2021 연구개발적립금사업 01

OPEN 공공누리 출처표시 상업용금지 변경금지 공공저작물 자유이용허락

빅데이터에 기반한 육아정책 쟁점과 향후 과제

김동훈·이재희·장현진

육아정책연구소
Korea Institute of Child Care and Education

2021 연구개발적립금사업 01

빅데이터에 기반한 육아정책 쟁점과 향후 과제

저 자 김동훈, 이재희, 장현진

연 구 진
연구책임자 **김 동 훈** (육아정책연구소 연구위원)
공동연구자 **이 재 희** (육아정책연구소 연구위원)
공동연구자 **장 현 진** (육아정책연구소 연구원)

육아정책연구소
Korea Institute of Child Care and Education

2021 연구개발적립금사업 01

빅데이터에 기반한 육아정책 쟁점과 향후 과제

발행일 2021년 12월
발행인 박상희
발행처 육아정책연구소
주소 서울시 중구 소공로 70 9층 육아정책연구소
전화 02) 398-7700
팩스 02) 398-7798
홈페이지 http://www.kicce.re.kr
인쇄처 (주)승림디엔씨 02) 2271-2581

보고서 내용의 무단 복제를 금함.
ISBN 979-11-6865-012-1 93330

육아정책연구소
Korea Institute of Child Care and Education

머리말

빅데이터에 대한 사회적, 정책적 활용 필요성이 높아지고 있으나, 육아정책과 관련하여 어떤 자료가 있고, 어떻게 활용될 수 있을지에 대해서는 논의가 부족한 실정이다. 육아관련 현안을 파악하여 정책을 수립·시행하는데 있어 빅데이터의 활용 가능성은 상당할 것으로 기대되고, 특히 육아지원정책에 대한 온 국민의 관심이 높은 상황에서 이들의 다양한 요구는 빅데이터를 이용하면 기존의 연구방법을 통해 나타나지 않았던 새로운 내용이나 현안들이 도출될 수 있다. 육아정책 분야에서 상대적으로 부족한 빅데이터가 어떻게 활용될 수 있는지에 대한 가능성을 보여주고, 빅데이터를 활용한 육아정책에 대한 아이디어 발굴이라는 새로운 접근방식에 대해 논의하고자 하였다.

본 연구 수행을 통해 유치원과 어린이집에 대한 종합적인 안전·환경평가 실시를 제안하였다. 또한 그동안 임신, 출산, 양육 등에 관하여 아이 중심으로 지원이 이루어져 왔고 향후에는 임신, 출산, 양육의 주체인 엄마아빠 등 양육주체들을 고려한 지원정책이 향후 좀 더 강화되고 보완될 필요가 있음을 강조하였다. 아울러 연구측면에서 향후에는 그동안 상대적으로 비중이 적었던, 보육이나 영아, 부모, 교사에 관련된 연구와 함께, 다문화, 장애인 등을 포함한 다양한 가족의 형태에 대한 더 많은 연구가 필요함을 제안하였다.

아직은 부족한 육아분에서의 빅데이터들이 향후 어떻게 연구에 활용되고 정책적 기여를 할 수 있는지에 대한 지속적인 연구가 필요해 보이며, 다양한 분석방법들을 활용하는 연구들이 많아지기를 기대한다.

본 연구 수행을 위해 도움을 주신 모든 분께 감사드린다. 본 보고서의 내용은 연구진의 의견이며, 육아정책연구소의 공식적인 입장이 아님을 밝혀둔다.

2021년 12월

육아정책연구소

소장 박 상 희

목차

요약 1

Ⅰ. 서론 11

1. 연구의 필요성 및 목적 13
2. 연구내용 14
3. 연구방법 14

Ⅱ. 빅데이터와 육아정책 17

1. 빅데이터 개념 19
2. 빅데이터 관련 법 및 주요 정책 추진 현황 23
3. 빅데이터 정책 활용·연계 사례 34
4. 육아정책 관련 빅데이터 39
5. 빅데이터 활용 육아정책 선행연구 41

Ⅲ. 공공 빅데이터 활용을 통한 영유아 안전·환경 분석 45

1. 육아정책 빅데이터 플랫폼 47
2. 빅데이터 분석사례: 영유아 안전·환경 공공 빅데이터 50
3. 소결 67

Ⅳ. 육아카페 빅데이터 활용을 통한 육아현안 분석 69

1. 분석 필요성 및 목적 71
2. 분석 내용 및 방법 72
3. 분석 결과 74
4. 소결 104

Ⅴ. 학술논문 빅데이터 활용을 통한 육아정책 연구동향 분석 107
1. 분석 필요성 및 목적 ····· 109
2. 분석 내용 및 방법 ····· 109
3. 분석 결과 ····· 112
4. 소결 ····· 124

Ⅵ. 정책제언 125
1. 육아 안전·환경 종합 평가항목 개발과 체계 마련 필요 ····· 127
2. 영유아의 부모를 위한 육아지원정책 강화 ····· 128
3. 육아지원정책을 위한 향후 연구과제 ····· 130

참고문헌 133

표 목차

〈표 Ⅱ-1- 1〉 빅데이터 분류 ····· 21
〈표 Ⅱ-1- 2〉 기존 데이터와 빅데이터의 비교 ····· 22
〈표 Ⅱ-1- 3〉 미래사회의 특성과 빅데이터의 역할 ····· 23
〈표 Ⅱ-2- 1〉 빅데이터 관련 법령 ····· 25
〈표 Ⅱ-2- 2〉 데이터3법 개정 주요 내용 ····· 25
〈표 Ⅱ-2- 3〉 빅데이터 TF ····· 30
〈표 Ⅱ-5- 1〉 육아분야 빅데이터 활용자료 및 분석내용 ····· 44
〈표 Ⅲ-1- 1〉 빅데이터 관련 플랫폼 ····· 48
〈표 Ⅲ-1- 2〉 통합데이터 현황 ····· 48
〈표 Ⅲ-1- 3〉 빅데이터 플랫폼(통합데이터지도) 내 육아관련 데이터 현황 ····· 49
〈표 Ⅲ-2- 1〉 시도별 대기질 ····· 51
〈표 Ⅲ-2- 2〉 대기질 하위 20개 시군구 ····· 52
〈표 Ⅲ-2- 3〉 미세먼지 PM10 30초과 어린이집 개소수 ····· 53
〈표 Ⅲ-2- 4〉 통합대기질 100 초과 유치원·어린이집 개소수 ····· 53
〈표 Ⅲ-2- 5〉 유치원·어린이집 인근 위험시설 ····· 57
〈표 Ⅲ-2- 6〉 유치원·어린이집 인근 500m 내 위험시설수
상위 40개 시군구 ····· 58
〈표 Ⅲ-2- 7〉 수질오염원 인근 아동수 현황 ····· 63
〈표 Ⅳ-2- 1〉 육아정책 관련 기사 제목 키워드 분석 ····· 72
〈표 Ⅳ-2- 2〉 육아카페 게시글 웹크롤링 결과 ····· 73
〈표 Ⅳ-3- 1〉 상위 100개 단어 빈도수 ····· 76
〈표 Ⅳ-3- 2〉 육아카페 토픽별 상위 30개 키워드 및 토픽명 ····· 78
〈표 Ⅳ-3- 3〉 Topic1 관련 원본문서 ····· 82
〈표 Ⅳ-3- 4〉 Topic2 관련 원본문서 ····· 83
〈표 Ⅳ-3- 5〉 Topic3 관련 원본문서 ····· 85
〈표 Ⅳ-3- 6〉 Topic4 관련 원본문서 ····· 87
〈표 Ⅳ-3- 7〉 Topic5 관련 원본문서 ····· 88
〈표 Ⅳ-3- 8〉 Topic6 관련 원본문서 ····· 90
〈표 Ⅳ-3- 9〉 Topic7 관련 원본문서 ····· 92

〈표 Ⅳ-3-10〉 Topic8 관련 원본문서 ····· 93
〈표 Ⅳ-3-11〉 Topic9 관련 원본문서 ····· 95
〈표 Ⅳ-3-12〉 Topic10 관련 원본문서 ····· 96
〈표 Ⅳ-3-13〉 Topic11 관련 원본문서 ····· 97
〈표 Ⅳ-3-14〉 Topic12 관련 원본문서 ····· 99
〈표 Ⅳ-3-15〉 Topic13 관련 원본문서 ····· 100
〈표 Ⅳ-3-16〉 Topic14 관련 원본문서 ····· 101
〈표 Ⅳ-3-17〉 Topic15 관련 원본문서 ····· 103
〈표 Ⅳ-4- 1〉 육아카페 토픽 범주화 ····· 104
〈표 Ⅴ-2- 1〉 키워드 별 문서 수 ····· 110
〈표 Ⅴ-3- 1〉 학술연구 상위 25개 단어 ····· 112
〈표 Ⅴ-3- 2〉 학술연구 토픽 모델링 결과 ····· 115
〈표 Ⅴ-4- 1〉 학술연구 토픽 범주화 ····· 124

그림 목차

[그림 Ⅱ-1- 1] 빅데이터 의미의 확장 ····· 19
[그림 Ⅱ-1- 2] 빅데이터의 특징 ····· 21
[그림 Ⅱ-2- 1] 빅데이터 관련 주요 정책 ····· 26
[그림 Ⅱ-2- 2] 빅데이터 마스터 플랜 비전 및 목표 ····· 28
[그림 Ⅱ-2- 3] 민간 및 공공분야의 빅데이터 추진체계 ····· 29
[그림 Ⅱ-2- 4] 공공데이터 정책의 추진방향 ····· 31
[그림 Ⅱ-3- 1] 오산시 다함께돌봄센터 최적설치지역 분석 ····· 35
[그림 Ⅱ-3- 2] 오산시 다함께돌봄센터 설치 우선지역 ····· 35
[그림 Ⅱ-3- 3] 지역별 초등돌봄 수요예측 사례-창원시, 청주시 ····· 36
[그림 Ⅱ-3- 4] 응답소 민원분석시스템 주요 화면-다양한 조건에 따른 민원 분석 및 동향 파악 ····· 37
[그림 Ⅱ-3- 5] 역학조사 지원시스템 ····· 38
[그림 Ⅲ-2- 1] 어린이집 및 유치원 위치 ····· 54
[그림 Ⅲ-2- 2] 시군구별 어린이집 및 유치원 위치 ····· 54
[그림 Ⅲ-2- 3] 대기질 지수: Cold Spots vs. Hot Spots ····· 55
[그림 Ⅲ-2- 4] 미세먼지: Cold Spots vs. Hot Spots ····· 55
[그림 Ⅲ-2- 5] 전국 영유아시설 인근 대기질 지수 분포 ····· 56
[그림 Ⅲ-2- 6] 전국 영유아시설 인근 미세먼지 분포 ····· 56
[그림 Ⅲ-2- 7] 영유아 위험시설 및 위험시설 인근 영유아 시설 ····· 59
[그림 Ⅲ-2- 8] 영유아 시설 인근 위험시설 분포 ····· 59
[그림 Ⅲ-2- 9] 영유아 시설 인근 위험시설 지역별 분포 ····· 60
[그림 Ⅲ-2-10] 수소이온농도: Cold Spots vs. Hot Spots ····· 61
[그림 Ⅲ-2-11] 잔류염소: Cold Spots vs. Hot Spots ····· 62
[그림 Ⅲ-2-12] 탁도: Cold Spots vs. Hot Spots ····· 62
[그림 Ⅲ-2-13] 수질오염원 및 수질오염원 인근 영유아수 ····· 65
[그림 Ⅲ-2-14] 전국 영유아시설 인근 수질오염원 분포 ····· 65
[그림 Ⅲ-2-15] 영유아시설 인근 수질오염원 수 ····· 66
[그림 Ⅳ-3- 1] 상위 1,156개 단어 워드클라우드 ····· 76
[그림 Ⅳ-3- 2] Intertopic Distance Map ····· 80

[그림 Ⅳ-3- 3] 육아카페 토픽 모델링 시각화 결과(토픽1) ····· 81
[그림 Ⅳ-3- 4] 육아카페 토픽 모델링 시각화 결과(토픽2) ····· 83
[그림 Ⅳ-3- 5] 육아카페 토픽 모델링 시각화 결과(토픽3) ····· 84
[그림 Ⅳ-3- 6] 육아카페 토픽 모델링 시각화 결과(토픽4) ····· 86
[그림 Ⅳ-3- 7] 육아카페 토픽 모델링 시각화 결과(토픽5) ····· 88
[그림 Ⅳ-3- 8] 육아카페 토픽 모델링 시각화 결과(토픽6) ····· 90
[그림 Ⅳ-3- 9] 육아카페 토픽 모델링 시각화 결과(토픽7) ····· 91
[그림 Ⅳ-3-10] 육아카페 토픽 모델링 시각화 결과(토픽8) ····· 93
[그림 Ⅳ-3-11] 육아카페 토픽 모델링 시각화 결과(토픽9) ····· 94
[그림 Ⅳ-3-12] 육아카페 토픽 모델링 시각화 결과(토픽10) ····· 96
[그림 Ⅳ-3-13] 육아카페 토픽 모델링 시각화 결과(토픽11) ····· 97
[그림 Ⅳ-3-14] 육아카페 토픽 모델링 시각화 결과(토픽12) ····· 98
[그림 Ⅳ-3-15] 육아카페 토픽 모델링 시각화 결과(토픽13) ····· 100
[그림 Ⅳ-3-16] 육아카페 토픽 모델링 시각화 결과(토픽14) ····· 101
[그림 Ⅳ-3-17] 육아카페 토픽 모델링 시각화 결과(토픽15) ····· 102
[그림 Ⅴ-2- 1] 키워드 별 문서 수 ····· 110
[그림 Ⅴ-3- 1] 학술연구 상위 단어 워드클라우드 ····· 113
[그림 Ⅴ-3- 2] 학술연구 토픽 모델링 시각화 결과(전체) ····· 117
[그림 Ⅴ-3- 3] 학술연구 토픽 모델링 시각화 결과(토픽1) ····· 118

요약

1. 서론

가. 연구의 필요성 및 목적

□ 육아정책 분야에서 상대적으로 부족한 빅데이터가 어떻게 활용될 수 있는지에 대한 가능성을 보여주고, 빅데이터를 활용한 육아정책에 대한 아이디어 발굴이라는 새로운 접근방식에 대해 논의하고자 함.

□ 육아관련 빅데이터에 대한 탐색과 빅데이터를 활용한 사례 분석을 통해 육아지원 사각지대나 정책적 요구의 파악 가능성과 향후 빅데이터를 활용하여 육아정책의 현안 파악과 과제 도출을 위한 시사점을 제안함.

나. 연구내용

□ 빅데이터에 대한 개념을 살펴보고 관련 법·제도·현황을 살펴봄.

- 빅데이터의 개념, 빅데이터 관련 법, 정부의 주요정책 추진현황, 빅데이터 활용사례 및 관련 선행연구 등을 살펴봄.

□ 육아정책 관련 빅데이터 활용·분석

- 유치원과 어린이집의 환경·안전과 관련 공공 빅데이터 분석
- 육아관련 인터넷카페 게시글에 토픽분석 실시
- 학술연구에 나타난 연구자들의 관심주제 분석

□ 육아관련 빅데이터를 활용한 육아정책 현안 및 과제 제안의 가능성을 모색하고, 향후 육아정책 관련 빅데이터 활용에 시사점을 제안함.

다. 연구방법

□ 문헌분석

- 빅데이터 개념, 빅데이터 관련 법 및 정책추진현황, 빅데이터 활용 선행연구 등에 대한 문헌분석 실시

□ 빅데이터 통계분석

- 어린이집과 유치원 인근 대기질, 위험시설, 수질 등과 관련한 안전·환경관련 공공 빅데이터 분석
- 인터넷 육아카페(맘스홀릭) 게시물과 KCI 논문초록 등 비정형데이터를 분석함. 비정형데이터는 텍스트마이닝을 통한 키워드분석, 추이분석 등 실시
- 빅데이터 분석도구로는 R과 파이썬을 사용하였고 빅데이터 분석 전문업체와의 협력하여 수행

□ 전문가 회의

- 육아정책분야 빅데이터의 활용가능성 전문가 자문, 빅데이터 분석을 위한 주요 키워드 도출 및 관련 전문가 자문

2. 빅데이터와 육아정책

가. 빅데이터 개념

□ 빅데이터는 대용량 데이터를 수집 · 저장 · 관리 · 분석 · 체계화를 위한 도구, 플랫폼, 인력, 조직 등을 포괄하는 용어로, 대용량 데이터를 통해 의미있는 정보를 수집하고 생성된 지식을 바탕으로 상황을 능동적으로 대응하거나 변화를 예측하기 위한 정보화 기술을 의미한다고 볼 수 있음.

□ 빅데이터의 특징으로는 일반적으로 규모(Volume), 다양성(Variety), 신속성(Velocity)으로 설명하고 있으나, 복잡성(complexity)과 가치(value)를 포함하기도 함.

나. 빅데이터 관련 법 및 주요 정책 추진 현황

□ 빅데이터 관련 법령은 데이터 활용 촉진, 데이터 활용 제한의 두 가지 측면이 있음.

- 데이터 활용 촉진 법으로는 공공기관의 정보공개에 관한 법률, 공공데이터의 제공 및 이용 활성화에 관한 법률, 데이터기반행정 활성화에 관한 법률 등이 있음.
- 빅데이터 활용 제한과 관련된 법률은 개인정보 보호법, 정보통신망 이용촉진

및 정보보호 등에 관한 법률 등이 있음.

□ 2012년에 발표한 '스마트 국가 구현을 위한 빅데이터 마스터 플랜'을 시작으로 빅데이터와 관련한 다양한 정책을 추진하고 있음.

[그림 1] 빅데이터 관련 주요 정책

다. 육아정책 관련 빅데이터

□ 육아정책에서 빅데이터는 크게 공공부문에서 제공하거나 공개하는 공공빅데이터와 이외 빅데이터가 있음.

- 대표적인 공공빅데이터로는 사회보장정보시스템(행복e음), 보육통합정보시스템, 나이스(NEIS), 유치원과 어린이집 정보공시자료 등이 있음.
- 패널자료로 한국아동패널, 전국보육실태조사, 한국복지패널 등이 있음.
- 공공이외 육아관련 정보를 가지고 있는 빅데이터로, 인터넷카페, 블로그, 트위터, 페이스북, 인스타그램 SNS로 대변되는 온라인 자료, 신용카드 자료, 스마트폰 등 모바일 데이터, 포털사이트 뉴스기사 등의 다양한 데이터가 있음.

라. 빅데이터 활용 육아정책 선행연구

□ 육아정책분야에서 빅데이터를 활용한 연구들은 육아정책연구소 등 연구기관과 학술논문을 중심으로 일부 수행되고 있음.

〈표 1〉 육아분야 빅데이터 활용자료 및 분석내용

빅데이터 활용자료	주요 분석 내용
SNS, 카페, 블로그 등 온라인 내 게시글 및 댓글	- 북한 영유아 교육·보육 이슈 - 4차 산업혁명 시대의 출산 및 양육 관련 이슈 - 저출산 정책에 대한 인식 - 영유아 부모의 자녀양육 고민 - 유아 웰빙에 대한 사회적 인식 - 조부모가 이야기되는 양상
인터넷 신문기사	- 북한 영유아 교육·보육 이슈 - 교육·보육 분야 사회갈등 이슈 - 4차 산업혁명 시대의 출산 및 양육 관련 이슈 - 저출산 정책에 대한 인식 - 육아에 대한 부정적 인식 - 유아 웰빙에 대한 사회적 인식
신용카드 사용내역	- 유아 가구의 업종별 카드결제액 추이 및 코로나19 전후의 영유아 가구 소비 변화
학술자료 (학위 논문, 학술지 논문 등)	- 유아 웰빙에 대한 사회적 인식

3. 공공빅데이터 활용을 통한 영유아 안전·환경 분석

가. 육아정책 빅데이터 플랫폼

□ 대표적인 빅데이터 플랫폼으로는 통합데이터지도((bigdata-map.kr)가 있음.

- 통합 데이터 지도 구축 사업은 플랫폼 16개소(문화, 통신, 유통소비, 헬스케어, 교통, 환경, 금융, 중소기업, 지역경제, 산림, 소방안전, 스마트치안, 해양수산, 농식품, 라이프로그, 디지털 산업혁신) 및 센터 150여개소를 모두 연계하는 사업임.
- 빅데이터 플랫폼 내 육아정책 관련 데이터는 총 241개로 나타남.

나. 빅데이터 분석사례: 영유아 안전·환경 공공빅데이터

□ 공공 데이터 중 유치원과 어린이집 등 영유아 보육·교육기관에 대한 대기질, 위험시설, 수질 등과 관련한 안전·환경관련 데이터를 활용하여 빅데이터를 분석함.

- 환경 빅데이터 플랫폼(envbigdata)에서 제공하는 유치원과 어린이집의 주소와 인근 대기질 지수, 미세먼지 지수, 위험시설, 수질지수, 수질오염원 인근

거주인구 정보 등의 자료를 활용함.

□ 유치원·어린이집 인근 대기 질(미세먼저 등)이 나빠졌을 경우 기관 내의 공기질도 동반하여 악화될 가능성이 높은 상황으로, 특히 수도권과 대도시, 공장 등 산업시설이 많은 곳에서 전반적으로 대기질이 나쁜 것으로 나타남.

□ 위험시설은 전국적으로 유치원은 평균 2.3개, 어린이집은 평균 2.4개의 위험시설이 반경 500미터 내에 있는 것으로 나타났고, 주유소 등을 포함한 위험시설이 최대 20개인 어린이집도 있었음.

□ 영유아 보육·교육시설의 수소이온농도, 잔류염소, 탁도 수치는 모두 법적 기준은 모두 충족하고 있었음.

4. 육아카페 빅데이터 분석을 통한 육아현안 분석

가. 분석 필요성 및 목적

□ 임신·출산부터 자녀양육에 이르기 까지 다양한 정부와 지자체의 정책과 서비스의 수요자이자 이해당사자자인 부모들의 경험과 인식에 대해 살펴볼 필요가 있음.

□ 육아 당사자인 부모들이 서로 자신의 경험과 인식들을 공유하는 일명 맘카페의 빅데이터 분석을 통해 육아정책에 대한 현안과 향후의 방향을 모색함.

나. 분석 내용 및 방법

□ 맘 카페에서의 육아관련 키워드에 대해 분석함.

- 임신, 결혼, 육아, 돌봄 등 맘카페 게시물을 추출하여 임신·출산 정책, 육아·돌봄, 보육·교육 등에 대한 여론을 분석함.
- 맘스홀릭 맘카페 게시글의 제목과 내용을 중심으로 웹 크롤링하였고, 수집기간은 2020년 1월부터 2021년 7월까지, 분석게시글 수는 69,836개중 48,155개임.

□ 분석과정은 데이터 전처리 과정과 형태소 분석 등 텍스트마이닝을 통한 토픽분석을 실시하였음. 분석도구는 Python의 한국어 자연어 분석 패키지인 KoNLPy 패키지와 R 4.1.1을 사용하였음.

다. 분석결과

□ 분석 시 사용한 단어 1,156개 중, 빈도수가 가장 많은 것은 ‘아이’였으며, 다음으로 ‘시간’, ‘아침’, ‘저녁’, ‘남편’, ‘자녀’, ‘날씨’, ‘커피’, ‘점심’, ‘주말’ 순임.

[그림 2] 상위 1,156개 단어 워드클라우드

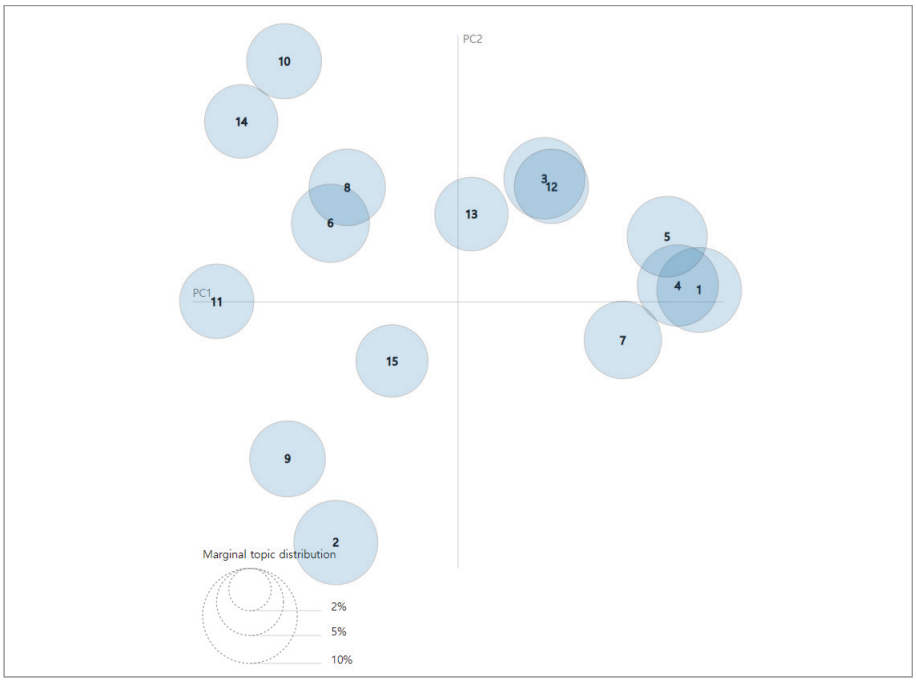
□ 맘카페 빅데이터 분석을 통하여 총 15개의 토픽을 도출하였음.

- Topic1은 ‘피부 및 건강제품 추천’, Topic2는 ‘식사 및 간식’, Topic3은 ‘질병 및 치료’, Topic4는 ‘제품 할인 및 온라인 구매’, Topic5는 ‘제품 추천 및 추천요청’, Topic6은 ‘아이와 엄마의 휴식’, Topic7은 ‘아이 선물 및 이유식’, Topic8은 ‘휴일과 외출’, Topic9는 ‘커피와 휴식 및 디저트’, Topic10은 ‘날씨’, Topic11은 ‘시간 사용’, Topic12는 ‘아이 및 엄마 공부’, Topic13은 ‘엄마 여가시간 및 아이 사진촬영’, Topic14는 ‘일상과 날씨’, Topic15는 ‘운동 및 다이어트’로 토픽별 상위 키워드와 게시판 내용을 토대로 작성함.

□ Intertopic Distance Map을 통해 토픽 비율과 토픽 간 파악한 결과,

- Topic1과 Topic4, Topic5가 매우 유사한데, Topic1은 ‘피부 및 건강제품 추천’, Topic4는 ‘제품 할인 및 온라인 구매’, Topic5는 ‘제품 추천 및 추천요청’에 관한 것임.
- 또한 Topic8과 Topic9는 매우 유사해 보이고, Topic8은 ‘휴일과 외출’, Topic9은 ‘커피와 휴식 및 디저트’인데 둘은 휴식이라는 공통점이 있음.

[그림 3] Intertopic Distance Map



5. 학술논문 분석을 통한 육아정책 연구동향 분석

가. 분석 필요성 및 목적

□ 학술연구 동향을 파악을 통해 육아정책 관련 분야 연구자들의 주요 관심이 어디에 있는지를 파악하고, 향후 연구주제 등에 시사점을 도출함.

나. 분석 내용 및 방법

□ KCI(https://www.kci.go.kr/kciportal/main.kci) 최근 5년 논문초록을 분석함.

- 자료 수집을 위해 임신, 출산, 어린이집, 유치원, 육아, 돌봄, 보육, 유아, 영아, 아동의 10개 키워드를 선정하였고, 기간은 2017년 1월 1일부터 2021년 11월 3일까지임.
- 웹크롤링을 통해 총 30,857개의 논문이 확인되었고 이 중 중복하여 검색된

학술지를 제거한 후 키워드별 논문 27,544개의 논문초록을 분석함.

□ 본 연구에서 수집된 자료는 형태소 분석 → 명사 추출 → 제외어(불용어), 유의어, 지정어 설정 → 수집된 자료에서 제외어 제거, 지정어와 유의어 변경 → 토픽모델링 → 결과해석 및 시각화의 과정을 거쳐 자료처리 및 분석을 실시함.

다. 분석 결과

□ 논문초록 상위 단위로는 '유아'가 가장 많았고, 다음으로 '교사', '아동', '프로그램', '관계', '어머니', '부모' 순이었음.

[그림 4] 학술연구 상위 단어 워드클라우드

□ LDA 분석 결과에 기초하여 총 15개의 토픽을 도출함.

- 1) 저출산 대응 및 돌봄정책, 2) 유아교사 연수, 3) 공동체 돌봄, 4) 다양성 교육, 5) 아동 놀이 및 예술, 6) 아동 장애 치료, 7) 유아 교과 교육, 8) 아동 인권, 9) 출산과 육아, 10) 교사 직무 만족도, 11) 아동 정서, 12) 부모의 양육과 정서, 13) 유아 정서 발달, 14) 교사 및 아동 정서, 15) 영유아 신체 및 정서 발달임.

□ Intertopic Distance Map을 통해 토픽 비율과 토픽 간 파악한 결과,

- 토픽1(저출산 대응 및 돌봄정책), 토픽3(공동체 돌봄), 토픽8(아동인권)이 밀

접한 관계를 가지고 있음.

- 토픽 10(교사 직무 만족도)과 토픽15(영유아 신체 및 정서 발달), 토픽2(유아 교사)와 토픽4(다양성 교육)가 일정수준 유사성을 가지고 있음.

[그림 5] 학술연구 토픽 모델링 시각화 결과

6. 정책제언

가. 육아 안전·환경 종합 평가항목 개발과 체계 마련 필요

□ 아이를 키우기 좋은 육아환경은 무엇이며, 안전을 기반으로 위험시설은 어떤 것이 있고, 영유아가 생활하는 생활권 단위에서 필요하고 개선되어야 하는 안전한 육아환경은 어떻게 되어야 하는 지에 대한 고민이 필요함.

□ 유치원과 어린이집이 속해 있는 해당 지역을 대상으로 종합적인 육아 안전·환경 평가 등을 실시할 필요가 있음.

나. 영유아의 부모를 위한 육아지원정책 강화

□ 그동안 임신, 출산, 양육 등에 관하여 아이 중심으로 지원이 이루어져 왔고 향후에는 임신, 출산, 양육의 주체인 엄마아빠 등 양육주체들을 고려한 지원정책이 향후 좀 더 강화되고 보완될 필요가 있음.

□ 부모가 함께하는 돌봄이 가능하도록 시간지원정책의 실질적 이용을 강화할 필요가 있음.

□ 양육의 주체인 부모 대상 지원 신체, 심리, 정신 건강 프로그램 강화가 필요함.

□ 국가의 공적 책임 강화를 통한 보육 및 돌봄서비스의 질 제고가 필요함.

- 전반적 서비스 질 제고를 위해 적절한 교사 대 아동수 등의 환경 개선, 안전한 실내·외 물적 환경, 지역사회 인프라 확보, 교사의 질 제고, 보육교사와 유치원교사 처우 개선, 프로그램 다양화 및 질 제고 등의 노력이 필요함.

다. 육아지원정책을 위한 향후 연구과제

□ 영아 및 부모 관련 연구 강화

□ 저출생 대응을 위한 돌봄 지원 확대 연구

□ 개정 누리과정에 대한 교사연수 및 협력 네트워크 지원 확대 연구

□ 교사 및 영유아 대상 문화다양성 교육 및 지원 강화 방안 연구

□ 영유아, 아동의 교육과 안전, 아동의 주거 환경, 아동의 신체적 건강, 교사와 부모의 배움, 소득, 힐링 등에 관련된 연구 필요

I

서 론

01 연구의 필요성 및 목적

02 연구내용

03 연구방법

Ⅰ. 서론

1. 연구의 필요성 및 목적

출산·양육·돌봄 등 육아와 관련된 다양한 데이터가 다양한 방식으로 실시간 생산되고 있다. 육아와 관련한 정보들이 블로그, 카페, SNS 등의 온라인 자료, 신용카드 자료, 스마트기기를 활용한 모바일 자료, 언론 자료 등 다양한 방식으로 실시간으로 축적되고 있다.

이들 정보는 성별, 연령, 계층, 지역 등 기초적인 분석뿐만 아니라 시점에 따른 패턴의 흐름 등으로 확대하여 새로운 가치를 지닌 분석이 가능하게 하고 있다. 또한 형평성 차원에서도 일반적인 자료나 정보로는 잘 나타나지 않는 육아지원 정책의 사각지대를 밝히는 데에도 활용될 수 있고, 미처 파악하지 못한 새로운 시각을 제공해 줄 수도 있다.

그러나 빅데이터에 대한 사회적, 정책적 활용 필요성이 높아지고 있음에도 불구하고, 육아정책과 관련하여 어떤 자료가 있고, 어떻게 활용될 수 있을지에 대해서는 논의가 부족한 실정이다. 또한 그간 육아정책 수립이나 평가를 위한 연구들은 주로 설문조사, 통계조사, 면담, 전문가조사, 행정통계 등을 활용하는 것이 대부분이다.

육아관련 현안을 파악하여 정책을 수립·시행하는데 있어 빅데이터의 활용 가능성은 상당할 것으로 기대되고, 특히 육아지원정책에 대한 온 국민의 관심이 높은 상황에서 이들의 다양한 요구는 빅데이터를 이용하면 기존의 연구방법을 통해 나타나지 않았던 새로운 내용이나 현안들이 도출될 수 있고, 이를 통해 수혜자가 필요로 하는 정책을 수립할 수 있을 것으로 기대된다.

이에 본 연구는 육아정책 분야에서 상대적으로 부족한 빅데이터가 어떻게 활용될 수 있는지에 대한 가능성을 보여주고, 빅데이터를 활용한 육아정책에 대한 아이디어 발굴이라는 새로운 접근방식에 대해 논의하고자 하였다. 이를 위해 특히 육아관련 빅데이터에 대한 탐색과 빅데이터를 활용한 사례 분석을 통해 육아지원 사각

지대나 정책적 요구의 파악 가능성과 향후 빅데이터를 활용하여 육아정책의 현안 파악과 과제 도출을 위한 시사점을 제안하고자 하였다.

2. 연구내용

연구목적에 따른 구체적 연구내용은 다음과 같다.

첫째, 빅데이터에 대한 개념을 살펴보고 관련 법·제도·현황을 살펴보았다. 먼저 빅데이터의 개념과 함께 빅데이터 관련 법 및 정부의 주요정책 추진현황을 살펴보았다. 다음으로 국내외 빅데이터 활용사례, 빅데이터와 육아정책과의 연관성을 파악하고 육아정책 관련 빅데이터 등을 탐색한다. 또한 육아분야에서의 빅데이터 활용 선행연구들을 살펴보았다.

둘째, 육아정책 관련 빅데이터를 활용·분석하였다. 우선 공공데이터를 활용하여 육아관련 사례분석이 가능한지를 파악하며, 여기에서는 유치원과 어린이집의 환경·안전과 관련한 빅데이터를 중심으로 분석하였다. 또한 임신·출산과 보육·교육·돌봄 등 육아에 대한 직접 수요자인 부모들이 이용하는 육아관련 인터넷카페의 게시글에 나타난 경험과 인식 등에 기반하여 토픽분석을 실시하였다. 미지막으로 학술연구에 나타난 연구자들의 관심주제를 분석함으로써 최근의 연구동향과 함께 향후 연구주제를 논의하였다.

셋째, 다양한 육아관련 빅데이터에 대한 사례 분석 결과 등에 기초하여 육아정책 현안 및 과제 제안의 가능성을 모색해 보고, 향후 육아정책 관련 빅데이터 활용 활성화와 관련 환경 개선 등에 대한 시사점을 도출하였다.

3. 연구방법

본 연구를 수행하기 위한 연구방법은 다음과 같다.

가. 문헌분석

빅데이터 개념, 빅데이터 관련 법 및 정책추진현황, 빅데이터 활용 선행연구 등

에 대한 문헌분석을 실시하였다. 정부 빅데이터 정책 및 활용·연계 관련 자료를 검색하여 다루고, 육아정책관련 공공데이터 현황, 빅데이터와 육아정책 관련 연구 등 육아정책 분야에서의 빅데이터 활용 사례 등에 대한 문헌을 검토하였다.

나. 빅데이터 통계분석

육아정책 관련 다양한 빅데이터를 수집하여 분석한다. 본 연구는 빅데이터는 크게 대기질, 위험시설, 수질 등과 관련한 안전·환경관련 공공 빅데이터 등 정형데이터와 인터넷 육아카페(맘스홀릭) 게시물과 KCI 논문초록 등 비정형데이터를 사용하여 분석한다.

비정형데이터는 대해서는 텍스트마이닝을 통한 키워드분석, 추이분석 등을 실시하였다. 육아와 관련된 중요 키워드들을 도출하고, 오픈소스와 빅데이터 분석기법을 이용하여 지역별·시기별 현안도 파악해 보고자 하였다.

일반적인 비정형데이터 텍스트마이닝 분석은 주요키워드를 통한 텍스트 데이터 수집단계, 불용어 제거 단어(명사) 추출 등 전처리단계, 분석 가능한 형태로 용어 행렬을 만드는 등의 정형화된 데이터로 변환, 정형화된 데이터로 빅데이터 분석 절차를 통해 수행한다(강창완, 조장식, 최승배, 2018). 빅데이터 분석도구로는 R과 파이썬을 사용하였고 빅데이터 분석 전문업체((주)한국정책리서치)와의 협력을 통해 수행하였다.

정형데이터에 대해서는 기초통계와 추이분석, 지역별 분석 등을 실시한다. 환경빅데이터 플랫폼(envbigdata)에서 제공하는 유치원과 어린이집의 주소와 인근 대기질 지수, 미세먼지 지수, 위험시설, 수질지수, 수질오염원 인근 거주인구 정보 등의 자료를 활용하였고, arcGIS를 통해 지도 및 그래프 등 안전·환경 관련 시각화를 도모하였다.

다. 전문가 회의

육아정책분야 빅데이터의 활용가능성에 대한 전문가 자문, 비정형데이터 분석을 위한 주요 키워드 도출 및 관련 전문가 자문을 실시하였다.

Ⅱ

빅데이터와 육아정책

01 빅데이터 개념

02 빅데이터 관련 법 및 주요 정책 추진 현황

03 빅데이터 정책 활용·연계 사례

04 육아정책 관련 빅데이터

05 빅데이터 활용 육아정책 선행연구

Ⅱ. 빅데이터와 육아정책

여기에서는 빅데이터의 일반적인 개념을 알아보고 우리나라 빅데이터 정책 및 주요 현황, 육아정책과 빅데이터의 연관성 및 관련 현황, 빅데이터 정책 활용·연계 사례, 빅데이터 활용 육아정책 선행연구를 살펴보았다.

1. 빅데이터 개념

빅데이터는 정보통신기술 발전, 스마트기기의 확산, 사물인터넷(IoT) 확대, 소셜 네트워크 서비스(SNS) 활성화에 따른 다양한 데이터의 폭발적 증가로 주목받기 시작했으며(최봉, 윤종진, 2018), 전 분야에 걸쳐 보다 의미있는 가치 창출을 위해 그 중요성이 날로 커지고 있다.

빅데이터(Big Data)의 개념은 데이터의 규모, 기술, 가치 등 접근하는 방식마다 약간의 차이는 있으나 기본적으로 규모 면에서 '대용량 데이터'라는 공통성을 갖는다. 이전에는 주로 규모와 기술의 관점에 주목하였으나, 이후 관리 · 분석 · 처리 · 축적하는 기술과 인력에 더해 데이터의 활용 효과까지도 아우르는 확장된 개념으로 사용되고 있다([그림 Ⅱ-1-1] 참조). 유사한 맥락에서, 이전에는 데이터를 저장, 관리하고 현황을 분석하는 대상으로 인식했다면, 데이터의 역할은 점차 분석과 추론(분석)의 방향으로 진화하였다. 빅데이터는 사회 환경변화를 신속하게 감지하고 대응하는 역량, 스마트한 지능형/개인화 서비스를 창출하는 원천요소로 인식되고 있다(정지선, 2011).

[그림 Ⅱ-1-1] 빅데이터 의미의 확장

자료: 교육과학기술부, 행정안전부, 지식경제부, 방송통신위원회, 국가과학기술위원회(2012). 스마트 국가 구현을 위한 빅데이터 마스터플랜, p.1.

Manyika 외(2011)는 Mckinsey 보고서에서 빅데이터를 '기존 데이터베이스 관리 도구를 통해 데이터를 수집·저장·관리·분석하는 역량을 넘어서는 새로운 개념의 데이터셋(Dataset)'으로 정의한 후 이를 주관성과 지속적 변화 가능성을 가진 개념으로 보았으며, 국제데이터협회(IDC: International Data Corporation)는 분석 비용 및 기술의 관점에서 '다양한 종류의 대규모 데이터로부터 가치를 추출하고 데이터의 초고속 수집·발굴·분석을 지원하도록 고안된 차세대 기술 및 아키텍처'로 정의하였다(Vesset et al,. 2012). 또한 Gartner(2012)는 빅데이터를 '향상된 시사점(insight) 발견과 더 나은 의사결정을 위해 사용되는 고효율, 대용량, 초고속 및 다양성의 특성을 가진 정보자산'으로 정의하였다.

국내에서도 다양한 접근의 정의가 이루어졌다. 삼성경제연구소의 함유근, 채승병(2012)은 빅데이터를 '기존의 방식으로는 관리와 분석이 매우 어려운 데이터 집합, 그리고 이를 관리·분석하기 위해 필요한 인력과 조직 및 관련 기술까지 포괄하는 용어'라고 말했으며, 정용찬(2013)은 '디지털 환경에서 생성되는 데이터로 규모가 방대하며 생성 주기도 짧고, 형태에 있어서도 수치 데이터뿐 아니라 문자와 텍스트 및 영상 데이터를 포함하는 대규모 데이터'로 정의했다. 또한, 한국정보통신기술협회(2015)는 빅데이터를 새로운 가치 추출을 위해 기존의 기술 또는 기법으로 처리하기 어려운 특징(규모, 신속성, 가변성 다양성, 진정성)을 갖는 데이터 모음으로 보았다. 한편, 빅데이터라는 용어가 처음으로 직접 언급되기 시작한 정부 정책에 해당하는 '스마트 국가 구현을 위한 빅데이터 마스터플랜'에서는 빅데이터를 '데이터 형식이 다양하고, 생성 속도가 매우 빨라 새로운 관리·분석 방법이 필요한 대용량 데이터'로 정의하였다(교육과학기술부 외, 2012).

종합적으로 볼 때, 빅데이터는 대용량 데이터를 수집 · 저장 · 관리 · 분석 · 체계화를 위한 도구, 플랫폼, 인력, 조직 등을 포괄하는 용어로, 대용량 데이터를 통해 의미있는 정보를 수집하고 생성된 지식을 바탕으로 상황을 능동적으로 대응하거나 변화를 예측하기 위한 정보화 기술을 의미한다고 볼 수 있다.

빅데이터는 정형의 정도에 따라 정형, 반정형, 비정형 데이터로 구분할 수 있다. 정형데이터는 고정된 형태로 저장되어 있으면서 연산이 가능한 데이터이고, 비정형 데이터는 영상, 음성 데이터뿐만 아니라 SNS에서 생성되는 소셜데이터까지 포함하는 개념으로, 일정한 형태로 저장되어 있지 않은 데이터이다. 반정형 데이터

는 로그 정보 등 데이터의 형태는 있으나 구조정보에 해당하여 연산은 불가능한 데이터를 말한다(〈표 Ⅱ-1-1〉 참조).

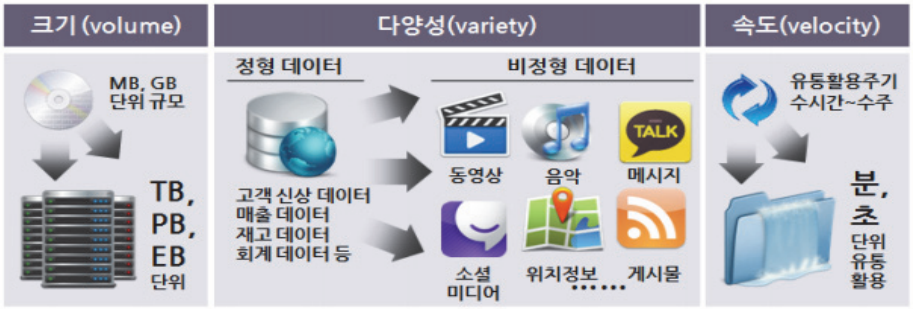
〈표 Ⅱ-1-1〉 빅데이터 분류

구분	내용
정형 (Structured)	고정된 필드에 저장된 데이터로, 지정된 행과 열에 의해 속성이 구별되는 형태 (예) 관계형 데이터베이스 및 스프레드시트 등
반정형 (Semi-structured)	메타데이터나 스키마 등을 포함하는 데이터로, 데이터의 구조에 대한 메타정보를 갖고 있는 형태 (예) XML이나, HTML, 텍스트 등
비정형 (Unstructured)	고정된 필드에 저장되어 있지 않은 데이터로, 사전에 정의된 구조가 없는 형태 (예) 텍스트 문서 및 이미지/동영상/음성 데이터 등

자료: 최봉, 윤종진(2018). 서울시 공공빅데이터 활성화방안. p.9의 내용을 발췌하여 정리함.

빅데이터의 특징에 대해서도 다양한 견해가 있으나 일반적으로는 3V, 즉 규모(Volume), 다양성(Variety), 신속성(Velocity)으로 설명한다([그림 Ⅱ-1-2] 참조).

[그림 Ⅱ-1-2] 빅데이터의 특징



자료: 채승병, 안신현, 전상인(2012). 빅데이터: 산업 지각변동의 진원. p.2.

규모(Volume)의 증가는 기술의 발전, IT의 일상화 등에 따른 데이터 크기의 증가를 의미한다. 전 세계적으로 생성된 디지털 정보량은 이미 ZB(제타바이트) 규모이며, 개별 데이터들의 양도 TB(테라바이트)를 넘어 PB(페타바이트), EB(엑사바이트) 등 기존의 관리 및 분석 기술로는 활용이 어려운 정도이다. 대규모성이라는 특징은 활용성 측면에서 다양한 이점이 있지만, 그 규모에 비례해 개인정보 등 민감도 높은 자료도 많아지게 되므로 개인정보보호 문제 등을 충분히 고려해야 한다(이

상윤, 2014). 다양성(Variety)은 데이터 유형과 종류의 범위가 넓어졌음을 뜻한다. 정형 데이터뿐만 아니라 사진, 오디오, 비디오 등 비정형화된 데이터 유형이 다양화되었고, 로그기록, 소셜, 위치, 소비 등 데이터의 종류가 증가하였다. 신속성(Velocity)은 데이터 생성, 수집, 분석의 실시간성을 의미한다. 데이터가 생성된 후 유통·활용되기까지 소요되는 시간이 분, 초 단위로 단축되었다. 센서, 모니터링 등의 사물정보를 포함하여 스트리밍 정보, 로그 정보 등 실시간성 정보가 증가했으며, 이에 대한 데이터 처리와 분석 역시 실시간으로 가능해졌다.

때로는 빅데이터의 주요 특성에 복잡성(complexity), 가치(value)등을 포함하기도 한다. 복잡성(complexity)은 데이터 관리 및 처리의 복잡성 심화로 대변된다. 비구조화된 데이터, 데이터 저장방식의 차이, 중복성 문제 등이 대표적이며 이에 따라 데이터 수집, 처리, 분석의 전 과정에서 복잡한 기술들이 계속적으로 요구된다. '가치(value)'를 강조하는 경우들도 있는데, 이는 빅데이터가 데이터를 수집하고 분석하는 수준을 넘어 진단, 예측 등을 통해 현상에 대한 의미를 발견하고 새로운 가치를 창출하는 것에 초점이 있음을 반영한다.

각 특징에 따라 기존의 데이터와 빅데이터를 구분하여 정리하면 〈표 Ⅱ-1-2〉와 같다.

〈표 Ⅱ-1-2〉 기존 데이터와 빅데이터의 비교

구분	기존 데이터	빅데이터
규모(volume)	MB(메가바이트), GB(기가바이트)	TB(테라바이트), PB(페타바이트)
신속성(velocity)	수시간 ~ 수주	실시간성으로 인한 데이터 생성과 이동 속도 증가
다양성(variety)	정형 데이터 위주 (매출, 재고, 인사 생산, 고객 데이터)	비정형화된 다양한 데이터 유형 (텍스트, 이미지, 동영상, 로그 등)
복잡성(complexity)	상대적으로 단순함	매우 복합적임
가치(value)	쓸모없는 자료로 판단되어 버려지는 데이터 존재	분석기술의 발전으로 새로운 가치를 부여하여 귀한 자료로 인식

자료: 김영옥(2016). 빅데이터 관련연구의 현황과 한국 유아교육학의 과제 p.7 〈표 1〉을 일부 수정함.

지금까지 살펴본 빅데이터에서의 핵심은 수집한 방대한 양의 데이터를 통해 어떤 의미있는 가치를 도출하는가이다. 이러한 맥락에서, 빅데이터는 불확실성, 리스크, 스마트, 융합 등 미래사회의 특성에 대응하는 다양한 역할을 수행하며 기회요

인을 창출하는 중요한 역할을 해낼 것으로 기대된다(정지선, 2011). 미래사회의 특성과 각 특성에 대응하는 빅데이터의 역할은 〈표 Ⅱ-1-3〉과 같이 정리할 수 있다.

〈표 Ⅱ-1-3〉 미래사회의 특성과 빅데이터의 역할

미래사회의 특성	빅데이터의 역할	
불확실성	통찰력	• 사회현상, 현실세계의 데이터를 기반으로 한 패턴분석과 미래전망 • 여러 가지 가능성에 대한 시나리오 시뮬레이션 • 다각적인 상황이 고려된 통찰력을 제시 • 다수의 시나리오로 상황 변화에 유연하게 대처
리스크	대응력	• 환경, 소셜, 모니터링 정보의 패턴 분석을 통한 위험징후, 이상 신호 포착 • 이슈를 사전에 인지/분석하고, 빠른 의사결정과 실시간 대응 지원 • 기업과 국가 경영의 투명성 제고 및 낭비요소 절감
스마트	경쟁력	• 대규모 데이터 분석을 통한 상황 인지, 인공지능 서비스 등 가능 • 개인화, 지능화 서비스 제공 확대 • 소셜(니즈)분석, 평가, 신용, 평판 분석을 통해 최적의 선택 지원 • 트렌드 변화 분석을 통한 제품 경쟁력 확보
융합	창조력	• 타 분야와의 결합을 통한 새로운 가치창출(의료, 자동차, 건물, 환경 등) • 인과관계, 상관관계가 복잡한 컨버전스 분야의 데이터 분석으로 안정성 향상, 시행착오 최소화 • 방대한 데이터 활용을 통한 새로운 융합시장 창출

자료: 정지선(2011). 신가치창출 엔진, 빅 데이터의 새로운 가능성과 대응전략. p.4.

2. 빅데이터 관련 법 및 주요 정책 추진 현황

가. 빅데이터 관련 법

현재 빅데이터에 대해서는 일반법적 성격을 갖는 법률이 제정된 바 없으므로 보다 포괄적인 의미인 정보 관련, 데이터 관련 법령을 따른다. 관련 법령은 데이터 활용 촉진, 데이터 활용 제한의 두 가지 측면으로 접근이 가능하다.

데이터 활용을 촉진하는 대표적인 법률은 1996년에 제정된 「공공기관의 정보공개에 관한 법률」(법률 제17690호, 이하 '정보공개법'), 2013년에 제정된 「공공데이터의 제공 및 이용 활성화에 관한 법률」(법률 제17344호, 이하 '공공데이터법'), 최근(2020.12.10.)에 제정된 「데이터기반행정 활성화에 관한 법률」(법률 제17370호, 이하 '데이터기반행정법') 등이 있다.

이 중 정보공개법은 공공기관이 보유·관리하는 정보는 국민의 알권리 보장 등

을 위하여 이 법에서 정하는 바에 따라 적극적으로 공개하여야 한다는 법이다. 정보공개는 공문서 성격을 가지는 일체의 기록물을 대상으로 한다.

또한 공공데이터법을 통해 국민의 공공데이터에 대한 이용권 보장 및 민간 활용을 보장하도록 하고 있다. 해당 법에 따르면, 공공데이터는 '데이터베이스, 전자화된 파일 등 국가기관, 지방자치단체 및 「국가정보화 기본법」제3조 제10호에 따른 공공기관이 법령 등에서 정하는 목적을 위하여 생성 또는 취득하여 관리하고 있는 광(光) 또는 전자적 방식으로 처리된 자료 또는 정보'로 정의된다.

데이터기반행정법은 데이터를 정책 수립과 의사결정에 활용하여 객관적이고 과학적으로 수행하는 행정을 하라는 국민의 요구와 시대적 요청에서 시작되었다(권헌영, 2020). 이는 행정·공공기관 간 데이터 공동 활용을 촉진하고 정책 수립에 데이터 활용을 활성화할 수 있도록 규정한 것으로, 여기서 데이터기반행정은 '공공기관이 생성하거나 다른 공공기관 및 법인·단체 등으로부터 취득하여 관리하고 있는 데이터를 수집·저장·가공·분석·표현하는 등의 방법으로 정책 수립 및 의사결정에 활용함으로써 객관적이고 과학적으로 수행하는 행정'을 뜻한다. 정부는 본 법률을 제정함으로 데이터기반행정에 대한 법적 근거를 마련하고, 올해 2월 '제1차 데이터기반행정 활성화 기본계획'을 발표하여 데이터기반행정의 정부혁신을 통한 지능형 정부로의 전환을 위한 첫걸음을 떼었다.

한편, 빅데이터 활용 제한과 관련된 법률은 2011년에 제정된 「개인정보 보호법」(법률 제10465호), 2001년에 제정된「정보통신망 이용촉진 및 정보보호 등에 관한 법률」(법률 제6360호, 이하 정보통신망법) 등이 있다.

개인정보 보호법은 국제수준의 통합적 개인정보 보호원칙을 규정하고, 개인정보 보호의 사각지대를 해소함으로써, 헌법상 사생활의 비밀과 자유 등을 보호하기 위한 목적으로 제정되어 운용되고 있다(이상윤, 2014). 정보통신망법은 정보통신망의 이용을 촉진하고 정보통신서비스를 이용하는 자를 보호함과 아울러 정보통신망을 건전하고 안전하게 이용할 수 있는 환경을 조성하여 국민생활의 향상과 공공복리의 증진에 이바지함을 목적으로 한다.

소개한 대표적인 법률 외에도, 〈표 Ⅱ-2-1〉과 같은 빅데이터 관련 법제가 있다.

〈표 Ⅱ-2-1〉 빅데이터 관련 법령

구분		내용
빅데이터 활성화	범부처	• 공공기관의 정보공개에 관한 법률(정보공개법) • 공공데이터의 제공 및 이용 활성화에 관한 법률(공공데이터법) • 전자정부법 • 국가정보화기본법 • 민원사무 처리에 관한 법률 • 행정절차법
	분야별	• 기상법/기상산업 진흥법(기상정보) • 통계법(통계정보) • 발명진흥법(특허정보) • 공간정보산업 진흥법(공간정보) • 교육관련기관의 정보공개에 관한 특례법
빅데이터 활용 제한		• 공공기관의 정보공개에 관한 법률(비공개대상정보) • 행정조사기본법(행정조사 관련정보 활용제한) • 국유재산법(국가재산처분제한, 비용산정 등) • 공유재산 및 물품관리법(지자체 재산의 처분제한 등) • 저작권법(권리처리 등) • 의료법

자료: 이상윤(2014). 빅데이터법제에 관한 비교법적 연구. p. 43의 표를 일부 수정함.

한편, 2020년 1월에 데이터3법, 즉「개인정보보호법」,「정보통신망법」, 「신용정보법」에 큰 변화가 있었다. 제4차 산업혁명 시대의 도래에 따른 세계 각국의 데이터경제 활성화 전략 추진, EU(유럽연합)의 일반 개인정보보호법(GDPR; General Data Protection Regulation) 시행(2018.5.), 국내에서의 개인정보의 효율적 활용을 위한 법 개정의 거센 요구 등에 근거하여 새로운 모습으로 개정되었다(강달천, 2020). 이는 2020년 2월 4일에 공표되었고, 당해 8월 5일부터 시행되었다. 데이터3법 개정의 주요 내용은 〈표 Ⅱ-2-2〉와 같다.

〈표 Ⅱ-2-2〉 데이터3법 개정 주요 내용

구분	주요 내용
개인정보보호법	• 개인정보보호위원회의 중앙행정기관화 • 개인정보 범위의 판단 기준 제시 • 가명정보의 제도화 • 수집목적과 합리적 관련 범위 내에서의 활용 확대
정보통신망법	• 「개인정보보호법」과 유사·중복 규정 삭제 • 삭제된 일부 규정을 「개인정보보호법」 내에 특례 규정으로 이관 • 「정보통신망법」에 존치하는 규정

구분	주요 내용
신용정보법	• 개인정보보호법」과 유사·중복 조항의 정비 등 • 가명정보의 개념 도입 등 • 정보활용 동의제도의 내실화 • 새로운 개인정보자기결정권의 도입 • 신용정보 관련 산업의 규제체계 변경 등

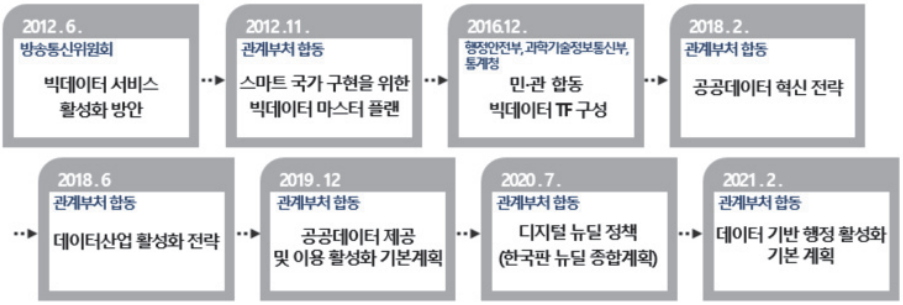
자료: 강달천(2020). 데이터 3법 개정의 주요 내용과 전망. pp.15~pp.18의 내용을 발췌하여 정리함.

데이터3법 개정을 통해 개인정보의 정의와 범위가 보다 명확해지고, 가명정보 이용의 통로가 제도화되었다. 이러한 변화는 관련 법률의 효율적·통일적·체계적 집행 및 규제 대상자의 법 준수 혼란 최소화 등에 기여하는 동시에, 데이터 활용폭 확장 및 빅데이터 이용 활성화 등을 통해 우리나라의 사회 및 경제 발전에 크게 기여할 것으로 예측한다(강달천, 2020).

나. 빅데이터 관련 주요 정책 추진 현황

정부는 2012년에 발표한 '스마트 국가 구현을 위한 빅데이터 마스터 플랜'을 시작으로 빅데이터와 관련한 다양한 정책을 추진하고 있으며, 이러한 정책들은 대부분 특정 부처, 개별 부서별로 추진되기 보다 범부처, 범정부 차원의 성격을 갖는다. 주요한 정책들을 시간의 흐름에 따라 정리하면 [그림 Ⅱ-2-1]과 같다.

[그림 Ⅱ-2-1] 빅데이터 관련 주요 정책



1) 빅데이터 서비스 활성화 방안

2012년 6월 21일에 방송통신위원회는 '빅데이터 서비스 활성화 방안'을 발표했다. 이는 세계 최고수준의 IT인프라와 빅데이터를 접목하여 향후 우리나라가 빅데

이터라는 새로운 IT 패러다임 시대에 뒤처지지 않고 글로벌 시장에서 경쟁할 수 있는 토대를 마련하기 위해 업계 및 학계의 의견을 34차의 위원회 회의를 통해 수렴한 결과에 해당하는 내용이다(방송통신위원회 보도자료, 2012.6.21.).

이 정책을 통하여 추진한 주요 과제는 다음 7가지로 정리할 수 있다. 첫 번째, 신규 서비스 발굴·확산을 위한 시범서비스를 추진하였다. 방송통신뿐만 아니라 교육, 의료 등 다양한 분야에 대한 시범 서비스를 공모방식으로 발굴하여 빅데이터에 대한 인식을 제고하고 이용자의 편익을 향상시키고자 하였다. 두 번째, 빅데이터 분석과정에서 필요한 핵심 기술들을 개발하고, 오픈소스 기반의 플랫폼을 개발함으로 빅데이터 기술 및 플랫폼 경쟁력을 강화하고자 하였다. 세 번째, R&D 및 시범사업을 대학과 연계하여 추진함으로 석·박사급 고급인력을 양성하고, 취업 예정자 및 실무자 재교육을 실시하는 등 전문인력을 확보하고자 하였다. 네 번째, 빅데이터 지원센터 설치·운영 및 정보공유 체계를 마련하고자 하였다. 여기에서 빅데이터 지원센터는 연구 개발을 위한 테스트베드, 시범서비스 수행, 중소기업 컨설팅 등을 제공하는 성격을 갖는다. 다섯 번째, 빅데이터 산업 및 활용실태 조사를 추진하였다. 이는 빅데이터 산업의 생태계 조성을 위한 기초자료로 활용하기 위한 목적이었으며, 국내 빅데이터 산업 관련 업체, 시장 규모, 빅데이터에 대한 인식 및 활용 현황 등의 내용을 포함하였다. 여섯 번째, 개인정보보호를 위한 제도적, 기술적 장치를 마련하고 빅데이터를 활용하는 기업의 개인정보 관리수준을 검증하는 개인정보 보호관리체계(PIMS) 인증제 개선을 추진하고자 하였다. 마지막으로, 빅데이터 서비스 및 산업 진흥을 위해 필요한 법제도 개선사항을 발굴하고 종합적인 정책방안을 검토하고자 하였다.

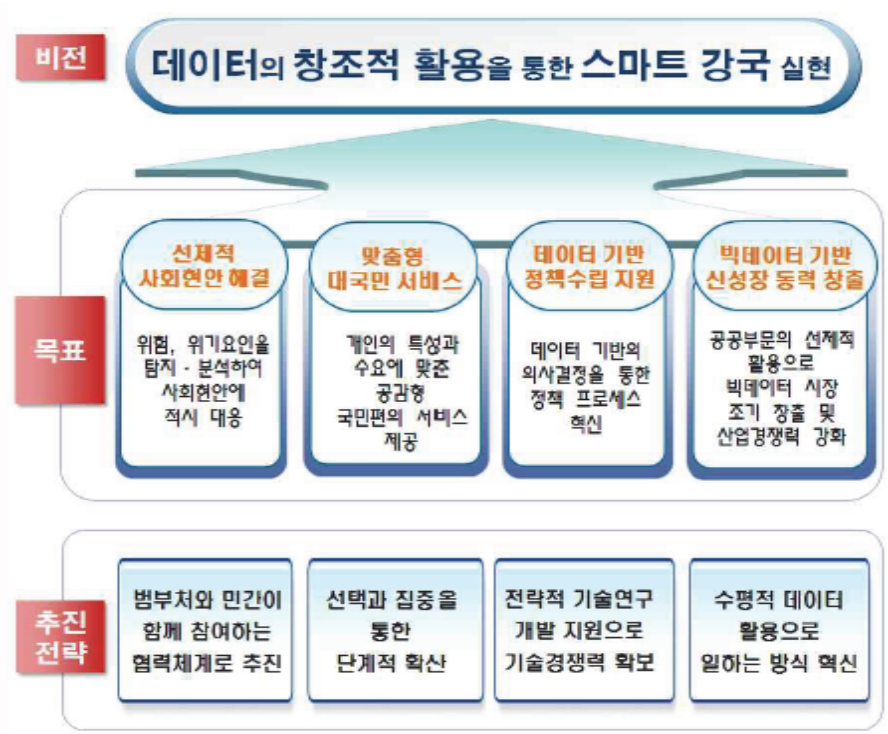
2) 스마트 국가 구현을 위한 빅데이터 마스터 플랜

빅데이터라는 용어가 직접적으로 언급되기 시작한 범정부 차원의 정책은 '스마트 국가 구현을 위한 빅데이터 마스터 플랜'이다. 이는 2012년 11월에 교육과학기술부, 행정안전부, 지식경제부, 방송통신위원회, 국가과학기술위원회 등 5개의 기관이 함께 수립한 정책으로, '데이터의 창조적 활용을 통한 스마트 강국 실현'이라는 비전 아래 창의적인 빅데이터 활용으로 생산성 향상, 기업 경쟁력 제고, 국가 미래전략 지원, 공공서비스 혁신을 도모하고자 하였다(교육과학기술부

외, 2012).

우리나라 현황과 각 부처의 의견, 외국사례 등을 참조하여 국민의 관심이 큰 6개 분야(사회안전, 국민복지, 국가경제, 국가 인프라, 산업지원, 과학기술)에서 빅데이터 대상과제 16가지를 제시하였으며, 대부분의 과제는 '주민참여형 교통사고 감소 체계 구축'과 같이 필요성과 기대효과가 명확하다. 이같이, 당시에는 빅데이터를 행정혁신의 수단으로 주목하여 빅데이터를 활용한 과학적 문제해결이 중점적으로 담긴 활용과제가 주를 이루었다(최봉, 윤종진, 2018). 활용과제와 병렬적으로 빅데이터 활용기반 조성을 위한 과제들이 제시되었는데, 빅데이터 공유활용 인프라 구축, 기술개발지원, 전문인력 양성 등의 내용을 포함하였다. 본 정책의 비전과 목표는 [그림 Ⅱ-2-2]와 같다.

[그림 Ⅱ-2-2] 빅데이터 마스터 플랜 비전 및 목표



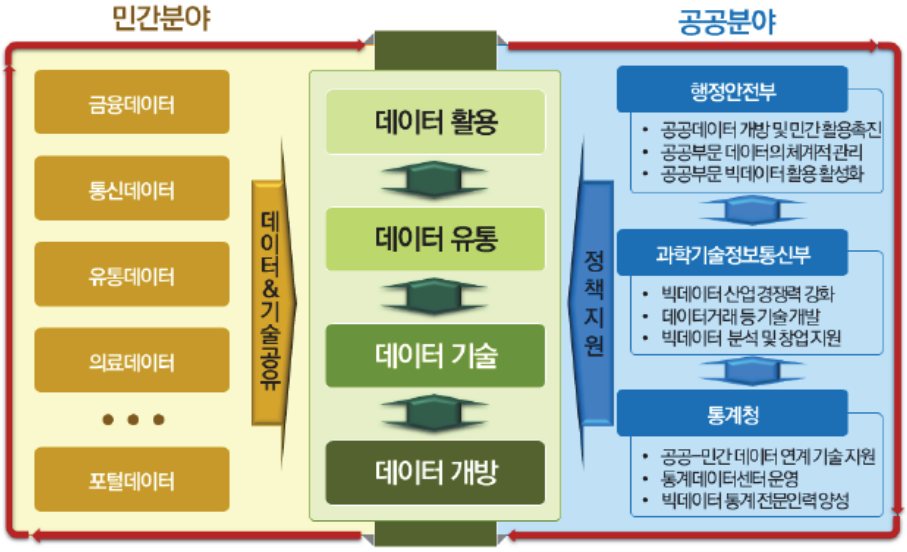
자료: 교육과학기술부 외(2012). 스마트 국가 구현을 위한 빅데이터 마스터플랜. p.17.

3) 민·관 합동 빅데이터 TF 구성

정부는 2016년 빅데이터 업무 추진 현황을 종합적으로 파악하고, 관련 공공과 민간 빅데이터 활성화를 지원하기 위하여 행정안전부, 과학기술정보통신부, 통계청 등 3개 기관이 참여하는 '민관 합동 빅데이터 TF'를 출범시켰다(관계부처 합동 보도자료, 2016.12.9.). 이전까지 각 부처는 공공 빅데이터협의회(당시 행자부), K-ICT 빅데이터협의회(당시 미래부), 빅데이터-통계 전략포럼(통계청) 등 개별적으로 빅데이터 관련 협의체를 운영해 왔으나, 빅데이터 TF 구성을 통해 범정부 차원의 종합적인 추진체계를 확립하게 되었다.

민관 합동 빅데이터 TF에서 행정안전부는 공공부문 빅데이터 관련 업무, 과학기술정보통신부는 민간부문 빅데이터 관련 업무, 통계청은 데이터 간 연계·분석 및 활용 지원업무를 수행하였으며, 민간의 데이터 전문기업과 학계 전문가, 데이터 보유 공공기관, 지자체 및 관계부처는 팀원으로 참여하였다. 민간 및 공공분야의 빅데이터 추진체계는 [그림 Ⅱ-2-3]과 같다.

[그림 Ⅱ-2-3] 민간 및 공공분야의 빅데이터 추진체계



자료: 행정안전부(2018). 빅데이터 추진체계 현황 및 향후과제. p.7.

민·관 합동 빅데이터 TF는 2016년 하반기부터 총 3차례의 회의를 통해 빅데이터에 대한 정부와 민간의 소통 및 협업 강화에 대한 논의를 진행하였다(〈표 Ⅱ-2-3〉 참조).

〈표 Ⅱ-2-3〉 빅데이터 TF

구분	일시	내용
1차 회의	2016. 12. 8.	• 부처의 빅데이터 주요 정책 및 과학기술정보연구원의 과학 데이터 활용 계획 공유
2차 회의	2017. 2. 24	• 부처별 17년 빅데이터 관련 추진계획 공유 • 빅데이터 활성화 방안 및 이를 위한 민관 협력체계 구축 방안 관련 논의
3차 회의	2017. 8. 30.	• 신정부 출범에 따라 각 부처별 빅데이터 추진방향 및 계획 공유 • 통계기반 빅데이터 산업활성화를 위한 각종 진흥정책과 규제개선 필요 사항 관련 논의

자료: 정도영, 김민창, 김재환(2018). 빅데이터 정책 추진 현황과 활용도 제고 방안. p.10 본문을 재정리함.

4) 공공데이터 혁신전략

2018년 2월 26일 제3기 공공데이터전략위원회가 공식적으로 출범하면서 공공데이터 혁신을 위한 '공공데이터 혁신전략'을 발표했다(관계부처 합동, 2018a). 이는 '국민이 주인인 정부'와 '혁신성장'을 목표로 하여 4대 추진전략, 15개 과제로 구성되어 있다.

이 정책은 데이터에 대한 인식과 기본방향에 상당한 변화가 있었음에 그 의의가 있다. 이전에는 공공데이터를 정부의 전유물로 인식하고 관련 내용들을 선택적으로 개방했다면, 이때부터는 공공데이터를 국민의 자산으로 인식하고 개인정보 등을 제외한 나머지 데이터는 개방하는 것을 원칙으로 보았다. 또한 데이터를 경제적 가치뿐만 아니라 사회적 가치도 구현할 수 있는 도구로 보았으며, 따라서 행정적인 부분 또한 데이터 기반, 증거 기반의 행정을 추진하고자 하였다. 한편, 빅데이터 정책을 추진함에 있어서도 정부가 주도적으로 하는 체계를 지양하고 민관 협치를 지향함으로, 공공데이터와 민간데이터가 자유롭게 흐르도록 하였다([그림 Ⅱ-2-4] 참조).

[그림 Ⅱ-2-4] 공공데이터 정책의 추진방향

	종전에는 (As-is)	앞으로는 (To-be)
(기본인식)	√ 공공데이터는 **정부의 전유물**	√ 공공데이터는 **국민의 자산**
(개방방식)	√ 선택적 개방	√ 개인정보 등 제외 **원칙적 개방**
(구현가치)	√ 경제적 가치	√ 경제적 가치+**사회적 가치**
(행정추진)	√ **직관·경험**에 의한 행정	√ **데이터·증거**에 의한 행정
(추진체계)	√ **정부 주도**	√ **민관 협치**

자료: 관계부처 합동(2018a). 공공데이터 혁신전략. p.7.

5) 데이터산업 활성화 전략

정부는 그간의 노력에도 불구하고 글로벌 스탠다드에 뒤처진 데이터 이용제도, 양질의 데이터 부족, 취약한 산업·기술 기반 등이 4차 산업혁명 확산의 걸림돌로 작용하고 있다고 판단했다. 동시에, 데이터 경제의 급속한 진전에 대응, 혁신성장과 국민 삶의 질 향상을 동시에 실현하면서 새로운 기회로 활용하기 위한 전략을 마련하는 것이 필요함을 인지하였다. 이에 정부는 '데이터를 가장 안전하게 잘쓰는 나라'라는 비전을 바탕으로 데이터 경제 활성화, 데이터 보호 및 활용을 위한 11개의 추진 과제를 내세웠다(관계부처 합동, 2018b). 구체적으로, 개인정보 자기통제권 확립, 데이터 안전활용 제도화 및 신기술 적용, 민·관 데이터 획기적 개방, 개방형 데이터 거래 체계 구축, 데이터 기반 산업경쟁력 제고, 빅데이터 활용 사회문제 해결 강화, 빅데이터 선도기술 확보, 미래수요 대응 전문인력 양성, 빅데이터 전문기업 성장 지원 등이 이에 해당한다.

정부는 해당 정책을 시행하므로 데이터 경제 활성화를 통한 산업과 기업의 혁신성장 가속화, 데이터 활용 기반 생활밀착형 서비스 제공을 통한 국민 개개인의 행복 증진, 데이터 기반 의사결정 및 디지털 행정서비스를 통한 정책의 신뢰성 확보 등의 효과가 있을 것으로 기대하였다.

6) 공공데이터 제공 및 이용 활성화 기본계획

공공데이터 제공에 대한 정책 수립은 이번이 세 번째로, 이는 공공데이터법 제7조에 근거하여 3년마다 국가와 각 지방자치단체의 부문 계획(안)을 종합하여 수립하고, 전략위원회의 심의·의결을 거쳐 확정한다.

우리나라는 1, 2차 공공데이터 제공 기본계획을 추진함에 따라 공공데이터 개방이 대폭 확대되고 공공데이터 민간 이용과 창업이 활성화 되었으며, 공공데이터 표준화 등 데이터 관리 체계를 마련하는 성과를 이루었다. 특히, 데이터 가용성, 접근성 및 데이터 활용에 대한 정부지원 등 3개 분야 24개 항목을 측정하는 OECD 공공데이터 평가에서 3회 연속 세계 1위를 달성하는 등 국제평가에서 대한민국 정부의 공공데이터 정책 구현 수준을 인정받은 바 있다.

그럼에도 불구하고 지난 정책 추진의 결과에 대해 몇 가지 한계점이 도출되었고, 수요자가 원하는 고가치 데이터 개방 확대 필요, 성장단계별 데이터 기업 지원 및 유통·거래 기반 조성 필요, 데이터기반 과학적 행정 및 데이터 거버넌스 간 협력 필요 등의 시사점이 도출되었다. 그리고 이를 출발점으로 삼아 '제3차('20~'22) 공공데이터 제공 및 이용활성화 기본계획'을 마련하였다(관계부처 합동, 2019).

3차 기본계획의 비전은 '사회적 가치 실현과 디지털 혁신 성장을 선도하는 데이터 강국'이다. 2차 시기가 '확산' 단계에 해당했다면, 3차 시기는 '성과 창출'의 시기로 보았다. 추진 전략은 1) 국민이 원하는 융합 및 개방 데이터 활성화, 2) 데이터가 안전하고 쉽게 유통되는 생태계 조성, 3) 사회 곳곳에 데이터 분석 및 활용 확대, 4) 데이터 추진기반 강화를 통한 글로벌 스탠다드화 이다.

3차 기본계획에서 다루는 추진과제는 총 15개로, 기존의 과제를 확산, 발전시키는 과제 8개, 신규 과제 7개로 구성되어 있다. 추진 전략에 따라 1) 비정형 데이터 개발 확대 및 융합 데이터 구축, 2) 데이터 구축가공·성장단계별 지원 강화를 통한 신산업 창출 및 데이터 유통·거래 생태계 조성, 3) 데이터에 기반한 과학적·체계적 행정 구현 및 사회 문제 해결을 촉진하는 데이터 분석·활용 확대, 4)민관 간, 공공간 범정부 데이터 거버넌스 확립 및 글로벌 선도를 위한 공동연구 추진 등 국제협력 다각화의 핵심 추진과제를 갖는다.

한편, 공통의 비전 및 추진과제 외에 중앙행정기관 45개, 광역자치단체 17개의 개별 기본계획이 별도로 마련되어 있다.

7) 디지털 뉴딜 정책(한국판 뉴딜 종합계획 중)

정부는 2020년 7월 14일, 대통령 주재 한국판 뉴딜 국민보고대회(제7차 비상경제회의)를 개최하여 「한국판 뉴딜 종합계획」을 확정·발표하였다(관계부처 합동, 2020). 이는 경제 패러다임 전환 추진 과정에서 코로나19 사태로 인한 극심한 경기침체 극복 및 구조적 대전환 대응이라는 이중 과제에 직면하였음을 출발점으로 삼는다.

한국판 뉴딜은 정책 방향을 2+1으로 잡았는데, 그중 하나가 경제 전반의 디지털 혁신 및 역동성 촉진·확산을 위한 '디지털 뉴딜'이다. 10대 대표과제 중 디지털 뉴딜에 대한 과제는 데이터 댐, 지능형(AI) 정부, 스마트 의료 인프라 총 3가지이고, 이중 빅데이터와 밀접하게 연관되는 것은 '데이터 댐' 과제로, 데이터 수집 · 가공 · 거래 · 활용기반을 강화하여 데이터 경제를 가속화하고 5G 전국망을 통한 全산업 5G·AI 융합 확산하는 것을 목표로 한다. 정부는 해당 과제 추진을 통해 분야별 데이터를 확충·연계·활용하여 5G·AI 융합서비스에 대한 글로벌 선도를 기대한다.

과학기술정보통신부와 한국지능정보사회진흥원은 디지털 뉴딜 핵심 프로젝트인 '데이터 댐'의 일환으로 6개 분야(농식품, 디지털 산업혁신, 라이프로그, 소방안전, 스마트치안, 해양수산)의 빅데이터 플랫폼을 구축하여 서비스를 개시하였다(과학기술정보통신부 보도자료, 2021.2.23.). 또한 이 6개 플랫폼의 데이터를 통합 데이터지도(www.bigdata-map.kr)와 연계시켜 이용자가 언제든지 쉽고 편하게 검색하고 활용할 수 있도록 하였다. 데이터지도에서는 기존의 10개 분야를 포함하여 총 16개의 플랫폼이 서비스된다. 이 플랫폼은 정부·공공·의료기관이 플랫폼의 운영을 맡아 민간기업과 협업하여 중소기업이나 연구자가 접근하기 어려웠던 다양한 데이터와 서비스를 공급한다.

8) 데이터 기반 행정 활성화 기본 계획

정부는 그동안 지속적인 투자로 세계 최고 수준의 디지털정부 인프라를 구축하고, 우수한 디지털 서비스를 제공하는 과정에서 방대한 공공데이터가 축적되었으나 정책 수립·의사결정 지원을 위한 공공데이터의 분석·활용은 미흡했음에 주목하였다. 이에 「데이터3법」개정, 「데이터기반행정법」 시행으로 마련된 법적 근거를

바탕으로 2021년 2월, 「제1차 데이터기반행정 활성화 기본계획('21~'23)」을 발표하여 세계적 수준의 전자정부 인프라와 공공데이터 개방정책 경험 등을 기초한 데이터기반의 지능형 정부로의 도약에 대한 첫걸음을 디뎠다(관계부처 합동, 2021). '데이터기반 과학적 행정으로 국민의 삶의 질을 개선하는 지능형 정부서비스 구현'을 비전으로 세웠으며, 1) 데이터 통합 기반 구축으로 데이터 공동활용, 2) 데이터기반 행정의 활성화를 위한 제도 확립, 3)지능형 서비스 제공을 위한 데이터 분석·지원, 4) 데이터기반의 일하는 방식으로 혁신 등 4가지 추진전략을 수립하였다.

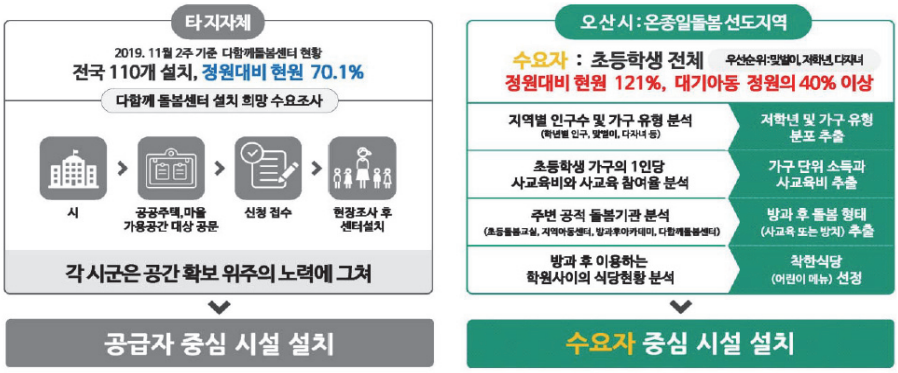
3. 빅데이터 정책 활용·연계 사례

서울시 심야버스 노선 지원 정책은 빅데이터 분석을 통해 공공서비스를 개선한 최초의 사례이자 성공사례로, 야간 유동인구가 많은 지역에 대한 분석을 바탕으로 노선을 결정하고 심야버스를 운영하여 시민들의 야간 귀가 문제를 효과적으로 해결한 정책이다(성지은, 박기량, 2014). 서울시는 서울시의 교통데이터와 통신회사 KT의 유동인구 데이터를 연계·분석하여 최적의 심야버스 노선을 구축하고자 하였으며 이를 통해 시민들의 야간통행 이동 비용 및 유류비 절감, 이산화탄소 배출 감소에 따른 환경 개선 등의 효과를 기대하였다(최봉, 2014). 구체적으로 살펴보면, 2013년 3월 한 달 동안 시민들이 자정부터 새벽 5시까지 사용한 KT 휴대폰 전화 데이터 30억여 건의 빅데이터를 분석하여 심야 유동인구가 가장 많은 지역을 확인했다. 또한, 심야 택시 승·하차 데이터, 카드사 데이터, 버스 승하차 데이터 등도 종합적으로 분석하여 교통수요량을 파악하였다(이승환 외, 2015). 수집된 정보들은 서울시의 '심야버스 노선수립 지원시스템'에서 분석되었는데 서울 전역을 일정 단위로 나누어 유동인구·교통수요량을 표시하고 기존의 버스노선과 시간·요일별 유동인구 및 교통수요 패턴을 분석했다. 이후 노선 부근 유동인구 가중치를 계산하는 등의 추가 분석 과정을 거쳐 최적의 노선과 배차간격을 도출했다(서울시, 2013).

오산시는 공공 빅데이터를 활용하여 수요자 돌봄 중심의 초등돌봄(온종일 돌봄) 정책을 추진하였다. 오산시는 지역별 인구수 및 가구 유형, 초등학교 가구의 1인당 사교육비와 사교육 참여율, 주변 공적 돌봄기관 등에 대한 분석을 통하여 수요자에

게 최적화된 돌봄시설 위치를 선정하였다([그림 Ⅱ-3-1] 참조). 나아가 방과후 이용하는 학원 사이의 식당현황을 분석하여 착한식당을 선정하였다(행정안전부, 한국정보화진흥원, 2019).

[그림 Ⅱ-3-1] 오산시 다함께돌봄센터 최적설치지역 분석



자료: 행정안전부, 한국정보화진흥원(2019). 2019년 공공 빅데이터 우수사례 경진대회 자료집. p.12.

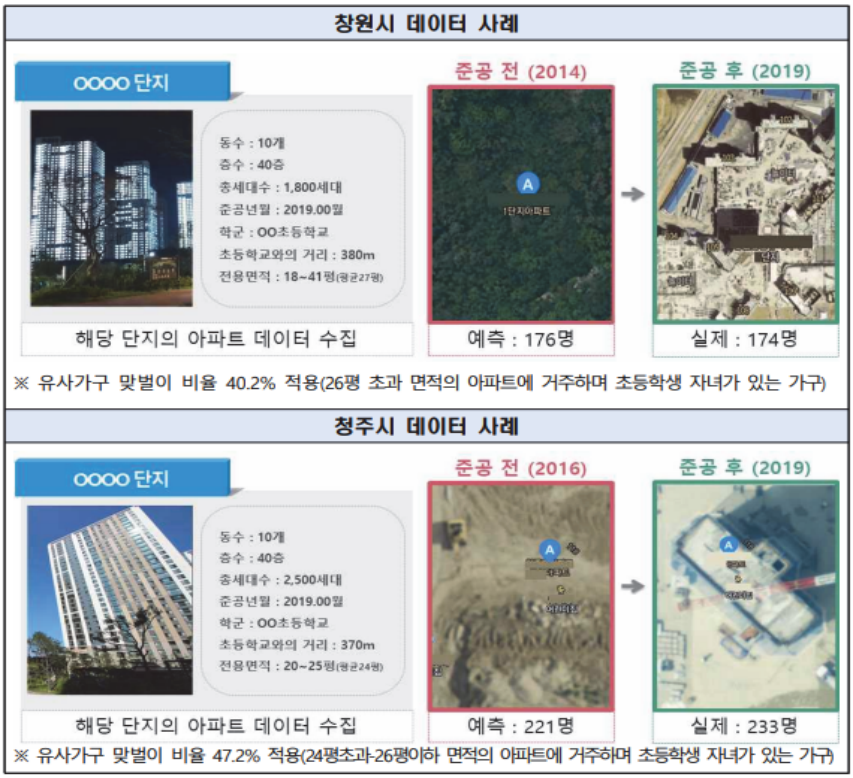
주민등록, 공동주택, 통계청 데이터를 주소매칭기술, 가구화 및 소득 추계 기법을 활용하여 수행한 후 가구단위 생활수준 데이터를 구축하였고, 이후 지역단위 가구 통계 데이터를 구축하여 초등학생 자녀가 있는 가구의 생활수준을 분석하였다(행정안전부, 한국정보화진흥원, 2019). 그리고 다함께돌봄센터 설치지역에 대한 기초분석을 토대로 다함께돌봄센터 설치 우선지역을 선정하였다([그림 Ⅱ-3-2] 참조).

[그림 Ⅱ-3-2] 오산시 다함께돌봄센터 설치 우선지역

자료: 행정안전부, 한국정보화진흥원(2019). 2019년 공공 빅데이터 우수사례 경진대회 자료집. p.16.

유사한 맥락에서 행정안전부, 교육부, 보건복지부는 아파트 시공 전에 돌봄 수요를 데이터 기반으로 예측하여 지역 내 돌봄 수요에 부합하는 다함께돌봄센터를 설치하는 방안을 추진한다고 발표하였다(관계부처 합동 보도자료, 2020. 7. 22.). 관련하여 2020년 1월부터 6개월간 창원시, 청주시의 데이터를 활용하여 신규 아파트 단지의 초등학생 돌봄수요를 예측하는 빅데이터 분석을 수행하였다([그림 Ⅱ-3-3] 참조). 초등 돌봄수요 예측 분석은 해당 아파트가 위치한 지역의 특성을 활용하고 관련 데이터에 기반한 알고리즘으로 분석하여 돌봄이 필요한 초등학생 수를 예측하는 것이다. 이에 전용면적, 세대당 주차대수, 건폐율 등 아파트의 구조와 관련된 데이터, 초등학교까지의 거리 등 아파트 주변의 환경 데이터 및 행정구역의 인구 데이터 등 총 290개의 변수가 활용되었다.

[그림 Ⅱ-3-3] 지역별 초등돌봄 수요예측 사례-창원시, 청주시



자료: 관계부처 합동 보도자료(2020. 7. 22.). 이제 빅데이터 분석으로 공동주택의 돌봄수요를 미리 예측한다. 붙임2.

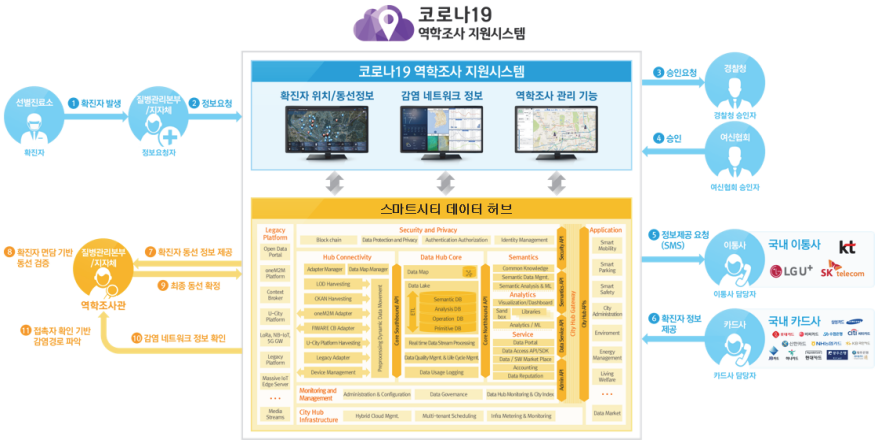
행정안전부와 인천광역시, 경찰청은 112 신고데이터와 신용카드 이용데이터 등을 결합·분석하여 범죄위험도를 분석했다(행정안전부 보도자료, 2020. 12. 10.). 구체적으로, 112 신고데이터 362만건, SKT 유동인구·카드매출 데이터 등을 결합하고 기계학습으로 월·일·시간 단위 무질서·범죄 위험도 예측모델을 구축하였다. 예측 결과에 근거하여 6주간 인천시 16개 지역에 순찰을 강화한 결과, 신고 건수와 범죄 발생 건수가 각각 23.7%, 9.7% 감소하였다. 이에 따라, 분석결과를 경찰청 범죄예방시스템에 반영하고 치안자원을 효율적으로 분배하여 범죄 발생의 사전예방에 활용하는 등 정책에 적극 활용하였다. 이외에도 다양한 정부부처와 지자체에서 빅데이터를 활용한 정책 수립이 활발하게 이루어지고 있다. 서울시는 2020년 9월, 스마트 민원분석 체계를 구축하고, 예측 기반의 선제적 행정서비스로의 패러다임을 전환하고자 'AI 기반 응답소 민원분석 시스템'을 구축했다(서울시 보도자료, 2020.9.18.). 2014년부터 누적된 1천만 건의 데이터뿐만 아니라 일평균 7,000여건에 달하는 접수 민원을 실시간 분석하기 위해 AI기술을 접목하였으며, 행정 및 정책에 활용하는 것을 넘어 누구나 쉽게 이 데이터를 분석·활용할 수 있도록 하였다. 서울시는 빅데이터를 활용한 민원 분석으로 주요 정책을 보완하고, 실시간 이슈 발굴을 통해 새로운 정책 수립의 기초자료로 활용하고자 하는 동시에 코로나19 확산 방역 정책에 따른 시민 피드백 등을 분석한다.

[그림 Ⅱ-3-4] 응답소 민원분석시스템 주요 화면-다양한 조건에 따른 민원 분석 및 동향 파악

자료: 서울시 보도자료(2020.9.18.). 서울시, 1천만 건 응답소 민원 빅데이터 AI로 분석. p.4.

한편, 코로나19 확산 등과 같이 긴급하고도 중요한 상황에 대한 정책적 대응에 있어 빅데이터가 크게 활약하고 있다. 일례로 국토교통부, 과학기술정보통신부, 질병관리본부는 '코로나19 역학조사 지원시스템'을 통하여 코로나19 확진자의 동선 파악 및 관리의 효율성과 효과성을 높였다([그림 Ⅱ-3-5] 참조). 이 시스템은 확진자 면접조사 결과를 보완하고, 빅데이터 실시간 분석을 통해 확진자 이동동선 및 시간대별 체류지점을 자동으로 파악하며, 대규모 발병지역을 분석하여 지역 내 감염원 파악 등 다양하게 활용된다(국토교통부 보도자료, 2020.3.26.).

[그림 Ⅱ-3-5] 역학조사 지원시스템



자료: 국토교통부 보도자료(2020. 3. 26.). 스마트시티 기술로 코로나19 확진자 동선 파악이 더 빠르고 더 정확해집니다. p.2.

성남시는 지역단위에서 코로나19 확산을 효율적으로 대응하기 위하여 빅데이터 기반 실시간 유동 인구 분석 기술을 도입했다[1]. 성남시는 SK텔레콤과의 협업을 통해 기지국 정보를 활용하여 지역별로 인구가 몰리는 정도를 성별, 연령별로 분석하여 시각화한 정보를 5분마다 업데이트하여 서비스하였다. 나아가 이 데이터를 활용하여 인구 밀집도가 높은 지역을 중심으로 방역 대책을 마련하고 소독 경로를 잡아 핀포인트 방역을 시행하였다. 현재는 다양한 지역에서 유사한 정책들이 시행되고 있다.

1) 서울경제, https://www.sedaily.com/NewsVIew/1Z2KYZABCW, 2021년 8월 2일 인출.

4. 육아정책 관련 빅데이터

육아정책에서 빅데이터는 크게 공공부문에서 제공하거나 공개하는 공공빅데이터와 이외 빅데이터가 있다.

육아정책분야의 대표적 공공 빅데이터로는 사회보장정보시스템(행복e음)에 집적되는 사회보장 데이터가 있다. 각종 사회복지 급여 및 서비스 지원 대상자의 자격과 이력에 관한 정보를 통합 관리하고, 지자체의 복지업무 처리를 지원하는 목적으로 개인별 가구별 DB로 중앙에 통합 구축한 데이터이다.[2] 또한 보육통합정보시스템은 전국 어린이집 원장, 복지부, 지자체 보육담당자, 유관기관 업무담당자, 보육교사가 사용하는 시스템으로, 보육바우처 및 행정지원 관련 DW, 대국민 포털 및 보육 유관기관 업무시스템에 대한 각종 데이터를 보유하고 있다.[3] 또한 유아교육 관련 기본 자료는 나이스(NEIS) 284종 중 학교현황(유치원 버스 운영 실태 현황, 유치원 사무직원 현황, 유치원 수료 현황, 유치원 신입 및 재취원 현황, 유치원 연령별 원아현황, 유치원 원아 수별 학급현황)과 학생현황(외국인 학생 및 졸업생 현황-유초중고, 의무취학대상 중 미취학 현황, 의무취학대상현황) 등이 있다(김영옥, 2016:193). 그러나 사회보장정보시스템과 보육통합정보시스템, 나이스(NEIS)는 관련 데이터를 일반에게 공개하지 않고 있는 실정이다.

공개되고 있는 공공데이터로는 유치원과 어린이집의 정보공시자료가 있다. 유치원알리미는 유치원 전반의 주요 정보를 객관적이고 투명하게 공개하는 제도로, 학급수, 유아수, 교사수 등 기본 현황과 유치원 규칙, 시설현황, 교육과정, 방과후과정, 수업일수, 원비 및 예결산, 급식, 식단표, 안전, 통학차량, 평가 등에 대한 다양한 데이터를 공개하고 있다.[4] 어린이집정보공개포털 또한 어린이집의 시설 설치 운영자, 보육료와 그 밖에 필요경비, 영유아의 건강, 영양 및 안전관리, 예결산, 어린이집 운영 관련 데이터를 공개하고 있다.[5]

이외에 전국단위 대규모 조사자료로서 한국아동패널, 전국보육실태조사, 한국복

2) 한국사회보장정보원(http://www.ssis.or.kr/lay1/S1T749C765/contents.do, 2021. 8.7. 인출)
3) 한국사회보장정보원(http://www.ssis.or.kr/lay1/S1T753C773/contents.do, 2021. 8.7. 인출)
4) 유치원알리미(https://e-childschoolinfo.moe.go.kr/introInfo/informList.do, 2021. 8.7. 인출)
5) 어린이집정보공개포털(http://info.childcare.go.kr/info/pnis/introduction/SummaryNursery.jsp, 2021. 8.7. 인출)

지패널이 있다. 한국아동패널은 출생부터 19세까지 20년간 장기적으로 추적·조사하는 데이터로 각 시기에 발생하는 아동의 성장·발달 특성, 양육실태와 요구, 육아지원기관의 기능과 효과, 지역사회와 육아정책의 영향, 정책 효과 등에 관한 종단적 자료를 국가 수준에서 수집하여 제공하고 있다.[6)]

전국보육실태조사는 3년단위 국가차원의 조사로 영유아 아동의 보육교육서비스 이용실태, 이용에 따른 만족도, 요구도 파악 및 보육시설 운영의 전반적 현황 파악을 통한 보육교육 정책 수립을 위한 수요자와 공급자에 대한 데이터이다. 세부적으로는 가구 특성, 아동부모 특성, 자녀양육, 보육교육정책, 어린이집·유치원관련 정책, 다문화가정특성, 생활시간조사, 돌봄, 사교육 등 가구와 아동 특성뿐만 아니라, 어린이집 설치·운영, 보육영유아 현황, 보육교직원 운영실태 및 요구, 보육프로그램 운영실태, 운영평가 및 관리 등 어린이집의 운영실태와 요구 등을 포괄하고 있다.[7)]

한국복지패널은 전 국민의 삶의 질 현황과 변화를 지역별·연령별·계층별로 파악하여 종단적 자료를 통한 계층별 소득변화 등을 장기간에 걸쳐 역동적으로 분석하기 위해 수집되는 데이터이다. 세부적으로 빈곤층, 차상위층 등의 규모 및 생활실태 변화를 동태적으로 파악하고, 인구집단별 생활실태 및 경제활동, 사회보험 · 퇴직(연)금 · 개인연금 가입, 주거, 생활비, 소득, 건강, 의료, 주거, 소득, 복지서비스 이용, 복지 요구 등을 포함하고 있는 데이터이다.[8)]

공공이외 육아관련 정보를 가지고 있는 빅데이터로, 인터넷카페, 블로그, 트위터, 페이스북, 인스타그램 SNS로 대변되는 온라인 자료, 신용카드 자료, 스마트폰 등 모바일 데이터, 포털사이트 뉴스기사 등의 다양한 데이터가 있다. 이러한 자료는 육아정책의 수요자 요구를 파악하는 다양한 경로가 될 수 있으며, 아직까지 육아정책에서 활용도는 낮지만, 온라인 자료, 뉴스기사 등 빅데이터들이 많이 있어 활용가능성이 높다.

---

6) 한국아동패널(https://panel.kicce.re.kr/panel/board/index.do?menu_idx=42&manage_idx=26, 2021. 8.7. 인출)
7) 통계청. 통계설명자료(https://meta.narastat.kr/metasvc/index.do?confmNo=331007&inputYear=2018, 2021. 8.7. 인출)
8) 통계청. 통계설명자료(https://meta.narastat.kr/metasvc/index.do?confmNo=331009&inputYear=2020, 2021. 8.7. 인출)

5. 빅데이터 활용 육아정책 선행연구

육아정책분야에서 빅데이터를 활용한 연구들은 육아정책연구소 등 연구기관과 학술논문을 중심으로 일부 수행되고 있다.

박창현 외(2020)는 통일을 대비하여 북한의 영유아 교육·보육 현황을 분석하고, 분석한 자료들에 기초하여 통일대비 영유아 교육·보육 정책의 방향을 수립하였다. 이때 온라인 상에 업로드된 북한 영유아 관련 비정형 데이터에 대한 추이분석, 워드 클라우드 분석, 관계망 분석, 시계열 분석 등을 통해 북한의 교육보육 관련 이슈를 도출하고자 하였다(박창현 외, 2020). 해당 연구는 빅데이터 분석 결과에 기초한 북한 영유아 교육보육 현장에 대한 이해를 바탕으로 통일대비 영유아 교육·보육 정책의 방향과 관련 정책 과제를 제안하였다.

양미선 외(2020)의 연구에서는 영유아 교육·보육 분야의 다양한 사회갈등을 진단하고 대응 방안을 모색하기 위한 방법 중 하나로 영유아 교육·보육 분야 이슈에 대한 빅데이터 분석을 실시했다. 갈등 양상과 쟁점에 대한 여론을 분석하기 위하여 41개 언론사 기사를 수집하였고, 수집된 데이터에 대하여 추이 분석, 감성 분석, 주요 이슈 분석을 실시하였다. 주요 이슈에 대한 심층분석은 무상교육, 누리과정 예산, 국공립 어린이집·유치원 확충, 맞춤형 보육, 유치원 3법, 사립유치원 파업으로 구분하여 분석하였으며, 이 결과를 근거로 영유아 교육·보육 분야 사회갈등을 관리 및 조정하고 선제적으로 대응하기 위한 방안을 제시하였다(양미선 외, 2020).

금융 빅데이터 플랫폼을 통하여 금융 빅데이터를 활용한 연구도 시도한 바 있다. 박진아, 김동훈, 김태우(2020)의 연구는 정부의 보편적 육아지원의 효과성을 분석하기 위한 '영유아 가구의 소비 실태 및 육아 물가 연구'의 3차 연도 연구로, 영유아 가구에서 소비하는 상품 및 서비스의 가격 변동과 영유아 가구의 체감물가 지표를 모니터링하는 것을 목적으로 했다. 이 연구는 비씨카드에서 구축한 금융 빅데이터 플랫폼을 통해 수집한 월별 온/오프라인 업종별 카드결제내역 데이터를 분석하여 영유아 가구의 업종별 카드결제액 추이와 코로나19 전후의 영유아 가구 소비 변화를 파악하였다(박진아 외, 2020). 이 연구에서 데이터 수집에 활용한 금융 빅데이터 플랫폼은 비씨카드 등 11개의 빅데이터 기업이 생산하는 데이터를 수집하여 유통하는 빅데이터 거래소로, 이용자로 하여금 금융 관련 빅데이터를 편리하게

이용할 수 있도록 지원하는 데이터 플랫폼을 말한다[9].

박창현(2019)의 연구에서는 출산 및 양육에 대한 소셜 빅데이터 분석을 통해 관련 이슈를 파악하고, 그 결과를 바탕으로 저출산 정책 해결에 대한 예측 모형 구축을 시도하였다. 구체적으로, 4차 산업혁명, 저출산, 출산, 양육, 유아교육, 보육, 육아 관련 비정형 데이터를 수집하여 추이 분석, 키워드 분석, 관계망 분석 등의 소셜 빅데이터 분석을 실시하였고, 빅데이터 분석을 통해 예측 모형을 도출하고 머신 러닝으로 주요 변수를 검증하고 추출하였다(박창현, 2019). 분석 결과에 기초하여 4차 산업혁명 시대 육아정책의 방향과 과제를 종합적으로 제안하였다. 이 연구는 빅데이터 분석을 단순히 현상을 파악하고 경향을 알아보는 수준을 넘어 예측모형을 구축하여 미래 육아정책의 방향성을 설정하는 수준까지 확장하여 활용했다는 것에 의의를 갖는다.

저출산 정책 및 출산율에 대한 빅데이터 분석 기반의 연구는 학술지 '통계연구'에서도 찾아볼 수 있다. 송태민(2016)은 소셜 빅데이터를 활용하여 저출산 정책의 수요를 예측하고 연관규칙을 파악하는 연구를 수행했다. 이 연구는 국내의 SNS, 온라인 뉴스 사이트 등 인터넷을 통해 수집된 소셜 빅데이터를 대상으로 하였으며, 총 183개의 온라인 채널을 통해 수집 가능한 버즈(buzz, 텍스트 기반의 웹문서)를 소셜 빅데이터로 보았다. 수집된 자료는 주제분석과 감성분석을 통해 정형화하였다. 이후 머신러닝의 연관분석과 의사결정나무 방법을 활용한 분석을 실시하였고, 그 결과에 기초하여 저출산 정책의 수요에 대한 예측 모형을 개발하였다.

2019년부터는 육아지원·보육 관련 학술지에서도 빅데이터를 활용한 연구들을 종종 찾아볼 수 있었다. 주세민, 옥성환, 황경태(2019)의 연구에서는 육아에 대한 부정적인 기사 분석을 통해 육아에 대한 공포 등 심리적 요인이 출산율에 미치는 영향을 파악하여 출산 및 육아 관련 불확실성 지표를 도출하고 이를 분석하였다. 구체적으로, 출산 및 육아에 대한 부정적 기사를 기반으로 키워드 선별 및 집계, 스케일 조정, 표준화의 과정을 거쳐 출산 및 육아 관련 불확실성 지표를 도출하였고, 이후 시계열 분석을 실시하였다. 이는 지표가 곧바로 출산율에 영향을 미칠 수 없다는 점에 근거한 것으로, 지표의 시차를 두면서 출산율과의 상관관계를 분석하

9) 금융 빅데이터 플랫폼 홈페이지, https://www.bigdata-finance.kr/platform/introPlatform.do (검색일: 2021. 8. 3.)

고, 상관관계가 높은 경우에 출산율과 회귀 계수를 도출하여 유의성 검정을 진행하였다(주세민 외, 2019). 이 연구는 빅데이터 분석 결과에 기초하여 산출한 지표를 바탕으로 시계열 분석을 실시하여 사회현상을 예측한 것을 특징으로 한다.

정옥경, 고온조, 김경근(2020)은 딥러닝 분석 중 하나인 계층적 주의 네트워크(hierarchical attention networks: HAN) 분석을 사용하여 온라인 육아 커뮤니티에 게시된 영유아 부모의 자녀양육 고민에 대한 연구를 수행했다. 연구진은 자녀양육 고민에 관한 장기간의 관찰 자료를 축적하고 있는 온라인 커뮤니티 중 활성화 수준이 높은 전국규모의 커뮤니티 두 곳을 선정하였고, 여러 게시판 중 육아 고민 관련 게시판의 게시글들을 수집하였다. 이후 연구목적에 어긋나는 일부 데이터를 제외하고 총 4,968,780개의 게시글을 분석대상으로 확정하였다. 학습 데이터 추출, 태깅, 데이터 전처리 등을 포함하는 분류기 학습과 게시글 분류 및 분석을 통해 0세부터 만 6세까지 각 연령별로 나타나는 여러 가지 고민 유형의 변화 추이를 파악하고 시사점을 도출하였다(정옥경 외, 2020).

최지은(2020)은 빅데이터를 활용하여 네트워크 분석을 실시함으로 유아 웰빙에 대한 사회적 인식을 탐색해보고자 하였다. 유아, 웰빙을 핵심 키워드로 검색하여 자료를 수집한 후 워드 추출과 정제작업, 네트워크 구축, 네트워크 분석 및 시각화의 과정으로 연구를 수행함으로 유아 웰빙에 대한 전반적인 사회적 인식은 어떠한지, 그 흐름 속에서 중심적인 위치에 있는 핵심적인 키워드들은 무엇이며, 이러한 키워드들 간의 관계 및 구조는 어떠한지에 대해 살펴보았다(최지은, 2020).

박하연, 이재림, 이강이(2021)의 연구에서는 온라인 카페의 육아 관련 게시글에서 조부모가 이야기되는 양상을 토픽모델링(topic modeling)을 통해 파악하고자 하였다. 육아 관련 소셜 빅데이터에서 조부모 관련 단어 67개를 포함하고 있는 온라인 카페 게시글 총 41,699건을 추출하여 분석대상으로 삼아 TF 및 TF-IDF 분석, 토픽모델링을 실시하였다. 이를 통해 조부모와 손자녀, 생애초기 양육, 육아와 육아지원, 결혼생활과 성역할 등 육아와 조부모에 관련된 다양한 시사점을 도출하였다.

육아분야 빅데이터 활용 선행연구들에서 사용한 빅데이터 자료와 주요분석 내용들을 정리하면 다음과 같다.

〈표 Ⅱ-5-1〉 육아분야 빅데이터 활용자료 및 분석내용

빅데이터 활용자료	주요 분석 내용
SNS, 카페, 블로그 등 온라인 내 게시글 및 댓글	- 북한 영유아 교육·보육 이슈 - 4차 산업혁명 시대의 출산 및 양육 관련 이슈 - 저출산 정책에 대한 인식 - 영유아 부모의 자녀양육 고민 - 유아 웰빙에 대한 사회적 인식 - 조부모가 이야기되는 양상
인터넷 신문기사	- 북한 영유아 교육·보육 이슈 - 교육·보육 분야 사회갈등 이슈 - 4차 산업혁명 시대의 출산 및 양육 관련 이슈 - 저출산 정책에 대한 인식 - 육아에 대한 부정적 인식 - 유아 웰빙에 대한 사회적 인식
신용카드 사용내역	- 유아 가구의 업종별 카드결제액 추이 및 코로나19 전후의 영유아 가구 소비 변화
학술자료 (학위 논문, 학술지 논문 등)	- 유아 웰빙에 대한 사회적 인식

주: 본 연구에서 검토한 선행연구에 기반하여 작성하였음.

한편, 유아교육분야의 빅데이터 연구 활성화를 위한 기초연구 성격의 연구들도 있다. 먼저, 김영옥(2016)은 연구를 통해 유아교육연구에서 빅데이터를 활용하고 적용하기 위한 한국유아교육학의 과제를 탐색하고자 하였다. 이를 위해 문헌분석을 실시하여 빅데이터의 활용과 전망, 교육분야에서의 현황, 유아교육분야의 현황을 파악하고 시사점을 도출한 후 한국유아교육학의 과제를 종합적으로 제시하였다. 빅데이터 관련 연구기반이 미흡한 유아교육학 분야에 남겨진 과제에 대한 연구자의 제안은 총 7가지로, 교육환경의 변화를 인지하고 수용하는 역량, 빅데이터의 음과 양에 대한 이해, 유아교육관련 데이터 개방과 표준화를 위한 노력, 데이터의 기록과 누적 및 활용, 전문가와 학회의 인프라구축 및 지원, 타 분야와의 통합시도 및 협력연구, 그리고 윤리와 가치에 대한 소통과 공유 등이다(김영옥, 2016).

유사한 맥락에서 이연승(2017)은 빅데이터의 개념, 빅데이터와 교육정책, 유아교육에서의 빅데이터 활용에 대한 문헌분석을 통해 유아교육분야에서의 빅데이터 활용 및 적용을 위한 방향을 모색하고자 하였다. 그 결과, 빅데이터 활용을 통한 유아교육정책을 유아 통합 데이터 시스템 구축, 빅데이터를 활용한 정보제공, 오피니언 마이닝(Opinion Mining) 기법을 활용한 교육과정 개발, 유아교사를 위한 데이터 리터러시(data literacy) 교육 등의 4가지로 제안하였다(이연승, 2017).

III

공공 빅데이터 활용을 통한 영유아 안전·환경 분석

01 육아정책 빅데이터 플랫폼

02 빅데이터 분석사례: 영유아 안전·환경 공공 빅데이터

03 소결

Ⅲ. 공공 빅데이터 활용을 통한 영유아 안전·환경 분석

1. 육아정책 빅데이터 플랫폼

육아정책연구와 연관된 빅데이터 활용방안을 모색하기 위해 먼저 대표적인 빅데이터 플랫폼에 대해서 살펴보았다.

가. 빅데이터 플랫폼과 데이터지도[10)]

빅데이터 플랫폼은 데이터를 다양한 경로로부터 수집·통합하고, 다양한 빅데이터 이용자 간의 공유와 활용의 중용한 기능을 담당한다(Pappas et al. 2018). 우리나라의 경우 과학기술정보통신부와 한국지능정보사회진흥원이 분야별 빅데이터 플랫폼 및 센터를 구축하여 데이터의 수집·유통을 활성화하고 있고, 특히 통합 데이터 지도를 통해 한 곳에서 누구나 양질의 데이터에 쉽게 검색·접근·활용할 수 있도록 지원하고 있다.[11)]

통합 데이터 지도 구축 사업은 플랫폼 16개소(문화, 통신, 유통소비, 헬스케어, 교통, 환경, 금융, 중소기업, 지역경제, 산림, 소방안전, 스마트치안, 해양수산, 농식품, 라이프로그, 디지털 산업혁신) 및 센터 150여개소를 모두 연계하는 사업이다[12)]. 따라서 서로 다른 플랫폼 간, 센터 간 데이터 연계를 통해 보다 전문적이고 경제적 가치가 높고 활용도가 높은 데이터를 만들 수 있고, 통합 데이터 지도에서는 이 데이터들을 통합적으로 모아서 사용자가 쉽게 접근할 수 있도록 지원을 하고 있고, 데이터 분야별, 지역별, 키워드별, 주제별로 상품수, 데이터별 조회수, 다운로드수, 인기데이터 정보를 제공하고 있다[13)].

---

10) 통합데이터지도(bigdata-map.kr) 홈페이지의 내용을 인용 또는 재구성하여 작성함.
11) https://www.bigdata-map.kr/intro/intro(2021.9.14. 인출)
12) https://www.bigdata-map.kr/board/faq(2021.9.14. 인출)
13) https://www.bigdata-map.kr/board/faq(2021.9.14. 인출)

〈표 Ⅲ-1-1〉 빅데이터 관련 플랫폼

분야	플랫폼 운영 주관	참여 센터
문화	한국문화정보원	국립중앙도서관 등 7개
통신	케이티	BC카드, 한국인터넷진흥원 등 15개
유통	매일방송	삼성카드, SK텔레콤 등 21개
헬스케어	국립암센터	삼성서울병원, 연세암병원 등 11개
교통	한국교통연구원	한국도로공사, 한국철도공사 등 7개
환경	한국수자원공사	한국환경정책평가연구원, 한국과학기술원 등 12개
금융	비씨카드	KT 등12개
중소기업	더존비즈온	NICE평가정보 등 17개
지역경제	경기도청	경기신용보증재단, 국토연구원 등 11개
산림	한국임업진흥원	한국산림복지진흥원 등 13개
소방안전	소방청	한국소방산업기술원, 한국화재보험협회 등 11개
스마트치안	경찰대학 치안정책연구소	한국토지주택공사, 연합뉴스 등 11개
해양수산	한국해양수산개발원	해양교통안전공단, 한국해양대학교 등 12개
농식품	한국농수산식품유통공사	한국무역통계진흥원 등 10개
라이프로그	원주연세의료원	고려대의료원, LG유플러스 등 13개
디지털산업혁신	한구산업기술시험원	한국기업데이터, 한국생산성본부 등 15개

출처: 통합데이터지도(https://www.bigdata-map.kr/intro)

통합데이터지도 내 데이터현황을 살펴보면 다음과 같다. 11개 분류기준으로는 총 21,337개의 데이터가 있고, 16개 분야기준으로는 17,349개의 빅데이터가 제공되고 있다.

〈표 Ⅲ-1-2〉 통합데이터 현황

단위: 건

번호	분류별 데이터 현황			분야별 데이터 현황			
	분류	데이터수	총조회수	분야	데이터수	총조회수	총다운로드수
1	행정법률	5,286	207,562	유통	9,679	744,401	16,928
2	제조소비	3,352	311,832	헬스케어	1,135	551	5,386
3	경제금융	2,859	369,382	문화	1,066	3,462	38,247
4	교통통신	2,396	230,887	환경	781	279,406	32,937
5	건강의료	2,138	110,632	소방안전	734	47,184	2,087
6	문화체육	1,457	51,829	지역경제	712	5,781	8,199
7	환경자원	1,100	281,352	산림	625	160,729	12,699
8	교육과학	1,089	25,990	통신	501	13,619	35,249
9	건설에너지	707	38,813	중소기업	484	81,106	4,838
10	농수산식품	631	46,093	디지털산업혁신	316	21,852	1,901
11	재난안전	322	36,990	교통	291	2,286	23,473

번호	분류별 데이터 현황			분야별 데이터 현황			
	분류	데이터수	총조회수	분야	데이터수	총조회수	총다운로드수
12				금융	272	321,889	27,365
13				라이프로그	218	15,906	3,593
14				농식품	206	200	3,516
15				해양수산	183	21,818	5,822
16				스마트치안	146	975	6,945
-	계	21,337	1,711,362	계	17,349	1,721,165	229,185

출처: 통합데이터지도(https://www.bigdata-map.kr/dashboard, 2021. 9.17 인출)

나. 육아정책 관련 빅데이터 현황

통합데이터지도 내 육아관련 데이터 현황을 살펴보면 다음과 같다. 먼저 육아정책 관련 주요키워드로는 임신, 출산, 육아, 보육, 돌봄, 영아, 유아, 어린이집, 유치원 등 9개를 설정하여 검색하였다. 이외에도 코로나, 학대, 아동, 교육, 결혼 등 추가 키워드 등으로 검색하였으나 데이터 중복이나 관련성이 적어 제외되었다.

〈표 Ⅲ-1-3〉 빅데이터 플랫폼(통합데이터지도) 내 육아관련 데이터 현황

단위: 건

분야	주요 키워드									계
	임신	돌봄	보육	어린이집	영아	유아	유치원	육아	출산	
교통				1			1			2
금융				1			2	1	1	5
데이터스토어	2	5	5	7		10	3	7	2	41
라이프로그	2			1						3
문화		2				2		3		7
산림						5				5
유통	3					9		55	56	123
중소기업							2	3		5
지역경제		3						4		7
통신						2	1			3
헬스케어	13									13
환경			3	3	3	4	3			16
AI-HUB		7				2				9
쿠콘	1								1	2
계	21	17	8	13	3	34	12	73	60	241

주: 1) '통합데이터지도-데이터검색'에서 해당 키워드 검색을 통해 조사한 자료임
2) 주요 키워드 검색결과에서 중복검색된 데이터를 제외하면 153개임.

2. 빅데이터 분석사례: 영유아 안전·환경 공공 빅데이터

본 절에서는 육아정책 관련 공공 빅데이터의 활용사례를 분석하고자 한다. 특히 공공 데이터 중 유치원과 어린이집 등 영유아 보육·교육기관에 대한 대기질, 위험시설, 수질 등과 관련한 안전·환경관련 데이터를 활용하여 빅데이터를 분석해보고자 하였다. 이를 위해 환경 빅데이터 플랫폼(envbigdata)에서 제공하는 유치원과 어린이집의 주소와 인근 대기질 지수, 미세먼지 지수, 위험시설, 수질지수, 수질오염원 인근 거주인구 정보 등의 자료를 활용하였다. 환경 빅데이터 플랫폼은 한국수자원 공사가 플랫폼 주관기관으로 하여 국민 생활과 밀접한 관련 있는 환경 정보에 대한 데이터를 제공해 주는 플랫폼이다[14]. 환경빅데이터 플랫폼을 통해 상하수도, 수자원, 기상/기후, 대기, 생태, 지질/지형, 사회경제, 소셜미디어, 폐기물, 화학물질, 환경법규, 신재생에너지 등 안전·환경 이슈와 관련된 다양한 데이터를 검색할 수 있고 관련 데이터를 제공받을 수 있다.

가. 대기질 분석 결과

유치원은 학교보건법, 어린이집은 실내공기질 관리법에 따라 원내 공기질을 알맞게 유지 관리해야한다. 유치원·어린이집 인근 대기 질이 나빠졌을 경우 원내 공기질도 동반하여 악화될 가능성이 높다. 본 분석에서는 지역별 대기 질 및 미세먼지 수치를 살펴보고 대기 질과 미세먼지 수치가 좋지 않은 장소에 위치해 있는 유치원과 어린이집을 추출하고자 하였다.

유치원과 어린이집 인근 대기질 및 미세먼지 데이터는 2020년 9월 기준 전국 유치원 8,672개소와 어린이집 54,568개소를 분석 대상으로 하였다. 대기질 빅데이터에서 제시한 지수는 미세먼지 수치를 나타내는 PM10(㎍/㎥)과 통합대기환경지수(CAI, Comprehensive Air-quality Index)였다. 통합대기환경 지수는 6개의 대기오염물질(아황산가스, 일산화탄소, 오존, 이산화질소, 미세먼지, 초미세먼지) 별로 통합대기환경지수 점수를 선정하여 가장 높은 점수를 통합하여 지수 값으로 사용하여 산출한다[15]. 미세먼지 PM10(㎍/㎥)은 50초과, 통합대기환경지수는

14) https://www.bigdata-environment.kr/user/content.do?sq=60(2021.9.16. 인출)
15) https://www.airkorea.or.kr/web/khaiInfo?pMENU_NO=129(2021.9.16. 인출)

100초과 되었을 때 공기질이 나쁜 것으로 판정한다[16][17]. 본 연구에서 분석한 미세먼지 및 대기질 지수는 상대적으로 대기질이 양호한 2020년 9월에 측정된 자료값이기 때문에 미세먼지 PM10(㎍/㎥) 50이 초과된 사례가 없어, 인근 대기질이 보통에 해당하는 30이 초과되고 통합대기 질이 나쁨에 해당하는 100 초과가 된 어린이집을 추출하였다.

유치원·어린이집 인근 미세먼지 수치, 통합대기질에 대한 평균값을 살펴보면 미세먼지의 경우는 경기가 22.4로 가장 높았으며 다음으로 서울 21.4, 대구 20.6, 충남 21.0, 경남과 인천이 19.8 순이었다. 통합대기질 지수 역시 경기가 110.7 가장 높았으며 다음으로 서울 105.8, 충남 103.9, 인천 98.1, 경남 97.9 순이었다.

〈표 Ⅲ-2-1〉 시도별 대기질

단위: ㎍/㎥, 개원

지역	미세먼지(PM10)			통합대기질(CAI)			유치원·어린이집		
	평균	최솟값	최댓값	평균	최솟값	최댓값	유치원	어린이집	합계
합계	20.2	9.9	34.1	100.0	49.0	168.8	8,672	54,568	63,240
강원	18.0	12.6	21.5	89.2	62.6	106.2	367	1,535	1902
경기	22.4	9.9	34.1	110.7	49.0	168.8	2136	16,450	18,586
경남	19.8	13.9	24.6	97.9	68.8	121.9	682	5,162	5,844
경북	19.1	10.6	26.8	94.5	52.4	132.6	703	2,792	3,495
광주	15.8	12.5	23.3	78.3	62.0	115.6	313	1,412	1,725
대구	20.6	17.4	23.3	101.9	86.1	115.2	347	1,874	2,221
대전	14.7	12.5	19.0	72.6	62.0	94.1	260	1,924	2,184
부산	19.2	15.2	23.1	95.1	75.0	114.3	404	2,300	2,704
서울	21.4	16.9	34.1	105.8	83.9	168.8	832	8,687	9,519
세종	16.9	15.0	20.4	83.9	74.4	101.0	60	4,22	482
울산	17.7	15.7	25.4	87.9	77.6	125.8	195	1,129	1,324
인천	19.8	16.2	25.0	98.1	80.3	123.9	375	2,682	3,057
전남	17.7	13.4	23.8	87.7	66.3	118.0	536	1,485	2,021
전북	17.3	12.7	21.2	85.4	62.9	105.0	517	1,921	2,438
제주	19.3	16.0	25.8	95.8	79.0	127.6	122	723	845
충남	21.0	15.1	26.2	103.9	74.6	129.9	498	2,605	3,103
충북	18.7	14.4	22.2	92.7	71.4	110.1	325	1,465	1,790

16) https://www.airkorea.or.kr/web/khaiInfo?pMENU_NO=129(2021.9.16. 인출)
17) https://www.airkorea.or.kr/web/contents/contentView/?pMENU_NO=132&cntnts_no=6(2021.9.16. 인출)

대기질이 가장 좋지 않은 시·군·구를 분석한 결과 부천시가 미세먼지가 평균 29.6으로 가장 좋지 않았으며 다음으로 고양시 26.5, 과천시 24.8, 용산구 24.1, 수원시 23.9 순이었다. 통합대기질 역시 부천이 146.7로 가장 좋지 않았으며 다음으로 고양시 131.0, 과천시 122.7, 용산구 119.3, 수원시 118.4 순이었다.

〈표 Ⅲ-2-2〉 대기질 하위 20개 시군구

단위: ㎍/㎥

순위	시/도	시/군/구	미세먼지			통합대기질		
			평균	최솟값	최댓값	평균	최솟값	최댓값
1	경기	부천시	29.6	17.1	34.1	146.7	84.5	168.8
2	경기	고양시	26.5	18.9	29.3	131.0	93.3	145.0
3	경기	과천시	24.8	18.9	26.4	122.7	93.3	130.5
4	서울	용산구	24.1	18.9	25.6	119.3	93.3	126.8
5	경기	수원시	23.9	18.5	27.2	118.4	91.4	134.8
6	서울	동작구	23.7	18.9	25.6	117.1	93.3	126.8
7	서울	관악구	23.5	18.9	26.4	116.5	93.3	130.5
8	서울	구로구	23.4	18.9	34.1	116.0	93.3	168.8
9	경기	평택시	23.4	18.8	30.2	115.8	93.1	149.7
10	경기	안양시	22.9	12.3	25.0	113.6	60.7	123.8
11	서울	영등포구	22.9	18.9	25.6	113.5	93.3	126.8
12	경기	의왕시	22.9	18.9	25.0	113.3	93.3	123.8
13	경기	군포시	22.9	18.9	25.0	113.3	93.3	123.8
14	서울	강북구	22.8	18.9	23.1	112.8	93.3	114.2
15	대구	서구	22.7	18.9	22.7	112.2	88.9	145.0
16	경기	파주시	22.7	18.0	29.3	112.2	93.3	112.3
17	경기	화성시	22.7	16.9	27.2	112.1	83.9	134.7
18	경기	김포시	22.5	16.7	25.8	111.4	82.6	127.7
19	인천	남동구	22.5	16.2	23.7	111.3	80.3	117.4
20	대구	중구	22.4	18.9	22.7	111.0	93.3	112.3

미세먼지 수치인 PM10이 30이 넘는 전국 어린이집을 산출한 결과 전국에 어린이집 413개소, 유치원 29개소 총 442개소가 추출되었다. 부천이 유치원 29개소 어린이집 413개소로 총 442개소, 평택이 유치원 5개소 어린이집 27개소로 총 32개소, 구로구가 유치원 1개소 어린이집 11개소 총 12개소로 추출되었다.

〈표 Ⅲ-2-3〉 미세먼지 PM10 30초과 어린이집 개소수

단위: 개원

지역		유치원	어린이집	합계
합계		29	413	442
경기	부천시	23	375	398
	평택시	5	27	32
서울	구로구	1	11	12

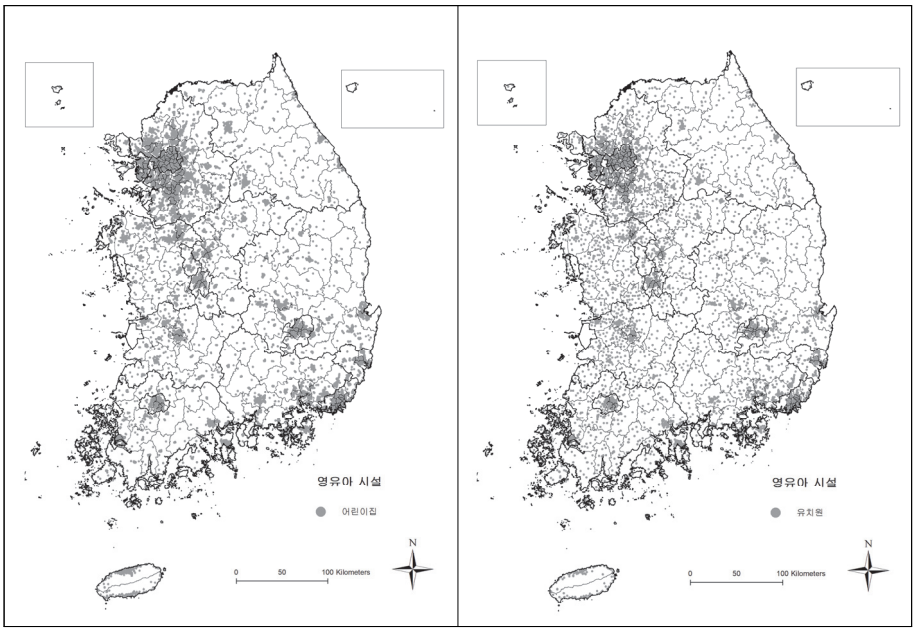
통합대기질 지수 100을 초과한 유치원·어린이집 개소수 지역별 순위를 산출한 결과 수원이 총 1,733개소로 가장 많았으며 다음으로 고양시 1,548개소, 창원 1,048 개소, 화성 920 개소, 안산 910 개소, 성남 855 개소 순이었다.

〈표 Ⅲ-2-4〉 통합대기질 100 초과 유치원·어린이집 개소수

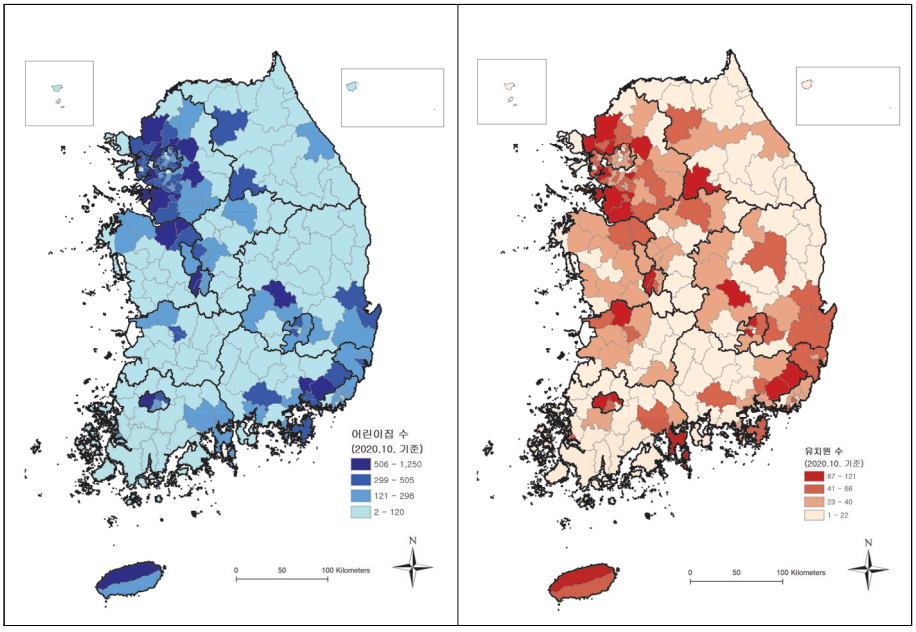
단위: 개원

순위	지역		유치원	어린이집	합계
1	경기	수원시	190	1543	1733
2	경기	고양시	165	1383	1548
3	경남	창원시	93	955	1048
4	경기	화성시	118	802	920
5	경기	안산시	89	821	910
6	경기	성남시	106	749	855
7	경기	용인시	81	678	759
8	경기	안양시	79	669	748
9	경기	부천시	45	679	724
10	충남	천안시	65	599	664
11	경남	김해시	44	616	660
12	경기	평택시	105	543	648
13	충남	아산시	55	590	645
14	서울	강서구	58	569	627
15	경기	의정부시	46	511	557
16	경기	파주시	61	488	549
17	서울	송파구	39	485	524
18	서울	양천구	42	450	492
19	인천	남동구	76	412	488
20	서울	구로구	35	439	474

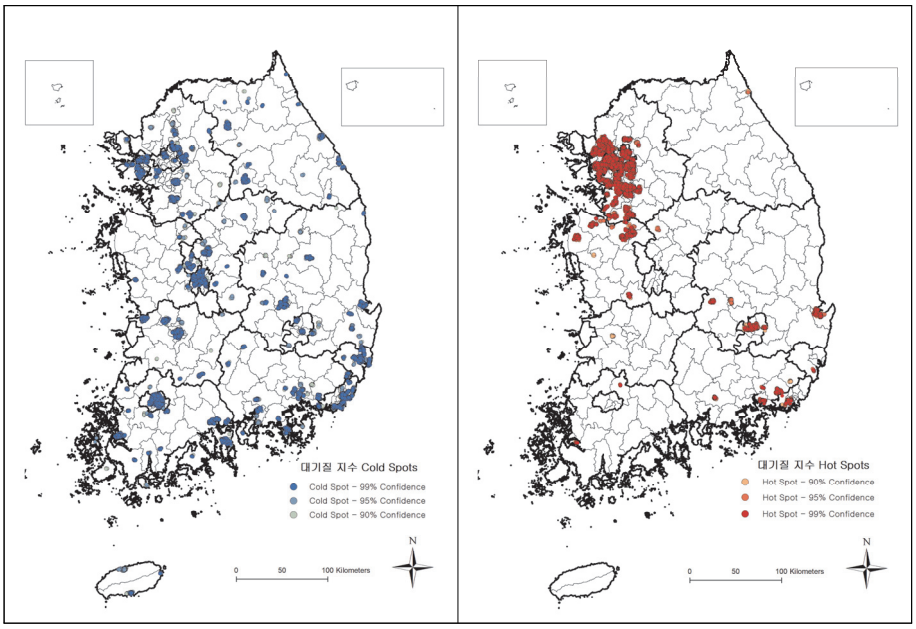
[그림 Ⅲ-2-1] 어린이집 및 유치원 위치



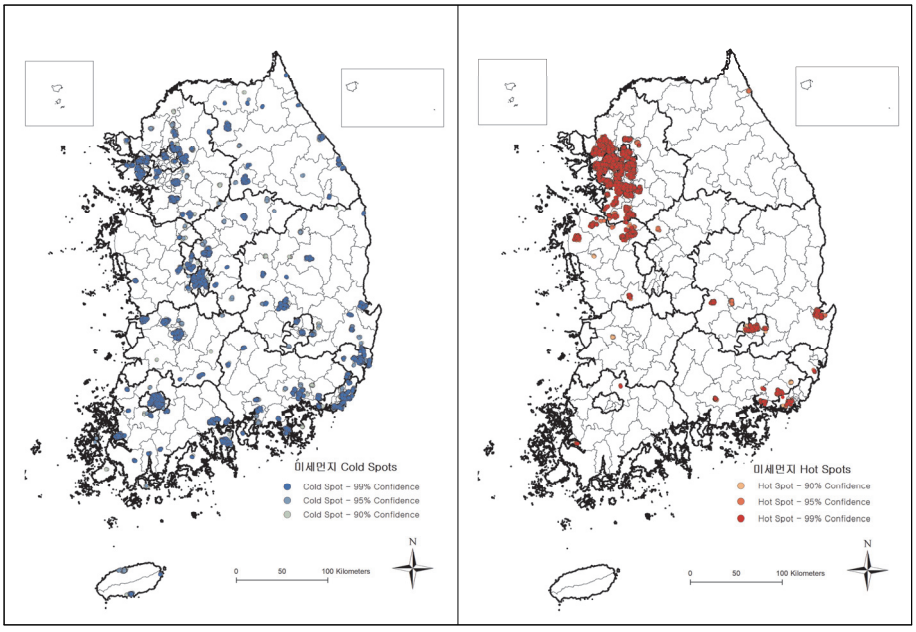
[그림 Ⅲ-2-2] 시군구별 어린이집 및 유치원 위치



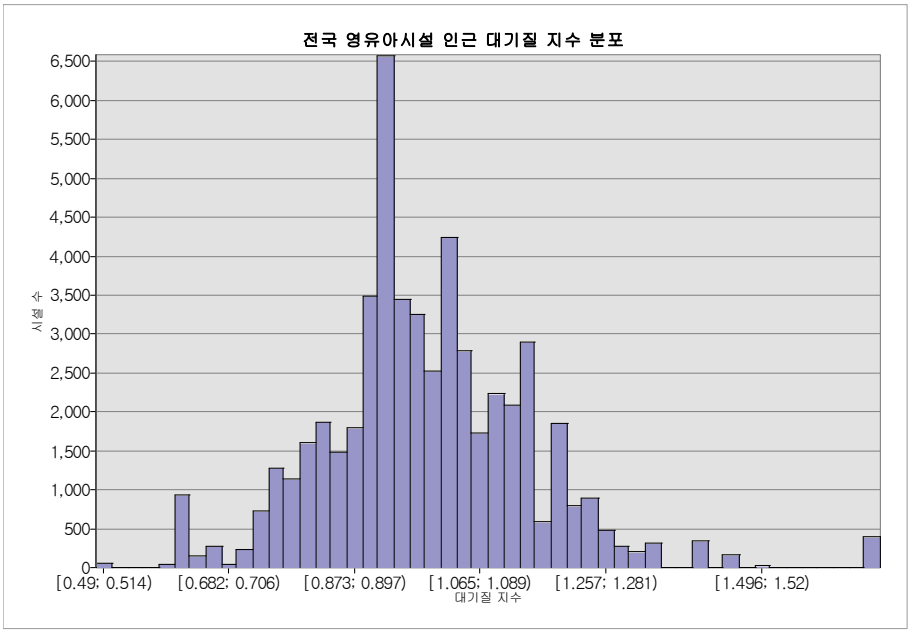
[그림 Ⅲ-2-3] 대기질 지수: Cold Spots vs. Hot Spots



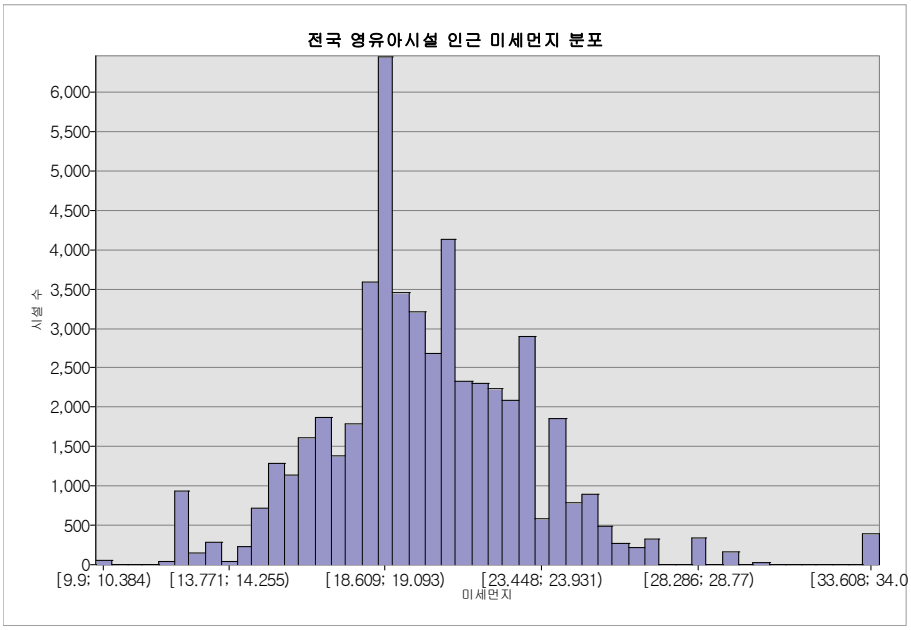
[그림 Ⅲ-2-4] 미세먼지: Cold Spots vs. Hot Spots



[그림 Ⅲ-2-5] 전국 영유아시설 인근 대기질 지수 분포



[그림 Ⅲ-2-6] 전국 영유아시설 인근 미세먼지 분포



나. 유치원·어린이집 인근 위험시설

전국 어린이집과 유치원의 반경 500m 내의 위험시설 정보에 대한 환경 빅데이터를 살펴보면 다음과 같다. 여기에서의 위험시설은 주유소, 가스충전소, 도료염료, 석탄광업, 연탄판매, 원유 채굴 및 정제처리업, 합성고무 제조업, 화학 살균·살충제 및 농업용 약제 제조업 등 화학물질이나 위험물 등 산업환경과 관련된 시설들이다.

유치원과 어린이집 인근 위험시설 관련 정보는 2020년 11월 기준 자료로, 전국적으로 87,741개에 달했다. 전국적으로 유치원은 평균 2.3개, 어린이집은 평균 2.4개의 위험시설이 반경 500미터 내에 있는 것으로 나타났고, 주유소 등을 포함한 위험시설이 최대 20개인 어린이집도 있었다.

시도별로는 대구와 대전의 위험시설 평균이 높고 세종이 낮았으며, 서울과 경기의 경우는 평균과 다소 비슷하였다. 전체 위험시설수 대비 시도별 비율은 경기 27.6%, 서울 18.5%, 경남 7.0% 순으로 높고, 세종 0.1%, 제주 1.1%, 울산 1.6% 순으로 낮았다.

〈표 Ⅲ-2-5〉 유치원·어린이집 인근 위험시설

단위: 개, %

지역	유치원			어린이집			인근 위험시설수		
	평균	최솟값	최댓값	평균	최솟값	최댓값	유치원	어린이집	합계
합계	2.3	1	19	2.4	1	20	10,360	77,381	87,741
서울	2.2	1	9	2.3	1	10	1,215	15,058	16,273
부산	2.7	1	13	2.8	1	13	787	4,443	5,230
대구	2.8	1	19	2.9	1	20	717	3,805	4,522
인천	2.3	1	10	2.6	1	12	469	4,195	4,664
광주	2.2	1	10	2.2	1	10	437	1,889	2,326
대전	2.8	1	9	2.9	1	12	457	3,465	3,922
울산	2.3	1	7	2.2	1	12	212	1,207	1,419
세종	1.6	1	4	1.6	1	4	22	107	129
경기	2.3	1	13	2.5	1	14	2,559	21,647	24,206
강원	2.1	1	9	2.4	1	13	322	2,257	2,579
충북	2.2	1	8	2.5	1	12	386	2,252	2,638
충남	1.9	1	10	2.5	1	18	357	2,895	3,252
전북	2.1	1	13	2.2	1	17	476	2,579	3,055
전남	2.2	1	8	2.2	1	8	473	1,706	2,179
경북	2.2	1	9	2.4	1	10	763	3,432	4,195
경남	2.0	1	11	2.1	1	12	627	5,546	6,173
제주	1.7	1	5	2.2	1	8	73	854	927
-*	1.0	1	1	1.0	1	1	8	44	52

* 데이터에 지역정보가 없는 경우임.

주: 1) 위험시설에는 LPG판매, 가스설비, 가스집단공급업, 가정용연료소매, 기타 기초 무기화학 물질 제조업, 도료염

료, 라이터 연소물 및 흡연용품 제조업, 복합비료 및 기타 화학비료 제조업, 산업용 가스 제조업, 석탄 광업, 석탄화학계 화합물 및 기타 기초 유기화학 물질제조업, 연료용 가스 제조 및 배관공급업, 연탄 및 기타 석탄 가공품 제조업, 연탄판매, 원유 및 천연가스 채굴업, 원유 정제처리업, 유기질 비료 및 상토 제조업, 윤활유 및 그리스 제조업, 인쇄 잉크 및 회화용 물감 제조업, 일반용 도료 및 관련제품 제조업, 주유소, 차량가스충전소, 차량연료소매-종합, 천연수지 및 나무 화학 물질 제조업, 합성고무 제조업, 합성수지 및 기타 플라스틱 물질 제조업, 화학 살균·살충제 및 농업용 약제 제조업이 포함된

2) 전체 위험시설 87,741개에서 주유소가 51,156개로 58.3%로 가장 많았고, 가정용연료소매 5657(6.4%), 차량가스충전소 5202개(5.9%)순으로 많았음.

어린이집과 유치원의 반경 500m 내의 위험시설수가 많은 상위 40개 시·군·구를 살펴보면 다음과 같다. 먼저 서울 중구가 2,427개로 가장 많았고, 경남 김해시 1,716개, 경기 부천시 1,486개, 경기 안양시 동안구 1,358개 순으로 나타났다. 위험시설수 상위 40개 시·군·구의 시도별 분포는 경기가 15개, 서울 7개, 인천 4개, 대구 3개, 대전과 경남이 각각 2개, 울산, 강원, 충남, 전북, 경북, 제주가 각각 1개로 나타났다.

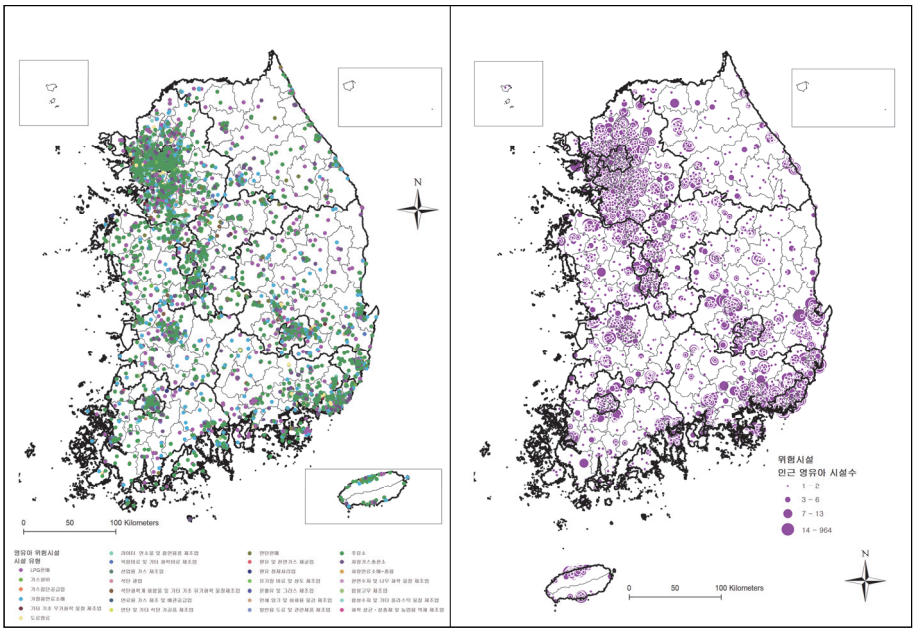
〈표 Ⅲ-2-6〉 유치원·어린이집 인근 500m 내 위험시설수 상위 40개 시군구

단위: 개

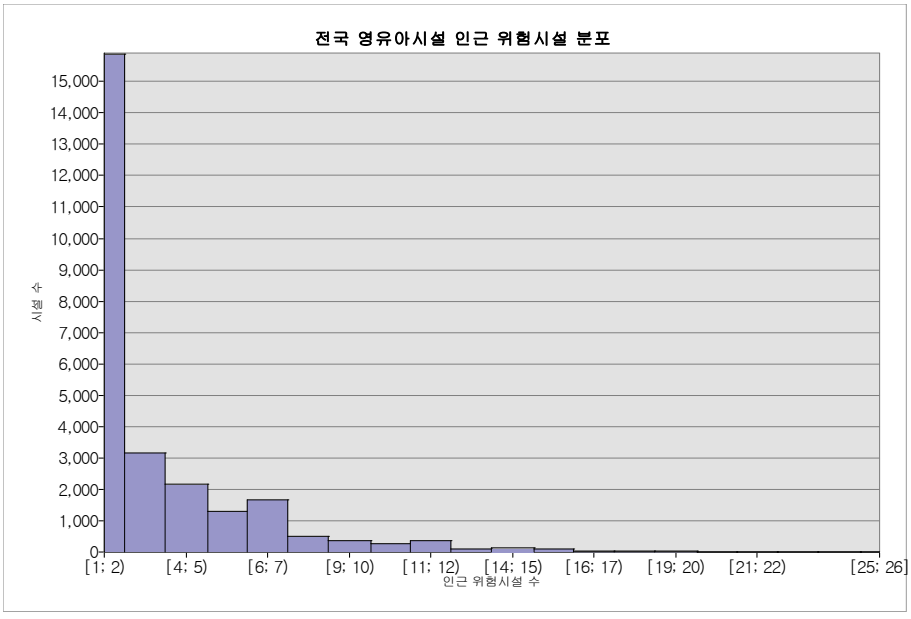
순위	시도	시군구	위험시설수	순위	시도	시군구	위험시설수
1	서울	중구	2,427	21	경기	용인시 기흥구	866
2	경남	김해시	1,716	22	서울	강서구	862
3	경기	부천시	1,486	23	경기	고양시 덕양구	860
4	경기	안양시 동안구	1,358	24	경기	군포시	823
5	대전	서구	1,355	25	경기	수원시 영통구	822
6	경기	수원시 권선구	1,243	26	제주	제주시	817
7	대구	달서구	1,210	27	대전	중구	806
8	서울	종로구	1,119	28	서울	양천구	789
9	경기	고양시	1,118	29	울산	남구	781
10	경남	창원시	1,099	30	서울	서초구	763
11	경기	의정부	1,060	31	광주	북구	759
12	경기	화성시	1,043	32	경기	안양시 만안구	747
13	충남	천안시 서북구	1,036	33	인천	계양구	746
14	서울	영등포	1,012	34	서울	구로구	745
15	인천	부평구	1,011	35	경북	구미시	744
16	경기	수원시 장안구	952	36	경기	안산시 단원구	742
17	강원	원주시	917	37	경기	평택시	742
18	전북	전주시 덕진구	904	38	대구	수성구	737
19	경기	수원시 팔달구	899	39	인천	서구	734
20	인천	남구	874	40	대구	북구	727

* 특별광역시 이외에는 시 단위로만 분류해서 보면, 경기 수원시, 서울 중구, 경기 안양시, 경남 창원시, 경기 고양시, 충북 청주시, 충남 천안시, 경기 성남시, 경남 김해시, 경기 용인시 순으로 상위 10위를 차지하고 있음.

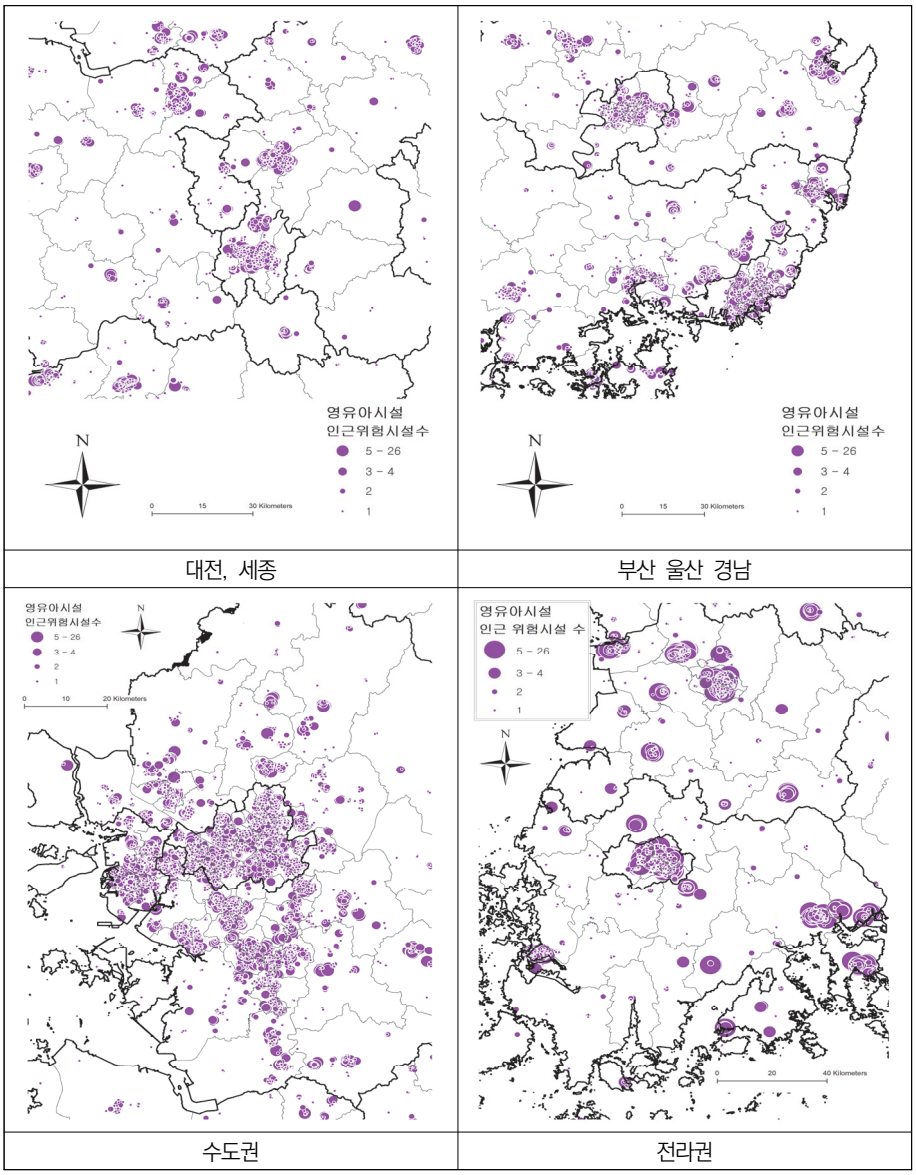
[그림 Ⅲ-2-7] 영유아 위험시설 및 위험시설 인근 영유아 시설



[그림 Ⅲ-2-8] 영유아 시설 인근 위험시설 분포



[그림 Ⅲ-2-9] 영유아 시설 인근 위험시설 지역별 분포



다. 유치원·어린이집 수질지수

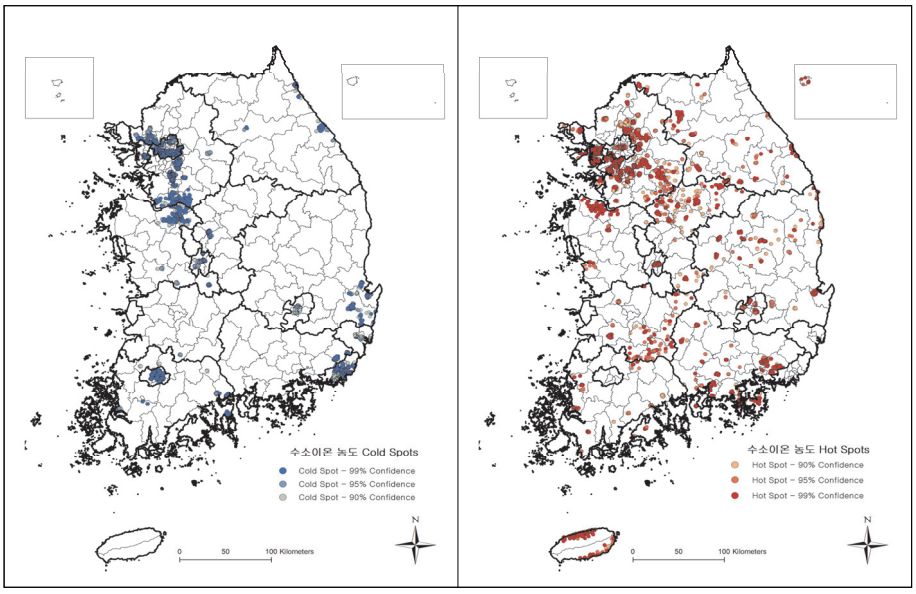
여기에서는 2020년 12월 기준 전국 유치원·어린이집 수질지수를 수소이온농도, 잔류염소, 탁도 수치를 살펴보았다.

수소이온농도지수(PH)는 수중의 산성 및 알칼리성의 세기를 나타내는 값으로 하수는 대개의 경우 7.0~7.2의 중성 또는 약알카리성이 일반적이나 공장폐수의 유입이 있으면 값이 크게 변동할 때가 있다. 잔류염소는 살균 때문에 물에 가한 염소가 잔류해서 유효하게 작용하고 있는 것을 말한다. 즉, 염소수 일부가 물과 반응하여 잔류 염소(HCIO 등)를 생성, 이것이 분해되어 내놓는 발생기 산소의 강한 산화 작용으로, 수도수의 소독제 등으로 쓰이고 있다.[18] 탁도는 수중의 부유물질 등에 의해 물이 혼탁한 정도, 물의 흐림정도를 나타낸다.

수도법 시행규칙 제22조3의 3항에 따르면, 저수조 청소후 수질 기준으로 잔류염소는 리터당 0.1밀리그램 이상 4.0밀리그램 이하, 수소이온농도(pH)는 5.8 이상 8.5 이하, 탁도는 0.5NTU(네펠로메트릭 탁도 단위, Nephelometric Turbidity Unit) 이하로 제시하고 있다. 탁도의 경우는 세계보건기구(WHO)가 권장하는 음용수의 탁도 기준인 5NUT보다 10배나 엄격한 기준이다.

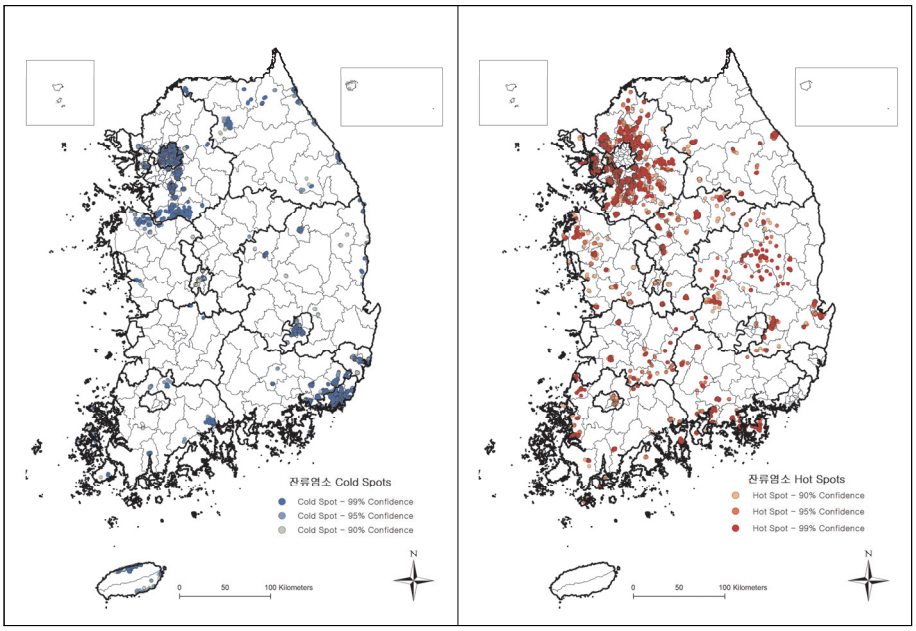
지역에 따라 유치원과 어린이집의 수소이온농도, 잔류염소, 탁도에 차이를 보이고 있으나, 법적 기준은 모두 충족하고 있는 실정이다.

[그림 Ⅲ-2-10] 수소이온농도: Cold Spots vs. Hot Spots

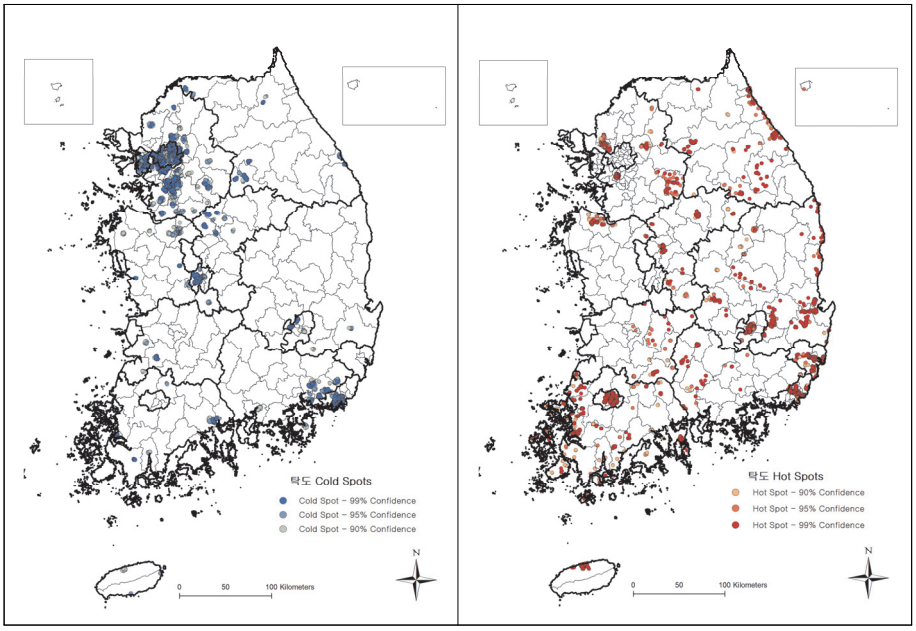


18) https://terms.naver.com/entry.naver?docId=1588411&cid=50314&categoryId=50314

[그림 Ⅲ-2-11] 잔류염소: Cold Spots vs. Hot Spots



[그림 Ⅲ-2-12] 탁도: Cold Spots vs. Hot Spots



라. 수질오염원 인근 아동수 현황

수질오염원 인근의 0-4세와 5-9세 아동수 현황을 살펴보았다. 수질오염원으로는 공업단지, 생활하수, 세차장, 쓰레기매립장, 유류취급소, 축산시설, 주유소, 폐기물처리, 하수종말처리장, 공동묘지, 광산 등이 포함되었다.

〈표 Ⅲ-2-7〉 수질오염원 인근 아동수 현황

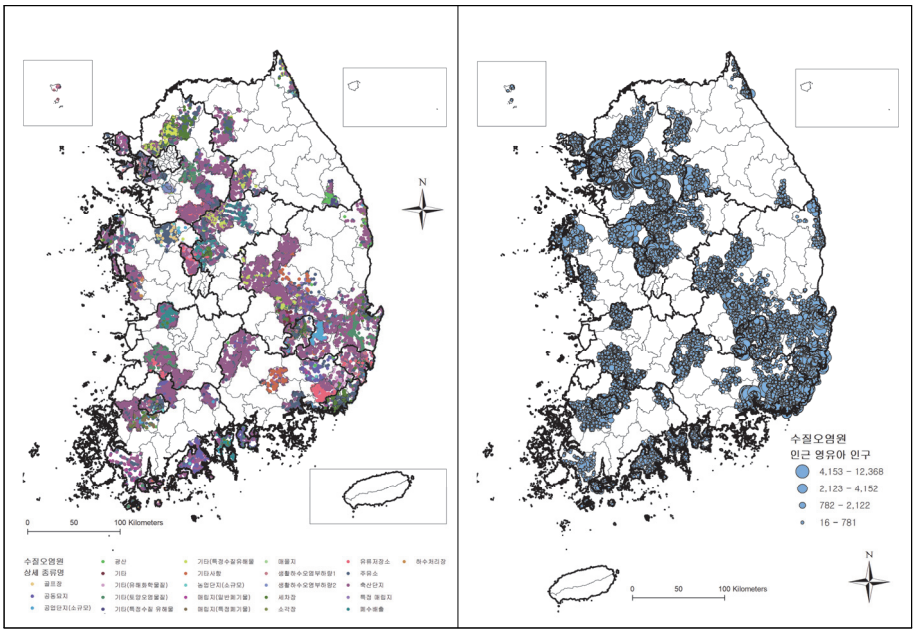
단위: 명

수질오염원	0 - 4세	5 - 9세	0-9세
하수처리장	140	110	250
공동묘지	56,618	74,453	131,071
공업단지	82,660	105,732	188,392
공업단지(소규모)	14	21	35
공장	55,779	78,466	134,245
광산	750	1,127	1,877
기타	1,116,977	1,391,449	2,508,426
기타(토양오염물질)	9,857	13,978	23,835
기타(특정수질 유해물질)	12,015	16,984	28,999
기타시설	4,068	5,390	9,458
기타토양오염물질	14,721	23,298	38,019
농공	423	593	1,016
농공단지	24,318	32,886	57,204
농업단지	15,866	21,586	37,452
농업단지 (소규모)	607	917	1,524
매립지	39,396	57,194	96,590
매립지 (일반폐기물)	76	81	157
매립지(일반폐기물)	1,402	2,052	3,454
묘지	5,202	7,969	13,171
병원	267,625	334,434	602,059
산업체	67,527	93,662	161,189
생활하수	79,638	113,166	192,804
생활하수 오염부하량	1,343	2,104	3,447
세차장	1,236,829	1,713,700	2,950,529
세탁소	190,575	236,787	427,362
소각장	11,154	15,519	26,673
쓰레기매립장	62	120	182
운수업	61,375	77,741	139,116
유류저장소	1,126,564	1,473,721	2,600,285

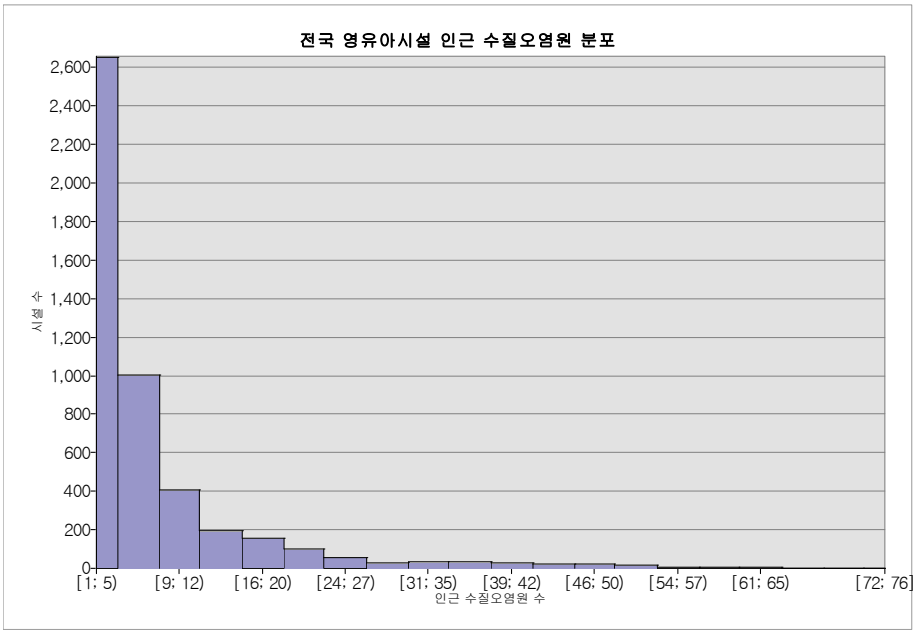
수질오염원	0 - 4세	5 - 9세	0-9세
유류취급	10,076	12,878	22,954
유해화학물질	344,000	460,028	804,028
인쇄업	4,837	6,463	11,300
일반 매립지	5,355	7,838	13,193
일반매립지	1,140	1,426	2,566
일반산업체	9,867	12,153	22,020
일반폐기물매	322	256	578
일반폐기물매립지	321	283	604
자동차정비	10,973	12,931	23,904
잠재오염원기타 (토양오염물질)	11,031	14,873	25,904
잠재오염원기타 (특정수질유해물질)	4,748	6,079	10,827
정비업	43,364	52,496	95,860
정수장	1,542	2,367	3,909
제조업	580,254	700,892	1,281,146
주유소	3,032,743	4,013,150	7,045,893
주유소 및 유류저장소	114,908	179,050	293,958
축사	306,093	436,641	742,734
축산단지	2,649,686	3,668,673	6,318,359
축산시설	305,302	398,065	703,367
토양오염	130,095	183,409	313,504
토양오염물질	75,097	101,662	176,759
토양오염유발시설	263,151	370,821	633,972
특정수질유해	49,755	63,266	113,021
특정수질유해물질	155,631	202,829	358,460
특정폐기물매립지	768	760	1,528
폐기물 매립지	737	921	1,658
폐기물매	1,300	2,045	3,345
폐기물매립장	810	1,584	2,394
폐기물매립지	9,064	11,530	20,594
폐기물처리	6,136	8,431	14,567
폐수배출	277,387	373,426	650,813
폐수배출시설	267,414	380,787	648,201
폐수처리장	596	468	1,064
하수종말처리장	2,235	2,946	5,181
하수처리장	375	609	984
환경기초시설	227	309	536

주: 수질오염원 개별시설별 인근 아동수로 아동수는 중복될 수 있음.

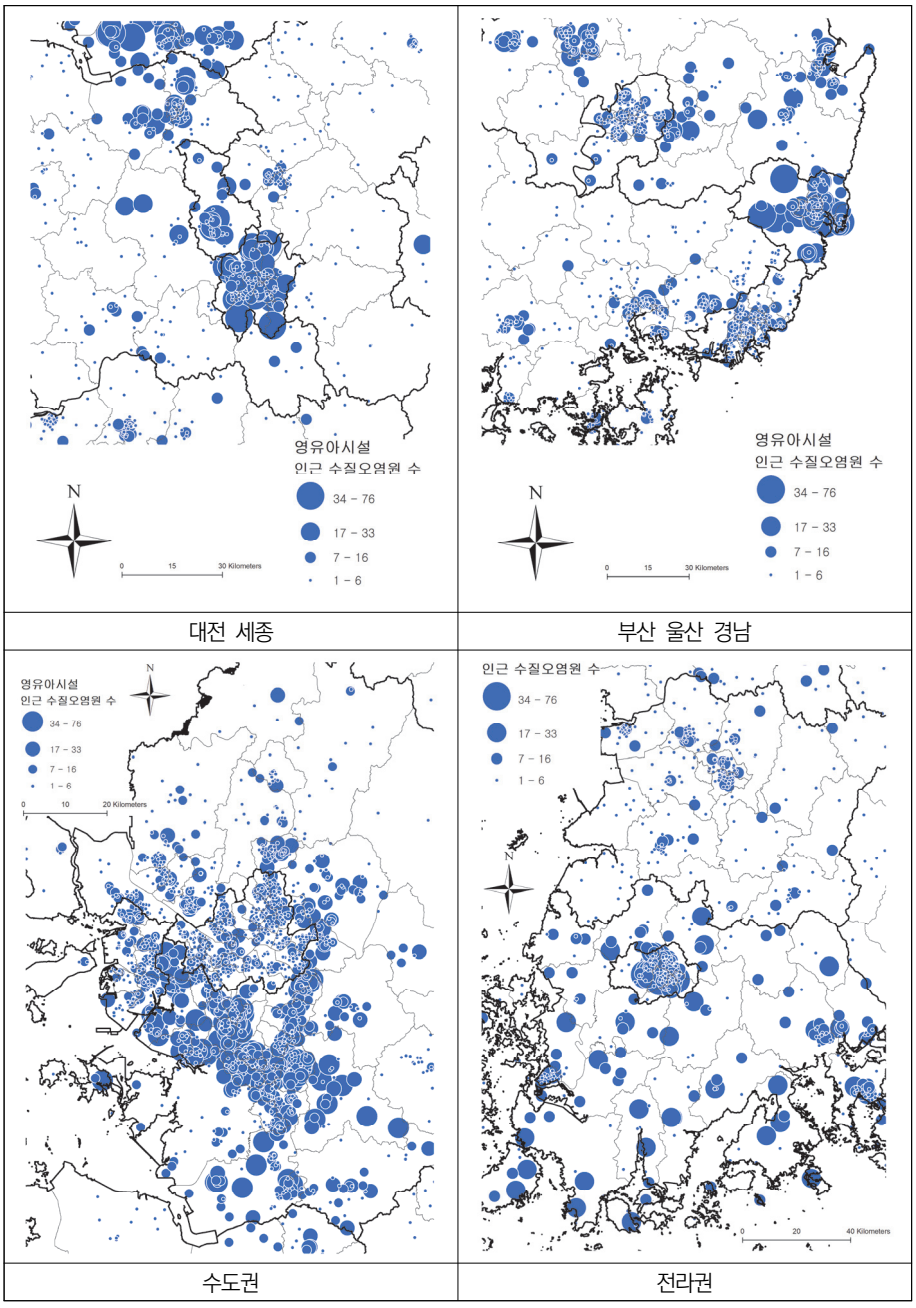
[그림 Ⅲ-2-13] 수질오염원 및 수질오염원 인근 영유아수



[그림 Ⅲ-2-14] 전국 영유아시설 인근 수질오염원 분포



[그림 Ⅲ-2-15] 영유아시설 인근 수질오염원 수



3. 소결

본 연구에서는 유치원과 어린이집 등 영유아 보육·교육기관에 대한 대기질, 위험시설, 수질 등과 관련한 안전·환경관련 데이터를 활용하여 빅데이터를 분석하였다. 이를 위해 유치원과 어린이집의 주소를 기반으로 하여 인근 대기질 지수, 미세먼지 지수, 위험시설, 수질지수, 수질오염원 인근 거주인구 정보 등을 분석하였다.

유치원·어린이집 인근 대기 질(미세먼저 등)이 나빠졌을 경우 기관 내의 공기질도 동반하여 악화될 가능성이 높은 상황이다. 특히 수도권과 대도시, 공장 등 산업시설이 많은 곳에서 전반적으로 대기질이 나쁜 것으로 나타났다. 위험시설은 전국적으로 유치원은 평균 2.3개, 어린이집은 평균 2.4개의 위험시설이 반경 500미터 내에 있는 것으로 나타났고, 주유소 등을 포함한 위험시설이 최대 20개인 어린이집도 있었다. 영유아 보육·교육시설의 수소이온농도, 잔류염소, 탁도 수치는 모두 법적 기준은 모두 충족하고 있었다.

Ⅳ

육아카페 빅데이터 활용을 통한 육아현안 분석

01 분석 필요성 및 목적

02 분석 내용 및 방법

03 분석 결과

04 소결

Ⅳ. 육아카페 빅데이터 활용을 통한 육아현안 분석

1. 분석 필요성 및 목적

저출생과 청년들의 비혼 추세, 출산을 포기하는 일련의 현상이 심화되면서 정부를 비롯한 지자체의 관심이 매우 높은 상황이다. 이에 정부 차원에서 영유아에 대한 무상 보육·교육, 아동수당 지급과 영아수당 도입(예정), 지자체별 출산장려금 등의 현금 및 서비스 지원 정책이 다양하게 추진되고 있다. 또한 육아휴직이나 유연근무제 도입 등 일·가정 양립을 위한 다양한 정책들도 마련되어 추진되고 있다.

그러나 임신·출산에 대한 지원부터 아이가 태어난 이후의 양육, 돌봄, 보육·교육에 대한 전반적인 지원이 강화되고 있음에도 불구하여 여전히 부모들의 수요를 충족하는 데는 한계가 있어 보이고, 저출생 현상도 쉽게 해결될 기미는 보이지 않는다.

이에 임신·출산부터 자녀양육에 이르기까지 다양한 정부와 지자체의 정책과 서비스의 수요자이자 이해당사자자인 부모들의 경험과 인식에 대해 살펴볼 필요가 있으며, 이를 당사자인 부모들이 서로 자신의 경험과 인식들을 공유하는 일명 맘카페의 빅데이터 분석을 통해 육아정책에 대한 현안과 향후의 방향을 모색하였다. 맘카페에서는 대부분 아이를 갖고자 하는 시기부터 자녀 출생 후 양육시기에 이르기까지 다양한 경험을 바탕으로 임신·출산·육아 관련 정보교환이 이루어지고 있으며, 결혼 생활에서 겪을 수 있는 다양한 고민을 공유하고 있기 때문에 좋은 정보원이 될 수 있다. 이를 통해 출산 당사자들의 니즈를 결혼을 비롯한, 임신·출산, 육아·돌봄 등에 있어 이들이 경험하고 있는 어려움이 무엇인지를 도출하고 이를 통해 육아정책에서의 사각지대도 찾아볼 수 있을 것이다.

2. 분석 내용 및 방법

가. 분석내용

분석내용은 먼저 맘카페에서의 육아관련 키워드에 대한 분석이다. 임신, 결혼, 육아, 돌봄 등 맘카페 게시물을 추출하여 임신·출산 정책, 육아·돌봄, 보육·교육 등에 대한 여론을 분석하였다. 맘스홀릭이라는 전국 최대단위 맘카페의 게시글의 제목과 내용을 중심으로 2020년 1월부터 2021년 7월까지의 게시글을 웹 크롤링 하였다.

분석에 앞서 사전에 육아정책과 관련된 최근의 이슈들을 전반적으로 파악하기 위하여 관련 언론기사 제목에 대한 키워드 분석을 실시하였다. 본 분석은 2020년 1월 1일부터 2021년 9월 7일까지 육아정책연구소 홈페이지의 육아정책 동향 게시판에 게시되고 있는 기사를 대상으로 하였으며, 단순 행사 홍보 기사 등을 제외한 후 총 2,447개의 기사 제목을 분석에 활용하였다. 형태소 분석을 통해 키워드를 도출하였고, 그 결과에 기초하여 TF분석으로 빈도를 파악하였다. 이때, 동일한 의미를 갖는 어절은 띄어쓰기에 관계없이 동일한 단어로 보았다. 분석 대상이 되는 기사에서의 주요 키워드는 총 140개였으며, 이중 높은 빈도를 차지한 50개 키워드는 다음과 같다.

〈표 Ⅳ-2-1〉 육아정책 관련 기사 제목 키워드 분석

순위	키워드	빈도	순위	키워드	빈도
1	코로나	128	26	시간	19
2	유치원	90	27	육아정책연구소	19
3	어린이집	78	28	학부모	19
4	아이	70	29	거리	18
5	부모	47	30	서울시	17
6	등교수업관련현황자료	37	31	자녀	17
7	학대	37	32	대응현황자료	16
8	아동학대	34	33	아이들	16
9	아동	33	34	가족돌봄휴가	15
10	학교	32	35	이유	15
11	초등	31	36	정인	15
12	정부	29	37	개월	14
13	등교	27	38	경찰	14
14	사회	25	39	종합	14

순위	키워드	빈도	순위	키워드	빈도
15	출산	24	40	아빠	13
16	엄마	23	41	우리아이	13
17	현황자료	23	42	필요	13
18	보육교사	22	43	강화	12
19	지원	21	44	결혼	12
20	교사	20	45	교육	12
21	안전	20	46	단독	12
22	저출산	20	47	마스크	12
23	교육분야코로나	19	48	신청	12
24	긴급돌봄	19	49	육아휴직	12
25	돌봄	19	50	학원	12

2020년 1월부터 2021년 7월까지 육아카페 맘스홀릭 게시글에 대한 웹크롤링 결과는 다음과 같다. 2020년에 40,978개의 게시글이 확인되었고, 2021년 1~7월까지는 28,858개의 게시글이 확인되어 전체 69,836개가 분석대상이 되었다. 이 중 상품광고, 체험단 모집, 체험단이 작성한 후기 등을 제거한 글 48,155개를 분석하였다. 분석대상 판단기준으로는 글자 수 200개 미만의 글 중 광고성으로 보이는 글을 제거하고 47,182개를 분석대상에 포함하였다. 또한 글자수 200개 이상의 글 중 상품 및 서비스광고, 체험단 모집, 체험단이 작성한 후기 등을 제거하고 973개 글을 분석대상으로 포함하였다.

〈표 Ⅳ-2-2〉 육아카페 게시글 웹크롤링 결과

단위: 개

구분	키워드	게시글 수	구분	키워드	빈도
2020년	1월	3,118	2021년	1월	1,906
2020년	2월	2,999	2021년	2월	1,575
2020년	3월	3,017	2021년	3월	4,496
2020년	4월	3,180	2021년	4월	5,582
2020년	5월	4,022	2021년	5월	5,487
2020년	6월	4,598	2021년	6월	5,491
2020년	7월	3,866	2021년	7월	4,321
2020년	8월	3,489			
2020년	9월	4,302			
2020년	10월	2,471			
2020년	11월	2,834			
2020년	12월	3,082	소계		28,858
소계		40,978	합계		69,836

나. 분석방법

수집된 육아카페 게시글은 비정형 텍스트 자료로 데이터 전처리 과정과 형태소 분석을 통해 분석이 가능한 데이터 형태로 가공하였다. 데이터 전처리 과정에서 분석에 불필요한 단어 및 어구는 삭제하였으며, 동일하거나 유사한 의미를 가진 단어들은 통일하였다.

먼저, 게시글 제목이나 내용에 '가입인사', '등업신청', '체험단' 등이 들어가는 게시글을 삭제하고, '오늘', '진짜', '부터', '근데' 등 연구목적에 맞는 분석에 필요하지 않은 단어와 '-', '·' 등 각종 특수기호 등 제외어를 설정하였다. 그리고 형태소 분석 과정에서 전문용어, 복합용어 등 분리돼서는 안 되는 단어가 분리되지 않도록 지정하였다. 마지막으로 유의어를 설정하여 의미가 비슷한 단어들을 하나의 단어로 통일하는 과정을 수행하였다. 예를 들어, 코로나, 코로나19, covid-19 등 동일 개념임에도 표현이 다른 경우에는 코로나19로 단어를 통일하였다. 문장을 의미 최소 단위인 형태소로 분절하는 형태소 분석에는 Python의 한국어 자연어 분석 패키지인 KoNLPy 패키지와 R 4.1.1(R에서 제공하는 'topicmodels' Package)을 사용하여 사용하였다.

이러한 절차를 통해 분석을 위해 사용할 최종 단어로 1,156개를 선정하였고, 제외어를 제거하고 유의어로 변경한 단어리스트에서 빈도수를 기준으로 빈도수 31개 이상 단어를 선택하였다.

3. 분석 결과

분석방법은 텍스트마이닝을 통한 토픽분석을 실시하였다. 텍스트마이닝이란 문서상의 의미있는 패턴과 관계를 찾고 이를 추출하는 것으로써 적절한 알고리즘을 용하여 비정형화된 데이터를 정형화된 데이터로 변환시켜 분석하는 일련의 과정이다(강창완 등, 2018). 텍스트마이닝은 웹 크롤링하여 얻어진 분석할 웹 데이터를 분석 가능하도록 원시 데이터를 정제하기 위한 정제단계를 거치고, 이후 정제된 웹 데이터를 각종 웹마이닝 기법을 적용하여 분석하였다.

가. 워드 클라우드 및 단어빈도표

분석 시 사용한 단어 1,156개를 살펴보면, 단어 '아이'가 가장 빈도수가 많고, 다음으로 '시간', '아침', '저녁', '남편', '자녀', '날씨', '커피' 순으로 빈도수가 많았다. 특히, 단어 '아이'는 다음으로 빈도수가 많은 '시간'에 비해서 2,500개 정도 많아 차이가 매우 컸다. 이는 '우리 아이가 예쁘다', '아이가 아파서 밥을 안먹어요.' 등과 같이 연령에 상관 없이 자신의 자녀에 대해 언급할 때 쓰이기 때문으로 보인다.

두 번째로 빈도수가 많은 단어 '아침'은 '입맛이 없어서 아침은 건너뛰고 싶다.', '아침 식사로 죽먹었어요.' '오늘 아침 늦게 일어났어요.' 등 '아침'이란 단어 자체가 '아침 식사'를 의미하거나, 하루 중 시간을 의미하는 두 가지 의미로 쓰이는 것으로 보인다. 세 번째로 빈도수가 많은 '저녁'도 '아침'과 동일하였다. 상위 단어 100개를 살펴보면, '아이', '남편', '친정'라는 단어가 등장하는 것으로 보아 글을 쓰는 주체는 주로 '엄마'들임을 알 수 있다. '아침', '저녁', '점심', '주말'이라는 단어가 3번째, 4번째, 9번째, 10번째 등장하는 것으로 보아 엄마들에게 굉장히 중요한 시간임을 알 수 있다. 특히 새벽이란 단어는 49번째 등장하는데 엄마들에게 새벽시간도 의미있는 시간일 것으로 추측된다.

커피(8번째), 한잔(21번째), 카페(50번째)이란 단어를 통하여 엄마들이 휴식을 취하는 방법을 알 수 있다. 엄마들은 카페나 집에서 커피를 마시고, 때론 술을 마시면서 휴식을 취하는 것을 알 수 있다. 또, 티비(39번째), 드라마(92번째), 방송(100번째) 이란 단어를 통하여 엄마들이 텔레비전을 보면서 휴식시간을 보내는 것을 알 수 있다.

생각(11번째), 고민(14번째), 걱정(16번째)이란 단어를 통하여 엄마들이 많은 고민을 하고 걱정을 하고 있음을 알 수 있다. 피곤(15번째), 육퇴(63번째)란 단어를 통하여 엄마들이 육아를 하면서 얼마나 피로를 느끼는지 알 수 있다. 주문(26번째), 선물(28번째), 제품(31번째), 마트(57번째), 배달(75번째) 단어를 통하여 엄마들이 제품을 온라인 주문하거나 마트에 직접 가고 제품을 배달시키는 것을 알 수 있다. 유산균(53번째), 영양제(80번째), 건강(85번째), 영양소(90번째) 등 단어를 보아서 엄마들이 아이와 본인 혹은 가족들의 건강을 건강보조식품 등으로 챙기고

있음을 알 수 있다. 낮잠(24번째), 자다(30번째) 등 단어를 통해서 엄마들이 낮잠을 자거나 아이를 재우기 위해 노력함을 알 수 있다. 빨래(32번째), 청소(36번째) 등을 통해서 육아와 동시에 집안일을 수행하고 있음을 알 수 있다. 과일(48번째), 요리(81번째), 라면(87번째), 이유식(88번째) 등 단어를 통하여 엄마들이 가족들을 위한 요리와 디저트 준비를 하고 있음을 알 수 있다. 피부(33번째), 화장품(64번째) 단어를 통하여 아이와 엄마 본인의 피부관리를 신경쓰고 있음을 알 수 있다. 또한, 운동(35번째), 신체(60번째), 다이어트(93번째)와 같이 건강과 몸매관리에도 신경쓰고 있다.

[그림 Ⅳ-3-1] 상위 1,156개 단어 워드클라우드

〈표 Ⅳ-3-1〉 상위 100개 단어 빈도수

단위: 개

순서	단어	빈도수	순서	단어	빈도수	순서	단어	빈도수	순서	단어	빈도수
1	아이	6,522	26	주문	971	51	내일	668	76	종일	523
2	시간	4,022	27	정리	963	52	감기	664	77	월요일	520
3	아침	3,937	28	선물	917	53	유산균	659	78	건조	514
4	저녁	3,387	29	기분	905	54	어린이집	656	79	에어컨	511
5	남편	3,276	30	자다	901	55	장난감	654	80	영양제	511
6	자녀	3,201	31	제품	892	56	사진	650	81	요리	511
7	날씨	2,702	32	빨래	878	57	마트	644	82	놀이	506
8	커피	2,548	33	피부	871	58	지역	640	83	친정	504

순서	단어	빈도수	순서	단어	빈도수	순서	단어	빈도수	순서	단어	빈도수
9	점심	1,985	34	친구	865	59	오후	638	84	정신	490
10	주말	1,846	35	운동	856	60	신체	635	85	건강	488
11	생각	1,816	36	청소	849	61	마음	627	86	오전	484
12	아기	1,742	37	우리	817	62	사용	620	87	라면	481
13	시작	1,648	38	큰애	779	63	육퇴	617	88	이유식	472
14	고민	1,628	39	티비	768	64	화장품	611	89	덥다	470
15	피곤	1,562	40	사람	766	65	학교	605	90	영양소	467
16	걱정	1,388	41	출근	766	66	머리	603	91	다행	463
17	추천	1,350	42	할인	751	67	소리	588	92	드라마	463
18	바람	1,338	43	반찬	737	68	퇴근	588	93	다이어트	461
19	준비	1,314	44	여름	727	69	따뜻	583	94	핸드폰	460
20	엄마	1,272	45	구매	724	70	공부	558	95	코로나19	458
21	한잔	1,157	46	아들	709	71	궁금	555	96	조심	455
22	마스크	1,147	47	가격	690	72	외출	542	97	효과	453
23	시원	1,076	48	과일	689	73	미세먼지	525	98	소식	446
24	낮잠	1,040	49	새벽	676	74	재미	525	99	필요	446
25	병원	989	50	카페	672	75	배달	523	100	방송	444

나. 토픽모델링

본 분석에서 토픽 수는 15개로 지정하였다. 다음의 표는 각 토픽별 가장 연관성이 높은 순서로 상위 30개 단어를 나타낸 것이고, '토픽비율'은 문서 전체에서 각 토픽이 차지하는 비중을 나타낸다. 즉, 다음의 표에서 나타난 것처럼 전체 문서에서 Topic1이 차지하는 비중은 7.97%, Topic2가 차지하는 비중은 7.83%, Topic3이 차지하는 비중은 7.31%, Topic4가 차지하는 비중은 7.31%임을 알 수 있다. '토픽 명'은 연구자가 토픽을 가장 잘 설명할 수 있도록 관련 단어를 참고하여 지정하였다.

Topic1은 '피부 및 건강제품 추천', Topic2는 '식사 및 간식', Topic3은 '질병 및 치료', Topic4는 '제품 할인 및 온라인 구매', Topic5는 '제품 추천 및 추천요청', Topic6은 '아이와 엄마의 휴식', Topic7은 '아이 선물 및 이유식', Topic8은 '휴일과 외출', Topic9는 '커피와 휴식 및 디저트', Topic10은 '날씨', Topic11은 '시간 사용', Topic12는 '아이 및 엄마 공부', Topic13은 '엄마 여가시간 및 아이 사진촬영', Topic14는 '일상과 날씨', Topic15는 '운동 및 다이어트'로 토픽별 상위 키워드와 게시판 내용을 토대로 토픽명을 결정하였다.

〈표 Ⅳ-3-2〉 육아카페 토픽별 상위 30개 키워드 및 토픽명

구분		토픽														
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
토픽명		피부 및 건강 제품 추천	식사 및 간식	질병 및 치료	제품 할인 및 온라인 구매	제품 추천 및 추천요청	아이와 엄마의 휴식	아이 선물 및 이유식	휴일과 외출	커피와 휴식 및 디저트	날씨	시간 사용	아이 및 엄마 공부	엄마 여가시간 및 아이 사진촬영	일상과 날씨	운동 및 다이어트
토픽비율(%)		7.97%	7.83%	7.31%	7.31%	7.15%	6.74%	6.65%	6.48%	6.35%	6.24%	6.12%	6.10%	5.98%	5.98%	5.78%
토픽별 상위 키워드	1	피부	저녁	병원	마스크	제품	아이	선물	시간	커피	날씨	시간	아이	시간	아침	운동
	2	유산균	점심	감기	할인	사용	시간	아이	날씨	한잔	바람	커피	공부	드라마	주말	다이어트
	3	제품	반찬	엄마	주문	추천	낮잠	아기	사람	아침	미세먼지	저녁	학교	사진	날씨	저녁
	4	추천	라면	걱정	쿠폰	가격	아침	이유식	남편	점심	에어컨	청소	수업	노래	빨래	머리
	5	영양제	아침	아이	가격	구매	엄마	장난감	주말	우유	아침	정리	학원	티비	시작	남편
	6	성분	치킨	우리	택배	생각	남편	엄마	마스크	저녁	창문	아침	온라인	카페	월요일	아침
	7	아기	배달	생각	구매	고민	새벽	사주	저녁	아이스크림	여름	점심	시간	영화	소식	맥주
	8	아이	남편	마음	배송	아기	아기	고민	아이	과일	환기	오후	수학	작성	출근	시간
	9	효과	떡볶이	전화	인터넷	매트	아빠	생일	버스	아이	감기	한잔	영어	내용	세탁기	냄새
	10	사용	김밥	아침	방송	아이	티비	친구	산책	딸기	하늘	준비	걱정	주말	기분	스트레스
	11	분유	메뉴	남편	사용	디자인	소리	아들	놀이터	과자	선풍기	낮잠	생각	핸드폰	시간	생각
	12	물티슈	고민	사람	쇼핑	이사	큰애	주문	마트	마트	공기	설거지	고민	육퇴	이불	피곤
	13	콜라겐	김치	시간	행사	브랜드	피곤	우리	시댁	수박	태풍	오전	시작	라디오	우산	시작
	14	크림	아이	아기	박스	소파	육퇴	생각	아침	고구마	걱정	집안	학습지	아이	내일	다리
	15	얼굴	시간	치료	제품	사이즈	어린이집	조카	휴가	카페	장마	퇴근	개학	우리	준비	음식
	16	건강	계란	친구	마트	냄새	주말	시작	바람	기분	햇빛	빨래	유치원	부부	아이	피로
	17	기저귀	마트	콧물	구입	카시트	저녁	언니	친정	날씨	보일러	정신	문제집	마지막	종일	야식

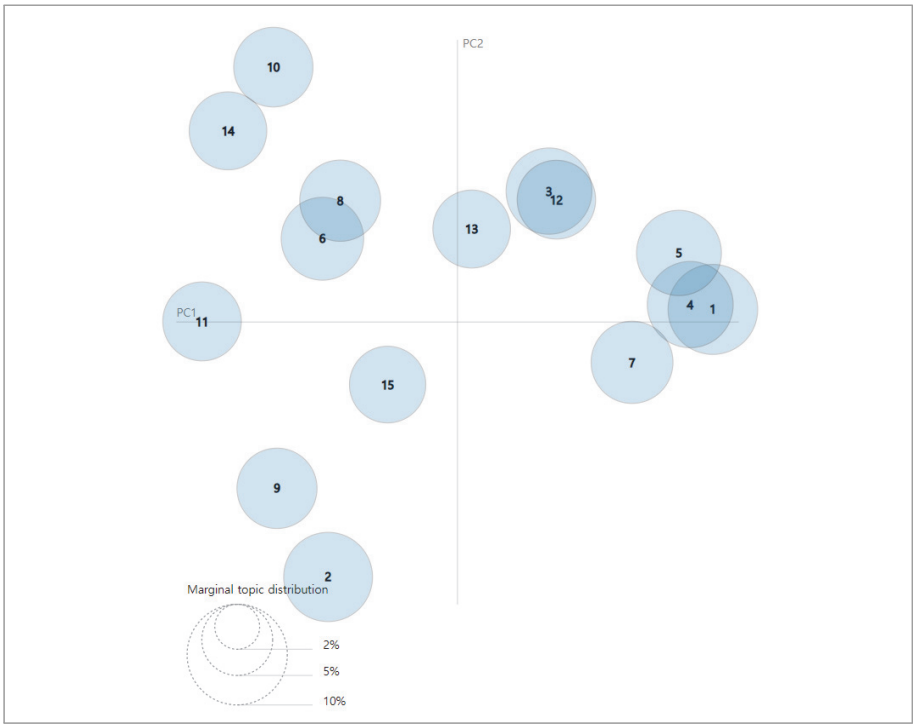
구분	토픽														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
18	관리	요리	어린이집	온라인	사워	아들	전집	공원	사과	청소	하원	교육	사람	등원	치킨
19	로션	음식	임신	정보	마음	거실	준비	연휴	바나나	시원	주말	자격증	촬영	오후	홈트
20	비타민	준비	수술	무료	업체	자유시간	방송	기분	생각	종일	남편	등교	마음	다행	주말
21	마스크	생각	진료	남편	선풍기	종일	추천	자전거	시간	외출	마무리	방학	재방	토요일	종일
22	걱정	냉장고	검사	생각	남편	늦잠	가격	외출	정신	소리	피곤	학습	채널	정리	그날
23	생각	해결	친정	카드	거실	우리	그림	운동	식사	기분	아이	한글	방송	정신	영양제
24	신경	재료	접종	신청	세제	준비	패키지	점심	얼음	저녁	여유	어린이집	마무리	피곤	점심
25	변비	만두	소리	가방	청소	짜증	사진	내일	오후	온도	큰애	준비	생각	바람	소화
26	트라블	주말	다행	이상	가능	할머니	인형	명절	샌드위치	새벽	화장실	큰애	유튜브	하늘	두통
27	여름	주문	예약	문자	사진	유치원	완성	동네	아이스	산책	점심시간	문제	남편	일요일	한잔
28	고민	볶음밥	비염	추가	유모차	생각	남편	아파트	음료	아이	자유시간	엄마	본방	장마	친구
29	도움	야채	이상	상품	설치	목욕	놀이	여행	냉장고	장난	컴퓨터	강의	이야기	외출	효과
30	화장품	걱정	백신	포인트	인테리어	덕분	구성	출근	남편	집안	출근	관심	세계	걱정	방법

다. 토픽 모델 시각화

Intertopic Distance Map은 토픽모델링 결과를 2차원 평면에 나타낸 것으로, 이것으로 각 토픽의 비율과 토픽 간 거리를 파악할 수 있다. 각 원의 크기는 〈표 Ⅳ-3-2〉에서 보여준 토픽비율로 전체 문서에서 각 토픽이 차지하는 비율이며, 원이 클수록 전체 문서에서 해당하는 토픽이 큰 비중을 차지하고 있음을 의미한다.

각 원들 사이의 거리는 토픽 간 거리로 원들 사이의 거리가 가까울수록 유사한 토픽으로 해석할 수 있다. 예를 들어, 그림에서 보면 Topic1과 Topic4, Topic5가 매우 유사한데, Topic1은 '피부 및 건강제품 추천', Topic4는 '제품 할인 및 온라인 구매', Topic5는 '제품 추천 및 추천요청'에 관한 것이다. 또한, 그림에서 보면 Topic8과 Topic9는 매우 유사해 보이고, Topic8은 '휴일과 외출', Topic9는 '커피와 휴식 및 디저트'인데 둘은 휴식이라는 공통점이 있다.

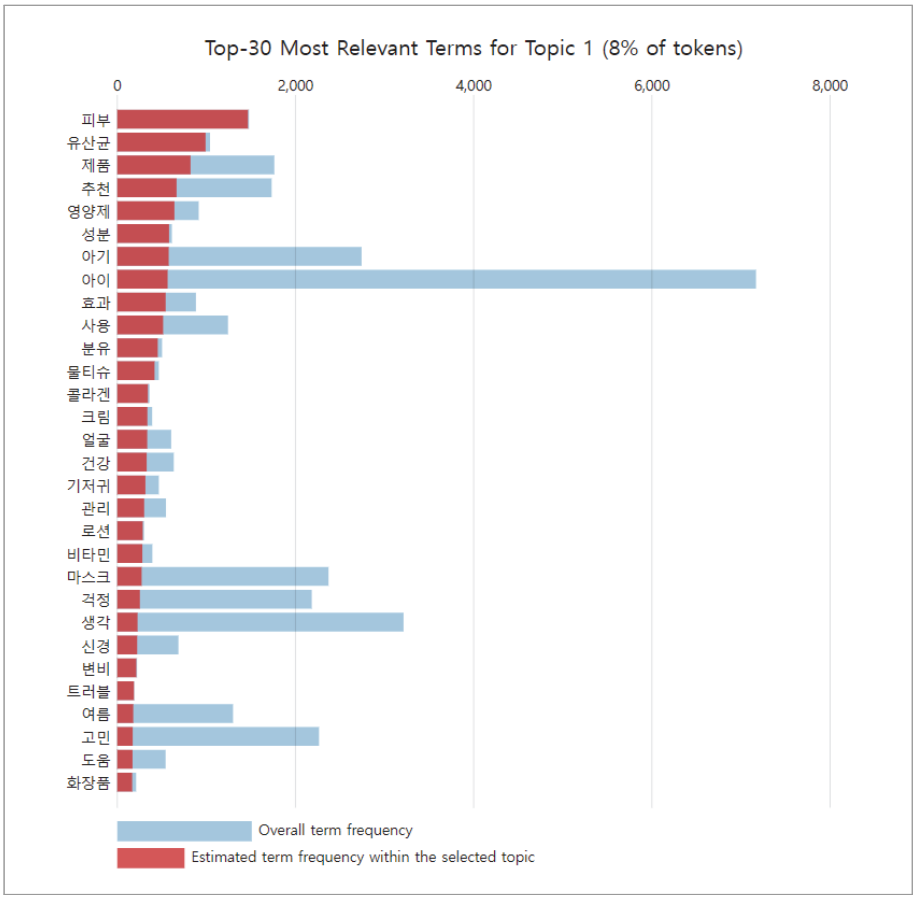
[그림 Ⅳ-3-2] Intertopic Distance Map



다음의 막대 그래프로 표현된 그림은 각 토픽에 해당하는 단어들 중 빈도수가 가장 많은 상위 30개 단어들을 나타내고, 맨 위에 위치한 단어일수록 토픽을 가장 잘 설명하고 있다고 볼 수 있다. 이 단어들을 참고하여 각 토픽이 무엇을 말하고 있는지 유추할 수 있고, Topic1과 관련성이 높은 맘카페 게시글을 참고하여 토픽을 결정하였다.

Topic1의 주제는 '피부 및 건강 제품 추천'으로 토픽명을 정하였고, 피부관련 제품(콜라겐, 화장품 등), 유산균, 영양제, 물티슈 등 건강 혹은 미용 관련 제품을 추천하거나 추천 요청하는 내용이다. 예를 들어, '화장품 추천해주세요.', '콜라겐 효과 있네요.', '아이 오메가3 추천해주세요.', '유산균 추천해주세요.' 등이다.

[그림 Ⅳ-3-3] 육아카페 토픽 모델링 시각화 결과(토픽1)

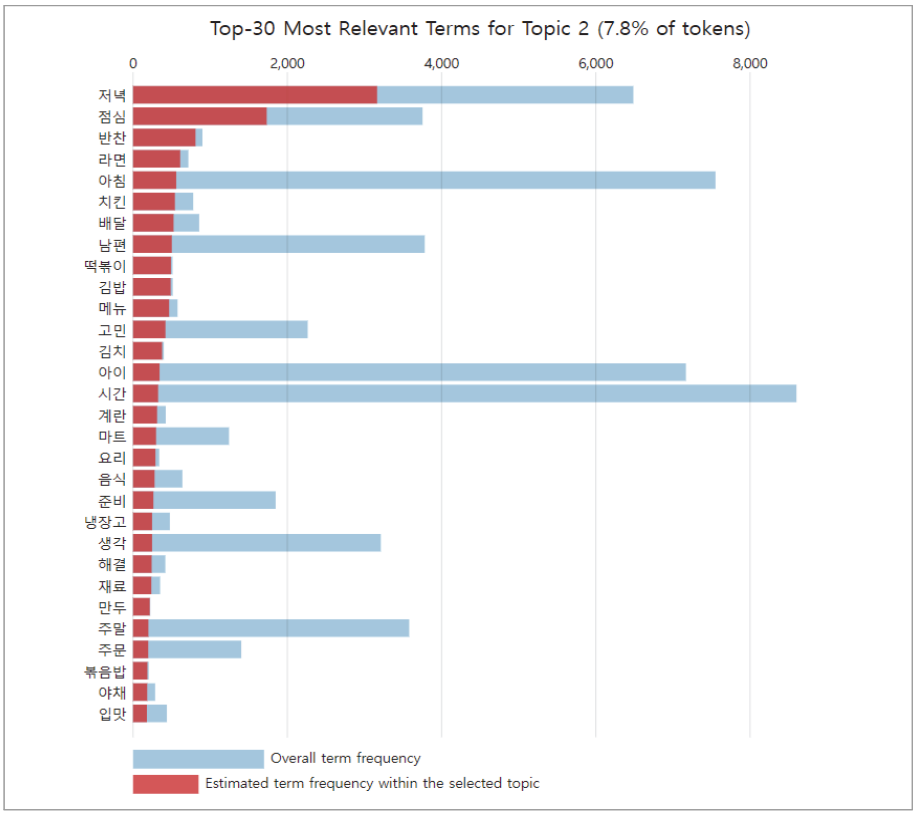


〈표 Ⅳ-3-3〉 Topic1 관련 원본문서

순서	제목	내용
1	애기 건조할 때 괜찮은 화장품 추천좀 해주세요	첫째 때는 안 그랬는데 둘째는 피부가 엄청 예민한지 화장품도 첫째가 쓰던 거 써도 문제가 생기네요 그렇다고 건조함이라도 잡히면 괜찮은데 그것도 안 잡혀요? 좀 예민한 피부인 거 같아서 순한 성분 브랜드나 제품이나 보습에 엄청 좋은 성분 있으면 추천 좀 해주세요
2	물티슈 가성비 괜찮아요	외출이 조금씩 늘어나면서 물티슈를 많이 사용하고 있어요 아무래도 실내보다는 야외로 나가니 휴지를 넉넉하게 챙겨가지고 다니는데 요즘에는 붕어빵 물티슈를 쓰고 있어요! 휴대용은 패밀리 라인으로 사용 중인데 원단도 좋고 피부 자극도 없어요 가성비 갑이라 부담 없이 쟁이고 사용해요!
3	김태희씨가 광고하는 분유 이름이 뭐에요?	TV에서 김태희씨가 광고하는 분유를 봤는데용 뭔가 우아하고 고급진게 좋아 보이더라구요ㅎㅎ 정확한 이름을 몰라서 혹시 김태희 씨가 광고하는 분유 이름 아시는 분이나 직접 쓰고 계신 분 있을까용?
4	뷰티 유튜브 많이들 보시나요??	알고리즘으로 디렉터 파이? 이 분 거 떴는데 레스트업 세럼 라인들 좋다고 추천해주셨는데 좋아 보여서요. 제가 요즘 피부가 다 뒤집어져서 진정 효과 있는 걸로 당분간 써보려고 하는데 케모시카? 성분 들어 있어서 괜찮다고 말해서 요~ㅠㅠ 시카랑 케모시카랑 다른 건가요??
5	콜라겐 먹으니까 효과있는거 같아요	콜라겐 다 상술인 줄 알았는데 먹으면 다르다고 하는 게 먼지 알거 같아요! 원래 마스크 쓰면 콧잔 등, 턱 주변 닿는 부위에 각질 뜨고 개기름도 심했는데 피부에 보충되서 그런가 훨씬 덜하고 좋아져서 만족 중 ㅠㅠ 이래서 다들 콜라겐 드시나 봐요ㅎ
6	아이 오메가3 꼭 먹어야하나요?	오메가는 비리다고 안 먹을려고 해서 매일 싸우면서 먹여서 제가 포기하게 되네요 거의 못 먹여요 초등학교 면서 먹이기 시작했는데 꼭 먹여야 하나 궁금해서 요지큐랩 유산균이랑 연세대 점업 종합영양제 먹고 있는데 오메가는 요런 거랑 별 게 인가요? 안 비린 오메가 추천해주세요?
7	유산균 추천해주세요!!	유산균에 대해 잘 모른답니다ㅠㅠㅠ

Topic2의 주제는 ‘식사 및 간식’으로, 점심밥, 저녁밥 준비나 식사를 하거나 간식을 먹은 내용 등과 관련이 높았다. 예를 들어, ‘저녁밥으로 계란찜해서 먹고 있다.’, ‘반찬으로 돈까스 튀겨 먹으려고 한다.’, ‘재난 지원금으로 회를 시켜먹었다.’ 등 내용이 있다.

[그림 Ⅳ-3-4] 육아카페 토픽 모델링 시각화 결과(토픽2)



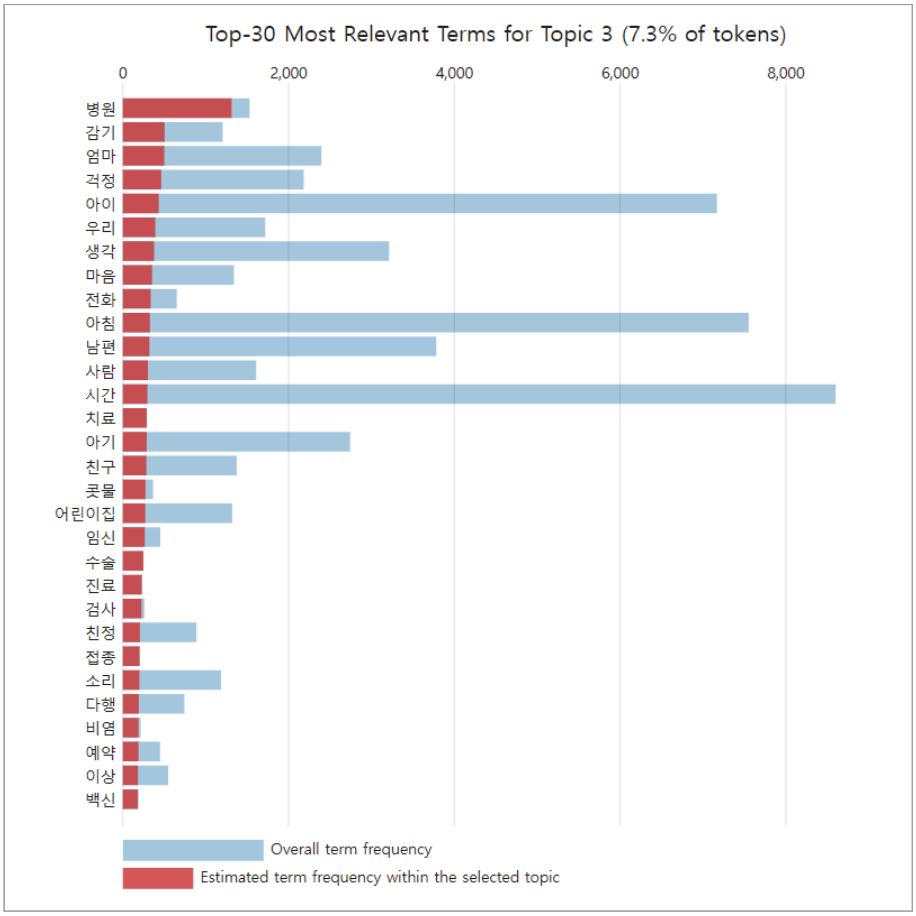
〈표 Ⅳ-3-4〉 Topic2 관련 원본문서

순서	제목	내용
1	저녁밥	간단하게 계란찜해서 지금 먹고 왔네요!! 설거지를 바로 해야 하는데... 살짝 귀찮아서 못하고 지금 쉬고 있네요!! 후식 좀 먹고 해야겠어요
2	밥 하네요	저녁을 일찍 먹는 편이라 밥 하네요. 있는 반찬에 돈까스 튀겨서 먹으려고요. 급 배고프네요
3	점심 먹고서	아이들이랑 점심 먹고 나서 아이들 그림 맞추는 카드 놀이 하다 보니 시간이 금방 지나갔네요 이제 컴터 좀 하고 애들 방치우라고 해야겠어요
4	어제오늘	어제 오늘 기분 좋네요 재난 기부금 4인 가족 나와서 그렇게 먹고 싶던 회... 드디어 시켜 먹었어요 ㅠㅠ감동감동 점심엔 형부가 콩국수 먹고프다고 했더니 다 같이 콩국수 먹으러 출동했네용 ㅋㅋ 가까운데 언니 사니 요래요래 조아영ㅎ
5	아이 반찬은..	주로 소고기 불고기 피망, 호박, 가지 야채볶음, 닭고기 야채볶음생선구이(고등어, 삼치) 생선 야채조림계란 후라이, 계란찜 이렇게 돌아가면서 주는데요.. 너무 똑같은 반찬만 매번 올라가는 것 같아서..?뭐 다른 반찬 해줄 만한 게 있을까요? 간편하면서도 영양 있는 거....?

순서	제목	내용
6	반찬도착..ㅎㅎ	택배기사분이 이 시간에 오셧네요. 내일 반찬거리..ㅋㅋㅋ 요즘에 반찬을 주문해서 먹거든요. 진미채, 콩자반, 오징어 볶음 등등..5개 왓어요~
7	저녁먹어요	치킨텐더 굽고, 리코타 치즈 샐러드에 추가해서 같이 먹어요

Topic3의 주제는 '질병 및 치료'로, 병원 치료한 내용 및 질병에 걸린 내용 등으로, 대상은 본인, 아이, 부모님 등 다양하였다. 예를 들어, '자녀 둘 다 감기에 걸렸다.', '관절염 치료 잘하는 한의원 정보 공유부탁합니다.' 등이 있었다.

[그림 Ⅳ-3-5] 육아카페 토픽 모델링 시각화 결과(토픽3)



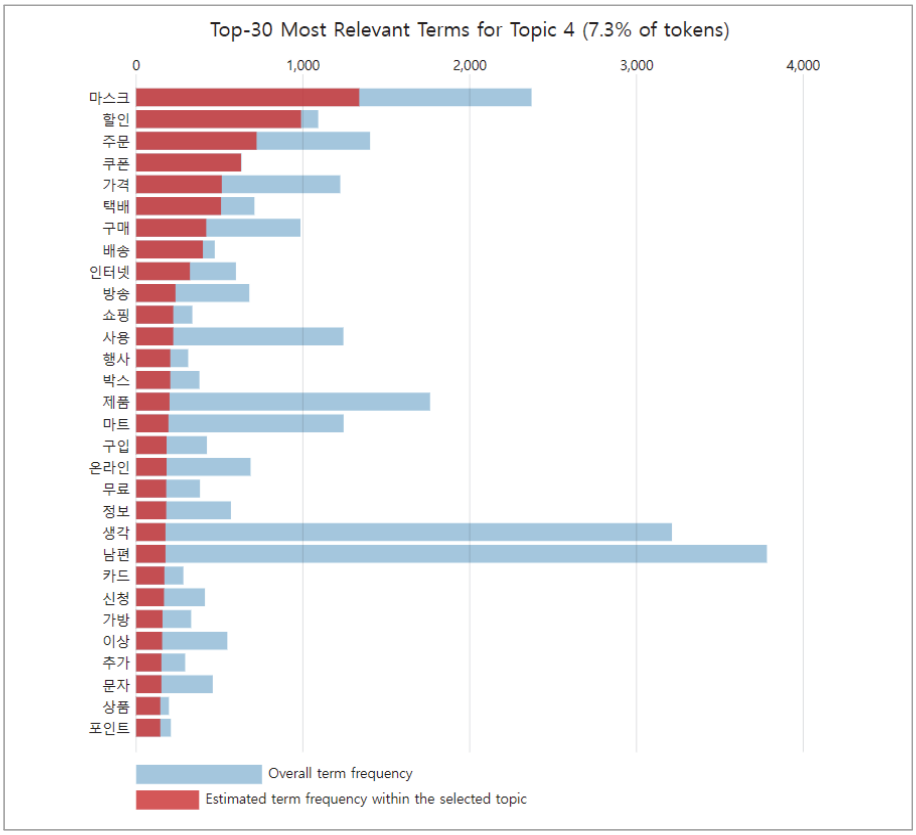
〈표 Ⅳ-3-5〉 Topic3 관련 원본문서

순서	제목	내용
1	두명 다 감기 걸렸어요	첫째가 시작으로 둘째까지 걸려버렸어요 ㅠ 남편도 간당 간당하는데. 저마저도 감기 걸릴까봐 마스크 쓰고 생활하네요..이번주는 집 밖을 못 나가네요..
2	어머니께서 퇴행성관절염 진단을 받으셨어요	올 초부터 자꾸만 무릎을 만지시고 한숨만 쉬시길래 날 잡고 여쭤보니 무릎이 예전 같지 않으시다고 많이 아프시다네요 ㅠ병원에 가봤더니 퇴행성이라고 하면서 MRI는 몇 장 찍더니 인공관절 수술을 해야 한다고 하네요 …딴 병원도 모셨는데 이번엔 부분치환인가 하는 걸로 권유를 하는데.. 혹시 수술을 하지 않고 한의원에서 치료하는 걸로는 관절염 치료하지 못할까요? 저는 저희 어머니께서 아직 젊다고 생각하지만 그래도 수술을 받으시게 되면 너무 고생이 심하셔서 몸도 마음도 많이 힘들어 하실까 봐요..? 한의원에서 하는 치료로 효과가 있을 지 궁금하고 은평구에 관절염 치료 잘하는 한의원도 아시면 정보 공유 부탁드려 봅니다ㅠㅠ
3	23주차 입체초음파하고 왔어요~	울감에 임태기를 여러 번 해봤었어요!! 임신기간 동안 입덧이 없어 임신을 실감하지 못했지만 입덧 대신 먹덧이 와서 엄청나게 먹다 보니 체중이 너무 많이 늘어났네요 ㅜㅜ 잦은 태동과 볼록해진 배를 보며 임신했다는 걸 실감하게 되네요~ 드디어 23주 3일 되는 날 정밀 초음파 했어요~ 두근두근~ 설레이는 마음으로~ 광안 ㅈㅁ병원에 도착!!? 전 7과 원장님 담당인데 정밀 초음파는 3과 과장님께 받았어요~??? 광안ㅈㅁ병원 선생님들은 다들 친절하신 거 같아요~?? 정밀 초음파 하시는데 꼼꼼히 잘 봐주셨어요 ㅎㅎ 토토가 꼬물꼬물 엄청 움직여서 시간이 조금 걸리긴 했어요~~ 차근차근 머리 둘레 기형아 여부 손가락 발가락 꼼꼼히 봐주시고... 이것저것 자세히 설명해주시더라구요~~ 기분 좋게 검사 받고 왔어요~~ 4D 입체초음파& 임신성 당뇨검사 예약하고 왔는데 ㅜㅜ 사실 벌써부터 걱정이 되네요~
4	광주 갑상선 병원 어디 없을까요..?	광주에 계신 친정 엄마 병원 알아보고 있어요 코로나 때문에 광주 방문도 어려운데 걱정이 많네요 갑상선으로 예전에 병원 치료 받은 적이 있었는데 광주에 갑상선 전문 병원 추천 좀 부탁드릴게요
5	오른쪽 나팔관? 있는쪽이 콕콕콕이요..병원을 언제 가는게 좋을까요?	3/9일 화유 되었구요.. 다낭성이 있고 배란 일이 정확하지 않은 터라.. 날짜 잡기 힘들어 병원을 다니구 있는데 다니고 있는 병원이 맘에 안 들어 옮기고 싶어요.. 지금 두 오른쪽 나팔관(?) 있는 위치 쪽인지 오른쪽 골반 밑에 쪽이라 해야 하나 ㅠ 몇 일 전부터 콕콕콕 아파요 3/25일 관계했는데 임신 가능성이 있을까요.. 착상통이라 생각하고 조금 기다려볼지, 아님 다음 임신기회를 위해 빨리 병원을 가봐야 할지 모르겠어요 일반 산부인과 진료보다 이제는 난임 병원으로 옮겨서 나팔관 조영술부터 받아봐야 하는데.. 이것도 생리 후에 가야 하는 건가요? ㅠ 하루라도 빨리 가는 게 좋을지, 기다렸다 생리터지고 가야 할지 알려주세요ㅠ
6	아버지 티눈 제거...	티눈이 진짜 크고 딱딱하고 딱 보기에도 깊어 보이세요. 거의 살이랑 합체되어서 굳은살처럼 보일 정도.. 걸을 때도 힘들어하시는 거 같은데 말을 안하셔서 병원 한 번 알아봐 드리려고 하거든요 요런 거는 약국 약으로도 해결 안 되겠죠? 무조건 병원 가야 되는 거죠?
7	더 심해졌어요... ㅡㅜ	지난 주 목요일부터 아이가 기 침끼가 있더라구요~~~콧물도 약간? 그래서 소아과 갔다가 가정보육하고~~ 주말까지 푹 쉬어서 이젠 감기가 다 나았겠구나~~ 했지만 그래도 아직 아이가 콧물이 약간 여전히 보이는 거 같길래 어제도 소아과 가서 약을 더 받아 오긴 했는데~증상이~~심하지 않고, 끝나가는 중인 것 같다며 약하게 주시는 거 같더라구요~~저도 아이 감기가 끝나 가는

순서	제목	내용
		것 같았고.....근데.....아이가 다 나아간다고 생각되 서 킥보드 태우고 다녀왔는데 그게 힘들었는지~~그 이후부터 누런 콧물이 보이고~~기침도 심해졌어요...ㅡㅜ 어젯밤에는 계속 기침하느라 또 잠을 잘 못 자서 가슴에 패치가 하나 남은 게 생각이 나서 그거 붙여줬더니 그제서야 그 때부터 잘 자더라구요....ㅜㅜ 왜 더 심해진 건지~~ㅜㅜ 계속 어린이집 안 보내고 집에서 쉬었는데...ㅜㅜ

Topic4의 주제는 '상품 할인 및 온라인 구매'로, 상품 할인정보 공유, 상품 할인받아 구매 혹은 택배 배송 등에 관련이 높았다. 예를 들어, '30% 할인받고 샀어요.', 'KF94마스크 사기 힘드네요. 인터넷에서 구매할 수 있는 정보 공유합니다.', '결국 반품처리 되었어요.' 등 상품 할인 및 온라인 상품 구매에 관련된 내용 및 '택배 때문에 멘붕 왔네요.' 등 물건 배송에 관련된 내용이 있었다.

[그림 Ⅳ-3-6] 육아카페 토픽 모델링 시각화 결과(토픽4)

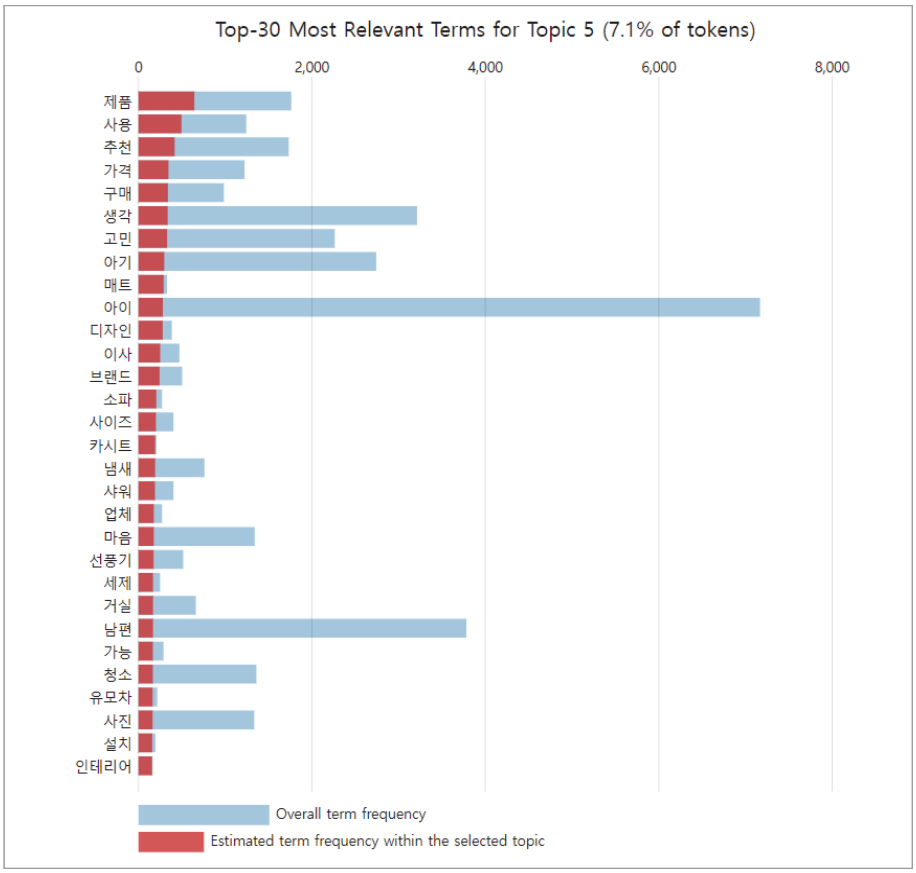


〈표 Ⅳ-3-6〉 Topic4 관련 원본문서

순서	제목	내용
1	택배 때문에 멘붕 왔네요	저희 집은 롯데 택배가 항상 밤에 오는데 한 박스 남편이 갖고 들어오길래 다 왔나 보다 택배 앱 보니 식품 한 박스를 완료 찍어놓고 박스 제품만 놓고 간 거죠ㅠㅠ 하필 식품을 완료 찍고 배송을 멈췄네요?
2	결국 반품처리 되었어요	택배 물건 하도 안 와서 문의하니 다른 지역에서 배송 완료 되었더라구요 판매자한테 문의 남기니 전화 왔는데 반품 처리 해준다네요 물건이 없어서 보내줄 수가 없다면서요 ㅠㅠ
3	30%할인받고 추가증정도..ㅎㅎㅎ	날 더워지니 배달이유식 소중함이 더 느껴지고 있는데 울 아이 먹일 이유식 주문할려고 홈페이지 들어가 보니 30% 할인과 함께 추가 증정 행사가 진행 중이라서 기분 좋게 결제 완료~2주분 구입했는데 6세트나 추가 증정 받게 됐네요..
4	마스크진짜 힘드네요 ㅠㅠ	오전부터 돌아댕기다가 줄 보고 그냥 오다 인터넷 써칭하네요 ㅠㅠ 약국에선 종일 줄서야 하고 인터넷은 전부 비싼 가격에도 품절이고 kf 94구하기 진짜 어렵네요 ㅠ 인터넷만 엄청 뒤지다 정보 공유합니다 kF94마스크 수시로 업데이트 되는 사이트네요
5	이런 기저귀가방 어때요?	안녕하세요!! 출산 앞두고 있는 예비맘입니닷!! ㅎㅎ 첫 출산이라 그런지 벌써부터 육아용품 이것저것 사들이 고 있네요 ㅎㅎ 설레발..ㅠㅎㅎㅎ 맨날 소셜 들어가보는데 위메프에서 특가 떠서 물어봐 용? 이런 가방들도 기저귀 가방으로 많이 사나요?? 기저귀 가방 브랜드는 아니고, 일반 가방 브랜드 같은데 음.. 여기 브랜드는 사본 적 없긴 한데, 거의 80 프로 이상 할인되 는 것 같아서요!? 이런 기저귀가방은 괜찮을까요? ? 뭔가 기저귀 아니어도 물건 담고 다니기에 좋을 것 같긴 한데.. 의견 부탁드려욤! 여기 다른 에코백도 있긴 하던데, 제가 보는 눈이 워낙 없고 선택 장애라서요 ㅠㅠ 다른 거 괜찮은 거 있으면 추천해주세요!!!!
6	홈*** 설선물세트 사전예약 할인~	홈***에서 설날 선물세트 사전 예약하면 19일부터 24일까지 100만원 당 10만원 상품권 최대 100만원까지 준대요~ 사전예약 기간 동안 혜택이 동일한 줄 알고 늦장부리다가,, 1차분을 놓쳣어요 ㅠㅠ 행사 카드로 구매하면 30% 할인도 되고, 사 전 예약 상품권도 받을 수 있으니 근처 홈플러스 익스프레스 한 번 방문해보세요~도움이 되실까 하고 공유해요~?

Topic5의 주제는 ‘제품 추천 및 추천요청’으로, 제품에 대해 구매 및 사용 평이나 구매를 위한 추천을 요청하는 글이 많았다. 예를 들어, ‘식기 세척기는 어떤거 쓰세요?’ 등 추천 요청 글이 있고, ‘가구 사려고 발품 팔고 다녔어요.’ 등 제품 구매에 관련된 글, ‘원목 식탁 후기 공유해요.’ 등 상품 사용 평 등이 있었다.

[그림 Ⅳ-3-7] 육아카페 토픽 모델링 시각화 결과(토픽5)



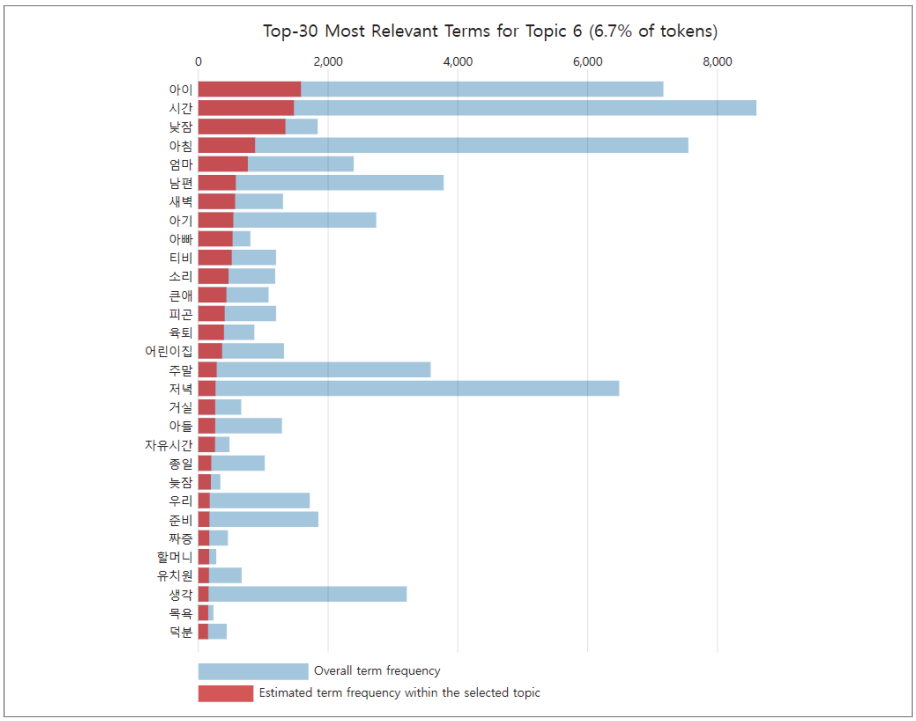
〈표 Ⅳ-3-7〉 Topic5 관련 원본문서

순서	제목	내용
1	식기세척기세제 어떤 거 쓰세요?	얼마 전에 식기세척기를 드디어 구입했어요~! 남편이랑 제가 먹은 그릇은 사은품으로 받은 일반 식기세척기 세제를 사용해서 써봤는데 너무 좋더라고요? 다만 아이 그릇을 닦을 때도 써야 될 거 같은데 ㅠㅠ식기세척기 세제 성분이 걱정이 되더라고요 다른 분들은 식기세척기 세제 어떤 거 쓰고 계세요?
2	비 뚫고 가구 발품 다녔어여...ㅠ	비 오는데.. 가구 매장 발품 팔고 왔어요! 일*, 한*, 알** 이렇게 다녀왔고요~ 1. 일* 먼저 일* 매장에서 책장 보러 갔는데 책장 말고 카페 장 보고 반해서 ㅎㅎ 쇼핑이 딴길로 좀 샜습니다 ;; 매장에서 컴퓨터로 직접 방 사이즈 맞춰 시뮬레이션을 돌려주시더라고요? 그래서 사이즈 잘 맞춰서 구매했어요! 학동 쪽에 있는 매장 갔었고 간 김에 진미 평양냉면도 먹고 왔지요 ㅋㅋㅋ 2. 한* 한* 디자인 파크 논현점 다녀왔는데, 여기 8층짜리 규모여서 가구 말고도 생활용품까지 다 팔더라고요. 어마어마.... 따로 상담받지 않고 구경하겠다 하고 돌아다녔는데 ㅠㅠ 한샘 디자인 파크 가실 분들은 미리 어떤 제품 볼 건지 정

순서	제목	내용
		해 놓고 가는 게 좋을 것 같았어요~ 저는 준비 없이 가기도 했고, 이미 예정에도 없던 카페 장을 구매 한지라 금방 나왔습니다. 3. 알** 논현가구거리에 있는 알** 매장 방문했어요! 저는 알**라는 브랜드를 몰랐는데 카페에서 추천받고 별 기대 없이 방문한 건데 ㅎㅎ 추천받고 갔어요. 매번 가구 종합 브랜드만 방문하다가, 소파만 있는 매장에 방문하니 눈 돌아가더라고요 ㅠ 이렇게 다양한 디자인의 소파가 존재 하구나에 1차 충격, 이렇게 다양하고 예쁜 컬러의 가죽도 있구나에 2차 충격. 원래 알로소는 담주에 다시 방문해보려고 했는데 지금 네이버에서 기획전 할인 때려서, 이 번 주 내로 제품 결정하고 결제하게요!!!! 사실 소파는 느긋하게 고민하고 싶었는데 기획전이 이번 주까지 라네여ㅠ
3	식기건조기 사려는데요 노지훈 건조기? 어떤가요	집에 주방이 좁아서 식기세척기 말고요 식기건조기 사려고 해요 노지훈 식기건조기 괜찮다는데 어떤가요? 클리쉐 65 말고 비슷한 걸로는 어떤 게 좋나요? 건조기는 한번도 안 써봐서요 아는 게 없어요 ㅠ
4	원목 식탁 튼튼하고 고급스러워요~	이번에 새롭게 구매한 원목 식탁 후기 공유해요~! 2년 전에 쇼파 구매하고, 너무 마음에 들어서 이번에 원목 식탁도 구매했어요~ 튼튼하고, 고급스럽네요! 역시 원목도 좋고, 시간이 지나도 유행이 안타는 스타일이라 딱 마음에 들어요. 컬러감이나 디자인, 마감 처리 그리고 원목까지 모두 마음에 듭니당!!
5	수박모양 커피카드 예쁘죠	스타*스 카드 인형과 선물로 왔는데 둘 다 너무 귀엽네요 카드인데 충전도 할 수 있고 5년 안에 쓰면 되나 봐요 전 커피 안 좋아하는데 한 잔 마시러 가야겠어요 빵순이라 빵으로 줬으면 좋았을 텐데 ㅎㅎ 망고 음료 파는 거 같던데 그런 거 마셔야겠네요
6	식기세척기 매립, 일반? 장단점, 금액	안녕하세요. 식기세척기에 관심이 많은 초보 줌마에요. ?근데 이사를 좀 종종 해야 될 것 같아서 매립은 좀 하려니 나중에 못 들고 간다는 생각에 매립을 못하겟더라구요 ㅠㅠ 그래서 질문 좀 드릴게요 1. 매립과 일반 차이가 있나요? 2. 렌탈과 매매 차이는 뭔가요? 3. 매립과 일반 금액이 달라요? 4. 4인 가족은 12인용이 좋을까요? 5. 제품 추천 좀 해주세요~ 넘 기능 많고 좋은 것보다 가성비로 기능 간단해도 잘 닦이고 고장 잘 안 나는 걸로요. 아니면 요런 걸 상담 문의할 만한 곳 알려주세용 식기 세척기 갖고 싶어용!!!
7	매트리스에서 벌레가 나왔어요ㅠㅠ	저 침대 깨끗하게 쓴다고 쓰는 편인데 이번에 얼룩이 생겨서 결국 돈을 써서 업체 불렀거든요 청소 후에 먼지 나온 걸 보여주셨는데, 와...먼지 나온 거 보이세요? 놀라 자빠질 뻔 했어요 ㅜㅜ 진드기 벌레도 보여주셨는데 벌레도 있더라구요 ㅠㅠ 예전에 다른 업체도 써봤는데 2~3일 자연 건조해야 한다고 하더라구요 혹시나 그 사이 또 곰팡이 생길까봐 찝찝하던데 이번에 맡긴 곳은 바로 건조차로 건조가 되어서 바로 사용 가능하더라구요? 세상 참 좋아졌어요...케어 받고 나니까 매트리스도 진짜 깨끗해지고 만족스러워요! 전 후 사진 한 번 비교해 보세요 ㅠㅠ 깨끗하네요ㅠㅠ

Topic6의 주제는 '아이와 엄마의 휴식'으로, 아기가 잘 때 혹은 아기가 장난감 놀이를 할 때 엄마가 쉬거나, 아기가 잠을 자면 엄마가 육퇴(육아퇴근, 즉, 하루 육아를 마침)를 하거나 밤에도 아이가 잠을 자다 깨면 엄마는 강제 기상을 할 수밖에 없는 상황을 담고 있다.

[그림 Ⅳ-3-8] 육아카페 토픽 모델링 시각화 결과(토픽6)



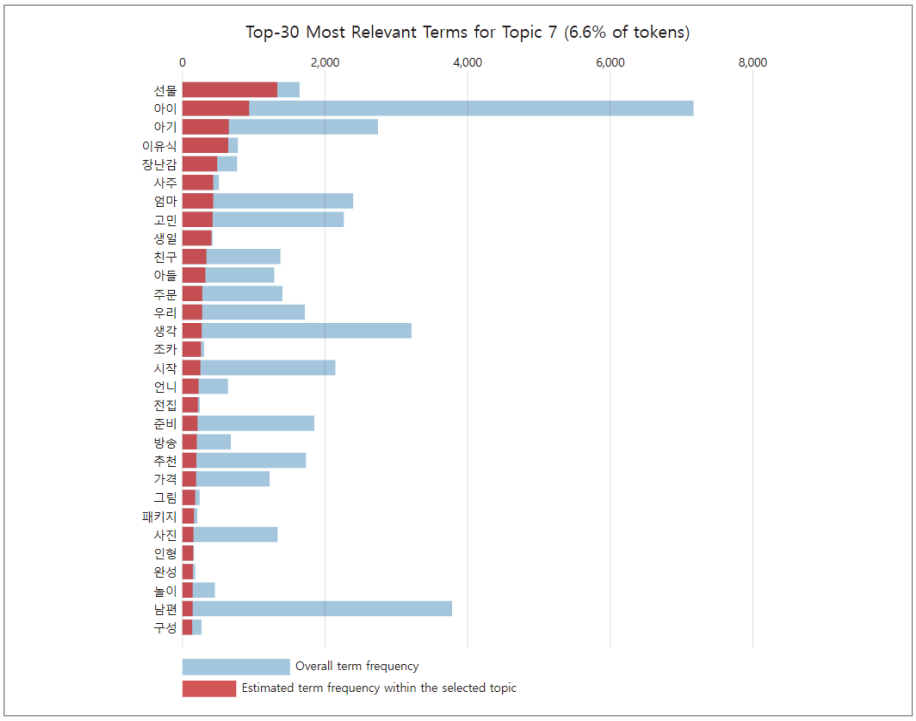
〈표 Ⅳ-3-8〉 Topic6 관련 원본문서

순서	제목	내용
1	아기 잠들었어요~	아기 잠들었어요~~거실에 아기가 놀던 아이큐베이비 장난감 있어서 정리하고 쉴려고요 아기가 좋아해서 놀아 줬는데 정리는 엄마 몫이에요 ㅋㅋ 놀이 해주면 이제는 혼자서도 잘 놀고 많이 컸어요 놀이 시간도 길어지고 재울 때도 수월해지고 크는 게 느껴져요 하루 종일 아기하고 놀아주고 재우고 했더니 피곤하지만 자유 시간은 즐기고 잘려고요
2	클레이놀이만 2시간~	플레이도우 국수 뽑는 장난감 예전에 사두었다가 요즘 클레이 놀이에 빠졌길래 꺼내줬더니 둘이서 2시간을 내리 노네요 ㅎㅎ 덕분에 저는 쉬구요~ 둘 찌가 약 먹는데 약 때문에 졸린지 힘이 없어요 ㅠ 얼른 저녁 먹이고 재우려구요~ 오늘도 이른 육퇴 기대해 봅니다!
3	정말 궁금해서 글올려봐요	저희 집 막둥이가 태어나서부터 저를 많이 힘들게 햇어요 거의 일주일에 4교대를 시켯죠 예를 들어 오늘 밤 9시에 잤다면 다음날은 밤 11시나 12시 그 다음날은 밤 1시에서 3시 이렇게 자는 시간이 일정하지 않고 계속 힘들게 햇어요 거의 17개월 동안.. 코로나도 있고 아직 어린 거 같아 얼집에 안 보낼려고 했는데 제가 미치겟더라구요 너무 힘들어서 어린이집에 보내면 활동량이 많아지니 잘 잘까 싶어 얼집에 보냈더니 밤 9시쯤 자서 새벽 1~2시에 일어나 4~5시에 자요 오늘은 6시 에 잤구 요매일 아침 9시나 10시에 깨워요 안 그럼 밤낮이 또 바뀔꺼 같아서 깨워서 얼집을 보내죠 얼집을 보내지 않으면 전

순서	제목	내용
		정말 잘 시간이 없어요 저희 집처럼 아기가 안자는 아기 있나요? 제 삶이 엉망 진창이예요 언제쯤 괜찮아질지..무섭네요
4	육퇴 했다고 생각했는데	육퇴했다고 생각하고 늦은 저녁 먹으려고 하는데 아기가 울어서 가보니 앉아서 울고 있더라고요 ㅠㅠ 안아서 토닥토닥 해주면서 백색 소음 내주고 하니 조용해져서 눕혔어용 ㅋㅋㅋ 아직 아기는 아기인가봐요~ 신생아 때의 백색 소음을 아직도 들려주니 말이죵ㅋㅋ
5	강제기상 이에요	아이가 자꾸 일어나래요 오늘 휴일 아닌가요 ㅎㅎ 엄마도 휴일엔 늦잠 자고 싶네요 아침 준비도 뭘해야 하나 꾸물거리고 있어요
6	몇번을 깼어요	자다가 새벽 두 시에 깼는데 그때까지 남편이 안 오고 애는 거실에서 자더라구요. 애를 안방 침대로 옮기고 다시 자려는데 잠이 안 와서 유튜브 듣다가 자 다 깨다 몇 번을 했네요.. 오늘 아빠 생신상 차려야 하는데 정신이 멍..
7	낮잠 시간	아이가 낮잠 자기 시작했는데~~아이의 낮잠 시간은 저의 자유시간인데~ 벌써 한 시간이나 지났네요...ㅡㅡ;; 뭐하느라 한 시간씩이나 보냈는지~~~~?

Topic7의 주제는 '아이 선물 및 이유식'으로, 돌 선물, 두돌 선물, 생일 선물 및 어린이날 선물 등 아이를 위한 선물에 대한 고민 등에 관한 내용이다. 또한, 아이에게 먹이고 있거나 먹일 이유식에 대한 내용도 포함하고 있다.

[그림 Ⅳ-3-9] 육아카페 토픽 모델링 시각화 결과(토픽7)

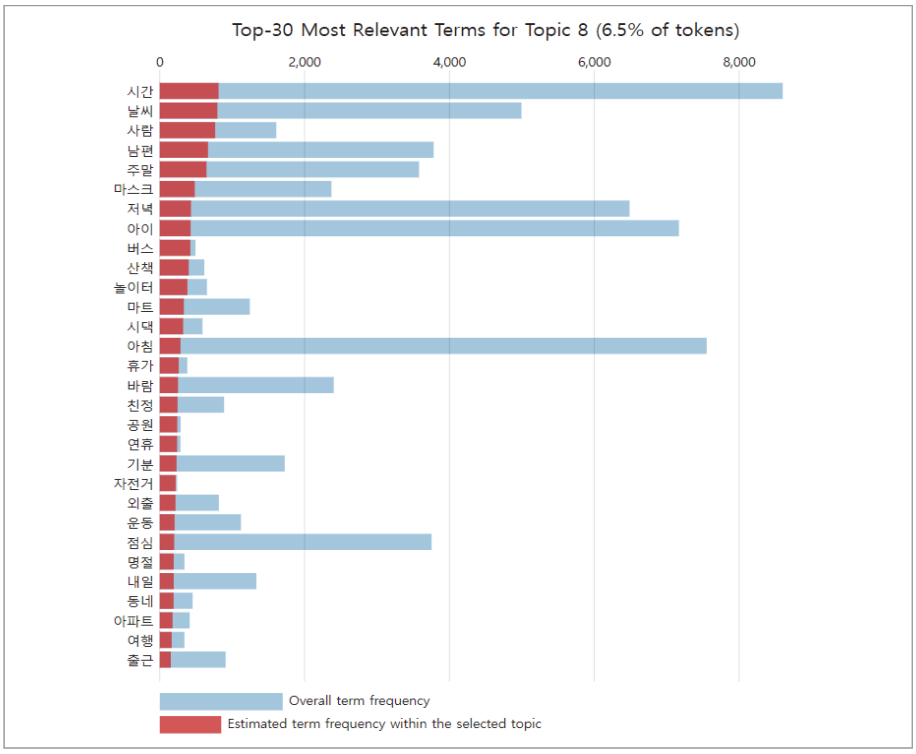


〈표 Ⅳ-3-9〉 Topic7 관련 원본문서

순서	제목	내용
1	돌 아기 선물로 전집 괜찮을까요?	돌 아기 선물로 시작 완성 풀 패키지 전집 구성이 좋은 거 같아서 사줄까 생각 중인데 이 책 사주신 분 있나요? 토끼펜 있고 시작 차트도 있는 거 풀 패키지로 선물 해줄까 고민 중인데 돌 아기 선물로 괜찮을까요?
2	동물 좋아하는 조카 선물로 사주려고요.	눈에 넣어도 안 아픈 첫 조카 이번에 두돌 되는데 선물 뭐 사줄까 하다가 동물 엄청 좋아하니 자연관찰 전집 사주려고요 금요일날 네이버 라이브 쇼핑에서 방송 한다고 해서 알림 설정 해놨네요~ 동물 좋아하는 아이니 좋아하겠죠? 조카 집에도 소전집 보드북 있는 것도 엄청 잘 보던데~~~ 블루래 빗 꺼 생생 자연관찰 사주면 정말 좋아하겠죠???
3	닌**	우리 둘째 아이 생일이네요 생일 선물은 닌** 주문했어요 언니도 생일 선물로 사줬더니 자기도 꼭 갖고 싶다고..ㅠㅠ 집에 두개가 꼭 필요할가 고민도 하고 가격도 비싸서 고민했는데 밥 잘 먹고 엄마 말 잘 듣기로 하고 샀네요 아직 어린 데 닌텐도 노래를 부르네요 ㅠㅠ
4	4월에 돌이라서 선물 사려고 해요	4월에 조카 돌이라서 선물 사려고요~ 뭐 사줘야 하나 고민 만땅 하다가 아기 책 사주려고 하는데 시작 완성 풀 패키지 홈쇼핑 방송 언제 하나요? 이 책 돌 선물로 괜찮을 거 같아서 토끼펜까지 있는 구성 사주려고 하거든요
5	첫 유산균 먹여요	"비**타 배배" 첫 유산균으로 샀어요 낙산균이라서 90도가 넘어도 균 손실 걱정 없고 안심하게 먹일 수 있는 유산균이라고 해요 분유에 타서 주셔도 되고 이유식에 넣어서 주셔도 되요 액상 유산균 살까 하다가 원료도 유기농이고 아연하고 비타민도 포함인 게 마음에 들어 선택했는데 유산균 먹고 황금변으로 봐서 유산균 선택 잘 했구나!!라고 생각 들어 요 신생아부터 36개월까지 먹일 수 있는 유산균인데 백일 전까지는 반포만 먹이면 된다고 하더라고요
6	이유식 배달 시키니 편하네요	이유식 배달 시켜서 먹이는 중인데 정말 편하고 좋네요 특히나 저가 먹이는 이유식은 무균이 유식인데 실온 보관 가능해서 외출할 때 따로 얼리거나 보냉 가방에 담아 나가지 않아도 되서는 편하게 먹일 수 있어 좋더라구요. 이유식 배달 시키는 거에 대해서 고민 많이 했었는데 괜한 고민을 했던 거 같아 요
7	아기 책 질문이요!!	어린이날도 다가오는데 첫 어린이날이라서...아기 선물 뭘로 해줄까~ 엄청 고민스럽네요~그러다가 장난감은 금방 잘 안 가지고 놀것 같고 책 선물이 의미도 있고 좋을 것 같아서 요~블루래빗 토이북 사주고 싶은데 홈쇼핑에서 사는 게 제일 혜택이 괜찮은 걸까요? 어디서 사야 할지...외출하기도 쉽지 않아서 요~그리고 몇 살까지 잘 보는지도 궁금하네요

Topic8의 주제는 ‘휴일과 외출’로, 휴일이나 여유시간을 활용하여 외출을 하거나 아이나 남편과 함께 시간을 보내는 내용, 날씨에 관한 내용도 포함되어 있다. 예를 들어, ‘남편과 드라이브를 나왔다.’, ‘주말은 아이들과 시간을 보낼 거다.’, ‘남편 퇴근하면 잠시 나갔다 올거다.’ 등 내용이 있었다.

[그림 Ⅳ-3-10] 육아카페 토픽 모델링 시각화 결과(토픽8)



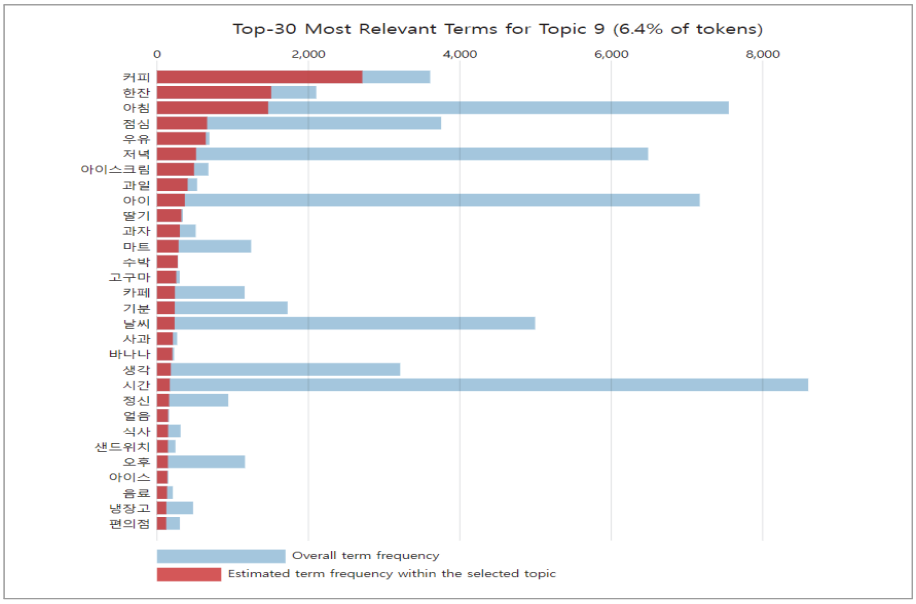
〈표 Ⅳ-3-10〉 Topic8 관련 원본문서

순서	제목	내용
1	집에서 따분하게 일상 보내다가 팔공산으로 드라이브 왔어요.	오늘 집에서 남편과 같이 점심에 떡국을 먹고 드라이브 나왔어요. 딸 아이가 집에서 자꾸 ㅜㅜ 심심하다고 떄떄거리네요,,, 하루 종일 휴대폰만 보는 딸아이 ,, 아기 상어를 무척 좋아하거든요 계속 휴대폰만 보는 거 같아 휴대폰을 뺏어버렷는데 ㅠㅠ 자꾸 심심하다 하고 남편도 따분해해서 팔공산으로 드라이브를 왔어요 오늘 낮에는 날씨가 참 좋네용 ㅎ 팔공산을 한 바퀴 정도 돌고 남편이 잠시 전산 작업을 해야 한다 해서 카페에 잠깐 들렸어요. 그래도 집에만 있는 것 보다는 이렇게 밖에 나오는 것도 정말 좋은 거 같아요.
2	후기하나쓰고	주말은 아이들과 시간을 보내기 위해 오늘 밀린 후기들 마무리하려 구요. 이제 하나 썼는데. 2시간 걸렸다는요.....ㅡㅜ
3	아침부터 할일이 많아서 ~	아침부터 집 정리하느라 여태 제대로 못했네요 ㅎㅎ 이제야 숨 한 번 돌립니다 날씨는 진짜 엄청 좋네요 ㅎ 미세먼지도 없구 ~ 막상 나가면 꽃가루 때문에 고생이겠지만 ㅠㅠ 이따가 남편 퇴근하면 잠시 나갔다 와야겠어요 ㅎㅎ 이런 날씨에 안 나가기 너무 아깝잖아요 ㅎㅎ
4	마스크~~!!	오늘 마트 갔는데 확실히 마스크를 많이 쓰고 계시네요~~ 아이들 마스크 쓰고 저도 마스크 쓰고서 마트 들어갔는데도 정말 어색하지 않고 마주치는 사람들과 일하는 분들이 마스크를 하고 계셨어요~~저희는 정말 시골이거든요, 미세먼지 잔뜩일 때도 마스크 쓰는 사람 별로 없었는데 이번에는 무지 심각하긴 하네요 ㅜ

순서	제목	내용
5	아무래도 연휴에 집콕모드~~	어머..연휴네요 이번 주 황금 연휴인데..전 알지도 못했어요.. 목요일도 쉬고..5월 1일도 쉬고.. 5월 5일은 어린이날 맞죠?? 왠지 뭐라도 해야 할 것 같은 생각은 뭘까요 이래서 제주도 여행 가나 봐요 ㅠ 마스크 끼고라도 나도 잠시 어디라고 다녀 올까 라는 생각 드네요.. 아직 아무런 계획도 없는데.. 그냥 있어야겠죠??? 혹시나 라는 생각도 들고... 저는 그냥 종식 될까까지 집콕 모드할까 싶어요 폐활량 떨어져서 마스크 끼고 돌아다니는 게 넘 힘드네요....
6	시댁엔 남편이...	답답한 남편.... 가장 위험한 사람이 어르신과 아이들이라는 데 정작 본인은 설마~란 생각 때문인지 추석 때 오라고.. 애들한테 과자사준다고... 애들은 마냥 좋아하고....;;; 추석이 고비라는데....좀....안 움직이고 싶은데...맨~~~~~~날 뉴스보는 남편은 가만히 있네요.....전 저희 친정에 이번 추석 때 안 간다고 말했는데.....시댁에는 남편이 말해야하는 거 아닌가요???!!
7	일하러 왔네요	벌써 불금이네요.. 오늘만 일하면 내일은 쉬는 날... 너무 좋네요.. 지금 왔는데요,, 벌써부터 집에 가 고 싶네요.. 내일 위해 오늘도 열심히 일해야겠어요,, 다들 오늘도 힘내세요

Topic9의 주제는 '커피와 휴식 및 디저트'로, 엄마들이 주로 커피를 마시며 휴식을 취하는 내용 및 가족들을 위한 과일 및 디저트에 관한 내용이다. 예를 들어, 아이가 낮잠을 자거나 아침시간, 혹은 잠시 시간이 날 때 커피를 즐겨 마시며 피로를 풀거나, 또한 가족들과 저녁식사 후 디저트를 먹거나 마트에서 과일을 샀다는 내용 등이 있었다.

[그림 Ⅳ-3-11] 육아카페 토픽 모델링 시각화 결과(토픽9)

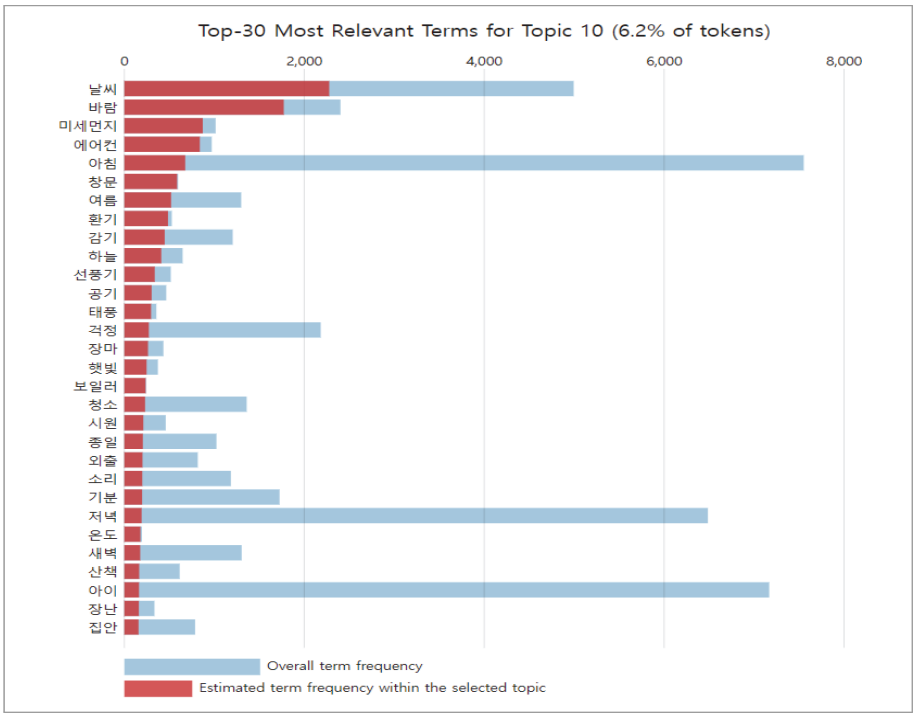


〈표 Ⅳ-3-11〉 Topic9 관련 원본문서

순서	제목	내용
1	시원한 아아 생각 납니다.	커피 한 잔 마실까 생각하고 있어요. 모닝 커피는 달달한 슈가 시럽 듬뿍 넣어서 마셨는데 오후 되니 나른 하고 잠도 오는 듯 하고 시원한 게 땅기네요 .얼음 가득 넘은 아아 한 잔 마셔야 겠어요 . 커피 한 잔 하셨나요 ?
2	사과~~	오전에 아침 먹이고 사과 배 깎아주었고 저녁 먹고 또 사과 두 개 잘랐는데 우리 아이들 맛나다며 잘 먹네요 저는 과일을 그리 잘 안 먹는데 신랑이랑 아이들은 완전 좋아해요ㅎ
3	아침커피는 다방커피	아침 먹고 커피로 다방 커피 헤이즐넛 마셔요 이제 뜨거운 믹스커피는 못 마시겠고 요런 컵커피 많이 사다가 냉장고에 쟁여놓고 있어요?? 오후에는 아이스 아메리카노 마시면서 후기 좀 많이 써야겠어요!! 생각보다 당첨이 많이 되서 쓸게 많네요 다들 화이팅입니다!!
4	카페인중독인가봐용 ㅜ	하 요즘 커피 위주로 마시는데 다른 건 생각이 안 나네요 ㅎㅎ 보이차도 많이 마셨는데 그냥 커피만 땡기네요 카페인 중독인가요? ㅜ 그래도 끊을 수가 없네요 커피는 ㅎㅎ 라이녹스 드 립주전자 덕에 커피 넘 맛있게 먹어요 ~ 오늘 날씨도 좋고 ~ 아이들은 낮잠 자구 ㅎㅎ 좋은 시간이예요
5	믹스커피	간단하게 샌드위치 만들어 점심 해결하고 이제 믹스커피 한 잔 마셔요~달달하니 따뜻하니 좋네요. 커피 다 마시면 이제 슬~슬~ 카페놀이도 하고 집 정리도 하고 해야겠어요~ 다들 맛점하세요 ~
6	오늘도 커피세잔	오전 오후 하루에 커피 두잔만 마시려는데 지키기 너무 힘들어요.. 아침과 저녁에 마시는 커피는 포기하기 힘드네요 오늘도 늦은 저녁 먹고 여덟 시 넘어 마지막 커피 마셨네요.
7	오렌지먹어요	저녁 먹고 후식으로 오늘은 오렌지요. 오렌지 새콤달콤 상콤하니 맛이 좋네요. 역시 과일은 먹어줘야 해요. 과일 좋아하는 우리 가족입니다.

Topic10의 주제는 ‘날씨’로, 태풍, 바람, 비, 무더위, 습기, 황사, 번개 등 날씨에 관련된 내용이었다. 예를 들어, ‘태풍이 온다.’, ‘비가 많이 온다.’ 등 날씨에 관련된 언급이 많았다.

[그림 Ⅳ-3-12] 육아카페 토픽 모델링 시각화 결과(토픽10)

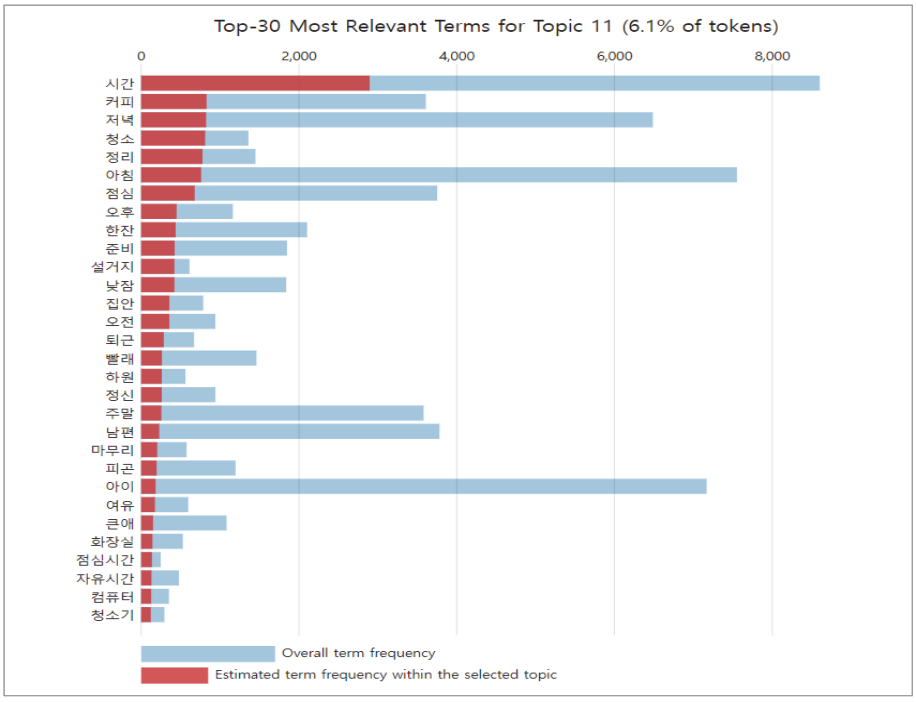


〈표 Ⅳ-3-12〉 Topic10 관련 원본문서

순서	제목	내용
1	다시 태풍이 온다니	다시 태풍이 온다고 하니까 이번 주말이 또 고비일 것 같아요 다들 마이삭 잘 넘기셨나요 아파트라 베란다 비들치는 문제가 있을 수 있겠으나 바람 소리에 좀 놀랬지요
2	바람이 어마무시했어요	아침 출근하는데 바람이 엄청나더라구요. 태풍 때문에 우산 뒤집힐 뻔..부산은 피해가 크던데.. 걱정되네요
3	화창한 휴일	어제는 비오더니 오늘은 날씨가 무척 좋네요 화창하니 오늘은 빨래 걱정 안해도 될 것 같아요 이 불 빨래 해야는데 말이죠 할 일이 많아지네요 ㅎㅎ
4	마스크 잘 쓰고 다시나요	여름에 마스크 끼려니 넘 답답하고 땀나서 찝찝하고 그러네요 ㅜㅜ 다들 마스크 잘 쓰고 다니시나 요 밖에 걸어 다닐 때는 마스크 쓰면 폭염이 따로 없어요 ㅜㅜ 이 제 여름 시작인데 7~8월에는 어떻게 쓰고 다닐지 걱정이네요
5	하늘에 구멍이 뚫렸나봐요	밖에 비가 내리는 게 아니고 쏟아져요 ㅎㅎㅎ 맞으면 아플 것 같아요 앞이 안 보이네요ㅠ
6	더워서 반팔	어제 오늘 환기를 못하니 집안 습도가 엄청나네요.해 뜨니 날씨도 더워서 반팔꺼내 입었어요~어제보다 황사가 덜하긴 한데 그래도 환기는 이따 하는 게 좋겠죠~
7	오늘 날씨 요란스러웠죠	오늘 천둥, 번개에 요란하게 비오고 쌀쌀하더니 지금은 날씨 왜 이리 좋은가요. 미세먼지도 싹 씻겨서 하늘 예쁜 거봐요. 내일도 날씨가 딱 지금처럼 시원하고 파란 하늘이었으면 좋겠네요

Topic11의 주제는 '시간 사용'으로, 엄마들이 시간을 어떻게 사용하는지 관한 내용이다. 예를 들어, 커피 마시는 시간, 청소하는 시간, 집안일 하는 시간, 아이 하원 시간, 점심시간 등 다양한 내용을 포함하고 있다.

[그림 Ⅳ-3-13] 육아카페 토픽 모델링 시각화 결과(토픽11)



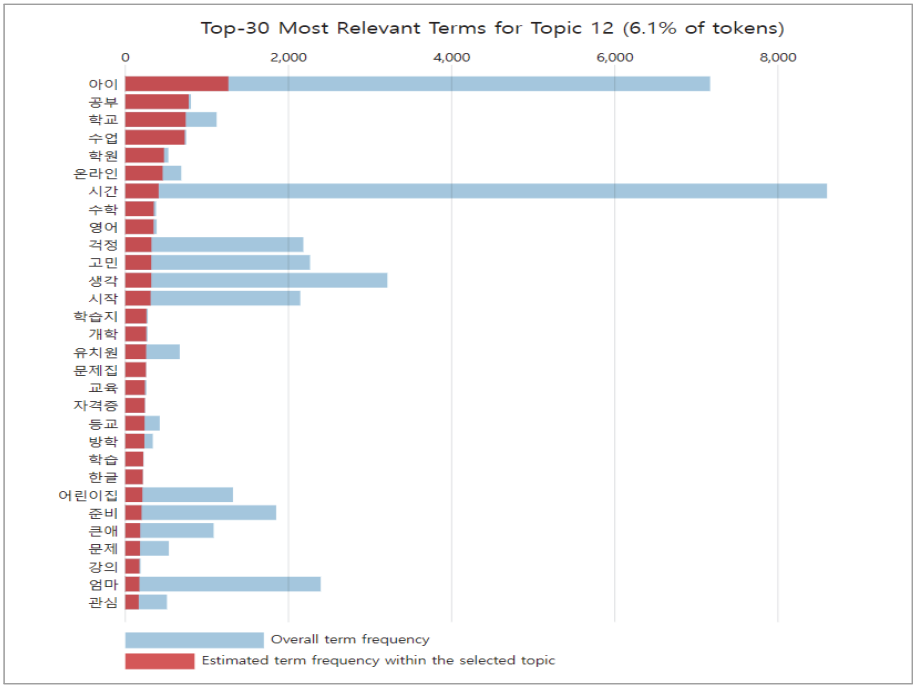
〈표 Ⅳ-3-13〉 Topic11 관련 원본문서

순서	제목	내용
1	오늘 대청소했어요.	애들 9시 등원시키고 쉴 틈 없이 대 청소했네요. 5년 전 이사 오면서 대충 정리해 두고 두아이 키우면 묵 혀뒀던 짐들 전부 꺼내서 버리고 정리하고 하니 하원 시간이였어요. ㅎㅎㅎㅎㅎ 온몸이 다 아픈데 디게 개운하네요. 이불빨래도 하고 진드기 패치도 붙여서 다 정리하고 지금 누워 있는데 너무 기분 좋아요. 침대 밑에 먼지가 너무 많아서 깜놀했습니다.ㄷㄷㄷ
2	커피마셔요	저녁 먹고 커피 한 잔 마시고 있어요. 설거지까지 다 마무리 하고 커피 마시며 드라마 보는데 이 시간이 너무너무 좋은 시간인데 윗층에서 쿵쿵 아이들이 신나게 뛰네요. 뛰다 말겠죠.. 참아 볼랍니다.ㅜㅜ
3	배고프네요 이시간에 ...	이 시간만 되면 왜 이렇게 배가 고픈 지ㅜㅜ만두 쪄 먹으려다가 그냥 뒷네요..휴 ㅠㅠ 배고파요.. 그냥 물 마시고 참는데 왤케 안참아지는지 ㅋㅋㅋㅋ
4	아침청소 완료	아침부터 부지런히 움직여 냉장고 정리하고 청소기까지 다 돌렸는데 이 시간~이 여유롭고 조용한 시간 너무 좋은 것 같아요~ 이제 커피 한 잔 하며 카페놀이 좀 해야겠어요~

순서	제목	내용
5	오후 시간도 훌쩍 지나가요	벌써 오후 시간도 4시가 다 되어 가네요 욕실 청소도 해야 하고 이리저리 집안일 남았는데 아이 하원 시간 다가오니 불안해지네요 ㅎㅎ 좀 다 하원 시키고 할까요
6	청소해야하는데 귀찮네요~!	아이 등원시키고 커피 한 잔 사와 서 컴퓨터에 앉아 있으니 아무것도 하기 싫으네요 ㅎㅎ 컴퓨터에 앉기 전에 청소부터 했어야 했나 봐요. 얼렁 청소부터 끝마치고 카페 좀 찬찬히 둘러봐야겠어요! 다들 좋은 시간 되세요!
7	벌써 오전이 다 지나가고 있네요	아직 아침밥도 못 먹었는뎁 빨리 먹어야겠어요 아침부터 너무 정신이 없다 보니...밥도 못 먹었네요 ㅠㅠ 애들 등 원시키다 보니 시간이 이렇게 훌쩍 가버리고...휴~ 이제 저만의 시간을 가져야 할 것 같아요 밥도 챙겨 먹고 커피 한 잔도 하고 그래야겠어요

Topic12의 주제는 '아이 및 엄마 공부'로, 엄마들의 아이 교육을 위한 고민이 담겨져 있고 엄마 본인의 미래 사회활동을 위한 공부에 관한 내용이다. 예를 들어, 아이 교육을 위해 학원, 학습지, 티비 프로그램 등을 활용하는 것을 고민하거나 정보를 공유하고 있다. 또 엄마 자신의 발전을 위해 인터넷 강의를 듣거나 자격증 공부를 한다는 내용들이 있었다.

[그림 Ⅳ-3-14] 육아카페 토픽 모델링 시각화 결과(토픽12)

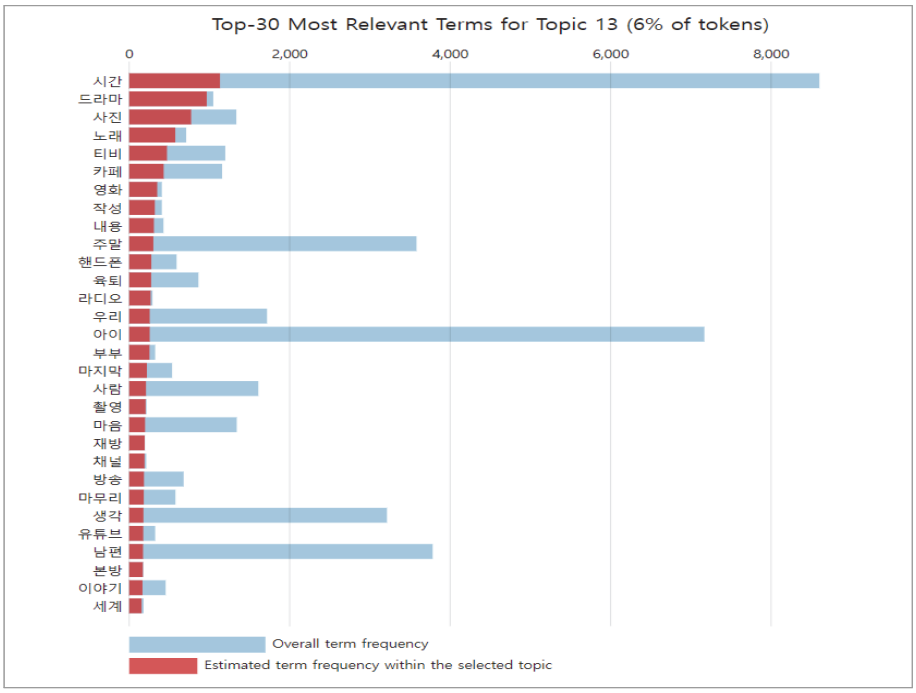


〈표 Ⅳ-3-14〉 Topic12 관련 원본문서

순서	제목	내용
1	영어학원 선생님놀이	공부방 차린 것처럼 아이랑 학원 놀이를 하고 있어요. 지금 영어 선생님으로 변신했는데 쉬는 시간이랍니다 ㅎㅎ영어 전집 하나 사서 한글 전집이랑 섞어서 공부시켜 주려 하는데 일단 첫날인데 아이가 잘 따라오는 거 같아요.이젠 매일 가능하면 오전 시간은 이렇게 한 시간이라도 놀아주려고요 ㅎㅎ 엄마는 영어 선생님, 할머니는 원장 선생님이라 말하는 아이의 표현이 귀엽습니다
2	인강이 아무래도 편하긴하네요 ㅎ	요새 인강으로 정리 수납 전문가 자격증 공부하고 있는데, 사실 인터넷 강의는 의지 부족으로 다 잘 들을 수 있을지 걱정하면서 시작했거든요 그런데 나름 장점이 있더라구요 아무 때나 편한 시간에 들을 수 있고 요새 핸드폰으로도 강의 다 들을 수 있고 하니 컴퓨터 앞에 앉아 있기 귀찮으면 핸드폰으로도 보고 ㅋㅋㅋ 그러면서 듣고 있어요 ㅎㅎ(강의 시간도 그렇게 길게 빠지지 않아서 그런가 ㅎㅎ) 빨리빨리 강의 다 듣고 마무리해야징~
3	키즈랜드에 있는 스콜라스틱으로 공부시키고 있어요	올*tv 키즈랜드에 키즈 컨텐츠가 6만 편이 넘어요 정말 다양한 교육 컨텐츠를 접할 수 있어서 다행이에요 ㅎㅎ 요즘은 영어는 거의 필수로 공부 가르치고 있다 보니까 저희 아이도 영어 공부로 고민이 많았어요 너무 스트레스 주는 것도 싫어서 자연스럽게 영어를 배울 수 있는 방법을 찾고 있었는데 스콜라스틱으로 흘려들으면서 자연스레 노출시킬 수 있으니까 좋네요~
4	수업 듣고 있어요	저희 지역 센터에서 무료 아이 교육 수업 무료로 있다고 해서 신청해서 듣고 있어요 어린이집 안 보냈으면 수업도 못 들을 뻔 했어요 ㅎㅎ
5	베이비 몬테소리 해야할까요??	아이는 잘 놀아야 한다고 생각했는데 막상 아이 낳으니까 다르네요 ㅜㅜ 베이비 몬테소리 할지 말지 너무 고민이 되어요 곧 12개월인데 지금 몇 날 며칠 그 고민만 계속 하네요 에휴 ㅎㅎ 프뢰벨이랑 몬테소리 둘 다 해보신 분들 계세요~~??
6	정리수납전문가 공부	정리수납 전문가 공부하고 있는데 우연히 티비 프로그램보다가 깜짝 놀랬어요 신박한 정리라고.. 우와 ~~ 연신 감탄하며 봤네요 정리 수납 잘하고 싶은 맘도 있었는데 비포 에프터를 티비로 보니까 확실하게 알겠더라구요 그래서 한국장학진흥원에서 정리 수납 전문가 자격증 공부를 하고 있다는 게 왠지 모를 뿌듯함이 들더라구요 끝까지 해서 저도 잘하고 싶네요
7	학습지	학습지 3월부터 시작하려고 했는데 결국 또 미뤘네요 ㅠㅠ 코로나 때문에 정말 아무것도 못하고 방학에 집안에만 있고 정말 심난 하네요

Topic13의 주제는 ‘엄마의 여가시간 및 아이 사진촬영’으로, 엄마들에게 하루에 주어지는 여가시간을 보내는 내용과 아이를 위한 사진 촬영에 관한 내용이다. 예를 들어, 엄마에게 주어지는 여가 시간은 육아퇴근(육퇴) 후, 아침에 눈뜬 시간, 살림퇴근(살퇴) 후 등이 있었다. 또한 이 시간을 활용하여 티비 드라마나 예능을 보거나 핸드폰을 하면서 시간을 보내거나, 돌잔치 같은 아이의 기념일에 사진을 찍는 내용도 포함되어 있었다.

[그림 Ⅳ-3-15] 육아카페 토픽 모델링 시각화 결과(토픽13)

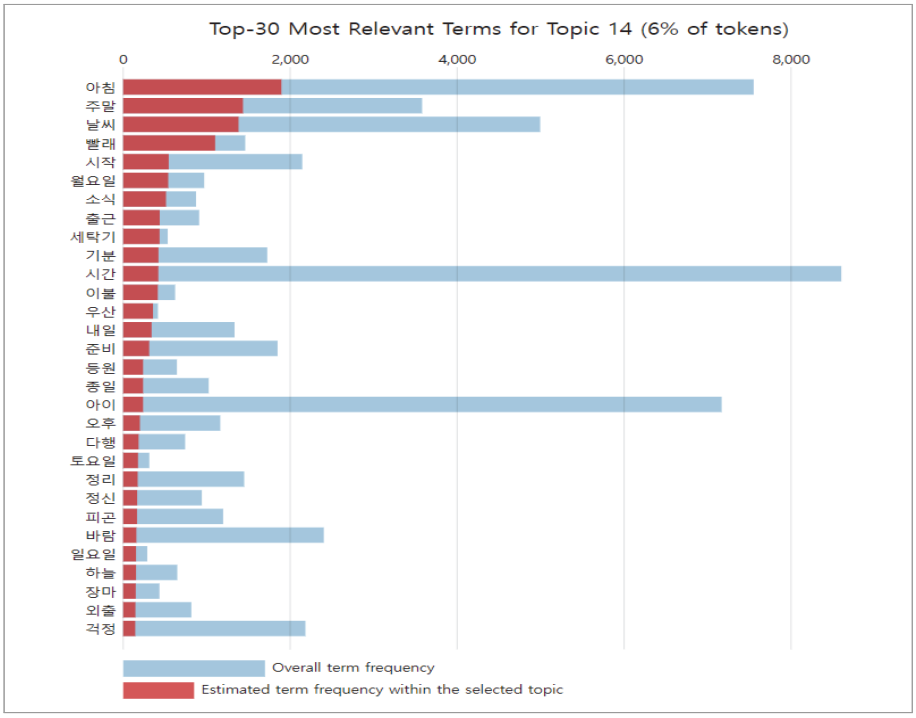


〈표 Ⅳ-3-15〉 Topic13 관련 원본문서

순서	제목	내용
1	육퇴했답니당	드이어 육퇴했어요~~~더 일찍 육퇴할 줄 알았는뎁 제 계획과는 다르게 늦게 육퇴했네요 낮잠을 얼마 안 잔 애들인데 왜 이렇게 늦게 자는지 원! 아무튼 두아이를 재워서 속이 시원하네요 이제 저만의 시간입니다!ㅋㅋㅋㅋㅋ과자나 먹으면서 자유를 만끽해야겠어요
2	본아트 촬영 후기〈**더하기강동〉	**더하기 강동점에서 본 아트 촬영 했어요!! 촬영 끝나고 나서 서비스로 더 많이 찍어주시고 감성 더하기 스탭 분들 너무 너무 친절하세요!! **더하기 강동점에 대한 정보가 없어서 궁금했는데 제가 직접 가보니 스튜디오도 깔끔하고 마음에 들고 무엇보다 아가를 소중하게 다루어 주시는 모습에 감동받았어요
3	좀 일찍 눈을 떠서요	뭐하지 하면서 핸드폰 보다 보니 30분이 훅 지나갔네요 ㅠㅠㅠ에휴... 눈팅은 시간이 참 빨리 가요
4	또 티비	또 티비 봐요 이따 나가기로 했는데 나가기 전까지 누워서 티비봐야겠어요 ㅎ
5	티비보며 휴식~	육퇴하고 늦은 저녁 먹으며 티비 봤어용ㅋㅋㅋ 오랜만에 봤는데 노래 들으면서 힐링도 하고 좋은데 여행쏭 특집인 만큼 여행 가 고 싶은 마음이 넘 컸네요! 마지막에 100점 도전 찬스도 성공하고 재밌었어요
6	아기 200일 촬영했어요	코로나로 집콕 중인.. 예쁜 사진 많이 못 찍어주는 아쉬움에 스튜디오 다녀왔어요. 전문가의 손길이 다르긴 하네요 ㅎㅎ 덕분에 아기 사진 잘 건졌어요. 앞으로 아기가 커가는 과정 사진으로 계속 남길려고 해요.
7	살퇴 후 티비 봐요	오늘 한번 다녀왔습니다 마지막회더라구요 재미있었는데 아쉽네요 할 일 다 하고 드라마 기다려요 여유로운 시간 좋아요

Topic14의 주제는 '일상과 날씨'로, 엄마들이 하루하루 살아가는 일상과 동시에 날씨에 관련된 내용이 많이 포함되었다. 예를 들어서, '비가 오다 멈췄으니 빨리 마트를 다녀와야겠다.' '아이들이 아침을 먹고 있다. 그런데 날씨가 흐리다.', '날씨가 너무 좋다. 빨래를 해야겠다.' 등이 있었다.

[그림 Ⅳ-3-16] 육아카페 토픽 모델링 시각화 결과(토픽14)



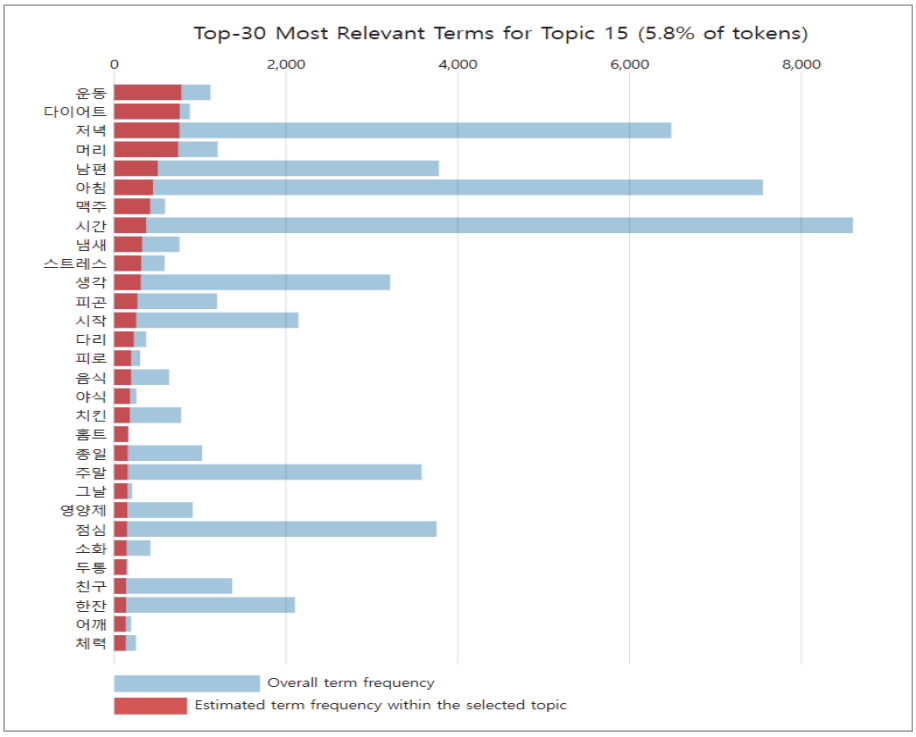
〈표 Ⅳ-3-16〉 Topic14 관련 원본문서

순서	제목	내용
1	즐거운 월요일	오늘은 날씨도 하늘도 햇살도 넘 좋네요~월요일 아침부터 이리 날씨가 좋으니 그냥 기분이 좋네요. 월요일이라 다운되는 하루가 될 꺼라 생각했는데 여유 있게 창 밖보며 즐겁게 월요일 보내보렵니다. 다들 즐거운 월요일 되셔요~
2	지금은 잠잠하네요	아침에는 비가 억수로 오구 바람도 미친 듯이 불더니 지금은 잠잠하네요~빨리 막둥이랑 마트 다녀와야겠어요
3	아침먹어요	오늘 아침은 왠일로 아이들이 늦게 일어나서 이제 아침 먹어요 비 소식이 있어서 날씨가 흐리네요
4	내일은 날씨 좋았으면 좋겠어요~	오늘은 날씨가 꿉꿉 하네요~ 비는 안 오는데 주말인데 날씨가 흐리니 주말 같지가 않네요~ 내일은 어린이집 등 원해야 하는데 포근했으면 좋겠어요~

순서	제목	내용
5	이불 빨기 좋은 날	오늘도 날씨가 너무 좋네요. 좋은 날씨 보니 이불이 너무 뽀송하게 잘 마를 것 같은 생각이 들더라고요. 슬슬 청소기 돌리고 이불 빨래 시작해 봐야겠어요~날씨 넘 좋네요~
6	아침	비도 온다 했는데 아직 흐리네요 아이 둘 빨리 아침 먹여야겠어요
7	어제 빨래하길 발한듯	주말 내내 비온 단 소리에 빨래를 어제 부랴부랴 해뒀거든요 아이 낮잠 이불도 있고 해서 밤새 바람 불고 하더니 다 말라서 정리했어요 ㅎㅎ

Topic15의 주제는 '운동 및 다이어트'로, 운동에 관한 내용 및 다이어트를 위한 식단 조절 등 내용을 포함하고 있다. 그 밖에 두피 건강 및 탈모에 관한 내용, 스트레스 풀기 위한 도구나 남편과 여가를 보낼 때 맥주 등 술을 마시는 내용 등이 있었다. 예를 들어, 운동 및 다이어트에 관한 내용으로 '남편이 헬스장을 다니며 운동을 하고 있다.', '다이어트를 위해 샐러드를 먹는다.', '주말마다 맥주를 먹어서 체중이 늘었다. 다이어트가 너무 어렵다.' 등이 있었다. 또, '머리카락이 너무 많이 빠진다.', '두통이 심하다.' 등 두피 건강 및 탈모에 관한 내용이나 '남편과 저녁에 맥주를 마셨다.' 등 내용이 있었다.

[그림 Ⅳ-3-17] 육아카페 토픽 모델링 시각화 결과(토픽15)



〈표 Ⅳ-3-17〉 Topic15 관련 원본문서

순서	제목	내용
1	남편이 다이어트 성공하면 가방사준다고 했어요!	제가 요새 뼈가 너무 아파서 병원 갔더니 살이 급하게 쪄서 뼈에 무리가 간거라고 하더라구요 ... 전 잘 몰랐는데 알게 모르게 회사 그만두고 집에서 쉬면서 살이 많이 쪘었나 봐요 ㅠ 남편한테 말하니 건강도 생각할 겸 10kg 빼면 원하는 가방 사준다고 하더라구요! 흥 제가 얼마짜리 고를 줄 알고 ㅎㅎ 절 너무 잘 다루는 남편이네요 ㅋ 그래도 마음이 고맙기도 하고 저도 몸이 너무 둔하고 아프긴 해서 슬슬 다이어트 시작해보려고 해요 우선은 무게 때문에 무리한 달리기 점핑 이런 건 안 좋다 해서 집 근처에서 다이어트 약 유명한 곳이라고 소문난 ㅂㅇㅌ에서 한 달치 처방받고 식이요법부터 시작하기로 했어요 이 방법으로 어느 정도 감량 되는 거 같으면 그때부터 pt랑 조깅이랑 같이 병행해서 다이어트 부스터 달아볼게요
2	이제 술 마셔도 진짜 예전같지가 않다는 ㅜㅜ	불금도 아닌데 마셨더니 속 쓰리네요. 요즘 좀 스트레스 받는 일도 있고 신랑한테 서운한 것도 있어서 마셨더니...좀 급하게 마셨나 봐요. 에효. 원래는 숙취 해소 젤리도 같이 먹으면서 천천히 마시는데... 어제 그냥 막 맥주를 들이 넣었더니.. 머리 두통이. 아침은 속이 안 좋아서 못 먹고 이제서야 국물 좀 먹고 굿** 구미 먹네요. 담에는 꼭 숙취 해소에 도움되는 것도 챙겨 먹음서 먹어야겠네요. 이제 술 마셔도 진짜 예전 같지가 않다는 ㅜㅜ
3	남편이 운동중인데	최근에 남편이 헬스로 운동을 하고 있는데 무리해서인지 상체 운동을 하다가 어깨 쪽에 담이 왔나 봐요 우선 한의원도 다녀오고 약도 먹고 있긴 한데 헬스를 계속 하고 싶어 하더라구요 아무래도 근육운동을 하면 무리가 될 것 같은데 ㅜㅜ 괜찮을지 걱정이네 요약 먹고 병원 병행하면서 해 두 상관 없을까요? 본인은 근육을 꼭 키우고 싶다네요.. ㅎㅎ
4	아침부터 고기를 먹자나여...? 다이어트 소화 둘다 잡았자나요..?	아침에 고기가 좋다고... 다이어터들은 다들 알고 계신 거 ㅋㅋ 어제 밤에 신랑이 족발을 시켰는데 대전대 천안 한방병원 한약 덕분인지 저는 한 입도 안 먹고 참아냈습니다 ...!! 스스로 대견 ㅋㅋ 그래서 아침에 일어나서 공복 요가 30분 가볍게 해주고 지금 먹으려고요 ㅋㅋ 얼마나 꿀 맛이게요~~~ 아침에 이렇게 먹어도 요즘 소화가 잘 되서 그런지 속도 편하고요.. 야식 안 먹고 아침에 단백질 섭취해줘서 그런지 체중도 조절 잘 되네요 ♡
5	몸무게	주말마다 맥주 먹어서 몸무게가 늘어서 다이어트 시작한 지 4일째네요 아침에는 정말 힘들더라구요 운동하는 걸 정말 싫어하는데 운동하고 하루에 한 끼 먹다가 어제부터는 저녁에 샐러드 먹고 힘드네요 이번에는 맘 먹고 운동 시작했는데 53키로에서 3키로 빼려고 생각 중인데 다이어트는 정말 힘드네요
6	이 시간 출출해서~~	이 시간 출출해서~~ 편의점 다녀왔어요. 맥주랑 편의점표 안주 먹으려구요. 신랑님한테 전에 같이 먹은 굿** 구미 젤리도 사오랬더니 왠일로 잘 찾아서 사왔어요. 뭐 시키면 잘못 사오는 신랑이라서 걱정 햇는데 ㅋㅋㅋ 제대로 사왔네요. 오늘은 이렇게 간단하게 심부름 잘 한 신랑님과 한 잔 하고 자야겠어요. 숙취젤리랑 같이 먹음 담날도 말짱하더라구요. 다들 굿밤 되세요.
7	집에서 두피마사지 할 수 있나요?	요즘 일 때문에도 너무 스트레스 받고 이래저래 짜증나네요.. 두피에서 열까지 나고...예전에 두피 마사지 받아보니깐 스트레스도 풀리고 시원하더라고요.. 헤어 토닉 아님 멘톨 컨디셔너 이런 걸로도 두피 마사지 가능하다던데 좋은 제품 추천 좀 ㅠㅠ 뭐가 없을까 용..ㅠㅠ

4. 소결

본 연구에는 맘카페 빅데이터 분석과정 통해 확인된 15개의 토픽을 내용에 따라 여가와 휴식, 육아 제품 공유, 일상과 날씨, 시간 사용, 공부 등 6개로 범주화하였다. 이러한 토픽범주에 따라 육아 관련 이슈를 찾을 수 있고 향후 육아정책연구 주제탐색에 도움이 될 것이다.

〈표 Ⅳ-4-1〉 육아카페 토픽 범주화

범주	여가와 휴식					육아제품 공유				건강		일상과 날씨		시간 사용	공부
토픽 번호	Topic 2	Topic 6	Topic 8	Topic 9	Topic 13	Topic 1	Topic 4	Topic 5	Topic 7	Topic 3	Topic 15	Topic 10	Topic 14	Topic 11	Topic 12
토픽 명	식사 및 간식	아이와 엄마의 휴식	휴일과 외출	커피와 휴식 및 디저트	엄마 여가 시간 및 아이 사진 촬영	피부 및 건강 제품 추천	제품 할인 및 온라인 구매	제품 추천 및 추천 요청	아이 선물 및 이유식	질병 및 치료	운동 및 다이어트	날씨	일상과 날씨	시간 사용	아이 및 엄마 공부
토픽 비율	7.83 %	6.74 %	6.48 %	6.35 %	5.98 %	7.97 %	7.31 %	7.15 %	6.65 %	7.31 %	5.78 %	6.24 %	5.98 %	6.12 %	6.10 %

‘여가와 휴식’에 범주는 토픽2, 토픽6, 토픽8, 토픽9와 토픽13이 포함되었고, 토픽2는 식사 및 간식, 토픽2는 아이와 엄마의 휴식, 토픽8은 휴일과 외출, 토픽9는 커피와 휴식 및 디저트, 토픽13은 엄마 여가시간 및 아이사진 촬영에 관련한 내용이었다. 이 범주에는 남편과 아이라는 단어가 자주 등장하면서 엄마들의 일상이 남편과 아이의 식사나 간식 등을 준비하는 것에 시간을 많이 소요한다는 것을 알 수 있다. 또한 엄마들은 하루 중 육아에서 해방되는 시간에 낮잠을 잠깐 자거나 혼자 커피를 마시거나, 드라마 등 TV를 즐겨보거나, 주말을 간절히 기다리고 주말에는 남편 혹은 아이들과 외출을 하기도 하였다. 반면, 범주 내 토픽내용들에는 아이가 낮잠을 잘 때나 육퇴후 쉴 수는 있으나 아이의 어린이집 등에서 돌아오면 쉬지 못하는 등 엄마들의 육아고충이 고스란히 반영되어 있었다. 동시에 육아뿐만 아니라 청소, 설거지 등 육아를 하면서 함께 살림을 꾸리는 엄마들의 고충도 알 수 있었다.

‘육아제품 공유’ 범주는 토픽1, 토픽4, 토픽5와 토픽7이 포함되었다. 토픽1을 보면 아이, 이기라는 단어가 다수 등장하고 피부, 유산균, 영양제, 물티슈, 분유,

콜라겐, 기저귀, 화장품 등의 단어가 등장하고 추천이라는 단어가 등장한다. 엄마들은 아이 및 본인의 피부 및 건강에 대한 관심이 많고 기저귀나 물티슈 등 아이를 위한 상품에 관심이 많았다. 이것을 구매하기 전에 맘카페를 통하여 추천을 받거나 이런 상품들을 사용하고 후기나 추천 글을 통해 정보를 공유하고 있음을 알 수 있다. 토픽4에는 할인, 주문, 쿠폰, 가격 등 단어가 등장하는 것을 보면, 이를 통하여 온라인 제품 정보 및 할인 정보를 공유하고 있음을 알 수 있다. 토픽7을 보면 선물, 아이, 아기라는 단어에서 알 수 있는 것처럼 아이를 위한 선물을 많이 하고 있음을 알 수 있고, 이유식이란 단어를 통하여 아이 이유식에 대한 고민을 이야기하고 관련 정보공유가 활발히 이뤄지고 있음을 알 수 있었다.

'건강' 범주에는 토픽3과 토픽15를 포함하였다. 토픽3은 질병 및 치료, 토픽15는 운동 및 다이어트에 관한 내용으로 '병원', '감기', '엄마', '걱정', '아이', '치료' 등으로 코로나19, 감기 등에 대해 엄마들이 많이 걱정하고 있으며, 질병 및 치료 관련 정보를 많이 공유하고 있는 것을 알 수 있다. 동시에 출산과 나이가 들어가면서 생기는 엄마들의 신체 및 몸매변화 등에 대한 스트레스 등으로 운동 및 다이어트 관련 내용들도 많이 있었다. 토픽명에는 포함되지 않았지만 '맥주', '스트레스', '피곤', '피로' 등 단어들은 엄마들이 육아와 관련하여 피로감과 스트레스를 많이 받고 있다는 것을 파악할 수 있다.

'일상과 날씨' 범주에는 토픽10와 토픽14를 포함하고 있다. 맘카페 특성상 엄마들의 일상적인 대화 속에서 날씨 관련 내용이 상당하며, '비가 온다', '너무 춥다' 등 단순한 내용과 '날씨가 좋으니 빨래를 해야겠다' 등 날씨와 관련한 일상생활을 수 있다.

'시간 사용' 범주에는 토픽11이 포함되어 있다. 엄마들이 육아를 하면서 시간을 어떻게 사용하고 있는지를 포함하는 내용들이 많고, '시간', '커피', '저녁', '청소', '정리', '아침', '점심', '오후', '한잔', '준비', '낮잠' 등의 단어가 많이 사용된 것으로 나타났다.

마지막 범주인 '공부'는 토픽12가 포함되어 있다. 토픽12에는 '아이', '공부', '학교', '수업', '학원', '온라인', '시간', '수학', '영어' 등 아이의 공부를 위한 학원, 학습지 등의 정보공유가 많았고, 동시에 엄마 본인들이 육아를 하면서 자격증 취득 등을 위한 공부와 관련된 내용이 많이 있었다.

V

학술논문 빅데이터 활용을 통한 육아정책 연구동향 분석

01 분석 필요성 및 목적

02 분석 내용 및 방법

03 분석 결과

04 소결

Ⅴ. 학술논문 빅데이터 활용을 통한 육아정책 연구동향 분석

1. 분석 필요성 및 목적

본 연구의 목적은 빅데이터 분석을 활용하여 육아정책과 관련한 다양한 이슈들을 탐색하고 새로운 내용이나 현안 모색에 대한 가능성을 탐색하는데 있다. 이러한 목적 아래 본 장에서는 학술연구 동향 파악을 통해 육아정책 관련 분야 연구자들의 주요 관심이 어디에 있는지를 파악하고, 향후 연구주제 등에 시사점을 도출하고자 하였다.

2. 분석 내용 및 방법

가. 분석내용

본 장에서는 육아 관련 분야 전문가들의 주요한 관심을 알아보기 위하여 KCI 웹사이트(https://www.kci.go.kr/kciportal/main.kci)에서 검색한 최근 5년의 논문 초록을 분석하였다. 자료 수집을 위해 임신, 출산, 어린이집, 유치원, 육아, 돌봄, 보육, 유아, 영아, 아동의 10개 키워드를 선정하였으며, 2017년 1월 1일부터 2021년 11월 3일까지 10개 키워드로 검색했을 때 확인되는 논문들을 분석대상으로 포함하였다.

웹크롤링 결과, 총 30,857개의 논문이 확인되었고 이 중 중복하여 검색된 학술지를 제거한 후 키워드별 논문 27,544개의 논문을 분석에 활용하였다(〈표 Ⅴ-2-1〉 참조).

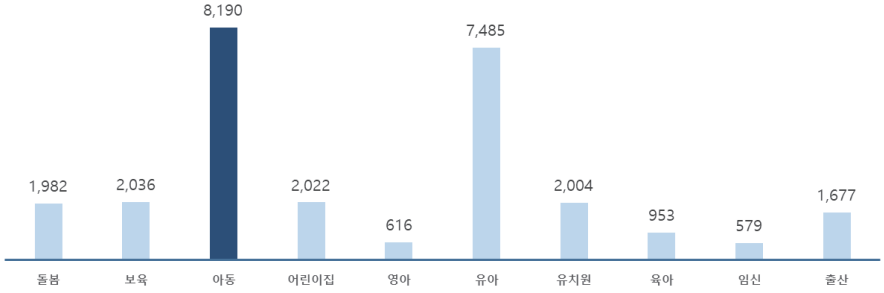
〈표 Ⅴ-2-1〉 키워드 별 문서 수

단위: 개

키워드	국내학술논문					빈도
	2017	2018	2019	2020	2021	
돌봄	296	344	399	472	471	1,982
보육	403	449	439	404	341	2,036
아동	1,603	1,703	1,791	1,727	1,366	8,190
어린이집	421	428	433	433	307	2,022
영아	124	146	131	120	95	616
유아	1,553	1,613	1,612	1,580	1,127	7,485
유치원	450	461	421	406	266	2,004
육아	213	191	219	183	147	953
임신	119	132	125	116	87	579
출산	347	347	373	331	279	1,677
합계	5,529	5,814	5,943	5,772	4,486	27,544

[그림 Ⅴ-2-1] 키워드 별 문서 수

단위: 개



나. 분석방법

본 연구에서 수집된 자료는 형태소 분석 → 명사 추출 → 제외어(불용어), 유의어, 지정어 설정 → 수집된 자료에서 제외어 제거, 지정어와 유의어 변경 → 토픽모델링 → 결과해석 및 시각화의 과정을 거쳐 자료처리 및 분석을 실시하였다.

먼저, 데이터 클리닝 과정의 일종으로 분석에 필요한 명사들을 추출하기 위해 데이터 전처리를 실시하였다. 형태소 분석을 통해 명사를 추출하였고, 추출된 명사들 중 불용어, 유의어, 지정어 등을 설정하였다. 각각 살펴보면, 불용어는 연구에 필요하지 않은 단어로, 다수의 문서에서 빈번하게 제시되어도 문서의 주제나 해당

연구와는 크게 관련되지 않은 단어를 뜻한다(유예림, 2017). 이 연구에서는 '~월', '때문', '무엇', '반면', '에서' 등이 이에 해당한다. 또한, '임신부', '임신여성', '임산부'를 '임부'로 통일하는 등 동일하거나 유사한 의미로 사용되는 여러 단어(유의어)는 하나의 단어로 통일하였다. 한편, 형태소 분석 과정에서 전문용어나 복합용어 등이 분리되어 추출되지 않도록 단어를 지정하였다. 예를 들면 '개정 누리과정'이 '개정'과 '누리과정'의 두 개 단어로 인식되지 않고 '개정누리과정'이라는 하나의 의미로 사용될 수 있도록 지정하는 것이다.

데이터 전처리 후 키워드 빈도 분석을 실시하였다. 키워드 빈도 분석은 문서 안에서 자주 등장하는 단어를 추출하고 등장하는 빈도를 중요도와 연결지어 분석하는 방법이다(황명화, 2014). 본 연구에서는 TF와 TF-IDF를 산출하여 결과 해석에 활용하였다. TF(Term Frequency: 단어 빈도)는 문서 내 특정 단어가 얼마나 자주 등장하는가를 나타낸 것으로, 기본적으로 빈도가 높으면 중요도도 높은 것으로 해석한다. TF값은 모든 문서에 자주 등장하는 상투어를 제외할 수 있는 장치가 없으므로 이러한 한계를 일정 수준 해결하기 위하여 TF-IDF값을 함께 본다. TF-IDF(Term Frequency-Inverse Document Frequency: 단어 빈도-역문서 빈도)는 TF에 전체 문서와 관계된 단어의 중요도를 나타내는 IDF를 곱한 값으로, 특정 문서 내에서 단어 빈도가 높을수록, 전체 문서들 중에서 그 단어를 포함한 문서가 적을수록 큰 값을 갖는다(유예림 2017).

이후 가장 보편적으로 활용되는 LDA(Latent Dirichlet Allocation) 모형을 활용하여 토픽모델링을 실시하였다. 토픽모델링은 관찰할 수 있는 문서들을 활용하여 그 기저에 있는 의미 구조를 탐색하려는 확률 모형으로, 위계적 베이지안 분석(hierarchical Bayesian analysis)의 한 종류이다(Blei & Lafferty, 2009). 토픽모델링을 통해 문서별 토픽 확률과 토픽별 단어 확률을 뽑으면 대상 문서가 어떤 단어들로 구성되었는지를 파악할 수 있다. 한편, LDA모형을 활용한 토픽모델링에서 도출된 단어를 바탕으로 토픽명을 결정하는 것은 연구자의 몫이므로, 보다 신중하고 전문성 있는 접근이 필요하다.

3. 분석 결과

가. 키워드 빈도 분석

최근 5년간의 논문 초록 분석을 통해 도출한 상위 빈도 단어는 〈표 V-3-1〉과 같다. 논문 초록의 경우, TF와 TF-IDF 기준 상위 빈도 단어들이 대체적으로 유사하게 나타났다. 논문 초록 상위 단위로는 '유아'가 가장 빈도수가 많고, 다음으로 '교사', '아동', '프로그램', '관계', '어머니', '부모' 순으로 빈도수가 많았다. 특히, 단어 '유아'는 다음으로 빈도수가 많은 '교사'에 비해서 12,476개 정도 많아 차이가 매우 컸고, 이는 크롤링한 총 문서 27,544개 중 키워드 '유아'로 수집된 논문이 7,485개로 매우 비중이 큰 것이 반영된 것으로 보인다. 두 번째로 빈도수가 많은 단어 '교사'는 대부분의 검색 키워드(돌봄, 보육, 아동, 어린이집, 유아, 유치원 등)와 관련이 있어 고르게 분포되었을 것으로 보이고, 세 번째로 빈도수가 많은 단어 '아동'은 '아동'이란 키워드로 8,190개 논문이 수집되어 가장 많은 논문 수가 분석된 것이 반영된 것으로 보인다.

〈표 V-3-1〉 학술연구 상위 25개 단어

순위	단어	TF	단어	TF-IDF
1	유아	31,906	유아	613.93
2	교사	19,430	교사	406.37
3	아동	19,141	어머니	386.08
4	프로그램	13,510	아동	371.96
5	관계	11,839	보육교사	311.07
6	어머니	11,718	프로그램	284.34
7	부모	10,938	부모	282.32
8	인식	10,848	유아교사	277.81
9	교육	9,850	인식	248.68
10	어린이집	9,546	관계	238.97
11	사회적	9,419	사회적	229.62
12	경험	8,096	어린이집	228.58
13	자녀	8,003	자녀	223.61
14	보육교사	7,860	상호작용	221.34
15	사용	7,522	교육	220.29
16	활용	7,411	영아	211.11
17	활동	7,080	경험	206.91

순위	단어	TF	단어	TF-IDF
18	유아교사	6,966	활동	205.55
19	변화	6,837	놀이	197.62
20	상호작용	6,410	아버지	186.23
21	유치원	6,208	유치원	173.38
22	수준	6,159	청소년	173.37
23	개발	5,762	영유아	173.01
24	특성	5,598	행복감	172.17
25	놀이	5,390	돌봄	167.63

TF에 기준한 상위 빈도 100개의 단어를 활용하여 워드클라우드를 작성하면 [그림 V-3-1]과 같다.

[그림 V-3-1] 학술연구 상위 단어 워드클라우드

나. 토픽모델링 및 시각화

LDA 분석 결과에 기초하여 총 15개의 토픽을 도출하였다. 본 연구의 목적이 빅데이터 분석을 통해 육아정책의 현안과 과제를 모색하는 것에 있고 본 분석이 관련 연구자들의 주요 관심을 탐색하는 것에 있다. 즉 최대한 넓은 수준에서 다양한 관점을 탐색하고자 하므로 복잡도와 토픽 응집성을 고려하여 토픽의 개수를 추출하기보다 내용상 의미가 구분되는 지점을 찾아 토픽 수를 결정하였다.

토픽모델링의 경우 LDA 분석 결과에 기초하여 토픽명을 결정하게 되는데, 이때 토픽명은 연구진이 부여하였다. 예를 들어 토픽9는 ‘여성’, ‘출산’, ‘남성’, ‘임신’, ‘결혼’ 등의 단어로 구성되어 있으므로, 단어들의 의미와 유사성, 맥락을 고려하여 토픽명을 ‘출산과 육아’로 정하였다.

동일한 방식으로 1) 저출산 대응 및 돌봄정책, 2) 유아교사 연수, 3) 공동체 돌봄, 4) 다양성 교육, 5) 아동 놀이 및 예술, 6) 아동 장애 치료, 7) 유아 교과 교육, 8) 아동인권, 9) 출산과 육아, 10) 교사 직무 만족도, 11) 아동 정서, 12) 부모의 양육과 정서, 13) 유아 정서 발달, 14) 교사 및 아동 정서, 15) 영유아 신체 및 정서 발달의 15개의 토픽명을 부여하였다(〈표 V-3-2〉 참조).

눈에 띄는 특징 중 하나는 토픽 중 다수가 교사와 관련된 내용이라는 점이다. 다만, 보다 심도있는 접근을 위해서는 각 토픽의 관계성이나 유사성 등에 대한 추가적인 분석 및 해석이 필요하다. 본 연구에서는 토픽들을 시각화하여 이들 관계를 파악하고자 하였다.

〈표 V-3-2〉 학술연구 토픽 모델링 결과

구분		토픽														
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
토픽명		저출산 대응 및 돌봄 정책	유아교사 인식과 경험	공동체 돌봄	다문화 교육	아동 놀이 및 예술, 문학	장애아동 및 치료	다양한 유아 교육	아동인권	출산과 육아	교사 직무 만족도	아동 사회성	부모의 양육과 정서	유아 상호 작용과 사회적 관계	예비유아 교사 정서	영유아 보육
토픽비율(%)		11.5%	10.3%	8.6%	8.2%	7.1%	7.0%	6.9%	6.5%	5.9%	5.6%	5.4%	5.1%	4.7%	4.1%	3.1%
토픽별 상위 키워드	1	정책	교사	경험	교육	놀이	아동	유아	아동학대	여성	보육교사	아동	어머니	유아	프로그램	영아
	2	필요	인식	돌봄	다문화	공간	중재	활동	아동	출산	교사	관계	자녀	상호작용	개발	영유아
	3	돌봄	어린이집	의미	유아교육	아동	사용	프로그램	권리	남성	유아교사	부모	부모	교사	예비유아 교사	교사
	4	지원	유치원	관계	수업	그림책	장애	활용	관련	임신	어린이집	우울	아버지	문제행동	자기 효능감	상호작용
	5	노인	교육	공동체	인식	의미	능력	창의성	필요	결혼	관계	학교적응	양육	어머니	향상	어린이집
	6	저출산	경험	이해	예비유아 교사	작품	발달	개발	문제	자녀	인식	휴대폰	양육 스트레스	사회적	스트레스	식생활
	7	증가	지원	자신	학생	이야기	행동	그림책	법률	가정	직무만족	초등학교	양육행동	관계	공감	영아교사
	8	관련	유아교사	가족	교육과정	교육	프로그램	향상	보호	필요	소진	사용	양육 효능감	또래	미술치료	부모
	9	변화	유아	사회	경험	어린이	일반	교육	규정	육아	전문성	활용	부모교육	행동	관계	발달
	10	고령	운영	인간	필요	디자인	특성	신체활동	학대	임부	원장	공격성	관계	자기조절	변화	영아반
	11	제공	필요	사회적	관련	문학	향상	과학	발생	육아휴직	직무 스트레스	사회적	유아기	유치원	활용	민감
	12	문제	요구	형성	내용	중심	관련	증진	성범죄	경험	행복감	경험	우울	유능성	집단 프로그램	영양
	13	활용	실행	가치	교사	표현	자폐	유아교육	보장	증가	심리	수준	양육태도	유능감	증진	놀이

구분		토픽														
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	14	사회적	놀이	실천	학습	특성	증가	관련	낙태	정책	수준	아동기	가정	정서조절	사용	적응
	15	중심	유아교육	개신교	교과	내용	언어	인성교육	범죄	변화	사용	변화	스트레스	놀이성	상담	가정
	16	확대	학부모	자기	활용	활용	참여	수학적	법원	지원	조직몰입	학교	참여	기질	아동	사용
	17	아동	영유아	존재	개발	그림	활용	기초	예방	가족	교수 효능감	학대	가족	사용	대인관계	놀이교수 효능감
	18	특성	보육	변화	유아교사	이해	변화	사용	개정	모성	영유아 교사	양육태도	양육참여	적응	심리	의사소통
	19	방안	원장	개념	역량	사용	치료	음악	인정	남편	리더십	한국아동	문제행동	발달	참여	관계
	20	운영	교육과정	인식	대학	인식	의사소통	음악교육	사회적	이용	영유아	중독	사용	어린이집	자기돌봄	애착

[그림 V-3-2]는 Intertopic Distance Map으로 토픽 15개의 구조적 관계를 2차원 평면에 나타낸 것이다. 각 원 안의 숫자는 〈표 V-3-2〉에서 제시한 토픽의 번호에 해당하며 원의 크기는 전체문서에서 해당 토픽이 차지하는 비율을, 각 원의 거리는 토픽이 포함한 단어 내용의 유사도를 나타낸다. 즉, 원의 크기가 클수록 해당 토픽이 전체 문서에서 차지하는 비율이 높고, 원의 거리가 가까울수록 비슷한 내용이라고 해석한다.

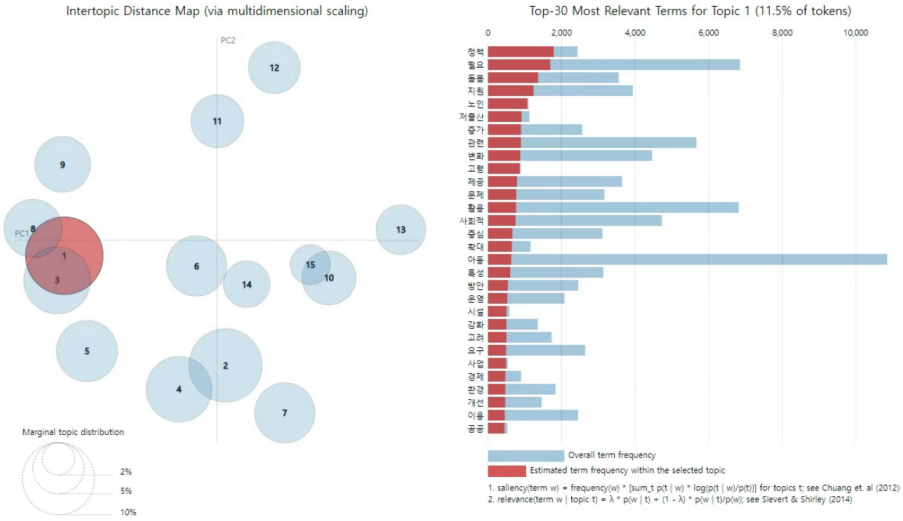
이를 바탕으로 살펴보면, 토픽1(저출산 대응 및 돌봄정책), 토픽3(공동체 돌봄), 토픽8(아동인권)이 밀접한 관계를 가지고 있고, 토픽 10(교사 직무 만족도)과 토픽 15(영유아 신체 및 정서 발달), 토픽2(유아교사)와 토픽4(다양성 교육)가 일정수준 유사성을 가지고 있음을 알 수 있다.

[그림 V-3-2] 학술연구 토픽 모델링 시각화 결과(전체)

전체 토픽의 구조뿐만 아니라 각 토픽에 대한 시각화 작업도 수행하였다. [그림 V-3-2]은 토픽1에 대한 시각화 결과로, 왼쪽에는 전체 토픽 구조 속에서 토픽1의 위치가 표시되었고 오른쪽에는 상위 단어 30개가 나열되었다.

빨간색으로 표기된 것은 토픽1에 대한 정보이고 파란색으로 표기된 것은 전체 토픽에 대한 정보이다. 오른쪽 막대그래프 해석 시 전체 문서 중 특정 토픽의 출현 빈도 비율이 높을수록 토픽에 대한 해당 단어의 기여도가 크다고 볼 수 있다. 즉, 전체 빈도와 특정 토픽 내 빈도수의 차이가 크지 않은 단어들은 해당 토픽에 집중적으로 출현한 것으로 해석이 가능하다(유예림, 2017). 토픽1의 시각화 결과에서 보면, '정책'의 경우 막대그래프에 빨간색으로 표기된 부분의 비중이 상당히 크므로 해당 단어는 토픽1에서 집중적으로 출현한 단어, 즉 토픽1에서만 많이 도출된 단어로 볼 수 있다.

[그림 V-3-3] 학술연구 토픽 모델링 시각화 결과(토픽1)



1) 토픽1: 저출산 대응 및 돌봄 정책 강화

토픽1은 '정책', '필요', '돌봄', '지원', '노인', '저출산' 등 키워드로 구성되어 있는 것으로 보아 저출산 대응 및 아동 돌봄 정책에 관한 주제로 이뤄졌음을 알 수 있다. 우리나라의 저출산 현상은 갈수록 심각해 짐에 따라 이와 관련하여 많은

연구가 진행되고 있고, 아동수당, 성평등, 육아휴직제도, 사회적 돌봄 등 국가정책 차원에서 문제를 해결하려는 연구가 많은 상황이다. 예를 들어, '사회의 지속발전을 위한 아동수당제도의 도입. 프랑스 법제의 시사점', '사회적 돌봄 부문으로의 남성참여 확대를 위한 시론적 연구: 유럽 국가들의 경험을 중심으로', '우리나라 저출산 대응정책 실효성에 대한 여성의 인식차이' 등이 있다.

2) 토픽2: 유아교사 인식과 경험

토픽2는 '교사', '인식', '어린이집', '유치원', '교육', '경험', '지원', '유아교사' 등 키워드로 구성되어 있는 것으로 보아 유아교사의 인식과 경험에 관한 주제로 이뤄졌음을 알 수 있다. 유아교사에 대한 다양한 연구 중 유아교사의 인식과 경험을 분석하는 연구가 다수 존재하고 있다. 예를 들어, '보육교사가 경험한 영유아의 놀이와 자율성 탐색', '유치원 교사가 인식하는 음악가 이미지 탐색', '영유아권리 침해 정도에 대한 영유아교사의 인식 개념도 분석', '보육교사가 지각하는 유아 친환경교육에 대한 인식 분석' 등이 있다.

3) 토픽3: 공동체 돌봄

토픽3은 '경험', '돌봄', '공동체', '사회' 등 키워드로 구성되어 있는 것으로 보아 공동체 돌봄과 기독교적인 돌봄에 관한 주제임을 알 수 있다.

예를 들어, '팬데믹 상황에서의 목회돌봄 패러다임 연구', '지역공동체 기반 돌봄 활동의 요인구성과 정책 방향: 노인과 아동의 돌봄공동체 사례를 중심으로' 등의 연구가 있다.

4) 토픽4: 다문화 교육

토픽4는 네 번째 토픽의 대표단어로는 '다문화', '교육', '수업', '예비유아교사' 등이 제시되었다. 다문화 수용성, 다문화 가족, 아동청소년 심리와 정서, 치료 등 '다문화'라는 주제로 다양한 연구들이 주를 이루고 있다. 예를 들어, '다문화가족 아동과 청소년에게 적용된 건강 관련 중재연구의 동향 : 간호교육에의 시사점', '아동의 다문화 수용성이 사회적 능력에 미치는 영향', '부모자녀관계와 다문화가족지

향성이 다문화가정 아동·청소년의 자존감 및 한국생활 적응에 미치는 영향' 등이 있었다.

5) 토픽5: 아동 놀이 및 예술, 문학

토픽5는 '놀이', '공간', '아동', '그림책', '작품', '문학' 등 키워드로 구성되어 놀이, 예술, 문학 등을 활용한 아동교육이 주제임을 알 수 있다. 아동 교육 방법으로 놀이를 통한 교육, 그림책, 동화책 등을 활용한 교육 등에 관한 연구가 진행되고 있다. 예를 들어, '아동문학을 통한 초등영어 지도방안 -주석기반 읽기활동을 중심으로-', '그림책을 활용한 유아 언어 교육 연구 동향 분석 -2007년 유치원 교육과정 개정 이후 논문-', '숲 자유놀이와 실내 자유선택활동에서의 유아 인지·사회적 놀이행동 비교', '그림책을 활용한 극놀이 프로그램 개발 및 효과' 등이 있었다.

6) 토픽6: 장애아동 및 치료

토픽6은 '아동', '장애', '프로그램', '자폐' 등 키워드로 구성되어 장애아동 및 치료에 관한 주제임을 알 수 있다. 정서행동 문제아동 혹은 자폐아 등 장애아동 치료에 관한 연구 혹은 발달장애인에 대한 움직임 연구 등에 관한 연구가 진행되고 있었다. 예를 들어, '발달장애인의 움직임 장애에 관한 고찰', '국내 학습장애아 급감 추세와 관련된 요인에 대한 소고', '퓨전 합창을 통한 발달장애학생의 공감 형성 과정 탐색', '가상현실 기반 작업치료프로그램이 학령기 지적장애 아동의 신체기능 및 학습능력에 미치는 영향' 등이 있다.

7) 토픽7: 다양한 유아교육

토픽7은 '유아', '활동', '프로그램', '창의성', '그림책', '신체활동', '인성교육', '수학적', '음악', '누리과정' 등 키워드로 구성되어 다양한 유아교육에 관한 주제임을 알 수 있다. 창의성 개발, 신체활동, 인성교육, 수학/음악 등 교육 관련 유아교육에 대한 다양한 연구가 존재한다. 예를 들어, '유아를 위한 극놀이 연계 과학교육 프로그램 개발 및 효과', ''Cre-Up 유아 창의프로그램'의 프로젝트 활동이 유아의 언어 및 도형 창의성 향상에 미치는 효과', '음악적 개념에 기초한 음악활동이 만

3세 유아의 음악 적성과 음악적 성향에 미치는 영향' 등이 있다.

8) 토픽8: 아동 인권

토픽8은 '아동학대', '아동', '권리', '법률', '학대', '성범죄', '낙태', '범죄', '법원' 등 키워드로 구성되어 아동 인권에 관한 주제임을 알 수 있다. 아동학대, 낙태 등 다양한 주제의 아동인권 관련 연구가 존재하고 아동인권보호를 관한 법률 혹은 학대피해 아동의 치료 등에 관한 연구가 있다. 예를 들어, '아동학대 범죄의 처벌 등에 관한 특례법 전후의 아동학대 인식 변화에 대한 고찰', '낙태죄의 허용 한계 사유와 허용 한계 시기에 대한 고찰', '아동학대에 대한 대응수단으로서의 형법 - 2020년의 아동복지법과 아동학대범죄 처벌 등에 관한 특례법의 개정을 중심으로 -', '놀이치료에 의뢰된 학대피해아동에 대한 인형심리진단평가 사례연구' 등이 있었다.

9) 토픽9: 출산과 육아

토픽9는 '여성', '출산', '남성', '임신', '결혼', '자녀', '육아', '임부(임산부)', '육아휴직' 등 키워드로 구성되어 출산과 육아에 관련한 다양한 주제를 포함하고 있음을 알 수 있다. 기혼여성, 미혼여성 등 임신과 출산 그리고 남편의 육아 등 육아와 관련된 다양한 연구가 존재하였다. 예를 들어, '일 대학병원에 의뢰된 결혼이주 임산부의 10년간 산전, 분만 특성 변화 연구', '임신기 부부의 성격, 부부친밀감, 가족지지가 태아애착에 미치는 영향', '출산 후 경과한 날에 따른 한국인 산모의 모유 단백체 분석', '미혼 여성의 임신능력 보존을 위한 난자 동결에 대한 인식' 등 연구가 있었다.

10) 토픽10: 교사 직무 만족도

토픽10은 '보육교사', '유아교사', '어린이집', '직무만족', '소진', '전문성', '직무스트레스', '행복감', '조직몰입' 등 키워드로 구성되어 교사의 직무만족 및 직무스트레스 등 교사의 정서에 관련한 주제임을 알 수 있다. 보육교사 특히 유아교사들의 직무 만족, 직무 스트레스, 심리적 소진, 행복감 등에 관한 연구가 존재하였

다. 예를 들어, ‘보육교사의 보육직무 스트레스 관리 상담 프로그램 개발 및 효과성 검증’, ‘보육교사의 공감능력 및 교수몰입이 행복감에 미치는 영향’, ‘원장의 여성 리더십, 어린이집의 협동적 조직문화, 보육교사의 감정노동 및 심리적 소진의 구조적 관계’, ‘유아교사의 직무스트레스와 스트레스 대처방식이 직무만족도에 미치는 영향’ 등의 연구가 있었다.

11) 토픽11: 아동 사회성

토픽11은 ‘아동’, ‘관계’, ‘부모’, ‘우울’, ‘학교적응’, ‘휴대폰(스마트폰)’, ‘초등학교’, ‘공격성’ 등 토픽으로 구성되어 아동의 사회적 관계와 정서에 관련한 주제임을 알 수 있다. 아동의 부모, 학교생활 등 아동의 사회적 관계와 우울, 공격성 등 정서에 관한 논문이나, 스마트폰 등 미디어 기기가 아동에게 미치는 영향에 관련된 논문도 다수 존재하였다. 예를 들어, ‘어머니의 행복감과 아동의 자존감 관계에서 아동의 학교적응의 매개효과’, ‘초등학교 고학년 학생들의 또래애착 분산요인과 성격 및 학교생활적응과의 관계: 사회관계모형의 적용’, ‘아동의 자아존중감, 주관적 행복감, 학교적응의 구조적 관계분석’ 등이 있었다.

12) 토픽12: 부모의 양육과 정서

토픽12는 ‘어머니’, ‘자녀’, ‘부모’, ‘아버지’, ‘양육’, ‘양육스트레스’, ‘양육행동’, ‘양육효능감’ 등 키워드로 구성되어 부모의 양육과 정서에 관한 주제임을 알 수 있다. 부모가 자녀를 양육할 때 느끼는 양육스트레스와 양육효능감 등 부모의 정서에 관한 연구가 존재하였다. 예를 들어, ‘어머니 양육스트레스와 유아 문제행동의 관계에서 어머니 우울의 종단적 매개효과’, ‘유아기 자녀를 둔 남편의 양육참여 및 맞벌이 부부의 양육스트레스가 결혼만족도에 미치는 영향: 종단적 자기-상대방 효과’ 등이 있었다.

또한 부모의 양육 태도가 유아의 정서에 미치는 영향에 관한 연구로, ‘부모양육태도가 유아의 회복탄력성과 자기조절력에 미치는 영향’, ‘어머니의 우울과 양육행동이 유아의 문제행동에 미치는 영향’ 등이 있었다.

13) 토픽13: 유아 상호작용과 사회적 관계

토픽13은 '유아', '상호작용', '교사', '문제행동', '어머니', '사회적', '관계', '또래', '행동', '자기조절', '유능성', '유능감', '정서조절' 등 키워드로 구성되어 부모나 교사와 맺는 사회적 관계와 정서에 관련한 주제임을 알 수 있다. 유아가 교사, 부모로부터 받은 영향과 교사, 부모 등과 맺는 사회적 관계가 유아에게 미치는 영향, 또래와의 관계 등에 관한 연구가 존재하였다. 예를 들어, '유아의 정서조절능력과 어머니의 감정코칭이 유아의 스트레스 대처행동에 미치는 영향', '유아의 정서지능, 셀프리더십, 놀이성이 또래놀이행동에 미치는 영향', '유아 정서조절력과 유치원 적응간의 관계에서 유아교사의 언어적 통제유형의 조절효과', '교사와 유아의 의도적 통제가 또래 놀이상호작용에 미치는 영향에서 교사-유아 갈등관계의 매개효과' 등이 있었다.

14) 토픽14: 예비유아교사 정서

토픽14는 '프로그램', '개발', '예비유아교사', '자기효능감', '스트레스', '공감', '미술치료' 등 키워드로 구성되어 주로 예비유아교사의 정서에 관한 주제임을 알 수 있다. 예비유아교사의 직무효능감, 스트레스 등 예비유아교사를 대상으로 한 다양한 연구가 존재하였는데, 예를 들어, '예비유아교사의 전공몰입이 직무수행효능감에 미치는 영향: 교육헌신을 매개효과로', '예비유아교사의 수학교과교육학지식과 수학교수효능감 간의 관계 내에서 수학정체성의 매개효과', '예비유아교사의 유머대처와 스트레스가 전공만족도에 미치는 영향' 등이 있었다.

15) 토픽15: 영유아 보육

토픽15는 '영아', '영유아', '교사', '상호작용', '어린이집' 등 키워드로 구성되어 영유아 보육에 관한 주제임을 알 수 있다. 영아 혹은 영유아의 교사, 부모와의 상호작용 등 영유아에 관한 다양한 연구가 존재하였다. 예를 들어, '보육교사의 영유아 권리인식과 임파워먼트가 교사-영유아 상호작용에 미치는 영향', '영아기 자녀를 둔 어머니들의 부모역할적응 프로그램에 관한 인식', '가을어린이집 만2세 영아반 실외놀이의 효율적 운영을 위한 실행연구' 등이 있었다.

4. 소결

최근 5년간의 논문 분석 결과를 토대로 토픽을 돌봄, 다문화, 장애아, 보육, 교사, 유아교육, 아동, 아동인권, 부모 등 9개 범주로 분류하였다.

〈표 Ⅴ-4-1〉 학술연구 토픽 범주화

범주	돌봄		다문화	장애아	보육	교사			유아교육		아동		아동인권	부모	
토픽 번호	Topic 1	Topic 3	Topic 4	Topic 6	Topic 15	Topic 2	Topic 10	Topic 14	Topic 7	Topic 13	Topic 5	Topic 11	Topic 8	Topic 9	Topic 12
토픽 명	저출산 대응 및 돌봄 정책	공동체 돌봄	다문화 교육	장애 아동 및 치료	영유아 보육	유아 교사 인식과 경험	교사 직무 만족도	예비 유아 교사 정서	다양한 유아 교육	유아 상호 작용 및 사회적 관계	아동 놀이 및 예술, 문학	아동 사회성	아동 인권	출산과 육아	부모의 양육과 정서
토픽 비율	11.5 %	8.6%	8.2%	7.0%	3.1%	10.3 %	5.6%	4.1%	6.9%	4.7%	7.1%	5.4%	6.5%	5.9%	5.1%

‘돌봄’ 범주에는 토픽1과 토픽3이 포함되어 있다. 토픽1은 저출산 대응 및 아동 돌봄 등 주제를 포함한 정책 관련 연구이고, 토픽3은 공동체, 종교 등을 포함한 민간 주체 ’돌봄‘에 관한 연구였으며, 특히 종교에 관한 언급은 기독교가 대부분이었다.

‘돌봄’에 등장하는 ‘아동’ 뿐만 아니라 ‘다문화’, ‘장애아’ 등 취약계층이나 사각지대에 놓일 가능성이 높은 집단 등 다양한 대상에 대한 연구가 진행되고 있었고 토픽 비중도 상위에 해당하였다.

영아, 유아, 아동 중에서는 유아에 관한 연구가 가장 활발하였고, 특히, 교육주체로서 예비교사를 포함한 유아교사에 관련 연구가 많이 이루어졌다. 세부적으로는 15개 토픽 중 유아 관련 토픽이 5개, 유아교사 관련 토픽이 3개로 나타났다.

아동 관련 토픽은 2개로 놀이, 예술, 문학 등을 통한 아동 교육, 아동의 학교적응이나 공격성 등 사회성과 감성에 관한 연구들도 일부 있었다. 또한 부모의 양육태도나 정서와 관련 연구도 주된 관심 중의 하나였다.

이외 유아와 유아교사 등과 관련한 연구 비중이 높은 반면, 영아 보육과 관련한 연구들은 상대적으로 적은 것으로 나타났다.

VI

정책제언

01 육아 안전·환경 종합 평가항목 개발과 체계 마련 필요

02 영유아의 부모를 위한 육아지원정책 강화

03 육아지원정책을 위한 향후 연구과제

Ⅵ. 정책제언

1. 육아 안전·환경 종합 평가항목 개발과 체계 마련 필요

그동안 영유아 보육·교육 관련 육아안전은 유치원과 어린이집을 중심으로 시설안전, 실내공기질, 급식안전, 놀이기구 안전, 교통안전 등 기관 실내외 안전을 중심으로 하고 있다. 향후에는 실제 유치원과 어린이집이 속해 있는 해당 지역을 대상으로 종합적인 육아 안전·환경 평가 등을 실시할 필요가 있다.

본 연구에서는 유치원과 어린이집 등 영유아 보육·교육기관 인근의 대기질 지수, 미세먼지 지수, 위험시설, 수질지수, 수질오염원 인근 거주인구 정보 등을 분석하였다. 유치원·어린이집 인근 대기 질(미세먼저 등)이 나빠졌을 경우 기관 내의 공기질도 동반하여 악화될 가능성이 높은 상황이다. 특히 수도권과 대도시, 공장 등 산업시설이 많은 곳에서 전반적으로 대기질이 나쁜 것으로 나타났다. 위험시설은 전국적으로 유치원은 평균 2.3개, 어린이집은 평균 2.4개의 위험시설이 반경 500미터 내에 있는 것으로 나타났고, 주유소 등을 포함한 위험시설이 최대 20개인 어린이집도 있었다. 영유아 보육·교육시설의 수소이온농도, 잔류염소, 탁도 수치는 모두 법적 기준은 모두 충족하고 있었다.

아이 하나를 키우기 위해서는 마을 하나가 필요하다는 아프리카 속담처럼 육아는 한 개인과 가정의 문제가 아닌 사회 공동체의 문제이다. 그러나 우리나라의 육아환경은 아이를 키우기에 많이 부족한 것이 현실이다. 육아환경은 사회·경제적 환경과 물리적 환경을 포함한다. 그럼에도 불구하고 사회·경제적 관점의 육아환경에 대한 연구에 비해 물리적 환경으로서의 육아환경에 대한 연구는 시설중심에 머물러 있다. 이에 물리적인 육아환경에 초점을 두되, 시설이외에도 아이를 키우기 좋은 육아환경은 무엇이며, 안전을 기반으로 위험시설은 어떤 것이 있고, 영유아가 생활하는 생활권 단위에서 필요하고 개선되어야 하는 안전한 육아환경은 어떻게 되어야 하는 지에 대한 고민을 가지고 육아환경 수준을 종합적으로 평가하는 평가항목 개발과 체계 마련이 필요하다.

2. 영유아의 부모를 위한 육아지원정책 강화

본 연구에서는 자녀 임신 전·후부터 하는 시기부터 자녀 출생 후 양육에 이르기까지 시기에 이르기까지 다양한 경험을 바탕으로 임신·출산·육아 관련 정보교환이 이루어지는 맘카페 게시글 분석을 통해 고 있으며, 결혼 생활에서 겪을 수 있는 다양한 고민을 공유하고 있기 때문에 좋은 정보원이 될 수 있다. 육아지원정책의 핵심 수요자인 엄마들의 관심들을 분석하였다. 이를 통해 향후에는 임신, 출산, 양육의 주체인 엄마아빠 등 양육주체들을 고려한 지원정책이 향후 좀 더 강화되고 보완될 필요가 있다.

가. 육아지원을 위한 부모 시간지원정책의 실질적 이용 강화

맘카페 빅데이터 분석을 통해 도출된 상위 키워드 100개를 보면 '남편'이란 단어는 상위 5번째에 위치할 만큼 엄마들에게 매우 중요한 존재로 볼 수 있다. 육퇴(육아퇴근), 퇴근, 휴식, 피곤 등의 키워드들은 엄마들의 자녀양육의 어려움을 고스란히 보여준다. 엄마들은 육퇴를 기다리고 남편의 퇴근을 기다리며 자녀양육의 피곤함을 잠깐의 휴식을 통해 해소하고자 한다고 볼 수 있다.

따라서 부모가 함께하는 돌봄이 가능하도록 시간지원정책의 실질적 이용을 강화할 필요가 있다. 법적으로 출산전후휴가나 배우자 출산휴가, 육아휴직, 육아기 근로시간 단축, 유연근무제 등 일하는 부모들에게 육아시간을 확보해 주는 다양한 시간지원정책들이 제도화되어 시행되고 있다(김동훈 외, 2020: 181). 그러나 부모들은 장시간 근로 등을 돌봄시간이 부족하다고 인식하고 있었고, 정부의 육아지원정책 중에서 현금이나 서비스지원 정책에 비해 시간지원 정책의 만족도가 5점 중 2.8점으로 낮게 나타났다(김동훈 외, 2020: 181). 특히 공공기관이나 대기업에 비해 각종 제도 이용이 상대적으로 어려운 중소기업이나 시간제 근로자 등을 포함한 비정규 근로자, 한부모나 장애아 등 취약계층 가정의 부모들의 실질적 이용을 강화할 수 있는 방안 마련이 필요하다. 예를 들어, 비정규 근로자 육아휴직 사업주에 대한 지원을 확대하거나, 정부차원의 유급 육아휴가제 및 육아휴가수당제 실시, 한부모 가정이나 장애아동 부모에게 추가적인 자녀 돌봄 휴가 제공, 취약계층 가구에 대한

육아휴직 급여 적용 기준 다양화 등도 검토할 필요가 있다(김동훈, 2018: 181).

또한 육아를 지원하는 정부나 지방자치단체의 정책을 잘 활용할 수 있도록 홍보를 활성화 하거나 각종 육아지원 인력(돌봄서비스 등)의 질을 높여 신뢰도를 높이는 등의 노력이 필요하다.

나. 양육의 주체인 부모 대상 지원 프로그램 강화

현재 정책은 경제적, 인적, 시간적 지원 등 물질적인 부분에 초점을 맞추고 있으나 출산 및 양육을 경험한 엄마를 포함해 가정에 대한 심리적 정신적 지원도 필요하다. 출산과 육아로 피로감이나 경력단절 등 부정적 경험이 아니라 사회적으로 존중받고 지지받는 경험을 할 수 있는 심리적 지원과 사회시스템이 필요하다.

또한 부모들을 위한 건강 의료비 지원 및 건강 프로그램 개발, 정신건강 프로그램 등을 더욱 강화할 필요가 있다. 또한 엄마들이 힐링할 수 있고 휴식을 취할 수 있는 공간이나 교육, 프로그램들을 정책적으로 더욱 육성할 필요가 있다. 이외에도 엄마들의 자기발전이나 취미생활, 미래 직업을 갖기 위한 교육프로그램과 취업프로그램도 더욱 개발될 필요가 있다.

맘카페 빅데이터를 보면 엄마들은 육아와 살림을 하며 매우 피로함을 느끼고 주로 혼자 있는 시간에 커피나 텔레비전을 보는 것으로 나타났다. 동시에 출산과 육아로 인해 변화한 몸매나 건강 등을 염려하고 있는 것으로 나타났다.

현재 정책은 임신, 출산, 양육 등에 관하여 아이 중심인 지원이 중심을 이루고 있다. 당연히 육아정책에서 아이를 최우선으로 양질의 육아지원, 돌봄, 보육·교육서비스를 제공하는 것이 중요하겠으나, 시각을 변화시켜 임신, 출산, 양육의 주체인 엄마들을 고려한 지원정책이 향후 좀 더 강화되고 보완될 필요가 있다.

다. 보육 및 돌봄서비스의 질 강화

우선적으로 국공립 어린이집과 유치원을 지속적으로 늘리면서 국가의 공적책임을 강화할 필요가 있다. 특히 전반적 서비스 질 제고를 위해 적절한 교사 대 아동 수 등의 환경 개선, 안전한 실내·외 물적 환경, 지역사회 인프라 확보, 교사의 질 제고, 보육교사와 유치원교사 처우 개선, 프로그램 다양화 및 질 제고 등의 노력이

필요해 보인다.

엄마들은 맘카페에서 ‘아이’, ‘공부’, ‘학교’, ‘수업’, ‘학원’, ‘온라인’, ‘시간’, ‘수학’, ‘영어’ 등의 단어빈도들이 높게 나타나 학원이나 공부 등과 관련된 정보에 관심이 많았다. 이는 그만큼 사교육에 대한 부모들의 관심이 높다고도 볼 수 있다. 이는 그동안 영유아 무상 보육·교육이나 누리과정 도입 등을 통해 출발선 평등과 차별없는 보육, 교육, 돌봄서비스를 제공받도록 노력하여 왔음에도 불구하고, 부모들은 빅데이터 분석 결과들은 부모들이 항상 더 높은 서비스를 요구한다는 것을 보여준다고도 볼 수 있다. 이에 어린이집과 유치원으로 대변되는 영유아 보육·교육 서비스의 질을 지속적으로 제고해야 할 것이다.

3. 육아지원정책을 위한 향후 연구과제

가. 영아 및 부모 관련 연구 강화 필요

그동안 연구자들의 연구주제 관심을 보면 상대적으로 보육(영아)과 부모에 관련된 연구가 부족한 것으로 나타났다. 영아, 유아, 아동 중에서는 유아에 관한 연구가 가장 활발하였고, 특히, 교육주체로서 예비교사를 포함한 유아교사에 관련 연구가 많이 이루어졌다. 아동과 관련해서는 놀이, 예술, 문학 등을 통한 아동 교육, 아동의 학교적응이나 공격성 등 사회성과 감성에 관한 연구들도 일부 있었다. 또한 부모의 양육태도나 정서와 관련 연구도 주된 관심 중의 하나였다. 또한 유아와 유아교사 등과 관련한 연구 비중이 높은 반면, 영아 보육과 부모 관련 연구들은 상대적으로 적은 것으로 나타나 이에 대한 연구가 보완될 필요가 있다.

이외에도 국민의 삶과 비교하여 보았을 때 영유아, 아동의 교육과 안전, 아동의 주거 환경, 아동의 신체적 건강, 교사와 부모의 배움, 소득, 쉼 등에 관련된 연구도 추가적으로 더 필요해 보인다.

나. 저출생 대응을 위한 돌봄 지원 확대 연구

저출생 기조 속에서 일·가정 양립 정책이나, 돌봄에 대한 국가의 책임이 더욱 강조되는 시대적 흐름 속에서 ‘저출생 대응을 위한 돌봄 지원 확대’ 연구가 강화될

필요가 있다. 이는 가장 많은 비중을 차지하기도 하지만 여전히 저출생 문제가 해소될 기미가 보이지 않는다는 점에서 좀더 면밀한 연구가 필요하다.

다. 개정 누리과정에 대한 교사연수 및 협력 네트워크 지원 확대 연구

누리과정이 유아와 놀이중심으로 개정되면서 교사들에게 놀이 중심 교육과정에 다양한 경험자료의 공유와 연수 등 변화된 교육과정에 대한 연수와 공유가 필요해 보인다. 이에 '개정 누리과정에 대한 교사연수 및 협력 네트워크 지원 확대' 연구가 강조될 필요가 있다.

라. 교사 및 영유아 대상 문화다양성 교육 및 지원 강화 방안 연구

갈수록 전형적인 가족의 형태와 달리 다문화 가정 등 다양한 형태의 가정이 증가하고 있는 실정이다. 이에 대학의 교사양성 학과(전공)에서 학생들을 대상을 다문화 등 다양성에 대한 교육의 강화와 유치원 및 어린이집 현장에서 교사들이 유아들을 대상으로 다름을 존중하고 이해하도록 교육이 필요하다고 보았다. 이에 '교사 및 영유아 대상 문화다양성 교육 및 지원 강화 방안' 등의 연구가 강화될 필요가 있다.

참고문헌

강달천(2020). 데이터 3법 개정의 주요 내용과 전망. KISA 한국인터넷진흥원.

강창완, 조장식, 최승배(2018). R과 RStudio를 이용한 빅데이터 분석 입문. 경기: 자유아카데미.

교육과학기술부, 행정안전부, 지식경제부, 방송통신위원회, 국가과학기술위원회(2012). 스마트 국가 구현을 위한 빅데이터 마스터 플랜.

과학기술정보통신부 보도자료(2021.2.23.). 빅데이터 플랫폼이 데이터 댐을 더욱 풍성하게 만듭니다.

관계부처 합동(2018a). 공공데이터 혁신전략.

관계부처 합동(2018b). 데이터 산업 활성화 전략.

관계부처 합동(2019). 제3차('20~'22) 공공데이터 제공 및 이용활성화 기본계획.

관계부처 합동(2020).「한국판 뉴딜」종합계획 - 선도국가로 도약하는 대한민국으로 대전환 -.

관계부처 합동(2021). 제1차 데이터기반행정 활성화 기본계획 ('21년~'23년).

관계부처 합동 보도자료(2016.12.9.). 민·관 합동 빅데이터 컨트롤타워 출범! - 행자부·미래부·통계청 빅데이터 TF 구성 및 제 1차 회의 개최 -.

관계부처 합동 보도자료(2020. 7. 22.). 이제 빅데이터 분석으로 공동주택의 돌봄 수요를 미리 예측한다.

국토교통부 보도자료(2020. 3. 26.). 스마트시티 기술로 코로나19 확진자 동선 파악이 더 빠르고 더 정확해집니다.

국회입법조사처(2020. 9). 빅데이터 플랫폼의 운영 실태와 개선과제, 입법·정책 보고서.

권연하, 박세진, 이현숙(2021). 텍스트 마이닝 기법을 활용한 코로나19 발생 이후 교육격차의 쟁점 분석. 학습자중심교과교육연구, 21(6), 625-644.

권헌영(2020). 데이터기반행정법과 데이터정책의 과제. KISO저널 제40호, 법제동향.

김동훈, 양미선, 김문정(2020). 포용적 사회를 위한 육아형평성 제고 방안. 육아

정책연구소 연구보고 2020-08.

김영옥(2016). 빅데이터 관련연구의 현황과 한국 유아교육학의 과제. 유아교육연구, 36(6), 181-206.

박진아, 김동훈, 김태우(2020). KICCE 육아물가지수 연구 (Ⅲ). 육아정책연구소 연구보고 2020-20.

박창현(2019). 빅데이터 분석을 통한 4차 산업혁명시대 육아정책의 방향과 과제. 육아정책연구소 이슈페이퍼 2019-08.

박창현, 김나영, 김지수, 최민수, 김정현(2020). 통일대비 북한의 영유아 교육· 보육 현황 분석 연구. 육아정책연구소 연구보고 2020-13.

박하연, 이재림, 이강이(2021). 육아와 조부모: 온라인 카페 게시글 토픽모델 분석. 가정과삶의질연구, 39(2), 91-107.

방송통신위원회 보도자료(2012.6.21.). 방통위, 빅데이터 서비스 활성화 적극 나선다.

서울시(2013). 민관융합 빅데이터와 공공데이터. 서울특별시 정보기획단.

서울시 보도자료(2020.9.18.). 서울시, 1천만 건 응답소 민원 빅데이터 AI로 분석… 선제적 행정에 활용.

성지은, 박기량(2014). 빅데이터를 활용한 정책 사례 분석과 시사점. 과학기술정책, 24(2), 94-106.

송태민(2016). 소셜 빅데이터 기반 저출산 정책 수요분석. 통계연구, (2016 특별호), 72-95.

권연하, 박세진, 이현숙(2021). 텍스트 마이닝 기법을 활용한 코로나19 발생 이후 교육격차의 쟁점 분석. 학습자중심교과교육연구, 21(6), 625-644.

양미선, 최은영, 김강민, 한재희, 김영민, 신하은(2020). 영유아 교육· 보육 분야 사회갈등 진단과 대응 방안. 육아정책연구소 연구보고 2020-15.

유예림(2017). 빅데이터 분석 기법을 활용한 2015 개정 교육과정 정책에 대한 언론보도 분석. 서울대학교 대학원 박사학위논문.

이상윤(2014). 빅데이터법제에 관한 비교법적 연구. 한국법제연구원.

이승환, 김인철, 권성수, 윤익준, 조희진(2015). 지방자치단체 정책지원을 위한 데이터 분석 운영체계 연구. 대구대학교 산학협력단.

이연승(2017). 유아교육분야에서의 빅데이터 활용을 위한 기초 연구. 유아교육연구, 37(4), 585-610.

정도영, 김민창, 김재환(2018). 빅데이터 정책 추진 현황과 활용도 제고방안. 국회입법조사처 입법·정책보고서.

정옥경, 고온조, 김경근(2020). 빅데이터를 사용한 온라인 육아 커뮤니티 영유아 부모의 자녀양육 고민 분석: 딥러닝 기법의 적용. 교육문화연구, 26(5), 1021-1043.

정용찬(2013). 빅데이터. 서울: 커뮤니케이션북스.

정지선(2011). 신가치창출 엔진, 빅 데이터의 새로운 가능성과 대응전략. 한국정보화진흥원. IT & Future Strategy, 18.

주세민, 옥성환, 황경태(2019). 빅데이터 기반의 출산율 변동 예측. [NIA] 정보화정책.

채승병, 안신현, 전상인(2012). 빅데이터: 산업 지각변동의 진원. CEO Information, 851(2). 삼성경제연구소.

최봉(2014). 공공부문의 빅데이터 활용과 시사점. 서울경제, 2014년 5월호, 서울연구원

최봉, 윤종진(2018). 서울시 공공빅데이터 활성화방안. 서울연구원 정책과제연구보고서, 1-147.

최지은(2020). 빅데이터를 활용한 유아 웰빙 (well-being) 에 대한 키워드 네트워크 분석. 육아지원연구, 15(4), 5-26.

한국정보통신기술협회(2015). 데이터 생애주기 기반 빅데이터 도입 및 활용지침.

함유근, 채승병(2012). 빅데이터, 경영을 바꾸다. 삼성경제연구소.

행정안전부(2018). 정부의 빅데이터 추진체계 및 정책방향. 빅데이터 추진체계 현황 및 향후 과제 세미나자료집. 행정안전부 공공데이터정책과.

행정안전부 보도자료(2020. 12. 10.). 데이터를 기반으로 한 과학적 행정시대 본격화.

행정안전부, 한국정보화진흥원(2019). 2019년 공공 빅데이터 우수사례 경진대회 자료집.

황명화(2014). 여론 모니터링을 위한 비정형 빅데이터 시공간분석 방법론 연구. 국토연구원, 국토연 2014-10.

Blei, D. M., & Lafferty, J. D. (2009). Topic models. In Srivastava, A. N., & Sahami, M. (Eds.). (2009). *Text mining: Classification, clustering, and applications*(pp.71-93). CRC Press.

Gartner(2012). Gartner IT Glossary, Accessed Nov 15. http://www.gartner.com/it-glossary/big-data.

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. Washington, District of Columbia.

Pappas, I. O., Mikalef, P., Giannakos, M. N., Krogstie, J., & Lekakos, G. (2018). Big data and business analytics ecosystems: Paving the way towards digital transformation and sustainable societies. Information Systems and e-Business Management, 16(3), 479-491.

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.

【관련 법령】

개인정보보호위원회, 개인정보 보호법 [시행 2020. 8. 5.] [법률 제16930호, 2020. 2. 4., 일부개정]. https://www.law.go.kr/법령/개인정보보호법 (2021. 8. 2. 검색)

금융위원회, 신용정보의 이용 및 보호에 관한 법률 [시행 2021. 7. 21.] [법률 제18124호, 2021. 4. 20., 일부개정]. https://www.law.go.kr/법령/신용정보의이용및보호에관한법률 (2021. 8. 2. 검색)

방송통신위원회·과학기술정보통신부, 정보통신망 이용촉진 및 정보보호 등에 관한 법률 [시행 2020. 12. 10.] [법률 제17358호, 2020. 6. 9., 일부개정]. https://www.law.go.kr/법령/정보통신망이용촉진및정보보호등에관한법률 (2021. 8. 2. 검색)

행정안전부, 공공기관의 정보공개에 관한 법률 [시행 2021. 6. 23.] [법률 제17690호, 2020. 12. 22., 일부개정]. https://www.law.go.kr/법령/공

공기관의정보공개에관한법률 (2021. 8. 2. 검색)

행정안전부, 공공데이터의 제공 및 이용 활성화에 관한 법률 [시행 2020. 12. 10.] [법률 제17344호, 2020. 6. 9., 타법개정]. https://www.law.go.kr/법령/공공데이터의제공및이용활성화에관한법률 (2021. 8. 2. 검색)

행정안전부, 데이터기반행정 활성화에 관한 법률 [시행 2020. 12. 10.] [법률 제17370호, 2020. 6. 9., 제정]. https://www.law.go.kr/법령/데이터기반행정활성화에관한법률 (2021. 8. 2. 검색)

【인터넷 자료】

서울경제, https://www.sedaily.com/NewsVIew/1Z2KYZABCW, 2021년 8월 2일 인출.

어린이집정보공개포털(http://info.childcare.go.kr/info/pnis/introduction/SummaryNursery.jsp, 2021. 8.7. 인출)

에어코리아(https://www.airkorea.or.kr/web/khaiInfo?pMENU_NO=129, 2021.9.16. 인출)

에어코리아https://www.airkorea.or.kr/web/contents/contentView/?pMENU_NO=132&cntnts_no=6, 2021.9.16. 인출)

유치원알리미(https://e-childschoolinfo.moe.go.kr/introInfo/informList.do, 2021. 8.7. 인출)

통계청. 통계설명자료(https://meta.narastat.kr/metasvc/index.do?confmNo=331007&inputYear=2018, 2021. 8.7. 인출)

통계청. 통계설명자료(https://meta.narastat.kr/metasvc/index.do?confmNo=331009&inputYear=2020, 2021. 8.7. 인출)

통합데이터지도(https://www.bigdata-map.kr/intro/intro, 2021.9.14. 인출)

통합데이터지도(https://www.bigdata-map.kr/board/faq, 2021.9.14. 인출)

통합데이터지도(https://www.bigdata-map.kr/dashboard, 2021. 9.17 인출)

한국사회보장정보원(http://www.ssis.or.kr/lay1/S1T749C765/contents.do, 2021. 8.7. 인출)

한국사회보장정보원(http://www.ssis.or.kr/lay1/S1T753C773/contents.do,

2021. 8.7. 인출)

한국아동패널(https://panel.kicce.re.kr/panel/board/index.do?menu_idx=42&manage_idx=26, 2021. 8.7. 인출)

혁신적 포용국가(http://www.inclusivekorea.go.kr/)

환경 비즈니스 빅데이터 플랫폼(envbigdata)(https://www.bigdata-environment.kr/user/content.do?sq=60, 2021.9.16. 인출)

환경 비즈니스 빅데이터 플랫폼(envbigdata)(https://www.airkorea.or.kr/web/khaiInfo?pMENU_NO=129, 2021.9.16. 인출)

빅데이터에 기반한 육아정책 쟁점과 향후 과제

93330
9 791168 650121
ISBN 979-11-6865-012-1

육아정책연구소
Korea Institute of Child Care and Education