



텍스트마이닝 기법을 활용한 한국·중국·일본의 관광연구 동향 분석

Analysis of Tourism Research Trend using Text Mining

저자 (Authors)	남한별, 전상우, 안영빈, 유다솔, 박병화 Nam, Hanbyeol, Jeon, SangWoo, An, Yeongbin, Yu, Dasol, Park, Byeonghwa
출처 (Source)	한국관광레저학회 학술발표대회 , 2017.12, 283-288 (6 pages)
발행처 (Publisher)	한국관광레저학회 The Korea Academic Society Of Tourism And Leisure
URL	http://www.dbpia.co.kr/Article/NODE07322977
APA Style	남한별, 전상우, 안영빈, 유다솔, 박병화 (2017). 텍스트마이닝 기법을 활용한 한국·중국·일본의 관광연구 동향 분석. 한국관광레저학회 학술발표대회 , 283-288.
이용정보 (Accessed)	성균관대학교 자연과학캠퍼스 115.***.238.89 2018/09/17 16:45 (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.

텍스트마이닝 기법을 활용한 한국·중국·일본의 관광연구 동향 분석

Analysis of Tourism Research Trend using Text Mining

남한별* · 전상우** · 안영빈** · 유다솔** · 박병화***

Nam, Hanbyeol · Jeon, SangWoo · An, Yeongbin · Yu, Dasol · Park, Byeonghwa

Abstract

관광산업에 대한 관심은 선진국이나 개도국, 후진국에서도 중요한 국가전략산업으로 인정되고 있다. 이에 따라 관광에 대한 학문적 연구가 괄목할만한 성장을 이루고 글로벌화에 따라 중국과 일본에 형향 또한 크다. 따라서 해외주요 등재지에 실린 한·중·일의 영문 초록 데이터를 이용하여 국내 관광연구 동향뿐만 아니라 중국과일본의 관광연구 동향을 정리해보고자 하였다. 텍스트마이닝 기법인 빈도분석, 사회연결망(SNA), 토픽모델링을 사용하였다. 분석 결과 국가별 특징적인 인기 연구주제를 파악할 수 있다. 공통적으로 선호되는 연구들도 있지만 국가별로 특징적인 연구들을 파악할 수 있었다. 본 연구는 기존의 키워드 중심 분석방법이 아닌 영문초록을 이용해 한·중·일을 비교분석 할 수 있고 연구주제들 간의 관계를 살펴볼 수 있어서 이 분야에서 새로운 주제를 선정하는데 참고할 유용한 정보로 활용될 수 있다.

주제어 : 초록(Abstracts), 텍스트 마이닝(Text Mining), 연구 동향(Research Trend), 관광(Tourism)

* 한남대학교 경제학과

** 한남대학교 비즈니스통계학과

*** 한남대학교 비즈니스통계학과, 교신저자, e-mail: bpark@hnu.kr

I. 서 론

21세기에 접어들면서 우리나라는 국민 소득향상, 주 5일제 근무, 직장 정년제도의 변화, 교통수단 발달, 여가시간의 증대와 여가활동의 변화 등으로 인해 관광에 대한 인식 변화와 관광산업의 발전이 폭발적으로 증가하고 있다(이지연, 2012). 이에 따라 국내의 관광학 영역도 점점 확대되고 성장해오고 있다.

관광연구동향에 대해서도 다양한 연구들이 꾸준히 이루어져 왔다. 하지만 한국 관광논문의 키워드를 이용한 연구 동향이거나 특정 광광분야의 연구 동향으로 학문적 의미가 특정 영역으로 한정되는 경우가 많다.

한·중·일 3국은 2010년 서울에서 한·중·일 FTA를 위한 공동연구가 진행됨에 따라 정치, 경제적으로 다양한 분야에서 협력과 경쟁을 통해 경제 교류가 활성화 되고 있다(조정윤, 2010). 따라서 3국의 관광부분을 비교분석하고자 한다.

학문은 연구자들이 생산한 것으로 그 분야의 세부 주제영역, 학파, 연구자들의 정보이용 행태 등이 반영된다(송혜지, 2013). 따라서 본 연구는 한·중·일 3국의 해외 관광 학술논문의 초록을 분석함으로써 최근 연구자들이 관심을 갖고 있는 주제의 동향을 파악할 수 있으며 앞으로 관광 연구 주제 선택에 있어서도 시사점을 나타내고자 한다. 또한 방대한 데이터를 텍스트 마이닝 기법을 통하여 객관적으로 분석한다.

II. 연구의 이론적 배경

세계 관광 기구(World Tourism Organization)는 관광이란 일반적으로 휴가활동에 국한된 것으로 여행, 여가, 사업, 혹은 다른 목적으로 1년 이하 동안 거주지가 아닌 다른 곳에 머무는 것이라고 정의했다. 관광산업에 대한 관심은 경기변동의 증가나 감소 상황뿐만 아니라, 선진국이나 개도국, 후진국에서도 중요한 국가전략산업으로 인정하고 관심을 기울이고 있는 실정이다(김창곤, 2007).

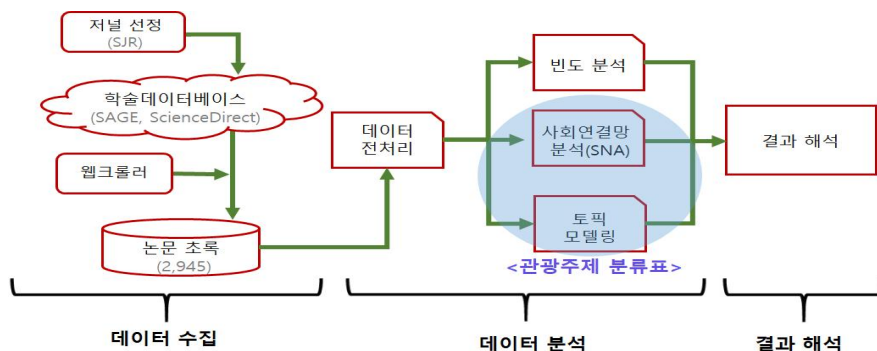
지금까지 진행된 관광에 대한 다양한 연구는 관광학 분야에서 ‘분석 지식’의 증대에 많은 공헌을 한 것이 사실이다. 그러한 의미 있는 축적된 선행연구를 기초로 관광 관련 연구들은 지식체계가 확장되어 왔다. 하지만 이 같은 학문적 기여에도 불구하고, 일부 선행연구는 추후 이루어지는 심화분석 과정에서의 유사 연구사례로만 적용되어 학문적 의미가 특정 영역으로 한정되는 경우도 많았다.(민웅기, 2016)

국내 관광연구의 키워드 데이터를 가지고 관광연구의 최근 연구동향을 분석하였다. 논

문의 키워드는 저자의 아이디어나 문제의식을 표현하기 위해 저자가 신중하게 선택하여 지식의 구조를 파악하는 좋은 방법으로 사용된다. 한 개 이상의 키워드를 통해 연구논문을 표현하여 각각의 키워드도 의미가 있지만 두 개 이상의 키워드 간 관계를 특정 텍스트 단위에서 공동으로 출현한 단어의 집합적 상호 연결을 표현하는 방식인 동시 출현 단어 분석(co-word analysis) 기법을 사용하였다. 국내 관광연구 주제들은 빠르고 다양하게 변화하는 양상을 보이고 지속적으로 선호되는 연구주제들도 있지만 연구대상의 내재적, 외재적 요인들이 점차 세분화되고 연구대상들도 종사 분야에 따라 다양하고 구체적으로 선정되어 진행되고 있는 것을 확인하였다.(김민선, 2016)

III. 연구 방법

1. 연구모형



〈그림 1〉 연구모형

본 연구의 절차는 <그림 1>의 연구모형으로 나타내었다.

2. 데이터 수집

본 연구에서는 해외 관광학 학술지 논문초록데이터를 수집하기 위해서 먼저 학술 저널의 영향력을 측정하는 웹사이트인 SJR(www.scimagojr.com)에서 관광학 관련 논문 상위 5개를 선택하였다. 한국, 중국, 일본에 관한 연구를 진행한 총 2945편의 논문이 수집되었다.

3. 데이터 분석 및 방법

1) 데이터 전처리

객관성 있는 분석을 위해 R 3.4.1을 이용하여 세 가지 데이터 전처리 과정을 거쳤다. 첫째 모든 데이터를 소문자화 하는 tm패키지를 이용하였다. 둘째 영어의 굴절어 문제를 해결하기 위하여 어근을 추출하는 snowballC패키지를 이용해 같은 의미의 단어는 동일어로 처리하였다. 셋째 기호, 숫자, 공백과 정관사, 전치사, 접속사 등은 제거하였다.

2) 빈도 분석(TF-IDF)

TF-IDF는 단어에 가중치를 부과하는 것으로 TF와 IDF를 곱한 값이다. TF(Term Frequency)값은 한 문서 내에서 특정 단어가 출현한 빈도수를 의미하고, IDF(Inverse Document Frequency)값은 특정 단어가 포함된 문서 빈도로 역수를 취한 값이다. 이는 보편적인 단어 일수록 IDF값이 작게 나타난다. 본 연구에서는 관광분야에서 한·중·일의 공통적으로 선호되는 연구 주제와 국가별 선호 연구주제를 파악하기 R 3.4.1을 이용해 TF-IDF를 적용해 실행하였다.

3) 사회연결망 분석(Social Network Analysis)

본 연구에서는 빈도분석에서 파악한 단어들을 점(노드, node)으로 만들고 관계를 선(링크, link)으로 연결한 결과를 통해 중심점을 이루고 있는 단어를 구획으로 나누었다. 빈도 분석보다 더 심도 있는 주제를 파악하기 위해 실행하였다.

IV. 분석 결과

1. 빈도 분석(TF-IDF)

수집된 논문은 한·중·일 동일하게 관광객(Tourist), 호텔(Hotel), 지역(Destination), 서비스(Service), 식당(Restaurant), 문화(Culture)의 단어가 가중치가 높은 것으로 보아 이러한 주제가 연구 분야로 선호되었다.

2. 사회연결망 분석(Social Network Analysis)

한·중·일의 사회연결망 분석을 시각화 하였을 때 호텔(Hotel), 이미지(Image), 관광

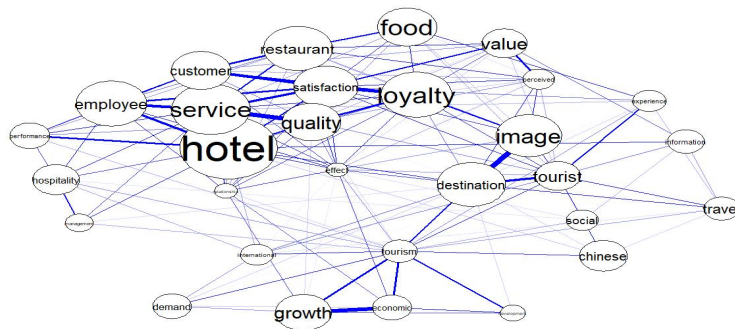
(Tourism)의 단어를 구심점으로 링크가 퍼져 있음을 알 수 있다. 따라서 이 단어들을 가지고 한·중·일을 비교분석 하였다.

1) 호텔(Hotel)

호텔과 관련된 연구들은 세 개 나라에서 차이를 보이고 있는데 중국은 직원·고객 만족 분야, 한국은 서비스품질 분야, 일본은 관광 목적지에 관한 연구가 중요한 연구 분야로 파악되었다.

2) 이미지(Image)

관광에서 현대는 이미지의 시대라 규정될 만큼 이미지는 생산적이며 생동적이고 창의적인 것으로 인정받고 있으며, 특히 정보와 미디어의 홍수 속에서 소비를 규명함에 있어서 이미지에 대한 분석은 마케팅에 있어서 초석을 제공한다(김영경, 2003)



〈그림 2〉 사회연결망분석 - 한국

V. 결 론

본 연구를 통해 한·중·일의 관광연구 동향을 파악 할 수 있었다. 빈도분석(TF-IDF)를 통해서 공통적으로 선호되는 주제와 국가별 선호되는 주제를 파악할 수 있었다. 한국과 일본에서는 중국단어의 가중치가 높은 것으로 중국에 대한 연구가 활발하였다. 반면 중국은 서비스업(Hospitality Industry)에 관한 연구가 활발하였다.

사회연결망(SNA)을 통해서 한·중·일 공통적으로 이미지(Image), 호텔(Hotel), 관광업(Tourism)을 구심점으로 여러 연구가 이루어짐을 알 수 있었다. 특이점으로 한국은 호

텔 온라인, 중국관광객연구가, 중국은 관광정책, 대만관광객연구가, 일본은 한국과중국관광객, 관광업의 경제, 성장, 개발에 대한 연구가 활발하였다.

참고문헌

- 김민선 (2016). 키워드 네트워크를 이용한 국내 관광연구의 최근 연구동향 분석. 『Journal of the Korea Academia-Industrial cooperation Society』 Vol. 17, No. 9 pp. 68-73
- 김영경 (2003). 커뮤니케이션 시대 이미지의 다원적 특성과 함의들. 『한국여성교양학회지』, 12 : 5-22
- 김창곤 · 신동주 · 지봉구 · 홍창식(2007). 관광학원론. 서울: 대왕사.
- 민웅기 (2016). 한국 관광사회학의 융합연구 생태계 구성 『관광연구논총』 제28권 제2호 pp. 3~25
- 송혜지 (2013). 텍스트 마이닝 기법을 활용한 한국의 경제연구 동향 분석. 『한국정보관리학회』, 20(8), 47-50
- 이지연 (2012). 관광분야의 새로운 분류체계 설계에 관한 연구. 『정보관리연구』. 43(3), 79-95
- 조정윤 (2010). 한중일 관광부문 인력시장 개방에 대한 변화요구와 대응방안. 『한국관광정책』, 42, 74-77