

소셜 빅데이터를 활용한 청년주거 이슈와 의미 네트워크 분석 : 코로나19 이후의 네이버 뉴스 기사를 중심으로

Youth Housing Issues and Semantic Network Analysis Using Social Big Data: Based on Naver News Articles Since an Outbreak of COVID-19

황윤서 · 이현정*

충북대학교 생활과학연구소 전임연구원 · 충북대학교 주거환경학과 교수*

Hwang, Yoon-seo · Lee, Hyun-Jeong*

Research Institute of Human Ecology, Chungbuk National University,
Department of Housing and Interior Design, Chungbuk National University*

Abstract

This study analyzed youth housing issues and changes after the COVID-19 pandemic and provided implications for future youth housing welfare plans using social big data. Data were collected by web crawling Naver articles published from February 2020 to February 2022. A total of 11,609 words were extracted, and term frequency-inverse document frequency values were used for analysis, resulting in a word cloud and network analysis. This study identified three main topics related to the youth housing support project: 'need factors and supply factors', 'economic factors', and 'social factors,' each with sub-topics. Step-by-step support measures are suggested for young people to become independent with education and publicity to promote participation and institutional foundation preparation as important factors.

Keywords: Youth housing, Housing issues, Big data, Network analysis, Naver article

I. 서론

코로나19 확산 방지 조치에 따른 비대면 환경은 어느 때보다 빠르게 성장하였고 소셜 미디어는 매우 중요한 소통 수단으로 자리매김하게 되었다. 그럼에 따라 실시간으로 떠오르는 소셜 빅데이터는 최근 쟁점이 되는 사회적 문제와 관련된 시사, 동향, 의견 등의 정보를 제공해주는 동시에 정량적 데이터를 보완할 수 있는 유용한 정성적 데이터 소스로도 활용되고 있다(김동

완, 2013; 정영철, 천미경, 2021; 최현주, 최연철, 2016).

이렇듯 소셜 미디어 플랫폼 및 블로그, 포럼, 온라인 커뮤니티와 같은 사회적 상호 작용의 디지털 소스에서 생성된 소셜 빅데이터는 사용자의 행태와 선호도에 대한 다양한 정보를 효율적으로 제공해 준다는 측면에서 정치, 사회, 경제, 문화, 과학 기술 등 전 영역에 걸쳐서 활용 가치가 높아지고 있다(황승구 외, 2013). 최근 정부에서도 2019년 3월 빅데이터 플랫폼 및 센터 구축사업을 시작하여 금융, 환경, 문화, 교통, 헬스케어

본 논문은 2019년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원(NRF-2019R1A2C1005122)에 의하여 연구되었음

이 논문은 2021년 (사)한국생활과학회 하계학술발표대회에서 발표한 논문을 수정·보완한 것임

* Corresponding author: Lee, Hyun-Jeong

Tel: +82-43-261-2740, Fax: +82-43-276-7166

E-mail: hlee@cbnu.ac.kr

© 2023, Korean Association of Human Ecology. All rights reserved.

등 10개 분야 빅데이터 플랫폼 및 94개소 센터를 운영하고(과학기술정보통신부, 2019) 장애인, 노인, 아동, 건강 보장 관련 키워드를 분석하여 정책연구 분야에 참고할 만큼 유용한 데이터로 활용되고 있다(정영철, 천미경, 2021).

한편 코로나19 사태로 가속화되는 주거불안은 청년의 삶을 더욱 불안하게 하고 청년의 주거비 부담은 지속적으로 증가하여 이들의 주거문제는 사회적으로도 우려가 커지고 있다(김진영, 이현정, 2020; 김혜은 외, 2021). 이미 코로나19 이전부터 오랜 기간 이어진 청년의 주거불안은 이들의 독립을 저해하고 부모로부터 물리적·경제적 의존을 높이지게 하는데 이는 비혼, 만혼 증가와 출산지연 및 포기 등 결혼, 출산, 양육의 생애단계 이행을 멈추게 하는 주요 원인으로 작용하고 있다(박미선, 2017).

그뿐만 아니라 이들을 지원하는 부모의 안정적인 노후 대비를 위협하는 등 복합적인 사회문제와 연관되어 있다는 점에서 심각성이 더욱 커지고 있다(변금선 외, 2021). 물론 청년주거 문제를 해결하기 위해 정부는 임대주택 공급, 금융지원 등 각종 주거정책을 운영하고 있지만, 실제 청년이 체감하는 정책의 실효성 측면에서 여전히 부족한 점이 지적됨에 따라 무엇보다 청년의 현실적인 요구를 반영한 개선방안 모색이 요구되고 있다(김강산, 2021; 변금선 외, 2021).

이러한 배경에서 본 연구는 현재의 주요 이슈와 인사이트(insight)를 알아볼 수 있는 소셜 빅데이터를 활용하여 코로나19 발생 후 청년주거 관련 주요 이슈 및 변화를 종합적으로 파악하고 청년주거복지 정책 방안 모색에 대한 시사점을 제공하기 위한 목적으로 진행되었다. 구체적으로는 코로나19 발생 후 국내 온라인 대표 포털 네이버 뉴스 기사에서 나타난 최근 청년주거 핵심 단어(keyword)와 관련 주요 토픽(topic)을 추출하고 각각의 연결관계와 주요 특성을 살펴보고자 한다.

II. 이론적 고찰

1. 청년주거와 코로나19

청년은 청년기본법상 만 19세 이상에서 만 34세 이하의 사람으로 정의하고 있다. 생애주기 관점에서 이 시기는 아동기에서 성인으로 가는 중간이행기로서 교육을 마친 후 사회에서 자리를 잡고 다음 생애단계로 넘

어가는 시기이다. 그러한 점에서 취학, 취업, 결혼, 출산 등 개인과 가족생애에 변화가 이어지는 청년주거는 그동안의 생활 장소 즉 주거 여건이 달라지고 성공적인 독립, ‘이행’(transition)이 주요 쟁점이라고 할 수 있다(임덕영 외, 2017).

물론 아직 독립할 여력이 없는 청년이 부모와 거주하는 것은 자연스러운 일이겠지만 다음 단계로의 이행 시간이 장기화하면서 결혼 및 출산 지연이나 아예 포기하게 되는 비중이 높아져 문제가 되고 있다(박미선, 2017; 박시내, 2021). 청년주거 문제는 코로나19 이전부터 열악한 경제상황, 고용여건 악화, 주거비 부담, 주거상향 사다리 붕괴 등의 대응과 주거지원 모색에 관한 연구와 정책적 논의가 꾸준히 제기되어 왔다(고함 20 청년연구소, 2016; 임덕영 외, 2017). 그러나 최근 코로나19는 청년의 여건을 더욱 악화시키고 청년에게 실직과 경제적 불안정을 불러오면서 재정적 어려움과 주거불안을 키우고 있다(변금선 외, 2021). 경제침체로 인한 청년의 고용불안과 일자리 상실에 따른 재정적 어려움은 청년에게 주거의 불안정 증가와 내 집 마련을 더 어렵게 하고 있는데 이는 결국 청년 부모의 부담으로까지 이어지고 있다(김진영, 이현정, 2020; 김혜은 외, 2021). 아직 결혼하지 않은 청년의 주거비 과부담이 부모로부터 경제적 독립과 물리적 독립 모두 지연시킬 뿐만 아니라 청년이 다시 부모의 집으로 되돌아오도록 하는 것이다(연합뉴스, 2023).

한국보건사회연구원(2020)에서 실시한 청년층(만 19세~34세) 생활실태 및 복지욕구조사에 따르면 부모와 동거하는 청년(56.7%)이 혼자 독립한 청년(15.5%)보다 3배가 넘는 높은 비율을 보였다(한국보건사회연구원, 2020). 또한 한국 대표 취업포털 잡코리아(2021)에서 직장인, 아르바이트생, 취업준비생 2,327명을 대상으로 실시한 설문조사 결과를 보면, ‘코로나19로 인한 경제적 압박감이 상승했다(79.2%)’고 응답하였다. 그리고 응답자 절반은 ‘나는 캥거루족이다(49.0%)’라고 하였으며 그중 3명 중 1명이 ‘코로나19 이후 캥거루족이 됐다(32.2%)’고 하는 등 청년의 주거실태가 더 어려워짐에 따른 비자발적인 부모의존 상태가 되었다는 것을 알 수 있다(잡코리아, 2021).

독립적으로 생활하는 청년가구도 임대료 및 대출금 상황에 대한 주거비 부담을 느끼는 비율(74.8%)이 일반가구(65%)에 비해 높았고(국토교통부, 2021) 대부분 임차(81.6%)로 거주하고 있었다. 일반가구와 비교했을

때 청년가구의 최저주거기준 미달 가구 비율(7.9%)이 일반가구 비율(4.5%)보다 높게 나타났다. 또한 1인당 주거면적은 일반가구(33.9㎡)보다 좁은 30.4㎡ 수준으로 나타나 상대적으로 열악한 주거환경에서 지내고 있었다(국토교통부, 2022).

반면 코로나19 확산방지 조치로 인해 사람들은 집에서 더 많은 시간을 보내게 되고 가족 구성원 간 함께 보내는 시간이 증가하면서 이전에 없었던 가족 간의 갈등이 빚어지고 있다. 2020년 긴급재난지원금 설문조사 결과에 따르면 응답자 절반 이상(55.8%)이 코로나19 이후 가족과 보내는 시간이 증가했다고 응답했는데 그 중 1/4은 가족관계 만족도가 하락한 것으로 나타났다(김지현, 2021). 그리고 코로나19로 인한 가구소득 감소, 가구원수 증가, 사회적 관계 단절, 가사노동, 돌봄부담 등은 가족관계를 불만족스럽게 하는 요인으로 나타났다(보건복지부, 2020).

그럼에도 불구하고 비대면을 통한 원격 활동이 증대됨에 따라 비싼 도시지역보다 밀도가 낮은 외곽지역의 저렴한 주택으로 주거이동이 가능하게 되었고 학업과 업무를 할 수 있는 비대면 공간은 물론 집 밖에서 하던 운동이나 취미활동이 가능한 새로운 주거공간에 대한 요구도 다양해지고 있다(R2Korea, 2019). 그러나 생애단계를 고려한 이행기 측면에서 바람직한 지역의 제한된 주택 경쟁과 외곽지역의 저렴한 옵션 부족 등은 여전히 해결해야 할 문제로 남아 있다(변금선 외, 2021).

정부도 이러한 청년의 주거문제 해결을 위해 임대주택 공급과 다양한 지원 제도 개선방안을 모색하고 있으나 여전히 지원 대상과 범위가 제한적인 부분이 있다(김다솜, 2023). 청년 주거정책 지원이 활성화되기 위해서는 무엇보다 주거 관련 법률 및 제도와 관련한 전문지식에 대한 정보를 제공해주는 교육 프로그램 개발과 제공 방법에 대한 지원이 중요하며(문소희, 이현정, 2019), 청년이 손쉽게 정보에 접근할 수 있는 모바일 기반 지원 루트 확대가 제기되었다(김강산, 2021). 더욱이 최근 코로나19로 더욱 어려워진 청년의 현실적인 생활 여건을 고려했을 때 사회 변화에 구조적으로 대응할 수 있는 선제적 지원방안이 요구되고 있다(변금선 외, 2021).

2. 소셜 빅데이터와 의미 네트워크 분석

소셜 빅데이터는 개인과 커뮤니티가 소셜 미디어 플랫폼, 웹 사이트, 모바일 앱 및 기타 온라인에서 매일 생성하는 텍

스트, 이미지, 비디오 및 사용자 생성 콘텐츠와 같은 광범위한 비정형 정보가 포함된 방대한 양의 데이터를 의미한다(황승구 외, 2013; Kitchin, 2014; Mayer-Schönberger & Cukier, 2013). 소셜 빅데이터의 규모와 내용 면으로 볼 때 소셜 미디어 사용자가 자기 생각과 의견을 자유롭게 생산하면서 쌓이는 다양하고 방대한 데이터와 일반 설문 조사에서는 확인할 수 없는 숨어있는 현상을 파악할 수 있다는 점에서 새로운 자료 수집 방법으로 의미가 있다(송태민, 2015; 이오준 외, 2014; 황승구 외, 2013).

빅데이터 분석은 생성된 데이터 수집, 처리, 분석과정과 같은 텍스트 마이닝(text mining) 과정을 거쳐 시각화로 이어지는데 인터넷 게시물이나 문서와 같은 비정형 데이터에 대한 마이닝 과정을 거친다. 마이닝은 데이터에서 통계적으로 의미가 있는 개념이나 특정 간 패턴이나 동향 등의 정보를 추출하는 과정이다(Hearst, 2003). 인터넷상의 웹 문서가 대표인 데이터이며, 하이퍼 링크(hyperlink)를 다운로드하는 웹 크롤러(web crawler) 프로그램을 통해 페이지에서 데이터(data)를 수집하는 웹 크롤링(web crawling) 작업으로 수집된다.

텍스트 마이닝의 하나인 의미 네트워크 분석(semantic network analysis)은 네트워크의 구조와 노드(node) 간의 관계 분석을 위해 다양한 척도를 적용할 수 있으며 측정 방법에 따라 연결 중심성, 매개 중심성, 고유 벡터 중심성, 모듈성 등으로 네트워크의 특성을 파악하고 시각화할 수 있다(이호, 2019; 황동열, 황고은, 2016). 연결 중심성은 노드와 연결된 연결 또는 에지(edge)의 수를 반영하고 있으며 중심성이 높은 노드는 네트워크 내에서 더 중요하고 많은 영향력을 가질 수 있다. 매개 중심성은 노드가 네트워크의 다른 노드 사이에서 최단 경로에 있는 정도를 나타내며 매개 중심성이 높은 노드는 네트워크 내의 서로 다른 클러스터 또는 커뮤니티 간의 매개 역할을 할 수 있다. 고유 벡터 중심성은 이웃 노드 영향을 반영하며 고유 벡터 중심성이 높은 노드는 네트워크 내에서도 중요한 다른 노드에 연결될 수 있다. 모듈성은 네트워크를 별개의 클러스터 또는 커뮤니티로 나눌 수 있는 정도를 반영하여 모듈성 값이 클수록 네트워크가 고도로 모듈화되고 고유한 하위 그룹 또는 클러스터를 포함할 수 있다. 이 중 연결 중심성은 네트워크에서 가장 많이 연결된 노드를 식별하는 데 사용되며 매개 중심성은 서로 다른 클러스터 간의 브리지 역할을 하는 노드를 식별하는 데 사용된다. 이러한 의미 네트워크 분석은 측정된 단어들의 연결구조를 시각적으로 나타내는 특성을 가짐으로써 개

념 간의 관계에 대한 통찰력을 제공하고 데이터의 패턴과 추세를 식별하는 데 도움이 될 수 있다(이호, 2019).

3. 관련 선행연구 동향

최근 주거분야에서 빅데이터를 활용한 사례를 살펴 보면 공유주택에 대한 인식(김도연, 2022), 아파트 단위 평면 분석(맹호영, 현경훈, 2021), 1인당 주거면적 추정(임재빈, 이상훈, 2020), 주거이동의 공간적 특성(이창효, 2020)과 같은 주거공간에 대한 인식과 주택 특성을 파악한 연구가 있었다. 그리고 주택의 경제적 특성과 관련한 주제로 서울시 성동구를 대상으로 한 코로나19와 지역경제와의 관계(신혜영 외, 2021), 사회 문제가 된 가계부채 이슈(박정민, 송태민, 2020), 아파트 거래량에 미치는 영향(김윤식, 최민섭, 2018), 전세 가격 예측(이종민 외, 2017) 등의 빅데이터 활용 연구가 있었다. 그 밖에 코로나19와 관련한 가족생활에 대한 이슈(박선영, 이재립, 2021), 노인 주거복지 연구 동향 분석 연구(황윤서, 이현정, 2022)와 주거 트렌드 및 디자인 동향을 살펴본 보건복지 분야의 정책관련 키워드 분석(정영철, 천미경, 2021), 국내 학술지에 게재된 주거연구 동향(이소연, 김영주, 2020) 및 국내외 주거디자인 동향(왕택우, 조지영, 홍이경, 2021), 부동산 시장 트렌드(R2Korea, 2019)등을 분석한 연구가 있었다.

청년가구와 관련하여 온라인 뉴스나 플랫폼을 활용한 연구로는 1인가구 인테리어 트렌드 분석 연구(카리사나타샤 외, 2021)와 코로나 19 발생 전후 청년 1인가구 주생활행위 변화를 분석한 연구(김태완, 장미선, 2021)가 있었으나 자료의 수집과 분석 과정에서 웹 크롤링이나 텍스트 마이닝 과정을 거치지 않고 연구자가 온라인에서 검색·수집한 자료를 내용분석하여 주요 특성을 살펴본 연구였다. 분석 내용도 원룸이나 오피스텔의 인테리어 트렌드와 주거공간 계획을 위한 주생활행위 변화에 맞춰져 있었다.

이와 같이 실제로 빅데이터를 활용한 청년주거 관련 이슈나 동향에 관한 연구는 찾아볼 수 없었다. 이에 본 연구는 청년주거와 관련된 소셜 빅데이터가 실시간으로 업로드되는 국내 대표 인터넷 포털 네이버를 통해 수집된 빅데이터의 관계 구조와 숨겨진 인사이트(insight)를 살펴보고자 한다.

III. 연구방법

1. 분석대상 및 자료수집

본 연구는 코로나19 발생 시점인 2020년 2월부터 2022년 2월까지 약 2년 동안 국내 대표 포털사이트 네이버의 온라인 뉴스에서 청년주거 관련 기사를 분석 대상으로 하였다. ‘청년주거’와 ‘청년월세’를 검색 키워드로 설정하여 해당 용어가 포함되는 모든 기사를 수집하였다. 자료를 수집할 때 웹 크롤링 방법을 적용하였으며 네이버 뉴스 제목과 기사 내용에 ‘청년주거’와 ‘청년월세’ 키워드가 들어간 문서에서 명사와 형용사를 추출하였다. 자료를 분석하기 전 불필요한 정보가 포함되지 않도록 불용어를 제거하고 필터링을 하였다. 제거한 불용어는 뉴스 기사에서 자주 등장하는 신문사명과 뉴스 용어(‘네이버’, ‘저작권자’, ‘서울경제’, ‘뉴스’, ‘무단’, ‘전재’, ‘재배포’, ‘채널’, ‘머니’, ‘금지’, ‘칼럼’, ‘구독’, ‘매일’, ‘매일경제’, ‘세계일보’, ‘일보’, ‘세계’, ‘조선’, ‘중앙’, ‘바로가기’, ‘뉴스시’, ‘언론’, ‘모바일’, ‘환경’, ‘속보’, ‘투데이’, ‘기자’, ‘메일’ 등)이다. 분석자료는 정부, 교육 및 연구기관 등 다양한 국내외 기업에서 AI 빅데이터 분석을 전문으로 하는 빅데이터 분석 전문가에 의뢰하여 필터링 되었으며 최종 집계된 11,609개의 단어를 자료 분석에 활용하였다.

2. 자료의 분석

최종 집계된 단어는 토픽모델링을 위한 문서(document)로 보고 각 문서별 추출된 명사를 이용하여 문서-단어 빈도 매트릭스를 구축하였다. LDA 토픽 모델에는 파이썬 gensim 패키지를 이용하였으며 학습 파라미터 값은 배치 사이즈 = 300, 학습반복횟수 = 50, gamma = 0.001로 설정하였다. 하이퍼파라미터로는 alpha = 0.03, eta = 0.01로 지정하였고 기타 다른 파라미터들은 파이썬 gensim 패키지에서 기본적으로 권장하는 파라미터를 사용하였다. 토픽개수를 지정하기 위해 perplexity 지수를 사용하여 분석을 실행하였다. 이렇게 정제된 단어는 단어의 빈도와 가중치를 함께 고려한 TF-IDF(term frequency-inverse document frequency, 단어 빈도·역문서 빈도) 값을 반영하였으며 분석 결과의 주요 특징을 시각적으로 파악할 수 있는 워드 클라우드(word-cloud) 표현 기법과 네트워크 분석을 시행하였다. 자료 분석의 주요 절차는 [그림 1]과 같다.

IV. 연구결과

1. '청년주거' 관련 TF-IDF 핵심 단어

네이버 뉴스에서 등장한 '청년주거' 관련 핵심 단어 순위를 사용 빈도와 중요도에 따라 가중치가 높게 나타난 TF-IDF 값으로 정리하면 <표 1>과 같다.

TF-IDF 값이 가장 큰 단어는 '청년', '지원', '주거', '월세', '주택', '사업', '대출', '소득', '신청', '급여', '정책', '서울' '대상', '기준', '보증금', '지급' 등으로 청년주거 지원사업 대상 기준을 연상하는 단어로 나타났다. 이어서 '거주', '주거비', '부부', '공급', '부모', '부담', '확대', '신혼', '임차', '안정', '지역', '전세', '임대', '대책', '코로나', '복지', '일자리', '창업',



[그림 1] 데이터 분석 절차

<표 1> 청년주거 관련 등장 단어의 TF-IDF 순위(상위 100개)

순위	단어	TF-IDF	순위	단어	TF-IDF	순위	단어	TF-IDF	순위	단어	TF-IDF
1	청년	295.5895	26	공급	47.3044	51	계획	34.02522	76	예산	28.09367
2	지원	206.8562	27	후보	47.26409	52	추진	33.88672	77	완화	27.98104
3	주거	131.5852	28	부모	47.21376	53	특별	33.20917	78	공간	27.35705
4	월세	119.1252	29	부담	44.76577	54	사회	32.9728	79	시행	27.03459
5	주택	115.1393	30	확대	43.67977	55	세대	32.74375	80	국민	26.56383
6	사업	92.88817	31	신혼	43.0735	56	제공	32.7246	81	무주택	26.01002
7	이하	91.57162	32	임차	42.58833	57	청년층	32.28782	82	발표	25.51752
8	가구	87.931	33	안정	41.84624	58	생활	32.0221	83	자녀	24.86843
9	대출	82.92175	34	올해	41.27755	59	분리	32.00158	84	도시	24.77357
10	소득	81.51909	35	지역	40.78442	60	센터	31.948	85	국토	24.71302
11	신청	77.15186	36	중위	40.69779	61	마련	31.11879	86	제도	24.50876
12	급여	76.78521	37	부산	40.60224	62	공공	30.68314	87	국회	24.26182
13	정책	72.23414	38	선정	40.40164	63	공사	30.66374	88	대표	23.93277
14	서울	71.816	39	전세	40.25115	64	시장	30.54226	89	이번	23.83307
15	대상	61.66542	40	임대	39.66688	65	이자	30.51792	90	부동산	23.53668
16	기준	60.99421	41	대책	38.60975	66	모집	30.41395	91	지난해	23.36603
17	서울시	58.819	42	코로나	37.53742	67	예정	30.07625	92	온라인	23.36306
18	보증금	55.30724	43	복지	37.44155	68	부산시	30.04991	93	평균	23.0776
19	지급	53.76838	44	대상자	37.13532	69	취업	30.03944	94	자금	22.90385
20	경제	53.20879	45	일자리	35.56888	70	가능	29.9037	95	의원	22.82537
21	거주	53.18314	46	창업	35.50249	71	수급	29.75998	96	확인	22.72586
22	주거비	52.49541	47	금융	35.42746	72	배포	29.45587	97	규모	22.64069
23	최대	51.41985	48	경우	35.3916	73	문제	29.00826	98	분야	22.40607
24	부부	49.58185	49	내년	35.04405	74	참여	28.49453	99	기간	22.3536
25	정부	48.84431	50	이상	34.28766	75	조사	28.17941	100	미혼	22.26647

‘금융’ 등과 같은 청년을 포함한 청년과 관계된 대상인 신혼부부, 부모의 주거불안에 대한 사회적 분위기를 보여주는 단어가 모여 있었다. 그 외에 ‘계획’, ‘추진’, ‘제공’, ‘센터’, ‘마련’, ‘모집’, ‘예정’, ‘취업’, ‘가능’, ‘수급’, ‘배포’, ‘참여’와 같은 주거지원과 관련된 단어가 나타났으며, 이와 유사한 특성을 보여주는 ‘예산’, ‘완화’, ‘공간’, ‘시행’, ‘무주택’, ‘제도’, ‘부동산’, ‘규모’, ‘분야’ 등의 주거지원 복지제도와 같은 맥락을 보이는 단어가 이어졌다.

TF-IDF 순위에 따른 단어를 워드클라우드 기법으로 시각화한 결과는 아래와 같다[그림 2]. TF-IDF 값이 클수록 글자의 이미지가 크게 나타나며 <표 1>에서 살펴본 것과 마찬가지로 ‘청년’, ‘지원’, ‘주거’, ‘월세’, ‘주택’, ‘사업’, ‘소득’, ‘급여’ 등의 청년주거 지원사업과 관련된 이미지가 두드러지게 나타났다.

2. ‘청년주거’ 관련 토픽 모델링 결과 및 네트워크 시각화

특정 노드가 일정 커뮤니티와는 가깝고 타 커뮤니티와는 거리가 먼 정도를 나타낸 지수로써 네트워크상의 커뮤니티를 구분하는 모듈성 지표를 기준으로 네트워크 분석을 실행하였다. 그 결과 ‘청년주거’와 관련된 핵심 단어는 총 399개의 단어가 연관된 것으로 나타났다. 각 단어 간의 연결정도와 단어 간 매개 축을 기준으로 상대적으로 연결관계가 높은 단어들을 통해 주요 토픽

의 특징을 파악한 후 이를 토대로 3개의 대분류와 13개의 소분류 토픽을 명명하였다<표 2>.

먼저 대분류 첫 번째 범주에 해당하는 1) ‘수요 및 공급적 요인’의 소분류 토픽 (1) ‘지원 기준’을 살펴보면, ‘사업’, ‘이하’, ‘이상’, ‘기준’, ‘최대’, ‘신청’, ‘주거비’, ‘보증금’, ‘중위’ 등의 단어들이 한 군집을 이루고 있었다. 주로 청년주거 지원사업을 반영한 내용으로 지원 기준과 관련된 단어가 포함되어 있었다. 토픽 (2) ‘지원 대상’은 ‘임차’, ‘선정’, ‘지급’, ‘시행’, ‘대상자’, ‘참여’, ‘수급’, ‘모집’ 등과 같은 청년주거 지원사업의 시행을 위한 대상과 관계된 단어가 주를 이루고 있었다. 토픽 (3) ‘지원 방안’은 ‘홈페이지’, ‘혜택’, ‘접수’, ‘방문’, ‘등록’ 등의 단어가 나타나 청년주거 지원사업 대상자에게 지원 제공하는 방법에 대한 것으로 확인되었다.

다음으로 1) ‘수요 및 공급적 요인’과 2) ‘경제적 요인’이 교차하는 토픽 (4) ‘주거문제’의 경우 ‘주거’, ‘지원’, ‘주택’, ‘월세’, ‘어려움’, ‘공간’, ‘준비’, ‘취약’ 등으로 청년의 주거 어려운 여건을 반영한 단어가 포함되어 있었다. 토픽 (5) ‘지원 활동’에서는 ‘최근’, ‘노력’, ‘상황’, ‘기관’, ‘국가’, ‘도입’, ‘적극’, ‘활동’, ‘기회’, ‘증가’, ‘이용’, ‘조정’, ‘방식’, ‘검토’, ‘활용’, ‘발굴’, ‘조성’ 등과 같은 청년주거 개선을 위한 정부의 지원 활동과 연관된 단어들이 많이 있었다.

이어서 2) ‘경제적 요인’ 범주에 포함되는 토픽 (6)



[그림 2] ‘청년주거’ 관련 TF-IDF 핵심 단어 워드클라우드

‘행정 관리 및 분석’은 ‘분석’, ‘행정’, ‘임대차’, ‘기반’, ‘구성’, ‘데이터’, ‘상승’, ‘협업’, ‘재산’, ‘합산’, ‘포털’ 등으로 주택시장 관련 자료 분석과 행정적 관리 요소가 포함된 단어가 주를 이루고 있었다. 토픽 (7) ‘금융’에 해당하는 핵심 단어는 ‘연간’, ‘계약’, ‘보증’, ‘수요’, ‘금리’, ‘상당’, ‘서류’, ‘신청자’, ‘은행’, ‘상품’ 등으로 주택과 관련된 금융 정책과 프로그램을 포함하여 재무 계획·관리와 연관된 단어가 많았다. 그리고

‘무이자’, ‘전월세’, ‘저소득’과 연결 지어 볼 때 취약 계층에 대한 지원에 맞춰져 있는 것으로 짐작해 볼 수 있다. 토픽 (8) ‘지원 공약’에서는 ‘프로그램’, ‘토지’, ‘비용’, ‘현실’, ‘청약’, ‘오피스텔’과 같은 부동산 정책이나 개발과 관련된 용어와 함께 ‘장관’, ‘공약’, ‘위원’, ‘발언’, ‘과제’, ‘방향’ 등의 정부, 정당의 공약이 반영된 단어가 포함되어 있었다. 이는 청년층의 주택 부족과 부적절한 주택 문제를 해결하기 위한 정부의

〈표 2〉 ‘청년주거’ 관련 네트워크 토픽 분류 및 핵심 단어

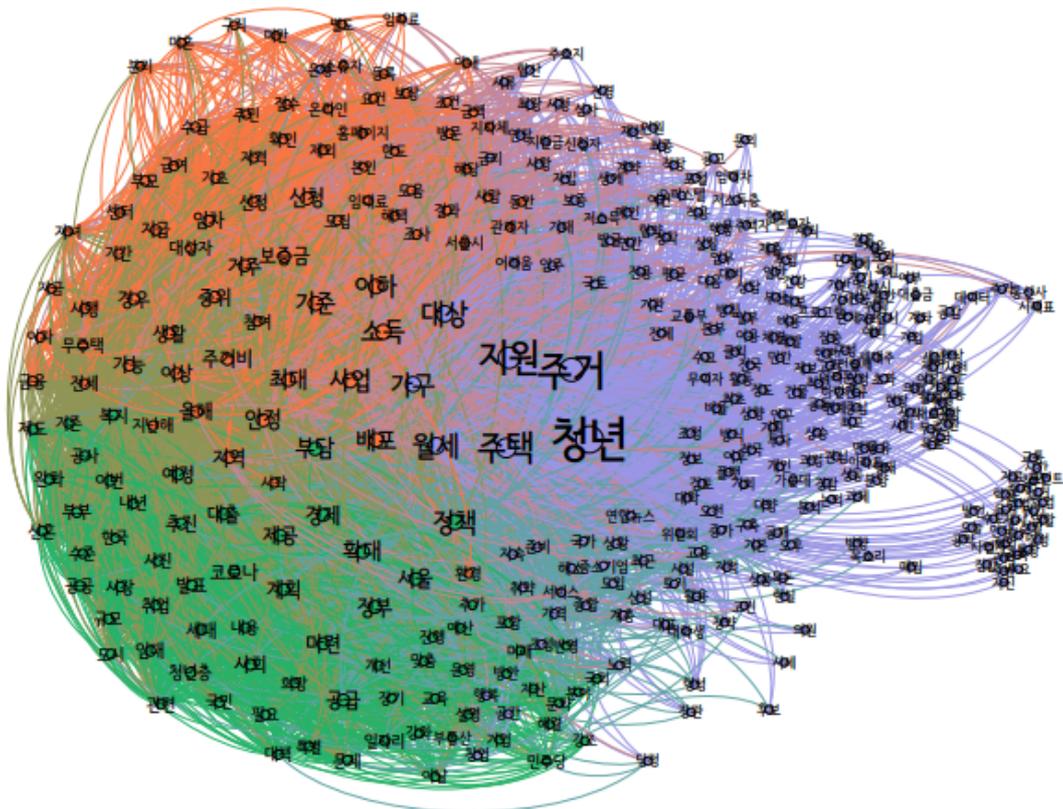
대분류	소분류	핵심 단어
1) 수요 및 공급적 요인	(1) 지원 기준	소득, 사업, 이하, 배포, 기준, 최대, 신청, 안정, 거주, 주거비, 지역, 보증금, 중위, 올해, 이상, 가능, 생활, 경우
	(2) 지원 대상	임차, 선정, 지급, 시행, 대상자, 참여, 부모, 전세, 센터, 지난해, 무주택, 기간, 금융, 수급, 확인, 급여, 모집, 조사, 시작, 서울시, 제외, 자격, 환경, 주민, 기초, 도움, 결과, 온라인, 이자, 본인
	(3) 지원 방안	홈페이지, 임대료, 자녀, 관계자, 혜택, 접수, 사람, 동안, 자금, 해당, 방문, 보장, 등록, 요건, 한도, 미혼, 분리, 미만, 조건, 금액, 구직, 별도, 소유자, 지자체
	(4) 주거 문제	주거, 청년, 지원, 주택, 월세, 대상, 가구, 어려움, 공간, 지속, 준비, 취약, 종합, 국토, 해소
	(5) 지원 활동	최근, 노력, 대표, 상황, 기관, 국가, 입주, 위원회, 전체, 사항, 행복, 계층, 자립, 정보, 교통부, 도입, 적극, 메인, 오진, 중소기업, 활동, 이후, 신실, 대학생, 기회, 대학, 지원금, 증가, 이용, 전용, 평균, 조정, 방식, 검토, 활용, 시설, 발굴, 가격, 서비스, 창업, 조성, 기대
2) 경제적 요인	(6) 행정 관리 및 분석	최종, 투입, 분석, 상황, 고려, 행정, 임대차, 기반, 후보, 구성, 단체, 건물, 시민, 데이터, 상승, 협업, 재산, 제출, 연계, 합산, 기사, 근로자, 관심, 현장, 집중, 저소득층, 예비, 재직, 하반기, 일정, 논의, 공고, 생각, 자리, 무주택자, 포털
	(7) 금융	전국, 연간, 비율, 계약, 보증, 수요, 제보, 저소득, 금리, 본부, 플랫폼, 부산, 여건, 상당, 생애, 시정, 다음, 최장, 기본, 서류, 다양, 신청자, 기획, 고용, 구축, 투자, 고민, 협약, 연장, 지적, 실질, 최초, 업무, 일부, 가운데, 방침, 의원, 대비, 공개, 심사, 과정, 이내, 형성, 오후, 정도, 적용, 민간, 은행, 인구, 임차료, 무이자, 한련, 개발, 직장, 확보, 공동, 체결, 최고, 전문, 공유, 신규, 상품, 문의, 주소지, 경기, 아파트, 중심, 의견, 목돈, 당정, 전월세
	(8) 지원 공약	개인, 최선, 정착, 프로그램, 토지, 건강, 비용, 현실, 청약, 전경, 오피스텔, 장관, 일반, 공감, 인원, 공약, 위원, 면적, 실제, 단계, 발언, 형태, 과제, 부족, 오늘, 이유, 세상, 부산시, 방향, 실시, 지급
	(9) 생애 전환	가족, 성공, 불안, 분양, 활성, 경감, 지방, 연령, 목표, 시세, 처음, 독립, 새로운, 가입, 제안, 수도, 결혼, 목소리, 사업자, 효과, 향후, 집값, 전환, 통신사, 학생, 세대주
	(10) 계획 및 의사결정	소통, 유지, 자료, 인상, 계좌, 여부, 회의, 시세표, 주가, 향상, 사전, 대출금, 관리, 현재, 체시, 방법, 주변, 초과, 평가, 자신, 충족, 교통, 시간, 주요, 참석, 선거, 통장, 선발, 위기, 포인트, 학업, 결정, 역할, 매입
3) 사회적 요인	(11) 위기대응	정책, 부담, 서울, 경제, 확대, 제공, 정부, 마련, 계획, 예정, 공급, 코로나, 추진, 사회, 대출, 이번, 세대, 복지, 발표, 사진, 청년층, 임대, 부부, 관련, 문제, 취업
	(12) 공공 지원	완화, 시장, 내년, 공공, 일자리, 규모, 공사, 제도, 신혼, 필요, 기존, 내용, 한국, 도시, 수준, 국민, 대책, 희망, 진행, 설명, 개선, 민주당, 운영, 미래, 교육, 추가, 특별, 부동산, 맞춤형, 해결, 예산, 반영, 장기, 방안, 강화
	(13) 민간 지원	기업, 포함, 자산, 강조, 국회, 문화, 분야

지원 공략에 맞춰진 것으로 이해할 수 있다. 토픽 (9) ‘생애 전환’에서는 ‘가족’, ‘성공’, ‘분양’, ‘연령’, ‘독립’, ‘결혼’, ‘전환’, ‘학생’, ‘세대주’ 등이 핵심단어로 나타났다. 주로 생애주기에서 주택시장에 처음 진입하는 청년의 독립과 결혼, 주택 마련에 대한 생애 전환기의 상황을 반영하는 단어가 주를 이루고 있었다. 토픽 (10) ‘계획 및 의사결정’에 따른 핵심 단어는 ‘소통’, ‘유지’, ‘회의’, ‘관리’, ‘제시’, ‘방법’, ‘평가’와 같은 계획에 필요한 관리와 운영에 관한 맥락을 보였으며 ‘자신’, ‘충족’, ‘주요’, ‘참석’, ‘선거’, ‘선발’, ‘결정’, ‘역할’과 같은 개인이나 공동체에서의 의사 결정과 관련된 단어가 포함되어 있었다.

마지막 대분류 3) ‘사회적 요인’은 3개의 소분류로 토픽이 세분화되었다. 순서대로 토픽 (11) ‘위기대응’에 포함된 핵심단어를 보면 ‘정책’, ‘부담’, ‘확대’, ‘제공’, ‘정부’, ‘마련’, ‘계획’, ‘예정’, ‘공급’, ‘코로나’, ‘추진’, ‘복지’, ‘문제’, ‘취업’ 등의 단어가 포함되어 있었다. 이는 코로나 확산에 따른 사회적 위기 상황에 대

응하기 위한 정부의 지원 대책이 반영된 것으로 보여진다. 토픽 (12) ‘공공 지원’에서는 ‘완화’, ‘공공’, ‘일자리’, ‘대책’, ‘개선’, ‘민주당’, ‘운영’, ‘맞춤’, ‘해결’, ‘예산’, ‘반영’, ‘장기’, ‘방안’, ‘강화’와 같은 정부의 공적 지원 제도와 관련이 있는 단어가 많이 나타났다. 그리고 ‘민주당’, ‘예산’, ‘장기’, ‘방안’과 같은 단어가 언급된 것은 정부의 정책 방안과 개선에 대한 부분에 좀 더 집중되어 있음을 알 수 있다. 토픽 (13) ‘민간 지원’에서는 ‘기업’, ‘포함’, ‘자산’, ‘국회’, ‘문화’, ‘분야’가 핵심단어로 나타났다. 민간 기업의 자산을 활용한 사회적 기여에 대한 맥락에서 볼 때 사회의 다양한 영역에서 정부, 소비자, 비영리조직에 이르기까지 기업과 상호 작용을 한다는 점에서 민간 지원으로 연결되었다.

이상의 ‘청년주거’ 관련 토픽 모델 분석을 네트워크 그래프로 시각화한 결과는 [그림 3]과 같다. 문자(노드) 크기는 해당 네트워크상에서 상대적인 고유 벡터 중심성의 값을 나타낸 것으로 문자 크기가 클수록 중요한



[그림 3] 네트워크 분석 결과 시각화

- 1) 수요 및 공급적 요인 2) 경제적 요인 3) 사회적 요인

단어임을 의미한다. 문자를 연결하는 선(edge)은 동시에 등장하는 단어 간의 관계를 표현한 것으로 선이 굵고 면적이 커질수록 높은 연결정도(Degree)를 가진다.

V. 결론

1. 결과의 요약 및 제언

본 연구는 소셜 빅데이터를 활용하여 최근 청년주거 관련 주요 이슈를 파악하고자 한 것으로서 코로나19 발생 후 최근 네이버 온라인 뉴스 기사에서 나타난 ‘청년주거’ 관련 핵심 단어와 그와 관련된 주요 토픽 특성을 살펴보았다. 주요 결과에 따른 시사점을 요약 정리하면 다음과 같다.

첫째, 코로나19 발생 후 ‘청년주거’ 관련 TF-IDF 값이 가장 큰 단어는 ‘청년’, ‘지원’, ‘주거’, ‘월세’, ‘주택’, ‘사업’, ‘대출’, ‘소득’, ‘신청’, ‘급여’, ‘정책’, ‘서울’, ‘대상’, ‘기준’, ‘보증금’, ‘지급’과 같은 ‘청년주거 지원사업’과 관련이 높은 것으로 나타났다. 그리고 청년과 가족의 ‘주거불안’을 연상하는 ‘주거비’, ‘부부’, ‘공급’, ‘부모’, ‘부담’, ‘확대’, ‘신혼’, ‘임차’ 등의 단어가 있었고, ‘청년주거 복지제도’와 관련된 ‘계획’, ‘추진’, ‘제공’, ‘예산’, ‘완화’, ‘공간’, ‘시행’, ‘무주택’, ‘제도’와 같은 단어가 상위 순위로 나타났다. 이는 청년의 열악한 주거상황을 보여주는 단어로 청년의 주거 불안이 청년 본인은 물론 부모에게도 부담을 주고 있음을 시사한다. 그리고 정부도 청년주거 개선을 위해 제도적 방안을 추진하고 있음을 보여주었다.

둘째, 네트워크 분석 결과 단어 간 연결 특성에 따라 나타난 코로나19 발생 후 최근 ‘청년주거’ 관련 토픽은 총 13가지로 도출되었다. 각 토픽은 크게 1) ‘수요 및 공급적 요인’, 2) ‘경제적 요인’, 3) ‘사회적 요인’으로 3개의 대분류 범주에 포함되며 ‘수요 및 공급적 요인’은 TF-IDF 순위와 유사한 청년주거 지원사업과 관계된 (1) ‘지원 기준’, (2) ‘지원 대상’, (3) ‘지원 방안’에 대한 것으로 나타났다. 주로 청년의 주거안정을 위한 재정적 지원에 대한 것으로 주거지원에 대한 수요와 공급적 맥락에서 공통된 특성을 보였다.

‘경제적 요인’은 (4) ‘주거 문제’, (5) ‘지원 활동’, (6) ‘행정 관리 및 분석’, (7) ‘금융’, (8) ‘지원 공약’, (9) ‘생애 전환’, (10) ‘계획 및 의사결정’으로 가장 많은 하위

토픽 간 연결관계를 보였다. 전반적으로 주택 및 재무관리, 임대가격, 보조금 및 재정 지원 프로그램과 같은 내용이 반영되어 있었고, 주택 금융 관련 법적·제도적 규제 측면에 초점이 맞춰져 있었다. 이는 청년주거와 관련된 정부의 행정 및 재정 계획과 취약한 계층의 청년주거 지원으로 공통된 연결관계를 갖고 있었다.

‘사회적 요인’은 (11) ‘위기대응 관리’, (12) ‘공공 지원’, (13) ‘민간 지원’으로 사회적으로 이슈가 되었던 코로나 팬데믹으로 인해 직면한 문제 해결과 청년 일자리 창출 및 주거복지 확대 등과 같은 정부와 민간의 노력과 협업에 관한 내용이 반영되어 있었다.

셋째, 이상의 결과를 종합했을 때 최근 네이버 뉴스 기사를 통해 나타난 ‘청년주거’의 주요 이슈는 청년주거 지원사업과 관련된 것으로서 청년의 주거지원을 위한 수요 및 공급 요인과 경제적·사회적 요인에 따른 청년주거 복지정책 대안 및 지원에 대한 것으로 확인되었다. 대체로 청년주거의 열악한 상황을 개선하기 위해 주거불안 완화를 위한 정부의 지원 노력과 관련된 단어들이 많았다. 이는 청년주거가 사회적으로도 심각한 문제로 여겨지고 있으며 청년 개인에게 직면한 어려움은 물론 생애주기 전환, 그리고 가족에게까지 미치는 영향을 포괄적으로 고려해야 함을 시사하고 있다(김진영, 이현정, 2020; 김혜은 외, 2021, 변금선 외, 2021). 치솟는 주택 가격과 증가하는 월세는 안정적인 주택을 찾는 청년에게 큰 장벽이 되고 부모나 정부의 임대 지원보조금 없이는 적절한 주거의 공급과 수요가 어려운 현실을 보여주고 있다(잡코리아, 2021; 한국보건사회연구원, 2020). 주거지원의 연속성 보장은 취약한 청년의 독립적 생활에 매우 중요하다고 할 수 있는데 이를 위해 정부는 다양한 공공·민간 기관 및 단체와의 협력에 더욱 노력이 필요할 것으로 여겨진다. 단순한 주거공간을 넘어 교육, 건강, 고용기회 등 청년의 전인적 요구를 해결하는 포괄적인 지원 프로그램 제공과 이의 지속된 지원이 청년의 독립 준비와 생애 전환에 무엇보다 주력해야 할 부분이라 할 수 있다. 또한 코로나19와 같은 팬데믹 상황으로 주거 위기에 처한 청년의 경우 거주 연속성을 보장받을 수 있는 적절한 주거시설 접근의 확대로 주택 선택권의 가용성을 높이는 방안도 요구되어 진다.

이처럼 청년주거 문제 및 개선을 위한 여러 토픽이 도출되었으나 청년의 주거 선호나 요구를 파악할 수 있는 정보는 상대적으로 미비한 측면이 있었다. 청년에

게 선호되었던 모바일로 접근 가능한 법률·제도 및 주택 거래 관련 교육, 홍보, 정보 교류 루트(김강산, 2021; 문소희, 이현정, 2019)등과 연관된 도픽 역시 찾아보기 힘들었다. 주거와 관련된 문제를 효과적으로 해결하기 위해서는 정부기관, 금융기관, 공공·민간 주택개발 및 공급 기관, 지역사회 간의 협력 인프라 구축은 필수적이라 할 수 있다. 그리고 정보 공유를 통해 지원가능한 자원을 제대로 파악하고 계획·활용함으로써 지역 경제 활성화에도 기여할 수 있을 것이다. 따라서 신뢰할 수 있는 정보 제공과 서로의 경험을 공유할 수 있는 온라인 플랫폼 구축 마련에 정부의 노력과 관심이 필요한 것으로 여겨진다. 더불어 청년의 요구를 현실적으로 반영한 정책 마련을 위해서는 청년의 참여 기회를 장려하고 독려하는 청년정책의 제도 개선도 함께 변화되어야 할 것이다.

2. 연구의 한계점 및 후속연구 제안

본 연구는 코로나19로 인해 더욱 어려워지고 있는 청년 주거복지의 주요 이슈를 분석하고 청년주거 대안 모색을 위한 방향 제시를 위한 것으로서 소셜 빅데이터를 활용하여 포괄적인 정보를 탐색·수집하고 체계화할 수 있는 의미 네트워크 분석기법을 적용하였다는 점에서 의의가 있다.

하지만, 온라인 뉴스 기사는 청년 주거문제에 대한 최근 이슈를 파악할 수 있다는 점에서 유용한 소셜 데이터가 될 수 있으나 데이터의 수집 범위에 따라 일부 편견과 제한이 있을 수 있으므로 청년주거에 대해 일반화하기는 어려울 수 있다. 따라서 청년주거 문제를 객관적으로 파악할 수 있는 인구 통계 및 주거실태 데이터와 청년주거 관련 다양한 소셜 데이터 소스로 자료를 비교·분석하여 더욱 명확한 청년주거 이슈를 파악할 수 있을 것이다.

주제어: 청년주거, 주거문제, 빅데이터, 네트워크 분석, 네이버 기사

REFERENCES

R2Korea(2019). 2019 부동산트렌드 전망 보고서. 서울: R2Korea.

고함20 청년연구소(2016). *청년 연구과제 개발을 위한 기초연구: 연구동향의 메타분석을 바탕으로*. 서울: 서울시 청년허브.

과학기술정보통신부(2019). 보도자료: 빅데이터 플랫폼 및 센터 구축 사업 22개소 센터 선정, <https://www.korea.kr/briefing/pressReleaseView.do?newsId=156349219>에서 인출.

국토교통부(2021). 2020년도 주거실태조사 연구보고서 (특성가구). 세종: 국토교통부.

국토교통부(2022). 보도자료: 2021년도 주거실태조사 결과 발표, <https://www.korea.kr/briefing/pressReleaseView.do?newsId=156543700>에서 인출.

김강산(2021). 청년 주거정책의 현황과 개선과제. *NARS 현안분석*, 216, 1-18.

김다솜(2023). 청년 주거정책 인지도 40% 이상.. 수혜 경험은? 10% 못 미쳐, <http://www.dailypop.kr/news/articleView.html?idxno=66996>에서 인출.

김도연(2022). 빅데이터 분석을 통한 공유주택의 인식에 대한 연구. *한국주거학회 학술대회논문집*, 34(1), 167-170.

김동완(2013). 빅데이터의 분야별 활용사례. *경영논총*, 34, 39-52.

김윤식, 최민섭(2018). 뉴스 빅 데이터와 구글 정보 검색이 아파트 거래량에 미치는 영향에 관한 실증 연구. *주거환경*, 16(2), 411-426.

김지현(2021). 이슈특집: 코로나19는 가족에 어떤 영향을 미쳤을까?, 복지국가연구센터. http://www.welfarestate.re.kr/beluxe_dzQZ85/7345에서 인출.

김진영, 이현정(2020). 독립한 자녀의 주거비 지원에 대한 부모세대의 경험 및 의식에 관한 연구. *한국주거학회논문집*, 31(3), 99-107.

김태완, 장미선(2021). 코로나19 발생에 따른 청년 1인가구의 주생활행위 변화. *한국주거학회논문집*, 32(6), 81-90.

김혜은, 정소현, 이강민, 황윤서, 이현정(2021). 청년의 부모의존동거에 대한 청년과 부모세대의 경험과 인식. *한국생활과학회지*, 30(6), 1017-1029.

맹호영, 현경훈(2021). 딥러닝 기반 아파트 단위 평면 빅데이터 분석 연구-1970년부터 2020년까지의 한국 아파트 중심으로. *한국실내디자인학회논문집*, 30(3), 65-76.

문소희, 이현정(2019). 청년의 주거 정보·교육·상담 프

- 로그랩 요구도 기초연구. *한국주거학회논문집*, 30(6), 115-125.
- 박미선(2017). 청년의 주거지원 필요성과 정책지원 방향. *국토*, 432, 26-32.
- 박선영, 이재림(2021). 소셜 빅데이터로 알아본 코로나 19와 가족생활: 토픽모델 접근. *한국콘텐츠학회 논문지*, 21(3), 282-300.
- 박시내(2021). 저(低)혼인 시대, 미혼남녀 해석하기. *KOSTAT 통계플러스*, 13, 24-47.
- 박정민, 송태민(2020). 소셜 빅데이터를 이용한 가계부채 트렌드 및 강신호와 미래신호의 탐색. *한국사회복지학*, 72(1), 35-61.
- 변금선, 김승연, 박민진, 이해림(2021). '2025 서울시 청년정책' 비전과 전략. 서울: 서울연구원.
- 보건복지부(2020). *국내외 사회보장 지원정책 분석 연구*. 세종: 보건복지부.
- 송태민(2015). 소셜 빅데이터 분석과 활용 방안: 메르스 정보확산과 위험예측을 중심으로. *보건복지포럼*, 9, 29-49.
- 신혜영, 박소현, 강성호, 이금숙(2021). 정형·비정형 빅데이터 분석을 통한 COVID-19와 지역 경제의 관계-서울시 성동구를 대상으로. *대한지리학회지*, 56(4), 421-435.
- 연합뉴스(2023). [이슈+] 청년 60% '캥거루족'... "독립계획 없다", <https://www.yonhapnewstv.co.kr/news/MYH20230308011100641>에서 인출.
- 왕택우, 조지영, 홍이경(2021). 동시출현단어 분석을 이용한 노인주거디자인 국내외 연구동향 비교-Web of Science Core Collection과 KCI-Korean 저널 데이터베이스를 중심으로. *한국주거학회논문집*, 32(6), 47-57.
- 이소연, 김영주(2020). 디지털 네트워크 분석을 통한 주거 연구 키워드 동향 분석. *한국주거학회논문집*, 31(3), 49-60.
- 이오준, 박승보, 정다울, 유은순(2014). 소셜 빅데이터를 이용한 영화 흥행 요인 분석. *한국콘텐츠학회 논문지*, 14(10), 527-538.
- 이종민, 이종아, 정준호(2017). 뉴스 빅데이터를 이용한 전세 가격 예측-토픽모형 분석을 중심으로. *부동산학보*, 69, 43-57.
- 이창효(2020). 공간 빅 데이터 분석을 활용한 연령계층간 주거이동의 공간적 특성(방향성) 차이. *국토계획*, 55(1), 98-111.
- 이호(2019). 빅데이터로 본 스마트 시티: 뉴스 네트워크 분석을 중심으로. *KISDI STAT Report*, 18-23, 1-18.
- 임덕영, 이경애, 진미윤, 이의중(2017). *청년 주거 문제와 정책 방안 연구*. 대전: 한국토지주택공사 토지주택연구원.
- 임재빈, 이상훈(2020). 공공 빅데이터를 활용한 1인당 주거면적 추정에 관한 연구-서울의 단독 및 다세대 주택을 중심으로. *지역연구*, 36(1), 51-67.
- 잡코리아(2021). 성인남녀 55.8% "나는 캥거루족". 잡코리아, https://www.jobkorea.co.kr/GoodJob/Tip/View?News_No=18634&schCtgr=0&schTxt=2566&Page=1에서 인출.
- 정영철, 천미경(2021). *2021년 소셜 빅데이터 기반 보건복지 이슈 동향 분석*. 세종: 한국보건사회연구원.
- 최현주, 최연철(2016). 빅데이터를 통해 바라본 유아 창의인성 교육 방안 연구. *어린이문학교육연구*, 17(4), 601-627.
- 카리사나타사, 최병숙, 박정아(2021). '오늘의 집' 앱을 중심으로 1인가구의 인테리어 트렌드 분석. *한국주거학회논문집*, 32(3), 11-20.
- 한국보건사회연구원(2020). *청년층 생활실태 및 복지욕구조사*. 세종: 한국보건사회연구원.
- 황동열, 황고은(2016). 빅데이터 기술을 활용한 인문콘텐츠 분야의 의미연결망 분석-2003년부터 2015년까지 인문콘텐츠학회 논문을 중심으로. *인문콘텐츠*, 43, 229-255.
- 황승구, 최완, 허성진, 장명길, 이미영, 박종열, 원희선, ... 김달(2013). *빅데이터 플랫폼 전략: 빅데이터가 바꾸는 미래 비즈니스 플랫폼 혁명*. 서울: 전자신문사.
- 황운서, 이현정(2022). 텍스트마이닝과 의미 네트워크 분석기법을 활용한 코로나 19 발생 이후 노인 주거복지 연구 동향. *한국주거학회논문집*, 33(5), 89-97.
- Hearst, M. A. (2003. 10. 17). What is text mining?, Retrieved from <https://people.ischool.berkeley.edu/~hearst/text-mining.html>
- Kitchin, R. (2014). Big data, new epistemologies and paradigm shifts. *Big data & society*, 1(1), 1-12.

Mayer-Schönberger, V., & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work, and think*. New York: Houghton Mifflin Harcourt.

Received 3 May 2023;
1st Revised 31 May 2023;
Accepted 6 June 2023