

코퍼스 분석을 활용한 지역재생정책 연구*

- 감천문화마을 관광활성화를 위한 정책제언 -

이 태 현** · 정 하 영*** · 이 새 미****

요 약

본 논문은 빅데이터 분석 기술을 도시재생정책의 평가 및 수립에 접목하기 위한 탐색적 연구로 감천문화마을의 관광활성화를 위한 정책제언을 목적으로 한다. 이를 위해 구글맵에 게시된 감천문화마을에 대한 리뷰 텍스트 전부인 총 1,503건을 웹 스크래핑(Web scraping)을 통해 수집하였다. 이 중 텍스트를 포함하고 있지 않거나, 한국어, 중국어, 영어 이외의 언어로 작성된 리뷰 98건을 제외한 1,405건을 분석 데이터로 활용하였다. 이후, 비정형 텍스트를 자연어 처리하는 과정인 형태소 분석을 거쳐 TF-IDF(Term Frequency-Inverse Document Frequency) 값을 추출하는 코퍼스 분석을 실시하였다. 분석결과에 관광산업의 체험경제 이론을 접목하여 관광정책에 대한 시사점을 제시하였다. 체험경제이론에 따른 텍스트 분류를 통해 알 수 있는 감천문화마을 관광서비스의 핵심적인 한계점은 오락적 체험요소가 사진, 벽화, 어린왕자 등 공공예술작품을 중심으로 나타난 점이다. 이와 같이 공공예술 작품이 중심이 되는 관광 콘텐츠는 지역에 방문객을 유입시키는 중요한 요인이 되지만 직접적인 수익을 발생시키지 않는다는 문제를 내포하고 있다. 수익을 창출하는 콘텐츠의 키워드가 나타나지 않는 사실을 통해 현재의 감천문화마을 관광서비스가 지역주민의 이익을 창출하는 단계로 발돋움하지 못하고 있는 문제의 원인을 알 수 있다. 이러한 문제를 해결하기 위해서는 연령별, 국가별, 지역별 방문객의 특성을 파악하고 산업관점의 서비스 마련과 마케팅 대상을 구체화 하는 과정이 필요하다. 마케팅 대상을 구체화 하여 특화된 서비스를 점차 늘려 감으로써 현재 자영업의 밀집 수준의 관광서비스를 탈피하고 체계적인 지역관광산업이 성장해 갈 수 있다. 본 연구는 지역 거버넌스를 통하여 지역재생정책을 개선하기 위한 논의에서 하나의 사례로 감천문화마을재생을 위한 향후의 정책 방향을 검토하였다. 이 과정에서 코퍼스 분석 방법론이 공공정책수립의 기초자료를 제공해 줄 수 있음을 본 탐색적 연구를 통해 밝혔다. 본 연구는 설문 등 기존의 조사방법론이 가졌던 한계를 넘어 관광정책검토에 있어서 빅데이터의 분석을 통해 방문객의 관광체험에 따른 핵심인식을 집약한 결과를 도출한 초기연구로서 의의를 가진다.

핵심용어 : 지역재생정책, 감천문화마을, 빅데이터 분석, 관광활성화, 코퍼스 분석, 거버넌스

* 이 논문은 2018년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2018S1A3A2075240)

** 제1저자, 동국대학교 인구조사와사회협동연구소 전임연구원, napsem@gmail.com

*** 교신저자, 경남대학교 LINC+ 사업단 조교수, jhayeong@gmail.com

**** 공동저자, 동아대학교 스마트 거버넌스 연구센터 전임연구원, emailme6@naver.com

◇ 논문접수일: 2020.01.21 ◇ 논문심사일: 2020.03.09 ◇ 논문게재 확정일: 2020.03.30

I. 서론

도심의 낙후지역을 대상으로 하는 도시재생정책 실행은 공공자금의 투입을 통해 이루어지며 이러한 사업은 대상지의 선정과 필요성에 대한 사회구성원의 합의가 요구된다. 또한 이를 실행하는 주체는 재생사업 과정과 결과에 대한 설명책임(accountability)을 가지게 됨에 따라 그 결과가 공공의 이익으로 확산되어야만 정책의 지속성을 가질 수 있다. 정책지속을 위한 사회적 합의의 맥락에서 감천문화마을은 도시재생사업을 대표하는 상징적 의미를 가진다. 감천문화마을은 인구성장기에 거주민 수가 약 3만 명에 달했으나 2017년 6월에는 7,847명으로 감소한 심각한 도심의 쇠퇴지역이었다(국가통계포털, 2017). 이러한 문제를 해소하기 위해 산복도로 르네상스 사업과 연계한 도시재생사업이 2009년부터 실시되어 2018년에는 한해 약 257만 명이 방문하는 관광지로 성장하게 되었다(관광지식정보시스템 통계, 2018). 이와 같이 마을의 원형을 보존하면서 지역공동체의 역량강화를 통해 새로운 모습으로 변모하고 성장해 가는 과정은 지역주민과의 거버넌스를 통해 도시를 재생하고자 하는 현재의 도시재생정책의 중요한 성공사례가 된다.

감천문화마을이 선도재생지역으로 주목받고 있음에도 거주민 수는 지난 10년간 여전히 감소하고 있다. 지역 거주민이 감소하는 가운데에서도 산복도로 르네상스 사업 등을 통하여 지속적인 공공의 지원을 이끌어 낼 수 있었던 것은 많은 방문객이 찾는 관광지로서의 가치를 가지고 있기 때문이다. 부산광역시 작성한 「2025 부산광역시 도시재생전략계획」은 도시재생의 학술적 정의를 “도시재생은 산업구조의 변화 및 신도시 위주의 도시 확장으로 인해 상대적 낙후된 기존 도시에 새로운 기능을 도입·창출함으로써 쇠퇴한 도시의 삶의 질을 향상시키고 물리적 정비와 함께 사회적, 경제적 재활성화를 통합적으로 추진하는 일”로 제시하였다(부산광역시, 2015).

이와 같은 도시재생의 정의를 통해 감천문화마을의 재생 현황을 해석하면 일반근린지역에 관광지로서의 새로운 기능을 도입·창출함으로써 주민의 삶의 질 향상과 물리적 정비 및 사회적, 경제적 재활성화를 도모하는 사례로 볼 수 있다. 그러나 감천문화마을의 관광기능은 관광객이 급격히 증가하였지만 마을체류시간을 늘릴 수 있는 새로운 콘텐츠의 부족, 접근성의 한계, 다른 관광지와 동일화되는 현상 등 관광지로서의 한계가 나타나고 있다(김가령, 2014). 따라서 감천문화마을의 관광활성화를 위한 지역재생정책의 개선이 요구된다. 관광활성화를 위한 지역재생정책을 위해서는 방문객의 관광수요 특성을 파악하기 위한 조사와 분석과정이 선행되어야 한다.

본 연구의 목적은 감천문화마을 관광활성화를 위한 및 방문객의 수요 특성 파악을 위해 방문객이 구글 지역 리뷰에 자발적으로 작성한 텍스트를 직접 활용하여 감천문화마을 관광경험에 대한 분석을 실시하고 이에 따른 지역재생정책개선의 시사점을 제시하는 것이다.

텍스트를 활용한 빅데이터 분석이 활성화 되면서 기업의 마케팅이나 수요예측 등 다양한 분야에서 활용되고 있으나 관광활성화 분야에서의 비정형 빅데이터 분석 연구는 많지 않다. 본 연구는 빅데이터 분석 기술을 부산의 도시재생을 위해 시작된 감천문화 마을의 지역관광활성화를 위한 방문객 수요 특성 파악에 접목하기 위한 탐색적 연구로 감천문화마을의 관광활성화 정책제언을 목적으로 한다. 이를 위하여 감천문화마을 방문경험자가 남긴 구글맵의 지역리뷰데이터 전체인 1,503건을 대상으로 코퍼스

분석을 실시한다. 대용량 텍스트 데이터의 수집을 통해 이루어지는 코퍼스 분석은 기존의 인터뷰조사 및 설문조사보다 폭넓은 표본을 기반으로 하며 조사자의 의도가 개입되지 않은 장점이 있다.

본 연구는 6장으로 구성되었다. 2장에서는 공공정책 분야에서 활용되어온 코퍼스 분석에 대한 선행연구와 정책개선의 검토 대상인 감천문화마을의 재생에 대한 선행연구를 소개하고 본 연구가 선행연구와 가지는 차이점을 설명하였다. 또한 감천문화마을이 가지는 지역적 특성과 지역재생의 핵심동력이 관광산업으로 이동하고 있는 현황을 살펴보았다. 3장에서는 수집된 데이터의 속성과 코퍼스 분석 방법을 설명하였다. 4장에서는 감천문화마을 방문객의 핵심인식을 나타내는 TF-IDF 분석 결과를 제시하였다. 5장에서는 관광정책에서 유용하게 적용되는 체험경제 이론을 바탕으로 TF-IDF 분석결과를 5가지 의미단위로 분류하여 본 연구의 목적인 감천문화마을 관광정책개선의 시사점을 도출하였다. 6장에서는 정책현황분석에 있어서 비정형 데이터의 분석을 자동화하여 결과를 도출한 초기연구로서 본 연구가 가지는 의미와 한계점을 기술하였다.

II. 선행연구 및 이론적 검토

1. 감천문화마을의 도시재생

감천문화 마을은 한국전쟁시기 낮은 산지에 형성된 집단주거지역으로 구도심의 재생과 같은 일반적인 도시재생사업과는 다른 지리적인 특징을 가지고 있다. 감천문화 마을은 주거용도로 순화된 지역으로 부산의 산업성장기에 도심 노동자의 저렴한 베드타운의 역할을 하며 유지되었다. 도심에 가깝지만 산복도로에 위치한 불편한 접근성, 좁은 필지, 골목길로 된 주택 진입로 등으로 인해 현재의 일반적인 생활환경 수준으로 주택을 정비하기에 어려움이 존재한다. 오수처리 시설 등의 기본적인 도시인프라조차 모든 주택에 공급하기 어려운 지형구조로 되어있다. 이러한 지리적 약점으로 인해 건물의 노후화와 거주민의 감소가 지속적으로 이어질 수밖에 없었으며 인구의 22%가 65세 이상인 고령화 마을이 되었다(국가통계포털, 2017). 그러나 저렴한 임대료, 공동체 유대감 등을 찾아서 입주예술가들이 전입해 오면서 감천문화마을이라는 명칭을 가지게 되었으며 입주예술가 집단의 전입으로부터 시작된 공공예술 작품설치 작업과 소형 갤러리 문화가 형성됨에 따라 지역재생을 위한 공공지원이 시작되어 사람들의 주목을 받게 되는 성공적인 도시재생지역이 되었다. 감천문화마을 재생에 관한 연구는 감천이 관광지로 주목 받기 시작한 2011년을 기점으로 활발히 이루어 졌다. 특히, 공동체 활성화를 통한 지역재생역량 강화에 주목한 지역사회학 분야와 관광활성화를 위한 분야의 연구가 많이 이루어 졌다.

김형균 외(2015)의 연구는 도시계획이론을 바탕으로 감천문화마을 사업을 통한 주민생활 및 상권의 변화를 조사하였다. 감천문화마을의 관광객 변화추이와 상권의 변화 양상을 2009년부터 시작된 도시재생사업으로 발생한 지역의 지가 및 건물임대료 상승, 상업시설의 증가 등의 변화를 구체적으로 조사한 자료를 제시하였다.

Baik(2012)의 연구는 감천문화마을의 공공미술 설치사업에 집중하였다. 이 연구에서는 마을에 설치되는 공공미술 작품이 지역의 정체성에 부합하기 위한 노력보다 작가들의 작품세계에 치우치는 것에 대

한 문제를 제기하고 있다. 사람들의 발길을 끌어들이기 위해 눈의 띄는 작품전시에 의존하는 정책이 지속되면, 장기적으로 주민생활의 개선과 새로운 문화 창출을 통한 지역의 자생력과 매력 강화와는 다른 근시안적인 이슈생성에만 집중하게 되는 문제로 이어질 수 있음을 지적하고 있다.

김가령(2014)의 연구는 감천문화마을의 관광활성화가 주민들의 소득증가로 이어지지 않는 문제를 다루고 있다. 공정관광행동의 개념을 통해 거주 지역에 관광객이 증가하는 가운데 거주민에게는 실질적 소득증가가 발생하지 않는 상황을 해소해야 한다고 주장하였다.

윤지영(2016)은 내·외부인에 대한 설문조사를 통해 마을의 질적 삶을 향상시키기 위한 연구를 진행하였다. 이를 위해 도시 브랜드력(brand-power)의 제고가 요구되며, 도시 브랜드력은 지역주민의 만족도가 방문객들에게 전달되며, 결과적으로 지역에 대한 친밀감과 호감에 직접적인 영향을 끼치는 요소라고 보았다. 이러한 브랜드력을 제고하기 위한 전략으로 골목경제 활성화를 위한 지역성 콘텐츠 개발, 정책 공조를 통해 살아있는 생활문화박물관으로 만드는 전략, 이미지 제고 정책을 통한 매력성 강화, 주변 마을과의 연계강화, 마을자산을 브랜드로 창조하는 전략 등을 제시하였다.

최지영 외(2017)의 연구는 감천을 방문한 관광객의 설문조사를 통해 마을의 문화마케팅이 마을 브랜드 자산구축에 미치는 영향을 분석하고, 차별화된 관광지가 되기 위한 구체적인 방안을 모색하였다. 감천문화마을이 가진 고유의 이미지를 상품화 하여 브랜드 자산을 구축하는 것이 문화마케팅을 강화하고 마을 관광산업의 가치향상에 기여한다고 주장하였다.

감천문화마을 재생과 관련한 많은 선행연구들이 관광활성화와 관광서비스 강화에 초점을 두고 진행되었으며 이와 같은 연구들은 감천문화 마을 재생동력이 관광에 있다는 것을 전제하고 있다. 감천문화마을은 문화관광지로서의 성공 및 관광객 증가로 인해 대규모 상업투자자본의 유입이 시작되었다. 이는 공동체 활성화를 통한 마을 만들기 사업에서 관광 산업화를 통한 도시재생으로 재생사업의 성격이 급하게 변화하고 있다는 것을 의미한다. 이러한 변화는 문화공동체 공간이었던 마을이 상업공간으로 점진적으로 변모해 가고 있으며 이로 인해 지역이 관광객들의 주목을 받게 된 원동력이 상실되어 가고 있다고 보는 견해(윤지영, 2016)와 도시재생의 주도권이 주민과 입주예술가의 자율성을 통한 결정에서 상업자본의 주도로 변해가는 과도기라는 분석도 존재한다(이정석, 2013).

그러나 감천문화마을에 부산의 다른 산복도로에 인접한 다른 마을과는 비교할 수 없이 많은 공공지원이 지속적으로 이루어질 수 있었던 이유는 관광·상업지로서 성공에 있다. 이에 대한 근거로 제시할 수 있는 수치가 마을의 인구추이다. 감천마을의 정주민구 숫자는 ‘산복도로르네상스 사업’을 시작한 2011년 10,110명에서 2017년 7,847명으로 12.8%가 줄어들었으나, 한해 관광객 수는 지속적으로 증가하여 약 150만명을 상회하고 있다. 즉, 감천문화마을 재생동력의 핵심은 관광활성화에 따른 방문객 수의 증가에 있다고 보는 것이 타당하다.

그러나 관광객의 인식과 방문경험에 대한 조사와 이를 통해 관광수요의 특성을 구체화 하는 연구는 아직 진행되지 않고 있다. 이 가운데 본 연구가 실시한 관광객의 인식이 담긴 비정형 데이터 수집과 방문경험을 집약한 핵심어의 계량적 도출은 감천문화마을의 관광활성화를 위한 지역재생사업 개선에 시사점을 제공하는 기초자료로 활용될 수 있을 것이다. 기존의 지역재생정책 수립을 위한 기초자료는

「2025 부산광역시 도시재생전략계획」 등과 같이 ‘인구통계’, ‘관광객 수의 증감량’, ‘노후 건축물 비율’, ‘지역총생산’ 등과 같은 정량적인 데이터를 토대로 정책을 수립하고 재생 전략을 제시하고 있다. 그러나 본 연구가 실시하는 관광객이 남긴 텍스트에 대한 분석은 정형화된 틀에 맞추어진 정

략적인 통계자료가 포함하지 않는 이해관계자(관광객)의 직접적인 의견을 수집하고 분석하는 자료로서 정성적인 내용을 포함하는 가치가 있다. 따라서 이를 활용한 재생전략은 관광활성화와 지역재생전략과의 연계의 필요성을 전달하는 기초자료로 활용 될 수 있으며, 관광산업활성화 계획 수립 과정에서 현장의 문제에 가깝게 대응한다는 점에서 의미를 가진다.

2. 관광산업과 체험경제이론

Pine & Gilmore(1999)은 관광산업의 사업성과 향상을 위한 새로운 패러다임으로 체험경제를 제안했다. 관광산업에서 체험은 관광 목적지를 방문하고, 보고, 배우고, 즐기고, 다른 생활방식으로 생활하는 것을 말한다(Stamboulis & Skayannis, 2003). 이 과정에서 경험하는 모든 것은 관광객들에게 행동적, 지각적, 인지적, 정서적 또는 표현적, 암시적일 수 있다(Oh et al., 2007). 특히 체험경제의 개념은 이동하고, 보고, 지식과 정보를 습득하고, 먹고, 마시는 등 현대 서비스산업의 모든 측면을 활용하여 보다 매력적인 경험을 제공하는 관광산업과 관련이 있다(Pine & Gilmore, 1999).

방문객들은 여행 동기와 경험, 개인적 인식, 행동, 여행 전에 숙지해야할 정보를 공유한다. 이러한 여러 요소는 체험으로 요약할 수 있다. 자신의 체험을 공유하고 타인의 체험을 참고 하는 이유는 모든 관광객들이 특정 여행 상황에서 보다 매력적이고 독특하며 기억에 남는 경험을 추구하기 때문이다(Ellis & Rossman, 2008). 관광객들은 지각한 체험요인을 기억하고 공유함으로써 감정적 애착과 충성도(재방문, 추천, 긍정적 입소문)에 영향을 미친다(Loureiro, 2014). 즉, 관광산업의 성공여부에는 관광객들에게 의미 있는 체험요인의 제공여부가 큰 영향을 미친다(Oh et al., 2007).

장서희(2017)는 체험경제 이론을 바탕으로 대전광역시가 제공하는 관광체험요소를 체험가치에 따른 지불의사를 나타내는 체험경제지수로 나타내었다. 장서희의 연구는 체험경제이론이 말하는 체험가치가 관광객의 여행지출금액 등에 직접적인 영향을 미치는 지를 파악하였다는 점에서 의미를 가진다. 분석 결과 중 주목할 만한 부분은 조사그룹 중 체험경제지수가 높은 그룹이 식음료비, 오락비, 교통비에서 월등히 많은 소비를 했다는 점이다. 이와 같은 분석결과는 관광객들이 주관적으로 판단하는 체험에 대한 가치를 최대화 하는 것이 관광객의 소비를 활성화하는데 기여한다는 시사점을 제공한다. 또한 장서희(2018)는 소비자가 인식하는 호스피탈리티에 영향을 미치는 구체적인 체험가치 요소에 대한 연구를 통해 소비자가 느끼는 감정적 이득이 지불의사를 높인다는 연구결과를 제시하여 관광산업에 있어서 체험경제이론의 유용성과 활용방안을 강조하고 있다.

천덕희(2013)는 크루즈여행객들을 대상으로 크루즈여행의 체험요인이 감정적 반응과 인지적 반응에 미치는 영향을 파악하였다. 분석에서는 크루즈여행의 체험요인을 현실도피적, 심미적, 오락적, 교육적 체험으로 구성하였다. 선박이라는 한정된 공간에서 고객의 체험을 통해 경제적 효과를 창출시키고 있는 크루즈산업에 대한 분석결과는 바다를 배경으로 하는 한정적인 공간인 감천문화마을의 체험요인에도 유사한 시사점을 제공할수 있다는 점에서 의미를 가진다. 이 연구에서 주목할만한 분석결과는 크루즈여행자들이 크루즈선박 내의 물리적 환경이나 주변 바다의 아름다운 풍경 등과 같은 심미적 체험을 긍정적으로 인식할수록 크루즈여행의 감정적인 반응인 흥분, 활기, 흥미, 들뜬기분을 나타내는 각성을 인지하는데 가장 큰 영향을 주고 있다는 점이다. 감천문화마을은 바다와 높은 조망을 배경으로하는 심미적 요소를 많이 갖추고 있는 지역이라는 점에서 본 연구에 참고할만한 시사점을 제시하고 있다.

장윤연과 서원석(2014)의 연구는 외래 관광객의 쇼핑에 대해 단순히 물건을 사는 과정이 아니라 하나의 체험활동이라는 관점을 제시하고 있다. 그리고 쇼핑체험요소와 관광의 만족도 및 정서적 몰입의 상관관계에 대한 분석을 실시하였다. 연구에서는 ‘만족도’를 기대하는 감정과 실제 경험과의 차이에서 오는 것으로 정의하고 재방문과 충성도에 영향을 미친다고 보았다. 연구에서는 쇼핑의 체험요소가 만족도에 정(+)의 영향을 주고, 상대방과 자신의 감정적 애착 상태를 의미하는 ‘정서적 몰입’으로 이어진다는 결과를 제시하고 있다. 이 연구는 감천문화마을이 상업자본의 유입을 통해 기존의 노후 주거시설이 쇼핑용도로 재구성되고 있는 현 시점에서 외래 관광객의 관광체험에서 쇼핑이 가지는 중요성을 제시하고 있다.

최수진 외(2020)의 연구는 관광객의 관광경험 평가 및 관광지에 대한 만족도가 형성하는 ‘충성도’에 대한 접근을 실시하였다. 온라인 설문 조사를 토대로 관광경험을 조절변수로 하여 만족도와 충성도 간의 영향 관계를 규명하였다. 연구결과는 관광경험을 통해 얻어지는 개인적 차원의 삶의 활력이나 즐거움이 긍정적인 관광만족도에는 긍정적인 영향을 미치지만 재방문 등으로 이어지는 충성도에는 영향을 주지 않으며, 재방문과 추천으로 이어지는 관광지 자체에 대한 충성도는 방문지역의 독특한 문화체험이나 감각적인 경험이 요구된다는 점을 시사하고 있다. 또한 재방문 경험이 많을수록 충성도가 점점 강화된다는 결과를 제시하였다. 최수진 외(2020)의 연구에서 다루는 관광경험 역시 체험경제 이론을 토대로 하고 있으며 오락적 요소, 교육적 요소, 심미적 요소, 일탈적 요소를 변수로 측정하였으며, ‘충성도’에는 심미적 요소와 일탈적 요소가 중요한 영향을 미친다는 사실을 시사하는 연구다.

본 연구는 선행연구들이 제시하는 관광산업에 있어서 체험경제이론 적용의 유용성을 토대로 감천문화 마을의 재생을 위한 관광활성화 과정에 체험경제이론을 적용하였다. 본 연구에서도 기존의 연구와 마찬가지로 체험경제요소를 오락성, 교육성, 일탈성, 심미성의 4가지 요소를 중심으로 파악하고 있다. 그러나 기존의 연구는 연구자가 작성한 정형화된 질문지에 대한 방문객의 응답을 통해 4가지 요소를 계량하는 방법론을 사용한 것에 반해, 본 연구는 방문객이 작성한 비정형 텍스트 데이터를 토대로 TF-IDF분석으로 정형화 하고 연구자가 의미단위로 분류하여 계량하였다는 점에서 방법론적인 차이를 가진다.

3. 공공정책 검토를 위한 코퍼스 분석

감천문화마을 재생사례와 같은 지역사회공동체의 역량강화로부터 출발하는 도시재생정책은 지역주민협의체와 공공정부의 협력을 중심으로 공동운영체제 구축과 자립적인 도시재생 동력생산을 목표로 한다는 점에서 거버넌스 개념을 기초로 한다. 이와같이 시민사회와의 협치를 도모하는 정책형성과정에서는 시민들의 의견으로부터 정책의 증거를 발견하고, 시민들과의 합의형성을 통해 결정에 이르는 과정이 중요한 의미를 가진다. 그러나 불특정 다수인 시민들로 부터 정책근거를 찾아내고, 합의형성을 위하여 시민들의 실질적인 공적의사(public concern)를 집약하는 데에는 많은 시간과 비용이 소요된다. 이와 같은 문제를 해결하기 위하여 코퍼스 언어학 기반의 내용분석 (Corpus Linguistic Based Content Analysis) 이 활용된다. 코퍼스 분석은 공공정책의 검토를 위하여 시민들의 공적의사를 집약하거나 사회의 거시적 토론을 분석하기 위해 대량의 텍스트로부터 타 연구자가 시도하더라도 동일한 결과를 도출 할 수 있으며 계량적으로 파악 가능한 정보를 추출하고자 하는 연구방법이다(정하영 외, 2017).

Jeong et al.(2008)은 TF-IDF, T-VSM(Term-based Vector Space Machine) 및 MDS(Multi-Dimensional Scaling)를 활용한 코퍼스 담화분석을 토대로 사업주체간 회의, 주민협의회, NGO단체와 지역행정과의 토론 등의 녹취록을 수집하여 계량화하는 분석을 실시하였다. 이들의 연구는 공공사업의 사회적 적절성을 판단하기 위하여 토론과정의 이해관계자 간의 인식의 차이와 정책 주장의 정통성을 검토하는 과정을 분석함으로써 코퍼스 분석을 활용한 공공정책검토의 유용성을 제시하고 있다.

최보경 외(2014)의 연구는 금연정책이라는 이슈에 대한 관련 기사의 속성과 댓글의 유형이 어떻게 달라지는지를 파악하기 위하여 ‘뉴스 프레임’, ‘정보원’, ‘하위 주제’로 분류된 기사속성 항목과 ‘이성·감성댓글’, ‘정책지지 댓글’로 분류한 댓글 속성 항목간의 발생빈도를 비교 분석하였다. 이들의 연구는 텍스트를 ‘기사’와 ‘댓글’로 분류하고 세부적인 하위분류를 실시하여 시사점을 도출하였다. 이러한 분석방법을 바탕으로 핵심단어를 의미단위로 분류하고자하는 본연구의 방법론에 활용하였다.

장미화 외(2016)는 정부의 캠핑정책에 대한 시민들의 인식을 분석하여 향후 정책수립을 위한 시사점을 제안하였다. 논문, 보고서 및 보도 자료를 기반으로 정책 분석을 실시하고, 정책에 대한 대중의 인식을 파악하기 위하여 소셜 미디어의 텍스트 225,279건을 분석하였다. 관광분야 정책의 문제점 파악과 향후 정책개선 방향을 소셜 미디어 빅데이터 분석을 통해 제시하고 있다는 점에서 본 연구의 방법론과 유사점을 가진다.

이태현 외(2017)의 연구에서는 해수담수화플랜트 가동을 통한 부산시 상수도 공급정책안에 대하여 국내 일간지 미디어 간에 인식차이가 존재하는 지를 파악하고, 이를 객관화하기 위한 연구를 진행하였다. 기장군의 해수담수화 플랜트 가동을 통한 상수도 공급이라는 사안에 대한 각 일간지 미디어가 가지는 인식을 집약한 핵심어를 추출하기 위한 코퍼스분석을 통해 대립이나 갈등을 해소할 수 있는 단서를 제공하는 것을 목적으로 하였다. 해수담수화 시설에 대한 기사에 나타난 핵심단어의 ‘의미단위’ 분류방법을 참고하였다.

정하영 외(2017)의 연구는 부산시 2015년 상반기 민원데이터의 내용을 효율적으로 집약하여 주요민원의 변화를 파악하기 위하여 코퍼스분석기법을 응용한 공기어구조맵을 작성하였다. “부산시에 바란다”라는 전자민원창구에 2015년 1월 1일 - 2015년 6월 31일에 공개 접수된 민원 1336건의 데이터에 대한 코퍼스분석을 통해 민원에서 주로 나타나고 있는 구체적인 문제에 대한 객관적이고 효율적인 요약 을 통하여 시민들의 이해관심을 파악하였다. 정책문제와 관련하여 텍스트 데이터를 대상으로 코퍼스 분석을 활용한 선행연구로 본 연구에 참고하였다.

Park & Yong(2017)은 텍스트 마이닝을 통하여 한국의 원자력 정책 변화를 파악하고자 한국 에너지 정책을 총괄하는 산업통상자원부 보도자료를 분석하였다. 또한 원자력 정책 보고서 Korea Energy Master Plan (KEMP)과 보도자료에서 나타나는 정책의 변화추이 간의 일관성을 확인하였다. 이를 통해 제3차 에너지 마스터 플랜(2018) 기간에 제기될 정책 이슈를 검토하였다.

Lazard et al.(2017)은 미국 FDA(Food and Drug Administration)의 전자담배 규제 발표 후 시민들의 반응을 살펴보기 위하여 트위터 데이터를 대상으로 토픽 분석을 수행하였다. 트위터를 통해 파악된 대중의 반응은 대체로 부정적이었다. 추출된 9개의 주제에서 FDA의 전자담배 규제가 공중 보건에 도움이 되는지, 건강에 해를 끼칠지 여부, 규제가 신흥 전자담배 시장에 미치는 영향 및 뉴스 공유 노력에 대한 규제 후 초기 반응을 보여주었다.

<표 1> 공공정책 검토를 위한 코퍼스 분석 선행연구

저자	대상정책	수집데이터	분석방법	분석결과
Jeong et al. (2008)	요도가와 댐 건설사업	공적토론 회의록	TF-IDF, T-VSM, MDS	토론 집단 간의 인식차이를 계량화
좌보경 외 (2014)	금연정책	온라인 미디어 기사 및 댓글	기사와 댓글 비교 분석	언론기사의 프레임에 따라 댓글의 유형이 상이
장미화 외 (2016)	캠핑 관련 공공정책	소셜 미디어 텍스트	빈도분석	정부의 캠핑정책에 대한 대중의 인식 파악
이태현 외 (2017)	기장군 해수담수화 시설 가동을 통한 상수도 공급안	온라인 미디어 기사	TF-IDF, 의미단위 분류, 시계열 분석	기장 해수담수화 상수도 공급에 대해 미디어 간에 존재하는 인식차이를 객관화
정하영 외 (2017)	부산시 공공정책 전반	부산시 온라인 민원게시판	TF-IDF, 공기어구조맵	6개월간의 부산시 민원을 대상으로 주요민원 파악
Park & Yong (2017)	Korean nuclear policy change	정책 보도자료	Latect Semantic Analysis(LSA)	Korea Energy Master Plan (KEMP, 2018)관련 보도자료를 분석하여 정책이슈 검토
Lazard et al. (2017)	e-cigarette regulations	Twitter	Topic Analysis	미국 FDA의 전자담배 규제에 대한 시민들의 반응을 파악

정책분야의 대용량 텍스트 분석에 대한 선행연구에서 나타나는 공통적인 문제점은 정책에 대한 시민들의 다양한 의견이 담긴 텍스트 데이터를 효율적으로 수집하기 어렵다는 점이다. 따라서 다양한 연구들이 특정 정책에 대한 시민의견을 집약하기 위하여 미디어 텍스트를 주로 활용하였다. 그러나 미디어 텍스트는 시민들의 여론을 형성하는 것을 목적으로 한다는 점에서 실질적인 시민들의 공적관심사를 반영한다고 보기 어렵다. 또한 직접적인 시민의견이 담긴 민원텍스트를 활용한 연구(정하영 외, 2017)의 경우 해당지역의 정책전반에 대한 시민의견이 집약되어있기 때문에 특정 정책에 초점을 맞추어 개선점을 검토하기에는 적합하지 않다.

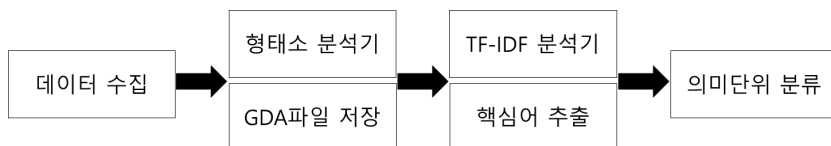
본 연구는 감천문화마을 도시재생정책이라는 특정 주제에 집중하여 다양한 시민의견을 집적 반영하고 있는 구글 지역리뷰데이터를 수집 분석하여 관광정책개선을 위한 결과를 제시하고자 한다는 점에서 미디어 데이터 및 민원데이터를 활용한 선행연구들과 차이점을 가진다.

III. 연구방법

1. 코퍼스 분석 방법론

코퍼스 분석은 자동화된 자연어 처리기술을 이용하여 제한된 시간에 인간의 인지능력으로 전부 읽을 수 없을 정도로 많은 양의 텍스트로부터 유용한 정보를 추출하는 것을 목적으로 한다. 이를 위해 하나의 일관된 텍스트가 아닌 불특정 다수 사용자의 다양한 비정형 텍스트를 집약하여 가치 있는 정보를 추출할 수 있다. 예를 들어 뉴스 콘텐츠 분석을 통해 추가예측의 중요정보로 활용하는 지능형투자 의사결정 모형에 관한 연구(김동영 외, 2014), 뉴스 투자주식의 가격지수를 이용한 미래예측(김유신 외, 2012), 트위터 데이터를 이용하여 국내대학의 인식 및 선호도 분석(양민혁 외, 2015), 2012년 한국대선에 관한 분석(배정환 외, 2013) 등 특정 사회현상 및 상품과 서비스의 평판을 분석하는 일에 많이 활용된다. 특히 상품과 서비스의 평판을 파악하는 것은 관광산업과 마케팅 분야에서는 중요한 과정이다. 따라서 감천문화마을의 관광정책 수립을 위한 시사점을 제시하는 데 있어서 코퍼스 분석 방법론은 매우 유효하다.

코퍼스 분석은 <그림 1> 과 같은 과정으로 실시되었다. 데이터 수집단계에서는 구글맵의 ‘감천문화마을’ 페이지에 작성된 평점과 리뷰데이터 1,503건을 전수조사 하였다. 이 중 한국어, 영어, 중국어로 작성된 1,405건의 데이터를 대상으로 동일한 코퍼스 분석을 위하여 479건의 중국어리뷰와, 384건의 영어리뷰를 한국어로 번역한 기초데이터를 작성하였다. 한국어분석기 사용에 앞서 영어와 중국어로 작성된 리뷰의 한국어 번역을 실시하고 한국어 문맥의 타당성을 파악하는 과정이 필요하다. 본 연구에서는 479건의 중국어 리뷰와, 384건의 영어 리뷰에 대한 번역과 역번역을 실시하였다. 역번역은 번역의 정확성과 타당성을 높이는 방법으로 역번역을 수행하여 번역된 테스트가 원문의 내용과 유사한지 검증할 수 있다(천버이 외, 2017). 분석을 위하여 각 언어에 능통한 전문번역사에 의뢰하여 한국어로 번역하고, 한국어를 다시 역번역하여 영어와 중국어로 나타내고 원문과 비교·수정하여 번역의 오류를 최소화하였다. 한글형태소 분석과 불용어 처리를 위한 전처리 단계는 오픈소스 형태소 분석기 MeCab를 자체 수정한 한글형태소 분석기를 이용하였다. 그 다음 펄(Perl) 언어로 작성된 TF-IDF분석기를 이용하여 핵심 단어를 추출하는 과정을 거쳤다.



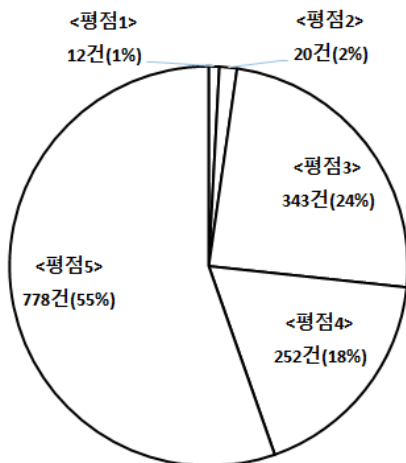
<그림 1> 데이터 수집 및 분석절차

2. 구글 지역리뷰 데이터 수집 및 기초분석

감천문화마을 관광객의 경험과 인식을 파악하기 위해 구글맵에 게시된 감천문화마을에 대한 리뷰 텍스트를 웹 스크래핑(Web scraping)을 통해 2018년 3월 10일부터 2018년 3월 31일까지 총 1,503건의 데이터를 수집하였다. 구글맵의 리뷰데이터는 1에서 5사이의 평점을 포함하고 있다. 전체 총량이 1,503건인 수집데이터 가운데 평점만 표시하고 텍스트를 포함하고 있지 않거나, 한국어, 중국어, 영어 이외의 언어로 작성된 리뷰 98건을 제외한 1,405건을 분석 데이터로 활용하였다. 구글맵이 제공하는 지역리뷰는 특정 지역에 대하여 다양한 국가의 방문객들이 자유롭게 방문경험을 작성하는 열린 플랫폼을 제공하고 있다. 구글맵이 제공하는 텍스트가 감천문화마을의 모든 방문자들의 경험을 대표한다고는 볼 수 없지만, 자발적으로 자유롭게 작성되었다는 점과 해당지역의 방문경험에 대한 1,000건 이상의 일관된 형식의 텍스트 데이터를 제공하고 있다는 점에서 분석 자료의 가치를 가진다.

<그림 2>는 수집된 데이터 1,405건의 지역에 대한 평점 비율을 원그래프로 나타낸 것이다. 지역리뷰 1,405건의 평점분포 중 좋은 관광경험을 나타내는 5점이 전체의 55%, 4점이 18%를 차지하고 있다. 반면, 3%가 1과 2의 부정적인 평점을 남겼으며, 24%가 중립적인 평가를 남겼다.

<그림 3>은 수집된 데이터의 리뷰 작성언어의 비율을 나타낸 것이다. 한국어 39%(542건), 영어 27%(384건), 중국어34%(479건)의 비율로 이루어져 있다. 수집된 데이터의 61%가 영어와 중국어를 사용하고 있다. 이는 표본의 추출이 하나의 웹서비스사용자로 이루어진 한정적인 데이터 이지만, 적극적인 체험경험을 남기는 방문객 중에 외국인이 국내 방문객 숫자를 넘어섰다는 점에서 향후 관광정책 수립에 중요한 지표가 될 수 있다. 특히 중국어권과 영어권 방문객의 인식의 차이를 파악하는 과정은 향후 감천문화마을의 관광정책 수립에 중요한 시사점이 될 수 있다.



<그림 2> 지역리뷰 평점비율



<그림 3> 리뷰작성에 사용된 언어 비율

수집된 데이터에서 사용언어에 따라 평점의 분포가 다르게 나타나는가를 파악하기 위하여 <표 2>에 서와 같이 교차분석을 실시하였다. 카이제곱 검정 결과 언어에 따른 평점의 분포가 서로 상관이 없고 독립적임을 파악할 수 있다(<표 3> 참조). 이 결과는 리뷰작성에 사용된 언어에 따라 감천문화마을 방문경험을 평가하는 평점의 차이가 있다는 것을 의미한다.

<표 2> 언어별 평점

평점	한국어	%	영어	%	중국어	%	합계	%
1	9	1.7	1	0.3	2	0.4	12	0.9
2	17	3.1	3	0.8	-	0.0	20	1.4
3	272	50.2	35	9.1	36	7.5	343	24.4
4	-	-	101	26.3	151	31.5	252	17.9
5	244	45.0	244	63.5	290	60.5	778	55.4
합계	542	100	384	100	479	100	1405	100

<표 3> 카이제곱 검정

	값	자유도	점근 유의확률(주요 특성양측검정)
Pearson 카이제곱	491.134a	16	.000
likelihood 수요 특성우도비)	568.880	16	.000
유효케이스 수	1405		

<표 2>에 나타난 바와 같이 한국어 사용자의 경우 높은 평점 4의 경우는 0건으로 나타났으며, 가장 높은 점수인 5점의 비율이 45%로 나타났다. 반면 중립적인 평가인 3점이 50.2%를 차지하고 부정적인 평가인 2점 이하가 4.8%를 차지하였다. 중국어 사용자는 92%, 영어사용자는 89.8%가 평점 4와5로 긍정적인 평가를 남기고 있다. 이를 비교해 봤을 때 한국어 사용자는 외국어 사용자에 비해 상대적으로 낮은 평가를 하고 있음을 알 수 있다.

3. 형태소 분석

한글 텍스트로 이루어진 비정형데이터를 대상으로 하는 코퍼스 분석은 자연어 처리기술을 토대로 이루어진다. 비정형데이터는 다양한 양식과 형태로 분산된 장소에 존재하는 정형화 되지 않은 데이터를 말한다. 본 연구에서는 구글 지역리뷰 페이지에서 수집한 비정형텍스트 데이터의 문장을 자연어처리하기 위한 형태소분석을 실시하였다. 형태소분석 과정에는 현대한국어분석을 위해 수정한 오픈소스 형태소분석기 MeCab를 사용하였다. 형태소 분석기는 텍스트 덩어리를 어노테이션(수요 특성Annotation)하는 과정에 사용된다. 어노테이션은 텍스트를 구성하는 각 단어를 ‘명사’, ‘조사’, ‘동사’ 등의 형

태소로 분해하여 각 형태소 단위마다 XML수요 특성(Extensible Markup Language)태그를 다는 과정이다. 본 연구에 사용된 형태소분석기는 명사를 고유명사, 의존명사, 보통명사로 구분하고 조사, 어간·어미, 접두·접미사를 구분하는 태그달기를 통해 기계적 분석이 가능한 메타텍스트 데이터로 변환한다. 한국어 542건, 중국어 479건, 영어 384건 으로 수집·분류·번역된 3개의 텍스트 덩어리에 대해 어노테이션을 실시하여 단어마다 태그가 포함된 XML을 기초로 한 GDA수요 특성(Global Document Annotation)파일 형태로 저장하였다.

4. TF-IDF 분석

감천문화마을을 방문하는 불특정 다수의 인식을 나타내는 핵심단어를 추출하기 위한 방법론으로 TF-IDF수요 특성(Term Frequency Inverse Document Frequency)분석을 실시하였다. TF-IDF법은 특정 텍스트의 단어사용 횟수와 일장에서 많이 쓰이는 평이한 단어의 집합인 비교텍스트와의 비교를 통하여 분석텍스트에서 특별히 많이 사용되는 단어 값을 분석의 요소로 사용하는 방법이다(수요 특성(Salton & McGill, 1983)). TF-IDF법을 통해 핵심단어 추출의 정확도를 높일 수 있다. <식 1>을 통해 수집된 텍스트를 형태소로 나눈 GDA메타데이터의 TF-IDF값을 계량한다.

$$TF-IDF(w, a) = TF(w, a) * IDF(w)$$

$$IDF(w) = \log \left(\frac{N}{DF(w)} + 1 \right)$$

TF(w, a) = Number of occurrences of term w in document a

DF(w) = Number of documents containing term w

N= Total number of documents

<식 1> TF-IDF값

IV. 감천문화마을 방문객 리뷰 코퍼스 분석결과

감천문화마을 관광객이 관심을 가지는 의미 있는 정보를 파악하기 위하여 형태소 분석을 통해 가공한 1,405건의 메타 데이터를 대상으로 TF-IDF값을 추출하였다. 데이터로부터 추출된 키워드의 TF-IDF값이 클수록 관광객들이 공통적인 인식을 가지고 있음을 의미한다. 외국인의 방문이 급증하고 있는 감천문화마을의 관광산업에서 향후 관광정책 수립을 위한 시사점을 도출하기 위하여 사용언어별로 감천문화마을에 대한 공통적인 인식을 파악하였다. 이를 위하여 한국어 542건, 중국어 479건, 영어 384건으로 분류·번역되어 형태소 분석을 완료한 3개의 GDA파일을 사용하여 사용언어에 따른 TF-IDF 분석을 실시하였다. <표 4>는 불용어 및 유사단어 처리를 실시하여 집계한 언어별 TF-IDF 분석 결과이다. <표 4>는 각 언어 사용자별로 나타나는 핵심단어사용을 통해 인식의 공통점과 차이점을 보여주고 있다.

〈표 4〉 사용언어별 코퍼스 키워드 분석결과

	한국어		영어		중국어	
	키워드	TF-IDF값	키워드	TF-IDF값	키워드	TF-IDF값
1	추억	250.3	카페	250.2	독특	196.4
2	블거리	230.9	버스	180.0	사진	178.1
3	사진	207.2	사진	178.1	상점	158.3
4	풍경	185.0	풍경	178.1	쇼핑	152.3
5	벽화	148.2	상점	177.3	풍경	139.6
6	어린왕자	133.0	다채	163.2	화려	132.5
7	언덕	120.1	언덕	153.5	산책	124.0
8	사람	108.6	골목길	149.6	어린왕자	120.9
9	동네	108.6	집	143.7	지도	106.4
10	추천	97.4	벽화	137.2	촬영	94.2
11	느낌	94.3	사랑	129.3	적합	93.2
12	관광객	78.9	예술	126.8	관광객	84.6
13	야경	74.3	재미	122.9	카페	82.6
14	버스	73.0	산책	101.5	색상	80.8
15	집	71.9	시간	99.6	시간	80.2
16	먹거리	68.8	가치	95.9	가치	79.9
17	카페	66.9	독특	89.0	예술	79.3
18	주차	64.0	지도	84.2	버스	68.1
19	골목길	59.9	화려	83.7	언덕	66.5
20	입구	56.4	어린왕자	76.1	벽화	59.5
21	경사	55.3	촬영	67.3	편안	45.1
22	주민	55.1	색상	60.6	사람	44.9
23	연인	54.4	먹거리	58.4	블록	44.6
24	한곳	52.0	사람	56.7	산	44.3
25	불편	46.6	계단	54.4	슈퍼	44.3
26	지도	44.3	착용	53.9	피곤	44.3
27	모습	44.1	수집	49.4	스타일	38.3
28	택시	39.5	홀롱	46.6	날씨	36.3
29	산책	39.5	신발	46.1	매력	34.6
30	가게	38.0	구입	45.3	느낌	34.4
31	친구	37.5	아무것	44.6	보물	31.7
32	소소	37.1	정류장	44.6	추천	30.8
33	줄	36.5	레스토랑	44.6	버전	30.2
34	상업적	34.6	역사	42.2	창조	29.1
35	색상	34.1	주변	42.0	재미	28.7
36	외국인	34.0	추천	41.0	수집	27.4

세 언어 사용자 리뷰의 핵심단어 추출로 부터 나타나는 공통적 인식은 세 언어에서 모두 높은 TF-IDF값을 나타내는 ‘사진’과 ‘풍경’이라는 단어를 통해 파악할 수 있다. 이는 다른 관광지와 비교해서 특별히 좋은 사진을 촬영하는 장소라는 인식을 보여주는 것으로 해석할 수 있다. 또한 감천문화마을의 주된 관광행동이 기념사진촬영이라는 것을 시사한다. 또한 어린왕자 조형물과 벽화와 같은 공공예술작품에 대한 집중된 관심은 모든 언어사용자로부터 공통적으로 나타난다. 한국어에서는 벽화가 4번째(TF-IDF 스코어 148.2), 어린왕자가 5번째로(TF-IDF 스코어 133.0)으로 높게 나타났으며, 영어에서는 벽화가 10번째(TF-IDF 스코어 137.2), 어린왕자가 20번째(TF-IDF 스코어 76.1) 중요단어로 나타났다. 중국어도 마찬가지로 어린왕자가 8번째(TF-IDF 스코어 120.9), 벽화가 20번째(TF-IDF 스코어 59.5) 중요 단어로 나타났다. 감천문화마을 리뷰에서 나타나는 ‘어린왕자’는 마을의 언덕 면에 접한 공공도로에 바다 쪽을 향해 설치된 어린왕자 조형물이다. 실제 감천문화마을 현장에서는 어린왕자 조형물에 나란히 앉아서 사진을 찍기 위해 차를 기다리는 긴 줄이 만들어지고 있다. 즉, 한국어에 나타나는 ‘줄’(33번째, TF-IDF 스코어 36.5)이라는 단어는 어린왕자와 관련된 핵심어이다. ‘벽화’는 감천문화마을 재생초기에 공공지금의 지원을 통해 그려진 벽화로 초기의 감천문화마을에 관광객의 관심을 끌게 된 계기가 되었다. 이러한 벽화와 공공예술작품은 국적에 관계없이 여전히 관광객의 중요한 관심사가 되고 있음을 알 수 있다.

‘버스’, ‘정류장’, ‘택시’, ‘도보’, ‘주차’, ‘신발’, ‘경사’, ‘언덕’ 등 접근성에 대한 단어들이 높은 점수를 보이고 있다. <표 4>에서 나타난 상위 36개의 핵심단어에 ‘주차’와 ‘택시’는 한국어에서만 나타나고 있는 점이 차이를 보인다.

감천문화마을에서 방문객들의 관심이 강하게 나타나는 단어 중에는 ‘집’(한국어 TF-IDF 스코어 71.9, 영어 TF-IDF 스코어 143.7)이 있다. 감천문화마을의 주택은 종교 공동체가 집단이주한 역사적 배경으로 인해 같은 크기와 비슷한 모양의 주택이 줄지어 건설되었다. <표 4>에서는 이러한 주택의 집합 형태에 국내외 방문객이 모두 강한 인상을 받고 있음을 알 수 있다. 즉, 집의 집합형태와 구조가 평범함을 벗어난 광경을 자아내는 일종의 관광컨텐츠가 되고 있는 것이다. 감천문화마을을 페루의 ‘마추픽추’에 비유하는 관광마케팅 역시 이러한 지형과 주택의 형태의 특이한 조화를 기반으로 이루어 졌다. 그러나 유사한 지형의 마을은 부산의 산복도로 주변에 많이 존재 하고 있으나 모두가 감천문화 마을과 같은 관심을 받는 것은 아니다. 감천 문화마을의 집들이 위에서 설명한 바와 같은 특징을 유지한 채 현대에 이르게 됨으로써 관광컨텐츠의 기능을 하고 있다고 보는 것이 타당하다.

핵심단어 추출에서 나타나는 세 언어사용자 간의 두드러진 인식의 차이점은 한국어의 ‘추억’(TF-IDF 스코어 250.3)이라는 단어와 영어 사용자(TF-IDF 스코어 89.0)와 중국어 사용자(TF-IDF 스코어 196.4)의 ‘독특’이라는 단어에서 나타난다. 이른 감천문화마을의 풍경을 접할 때 한국어 사용자는 익숙함을 내포한 추억과 향수, 과거의 생활을 느끼는 반면, 영어와 중국어 사용자에게는 생소함을 내포하는 ‘독특’하다는 느낌을 주고 있다. 즉, 감천문화마을은 한국인에게는 익숙함을 외국인에게는 생소함을 주는 관광지임을 알 수 있다. 이러한 점은 문화를 기반으로 하는 민속촌과 같은 관광지에서 흔히 나타나는 속성이다. 감천문화마을은 역사마을이나 민속촌과 같이 인위적으로 조성된 관광지가 아님에도 유사한 관광지 속성을 가지게 되었음을 알 수 있다.

한국어 사용자의 핵심단어 중에 ‘한곳’(24번째, TF-IDF 스코어 52.0)이라는 단어가 있다. 이 단어가 사용되는 맥락이 내포하는 의미는 감천문화마을의 관광서비스가 ‘좁은 한 곳’에 집중되어 있다는

것이다. 같은 모양으로 줄지어 있는 집들과 바다가 어우러지는 풍경은 가장 높은 한 지점에서 모두 감상이 가능하다. 따라서 감천 문화마을의 범위는 넓지만 대부분의 관광서비스는 전체를 조망하기 편리한 시설이 주로 위치한 마을입구로부터 500미터 내외에 집중되어 있다. 힘들게 감천문화마을을 방문한 방문객 중에는 마을을 폭넓게 둘러보며 시간을 보내고 싶지만 관광서비스는 한곳에 집중되어 있어 불편함을 느꼈음을 집약하는 단어이다.

한국어 사용자의 핵심단어에는 ‘외국인’을 언급하고 있다. 외국인 관광객이 대폭증가 하였음을 인지하는 맥락으로 사용된 단어다. 영어사용자의 경험을 나타내는 단어인 ‘신발’, ‘착용’은 가급적 편안한 신발을 착용한 상태로 방문할 것을 권하는 맥락에서 사용되었다. 영어와 중국어에서 공통적으로 나타나는 ‘지도’는 입구에서 외부인들 대상으로 의무적으로 구입하도록 요구하는 마을지도에 대한 각자의 생각을 적고 있다. 감천문화마을의 지도판매는 지역재생의 결과를 대다수의 주민들과 나누기 위한 정책으로 운영되는 중요한 수익창출모델이므로 지역관광이 활성화 되는 과정에서 주민들이 겪는 불편과 어려움을 잘 설명하고 지도판매의 취지를 설명하는 노력이 더욱 요구된다.

V. 감천문화마을 관광정책의 시사점

1. 감천문화마을 관광체험요인 파악을 위한 의미단위 분류

감천문화 마을의 관광산업은 매년 가파르게 성장하고 있다. 부산시는 감천문화마을 방문객 수를 2017년 150만 명, 2018년 약 250만 명 규모로 추산하였으며, 2020년에도 지속적인 증가를 예상하고 있다. 해당 자치구청의 추산에 따르면 한해 약 100만 명의 관광객이 증가했다. 방문객수의 증가만으로도 현재의 감천문화마을은 원활한 지역관광운명을 위한 지역재생정책의 개선이 시급히 요구되는 시점에 이르렀다. 국내관광산업의 체험경제이론 적용에 대한 장서희(2017,2018), 천덕희(2018), 장운연과 서원석(2014) 등의 선행연구 들은 체험에 대한 가치 증가가 관광객의 소비를 활성화하는데 기여한다는 분석결과와 같은 의미 있는 시사점을 제공하고 있다. 즉, 체험경제이론은 관광활성화 정책을 수립하는데 있어서 유용한 분석의 틀을 제공한다. 그러나 기존 관광분야의 체험경제연구는 연구자가 작성한 정형화 된 질문지에 대한 조사대상의 응답을 통해 체험경제 요소를 계량하고 다른 지불의사, 만족감, 몰입감과 같은 다른 관광의 중요 요소와의 상관관계를 분석하는데 중점을 두고 있다.

방문객 리뷰를 대상으로 TF-IDF분석을 통해 도출한 핵심단어는 방문객이 강한 인상을 받은 주요 체험요인을 파악하기는 용이하지만, 다양한 인식과 관심사로 인해 관광산업정책 도출을 위한 결과해석에는 어려움이 따른다. 따라서 분석결과를 체험경제이론의 4가지 구체적인 체험요소를 바탕으로 의미단위로 분류함으로써 관광활성화 정책시사점 도출에 활용하였다. 체험경제요인에 따른 의미단위 분류는 감천문화마을 방문객이 경험하는 구체적, 개별적인 하위 요소들을 파악하게 함으로써 관광활성화 정책의 세부적인 개선사항과 방향을 제시할 수 있다. 더불어 선행연구 들이 제시하고 있는 체험경제이론의 시사점과 접목하여 지불의사 및 만족감, 몰입감을 제고하기 위한 정책으로 이어질 수 있다.

TF-IDF 분석결과를 의미단위로 분류하기 위한 관광체험요소를 구성하기 위하여 Pine &

Gilmore(1999)의 체험경제에서 사용된 오락적, 교육적, 심미적, 일탈적 요인의 4가지에 부정적인 체험을 추가하여 5가지 체험요인으로 구분하였다. 설문지 설계를 위해 주로 사용되었던 4가지 체험요인에는 부정적 체험에 대한 조사가 포함되어 있지 않다. 그러나 방문객의 리뷰 데이터와 같은 조사 설계과정이 없는 비정형데이터에는 부정적인 체험이 포함되어 있다는 점을 감안하였다.

2. 의미단위 분류결과와 시사점

코퍼스 분석결과로 나타난 <표 4>의 핵심단어가 리뷰 내에서 주로 사용된 맥락에서 의미를 파악하여 5가지 체험요인으로 구분하였다. 그리고 각 단어가 갖는 TF-IDF 값의 언어별 상위 36개 단어를 대상으로 의미단위별 합산을 통해 감천문화마을 방문객 리뷰에서 나타나는 각 체험요인을 계량화 하였다. <표 5>는 <표 4>의 TF-IDF 값으로 정렬된 단어를 기준으로 각 단어가 사용된 맥락을 파악하여 체험요인 분류에 따른 의미단위 분류를 실시한 결과를 나타낸 것이다. 상위 36개 단어 중 체험경제요인 4가지와 부정적 체험 중 어디에도 분류할 수 없는 단어는 미분류로 처리하였다.

의미단위의 분류과정은 핵심단어가 포함된 문장을 수집·번역한 1,405건의 텍스트로부터 검색 추출하여 문장을 5가지 체험요소의 하나로 분류하는 과정을 통하여 실시하였다. 문장의 분류는 감천문화마을 지역재생 관련 연구자 3명의 일치되는 의견체택을 통해 실시하였다. 오락적 체험은 마을에서 오감을 자극하거나 촉진하여 유희와 재미를 느끼게 하는 요소를 분류하였다. 심미적 체험은 풍경, 언덕, 야경 등과 같이 친근히 흡수되며 수동적 반응을 나타내는 요소를 분류하였다. 일탈적 요소는 기존 일상의 경험을 벗어난 체험과 일상의 경관을 벗어난 독특한 경관, 그리고 관광객의 예상을 벗어난 체험을 분류하였다. 부정적 체험은 ‘택시’, ‘버스’등과 같이 방문객이 불편함을 호소하며 리뷰에서 부정적인 경험으로 언급하고 있는 사안을 분류하였다. 교육적 체험은 정보를 습득함으로써 자극되는 인식을 분류하였다. 단어의 분류를 통해 각 체험요인을 구성하는 핵심단어를 파악할 수 있으며, 체험요인별로 계량된 수치를 통해 각 언어별로 우세한 체험요인과 부족한 체험요인을 파악할 수 있다.

<표 5> 체험요인 의미단위별 분류

체험요인 의미단위	한국어 TF-IDF 값 상위 36개 단어	영어 TF-IDF 값 상위 36개 단어	중국어 TF-IDF 값 상위 36개 단어			
오락적 체험	불거리	230.9	카페	250.2		
	사진	207.2	사진	178.1	사진	178.1
	벽화	148.2	상점	177.3	상점	158.3
	어린이왕자	133.0	벽화	137.2	쇼핑	152.3
	추천	97.3	재미	122.9	산책	124.0
	먹거리	68.8	산책	101.5	어린이왕자	120.9
	카페	66.9	어린이왕자	76.1	촬영	94.2
	연인	54.4	촬영	67.3	카페	82.6
	산책	39.5	먹거리	58.4	벽화	59.6
	가게	38.0	작용	53.9	슈퍼	44.3
	친구	37.5	수집	49.4	재미	28.7
	소소	37.1	구입	45.3	수집	27.4
			레스토랑	44.6		
		1061.5	1362.2	1070.4		
심미적 체험	추억	250.3	풍경	178.1	풍경	139.6

	풍경 185.0 언덕 120.1 동네 108.6 느낌 94.3 야경 74.3 주민 55.1 모습 44.1 색상 34.1	다채 163.2 언덕 153.5 사랑 129.3 예술 126.8 가치 95.9 화려 83.7 색상 60.6 계단 54.4 홀름 46.6	화려 132.5 색상 80.8 가치 79.9 예술 79.3 언덕 66.5 산 44.3 스타일 38.3 매력 34.6 느낌 34.4 보물 31.7 창조 29.1
	965.9	1092.1	791
일탈적 체험	집 71.9 골목길 59.9 외국인 34.0	골목길 149.6 집 143.7 독특 89.0	독특 196.4 블록 44.6
	165.8	382.3	241
부정적 체험	버스 73 주차 64.0 불편 46.6 경사 55.3 한곳 52.0 택시 39.5 상업적 34.6 지도 44.3 줄 36.5	시간 99.6 지도 84.2 버스 68.1 신발 46.1 아무것 44.6 정류장 44.6	지도 106.4 시간 80.2 버스 68.1 피곤 44.3
	445.8	387.2	299
교육적 체험		역사 42.2	
	0	291.2	0
미분류	사람 108.6 추천 97.3 관광객 78.9 입구 56.4	사람 56.7 주변 42 추천 41	적합 93.2 관광객 84.6 편안 45.1 사람 44.9 날씨 36.3 추천 30.8 버전 30.2
	341.2	139.7	365.1
합계	2980.2	3405.7	2766.5

관광체험경제이론에 따른 의미단위 분류와 TF-IDF 값 계량을 정리한 결과를 기반으로 도출한 감천 문화마을의 관광활성화정책 시사점은 다음과 같다.

첫째, 감천문화마을의 관광은 오락적 체험이 강한 반면 오락적 체험요소를 구성하는 요인은 공공미술, 기념사진촬영, 카페, 먹거리, 산책 등으로 한정적이며 지역재생을 위한 재정마련에 도움이 되는 요소가 부족하다. 지역 관광산업을 활성화하기 위해서는 재정사업의 발굴과 재정효율성의 검토가 중요한 사안이다. 따라서 의미단위 분류를 통해 나타나는 오락적 체험요소를 토대로 재정사업 확충과 관광객의 소비지출을 위한 사업다각화 전략이 요구된다.

둘째, 심미적 체험이 오락적 체험 다음으로 높은 점수를 보이고 있다. 심미적 체험을 만드는 요인은 풍경, 색상, 언덕, 야경, 다채와 같은 지역의 지리 환경적 조건과 마을의 도시구조원형에서 부터 나타난다. 따라서 감천문화 마을의 매력을 유지하기 위해서는 기존 건축물의 구조와 형태를 보존하기 위한 세부적인 규제정책과 마을에 포함된 집의 색상을 다양하고 화려하게 유지할 수 있는 외관 지원정책이 병행되어야 한다.

셋째, 일탈적 체험에서는 외국인 방문객의 점수가 높게 나타나고 있다. 이는 감천문화마을의 집, 건축물, 골목길, 취락형태가 외국인에게는 독특한 인상을 주기 때문이다. 특히 같은 크기 '집'의 집합형태가 중요한 콘텐츠로 작용하고 있다. 따라서 기존 주택의 형태 보존 및 유지보수를 위한 지원과 증개축과정에서 보존되어야만 하는 요소를 지정하여 변경을 제한하는 규제정책이 요구 된다.

넷째, 상위 36개 단어에서는 교육적 체험을 적고 있는 텍스트와 단어가 영어에서 '역사'하나만 나타나고 있다. 감천문화마을에서는 작은 공방들을 중심으로 운영되는 만들기, 체험교실 등의 콘텐츠가 많지만 이에 대한 언급은 분석과정에서 나타나지 않았다. 교육적 체험요소역시 관광활성화 전략에서는 중요한 구성요소이므로 이에 대한 구상과 정책보완이 필요하다.

다섯째, 부정적인 체험이 세 번째로 높은 점수를 나타내고 있다. 해당 단어가 의미하는 바를 파악하고 이에 대한 개선이 실행되어야 한다. 한국어에서 부정적 체험을 언급하는 핵심단어가 외국어에 비해 많이 나타나고 있으며 TF-IDF 값 또한 높게 나타났다. 외국어에서 나타나지 않는 '주차'에 대한 부정적 체험이 많으며, 마을이 '상업적'이라는 인식이 나타나고 있다. 또한 마을의 관광 콘텐츠가 '한 곳'에 집중되어 있는 것과 사진촬영을 위한 '줄'이 긴 것이 부정적 체험으로 나타났다. 외국어의 부정적 요인은 교통, 짧은 체류 시간에서 주로 나타나고 있다. 따라서 교통개선을 위한 공공정부와의 협의가 요구된다. 짧은 체류 시간은 관광콘텐츠를 마을 전반으로 확산시켜가는 장기적인 빈집 활용 지원 정책 등을 통해 개선이 가능할 것이다. 또한 장기적으로 체류시간을 늘이기 위한 정책과제공모 등을 추진할 필요가 있다.

VI. 결 론

지역정책은 동태적 속성을 가진다. 인구의 변화, 생활패턴의 변화, 산업의 변화, 도시구조의 변화는 멈추지 않고 매일 매일 조금씩 이어지기 때문이다. 본 연구의 대상인 감천문화마을과 같은 도심의 작은 지역은 작은 변화도 눈에 두드러지게 나타난다. 따라서 감천문화마을과 같이 역동적인 지역의 재생정책을 효율화하기 위해서는 빠른 정책수요파악과 정책형성과정이 필요하다. 앞서 설명한 바와 같이 감천문화마을 재생은 관광활성화 정책의 성공여부가 중요한 영향을 미친다. 마을을 찾는 외래 관광객의 증가를 목적으로 하는 공공지원과 마을을 방문한 관광객의 소비를 통한 관광산업의 이익 창출이 점점 중요성을 가지게 될 것이다. 따라서 본 연구에서 논하는 관광활성화 정책의 일차적인 수혜대상은 지역 관광을 운영하고 있는 지역주민이다. 이를 위하여 국내방문객과 외국인 방문객의 관광경험을 파악하고 관광정책을 개선하기 위한 시사점을 제공하고자 하였다.

본 연구는 관광활성화를 위한 정책수요 파악의 관점에서 감천문화마을 방문객의 체험과 인식을 집약하기 위하여 방문객이 자발적으로 작성한 구글 지역리뷰 텍스트를 수집하고 분석하였다. 이러한 분석

결과를 관광정책에 효율적으로 반영하기 위한 이론적 토대로 체험경제 이론을 활용하였다.

관광산업의 체험경제 이론과 TF-IDF 값을 기반으로 도출한 향후 감천문화마을의 관광정책에 대한 시사점은 다음 네 가지로 요약가능하다. 첫째, 장기적으로 오락적 체험요소를 강화할 수 있는 콘텐츠 마련이 우선적으로 이루어져야한다. 둘째, 마을의 매력을 지속적으로 보존하기 위해서는 기존 건축물의 구조와 형태를 보존하기 위한 세부적인 규제정책과 마을에 포함된 집의 색상을 다양하고 화려하게 유지할 수 있는 외관 지원정책이 병행되어야 한다. 셋째, 집과 골목길이 외국인에게는 일탈적 체험요인으로 작용하므로 이를 활용하는 콘텐츠 및 기념품 개발이 필요하다. 넷째, ‘주차’문제, ‘상업적’ 장소로 변하고 있다는 인식의 확산, 마을의 관광 콘텐츠가 ‘한 곳’에 집중되어 있는 것, ‘교통’의 불편 등 부정적인 의견에 대한 개선이 이루어져야 한다.

분석을 통해 알 수 있는 감천문화마을 관광서비스의 가장 핵심적인 한계점은 사진, 벽화, 어린왕자 등 공공예술작품을 중심으로 나타난 오락적 체험요소로 부터 파악가능하다. 이와 같이 공공예술 작품이 중심이 되는 관광 콘텐츠는 지역에 방문객을 유입시키는 중요한 요인이 되지만 직접적인 수익을 발생시키지 않는다는 문제를 내포하고 있다. 수익을 창출하는 콘텐츠의 키워드가 나타나지 않는 사실을 통해 현재의 감천문화마을 관광서비스가 지역주민의 이익을 창출하는 단계로 발돋움하지 못하고 있는 문제의 원인을 알 수 있다. 이러한 문제를 해결하기 위해서는 연령별, 국가별, 지역별 방문객의 특성을 파악하고 산업관점의 서비스 마련과 마케팅 대상을 구체화 하는 과정이 필요하다. 마케팅 대상을 구체화 하여 특화된 서비스를 점차 늘려 감으로써 현재 자영업의 밀집 수준의 관광서비스를 탈피하고 체계적인 지역관광산업이 성장해 갈 수 있다.

본 연구는 지역거버넌스 중심의 도시재생정책을 개선하기 위한 논의에서 하나의 사례로 감천문화마을 재생을 위한 향후의 정책 방향을 검토하였다. 이 과정에서 코퍼스 분석 방법론이 공공정책수립의 기초자료를 제공해 줄 수 있음을 본 탐색적 연구를 통해 밝혔다.

본 연구는 감천문화마을 관광객의 인식을 사용 언어로 세분화하여 대표적인 관심사를 파악함으로써 감천문화마을 관광서비스 개선을 위한 시사점을 제공하는 실무적인 기여를 하였다. 구체적으로 정량적인 데이터의 한계를 보완할 수 있는 이해관계자(관광객)의 직접적인 의견을 반영한 정성적인 내용을 토대로 체험경제요인을 분류하고 계량한 데이터를 제시하고 이를 활용하여 도시재생전략과 관광산업활성화 정책과의 연계성을 전달하는 기초자료를 제시하였다. 이는 「2025 부산광역시 도시재생전략계획」 등에 포함된 관광산업활성화 계획이 ‘인구통계’, ‘관광객 수의 증감’, ‘노후 건축물 수’, ‘사업체 수 증감’, ‘지역총생산’ 등과 같은 정성적인 데이터를 근거로 하는 점에 비교하여 현장의 문제를 직접적으로 반영하는 시사점을 제시한다는 점에서 실무적인 의미를 가진다. 또한 관광객이 직접 남긴 관광 수요의 특징을 분석하는 과정을 통해 감천문화마을의 관광활성화를 위한 전략적 접근과 현재의 관광현황 개선을 위한 정책적 제언을 제공하고 있다.

외국인 관광객이 많이 찾는 지역에 대하여 다양한 언어 사용자대상의 표본 추출과 데이터 수집은 많은 시간과 비용이 소모된다. 그러므로 본 연구에서 실시된 다양한 언어 사용자가 자발적으로 남긴 텍스트의 분석은 의미를 갖는다. 또한 관광객 리뷰를 수집하고 코퍼스 분석을 실시함으로써 기존의 조사방법론으로 도출할 수 없었던 지역 활성화를 위한 의미 있는 시사점을 도출하였고, 향후 유사분야에서 본 방법론이 활용되는 후속연구를 유발하였다는 학술적인 공헌도가 있다.

본 연구는 다양한 토론지점이 존재한다. 첫째, 최근 감천문화 마을에서 오버투어리즘의 문제제기가

발생하고 있는 시점에서 관광활성화를 위한 정책 제고는 다양한 관점의 차이가 있을 수 있으며 토론의 여지가 있다. 하지만 본 연구는 관광활성화가 지역의 지속적인 재생동력을 생산하는 가치를 가진다는 관점에서 연구를 진행하였다.

둘째, 분석 데이터로 사용된 구글맵의 지역리뷰는 다양한 외국어 사용자의 리뷰를 생생하게 수집할 수 있다는 장점이 있는 반면에 적극적으로 리뷰를 작성하는 일부 이용자의 의견을 반영한다는 점에서 다양한 의견을 수렴하는데 한계가 있다. 또한 감천문화마을의 관광객에 대한 인식파악이 대표관심사를 나타내는 단어도출에 그침으로써 인식을 구성하는 단어 간의 맥락을 원활하게 파악하기 어렵다는 한계를 가진다. 향후 연구에서는 SNS 등을 통해 데이터 수집을 다각화하기 위한 방법론의 검토와 단어의 연관성을 파악하여 인식을 대표하는 단어들 간의 인과구조를 도출하기 위한 방법론에 대한 연구가 요구된다.

참고문헌

- [1] 국가통계포털, 읍면동별 연령(각세) 및 성별 한국인 현황_ 사하구 (2017)
- [2] 김가령 (2014), 「문화마을 관광체협의 지각된 가치와 공정관광태도, 공정관광행동의도 간 영향관계에 관한 연구」. 『문화정책논총』 28(2), pp.174-196.
- [3] 김동영·박재원·최재현 (2014), 「SNS와 뉴스기사의 감성분석과 기계학습을 이용한 주가예측 모형 비교 연구」. 『한국IT서비스학회지』 13(3), pp.221-233.
- [4] 김유산·김남규·정승렬 (2012), 「뉴스와 주가 : 빅데이터 감성분석을 통한 지능형 투자 의사결정모형」. 『지능정보연구』 18(2), pp.143-156.
- [5] 김형균·김종욱·박상필·서정렬 (2015), 『도시재생 사업지역의 주민생활 및 상권변화 연구』. 부산:부산연구원
- [6] 관광지식정보시스템, 주요관광지점 입장객 통계 (2018)
- [7] 배정환·손지은·송민 (2013), 「텍스트 마이닝을 이용한 2012년 한국대선 관련 트위터 분석」. 『지능정보연구』 19(3), pp.141-156.
- [8] 부산광역시 (2015), 『2025 부산광역시 도시재생전략계획 요약보고서』
- [9] 양민혁·정인선·김용채·조완섭 (2015), 「SNS 데이터를 활용한 국내대학 인식 및 선호도 분석」. 『한국빅데이터서비스학회 논문지』 1(1), pp.1-13.
- [10] 윤지영 (2016), 「도시재생마을 브랜딩 전략 방안」. 『BUSAN DEVELOPMENT FORUM』 159, pp.127-129.
- [11] 이정석 (2013), 「주민주도 지역혁신의 필요성과 성공조건」. 『BDI 정책포커스』 1921-12.
- [12] 이태현·홍순구·정하영·홍혜승 (2017), 「지방정부 공공사업에 관한 일간지 미디어자료의 시계열 코퍼스 분석: 기장 해수담수화플랜트사업을 중심으로」. 『지방정부연구』 20(4), pp.1-23.
- [13] 장미희·윤영일 (2016), 「소셜 미디어 빅데이터 분석을 통한 캠핑에 대한 정부 정책과 국민들의 인식 변화 연구」. 『관광연구』 31(1), pp.91-112.
- [14] 장서희 (2017), 「체험경제지수 수준에 따른 그룹별 인식 및 행동 차이: 대전광역시 방문객을 중심으로」. 『관광학연구』 41(3), pp.85-104.
- [15] 장서희 (2018), 「호스피탈리티와 관광 상품 이득가치와 체험손실가치에 영향을 주는 체험가치 요소들」. 『관광학연구』, 42(8), pp.147-163.

- [16] 장윤영·서원석 (2014), 「외래관광객의 쇼핑 체험요소가 만족도, 정서적 몰입에 미치는 영향: Pine & Gilmore의 체험경제이론을 중심으로」. 『관광학연구』 38(10), pp.199-219.
- [17] 정하영·이태현·홍순구 (2017), 『코퍼스분석과 공적토론연구: 문제 발굴과 의사결정을 위한 빅데이터과학』. 부산:만타.
- [18] 정하영·이태현·홍순구 (2017), 「공공서비스의 수요대응성 향상을 위한 전자민원 코퍼스분석: 부산시 사례를 중심으로」. 『지방정부연구』 21(1), pp.423-436.
- [19] 좌보경·백혜진·서필교 (2014), 「금연정책관련 온라인 뉴스와 댓글 유형의 내용 분석」. 『Journal of Public Relations』 18(3), pp.13~43.
- [20] 천덕희 (2013), 「크루즈여행 체험요인이 여행자의 감정적 반응과 인지적 반응에 미치는 영향: 체험경제 이론과 PA 모델을 중심으로」. 『관광학연구』 37(9), pp.185-206.
- [21] 천버이·한준영 (2017), 「스포츠사회과학 분야 학술지 설문연구와 역번역 사용 동향」. 『한국체육측정평가학회지』 19(3), pp.13-23.
- [22] 최수진·김빈·가정혜 (2020), 「관광경험이 관광지 만족과 충성도에 미치는 영향: 방문경험에 따른 조절 효과를 중심으로」. 『관광경영연구』 24(1), pp.187-211
- [23] 최지영·김현지 (2017), 「감천문화마을의 지역 문화마케팅이 지역 브랜드 자산과 만족도에 미치는 영향」. 『관광연구』 32(1), pp.167-191.
- [24] Baik, Y. (2012), "The significance and effect of the aesthetic experience through public art". Korean Journal of Culture and Arts Education Studies, 7(4), pp.219-241.
- [25] Ellis, G., & Rossman, J. (2008), "Creating value for participants through experience staging: Parks, recreation, and tourism in the experience industry", Journal of Park and Recreation Administration, 26(4), pp.1-20.
- [26] Jeong, H., Shiramatsu, S., Kobayashi, K., & Hatori, T. (2008), "Discourse analysis of public debates using corpus linguistic methodologies", Journal of Computers, 3(8), pp.58-68.
- [27] Lazard, A. J., Wilcox, G. B., Tuttle, H. M., Glowacki, E. M., & Pikowski, J. (2017), "Public reactions to e-cigarette regulations on Twitter: A text mining analysis", Tobacco control, 26(e2), pp.112-116.
- [28] Loureiro, S. (2014), "The role of the rural tourism experience economy in place attachment and behavioral intentions", International Journal of Hospitality Management, 40, pp.1-9.
- [29] Oh, H., Fiore, A. M., & Jeoung, M. (2007), "Measuring experience economy concepts: tourism applications", Journal of Travel Research, 46(2), pp.119-132.
- [30] Park, C., & Yong, T. (2017), "Prospect of Korean nuclear policy change through text mining", Energy Procedia, 128, pp.72-78.
- [31] Pine, B. J., & Gilmore, J. H. (1999), *The experience economy: Work is theatre & every business a stage*. Boston: Harvard Business Press.
- [32] Salton, G., & McGill, J. (1983), *Introduction to Modern Information Retrieval*. NY: McGraw-Hill.
- [33] Stamboulis, Y., & Skayannis, P. (2003), "Innovation strategies and technology for experience-based tourism", Tourism Management, 24(1), pp.35-43.
- [34] 구글맵 감천문화마을 지역리뷰, <http://www.google.co.kr/maps/place/%EA%B0%90%EC%B2%9C%EB%AC%B8%ED%99%94%EB%A7%88%EC%9D%84/@35.0973948,129.0084037,17z/data=!3m1!4b1!4m5!3m4!1s0x3568e9b7a3113a51:0x977e01d7224cf6a3!8m2!3d35.0973904!4d129.0105924?hl=ko>, 검색일자: 2018년 3월 27일 최종방문)

A Study on Urban Regeneration Policy Using Corpus Analysis : Gamcheon Culture Village Tourism Promotion Plan

Lee, Taihun · Jeong, Hayeong · Lee, Sae-Mi

Abstract

The activation of big data analysis is useful in various fields, such as marketing and demand forecasting of companies, but there are a few studies, in the policy evaluation of public institutions. This research is an exploratory study intended to integrate big data analysis technology into the evaluation of urban revitalization policies, and the research aims to propose policies for tourism facilitation in Gamcheon Culture Village. To this end, a total of 1,503 review texts, posted on Google Maps, on Gamcheon Culture Village were collected through web scraping. Among them, 1,405 cases were used as analytical data—except for 98 reviews written in languages other than Korean, Chinese, or English. Thereafter, to extract TF-IDF(Term Frequency-Inverse Document Frequency) values, corpus analysis was performed after carrying out morpheme analysis. The implications for tourism policies were presented by grafting the experience economic theory of the tourism industry onto the results of the corpus analysis. This study is meaningful, as it is an initial study that derived the result that aggregated the core perceptions of visitors according to their tourism experiences through the big data analysis of tourism policy reviews beyond the limitations of existing survey methods.

Keywords: Regional revitalization, Gamcheon Culture Village, Big data analysis, Tourism facilitation, Corpus analysis