

# 구글맵리뷰 텍스트마이닝을 활용한 공원 이용자의 인식 및 평가

- 서울숲, 보라매공원, 올림픽공원을 대상으로 -

이주경\* · 손용훈\*\* · \*\*\*

\*서울대학교 환경대학원 협동과정 조경학 박사과정 ·

\*\*서울대학교 환경대학원 환경조경학과 부교수 · \*\*\*서울대학교 환경계획연구소 겸무연구원

## Perception and Appraisal of Urban Park Users Using Text Mining of Google Maps Review

- Cases of Seoul Forest, Boramae Park, Olympic Park -

Lee, Ju-Kyung\* · Son, Yong-Hoon\*\*

\*Interdisciplinary Program in Landscape Architecture, Seoul National University

\*\*Associate Professor, Graduate School of Environment Studies, Seoul National University

\*\*\*Environmental Planning Institute Adjunct Researcher, Seoul National University

### ABSTRACT

The study aims to grasp the perception and appraisal of urban park users through text analysis. This study used Google review data provided by Google Maps. Google Maps Review is an online review platform that provides information evaluating locations through social media and provides an understanding of locations from the perspective of general reviewers and regional guides who are registered as members of Google Maps. The study determined if the Google Maps Reviews were useful for extracting meaningful information about the user perceptions and appraisals for parks management plans. The study chose three urban parks in Seoul, South Korea; Seoul Forest, Boramae Park, and Olympic Park. Review data for each of these three parks were collected via web crawling using Python. Through text analysis, the keywords and network structure characteristics for each park were analyzed. The text was analyzed, as were park ratings, and the analysis compared the reviews of residents and foreign tourists. The common keywords found in the review comments for the three parks were "walking", "bicycle", "rest" and "picnic" for activities, "family", "child" and "dogs" for accompanying types, and "playground" and "walking trail" for park facilities. Looking at the characteristics of each park, Seoul Forest shows many outdoor activities based on nature, while the lack of parking spaces and congestion on weekends negatively impacted users. Boramae Park has the appearance of a city park, with various facilities providing numerous activities, but reviewers often cited the park's complexity and the negative aspects in terms of dog walking groups. At Olympic Park, large-scale complex facilities and cultural events were frequently mentioned, emphasizing its entertainment functions. Google Maps Review can function as useful data to identify parks' overall users' experiences and general feelings. Compared to data from other social media sites, Google Maps Review's data provides ratings and understanding factors, including user satisfaction and dissatisfaction.

*Key Word: Online Reviews, Google Maps, Social Network Service, Text Analysis, Usage Behavior*

---

**Corresponding author:** Yong-Hoon Son, Associate Professor, Dept. of Landscape Architecture, Graduate School of Environmental Studies, Seoul National University, Seoul 08826, Korea, Tel.: +82-2-880-8107, E-mail: [sonyh@snu.ac.kr](mailto:sonyh@snu.ac.kr)

## 국문초록

본 연구의 목적은 Google Maps에서 제공하는 장소에 대한 리뷰를 활용하여 실제로 공원을 방문한 이용자의 인식과 평가를 파악하는 것이다. 구글맵리뷰는 Social Network Service(SNS)를 통해 장소에 대한 인식과 평가에 관한 정보를 얻는 온라인 리뷰이며, 일반 리뷰어와 구글맵의 회원으로 등록된 지역 가이드의 관점에서 장소에 대한 이해를 볼 수 있는 서비스이다. 본 연구에서는 구글맵리뷰 분석이 공원 관리에 필요한 이용자들의 인식과 평가를 추출하는데 활용될 수 있는지를 살펴보고자 하였다. 서로 다른 공간특징과 시설을 가지는 3개의 공원(서울숲, 보라매공원, 올림픽공원)을 대상으로 파이썬을 활용한 웹 크롤링을 통해서 구글맵리뷰 내용을 수집하였다. 그리고 텍스트 분석을 통해 공원별 주요 키워드 분석과 네트워크 구조에 따른 특성을 분석하고, 이와 함께 구글맵리뷰에서 제공하는 별점 평점값과 외국인 리뷰 데이터에 대한 분석도 수행했다. 연구 결과, 3개의 공원에서 공통으로 나타나는 특성으로는 이용목적으로 ‘산책’, ‘자전거’, ‘휴식’, ‘피크닉’이 있었으며, 동반유형으로 ‘가족’, ‘아이’, ‘애견’이, 인프라로는 ‘놀이터’, ‘산책로’가 있었다. 공원별 특색을 보면 서울숲은 자연을 기반으로 하는 야외활동이 많이 나타났고 반면, 주차공간 부족과 주말 혼잡은 공원 이용자에게 부정적인 영향을 미치고 있었다. 보라매공원은 수많은 활동을 제공하는 다양한 시설을 갖춘 도시공원의 모습을 가지고 있었다. 리뷰어들은 반려견을 동반하는 이용자 그룹과 그렇지 않은 다른 이용자 그룹 간의 갈등과 공원의 복잡함에 대한 부정적인 측면을 언급했다. 올림픽공원에는 대형 복합시설이 있으며, 커뮤니티, 문화예술공연과 같은 대규모 문화 이벤트가 많이 언급되었고, 레크리에이션 기능이 강조되었다. 구글맵리뷰는 공원에 대한 이용자의 전반적 경험과 이미지에 대한 특징을 파악하는 유용한 자료라고 할 수 있다. 또한, 다른 소셜미디어 데이터와 비교할 때 특히 구글맵리뷰는 공원에 대한 이용자 평점값과 만족 및 불만족 요인을 이해할 수 있는 데이터를 제공한다.

주제어: 온라인 리뷰, 구글맵, 소셜 네트워크 서비스, 텍스트 분석, 이용행태

## 1. 서론

현대사회에서 공원은 도시환경 개선, 사회문화 다양성의 포용, 경제성장, 역사성 보존의 기능을 가지며(CABE, 2010), 도시를 더 살기 좋고 지속 가능한 도시로 만들기 위한 혜택을 제공하고 있다. 단, 공원이 주는 이러한 혜택은 공원 및 녹지공간이 잘 관리되고 유지될 때 기대할 수 있는 기능이다(Nam and Dempsey, 2019). 따라서 공원의 양적 확장뿐만 아니라, 공원의 질적 관리 측면에서 지속가능한 공원의 운영 및 관리에 관한 논의가 증가하고 있다. 특히 변화하는 공원 이용자들의 수요를 반영한 공원 관리 및 경영이 요구되고 있으며, 공원의 운영관리에 있어 실제 이용자의 경험과 요구를 파악하는 것은 주요한 관점 중 하나이다(Warnock and Griffiths, 2015). 이를 위해 장기간의 다양한 공원 이용자의 관점을 포함한 데이터 수집이 필요하며, 지속적인 조사와 분석이 필요하다.

공원 운영관리에 이용자의 관점을 반영하기 위해서는 실제로 공원을 방문한 경험을 바탕으로 한 이용자의 평가와 인식을 파악하는 것이 필요하다. 현재 세계는 온라인 Social Network Service(SNS)를 통해 연결되고 있으며, 이곳에서 경험과 정보를 공유하고, 결과적으로 서로의 인식에 영향을 미친다(Lund *et al.*, 2018). 소셜미디어 콘텐츠를 분석하면 도시 녹지 공간이 제공하는 사회적 이익을 평가하고, 공원 이용자의 행태와 선호도를 이해할 수 있으며(Richards and Friess, 2015), 이러한 정

보는 공원의 관리로 연결할 수 있다(Guerrero *et al.*, 2016). SNS 데이터는 객관성, 신뢰성, 전문성 등 본질적인 문제에 대한 과제가 있지만(Chang and Nam, 2012), 특별한 이해관계가 없는 방대한 데이터가 자발적으로 생성되고, 시간 및 장소의 제약에 자유롭기 때문에 데이터 취득 및 가공에 있어서 효용 가치가 높다.

소셜 네트워크를 통해 장소에 대한 인식과 평가에 관한 정보를 얻는 방법 중 하나는 온라인 리뷰이다. 온라인 리뷰는 필요한 소비자 지향에 초점을 맞추며, 최종 이용자의 관점에서 장소에 대한 이해를 볼 수 있고(Kavaratzis *et al.*, 2005), 리뷰에 담긴 가치를 활용하면 품질 향상과 마케팅의 중요한 정보로 활용될 수 있다(Litvin *et al.*, 2008). 또한, 온라인 리뷰는 정보를 제공하고 취득하기 위한 활동으로 다른 SNS에 비해서 비교적 목적이 뚜렷한 서비스라고 할 수 있다.

본 연구의 목적은 Google Maps에서 제공하는 장소에 대한 리뷰를 활용하여 공원 이용자의 인식과 평가를 파악하는 것이다. 구글맵리뷰는 Google 지도에서 방문한 장소에 대한 정보를 공유하는 공간으로 경험한 내용을 리뷰하거나, 사진 및 동영상 등을 공유할 수 있으며, 별점을 통해 장소를 평가한다. 또한 지역 가이드 서비스를 통해 활발하고 지속적인 참여와 비교적 성의 있는 리뷰작성을 유도한다. 지역 가이드는 글로벌 커뮤니티 서비스로 지역 가이드를 등록하면 리뷰를 포함한 Google 지도 콘텐츠 참여에 따른 레벨 등급이 주어지고, 포인트가 적립된다.

등급이 높아지면 지역 가이드 배지, Google 신제품 사전 체험관 이용 등의 보상이 주어지기도 한다. 즉, 생활형 놀이와 같은 재미있는 콘텐츠들로 형성되어 있어, 지역 가이드는 비록 금전적 보상은 적지만 자신이 작성한 리뷰에 대해 반응이 증가할수록 재미를 느끼고, 성취감을 느끼기도 한다. 또한, 리뷰에 담긴 정보의 사실관계를 위해 여러 사람이 검증하게 할 수 있는 시스템이 갖추어져 있어 비교적 정확한 정보를 얻을 수 있다.

구글맵리뷰는 유럽 등 해외 공원에서는 활성화되어 사용되고 있지만, 아직 국내 이용률은 낮다. 데이터 양이 많지 않기 때문에 특정 계층, 연령, 국적, 지역 등으로 편중될 가능성이 있어서 국내 연구에서는 구글맵리뷰 결과만으로 이용자들의 인식을 전부 파악하기에는 한계가 있다.

하지만, 구글맵리뷰는 소셜미디어 데이터로서 가지는 장점도 분명히 존재한다. Matsuoka(2017)에 의하면 구글맵 이용자들은 지역, 지구, 국가와 같은 다양한 스케일의 공간에 대해 참여형 지도를 만들면서 서로 공유하여 글로벌화에 적합한 플랫폼이라는 평을 받고 있다. 구글맵은 많은 나라에서 이용하는 서비스로 네이버 블로그와 같이 국내 장소만 분석 가능한 것이 아니라, 해외 장소 연구에도 활용될 수 있으며, 국가 간 비교도 할 수 있는 플랫폼이다. 데이터로서 구글맵리뷰의 또 다른 장점은 다른 소셜미디어와 다르게 별점을 사용하여 방문한 공원에 대한 직접적인 평가 점수가 제공된다는 점이다. 이를 통해서 감성 분석 등 문서의 맥락에 의해서 작성자의 인식 및 만족도를 평가하는 것이 아니라, 이용자가 제공한 평가 값을 활용하여 이용만족도를 분석할 수 있다. 또한, 구글맵리뷰는 지도에 있는 장소를 특정하여 리뷰 글을 작성하므로, 장소 검색어를 통한 데이터 크롤링보다 더 명확하게 연구 대상지에 대한 데이터를 수집할 수 있다. 많은 SNS가 상업성으로 활용되면서 광고성 글이 많이 게시되며, 연구 대상과 관련이 없거나 적은 글이 게시되기도 하지만, 구글맵리뷰는 지역 가이드의 활동 및 시스템적으로 광고 및 허위정보가 적어 데이터를 신뢰할 수 있고, 분석과정에서 데이터 전처리 작업이 많이 필요하지 않다는 것도 중요한 장점이라고 할 수 있다.

이러한 이유로 본 연구에서는 구글맵리뷰를 통해 공원 이용자들이 경험하는 공원 서비스에 대한 수요를 리뷰 텍스트, 별점, 이용자층에 대해 다각적으로 파악하고, 공원 관리에 참고할 수 있는 이슈를 고찰하였다. 본 연구를 통해서 구글맵리뷰와 같은 장소 기반 온라인 리뷰를 이용하여, 도시공원의 운영관리에서 적용 가능한 이용자 평가 자료를 얻을 수 있으리라 기대한다.

## II. 이론적 고찰

### 1. 온라인 리뷰데이터 연구 동향

생활 속 많은 소비공간에서 온라인 리뷰는 필수적인 요소로

자리 잡음에 따라 다양한 분야에서 온라인 리뷰데이터에 담긴 사용자의 인식을 파악하는 연구가 이루어지고 있다. 대부분의 온라인 리뷰를 분석하는 연구는 제품, 영화, 호텔과 같은 상품 서비스를 대상으로 하고 있으며, 소비자의 성향 및 이용 후 만족과 지속적인 이용 의도를 파악하고자 하는 연구가 많이 수행되고 있다.

Bae(2016)는 소셜커머스에 나타난 온라인 리뷰에서 후기수용 신념변수(post-adoption beliefs)를 이용하여 소비 가치 요인을 도출하고, 이 요인이 이용 후 만족과 지속적인 이용 의도에 어떠한 영향을 미치는지를 분석했다. Lee(2013)는 온라인상에 남겨진 호텔 이용 후기를 자료 범주화, 어휘단어 분리 등을 통해 질적으로 분석하여, 호텔이용자들이 지각한 고객가치의 차원을 분석했다. 또한, 리뷰 데이터는 대부분 비정형 또는 반정형의 텍스트로(Buneman, 1997; Byun *et al.*, 2016), 이를 정량화하여 분석하는 텍스트 마이닝(Text Mining)을 활용한 연구들이 지속적으로 이루어지고 있다. Cho *et al.*(2014)은 영화 관련 온라인 리뷰 텍스트 마이닝(Text Mining)을 진행하고, 판별분석을 통해 영화별 최초 흥행성과 예측모형을 개발했다. Byun *et al.*(2016)은 사용자 리뷰를 평가 기준으로 자동 분류하고, 각 평가 기준별 주요 이슈를 요약하여 제공할 수 있는 방법론을 제시하여 온라인 호텔리뷰에 적용해 상세 이슈를 일목요연하게 파악하도록 하였다.

그 외에도 온라인 리뷰를 통해 서비스 이용 경험을 효과적으로 분석할 수 있는 방법론을 제시하기 위해 Latent Dirichlet allocation(LDA) 토픽모델링을 통한 빅데이터 연구(Jin *et al.*, 2013; Xianghua *et al.*, 2013; Chae *et al.*, 2015; Park, 2015), 리뷰의 유사도를 산출하기 위해 색인어의 빈도 산출 및 Term Frequency - Inverse Document Frequency (TF-IDF) 가중치를 분석한 연구(Jeon and Ahn, 2015), 온라인 댓글들을 분류하고, 이를 시각화하는 시스템을 제안한 연구(Lee *et al.*, 2009) 등이 있다. 또한, 온라인리뷰 텍스트에 관한 감성분석(sentiment analysis) 연구의 사례도 있다(Choi *et al.*, 2016; Lee *et al.*, 2016; Lee *et al.*, 2017).

구글맵리뷰 데이터를 활용한 선행연구로는 Kim *et al.*(2019)이 있다. 본 연구는 공원을 대상으로 인스타그램과 구글맵리뷰에 나타나는 텍스트를 이용행태 범주로 분류하는 분석이 이루어졌고, 구글 방문 데이터를 활용하여 공원별 이용 시간대 특성을 파악했다.

해외연구의 사례로 Munawir *et al.*(2019)은 구글맵리뷰에 대해 텍스트 마이닝을 수행하여 공원 방문객들의 인식과 평가를 파악했다. 구글맵리뷰에서 제공되는 별점을 공원별로 파악하고, 리뷰 내용에 대해 TF-IDF 방법을 활용하여 공원별 주요 키워드와 키워드 간의 네트워크 관계를 표현했다.

## 2. SNS를 활용한 공원의 이용행태 및 특성 연구

최근에는 SNS를 활용한 공원 이용행태 및 특성 연구가 지속적으로 수행되고 있다. 조경, 건축 및 도시 분야에서 초기의 연구 중 Bark and Kim(2011)은 블로그에 나타나는 태그 키워드를 활용하여 선유도공원의 경관의 특성 및 향유방식을 해석했다. Lee(2013)는 1인 미디어에 나타난 장소감과 미시담론으로서 갖는 의미를 밝히기 위해 블로그와 트위터를 대상으로 질적 연구를 수행하였다. 소셜미디어에 나타난 도시공원의 장소감 연구의 필요성을 제안한 초기의 연구로서 빅데이터 분석 및 해석의 한계점이 있었다.

데이터로서 SNS는 특성상 방대한 양의 빅데이터이며, 비정형의 데이터로서의 특징을 가진다. 따라서 최근에는 텍스트 마이닝을 통해 텍스트 내용을 정량적으로 분석하고 의미를 추출하는 연구들이 많이 수행되고 있다. Woo and Suh(2018)는 블로그에 나타난 올림픽공원을 대상으로 텍스트 마이닝을 활용하여 공원의 시계열적 특성과 사회연결망 분석을 통해 키워드 간의 연결망 구조를 파악했다. 텍스트 마이닝은 과거에서 발생한 대량의 텍스트 데이터 분석이 가능하고, 장기적인 관점에서 분석이 가능하다는 장점을 언급하며, SNS가 행태분석의 연구영역을 확대시키며, 새로운 가치가 추출될 것이라고 논의했다. 또한, 테마공원 36곳의 트위터 텍스트를 이용하여 감성 분석을 수행하고, 장소 선호도를 도출하는 방법을 제안한 연구(Chae, 2017), 인스타그램과 구글 콘텐츠를 활용하여 공원 이용행태를 파악한 연구(Kim *et al.*, 2019), 경의선숲길 블로그 글을 대상으로 텍스트 마이닝 기법을 사용하여 공원 조성 전후에 대한 방문자의 경험 변화를 분석한 연구(Kim *et al.*, 2019) 등이 수행되었다.

SNS를 활용한 국내 연구는 주로 텍스트를 데이터로 활용하여 텍스트 마이닝이라는 정량적 방법을 통해 공원의 인식 및 이용 특성을 객관적으로 파악하고자 했다. 텍스트 마이닝 기법으로 빈도분석(frequency analysis), 중심성분석(centrality), 의미네트워크분석(semantic network analysis)이 주로 나타났다.

공간에 대한 이용행태연구는 공원의 이용행태 및 특성을 분석함에 유리하다는 연구 결과들이 도출되었고, 실증적인 결과 도출을 위한 기술적인 방법 및 입체적인 해석이 보완되어야 한다고 했다.

국내에서 블로그, 인스타그램, 트위터를 활용하여 텍스트마이닝 기법을 통해 공원의 다양한 특성을 파악하는 연구는 지속적으로 수행되고 있다. 반면에, 구글맵리뷰를 대상으로 텍스트마이닝을 수행한 국내연구는 아직 찾아보기 힘들므로, 본 연구는 소셜미디어 중 비교적 새로운 데이터로서 구글맵리뷰의 활용성을 탐색하는 데 의의가 있다고 하겠다.

## III. 연구방법

### 1. 연구대상지

본 연구는 규모, 운영 주체, 공간 특성, 방문객 특성을 고려하여, 가능한 서로 다른 성격의 도시공원들을 연구 대상지로 선정했으며, 또한 데이터 분석을 고려해서 최소 1,000개 이상의 리뷰 데이터가 있는 공원을 선택했다. 최종적으로 서로 다른 운영 주체와 공원 환경을 가지고, 지역민 및 외지인의 방문 행태가 다르게 나타난 서울숲, 보라매공원, 올림픽공원을 연구 대상으로 선정하였다(Table 1 참조).

서울숲은 '뚝섬 숲 조성 기본계획'에 따라 기존의 뚝섬체육공원 일대를 대규모 도시숲으로 만들기 위해 조성되었으며, 문화예술공원, 생태숲, 습지생태원, 자연체험학습원, 한강수변공원으로 특성 있는 5개 지구로 나뉜다. 다른 지역에 비해 공원이 부족했던 서울 동북부 지역의 시민들에게 자연 휴식공간을 제공하며, 다양한 자연체험 기회와 다채로운 시설을 제공한다. 지역 주민 외에도 외지인의 방문이 높은 공원이다. 또한, 2016년부터 (재)서울그린트러스트에 민간위탁되어 공원을 운영하고 있다.

보라매 공원은 서울 서남권의 대표적인 서울시 직영 공원으로서 서울시 동부 공원녹지사업소가 위치하고 있다. 과거 공군사

Table 1. Characteristics of research site

Park name	Scale	Operating agent	Spatial characteristics	Visitor characteristics
Seoul Forest	1,156,498m <sup>2</sup>	Contracting-out (Seoul green trust)	- 5 Districts (culture & art park, eco-forest, wetland ecology park, educational experience park, Hangang riverside park). - Provides a variety of nature experience programs.	Residents and outside visitors
Boramae Park	424,106m <sup>2</sup>	Eastern park and greenery office	- A park for citizens' rest and health promotion. - Provide many facilities and places of activity.	Residents usage rate ▲
Olympic Park	1,447,122m <sup>2</sup>	National sports promotion agency	- Large park / recreational park. - 6 stadiums and multi-purpose cultural facilities and convenience facilities. - Provides large-scale performance and cultural events.	Outside visitors usage rate ▲

관학교가 있던 부지로 에어파크와 관련된 전투기 모형들이 전시되어 있다. 주변 일대 영·유아부터 노인까지 거의 모든 연령대가 이용하고 있으며, 체육시설, 체험시설, 교양시설 등 많은 시설이 제공되어 다양한 유형의 이용자가 이용하고 있다.

올림픽공원은 1988년 서울올림픽 유치 기념으로 조성되어 현재 국민체육진흥공단이 관리하고 있다. 면적 1,447,122m<sup>2</sup>의 대규모 공원으로 체육, 역사, 교육, 휴식 등 다양한 기능을 갖는 종합공원이다. 6개의 경기장과 다목적 복합문화시설, 편의시설들이 있으며, 공연문화행사 및 다양한 프로그램을 제공하는 위락형 성격이 강한 공원이다. 이용률은 근린 주민 외에도 서울 전역에서 높다. 또한, 공원 내에 사적 제297호인 몽촌 토성이 있으며, 우수한 자연경관을 바탕으로 한 사진 촬영장소로 유명하다.

## 2. 연구방법

### 1) 데이터 수집 및 정제

본 연구에서는 Beautiful Soup 4, Selenium 파이썬 패키지를 사용하여 Python web crawler Code를 제작하여 직접 Google 지도에서 구글맵리뷰 데이터를 수집하였다. 데이터는 2020년 4~5월 기준 전체 데이터 수집을 목표로 했으며, 오류 데이터 및 중복 데이터 삭제 후 최종적으로 공원별로 수집한 데이터는 Table 2와 같다. 데이터 정보는 연도별 편차가 존재했다. 2020년을 제외하고, 최신 연도일수록 수집 건수가 많았다. 가장 오래된 데이터로는 2011년도부터 파악할 수 있었지만, 2016년 이전은 데이터 수가 현저히 적었다(Table 2 참조). Google 지도에서 위치 기준 공원명을 검색한 후에 사용자 이름, 공원에 대한 평가(별점), 날짜, 의견을 포함한 모든 리뷰 데이터를 수집했으며, 구글맵리뷰 데이터의 형태는 Figure 1과 같다.

장소에 대한 리뷰 글은 자유로운 형식과 일상 어투로 대부분 작성되어 있다. 이에 의미 있는 결과 도출을 위해 리뷰 데이터 텍스트 분석에 앞서 지정어(defined word), 유의어(thesaurus), 제외어(exception) 사전을 만드는 작업을 수행했다. 먼저 사전 등록 없이 데이터를 불러와 단어와 Document의 구조를 파악한 후에, 공원별로 빈도가 높거나 의미가 있는 이용행태, 시설명, 프로그램명 등을 중심으로 지정어를 탐색하여 추출했다. 또한, 띄어쓰기와 축약어를 고려하여 주요 용어를 대상으로 유의어

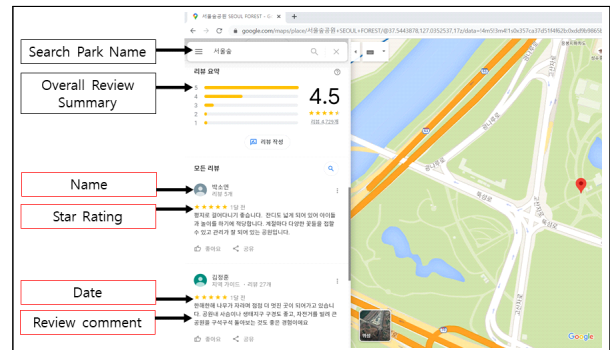


Figure 1. Structure of Google maps review

사전을 만들고, 불필요한 기호나 분석에 방해되는 단어를 고려한 제외어 사전을 만들었다. 텍스트 분석은 사회연결망 분석 소프트웨어인 넷마이너(Netminer) 4.3.3을 활용하였다.

### 2) 텍스트 마이닝

텍스트 마이닝(text mining)은 데이터 마이닝(data mining)의 한 분야로, 기계적 알고리즘을 활용하여 비정형의 텍스트 데이터를 모델링하고 구조화함으로써 유의미한 데이터를 추출하는 방법이다(Feldman and Sanger, 2007; Hearst, 2003; Daniel, 2015). 또한, 텍스트 마이닝은 이전에 알려지지 않는 유용한 정보를 발견하는 것으로 탐색적 데이터 분석(exploratory data analysis) 프로세스라고 할 수 있다(Tukey, 1977; Hoaglin *et al.*, 2000; Hearst, 1999). 즉, 대용량의 자료로부터 입력 텍스트를 구조화하고, 구조화된 데이터 안에서 다양한 기계적 알고리즘을 통해 관계, 패턴 등을 유도하여, 데이터를 축소(reduction)해 가면서 의미를 파악하는 과정이라고 할 수 있다.

본 연구의 텍스트 마이닝 분석 기법은 TF-IDF를 이용한 단어의 빈도분석(frequency analysis), Degree Centrality를 활용한 단어의 중심성 분석(centrality analysis), Community(betweenness)를 활용한 단어 간 응집성 분석(cohesion analysis)을 수행하였다. 연구의 흐름은 Figure 2와 같다.

Term Frequency - Inverse Document Frequency (TF-IDF)는 Term Frequency(TF)와 Inverse Document Frequency(IDF)를 곱한 값으로 특정 문서 내에서 키워드의 관련성 및 중요성을 나타내는 통계적 수치를 말하며 즉, 모든 데이터에서 나오는 단

Table 2. Data collection status

Park name	Collection date	Number of reviews collected (Number of reviews)	2020	2019	2018	2017	Below 2016
Soul Forest	2020. 04. 28	3,403(6,430)	289	2,515	404	128	67
Boramae Park	2020. 05. 08	2,470(4,621)	549	1,625	229	58	8
Olympic Park	2020. 05. 13	1,725(1,832)	167	1,053	282	148	71

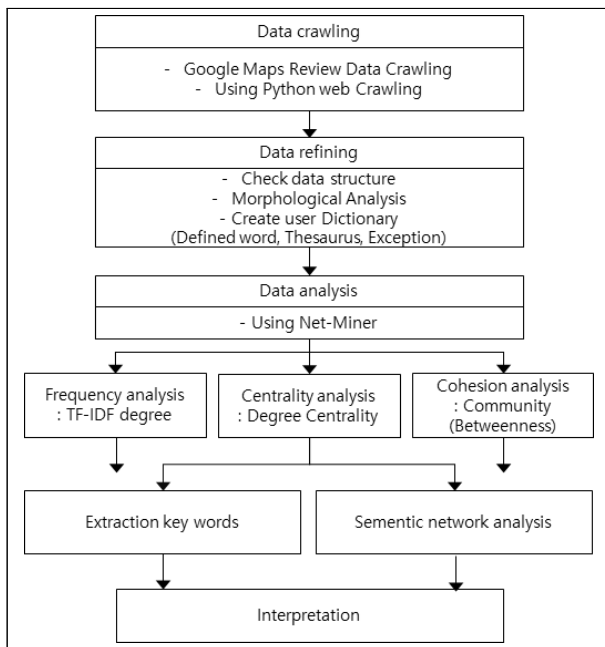


Figure 2. Study process

어보다 특정 문서 내에서 나오는 단어가 중요하다고 하는 개념으로, 상투어를 제외할 수 있다. 본 연구에서는 기존의 텍스트 네트워크 분석 연구에서 사용되는 기준값을 참고하여 TF-IDF 값이 0.5 이상의 단어만을 추출해(Ko, 2019; Son and Kim, 2019) 좀 더 유의미한 단어들로 축소해서 분석을 진행했다.

중심성 분석(centrality analysis)은 네트워크 안에서 키워드의 연결 정도를 측정하면서 키워드가 어떤 역할을 하는지를 분석하는 방법이다. 중심성의 종류 중 연결중심성(degree centrality)을 활용하였으며, 연결 중심성(degree centrality)은 직접 연결된 이웃의 수로 측정하여 이웃 노드가 많을수록 연결중심성이 높다고 할 수 있고, 직접적인 영향력의 크기를 측정하는데 용이하다.

응집성 분석은 키워들 간 맥락에 따라 그룹을 형성하고, 하위 그룹 간의 관계를 통해 네트워크의 특징을 파악할 수 있는 분석이다. 어떤 노드들이 그룹을 형성하고, 그룹에 속한 노들은 얼마나 응집되어 있는지, 그룹의 특성은 무엇인지 파악할 수 있다. 또한 네트워크 연결 관계를 시각적으로 도출하기에 용이하다. 본 연구에서는 응집성 종류 중 Community(betweenness)를 사용하여 알고리즘에서 도출된 적절한 Cluster 개수를 선택하여 Map을 작성했다. 의미 네트워크 분석의 결과는 연결 중심성 값이 상위 10개(0.055~0.038) 단어를 중심으로 해석하였다.

이와 같은 분석개념을 활용하여 공원별 중심 키워드 분석은 연결 중심성(degree centrality) 값을 기준으로 정리하였고, 중심성 값이 같으면 TF-IDF가 높은 것을 우선순위로 하였다. 또한, 공원별로 연결 중심성 값이 상위 10개인 단어가 어떤 하위 단어들과 연결되어 그룹을 만드는지를 중심으로 분석했다. 분

석 결과를 바탕으로 리뷰의 원문을 재참고하여 최종적으로 공원의 이슈들을 해석했다.

## IV. 결과 및 고찰

### 1. 공원별 속성 데이터 현황

구글맵리뷰는 리뷰어가 1~5등급의 별점으로 공원을 평가한 내용을 함께 제공한다. 공원은 기본적으로 도시의 다른 공간과 비교하면 긍정적 답변이 많고, 따라서 대부분의 공원은 평균적으로 4점 이상의 점수를 받는다. Table 3을 보면 세 공원 모두 별점 5가 50%를 넘고 있으며, 별점 4~5의 비율이 서울숲 88%, 보라매공원 74%, 올림픽공원 93%를 차지하고 있다. 올림픽공원은 면적이 크고, 경관이 아름다우며, 다양한 프로그램이 운영되고 있어서 제일 높은 평가를 받고 있었다.

실제 별점 3 이하의 리뷰를 읽어보면, 부정적인 의견이 있었다. 공원은 기본적으로 긍정적인 평가를 받는 곳이지만 일부 부정적인 감정이 있을 때, 리뷰어들은 별점 3 혹은 그 이하로 평가하는 경향이 있음을 알 수 있었다. 따라서 본 연구에서는 특별히 별점 3개 이하의 리뷰 글만을 모아서 각 공원에 어떠한 부정적인 의견이 있는지 분석하였다.

또한, 구글맵리뷰의 특성 중 하나는 전 세계의 사람들이 자유롭게 리뷰를 작성할 수 있고, 원문과 번역된 내용이 함께 제공된다는 점이다. 연구에서는 번역되어서 제공되는 리뷰를 외국인이 작성한 글이라고 판단하고 추출 및 분석하였다. Figure 3은 외국인 리뷰 데이터의 비율이며, 외국인 평가 비율은 서울

Table 3. Number of park reviews by star point

Park name	5 stars	4 stars	3 stars	2 stars	1 star
Seoul Forest	1,926 (56.6%)	1,069 (31.4%)	289 (8.5%)	54 (1.6%)	65 (1.9%)
Boramae Park	1,330 (53.9%)	500 (20.2%)	429 (17.37%)	75 (3.0%)	136 (5.51%)
Olympic Park	1,154 (66.9%)	449 (26.0%)	78 (4.5%)	18 (0.9%)	26 (1.5%)

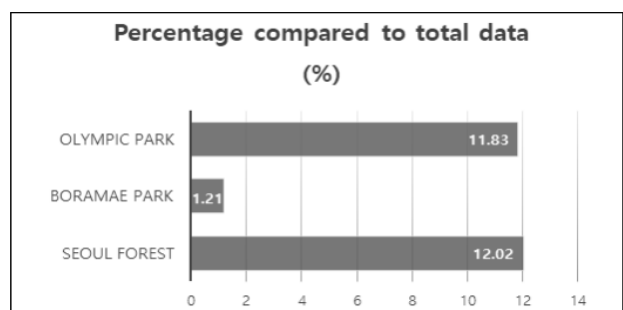


Figure 3. Foreign data's compared to total data

숲 12.02%(409개), 올림픽공원 11.83%(204개)로 전체 데이터의 약 10% 비율을 차지하였다. 반면, 보라매공원은 지역 주민의 이용이 많으며, 외국인 평가 비율도 1.21%로 매우 낮게 나타났다.

## 2. 공원별 중심성 분석 및 중요 키워드 분석

### 1) 전체 리뷰 데이터에 대한 중요 키워드

서울숲은 Table 4를 보면 '산책', '숲', '가족', '아이', '사슴', '자전거', '놀이터', '다양', '나무', '주차' 등의 키워드가 높은 중심성 값을 가진다. 전체 리뷰 데이터의 주요 키워드를 살펴보면 이용목적으로 '산책', '자전거', '피크닉', '사진'이 있었고, 동반유형으로 '가족', '아이', 인프라는 '사슴', '나비정원'이 등장

했다. '사슴', '나비정원'은 서울숲의 대표 프로그램으로 공원 이용자들이 인식하고 있다. 전체적으로 서울숲의 주요 키워드로 '숲', '사슴', '나무', '꽃', '동물', '나비정원', '자연'과 같은 자연 및 자연체험과 관련된 키워드들이 많이 추출되었다. 서울숲은 방문객들에게 녹지와 자연을 기반으로 하는 체험활동을 주로 제공하며, 이에 대한 이용자 수요도가 높은 것을 알 수 있다. 또한, 이용 형태에 관해서 '주차' 키워드도 높은 값을 보였다.

Table 5는 보라매공원에서의 주요 키워드를 나타낸다. '산책', '운동', '공사', '아이', '음악분수', '가족', '놀이터', '시설', '전투기', '애견'이 높은 중심성 값을 보였다. 이용목적과 관련된 키워드로 '산책', '운동', '조깅', '휴식', '자전거'가 있었으며, 동반유형은 '아이', '가족', '애견', 인프라는 '음악분수', '놀이터', '전투기', '운동장', '호수', '암벽등반'이 나타났다. 보라매공원은

Table 4. Key words in Seoul forest (Top 20)

Word	Cen.	Word	Cen.
Stroll (산책)	0.1715	Picnic (피크닉)	0.0474
Forest (숲)	0.1387	Flower (꽃)	0.0401
Family (가족)	0.1022	Shade (그늘)	0.0401
Child (아이)	0.1022	Air (공기)	0.0365
Deer (사슴)	0.0985	Animal (동물)	0.0365
Bicycle (자전거)	0.0985	Butterfly garden (나비정원)	0.0365
Playground (놀이터)	0.0766	Nature (자연)	0.0328
Diversity (다양)	0.073	Weekend (주말)	0.0328
Tree (나무)	0.0657	Photo (사진)	0.0328
Parking (주차)	0.0547	Cafe (카페)	0.0328

\*Cen.: Degree centrality.

Table 5. Key words in Boramae park (Top 20)

Word	Cen.	Word	Cen.
Stroll (산책)	0.1946	Track (트랙)	0.0649
Exercise (운동)	0.1892	Clean (깨끗)	0.0541
Construction (공사)	0.1297	Rest (휴식)	0.0486
Child (아이)	0.1189	Tree (나무)	0.0486
Musical fountain (음악분수)	0.1027	Playing field (운동장)	0.0486
Family (가족)	0.0919	Lake (호수)	0.0486
Playground (놀이터)	0.0757	Rock climbing (암벽등반)	0.0432
Facility (시설)	0.0703	Diversity (다양)	0.0378
Aircraft (전투기)	0.0703	Jogging (조깅)	0.0378
Dog (애견)	0.0649	Bicycle (자전거)	0.0378

\*Cen.: Degree centrality.

이용자와 관련된 다양한 시설 키워드가 많이 추출되었다. 다양한 시설 이용이 보라매공원의 이용 특성으로 판단된다. 중요 20개 키워드에는 포함되지 않았지만, 인라인스케이트장, X게임장, 안전체험관, 애견놀이터 등의 시설 관련 단어가 많이 나타났다. 또한, 보라매공원 인근 지역주민의 이용률이 높은 공원으로 산책뿐만 아니라, 일상적 운동 및 조깅과 같은 활동이 많이 이루어지고 있었다. '공사'가 높은 값을 보이는 이유는 2017년부터 착수된 신림선 경전철 공사로 인한 이슈 때문으로 판단된다.

Table 6은 올림픽공원의 리뷰 키워드를 나타낸다. '산책', '운동', '문화', '다양', '올림픽', '시설', '스포츠' 등 키워드가 높은 중심성 값을 보였다. 전체 리뷰 데이터의 주요 키워드를 살펴보면 이용목적으로 '산책', '운동', '행사', '자전거', '공연', '사진'이 나타났으며, 동반유형은 '아이', '가족', '연인', 인프라로는 '올림픽', '박물관', '경기장', '나홀로나무', '몽촌토성'이 있었다. '스포츠' 키워드는 행위와 시설에 대한 내용이 함께 포함되어 이용목적이자 시설 두 가지 특징을 모두 나타낸다. 올림픽공원은 크게 운동 관련 키워드인 '운동', '시설', '스포츠', '경기장'과 문화 관련 키워드인 '문화', '박물관', '행사', '공연'을 중심으로 나타나고 있는 것을 알 수 있다. 운동과 관련된 이용행태 및 시설과 문화 향유를 위한 공연, 행사의 내용이 구분되어 작성된 리뷰들도 있지만, 함께 등장하는 리뷰들이 많았다. 올림픽공원은 규모가 큰 공간들이 있고, 대형행사가 많이 이루어져 서울숲과 보라매공원과 다른 특징적인 이용 키워드들이 많이 추출되었다. 또한, 백제 시대 토성유산인 몽촌토성이 보존되고 있어 '몽촌토성'과 관련한 박물관, 미술관 등의 내용이 있었고, 또한 올림픽공원의 랜드마크인 '나홀로나무'는 유명한 포토 스팟으로 이와 관련된 내용도 많았다. 공원의 조성 목적에 맞게 '올림픽' 관련 시설과 기념적인 내용도 있었다.

## 2) 부정적 의견을 담은 별점 3 이하 리뷰 분석

별점 3 이하의 리뷰들을 살펴보면 대체로 공원 이용에 대한 부정적인 의견들이었다. 공원별로 이용자들이 인식하는 문제점을 파악하기 위해 별점 3 이하 리뷰만을 따로 추출하여 분석하였다.

Table 7에서 서울숲의 별점 3 이하 리뷰 데이터를 보면 '주차'와 '부족'이 높은 값을 나타내고 있다. 대체로 별점 3 이하의 데이터는 부정적인 의견이 높기 때문에 즉, 서울숲은 주차 부족 문제가 압도적으로 큰 문제로 인식됨을 알 수 있다. '주차'가 포함된 문장을 살펴보면 중요 키워드로 같이 추출된 '부족', '주말', '복잡'이라는 키워드와 함께 등장한다. 주말의 높은 공원 이용률에 따라 주차 부족 문제와 함께 복잡함에 대한 불만들이 많았다. 많은 사람으로 여유 공간이 부족하고 휴식하기 어렵다는 부정적인 의견이 있었다. '숲' 키워드에는 숲을 기대했는데 평범한 공원이어서 실망했다는 내용이 있었다. '동물' 키워드는 주로 반려견 관련 내용이었으며, 공원 내 배변 문제와 반려견이 많아 복잡했다는 부정적인 의견들이 있었다. '편의시설' 키워드는 공원 규모에 비해 내부 편의시설이 부족하다는 부정적인 의견이었다.

Table 7에서 보라매공원의 별점 3 이하 리뷰데이터를 보면 '공사' 키워드가 압도적으로 많았다. 원문에서 '공사'는 '불편', '경전철', '복잡', '좁다'라는 단어와 함께 나타나고 있다. '공사'는 전체 리뷰 데이터에서도 높은 값을 나타냈는데, 신림선 경전철 공사가 공원 이용에 있어 크게 부정적인 영향으로 작용한 것을 알 수 있다.

보라매공원은 지역 주민의 이용률이 높고, 애견놀이터와 애견 산책로가 잘 정비되어 있어 '애견'과 관련된 다양한 이슈가 있다. 별점 3 이하의 리뷰데이터에서 '애견'키워드가 들어간 내

Table 6. Key words in Olympic Park (Top 20)

Word	Cen.	Word	Cen.
Stroll (산책)	0.1688	Event (행사)	0.0375
Exercise (운동)	0.0625	Stadium (경기장)	0.0375
Culture (문화)	0.0625	Bicycle (자전거)	0.0313
Diversity (다양)	0.0563	Rest (휴식)	0.0313
Olympic (올림픽)	0.05	Alone Tree (나홀로나무)	0.0313
Facility (시설)	0.05	Performance (공연)	0.0313
Sports (스포츠)	0.05	Couple (연인)	0.0313
Child (아이)	0.0437	Mongchontoseong (몽촌토성)	0.0313
Museum (박물관)	0.0437	Dog (애견)	0.0313
Family (가족)	0.0375	Photo (사진)	0.025

\*Cen.: Degree centrality.



Table 7. Key words for rating data below 3

Seoul Forest		Boramae Park		Olympic Park	
Word	Cen.	Word	Cen.	Word	Cen.
Parking (주차)	0.1538	Construction (공사)	0.3226	Facility (시설)	0.0652
Lack (부족)	0.1538	Exercise (운동)	0.0968	Problem (문제)	0.0543
People (사람)	0.0769	Dog (애견)	0.0968	Convenience (편의점)	0.0435
Forest (숲)	0.0769	Facility (시설)	0.0968	Photo (사진)	0.0435
Weekend (주말)	0.0769	Parking (주차)	0.0968	Paving (포장)	0.0435
Animal (동물)	0.0769	Stroll (산책)	0.0645	Photograph (촬영)	0.0435
Complexity (복잡)	0.0769	Discomfort (불편)	0.0323	Wedding (웨딩)	0.0435
Convenience (편의시설)	0.0769	LRT (경전철)	0.0323	Stroll (산책)	0.0435
Couple (연인)	0.0769	Complexity (복잡)	0.0323	Exercise (운동)	0.0326
Skyscraper (고층빌딩)	0.0769	Narrow (좁다)	0.0323	Restaurant (식당)	0.0326

\*Cen.: Degree centrality.

용을 살펴보면, 공원 내 배변 문제와 애견이 너무 많은 것에 대한 불만들이었다. 반대로 애견놀이터가 있어 좋다는 의견도 보였다. 반려견에 대해서는 견주와 그 외 이용자들 사이에 의견이 나누어짐을 알 수 있었다. 또한 ‘운동’, ‘시설’ 키워드에 대해서는 운동 시설의 노후 및 안정성에 대한 불만이 있었고, 많은 이용자로 인한 조깅 및 운동 시의 불편함이나, 산책코스와 자전거 도로의 미분리로 인한 안전 문제, 이용 질서에 대한 문제 등이 지적되었다. 또한, 너무 많은 시설에 비해 녹지가 적어서 시설로만 들어갈 것 같다는 우려의 의견도 있었다. 이와 함께 주차 부족에 대한 의견도 있었다.

올림픽공원의 리뷰가 있는 별점 3 이하의 데이터는 44개로 매우 적어서 분석에 한계가 있었다. Table 7에서 별점 3 이하의 리뷰 데이터를 보면 10개 키워드 모두 중심성 값이 비슷하며, ‘시설’과 ‘문제’ 키워드가 조금 높은 값을 나타내고 있다. 시설이 너무 많아 시설 활용성이 적다는 내용과 시설물 노후화와 관련된 내용이 있었다. 하지만 데이터 수가 적어 올림픽공원에 대한 부정적 의견으로 대표하기는 어렵다.

### 3) 외국인 리뷰데이터에 대한 키워드 분석

Table 8에서 서울숲 외국인 리뷰 데이터를 보면 ‘자전거’, ‘사슴’, ‘나무’가 높은 중심성 값을 나타낸다. 추출된 중요 키워드들이 전체 리뷰데이터와 비슷한 결과를 보이지만, 전체 리뷰데이터와 다르게 외국인 리뷰에서는 이용목적이 ‘산책’보다 ‘자전

거’가 높은 것을 유추할 수 있다. 또한, 전체 리뷰데이터 분석과 같이 ‘사슴’, ‘나무’, ‘나비’, ‘숲’, ‘정원’과 같은 자연과 관련된 단어들이 많이 추출되었다. 외국인들에게도 서울숲의 이미지는 자연이 강조되며 관련된 수요가 높은 것을 알 수 있다.

보라매공원의 외국인 리뷰 데이터는 30개로 수가 적어 분석에 한계가 있었다. Table 8을 보면 ‘애견’, ‘조깅’, ‘시설’이 높게 나타나고 있다. 하지만 마찬가지로 분석 데이터의 수가 적어 외국인 이용 특성을 대표한다고 말하기 어렵다.

Table 8에서 올림픽 공원의 외국인 리뷰 데이터를 보면 ‘올림픽’이 가장 높은 값을 보이며, 뒤에 이어 ‘산책’, ‘자전거’, ‘스포츠’가 나타난다. ‘올림픽’ 키워드는 서울올림픽과 관련하여 기념비적인 상징과 관련된 의견들이 많았다. ‘산책’과 ‘자전거’를 타기 좋은 공원으로 인식하고 있고, ‘스포츠’와 ‘예술’이 있는 문화공원으로 인식하고 있었다. 올림픽공원은 특히 세 공원 중 외국인 리뷰의 비율이 가장 높은 공원이었고, 외국인들은 올림픽공원을 대체로 “서울에서 나무가 많고, 편안하고 걷기 좋은 장소”로 인식하고 있었다.

## 3. 공원별 네트워크 구조에 따른 특성 분석

### 1) 서울숲

응집성 분석(cohesion analysis)을 활용해 공원별 네트워크 구조를 Clustered Map으로 시각적으로 도출하였다. Clustered

Table 8. Key words for foreign review data

Seoul Forest		Boramae Park		Olympic Park	
Word	Cen.	Word	Cen.	Word	Cen.
Bicycle (자전거)	0.1026	Dog (애견)	0.0746	Olympic (올림픽)	0.1818
Deer (사슴)	0.0940	Jogging (조깅)	0.0746	Stroll (산책)	0.0727
Tree (나무)	0.0940	Facility (시설)	0.0746	Bicycle (자전거)	0.0727
Picnic (피크닉)	0.0855	Teenager (청소년)	0.0597	Sports (스포츠)	0.0727
Family (가족)	0.0855	Library (도서관)	0.0597	Art (예술)	0.0545
Child (아이)	0.0855	Old people (노인)	0.0597	Subway (지하철)	0.0545
Butterfly (나비)	0.0684	Comfort (편안)	0.0448	Work (작품)	0.0545
Forest (숲)	0.0598	People (사람)	0.0448	Family (가족)	0.0364
Garden (정원)	0.0598	Lawn (잔디밭)	0.0448	Tree (나무)	0.0364
Stroll (산책)	0.0513	Construction (공사)	0.0299	Comfort (편안)	0.0364

\*Cen.: Degree centrality.

Map의 노드 크기는 연결 중심성(degree centrality) 값을 기준으로 설정했으며, 중심성 값이 상위 10개의 단어를 검은색으로 표시하고, 어떤 하위 단어들과 연결되는지를 분석했다.

Figure 4는 서울숲의 네트워크 구조로 그룹1은 ‘가족’, ‘산책’, ‘자전거’, ‘주차’를 중심으로 연결망이 형성되어 있다. ‘가족’, ‘산책’을 중심으로 한 일반적인 공원 이용행태를 나타낸다고 할 수 있다. ‘자전거’ 산책, ‘한강’ 수변공원과 연계된 활동 등이 이용 특징으로 보인다. 또한 ‘주말’, ‘주차’ 키워드가 높은 중심성을 가지고 있어서, 앞서 말한 바와 같이 서울숲에서 주말의 주차 문제가 공원 이용에 관한 큰 이슈임을 알 수 있다.

그룹 2는 ‘아이’, ‘숲’, ‘사슴’, ‘놀이터’를 중심으로 연결망이 형성되어 있으며, 자연을 기반으로 한 체험과 시설들이 그룹을 형성하고 있다. 하위 키워드로 ‘곤충식물원’, ‘나비정원’과 같은 시설들이 있고, ‘먹이체험’, ‘힐링’과 같은 체험 키워드도 보인다. 그룹 2에서 보이는 ‘자연체험’은 서울 숲의 대표적인 이용행태이다.

그룹 3은 ‘공원 주변 상권을 포함한 다양한 체험’을 나타낸다. ‘다양’이라는 키워드를 중심으로 카페, 식당과 같은 편의시설 및 볼거리들이 연결망으로 형성되어 있다. 앞서 분석한 부정적 견해에서 서울숲은 공원 내부에 편의시설이 부족하다는 점을 이야기하였으나, 공원 외부의 성수동과 갤러리아포레 등 주변에는 다양한 카페, 레스토랑이 있어서 함께 이용하는 패턴이 많이 있음을 알 수 있다.

그룹 4는 ‘나무’ 키워드를 중심으로 ‘꽃’, ‘튤립’, ‘자연’과 같은 자연적 요소의 그룹이다. 서울숲은 나무와 꽃 등 자연적 요소를 체험하는 프로그램을 다수 제공하며, 이를 통해서 이용자들은 ‘나무’, ‘초화류’ 등 식생을 탐방하는 기회가 많이 있음을 알 수 있다.

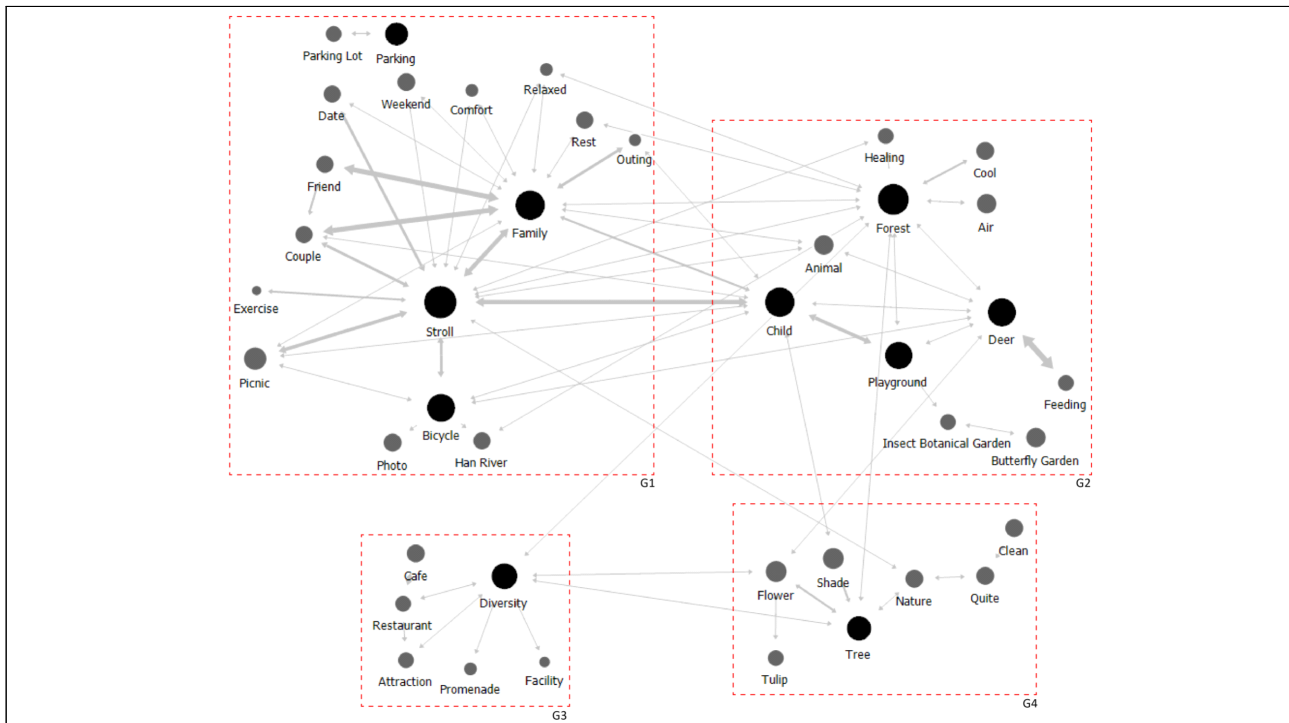
## 2) 보라매공원

Figure 5는 보라매공원의 네트워크 구조로 그룹 1은 ‘산책’, ‘가족’, ‘음악분수’, ‘애견’을 중심으로 연결망이 형성되어 있고, 서울숲의 그룹 1과 마찬가지로 주로 일반적인 공원 이용행태들이 나타난다.

여기서 ‘음악분수’는 보라매공원의 주요 시설로 인식되며, ‘산책’과 ‘호수’, ‘애견놀이터’와 연결되어 있다. 또한, 중요 키워드 분석과 같이 보라매공원에서 ‘애견’, ‘애견놀이터’는 보라매공원을 특징짓는 중요한 공간이 되고 있음을 알 수 있다.

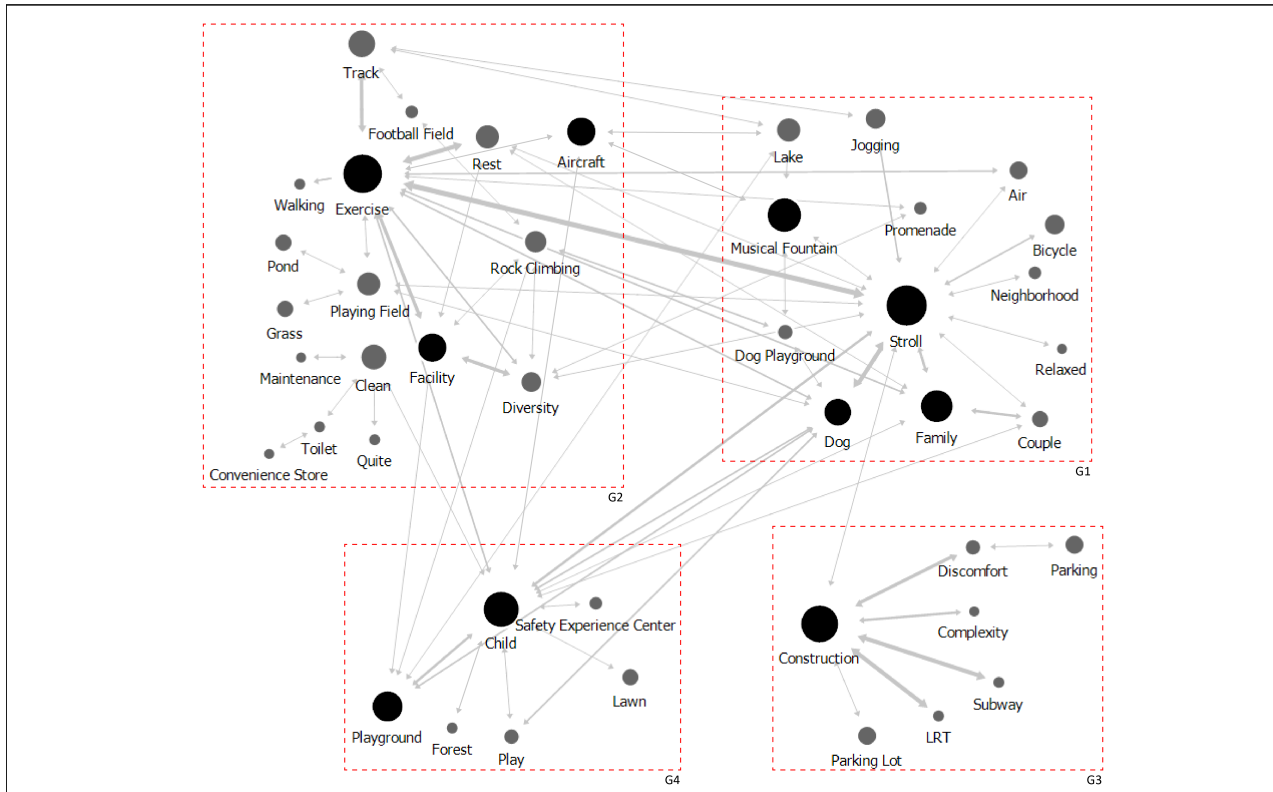
그룹 2는 ‘다양한 운동장소’를 나타낸다. ‘운동’, ‘시설’, ‘전투기’를 중심으로 연결망이 형성되어 있으며, 시설 관련 키워드들이 하위 키워드로 많이 나타난다. ‘트랙’, ‘축구장’, ‘운동장’, ‘암벽등반’과 같은 운동과 관련된 시설 및 활동들이 있다. ‘시설’과 ‘깨끗’ 키워드가 높은 중심성을 가지며 서로 연결되어 있는데, 보라매공원은 많은 이용자에게 깨끗한 시설관리 이미지를 가진 것을 알 수 있다.

그룹 3은 ‘공사’를 중심으로 전철 공사 관련 연결망이 형성되



●: Top 10 words for degree centrality

Figure 4. Cohesion analysis about Seoul Forest



●: Top 10 words for degree centrality

Figure 5. Cohesion analysis about Boramae Park

어 있고, 경전철 공사로 인한 ‘불편’, ‘복잡’이 하위 키워드로 연결되어 있다. 일시적인 현상일지라도, 보라매공원 이용에 있어서 현재 경전철 공사는 큰 이슈로 인식됨을 알 수 있다.

그룹 4는 아이들을 위한 놀이 및 체험시설 그룹이다. ‘아이’, ‘놀이터’를 중심으로 ‘안전체험관’이 하위 키워드로 나왔으며, ‘숲’, ‘놀이’, ‘잔디밭’과 연결성을 가진다.

### 3) 올림픽공원

Figure 6에서 올림픽공원의 그룹 1도 앞서 두 공원과 마찬가지로 ‘산책’을 중심으로 연결망이 형성된 일반적인 공원 이용행태 그룹이다. ‘피크닉’, ‘데이트’, ‘휴식’과 같은 이용행태와 ‘가족’, ‘연인’, ‘애견’, ‘아이’와 같은 동반 유형들이 있다. 또한 ‘편안’, ‘평화’, ‘아름답다’, ‘여유’와 같은 긍정적인 감정 키워드도 함께 나타났다. 특히 ‘산책’은 올림픽 공원의 다른 키워드들보다 Centrality 값이 현저하게 높았다. 이에 따라 ‘산책’을 중심으로 많은 하위 키워드들과 함께 그룹이 형성되었고, 올림픽공원에서 가장 보편적이면서 중요한 이용행태라고 볼 수 있다.

그룹 2는 ‘문화’, ‘운동’, ‘다양’, ‘스포츠’, ‘시설’을 중심으로 연결망이 형성되어 있고, 다양한 스포츠 시설 및 문화예술과 관련된 그룹이다. ‘경기장’, ‘시설’과 같은 인프라와 ‘행사’, ‘콘서트’, ‘예술’, ‘공연’이 나타났다. 문화 및 스포츠 이벤트는 올림픽공원을 대표하는 이용 특성이다.

그룹 3은 올림픽공원의 랜드마크이자 포토존인 나홀로나무 촬영과 관련된 그룹이다. ‘나홀로나무’를 기점으로 ‘올림픽’과

‘촬영’, ‘사진’이 연결성을 가지며 그룹을 형성하고 있다.

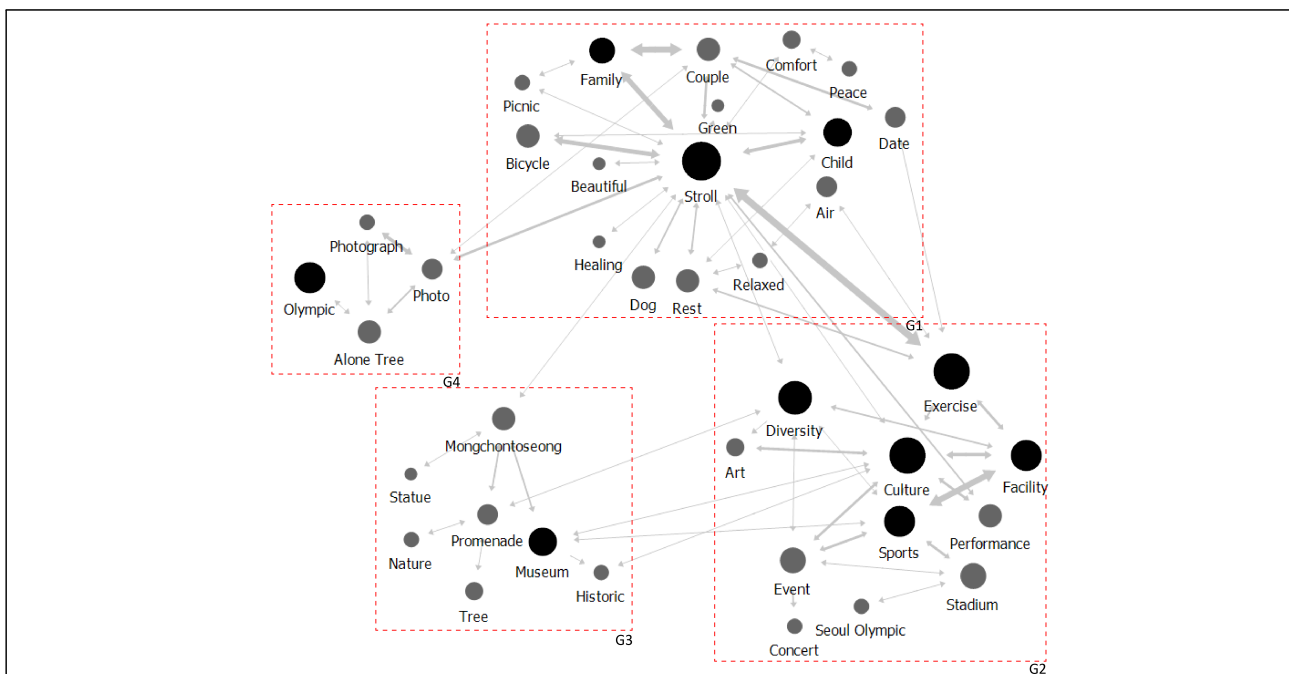
그룹 4는 ‘몽촌토성’, ‘박물관’과 연결된 공원이 가지는 장소성 및 역사성 맥락을 나타내는 그룹이다. 올림픽공원 내에 있는 백제 문화재인 몽촌토성과 박물관을 중심으로 ‘산책로’와 ‘조각상’ 등이 포함되어 있다.

### 4) 분석 종합

공원별 네트워크 구조에 따른 특성 분석을 종합하면 세 개의 공원은 각각 4개의 그룹으로 특성들이 나타났고, 특히 각 공원의 그룹 1은 모두 일반적인 공원 이용행태와 관련된 그룹으로, 공원이 가지는 대표적인 활동을 나타낸다. 가족, 연인 등 편한 상대와 함께 하는 ‘산책’은 세 공원에서 나타나는 대표적인 활동이다. 일상적 공원 이용 그룹 안에서 유아, 자전거, 반려견 등 각 공원마다 특성도 함께 하위 키워드로 나타났다. 그 외 그룹 2, 3, 4는 각 공원의 공간 환경 및 시설의 특수성과 현재의 특수한 이용환경이 잘 나타났다. 공원별로 나타난 그룹의 특성 및 주요 키워드는 Table 9와 같이 정리하였다.

## V. 결론

구글맵리뷰의 텍스트마이닝을 활용하여 공원별로 전체 리뷰, 별점 3 이하 리뷰(부정적 인식), 외국인 리뷰에 나타나는 중요 키워드를 추출했고, 키워들 간의 네트워크 구조를 파악했다. 이



\* ●: Top 10 words for degree centrality

Figure 6. Cohesion analysis about Olympic Park

Table 9. Group characteristics by park

Park name	Seoul Forest	Boramae Park	Olympic Park
G1	Stroll and family activity		
	'Stroll', 'family', 'bicycle', 'parking'	'Stroll', 'musical', 'fountain', 'family', 'dog'	'Stroll', 'family', 'child', 'bicycle'
G2	Nature-experience program	Exercise-related activities and facilities	Sports and cultural arts
	'Forest', 'child', 'deer', 'playground'	'Exercise', 'facility', 'aircraft', 'track'	'Exercise', 'culture', 'sports', 'diversity'
G3	Adjacent convenience and attractions	Inconvenience caused by construction	Alone tree photography
	'Diversity', 'cafe', 'restaurant', 'attraction'	'Construction', 'discomfort', 'complexity'	'Alone tree', 'olympic', 'photo'
G4	Natural elements	Play and experience facilities	Mongchontoseong history
	'Tree', 'flower', 'nature'	'Playground', 'child', 'play'	'Mongchontoseong', 'museum'

를 통해 공원 이용자들이 경험하는 공원 서비스에 대한 수요를 파악하였다. 연구에 사용된 구글맵리뷰는 장소 기반 리뷰 데이터이고, 공원을 직접 방문한 경험을 기술한 것으로, 연구 대상에 해당하는 각 공원 이용자들의 인식을 비교적 명확히 파악할 수 있었다. 또한, 구글맵리뷰는 리뷰와 함께 별점을 활용한 공원 이용 평가와 외국인의 이용 평가를 파악할 수 있어 네이버 블로그 글 등 기존 데이터에서 파악하기 어려운 공원에 대한 이용자 인식을 분석할 수 있었다.

리뷰어들은 공원을 대부분 긍정적인 공간으로 인식하고 있음을 알 수 있었고, 부정적인 의견에 대해서는 별점 3 혹은 그 이하로 평가하는 경향이 있음을 알 수 있었다. 이를 통해 이용에 부정적인 영향을 끼치는 요소와 이슈들을 파악할 수 있었다. 분석에서 나타난 주된 부정적인 내용은 공원의 혼잡, 반려견 문제, 산책로, 시설 노후 및 활용성 등이었다.

본 분석에서 외국인 리뷰 비율은 데이터의 약 10% 정도였으며, 전반적으로는 국내 이용자들의 인식과 비슷한 경향을 보였지만, 공원별로 외국인들의 주로 인식하는 주된 정보를 파악할 수 있었다. 또한, 공원의 네트워크 분석을 통해서 공원이 제공하는 서비스와 공원 이용자의 행태와 선호도를 이해할 수 있었다. 이러한 정보는 공원 서비스 질의 개선 및 관리로 연결할 수 있다.

공원 운영관리계획은 공원의 지형적 상황, 관리인 관점, 이해관계자 인식, 이용자 수요, 건강효과 등을 고려하여 정책(policy), 자금(funding), 거버넌스(governance), 평가(evaluation), 유지관리(maintenance)에 대한 계획을 세우는 것이다(Nam, J. and Dempsey, N., 2019.). 이용자 수요를 파악하기 위해 설문조사 형식의 이용자 만족도 조사를 매년 실시하고 있으며 공원의 전반 만족도, 시설관리 상태, 조경 및 경관, 안전성, 친절도, 건의 사항을 조사하고 있다. 구글맵리뷰는 이용자의 주관성이 강하여 편차가 컸으므로, 직접 평가지표로는 활용하긴 어렵다. 그러나 구글맵리뷰는 의도적인 평가에 대한 답변이 아니고, 누군가에

게 정보를 제공한다는 자발적이며, 이해관계가 개입되지 않은 '선의'의 의견이므로 공원에 대한 총평과 이미지, 주된 만족 및 불만족 요인을 모니터링하는 데 유용한 데이터가 될 수 있다.

본 연구는 공원 서비스에 대한 이용자 인식 분석의 새로운 수단으로 구글맵리뷰 데이터의 이용 가능성에 대해서 연구하였다. 공원 서비스에 대한 수요는 다양한 데이터 소스를 활용하여 파악할 필요가 있으므로, 구글맵리뷰 데이터만을 사용하여 공원 이용자의 인식 및 평가를 분석한 것은 본 연구의 한계점이다. 따라서 후속 연구로 네이버 블로그 등 다양한 소셜미디어 데이터와 구글맵리뷰 데이터 간 차이에 대한 연구가 필요하다고 하겠다.

또한 아직 국내에 구글맵리뷰 서비스 이용이 활발하지 않아 기간이 짧고 데이터 수가 적어 상대적으로 데이터가 많은 공원만을 분석할 수 있었으며, 연도별 편차가 존재했다. 하지만 지속적으로 축적한 데이터는 더 큰 신뢰성을 가진 데이터가 되므로, 향후 꾸준히 데이터를 수집해 간다면 구글맵리뷰가 가지는 의미도 더 커지리라 생각한다. 데이터가 충분히 누적되면, 공원 이미지에 대한 이용자 인식의 변화를 동적으로 모니터링할 수 있다.

마지막으로 빅데이터는 어떤 현상에 대한 배경과 맥락에 대한 포괄적 정보를 제공해 주지 못하므로, 기존 공원 및 경관에 대한 이론과 사회문화적 이슈들을 결합한 방법론 연구가 추가적으로 필요하다. 또한, 구글서비스를 포함한 SNS 데이터 모두는 작성자의 기억과 인식이 증폭되거나 왜곡될 가능성이 있기 때문에, 이에 대해서도 주의 깊은 해석이 필요하며, 더 나아가 공원의 운영관리에 활용할 수 있는 다양한 소셜미디어를 활용한 공원 평가에 대한 프로세스 개발이 필요하다.

## References

1. Bae, J. K.(2016) The effects of post-adoption beliefs of Chinese social

- commerce consumers on user satisfaction and continuous usage intention : Focused on the three major social commerce services in China, The Korean Association of Logos Management 14(2): 115-134.
2. Bark, S. H. and Y. G. Kim(2011) The landscape of Seonyoo-do park captured in one-person media focusing on blogs, Journal of the Korean Institute of Landscape Architecture 39(3): 64-73.
  3. Buneman, P.(1997) Semistructured data, In Proceedings of the Sixteenth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems (pp. 117-121).
  4. Byun, S., D. Lee and N. Kim(2016) Methodology for identifying issues of user reviews from the perspective of evaluation criteria: Focus on a hotel information site, Journal of Intelligence and Information Systems 22(3): 23-43.
  5. CABE(2010), Urban Green Nation: Building the Evidence Base.
  6. Chang, B. H., and S. H. Nam(2012) Social media and network journalism: Focusing on audiences' evaluations on journalism functions of social media, Korea Regional Communication Research Association 12(4): 457-496.
  7. Chae, S. H., J. I. Lim and J. Kang(2015) A comparative analysis of social commerce and open market using user reviews in Korean mobile commerce, Journal of Intelligence and Information Systems 21(4): 53-77.
  8. Chae, I. Y.(2017) A method for analysis of preferences of places based on sentimental analysis using SNS data: Case study on theme parks in Seoul, South Korea, Seoul University, 1-92.
  9. Chae, I. Y., J. Y. Kim, and K. Y. Yu(2017) Place Preference Analysis Using Twitter Data and Google Reviews.
  10. Cho, S. Y., H. K., Kim, B., Kim, and H. W. Kim(2014) Predicting movie revenue by online review mining: Using the opening week online review, Information Systems Review 16(3): 113-134.
  11. Choi, J., H. Ryu, D. Yu, N. Kim and Y. Kim(2016) System design for analysis and evaluation of e-commerce products using review sentiment word analysis, KIISE Transactions on Computing Practices 22(5): 209-217.
  12. Daniel, B.(2015) Big data and analytics in higher education: Opportunities and challenges, British Journal of Educational Technology 46(5): 904-920.
  13. Feldman, R. and J. Sanger(2007) The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data, Cambridge University Press.
  14. Guerrero, P., M. S. Møller, A. S. Olafsson, and B. Snizek(2016) Revealing cultural ecosystem services through Instagram images: The potential of social media volunteered geographic information for urban green infrastructure planning and governance, Urban Planning 1(2): 1-17.
  15. Hearst, M. A.(1999) Untangling text data mining, In Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics (pp. 3-10).
  16. Hearst, M.(2003) What is Text Mining, SIMS, UC Berkeley, 5.
  17. Hoaglin, D. C., F. Mosteller, and J. W. Tukey(2000) Understanding Robust and Exploratory Data Analysis (No. Sirsi) i9780471384915.
  18. Jeon, B. and H. Ahn(2015) A collaborative filtering system combined with users' review mining: Application to the recommendation of smartphone apps, Journal of Intelligence and Information Systems 21(2): 1-18.
  19. Jin, S. A., G. E. Heo, Y. K. Jeong, and M. Song(2013) Topic-network based topic shift detection on twitter, Journal of the Korean Society for Information Management 30(1): 285-302.
  20. Kavaratzis, M., and G. J. Ashworth(2005) City branding: An effective assertion of identity or a transitory marketing trick?, Tijdschrift voor economische en sociale geografie 96(5): 506-514.
  21. Kim, J. E., C. Park, A. Y. Kim, and H. G. Kim(2019) Analysis of behavioral characteristics by park types displayed in 3rd generation SNS, Journal of the Korean Institute of Landscape Architecture 47(2): 49-58.
  22. Kim, S. R., Y. Choi, and H. Yoon(2019) The analysis of the visitors' experiences in Yeonnam-dong before and after the Gyeongui line park project-a text mining approach, Journal of the Korean Institute of Landscape Architecture 47(4): 33-49.
  23. Ko, H. J.(2019) Evaluation of Ecosystem Services of Urban Green Spaces Using Text Mining Techniques: Focused on Cultural Ecosystem Services in Gwacheon and Ansan, Republic of Korea, Doctoral Dissertation, Seoul National University.
  24. Lee, J. I.(2013) A Study on the Micro-Discourse about Seoul Forest in Personal Media: Focused on Sense of Place, Master's Thesis, Seoul National University.
  25. Lee, J. H.(2013) Understanding customer values by analyzing the contents of online hotel reviews, The Journal of the Korea Contents Association 13(10): 533-546.
  26. Lee, S. H., J. Cui, and J. W. Kim(2016) Sentiment analysis on movie review through building modified sentiment dictionary by movie genre, Journal of Intelligence and Information Systems 22(2): 97-113.
  27. Lee, Y. J., J. H. Ji, G. Woo and H. G. Cho(2009) TRIB: A clustering and visualization system for responding comments on blogs, The KIPS Transactions: PartD 16(5): 817-824.
  28. Lee, Y. M., P. Kwon, K. Y. Yu, and J. Y. Kim(2017) Method for spatial sentiment lexicon construction using Korean place reviews, Journal of Korean Society for Geospatial Information System 25(2): 3-12.
  29. Litvin, S. W., R. E. Goldsmith and B. Pan(2008) Electronic word-of-mouth in hospitality and tourism management, Tourism Management 29(3): 458-468.
  30. Lund, N. F., S. A. Cohen, and C. Scarles(2018) The power of social media storytelling in destination branding, Journal of Destination Marketing & Management 8: 271-280.
  31. Matsuoka Keisuke, 「Google Maps, The Birth of a New World」, Hong Sung Min, Wisdom House(2017), p165.
  32. Munawir, M., D. Koerniawan and B. J. Dewancker(2019) Visitor perceptions and effectiveness of place branding strategies in thematic parks in Bandung City using text mining based on Google maps user reviews, Sustainability 11(7): 2123.
  33. Nam, J., and N. Dempsey(2019) Place-keeping for health? Charting the challenges for urban park management in practice, Sustainability 11(16): 4383.
  34. Park, J. D.(2015) A study on mapping users' topic interest for question routing for community-based Q&A service, Journal of the Korean Society for Information Management 32(3): 397-412.
  35. Richards, D. R. and D. A. Friess(2015) A rapid indicator of cultural ecosystem service usage at a fine spatial scale: Content analysis of social media photographs, Ecological Indicators 53: 187-195.
  36. Son, Y. H., and Y. J. Kim(2019) The image of Ruralism in Korea through a text mining for online news media analysis, Journal of Korean Society of Rural Planning 25(4): 13-26.
  37. Tukey, J. W.(1977) Exploratory Data Analysis 2: 131-160.
  38. Warnock, S., and G. Griffiths(2015) Landscape characterisation: The living landscapes approach in the UK, Landscape Research 40(3): 261-278.
  39. Woo, K. S., and J. H. Suh(2018) Time series analysis of park use behavior utilizing big data-targeting Olympic Park, Journal of the Korean Institute of Landscape Architecture 46(2): 27-36.

40. Xianghua, F., L. Guo, G. Yanyan, and W. Zhiqiang(2013) Multi-aspect sentiment analysis for Chinese online social reviews based on topic modeling and HowNet lexicon, Knowledge-Based Systems 37, 186-195.

---

Received : 21 May 2021

Revised : 7 July 2021

(1st)

Accepted : 7 July 2021

3인익명 심사필