on text categorization based on LDA. Master's Degree Dissertation. Xi'an University of Technology, Xi'an, China.
- Steyvers, M., & Griffiths, T. (2007). Probabilistic topic models. In Landauer, T. K., McNamara, D. S., Dennis, S. & Kintsch, W. (Eds.), *Latent Semantic Analysis: A Road to Meaning*. (pp. 427–448). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.
- Stufflebeam, D. L. (1968). *Evaluation as enlightenment for decision making*. OH: Ohio State University.
- 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.
- Tsytsarau, M., & Palpanas, T. (2012). Survey on mining subjective data on the web. *Data Mining and Knowledge Discovery*, *24*(3), 478–514.
- Tyler, R. W. (1949). *Basic principle of curriculum and instruction*. Chicago: The University of Chicago Press.
- Vesset, D., Woo, B., Morris, H. D., Villars, R. L., Little, G., Bozman, J. S., & Yezhkova, N. (2012). *Worldwide big data technology and services 2012–2015 forecast*. IDC Report, 233485.
- Wallach, H. M., Mimno, D. M., & McCallum, A. (2009). Rethinking LDA: Why priors matter. *In Advances in neural information processing systems* (pp. 1973–1981). Retrieved August 30, 2016, from [http://machinelearning.wustl.edu/mlpapers/paper\\_files/NIPS2009\\_0929.pdf](http://machinelearning.wustl.edu/mlpapers/paper_files/NIPS2009_0929.pdf)
- Wang, C., Paisley, J. W., & Blei, D. M. (2011). Online Variational

- Inference for the Hierarchical Dirichlet Process. In AISTATS 2(3), 4, Retrieved July 17, 2016, from <http://jmlr.csail.mit.edu/proceedings/papers/v15/wang11a/papers/v15/wang11a/wang11a.pdf>.
- Yu, S. Y. (2014). Exploratory study of developing a synchronization-based approach for multi-step discovery of knowledge structures. *Journal of Information Science Theory and Practice*, 2(2), 16-32.
- Zhou, Y., & Moy, P. (2007). Parsing Framing Processes: The Interplay Between Online Public Opinion and Media Coverage. *Journal of Communication*, 57, 79-98.
- 네이버 뉴스 홈페이지 <http://news.naver.com>(검색일: 2016. 5. 26.)
- 네이버 데이터랩 홈페이지 <http://datalab.naver.com>(검색일: 2016. 6. 1.)
- 뉴시스(2015. 7. 1). 초등교과서 한자병기 반대 운동본부 발족. [http://www.newsis.com/ar\\_detail/view.html/?ar\\_id=NISX20150701\\_0013763935&cID=10201&pID=10200](http://www.newsis.com/ar_detail/view.html/?ar_id=NISX20150701_0013763935&cID=10201&pID=10200) (검색일: 2016. 9. 31.)
- 뉴시스(2016. 8. 6). 내년 중학교 자유학기제 전면 시행. [http://www.newsis.com/ar\\_detail/view.html/?ar\\_id=NISX20150701\\_0013763935](http://www.newsis.com/ar_detail/view.html/?ar_id=NISX20150701_0013763935) (검색일: 2016. 9. 30.)
- gensim 홈페이지 <https://radimrehurek.com/gensim>(검색일: 2016. 9. 30.)
- KOMORAN 2.4 파이썬 버전 깃허브 [https://github.com/shineware/komoranPy\\_2.0](https://github.com/shineware/komoranPy_2.0) (검색일: 2016. 8. 18.)
- Python 홈페이지 <https://www.python.org>(검색일: 2016. 6. 27.)
- pyLDAvis 홈페이지 <http://pyldavis.readthedocs.io/en/latest/readme.html> (검색일: 2016. 8. 2.)
- R 프로그램 홈페이지 <https://www.r-project.org>(검색일: 2016. 6. 1.)

## 부 록

1. 정책 단계에 따른 분석 자료의 빈도수 ..... 171
2. 데이터 전처리에서 활용된 불용어 목록 ..... 173
3. 형태소 분석 단계에서 활용된 사용자 사전의 단어 목록 174
4. 토픽명 결정을 위한 전문가 협의회 자료 및 결과 .. 177



## < 부 록 1 >

### 정책 단계에 따른 분석 자료의 빈도수

구분	매체 유형	정책 형성		정책 결정		정책 집행		전체	
		빈도 (건수)	비율 (%)	빈도 (건수)	비율 (%)	빈도 (건수)	비율 (%)	빈도 (건수)	비율 (%)
JTBC	TV	0	0.00	0	0.00	4	0.48	4	0.26
KBS 뉴스	TV	5	7.46	15	2.42	17	2.06	37	2.45
MBC뉴스	TV	1	1.49	1	0.16	0	0.00	2	0.13
MBN	TV	0	0.00	3	0.48	14	1.70	17	1.12
SBS	TV	1	1.49	17	2.74	3	0.36	21	1.39
SBS 뉴스	TV	0	0.00	1	0.16	13	1.58	14	0.93
TV리포트	인터넷	0	0.00	0	0.00	1	0.12	1	0.07
TV조선	TV	0	0.00	3	0.48	0	0.00	3	0.20
YTN	TV	8	11.94	39	6.29	23	2.79	70	4.63
ZDNet Korea	인터넷	1	1.49	0	0.00	1	0.12	2	0.13
강원일보	신문·잡지	0	0.00	3	0.48	0	0.00	3	0.20
경향신문	신문·잡지	0	0.00	22	3.55	41	4.97	63	4.17
국민일보	신문·잡지	3	4.48	29	4.68	27	3.27	59	3.90
노컷뉴스	인터넷	1	1.49	22	3.55	52	6.30	75	4.96
뉴스1	인터넷	2	2.99	61	9.84	51	6.18	114	7.54
뉴스스	인터넷	6	8.96	41	6.61	60	7.27	107	7.08
더불어민주당	기타	0	0.00	2	0.32	3	0.36	5	0.33
데일리안	인터넷	0	0.00	0	0.00	6	0.73	6	0.40
동아일보	신문·잡지	3	4.48	11	1.77	21	2.55	35	2.31
디지털데일리	인터넷	0	0.00	0	0.00	3	0.36	3	0.20
디지털타임스	신문·잡지	2	2.99	1	0.16	3	0.36	6	0.40
레이디경향	신문·잡지	0	0.00	3	0.48	2	0.24	5	0.33
매일경제	신문·잡지	2	2.99	16	2.58	16	1.94	34	2.25
매일신문	신문·잡지	1	1.49	2	0.32	10	1.21	13	0.86
머니S	신문·잡지	0	0.00	2	0.32	7	0.85	9	0.60
머니투데이	신문·잡지	2	2.99	22	3.55	22	2.67	46	3.04
문화일보	신문·잡지	0	0.00	5	0.81	7	0.85	12	0.79
미디어오늘	신문·잡지	0	0.00	2	0.32	6	0.73	8	0.53

구분	매체 유형	정책 형성		정책 결정		정책 집행		전체	
		빈도 (건수)	비율 (%)	빈도 (건수)	비율 (%)	빈도 (건수)	비율 (%)	빈도 (건수)	비율 (%)
부산일보	신문·잡지	1	1.49	7	1.13	8	0.97	16	1.06
블로터	인터넷	0	0.00	1	0.16	0	0.00	1	0.07
서울경제	신문·잡지	1	1.49	13	2.10	10	1.21	24	1.59
서울신문	신문·잡지	1	1.49	11	1.77	17	2.06	29	1.92
세계일보	신문·잡지	1	1.49	17	2.74	21	2.55	39	2.58
스포츠서울	신문·잡지	0	0.00	0	0.00	1	0.12	1	0.07
시사IN	신문·잡지	0	0.00	2	0.32	3	0.36	5	0.33
아시아경제	신문·잡지	1	1.49	15	2.42	41	4.97	57	3.77
아이뉴스24	인터넷	2	2.99	0	0.00	4	0.48	6	0.40
여성신문	신문·잡지	1	1.49	2	0.32	1	0.12	4	0.26
연합뉴스	TV	4	5.97	56	9.03	84	10.2	144	9.52
연합뉴스TV	TV	1	1.49	2	0.32	6	0.73	9	0.60
오마이뉴스	인터넷	1	1.49	24	3.87	15	1.82	40	2.65
이데일리	신문·잡지	4	5.97	7	1.13	18	2.18	29	1.92
전자신문	신문·잡지	2	2.99	2	0.32	2	0.24	6	0.40
정책브리핑	기타	2	2.99	20	3.23	40	4.85	62	4.10
조선비즈	신문·잡지	0	0.00	1	0.16	1	0.12	2	0.13
조선일보	신문·잡지	0	0.00	0	0.00	12	1.45	12	0.79
조세일보	신문·잡지	0	0.00	1	0.16	0	0.00	1	0.07
주간경향	신문·잡지	0	0.00	0	0.00	1	0.12	1	0.07
주간동아	신문·잡지	0	0.00	1	0.16	0	0.00	1	0.07
주간조선	신문·잡지	0	0.00	0	0.00	1	0.12	1	0.07
중앙일보	신문·잡지	0	0.00	0	0.00	10	1.21	10	0.66
참세상	인터넷	1	1.49	0	0.00	0	0.00	1	0.07
청와대	기타	0	0.00	1	0.16	0	0.00	1	0.07
파이낸셜뉴스	신문·잡지	1	1.49	10	1.61	22	2.67	33	2.18
프레시안	인터넷	0	0.00	3	0.48	3	0.36	6	0.40
한겨레	신문·잡지	1	1.49	44	7.10	30	3.64	75	4.96
한겨레21	신문·잡지	0	0.00	2	0.32	2	0.24	4	0.26
한국경제	신문·잡지	2	2.99	10	1.61	7	0.85	19	1.26
한국경제TV	TV	1	1.49	0	0.00	1	0.12	2	0.13
한국일보	신문·잡지	0	0.00	23	3.71	26	3.15	49	3.24
헤럴드경제	신문·잡지	1	1.49	22	3.55	25	3.03	48	3.17
Total		67	100	620	100	825	100	1,512	100

## < 부 록 2 >

### 데이터 전처리 단계에서 활용된 불용어(stopwords) 목록

가람, 가운데, 가지, 가톨릭대, 각각, 각자, 각종, 강남, 강남구, 강원, 강원대, 강원도, 개개인, 개별, 개수, 개인, 개정, 개정안, 개편, 겨울, 결과, 결국, 경기, 경기도, 경기도, 경남, 경북, 경북대, 경상대, 경신고, 경우, 경인교대, 경주, 경희, 계열, 고등학교, 고려대, 고려대학교, 고양, 곳곳, 공보관, 공정위, 공주, 공화, 과정, 과천, 관련, 관련자, 광화문, 교과, 교과서, 교보문고, 교수, 교안, 교육, 교육감, 교육과, 교육과정, 교육부, 교육장, 구청, 그간, 그동안, 그때, 그랜드, 그룹, 그림, 그중, 금년, 금요일, 금호, 기간, 기사, 기자, 기점, 기존, 기획, 김원, 김현, 나머지, 나중, 날짜, 남부, 내년, 내달, 내역, 내외, 내일, 내정자, 내후년, 누구, 뉴욕, 다수, 다음, 다음날, 다음달, 단계, 단양, 단양군, 단체, 단체장, 담당, 담당자, 당시, 당장, 당초, 대건, 대구, 대다수, 대부분, 대사, 대상, 대성, 대신, 대전, 대전대, 대책위, 대표, 대회, 덕분, 덕성여대, 도내, 도민, 도중, 동국대, 동국대학교, 동시, 동안, 드림, 마음, 마찬가지로, 마트, 만약, 만일, 말씀, 매주, 며칠, 모두, 모습, 무엇, 문과, 문화관, 물론, 민정, 박명, 박백범, 반면, 발표, 방안, 법인, 별개, 병우, 보드, 보성, 본래, 본부장, 본청, 본회의장, 부문, 부분, 부산, 부산대, 부서, 부설, 부장, 부처, 부총리, 비롯, 비서관, 사대, 사람, 사무국장, 사무총장, 사설, 사실, 사실상, 사이, 사이트, 사회자, 산하, 상반기, 상임, 상하, 새해, 생각, 서강대, 서대문구, 서로, 서비스, 서울, 서울교대, 서울대, 서울대학교, 서울시, 서울특별시, 서원, 서정, 서초구, 선수, 성균관대, 성현, 성호, 세종, 세종로, 센터, 소속, 소연, 소장, 수개월, 수원, 수차례, 순간, 스스로, 시간, 시간대, 시군, 시절, 신성, 신승, 실제, 아래, 아무것, 아산, 아이, 아주대, 안녕, 안산, 앞뒤, 애초, 야간, 약간, 양쪽, 어른, 어제, 어제오늘, 업체, 여고, 여기, 여기저기, 여럿, 여름, 여부, 여사, 여의도, 연간, 연구관, 연구사, 연구소장, 연내, 연도, 연말, 연세대, 연일, 연초, 옛새, 영등포구, 영수, 영암, 영역, 영일, 예년, 예전, 예정, 오늘, 오늘날, 오래전, 오른쪽, 오바, 오전, 오후, 올해, 왼쪽, 우드, 우리, 우선, 우성, 우여, 원래, 원장, 원주, 월간, 위원장, 위의원, 위주, 은정, 은혜, 은희, 의장, 이진, 이과, 이날, 이남, 이내, 이달, 이듬해, 이때, 이름, 이미, 이번, 이병, 이사, 이사장, 이상, 이선, 이승, 이야기, 이영, 이외, 이자, 이제, 이전, 이정, 이제, 이종, 이준, 이중, 이지, 이차, 이틀, 이하, 이화여대, 이화여자대학교, 이후, 인천, 일명, 일반, 일부, 일수, 일일, 일자, 일주일, 일차, 자리, 자문관, 자신, 자체, 작년, 장관, 장소, 장학관, 재단, 재우, 저녁, 전국, 전날, 전남, 전남대, 전년, 전라남도, 전라북도, 전북, 전북도, 전월, 전직, 전주, 전차, 전체, 절대, 정기, 정도, 정우, 정작, 정책실장, 정해주, 정현, 제일, 제주, 제주도, 제주시, 조금, 조서관, 조원, 조차, 종로, 종로구, 주간, 주당, 주변, 주장, 주한, 중순, 중심, 중앙, 중앙대, 중학교, 지점, 지금, 지난날, 지난달, 지난번, 지난주, 지난해, 지방, 지부, 지역, 지역구, 직속, 직전, 직후, 진재, 차관, 차례, 차원, 차후, 창원, 처음, 처장, 천안, 첫날, 첫째, 청사, 청주, 청주시, 체계, 체제, 초기, 초등학교, 최근, 추후, 출신, 충남, 충남대, 충북, 충북대, 충청, 취재진, 토요일, 통합, 특정, 페이지, 평소, 포럼, 포함, 하나, 하늘, 하늘교육, 하루, 하반기, 하순, 하영, 하이, 학교, 학교정책관, 학기, 학년, 학년도, 학생, 한국, 한국교원대, 한남대, 한마디, 한성, 한양대, 한일, 한쪽, 한창, 한철, 한편, 해남, 해당, 핸디, 향후, 현우, 현재, 현행, 협의회, 협회, 호텔, 홈페이지, 홍익대, 회관, 회장단(이상 472개, 가나다순)

## < 부 록 3 >

### 형태소 분석 단계에서 활용된 사용자 사전의 단어 목록

AI, EBS, HIV, ICT, IT, KICE, KMOOC, NCS, NQF, NEAT, OECD, PC, PC방, PISA, SAT, SNS, STEAM, TF, UCC, UN, A형, B형, 가중치, 가치, 각계각층, 각주, 각론, 각론조정위원회, 간소화, 간첩, 간체, 갈등, 감감무소식, 감수본, 감수성, 감신정변, 갑오개혁, 강경, 강남, 강석화, 감염, 강원, 강화, 개고본, 개념, 개발, 개방형, 개신교, 개인, 개편, 개헌, 개혁, 거꾸로교실, 거점학교, 거창, 거창양민학살, 건강한사회를위한국민연대, 건국일, 건국절, 검정, 검정기준, 격랑, 결과물, 결정, 경기, 경상, 경우의 수, 경인교대, 경제, 경제수학, 경찰, 경축사, 계열성, 계획, 고고학, 고구려, 고구려연구재단, 고노 담화, 고대사, 고등학교, 고령화, 고시, 고유어, 고입, 골든타임, 공격, 공고, 공교육살리기 학부모연합, 공교육정상화법, 공과, 공급, 공동체, 공론, 공모, 공부, 공산주의, 공통교과, 과거사, 과정중심평가, 과포자, 과학, 과학기술기분법, 과학중점, 과학탐구, 과학탐구실현, 관련, 관리, 관리지침, 관찰, 광복군, 광장, 광주, 광화문, 교과, 교과교실제, 교과서, 교과중점과정, 교과중점학교, 교대, 교수, 교수역량, 교원단체, 교원대, 교원양성, 교원양성기관, 교원연구회, 교원연수, 교육감, 교육개혁추진협의회, 교육과정, 교육문화체육관광위원회, 교육부, 교육을바꾸는사람들, 교육인적자원부, 교육전문직, 교육지원청, 교육청, 교육행정, 교육현장, 교재, 교학사, 구글, 구성원, 구조, 구조조정, 국가경쟁력, 국가자격증, 국가주의, 국교, 국립특수교육원, 국민, 국민공통기본교육과정, 국민안전처, 국민통합, 국사편찬위원회, 국어, 국자, 국제, 국제성인역량평가, 국정, 군국주의, 권위주의, 군정, 규정, 근현대사, 금성, 기기, 기대, 기독교, 기술, 기술중점, 기업가정신, 기준, 기초교과, 기초과학연구원, 기초과학학회협의체, 기초소양, 기초학력, 기획재정부, 김구, 김근태, 김대중, 김대원, 김무성, 김성수, 김성철, 김일성, 김학중, 깜깜이, 나치, 날림, 남경순, 남북공동선언, 남북공동성명, 남북기본합의서, 남침, 남한, 내년, 내부, 내신, 내정, 내정자, 노근리, 노동, 노동기본법, 노무현, 녹색성장, 녹색성장위원회, 논술형, 논어, 농가, 농무성, 농성, 농축, 누리과정, 뉴라이트, 다문화, 단계, 단과대학, 담화서, 답정너, 당정, 당정청, 대강주의, 대규모, 대안, 대안교재, 대응, 대입전형, 대전대, 대주제, 대통령, 대학별고사, 대한민국인, 대한민국, 대한민국임시정부기념사업회, 대한수학회, 던지기, 도덕성, 도서, 도입, 도종환, 독도, 독립운동, 독일, 독재, 돌봄, 동기, 동독, 동물, 동북공정, 동북아역사재단, 동아시아사, 두산동아, 뒷북, 뒷전, 득세, 로드맵, 리베르, 마녀사냥, 마스터플랜, 마지막, 만족도, 맞춤형, 매주, 멀티미디어, 모바일, 모집단위, 무계중심, 무계열, 무리수, 무비판적, 무상급식, 무상물수, 무상분배, 무용지물, 무학년제, 문과, 문구, 문명, 문부과학성, 문자, 문재인, 문제, 문제은행, 문제풀이, 문제해결능력, 문학, 문해, 물리, 물수능, 뮤지컬, 미국, 미래부, 미래엔, 미래형, 미스매치, 미적분, 미주, 미취학, 민족문제연구소, 민족운동, 민족혁명당, 민족주의, 민주노총, 민주주의, 민주화, 민주화운동, 바른생활, 바칼로레아, 박근혜, 박병기, 박백범, 박수, 박정희, 반대, 반면교사, 반민족행위처벌법, 반민특위, 받아쓰기, 발달, 발달단계, 발행체제, 방과후학교, 배경식, 배희철, 백범, 백분위, 백년대계, 범교과, 범과정치, 범주법인, 베이징, 베트남, 변천사, 변화, 병기, 보도자료, 보수, 보통교과, 본질, 봉영식, 부분, 부총리, 부트캠프, 북벌, 북한, 분야, 분열, 분할모집, 불복종, 불안, 불일치, 불친절, 불통, 불룩타임, 비교과, 비공개, 비상교육, 비판, 빅데이터, 사고, 사교육걱정없는세상, 사교육비,

사교육총연합회, 사대주의, 사료, 사범대, 사사오입, 사상, 사실, 사일구혁명, 삼일운동, 삼차원, 사회, 사회문제탐구, 사회성, 사회적교육과정위원회, 사회주의, 사회중점, 사회탐구, 삭제, 산업화, 산학일체, 삼성전자, 상고사, 상대적, 상복, 상봉, 상상력, 상위권, 상하이, 상황, 새교육개혁포럼, 색깔론, 생활경제, 생활수준, 생활외국어, 서강대, 서남수, 서독, 서석진, 서술기준, 서술형, 서양, 서울, 서울교대, 서울대, 서울문화재단, 서인순, 서울중앙지법, 서책형, 선거, 선다형, 선도학교, 선박, 선율, 선택교과, 선택권, 선행학습, 설동호, 설명회, 설문조사, 성명, 성시윤, 성장, 성추행, 성취기준, 성취도, 성취평가, 성취평가제, 세계사, 세계시민교육, 세계적, 세계화, 세대, 세부능력및특기사항, 세습, 세월호, 세종, 세종대왕, 소논문, 소련, 소방안전교부세, 소수, 소양, 소재목, 소주제, 소프트웨어, 속전속결, 손질, 수능, 수립, 수시모집, 수업, 수업지원단, 수월성, 수요, 수정본, 수포자, 수학과제탐구, 수학여행, 수학익힘책, 수학지도법, 수행평가, 숙제없는학교, 순서, 순화어, 승배, 스마트교육, 스마트폰, 스웨덴, 스토리텔링, 스포츠클럽, 슬기로운생활, 시너지, 시대, 시대착오, 시매쓰, 시민단체, 시범학교, 시안, 시위, 시행령, 식민지근대화론, 식품의약품안전처, 신자유주의, 신형식, 실무교과, 실업, 실용수학, 심사, 심폐소생술, 실생활, 심미적감성, 심의, 심의기준, 심화선택, 씨마스, 아동, 아두이노, 아래, 아르, 아우성, 안정맞춤, 안양옥, 안전기준심의회, 안전종합대책, 안전처, 안전한생활, 안중근기념사업회, 안착, 알고리즘, 알파고, 야마토, 양성, 양정호, 양정현, 애국단체총협의회, 어문, 어문정책정상화추진회, 에듀톤, 여의도연구원, 여행지리, 역기능, 역사, 역사교육연구수, 역사관, 역사부도, 역사정상화추진단, 역사정의실천연대, 연구기관, 연구위원회, 연구원, 연구진, 연구학교, 연기, 연세대, 연임제, 연합형, 열사, 영광, 영국, 영어독해와작문, 영어회화, 영유권, 영토, 예고제, 예비고사, 예산결산특별위원회, 예술중점, 예이즈, 오수창, 올림피아드, 외교, 외국어중점, 외래어, 완성도, 완전, 왕조, 왜곡, 요구, 요소, 용대, 용산, 우경화, 우라늄, 우수사례, 우여곡절, 우파, 우편향, 운필력, 울산, 워크북, 워크숍, 원년, 원인, 위시, 위안부, 위원장, 위원회, 유관순, 유기홍, 유럽, 유발, 유상물수, 유상분배, 유신, 유튜브, 융합, 의열단, 의회, 응답하라 1988, 이과, 이공, 이기주의, 이덕환, 이명박, 이부섭, 이산가족, 이상, 이석문, 이성, 이세돌, 이수단위, 이승만, 이영, 이완구, 이의, 이익집단, 이차원, 이차함수, 이찬승, 이해당사자, 이화여대, 은수미, 응급처치, 인명, 인성, 인성교육진흥법, 인재, 인재상, 인정, 인정도서심의회, 인지, 인터넷, 인프라, 일반사회, 일반선택, 일방통행, 일본, 일본군, 일일, 일학습병행제, 일제강점기, 일주일, 임성호, 임법예고, 임베디드, 임시정부, 임용시험, 자가당착, 자격고사, 자기관리, 자기소개서, 자기주도학습, 자기주도학습전형, 자기평가, 자기성찰평가, 자문위원회, 자본주의, 자신감, 자위대, 자유민주주의, 자유발행제, 자유학기제, 자율교육학부모연대, 자체, 작품, 잘못, 장관, 장치, 재공모, 재구조화, 재난, 재난안전특별교부세, 재직자, 재활용, 쟁점, 저출산, 적정화, 전경련, 전국교수협의회, 전국수학교원모임, 전국역사교원모임, 전국한자교육추진총연합회, 전근대사, 전남, 전문가, 전문교과, 전문담임제, 전문성, 전용, 전자기학, 전자도서, 전적, 전체주의, 전태일, 전형, 절대평가, 절반, 절차, 점역, 정권, 정도, 정변, 정보, 정보기술과활용, 정보화사회, 정보통신기술, 정보통신산업진흥원, 정부, 정부세종청사, 정상화, 정시모집, 정오표, 정의, 정쟁, 정책, 정체성, 정치, 정통성, 정형화, 제기, 제도, 제시, 제작, 제주도, 제헌헌법, 조봉암, 조선, 조선민주주의인민공화국, 존재, 종교, 종로구, 중복, 종합캠퍼스, 좋은교과서만들기시민연대, 좋은교원운동, 좌경화, 좌파, 좌편향, 주도, 주목, 주진우, 주체사상, 중국, 중도, 중립, 중학교, 즉답, 즐거운생활, 증진, 지구과학, 지도, 지도서, 지도안, 지사, 지속가능발전, 지식정보처리, 지원, 지침, 지학사, 직선제, 직업탐구, 진로선택, 진보, 진재관, 김중이수, 김중학기, 김필기준, 김필진, 김희, 차관, 차지, 참교육학부모회, 참사, 창의, 창의체험활동, 창제, 천리마운동, 천안함, 천재교육, 청년, 청소년금융교육협의회,

청주, 체제, 초등학교, 초빙, 초중등교육법, 총괄평가, 총선거, 총학생회, 최고가격제, 최문순, 최상위권, 최저학력기준, 최종안, 중복, 친일, 컨설팅, 컴퓨터, 컴퓨터일반, 코딩, 콘텐츠, 쿠데타, 큰그림, 클라우드, 태극기, 태백, 태블릿, 태블릿PC, 데스크포스, 토크 콘서트, 통일, 통일교육, 통일교육원, 통폐합, 통합, 통합과학, 통합교과, 통합사회, 통합학급, 통합진보당, 투자, 트렌드, 특성화, 특수교육, 특수학교, 특수학급, 파병, 판단, 패러다임, 패자부활전, 페이스북, 편사부장, 편수용어, 편찬심의위원, 편찬심의회, 편향, 평등교육실현학부모회, 폐지, 포트폴리오, 표결, 표의문자, 표음문자, 표준점수, 풍선효과, 프랑스, 프로그래밍, 프로그램, 프로젝트, 핀란드, 필수교과, 하늘교육, 학교안전계획, 학교안전관리사, 학교안전관리위원회, 학교운영위원회, 학교정책관, 학교폭력, 학기, 학년, 학년도, 학력, 학력고사, 학령인구, 학벌, 학생부, 학생부교과전형, 학생부종합전형, 학생회, 학습, 학습권, 학습동기, 학습량, 학습부담, 학원, 학자, 한국, 한국고대사학회, 한국과학교육학회, 한국과학기술단체총연합회, 한국과학기술원, 한국과학기술한림원, 한국과학창의재단, 한국공학한림원, 한국교원대, 한국교원단체총연합회, 한국교육개발원, 한국교육고용패널, 한국교육과정학회, 한국교육정책교원연대, 한국교육정책연구소, 한국교육학술정보원, 한국교회연합, 한국근현대사학회, 한국기독교역사교과서공동대책위원회, 한국대학교육협의회, 한국독립당, 한국문화예술교육진흥원, 한국사, 한국사교과서국정화저지네트워크, 한국사능력검정시험, 한국사회의이해, 한국역사연구회, 한국예술영재교육연구원, 한국장로교총연합회, 한국전자통신연구원, 한국전쟁, 한국정보과학연합회, 한국정보교육학회, 한국정보과학회, 한국직업능력개발원, 한국청소년정책연구원, 한국초등국어교육학회, 한국컴퓨터교육학회, 한국학대학원, 한국학중앙연구원, 한국해운조합, 한글, 한글문화연대, 한나라당, 한자, 한자교육추진연합회, 한자수, 항일운동, 해결, 해법, 해축, 해커튼, 핵심개념, 핵심역량, 행복, 행정예고, 행정절차법, 허위, 헌법, 헌법재판소, 혐로, 헬조선, 혁신, 현수막, 현장, 현장검토본, 협력교원제, 협력종합예술활동, 형성평가, 호도, 호텔, 혼용, 홍익대, 화이부동, 확정, 환경, 활동지, 황교안, 황우여, 회원국, 획일, 효과, 후기, 휴머노이드(이상 953개, 가나다순)

## < 부 록 4 >

### 토픽명 결정을 위한 전문가 협의회 자료 및 결과

#### 1. 2015 개정 교육과정 정책 개요

##### 가. 정책 단계별 주요 추진 사안

※ 관련 제시 자료는 'II. 이론적 배경'의 [그림 II-1]의 내용(p. 20)과 동일함.

##### 나. 정책 주요 내용

※ 관련 제시 자료는 'II. 이론적 배경'의 <표 II-9>의 내용(pp. 45-46)과 동일함.

#### 2. 2015 개정 교육과정 정책의 토픽명(쟁점) 도출 과정

##### 가. 분석 방법

**\* 토픽 모델링**

- 텍스트 자료 내 단어들의 사용 빈도를 통계적으로 분석해 전체 자료를 관통하는 잠재적 주제(topic)들을 자동으로 추출하는 기법
- **잠재 디리클레 할당**(Latent Dirichlet allocation: LDA)을 활용해 2015 개정 교육과정 정책의 **주요 쟁점** 도출
  - 전제: 문서들은 다양한 단어들로 표현되는 토픽들로 구성되고, 토픽의 구성 비율은 베이지안 확률모형의 사후 분포를 통해 추정
    - 추정 방법: VEM(variational expectation-maximization), **깁스 표집**(Gibbs sampling)
  - 개수 설정 방법: ① 여러 가지 경우의 수를 고려해 분석한 후 **용어 분류의 정확성이 가장 높은 결과** 선택, ② 일종의 모형 적합도 지수인 **perplexity**와 **topic coherence** 지수 활용, ③ 확률 모델링을 최적화하여 비모수적 추정
- Python 프로그램의 **gensim**(Rehurek & Sojka, 2010) 패키지 활용
  - TF를 사용하되, 5회 미만 출현 단어는 제외

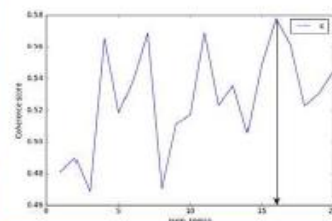
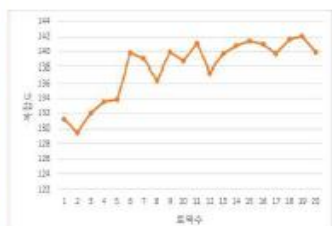
$$\text{Perplexity}(W) = \sqrt[N]{\prod_{i=1}^N \frac{1}{P(w_i|w_1 \dots w_{i-1})}}$$
$$\text{Coherence}_{Sim}(T) = \frac{\sum_{1 \leq i \leq n-1} \sum_{i+1 \leq j \leq n} \text{Sim}(w_i, w_j)}{\binom{n}{2}}$$

3

## 나. 분석 결과

### 3. 2015 개정 교육과정 정책의 쟁점 분석

#### 가. 분석 기간 전체



구분	1	2	3	4	5	6	7	8
	역사 교과서 국정화 관련 정치적 대립	역사 교과서 국정화 찬반 논란	역사 교육 내용	교육과정 운영 지원 체제 구축	신설 교과목의 특성	교육과정 개정의 필요성	역사 교과서 집필진 구성	역사 교과서 국정화 고시
문서 수	87	86	101	60	168	65	85	143
구분	9	10	11	12	13	14	15	16
	수능 대입제도 개선	교수 학습 방법 개선	초등 교과서 한자 명기	현장 교원의 의견 수렴	자유 학기제 운영	소프트웨어 교육 강화	안전 교육 강화	인성 교육 강화
문서 수	103	85	118	139	41	100	63	68

4

#### (1) 토픽1: 자유학기제 운영

〈기사 예시〉

내년 중학교 자유학기제 전면 시행(뉴시스, 2015-08-06)

내년부터 중학교에서 한 학기 동안 시험 없이 진로탐색 활동을 하는 '자유학기제'가 전면 시행된다. (중략)  
 '교과통합 진로교육'을 활성화해 학습 동기를 부여하고, 현장직업체험 등 다양한 진로 체험활동을 학교별로 2회 이상 실시한다. 진로활동 기록 사항을 상급 학교에 연계해 진로 상담·지도를 체계적으로 관리할 수 있도록 했다. 주제선택 활동과 동아리 활동은 학습동기 유발과 자율적 문제해결력 함양을 위해 학생 주도의 프로그램을 운영한다. (중략)  
 교육당국은 내년 전면 시행 시 필요한 체험처 4만7300개와 체험 프로그램 9만4600개를 확충한다는 목표다. (중략) 기업의 각종 시설을 개방하고, 산업현장 체험 프로그램 운영도 확대한다. (중략)  
 또 원격영상 진로멘토링과 진로체험 버스 방방곡곡 프로그램, 찾아가는 진로 체험단을 확대 운영해 농어촌 학생의 체험활동을 지원한다. (중략)  
 교육부는 이달 중으로 자유학기제의 법적 근거를 마련하기 위한 '초·중등교육법 시행령'을 개정하고, '2015 개정 교육과정'에서 자유학기제 편성·운영방안을 제시할 예정이다.  
 시안에 대한 의견수렴 절차를 거쳐 오는 10월 자유학기제 시행계획을 확정할 방침이다.

구분	13
1	운영
2	교육청
3	계획
4	자유학기제
5	지원
6	확대
7	마련
8	실시
9	협력
10	시행
문서수	41

문서No.	토픽1	토픽2	토픽3	토픽4	토픽5	토픽6	토픽7	토픽8	토픽9	토픽10	토픽11	토픽12	토픽13	토픽14	토픽15	토픽16
309	0.0178	0.0178	0.0221	0.0307	0.0178	0.0652	0.0135	0.0264	0.0221	0.0307	0.0135	0.0221	0.6212	0.0135	0.0393	0.0264



## (2) 토픽2: 교육과정 운영 지원체제 구축

구분	4
	교육과정 운영 지원체제 구축
1	대학
2	지원
3	추진
4	확대
5	개원
6	사회
7	확산
8	재정
9	직업
10	강화
문/수	60

### 〈기사 예시〉

2022년까지 고교 한 학급 학생 24명으로 줄인다(정책브리핑, 2016-04-25)

오는 2022년까지 고등학교 한 학급 학생 수를 OECD 수준인 24명으로 줄이는 방안이 추진된다. (중략)

교육부는 25일 시도 부교육감 회의에서 이런 내용이 포함된 '고교 맞춤형 교육 활성화 계획'을 발표했다.

이번 계획은 자기주도 학습과 학생 참여 수업이 중시되는 **2015 개정 교육과정**이 2018학년도부터 적용되고 중학교 자유학기제에서 진로·적성 교육을 경험한 학생들이 고교에 진학하는 점을 고려해 마련됐다. (중략)

학령인구 감소에 대비하고 중학교 자유학기제와 연계해 협력·토론형 수업이 **확산**할 수 있도록 교육부는 2022년까지 고교 학급당 학생 수를 24명으로 줄이기로 했다. 교사 1인당 학생 수 역시 같은 기간 16.6명에서 13.3명으로 개선을 방침이다. (중략)

**직업계** 고교생 비중은 현재 19% 수준에서 2022년까지 30%로 늘린다. (중략) 일반고 중에서 과학, 예술, 체육 교과를 중점적으로 가르치는 '교과 중점 학교'는 지난해 130곳에서 올해 200곳, 내년 300곳으로 **확대**한다.

올해부터 전체 중학교로 전면 확대된 자유학기제에 맞춰 고교 선택시험은 폐지가 추진된다.

아울러 '진로교육 집중학기제' 운영을 통해 37곳에서 2018년 100곳으로 늘리고, **직업계고(특성화고, 마이스터고)** 학생 비중은 2022년까지 30%로, 취업률은 65%로 끌어올릴 계획이다.

7

## (3) 토픽3: 안전 교육 강화

구분	15
	안전 교육 강화
1	안전
2	강화
3	예산
4	교육청
5	신설
6	재난
7	조직
8	대응
9	업무
10	정책
문/수	63

### 〈기사 예시〉

내년 초등 1~2학년부터 '4대 안전교육' 대폭 강화(연합뉴스, 2016-08-28)

내년 3월 새학기부터 초등학교에서 1~2학년 저학년층을 대상으로 생활, 교통, 신변, 재난 등 '4대 안전교육'이 대폭 **강화**된다.

교육부는 **2015 개정 교육과정**에 맞춰 내년부터 초등 1~2학년에 '안전한 생활'이 **신설**됨에 따라 교과서 현장검토본 개발을 거의 마치고 다음달 중순부터 의견 수렴을 한다고 28일 밝혔다.

'안전한 생활' 교과서에서 다루게 될 영역은 생활안전, 교통안전, 신변안전, **재난안전** 등 크게 4가지다. 4개 영역에는 학교·가정·사회에서의 각종 사고 예방, 화재나 천재지변시 대처법 등 각종 안전교육이 망라돼 있다. (중략)

특히 최근 잇따른 아동학대 사건으로 관심이 높아진 가정폭력 등 신변안전 교육은 사실상 신설되는 것이나 마찬가지다. 화재, 지진, 황사, 미세먼지 등 각종 재난·자연 현상에 대처하는 교육 역시 대폭 강화됐다.

다만 교육부는 '안전한 생활'을 기존의 통합 교과와 같은 독립 교과로는 신설하지 않고, 창의적 체험활동 시간을 이용해 수업을 하게 했다. 독립 교과로 만들면 교사나 학생의 학습 부담이 커지는 결과로 이어지기 때문이다.

교육부 관계자는 "세월호 사건 이후 안전교육을 강화해야 한다는 지적에 따라 '안전한 생활'을 독립 교과로 신설하자는 논의가 있었으나 공평성 등 의견 수렴 과정에서 창의적 체험활동에 편성하는 것으로 결론이 났다"고 말했다. (후략)

8

#### (4) 토픽4: 교육과정 개정의 필요성

구분	6
	교육과정 개정의 필요성
1	사회
2	문제
3	용한
4	창의
5	필요
6	학습
7	내용
8	인재
9	지식
10	해결
문서수	65

#### 〈기사 예시〉

국민행복·창의인재 양성, 교육개혁(정책브리핑, 2015-08-27)

박근혜 정부는 출범 후 교육개혁을 4대 개혁의 하나로 설정하고 지난 2년 반 동안 '행복교육과 창의인재 양성'을 목표로 다양한 교육개혁 과제를 추진하고 있다. (중략)

교육개혁을 성공적으로 추진해 국민행복과 창의인재 양성에 밑거름이 될 수 있도록 ▶공과 끼를 키우는 교육 ▶사회 수요에 부응하는 교육 ▶능력 중심 사회 구현이라는 3대 목표를 되새기고, 전문가·학부모·교원 등으로 구성된 교육개혁추진협의회를 운영해 핵심 개혁과제인 자유학기제 확산, 공교육 정상화 추진, 지방교육재정 개혁, 산업수요 맞춤형 인력 양성, 일·학습병행제 추진을 강화하기로 했다. (중략)

또한 자유학기제 기간에 학생들이 공·끼 교육을 받을 수 있도록 교육과정을 혁신한다. (중략)

교육부는 학교가 유발하는 사교육 요인을 제거해 공교육 정상화 토대를 도모한다. 지난해 9월부터 '공교육 정상화법'이 시행됨에 따라 사교육을 유발하는 학교의 선행 출제가 금지될 수 있도록 유도하고, 학교 교육과정 운영 및 선행 출제 여부 등을 주기적으로 점검한다. 아울러 공교육 정상화법을 보완하기 위해 방과 후 학교 선행교육 제한을 완화하는 내용의 법 개정을 추진한다.

중·고교 교육과정 시간 운영회, 교육과정심의회의 등을 통한 의견을 수렴해, 9월에 2015 개정 교육과정(문·이과 통합형)도 확정해 고시한다. 이로써 문·이과 구분 없이 기초 소양을 함양할 수 있는 공통과목을 개설하고 진로를 반영할 수 있는 다양한 선택과목을 마련한다. 즉 국어, 수학, 영어, 한국사, 통합사회, 통합과학, 과학탐구실험 등 교과별 핵심 개념을 중심으로 학습량을 적정화하고, 협력·토론 수업과 안전교육 소프트웨어(SW) 교육을 강화한다. (후략)

9

#### (5) 토픽5: 인성교육 강화

구분	16
	인성 교육 강화
1	활동
2	예술
3	연극
4	체육
5	체험
6	인성
7	강화
8	음악
9	교육지원청
10	문화
문서수	68

#### 〈기사 예시〉

서울 중학생들 정규수업으로 유지철·연극 배운다(연합뉴스, 2016-09-08)

(전략)

서울시교육청은 8월 학생들의 예술적 감수성과 창의력, 협력적 인성을 함양하기 위한 '중학교 협력종합예술활동' 운영 기본계획을 하반기부터 단계적으로 추진한다고 밝혔다.

'협력종합예술활동'은 중학교 3년 과정 중에 최소 한 학기(총 17시간) 이상 교육과정 내에서 유지철·연극·영화 등 종합예술활동에 모든 학생이 역할을 분담해 참여하고 느낀 점들을 발표하도록 하는 내용의 예술 체험교육이다.

교육청은 '서울시 중학교 교육과정 편성 운영지침'에 이번 계획을 반영, 교과 수업이나 창의적 체험활동 시간에 협력종합예술활동을 편성해 교육하도록 할 방침이다.

교육청은 서울시와 다양한 문화예술기관, 예술단체 등과의 협력방안을 마련해 교육의 시너지효과를 높이는 한편, 연습과 공연장소를 확보하고 운영학교에 예산 지원을 추진한다.

(중략)

교육청은 내년에는 희망 학교를 중심으로 협력종합예술활동을 운영하고, 2015 개정 교육과정이 적용되는 2018년부터 2020년까지는 단계적으로 늘려 모든 중학교로 확대할 계획이다.

교육청은 "종합예술 교육은 상호존중 역량을 가진 민주시민을 기르는 데 목적이 있다"며 "지식과 기능에 치우친 초·중등학교 예술교육을 표현력과 협력 중심의 과정으로 전환할 수 있도록 지속해서 문화예술교육 활성화와 정책적 뒷받침 계획"이라고 밝혔다.

10

## (6) 토픽6: 교수·학습방법 개선

구분	10
	교수·학습 방법 개선
1	교원
2	수업
3	참여
4	개선
5	연수
6	방법
7	환경
8	현장
9	교실
10	토론
문/수	85

### 〈기사 예시〉

학습량 줄여 핵심 개념 중심으로 사고력 기르는 독서·토론 활성화(한국일보, 2015-09-17)

교육부의 '2015 교육과정 개정'의 골자는 핵심 개념 중심으로 한 학습량 경감, 학생 참여 중심의 수업, 창의적 사고 과정 유도다. 구체적으로 수업현장에서 독서 활동을 적극 활용할 계획이다. (중략)

실생활에 기반한 토론 학습도 도입할 예정이다. 수학과목에서는 기존에 복잡한 수식으로 수업을 진행했던 것과 달리 실생활에서 발생하는 사례를 토론 과제로 제시하기로 했다. 통계청 자료를 활용, 한국의 사회·자연환경·경제 등의 주제에 대해서도 실용적인 접근을 시도한다. 고교 과정에서 신설되는 '통합과학'에선 문·이과 학생들이 모두 쉽게 받아들일 수 있는 '지속가능한 친환경 에너지 도시 설계하기' 등을 학습 주제로 제시할 계획이다.

이밖에도 과다한 학습량을 진도 맞추기 식으로 가르치거나, 어려운 이론을 암기하게 하는 기존 방식을 지양하고 학생들의 창의력과 융·복합적 사고력을 신장시키는 방향으로 교육과정을 개편할 방침이다. 교육부 관계자는 "학습량을 줄여 확보된 시간에 학생들이 참여할 수 있는 다양한 형태의 수업을 진행하도록 하겠다"고 말했다. (중략)

교육부 관계자는 "중단원을 통합하거나 폐지하는 등 방법으로 전과목 합쳐 기존 학습량 대비 15% 가량 경감하는데 성공했다"며 "교육과정이 실제 적용될 때까지 교사 연수, 매뉴얼 배포 등을 통해 수업 현장의 실질적 변화를 유도할 계획"이라고 말했다.

11

## (7) 토픽7: 역사교과서 집필진 구성

구분	7
	역사교과서 집필진 구성
1	집필진
2	개발
3	국사편찬위
4	구성
5	역사
6	집필
7	집필기준
8	공개
9	검토
10	확정
문/수	85

### 〈기사 예시〉

국정교과서, 심의위원회 '깜깜이'(경향신문, 2015-11-30)

교육부가 국정 역사교과서 편찬준거와 집필 원고 심의할 편찬심의위원회 숫자만 밝히고 명단을 공개하지 않았다. 집필진에 이어 심의도 '깜깜이'로 진행되고 있다.

교육부는 30일 '2015 개정 교육과정'에 따른 '중등 역사과 교과용도서 편찬심의회'를 초빙과 공모를 통해 총 16명으로 구성한다고 발표했다. 교육부는 "편찬심의회에 학자와 현장 교사, 학부모가 포함돼 있다"고 했으나 각각의 숫자나 초빙·공모가 몇 명씩인지 밝히지 않았다. 교육부는 "초등 국정도서의 경우 교과서 집필이 완료된 후 편찬심의회 명단을 공개해왔으며 김정심의위원회 김정심사 종료 후 공개해왔다"면서 "그간의 전례에 따라 교과서 집필이 끝난 후 현장 검토 과정에서 편찬심의위원회 명단을 공개할 방침"이라고 밝혔다.

국정 역사교과서의 편찬준거 발표 일정도 불투명하다. 국편은 국정과 확정고시 다음날인 지난 4일 기자회견에서 편찬준거를 11월 말 김정배 국편위원장이 직접 발표할 것이라고 밝혔으나, 30일까지도 분명한 답을 내놓지 못했다. 국편 관계자는 "12월 7일 전후로 예정하고 있으나 교육부와 조율이 끝나지 않았다"고 말했다. 교육부 관계자는 "지금부터 정해야 한다"며 날짜를 밝히지 않았다. 편찬준거 발표 약속이 늦춰지면서 11월 말 착수하겠다는 국정교과서 집필도 미뤄지게 됐다.

12



## (8) 토픽8: 역사 교과서 국정화 찬반 논란

구분	2
	역사 교과서 국정화 찬반 논란
1	역사
2	국정
3	한국사
4	국민
5	반대
6	대한민국
7	사회
8	국가
9	교원
10	정부
문/수	86

### 〈기사 예시〉

‘교과서 개발’ 공정회... 국정화 찬반 의견으로 갈등(MBN, 2015-09-11)

오늘(11일) 오후 국사편찬위원회에서 **2015 개정 교육과정**에 따른 역사 교과서 편찬 준비 개발사인 공정회가 열렸습니다.

공정회에서는 중학교 **역사**와 고등학교 **한국사** 등 4개 교과서의 집필 기준을 개발하는 연구진이 참석해 현재까지 나온 집필 기준을 발표하고, 이에 대한 각계의 의견을 수렴했습니다.

이 자리에서 한국사교과서 집필 기준 연구에 참여한 학자 일동은 공정회 중간, 역사 교과서 **국정화**에 대해 **반대**한다는 뜻을 밝히기도 했습니다.

이에 공정회에 참석한 보수 학부모 단체는 교과서 국정화에 찬성한다며 맞불 기자회견을 했습니다.

공정회에 앞서 전교조를 비롯한 교육시민단체 일부는 황우여 교육부장관이 이번 공정회에 대한 안내를 2주 전에 하지 않아 행정절차법을 어겼다며 검찰에 고발장을 제출했습니다.

13

## (9) 토픽9: 역사 교과서 국정화 관련 정치적 대립

구분	1
	역사 교과서 국정화 관련 정치적 대립
1	야권
2	정부
3	국회의원
4	대통령
5	국회
6	박근혜
7	여권
8	정치
9	정권
10	비판
문/수	87

### 〈기사 예시〉

‘교과서 전쟁’ 교문위 국감, 여야 충돌로 정회(머니투데이, 2015-09-10)

**국회** 교육문화체육관광위원회의 교육부 국정감사가 한국사 교과서 국정화 논란으로 시작 한 시간여 만에 정회했다. (중략)

먼저 유기를 새정치민주연합 **의원**은 황우여 사회부총리 겸 교육부 장관의 증인선서가 끝난 직후 의사전행 발언을 통해 **정부가** 한국사 교과서 국정화를 강행하려는 것은 **박근혜 대통령**의 집착 때문"이라면서 황 부총리의 명확한 입장 표명을 촉구했다. (중략)

이에 대해 황 장관은 "주요 업무보고 3쪽에서 15쪽에 걸친 **2015 개정 교육과정**과 교과서 개선 부분에 간략하나마 보고를 드렸다"며 "지면관계상 충분하지는 못하지만 보고는 드렸다"고 말했다.

신성범 새누리당 의원도 "업무보고 15페이지에 부족하긴 하지만 확정고시 예고 등 업무보고의 요건을 가지고 있다고 생각한다"며 "일단 업무보고를 받아보고, 이 내용을 토대로 질의를 통해서 교육부의 입장 듣는 게 온당하다"고 말했다.

이후 **여야** 의원들의 의사전행발언이 계속됐고, 박주선 교문위원장은 결국 정회를 선언했다.

14

# (10) 토픽10: 역사교육 내용

구분	3
	역사교육 내용
1	대한민국
2	북한
3	정부
4	수립
5	근현대사
6	서술
7	일본
8	내용
9	역사
10	주제사상
문/수	101

## 〈기사 예시〉

대한민국 정부 수립' 표현, 건국절 연상케 '대한민국 수립'으로 바뀌(경향신문, 2015-09-22)

(전략)

새정치민주연합 '한국사 교과서 국정화 저지 특별위원회'(위원장 도종환)가 22일 교육부에서 제출받은 '2015 개정 교육과정' 역사과 교육과정을 보면, 중·고교 모두 현대사 부분에서 2009 교육과정의 '대한민국 정부 수립'이라는 표현이 '건국'을 연상하게 하는 '대한민국 수립'으로 변경돼 있다. '대한민국 정부 수립'과 '대한민국 수립'은 앞서 1955년 1차 교육과정부터 1997년 7차 교육과정까지는 혼용되다 2009 교육과정부터는 '대한민국 정부 수립'으로 통일됐다.

하일식 연세대 사학과 교수는 "1948년 8월15일을 정부 수립일이 아니라 건국일로 보는 뉴라이트 학자들의 주장을 반영한 것"이라며 "건국"이라는 표현을 집어 넣으면 논란이 생길 것을 우려해 우회적인 표현을 사용한 것 같다"고 말했다. 이종식 민족문제연구소 연구위원은 "뉴라이트 학자들의 주장은 대한민국이 임시정부 법통을 계승한다고 명시한 헌법을 위배하는 것"이라며 "이처럼 민감한 문제를 왜 이런 식으로 처리했는지에 대해 교육부가 해명해야 한다"고 말했다. (중략) 중학교 역사 교육과정에서는 북한 관련 내용이 빠진 것으로 드러났다. 2009 교육과정에서는 '1970년대 이후 탈북자 문제'와 북한의 실상이 내용요소로 포함돼 있었지만 2015 교육과정에서는 '탈북자 문제를 위한 노력'이라고만 나와 있다.

15

# (11) 토픽11: 역사 교과서 국정화 고시

구분	8
	역사 교과서 국정화 고시
1	국정
2	검정
3	한국사
4	고시
5	정부
6	확정
7	논란
8	인정
9	수정
10	전환
문/수	143

## 〈기사 예시〉

역사교과서 국정화 확정고시...황우여 "국정교과서로 국민 통합 이룰 것(동아일보, 2015-11-03)

2017년부터 중·고교 역사교과서가 검정에서 국정으로 바뀐다. 황우여 사회부총리 겸 교육부 장관은 이러한 내용의 중·고등학교 교과용도서 국·검·인정 구분(안)을 3일 확정해 고시했다.

황 부총리는 "현행 역사교과서의 검정 발행 제도로는 올바른 역사 교과서를 만드는 것이 현실적으로 불가능하다는 것이 정부의 판단"이라며 "편향된 표현을 부분적으로 수정한다 해도 근본적으로 고칠 수 없고 학교의 자율적인 교과서 선택권마저 원천적으로 배제되고 있어 검정의 장점이랄 할 수 있는 다양성을 훼손하고 있다"고 말했다.

이어 "역사교과서가 이념 논쟁의 도구가 되어 국론 분열과 사회적 갈등을 야기해 왔다"며 "더 이상 역사교과서로 인한 사회적 논란을 막고 역사교육을 정상화해 국민통합을 이루기 위해 국가의 책임으로 올바른 역사교과서를 발행하기로 했다"고 강조했다. (중략)

황 부총리는 "새 국정 교과서가 친일, 독재 미화과 같은 역사 왜곡을 위한 것이라는 우려를 불식하기 위해 교과서 개발 전 과정을 전문가 감수와 전문가 교사연구회 검토, 입 전서 등으로 집필부터 발행까지 투명하게 운영하겠다"고 밝혔다.

▼ 다음은 황우여 사회부총리·교육부 장관 총리 말화 전문 ▼ (하략)

16

## (12) 토픽12: 소프트웨어 교육 강화

구분	14
	소프트웨어 교육 강화
1	소프트웨어
2	과학
3	기술
4	개발
5	정보
6	디지털
7	미래부
8	활용
9	필수
10	진행
문/수	100

### 〈기사 예시〉

내년 중학교 입학생부터 SW교육 의무화(디지털타임스, 2014-07-23)

정부가 교과과정을 개편, 바르면 내년부터 초·중등학교에서 **소프트웨어** 교육이 **필수**과목이 된다. (중략)

**미래창조과학부**는 SW중심사회 실현을 위한 청사진을 제시하고, 교육부는 초·중등 SW교육 활성화방안을, 산업부는 제조업의 SW융합 및 **활용**전략을, 문체부는 SW저작권 보호·이용기반 확산방안을 각각 발표했다. (중략)

이에 따르면 미래부와 교육부가 협의, 오는 2017년부터 적용되는 **문이과 통합형 교육과정** 수립에 SW 조기교육을 대대적으로 강화하는 방안이 추진된다. 초·중등학교는 SW 소양교육이 필수과목으로 도입되고, 고등학교에서 그간 심화선택과목으로 시행되어온 **정보** 과목을 SW 과목으로 변경하고 이를 일반선택과목으로 전환한다. (중략)

교육부 박재운 차의인재정책관은 "이같은 교과과정 개편이 2017년부터 적용될 예정이나 하반기에 이를 시행지침 형태로 각급학교에 하달, 바르면 2015년부터 이를 적용하는 방안을 논의 중"이라고 밝혔다.

이에 더해 전국 4개 권역의 대학교 부설로 정보보호 영재교육원을 설치하는 한편 미래부가 72개 초·중등학교를 하반기에 SW 교육 시범학교로 지정하고 내년에는 미래부·교육부가 이를 130개고 이상으로 확대한다. (하략)

17

## (13) 토픽13: 수능·대입제도 개선

구분	9
	수능·대입제도 개선
1	수능
2	평가
3	대입
4	시험
5	영어
6	제도
7	출제
8	학부모
9	부담
10	반영
문/수	103

### 〈기사 예시〉

수능 7개 과목 필수... 문항 수 시험 시간 조정 불가피(국민일보, 2015-09-25)

'2015 개정 교육과정'이 확정되면서 관심은 **대학수학능력시험** 등 **대입제도**의 변화로 모아지고 있다. (중략)

새 교육과정은 국어, 수학, **영어**, 한국사, 통합사회, 통합과학, 과학탐구실험을 고교 공통과목으로 정했다. 문·이과를 막론하고 이 7개 과목을 반드시 배워야 한다. 이 과목들은 수능 필수과목이 될 것으로 보인다. (중략)

새 수능에서 필수과목이 확대되면 **시험시간**이 대폭 늘어날 가능성이 높다. 수험생은 하루에 7개 과목을 치르는 **부담**을 안게 된다. (중략)

새 교육과정은 고교 교과를 심화 수준에 따라 '공통과목' '일반선택' '진로선택' '전문교과'로 구분했다. 수능이 어느 수준까지 심화될지는 **학부모·학생**의 관심사다. (중략)

김두용 교육부 대입제도과장은 "현재도 수능이 전체 대입에서 차지하는 비중이 생각보다 강하지 않다"며 "(진로선택 등 선택과목 같은 경우에 수능 범위에서 제외된다면) 대학의 모집단위 특성에서 요구하는 부분이 있으면 학생부를 통해 반영되도록 유도하는 게 바람직하다"고 말했다. 수능보다는 학생부 강화를 암시한 발언으로 풀이된다.

변수는 내신 **평가** 방식이다. 교육부는 2017년 수능 개편안과 함께 성취평가제 **반영 여부**도 발표하겠다고 예고했다. (하략)

18



#### (14) 토픽14: 초등 교과서 한자 병기

구분	11
	초등 교과서 한자 병기
1	한자
2	병기
3	한글
4	학부모
5	정책
6	방법
7	반대
8	국어
9	연구
10	학습
문사수	118

##### 〈기사 예시〉

초등교과서 한자병기 찬반 양론 격돌(뉴스1, 2015-08-24)

(전략) 김정자 국가교육과정개정연구위원장은 "초등학교 한자교육은 2009 개정 교육과정 이전에도 시도 교육청이나 학교 단위에서 자율적으로 다양한 방법과 수준으로 실시되고 있었다"며 "이번 2015 개정 교육과정과 관련해 한자교육 활성화 방안이 논의되고 있는 것은 시의적절한 것으로 보인다"고 발언을 엮었다. (중략)

2010년 설문조사에서는 어휘력 신장 차원에서 교사 77%, 학부모 89%가 한자교육을 찬성한다는 결과가 나오기도 했다. (중략)

초등학교 교과서 한자병기 정책이 '사교육 활성화 정책'이라는 비판은 여전했다. 한희정 서울 유현초등학교 교사는 "교과서에 한자를 병기하게 되면 현실적으로 교사나 학생 모두에게 학습 부담으로 작용할 수밖에 없다"며 "77%가 초등학교 한자교육에 찬성한다고 했던 바로 그 설문에서 교사들이 초등학교 한자교육을 반대하는 이유로는 학생의 학습부담(39.2%), 충분히 가르치고 배울 수 있는 시간 확보 어려움(29.5%)을 꼽았다"고 지적했다. "얼마 벗속에서부터 영어 사교육을 위해 영어 태교를 하고 0세부터 한글 사교육이 시작되는 나라, 그런데 정규 국어 교육과정에서 한글을 가르치는 시간은 고작 24차시에 불과한 나라에서 태어나서 자라나는 미래세대에게 정말 교육부가 해야 할 일은 무엇인가'라는 질문을 던진 한 교사는 "더 이상 우리 아이들을 학습하는 기계, 괴물 같은 존재로 키우고 싶지 않다"고 강조했다. (하략)

19

#### (15) 토픽15: 현장교원 의견 수렴

구분	12
	현장교원 의견 수렴
1	공청회
2	의견
3	고시
4	현장
5	총론
6	논의
7	국가
8	연구
9	지적
10	전교조
문사수	139

##### 〈기사 예시〉

내달 확정 '교육과정 개정안' 2차 공청회서도 숭한 문제제기(경향신문, 2015-08-18)

교육부가 추진 중인 '2015 교육과정 개정'의 총론·각론 확정 고시가 한 달 앞으로 다가왔다. (중략) 하지만 7~8월 교과목별로 진행된 2차 시안 공청회도 숭한 쟁점과 의문점이 제기되고 있다. 역사과 교육과정에서는 세계적 추세와 반대로 근현대사 비중을 축소하고, 한국사 교과서의 국정화 논란도 진행 중이다. 교육부는 '수포자'(수확포기자)를 없애기 위해 학습량을 크게 줄였다고 자평했지만 교육현장에서는 고개를 젓고 있다. 문·이과 구분 없이 모든 학생들이 배워야 하는 통합사회와 통합과학도 학습량이 줄어들지 않은 데다 현장 교사들의 준비도 부족해 실효성에 물음표가 쳐진 상황이다. (중략)

통합형 교육과정 취지에 반대하는 이들은 거의 없다. 그러나 대입 수능시험 개편 방안이 나오지 않아 현장 혼란을 초래할 우려가 높다는 지적이 제기된다. (중략)

현장 교사들이 통합사회와 통합과학 등 통합과목을 가르칠 준비가 돼 있지 않다는 점도 2015 교육과정 개정의 문제점으로 꼽힌다. 현직 고교 교사인 김옥훈 역사교육연구소장(독산고)은 "통합과목이 생기더라도 현재 각 과목 교사들이 가르칠 수밖에 없는 것이 교육현장의 판단과 현실"이라며 "준비가 안된 현장에서는 기가 막힌 상황이고 결국 아무렇게 가르쳐야 한다는 얘기"라고 말했다.

20

## (16) 토픽16: 신설 교과목의 특성

구분	5 신설 교과목의 특성
1	수학
2	과학
3	사회
4	영어
5	국어
6	내용
7	학습
8	신설
9	공통교과
10	수업
문서수	168

### 〈기사 예시〉

2018년 고1부터 문·이과 공통 통합사회·통합과학 등 신설(헤럴드경제, 2015-09-22)

2018학년도부터 고등학교 문·이과 구분 없이 공통으로 배우는 통합사회, 통합과학 등의 과목이 **신설**되고 연극, 소프트웨어교육 등 기초소양 교육이 강화된다. 이에 따라 초·중·고등학교 **수업**에서 수학, 영어를 비롯한 교과별 **학습** 부담이 줄어든 전망이다. 교육부는 23일 이 같은 내용을 골자로 한 '2015 개정 교육과정'을 확정·고시할 예정이라고 22일 밝혔다. (중략)

이번 교육과정 개정의 핵심은 무엇보다 고등학교 수업의 변화를 가져오는데 있다. 고등학생이 문·이과 구분없이 배우는 '**공통과목**'을 도입하고 진로, 적성에 따른 맞춤형 교육을 강화한 것이 특징이다.

고등학생이 필수로 듣는 공통과목으로 **국어, 수학, 영어, 통합사회, 통합과학**, 한국사, 과학탐구실험 등 7개가 신설된다.

현재 고교 교육과정이 모두 선택과목으로 돼 있는 상황에서 일부 학생들의 '지식편중' 현상을 개선하고 기초소양 교육을 제대로 하자는 취지다.

통합사회는 기존 지리, 일반사회, 윤리, 역사의 기본적인 **내용**을 담고 통합과학은 현재 물리학, 화학, 지구과학, 생명과학 등의 과목을 핵심개념 위주로 통합하는 것이다. 교육부는 통합사회나 통합과학을 여러 명의 교사가 내용을 나눠서 가르치는 것이 아니라 교사 1명이 교육하는 방안을 구상하고 있다. (하략)

21

## 나. 정책 단계별

### 정책 형성 단계

구분	1 과학교육 축소	2 수능·대입제도 개선	3 소프트웨어 교육 강화	4 공통 교과 신설	5 역사 교과서 국정화
1	과학	수능	소프트웨어	사회	교원
2	기술	도입	정책	이수단위	문제
3	수학	대입	미래부	과학	검정
4	학계	정부	추진	필수	국정
5	논의	평가	확대	영어	내용
6	의견	확정	정보	수업	수정
7	연합회	융합	개발	국어	연구
8	요구	개발	지원	공통교과	전환
9	국가	선택	인재	내용	현장
10	포럼	한국사	강화	신설	고시
문서수	16	12	12	16	11

22



# (1) 토픽1: 과학교육 축소

구분	1
	과학교육 축소
1	과학
2	기술
3	수학
4	학계
5	논의
6	의견
7	연합회
8	요구
9	국가
10	포럼
문서수	16

## 〈기사 예시〉

문이과 구분 폐지에 과학계 폭발... "통합하고 과학교육 축소?" (머니투데이, 2014-09-03)

문이과 구분 폐지를 주요 골자로 한 교육부의 국가교육과정 개정안이 과학기술계의 격렬한 저항에 부딪혔다. (중략) 3일 한국과학기술단체총연합회를 비롯해 한국과학기술한림원, 한국공학한림원 등 11개 단체는 서울 여의도 중소기업중앙회 대회의실에서 기초과학학회협의체 주관으로 열린 '창조경제 시대의 미래인재양성교육 국민대토론회'에서 미래 인재상 제시가 명확치 않고, 10명 정도로 소수인 '국가교육과정 개정연구위원회(이하 개정위)'의 설문을 결정이 우려되는 현 교육과정 개정 작업을 즉시 중단하라고 촉구했다. (중략) 과학교육 교육 비중은 2009년에는 15.1%였으나 개정안의 비중은 10.8%에 지나지 않는다. 과학이 물리·화학·생물·지구과학의 4과목으로 구성되는 만큼, 과목당 필수 교과시간은 2.5단위에 불과하다. 이대로라면 최소한의 과학교육도 불가능한 상황이란 게 과기계 설명이다. (중략) 다만, '반드시 가르쳐야 할 과목'으로 과학은 현 '축소'보단 '강화'로 선회해야 한다는 주장을 펼치며, 교육부와 평평한 대립양상을 보이고 있다. 교육부는 과기계가 반발하자 각계 의견을 들었으며, 급세 입장을 바꿔 '국가교육과정개정 자문위원회'를 급조했으나, 자문위원회가 정책 결정에 영향력을 크게 미칠 수 없는 구조인 데다 이는 순서가 완전히 뒤바뀐 것이라고 질타하며, 과기계는 사법계 출신들로만 구성된 개정위를 과학계와 산업계를 포함한 여러 분야의 전문가들이 참여할 수 있도록 재구성할 것을 요구하고 있다. (하략)

23

## | 정책 결정 단계(I)

구분	1	2	3	4	5	6
	역사교과서 국정화	초등 교과서 한자 병기	소프트웨어 교육	현장교원 의견 수렴	수학교육 내용 적정화	수능·대입 제도 개선
1	국정	한자	수업	교원	수학	수능
2	역사	병기	소프트웨어	현장	내용	평가
3	한국사	한글	시안	의견	학습	대입
4	검정	방법	운영	고시	학습량	영어
5	반대	반대	체육	공청회	시안	문제
6	개발	정책	활동	적용	학습부담	시험
7	정부	공청회	자유학기제	국가	사항	제도
8	인정	제시	정보	충론	국어	대학
9	검토	활성	교원	논의	사고	도입
10	학계	연구	체험	추진	성취기준	공교육
문서수	66	71	37	57	72	50

24

## | 정책 결정 단계 (2)

구분	7	8	9	10	11
	역사 교육 내용	교육과정 개정의 필요성	신설 교과목의 특성	역사 교과서 국정과 관련 정치적 대립	안전 교육 강화
1	역사	창의	사회	정부	안전
2	대한민국	융합	과학	야권	계획
3	국가	사회	신설	국회의원	지원
4	일본	과학	국어	박근혜	추진
5	근현대사	인재	공통교과	국회	대학
6	정부	기술	진로	대통령	재난
7	비중	필요	강화	국감	개혁
8	내용	개발	이수단위	규정	확대
9	축소	변화	선택교과	여권	현장
10	수립	환경	적용	진행	정부
문서수	39	67	86	37	38

25

## (1) 토픽1: 수학교육 내용 적정화

구분	5
	수학교육 내용 적정화
1	수학
2	내용
3	학습
4	학습량
5	시안
6	학습부담
7	사항
8	국어
9	사고
10	성취기준
문서수	72

### 〈기사 예시〉

교사들 "수학 교육과정 시안, 학습부담 경감 노력 미흡" (SBS, 2015-07-01)

현직 교사들이 2015 개정 수학 교육과정 시안에 대해 "학습부담 경감" 노력이 미흡하다"고 지적했습니다.  
한국교원단체총연합회 새교육개혁포럼이 교총회관에서 개최한 '현장중심 교육과정포럼'에서 수학 분야 토론자로 참석한 경기 청덕중 배숙 수석교사는 "예전에 삭제된 부분이 오히려 추가된 것도 있어 학습 부담 경감의 의지가 제대로 실행됐다고 보기 어렵다"고 평가했습니다.  
이어 "2015 개정 교육과정"이 추구하는 것을 모두 담으려면 앞은 수학교과서를 기대하는 것은 더욱 어렵게 못하다"고 지적했습니다.  
앞서 지난 5월 박경미 홍익대 수학교육과 교수는 수학과 교육과정 연구진을 대표해 '2015 개정 수학과 교육과정' 시안의 윤곽을 발표했습니다.  
당시 박 교수는 "수학교과목의 내용 감축이 필요하다는 점에 어느 정도 사회적 공감대가 형성돼 있지만, 수학과 교육과정의 국제적 표준에 비추어 지나친 내용 감축도 문제를 야기한다"고 설명한 바 있습니다.

26

## 정책 집행 단계

구분	1 교원의 전문성 개발	2 역사교과서 국정화 고시	3 역사교과서 집필진 구성	4 지원체제 구축	5 공교육 정상화	6 역사교과서 국정화 찬반 논란	7 역사 교육 내용	8 인성교육 강화	9 교수·학습 평가 개선
1	교원	역사	집필진	지원	적용	국정	정부	운영	수능
2	수업	검정	국사편찬위	강화	학부모	역사	대한민국	계획	사회
3	소프트웨어	국정	역사	예산	교원	정부	북한	교육청	문제
4	연수	고시	집필기준	정책	평가	한국사	수립	확대	과학
5	참여	확정	집필	사회	한글	반대	내용	자유학기제	학습
6	창의	한국사	구성	안전	수학	야권	주체사상	예술	영어
7	교실	발행	개발	교육청	정책	대통령	서술	지원	평가
8	진행	사회	공개	조직	수업	박근혜	기술	수업	디지털
9	미래	교원	확정	대학	공교육	국회	문제	활동	필요
10	방법	국민	검토	환경	지적	정치	총리	연극	내용
문서수	157	110	86	79	117	27	92	77	80

27

## (1) 토픽1: 교원의 전문성 개발

구분	1 교원의 전문성 개발
1	교원
2	수업
3	소프트웨어
4	연수
5	참여
6	창의
7	교실
8	진행
9	미래
10	방법
문서수	157

### 〈기사 예시〉

학생참여형 교실수업 위한 대규모 교원연수 추진(아시아경제, 2016-06-08)

교육부가 창의융합형 인재 양성을 추구하는 '2015 개정 교육과정'이 일선 학교에 바르게 정착될 수 있도록 교원 참여형 현장 밀착형 교원 연수를 추진한다. 초·중·고등학교 교원과 전문직 43만명 중 절반이 넘는 23만명이 교실수업 개선을 위한 연수에 참여하게 된다. (중략)

일방적인 강의식 전달이 아닌 토론과 실습 중심의 참여형 연수도 강화했다. 교사들이 먼저 토론하고 직접 탐구하거나 실험·실습 방식의 연수에 참여함으로써 실제 교실에서도 학생 참여형 수업으로 전환을 꾀하기 위해서다. 연수 방법 면에서는 원격연수와 집합연수를 결합하는 방식을 도입했다. 먼저 원격연수 과정에서 새롭게 바뀐 내용에 대해 학습하고 1박2일 집합연수 과정에서 새 교육과정의 교실 현장 적용 방안에 대한 토론, 실제 교수·학습 설계 실습 등을 진행한다. 원격연수와 집합연수를 결합한 연수는 초등학교 1~2학년 교육과정 전반, 고등학교 통합사회 및 통합과학 과목에 우선 적용된다. 내용 면에서는 2015 개정 교육과정의 개정 내용을 포함해 실제 교실수업 변화를 위한 수업 및 평가 방법의 개선에 초점을 맞췄다. 이는 2015 개정 교육과정이 일방적 강의식 수업이 아닌 프로젝트학습, 토의·토론학습, 탐구학습 등 학생 참여식 수업을 강화한 것과 맥을 같이 한다. (하략)

28

### 3. 토픽명에 대한 타당도 검토

#### 가. 분석 기간 전체

다음은 분석 결과 도출된 16개 토픽과 그에 따른 주요어를 바탕으로, 연구자가 부여한 '2015 개정 교육과정 정책'의 토픽(쟁점)명의 타당성을 묻는 질문지입니다. 각 토픽별 명칭이 적절한지 검토해 주시고 해당하는 곳에 **√**표 해 주십시오.

구분	1	2	3	4	5	6	7	8	9	10	전혀 타당 하지 않다	타당 하지 않다	보통 이다	타당 하다	매우 타당 하다
1 역사교과서 국정화 관련 정치적 대립	야권	정부	국회의원	대통령	국회	박근혜	여권	정치	정권	비판	①	②	③	④	⑤
2 역사교과서 국정화 찬반 논란	역사	국정	한국사	국민	반대	대한민국	사회	국가	교원	정부	①	②	③	④	⑤
3 역사교육 내용	대한민국	북한	정부	수립	근현대사	서술	일본	내용	역사	주제사상	①	②	③	④	⑤
4 교육과정 운영 지원체제 구축	대학	지원	추진	확대	개혁	사회	확산	재정	직업	강화	①	②	③	④	⑤
5 신설 교과목의 특성	수학	과학	사회	영어	국어	내용	학습	신설	공통교과	수업	①	②	③	④	⑤
6 교육과정 개정의 필요성	사회	문제	융합	창의	필요	학습	내용	인재	지식	해결	①	②	③	④	⑤
7 역사교과서 집필진 구성	집필진	개발	국사 편찬위	구성	역사	집필	집필 기준	공개	검토	확정	①	②	③	④	⑤
8 역사교과서 국정화 고시	국정	검정	한국사	고시	정부	확정	논란	인정	수정	전환	①	②	③	④	⑤
9 수능·대입 제도 개선	수능	평가	대입	시험	영어	제도	출제	학부 모	부담	반영	①	②	③	④	⑤
10 교수·학습 방법 개선	교원	수업	참여	개선	연수	방법	환경	현장	교실	토론	①	②	③	④	⑤
11 초등 교과서 한자 병기	한자	병기	한글	학부 모	정책	방법	반대	국어	연구	학습	①	②	③	④	⑤
12 현장교원 의견 수렴	공청회	의견	고시	현장	총론	논의	국가	연구	지적	전교조	①	②	③	④	⑤
13 자유학기제 운영	운영	교육청	계획	자유 학기제	지원	확대	마련	실시	협력	시행	①	②	③	④	⑤
14 소프트웨어 교육 강화	소프트 웨어	과학	기술	개발	정보	디지털	미래 부	활용	필수	진행	①	②	③	④	⑤
15 안전교육 강화	안전	강화	예산	교육청	신설	재난	조직	대응	업무	정책	①	②	③	④	⑤
16 인성교육 강화	활동	예술	연극	체육	체험	인성	강화	음악	교육 지원청	문화	①	②	③	④	⑤

※ 토픽명에 대한 기타 의견은 공란에 기재해 주시기 바랍니다.

## 나. 정책 단계별

### (1) 정책 형성 단계

다음은 이 연구에서 설정한 2015 개정 교육과정의 정책 ‘형성’ 단계에 따른 토  
픽명의 타당성을 묻는 질문지입니다. 각 토픽별 명칭이 적절한지 검토해 주시고  
해당하는 곳에 ☒표 해 주십시오.

구분		1	2	3	4	5	6	7	8	9	10	전혀 타당 하지 않다	타당 하지 않다	보통 이다	타당 하다	매우 타당 하다
1	과학교육 축소	과학	기술	수학	학계	논의	의견	연합 회	요구	국가	포럼	①	②	③	④	⑤
2	수능·대입 제도 개선	수능	도입	대입	정부	평가	확정	융합	개발	선택	한국 사	①	②	③	④	⑤
3	소프트웨어 교육 강화	소프트 웨어	정책	미래 부	추진	확대	정보	개발	지원	인재	강화	①	②	③	④	⑤
4	공통교과 신설	사회	이수 단위	과학	필수	영어	수업	국어	공통 교과	내용	신설	①	②	③	④	⑤
5	역사교과서 국정화	교원	문제	검정	국정	내용	수정	연구	전환	현장	고시	①	②	③	④	⑤

※ 토픽명에 대한 기타 의견은 아래 공간에 기재해 주시기 바랍니다.

## (2) 정책 결정 단계

다음은 이 연구에서 설정한 2015 개정 교육과정의 정책 ‘결정’ 단계에 따른 토  
픽명의 타당성을 묻는 질문지입니다. 각 토픽별 명칭이 적절한지 검토해 주시고  
해당하는 곳에 √표 해 주십시오.

구분	1	2	3	4	5	6	7	8	9	10	전혀 타당 하지 않다	타당 하지 않다	보통 이다	타당 하다	매우 타당 하다
1 역사교과서 국정화	국정	역사	한국 사	검정	반대	개발	정부	인정	검토	학계	①	②	③	④	⑤
2 초등 교과서 한자 병기	한자	병기	한글	방법	반대	정책	공청 회	제시	활성	연구	①	②	③	④	⑤
3 소프트웨어 교육	수업	소프트 웨어	시안	운영	체육	활동	자유 학기제	정보	교원	체험	①	②	③	④	⑤
4 현장교원 의견 수렴	교원	현장	의견	고시	공청 회	적용	국가	총론	논의	추진	①	②	③	④	⑤
5 수학교육 내용 적정화	수학	내용	학습	학습 량	시안	학습 부담	사항	국어	사고	성취 기준	①	②	③	④	⑤
6 수능·대입 제도 개선	수능	평가	대입	영어	문제	시험	제도	대학	도입	공 교육	①	②	③	④	⑤
7 역사 교육 내용	역사	대한 민국	국가	일본	근현 대사	정부	비중	내용	축소	수립	①	②	③	④	⑤
8 교육과정 개정의 필요성	창의	융합	사회	과학	인재	기술	필요	개발	변화	환경	①	②	③	④	⑤
9 신설 교과목의 특성	사회	과학	신설	국어	공통 교과	진로	강화	이수 단위	선택 교과	적용	①	②	③	④	⑤
10 역사교과서 국정화 관련 정치적 대립	정부	야권	국회 의원	박근혜	국회	대통령	국감	규정	여권	진행	①	②	③	④	⑤
11 안전교육 강화	안전	계획	지원	추진	대학	재난	개혁	확대	현장	정부	①	②	③	④	⑤

※ 토픽명에 대한 기타 의견은 아래 공란에 기재해 주시기 바랍니다.

### (3) 정책 집행 단계

다음은 이 연구에서 설정한 2015 개정 교육과정의 정책 ‘집행’ 단계에 따른 토  
픽명의 타당성을 묻는 질문지입니다. 각 토픽별 명칭이 적절한지 검토해 주시고  
해당하는 곳에 ☒표 해 주십시오.

구분	1	2	3	4	5	6	7	8	9	10	전혀 타당 하지 않다	타당 하지 않다	보통 이다	타당 하다	매우 타당 하다
1 교원의 전문성 개발	교원	수업	소프트 웨어	연수	참여	창의	교실	진행	미래	방법	①	②	③	④	⑤
2 역사교과서 국정화 고시	역사	검정	국정	고시	확정	한국 사	발행	사회	교원	국민	①	②	③	④	⑤
3 역사교과서 집필진 구성	집필 진	국사편 찬위	역사	집필 기준	집필	구성	개발	공개	확정	검토	①	②	③	④	⑤
4 교육과정 운영 지원체제 구축	지원	강화	예산	정책	사회	안전	교육 청	조직	대학	환경	①	②	③	④	⑤
5 공교육 정상화	적용	학부 모	교원	평가	한글	수학	정책	수업	공교육	지적	①	②	③	④	⑤
6 역사교과서 국정화 찬반 논란	국정	역사	정부	한국 사	반대	야권	대통 령	박근 혜	국회	정치	①	②	③	④	⑤
7 역사 교육 내용	정부	대한 민국	북한	수립	내용	주체 사상	서술	기술	문제	총리	①	②	③	④	⑤
8 인성교육 강화	운영	계획	교육 청	확대	자유 학기제	예술	지원	수업	활동	연극	①	②	③	④	⑤
9 교수·학습·평가 개선	수능	사회	문제	과학	학습	영어	평가	디지털	필요	내용	①	②	③	④	⑤

※ 토픽명에 대한 기타 의견은 아래 공란에 기재해 주시기 바랍니다.

#### 4. 토픽명에 대한 타당도 분석 결과

##### 가. 분석 기간 전체

최초 토픽명		타당도 평정 결과 (5점 만점 기준)		제안된 토픽명	최종 토픽명
		평균	표준편차		
1	역사교과서 국정화 관련 정치적 대립	4.27	0.90	-	(변동 없음)
2	역사교과서 국정화 찬반 논란	4.45	0.52	-	(변동 없음)
3	역사교육 내용	4.27	0.79	역사교육 세부 내용, 역사교과서 내용	<b>역사교육 세부 내용</b>
4	교육과정 운영 지원체제 구축	4.18	0.40	교육과정 운영 지원	(변동 없음)
5	신설 교과과의 특성	<u>3.55</u>	1.04	공통 과목 운영, 공통·신설 과목 도입, 고교 공통 과목 소개	<b>고교 공통 과목 운영</b>
6	교육과정 개정의 필요성	4.73	0.47	-	(변동 없음)
7	역사교과서 집필진 구성	4.27	0.79	-	(변동 없음)
8	역사교과서 국정화 고시	4.55	0.69	-	(변동 없음)
9	수능·대입 제도 개선	4.36	0.50	-	(변동 없음)
10	교수·학습 방법 개선	4.55	0.69	-	(변동 없음)
11	초등 교과서 한자 병기	4.73	0.47	-	(변동 없음)
12	현장교원 의견 수렴	4.73	0.47	-	(변동 없음)
13	자유학기제 운영	4.64	0.67	-	(변동 없음)
14	소프트웨어 교육 강화	4.73	0.47	-	(변동 없음)
15	안전교육 강화	4.91	0.30	-	(변동 없음)
16	인성교육 강화	<u>3.45</u>	1.04	체험·인성교육 강화, 예술·체육 활동 강화	<b>체험·인성교육 강화</b>



## 나. 정책 단계별

### (1) 정책 형성 단계

최초 토픽명		타당도 평정 결과 (5점 만점 기준)		제안된 토픽명	최종 토픽명
		평균	표준편차		
1	과학교육 축소	<b>3.64</b>	0.81	과학교육 축소 논란, 과학교육 축소 우려	<b>과학교육 축소 우려</b>
2	수능·대입 제도 개선	4.55	0.52	-	(변동 없음)
3	소프트웨어 교육 강화	4.64	0.50	소프트웨어 교육 확대	<b>소프트웨어 교육 확대</b>
4	공통교과 신설	<b>3.82</b>	1.17	공통 과목 신설, 고교 공통 과목 도입	<b>고교 공통 과목 도입</b>
5	역사교과서 국정화	4.27	0.79	-	(변동 없음)

### (2) 정책 결정 단계

최초 토픽명		타당도 평정 결과 (5점 만점 기준)		제안된 토픽명	최종 토픽명
		평균	표준편차		
1	역사교과서 국정화	4.45	0.52	역사교과서 국정화 찬반 논란	<b>역사교과서 국정화 찬반 논란</b>
2	초등 교과서 한자 병기	4.73	0.47	-	(변동 없음)
3	소프트웨어 교육 강화	4.09	0.83	-	(변동 없음)
4	현장교원 의견 수렴	4.73	0.47	-	(변동 없음)
5	수학교육 내용 적정화	4.55	0.52	-	(변동 없음)
6	수능·대입 제도 개선	4.45	0.52	-	(변동 없음)
7	역사 교육 내용	4.55	0.69	역사 교육과정 내용, 역사 교육 세부 내용	<b>역사 교육 세부 내용</b>
8	교육과정 개정의 필요성	4.64	0.67	-	(변동 없음)
9	신설 교과의 특성	4.00	0.89	공통 과목 운영, 고교 공통 과목 운영	<b>고교 공통 과목 운영</b>
10	역사교과서 국정화 관련 정치적 대립	4.27	0.79	-	(변동 없음)
11	안전교육 강화	4.73	0.47	-	(변동 없음)

### (3) 정책 집행 단계

토픽명		타당도 평정 결과 (5점 만점 기준)		제안된 토픽명	최종 토픽명
		평균	표준편차		
1	교원의 전문성 개발	4.36	0.67	교원의 교수역량 개발	(변동 없음)
2	역사교과서 국정화 고시	4.73	0.47	-	(변동 없음)
3	역사교과서 집필진 구성	4.55	0.69	-	(변동 없음)
4	교육과정 운영 지원체제 구축	4.45	0.52	-	(변동 없음)
5	공교육 정상화	4.00	1.00	-	(변동 없음)
6	역사교과서 국정화 찬반 논란	4.64	0.50	-	(변동 없음)
7	역사 교육 내용	4.55	0.52	역사 교육과정 내용, 역사 교육 세부 내용	역사 교육 세부 내용
8	인성교육 강화	3.55	0.82	예술·체육활동 강화, 체험·인성교육 강화	체험·인성교육 강화
9	교수·학습·평가 개선	3.82	0.98	교실 수업·평가 개선	교실 수업·평가 개선

Analysis of media coverage on 2015 revised  
curriculum policy using Big Data Analysis

A Dissertation for the Degree of  
Doctor of Philosophy in Education

by

Ye-Lim Yu

Major Advisor: Sun-Geun Baek, Ph.D.

Department of Education  
Graduate School  
Seoul National University

2017

## Abstract

# Analysis of media coverage on 2015 revised curriculum policy using Big Data Analysis

Ye-Lim Yu

Department of Education

The Graduate School

Seoul National University

In the Big Data era, where various kinds of information are generated and circulated in real time, it is important to analyze the media coverage of major education policies by using big data analysis techniques. The purpose of this study is to analyze the media coverage of the 2015 revised curriculum policy, which is one of the core of Korea's primary and secondary education policy, using big data analysis technique. In this study, the following research questions were proposed.

1. Does the interest of the media in the 2015 revised curriculum policy change with the formation, making and implementation of policy?
2. What are the keywords of the news articles about the 2015 revised curriculum policy and how does it change with the formation, making and implementation of policy?
3. What are the key issues of the news articles about the 2015 revised curriculum policy and how does it change with the formation, making and implementation of policy?

In order to investigate these research questions, 1,598 news articles related to the 2015 revised curriculum policy reported on the Naver News homepage, which has the greatest power to influence public opinion in Korea, were collected using web scraping technique with Python. Preprocessing and morphological analysis were conducted on a total of 1,512 final analysis data except articles that did not meet the purpose of analysis. Then, buzz analysis, keyword frequency analysis, and Latent Dirichlet Allocation (LDA) which was the most common method of topic modeling were performed. The results of this study are summarized as follows.

First, the interest of Korea's media in relation to the 2015 revised curriculum policy changed according to the formation, making and implementation of policy. That is, in the policy formation phase, media had shown little interest on the 2015 revised curriculum. Just before the policy was decided and implemented, and at the time when the controversy over the nationalization of the history textbook was maximized, the contents were reported intensively.

Secondly, when analyzed based on the Term Frequency (TF) during the whole analysis period, the keywords of the 2015 revised curriculum policy appeared in the news articles of Korea's media were 'textbook', 'department of education', 'history', 'student', 'government-issued', 'subject', 'high school', 'teacher', 'elementary school', and 'junior high school'. On the other hand, based on the TF-Inverse Document Frequency (TF-IDF), the keywords of the policy were changed such as 'history', 'government-issued', 'Chinese character', 'software', 'mathematics', 'government', 'science', 'classes', 'College Scholastic Ability Test (CSAT)', and 'metropolitan office of education'.

Thirdly, the keywords of the news articles about the 2015 revised curriculum policy have changed with the formation, making and implementation of policy. In the policy formation phase, TF-IDF is higher in order of 'software', 'CSAT', 'science', 'technology', and 'unit', whereas in the

policy-making phase, 'chinese character', 'government-issued', 'briefing', 'government', and 'mathematics' showed the highest frequency. In the policy implementation phase, 'history', 'government-issued', 'software', 'metropolitan office of education', and 'Hangul' showed the highest frequency.

Fourthly, during the whole analysis period, sixteen issues of the news articles about the 2015 revised curriculum policy were extracted, including ① implementation of common subjects in high school, ② notification for the government-issued history textbook, ③ feedback from field teachers, ④ chinese character parallel written in the elementary textbook, ⑤ improvement of CSAT and college entrance system, ⑥ contents of history education, ⑦ reinforcement of software education, ⑧ political conflict on the government-issued history textbook, ⑨ pros and cons of the government-issued history textbook, ⑩ authors of the history textbook, ⑪ improvement of teaching and learning methods, ⑫ reinforcement of character education, ⑬ necessity of curriculum revision, ⑭ reinforcement of safety education, ⑮ development of the support system for curriculum, and ⑯ operation of the free semester system.

Fifthly, the key issues of the news articles about the 2015 revised curriculum policy have changed with the formation, making and implementation of policy. In the policy formation phase, five issues were extracted such as 'concerns about reducing science education', 'introduction of common subjects in high school', 'improvement of CSAT and college entrance system', and 'enhancement of software education', whereas in the policy-making phase, eleven issues were extracted such as 'implementation of common subjects in high school', 'chinese character parallel written in the elementary textbook', and 'pros and cons of the government-issued history textbook'. In the policy implementation phase, nine issues were extracted such as 'development of teachers' expertise', 'normalization of public education', and 'notification for the government-issued history textbook'.

This study can be meaningful in that it has attempted new analysis of major education policy by using big data analysis technique. As the big data analysis technique used in this study suggests the possibility of the new policy evaluation study in education, it is necessary to expand the analysis target with various education policies through the follow-up study and develop the related analysis technique more sophisticatedly.

\* **keywords** : big data analysis, 2015 revised curriculum, online news, topic modeling, Latent Dirichlet Allocation (LDA), policy evaluation in education

\* ***Student Number*** : 2008-30394