

토픽 모델링을 이용한 도시재생사업 동향 분석

정정운*, 정은희**

*강원대학교 도시재생융합학과

**강원대학교 글로벌인재학부 지역경제전공

e-mail:joungun09@naver.com

The Trend Analysis of Urban Regeneration Project using Topic Modeling

Jeong-Un Jeong*, Eun-Hee Jeong**

*Dept. of Urban Regeneration Convergence, Kangwon National University

**Dept. of Regional Economics, Kangwon National University

요약

본 논문에서는 도시재생의 동향을 분석하기 위해 정선군 재생과 마을호텔을 키워드로 2016년 1월부터 2020년 10월까지 중앙지, 경제지, 지역종합지, 전문지 등에서 377건의 기사를 수집하고, 각각의 기사에 대한 키워드를 추출하여 연관성을 분석하여 워드클라우드로 시각화하였다. 그 결과, 도시, 재생, 사업, 지역, 마을, 주민, 정선군, 지원, 정신, 호텔, 추진, 센터, 강원랜드, 조성, 관광, 문화, 폐광 등의 순으로 나타났으며, 정선군의 도시재생사업이 주민, 마을, 호텔 분야가 핵심을 이루고 있다는 것을 알 수 있다. 또한, LDA 알고리즘을 통한 토픽모델링을 진행하여 10개의 토픽에 대해 15개의 단어를 추출하였으며 연관성 중심으로 분석한 결과 마을호텔, 주민 공동체, 강원랜드 연계, 청년창업 등으로 정선군의 도시재생 동향을 파악할 수 있었다.

1. 서론

최근 도시재생 뉴딜사업에 대한 관심이 높아지고 있다. 전국 시·도 단위에서 초고령사회 및 지역쇠퇴가 가속화되며 2016년 최초 시범사업 지역을 시작으로 현재까지 330개의 사업이 진행되며 삶의 질 향상, 도시 활력 회복, 공동체 회복 및 사회 통합을 목적으로 활성화 계획이 수립되었다.

하지만 도시 재생 사업을 진행하면서 우리나라의 고질적 문제점도 드러났다. 그중 하나는, 도시재생 사업이 지나친 정부 주도로 이루어진다는 부분이다. 이에 민간 영리 부문 및 주민공동체 등의 참여가 제한되며, 도시재생의 주체가 되는 주민들의 의견이 충분히 반영되지 않는다는 비판이 제기되고 있다[1].

따라서 본 논문에서는 정선군의 마을호텔에 대한 도시재생 사업에 관련된 기사를 중앙지, 경제지, 강원도 지역종합지, 주요 방송사, 전문지 등에서 기사를 수집하여 토픽 모델링 기법으로 정선군의 도시재생사업 동향을 분석하고자 한다.

2. 선행연구

본 연구에서는 문헌 자료를 사용한 분석기법 중 토픽모델

링 기법을 사용하며, 이 기법은 수집된 로우데이터의 빅데이터화를 통한 자료구축 방식이다. 나아가 통계화된 데이터를 재가공하여 단어 네트워킹, 텍스트 마이닝 등으로 활용되어 광범위한 분야에 사용되는 해당 머신러닝 기술은 정보의 관계, 의미 등을 구조화하는 데 유용하다는 장점이 있다.

조경원, 우영운의 텍스트마이닝을 이용한 4차 산업 연구 동향 토픽 모델링에서는 한국학술지인용색인(KCI)에서 웹 스크래핑 프로그램을 통한 “4차산업” 키워드를 토대로 LDA알고리즘을 R언어로 구현하여 4차산업과 관련된 넓은 분야에서 변화에 영향을 미친 주요 기술들을 연구하였다[2]. 김용진, 손용훈의 뉴스데이터를 통한 LDA 토픽 분석으로 장수군 농촌 지역의 활성화 사업의 특징[3]에 대한 연구를 살펴보고, 국내 기록관리학 연구 동향 분석 등 이외에도 보건학, 컴퓨터공학, 관광학, 예술학, 정치학 등 다양한 분야에서 정보를 추출하는 데 사용되고 있다.

3. 연구방법

3.1 자료수집

정선군을 대상으로 한 데이터는 뉴스데이터 분석을 바탕으

로, 토픽모델링 기법 중 LDA 토픽모델링을 활용하여 정선군의 도시재생사업 동향을 분석할 것이며, 해당 알고리즘 기법은 자료 내 숨어있는 주제어를 찾아내며, 관찰된 변수를 추론하는 것을 목적으로 한다.

본 논문에서는 빅카인즈(BIGkinds)[4]를 이용하여 중앙지, 경제지, 지역종합지, 방송사, 디지털타임스, 전자신문 등 총 54곳에서 뉴스 데이터를 수집하였다. 데이터 수집 기간은 2016년 1월에서 2020년 10월까지 약 5년으로 지정하였고, 검색어는 “정선군 재생, 마을호텔”로 설정하였다. 그 결과 총 377건의 기사를 수집하였다. 377건의 기사 중에서 중복되거나 불필요한 기사 12건을 제외한 365건을 분석하였다.

3.2 데이터 처리 및 분석

2016년 1월부터 2020년 10월까지의 “정선군 재생, 마을호텔”과 관련된 기사 총 365건에서 키워드들에 대한 전처리과정을 거쳐 최종적으로 선정된 키워드는 도시, 재생, 사업, 지역, 마을, 주민, 정선군, 지원, 정선, 호텔, 추진, 센터, 강원랜드, 조성, 관광, 문화, 폐광 등의 순으로 나타났다. 정선군의 도시재생사업이 주민, 마을, 호텔 분야가 핵심을 이루고 있다는 것을 알 수 있다.

그림 1은 키워드에 대한 빈도수가 상위 200개를 워드 클라우드를 이용하여 시각화한 결과이다.



[그림 1] Word cloud by top 200 keywords

표 1은 도시재생과 관련된 토픽과 키워드를 추출하기 위한 LDA 알고리즘의 결과이다. 각 토픽의 비중 순으로 10개의 토픽들을 추출하였고, 각 토픽별 연관성을 중심으로 토픽의 이름을 정하였다. 토픽 분석 결과 마을호텔의 계획, 운영, 조성 그리고 주민들의 공동체 구성, 폐광자원을 활용한 관광 콘텐츠와 지역의 특성을 살린 생태 관광 콘텐츠, 18번가를 활성화하기 위한 강원랜드와의 연계, 청년창업 등에 대한 기사가 주를 이루고 있었음을 알 수 있었다.

[표 1] LDA 토픽 분석 결과

토픽이름	상위 15개
일자리 창출	센터, 지역, 공모, 인구, 방안, 폐광, 추진, 일자리, 자원, 청년, 거점, 참여, 추리, 국비, 상생
마을호텔 계획	지원, 산업, 선정, 관광객, 유산, 호텔, 골목길, 체험, 주택, 참여, 감소, 공동, 계획, 골목, 정비
마을호텔 운영	주민, 도시, 정선, 고한읍, 운영, 변화, 공간, 생활, 호텔, 시설, 강원랜드, 올림픽, 콘텐츠, 구축, 지자체
주민 공동체	사업, 주민, 지원, 사북, 건물, 확보, 역할, 연계, 공동체, 참여, 시행, 이용규, 기업, 정선, 구성, 전통
마을호텔 조성	마을, 호텔, 조성, 추진, 교육, 사북읍, 정부, 지역주민, 뉴딜, 골목길, 여행, 삼탄, 관심, 함백산, 숙박
강원랜드 연계	경제, 주민, 강원랜드, 미래, 버스, 복지, 탄광, 호텔, 사업, 센터장, 협력, 아리랑, 프로젝트, 군민, 18,
관광콘텐츠 (폐광자원)	정원, 도시, 개최, 시설, 공공, 정선, 문화, 아트마, 폐광, 강원랜드, 전문가, 광부, 식당, 야생화, 프로그램,
관광콘텐츠 (생태)	재생, 관광, 진행, 문화, 창출, 야생화, 도시, 대표, 기업, 시장, 서울, 행사, 커뮤니티, 강원랜드, 생태
청년창업	재생, 지역, 창업, 청년, 동네, 18, 활성, 환경, 센터, 폐광, 임대, 방문, 위원회, 선정, 리모델링

4. 결론

본 논문에서는 정선군의 마을호텔에 대한 도시재생사업에 관련된 기사를 중앙지, 경제지, 강원도 지역종합지, 주요 방송사, 전문지 등에서 기사를 수집하여 토픽 모델링 기법으로 정선군의 도시재생 산업 동향을 분석하였다.

도시재생이 이루어지는 전국의 사업지 중 ‘폐광촌의 기적’, ‘18번가의 기적’이라 불리는 정선군의 ‘마을호텔 18번가’ 사업은 관이 아닌 주민주도 자생적 도시재생의 모범모델로 알려져 이목을 끌고 있음에, 도시재생사업이 추구하는 지역의 자생력 성장과 주민참여라는 근원적 목표에 걸맞은 사례로써 긍정적 기대효과가 과급되리라 예상한다.

참고문헌

[1] 예시제, 10조 원 들어간 ‘도시재생’ 공공성 집착 버리고 기업 투자 받아들여야, <https://www.yeosijae.org/research/1017> (2020.10.07. 검색)

[2] 조경원, 우영운. (2019). 텍스트 마이닝을 이용한 4차 산업 연구 동향 토픽 모델링. 한국정보통신학회논문지, 23(7), 764-770.

[3] 김용진, 손용훈. (2018). 뉴스데이터의 LDA 토픽 분석을 통한 장수군 농촌지역 활성화 사업의 특징- 관광·생활 키워드를 중심으로 -. , 24(4), 69-80.

[4] 빅카인즈 <https://www.bigkinds.or.kr/>