

빅데이터 분석을 활용한 사용자 경험 평가 방법론 탐색 : 아마존 에코에 대한 온라인 리뷰 분석을 중심으로

Exploration of User Experience Research Method with Big Data Analysis : Focusing on the Online Review Analysis of Echo

황해정, 심혜린, 최준호
연세대학교 정보대학원

Hae Jeong Hwang(haejeong066@gmail.com), Hye Rin Shim(leia01@yonsei.ac.kr)
Junho Choi(junhochoi@yonsei.ac.kr)

요약

이 연구는 이미 실생활에서 사용되고 있으나 이에 대한 실증적 사용자 경험 조사가 부족한 사물인터넷 기반 제품에 대한 새로운 사용자 경험 방법론을 탐색해보고자 진행되었다. 지금까지의 사용자 경험에 대한 연구가 주로 설문이나 관찰 방법 등을 통해 이루어져 온 것과 달리 본 연구에서는 사물인터넷 기반 제품 중 지능형 에이전트인 아마존 에코(Echo)를 대상으로 사용자들의 온라인 리뷰를 분석하는 빅데이터 분석 기법을 활용하여 사용자 경험을 살펴보았다. 토픽 모델링 분석 결과, 에코의 기능, 음성 인터랙션, 지속적인 기능 개선과 관련된 사용 경험 요인들이 도출되었다. 또한 회귀분석결과 지속적인 기능 개선이 만족도에 가장 큰 영향을 미치는 것으로 나타났다. 연구의 의의는 사용자 경험을 제고할 수 있는 지능형 사물인터넷 제품 연구방법으로서 빅데이터 분석방법론 활용 가능성을 제시한 점이다.

■ 중심어 : 빅데이터 분석 | 토픽 모델링 | 온라인 리뷰 분석 | 사용자경험(UX) | 지능형 에이전트 |

Abstract

This study attempted to explore and examine a new user experience (UX) research method for IoT products which are becoming widely used but lack practical user research. While user experience research has been traditionally opted for survey or observation methods, this paper utilized big data analysis method for user online reviews on an intelligent agent IoT product, Amazon's Echo. The results of topic modelling analysis extracted user experience elements such as features, conversational interaction, and updates. In addition, regression analysis showed that the topic of updates was the most influential determinant of user satisfaction. The main implication of this study is the new introduction of big data analysis method into the user experience research for the intelligent agent IoT products.

■ keyword : Big Data Analysis | Topic Modeling | Online Review Analysis | User Experience (UX) | Intelligent Agent |

I. 서론

가트너는 ‘2016년도 전략 기술 트렌드 10(strategic technology trends for 2016)[1]’을 발표하면서, 사물인터넷, 인공지능, 자율성을 지닌 로봇이나 자동차 등 새로운 기술이 적용된 포스트 모바일(post-mobile) 제품이 각광받을 것으로 전망했다. 구글은 2016 개발자 회의(I/O)를 통해 구글 어시스트와 구글 홈을 발표하면서 지능형 에이전트와 사물인터넷 시장으로의 본격적인 진출을 예고했다[2].

사물 인터넷 기반 제품으로 지능형 에이전트인 아마존 에코와 같은 제품이 이미 상용화되었다. 이러한 제품은 사용자 개인 특성이나 주변 정보를 학습하고[3] 물리적 공간이나 환경과 밀접하게 연결되는 특성이 있어[4][5] 이에 대한 사용자 경험과 상호작용방식이 다른 제품군에서와 다른 양상으로 나타난다[6-8]. 따라서 새로운 연구방법론의 모색이 필요한 시점이다.

HCI(Human-computer interaction)분야에서는 제품을 사용하는 사용자 경험을 파악하기 위하여 주로 설문이나 인터뷰, 실험 방법을 활용해 왔다[9][10]. 그러나 설문이나 실험은 실제 제품 사용 맥락을 반영하지 못하는 한계점이 있다[10]. 이에 비해 사용 맥락을 고려하는 장기 관찰 및 인터뷰 방법을 이용하여 환경이 제품의 사용 경험에 영향을 크게 미치는 사물인터넷 기반 제품들의 사용 경험 연구에 활용되기도 하였다[4][5][7][11][12]. 하지만 장기 관찰 및 인터뷰는 데이터 수집 및 분석에 적지 않은 시간과 자원이 소요되기 때문에 실무에서 사용하는데 제약이 있다[9].

본 연구에서는 사물 인터넷 기반 제품에 대한 사용자 경험을 파악하는 방법으로 실제 사용자들의 온라인 리뷰를 대상으로 빅데이터 분석을 탐색적으로 적용해 보고자 한다. 온라인 리뷰에는 사용자 생각이 직접적으로 반영되며, 대규모로 수집될 수 있는 장점이 있다[13][14]. 또한 실제 제품 사용 후 작성되므로 때로는 사용자가 자각하지 못한 멘탈 모델을 드러내기도 한다[15]. 과거에 비해 비정형 데이터를 수집하고 처리하는 컴퓨팅 기술이 발전하면서 비정형 데이터 핵심 내용을 추출하는 것이 용이하다. 뿐만 아니라 이를 통계적 분

석에 활용할 수 있는 정량 데이터로 변환하는 것이 가능해진 것도 온라인 리뷰를 활용하는데 유리한 여건이 된다.

이 연구에서는 사용자 경험을 구성하는 하위 요소 도출을 위해 텍스트 마이닝 중 토픽 모델링 방법을 적용하였다. 토픽 모델링은 통계적으로 문서 내의 하위 주제들을 자동적으로 생성하는 특징이 있다[16]. 도출된 하위 주제는 세부 요소들을 발굴하는 동시에, 해당 요소와 관련된 평가나 감성어휘를 자동 추출하기 때문에 [17] 온라인 리뷰를 대상으로 한 연구에서 많이 사용되어 왔다. 사용자 경험에 대한 평가에서는 사용자가 중요하게 생각하는 제품 하위 요소나 개념을 밝히고자 하므로[18][19] 사용자 경험 평가에 토픽 모델링을 적용하려고 한다.

요컨대 본 연구의 목적은 사물 인터넷 기반 제품과 같은 새로운 종류의 제품에 대한 사용자 경험을 평가하기 위한 방법으로 빅데이터 분석 방법론 중에서 토픽 모델링을 제안하고 이 방법론의 효용을 검증하는 것이다. 이를 위하여 사물 인터넷 기반 제품인 지능형 에이전트 에코를 사례로 선정하여 연구를 진행했다. 구체적으로는 사용자가 지각하는 에코의 하위 특성을 도출하고 해당 특성이 사용만족도에 미치는 영향을 통계적으로 살펴보았다.

II. 문헌 연구

1. 기존 사용자경험 조사방법론의 문제

사용자 경험(user experience)은 사용자가 제품이나 서비스를 사용하면서 느끼는 정서적 반응이나 가치, 태도 등 전반적인 경험을 포괄하는 개념이다[9][13][20][21]. 연구에 따라 다르지만, 사용자 경험은 제품 사용 전 기대나 사용 시 경험하게 되는 전반적인 행동이나 느낌, 사용 후 태도 및 정서적, 인지적 반응을 포함한다. 사용자 경험 연구를 통해 연구자는 사용자가 제품에 대해 무엇을 가치 있다고 여기는지, 감성적 측면의 반응이나 상호작용은 어떠한지 알아보고자 한다[9][13][22][23].

이 때 사용자가 느끼는 감정이나 가치는 제품을 사용

하는 맥락에 따라 달라진다. 그렇기 때문에 Law et al. [23]은 사용자 경험 평가의 세 요소로 사용자 내적 상태나 제품 기능, 특징과 함께 사용자와 제품이 상호작용하는 환경 맥락(context)을 꼽았다. 즉, 사용자 경험을 평가한다는 것은 사용자와 제품이 상호작용하는 특정 맥락에서, 사용자가 가치 있다고 느끼는 제품 요소나 개념은 어떤 것인지, 해당 요소에 대한 사용자 감정이나 반응은 어떠한지를 살펴보는 것이다.

HCI분야에서 이를 위해 사용해 온 방법들이 갖는 가장 큰 문제점은 낮은 실용성(practicability)이다. Vermeeren et al.[9]은 현재 학술적 혹은 실무적으로 활용되고 있는 사용자 경험 평가 방법의 특징을 정리하였다. 그 결과, 맥락이 중요한 요소로 작용하는 제품이나 서비스에 대한 사용경험 평가에서는 장기 조사나 현장 실험을 활용하는 경우가 많았다. 그러나 이러한 방법은 비용과 시간이 많이 들고 분석하기에도 어렵다는 단점이 있다고 지적하였다.

전반적으로 현재 사용되고 있는 사용자 경험 평가 방법들은 깊이 있는 사용 경험 연구를 위하여 정성 데이터를 수집하고 활용하는 경우가 많기 때문에 이러한 문제점들이 두드러졌다. 이기호 외[13]는 이와 같은 문제점에서 출발하여 실무에서 사용할 수 있는 효율적인 사용자 경험 평가 방법을 개발하고자 하였다. 맥락 질문법과 설문을 통하여 다양한 제품군에 대해 범용적인 평가 방법을 개발하고 적용해 보았다. 그러나 이 연구 역시 장기간에 걸친 현지 조사와 정성 연구를 기반으로 요인을 도출하고 대규모 설문을 통하여 프로세스를 개발하였다는 점에서 효율성 문제가 있다.

사용자 경험 평가 방법론은 제품에 따라 새롭게 개발되거나 개선돼 왔다. 변대호[24]는 스마트 TV는 기존 TV나 IPTV와 다른 특성을 갖기 때문에 사용자가 만족을 느끼는 요인도 달라질 수 있다고 보았다. 관련 요인을 밝히기 위해 상품 후기를 분석하여 만족에 영향을 주는 요소들을 발굴하였다. Kjeldskov et al.[10]은 모바일과 관련한 사용자 경험 평가 방법들을 정리하였다. 이처럼 제품에 따라 사용자 경험에 영향을 미치는 요소나 사용 맥락, 특징적인 인터랙션 등이 달라지기 때문에 적절한 사용자 경험 평가 방법이 필요하다.

이 연구에서는 기존 사용자 경험 평가 방법의 낮은 실용성을 극복하고 사용맥락이 사용경험에 밀접한 영향을 주는 사물 인터넷 기반 제품에 대한 사용자 경험 평가에 적절한 방법론으로 온라인 리뷰를 활용한 토픽 모델링을 제안한다.

2. 사용자 경험 도출을 위한 온라인 리뷰 분석

온라인 리뷰는 서비스나 제품에 대한 사용자 생각과 의견을 알 수 있는 정보원이다[17][25]. 온라인 리뷰를 통하여 서비스나 제품 개선점을 파악하거나[14][26] 의사 결정에 활용하기도 한다[15][27].

온라인 리뷰 연구는 크게 두 종류로 나눌 수 있다. 첫 번째는 온라인 리뷰에 나와 있는 사람들 의견이나 정보를 추출하는 연구이다. 자동차 포럼에서 리뷰를 크롤링하여 자동차 브랜드와 개별 모델에 대한 사람들 생각을 구조적으로 파악하고자 한 연구[15]가 대표적이다. 온라인 리뷰에서는 사람들이 제품을 어떻게 생각하고 느끼는지, 많은 사람들이 공통적으로 생각하는 제품 장단점은 무엇인지[28] 등을 알 수 있기 때문에, 온라인 리뷰를 통해 사용자 경험을 파악하고 측정하려는 연구들이 진행되어 왔다[14][24]. 이 연구들은 오픈이언 마이닝이나 텍스트 마이닝을 통한 키워드 추출 방법을 사용하였다. 예를 들어 채승훈[29]은 소셜 커머스 어플리케이션과 오픈마켓 어플리케이션 사용 경험을 비교하기 위하여 텍스트 마이닝 기법을 활용하여 리뷰를 하위 주제로 분류하고, 해당 주제가 긍정적인지 부정적인지를 계산하였다.

이들 연구에서 온라인 리뷰를 사용한 이유는 설문이나 실험에 비해 사용자의 솔직한 생각이 반영되기 때문이다. 설문이나 실험에서는 연구 참여자가 연구 진행자를 의식해 사회적으로 바람직한 응답을 하거나, 실험 상황에서 편향들이 발생하기도 한다[30]. 그러나 온라인 리뷰는 자연스러운 일상에서 작성되기 때문에 사용자 의견이나 생각이 그대로 반영되며 때로는 사용자가 의식하지 못한 멘탈 모델을 드러내기도 한다[15]. 제품이나 서비스를 사용하는 사용자 경험 측면에서는 실험이나 설문에 비해 온라인 리뷰에서 실제 맥락을 반영한 사용자 피드백이 나타난다.

두 번째는 온라인 리뷰 정보를 바탕으로 수요를 예측하는 모델을 세우거나[31], 리뷰에서 사용자 특성을 도출하여 추천 시스템에 활용하는 연구[32]이다. 이 유형의 연구들은 제품 판매 데이터나 사용자 특성 등 다른 데이터와 온라인 리뷰 사이에 통계적 관계를 검증한다. 또는 다른 데이터를 설명하는데 온라인 리뷰 데이터를 활용한다. 즉 온라인 리뷰를 종속 변수로 하여 다른 독립 변수를 설명하고자 한다.

두 번째 유형의 연구에서 확인할 수 있는 온라인 리뷰의 장점은 통계적 유용성이다. 연구 대상이 특수하거나 서비스 및 제품의 사용자가 적은 경우 설문이나 실험을 위한 표본을 모집하기 어렵다. 설문이나 실험을 진행하더라도 살펴보고자 하는 모든 측정 항목이나 개념에 대해 조사를 진행하는 것은 거의 불가능하다. 외부 데이터와 이를 결합할 경우 통계적인 신뢰도에 문제가 발생할 수 있다. 이에 비해 온라인 리뷰는 대규모 표본을 비교적 쉽게 구할 수 있다. 리뷰 특성상 구매자 특징이나 제품의 판매 데이터, 별점으로 나타나는 만족도 등 연관 지어 분석할 수 있는 관련 데이터가 풍부하다는 점도 온라인 리뷰가 갖는 장점이다.

3. 토픽 모델링

토픽 모델링은 하나의 문서를 주제(topic) 혼합체로 가정하는 데서 출발한다[33]. 문서에 나타나는 주제들을 알아내기 위하여 통계적인 추론 기법을 활용하는데, LDA(Latent dirichlet allocation) 알고리즘이 대표적이다[32]. LDA 알고리즘은 문서 내에 출현하는 단어들이 독립적이지 않다고 가정하여 단어들의 다항분포를 생성한 조건을 사후적으로 추론하여 주제를 발견한다. 즉, LDA에서 주제는 단어들에 대한 다항분포로 정의될 수 있다[16][17].

LDA 알고리즘은 구조화되지 않은 비정형 데이터에서 주제를 자동적으로 발굴한다는 장점이 있으며[34], 해당 주제와 관련된 단어들을 확인할 수 있기 때문에 [17] 텍스트 마이닝 연구나 온라인 리뷰를 대상으로 한 연구들에서 활용되어 왔다. Jo et al.[17]은 LDA 기반의 새로운 알고리즘을 제안하면서 아마존 리뷰를 대상으로 알고리즘 평가를 진행하였다. 연구 결과 카메라에

대한 리뷰 내에서도 렌즈나 줌 등 다양한 하위 평가 요소들이 발굴되었고 토픽 내에 이에 대한 평가어가 같은 토픽 내에서 출현했다. 채승훈 외[14]는 쇼핑 어플리케이션에 대한 사용자 경험 비교 분석을 위하여 LDA를 활용하였다. 그 결과 쿠폰이나 할인 등 마케팅 측면에서 뿐만 아니라 로그인이나 결제 등 어플리케이션 기능에 대한 사용성에서 나타나는 하위 요소들이 리뷰에서 나타났다. 또한 해당 토픽의 비율과 리뷰에서 나타나는 감정값을 계산하였다.

이처럼 LDA는 LSA(Latent semantic analysis)등 다른 토픽 모델링 방법에 비해 결과 해석이 용이하고[16] 방대한 비정형 데이터의 차원을 축소해 여러 주제들을 도출하는데 장점이 있다[17]. 보편적으로 정성 연구에서는 연구자가 데이터를 분석하기 위해 상당한 시간과 노력을 투입한다. LDA에서는 그러한 과정이 컴퓨터 연산 과정에 의해 이루어져 특히 데이터가 방대한 경우 유용할 수 있다.

사용자 경험 평가에서는 사용자에게 지각되는 기능이나 하위 요소들을 도출하고 그에 대한 사용자 감정을 파악하는 것이 중요하다. LDA는 문헌 내에서 하위 주제를 도출하는 동시에 단어들을 함께 확인할 수 있기 때문에 하위 요소에 대한 감정 단어나 평가 단어를 도출하여 사용자 경험을 평가하는데 적용될 수 있다.

III. 사례 연구

1. 연구 대상 및 방법론

연구 대상으로는 아마존(Amazon.com)에서 출시한 에코(Echo)를 선정하였다. 에코는 지능형 에이전트를 탑재한 제품으로 스피커, 스마트 홈, 물건 주문 등 다양한 기능을 갖추고 있다[35]. 에코는 원통형의 단순한 기기 디자인으로 제품에 스크린이 없으며 음성을 통한 대화형 인터랙션을 주로 사용하도록 되어 있다.

에코는 알렉사(Alexa)라고 불리는 지능형 에이전트를 탑재하고 있어 사용자 질문에 대답하거나 사용자 요청에 따라 음악을 틀어주는 기능을 수행한다. 사용자가 스마트 전구나 스마트 홈 제품을 설치한 경우 이 제품

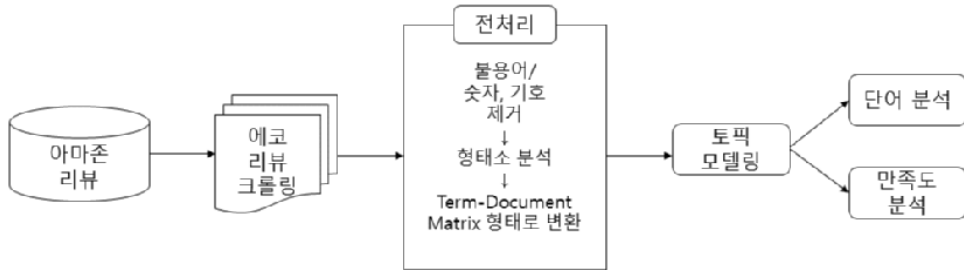


그림 1. 연구 절차

들과 연결되어 집안 조명을 에코를 통해 음성으로 제어할 수 있으며 최근에는 온라인 장바구니에 물건을 추가하거나 배달 음식을 주문하는 기능도 추가되었다. 아마존 에코 판매 페이지에는 약 3만 건이 넘는 리뷰(2016년 3월 기준)가 게시되어 있어 본 연구에서 제안하고자 하는 온라인 리뷰 분석 방법을 적용하기에 충분하다고 판단하였다.

2. 데이터 수집 및 전처리

본 연구는 위 [그림 1]과 같이 데이터 수집 및 전처리, 분석의 과정을 거쳐 진행되었다. 우선 리뷰 수집은 오픈소스 프로그램인 파이썬(Python)을 이용하여 크롤링하였다. 에코 온라인 리뷰 페이지에서 리뷰 작성 날짜, 작성자 ID, 제품을 평가한 별점, 리뷰를 가져오도록 코드를 작성하였다. 수집 결과, 2015년 6월 19일 ~ 2016년 3월 30일까지 작성된 리뷰 34,548건이 수집되었다.

데이터 전처리와 분석에는 오픈 소스인 R ver3.4.3.을 사용하였다. 리뷰 작성자, 제품에 대한 별점, 작성 날짜를 해당 내용에 알맞은 형태로 변경하였다. 특히 리뷰 데이터는 최종적으로 분석에 알맞은 벡터 데이터 형태로 변경하였다. 벡터 데이터는 리뷰에 등장하는 단어와 해당 단어 출현 빈도로 구성된다[36]. 이 과정은 텍스트 분석을 위하여 개발된 R의 tm(text mining) 패키지[37]를 사용하였다. tm 패키지는 텍스트 데이터를 받아 말뭉치(corpus)를 만들고, 말뭉치에서 문장 부호, 숫자 등

을 제거하는 등 텍스트 마이닝을 위한 기능을 제공하고 있다.

구체적으로 시행된 전처리 과정은 다음과 같다. 텍스트 문서를 말뭉치로 만든 뒤, 크롤링 과정에서 삽입된 html 태그, 관사나 지시 대명사 등 분석 시 유의미하지 않은 단어들을 불용어(stopwords)로 처리하여 삭제하였다. 이 과정에서 이미지, 영상만 있거나, 한 문장밖에 없는 짧은 리뷰들은 삭제되었다. 정제된 텍스트 데이터는 문서-단어 매트릭스(document-term matrix)로 변환하여 한 리뷰 안에 등장하는 단어들과 빈도를 가진 매트릭스를 구성하였다. 매트릭스 구성 시 단어 빈도는 TF-IDF(term frequency - inverse document frequency) 가중치로 계산하였다. 전처리를 거쳐 분석에 활용된 리뷰는 총 16,861건이었다. 리뷰의 평균 별점은 4.3점, 리뷰의 평균 길이는 653자였다.

3. 분석 방법

본 연구에서는 리뷰에 어떠한 내용이 분포되어 있는지 살펴보기 위하여 토픽모델링을 활용하였다. 그 중에서도 Blei[33]가 제안한 알고리즘에 따라 LDA기법을 사용하였다. 분석에는 오픈소스 R의 lda패키지를 사용하였다[38].

LDA 분석 시 지정할 수 있는 파라미터로는 토픽 수, 계산을 시행할 샘플링 횟수, 샘플링 결과에서 초기 결과 중 사용하지 않을 계산 결과 수(burn-in), 문서 내 주제의 비율(alpha), 주제의 잠재 다항분포(eta) 값이 있다. LDA는 확률에 기반한 모델이기 때문에 주제의 수가 주어지지 않은 경우, 샘플링의 반복 횟수와 주제의

1 에코 온라인 리뷰 페이지 주소 :

https://www.amazon.com/Amazon-Echo-Bluetooth-Speaker-with-WiFi-Alexa/product-reviews/B00X4WHP5E/ref=cm_cr_dp_sec_al_summary?ie=UTF8&showViewpoints=1&sortBy=helpful

수를 조정하여 가장 적절하다고 판단된 수로 지정한다 [34][39]. 본 연구에서는 샘플링 횟수를 7,000회로 지정하고, 총 5개 토픽을 지정하였다. 샘플링을 반복할수록 결과가 점차 개선되기 때문에 총 샘플링 횟수 중 초기 연산 결과를 버리고 나머지 결과만으로 결과를 도출하도록 지정할 수 있다. 이를 번인(burn-in)이라 하는데, 이 값은 1,000으로 지정하였다. 온라인 리뷰는 리뷰 내에 다양한 주제들이 등장할 수 있기 때문에 문서 내 주제 비율을 0.1로 높게 설정하였다. 단어 출현 빈도와 관련된 잠재 다항분포 값 역시 0.1로 정하였다.

LDA 결과에서 활용된 데이터는 두 종류이다. 첫 번째는 각 토픽에 해당하는 상위 30개 단어를 뽑아, 어떠한 하위 주제들이 도출되었는지 살펴보았다. 두 번째는 도출된 하위 주제가 만족도에 미치는 영향을 살펴보기 위하여 문서 내 토픽의 비율을 계산하였다. 문서 내 토픽 비율은 한 온라인 리뷰 내에서 각 토픽이 차지하는 비율로, 앞서 도출한 토픽 단어와 리뷰 단어를 비교하여 계산하게 된다. 각 온라인 리뷰마다 사용자들이 평가한 별점(5점만점)을 종속 변수에 해당하는 만족도로, 각 토픽의 비율을 독립 변수로 하여 다중 회귀 분석을 실시하였다.

IV. 연구결과

1. 토픽 도출

토픽모델링 결과, 총 5개의 토픽이 도출되었다. 전반적으로 에코를 부르는 이름인 “alexa”가 여러 토픽에 분포되어 있고, good, love 등 에코에 대한 긍정적인 감정 평가도 나타났다.

Topic 1은 update, keep, better 등 지속적인 업데이트와 관련된 내용이 주를 이루었다. 또한 가족이나 친구 등 에코를 사용하는 다른 사람에 대한 언급도 함께 나타났다. Topic 1은 특정 기능에 대한 주제라기보다 제품을 사용하면서 업데이트되는 기능이나 성능에 대한 내용이라 할 수 있다. 에코는 출시 후 지속적으로 업데이트를 진행하면서 새로운 기능을 추가하고 연동되는 제품이나 서비스를 확장해 왔다. 대표적으로 스마트 홈

기능이나 쇼핑 리스트 관련 기능이 있다. 이러한 업데이트 측면이 리뷰에서도 나타난 것을 알 수 있다. 또한 다른 토픽에 비해 amazing, enjoy, best 등 감정적인 단어가 많이 나타났다.

Topic 2는 speaker, sound, quality, bluetooth 등 스피커와 관련된 내용이었다. 특히 제품 물성을 지칭하는 단어들(device, product)과 그에 대한 세부 평가 요소(function, quality, bluetooth)가 함께 나타나 스피커에 대한 스펙을 평가하는 방식으로 내용이 구성된 것을 알 수 있다.

표 1. LDA에 의한 토픽 분류 결과

	topic1 업데이트	topic2 스피커	topic3 대화형 UI	topic4 스마트 홈	topic5 기타 기능
1	love	great	question	device	list
2	great	voice	ask	light	love
3	alexa	speaker	answer	connect	music
4	music	good	alexa	control	great
5	family	like	love	work	play
6	update	device	time	turn	use
7	keep	sound	understand	product	shop
8	feature	product	siri	alexa	alexa
9	get	use	know	home	listen
10	better	really	tell	house	radio
11	use	quality	get	unit	timer
12	purchase	love	like	room	get
13	amazing	better	find	get	device
14	new	feature	really	purchase	thing
15	friend	function	say	remote	easy
16	time	work	music	addition	ask
17	list	music	good	star	feature
18	much	nice	play	system	time
19	become	pretty	weather	issue	speaker
20	help	integrate	alway	wifi	song
21	best	command	great	app	add
22	function	much	word	like	prime
23	product	well	improve	try	book
24	part	price	name	voice	read
25	buy	capable	wake	smart	alarm
26	continue	still	search	volume	station
27	device	see	many	people	good
28	really	bluetooth-h	device	though	make
29	ad	google	give	integrate	tell
30	enjoy	new	song	able	like

Topic 3에서는 question, siri, answer 등 음성 인터랙션 관련된 내용이 나타났다. 이 토픽에서는 말과 관련된 단어(question, ask, answer)들이 주로 나타났고, 음성 UI를 통해 실행되는 내용들(music, weather, search)도 함께 나타났다.

Topic 4에서는 light, control, house 등 스마트 홈 기능과 관련된 단어들이 하나의 주제로 묶였다. 주제 특성 상 공간과 관련된 단어들(home, house, room)이 가장 많이 나타났다. 또한 light, unit 등 스마트 홈 시스템과 관련한 하위 제품들의 명칭들도 나타났다.

마지막으로 Topic 5에는 에코의 스피커 기능으로 할 수 있는 기타 기능들이 함께 나타났다. 음악이나 라디오 듣기(music, radio, station, song), 책 읽기(book, tell), 쇼핑(shop, list), 타이머 등 에코로 할 수 있는 기능을 나타내는 단어들이 함께 묶였다.

토픽은 크게 기능과 관련한 토픽(topic 2, topic 4, topic 5), 인터랙션 방식과 관련한 토픽(topic 3), 업데이트를 통한 기능 개선과 관련한 토픽(topic1)으로 분류할 수 있었다. 분류와 상관없이 분석에서 나타난 평가, 감정 관련 어휘로는 easy, new, love, amazing 등의 형용사가 있었다. 또한 가족이나 친구 등 제품과 관련된 사람들을 나타내는 명사가 나타났으며, 제품이 사용되거나 기능이 영향을 미치는 공간에 대한 어휘들도 있다.

2. 토픽 속성이 만족도에 미치는 영향

앞에서 도출된 LDA 결과를 바탕으로 각 토픽이 제품 만족도에 미치는 영향을 통계적으로 살펴보았다. 앞서 분류한 세 종류 토픽 중 어떠한 토픽이 만족도에 가장 큰 영향을 미치는지 살펴보기 위하여 모델을 비교해 보았다. 모델 비교를 위하여 상대적인 모델의 간명성과 적합도를 평가하는 데 사용되는 AIC (Akaike information criterion) 값을 활용하였다[40][41]. 만족도에 영향을 미치는 변수들에 대한 회귀모델 결과는 다음 [표 2]와 같다.

먼저 모델 1은 스피커, 스마트 홈, 기타 기능이 만족도에 미치는 영향을 보여준다. 기타 기능의 경우 리뷰에서 등장 빈도가 높을수록 만족도도 높아지는 것으로 나타났으나($t=5.87, p<.001$), 스피커 관련 기능($t=-3.38,$

표 2. 만족도에 영향을 미치는 변수에 대한 회귀분석 모델

	model 1	model 2	model 3
업데이트			0.686*** (0.05)
대화형 UI		-0.521*** (0.048)	-0.412*** (0.048)
기타 기능	0.294*** (0.050)	0.239*** (0.050)	0.275*** (0.025)
스마트 홈	-0.215*** (0.046)	-0.331*** (0.047)	-0.252*** (0.047)
스피커	-0.173*** (0.051)	-0.204*** (0.051)	-0.129** (0.051)
AIC	39925.43	39809.74	39625.25
Adjusted R ²	0.005	0.013	0.026
Residual Std. Error	1.002	0.998	0.991
F Statistic	25.442***	48.776***	77.070***

* : $p<.05$ ** : $p<.01$ *** : $p<.001$ 에서 유의함

$p<.001$)과 스마트 홈 기능($t=-4.62, p<.001$)은 적게 나타날수록 만족도가 높아지는 것으로 나타났다.

모델 2는 모델 1에서 대화형 UI를 추가로 회귀시킨 것으로 모델 1에 비해 설명력이 좀 더 올라간 것을 알 수 있다. 기능과 대화형 UI 모두 통계적으로 유의하게 영향을 미치는 것으로 나타났다. 스피커, 스마트 홈 기능과 유사하게 대화형 UI($t=-10.869, p<.001$)는 리뷰에서 많이 등장할수록 만족도가 낮아지는 것으로 나타났다.

모델 3은 모델 2에서 업데이트 관련 항목을 추가하여 회귀시킨 결과이다. 결과를 살펴보면 만족도를 매우 낮은 수준에서 설명하고 있으나, 세 모델 중 AIC값이 가장 낮아 간명성과 적합도 측면에서 가장 뛰어난 모델이라 볼 수 있다. 스피커, 대화형 UI, 스마트 홈 기능은 리뷰에 적게 등장할수록 만족도에 긍정적 영향을 주는 것으로 나타났고, 지능 측면($t=13.69, p<.001$)과 기타 기능은 리뷰 등장 빈도가 만족도에 긍정적인 영향을 미치는 것으로 밝혀졌다.

만족도에 영향을 미치는 변수들 간 상대적인 영향력을 평가하면, 업데이트 측면($\beta=.686$)이 가장 큰 영향을 보이는 변수로 나타났다. 대화형 UI($\beta=-.412$)가 두 번째로 큰 영향력을 보이는 변수였다. 대화형 UI의 등장 빈도가 높을수록 부정적인 영향을 미치는 이유는 검색 기능의 한계 때문인 것으로 추정된다. 해당 토픽과 관련된 리뷰들을 살펴보면 음성인식을 자체는 높으나 북

잡한 질문에 대한 답을 하지 못한다는 내용이 있었다. 특정한 사건이나 정보에 대해 질문했을 때에는 구글 검색 결과를 그대로 읽어 불필요한 정보까지 제공하는 것이다.

스피커와 스마트 홈 기능도 등장 빈도가 높을수록 만족도에 부정적인 영향을 미치는 것으로 나타났다. 블루투스나 와이파이의 연결 불안정으로 인한 문제이거나 다른 서비스나 제품과의 통합 측면에서 발생한 문제로 추정된다.

업데이트 측면과 기타 기능은 모두 만족도에 긍정적인 영향을 주는 것으로 나타났다. 온라인 쇼핑이나 라디오 플레이, 책 읽기 등 예코가 업데이트 되면서 추가된 기능들이 만족도에 긍정적인 영향을 주는 것으로 보인다.

V. 토의

본 연구는 온라인 리뷰 분석 방법을 사용자 경험 분석 방법으로 적용해 보고자 하였다. 아마존의 지능형 에이전트 제품인 예코에 대한 온라인 리뷰를 바탕으로 토픽 모델링 분석을 실시한 결과, 총 5개 토픽이 도출되었다. 5개 토픽은 다시 기능, 인터랙션, 업데이트라는 3개의 카테고리로 나뉘었다. 각 카테고리 중 사용자가 느끼는 만족도에 가장 큰 영향을 미치는 카테고리는 업데이트 측면이었다. 연구 결과를 하위 요소 측면과 어휘 측면으로 나누어 기존 연구와 함께 살펴보고자 한다.

우선 사용자들은 지능형 에이전트가 탑재된 제품을 크게 기능, 인터랙션 방식, 기능개선의 측면으로 나누어 평가하는 모습을 보였다. 기능과 인터랙션 방식은 기존의 사용성 평가 연구에서도 평가 대상이 되었던 요소들이다[13][20][21]. 기존 정성 연구 결과와 비교해 보면 상당히 개략적으로 나타난다는 점을 알 수 있다. 그러나 LDA분석으로 하위 요소와 만족도와의 상관관계 분석을 통해 비정형 데이터로도 정량적 분석을 시행할 수 있었다.

토픽 모델링에서 나타난 평가 및 감정 어휘들은 기존 사용자 경험 평가를 위한 설문에서 사용되는 어휘들과

유사한 어휘들도 있었으나, 그렇지 않은 어휘들도 나타났다. 대표적으로 easy나 amazing과 같은 어휘는 각각 사용용이성, 새로움(novelty)과 같은 개념을 측정할 때 사용되는 단어들이다. 반면 love나 great, help, understand와 같은 단어들은 기존 사용성 평가 연구에서 잘 사용되지 않거나 개념적으로 아직 정의가 불분명한 단어들이라 할 수 있다. 이 단어들은 지능형 에이전트나 스마트홈 IoT 기기가 갖는 특성 혹은 상호작용 시 발생하는 감정들을 표현하는 것으로 이에 대한 지속적인 개념화와 연구가 필요하다.

VI. 결론

1. 연구 요약 및 함의

본 연구는 온라인 리뷰를 이용하여 사물 인터넷 기반 제품에 대한 사용자 경험을 살펴보고자 하였다. 아마존에서 출시한 예코를 사례로 분석한 결과, 총 3가지 하위 카테고리에 대해 5가지 토픽이 추출되었다.

본 논문의 실무적 의의는 효율적인 사용자 경험 평가를 위하여 온라인 리뷰와 텍스트 마이닝 방법론을 실제로 적용해 보았다는데 있다. 온라인 리뷰를 이용한 텍스트 마이닝 방법론은 상대적으로 적은 비용과 시간으로 분석을 진행할 수 있다. 예컨대 가정 내 로봇에 대한 사용자 경험 연구[4][11]에서는 정성 연구 방식으로 6개월 간 연구를 진행하였다. 그러나 온라인 리뷰를 텍스트 마이닝으로 분석하는 경우 짧게는 1주일이면 대략적인 결과를 도출할 수 있어 시간 효율 측면에서 우위를 갖는다. 관련 제품을 기획하고 판매하는 회사에서는 온라인 리뷰 분석을 통해 사용자 경험을 상대적으로 쉽고 빠르게 파악할 수 있을 것이다.

온라인 리뷰와 텍스트 마이닝 방법은 특히 다음 두 가지 경우에 유용하게 적용될 수 있다. 첫 번째는 개선이 필요한 문제를 발굴하고, 의사결정을 내릴 때이다. 문제 해결을 위한 자원과 시간을 배분하기 위한 의사결정 시에는 만족도에 대한 회귀 분석 등 정량적 분석이 도움이 될 수 있다. 현재 사용자들이 제품에 대해 가장 중요하다고 느끼는 기능은 무엇인지, 해당 기능이 만족

도에 어떠한 영향을 미치는지를 정량적으로 확인할 수 있다. 예컨대 이 연구에서처럼 에코 분석 결과 대화형 UI나 스피커가 만족도에 부정적인 영향을 미치는 것으로 나타난 경우 해당 기능과 관련해서 문제를 파악하고 개선해야 한다는 목표를 세울 수 있다. 두 번째는 시간 흐름에 따라 제품에 대한 사용자 경험을 비교할 수 있다는 점이다. 시기에 따라 리뷰를 분석하거나 업데이트 전후를 기준으로 내용을 분리하여 비교할 경우, 제품사용주기에 따라 사용자 반응이 어떻게 달라지는지를 확인할 수 있을 것이다. 이를 통해 제품이나 서비스가 사용자 경험에 초점을 둔 기획 의도나 방향에 따라 적절하게 개선되고 있는지, 사용자가 긍정적인 반응을 보이는 기능이나 특징에는 어떤 것들이 있는지 확인할 수 있다.

이 연구가 갖는 이론적인 시사점으로는 지능형 에이전트와 관련하여 사용자가 느끼는 요소나 감정과 관련하여 새로운 어휘들을 발견하였다는 것이다. 새로운 어휘를 발견하는 과정은 사용자 경험과 관련된 새로운 요소를 발굴하고 개념화하기 위한 가장 기초적인 단계가 될 수 있다. 예컨대 대화형 인터랙션은 주로 알고리즘 측면에서 논의가 이뤄져 왔다. 이에 대한 사용자 경험을 평가하는 설문 항목이나 측정 항목을 만들 때, ‘잘 이해한다’고 느끼는지 ‘질문에 대한 대답이 적절한지’와 같이 본 연구에서 나타난 단어들이 활용될 수 있다. 나아가 이 단어들을 중심으로 사물 인터넷 제품이나 대화형 인터랙션 중심의 서비스에서 사용자 경험 평가 시 고려해야 하는 주요 개념을 새롭게 발굴하고 정의해 나갈 수 있을 것이다.

2. 연구의 한계 및 향후 연구 방향

본 연구가 지닌 한계점은 다음과 같다. 첫 번째는 온라인 리뷰를 활용함으로써 발생하는 한계점이다. 기존 연구에 따르면, 온라인 리뷰는 상대적으로 호의적인 의견이 주를 이루기 때문에 부정적 의견이 반영되지 않거나 자료 자체가 편향될 수 있다[37]. 본 연구에서도 대부분 리뷰가 긍정적으로 제품을 평가하고 있었고, 부정적인 성향의 리뷰는 훨씬 적었다. 리뷰 내에서 긍정적 내용과 부정적 내용을 구분하여 분석에 활용하거나, 판

매량이나 로그데이터 등 별점 이외의 지표를 분석에 활용하는 방법을 고려해 볼 수 있을 것이다.

두 번째는 분석 방법상 발생하는 한계점이다. 본 연구에서 사용한 LDA 알고리즘은 특정한 기준을 설정하지 않고 통계적 원리를 활용하여 자동적으로 주제를 발굴한다. 따라서 기준에 밝혀지지 않은 내용을 탐색하는데에 장점이 있으나, 반대로 도출된 내용에 대한 과학적 엄밀성(scientific quality)이 낮다는 한계가 있다. 이를 보완하기 위해서는 LDA 결과를 해석할 때에 이 분야 전문가들과 함께 분석함으로써 외적 타당성을 높이는 방법, 사후 설문이나 통계적 검증을 통해 내적 타당성을 높이는 방법 등을 생각해 볼 수 있다. 아울러 본 연구에서 활용된 알고리즘은 비교적 단순한 수준이므로 향후 연구에서는 검증하고자 하는 제품 사용성 목표에 기반하여 새롭게 알고리즘을 구현할 수도 있을 것이다.

참고 문헌

- [1] <http://www.gartner.com/newsroom/id/3143521>.
- [2] <https://googleblog.blogspot.kr/2016/05/io-buildi-ng-next-evolution-of-google.html>
- [3] 조광수, “연결하고 운영하는 통합서비스가 참된 IoT인터넷 들어간 기계는 Smart Things일 뿐,” 동아비즈니스리뷰(DBR), 제184권, 제1호, 2015.
- [4] A. Brush, B. Lee, R. Mahajan, S. Agarwal, S. Saroiu, and C. Dixon, “Home automation in the wild: challenges and opportunities,” Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, 2011.
- [5] J. Forlizzi, “How robotic products become social products: an ethnographic study of cleaning in the home,” in Proceedings of the ACM/IEEE International Conference on Human-robot Interaction, ACM, 2007.
- [6] C. Breazeal, “Social interactions in HRI: the robot view. Systems, Man, and Cybernetics,

- Part C: Applications and Reviews,” IEEE Transactions, Vol.34, No.2, pp.181-186, 2004.
- [7] J. Sung, R. E. Grinter, and H. I. Christensen, “Domestic robot ecology,” International Journal of Social Robotics, Vol.2, No.4, pp.417-429, 2010.
- [8] E. Welbourne, L. Battle, G. Cole, K. Gould, K. Rector, S. Raymer, M. Balazinska, and G. Borriello, “Building the internet of things using RFID: the RFID ecosystem experience,” Internet Computing, IEEE, Vol.13, No.3, pp.48-55, 2009.
- [9] A. P. Vermeeren, E. L. C. Law, V. Roto, M. Obrist, J. Hoonhout, and K. Väänänen-Vainio-Mattila, “User experience evaluation methods: current state and development needs,” Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries, ACM, 2010.
- [10] Jesper Kjeldskov and Connor Graham, “A review of mobile HCI research methods,” Human-computer interaction with mobile devices and services, Springer, pp.317-335, 2003.
- [11] J. Fink, V. Bauwens, F. Kaplan, and P. Dillenbourg, “Living with a vacuum cleaning robot,” International Journal of Social Robotics, Vol.5, No.3, pp.389-408, 2013.
- [12] J. Forlizzi and C. DiSalvo, “Service robots in the domestic environment: a study of the roomba vacuum in the home,” Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-Robot Interaction. ACM, 2006.
- [13] 이기호, 이인성, 전석원, 양승화, 최지웅, 김진우, 박승용, 한명희, “사용자 경험 측면에서 제품을 평가하는 방법,” 한국HCI학회, 2008.
- [14] 채승훈, 임제익, 강주영, “사용자 리뷰를 통한 소셜커머스 와 오픈마켓의 이용경험 비교분석,” 지능정보연구, 제21권, 제4호 pp.53-77, 2015.
- [15] O. Netzer, R. Feldman, J. Goldenberg, and M. Fresko, “Mine your own business: Market-structure surveillance through text mining,” Marketing Science, Vol.31, No.3, pp.521-543, 2012.
- [16] D. M. Blei, “Probabilistic topic models,” Communications of the ACM, Vol.55, No.4, pp.77-84, 2012.
- [17] Y. Jo and A. H. Oh, “Aspect and sentiment unification model for online review analysis,” Proceedings of the fourth ACM International Conference on Web Search and Data Mining, ACM, 2011.
- [18] 문지현, 임성택, 박차라, 이인성, 김진우, “사용자 경험에 대한 HCI 적 관점에서의 개념적 고찰,” 한국 HCI 학회 논문지, 제3권, 제1호, pp.9-17, 2008.
- [19] 이인성, 최보름, 김진우, 이기호, 정승기, “가치 중심적 HCI 를 위한 새로운 방법론의 개발 및 적용,” 한국 HCI 학회 논문지, 제2권, 제1호, pp.13-24, 2007.
- [20] 장진철, 김현지, 박유경, 배동환, 이문용, “경험 샘플링 방법을 이용한 스마트폰 사용자 경험 요소 도출,” 한국 HCI 학회 학술대회, pp.249-252, 2014.
- [21] J. A. Bargas-Avila and K. Hornbæk, “Old wine in new bottles or novel challenges: a critical analysis of empirical studies of user experience,” Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, 2011.
- [22] M. Hassenzahl and N. Tractinsky, “User experience—a research agenda,” Behaviour & Information Technology, Vol.25, No.2, pp.91-97, 2006.
- [23] E. L. C. Law, V. Roto, M. Hassenzahl, A. P. Vermeeren, and J. Kort, “Understanding, scoping and defining user experience: a survey approach,” Proceedings of the SIGCHI Conference on Human Factors in Computing

- Systems, ACM, 2009.
- [24] 변대호, “감정표현어를 이용한 스마트 TV 의 사용자경험 평가,” 한국콘텐츠학회논문지, 제15권, 제5호, pp.132-141, 2015.
- [25] B. Pang and L. Lee, “Opinion mining and sentiment analysis,” *Foundations and Trends in Information Retrieval*, Vol.2, No.1-2, pp.1-135, 2008.
- [26] 최민영, 장혜진, 이호승, “생활가전제품 디자인을 위한 온라인 리뷰 분석의 효용성 연구,” 디지털디자인학연구, 제11호, 제4권, pp.185-194, 2011.
- [27] J. H. Kim, H. S. Byeon, and S. H. Lee, “Enhancement of User Understanding and Service Value Using Online Reviews,” *The Journal of Information Systems*, Vol.20, No.2, pp.21-36, 2011.
- [28] 윤상혁, 손지현, 고민삼, 김영걸, “SNS 온라인 리뷰를 활용한 TV 프로그램 품질평가 연구,” 방송통신연구, 제90권, 제4호, pp.42-73, 2015.
- [29] 변대호, “감정표현어를 이용한 스마트 TV 의 사용자경험 평가,” 한국콘텐츠학회논문지, 제15권, 제5호, pp.132-141, 2015.
- [30] Anol Bhattacharjee, *Social science research: principles, methods, and practices*, Textbooks Collection, 2012.
- [31] 조승연, 김현구, 김범수, 김희웅, “영화 흥행성과 예측을 위한 온라인 리뷰 마이닝 연구: 개봉 첫 주 온라인 리뷰를 활용하여,” *Information Systems Review*, 제16권, 제3호, pp.113-134, 2014.
- [32] 조승연, 최지은, 이규현, 김희웅, “고객 온라인 구매후기를 활용한 추천시스템 개발 및 적용,” 한국경영정보학회 학술대회, pp.157-164, 2015.
- [33] D. M. Blei, A. Y. Ng, and M. I. Jordan, “Latent dirichlet allocation,” *the Journal of Machine Learning Research*, Vol.3, No.4, pp.993-1022, 2003.
- [34] 강범일, 송민, 조화순, “토픽 모델링을 이용한 신문 자료의 오피니언 마이닝에 대한 연구,” 한국문헌정보학회지, 제47권, 제4호, pp.315-334, 2013.
- [35] <http://www.amazon.com/Amazon-SK705DI-Echo/dp/B00X4WHP5E#compare>.
- [36] 김승우, 김남규, “오피니언 분류의 감성사전 활용효과에 대한 연구,” *지능정보연구*, 제20권, 제1호, pp.133-148, 2014.
- [37] D. Meyer, K. Hornik, and I. Feinerer, “Text mining infrastructure in R,” *Journal of Statistical Software*, Vol.25, No.5, pp.1-54, 2008.
- [38] J. Chang and M. J. Chang, “Package ‘lda,’” Citeseer, 2010.
- [39] J. Chang, S. Gerrish, C. Wang, J. L. Boyd-Graber, and D. M. Blei, “Reading tea leaves: How humans interpret topic models,” in *Advances in Neural Information Processing Systems*, 2009.
- [40] K. P. Burnham and D. Anderson, *Model selection and multi-model inference: A Practical information-theoretic approach*, Springer, 2003.
- [41] H. Akaike, “A new look at the statistical model identification,” *IEEE Transactions On Automatic Control*, Vol.19, No.6, pp.716-723, 1974.

저자 소개

황해정(Hae Jeong Hwang)

준회원



- 20013년 8월 : 연세대학교 철학 전공(학사)
- 2016년 8월 : 연세대학교 정보대학원 UX/콘텐츠(석사)

<관심분야> : HCI, UX 디자인, 사용자 경험 평가, 데이터 분석

심혜린(Hye Rin Shim)

정회원



- 1996년 : 이화여자대학교 신문방송학 전공(학사)
- 2002년 : 플로리다 주립대학(게인스빌) PR전공(석사)
- 2015년 ~ 현재 : 연세대학교 정보대학원 UX 트랙(박사과정)

<관심분야> : 언어인지 UX디자인, 융합 콘텐츠 기획

최준호(Junho Choi)

정회원



- 1995년 : 연세대학교 신문방송학과(석사)
- 2002년 : 뉴욕주립대학(버팔로) 커뮤니케이션학(박사)
- 2002년 ~ 2006년 : Rensselaer Polytechnic Institute. Department

of Language, Literature, & Communication 조교수

- 2006년 9월 ~ 2009년 2월 : 광운대학교 미디어영상학부 부교수
- 2009년 3월 ~ 현재 : 연세대학교 정보대학원 UX 트랙 부교수

<관심분야> : 융합 콘텐츠 기획, HCI, UX 디자인