



가족 관련 이슈 및 연구동향, 한가원의 주요연구 공유를 하고자 발간된 한국건강가정진흥원의 발간물입니다.

관련연구 SNS 빅데이터를 활용한 가족 변화 양상 분석

연구기관 RSN

연구담당 한국건강가정진흥원 사업기획부

SNS가 나타내는 한국가족의 변화 양상

최근 가족 유형 다양화 및 SNS 사용량 증가, 코로나로 인해 가족끼리 함께하는 시간이 증가하는 등의 원인으로 인해 가족 관련 정보량은 매년 증가하는 추세를 보이고 있습니다. 또한 SNS 내 가족의 의미 및 행태 변화 양상을 파악하여 변화한 내용에 적합한 가족 서비스 발전 방향성을 모색하고, 맞춤형 제공 방안을 마련하고자 SNS 빅데이터를 활용한 가족 변화 양상 분석 연구를 진행하였습니다.

주요내용

해당 연구는 2017년부터 2022년까지 약 6개년의 커뮤니티, 카페, 인스타그램, 유튜브 채널의 '가족' 관련 소셜데이터 약 2,700만건의 데이터를 성별, 연령층 등의 계층, 언급되는 가족 형태 등 다양한 방법으로 카테고리화하여 시계열, 감성, 연관어, LDA, 네트워크 분석 등의 분석 방법을 통해 가족의 의미 및 개념에 대한 변화 양상과 가족생활의 변화 및 특이점을 확인하였습니다.

본 브리프에서는 해당 연구의 '가족' 관련 소셜데이터 분석 결과, 제언 등 주요 내용을 발췌하여 공유하고자 합니다.

목차

- I. 서론
- II. 분석방법
- III. 분석결과
- IV. 제언

I. 서론

가족과 관련하여 소셜미디어 상의 다양한 데이터의 변화를 확인하고 분석하여, 가족의 의미 및 행태 변화 양상을 파악하여 변화한 내용에 적합한 가족 서비스 발전 방향성을 모색하고, 맞춤형 제공 방안을 마련하고자 본 연구를 추진하게 되었습니다.

효과적이고 의미 있는 분석을 위해 가족, 가정 등과 관련된 주요 키워드를 선정하여 2017년 1월부터 2022년 6월까지 소셜데이터(커뮤니티, 카페, 인스타그램, 유튜브) 26,994,344건을 분석하였습니다.

분석 채널을 위와 같이 설정한 이유는 연령, 성별, 가족 구성원 등 다양한 이용자들이 언급한 ‘가족’과 관련된 데이터가 필요하기 때문이며, 본 연구가 일반 국민들이 생각하는 가족에 대한 의미 변화 분석이 주요 목적이기 때문에 ‘언론’ 채널은 제외하게 되었습니다.

또한 ‘블로그’ 채널은 원문 데이터의 양이 많고, 가족의 의미를 주제로 다루기보다 타 주제를 언급하는 게시물에 가족 관련 키워드가 포함되는 경우가 많아, 전제 분석 내용에 필요하지 않은 키워드가 포함될 수 있으므로 제외하였습니다. 각 채널 게시물의 이미지나 동영상은 분석에 사용되지 않고, 텍스트 분석만 진행하였으며, 게시물의 댓글 또한 분석에 사용하지 않았습니다. 분석 과정은 아래 [그림 1]과 같습니다.

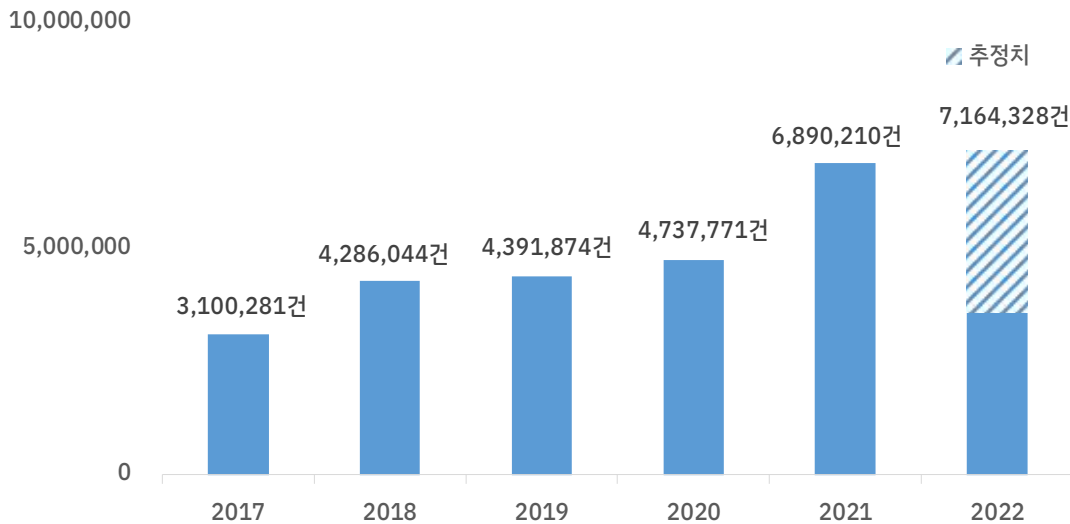


[그림 1] 데이터 분석과정

II. 분석방법

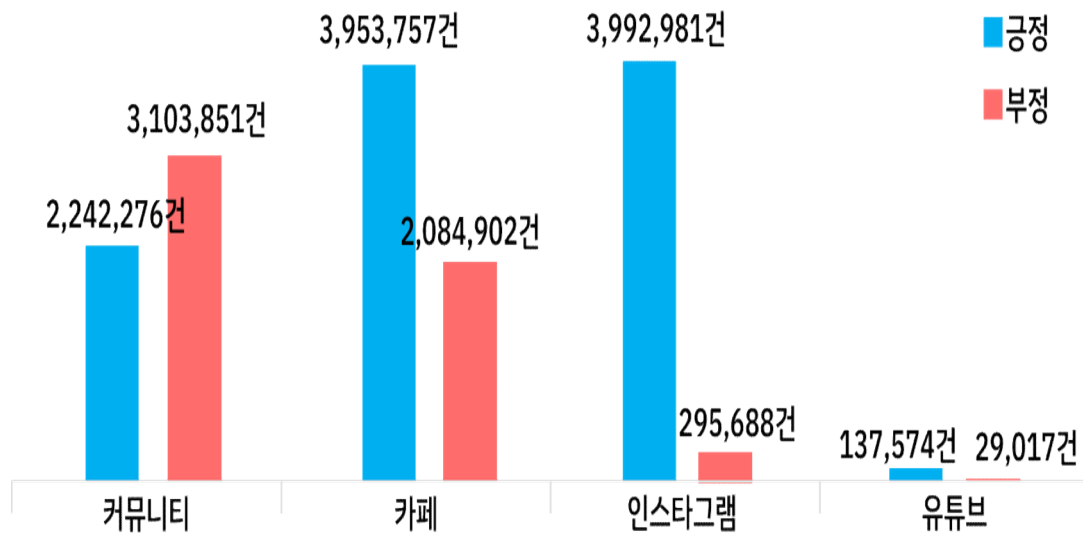
분석방법으로는 ① Volume & Sentiment 분석, ② 연관어 분석 (Word Cloud 분석), ③ 중요 키워드 분석 (TF-IDF 분석), ④ 토픽모델링 분석 (LDA), ⑤ 의미연결망 분석 (Network 분석), ⑥ 의미연결망 분석 (CONCOR 분석), ⑦ 연속 문자 추출 분석 (N-gram 분석) 등이 사용되었습니다.

III. 분석결과



[그림 2] 연도별 전체 정보량 변화

전체 가족 관련 정보량은 매년 증가하는 추세를 보입니다. 21년도 1분기를 기점으로 정보량이 급상승하였는데 이는 엔데믹이 언급될 만큼 코로나19의 확산이 잦아들면서 가족 모임 및 여행, 외식 등 가족이 함께하는 다양한 활동과 관련된 데이터가 증가하였기 때문입니다.



[그림 3] 채널별 감성 정보량 비교

채널별로는 <그림3>에 보이는 것처럼 인스타그램과 유튜브는 긍정적 언급(추억, 일상, 여행, 재미있는 이야기 등), 카페와 커뮤니티는 위의 두 채널에 비해 부정적 언급량(가족 간 갈등, 고민, 불만 등)이 많음을 알 수 있습니다. 가족 관련 전체 연관어는 아래의 <그림4>처럼 ‘일상생활 언급’, ‘가족 생애주기 및 돌봄’, ‘가족 관련 감성어’ 등 세 가지 유형으로 분류되었으며, 육아와 관련된 연관어가 눈에 띄는 것이 특징입니다.



- “가족관련 일상 생활 관련 연관어”**

 - 주요 키워드: 일상 여행, 친구, 친정 생활, 양육, 공부, 캠핑, 대하, 활동 등
 - 가족 여행, 재테크, 친구 관계, 가족 간의 대하, 활동 등 일상 속 가족 구성원 연대사
- “가족 관련 생애 주기 및 돌봄 관련 연관어”**

 - 주요 키워드: 신혼기, 신혼부부, 양육기, 육아, 행복, 이기 등
 - 가족 생애 주기 중 신혼부부, 행복, 신혼여행, 신혼집 등 신혼기 관련 연대사
 - 육아, 이기 임신 출산 등 양육기 관련 연대사 많음
- “가족 관련 감성_행복, 고민, 어려움 등 다양한 상황 관련 감성어”**

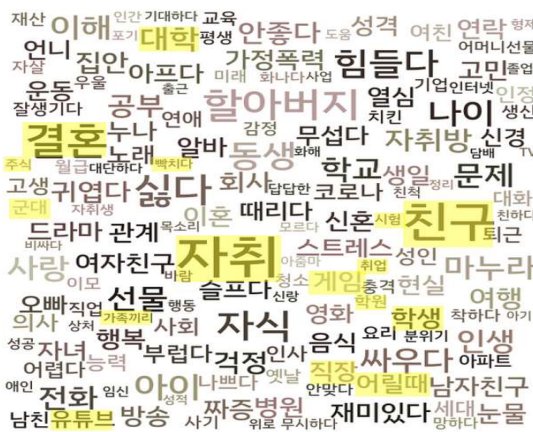
 - 주요 키워드: 힘들다, 맞선, 행복, 립스타그램, 이쁘다, 데이트, 즐겁다, 싸우다, 코로나, 스트레스, 불행 등
 - 가족 및 연인 간 데이트, 맞선부터 관련 긍정 연대사
 - 가족 간 불화, 불통, 이혼 등 위 상황과 코로나로 인한 건강 관련, 스트레스, 자립 등의 부정 연대사 많음

[그림 4] 가족 관련 전체 연관어 Top 200



[그림 5] 여성(왼쪽)과 남성(오른쪽)이 언급한 가족 관련 전체 연관어 Top 150

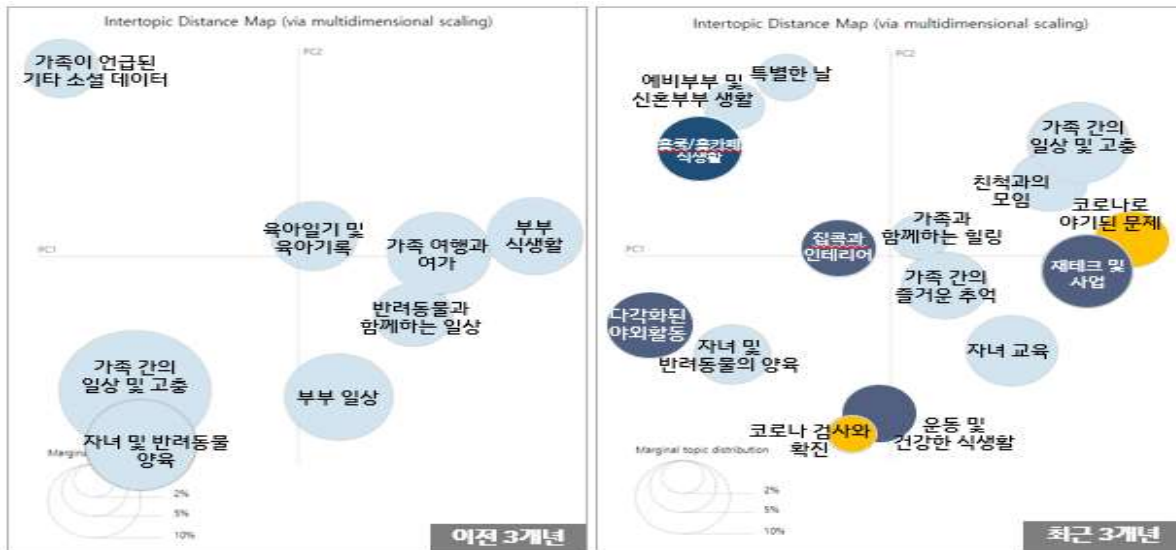
성별에 따라 가족 관련 연관어의 차이가 나타났습니다. 여성은 ‘남편, 결혼, 서운하다, 시어머니, 친정엄마, 생활비, 설거지, 어린이집, 힘들다, 스트레스’와 같이 가족 관계, 가족과의 삶에 대한 고충 등 가족의 일원으로 느끼는 감정, 표현 등의 언급을 주로 하였습니다. 반면, 남성은 ‘아내, 결혼, 취업, 운동, 게임, 보험, 재산’으로 가족보다는 개인과 관련된 키워드를 주로 언급하였습니다.



[그림 6] 청년(왼쪽)과 장년(오른쪽)이 언급한 가족 관련 전체 연관어 Top 150

또한 연령대에 따라서도 차이가 나타났는데 ‘청년층’은 가족과 친구 등 본인과 관련된 제 3자와 관련된 내용이 많으며, 청년층의 나이대에 누릴 수 있는 현실적 상황과 개인의 취미 등을 주로 언급하였습니다.

반면, ‘중장년층’의 가족 관련 주요 연관어는 ‘자식, 교육, 어린이집, 임신, 반찬, 청소, 부동산, 경제’ 등이었으며, 자녀의 육아와 교육, 가사 노동, 자산 관리 등 부모의 관점이 많이 담긴 키워드가 주로 언급되었습니다.

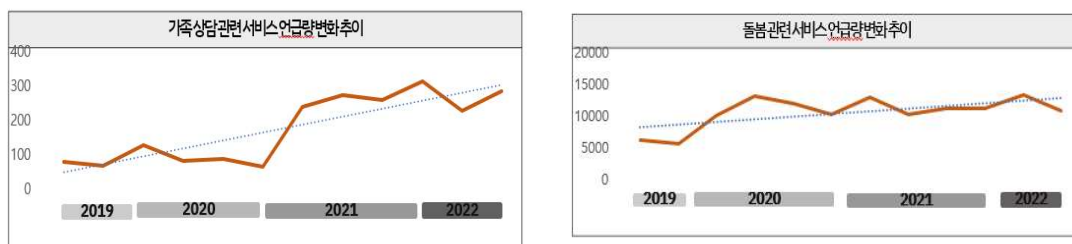


[그림 7] 연도별 주요 Topic 변화 분석

한편, 코로나19는 가족생활 내 변화를 가져왔습니다. 코로나19가 발생한 2020년을 기점으로 소셜데이터 가족 관련 토픽 분석 결과, 여행, 외식, 여가 등의 가족 활동 비중이 늘어났으며, 가족 구성원의 개인적 취향이 반영된 활동을 지향하는 양상이 나타나는 것을 확인할 수 있었습니다.

이는 가족의 의미가 가구 내 함께 생활하는 단순 공동체에서 가족 구성원 개개인의 취향을 존중하고 다양성이 존재하는 개별적 주체의 집합체로 변화되어가고 있음을 나타내고 있습니다. 또한 가족의 범위 내 반려동물·식물을 재배나 사육의 대상이 아닌 가족구성원으로 받아들이는 비혈연관계 연관 키워드 언급도 증가하였음을 확인하였습니다.

마지막으로, 가족 서비스 관련 언급량을 비교한 결과, 돌봄관련서비스와 가족상담관련서비스의 언급량이 지속적으로 증가하고 있으며, 특히 가족상담관련서비스는 2020년 하반기부터 급격히 증가하였음을 알 수 있었습니다.



[그림 8] 가족상담 및 돌봄 관련서비스 연도별 언급량 변화추이

III. 제언

본 연구의 분석 결과 및 한계점을 토대로 후속 연구에 대한 제언을 하고자 합니다.

SNS 빅데이터를 활용한 가족변화 양상을 분석한 결과, 가족 형태 및 관련 행태의 양상이 점차 세분화되고, 각각의 고충, 관심사, 의식주 등 전반적인 생활 및 삶의 방식의 차이가 생겨나고 있는 것으로 보여집니다.

따라서 전통적으로 분류되는 가족 형태가 아닌, ‘반려동식물을 키우는 가족, 1인가구, 딩크족, 공동체가족, 비혼’ 등 새롭게 등장한 가족 형태를 중심으로 심층 분석을 한다면 보다 다양한 범위의 가족들의 특성을 파악하고 해당 가족군의 니즈에 부합하는 서비스를 제공하는 데에 도움이 될 것으로 생각합니다.

본 연구는 6개년이라는 긴 기간과 가족과 관련된 광범위한 키워드를 데이터 추출용으로 활용하여 수집하고 분석하여 거시적인 관점에서 가족의 의미 변화를 살펴볼 수 있었습니다. 그러나, 시간적, 기술적 한계에 의해 세부적인 가족 형태 및 특성 관련 미시적인 분석에는 어려움이 있었습니다.

따라서 후속 연구를 수행할 시, 가족의 형태 및 특성, 키워드, 대상 수집원 등을 구체적으로 설정하여 인사이트가 드러나는 연구를 수행할 수 있을 것으로 보입니다.

※ 위 연구보고서 전문은 ALIO : 공공기관 경영정보 공개시스템(https://alio.go.kr/organ/organDisclosureDtl.do?apbaId=C0845&apba_id=C0845)에서 다운로드 받으실 수 있습니다.