

텍스트 마이닝과 콘텐츠 분석을 이용한 COVID 19 전/후 한국에 방문하는 여행객들의 리뷰 데이터 내 프라이버시 비교

우성민 (부산대학교 경영학과 석사생, 주저자 smwoo55@pusan.ac.kr)
김종기 (부산대학교 경영학과 교수, 교신저자 jkkim1@pusan.ac.kr)

… Abstract …

COVID-19 발생은 우리 삶의 모든 영역에 변화를 가져왔다. 봉쇄 기간 동안의 격리와 관련된 조건으로 인해 팬데믹은 관광객의 습관과 행동에 변화를 가져왔다. 본 연구는 관광객들이 가장 많이 거론하는 프라이버시 카테고리, 테마, 속성의 차이를 알아보기 위해 코로나19 발생 전후 한국 명소 리뷰를 기반으로 관광객의 프라이버시 우려를 탐색하는 것을 목적으로 한다. 이 연구는 자동화된 텍스트 분석과 콘텐츠 분석을 결합했다.

데이터는 트립어드바이저(TripAdvisor.com)에서 수집했으며, 한국을 여행하고 14개의 다른 명소를 방문한 관광객들이 영어로 작성한 36,708개의 명소 리뷰로 구성되었다. 자동화된 텍스트 분석과 검증된 개인 정보 사전을 사용하여 개인 정보 보호 문제의 범주, 주제 및 속성을 탐색하여 리뷰를 조사했다. 또한 코로나19 팬데믹 기간 동안 관광객들의 인식 변화를 알아보기 위해 감성과 감정 분석을 진행했다

Key Words : Privacy, Online Tourist Reviews, COVID-19, LIWC, Privacy Concern.

I. 서 론

코로나-19(COVID-19)는 새롭게 발견된 코로나바이러스로 2020년 3월 11일 WHO의 팬데믹 선언 이후 전 세계로 확산되어 팬데믹을 선언했다. 사스(SARS)나 메르스(MERS)와 같은 기존의 전염병과 달리 코로나19는 장기화가 이어지고 있다. 한국 정부 또한 위기경보를 2020년 1월 20일 주의 단계로 격상(국가위기 관리 4단계 중 2단계) 되었고, 이후 2020년 2월 23일 ‘심각’ 단계로 격상(국

가위기 관리 4단계 중 4단계)된 이후 26개월 간 심각 단계를 유지하며 코로나-19로 인해 우리의 일상생활에 많은 변화를 가지고 오게 되었다. 이 변화는 관광에도 직·간접적인 영향을 주게 되었다. 관광객들이 위험을 회피하고 안전과 위생을 우선시하는 경향이 지배적인 현상이 되었다. 안전과 위생은 관광 활동의 기본적 가치이나 코로나19로 인해 관광객들에게 상당히 중요한 의사결정 요소로 자리잡았다. 안전과 위생을 중시하는 심리적·행태적 변화는 코로나-19 이

후에도 지속될 것으로 전망된다. 그리고 인해서 관광객들의 관광 활동의 개별화와 소규모화와 같은 행태 변화를 가져왔다.

이러한 변화로 인해 잠재 관광객들은 방문 전에 온라인 환경에서 얻을 수 있는 정보를 이용하여 해당 관광지를 불확실성을 낮추고 양질의 정보를 얻으려는 경향이 더 커질 것이다. 온라인 구전 정보의 한 형태인 온라인 리뷰는 잠재 관광객들에게 실제 관광경험을 바탕으로 한 관광객들의 만족, 불만족 이유를 알려주는 좋은 정보원천이 되고 있다(이현주, 2017; 오홍철, 2018), 잠재 관광객들은 호텔 및 여행사 등에서 상업적으로 제공되는 정보보다 온라인 리뷰와 같은 소비자 리뷰를 더 선호하고, 신뢰하는 것으로 나타났다(Liu & Park, 2015; Fang, Kucukusta & Law, 2016).

본 연구에서는 온라인 플랫폼에서 제공하는 사용자들의 관광명소 방문 리뷰를 활용하여 각 리뷰가 가지고 있는 속성을 프라이버시, 감정 그리고 감정 분석을 통해서 코로나19가 가져온 관광객들의 프라이버시 염려의 변화와 감정 그리고 감정의 변화를 알 수 있다.

본 연구는 다음과 같은 목적을 가진다.

첫째, 관광객이 작성한 고객 리뷰 데이터를 이용하여 개인정보에 대한 염려가 코로나-19 이후 증가하였는지 파악한다.

둘째, 관광객이 작성한 실제 고객 리뷰 데이터를 이용하여 자동화된 텍스트 분석과 검증된 개인 정보 사진을 사용하여 개인 정보 보호 문제의 범주 (privacy, Norms requisites, Outcome state,

Private secret, Intimacy, Law, Restriction, Open visible)에서 코로나-19 전/후 주요 범주가 변화되었는지 파악한다.

II. 선행 연구

1. 개인정보 및 프라이버시의 개념

Warren과 Brandeis(1890, p. 193)가 “혼자 남겨질 권리”로 개념화한 프라이버시의 정의는 문명화된 사람들이 가장 가치 있는 권리이자 가장 포괄적인 권리로 묘사된다. 여러 연구에서 이 구조의 다양한 측면을 문서화하였다(Pedersen, 1979; Smith et al., 1996; Westin, 1970). Pedersen(1979)은 익명성(군중에서 눈치 채지 못하는 것), 고독성(혼자 있고 타인의 관찰로부터 자유로운 것), 독립성(혼자 있고 타인과 떨어져 있기를 원하는 것), 내성적(남과 이야기하고 싶지 않은 것), 친밀감(가족이나 친구와 단둘이 있는 것)의 프라이버시 차원을 개념화했다.

또한 학자들은 고객의 행동에 영향을 미치는 개인적 심리적 변인으로 프라이버시를 조사하였다(Bitner, 1992; Eastlick et al., 2006). 실제로 잠재적 소비자의 프라이버시 염려를 적시에 평가하는 것은 사업 성공과 장기적인 소비자 만족에 영향을 미친다(Phelps et al., 2001).

초기에는 프라이버시를 “외부의 간섭

이나 침해로부터 벗어나 자유롭게 혼자 있을 수 있는 권리”로 개인에 대한 불가침의 권리로 정의하였고 정보통신기술의 발달로 온라인에서 사용되는 개인정보가 본인의 의지와 관계없이 수집되고 이용되는 프라이버시 침해문제가 생겨남에 따라서 프라이버시가 정보에 대한 프라이버시 개념으로 확대되었다.

2. 프라이버시 염려 개념

일반적으로 프라이버시 염려(Privacy Concern)는 프라이버시 침해로 인한 손실에 대한 개인의 염려를 표현하는 개념이다 (김중기 · 김상희, 2014). 프라이버시 관련 연구에서는 프라이버시 염려를 핵심적인 개념으로 보고, 이의 결정 요인 및 효과 등을 규명하기 위한 연구뿐 아니라 개인의 행동을 이해하기 위한 프라이버시와 관련된 태도 및 행동에 관한 연구가 활발히 진행되어 왔다.

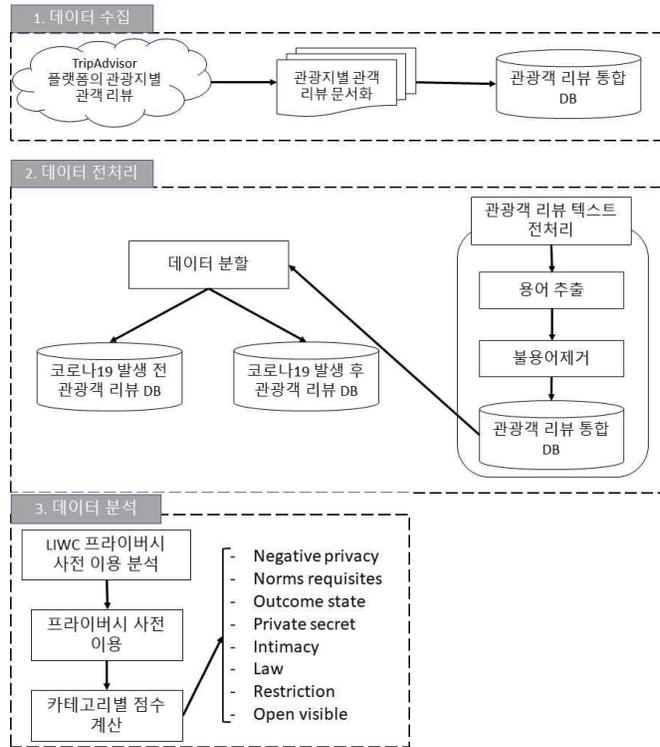
먼저 전통적인 이론으로 볼 수 있는 것으로, 개인의 프라이버시에 대한 염려가 높을수록 프라이버시를 보호하는 행동이 증가하거나 정보를 제공하는 행동이 감소하는 경향이 있다고 간주하는 것이다. 이러한 전통적인 프라이버시 연구는 개인의 행동을 체계적으로 설명하기 위해 합리적 행동이론(TRA, Theory of Reasoned Action)을 제안한 Fishbein & Ajzen (1975)의 연구를 반영하여 ‘태도와 행동의 일관성’을 전제로 하고 있다. 합

리적 행동이론에 따르면, 행동은 행동 의도에 의해 결정되고 행동 의도는 행동에 대한 긍정적인 또는 부정적인 태도에 의해 결정되기 때문에 태도와 행동 간에는 일관성이 존재한다는 것이다.

Ⅲ. 연구모형과 가설

1. 프레임워크 구성

본 연구는 코로나-19 전후 관광객들의 리뷰 데이터를 활용하여 자동화된 텍스트 분석을 위해 개발된 검증된 프라이버시 사전, 감성 분석 및 감정 분석을 이용한 프레임워크를 제안한다. 데이터 프레임워크는 그림 1에 나와 있다. 관광객이 다녀 온 관광명소의 리뷰데이터를 수집하고 수집한 데이터를 바탕으로 코로나-19 전후로 리뷰 데이터 내에서 단어와 구문으로 프라이버시를 측정하였으며 감성 분석 및 감정 분석을 통해서 관광객들의 인식 변화를 측정하였다. 본 연구의 프레임워크는 ‘데이터 수집’, ‘데이터 전 처리’의 2단계를 수행하고 국내 코로나-19 국가위기 관리 주의 단계가 발효된 시점을 기준으로 데이터를 분할하여 그 데이터를 기반으로 ‘데이터 분석’을 하는 것으로 구성된다.



<그림 1> 연구 모형

IV. 실험결과

첫 번째 단계인 ‘데이터 수집’은 온라인 플랫폼 사이트로부터 코로나19 발생 전후의 외국인 방문객들의 리뷰를 크롤링하여 데이터를 수집하는 단계이다. 본 연구에서는 서울, 부산, 경주 등의 주요 도시의 관광명소의 리뷰 데이터를 수집하였다. 수집된 리뷰 데이터를 문서화하고, 문서화된 리뷰 데이터를 모두 통합하여 코로나19 팬데믹 이전과 이후로 데이터를 분할하였다.

전체 데이터를 분석 결과는 <표1>에서 볼 수 있다. 분석 결과를 확인 하였을 때 결과 상태(Outcome state)가 5% 유의수준으로 유의미하다고 볼 수 있다. 결과 상태는 프라이버시를 통해 제공되는 정적 행동 상태와 결과와 관련이 있다(Gill et al., 2011). 코로나-19 전후를 비교 하였을 때 코로나-19 팬데믹 이후 한국을 방문한 관광객이 개별적인 공간이나 바람직하지 않은 타인으로부터의 분리로 심리적으로 자유감을 얻었다고

볼 수 있다.

Scores		Mean scores		T-test
		Before	After	P-value
Total	Outcome State	0.084	0.135	0.045
Family	Norms Requisites	0.104	0.299	0.089
	Negative privacy	0.016	0.128	0.01
	Open Visible	0.077	0.406	0.007
Friends	Norms Requisites	0.103	0.243	0.067
부산	Negative privacy	0.037	0.069	0.099
	Outcome State	0.062	0.228	0.003
	Open Visible	0.111	0.013	0.063
제주	Open Visible	0.115	0.261	0.078
경주	Outcome State	0.063	0.276	0.062

<표 1> 연구변수의 측정항목

여행 형태의 경우 비즈니스, 개인, 커플, 가족, 친구 중 친구와 가족에서 코로나-19 이후 프라이버시의 변화가 나타났다. 가족의 경우 규범적 요건, 부정적인 프라이버시 경험, 개방된 가시성에서 변화가 나타났으며 친구와 관광을 한 경우 규범적 요건(Norms Requisites)이 변화가 나타났다. 관광명소의 경우 부산에서는 부정적인 프라이버시(Negative privacy), 결과 상태(Outcome state), 개방된 가시성(Open visible)에 변화가 있었다고 나타났다. 경주를 방문한 관광객은 결과 상태(Outcome state)가 개별적인 공간이나 바람직하지 않은 타인으로 부터의 분리로 심리적으로 자유감을 얻었다고 볼 수 있다. 제주를 방문한 관광객은 개방된 가시성(Open visible)이 나

타났으며 이는 다른 사람들에게 공개되고, 타인이 공개적인 접근이 가능한 상태로 프라이버시가 침해되었다고 볼 수 있다.

참 고 문 헌

- 김종기; 김상희. 온라인 사용자의 프라이버시 보호행동에 대한 연구: 프라이버시 역설 관점에서. 인터넷전자상거래연구, 2013, 13.1: 41-64.
- WARREN, Samuel; BRANDEIS, Louis. The right to privacy. In: Killing the Messenger. Columbia University Press, 1989. p. 1-21.
- PEDERSEN, Darhl M. Dimensions of privacy. Perceptual and motor skills, 1979, 48.3_suppl: 1291-1297.

FISK, Raymond P.; BROWN, Stephen W.; BITNER, Mary Jo. Tracking the evolution of the services marketing literature. *Journal of retailing*, 1993, 69.1: 61-103.

D'ACUNTO, David; VOLO, Serena; FILIERI, Raffaele. "Most Americans like their privacy." Exploring privacy concerns through US guests' reviews. *International Journal of Contemporary Hospitality Management*, 2021.

VASALOU, Asimina, et al. Privacy dictionary: A new resource for the automated content analysis of privacy. *Journal of the American Society for Information Science and Technology*, 2011, 62.11: 2095-2105.

WORLD HEALTH ORGANIZATION, et al. Coronavirus disease 2019 (COVID-19): situation report, 73. 2020.

BBC. Coronavirus Privacy: Are South Korea's Alerts Too Revealing? (2020).

Zastrow M. South Korea is reporting intimate details of COVID-19 cases: has it helped?. *Nature*. 2020 Mar 18.

Wang CJ, Ng CY, Brook RH. Response to COVID-19 in Taiwan: big data analytics, new technology, and proactive testing. *Jama*. 2020 Apr 14;323(14):1341-2.

Kuhn A. South Korea's Tracking of

COVID-19 Patients Raises Privacy Concerns. NPR. 2020 May.