

계량적 모델을 통한 지리학 연구의 최신동향 및 토픽 분석

An Analysis of the latest Trends and Topics in Geography Research Using Topic Modeling

저자 (Authors)	이유빈, 이영호, 성정창, 애나 스타네스쿠, 지상훈, 황철수 Yubin Lee, Youngho Lee, Jeongchang Seong, Ana Stanescu, Sanghoon Ji, Chul Sue Hwang
출처 (Source)	대한지리학회지 55(6) , 2020.12, 589-599 (11 pages) Journal of the Korean Geographical Society 55(6) , 2020.12, 589-599 (11 pages)
발행처 (Publisher)	대한지리학회 The Korean Geographical Society
URL	http://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE10517326
APA Style	이유빈, 이영호, 성정창, 애나 스타네스쿠, 지상훈, 황철수 (2020). 계량적 모델을 통한 지리학 연구의 최신동향 및 토픽 분석. 대한지리학회지 , 55(6), 589-599.
이용정보 (Accessed)	경희대학교 163.180.5.*** 2021/05/21 22:05 (KST)

저작권 안내

DBpia에서 제공되는 모든 저작물의 저작권은 원저작자에게 있으며, 누리미디어는 각 저작물의 내용을 보증하거나 책임을 지지 않습니다. 그리고 DBpia에서 제공되는 저작물은 DBpia와 구독계약을 체결한 기관소속 이용자 혹은 해당 저작물의 개별 구매자가 비영리적으로만 이용할 수 있습니다. 그러므로 이에 위반하여 DBpia에서 제공되는 저작물을 복제, 전송 등의 방법으로 무단 이용하는 경우 관련 법령에 따라 민, 형사상의 책임을 질 수 있습니다.

Copyright Information

Copyright of all literary works provided by DBpia belongs to the copyright holder(s) and Nurimedia does not guarantee contents of the literary work or assume responsibility for the same. In addition, the literary works provided by DBpia may only be used by the users affiliated to the institutions which executed a subscription agreement with DBpia or the individual purchasers of the literary work(s) for non-commercial purposes. Therefore, any person who illegally uses the literary works provided by DBpia by means of reproduction or transmission shall assume civil and criminal responsibility according to applicable laws and regulations.

계량적 모델을 통한 지리학 연구의 최신동향 및 토픽 분석

이유빈*·이영호**·성정창***·애나 스타네스쿠****·지상훈*****·황철수*****

An Analysis of the latest Trends and Topics in Geography Research Using Topic Modeling

Yubin Lee* · Youngho Lee** · Jeongchang Seong*** · Ana Stanesco**** ·
Sanghoon Ji***** · Chul Sue Hwang*****

요약 : 지리학의 연구 주제와 방법론이 점차 확장되고, 방대한 양의 연구 결과가 도출되면서 과거의 방식으로는 최근 지리학 관련 연구들의 연구 경향을 한 눈에 파악하기 어려워졌다. 따라서 본 연구에서는 토픽 모델링 기반의 LDA 방법론을 활용하여, 2018년부터 2020년까지 AAG 학회에 제출된 초록들을 분석해 최근 지리학의 연구 토픽을 분석하고자 하였다. 이를 통해 기존 지리학 연구동향 분석의 결과와 비교하여 현재의 지리학 연구동향을 파악하고자 하였다. 이를 위하여 초록 내의 모든 단어를 토큰화하고 표제어추출을 한 뒤, 명사와 형용사에 해당하는 단어만 추출하였다. 이후 LDA를 통하여 총 8개의 토픽(정치생태학, 지리공간 과학, 도시 개발, 지리공간 시각화, 토지 피복 변화, 보건 지리학, 도시 공동체, 정치 지리학)을 추출하였다. 본 연구는 지리학 연구자들이 지리학의 전반적인 연구동향을 이해하고 최근에 주목받고 있는 연구 주제를 탐색하는데 기여할 수 있을 것이다.

주요어 : 토픽 모델링, 토픽 분석, 잠재 디리클레 할당법, 텍스트 마이닝, 빅 데이터

Abstract : With the gradual expansion of the scope of research in geography and the vast amount of research results being produced, it has become difficult to grasp the trend of research in recent geography studies. In this study, we tried to analyze the recent research trend and topics of geography by analyzing the abstracts submitted to the AAG Conference from 2018 to 2020, using the LDA-based Topic Modeling. All words in the abstract were tokenized and lemmatized to extract only nouns and adjectives. Eight topics; 'Political Ecology,' 'Geospatial Science,' 'Urban Development,' 'Geospatial Visualization,' 'Land Cover Change,' 'Health Geography,' 'Urban Community' and 'Political Geography' has been generated through the LDA. This study could contribute to geography researchers understanding the overall research trends in geography and exploring research topics that have recently drawn attention.

Key Words : Topic Modeling, Topic Analysis, LDA, Text Mining, Big Data

1. 서론

지리학은 지표면에 나타나는 모든 현상에 관심을 두기 때문에(Schaefer, 1953), 지리학이 다루는 연구 범위는 도

시, 기후, 공간구조, 지형, 경관 등으로 매우 다양하다(Murphey, 1982). 최근에는 GIS(Geographical Information System) 뿐만 아니라 머신러닝(원석환·황철수, 2018; 이영호·홍성연, 2019), 딥러닝(조원호 등, 2019)과 같은 빅데이터 분석 기법도 지리학 분야에서 활발히 활용되기 시작했

* 경희대학교 지리학과 대학원(Department of Geography, Kyung Hee University), leeyubin@khu.ac.kr

** 경희대학교 지리학과 대학원(Department of Geography, Kyung Hee University), emfo0124@khu.ac.kr

*** 웨스트 조지아 대학교 지리학과(University of West Georgia), jseong@westga.edu

**** 웨스트 조지아 대학교 컴퓨터학과(University of West Georgia), astanesco@westga.edu

***** 경희대학교 지리학과 대학원(Department of Geography, Kyung Hee University), jish@khu.ac.kr

***** 경희대학교 지리학과 (Department of Geography, Kyung Hee University, hcs@khu.ac.kr

다(Tian *et al.*, 2008). 이와 같이 지리학의 연구 주제와 방법론이 계속 확장되면서 지리학이 다루는 분야는 더욱 다양해졌지만, 그만큼 연구자들이 지리학의 전반적인 연구 경향을 파악하는 것은 더 어려워지게 되었다(Gorraiz *et al.*, 2016). 이에 따라 지리학 관련 연구들의 최신 연구 흐름을 요약하고 종합하는 연구가 필요하다.

연구동향 파악을 위한 기존 연구들은 전문가 평가, 델파이(Delphi) 기법 등의 정성적 연구 방법을 활용하였다(노승용, 2006; 최용석 등, 2008). 하지만 정성적 연구방법론들은 시간 및 경제적 비용 증가를 비롯하여 전문가의 주관적 견해가 반영될 수 있는 등 많은 단점이 존재한다(서창교 등, 2001; 나상태 등, 2016). 이러한 단점을 보완하기 위하여 정량적 연구방법인 계량지리학 방법론이 등장했다(Nederhof, 2006). 최근 연구동향 분석을 위한 계량지리학 기법 중에서는 토픽 모델링 방법론이 주로 사용되고 있다(김갑철·노현중, 2019; 박종도, 2019; 이찬호 등 2019).

토픽 모델링은 여러 문서의 텍스트를 분석하여 주제를 파악하고 유사한 단어들을 군집시키는 방법으로(이원상·손소영, 2015; 박주섭 등, 2017), 토픽 모델링 기법 중 최근 활발하게 활용되고 있는 기법은 LDA(Latent Dirichlet Allocation, 잠재 디리클레 할당법)이다(Griffiths and Steyvers, 2004). LDA는 텍스트에 내포된 의미를 정교하게 도출할 수 있으며(남승주·이현철, 2019), 이를 통해 질적 분석에 준하는 의미와 해석을 수행할 수 있다는 장점을 가지고 있다(Jacobs and Tschötschel, 2019). 이러한 장점을 바탕으로 LDA는 문헌정보학, 인공지능 등 다양한 분야에서 기술 및 연구동향 분석을 위해 활용되고 있다(박자현·송민, 2013; 정명석·이주연, 2018; 황서이·김문기, 2019). 지리학에서는 최근에 계통지리적 특정 분야로 한정하여 LDA 활용이 시도된 바 있지만 지리학의 전반적 연구동향의 분석에 활용된 사례는 미비하다. 본 연구에서는 2018~2020년도 AAG(American Association of Geographers) 학회에 제출된 초록들을 바탕으로, 토픽 모델링 기반의 LDA를 활용하여 지리학의 최근 연구동향을 분석하고자 하였다.

2. 선행연구

지리학 연구동향 분석과 관련된 기존 연구들은 지리학 전반적인 연구동향을 파악하기보다 세부적인 분야들을

중심으로 진행되어왔다. 국내 연구의 경우, 도시지리학(남영우, 1996; 남영우·이인용, 2002; 최재현·남영우, 2012; 최재현·김숙진, 2017), 관광지리학(문영철, 1997; 노시학·정은혜, 2004; 윤병국, 2012), 경제지리학(손용택, 2007), 지역지리학(최병두, 2016), 통일지리학(이민부·김걸, 2016), 환경지리학(유영준·홍충렬, 1997), 지형학 및 수문지리학(장희준·김창환, 2004) 등의 분야에서 정성적으로 연구동향 분석이 수행되었다. 반면, GIS(임시영 등, 2014; 김병선 등, 2015; 이유빈 등, 2019)와 지리교육(이동민, 2017; 김갑철·노현중, 2019; 이의한·배선헌, 2019) 분야에서는 정량적인 방법론을 활용하여 연구동향 분석이 수행되었다. 지리학의 전반적인 연구동향을 분석하고자 하는 연구는 박삼옥(2005)에 의해 정성적으로 수행되었으며, ‘지형’, ‘기후’, ‘환경’, ‘역사/문화’, ‘경제’, ‘도시’, ‘인구/사회’, ‘정치’, ‘지역’, ‘지도/GIS’, ‘지리교육’, ‘기타’로 분류하여 시기별 논문수를 비교하였고, 1970년대 이전부터 1990년대까지의 분야별 연구동향과 21세기 지리학 연구의 전망을 분석했다. 그러나 지리학 분야에서 전반적인 연구동향을 분석하는 연구 자체가 미흡한 실정이며, 기존 연구들은 정성적인 방법론에 의존하여 대량의 연구 문서들을 고려하기 어렵다는 한계를 가지고 있다.

해외 연구의 경우에도 도시지리학(Abe, 1996; Xue and Wang, 2014), 관광지리학(Nepal, 2009; Evren and Kozak, 2014; Merigó *et al.*, 2019), GIS(Knowles, 2005; Liu *et al.*, 2016), 경제지리학(Dadao, 2017), 교통지리학(Wang and Jin, 2005; Derudder *et al.*, 2019) 등의 분야에서 세부적인 지리학 연구동향을 분석해왔다. 가장 최근에 지리학 전반의 동향을 분석한 연구는 Zhuang *et al.* (2020)에 의해 수행되었다. 이 연구에서는 미국 AAG 학회와 중국 CCG(China Conference on Geography) 학회의 데이터를 대상으로 동시출현기법을 적용한 키워드 및 저자 네트워크 분석을 수행하여 두 학회에서 나타난 차이를 발견하였다. 해외 연구에서도 전반적인 지리학 연구동향을 분석한 연구는 적었으며, 비교적 다양한 세부분야에서 정량적인 방법론을 사용한 연구가 진행되었다(Evren and Kozak, 2014; Liu *et al.*, 2016; Derudder *et al.*, 2019; Merigó *et al.*, 2019; Zhuang *et al.*, 2020). 이런 배경에서 국내에서도 최근 활용되고 있는 정량적 방법론으로 지리학의 전반적인 연구동향이나 주요 연구 토픽을 파악하는 연구가 시도될 필요가 있다.

3. 데이터 및 연구방법

1) 데이터 획득 및 전처리

본 연구에서 활용한 데이터는 2018년부터 2020년까지 미국지리학회 연례학술대회(AAG Annual Meeting)에 제출된 초록이다. AAG 학회는 세계에서 가장 큰 지리학 학회 중 하나이며, 매년 지리학 내에서 다양한 분야의 수천 개의 초록이 발표되고 있다(이유빈 등, 2019). 예컨대, 2017년 AAG 학회를 기준으로 90개 국가 출신의 참여자들이 다양한 분야로 총 6,801개의 초록을 발표하였기 때문에, AAG 학회 초록 데이터는 국제적 수준에서 지리학 연구의 흐름을 보여줄 수 있는 자료로 판단하였다. 본 연구에서는 python을 통해 AAG 웹사이트의 Abstract Gallery를 크롤링하여 2018, 2019, 2020년 초록 각 6,197, 6,026, 4,897개를 추출하였고, 이를 병합하여 총 17,114개의 초록을 취득하였다(<http://www.aag.org/>). 수집한 초록 내에는 제목, 저자, 토픽, 키워드, 세션 타입, 발표 파일, 초록 내용 등이 포함되어 있으며, 분석을 위해 제목, 토픽, 키워드, 초록 내용만을 추출 및 병합하여 데이터를 구축하였다. 이때 제목, 키워드, 초록 내용 중에서 결측치가 존재한다면 연구 대상에서 제외하여, 총 17,098개의 초록만을 분석에 사용하였다. 참고로 토픽은 초록 제출 시 필수 기입 사항이므로, 결측치가 존재하지 않는다.

텍스트 마이닝에서 가장 중요한 단계는 데이터를 정제하는 전처리 단계이며, 형태소 분석, 불용어 처리 등이 이에 해당한다(감미아·송민, 2012; 박수경·이봉규, 2019). 먼저 데이터에 특수문자가 포함되어 있을 경우, python에서 데이터를 불러올 때 인코딩 문제가 발생할 수 있으므로, 초록 내에 포함된 모든 특수문자를 삭제하였다. 이후, 문장 내의 각 단어를 추출하기 위해 모든 단어를 토큰화하였고, 단어의 통일성을 위하여 모두 소문자로 변환하였다. LDA 모델에서는 불용어를 수동으로 지정해야하므로(Alghamdi and Alfalqi, 2015), 본 연구에서는 빈도는 높지만 큰 의미가 없는 geography, geographically 등의 단어를 불용어로 지정하여 삭제하였다. 다음으로 표제어추출(lemmatization) 과정을 통하여 모든 단어를 기본 사전형으로 바꾸었고, 단어 토큰에 품사를 부착하여 명사와 형용사에 해당하는 단어만 추출하였다. 이 과정은 python의 nltk 패키지를 기반으로 진행되었으며, 최종적으로 얻은

고유한 단어는 총 58,870개이다.

2) 연구 방법

토픽 모델링은 많은 양의 텍스트 데이터를 분석하는 데에 널리 사용되며, 토픽모델링 내에는 LSI(Latent Semantic Indexing), pLSI(probabilistic Latent Semantic Indexing), LDA(Latent Dirichlet Allocation) 등 다양한 기법들이 존재한다(이원상·손소영, 2015). 이 중 LDA는 pLSA의 불완전성을 베이지안 접근법으로 보완 및 확장시킨 방법이며, 문서 말뭉치(corpus)에서 잠재적인 토픽을 형성시키는 데에 가장 많이 사용되는 방법 중 하나이다(Sukhija *et al.*, 2016; 김수현 등, 2018; 음수민 등, 2019; Benzel and Stanescu, 2020). LDA 방법론으로 텍스트에 내포되어 있는 의미를 정교하게 도출할 수 있으며(남승주·이현철, 2019), 이를 통해 질적 분석에 준하는 의미와 해석을 수행할 수 있다는 장점을 가지고 있다(Jacobs and Tschötschel, 2019).

토픽 모델링에서 문서 안의 단어들은 벡터로 표현되고, 이 벡터들을 조합하여 문서의 잠재적인 토픽을 형성한다(최재웅 등, 2019). LDA의 확률적 그래프 모델은 그림 1과 같으며(Blei, 2012; Lee *et al.*, 2018), 이는 단어가 어떤 토픽에 포함될 확률을 계산하는 과정을 나타낸다(은병원, 2015). 그림에서 D는 전체 문서 말뭉치 개수, N은 d번째 문서의 단어 개수, W는 단어, Z는 토픽을 의미한다. β_k 는 토픽별 단어의 분포를 나타내며 η 값에 따라 β 의 디리클레 분포(Dirichlet distribution)가 결정된다. α 는 토픽에 대한 디리클레 분포로, 토픽들의 패턴을 나타내며 θ 를 결정하는 매개변수이다. θ_d 는 d번째 문서에서 토픽의 비율, $Z_{d,n}$ 은 d번째 문서에서 n번째 단어에 할당된 토픽을 의미한다. 마지막으로 $W_{d,n}$ 은 d 문서에서 n번째 단어이다. 즉, LDA 모델은 θ 값에 따라서 토픽인 Z가 결정되며, 각 단어의 토픽인 Z와 토픽별 단어 비율인 β 에 따라 단어 W가 결정되므로, 매개변수에 따라 결과값이 다르게 나타나는 프로세스임을 알 수 있다(Blei, 2012; 박자현·송민, 2013; 박준형·오효정, 2017; 심준식·김형중, 2017, 박정열·안희자, 2019).

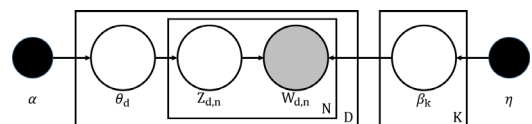


그림 1. LDA의 문서 생성 프로세스(Blei *et al.*, 2003)

토픽 모델링에서는 문서 말뭉치가 잠재적으로 갖는 토픽의 개수를 설정하는 것이 중요하다(전형진 등, 2018; 박수경·이봉규, 2019). 보통 토픽 개수를 선택할 때 혼잡도(Perplexity)가 낮은 모델을 선택하지만, 이는 학습 성능 정도를 의미할 뿐 결과를 해석하기 난해하다는 단점이 있다. 최근에는 혼잡도가 모델 평가를 위한 최고의 방법이 아니라는 연구 결과가 발표되었으며, 이러한 문제를 해결하기 위한 방법 중 하나로 일관성(Coherence) 값을 이용할 수 있다(최선영·고은지, 2019). 일관성을 적용하는 방법은 Newman *et al.*(2010)에 의해 제안되었으며, 토픽 모델링으로 만든 토픽들의 상위 단어들이 높은 유사도를 가질수록 높은 일관성 값이 계산된다. 토픽 모델링은 확률적 생성 알고리즘을 기반으로 하므로 모델을 만들 때마다 그 결과 값이 조금씩 다르게 나타나는데, 이러한 확률적 변동성 문제 해결을 위해 일관성을 최대화하는 모델을 선택하는 것이 바람직하다(이태준 등, 2017; 윤효준 등, 2019).

이런 배경에서 본 연구에서도 토픽 개수 선정에 일관성 지수를 활용하였다. 토픽의 개수가 2일 때부터 10일 때까지의 경우를 계산한 결과, 그림 2와 같이 토픽 개수가 8일 때 가장 높은 일관성 값인 0.528이 나왔다. 토픽 수를 설정하여 LDA 모델을 만든 후에는 IDM(Intertopic Distance Map)을 작성하여, 추출된 토픽들의 비중과 유사도를 파악할 수 있다. 추출이 잘 된 토픽의 경우 원의 크기가 크고, 토픽들이 사분면 전체에 고르게 배치된다(전형진 등, 2018; 최선영·고은지, 2019). 본 연구에서는 토픽 개수를 8개로 지정한 뒤 총 20개의 무작위 LDA 모델을 만들었고, 그 중 IDM에서 최대한 토픽들의 중복 영역이 없는 경우의 모델을 선택하였다.

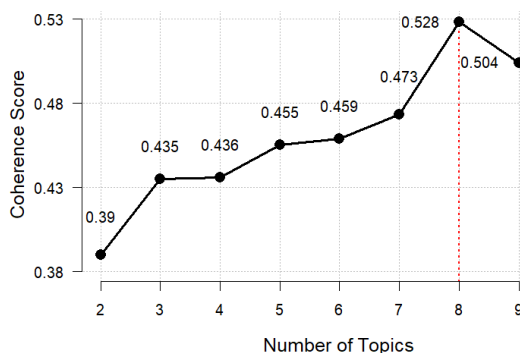


그림 2. 일관성 값 선정

4. 연구결과 및 토의

2018년부터 2020년까지 AAG 학회에 제출된 논문 초록에는 58,870개의 고유 단어가 존재하며, 최빈 상위 300개 단어로 작성한 워드 클라우드는 그림 3과 같다. 워드 클라우드는 대표적인 텍스트 시각화 기법 중 하나로, 단어의 출현 빈도가 높을수록 워드 클라우드 내에서 단어가 크게 표현된다(이은영 등, 2019). 데이터 전체에서 최빈 단어는 표 1과 같으며, “도시”, “토지”, “공간”, “사회”, “이용” 등의 순서로 나타났다. GIS를 활용한 연구가 증가함에 따라 “GIS”의 빈도도 높을 것으로 기대하였으나, 86번째 최빈 단어로 나타났다. 이는 “GIS”가 제목이나 키워드에서만 강조되는 경향이 있기 때문에 제목, 토픽, 키워드, 초록 내용 내의 모든 단어로 빈도 분석을 하면 상대적으로 빈도가 낮게 나타나기 때문이다. 그림 3과 표 1에 나타난 단어에서 합성어는 단일 단어로 분리되어 나타났다 때문에, 해당 키워드가 실제 초록에서는 어떤 의미로 사용되었는지 확인할 수 없다는 한계가 있다. 그러나 최빈 단어를 조합하여 살펴보면, 도시에서 발생하는 다양한 현상 및 담론, 토지이용 패턴과 토지의 변화 양상 파악과 관련된 연구, GIS 기반의 공간분석 연구 등이 활발히 이루어졌음을 알 수 있다.

토픽 모델링 분석 결과는 표 2와 같다. 토픽별 관련 단어와 해당 토픽을 대표하는 논문들의 초록 내용을 토대로 8개의 토픽 상세 이름을 지정하였다. 토픽 1은 8개의 토픽 중 6.8%를 차지하며, “토지”, “에너지”, “토착”, “농업”, “생산” 등의 단어를 높은 확률로 가진다. 실제로 해당 토픽을 갖는 대표적인 논문은 환경 변화에 따른 토착주민의 생활



그림 3. 워드 클라우드

표 1. 최빈 상위 20개 키워드

순위	키워드	빈도	순위	키워드	빈도
1	urban (도시)	17683	11	based (기반)	7771
2	land (토지)	10693	12	water (물)	7567
3	spatial (공간)	10480	13	environmental (환경)	6941
4	social (사회)	9878	14	political (정치)	6898
5	research (연구)	9859	15	using (활용)	6818
6	analysis (분석)	8516	16	climate (기후)	6765
7	use (사용)	8165	17	city (도시)	6765
8	change (변화)	8016	18	economic (경제)	6555
9	new (새)	7904	19	local (지역)	6313
10	development (개발)	7877	20	human (인간)	5510

표 2. 8개의 토픽 이름, 비율, 포함된 키워드

Topic	Theme	Proportion	Keywords
1	Political Ecology (정치 생태학)	6.8%	land, energy, indigenous, agricultural, production, farmers, agriculture, rural, political, environmental, region, mining, development, ecology, oil, communities, peoples, new, use, gas
2	Geospatial Science (지리공간 과학)	12.3%	information, spatial, based, research, analysis, science, gis, education, using, learning, students, social, model, systems, time, methods, network, disaster, use, used
3	Urban Development (도시 개발)	16.4%	urban, city, economic, cities, housing, spatial, development, china, food, regional, rural, land, new, population, areas, social, analysis, use, income, market
4	Geospatial Visualization (지리공간 시각화)	6.4%	digital, maps, historical, new, map, place, landscape, cultural, public, media, technology, space, mapping, technologies, art, history, world, use, visual, design
5	Land Cover Change (토지 피복 변화)	17.1%	land, change, water, climate, cover, use, forest, using, sensing, river, remote, changes, area, vegetation, tree, fire, species, analysis, results, soil
6	Health Geography (보건 지리학)	7.2%	health, spatial, risk, air, analysis, level, heat, exposure, using, factors, results, environmental, pollution, disease, vulnerability, weather, sea, population, high, medical
7	Urban Community (도시 공동체)	14.7%	environmental, development, climate, local, urban, water, change, planning, community, policy, social, research, management, governance, tourism, communities, conservation, public, infrastructure, human
8	Political Geography (정치 지리학)	19.1%	political, social, urban, work, space, food, spaces, cultural, new, politics, practices, women, people, research, place, within, ways, city, black, power

변화, 에너지에 대한 정치적 논의 등의 내용을 담고 있으므로, 토픽명을 ‘정치 생태학’으로 지정하였다.

토픽 2는 12.3%의 비율을 차지하며, “정보”, “공간”, “기반”, “연구”, “분석” 등의 단어 순으로 높은 가중치를 가졌다. 이 토픽을 갖는 대표적 논문에서는 시공간 분석, 트윗을 활용한 재난 공간 분석 등 다양한 주제의 공간 분석이 이루어졌다. “학생”, “교육”, “학습”과 같이 공간분석 교육과 관련된 논문도 높은 기여도를 갖고 있어, 공간분석 교육에 많은 관심이 있음을 알 수 있다. 이 토픽의 대부분의 단어는 공간 분석과 관련되어 있으므로 토픽명을 ‘지리공간 과학’으로 지정하였다.

학’으로 지정하였다.

토픽 3은 16.4%를 차지하며 “도시”, “경제”, “주택”, “공간”, “개발” 등의 단어를 포함하고, 대부분 도시 개발과 관련된 단어이므로 ‘도시 개발’로 지정하였다. 실제로 주택 단지와 인종 그룹의 관계, 인구 증가와 도시 확장, 도시 내 교통 선 호도, 주택 경제, 주택시장규제 등 다양한 도시 개발 관련 논문이 포함되어 있다. 이 토픽에는 “도시”와 반대되는 “시골” 단어도 포함되어 있으며, “시골”을 포함한 대표적인 논문에서는 교외 지역의 빈곤, 도시화 및 농촌 접근성 등 농촌 지역의 경제 및 도시화와 관련된 연구와 관련 있다.

토픽 4는 6.4%를 차지하며 “디지털”, “지도”, “역사”, “장소”, “경관” 등의 단어를 포함하고 있다. 이 토픽을 갖는 대표적 논문은 웹 GIS, 미국 의회 지구 재구획 등의 게리맨더링, 지도 가독성 신뢰도 등의 내용을 담고 있다. 또, “역사”와 관련된 단어들이 포함되어 있는데, 이러한 논문에는 역사와 정치에 따른 영토화 문제 등을 다룬다. 이 토픽의 대부분 단어는 시각화와 관련되어 있으므로 토픽명을 ‘지리 공간 시각화’로 지정하였다.

토픽 5는 전체 토픽 중에서 17.1%를 차지하며 “토지”, “변화”, “물”, “기후”, “피복” 등 토지 피복 변화와 관련된 단어로 이루어져 있어, 토픽명을 ‘토지 피복 변화’로 지정하였다. 대표적인 논문은 딱정벌레로 인한 나무 사망, 위성영상을 활용한 토지 피복 변화 감지, NDVI 복원, 토지피복변화 시계열 분석 등의 연구 내용을 담고 있다.

토픽 6은 7.2%를 차지하며 “건강”, “공간”, “위험”, “공기”, “분석” 등의 환경 및 건강과 관련된 단어로 이루어져 있으므로 토픽명을 ‘보건 지리학’으로 지정하였다. 이 토픽의 대표적인 논문에서는 임산부의 오염 물질 접촉과 아기 저체중의 연관성, 해안 굴곡 지역에서 심혈관 및 호흡기 질환 발병 영향, 만성 질병의 공간적 분포, 특정 질병의 사회경제적 및 환경적 요인 등 환경에 따른 건강 및 질병에 대한 연구를 진행했다.

토픽 7은 14.7%를 차지하며 “환경”, “개발”, “기후”, “지역”, “도시” 등의 단어로 이루어져 있다. 해당 토픽의 대표적 논문은 지속 가능한 개발, 관광 개발과 자연 보호의 시너지 효과를 위한 정책 제안, 국가 해양 관할권의 거버넌스, 해양생물 보호를 위한 NGO와 정부의 협력, 환경 관련 거버넌스 등이며, 도시 환경문제 해결 및 도시 공동 사회와 관련된 연구 내용이 담겨 있어 ‘도시 공동체’로 지정하였다.

마지막으로 토픽 8은 8개의 토픽 중에서 가장 많은 비율인 19.1%를 차지하며, “정치”, “사회”, “도시”, “노동”, “공간” 등의 정치·문화·사회 관련 단어로 이루어져 있어 ‘정치 지리학’으로 지정하였다. 대표적 논문에서는 위안부 문제와 페미니즘, 국가 개입 권력에 대한 논의, 유흥 산업 규제에 대한 담론, 북미 경찰 노조, 시간과 실향민의 이주 및 시민권에 대한 논의 등의 다양한 주제의 정치 지리학 연구 내용을 확인할 수 있다.

이 중 토픽 1, 7, 8은 ‘정치’라는 공통된 주제가 있으며, 정치 활동을 통해 환경 또는 주민의 삶 보전을 연구한다는 점이 비슷하다. 또, 토픽 5, 6, 7은 환경 자체를 연구하거나 환경과 인간의 관계를 연구하는 등 환경을 주제로 연구한다는 점에서 유사점이 있다. 토픽 3, 7, 8은 도시를 주로 대상으로 연구한다는 공통점이 있었으며, 마지막으로 토픽 2, 4는 지리공간을 분석하거나 시각화하는데에 초점을 두었

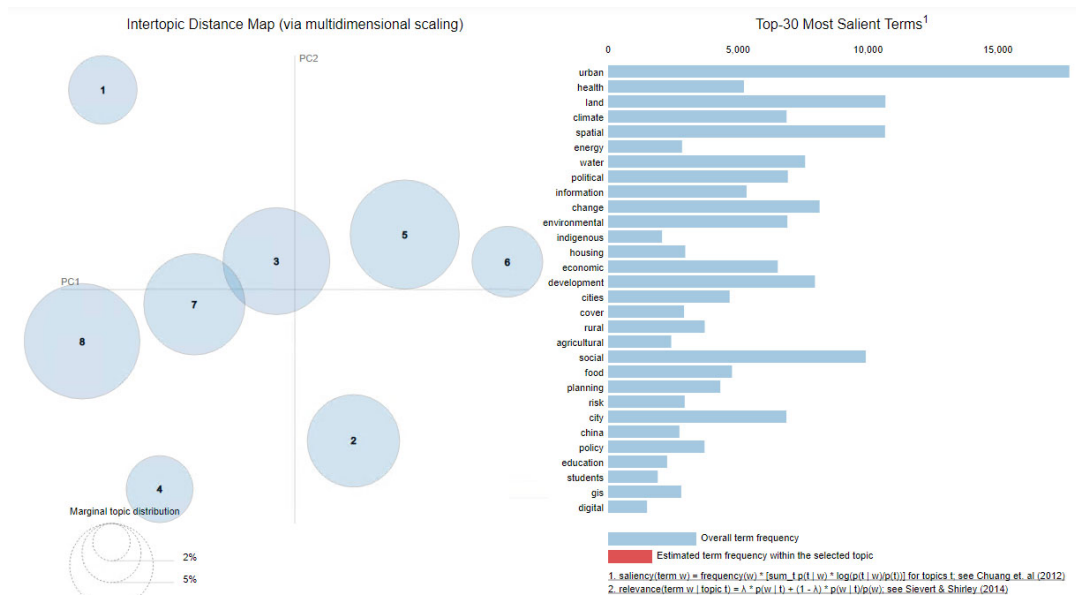


그림 4. 토픽 간 거리 지도 (Intertopic Distance Map, IDM)

다는 점에 유사성이 있다.

그림 4는 토픽 모델링의 결과인 토픽들을 주성분 분석을 통해 2차원 척도로 표시한 IDM(Intertopic Distance Map)이다. 그림의 오른쪽 부분에서는 전체 토픽에서 핵심적인 역할을 하는 단어 30개를 확인할 수 있다. 왼쪽 부분에서는 토픽들의 구조적 관계 및 연관성을 파악할 수 있으며, 같은 사분면에 속하거나 토픽의 거리가 가까울수록 유사한 의미를 가진다(홍영희, 2019). 본 연구결과로 도출된 IDM은 여덟 개의 토픽이 사분면에 골고루 분포되었으므로 비교적 적절하게 추출되었다고 볼 수 있다. 사분면 상에서 토픽 3과 토픽 7은 일부 겹쳐 보이며, “개발”, “사회”, “도시” 등의 단어가 함께 나타났고, 모두 사회 및 개발과 관련된 주제이다. 그러나 토픽 3은 도시 개발에 초점을 맞추고 있고, 토픽 7은 기후 및 환경에 초점을 두고 있기 때문에 각각 다른 토픽으로 분류할 수 있다.

지리학의 전반적인 연구 동향은 그림 5와 같이 가로 선인 환경 축과, 세로 선인 기술 축으로 설명될 수 있다. 먼저, 가로 선은 인문환경이나 자연환경의 스펙트럼을 나타내는 환경 축으로 이해할 수 있다. 가로 선을 기준으로 살펴보면 중간 지점에 ‘도시 개발’ 토픽이 위치하고 있으며 이를 중심으로 왼쪽으로는 인문 환경, 오른쪽으로는 자연 환경과 관련된 토픽이 위치하고 있다. ‘정치 지리학’과 ‘보건 지리학’은 사분면 상에서 가장 먼 거리로 표현되었는데, 이는 가장 왼쪽에 있는 ‘정치 지리학’에는 “정치”, “사회”, “도시” 등 인문지리와 가까운 단어들이 포함되어 있고, 반대로 가장 오른쪽에 있는 ‘보건 지리학’에는 “열”, “환경”, “기상” 등 자연지리와 가까운 단어들이 포함되어 영향을 미쳤다고 볼 수 있다. ‘정치 생태학’이 왼쪽에 위치하고 있어 지리학 내에

서는 생태학이 의사결정 및 인문환경의 관점에서 정치학으로 다루어지고 있음을 알 수 있었고, ‘지리공간 과학’ 토픽이 척도의 오른쪽에 위치하고 있어, GIS 기술이 자연지리 분야에서 더 많이 활용되고 있음을 알 수 있었다.

세로 선은 다양한 기술의 수용 정도를 나타내는 기술 척도로 이해할 수 있다. 세로 선을 기준으로 살펴보면 중간 지점에 ‘도시 개발’이 위치하고 있으며 이를 중심으로 위쪽으로는 첨단 기술 혹은 방법론, 아래쪽으로는 저수준의 기술 혹은 방법론과 관련된 토픽이 위치하고 있다. ‘정치 생태학’과 ‘지리공간 시각화’는 가장 먼 거리로 표현되었다. 가장 위에 위치한 지리공간 시각화에는 “디지털”, “미디어”, “기술” 등과 같은 컴퓨터 기술과 관련 단어가 포함되어 있는 반면, 가장 아래에 있는 ‘정치 생태학’에는 “토지”, “에너지”, “농업” 등 낮은 기술과 관련된 단어가 포함되어 영향을 미친 것으로 보인다. 그래프에서 세로 선 위쪽은 방법론과 시각화와 관련된 토픽들이 분포되어 있고, 가운데에는 ‘보건 지리학’이나 ‘정치 생태학’과 같이 환경과 인간을 함께 다루는 토픽들이 분포되어 있으며, 아래쪽에는 토지와 관련된 단어를 다루는 토픽들이 분포하고 있다. 자연 환경 관련 토픽에 비해 인문환경 관련 토픽들이 비교적 척도 위쪽으로 올라가 있어, 인문환경 분야에서 컴퓨터 기술의 활용 및 시각화가 비교적 더 활발하다는 것을 알 수 있었다.

5. 결론

지리학의 연구 분야가 다양한 주제와 방법론으로 확장되면서, 연구자들은 지리학 분야에서 어떤 연구들이 이루어지는지 한눈에 파악하기 어려워졌다. 이로 인해 지리학의 전반적인 연구동향을 파악하는 연구가 필요한 상황이지만, 지리학에서는 전반적인 연구동향을 파악하기보다는 도시지리, 환경지리, GIS 등 세부적인 분야들을 중심으로 연구동향 및 토픽 파악 연구가 진행되어왔다. 지리학의 전반적인 연구동향을 분석한 기존 연구가 존재하지만, 정성적인 방법론을 활용하여 연구자의 주관적 가치가 개입되거나 대량의 문서를 활용하기 어렵다는 한계가 존재한다. 따라서 정량적인 방법론으로 대량의 문서를 분석하여, 최근 지리학 관련 연구들은 어떻게 이루어지는지 요약하고 종합하는 연구가 빈번하게 이뤄질 필요가 있다.

본 연구에서는 토픽 모델링의 LDA 방법론으로 AAG 학

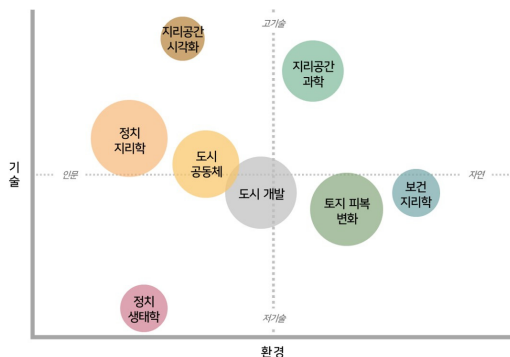


그림 5. 지리학 최신 연구 토픽

회에 제출된 초록을 분석하여 지리학의 전반적인 연구 토픽들을 알아보고자 하였다. 분석을 위하여 2018년부터 2020년까지 AAG 학회에 제출된 초록 중에서 분석 가능한 초록 17,098개를 추출하였으며, 전처리 과정을 통하여 모든 단어를 토큰화하고 기본 사전형으로 정제하였다. 이후 빈도는 높지만 큰 의미가 없는 단어는 불용어로 처리하고, 모든 단어 토큰에 품사를 부착하여 명사와 형용사만 추출하였다. 토픽 모델링 방법론은 연구자가 토픽 개수를 직접 지정해야 하며, 본 연구에서는 일관성 값을 기준으로 토픽 개수를 8개로 선정하였다.

2018년부터 2020년까지 AAG 학회에 제출된 초록에는 총 58,870개의 고유 단어가 있으며, 최빈 단어는 “도시”, “토지”, “공간”, “사회”, “이용” 등의 순서로 나타났다. 기여도가 높은 단어와 초록의 내용을 바탕으로 토픽명을 지정하였으며, 이를 토대로 토픽 1은 ‘정치 생태학’, 토픽 2는 ‘지리공간 과학’, 토픽 3은 ‘도시 개발’로 지정하였다. 토픽 4는 ‘지리공간 시각화’, 토픽 5는 ‘토지 피복 변화’, 토픽 6은 ‘보건 지리학’으로 지정하였고, 토픽 7은 ‘정치 생태학’, 토픽 8은 ‘정치 지리학’으로 지정하였다. 토픽 1, 7, 8은 정치라는 공통주제가 있었고, 토픽 5, 6, 7은 환경을 주제로 연구한다는 유사점이 있었다. 토픽 3, 7, 8은 도시를 주로 대상으로 하며, 토픽 2, 4는 지리공간 분석 및 시각화를 다룬다는 점이 유사하였다.

IDM에서 가로 선은 인문 및 자연 환경 축, 세로 선은 기술의 수용정도를 나타내는 축으로 이해할 수 있으며 이를 바탕으로 2018~2020년의 지리학 전반적 연구 동향을 설명할 수 있다. 가로 선을 기준으로 보면 ‘정치 지리학’ 토픽과 ‘보건 지리학’ 토픽은 가장 먼 거리로 표현되었으며, 이는 ‘정치 지리학’ 토픽은 인문지리 관련 단어들, 반대로 ‘보건 지리학’ 토픽은 자연지리 관련 단어들의 영향을 받은 것으로 보인다. 세로 선을 기준으로 보면 ‘정치 생태학’ 토픽과 ‘지리공간 시각화’ 토픽이 가장 먼 거리로 표현되었으며, 이는 ‘정치 생태학’ 토픽에는 낮은 기술 관련 단어들이, ‘지리공간 시각화’ 토픽에는 첨단 기술 관련 단어들이 포함된 것이 영향을 미친 것으로 파악된다.

본 연구는 국내외로 활발히 진행되지 않고 있던 지리학의 연구동향 및 토픽 분석 연구를 진행하여 전반적인 연구 동향을 요약하고, 기존 연구들과의 유사점과 차이점을 제시하는 것을 통해 최근 지리학 연구 토픽의 차별성을 확인했다는 점에서 의의가 있다. 지표 공간에서 발생하는 거의 모든 현상들을 다루는 지리학의 특성상 지리학에는 계속

해서 새로운 방법론들이 적용되고 연구 범위가 확장될 것으로 예상되므로, 이러한 연구는 지리학 연구자들이 지리학의 전반적인 연구동향을 이해하고 최근에 주목받고 있는 연구 주제를 탐색하는데 기여할 수 있을 것이다. 본 연구의 분석에서는 2018년부터 2020년까지 AAG 학회에 제출된 초록만을 데이터로 사용하였기 때문에, 추후에는 시간적 범위를 확장하고 다양한 학회 및 학회지의 연구들을 고려하는 것을 통해 지리학 연구동향의 역사와 보다 포괄적인 연구동향을 파악할 수 있을 것으로 기대한다.

사사

이 논문은 정보통신기획평가원 글로벌핵심인재양성(IITP-2020-0-01593) 사업의 지원을 받아 연구되었습니다. This research was supported by the MSIT (Ministry of Science, ICT), Republic of Korea, under the High-Potential Individuals Global Training Program (IITP-2020-0-01593) supervised by the IITP (Institute for Information & Communications Technology Planning & Evaluation). We also thank AAG Headquarter and Board Members for supplying abstract data for this research.

참고문헌

- 감미아·송민, 2012, “텍스트 마이닝을 활용한 신문에 따른 내용 및 논조 차이점 분석,” *지능정보연구*, 18(3), 53-77.
- 김갑철·노현중, 2019, “토픽모델링을 활용한 국내 지역지리교육 연구 동향 분석,” *사회과학교육*, 58(4), 49-67.
- 김병선·정민우·전상은·신동빈, 2015, “키워드 네트워크 분석을 이용한 지리공간정보의 글로벌 연구 동향 분석,” *한국공간정보학회지*, 23(1), 69-77.
- 김수현·박해정·이정혜, 2018, “Word2vec based Latent Semantic Analysis (W2V-LSA): 새로운 토픽 모델링을 통한 블록체인 기술 연구 트렌드 분석,” *대한산업공학회 추계학술대회 논문집*, 2296-2304.
- 나상태·김자희·정민호·안주언, 2016, “토픽 모델링을 이용한 시뮬레이션 연구 동향 분석,” *한국시뮬레이션학회논문지*, 25(3), 107-116.
- 남승주·이현철, 2019, “LDA 토픽 모델링을 활용한 항공승객 유형 별 특성 분석,” *경영과학*, 36(3), 67-85.

- 남영우, 1996, “한국 도시지리학 35년사,” *대한지리학회지*, 31(2), 198-212.
- 남영우·이인용, 2002, “한국 도시지리학 반세기의 회고와 전망,” *한국도시지리학회지*, 5(1), 1-12.
- 노승용, 2006, “델파이 기법 (Delphi Technique): 전문적 통찰로 미래예측하기,” *국토*, 53-62.
- 노시학·정은혜, 2004, “국내 관광지리학의 연구동향 분석,” *관광연구저널*, 18(2), 323-336.
- 문영철, 1997, “한국 관광지리학의 연구 동향과 과제,” *한국지역지리학회지*, 3(2), 89-103.
- 박경열·안희자, 2019, “텍스트 마이닝을 활용한 DMZ 관광 이슈의 토픽 모델링 분석,” *관광레저연구*, 31(4), 143-159.
- 박삼욱, 2005, “한국의 지리학연구 60년 회고와 전망,” *대한지리학회지*, 40(6), 770-788.
- 박수경·이봉규, 2019, “키즈 크리에이터 (kids creator)에 대한 사회적 이슈 분석: 뉴스 데이터 토픽 모델링의 적용,” *문화와융합*, 41(4), 549-578.
- 박자현·송민, 2013, “토픽모델링을 활용한 국내 문헌정보학 연구동향 분석,” *정보관리학회지*, 30(1), 7-32.
- 박종도, 2019, “토픽 모델링을 활용한 다문화 연구의 이슈 추적 연구,” *한국문헌정보학회지*, 53(3), 273-289.
- 박주섭·홍순구·김종원, 2017, “토픽모델링을 활용한 과학기술 동향 및 예측에 관한 연구,” *한국산업정보학회논문지*, 22(4), 19-28.
- 박준형·오효정, 2017, “국내 기록관리학 연구동향 분석을 위한 토픽모델링 기법 비교: LDA와 HDP를 중심으로,” *한국도서관정보학회지*, 48(4), 235-258.
- 서창교·김은진·이영숙, 2001, “웹에 기반한 델파이 의사결정지원 시스템의 구현,” *정보시스템연구*, 10(1), 5-25.
- 손용택, 2007, “한국 경제지리학의 연구동향과 과제,” *정신문화연구*, 30(3), 5-34.
- 심준식·김형중, 2017, “LDA 토픽 모델링을 활용한 판매 검색 및 분류 방법,” *전자공학회논문지*, 54(9), 67-75.
- 온병원, 2015, “국내 휴대폰의 진화패턴 규명을 위한 텍스트 마이닝 방안 제안 및 사례 연구,” *한국컴퓨터정보학회논문지*, 20(2), 29-45.
- 원석환·황철수, 2018, “의사결정나무와 서포트 벡터 머신 모델을 활용한 토지이용 변화 시뮬레이션: 통일 후 북한 도시를 대상으로,” *국토연구*, 41-56.
- 유영준·홍충렬, 1997, “국내 환경지리학의 연구 동향,” *국토지리학회지*, 29, 123-137.
- 윤병국, 2012, “2000년 이후 국내 관광지리의 연구동향과 향후 과제,” *관광연구저널*, 26(1), 131-148.
- 윤효준·박재현·윤지운, 2019, “비정형 텍스트 자료에서 잠재정보 추출을 위한 토픽모델링 소개: 치매관련 신체활동 뉴스 기사의 이슈 분석,” *체육과학연구*, 30(3), 501-512.
- 음수민·이수길·조성원·이철용, 2019, “LDA 기반의 토픽모델링을 이용한 철도차량용 무선급전시스템 연구 동향 분석,” *대한산업공학회지*, 45(4), 284-301.
- 이동민, 2017, “세계 지리교육 연구의 최근 동향에 대한 언어네트워크 분석,” *한국지리학회지*, 6(3), 355-368.
- 이민부·김길, 2016, “통일지리학의 연구동향과 과제,” *대한지리학회지*, 51(6), 873-892.
- 이영호·홍성연, 2019, “머신러닝을 활용한 개인의 교통수단 선택 예측모형 구축,” *한국데이터정보과학회지*, 30(5), 1011-1024.
- 이원상·손소영, 2015, “공간빅데이터 연구 동향 파악을 위한 토픽모형 분석,” *대한산업공학회지*, 41(1), 64-73.
- 이유빈·이영호·성정창·황태건·황철수, 2019, “AAG 2000-2019 학회발표논문 분석을 통한 미국의 GIS 연구동향 분석,” *국토지리학회지*, 53(4), 495-508.
- 이은영·주경희·이두희, 2019, “워드 클라우드 기법을 이용한 최근 소비자학 연구 트렌드 분석,” *상품학연구*, 37, 1-7.
- 이의한·배선학, 2019, “데이터 마이닝을 활용한 지리교육의 연구 동향 분석: 텍스트 마이닝의 적용 가능성을 중심으로,” *한국지리학회지*, 8(3), 363-375.
- 이찬호·김민승·이정희·성태웅, 2019, “토픽 모델링 기반의 LDA 기법과 웹 검색 트래픽을 활용한 기술개발 트렌드의 비교 연구,” *한국정보과학회 학술발표논문집*, 1767-1769.
- 이태준·이승배·오창동, 2017, “원자력 이슈에 대한 정부와 언론의 커뮤니케이션 전략 비교연구,” *언론과학연구*, 17(3), 172-229.
- 임시영·이미숙·진기호·신동빈, 2014, “텍스트마이닝 기술을 이용한 공간정보 분야의 연구 동향에 관한 고찰 - 국가연구개발사업 보고서 및 논문을 중심으로,” *한국공간정보학회지*, 22(4), 11-20.
- 장희준·김창환, 2004, “미국에서의 지형학과 수문지리학의 최근 연구동향,” *대한지리학회지*, 39(6), 873-887.
- 전형진·김도연·한국진·김동우·손승우·이철민, 2018, “토픽모델링을 활용한 실내환경 분야 연구동향 파악: 실내환경학회지 초록 사례연구,” *실내환경 및 냄새 학회지*, 17(4), 322-329.
- 정명석·이주연, 2018, “Latent Dirichlet Allocation(LDA) 모델

- 기반의 인공지능(AI) 기술 관련 연구 활동 및 동향 분석,” 한국산업정보학회논문지, 23(3), 87-95.
- 조원호·임용호·박기호, 2019, “합성곱 신경망을 이용한 딥러닝 기반의 토지피복 분류: 한국 토지피복을 대상으로,” 대한지리학회지, 54(1), 1-16.
- 최병두, 2016, “한국의 신지역지리학:(2) 지리학분야별 지역 연구 동향과 과제,” 한국지역지리학회지, 22(1), 1-24.
- 최선영·고은지, 2019, “메타데이터를 활용한 1960~2018 〈한국 언론학보〉 논문 분석: 다이나믹 토픽 모델링(Dynamic Topic Modeling) 방법을 중심으로,” 한국언론학보, 63(4), 7-42.
- 최용석·백승철·권혁인, 2008, “텔파이기법을 이용한 U-city 사업의 핵심성공요인 도출,” 인터넷전자상거래연구, 8(3), 183-209.
- 최재웅·장재혁·김대환·윤장혁, 2019, “인문사회 과학기술 분야 연구의 학제적 동향 분석: 토픽 모델링과 네트워크 분석의 활용,” 산업경영시스템학회지, 42(1), 74-86.
- 최재현·김숙진, 2017, “한국도시지리학회지 게재 논문으로 본 도시지리 연구의 주제와 과제: 1998~ 2016년,” 한국도시지리학회지, 20(1), 1-26.
- 최재현·남영우, 2012, “한국 도시지리학 연구동향: 1960~2012년,” 대한지리학회지, 47(4), 541-553.
- 홍영희, 2019, “소프트웨어(SW)교육 정책에 대한 토픽 분석: 부산지역 중합지를 중심으로,” 통계연구, 24(2), 52-77.
- 황서이·김문기, 2019, “국내 인공지능분야 연구동향 분석-토픽 모델링과 의미연결망분석을 중심으로,” 한국디지털콘텐츠학회 논문지, 20(9), 1847-1855.
- Abe, K., 1996, Urban geography in postwar Japan, *Geographical review of Japan*, 69(1), 70-82.
- Alghamdi, R. and Alfalqi, K., 2015, A survey of topic modeling in text mining, *International Journal of Advanced Computer Science and Applications*, 6(1), 147-153.
- Benzel, S. and Stanescu, A., 2020, Histogram Methods for unsupervised clustering, *Proceedings of the 2020 ACM Southeast Conference*, 218-251.
- Blei, D. M., Ng, A. Y. and Jordan, M. I., 2003, Latent dirichlet allocation, *Journal of machine Learning research*, 3, 993-1022.
- Blei, D. M., 2012, Probabilistic topic models, *Communications of the ACM*, 55(4), 77-84.
- Dadao, L., 2017, The changing humanistic and economic geography in China, *Scientia Geographica Sinica*, 37(5), 641-650.
- Derudder, B., Liu, X., Hong, S., Ruan, S., Wang, Y. and Witlox, F., 2019, The shifting position of the journal of transport geography in ‘transport geography research’: A bibliometric analysis, *Journal of Transport Geography*, 81, 102538.
- Evren, S. and Kozak, N., 2014, Bibliometric analysis of tourism and hospitality related articles published in Turkey, *Anatolia*, 25(1), 61-80.
- Gorraiz, J., Gumpenberger, C. and Glade, T., 2016, On the bibliometric coordinates of four different research fields in Geography, *Scientometrics*, 107(2), 873-897.
- Griffiths, T. L. and Steyvers, M., 2004, Finding scientific topics, *Proceedings of the National Academy of Sciences*, 101(1), 5228-5235.
- Jacobs, T. and Tschötschel, R., 2019, Topic models meet discourse analysis: A quantitative tool for a qualitative approach, *International Journal of Social Research Methodology*, 22(5), 469-485.
- Knowles, A. K., 2005, Emerging trends in historical GIS, *Historical Geography*, 33, 7-13.
- Lee, H., Seo, H. and Geum, Y., 2018, Uncovering the topic landscape of product-service system research: From sustainability to value creation, *Sustainability*, 10(4), 911.
- Liu, F., Lin, A., Wang, H., Peng, Y. and Hong, S., 2016, Global research trends of geographical information system from 1961 to 2010: A bibliometric analysis, *Scientometrics*, 106(2), 751-768.
- Merigó, J. M., Mulet-Forteza, C., Valencia, C. and Lew, A. A., 2019, Twenty years of tourism geographies: A bibliometric overview, *Tourism Geographies*, 21(5), 881-910.
- Murphey, R., 1982, *The Scope of Geography*, Methuen & Co Ltd., London.
- Nederhof, A. J., 2006, Bibliometric monitoring of research performance in the social sciences and the humanities: A review, *Scientometrics*, 66(1), 81-100.
- Nepal, S. K., 2009, Traditions and trends: A review of geographical scholarship in tourism, *Tourism Geographies*, 11(1), 2-22.
- Newman, D., Lau, J. H., Grieser, K., and Baldwin, T., 2010,

- Automatic evaluation of topic coherence, *Human language technologies: The 2010 annual conference of the North American chapter of the Association for Computational Linguistics*, 100-108.
- Schaefer, F. K., 1953, Exceptionalism in geography: A methodological examination, *Annals of the Association of American geographers*, 43(3), 226-249.
- Sukhija, N., Tatineni, M., Brown, N., Moer, M. V., Rodriguez, P. and Callicott, S., 2016, Topic modeling and visualization for big data in social sciences, *2016 Intl IEEE Conferences*, 1198-1205.
- Tian, Y., Wen, C. and Hong, S., 2008, Global scientific production on GIS research by bibliometric analysis from 1997 to 2006, *Journal of Informetrics*, 2(1), 65-74.
- Wang, C. and Jin, F., 2005, Research history and developing trend about geography of transportation in China, *Progress in Geography*, 24(6), 66-78.
- Xue, D. and Wang, L., 2014, Progress of urban geography research in China since the 1980s, *Acta Geographica Sinica*, 69(8), 1117-1129.
- Zhuang, L., Ye, C. and Lieske, S. N., 2020, Intertwining globality and locality: bibliometric analysis based on the top geography annual conferences in America and China, *Scientometrics*, 122(2), 1075-1096.
- 교신: 황철수, 02447, 서울특별시 동대문구 경희대로 26, 경희대학교 지리학과(이메일: hcs@khu.ac.kr)
- Correspondence:: Chul Sue Hwang, Department of Geography, Kyunghee University, 26, Kyungheedaero-ro, Dongdaemun-gu, 02447, Seoul, Republic of Korea(e-mail: hcs@khu.ac.kr)
- 최초투고일 2020. 11. 29
수정일 2020. 12. 16
최종접수일 2020. 12. 17