

소셜 빅데이터로 알아본 코로나19와 가족생활: 토픽모델 접근

COVID-19 and Korean Family Life on Social Media: A Topic Model Approach

박선영*, 이재림**

서울대학교 아동가족학과*, 서울대학교 아동가족학과 및 생활과학연구소**

Sunyoung Park(s2with@snu.ac.kr)*, Jaerim Lee(jrlee@snu.ac.kr)**

요약

본 연구의 목적은 코로나19 확산으로 가족생활에서 급격한 변화가 일어난 1차 확산기에 블로그와 온라인 카페에 게시된 소셜 빅데이터를 분석하여 키워드를 파악하고, 게시글에 잠재된 주요 토픽을 발견하는 것이다. 강화된 사회적 거리두기가 처음 시행되었던 2020년 2월 23일부터 4월 19일까지 네이버와 다음의 블로그 및 카페에 게시된 글 중 '코로나'와 '가족' 또는 '코로나'와 '가정'이 함께 언급된 문서 총 351,734건을 분석하였다. 수집된 데이터는 전처리를 거쳐 텍스트 마이닝 기법으로 분석하였다. TF-IDF 가중치 값을 바탕으로 상위 100개 단어를 살펴보았으며, 잠재디리클레할당 방식의 토픽모델 분석을 통해 총 22개 토픽을 도출하고 토픽명을 부여하였다. 연구결과, 코로나19가 가족의 일상생활에 미친 전방위적 영향이 나타났으며, 특히 식생활, 주거생활, 여가생활, 종교생활, 자녀돌봄, 자녀교육, 가족관계, 가족의례 등에서 변화가 두드러졌다. 더불어, 가족 관련 국내 문헌에서는 잘 논의되지 않던 건강공동체로서의 가족을 시사하는 토픽도 등장하였다.

■ 중심어 : | 코로나19 | 가족생활 | 소셜미디어 | 빅데이터 | 블로그 | 온라인카페 | 토픽모델링 |

Abstract

The purpose of this study was to explore what social media posts tell us about family life during the COVID-19 pandemic by examining the keywords and topics underlying posts on blogs and online forums. Our criteria for web crawling were (a) blog and forum posts on Naver and Daum, the top portal sites in Korea, (b) posts between February 23 and April 19, 2020, the period of the first heightened social distancing orders, and (c) inclusion of "COVID" and "family" or "COVID" and "home." We analyzed 351,734 posts using TF-IDF values and topic modeling based on latent Dirichlet allocation. We identified and named 22 topics including COVID-19 prevention, family infection, family health, dietary life and changes, religious life, stuck at home, postponed school year, family events, travel and vacations, concerns about family and friends, anxiety and stress, disaster and damage, COVID-19 warning text messages, family support policies, Shin-cheon-ji and Daegu. The results show that COVID-19 impacted various domains of family life including health, food, housing, religion, child care, education, rituals, and leisure as well as relationships and emotions.

■ keyword : | COVID-19 | Family Life | Big Data | Social Media | Blog | Online Forum | Topic Modeling |

I. 서론

코로나바이러스감염증-19(이하 코로나19)의 확산에 따른 강도 높은 사회적 거리두기 조치는 한국인의 가족 생활에 다양한 변화를 가져왔다[1][2]. 어린이집 및 유치원 휴원과 초·중·고등학교 개학연기에 따라 가사 및 양육의 어려움이 확대되었고, 재택근무 확대 등으로 가족과 보내는 시간이 증가하였다. 무급휴가, 임금삭감, 휴업 및 폐업, 실직 등은 가정경제의 어려움으로 이어졌으며[1], 홈쇼핑, 인터넷·모바일 쇼핑 이용 확대 등 소비패턴이 변화하였다. 외출과 다중이용시설 출입이 제한되면서 직접 요리를 하는 사람들이 증가하였으며, 가정간편식이나 배달음식, 포장음식의 소비도 늘어났다[3]. 집에서 가족과 함께 보내는 시간이 증가하면서 가족관계가 개선되고 유대가 강화된 가족도 있으나[4], 가족 간 갈등이 오히려 증가하였다고 호소하는 경우도 나타나고 있다[5]. 해외에서는 코로나19 이후 이혼율이 증가함에 따라 '코로나 이혼(Covidivorce)'이라는 신조어가 등장하기도 하였다[6].

이처럼 코로나19 이후 가족생활의 변화는 매우 다양한 영역에 걸쳐 일어나고 있어 통합적인 접근이 필요하다. 그러나, 보건의로 및 제약 분야, 산업 및 경제 분야, 교육 분야 등에 비해 코로나19와 가족생활에 관한 연구는 부족하다. 코로나19 이후 가족의 변화를 다룬 연구들은 가족돌봄[7][8], 가사 및 양육 분담[9][10], 성별 불평등[11][12], 가정폭력 및 아동학대[13-16] 등 특정 측면을 구체적으로 다루는 경향이 있다. 특히 국내 연구는 대부분 브리프, 이슈페이퍼 등의 형식으로 자료분석 없이 이루어지고 있어 제한적이다. 최근에는 코로나19가 가족에 미친 전반적 영향을 알아보기 위해 사회조사 방법을 활용한 연구[1]도 등장하였으나, 코로나19와 가족생활 관련 연구는 초동적인 수준에 그치고 있다는 점에서 가족생활의 다양한 측면을 고려하는 연구가 필요하다.

기존의 사회조사 연구는 인과관계를 분석할 수 있다는 장점이 있으나, 특정 변수를 선정하고 설문지를 개발하게 되므로 제한적일 수 있다. 코로나19 사태와 같이 유래를 찾기 어려운 상황에서의 가족생활을 연구할 때는 선행연구나 학문적 통찰을 토대로 한 사회조사도

필요하지만, 코로나19 시대를 살아가고 있는 일반인들의 글이나 이야기를 살펴보는 것이 도움이 된다. 그러나 인터뷰 등의 질적 연구방법은 많은 사람의 이야기를 듣기 어렵다는 점에서 제한적이다. 이에 최근에는 다양한 채널의 정형, 비정형 데이터를 포괄하는 소셜 빅데이터 연구방법이 새로운 대안으로 제시되고 있다[17].

소셜 빅데이터란 일반 대중이 블로그, 온라인 카페, 소셜 네트워크 서비스(SNS) 등의 온라인 채널에서 생성한 대규모 데이터로서 개인의 경험, 생각, 감정 등이 담겨 있는 자료이다[18]. 소셜 미디어는 흔히 SNS와 혼재되어 사용되곤 하지만 하나의 개별 미디어 서비스가 아니라 다양한 미디어 형태를 포괄하는 개념이다[19]. 소셜 빅데이터는 다양한 채널에서 수많은 사람이 만든 어넌 방대한 양의 데이터에 기반하여 가치 있는 지식을 추출할 수 있다는 점에서 유용하다[20].

코로나19의 경우 확산 추이에 따라 가족생활 등 사회 전반의 변화가 빠르게 이루어지는 경향이 있어 이를 신속히 반영하는 매체를 활용하는 것이 중요하다. 또한 예측이 불가능한 재난상황이라는 점에서 가족생활의 전방위적 측면을 포괄해야 한다. 이러한 연구를 위해서는 일상과 경험을 공유하고 상호작용하는 것이 주요 목적인 소셜 미디어 채널을 선택하는 것이 효과적이다. 다양한 소셜 미디어 채널 가운데 블로그와 온라인 카페는 기록과 정보 교류의 목적이 큰 매체[21]이다. 다양한 사람들의 경험, 생각, 감정, 의견이 생생하게 담겨 있으므로 가족생활과 같은 일상에 관한 연구의 풍부한 정보원이 된다. 또한 인스타그램, 페이스북, 트위터 등 SNS 사이트의 게시물에 비해 글의 길이가 길어서 토픽 모델링 등의 분석방법을 적용하기 적절하다는 특징도 있다.

본 연구의 목적은 블로그와 카페 게시글 형태의 소셜 빅데이터를 통해 코로나19 시기 가족생활을 설명하는 키워드와 토픽을 발견하는 것이다. 키워드는 특정 주제와 관련된 전반적 경향이나 양상을 담고 있으며, 토픽은 거시적 맥락 속에 숨겨진 중요한 구조나 의미를 담고 있다. 따라서 소셜 빅데이터를 활용하여 키워드 및 토픽을 발견할 경우, 코로나19 이후 가족생활의 전반적 양상과 구체적 측면을 동시에 알아볼 수 있을 것이다. 구체적인 연구문제는 다음과 같다.

[연구문제 1] 소셜 빅데이터에 나타난 코로나19와 가

족생활 관련 키워드는 무엇인가?

[연구문제 2] 소셜 빅데이터에 나타난 코로나19와 가족생활 관련 토픽은 무엇인가?

II. 선행연구 고찰

1. 코로나19와 가족생활 관련 연구동향

코로나19는 개인과 가족의 일상생활과 삶의 질에 전방위적인 영향을 미쳤다[22][23]. 학자들은 코로나19가 가족생활 등 삶의 근본을 변화시키고 지속적인 영향을 미치고 있다고 지적한다[24]. 코로나가 가족생활에 미친 영향을 논의한 Lee et al. (2020)은 한국인에게 수면공간에 불과했던 집이라는 공간이 코로나19 이후 양육과 교육, 가사노동, 소비, 여가, 종교활동 등 삶의 핵심이 되는 다양한 활동이 이루어지는 새롭고 복합적인 의미의 공간이 되었다고 하였다.

코로나19와 가족생활을 연결한 선행연구는 자녀돌봄, 가사노동, 가정폭력, 일과 가족 등 특정 주제에 초점을 맞추는 경향이 있다. 코로나19 시기 자녀돌봄에 대한 국내 연구에서는 부부가 양육을 분담하는 것이 중요하며, 돌봄지원 서비스가 필요함을 강조하였다[7][8]. 가사 및 양육 분담을 다룬 연구들은 코로나19 시기 여성의 역할과중과 그로 인한 고용상태 위협 등을 논의하였다[9-12]. 이외에 아동학대, 배우자폭력 등 코로나19 시기 가정폭력에 대해 논의한 연구도 있다. 이러한 선행연구는 코로나19로 인한 스트레스 상황이 폭력과 학대의 위험성을 높이며, 피해자들이 가해자인 파트너나 부모와 온종일 함께 시간을 보내면서 높은 수준의 위협에 노출된 반면 사회적 거리두기로 인해 도움요청 및 외부개입이 어려워짐으로써 집이 위험한 공간이 되었다고 하였다[15][16].

일과 가족의 측면에서는 코로나19 이후 재택근무, 유연근무, 강제휴가, 해고 등의 조치에 따라 근로의 양상이 변화하면서, 일과 가족생활의 양립에 중요한 영향을 미치고 있다[25][26]. 코로나19 상황에서도 경제활동과 자녀양육을 전담해야 하는 한부모의 어려움을 살펴본 연구도 있다[27]. 이외에는 코로나19로 인한 가족의 상실[28], 임신과 출산[29] 등에 대한 논의도 이루어지고

있다. 코로나19로 대면 서비스가 어려워진 상황에서 가족생활 관련 교육, 상담 등 비대면 서비스 전달체계에 대한 고민을 담은 연구도 등장하고 있다[30][31]. 이론적으로는 코로나19의 영향을 생애과정 관점에서 논의한 연구도 있다[32].

코로나19와 가족생활에 관한 이상과 같은 선행연구의 대부분은 초동적 수준에서 연구자의 통찰이나 국가 통계, 뉴스기사 등에 의지하는 경향이 있다. 사회조사 방법을 활용하여 코로나19가 가족생활에 미친 영향을 파악한 연구[1][33]도 등장하고 있으나, 코로나19와 가족 관련 연구는 시작단계에 불과하다. 코로나19 사태가 장기화되고 있는 상황에서 가족의 일상 전반에 일어난 변화 또한 막대할 것이다. 이에 코로나19 맥락에서의 가족생활에 대해 다양한 연구방법을 사용하여 살펴볼 필요가 있다.

2. 코로나19 관련 소셜 빅데이터 연구동향

전염병 재난은 삶의 거의 모든 분야에 영향을 미치므로[34] 코로나19가 가족생활에 가져온 변화를 분석할 때도 다양한 측면을 포괄적으로 살펴보는 연구가 필요하다. 기존의 사회조사 방법은 특정 측면의 변화를 구체적으로 파악할 수 있다는 장점이 있으나, 일상생활 전반의 변화를 빠르게 알아보기에는 어려움이 있다. 이에 본 연구는 코로나19가 가족생활에 미친 다양한 영향을 탐색하기 위해 대규모의 온라인 게시글을 분석하는 빅데이터 분석방법을 사용하고자 한다. 온라인에서 수집한 빅데이터는 생활의 모든 측면에서 생산되고 저장된 방대한 정보를 활용한다는 점에서 정보의 양과 질이 기존의 사회조사 자료와 차별된다[35].

코로나19와 관련하여 자연어 형태의 온라인 빅데이터 분석방법을 활용한 연구도 속속 등장하고 있다[36-44]. 해외에서는 중국의 코로나19 이후 뉴스 빅데이터를 수집하여 토픽을 추출한 연구[41], 중국의 주요 포털 사이트인 위챗과 웨이보의 게시물을 분석하여 코로나19가 농촌경제에 미친 영향과 농촌 지원 정책과 관련된 토픽을 추출한 연구[42]가 있다. 이외에도 트위터 게시글 중 코로나19와 외로움(loneliness)이 동시에 등장한 문서에서 토픽을 추출한 연구[43], 코로나19 이후의 관광산업에 관심을 갖고 트립어드바이저

(TripAdvisor)의 게시글을 분석하고 토픽을 추출한 연구[44]도 있다. 이상을 토대로 보면, 해외의 코로나19 관련 온라인 빅데이터 연구는 상대적으로 다양한 채널, 다양한 주제가 특징적이다.

반면, 현재까지 국내에서 이루어진 코로나19 관련 온라인 빅데이터 연구는 대부분 뉴스 등 언론보도를 수집, 분석하여 핵심 키워드 및 토픽을 파악하고 시기별 변화 동향을 알아보았다[36-38]. 트위터 데이터를 분석하여 키워드 및 토픽을 파악하고 토픽별 정보 확산 네트워크 및 속도를 비교, 분석한 연구[39]와 코로나19 관련 네이버 지식인 질문을 분석하여 토픽을 파악하고, 질문에 나타난 불안, 걱정 관련 단어를 네트워크 분석한 연구[40] 등도 존재하지만 매우 소수에 불과하다.

국내 선행연구들은 동일하게 ‘코로나19’, ‘신종코로나’, ‘우한폐렴’ 등의 검색어를 사용하여 데이터를 수집, 분석하였음에도 결과의 양상, 연구의 시사점 및 제언에 상당한 차이가 있었다. 단순히 비교하기는 어렵지만 이러한 차이는 분석한 데이터의 채널별 특성 차이에 기인하는 것으로 추측된다. 뉴스 데이터를 통해서만 공식적 논의와 사회적 관심 등을 알아볼 수 있으나, 일반 대중의 개인적인 상호작용을 파악하기에는 어려움이 있다. 반면 소셜 빅데이터는 대중들이 블로그, 온라인 카페, SNS 등의 온라인 채널에서 공유한 대규모 데이터로서 개인의 경험, 생각, 감정 등이 담겨 있다[18]. 따라서 소셜 빅데이터를 활용한 연구는 코로나19 이후 가족생활을 일반 대중의 목소리를 통해 생생하게 파악할 수 있으며, 다양한 사람들의 논의를 포괄할 수 있다.

일반적으로 소셜 미디어¹의 종류는 블로그, SNS, 위키스, 콘텐츠 커뮤니티(예: 유튜브), 팟캐스트, 카페, 트위터 등이 있다[45]. 소셜 미디어를 구성하는 채널이 다양하다 보니 이용자의 이용 동기에 따라 소셜 미디어의 종류 및 이용 정도가 달라진다[46]. 효과적인 소셜 빅데이터 연구를 위해서는 데이터를 생성, 사용하는 이용자의 특성과 채널별 특성을 고려하는 것이 필요하다[47]. 국내에서 소셜 미디어의 양방향적 소통, 신속한 커뮤니티 구성, 공통 관심사의 공유 등이 주로 나타나는 채널

에는 블로그, 온라인 카페, 페이스북, 트위터 등이 있다[48]. 채널별 특성을 보면, 블로그는 사용자의 관심사에 따라 일기, 칼럼, 기사 등 자유로운 형태의 기록이 가능한 개인이나 집단의 홈페이지이다[49]. 온라인 카페는 유사성을 공유한 회원들 간의 관계 형성을 돕고, 자유롭게 의견을 논의, 교환할 수 있도록 하는 장이다[50]. 트위터는 단문 위주의 콘텐츠 구성으로 관심사와 이슈의 확산 및 공유가 용이하며, 팔로우라는 독특한 구조를 통한 관계 형성 및 유지가 가능하다[51]. 페이스북은 간단한 글과 댓글 작성, 좋아요 버튼 등을 통해 쉽고 빠르게 상호작용을 할 수 있어 네트워크를 유지, 확장하기에 용이하다[52]. 즉, 블로그와 카페는 자유로운 의견 공유와 개인적 표현이 주요 기능이므로 게시글의 내용이 상대적으로 풍부할 가능성이 높은 반면, 트위터와 페이스북은 간단한 콘텐츠를 바탕으로 한 확산 및 공유, 네트워크가 주요 기능이므로 게시글의 내용이 상대적으로 제한적일 가능성이 있다.

코로나19 시기 가족생활에 대해 알아보기 위해서는, 다양한 연령과 성별이 자유롭게 의견을 나눌 수 있고 문서 내에 여러 논의가 담길 수 있는 채널을 선택하는 것이 좋다. 이에 본 연구에서는 국내 포털 사이트 중 이용자 수가 가장 많은 네이버와 다음의 블로그 및 카페 데이터를 활용하여 코로나19가 가족에 미친 영향을 알아보고자 한다. 블로그와 카페는 게시글의 길이가 상대적으로 긴 경우가 많아서 본 연구에서 활용하고자 하는 빅데이터 분석방법인 토픽모델 분석에도 장점이 있다. 다수의 비정형 문서에서 글의 주제를 바탕으로 단어를 분류하는 토픽모델링을 활용할 경우, 블로그와 카페 게시글에 잠재된 주요 토픽을 파악할 수 있다. 토픽은 코로나19가 가족생활에 가져온 다양한 변화의 영역과 내용을 드러내므로, 코로나19 이후 가족생활의 전반적 현황을 파악할 때 도움이 될 것이다.

III. 연구방법

본 연구의 데이터 수집, 전처리, 분석의 절차는 [그림 1]과 같다. 단계별 구체적인 내용은 다음과 같다.

1 소셜 빅데이터라고 할 때 뉴스기사를 포함하는 연구도 있으나, 소셜 미디어는 기자, 학자 등 특정 집단이 아닌 일반 대중이 생성한 자료를 의미하는 것이 일반적이기 때문에, 본 연구에서도 소셜 빅데이터에 뉴스기사를 포함하지 않았다.

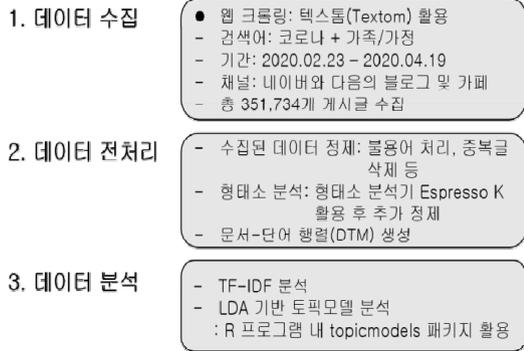


그림 1. 데이터 수집, 전처리, 분석 절차

1. 데이터 수집

본 연구의 데이터 수집기간은 국내 코로나19의 1차 대유행 기간이라고 볼 수 있는 2020년 2월 23일부터 2020년 4월 19일까지로 총 57일이다. 2월 23일은 정부가 위기단계를 경계에서 '심각'으로 격상한 날이며, 3월 22일부터 4월 19일까지는 '강화된 사회적 거리두기'가 시행되었다. 4월 20일에는 두 달 만에 확진자 수가 한 자릿수를 기록하면서 '완화된 사회적 거리두기'로 전환되었다[53].

데이터 수집은 수집기간 직후인 2020년 4월 23일에 이루어졌으며, 소셜 미디어 게시글 수집을 위해 빅데이터 분석 솔루션인 텍스트(Textom)를 활용하였다. 텍스트는 웹과 소셜미디어 상의 비정형 데이터를 수집하여 데이터세트(dataset)를 형성하며, 단계적 처리방식을 통해 데이터 큐레이션의 효율성이 높다는 특징이 있다. 텍스트 프로그램에서 데이터를 수집하는 경우 제목과 원글만 수집되므로, 댓글은 제외되었다. 본 연구의 목적은 코로나19 사태 이후 가족생활의 전반적 경향을 파악하는 것이므로 '코로나'가 반드시 포함되면서 '가족' 또는 '가정'이라는 단어가 언급된 게시글을 수집하였다. 텍스트에서 '코로나 + 가족', '코로나 + 가정' 각각의 검색어로 두 세트의 원자료를 수집하여 통합한 후 중복 게시글을 제거한 상태에서 전처리 및 분석을 하였다. 텍스트 솔루션은 유료서비스의 경우에도 채널별로 하루에 게시된 글을 최대 1,000건까지 수집가능하다는 제한점이 있어서 최대한 많은 데이터를 수집하기 위해 '코로나 + 가족', '코로나 + 가정'의 두 세트의 데이터를

별로 수집하였다.

데이터 수집 채널은 네이버와 다음의 블로그, 카페였으며, 채널을 선택한 이유는 다음과 같다. 첫째, 본 연구의 목적이 코로나 19와 관련된 가족생활의 전반적 경향을 알아보는 것이므로, 경험 및 의견의 공유가 자유롭게 이루어지면서 상대적으로 내용이 풍부한 블로그와 카페를 선택하였다. 둘째, 네이버와 다음을 선택한 이유는 국내 검색포털 가운데 점유율 1, 3위를 차지하고 있기 때문이다[54]. 점유율 2위인 구글의 경우 블로그나 카페를 운영하지 않아 제외되었다. 페이스북, 유튜브, 인스타그램, 트위터 등의 SNS 채널은 텍스트를 활용하여 테스트해 본 결과 수집되는 게시글의 수가 다른 채널에 비해 적어 분석에 어려움이 있다고 판단하였다. 또한 페이스북, 인스타그램, 트위터 등의 SNS 채널은 채널별로 이용자의 특성이 매우 다른 경향[55][56]이 있어, 소셜 미디어 채널 중 이용자 특성이 유사한 블로그와 카페 데이터를 분석 대상으로 한정하였다. 셋째, 뉴스 채널을 제외한 이유는 코로나19 확진 및 확진자 동향, 관련 정책 소개 등에 대한 게시글이 과도하게 수집될 가능성 때문이었다. 즉 뉴스기사를 포함할 경우, 토픽모델 분석의 결과로 코로나19 현황에 대한 주제가 다수 도출되어 가족생활에 대한 토픽이 부각 되지 않을 우려가 있었다.

최종적으로 분석한 게시글은 총 351,734건으로 수집기간이 짧았음에도 방대한 자료이다. 블로그 총 216,230건(다음 102,230건, 네이버 114,000건), 카페 총 135,504건(네이버 86,502건, 다음 49,002건)이었다. 텍스트 프로그램의 경우 채널별 일별 최대 수집량이 1,000건이므로 데이터 수집기간인 57일간 매일 최대 건수로 수집할 경우 57,000건이 된다. 본 연구에서는 '코로나 + 가정'과 '코로나 + 가족'의 검색어 각각의 데이터를 수집하였으므로 $57,000 \times 2 = 114,000$ 건이 수집될 경우 57일간 매일 일별 최대 수집량을 수집한 셈이 된다. 본 연구에서는 네이버 블로그만이 여기 에 해당하였다.

2. 데이터 전처리

데이터 전처리는 크게 4단계로 실시하였으며, 구체적인 절차는 다음과 같다. 1단계로, 텍스트 프로그램의 정

제·형태소 분석 기능을 활용하여 전처리를 하였다. 게시글의 제목과 본문을 함께 분석하기 위해 통합한 후 광고 관련 단어를 불용어로 처리하였다. 형태소 분석을 위해 일반명사, 고유명사, 복합명사 등을 한 단어로 추출하는 형태소 분석기인 Espresso K를 활용하여 명사와 동사, 형용사를 추출하였다. 한국어는 영어와 달리 단순히 공백을 기준으로 단어를 구분할 수 없기 때문에 의미를 가진 명사, 동사, 형용사를 추출하여 키워드를 알아보는 것이 적절하다고 알려져 있다. 2단계로는, 빈도 및 n-gram 분석결과를 바탕으로 코로나19 관련 용어, 정책 용어, 전문용어 등을 전처리하였다. 예를 들어, Espresso K에서 별개의 단어로 각각 추출되었던 '재난 + 지원금', '긴급 + 생계비', '가족 + 돌봄 + 휴가', '가정 + 폭력', '부부 + 상담', '양성 + 판정', '음성 + 판정', '재택 + 근무', '사회적 + 거리두기', '역학 + 조사' 등은 한 단어로 취급될 수 있도록 붙여쓰기 하였다. 3단계로는, 출현빈도가 높은 동의어가 키워드로 지나치게 자주 나타나지 않도록 한 단어로 통합하였다. 예를 들어 '코로나', '코로나바이러스', '신종코로나바이러스', '코로나19', '우한폐렴' 등의 단어는 '코로나19'로, '선생님', '선생', '교사', '쌤' 등의 단어는 '선생님'으로 통합하였다. 4 단계로는, 통계 소프트웨어 R의 'tm' 패키지와 'KoNLP' 패키지를 이용하여 각 단어의 출현빈도(term frequency)를 나타내는 문서-단어 행렬(document-term matrix)을 생성하였다. 이때 숫자, 날짜 및 시간을 의미하는 단어, 의미가 없는 한 단어 등을 삭제하였다.

3. 데이터 분석

코로나19와 가족생활에 관한 블로그, 카페 게시글의 주요 키워드와 토픽을 도출하기 위해 텍스트마이닝 기법 중 TF-IDF(term frequency-inverse document frequency)와 토픽모델 분석(topic modeling)을 활용하였다. 온라인 게시글은 특성상 데이터의 양이 방대하므로 빅데이터 분석기술을 활용하는 것이 효과적이다[56]. 특히, 텍스트 빅데이터 분석기술은 기존의 문헌 분석이 갖는 비효율성을 줄이고 자료에 잠재된 의미를 찾아 유용한 지식을 추출할 수 있다는 장점이 있다[57].

먼저 코로나19와 가족생활 관련 문서의 전반적 경향을 알아보기 위해 TF-IDF 가중치 값을 바탕으로 상위

100개 단어를 파악하였다. TF-IDF 가중치 모델은 TF(term frequency)와 IDF(inverse document Frequency)를 곱한 것이다. 특정 게시글에는 자주 출현하지만, 전체 게시글의 집합에서 그 단어가 출현한 문서의 수가 적을 경우 TF-IDF 값이 커진다. TF-IDF 값이 큰 단어는 그 단어가 포함된 문서의 주제 및 의미를 결정지을 가능성이 크기 때문에 주요 키워드가 될 수 있다[58].

이어서, 수집된 블로그, 카페의 게시글에 잠재된 주요 토픽을 파악하기 위해 토픽모델 분석(topic modeling) 기법을 활용하였다. 토픽모델 분석이란 다수의 비정형 문서에서 글의 주제와 관련된 정보를 바탕으로 문서를 분류하는 확률모형이다. 토픽모델링은 '의미는 관계적이다'라는 전제를 가지고 있으며, 의미가 단어 자체에 내재된 것이 아니라 어떤 단어의 군집에 속해있는가에 따라 결정된다고 본다. 토픽모델링은 문서별로 각 토픽에 분류될 확률, 토픽별로 각 단어가 속할 확률을 자동으로 산출하기 때문에 연구자의 사전지식이나 분류작업 없이 대량 문서를 빠르고 정확하게 처리할 수 있다. 즉, 토픽을 추출하는 방법이 명료하고, 의미의 관계성을 찾아내기 용이하다는 점에서 온라인 텍스트 분석에 매우 유용하다[59]. 본 연구는 토픽모델링 기법 중에서도 가장 자주 사용되는 잠재디리클레할당(latent Dirichlet allocation; LDA)을 사용하였다. LDA 기법은 문서 집합 내에서 동시 출현 빈도(co-occurrence frequency)가 높은 단어들을 기준으로 문서를 군집하여 토픽을 추출하는 통계적 텍스트 처리 기법으로, 의미상 일관성 있는 주제를 파악할 수 있다[60]. 이상의 분석은 통계 소프트웨어 R의 'topicmodels' 패키지를 사용하였다.

토픽모델링에서 연구자의 주요 역할은 토픽의 개수를 정하고, 토픽에 주제명을 붙이는 것이다. 토픽의 개수의 경우, 최적의 수에 대한 정답은 없으며 토픽의 해석가능성, 연구문제에 기반한 토픽의 유용성 등을 검토하여 결정한다[59]. 특히, 토픽모델링은 예외적 단어를 소수의 토픽으로 묶어 나머지 토픽의 해석가능성을 높이는 기법이므로, 전체 토픽의 적절성보다는 얼마나 많은 수의 의미 있고 유용한 토픽을 산출하는가를 고려하는 것이 중요하다[61]. 코로나19와 관련하여 토픽모델

링을 실시한 선행연구의 경우에서도 토픽 수가 다양하게 나타났다. 코로나19 관련 뉴스 기사를 분석한 연구에서는 20개 토픽을 추출하였고[62], 코로나19와 관련된 걱정에 관한 내용을 담은 문서를 분석한 연구에서는 글의 길이에 따라 짧은 문서는 15개 토픽, 긴 문서는 20개 토픽으로 추출하였다[63]. 네이버의 코로나19 관련 질문과 답변을 분석한 연구에서는 50개 토픽을 추출한 후 6개로 군집화하여 설명하였다[40].

본 연구에서는 적절한 토픽의 개수를 결정하기 위해, 토픽 수를 점진적으로 늘려가면서 토픽모델링을 실시하고 결과를 비교해 나갔다. 각각의 토픽이 가지는 의미와 차별성을 검토한 결과 20-25개 사이가 적절하다고 판단하고, 토픽이 20개인 경우부터 25개인 경우까지 도출된 각 토픽에 이름을 붙여보고 해석가능성과 배타적 의미를 치밀하게 살펴보았다. 그 결과 토픽의 수가 22개일 때 토픽별 의미가 명료하고 유용한 토픽이 많이 산출되어, 최종적으로 토픽의 수를 22개로 결정하였다. 이어서 각 토픽에 속할 확률이 높은 단어와 각 토픽으로 분류될 확률이 높은 게시글의 내용을 검토하여 22개 토픽의 주제명을 결정하였다. 마지막으로, 22개 토픽의 내용을 좀 더 간명하게 이해하기 위해, 각 토픽의 내용을 토대로 6개의 상위범주를 도출하고 상위범주를 명명하였다.

IV. 연구결과

1. 코로나19와 가족생활 관련 게시글의 키워드

코로나19와 가족생활 관련 게시글의 키워드를 TF-IDF 값을 통해 알아보면 [표 1]과 같다.

표 1. 코로나19와 가족생활 관련 게시글의 TF-IDF 상위 100개 단어

순위	단어	TF-IDF	순위	단어	TF-IDF
1	코로나19	209264.90	51	한국	24767.73
2	가정	166776.80	52	방법	24765.37
3	가족	103697.70	53	방문	24396.54
4	아이	67533.63	54	사용	24067.85
5	마스크	63607.46	55	휴원	23723.65
6	바이러스	61584.35	56	정부	23613.14
7	지원	59851.13	57	미국	23568.98
8	감염	54642.66	58	출음	23560.14

9	확진	53757.38	59	가정예배	23328.69
10	사회적거리두기	51243.42	60	부모	23281.32
11	확진자	50925.98	61	힘들	23232.58
12	사태	49763.81	62	교회	22943.21
13	상황	46052.39	63	사진	22641.33
14	사람	44734.73	64	엄마	22348.55
15	가정보육	43536.31	65	집	22048.79
16	건강	42693.75	66	필요	21955.60
17	소독	41560.71	67	안전	21459.83
18	시간	41180.20	68	사랑	21321.48
19	어린이집	40629.52	69	가정학습	20701.20
20	개학	39627.94	70	경제	20675.10
21	온라인	37548.46	71	어렵	20668.80
22	확산	36944.44	72	카페	20610.27
23	생각	36543.66	73	진행	20552.16
24	마음	36384.46	74	추가	20399.59
25	예방	35479.80	75	주말	20286.23
26	환자	34748.35	76	준비	19903.66
27	안내	32836.82	77	긴급	19594.07
28	연기	32503.94	78	유치원	19389.50
29	발생	32360.04	79	기도	19223.69
30	친구	30569.47	80	외출	19204.27
31	걱정	30453.35	81	일	18834.93
32	방역	29734.59	82	추천	18800.03
33	자가격리	29305.92	83	밥	18682.78
34	신청	29233.99	84	총식	18420.68
35	극복	29152.80	85	남편	17795.62
36	집콕	28532.32	86	가족돌봄휴가	17681.15
37	뉴스	27648.52	87	고민	17450.65
38	행복	27539.78	88	시장	17364.56
39	여행	27480.87	89	주일	17266.03
40	신청지	27157.43	90	활동	17254.76
41	놀이	27119.21	91	중국	17243.95
42	학교	27066.48	92	수업	17150.50
43	대응	26398.50	93	소득	16978.01
44	생활	25997.66	94	동선	16913.80
45	시작	25888.94	95	학생	16883.47
46	지원금	25615.17	96	취소	16754.99
47	검사	25589.97	97	자녀	16686.31
48	예배	25567.60	98	회사	16659.49
49	하루	25135.48	99	확인	16620.35
50	대구	24902.52	100	직장	16446.26

주: 단어별 단순빈도는 교신저자의 별도 연구[2] 참고

TF-IDF 기준으로 1-3위에 해당하는 ‘코로나19’, ‘가정’, ‘가족’은 본 연구에서 데이터 수집에 사용한 검색어이다. 이외의 단어들을 살펴보면 아이(5위), 가정보육(15위), 어린이집(19위), 개학(20위), 온라인(21위), 연기(28위), 놀이(36위), 집콕(37위), 여행(40위), 학교(42위), 생활(44위), 휴원(55위), 부모(60위), 엄마(64위), 집(65위), 가정학습(69위), 주말(75위), 유치원(79위) 등 자녀돌봄, 자녀교육, 부모역할과 관련된 단어가 중요하게 등장하였다. 또한 예배(48위), 가정예배(59위), 교회(62위), 기도(79위), 주일(89위) 등 코로나19 이후의 종교생활 관련 단어도 주요 키워드였다. 사회적 거리두

토픽 1은 '마스크', '예방', '소독', '방역', '손', '감염', '가족', '극복', '건강', '착용', '사용', '예방수칙', '구입', '질병관리본부', '확산' 등이 상위단어였다. 가족의 건강을 위해 코로나19 감염 및 확산을 예방하고자 마스크를 구입, 착용하고 손을 씻는 내용과 가정 내 소독 및 방역을 위한 질병관리본부 예방수칙에 대한 내용, 여타 기관의 소독 및 방역 관련 안내 등을 포함되었다. 게시글의 예로는 "2월 초만 해도 주출할 것 같았던 코로나19 근데 갈수록 감염자가 늘어가는 상황에 (중략) 아이들을 위해서라도 이제 마스크 착용이랑 손소독제 사용이 필수가 되었네요", "코로나 가정 소독 방역 락스나 알콜로도 됩니다. 질병관리본부 가이드라인입니다", "일산○○은 코로나19 예방을 위하여 원내 전체 소독 방역을 하루 2번 진행하고 있습니다" 등을 통해 파악할 수 있다. 이에 토픽 1을 '코로나19 예방'이라고 명명하였다.

토픽 2는 '확진', '가족', '확진자', '발생', '자가격리', '추가', '이동경로', '거주', '여행', '확진동선', '방콕', '나들이', '자택', '주말', '방문' 등이 상위단어였다. 추가 확진 및 확진자 발생과 이들의 이동경로에 대한 내용이 중심이었으며, 이에 따라 '거주', '자가격리', '자택', '방문' 등의 단어가 포함되었다. 이외에 '여행', '나들이', '주말', '방콕' 등은 추가 확진자 발생에 따라 사회적 거리두기가 강화되면서 나타난 가족의 변화를 반영하는 단어이다. 수집된 게시글 중 "구미 코로나19 확진동선 안내", "주말 나들이 못나가요. (중략) 코로나 확진자가 계속해서 많아져서 나들이 나가보려고 해도 무서워서 엄두가 안 나네요. 일요일인데 그냥 집에 있을 예정이에요. 주말 나들이도 못 나가고, 매일 불안하게 방콕하는 것도 너무 힘드네요" 등을 보면 이러한 양상을 파악할 수 있다. 이에 해당 토픽 2를 '코로나19 확진과 가족'이라고 명명하였다.

토픽 3은 '코로나19', '가족', '감염', '사람', '상황', '대구', '걱정', '환자', '한국', '생각', '사태', '확진자', '뉴스', '격리', '확진' 등이 상위단어였다. 한국의 코로나19 상황에 대한 걱정을 담고 있으며, 뉴스 등의 매체를 활용하여 확진 현황에 대한 정보를 얻고 있음이 나타난다. 특히, 이러한 걱정과 정보 획득의 동기가 가족감염인 것으로 나타났는데, 가족감염은 두 가지 의미로 해석할

수 있다. 가족 내 감염이 발생할 것을 우려하는 것과 가족 간 감염의 확산을 우려하는 것이 해당된다. 수집된 게시글 중 "코로나 특집, 집순이 모드 주말 사이 코로나가 심각단계로 격상됐다 원래는 미세먼지에도 마스크를 잘 안 썼는데, 신천지 사태 때문에 상황이 많이 나빠지고 있다", "코로나19 가족격리 저만 불안한가요? (중략) 대구서 가족 감염이 잇따라 이루어지고 있다고 하네요"에서 이러한 양상을 파악할 수 있다. 이에 토픽 3을 '가족감염'이라고 명명하였다.

토픽 4는 '가족', '검사', '면역력', '환자', '돌봄', '전화', '음성', '여성', '결과', '서비스', '방문', '생일', '업무', '실시', '확인' 등이 상위단어였다. 가족의 면역력을 증진하기 위한 노력과 가족구성원의 코로나19 검사 및 결과 확인에 대한 내용을 담고 있다. 또한 환자 돌봄 서비스, 전화 및 방문 서비스 등에 대한 내용도 담겨 있다. 수집된 게시글 중 "우리 가족 면역력 건강을 노니영양제로 챙기고 있어요", "아이가 아파서 급하게 병원 갔는데 어떤 가족이 코로나 음성이라 어찌고 하더니 쫓겨나갔어요. 괜찮은 거겠죠? 짹짹하네요" 등에서 이러한 내용이 나타난다. 이에 토픽 4는 '가족건강'이라고 명명하였다.

토픽 5는 '가족', '부산', '산책', '심각', '수원', '아파트', '마트', '단계', '격상', '해외', '인사', '입국금지', '두려움', '강아지', '평화' 등이 상위단어였다. 이는 위기단계가 심각단계로 격상된 이후 변화된 가족의 양상을 나타낸다. 일상적 생활공간인 아파트, 마트와 일반적인 활동인 강아지, 산책, 해외, 입국금지에 대한 내용이나 두려움, 평화 등의 단어를 통해 일상생의 변화에 대한 정서적 반응을 유추할 수 있다. 수집된 게시글 중 "코로나 터지고 구정부서 가정보육 시작했고 생필품 사러 어쩔 수 없이 마트 정도는 갔구요. 꼭 가야할 때 병원, 애 미칠 거 같아서 사람 없는 시간에 산책, 단지 놀이터 정도?" 등을 통해 이를 파악할 수 있다. 이에 토픽 5를 '일상생활'로 명명하였다.

토픽 6은 '조심', '식사', '저녁', '코로나19', '음성판정', '맛집', '집단', '요리', '불안감', '며칠', '원료', '바람', '추천', '엄마표', '나날' 등이 상위단어였다. 식사와 관련된 내용을 다루고 있으며, 크게 두 가지 양상을 포함하였다. 첫째는 코로나19에 대한 불안감으로 외식을 조심하며 가정에서 요리하는 양상이며, 이로 인해 '엄마표',

‘추천’ 등의 단어가 나타났다. 수집된 게시물 중 “엄마표 소불고기 저녁 식사 메뉴로 추천해요. 원래 요리를 되게 못하는데 코로나 땀에 집안에만 박혀있으니 뭐라도 만들게 되네요” 등으로 나타났다. 두 번째는 식사를 통한 코로나19 감염과 관련된 내용이었다. ‘식사’, ‘음성판정’, ‘완료’ 등의 단어가 동시에 포함된 것을 통해 유추할 수 있다. 수집된 게시물 중 “자택에서 식사 같이한 지인 1명 검사 중”, “대구경북 지역의 신천지 사람들 외에도 확진자들이 대거 나오는 건 대부분 가족이나 친구가 같이 식사 하거나 그래서일 가능성이 백퍼임” 등을 통해 파악할 수 있다. 이에 토픽 6을 ‘식생활’이라고 명명하였다.

토픽 7은 ‘밥’, ‘외식’, ‘집’, ‘음식’, ‘배달’, ‘가족’, ‘돌봄’, ‘전파’, ‘동거인’, ‘대화’, ‘브리핑’, ‘점심’, ‘증언’, ‘질문’, ‘수원시’ 등이 상위단어였다. 이는 코로나19 이후 외식보다는 ‘집, 밥, 배달, 음식’ 등으로 식사가 이루어지고 있음을 보여준다. 수집된 게시물 중 “코로나 때문에 집콕인 요즘 나가지도 않고 오는 손님도 없이 우리 가족 셋이서 오붓하게(?) 지내는 중이다. 평소에는 친정엄마와 시어머님이 왔다 갔다 하시면서 음식들을 가져다 주시는데, “이놈의 코로나 땀시 삼시세끼 힘드네요”, “올 가족만 먹으면 되니 배달 좀 시켜 먹을랍니다” 등을 통해 알 수 있다. 이에 토픽 7을 ‘식생활 변화’라고 명명하였다.

토픽 8은 ‘가족’, ‘건강’, ‘마음’, ‘사랑’, ‘행복’, ‘기도’, ‘교회’, ‘회원’, ‘예배’, ‘주일’, ‘사태’, ‘세상’, ‘고통’, ‘모임’, ‘기원’ 등이 상위단어였다. 가족의 건강, 사랑, 행복을 기도, 기원하고 마음을 위로하는 활동에 대한 내용을 담고 있었으며, 이에 따라 세상, 고통, 사태 등의 단어와 교회, 예배, 주일, 회원 등의 단어가 동시에 나타났다. 수집된 게시물 중 “4월 5일 종려주일 가정예배 순서: 10시에 알람을 맞춰놓고, 알람이 울리게 되면, 함께 기도하겠습니다(기도제목은 코로나19, 나라와 민족, 선거, 교회, 가족 건강과 가정을 위해서입니다)”, “부처님 오신날 연등접수 안내: 등을 밝히시어 가족 건강, 가정 평화와 소원성취를 이루십시오”, “프란치스코 교황님이 경제적 문제에 직면한 가정을 위해 기도하는 모습을 보고 저도 마음의 위로를 받았네요” 등을 바탕으로 토픽 8을 ‘종교생활’이라고 명명하였다.

토픽 9는 ‘가족’, ‘아이’, ‘집콕’, ‘카페’, ‘엄마’, ‘어린이집’, ‘주말’, ‘남편’, ‘시간’, ‘아빠’, ‘집’, ‘판정’, ‘하루’, ‘오랜만’, ‘단위’ 등이 상위단어였다. 코로나19 이후 가족생활을 반영하며, ‘집콕, 카페, 어린이집, 집’ 등의 단어와 ‘엄마, 아빠, 남편’, ‘주말, 시간, 하루, 오랜만’ 등의 단어를 통해 어린이집 휴원이 장기화됨에 따라 가족과 보내는 시간이 증가하였음을 나타낸다. ‘카페’라는 단어의 경우 중의적 의미를 내포하고 있었는데, 가족원과의 활동을 찾기 위해 다양한 카페(온라인)에서 정보를 얻는 내용과 가족들이 카페(오프라인)에 방문하는 내용을 모두 포함하였다. 수집된 게시물 중 “4인 가족 집콕... 아빠는 재택근무, 엄마는 무급휴직”, “슬기로운 집콕 생활. 코로나 19로 많은 학부모님들께서 걱정으로 하루를 보내는 것 같습니다”, “코로나19, 방역, 마트, 그리고 신천지. (중략) 오랜만에 찾은 단골 카페에서 당분간 테이크아웃만 가능하다고 했다” 등을 통해 알 수 있다. 더불어 “코로나 때문에 집콕 하는 시간이 길어지니까 남편이랑 홈트를 같이 시작했는데ㅋㅋㅋ 요즘 덕분에 사이가 좀 좋네요(부끄)”, “코로나 때문에 부부싸움 ㅠㅠ (중략) 아이들과 집에만 있으니 너무 답답하고 지쳐서 주말에 애들이랑 나들이라도 가자고 말 꺼냈다가 남편한테 괜한 소리만 들었네요ㅠㅠ” 등의 게시물을 통해 코로나19 이후 가족관계가 개선되거나 악화된 가족이 동시에 나타나고 있음을 파악할 수 있었다. 이에 토픽 9를 ‘집콕생활’이라고 명명하였다.

토픽 10은 ‘연기’, ‘개학’, ‘안내’, ‘학교’, ‘아이’, ‘학원’, ‘유치원’, ‘파악’, ‘확산’, ‘건강’, ‘온라인’, ‘공지’, ‘학생’, ‘휴업’, ‘수업’ 등이 상위단어였다. 코로나19 확산에 따른 ‘유치원, 학교, 학원’의 ‘개학연기, 휴업 조치, 온라인 수업’ 등에 대한 공지 및 안내에 대한 내용을 담고 있었다. 수집된 게시물 중 “큰아이 개학이 또 연기되었네요”, “코로나19 지역 전파 상황의 심각으로 학교 개학연기와 지역 학원, 도장에도 자체 휴관 결정 지침 사항이 내려와 저희 ○○태권도에서도 아이들의 건강이 가장 최우선임을 알기에 6주간 휴관을 결정하였습니다”, “처음 시행하는 온라인 개학인 만큼 학부모님들과 학생분들은 궁금한 점이 정말 많을 것 같아요” 등을 통해 나타난다. 더불어, “코로나 덕에 강제 이산가족이네요. 개학연기로 한 달은 어찌어찌 넘겼는데 당장 말길 곳이

없어 시골친정에 아이를 보냈어요. 얼른 코로나 종료되어서 일상으로 돌아가고 싶어요”, “코로나로 집에 있는 온갖 것들을 다시 꺼내 놀다보니 아이가 좋아하는 것과 좋아하지 않는 것을 자연스럽게 알게 되고 좋은 것 같아요” 등의 게시글을 통해 개학 및 개원 연기 조치 이후의 가족생활 및 가족관계도 파악할 수 있다. 이에 토픽 10을 ‘개학연기’로 명명하였다.

토픽 11은 ‘가족’, ‘취소’, ‘사회적거리두기’, ‘예약’, ‘돌잔치’, ‘결혼식’, ‘행사’, ‘휴원’, ‘휴관’, ‘마음’, ‘결혼’, ‘코로나19’, ‘고민’, ‘예정’, ‘진행’ 등이 상위단어였다. 코로나19 시기 가족행사 관련 내용을 담고 있었다. 돌잔치, 결혼식, 결혼 등의 가족행사를 진행할지에 고민, 예정대로 하거나 예약을 취소하는 등의 양상이 나타났다. 수집된 게시글 중 “코로나19 때문에 돌잔치 예약한 거 다 취소하고 우리 세 가족이 소소하게 집에서 ○○이의 첫 생일을 보내려고 마음먹었지만 (중략) 생일날이 다가오면 다가올수록 ○○이한테 미안한 감정과 함께 마음이 헛헛하더라고요”, “코로나 때문에 결혼식도 취소하고 날 짜를 미룰지 계속 고민하던데” 등을 통해 이러한 부분이 나타난다. 이에 토픽 11을 ‘가족행사’라고 명명하였다.

토픽 12는 ‘가족’, ‘여행’, ‘가족돌봄휴가’, ‘혼자’, ‘후기’, ‘취소’, ‘인생’, ‘용인’, ‘제주도’, ‘영화’, ‘마지막’, ‘종식’, ‘이탈리아’, ‘제주’, ‘문의’ 등이 상위단어였다. 코로나19 이후 여행 및 휴가 등에서 나타난 변화에 대한 내용을 담고 있었다. ‘가족, 혼자, 여행, 가족돌봄휴가’, ‘용인, 제주도, 이탈리아, 제주’, ‘마지막, 취소, 문의, 종식’ 등의 단어가 동시에 나타난 점을 통해 유추할 수 있다. 수집된 게시글 중 “코로나 터져서 이게 마지막 여행이 되었지”, “코로나 심각성을 모르던 설 연휴에 혼자 교토 여행을 다녀온 것을 마지막으로 집을 떠나지 못했다. (중략) 2월 가족 여행도 취소” 등을 통해 구체적인 내용을 파악할 수 있다. 이에 토픽 12를 ‘여행-휴가’라고 명명하였다.

토픽 13은 ‘가족’, ‘친구’, ‘걱정’, ‘아들’, ‘시국’, ‘모습’, ‘사진’, ‘아이’, ‘고민’, ‘파이팅’, ‘생각’, ‘친정’, ‘함께’, ‘마음’, ‘조심’ 등이 상위단어였다. 이는 코로나19 이후 가족과 친구에 대한 걱정을 담고 있었으며, ‘시국, 고민, 생각, 파이팅’ 등의 단어와 ‘아들, 친정, 함께, 조심’ 등의

단어가 함께 나타났다는 점으로 유추할 수 있다. 수집된 게시글 중 “코로나19 때문에 군인은 민간인 걱정, 민간인은 군인아들 걱정”, “이 어려운 시국 함께 잘 이겨내 봅시다! 우리 가족 화이팅! 대한민국 파이팅!” 등의 표현을 통해서도 알 수 있다. 이에 토픽 13을 ‘가족-친구 걱정’이라고 명명하였다.

토픽 14는 ‘불안’, ‘직원’, ‘경북’, ‘접촉자’, ‘사망’, ‘참석’, ‘아버지’, ‘말씀’, ‘역학조사’, ‘미세먼지’, ‘강제’, ‘졸업식’, ‘폐렴’, ‘폐쇄’, ‘소속’ 등으로 코로나19 시기의 불안을 담고 있었다. ‘소속, 직원, 접촉자, 참석’, ‘사망, 역학조사, 강제, 폐쇄’ 등의 단어를 통해 코로나19 접촉자로 인한 역학조사 및 강제폐쇄와 관련된 불안이 나타났으며, ‘폐렴’ 등에 대한 불안도 드러났다. 수집된 게시글 중 “코로나19는 아닌지 아니면 1월 2월 미국에서 아팠을 때 코로나19나 폐렴이었는데 그때는 모르고 지나갔다가 혹시 폐에 문제가 생긴 건 아닌지 (중략) 이런 저런 걱정과 불안이 많아서 고민” 등으로 파악할 수 있다. 이에 토픽 14를 ‘코로나19와 불안’이라고 명명하였다.

토픽 15는 ‘스트레스’, ‘이웃’, ‘교인’, ‘동네’, ‘고생’, ‘마음’, ‘친척’, ‘출발’, ‘코로나19’, ‘유급’, ‘나름’, ‘전원생활’, ‘자연텃밭’, ‘부담’, ‘오픈’ 등이 상위단어였다. ‘스트레스’, ‘마음’, ‘고생’, ‘부담’ 등의 단어가 동시에 나타난 것을 통해 코로나19 시기의 다양한 스트레스와 관련된 주제를 파악할 수 있다. 특히, 동네에서 코로나19 확진자가 발생하는 경우 이러한 스트레스가 더욱 커지는 양상이 나타났으며, 스트레스 대처를 위해 ‘전원생활, 자연텃밭’ 등을 이용하고 있었다. 코로나19로 인해 생계가 어려워지거나 자녀돌봄 공백이 발생한 경우 친척에 의존하거나 가족돌봄휴가를 사용한 모습도 엿볼 수 있었다. 본 연구의 경우 카페 및 블로그 데이터를 수집하였기 때문에 ‘이웃’이라는 단어가 중의적으로 사용되었는데, 실제 대면 관계의 이웃과 온라인상의 이웃을 의미한다. 수집된 게시글 중 “코로나19로 가정보육, 혹은 긴급보육 보내시는 이웃분들 마음 고생 심하실 텐데 모두 모두 힘내세요”, “지금의 상황을 보면 확진자도 문제이지만 의료진과 간호사와 스텝, 공무원, 감염자 가족 등 수많은 사람들이 심리적인 충격과 스트레스로 상처받고 있다” 등으로 표현되었다. 이에 토픽 15를 ‘코로나19와 스트레스’라고 명명하였다.

토픽 16의 상위단어는 ‘난리’, ‘가족’, ‘구매’, ‘의료진’, ‘도시’, ‘위생’, ‘가격’, ‘선물’, ‘선별진료소’, ‘가게’, ‘대란’, ‘비행기’, ‘졸업’, ‘진료’, ‘항공권’ 등으로, 코로나19가 확산되며 나타난 다양한 ‘대란’과 관련된 내용을 담고 있었다. ‘난리, 대란’과 ‘구매, 위생, 가격, 가게, 비행기, 항공권’, ‘졸업, 선물, 도시, 의료진, 선별진료소, 진료’ 등의 단어가 동시에 나타난 점에서 이를 유추할 수 있다. 수집된 게시글 중 “코로나19 바이러스 때문에 주위에서 난리다. 아니 대한민국 아니 전 세계가 난리다. 우리도 4년만에 처음으로 휴가를 잡아서 베트남에 가려고 비행기표 까지 끊고 준비했건만, “마스크 대란” 등으로 언급되고 있다. 이에 토픽 16을 ‘코로나19 시기 대란’이라고 명명하였다.

토픽 17은 ‘인천’, ‘아내’, ‘피해’, ‘신경’, ‘신랑’, ‘직장’, ‘중국인’, ‘공원’, ‘네이버’, ‘질병’, ‘주민’, ‘이해’, ‘유럽’, ‘자영업자’, ‘피해자’ 등이 상위단어였다. 코로나19로 인한 다양한 피해 양상을 반영하는 것으로 아내와 신랑, 직장, 자영업자, 주민, 유럽, 피해자 등의 단어가 동시에 나타난 점에서 유추할 수 있다. 수집된 게시글 중 “배스킨라빈스는 아내분이 점주시고 인천공항은 유럽 쪽 유학하는 딸 마중 가신 거라는데 코로나만 아니었으면 화목했을 한가정의 모습이 그려져서 너무너무 안타깝네요”, “코로나19 최대 피해자라고 할 수 있는 소상공인·자영업자” 등이 있었다. 이를 토대로 토픽 17을 ‘코로나19의 피해’라고 명명하였다.

토픽 18은 ‘동선’, ‘가족’, ‘지인’, ‘보건소’, ‘주변’, ‘운동’, ‘이동’, ‘방문’, ‘우한’, ‘남성’, ‘손씻기’, ‘문자’, ‘조심’, ‘안심’, ‘연인’ 등이 상위단어였다. ‘동선, 주변, 이동, 방문’과 ‘가족, 지인, 연인’, ‘보건소, 손씻기, 조심, 문자’ 등의 단어가 동시에 등장한 것으로 보아 코로나19 확진자의 이동 동선 관련 내용임을 파악할 수 있다. 수집된 게시글 중 “양천구 코로나19 4번째 확진자 발생 이동 경로 발표 긴급문자발송”, “불일이 있어서 수원 갔는데 긴급문자가” 등을 통해서 파악할 수 있다. 이에 토픽 18을 ‘긴급재난문자’라고 명명하였다.

토픽 19는 ‘가족’, ‘지원’, ‘신청’, ‘긴급’, ‘울산’, ‘대응’, ‘정부’, ‘안내’, ‘휴원’, ‘대상’, ‘사용’, ‘신도’, ‘상황’, ‘지원금’, ‘근로자’ 등이 상위단어였다. 코로나19로 인해 어려움을 겪는 가족을 지원하기 위한 정부 차원의 대응에

대한 내용을 담고 있었다. ‘긴급재난지원금, 신청, 대상, 사용처, 안내’ 등을 통해 파악할 수 있다. 수집된 게시글 중 “가족돌봄휴가 유급 지원 코로나로 인해 정부에서 일시적으로 유급지원 한다고 하니 참고하세요”, “코로나19 관련 가족돌봄비용 긴급지원에 대해 알려드립니다. (중략) 비상상황 종료 시까지 다음의 사유로 가족돌봄휴가를 사용한 근로자에 한해서 지원합니다.” 등으로 언급되고 있었다. 이에 토픽 19를 ‘가족지원정책’이라고 명명하였다.

토픽 20은 ‘확진판정’, ‘사망자’, ‘공개’, ‘오후’, ‘신천지’, ‘천안’, ‘금지’, ‘마중’, ‘연장’, ‘대전’, ‘구축’, ‘대구교회’, ‘의료원’, ‘심각단계’, ‘출입’ 등이 상위단어였다. 신천지 대구교회에서 나타난 집단감염 사태 이후 활발하게 논의된 코로나19 환자 현황 관련 내용을 포함하고 있었다. 수집된 게시글 중 “대구 쪽에 확진자가 집중되는 현상은 대구 신천지 교회가 주도적인 역할을 했는데요”, “원래는 미세먼지에도 마스크를 잘 안썼는데 신천지 사태 때문에 상황이 많이 나빠지고 있다.” 등으로 나타난다. 이에 토픽 20을 ‘코로나19와 신천지’라고 명명하였다.

토픽 21은 ‘가족’, ‘대구’, ‘시장’, ‘부탁’, ‘입원’, ‘종교’, ‘베트남’, ‘시대’, ‘한걸음’, ‘자제’, ‘관계’, ‘국가’, ‘사태’, ‘퇴원’, ‘확진자수’ 등이 상위단어였다. 토픽 20의 연장선으로 신천지 대구교회의 집단감염 사태 이후 대구를 중심으로 지역사회 전파가 확산되면서 나타난 논의를 담고 있었다. 대구시장이 입원, 퇴원, 확진자 수 등에 대해 브리핑하며 국가 위기 사태에 대응하기 위해 종교활동 및 단체활동의 자제를 부탁하는 내용을 담고 있다. 수집된 게시글 중 “대구시장 신천지 교인, 오늘 중 검사 받아라 진단검사 행정명령”, “권영진 대구시장은 자기 격리 중인 확진자의 생활치료센터 이송은 국가 전체적으로도 병상이 턱없이 부족한 상황 속에서 시민의 가족과 이웃의 소중한 생명을 지키기 위한 결정을 해야 할 것이라며” 등을 통해 파악할 수 있다. 이에 토픽 21을 ‘코로나19와 대구’라고 명명하였다.

토픽 22는 ‘회사’, ‘주문’, ‘식당’, ‘국내’, ‘근무’, ‘사항’, ‘직계가족’, ‘손님’, ‘불편’, ‘필수’, ‘실시간’, ‘정리’, ‘검체’, ‘예전’, ‘독감’ 등이 상위단어였다. 해당 토픽에 나타난 단어들을 바탕으로 수집된 자료를 검토한 결과 크게 3

가지 내용이 동시에 담겨 있는 것으로 나타났다. 하나는 코로나19 이후의 근무와 관련된 내용이며, “코로나로 인해 바뀐 회사 근무”, “회사 근무도 아이 양육도 힘든 요즘 (중략) 재택근무를 하며 아이를 돌보는 어떤 아빠의 모습을 뉴스를 통해서도 볼 수 있었다. 음... 얼마나 많은 회사, 또 얼마나 많은 직종이 가능할까?” 등을 통해 확인할 수 있다. 다음은 식당 관련 내용으로 “식료품, 식당은 영업을 금지하는 대신 손님을 제한해서 받는 방법으로 코로나를 대비한다고 해요”, “어머니가 식당을 하시는데 방역도 필수로 하고 이것저것 불편하지만 열심히 하고 있으시다는데 손님이 많이 끊겼다고 하네요” 등의 내용이 있었다. 마지막은 직계가족 관련 내용으로 “코로나19로 여러 지역에서 모이는 게 크게 무리여서 직계가족끼리 하기로 했어요” 등이 있었다. 토픽 22는 하나의 명칭으로 지칭하는 것이 부적합하다고 판단하여 ‘기타’라고 명명하였다.

이상과 같이 코로나19와 가족생활 관련 블로그와 온라인 카페 게시글에서 도출된 토픽 22개의 주제를 좀 더 간명하게 이해하기 위해, 유사한 토픽끼리 묶어서 상위범주를 도출하였다. [표 3]과 같이 22개 토픽은 6개의 상위범주로 분류할 수 있다.

표 3. 코로나19와 가족생활 토픽 범주화

번호	범주명	토픽 번호(토픽명)
1	코로나19라는 질병과 가족	토픽 1(코로나19 예방) 토픽 2(코로나 확진과 가족) 토픽 3(가족감염) 토픽 4(가족건강)
2	코로나19로 인한 가족생활 변화	토픽 5(일상생활) 토픽 6(식생활) 토픽 7(식생활 변화) 토픽 8(종교생활) 토픽 9(집콕생활) 토픽 10(개학연기) 토픽 11(가족행사) 토픽 12(여행, 휴가)
3	코로나19 관련 심리적 이슈	토픽 13(가족, 친구 걱정) 토픽 14(코로나19와 불안) 토픽 15(코로나19와 스트레스)
4	코로나19의 타격과 피해	토픽 16(코로나19 시기 대란) 토픽 17(코로나19의 피해)
5	코로나19 대응 정책	토픽 18(긴급재난문자) 토픽 19(가족지원정책)
6	코로나19 지역확산 및 기타	토픽 20(코로나19와 신천지) 토픽 21(코로나19와 대구) 토픽 22(기타)

범주 1은, ‘코로나19라는 질병과 가족’으로 코로나19의 예방, 감염, 확진 등의 이슈와 함께 가족건강에 대한 내용을 담은 토픽으로 구성되어 있다. 범주 2는, ‘코로나19로 인한 가족생활 변화’이며, 코로나19 이후 나타난 다양한 일상 및 가족생활의 변화를 담은 토픽으로 구성되어 있다. 범주 3은, ‘코로나19 관련 심리적 이슈’로 코로나19 이후의 걱정, 불안, 스트레스 등의 내용을 담은 토픽으로 구성되어 있다. 범주 4는, ‘코로나19의 타격과 피해’로 코로나19 시기 발생한 어려움에 대한 토픽으로 구성되어 있다. 범주 5는, ‘코로나19 대응 정책’으로 관련 정책에 대한 내용을 담은 토픽으로 구성되어 있다. 범주 6은, ‘코로나19 지역확산 및 기타’로 코로나19의 1차 확산기에 발생한 지역확산에 대한 내용을 담은 토픽과 기타 토픽으로 구성되어 있다.

V. 결론 및 논의

본 연구에서는 코로나19와 가족생활 관련 블로그, 카페 게시글의 키워드를 살펴보고 토픽을 도출하였다. 분석대상은 2020년 2월 23일부터 4월 19일 사이에 네이버와 다음의 블로그, 카페에 게시된 문서 총 351,734건이었다. TF-IDF 값을 기준으로 키워드를 살펴보았으며, 토픽모델링을 통해 총 22개의 토픽을 도출하고 토픽명을 부여하였다.

코로나19와 가족생활 관련 게시글의 주요 키워드를 통해 코로나19 이후 가족생활의 다양한 차원에서 여러 변화가 있었음이 드러났다. 구체적으로는 코로나19 이후의 자녀돌봄, 자녀교육, 종교생활, 근로와 관련된 키워드가 다수 등장하였다. 더불어 확진현황, 개인위생용품 및 소독·방역, 코로나19 관련 정책 등 다양한 영역과 관련된 키워드도 함께 발견되었다. 블로그 및 온라인 카페의 특성상, 상위 키워드를 통해 코로나19 이후 사람들의 생각과 감정을 생생하게 파악할 수 있다. 특히, 긍정 정서를 의미하는 ‘좋음’과 부정 정서를 의미하는 ‘힘들’, ‘어렵’ 등의 키워드는 토픽에서는 나타나지 않았지만 TF-IDF 기준 키워드로 나타나 코로나19 이후의 정서적 측면을 보여주었다.

토픽모델링의 결과로 도출된 토픽들은 가족생활의

매우 다양한 측면을 아우른다. 코로나19라는 질병의 예방, 감염, 확진과 가족건강 등의 주제가 토픽 1(코로나19 예방), 토픽 2(코로나 확진과 가족), 토픽 3(가족감염), 토픽 4(가족건강)에서 나타났다. 코로나19로 인해 일상생활 전반뿐만 아니라 식생활, 주거생활, 여가생활, 종교생활, 자녀돌봄, 자녀교육, 가족의례 등 가족생활의 구체적인 영역에서 나타난 변화에 대한 내용은 토픽 5(일상생활), 토픽 6(식생활), 토픽 7(식생활 변화), 토픽 8(종교생활), 토픽 9(집콕생활), 토픽 10(개학연기), 토픽 11(가족행사), 토픽 12(여행, 휴가)에서 나타났다. 코로나19의 걱정, 불안, 스트레스 등 심리적 이슈는 토픽 13(가족, 친구 걱정), 토픽 14(코로나19와 불안), 토픽 15(코로나19와 스트레스)에서 드러났다. 코로나19 상황이 '대란'으로 비유될 정도로 가족생활에 커다란 타격과 피해를 주었음은 토픽 16(코로나19 시기 대란)과 토픽 17(코로나19의 피해)에서 나타났다. 코로나19 관련 정부의 대응이나 가족지원정책에 대한 주제는 토픽 18(긴급재난문자)과 토픽 19(가족지원정책)를 통해 파악할 수 있다. 자료수집 시기에 코로나19가 대구지역을 중심으로 급속히 확산된 것을 반영하는 내용도 토픽 20(코로나19와 신천지), 토픽 21(코로나19와 대구)에서 나타났다.

이상과 같은 본 연구의 결과를 통해 코로나19가 가족생활에 미친 전방위적 영향을 확인할 수 있다. 특히, 식생활, 주거생활, 여가생활, 종교생활, 가족돌봄, 자녀교육, 가족의례 등의 영역에서의 변화와 영향이 두드러짐을 파악할 수 있다. 이 중 식생활과 주거생활은 인간이 살아가는데 가장 기초적이고 필수적인 영역이다. 식생활의 경우, 코로나19 이후 외식을 자제하고 집에서 직접 요리를 하거나, 식당에서 음식을 포장, 배달하는 등 변화가 나타났다. '엄마표 요리' 등의 단어를 통해 코로나19 시대에도 여성이 식생활의 주된 역할을 수행하고 있음도 알 수 있다.

주거생활의 경우, 사회적 거리두기로 외출을 자제하게 되면서 '집콕', '방콕' 등의 생활양식이 우리 사회에서 흔히 보이는 양상이 되었다. 코로나19 이전의 한국인에게 집은 수면과 같은 제한된 기능만 수행했던 공간이었다면, 코로나19 이후의 집은 일, 돌봄, 교육, 여가, 종교 활동 등을 위한 복합적 공간으로 변모[2]했음을 알 수

있다. 사회적 거리두기에 따라 집 밖의 사람들과의 접촉이 감소한 반면, 집 안의 사람들과의 상호작용이 증가했다는 점은 코로나19가 가져온 중대한 변화이다. 본 연구의 토픽모델링 결과 토픽 9(집콕생활)에는 코로나19로 인해 오랜만에 가족과 함께 시간을 보내게 되었다는 내용과 가족의 공유시간이 증가하여 가족관계에 변화가 나타났다는 내용이 함께 담겨 있어, 코로나19 이후 가족생활 및 가족관계를 파악할 수 있다. 토픽 10(개학연기)에서도 원격수업 이후 부모-자녀 간의 공유시간이 증가하자 관계가 긍정적 또는 부정적으로 변화하였다는 내용이 게시글을 통해 나타났다. 가족 행사 및 의례, 가족 여행과 휴가 등에의 변화가 주요 토픽으로 도출된 점도 주목할 만하다.

본 연구의 자료수집 시점인 2020년 2월 23일부터 4월 19일은 코로나19의 확산이 가족생활에 급격한 변화를 가져온 초기에 해당한다. 이러한 시기에 식생활, 주거생활, 여가생활, 종교생활, 자녀돌봄, 자녀교육, 가족의례 영역의 변화가 소셜 미디어에서 부각된 것은 코로나19 확산기에 일반인이 체험한 변화를 파악할 수 있다는 점에서 의의가 있다. 최근의 선행연구에서는 돌봄, 가사노동 등의 영역에 상대적으로 더 관심을 기울인 경향이 있는데, 코로나19가 가족생활에 미친 영향을 파악하기 위해서는 일상생활의 다양한 차원에 관심을 두는 연구가 필요함을 시사한다.

코로나19라는 질병과 가족의 연결방식을 파악할 수 있었다는 점도 본 연구에서 흥미로운 지점이다. 사회적 거리두기를 강화하더라도 가족 등 거주공간을 공유하는 사람들은 코로나19의 감염에서 자유로울 수 없다. 실제로 코로나19 확진자와 접촉한 가족은 직장 등에서 접촉한 사람보다 감염률이 높다고 한다[64]. 이렇듯 가족이 건강공동체라는 점이 본 연구의 결과에서도 나타난다. 예를 들면 가족 간 코로나19의 감염과 이에 대한 우려를 주제로 한 토픽 3이나, 코로나19 가족건강에 대한 토픽 4가 발견된 점이 그러하다. 건강공동체로서의 가족을 시사하는 내용은 코로나19와 직결된 신체적 건강 이외에도, 코로나19로 인한 심리정서적 어려움 등 정신적 건강과 연결된 토픽을 통해서도 알 수 있다. 토픽 13은 가족이나 친구에 대한 걱정, 토픽 14는 불안, 토픽 15는 스트레스를 주제로 하여 모두 코로나19 시

기 가족생활과 직간접적으로 관련된 심리적 어려움을 드러낸다. 국내 가족 관련 선행연구에서는 가족을 건강 공동체로서 바라보는 시각이나 관련 연구가 활발하지 않은데, 본 연구의 결과를 토대로 건강공동체로서의 가족에 관한 연구가 증가하기를 기대한다.

코로나19 시기 재택근무의 증가와 같은 일과 가족생활에서의 변화[25][26]나 코로나19의 경제적 영향[65][66]은 코로나19 관련 선행연구에서 강조된 주제이지만, 본 연구의 결과에서는 관련 내용이 상대적으로 두드러지지 않았다. 이는 해외에 비해 국내에서는 코로나19 시기 근무여건 변화의 폭이 상대적으로 적었고 기업 규모 및 직종에 따라 근무여건의 편차가 컸기 때문에 소셜 미디어에서 관련 내용이 적게 나타났을 가능성이 있다. 실제로 코로나19 이후 근무여건 변화를 국제 비교한 자료에 따르면, 한국은 미국, 일본, 영국, 중국, 이탈리아 가운데 가장 근무환경이 적게 변화한 국가였다[67]. 한편, 코로나19의 경제적 측면이 본 연구에서 별도의 토픽으로 도출되지는 않았지만, TF-IDF 상위 단어에 '지원금', '힘들', '경제', '어렵', '소득' 등 경제적 상황과 관련된 단어는 다수 등장하였다. 본 연구에서 경제적 측면이 하나의 토픽으로 묶이지 않은 이유는 여러 토픽에 경제적 측면이 포함되었기 때문일 수도 있다. 실제로 토픽 16(코로나19 시기 대란)에 소비생활의 이슈가, 토픽 17(코로나19의 피해)에 자영업자 등이 경험하는 경제적 피해와 어려움이 드러났다. 토픽 14(코로나19와 불안), 토픽 15(코로나19와 스트레스) 등에도 경제적 어려움으로 인한 심리적 어려움이 포함되었다.

본 연구의 한계점과 후속연구를 위한 제언은 다음과 같다. 첫째, 본 연구는 코로나19가 가족생활에 미친 전반적 영향을 알아보고자 하였으므로 다양한 차원을 개괄적으로 탐색하는 수준에 그쳤다. 그러나, 연구결과 나타난 토픽들이 코로나19가 영향을 미친 삶의 중요한 영역을 보여주고 있다는 점에서 개별 토픽의 구체적 양상이나 인과관계를 알아보는 후속연구가 필요하다. 둘째, 코로나19가 가족의 심리적 측면이나 가족관계에 미치는 영향이 크다는 것은 파악할 수 있었으나, 구체적으로 어떤 부분에서 어떤 영향을 미치고 있는지는 파악할 수 없었다. 후속연구에서는 가족스트레스 관점의 선

행연구[1]처럼 가족학적 이론을 토대로 코로나19 팬데믹이라는 특수한 상황이 가족생활에 미친 영향을 세밀하게 분석하는 연구가 필요할 것이다. 셋째, 소셜 빅데이터를 추출할 때 사용한 검색어가 '가족', '가정'으로 제한적이었다. 가족 관련 선행연구에서도 다양한 검색어를 사용하여 소셜 빅데이터를 추출한 연구가 소수 등장하고 있으나[68][69], 아직까지 대다수의 연구는 기술적 한계로 매우 적은 수의 검색어를 사용하는 경향이 있다. '가족', '가정'이라는 검색어가 본 연구의 문제의식을 담기에 충분히 포괄적이기는 하지만, 가족생활의 다양한 구성요소를 고려하여 후속연구에서는 보다 다양한 검색어를 활용하여 데이터를 수집할 필요가 있다. 넷째, 본 연구는 코로나19 확산 초반기에 블로그, 카페에 게시된 글만 수집하고 분석하였다. 코로나19 사태가 장기화되고 있다는 점에서 후속연구에서는 더 오랜 기간의 데이터를 수집하여 토픽의 시기별 동향을 비교 분석하는 것이 도움이 될 것이다.

참 고 문 헌

- [1] 진미정, 성미애, 손서희, 유재연, 이재림, 장영은, “코로나 19 확산에 따른 가족생활 및 가족관계의 변화와 스트레스,” *Family and Environment Research*, 제58권, 제3호, pp.447-461, 2020.
- [2] J. Lee, M. Chin, and M. Sung, “How Has COVID-19 Changed Family Life and Well-Being in Korea?,” *Journal of Comparative Family Studies*, Vol.51, No.3-4, pp.301-313, 2020.
- [3] 한국리서치, “위축된 삶, 지속될 영향, 실용적 적응: 코로나19 시대, 소비자생활과 소비활동에 대한 영향과 전망,” 2020.08.10.
- [4] 원우식, “코로나가 식구 찾아줬네,” *조선일보*, 2020.08.12.
- [5] 이미나, 김수영, “[법알못] 코로나 재택에 갇혀진 갈등…남편과 이혼위기입니다,” *한국경제*, 2020.08.15.
- [6] D. Bilefsky and C. Yeginsu, “Of ‘Covidivores’ and ‘Coronababies’: Life During a Lockdown,” *The New York Times*, 2020.08.12.
- [7] 정익중, “코로나 19 로 인한 아동돌봄 문제에 대한 해외 대응과 그 시사점,” *국제사회보장리뷰*, pp.47-59,

- 2020.
- [8] 최윤경, “코로나 19 와 아동 돌봄의 쟁점,” 젠더리뷰, 제57권, pp.63-69, 2020.
- [9] H. J. McLaren, K. R. Wong, K. N. Nguyen, and K. N. D. Mahamadachchi, “Covid-19 and Women’s Triple Burden Vignettes from Sri Lanka, Malaysia, Vietnam and Australia,” *Social Sciences*, Vol.9, No.5, p.87, 2020.
- [10] 장은하, 김희, “코로나 19와 성 불평등: 해외 사례를 중심으로,” 서울: 한국여성정책연구원. KWDI Brief, 제54호, pp.1-9, 2020.
- [11] T. M. Alon, T. M. Doepke, J. Olmstead-Rumsey, and M. Tertilt, “The Impact of COVID-19 on Gender Equality,” *National Bureau of Economic Research*, No.26947, 2020.
- [12] C. Collins, L. C. Landivar, L. Ruppanner, and W. J. Scarborough, “COVID-19 and the Gender Gap in Work Hours,” *Gender, Work & Organization*, 2020.
- [13] S. M. Brown, J. R. Doom, S. Lechuga-Peña, S. E. Watamura, and T. Koppels, “Stress and Parenting During the Global COVID-19 Pandemic,” *Journal of pre-proof, Child Abuse & Neglect*, 2020.
- [14] L. Cluver, J. M. Lachman, L. Sherr, I. Wessels, E. Krug, S. Rakotomalala, and A. Butchart, “Parenting in a Time of COVID-19,” *The Lancet*, Vol.395, Apr. 2020,
- [15] A. Mantovani, C. D. Byrne, M. H. Zheng, and G. Targher, “Diabetes as a Risk Factor for Greater COVID-19 Severity and In-hospital Death: A Meta-analysis of Observational Studies,” *Nutrition, Metabolism and Cardiovascular Diseases*, Vol.30, No.8, pp.1236-1248, 2020.
- [16] C. Mazza, E. Ricci, S. Biondi, M. Colasanti, S. Ferracuti, C. Napoli, and P. Roma, “A Nationwide Survey of Psychological Distress among Italian People During the COVID-19 Pandemic: Immediate Psychological Responses and Associated Factors,” *International Journal of Environmental Research and Public Health*, Vol.17, No.9, p.3165, 2020.
- [17] 한윤선, 김하영, 송주영, 송태민, “소셜빅데이터 수집 및 분석을 위한 아동청소년 학교폭력 온톨로지 개발,” *한국콘텐츠학회논문지*, 제19권, 제6호, pp.10-23, 2019.
- [18] H. Ishikawa, *Social Big Data Mining*, Florida: CRC Press, 2015.
- [19] 황유선, *소셜미디어의 이해*, 서울: 미래인, 2012.
- [20] Y. Demchenko, C. De Laat, and P. Membrey, “Defining Architecture Components of the Big Data Ecosystem,” *International Conference on Collaboration Technologies and Systems*, pp.104-112, 2014, May.
- [21] 윤명희, “블로그의 사회적 유형분석: 1 인 커뮤니티의 다층화,” *한국사회학*, 제41권, 제1호, pp.156-193, 2007.
- [22] J. M. Fegert and U. M. Schulze, “COVID-19 and Its Impact on Child and Adolescent Psychiatry—a German and Personal Perspective,” *Irish Journal of Psychological Medicine*, pp.1-3, 2020.
- [23] H. Prime, M. Wade, and D. T. Browne, “Risk and Resilience in Family Well-Being During the COVID-19 Pandemic,” *American Psychologist*, 2020.
- [24] D. Perlman, “A Finale for JCFS’ Global Family Perspectives on COVID-19 Special Issue: Themes and Reflections,” *Journal of Comparative Family Studies*, Vol.51, No.3-4, pp.445-453, 2020.
- [25] A. Bick and A. Blandin “Real-time Labor Market Estimates During the 2020 Coronavirus Outbreak,” 2020.
- [26] E. Brynjolfsson, J. J. Horton, A. Ozimek, D. Rock, G. Sharma, and H. Y. TuYe, “COVID-19 and Remote Work: An Early Look at US Data,” *National Bureau of Economic Research*, No.w27344. 2020.
- [27] R. Hertz, J. Mattes, and A. Shook, “When Paid Work Invades the Family: Single Mothers in the COVID-19 Pandemic,” *Journal of Family Issues*, 2020.
- [28] A. M. Verdery and E. Smith-Greenaway,

- “COVID-19 and Family Bereavement in the United States,” *Applied Demography Newsletter*, Vol.32, No.1, pp.1-2, 2020.
- [29] L. Stone, “Short-Run Fertility Responses to Mortality Events: A Look to the Past,” *Applied Demography Newsletter*, Vol.32, No.1, pp.18-20, 2020.
- [30] T. Jurcik, G. E. Jarvis, J. Zeleskov Doric, Y. Krasavtseva, A. Yaltonskaya, K. Ogiwara, and K. Grigoryan, “Adapting Mental Health Services to the COVID-19 Pandemic: Reflections from Professionals in Four Countries,” *Counselling Psychology Quarterly*, pp.1-27, 2020.
- [31] V. Varea and G. González-Calvo, “Touchless Classes and Absent Bodies: Teaching Physical Education in Times of Covid-19,” *Sport, Education and Society*, pp.1-15, 2020.
- [32] R. A. Settersten Jr, L. Bernardi, J. Härkönen, T. C. Antonucci, P. A. Dykstra, J. Heckhausen, and C. H. Mulder. “Understanding the Effects of Covid-19 Through a Life Course Lens,” *Advances in Life Course Research*, Vol.45, 2020, September.
- [33] K. Hank and A. Steinbach, “The Virus Changed Everything, Didn’t it? Couples’ Division of Housework and Childcare Before and During the Corona Crisis,” *Journal of Family Research. Early View*, pp.1-16, 2020.
- [34] 조재희, 조인호, “2018 메르스 해외 재유입에 대한 주요 온라인 이슈 탐색: 토픽모델링 분석과 감성 분석을 중심으로,” *한국디지털콘텐츠학회논문지*, 제20권, 5호, pp.1051-1060, 2019.
- [35] 한신갑. “빅데이터와 사회과학하기: 자료기반의 변화와 분석전략의 재구상,” *한국사회학*, 제49집, 제2호, pp.161-192, 2015.
- [36] 김상미, “코로나 19 관련 온라인 교육에 관한 국내 언론보도기사 분석,” *한국디지털콘텐츠학회논문지*, 제21권, 제6호, pp.1091-1100, 2020.
- [37] 김태종. “뉴스 빅데이터를 활용한 코로나 19 언론보도 분석: 토픽모델링 분석을 중심으로,” *한국콘텐츠학회논문지*, 제20권, 제5호, pp.457-466, 2020.
- [38] 오미애, 전진아, “코로나바이러스감염증-19 소셜 빅데이터 기반 주요 이슈 분석,” *보건·복지 Issue & Focus*, 제376권, pp.1-12, 2020.
- [39] H. W. Park, S. Park, and M. Chong, “Conversations and Medical News Frames on Twitter: Infodemiological Study on Covid-19 in South Korea,” *Journal of Medical Internet Research*, Vol.22, No.5. pp.e18897, 2020.
- [40] W. Jo, J. Lee, J. Park, and Y. Kim, “Online Information Exchange and Anxiety Spread in the Early Stage of the Novel Coronavirus (COVID-19) Outbreak in South Korea: Structural Topic Model and Network Analysis,” *Journal of Medical Internet Research*, Vol.22, No.6, pp.e19455, 2020.
- [41] J. X. Koh and T. M. Liew, “How Loneliness is talked About in Social Media During COVID-19 Pandemic: Text Mining of 4,492 Twitter Feeds,” *Journal of Psychiatric Research*, 2020.
- [42] Q. Liu, Z. Zheng, J. Zheng, Q. Chen, G. Liu, S. Chen, and C. J. Zhang, “Health Communication Through News Media During the Early Stage of the COVID-19 Outbreak in China: Digital Topic Modeling Approach,” *Journal of Medical Internet Research*, Vol.22, No.4, 2020.
- [43] D. Pan, J. Yang, G. Zhou, and F. Kong, “The Influence of COVID-19 on Agricultural Economy and Emergency Mitigation Measures in China: A Text Mining Analysis,” *PloS one*, Vol.15, No.10, 2020.
- [44] N. G. Uğur and A. Akbıyık, “Impacts of COVID-19 on Global Tourism Industry A Cross-Regional Comparison,” *Tourism Management Perspectives*, Vol.36, 2020.
- [45] 황유선, 박남기, *미디어기업의 소셜미디어 활용*, 서울: 한국언론진흥재단, 2010.
- [46] 김영민, “소셜미디어 이용자의 이용행태와 사회적 지각 인식 성격, 이용동기, 이용방식을 중심으로,” *한국콘텐츠학회논문지*, 제15권, 제4호, pp.407-419, 2015.
- [47] P. DiMaggio, “Adapting Computational Text Analysis to Social Science (and vice versa),” *Big Data & Society*, Vol.2, No.2, pp.1-5, 2015.

- [48] 오은혜, “사용자 만족에 영향을 미치는 소셜미디어의 특성과 상호작용에 관한 실증연구,” e-비즈니스연구, 제16권, 제6호, pp.405-424, 2015
- [49] G. Lawson-Borders and R. Kirk, “Blogs in Campaign Communication,” American Behavioral Scientist, Vol.49, No.4, pp.548-559, 2005.
- [50] 구교태, “가상 커뮤니티 이용 욕구(needs)와 행위 (behavior)에 관한 연구: 대학생들의 인터넷 카페 이용을 중심으로,” 한국언론정보학보, Vol.30, pp.7-33, 2005.
- [51] 조화순, 김정연, “소셜미디어의 매체 특성과 참여의 커뮤니케이션: 반값등록금 관련 블로그와 트위터 내용 분석,” 사이버커뮤니케이션학보, 제29권, 제2호, pp.95-130, 2012.
- [52] 박은진, *페이스북 이용자의 네트워크 특성과 이용행태의 관계에 관한 연구*, Doctoral dissertation, 서울대학교 대학원, 2014.
- [53] 정부 24, “‘사회적 거리두기’ 완화된 형태로 지속…2주마다 평가해 수위 조절,” 정책브리핑. <https://www.gov.kr/portal/ntnadmNews/2144351>, 2020. 08. 10.
- [54] 인터넷트렌드, 2020. 08. 09 <http://www.internet.trend.co.kr/trendForward.tsp>.
- [55] G. Blank and C. Lutz, “Representativeness of Social Media in Great Britain: Investigating Facebook, LinkedIn, Twitter, Pinterest, Google+, and Instagram,” American Behavioral Scientist, Vol.61, No.7, pp.741-756, 2017.
- [56] 이수연, 김현정, 정수연, “텍스트마이닝 기반 토픽 분석을 통한 온라인 성차별성의 이해,” 사이버커뮤니케이션학보, 제33권, 제3호, pp.159-199, 2016.
- [57] 박현정, 김한나, 홍유정, “토픽모델링을 활용한 학생 인권조례의 사회적 이슈 분석,” 아시아교육연구, 제18권, 제4호, pp.683-711, 2017.
- [58] 이성직, 김한준, “TF-IDF의 변형을 이용한 전자뉴스에서의 키워드 추출 기법,” 한국전자거래학회지, 제14권, 제4호, pp.59-73, 2009.
- [59] 남춘호, “일기자료 연구에서 토픽모델링 기법의 활용 가능성 검토,” 비교문화연구, 제22집, 제1호, pp.89-135, 2016.
- [60] D. M. Mimno and A. McCallum, “Topic Models Conditioned on Arbitrary Features with Dirichlet-multinomial Regression,” UAI, Vol.24, pp.411-418, Jul. 2008,
- [61] P. DiMaggio, M. Nag, and D. Blei. “Exploiting Affinities between Topic Modeling and the Sociological Perspective on Culture: Application to Newspaper Coverage of US Government Arts Funding,” Poetics, Vol.41, No.6, pp.570-606, 2013.
- [62] Y. Bai, S. Jia, and L. Chen, “Topic Evolution Analysis of COVID-19 News Articles,” In Journal of Physics: Conference Series, Vol.1601, No.5, Vol.1601, No.5, pp.1-6, 2020.
- [63] B. Kleinberg, I. van der Vegt, and M. Mozes, “Measuring emotions in the Covid-19 real world worry dataset,” arXiv preprint arXiv:2004.04225, 2020.
- [64] Y. J. Park, Y. J. Choe, O. Park, S. Y. Park, Y. M. Kim, J. Kim, and J. Lee, “Early Release-Contact Tracing during Coronavirus Disease Outbreak South Korea, 2020,” Emerging Infectious Diseases Journal-CDC, Vol.36, No.10. pp.2465-2468, 2020.
- [65] R. Baldwin and B. Weder di Mauro. “Economics in the Time of COVID-19,” London: CEPR Press, A VoxEU.org Book, Centre for Economic Policy Research. 2020.
- [66] N. Fernandes, “Economic Effects of Coronavirus Outbreak (COVID-19) on the World Economy,” Available at SSRN 3557504. 2020.
- [67] M. Belot, S. Choi, J. C. Jamison, N. W. Papageorge, E. Tripodi, and E. van den Broek-Altenburg. “Six-Country Survey on Covid-19,” 2020.
- [68] 송태민, 이기호, 진달래, 천미경, 서동철, 박현애, *2016년 소셜 빅데이터 기반 보건복지 이슈 동향 분석*, 서울:한국보건사회연구원, 2016(2).
- [69] 안수빈, 이강이, 이재림, 김은경, “소셜 빅데이터를 활용한 양육스트레스의 의미 네트워크 분석,” 한국가정관리학회지, 제38권, 제1호, pp.61-77, 2020.

저 자 소 개

박 선 영(Sunyoung Park)

준회원



- 2019년 8월 : 경희대학교 아동가족학과 (생활과학사)
- 2020년 3월 ~ 현재 : 서울대학교 아동가족학과 석사과정

〈관심분야〉 : 청년과 가족, 커플관계, 소셜미디어 빅데이터

이 재 림(Jaerim Lee)

정회원



- 2001년 2월 : 서울대학교 아동가족학과(생활과학 석사)
- 2009년 8월 : 미국 University of Minnesota, Family Social Science (Ph.D.)
- 2015년 3월 ~ 현재 : 서울대학교 아동가족학과 교수 및 생활과학연

구소 겸무연구원

〈관심분야〉 : 청년과 가족, 세대관계, 가족생활교육