

도시 브랜드의 전략적 활용을 위한 빅데이터 분석 : 서울시 도시 브랜드 "I SEOUL U" 사례

Big Data Analysis for Strategic Use of Urban Brands: Case Study Seoul city brand "I SEOUL U"

임혜원

건국대학교 도시재생학과

Haewen Lim(imcharlotte@konkuk.ac.kr)

요약

본 연구에서는 서울시 도시 브랜드 I SEOUL U에 대한 인식과 평가를 분석하기 위하여 온라인 빅데이터를 대상으로 한 텍스트마이닝 분석을 수행하였다. 이를 위하여 데이터 수집 및 분석을 위한 처리 프로그램인 텍스트톰(Textom)을 사용하였고 'I SEOUL U' 키워드를 분석키워드로 선정하였다. 키워드 분석 결과 I SEOUL U와 관련된 키워드는 첫째, 비즈니스와 마케팅 관련 용어로서 팝업 스토어, 갤러리, 공동 브랜드, (축제 등) 개최, 상품, 민간기업, 온라인 등이다. 둘째, 이벤트 관련 용어로서 한강, 식목일, 나무 심기, 홍대, 크리스마스, 마포구, 중구, 세종대, 축제거리 등이다. 셋째는 홍보 관련 용어로서 로봇공학박사 데니스 홍, Government, 조형물, Korea 등이었다. N gram 분석 결과에서는 서울시 브랜드로서 공익적 성격의 도시 브랜드인 I SEOUL U의 경우에도 민간 기업의 상업 활동에 많은 기여를 하는 것으로 밝혀졌다. 연결 중심성 분석에서는 비즈니스 및 마케팅, 이벤트, 홍보 등의 범주가 도출되었다. 매트릭스 분석에서는 제품 판매와 관련하여 주로 팝업 스토어의 아이템들이 많고 공동 브랜드 형태의 제품들이 개발되는 것으로 나타났다. 토픽 모델링에서는 총 10개의 토픽이 추출되었고 상업적 활용과 이벤트 축제에 관한 정보 니즈가 많은 것으로 나타났다.

■ 중심어 : | 도시브랜드 | 서울 | ISEOUU | 빅데이터 | 텍스트마이닝 | 텍스트톰 |

Abstract

In this study, text mining analysis was performed on online big data for recognition and assessment of urban brand I Seoul U. To this end, TEXTOM, a processing program for data acquisition and analysis was used, and the 'I SEOUL U' keyword was selected as an analysis keyword. Keyword analysis shows the keywords associated with I Seoul U to be as follows: First, as a business and marketing term, keywords include pop-up store, gallery, co-branding, (festival, etc.), commodities, private companies and online. Second, as an event-related term, keywords include Han River, tree-planting day, tree planting, Hongdae, Christmas, Mapo, Jung-gu, Sejong University, and festival. Third, as a promotional term, keywords include robotics engineer Dr. Dennis Hong, Government, Art and Korea. In the N Gram analysis, as the city brand of Seoul, I Seoul U, in the public interest, was found to contribute to the commercial activities of private companies. In connection-oriented analysis, business and marketing, events, and promotions have been derived as categories. In matrix analysis, it was found that the products of the pop-up store are mainly developed, and products in the form of co-branding were being developed. In the topic modeling, a total of 10 topics were extracted and needs for commercial utilization and information for event festivals were mostly found.

■ keyword : | City Brand | Seoul | Seoul U | Big Data | Text Mining | Textom |

I. 서론

관광객, 투자자 및 재능 있는 사람들을 전 세계적으로 유치하기 위해 점점 더 많은 도시들이 도시 브랜드에 투자하고 있다. 도시 브랜드는 이론적이고 실용적인 맥락에서 도시 자치 정부의 경영에 대한 중요한 연구 주제이다. 도시 브랜드는 도시의 이야기를 전 세계에 전할 수 있는 차별화된 도시 경영의 전략적 도구이다. 이러한 차별화를 달성하기 위해 도시 브랜드는 문화 행사, 복원 및 유산 프로모션, 상징적인 건물의 구성[1] 등의 세 가지 주요 접근법을 활용한다. 이러한 세 가지 접근법 외에도 도시 브랜드의 경제적 활용은 매우 중요한 이슈이다. 본 연구는 다양한 문화적 내용부터 상업적 활용 가치까지 광범위한 도시 브랜드 효과를 분석하고 있다. 특히 도시 브랜드는 도시에 대한 인식의 변화뿐만 아니라 민간기업들에 새로운 비즈니스의 기회를 제공한다. 따라서 도시 브랜드에 대한 이용자들의 태도는 브랜드의 명성에 따른 새로운 도시 문화 창출과 미래의 경제적 가치를 끊임없이 창출해 내는 자산이다.

이러한 도시 브랜드의 다양한 효과를 분석하기 위하여 본 연구에서는 서울시의 도시 브랜드에 대한 이용자들의 인식과 태도 분석을 수행하고자 한다. 서울시 도시 브랜드의 대표적인 사례로서 I SEOUL U는 그동안 많은 성과를 나타내고 있다. 서울은 약 1천만 명의 주민들과 함께 하는 메가폴리스이며 600년 동안 한국의 수도이다. 이 도시는 어둡고 복잡한 역사적 배경에서도 급속한 경제 성장과 도시 개발을 위한 "한강의 기적"으로 국제적인 명성을 얻고 있다. 최근에 서울은 세계적인 관광도시를 꿈꾸며 도시 브랜드를 재배치하려고 노력하고 예술 및 문화 산업을 신속하게 확장하고 있다. 2000년대 이후 서울에서는 문화 행사의 수가 급속히 증가했으며 2007년 이후 서울시 정부의 마케팅 예산이 눈에 띄게 증가하고 있다. 그러나 많은 노력과 도시 당국의 개발에도 불구하고 서울은 도시 브랜드와 문화 행사에 대해서 아직 전략적인 방향을 잡지 못하고 있는 것으로 보인다. 따라서 서울시의 대표적인 도시 브랜드인 I SEOUL U에 대한 사용자들의 인식과 태도를 분석한 결과는 향후 도시 브랜드의 전략적 방향에 대해서 많은 시사점을 제공할 수 있을 것이다.

그동안 서울시의 도시 브랜드에 대한 실적을 살펴보면 서울시 정부는 34개 축제를 주최했고 현재는 서울의 25개 자치구에서 74개 축제를 관리하고 있다[2]. 서울시는 2002년 일본과 함께 FIFA 월드컵을 공동 주최한 이후 다음 해에 하이서울 페스티벌을 시작하고 이때부터 축제 문화를 근거로 하는 최초의 도시 브랜드 정책을 시행하였다. 그 이후 이 축제는 한국 문화를 바탕으로 도시 브랜드 개발을 촉진하는 계기가 되었고 서울 문화재단 (Seoul Foundation for Arts and Culture) 이 주최하는 연례행사가 되었다. 2003년부터 매년 5월 개최된 하이서울 페스티벌은 2008년부터는 사계절 축제로 진행됐으며 대한민국 공공행정 대상 '문화축제부문', 행정안전부 산하 '멋진 생각 상', 서울 관광대상 '우수관광 상품상'을 수상했다. 2009년에는 국제적인 경제위기의 경제난 극복에 동참하며 '나눔'과 '희망'을 주제로 한 시민 참여형 축제로 거듭났다. 개최 일정도 5월 첫째 주로 시기를 정례화하여, 매년 5월이 오면 서울 시민뿐만 아니라 전 세계인들이 축제를 기다리며 서울을 다시 찾게 되는 계기를 마련했다. 2015년부터는 10월에 개최됐고 2016년도부터 서울시의 슬로건이 I SEOUL U로 변경되면서 서울거리예술축제로 명칭이 바뀌었다. 이후에는 매년 가을 9~10월에 열리고 있다.

그러나 이렇게 다양한 서울시의 도시 브랜드 노력에도 불구하고 서울시의 이미지와 경제적 효과에 미치는 영향력에 대해서는 아직 연구된 바가 없다. 그 이유는 실제 이용자인 서울 시민들의 인식과 기업들의 경제적 목적 등에 관한 광범위한 자료를 수집하기 어렵기 때문이다. 따라서 본 연구는 도시 브랜드 연구의 새로운 접근법으로서 도시 브랜드 이용자들의 인식과 활용에 대한 광범위한 자료 수집을 위하여 빅데이터를 활용하였다. 본 연구에서는 I SEOUL U에 대한 온라인 텍스트 자료를 수집하고 텍스트마이닝을 이용하여 I SEOUL U에 대한 인식과 관련 활동에 대한 시민들의 정보검색 자료를 분석하였다.

빅데이터 분석은 과거에도 사용되었으나 최근 모바일 기기의 발전으로 인해 소셜 네트워크 서비스상에서 유통되는 데이터의 양이 증가하면서 그 활용성이 더욱 높아지고 있다[3]. 소비자들의 인식 조사는 일반적으로 설문조사를 통해 시행됐으나, 소셜 데이터 자료는 자발

적으로 표현되고 실시간으로 확보할 수 있으므로 기존에 인위적인 실험 환경이나 구조화된 설문 방식을 보완한 방법이다[4]. 최근 소비자들은 블로그, 커뮤니티, SNS 등 다양한 온라인 환경에서 제품과 서비스를 사용한 경험, 가치, 니즈를 표현하고 공유하고 있으므로 고객의 경험과 니즈, 가치를 텍스트의 문맥 관계 속에서 탐색할 수 있다[5].

빅데이터는 소비자들의 인식 및 트렌드를 분석하기에 적합하여 개인의 소비 행동, 선택분만 아니라 감정 등을 분석할 수 있으며, 다양한 분석을 통해 가까운 미래의 전망을 예측하는데 활용할 수 있다[6]. 따라서 본 연구에서는 I SEOUL U를 검색어로 하여 포털, SNS, 유튜브 등의 다양한 온라인 매체상에서 실제로 검색이 이루어진 댓글과 웹 페이지 등을 수집하였다. 또한, 이 페이지들을 대상으로 I SEOUL U가 포함된 텍스트에 관한 텍스트마이닝을 수행하였다. 텍스트마이닝은 주로 주제어 빈도수, 연관성 및 연결 중심성, 네트워크 분석, 토픽모델링 등의 도시 브랜드에 대한 인식과 연관 검색어 등의 상호작용을 분석하였다. 분석 결과는 서울시 도시 브랜드에 대한 평가와 활용방안 등에 관한 시사점을 찾고 서울시 도시 브랜드의 미래 활용방안을 제시하고자 하였다.

II. 이론적 배경

1. 도시 브랜드의 목표

도시의 주요 목표 중 하나가 도시의 매력을 높이기 위해 도시 마케팅 및 도시 브랜드 전략을 수행하는 것이다[7][8]. 도시 브랜드는 도시 주민의 행복을 약속하는 인본주의 철학을 기반으로 도시에 투자한 사람들에게 대한 새로운 투자 기회와 전략적 활용 기회를 제공한다[9]. 도시 브랜드는 도시에 대한 인식의 변화와 민간 기업들에게 새로운 비즈니스의 기회를 제공하고 명성 확보에 따른 새로운 도시 문화와 경제 활동을 끊임없이 창출해 내는 자산이다[10].

또한, 도시 브랜드는 도시의 핵심 가치와 활동을 식별하고 구축하는 것을 포함한다. 도시 브랜드는 국가의 효과적인 브랜드를 구축을 통하여 인바운드 투자 및 관

광자원 증대, 도시 투자자의 신뢰성 증가, 타 도시 및 공공 또는 민간 연구 및 대학 기관의 연계, 민간기업과의 협력을 통한 관련 제품 또는 서비스 개발, '원산지'효과, 시민 자부심 형성, 생산적인 글로벌 파트너십 등을 달성할 수 있다[10].

2. 도시 브랜드의 구축

도시 브랜드는 도시 브랜드라는 용어와 구별된다. 그것은 도시 브랜드의 명성을 확립하거나 관리하기 위해 브랜드 이름과 정체성을 설계, 계획하고 전달하는 과정이다[11]. 도시 브랜드는 도시의 정체성을 확립하는 것이고[12], 도시의 정체성의 핵심 요소에는 도시의 미래를 향상시키는 데 도움이 되는 모든 이해 관계자가 포함되어야 한다. 이해 관계자들은 도시의 미래에 관심이 있는 사람들, 그룹 및 조직의 이미지와 관광, 민간 부문, 외국 및 국내 정책, 투자 및 이민, 문화와 관련된 도시에 관한 이야기를 전하는 모든 중요한 요소를 포함해야 한다[13][14].

도시 브랜드의 중요한 도구는 주로 문화 행사이다. 도시 정부가 문화 행사의 성격을 규정하고 도시 브랜드에 대한 그 역할을 정하며 문화 행사에 대한 지원책을 마련하는 것이 매우 중요하다. 문화 행사가 갖는 도시 브랜드에 대한 역할은 경제적, 문화적, 사회적, 환경적, 정치적 요인 등 4개의 범주로 구분된다[15-17]. 도시 정부가 주최하는 이벤트 행사에서는 이벤트의 역할과 영향이 구체적으로 정해져야 한다. 이벤트 행사의 역할에는 "매력", "이미지 제작자", "정적 매력의 애니메이터", 경제적 의미에서 "다른 개발을 위한 촉매" 등이 포함된다[16]. 그러나 이러한 역할은 일반 대중이나 여행자가 아닌 그 도시의 정치적 목적과 경제적 목적 등에 의해서 포괄적으로 검토되고 정의되어야 한다[16]. 이러한 역할은 도시의 고유성에 기초한 브랜드 개발에 도움이 된다[18]. 또한, 문화 행사는 대부분의 관광 목적의 도시 브랜드에서 마케팅 전략 중 하나로 중요한 역할을 한다[19]. 따라서 많은 도시들은 도시의 관광 이미지를 향상시키기 위해 이벤트 행사를 고려하는 경우가 많다[20].

3. 서울시의 도시 브랜드의 개발

서울시는 2002년 일본과 함께 FIFA 월드컵을 공동 주최한 이후 하이 서울 페스티벌을 시작하고 도시 브랜드의 일환으로서 축제 문화를 근거로 하는 최초의 도시 브랜드 정책을 시행하였다. 그 이후 이 축제는 한국 문화를 바탕으로 도시 브랜드를 촉진하는 계기가 되었고 서울 문화 재단이 주최자가 되는 연례행사가 되었다. 2003년부터 매년 5월 개최된 하이서울페스티벌은 2008년 세계절 축제로 바뀌 진행됐으며 그 해에 대한민국 공공행정대상 '문화축제부문', 행정안전부 산하 '멋진 생각 상', 서울 관광대상 '우수관광 상품상'을 수상했다.

이렇게 해서 서울시 최초의 도시 브랜드가 선정되었고 2008년 10월경에는 서울의 이미지를 홍보하기 시작했는데 서울 최초의 도시 브랜드 전략으로 평가받고 있다[21]. 오세훈 시장은 이 도시 브랜드가 여전히 존재하고 있는 상태에서 2006년 시장으로 선출되었다. 그 이후에는 "아시아의 영혼"이 슬로건의 하단에 추가되었고 이 브랜드는 2002년부터 2006년까지 이명박 시장의 재임 동안 활용했다.

2014년에 "거리에서 놀자"라는 축제 이름으로 명칭이 바뀌었는데 이는 축제의 주제와 "서울 스트리트 아트 축제"로 축제의 테마를 설명하는 단어만 변경한 거에 불과하였다. 한편, 2016년에 예술과 문화를 위한 행사로써 공식적으로 축제의 이름이 바뀌었다[22]. 그 이후 "Hi Seoul"이라는 단어는 축제 이름에서 사라졌으며 2015년 축제의 테마를 대표하는 "서울 스트리트 아트" 축제로 그 명칭이 변경되었다.

그 이후 박원순 시장이 재임된 2015년부터 새로운 도시 브랜드가 시작되었고 이때부터 그 이름은 "I SEOUL U"로 불리게 되었다.

4. 도시 브랜드 개발을 위한 빅데이터의 중요성

4.1 빅데이터의 개념과 유형

빅데이터는 방대한 크기를 가지는 비정형 데이터로서 생성 주기가 짧아 이전에는 처리하기 어려운 데이터 집합을 의미한다. 빅데이터에 대한 생성, 수집, 분석, 표현의 발전은 현대 사회를 더욱 정확하게 예측하여 효율적으로 작동시키게 하며, 개인마다 맞춤형 정보를 제공, 관리, 분석이 가능해 과거에는 불가능했던 기술을 실현

하기도 한다.

빅데이터 분석의 장점으로는 수집된 대량의 비구조화된 데이터를 수집 및 분석하여 소비자가 요구하는 의견과 생각을 실시간으로 파악함으로써 상품과 서비스에 대한 예측을 통한 전략 수립과 새로운 가치를 창출하는 데 있다[23].

SNS 사용자들의 이용 동기를 분석한 연구들에서는 페이스북 이용 동기를 지인에 대한 정보를 찾고자 하는 사회적 검색 (social searching) 동기와 새로운 친구를 찾고자 하는 사회적 탐색 (social browsing) 동기로 구분하고, 사회적 탐색보다는 사회적 검색의 동기가 더 크다는 것을 밝혀냈다[24]. 또한, 사용자들이 사회적 연결, 정체성 공유, 콘텐츠, 사회적 탐색, 소셜 네트워크 탐색, 상태 업데이트 등의 동기를 가지고 페이스북을 이용하고 있다고 주장했다[25]. D. Zhao, and M. B. Rosson, (2009)는 마이크로블로그 서비스 중의 하나인 트위터의 사용자들을 분석한 결과, 사용자들이 사회적 네트워크상에서의 가시성 확보, 정보 획득, 스트레스 해소 등의 목적으로 가지고 이용하고 있다고 보았다[26].

매체 특성에 관한 연구에서는 SNS를 개방형 SNS (페이스북, 트위터)와 폐쇄형 SNS (네이버 밴드, 카카오톡)으로 구분한 후, 개방형 SNS의 경우에는 정보 추구 동기가, 폐쇄형 SNS의 경우에는 재미 추구의 동기가 더 높다는 것을 밝혀냈다[27]. 매체에 대해서는 페이스북과 인스타그램에 대한 이용 동기의 차이를 분석하였으며, 그 결과 페이스북 사용자는 지인과의 소통 동기가 더 높지만, 인스타그램 사용자는 자기표현, 새로운 인맥 추구, 탐색 및 둘러보기 동기가 더 높은 것으로 나타났다[28].

그동안 국내에서 성공한 SNS를 도입 시기별로 살펴보면 싸이월드, 네이버 블로그, 다음 카페와 같이 네트워크 플랫폼을 기반으로 기존의 오프라인 인맥을 온라인으로 연결하는 1세대 SNS를 시작으로, 페이스북과 트위터와 같이 불특정 다수와의 관계를 확대하고 관심 있는 콘텐츠를 중심으로 네트워크를 강화하는 2세대 SNS를 거쳐, 제한된 네트워크를 기반으로 하여 특정 주제에 대한 관심사를 공유하는 인스타그램과 같은 3세대 SNS로 진화되어 온 것을 볼 수 있다[29].

온라인 커뮤니케이션을 활성화하는 데 있어서 콘텐츠

츠 개발은 매우 중요한 요소이다. 3세대 SNS는 사진과 영상을 중심으로 관심 분야에 관한 공유가 중심이다. 이는 사회관계망을 우선하는 기존의 SNS와는 차이가 있다. 특히 대표적인 3세대 SNS인 인스타그램은 사진과 동영상과 같은 시각 이미지를 중심으로 구현되는 플랫폼으로, 전 세계적으로 일 사용자가 5억 명을 넘어섰고, 매일 약 9,500만 개의 사진이 업로드되고 일일 평균 약 40억 장의 사진이 전파될 정도로 폭발적으로 성장하고 있다[30].

이 밖에도 온라인 정보검색 기법의 발전은 온라인 커뮤니케이션을 활성화하는데 많은 기여를 하였다. 그중에서도 해시태그는 웹 문서에서 검색을 위해 활용되는 미터 태그와 같이 특정 키워드를 편리하게 검색할 수 있도록 도입된 기능으로서 온라인 텍스트 형성에 중요한 원천이 되고 있다.

일반적으로 SNS상에서 사용되는 해시태그는 콘텐츠의 북마크로서의 역할과 가상 커뮤니티의 심볼로서의 역할을 수행하고 있다고 알려져 있다[32]. SNS 매체의 대표적인 인스타그램의 경우, 이 두 가지 기본적인 역할 이외에 자신만의 독특한 언어 사용으로 자의적 범주화를 구축하는 등의 '언어 유희적 특성'을 지닌 예외적인 기능도 보이고 있음이 밝혀졌다[31].

특히 SNS에서 발생한 데이터는 개개인의 특성을 알아낼 수 있는 의미 있는 정보를 가지고 있으며, 이를 이용하여 기업의 마케팅 전략이나 의사결정 등 여러 가지 유용한 정보로 활용하고 있다[33]. SNS 빅데이터 분석은 소비자의 실시간 목소리를 듣는 '리스닝' 기능과 흐름을 관찰할 수 있는 '모니터링'을 가능케 도와주는 도구로써, 기존에 기업이 소비자 목소리를 듣기 위해 시행했던 설문조사에서는 기업에 대한 자유로운 고객 목소리를 듣기 힘들었던 점을 보완하여 현실에 가까운 소비자들의 목소리를 들을 수 있는 방법이다[4].

온라인 빅데이터의 가장 중요한 소스로서 SNS는 일반적으로 제품에 대해 감성을 잘 표현하기 때문에 제품 사용 전이나 사용 후 대중의 관심과 소비자의 반응을 파악하는 데 유용하다. 특히 소비자의 심리적 가치는 제품이나 서비스를 이용함으로써 얻을 수 있는 느낌이나 감정을 나타낸다[34]. 이 외에도 웹과 소셜 미디어 게시물, 논문, 신문, 잡지 기사, 그리고 이메일 등 실생

활에서 생산되는 자료 대부분이 바로 텍스트마이닝 분석대상이 될 수 있다. 의미 네트워크 분석은 소셜 네트워크 분석의 원리를 적용한 텍스트 분석방법으로, 특정 이슈에 대해 어떤 방식으로 단어들 사용되어 의제를 구성하는지 그 구조와 관계에 대해 분석할 수 있다[35].

4.2 레저 관광 산업의 텍스트마이닝의 활용

텍스트 마이닝은 비정형 텍스트 데이터로부터 의미 있는 패턴과 가치를 추출함으로써 새로운 지식을 발견하는 분석 기법을 의미한다[36]. 텍스트마이닝을 활용한 소비자 트렌드에 대한 연구들은 다양한 분야에서 시도되고 있으며 특히 레저, 관광 분야에서의 텍스트 마이닝을 활용한 시장 분석과 경영 전략 개발 부분에서 매우 활발한 연구가 진행되고 있다.

특히 스포츠 관련 온라인 텍스트 분석은 스포츠 마케팅의 핵심적인 도구로서 그 역할이 증대되고 있다. 코로나19 시대의 프로야구 시청 품질에 대한 시청자들의 문제점과 이슈를 파악하기 위해 '코로나19와 프로야구', '코로나19와 프로야구 무관중'과 관련된 키워드를 중심으로 텍스트(textom) 프로그램의 텍스트마이닝과 소셜 네트워크 분석을 활용한 연구를 통해 시청 품질의 개선과 전략을 제시하기도 하였다[37]. 프로축구와 관련해서는 키워드 데이터를 텍스트 프로그램을 이용하여 수집하고 텍스트마이닝 분석을 통하여 프로축구 소비자 인식 변화를 분석하고 프로축구의 발전 전략을 제시하였다[38]. 스포츠산업의 서비스화 방안을 제시하기 위하여, 스포츠산업의 빅데이터 활용사례도 점점 늘어나고 있다. 특히 이러한 연구에서는 빅데이터 분석 플랫폼인 텍스트와 웹트렌드 솔루션을 사용하여 포털·소셜 네트워크에서 스포츠와 연관된 키워드를 분석하고 있다[39].

스포츠 관련 분야의 텍스트마이닝 활용 가치는 특히 관광, 레저 등의 다양한 분야까지 그 적용 범위가 확대되고 있다. 특히 관광 분야에서는 뉴트로 트렌드 현상을 효과적으로 분석하기 위해 네이버 블로그와 카페를 통해 수집한 소셜 빅데이터를 분석하기도 한다. 관광 주제의 다양한 텍스트에서 주요 키워드의 빈도와 TF-IDF를 추출하고 다양한 차원의 이미지와 뉴트로 관광에 대한 감성적 차원을 분석하기도 한다. 최근의 연구 동향을 살

퍼보면 관광 브랜드에 대한 뉴트로 관광 이미지와 감성의 현황을 규명하고 현대산업 분야에 빅데이터 활용방안과 마케팅 전략 방향성을 연구하였다[40].

김태호 외(2019)는 빅데이터 분석을 활용하여 체류형 울산관광 활성화 방안을 연구하였다[49]. “울산관광” 관련 키워드에서 추출된 주요 단어를 중심으로 체류형 울산관광 활성화 방안에 대하여 음식 관광, 역사 문화관광, 체험관광, 레저스포츠관광, 해양관광, 생태관광, 산업관광 등을 제안하였다.

윤혜진 외(2017)는 웹오브사이언스 데이터베이스에서 1990년부터 2017년까지 발표된 논문 초록 594개를 대상으로 하여 텍스트마이닝 기법 및 시계열분석을 적용하였다[50]. 텍스트마이닝 결과 총 20개의 토픽이 도출되고 시계열분석을 통하여 관광 테마의 트렌드를 분석하였다.

최근에는 언택트 시대 새로운 여행의 테마로 자리 잡은 캠핑에 대한 트렌드 변화를 분석하기 위해 코로나 19 발생 이전과 이후의 온라인 자료에 대한 텍스트마이닝 기법이 많이 활용되고 있다. 이러한 연구에서는 네이버와 다음의 구글의 블로그 및 카페의 글을 통해 검색어 기능을 이용하여 관련 텍스트를 대량을 수집한 후에 텍스트를 이용하여 단어의 TF-IDF 분석 및 N-gram 네트워크 시각화를 통한 분석 등을 실시하였다. 이를 바탕으로 관광목적지, 관련 기업과 캠핑용품의 생산, 유통 기업, 관련 서비스업의 마케팅 전략과 정책 등에 대해 효율적인 방안을 모색하기도 한다[41].

관광 산업의 비즈니스호텔의 경우에도 텍스트 마이닝은 매우 활용도가 높다. 호텔에 대한 소비자들의 인식 동향을 파악하기 위해 소셜 미디어상의 텍스트 데이터를 분석대상으로 하여 텍스트마이닝을 실시한 연구들이 등장하고 있다. 특히 이러한 연구에서는 소셜 네트워크 분석 방법을 적용하여 핵심단어의 중심성을 분석하고, 단어들 간에 존재하는 네트워크를 파악하여 도출된 핵심어 분석을 통해 긍정적인 인식을 바탕으로 고객 중심의 대안을 제시하고, 비즈니스호텔의 현재 위치에 입각한 마케팅 방안을 지원하고, 질적 성장을 위한 구체적 방안들을 제시하고 있다[42].

또한, 지역 관광 산업에 대한 지원을 위하여 컨벤션 육성 사업이 해당 지역의 사업 계획에 부합하여 육성

및 실행되고 있는지를 파악하고 미디어에 투영된 지역 특화 컨벤션 실행현황을 확인하고자 텍스트마이닝을 활용하기도 한다. 해당 지역의 온라인 기사를 수집하여 텍스트(Textom)을 통한 언어 네트워크분석을 진행하고 이를 바탕으로 지역특화컨벤션 육성 사업을 더욱 발전시키는 방안을 모색하기도 한다[43].

III. 연구 방법론

1. 데이터 수집 및 전처리

본 연구에서는 서울시 도시 브랜드 I SEOUL U에 대한 인식과 평가를 위하여 온라인 빅데이터를 대상으로 한 텍스트마이닝 분석을 수행하였다. 이를 위하여 데이터 수집 및 분석을 위한 처리 프로그램인 텍스트(Textom)을 사용하였다.

텍스톰은 인터넷상의 데이터를 원하는 채널 별로 수집하며, 수집된 데이터를 정제하고 매트릭스 생산까지 처리해주는 빅데이터 처리 솔루션으로 데이터의 연관성과 정확도에 따라서 추출하는 것이 가능하다. 텍스톰은 빅데이터 텍스트 마이닝 기법을 사용하여 자료를 수집하는데 비정형 데이터 자료에서 동시 출현하는 단어 빈도를 확인하여 단어의 중요도를 파악할 수 있다[44]. 또한, 비정형 데이터 자료로부터 구조화된 형태의 정보를 추출하여 전체 데이터 구조를 시각화 자료를 통해 확인하게 되며 핵심단어의 빈도와 단어 간 관계를 분석하여 데이터 및 연관 키워드 순위를 제공하고 검색 키워드의 공동 출현 빈도에 따른 매트릭스 정보를 제공하고 있어 의미 체계를 파악하는 것이 가능하여 네트워크 분석 시 유용한 소프트웨어이다[44].

본 연구에서는 서울시 도시 브랜드 I SEOUL U에 대한 분석을 위하여 ‘I SEOUL U’ 키워드를 분석키워드로 선정하였다. 데이터 수집과 분석을 위하여 온라인 프로그램인 텍스톰을 사용하였으며, 수집 기간은 2020년 9월부터 2021년 9월까지이고 수집 대상은 네이버(웹페이지, 블로그, 카페, 지식IN), 구글(웹페이지), 페이스북, 유튜브, 트위터 등의 댓글 등을 포함했다.

2. 텍스트마이닝(Textmining) 분석 절차

텍스트마이닝은 비정형의 텍스트 데이터를 자연어 처리와 형태소 분석기술로 수집어를 정제하고, 단어를 추출해 빈도수를 제시하여 순위나 인식의 유사성, 일반성을 찾아내는 데 사용된다[45]. 본 연구에서는 'I SEOUL U' 키워드를 중심으로 텍스트에서 추출된 텍스트를 텍스트의 정제/형태소 분석의 직접 선택을 사용하여 정제하였다. 'I SEOUL U' 키워드가 포함된 온라인 문서에서 앞뒤 단어 3개를 기준으로 6개의 단어를 분석대상 범위로 지정하고 명사, 대명사, 형용사의 단어를 추출하여 분석 데이터로 활용한다. 또한, 키워드들의 관련성 분석을 위하여 연결 중심성 분석을 하였다. 연결 중심성이 높은 단어는 연구 주제의 핵심 주제어이기 때문에 연구 분야에서 중심이 되는 주제어라고 볼 수 있다. 따라서 빈도수가 높고 연결 중심성까지 높다면 그 노드는 영향력이 큰 핵심적인 키워드라고 볼 수 있다. 다시 말해 I SEOUL U 도시 브랜드 분야의 연구 동향을 논의할 때 반드시 고려해야 할 필수적인 연구 영역이라고 할 수 있다.

본 연구는 텍스트(TEXTOM) 프로그램을 통해 데이터 정제 작업을 진행하고, 텍스트마이닝의 단어 빈도수 분석, N-gram 분석, TF-IDF 분석, 바 차트와 워드 클라우드 작업에 이어서 연결 중심성 분석, 매트릭스 분석, 매트릭스 차트, 토픽모델링 분석을 진행하였다.

N-gram은 n개 단어가 연속적으로 나오는 빈도를 의미한다. 연속되는 두 각 단어의 빈도가 높게 나타난 것은 두 단어가 나란히 있을 빈도가 높음을 의미한다.

TF (단어 빈도수, term frequency)는 어떠한 단어가 특정 범위 내에서 얼마나 중요한지를 빈도로 표시하는 통계치이다. DF (문서 빈도수, document frequency)는 특정한 단어가 문서 내에서 사용되는 빈도를 나타내며, 특정 단어가 문서 빈도수가 높을수록 그 단어는 보편적 단어라는 의미로 해석할 수 있다. 또한, IDF (inverse document frequency)는 DF의 역수를 말하고, 특히, 문서 간의 역수를 취함으로써 DF가 커질수록 IDF는 감소한다.

중심성은 연결 중심성과 위세 중심성과 같이 노드가 네트워크의 중심에 위치하는 정도를 나타내는 값이다. 연결 중심성은 다른 단어와 직접 연결된 정도를 의미하며, 연결된 인접 노드들의 총 개수로 측정된다. 따라서

연결 중심성이 높은 단어일수록 연구의 중심이 되는 대상이 될 가능성이 크다. 매트릭스 분석은 정제된 단어 간의 동시 출현빈도(공출현빈도: co-occurrence)로 작성된 매트릭스 데이터를 토대로 키워드 간의 연관성 분석 정보를 제공한다. 동시 출현빈도 즉 공출현(co-occurrence)빈도가 높다는 것은 일정 범위 내에서 노드들이 동시에 출현했을 경우 그 범위 내 노드들 간에 의미에 따라 상호연관성이 높음을 의미한다. 특히 본 연구에서는 구조적 등위성 분석(CONCOR, CONvergence of interation CORrelation)을 통하여 단어 간의 동시 출현 매트릭스의 피어슨 상관관계 분석을 토대로 단어들의 블록을 식별해내고 블록들 간의 관계를 파악하고자 하였다. 토픽모델링은 텍스트 문서의 집합으로 구성된 단어를 분석하여 잠재적인 주제를 도출해내는 확률 기반의 모형으로 텍스트마이닝 기법의 하나다. 이 기법은 텍스트를 구성하는 단어들을 확률 통계적으로 분석하고 자료를 관통하는 주제와 주제의 연결성 주제들을 시간의 흐름에 따라 변화를 추출한다[47].

잠재 디리클레 할당은 대표적인 토픽모델링 기법으로 사람이 문서를 작성한다는 가정에서 출발하는 알고리즘이다. 그런 다음, 디리클레 분포를 기반으로 하는 생성 모형으로 문서의 잠재적 주제를 찾아내고 해당 문서 범위에서 토픽들이 점유하는 비율과 문서의 개별단어가 토픽에 해당하는 확률을 보여준다[47][48].

본 연구에서는 토픽 분석을 위하여 텍스트의 word level semantic clustering(=문서 내의 단어들의 공출현관계를 토대로 벡터화하여 인접단어를 같은 그룹으로 묶어주는 기법)을 활용하였으며, 토픽 수 10개, 토픽별 단어 수 5개를 분석하였으며, 각각의 토픽을 선택할 시 나오는 단어키워드는 클릭한 키워드와 공출현관계가 형성된 단어키워드이며 워드벡터상으로 볼 경우에도 인접률이 높아서 연관성이 높다고 해석할 수 있다.

IV. 분석 결과

1. 데이터 수집 및 전처리 결과

최근 1년간(2020년 9월 20일~2021년 9월 20일) 'I SEOUL U' 키워드가 포함된 네이버(블로그, 카페), 다

음(블로그, 카페), 구글(웹페이지), 페이스북, 트위터의 온라인 문서를 추출하여 분석하였다. 추출된 키워드는 52,594개를 대상으로 하였다. 정제/형태소 분석으로 분리정제 작업을 거쳐 최종적으로 분석에 사용된 수집 양은 상위 100위 안에 있는 16,169개의 키워드를 대상으로 분석하였다.

아래 [그림 1]은 각 자료 수집처 별 단어 빈도수를 나타내고 있다. 빈도수는 다음의 블로그와 카페를 합칠 때 가장 많고 그다음은 네이버 블로그인 것으로 나타났다. 이 외에도 구글 페이스북, 트위터 등의 SNS 댓글에 대한 자료도 수집되어서 온라인 텍스트는 국내 포털과 주요 SNS 등이 모두 포함되어 있다.

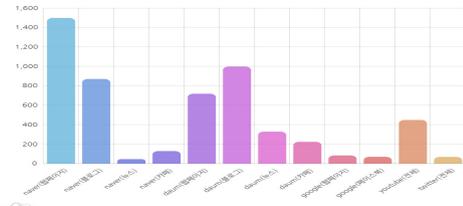


그림 1. ISEOULU 관련 주요 단어 검색 출처별 분포

2. 단어 빈도수 분석

I SEOUL U 키워드 분석은 I SEOUL U 검색어와 관련하여 텍스트상에서 가장 자주 등장하는 단어들의 빈도수를 기준으로 명사와 외국어 단어 중심으로 추출하였다. [표 1]은 I SEOUL U를 포함한 텍스트에 대한 키워드 분석을 시행한 결과 상위 50개 단어의 빈도수를 보여주는 결과표이다.

상위에 랭크된 50개 키워드는 파트너스, 한강, 달력, 기업, 팝업 스토어, 갤러리, 식목일, 제품, 공동 브랜드, 개최, 나무 심기, 판매, 활용, 상품, 보라빛, 안전도, Metropolitan, 이벤트, 상징물, 공지사항, 흥대, 다양, 테니스, Christmas, 민간기업, 제작, 사람, 시장, 마포구, 중구, 진행, 온라인, 기대, 사이즈, 장식, 책상 앞, rights, 세종대, 축제 거리, Government, 안전, 진행편, 코칭, 사진, 마음, 준비, 조형물, 로봇공학박사, Korea 등으로 나타났다. 일반어를 제외할 경우 I SEOUL U와 관련된 키워드는 첫째, 비즈니스와 마케팅 관련 용어들이다. 이 범주에 속하는 단어들은 파트너스, 기업, 팝업스토어, 갤러리, 제품, 공동 브랜드, (축제 등) 개최, 판매, 상품, 민간기업, 제작, 시장, 온라인 등이다.

표 1. I SEOUL U 상권에 대한 상위 50개 단어의 빈도수분포

순위	키워드	빈도수	비중	순위	키워드	빈도수	비중
1	파트너스	896	1.28%	26	제작	240	0.34%
2	한강	819	1.17%	27	사람	236	0.34%
3	달력	770	1.10%	28	시장	232	0.33%
4	기업	496	0.71%	29	마포구	232	0.33%
5	팝업스토어	446	0.64%	30	중구	229	0.33%
6	갤러리	443	0.63%	31	진행	228	0.32%
7	식목일	432	0.62%	32	온라인	228	0.32%
8	제품	428	0.61%	33	기대	223	0.32%
9	공동브랜드한	400	0.57%	34	사이즈	222	0.32%
10	개최	382	0.54%	35	장식	221	0.31%
11	나무심기	378	0.54%	36	책상앞에	220	0.31%
12	판매	373	0.53%	37	rights	220	0.31%
13	활용	357	0.51%	38	세종대	219	0.31%
14	상품	346	0.49%	39	축제거리에서	218	0.31%
15	보라빛	324	0.46%	40	Government	216	0.31%
16	안전도	323	0.46%	41	안전	214	0.30%
17	Metropolitan	316	0.45%	42	진행편	208	0.30%
18	이벤트	312	0.44%	43	코칭	208	0.30%
19	상징물	285	0.41%	44	등록	206	0.29%
20	공지사항	274	0.39%	45	사진	202	0.29%
21	흥대	259	0.37%	46	마음	197	0.28%
22	다양	254	0.36%	47	준비	194	0.28%
23	테니스	250	0.36%	48	조형물	182	0.26%
24	Christmas	241	0.34%	49	로봇공학박사	182	0.26%
25	민간기업	241	0.34%	50	Korea	179	0.26%

4. TF-IDF분석

TF (단어 빈도수, term frequency)는 특정한 단어가 1개 문서 내에서 얼마나 자주 등장하는지를 나타내는 단어 빈도수 값을 나타낸다. DF(문서 빈도수, document frequency)는 특정한 단어가 나타나는 문서 수를 나타내며, 일반적으로 특정 단어가 여러 문서에서 빈도가 높은 경우, 그 단어는 흔한 보편적 단어라는 의미로 해석한다. IDF(inverse document frequency)는 DF의 역수를 말하고, 특히, 문서 간의 역수를 취함으로써 DF가 커질수록 IDF는 감소한다. 따라서 $TF - IDF = TF \times 1/DF$ 이므로 특정 범위 내에서 모든 단어의 빈도수와 단어가 포함된 문서(정보)들의 빈도수를 구한 후 역수를 취해 곱하여 문서의 중요도를 찾는 방법이다.

분석 결과 달력, 파트너스, 한강, 식목일, 나무심기, 팝업스토어, 제품, 보라빛, 공동 브랜드, 기업, 활용, 개최, 안전도, 갤러리, 상품, 상징물, 판매, Christmas, 이벤트, 온라인, 데니스, 안전, 진행편, 코칭, Metropolitan 등이 상위에 링크되어 있다[표 3].

5. 연결 중심성 분석

텍스트 네트워크 분석에서 사용되는 '연결 중심성 (degree centrality)' 지표는 노드가 네트워크의 중심에 위치하는 정도를 나타내는 값이다. 연결 중심성은 다른 단어와 직접 연결된 정도를 의미하며, 연결된 인접 노드들의 총 개수로 측정된다. 따라서 연결 중심성이 높은 단어일수록 I SEOUL U와 관련된 온라인 커뮤니케이션의 중심이 될 가능성이 크고 이 값이 큰 노드는 연결된 대상들이 많아서 이 노드를 제거하면 수많은 온라인 댓글의 구성이 어렵게 된다[46].

I SEOUL U의 연결 중심성은 분석키워드 'I SEOUL U'와 주요 키워드들의 직접적인 연결 관계를 기반으로 해 수치화한 것으로 연결성이 높은 단어는 분석키워드와의 관계에서 영향력이 높은 단어들이라고 볼 수 있다.

연결 중심성 분석을 위하여 텍스트 매트릭스에서 생성된 빈도 매트릭스를 기반으로 'I SEOUL U' 키워드 중심성 분석 결과 한강, 사진, 사람, 아이, 조형물, 여의 연결 중심성 지수를 분석하였다[표 4].

표 3. I SEOUL U 관련 주요 단어에 대한 TF-IDF 분석 결과

순위	키워드	TF_IDF	순위	키워드	TF_IDF
1	달력	2285.78	26	공지사향	818.41
2	파트너스	2124.16	27	시장	743.12
3	한강	1848.80	28	홍대	740.98
4	식목일	1589.78	29	다양	732.97
5	나무심기	1391.06	30	사람	708.21
6	팝업스토	1296.46	31	민간기업	697.48
7	제품	1220.99	32	제작	692.57
8	보라빛	1192.33	33	마포구	680.42
9	공동브랜드한	1183.80	34	진행	674.77
10	기업	1169.71	35	중구	670.62
11	활용	1164.64	36	기대	659.97
12	개최	1159.10	37	사이즈	657.01
13	안전도	1131.26	38	장식	655.05
14	갤러리	1130.11	39	책상앞에	653.08
15	상품	1120.42	40	rights	653.08
16	상징물	936.80	41	세종대	651.11
17	판매	917.41	42	축제거리에서	650.14
18	Christmas	895.99	43	사진	648.19
19	이벤트	891.34	44	Government	645.17
20	온라인	885.74	45	로봇공학박사	643.16
21	데니스	883.47	46	Dennis	643.16
22	안전	865.56	47	등록	627.08
23	진행편	855.65	48	조형물	626.48
24	코칭	855.65	49	마음	609.58
25	Metropolitan	825.64	50	서울	606.03

표 4. I SEOUL U 관련 주요 단어에 대한 주요 단어의 연결

순위	키워드	연결중심성	순위	키워드	연결중심성
1	한강	0.0191	26	오세훈	0.0049
2	사진	0.0191	27	여행	0.0048
3	사람	0.0161	28	시청	0.0048
4	아이	0.0139	29	출산	0.0048
5	조형물	0.0122	30	날씨	0.0046
6	여의도	0.0104	31	시작	0.0046
7	신청	0.0092	32	설치	0.0046
8	Korea	0.0082	33	진행	0.0044
9	시장	0.0078	34	코로나	0.0043
10	시민	0.0078	35	대표	0.0043
11	서울숲	0.0075	36	나들이	0.0042
12	산책	0.0069	37	다양	0.0042
13	한강공원	0.0066	38	사용	0.0042
14	슬로건	0.0066	39	University	0.0042
15	도시	0.0064	40	euro	0.0041
16	영상	0.0062	41	방문	0.0040
17	자전거	0.0061	42	가을	0.0040
18	생각	0.0060	43	이벤트	0.0039
19	로고	0.0059	44	마음	0.0039
20	Korean	0.0057	45	길	0.0038
21	노들섬	0.0056	46	홍보	0.0038
22	기자	0.0055	47	도착	0.0038
23	공원	0.0053	48	광장	0.0038
24	시간	0.0050	49	하늘	0.0037
25	사랑	0.0050	50	사이트	0.0036

중심성 분석 결과 한강, 사람, 아이, 조형물, 여의도, 신청, Korea, 시장, 시민, 서울숲, 산책, 한강공원, 슬로건, 도시, 영상, 자전거, 생각, 로고, Korean, 노들섬, 기자, 공원, 시간, 사랑 등의 키워드들의 연결 중심성이 매우 높은 것으로 나타났다. 앞에서 단어 빈도 분석에서 3개의 카테고리가 도출되었다. 크게 비즈니스 및 마케팅, 이벤트, 홍보 등의 범주로 도출되었고 연결 중심성 분석 결과에서는 이러한 범주별로 중심 주제어들이 잘 나타나 있다. 그러나 비즈니스 및 마케팅 범주에서는 단어 빈도 분석에서 상위 랭크 되었던 단어들이 거의 있지 않아서 비즈니스와 마케팅 범주에서는 모든 단어의 연결 중심성이 낮은 것으로 나타나 비즈니스와 마케팅은 관심도와 비교하면 실제로 온라인 소통과 커뮤니케이션 수준은 매우 낮은 것으로 나타났다. 두 번째 범주인 이벤트의 경우에는 한강, 여의도, 시민, 서울숲, 한강공원, 도시, 자전거, 노들섬, 공원 등의 이벤트 관련 단어 이들의 연결 중심성이 높은 것으로 나타나 이벤트 관련 주제어들이 온라인상에서 활발한 소통이 일어나는 것으로 보인다. 홍보 관련 단어들의 경우에는 조형물, 슬로건, 영상, 로고, Korean 등의 연결 중심성이 높은 것으로 나타나 홍보 관련 단어들이 I SEOUL U

관련하여 온라인상에서 많은 커뮤니케이션이 일어나고 있는 것으로 나타났다.

6. 매트릭스 분석 및 매트릭스 차트

매트릭스 분석은 정제된 단어 간의 동시 출현 빈도(공출현빈도: co-occurrence)로 작성된 매트릭스 데이터를 토대로 키워드 간의 연관성 분석 정보를 제공한다. 공출현이란 전체 텍스트 내(한 행) 특정 범위에서 노드들이 같이 출현하였을 때 이 범위 내에 있는 모든 노드 간에 의미론적으로 상호 연관되는 관계가 있다고 추정할 수 있다. 수학적 분석을 위한 소셜 네트워크 데이터의 기본 형태는 테이블 형태의 매트릭스로서 동시 출현 빈도 행렬을 산출하여 분석하고자 하는 해당 데이터의 공출현 빈도, 근접 관계 등을 통해 네트워크로 시각화 정보를 제공해준다.

본 연구에서는 텍스트에서 제공하는 매트릭스를 이용하여 I SEOUL U와 관련된 주요 키워드 간의 상관관계 분석을 시행하고 각 단어의 동시 출현 가능성에 대한 정보를 토대로 I SEOUL U에 대한 인식을 분석하고자 하였다. 상관성이 높은 연관어들은 다음과 같다.

모형으로 텍스트마이닝 기법의 하나다. 이 기법은 텍스트를 구성하는 단어들을 확률·통계적으로 분석하고 자료를 관통하는 주제와 주제의 연결성 주제들을 시간의 흐름에 따라 변화를 추출한다[47]. 잠재 디리클레 할당은 대표적인 토픽모델링 기법으로 사람이 문서를 작성한다는 가정에서 출발하는 알고리즘이다. 그런 다음, 디리클레 분포를 기반으로 하는 생성 모형으로 문서의 잠재적 주제를 찾아내고 해당 문서 범위에서 토픽들이 점유하는 비율과 문서의 개별단어가 토픽에 해당하는 확률을 보여준다[47]. 본 연구에서는 토픽 분석을 위해 텍스트의 Word level semantic clustering을 활용하였으며, 토픽 수 14개, 토픽별 키워드 수 5개를 선정하여 분석하였으며, 그 결과는 아래의 [표 6]과 같다.

각 전체 토픽을 클릭 시 나오는 키워드들은 클릭한 키워드와 공출현 사례가 많이 발생한 단어들이며, 실제 워드 벡터 상으로도 인접률이 높아 연관성을 가진다고 해석할 수 있다.토픽 1은 식목일 나무 심기와 관련하여 이벤트 행사와 관련된 토픽이다.토픽 2는 서울시장 중심의 유튜브 채널을 통한 다양한 홍보물에 관련된 토픽이다.

토픽 3은 상징물의 개발과 민간기업과의 협업에 관한 토픽이다. 토픽 4는 서울시 관광 지역 등에서 다양한 이벤트와 홍보 등의 활동에 관련된 토픽이다. 토픽 5는 시민 참여와 지원을 위한 행정적 지원 활동에 관한 토픽이다.

토픽 6은 도시 안전과 관련된 토픽인데 박물관과 연관성이 있는 것으로 보아 I SEOUL U의 활용 측면과 관련된 토픽으로 추정된다. 토픽 7은 홍보의 도구 제작에 관한 주제로 서울시 관광명소를 활용한 달력과 캘러리 등의 제작 관련 토픽이다. 토픽 8은 민간기업의 판매와 관련하여 유통채널로서 팝업스토어 관련 제품과 판매장소 등에 관한 토픽이다. 토픽 9는 코로나와 관련하여 마스크에 새겨진 I SEOUL U 로고에 관한 토픽인 것으로 추정된다. 토픽 10은 조형물 홍보물 제작과 설치에 관한 주제로 대학과의 연계된 활동인 것으로 추정된다.

V. 결론 및 논의

1. 분석 결과 요약

본 연구에서는 서울시 도시 브랜드 I SEOUL U에 대한 인식과 평가를 위하여 온라인 빅데이터를 대상으로 한 텍스트마이닝 분석을 수행하였다. 이러한 연구의 결과는 도시 브랜드의 효과가 주로 어떤 분야에서 나타나는지 정보를 제공해주고 그 결과 도시 브랜드의 효과를 극대화할 방안들에 대한 정책적 시사점들을 제공해 줄 것이다.

표 6. 토픽모델링 분석결과

토픽	중요도1	중요도2	중요도3	중요도4	중요도5
1	식목일 0.029	나무심기 0.025	자전거 0.022	Junior 0.021	준비 0.019
2	유튜브 0.019	시장 0.019	설치 0.018	오세훈 0.014	조형물 0.013
3	상징물 0.068	활용 0.032	모집 0.025	협업 0.021	개발 0.021
4	관광특구 0.024	홍보영상 0.02	다양 0.017	안전 0.017	광화문 0.013
5	시민 0.036	기대 0.021	호감 0.02	주제 0.018	지원 0.018
6	자체 0.02	안전 0.018	Museum 0.015	Security 0.011	도시 0.01
7	달력 0.059	갤러리 0.035	한강 0.028	제작 0.025	마음 0.023
8	팝업스토어 0.07	제품 0.058	판매 0.054	민간기업 0.051	홍대 0.047
9	높이 0.027	마스크 0.024	로고 0.022	모습 0.017	등장 0.012
10	조형물 0.038	설치 0.02	위로 0.016	University 0.015	Daegu 0.014

본 연구에서는 데이터 수집 및 분석을 위한 처리 프로그램인 텍스트(Textom)를 사용하였다. 이를 위하여 서울시 도시 브랜드 I SEOUL U에 대한 분석을 위하여 'I SEOUL U' 키워드를 분석키워드로 선정하였다. 데이터 수집과 분석을 위하여 온라인 프로그램인 텍스트를 사용하였으며, 수집 기간은 2020년 9월부터 2021년 9월까지이고 수집 대상은 네이버(웹페이지, 블로그, 카페, 지식IN), 구글(웹페이지), 페이스북, 유튜브, 트위터 등의 댓글 등을 포함시켰다.

단어 빈도수는 일반어를 제외할 경우 I SEOUL U와 관련된 키워드는 첫째, 비즈니스와 마케팅 관련 용어들이다. 이 범주에 속하는 단어들은 파트너스, 기업, 팝업 스토어, 갤러리, 제품, 공동 브랜드, (축제 등)개최, 판매, 상품, 민간기업, 제작, 시장, 온라인 등이다. 둘째, 이벤트 관련 용어들이다. 한강, 달력, 식목일, 나무심기, Metropolitan, 이벤트, 상징물, 홍대, 크리스마스, 마포구, 중구, 안전, 진행, 장식, 세종대, 축제거리 등이 여기에 속한다. 셋째는 홍보 관련 용어들이다. 활용, 공지 사항, 로봇공학박사 데니스 흥, rights, Government, 등록, 사진, 조형물, Korea 등이 이 범주에 속하는 단어들이었다.

N gram의 분석 결과에서는 I SEOUL U와 관련된 N gram의 첫 번째 연관어는 식목일 나무 심기이며 I SEOUL U는 주로 식목일 이벤트 행사로 가장 많이 활용되는 것으로 보인다. 그다음은 제품 판매로서 I SEOUL U 관련 제품의 개발과 판매 관련 텍스트가 매우 많은 것으로 나타났다. 그다음은 민간기업의 공동 브랜드로서 빈도수에 나타난 바와 같이 I SEOUL U는 민간기업의 상업적 활용 도구로서 매우 빈번하게 사용되고 있다. 따라서 서울시 브랜드로서 공익적 성격의 도시 브랜드는 I SEOUL U도 민간기업의 상업 활동에 많은 기여를 하는 것으로 추정된다.

TF-IDF 분석 결과에서는 달력, 파트너스, 한강, 식목일, 나무 심기, 팝업 스토어, 제품, 보랏빛, 공동 브랜드, 기업, 활용, 개최, 안전도, 갤러리, 상품, 상징물, 판매, Christmas, 이벤트, 온라인, 데니스, 안전, 진행편, 코칭, Metropolitan 등이 상위에 링크되어 있다.

연결 중심성 분석에서도 단어 빈도수 분석 결과와 마찬가지로 3개의 카테고리가 도출되었다. 크게 비즈니스

및 마케팅, 이벤트, 홍보 등의 범주로 도출되었고 연결 중심성 분석 결과에서는 이러한 범주별로 중심 주제어들이 잘 나타나 있다. 그러나 비즈니스 및 마케팅 범주에서는 단어 빈도 분석에서 상위 랭 되었던 단어들이 거의 있지 않아서 비즈니스와 마케팅 범주에서는 모든 단어의 연결 중심성이 낮은 것으로 나타나 비즈니스와 마케팅은 관심도와 비교하면 실제로 온라인 소통과 커뮤니케이션 수준은 매우 낮은 것으로 나타났다. 두 번째 범주인 이벤트의 경우에는 한강, 여의도, 시민, 서울숲, 한강공원, 도시, 자전거, 노들섬, 공원 등의 이벤트 관련 단어들의 연결 중심성이 높은 것으로 나타나 이벤트 관련 주제어들이 온라인상에서 활발한 소통이 일어나는 것으로 보인다. 홍보 관련 단어들의 경우에는 조형물, 슬로건, 영상, 로고, Korean 등의 연결 중심성이 높은 것으로 나타나 홍보 관련 단어들이 I SEOUL U 관련하여 온라인상에서 많은 커뮤니케이션이 일어나고 있는 것으로 나타났다.

매트릭스 분석 결과에서는 상관계수의 값으로 추정된 결과에서 I SEOUL U 관련 상품과 판매에 대한 연관성이 대부분 매우 높게 나타나고 있다. 제품 판매와 관련하여 주로 팝업 스토어의 아이템들이 많고 공동 브랜드 형태의 제품들이 개발되는 것으로 나타났다. 제품 테마는 주로 I SEOUL U를 상징하는 상징물들이 많은 것으로 보이며 I SEOUL U의 팝업 스토어 매장과 이벤트 관련하여 홍대 근처 지역이 매우 큰 비중을 차지하는 것으로 나타났다.

마지막으로 토픽모델링 분석 결과에서는 총 10개의 토픽이 추출되었다. 토픽 1은 식목일 나무 심기와 관련하여 이벤트 행사와 관련된 토픽이다. 토픽 2는 서울시장 중심의 유튜브 채널을 통한 다양한 홍보물에 관련된 토픽이다. 토픽 3은 상징물의 개발과 민간기업과의 협업에 관한 토픽이다. 토픽 4는 서울시 관광 지역 등에서 다양한 이벤트와 홍보 등의 활동에 관련된 토픽이다. 토픽 5는 시민 참여와 지원을 위한 행정적 지원 활동에 관한 토픽이다. 토픽 6은 도시 안전과 관련된 토픽인데 박물관과 연관성이 있는 것으로 보아 I SEOUL U의 활용 측면과 관련된 토픽으로 추정된다. 토픽 7은 홍보의 도구 제작에 관한 주제로서 서울시 관광명소를 활용한 달력과 갤러리 등의 제작 관련 토픽이다. 토픽

8은 민간기업의 판매와 관련하여 유통채널로서 팝업 스토어 관련 제품과 판매장소 등에 관한 토픽이다. 토픽 9는 코로나와 관련하여 마스크에 새겨진 I SEOUL U 로고에 관한 토픽인 것으로 추정된다. 토픽 10은 조형물 홍보물 제작과 설치에 관한 주제로서 대학과의 연계된 활동인 것으로 추정된다.

2. 시사점 및 논의

본 연구에서는 서울시 도시 브랜드 I SEOUL U에 대한 인식과 평가를 위하여 온라인 빅데이터를 대상으로 한 텍스트마이닝 분석을 수행하였다. I SEOUL U와 관련된 키워드는 주로 비즈니스와 마케팅 관련 용어들이다. 따라서 서울시 브랜드로서 공익적 성격의 도시 브랜드는 I SEOUL U도 민간기업의 상업 활동에 많은 기여를 하는 것으로 추정된다. 매트릭스 분석에서는 제품 판매와 관련하여 주로 팝업 스토어의 아이템들이 많고 공동 브랜드 형태의 제품들이 개발되는 것으로 나타났다. 제품 테마는 주로 I SEOUL U를 상징하는 상징물들이 많은 것으로 보이며 I SEOUL U의 팝업 스토어 매장과 이벤트 관련하여 홍대 근처 지역이 매우 큰 비중을 차지하는 것으로 나타났다. 이 밖에 토픽 모델링에서 총 10개의 토픽이 추출되었고 여기서도 상업적 활용과 이벤트 축제를 위한 정보에 대한 니즈가 많은 것으로 나타났다.

매트릭스 분석 결과에서는 이벤트 관련 용어들이 많은 것으로 나타났는데 I SEOUL U 도시 축제 브랜드가 이벤트 중심의 거리 축제로서 많은 사람에게 각인된 결과로 보여진다. 이와 관련하여 한강, 달력, 식목일, 나무심기, Metropolitan, 홍대, 크리스마스, 세종대, 축제거리 등의 단어들이 연관된 것으로 나타나 이러한 단어들에 대한 추가적인 조사를 통하여 이벤트 중심의 거리 축제에 대한 다양한 기획안들이 마련되어야 할 것이다.

본 연구 결과를 살펴보면 도시 브랜드의 주 효과는 주로 경제적인 가치의 향상이라는 점을 알 수 있다. 단어 빈도수 분석에서 나타났듯이 I SEOUL U와 관련된 키워드는 주로 비즈니스와 마케팅 관련 용어들이다. N gram 분석 결과에서도 I SEOUL U 관련 제품의 개발과 판매에 관련된 사례가 많은 것으로 나타났다. 또한, I SEOUL U는 민간기업의 홍보 등의 상업적 활용 도구

로서 빈번하게 사용되는 것으로 나타났다. 따라서 지방 정부의 공익적 성격의 도시 브랜드가 성공할 경우 오히려 민간기업의 상업 활동에도 많은 도움을 준다는 사실이 밝혀졌다. 그러나 연결 중심성 분석 결과에 나타났듯이 비즈니스 및 마케팅 관련 단어들의 연결 중심성이 낮은 것으로 나타나 도시 브랜드의 상업적 활용은 관심도와 비교하면 활성화되지 못하는 것으로 나타났다. 따라서 도시 브랜드의 상업적 활용과 이로 인한 도시 브랜드의 경제적 가치 향상을 지원하기 위한 정책적 방안들이 절실하다고 할 수 있다. 시민과 기업들의 도시 브랜드에 관한 관심이 실제로 경제적 가치 향상으로 나타날 수 있도록 도시 브랜드의 경제적 활용과 그 효과에 대한 다양한 전략 도구들을 정책적으로 마련해야 할 것이다. 본 연구의 한계점으로는 자료 수집 기간이 코로나 발생 이후로서 코로나 이전의 자료는 반영하지 못하고 있는 점을 들 수 있다. 따라서 본 연구 결과는 최근의 I SEOUL U 관련 주요 단어들의 장기간의 트렌드를 모두 반영하고 있지 못하다. 이러한 점에서 향후 연구에서는 좀 더 장기적인 데이터 수집을 통하여 코로나 이전과 이후의 차이를 비교하는 것도 의미가 있을 것이다.

참고 문헌

- [1] M. Kavaratzis, "Place branding: A review of trends and conceptual models," *The marketing review*, Vol.5, No.4, pp.329-342, 2005.
- [2] The Seoul Research Data Service, *Festivals of Seoul in 2013*, <http://data.si.re.kr/node/73>
- [3] 박수근, 박태승, "소셜매트릭스 기법을 이용한 태권도복의 트렌드 분석," *한국체육과학회지*, 제28권, 제1호, pp.199-510, 2019.
- [4] 임광혁, "SNS 빅데이터 분석 기술 동향 및 발전방향," *한국콘텐츠학회지*, 제15권, 제1호, pp.38-43, 2017.
- [5] 정풀잎, 안현철, 곽기영, "텍스트 마이닝과 소셜 네트워크 분석을 이용한 스마트폰 디자인의 핵심속성 및 가치 식별," *대한경영학회지*, 제32권, 제1호, pp.27-47, 2019.
- [6] 김상락, 강만모, "빅데이터 분석 기술의 오늘과 미래," *정보과학회지*, 제32권, 제1호, pp.8-17, 2014.
- [7] N. Morgan, A. Pritchard, and R. Pride, *Destination*

- branding: creating the unique destination proposition. Destination branding: creating the unique destination proposition*, Oxford: Elsevier, 2002.
- [8] A. Insch, *Branding the city as an attractive place to live*, In *City Branding*, Palgrave Macmillan, London, pp.8-14, 2011.
- [9] K. M. Lee and C. D. Kim, *Branding Seoul Strategy*, Seoul: Seoul Development Institute, 2010.
- [10] A. C. Middleton, *City branding and inward investment*. In *City Branding*, UK: Palgrave Macmillan, pp.15-26, 2011.
- [11] S. Anholt, *What is competitive identity?*, In *Competitive identity*, Palgrave Macmillan, London, pp.1-23, 2007.
- [12] Y. S. Yoon and W. H. Kim, *City Branding*, Seoul: Unitas Brand, 2010.
- [13] J. P. Houghton and A. Stevens, City branding and stakeholder engagement, In *City branding*, Palgrave Macmillan, London, pp.45-53, 2011.
- [14] S. V. Gelder, *City Brand Partnership, City Branding, Theory and Cases*, Palgrave Macmillan, New York, pp.37-38, 2011.
- [15] J. Carlsen, T. D. Andersson, J. Ali-Knight, K. Jaeger, and R. Taylor, "Festival management innovation and failure," *International Journal of Event and Festival Management*, Vol.1, No.2, pp.121-131, 2010.
- [16] D. Getz, *Event studies: theory, research and policy for planned events*, London: Routledge, 2012.
- [17] G. Richards and R. Palmer, *Eventful cities*, Routledge, 2012.
- [18] S. S. Grumwell, I. S. Ha, and B. S. A Martin, "comparative analysis of attendee profiles at twourban festivals," In *Journal of Convention & Event Tourism*, Vol.9, No.1, pp.1-14, 2008.
- [19] D. Getz, *Event tourism: Definition, evolution, and research*, *Tourism Management*, pp.403-428, 2008.
- [20] B. Quinn, "Arts festivals and the city," *Urban studies*, Vol.42, No.5-6, pp.927-943, 2005.
- [21] https://www.seoul.go.kr/seoul/brand_old.do, 2021.09.28.
- [22] C. E. Han, HiSeoul Festival changed name as Seoul Street Arts Festival, 2016, <https://www.thefestival.co.kr/news/all/2634/?mcurpage=71>.
- [23] 백봉현, 하일규, 안병철, "SNS상의 비정형 빅데이터로부터 감성정보 추출 기법," *멀티미디어학회논문지*, 제17권, 제6호, pp.671-680, 2014.
- [24] 이애리, 이주원, "소셜 빅 데이터를 이용한 상권 확장 트렌드 및 소비 트렌드 분석: 인스타그램 해시태그 데이터 기반," *e-비즈니스연구*, 제19권, 제6호, pp.401-413, 2018.
- [25] 이종화, 이윤재, 이현규, "SNS의 해시태그를 이용한 감정 단어 수집 시스템 개발," *정보시스템연구*, 제27권, 제2호, pp.77-94, 2018.
- [26] D. Zhao, and M. B. Rosson, "How and why people Twitter: the role that micro-blogging plays in informal communication at work." In *Proceedings of the ACM 2009 international conference on Supporting group work*, pp.243-252. ACM, 2009.
- [27] 이정권, 최영, "소셜미디어 이용 동기 연구," *한국언론학보*, 제59권, 제1호, pp.115-148, 2015.
- [28] 신현희, 김경란, "SNS의 형태적 진화에 따른 이용동기 및 사용행태 차이-대학생의 페이스북과 인스타그램 이용을 중심으로," *디지털융복합연구*, 제15권, 제6호, pp.155-164, 2017.
- [29] 한국정보화진흥원, "빅데이터 시대! SNS의 진화와 공공정책," *IT & Future Strategy*, 제13권, 2012.
- [30] 태그라이브 전략기획팀, "마케터가 꼭 알아야 할 인스타그램 통계 2019," [taglive.net/forum trend/020.html](http://taglive.net/forum/trend/020.html), 2018.
- [31] 김정인, 정혜욱, "인스타그램 유저들의 예외적 해시태그 (hashtag) 이용 행태 연구," *디자인융복합 연구*, 제15권, 제2호, pp.136-152, 2016.
- [32] L. Yang, S. Tao, Z. Ming, and M.Qiaozhu, "We know what@ you# tag: does the dual role affect hashtag adoption?," In *Proceedings of the 21st international conference on World Wide Web*, pp.261-270, ACM, 2012.
- [33] S. Peng, G. Wang, and D. Xie, "Social influence analysis in social networking big data: opportunities and challenges," *IEEE Network*, Vol.31, No.1, pp.11-17, 2016.
- [34] J. C. Sweeney and G. N. Soutar, "Consumer

perceived value: The development of a multiple item scale,” Journal of retailing, Vol.77, No.2, pp.203-220, 2001.

[35] 김봉재, “텍스트마이닝을 활용한 학부모 기사 분석,” 학부모연구, 제4권, 제2호, pp.113-146, 2017.

[36] R. Feldman and I. Dagan, “Knowledge Discovery in Textual Databases (KDT),” In KDD, Vol.95, pp.112-117, 1995.

[37] 백승현, 김기탁, “코로나 19에 따른 프르야구 무관중 시청률요인의 중요도, 만족도 분석,” 한국엔터테인먼트 산업학회논문지, 제15권, 제2호, pp.123-135, 2021.

[38] 김태훈, *소셜 미디어 빅데이터 기반 프로축구 관력 소비자 인식 분석*, 동아대학교 대학원 박사학위논문, 2020.

[39] 최재석, *빅데이터 분석을 활용한 스포츠산업 서비스화 방안*에 관한 연구, 충남대학교 대학원, 박사학위논문, 2021.

[40] 이소희, *소셜 빅데이터를 이용한 뉴트로(New-tro) 관광 트렌드 및 이미지 인식 연구*, 경희대학교 대학원, 석사학위논문, 2020.

[41] 최기태, 김지연, “빅데이터를 활용한 캠핑 트렌드 분석-코로나19 이후 변화를 중심으로,” 호텔리조트연구, 제20권, 제1호, pp.279-294, 2021.

[42] 윤영일, 하동현, “소셜미디어 빅데이터 분석을 통한 비즈니스호텔의 인식 연구,” 경영교육연구, 제32권, 제3호, pp.379-398, 2017.

[43] 김미성, 서유빈, 김현정, 이혜련, “언어네트워크분석을 통한 지역특화컨벤션 육성 사업 비교-대구와 부산 사례를 중심으로,” 관광학연구, 제43권, 제7호, pp.183-200, 2019.

[44] 구영석, “빅데이터를 통한 소비자의 의복관리방식 트렌드 분석,” 한국의류산업학회, 제5호, pp.639-649, 2020.

[45] 박혜진, “뉴 노멀 시대의 빅 데이터 분석을 통한 화장품 트렌드에 대한 연구,” 한국화장품미용학회지, 제10권, 제3호, pp.465-480, 2020.

[46] 오정심, “빅데이터 토픽모델링 및 네트워크 분석을 통한 문화콘텐츠학 지식 구조 연구,” 문화정책논총, 제34권, pp.35-70, 2020.

[47] D. M. Blei, “Probabilistic topic models,” Communications of the ACM, Vol.55, No.4, pp.77-84, 2012.

[48] D. M. Blei, A. Y. Ng, and M. Jordan, “Latent dirichlet allocation,” the Journal of machine

Learning research, Vol.3, pp.993-1022, 2003.

[49] 김태호, 김학선, “빅데이터 분석을 활용한 체류형 울산관광 활성화 방안에 대한 제언,” 산업혁신연구, 제35권, 제1권, pp.47-76, 2019.

[50] 윤혜진, 김창식, 광기영, “지역관광 연구트렌드 분석: 텍스트마이닝 활용,” 한국관광레저학회 학술발표대회, pp.275-282, 2017.

저 자 소 개

임 혜 원(Haewen Lim)

정희원



- 2009년 2월 : Ritsumeikan APU 사회과학부(학사)
- 2011년 2월 : University of Surrey 관광개발(석사)
- 2017년 1월 : University of Westminster 건축환경(박사)
- 2018년 9월 ~ 현재 : 건국대학교

도시재생학부 겸임교수

<관심분야> : 도시브랜드, 도시재생, 이벤트, 관광