



# 패션 챗봇의 품질이 챗봇 신뢰 및 브랜드 신뢰에 미치는 영향

정 훈 실 · 김 영 인<sup>+</sup>

연세대학교 대학원 생활디자인학과 박사과정 · 연세대학교 생활디자인학과 교수<sup>+</sup>

## The Effect of Chatbot Quality on Chatbot Trust and Brand Trust

Jeong, Hoon Sil · Kim, Young In<sup>+</sup>

Doctoral Course, Yonsei University, Dept. of Human Environment & Design

Professor, Yonsei University, Dept. of Human Environment & Design<sup>+</sup>

(received date: 2018. 8. 16, revised date: 2019. 2. 11, accepted date: 2019. 2. 12)

### ABSTRACT

This study aims to develop a quality assessment tool to evaluate fashion chatbots' quality and factors affecting brand trust. The results of this study are summarized as follows. First, two qualities of fashion chatbots are classified: the quality of the process of interaction and information provision and the quality of the system and design environment. To assess the quality of each fashion chatbot, an assessment tool consisting of 19 questions was developed and shown to have a high reliability. Second, the elements that affect the quality of the fashion chatbot, such as interactions, information, system, and design quality, had a positive effect on chatbot trust and design quality exerted a positive influence on brand trust. In addition, chatbot trust had a positive effect on brand trust. Third, the analysis found a fashion chatbot's quality to have different impacts on chatbot trust and brand trust, which was affected by 'the purpose of using a fashion chatbot' and 'the user's perception on the fashion chatbot'. The study is meaningful in that it provides a scientific basis for fashion chatbots.

Key words: brand trust(브랜드 신뢰), chatbot trust(챗봇 신뢰), chatbot quality(챗봇 품질), fashion chatbot(패션 챗봇)

---

본 논문은 박사학위 청구논문의 일부임

본 연구는 2017년 연세대학교 대학원 연구장학금 지원에 의한 것임

This research was partially supported by the Graduate School of YONSEI University Research Scholarship Grants in 2017.

Corresponding author: Kim, Young In, e-mail: [youngin@yonsei.ac.kr](mailto:youngin@yonsei.ac.kr)

## I. 서론

챗봇(Chatbot)이란 사람들이 일상에서 쓰는 언어로 동작하는 대화형 시스템 혹은 서비스를 의미한다. 2014년부터 구글과 페이스북과 같은 거대 IT 기업들은 챗봇을 차세대 주력산업으로 삼아 치열한 플랫폼 경쟁을 시작했다.

변화에 민감한 패션 분야는 이와 같은 흐름을 놓치지 않고 챗봇을 개인화된 고객 점점 채널로 활용하려는 움직임을 보이고 있다. 2016년 타미 힐피거가 패션 브랜드 최초로 챗봇을 선보인 이래로 버버리, 루이비통과 같은 명품 브랜드들까지 가세하여 패션 분야 전반으로 퍼지고 있다.

온라인 상거래 분야도 패션과 챗봇을 결합하는 것에 대한 관심이 지대하다. 예를 들면 중국의 온라인 상거래 업체인 알리바바는 2017년 11월 서비스를 시작한 'Fashion AI'를 통해 온라인 쇼핑물과 오프라인 매장을 연결하여 주문·결제를 지원하고 있으며, 국내의 경우 롯데백화점의 인터넷 쇼핑몰 엘롯데닷컴이 패션 용어를 학습하고 사용자의 취향에 맞는 패션 제품을 추천하는 '로사'를 2018년 1월부터 정식으로 서비스하기 시작했다.

그러나 현재 대부분의 챗봇은 사용자의 기대 수준을 만족시키지 못하고 있다(Coniam, 2014). 이는 챗봇에 대해 양적인 팽창보다 질적인 검증이 필요하다는 것을 의미한다. 고객 개개인에 대해 맞춤형 추천 서비스를 제공하는 패션 챗봇은 높은 수준의 품질이 더욱 필수적이다.

챗봇에 대한 연구는 자연스러운 대화를 구현하기 위한 기술에 초점을 맞추거나(Cho, Choi & Park, 2009; Deshpande et al., 2017) 챗봇과의 상호작용이 미치는 영향을 규명하기 위한 연구가 주로 이루어지고 있다(Hong, 2016; Koo et al., 2016). 챗봇의 품질에 대한 연구는 소수이며, 대부분 단편적인 측면으로 품질을 평가할 것을 제안하고 있다(Goh, Ardil & Fung, 2007; Vetter, 2002). 하지만 챗봇은 사용자의 목적을 달성하기 위해 정

보, 환경 등 다양한 조건들이 관여하기 때문에 다차원적인 품질평가도구가 필요하다. 특히 디자인을 제안하고 추천하며, 구입의 과정을 돕는 등 사용자와 밀착된 서비스를 제공하는 패션 챗봇의 경우 다차원적인 품질평가도구를 통한 품질 관리가 더욱 중요하다.

따라서 이 연구는 패션 챗봇의 품질평가도구를 개발하고, 챗봇 신뢰와 브랜드 신뢰에 영향을 주는 챗봇의 품질과 원인을 규명하고자 하였다.

## II. 이론적 배경

### 1. 챗봇 품질의 정의

챗봇의 품질에 대한 선행연구들은 챗봇과 사용자의 대화의 품질 및 정보, 사용성과 챗봇의 속성을 기준으로 챗봇의 품질을 평가하였다. 챗봇과의 대화를 기준으로 품질을 정의한 Vetter(2002)는 챗봇과의 커뮤니케이션의 성공 및 실패에 대한 만족도를 기준으로 챗봇의 품질을 평가해야 한다고 주장하였으며, Goh, Ardil & Fung(2007)과 Kosir(2013)은 사용자의 질문에 챗봇이 얼마나 유용한 대답을 제공하는가를 기준으로 품질을 평가해야 한다고 주장하였다. Coniam(2014)은 대화, 혹은 언어의 가장 기본적인 품질이라고 할 수 있는 문법과 의미의 정확도로 챗봇의 품질을 평가해야 한다고 주장하였다. Meira & Canuto(2015)는 응집성, 결합성 등의 평가지표로 챗봇의 정보 구조 설계를 평가하며, 챗봇과의 상호작용에서 호감도, 지적인 능력을 기준으로 챗봇의 품질을 평가할 것을 제안했다. Kuligowska(2015)는 챗봇의 외관, 개성과 같은 속성과 사용성 등 10가지의 품질평가요인을 제시하였다. Nicole & Morgan(2017)은 챗봇의 윤리성을 중요한 품질평가요인으로 제안하였다.

그러나 패션 챗봇은 대화뿐만 아니라 정보, 시스템 환경, 디자인 환경 등 여러 조건들이 관여하기 때문에 단편적인 측면으로 품질을 평가하기에는 한계가 있다. 이에 이 연구는 챗봇을 '서비스를

제공하는 소프트웨어'의 관점에서 바라보고 이와 관련된 선행연구를 함께 고찰하였다.

서비스의 품질은 소비자가 기대한 품질과 실제로 지각한 품질간의 차이에서 인식되는 품질을 의미한다(Gronroos, 1984). 서비스 품질을 평가할 때 가장 널리 사용되는 연구는 Parasuraman, Zeithaml & Berry(1988)의 5차원 모형이다. 5차원 모형은 유형성, 신뢰성, 반응성, 공감성, 확신성의 성과와 기대 간의 차이로 서비스 품질을 평가한다. Gronroos(1984)는 고객이 서비스를 통해 얻게 되는 결과를 의미하는 기술적 품질과 고객들이 어떠한 과정을 거쳐서 결과를 얻게 되는가를 의미하는 기능적 품질의 2차원 모형을 제시하였다. Brady & Cronin(2001)은 기술적 품질과 기능적 품질과 함께 서비스가 전달되는 환경의 품질을 추가하여 3차원 모형을 제시하였다.

패션 챗봇이 고객의 기호에 맞는 디자인을 제시하고 선택을 도우며, 이를 구입할 수 있도록 도울 때 제공하는 정보의 질적 수준은 '정보 품질'의 관점에서 바라보아야 한다. 정보의 품질이란 '정보 시스템의 결과물'로서의 품질이다(DeLone & McLean, 1992). Bailey & Pearson(1983)은 정확성, 적시성, 간결성, 완전성, 현재성, 구성 방식, 관련성, 신뢰성으로 정보의 품질을 평가하였다. Srinivasan(1985)은 정확성, 적시성, 이해성, 완전성으로 정보의 품질을 평가하였다. Parker & Case(1991)은 정확성, 적시성, 이용가능성, 이해성, 관련성, 유용성, 신뢰성, 일관성으로 정보의 품질을 평가하였다.

소프트웨어의 품질은 '시스템'과 '인터페이스 디자인', 그리고 '소프트웨어와 사용자의 인터랙션'에 대한 품질로 구분할 수 있다. 챗봇과 사용자는 대화를 통해 인터랙션 하기 때문에 인터랙션을 제외하고 시스템과 인터페이스 디자인의 품질에 대한 선행연구를 추가로 고찰하였다.

시스템 품질은 시스템을 사용하는 과정에 관련된 품질이다(Etezadi-Amoli & Farhoomand, 1996). Bailey & Pearson(1983)은 시스템의 정확성, 신뢰

성, 적시성, 관련성, 접근성으로 품질평가요인을 정의했다. Etezadi-Amoli & Farhoomand(1996)는 시스템의 도움말, 사용성, 기능성, 결과물의 품질, 지원 정도, 보안성이 사용자의 수행 능력을 높이는 직접적인 요인임을 규명하였다. DeLone & McLean(2003)은 정보시스템의 바람직한 특성을 정확성, 신뢰성, 사용성, 직관성, 학습성, 융통성, 유용성, 응답 시간으로 정의하였다.

인터페이스는 기능적 인터페이스와 감성적 인터페이스로 분류된다. 기능적 인터페이스는 시스템 인터페이스와 사용성과 관계되므로 감성적 인터페이스에 대한 선행연구를 추가로 고찰하였다. Lavie & Tractinsky(2004)는 웹사이트 인터페이스의 미학적인 측면을 측정하는 척도를 제안하였다. 그 결과 질서 있고 명확한 아름다움을 의미하는 '고전적 미학'과 디자이너의 창의성과 독창성에서 비롯된 '표현적 미학'으로 나누어 감성적 인터페이스를 평가하는 척도를 개발하였다.

이상의 선행연구들에서 챗봇의 품질은 '서비스가 전달되는 과정에 대한 품질'과 '서비스의 결과에 대한 품질', '서비스가 전달되는 배경에 대한 품질'로 분류할 수 있다는 것을 확인하였다. 서비스가 전달되는 과정은 챗봇과의 대화 인터랙션을 통한 '상호작용'과 '제공정보'의 품질에 대한 평가가 필요하며, 챗봇을 사용하는 물리적 환경으로 '시스템'과 '감성적 인터페이스 디자인'의 품질에 대한 평가가 필요하다는 것을 확인하였다.

상호작용의 품질을 평가하기 위한 세부적인 평가요인은 대화 능력, 대화의 효용성 및 효과성, 언어의 정확성이며, 제공정보의 품질을 평가하기 위한 세부적인 평가요인은 정보의 정확성, 적시성, 이해성, 완전성, 신뢰성, 사용자 목표와의 관련성으로 분석하였다. 시스템의 품질을 평가하기 위한 세부적인 평가요인은 정확성, 신뢰성, 사용성이며, 감성적 인터페이스 디자인의 품질을 평가하기 위한 세부적인 평가요인은 고전적 미학과 표현적 미학으로 분석하였다.

그러므로 이 연구는 챗봇의 품질을 상호작용, 제공정보, 시스템, 감성적 인터페이스 디자인을 분류하여 세부적으로 분석하였다.

## 2. 신뢰의 정의 및 특성

챗봇 신뢰(Chatbot trust)와 브랜드 신뢰(Brand trust)를 정의하기 위해 신뢰에 대한 선행연구를 고찰하였다. 선행연구들은 신뢰를 '개인의 심리적 상태'로 규정짓고 있다는 공통점이 있다(Moorman et al., 1993; Sirieix & Dubois, 1999). 개인의 심리적인 상태로 보는 관점의 연구 중 가장 많이 활용되고 있는 정의는 Mayer, Davis & Schroorman (1995)의 정의이므로 이 연구는 그의 정의를 맥락에 맞게 수정하여 챗봇 신뢰를 '챗봇이 자신에게 중요한 행동을 할 것이라는 긍정적인 기대를 바탕으로 챗봇을 믿고자 하는 의도'로 정의하였다.

브랜드 신뢰에 대한 선행연구들은 공통적으로 브랜드 신뢰를 '브랜드에 대한 소비자의 인지적 믿음'으로 정의하고 있으며, 브랜드 신뢰가 브랜드에 대한 몰입을 유도하여 더욱 안정적이고 지속적인 관계를 유지한다는 공통된 결론을 도출하고 있다(Bradach & Eccles, 1989; Everard & Galletta, 2006). 그러므로 이 연구는 브랜드 신뢰를 '브랜드에 대한 인지적 믿음'으로 정의하였다.

다수의 연구에서 품질이 신뢰에 긍정적인 영향을 미치고 장기적인 관계를 유지하는데 중요한 역할을 하고 있음을 규명하였다(Swan et al., 1985; Doney & Cannon, 1997). 또한 온라인 서비스의 품질이 고객의 신뢰에 영향을 미친다는 연구도 상당수 있다(Yoo, Na & Choi 1999; Gefen et al., 2003). 즉 신뢰를 형성하고 유지하는데 제품이나 서비스의 품질이 중요한 역할을 한다.

그리고 신뢰는 전이될 수 있다(Jacoby et al., 1971; Stewart, 2003). Stewart(2003)은 온라인 웹사이트에서 링크를 통해 조직 간에 신뢰가 전이될 수 있다는 것을 증명하였으며, Lee et al.(2005)는 온라인과 온라인, 오프라인과 오프라인처럼 동일

한 채널은 물론 온라인과 오프라인 사이의 상이한 채널에서도 신뢰가 전이될 수 있음을 밝혔다.

그러므로 이 연구에서는 챗봇의 품질이 챗봇 신뢰와 브랜드 신뢰에 영향을 미치는지를 확인하며, 챗봇 신뢰가 브랜드 신뢰로 전이되는지를 확인하고자 하였다.

## Ⅲ. 연구방법

### 1. 패션 챗봇의 품질평가도구 개발 및 검증 방법

이 연구는 Netemeyer et al.(2003)의 4단계 척도 개발 모형을 기반으로 패션 챗봇의 품질평가도구를 개발하였다. 첫 번째는 '구인 정의 및 내용 영역 확인' 단계이며 패션 챗봇의 품질평가요인을 도출하기 위해 챗봇, 서비스, 소프트웨어의 품질에 대한 선행연구를 종합적으로 고찰하였다.

두 번째는 '측정 문항 개발 및 검토' 단계이며 패션 챗봇의 품질을 평가하는 문항을 도출하기 위해 지각된 품질을 상황별로 구분하였다. 상황별로 구분한 품질평가요인을 바탕으로 챗봇, 서비스, 소프트웨어의 품질에 관한 선행연구를 재분류하고 품질을 평가하는 문항을 추출하여 정제하였다. 개발된 문항은 교수 및 챗봇과 관련된 일을 하고 있는 경력 10년차 이상의 전문가 2인의 의견을 수렴하여 수정 및 보완하였다.

세 번째는 '척도개발 및 개선'의 단계로 설문조사를 실시하고 응답 결과를 통계적으로 분석하여 평가도구를 점검하였다. 그리고 요인분석을 통해 신뢰도를 확인하였다.

네 번째는 '척도 점검 및 확장'의 단계로 패션 챗봇의 품질을 평가하는 영역이 적절한지, 각 영역을 평가하기 위한 문항이 타당한지를 판단하고자 교수 및 챗봇과 관련된 일을 하고 있는 경력 10년차 이상의 전문가 5인의 내용타당성 검사를 실시하였다.

## 2. 패션 챗봇의 품질이 브랜드 신뢰에 작용하는 영향 규명 방법

이 연구는 온라인에서 제공하는 서비스의 품질이 신뢰에 긍정적인 영향을 준다는 선행연구를 바탕으로, 패션 챗봇의 품질이 챗봇 신뢰와 브랜드 신뢰에 긍정적인 영향을 줄 것이라는 가설을 세웠다. 또한 신뢰가 전이된다는 선행연구를 바탕으로 챗봇을 통해 형성된 신뢰가 브랜드 신뢰에 긍정적인 영향을 미칠 것이라는 가설을 정립하였다. 선행연구 고찰을 통해 챗봇의 품질을 상호작용, 제공정보, 시스템, 감성적 인터페이스 디자인의 측면에서 분석하기로 하였으므로 이 연구에서 규명하고자 하는 가설을 정리하면 다음과 같다.

### 가설 1. 챗봇의 품질은 챗봇 신뢰에 긍정적인 영향을 줄 것이다.

- 1-1. 챗봇의 상호작용 품질은 챗봇 신뢰에 긍정적인 영향을 줄 것이다.
- 1-2. 챗봇의 정보 품질은 챗봇 신뢰에 긍정적인 영향을 줄 것이다.
- 1-3. 챗봇의 시스템 품질은 챗봇 신뢰에 긍정적인 영향을 줄 것이다.
- 1-4. 챗봇의 디자인 품질은 챗봇 신뢰에 긍정적인 영향을 줄 것이다.

### 가설 2. 챗봇의 품질은 브랜드 신뢰에 긍정적인 영향을 줄 것이다.

- 2-1. 챗봇의 상호작용 품질은 브랜드 신뢰에 긍정적인 영향을 줄 것이다.
- 2-2. 챗봇의 정보 품질은 브랜드 신뢰에 긍정적인 영향을 줄 것이다.
- 2-3. 챗봇의 시스템 품질은 브랜드 신뢰에 긍정적인 영향을 줄 것이다.
- 2-4. 챗봇의 디자인 품질은 브랜드 신뢰에 긍정적인 영향을 줄 것이다.

### 가설 3. 챗봇 신뢰는 브랜드 신뢰에 긍정적인

## 영향을 줄 것이다.

가설을 증명하기 위해 챗봇에 대해 긍정적이고 적극적이며 호의적인 태도를 가지고 있으며, 향후 챗봇의 주 사용자로 전망되는 20-39세 성인 남녀를 대상으로 2018년 5월 4일에서 5월 23일까지 온라인으로 설문조사를 진행하였다. 설문조사는 연구대상으로 선정된 패션 챗봇으로 제시된 과제를 수행한 뒤 질문에 답을 하도록 하였다.

설문조사 결과 국내 패션 챗봇 2가지에 대한 응답을 각각 158케이스, 국외 패션 챗봇 2가지에 대한 응답을 각각 56케이스씩 총 428케이스를 확보하였다. 설문결과는 통계 프로그램 SPSS 21.0을 이용하여 요인분석, 다중회귀분석과 위계적 회귀분석을 실시하였다.

연구대상 패션 챗봇을 선정하기 위해 패션 챗봇이 제공하는 중요한 기능인 개인화 추천 기능을 제공하는 챗봇 중 사용자 수가 많은 국내 챗봇과 국외 챗봇을 2개씩 선정하였다. 사용자 수는 '회원수'나 '좋아요 수'를 기준으로 하였다. 그 결과 국내 챗봇은 2015년 기준 회원수가 2000만명인 인터파크의 '특집사'와 2016년 기준 회원수 1700만명인 엘롯데의 '로사'를 선정하였고, 국외 챗봇은 2018년 5월 기준 좋아요 수 3.4만명인 eBay의 'Shopbot'과 2018년 5월 기준 좋아요 수 925명인 'Mode.ai의 챗봇'을 선정하였다.

챗봇 신뢰를 측정하기 위한 항목은 다양한 서비스 산업을 대상으로 높은 신뢰도와 타당도를 가진 연구 결과를 낸 Sindeshmukh, Singh & Sabol (2002)의 측정도구를 사용하여 '이 챗봇은 유능하다', '이 챗봇은 믿음직하다', '이 챗봇은 문제없이 잘 작동한다', '이 챗봇은 고객의 문의에 잘 반응한다'로 조작하였다.

브랜드 신뢰를 측정하기 위한 항목은 Chaudhuri & Holbrook(2001)의 문항이 여러 후속연구에 의해 신뢰도가 검증되고 있으므로, 측정 문항을 연구의 맥락에 맞게 수정하여 '나는 이 브랜드를 신

뢰한다', '나는 이 브랜드를 믿는다', '나는 이 브랜드가 정직하다고 생각한다', '나는 이 브랜드를 안심하고 이용할 수 있다'로 조작하였다.

## 2. 브랜드 신뢰에 작용하는 품질의 차이와 원인 규명 방법

대표 패션 챗봇별로 품질이 브랜드 신뢰에 작용하는 영향에 차이가 있는지를 비교 분석하기 위해 통계 프로그램 SPSS 21.0을 이용하여 빈도분석, t-test, ANOVA, 다중회귀분석을 실시하였다. 그 다음 패션 챗봇별로 브랜드 신뢰에 영향을 주는 품질이 다르게 나타나는 이유를 분석하기 위해 수정델파이 조사(Modified Delphi)를 실시하였다.

이 연구는 사용자와 공급자의 의견을 고루 반영하기 위해 챗봇을 사용한 경험이 있는 20-39세의 사용자 3명과 현재 챗봇과 관련된 서비스를 제공하는 기업에서 일하고 있는 전문가 3명으로 핵심 패널과 전문가 그룹을 구성하였다. 전문가는 해당 분야에서 10년 이상의 경력을 가진 사람을 선정하였다.

수정 델파이는 총 2회에 걸쳐 진행하였다. 1차 조사는 2018년 6월 10일 실시하였으며 전문가들이 한자리에 모여 토론하는 방식으로 진행하였다. 2차 조사는 2018년 6월 18일에 실시하였으며, 사용자와 전문가가 모두 참여하였다. 조사는 모든 패널과 전문가들이 한자리에 모여 패션 챗봇에 대해 자유롭게 토론하는 방식으로 진행하였다.

## IV. 연구결과

### 1. 패션 챗봇의 품질평가도구

Netemeyer et al.(2003)의 4단계 척도 개발 모형을 바탕으로 패션 챗봇의 품질평가도구를 개발하였다. 첫 번째 '구인 정의 및 내용 영역 확인' 단계에서 챗봇의 품질, 서비스 품질, 소프트웨어의 품질에 대한 선행연구를 종합적으로 분석하였다.

그 결과 패션 챗봇의 품질은 '과정의 품질'과 '환경의 품질'로 정리할 수 있으며, 과정의 품질은 '상호작용 품질'과 '제공정보의 품질'로 구성되고, 환경의 품질은 '시스템의 품질'과 '디자인의 품질'로 구성하였다.

두 번째 '측정 문항 개발 및 검토' 단계에서는 챗봇의 품질과 관련된 선행연구들을 상호작용, 제공정보, 시스템, 디자인으로 재분류하여 해당 품질을 측정할 수 있는 문항을 추출하여 정제하였다.

세 번째 '척도 개발 및 개선'의 단계를 진행하기 위해 설문조사를 진행하고 요인 및 신뢰성 분석을 실시하였다. 패션 챗봇의 품질을 평가하기 위해 개발된 문항의 내적 일치도를 검증한 결과는 <Table 1>과 같다.

요인 1은 상호작용으로 총 5문항이며 설명변량 67.90%, Cronbach's  $\alpha$  값 0.87으로 나타났고, 요인 2는 제공정보으로 총 5문항이며 설명변량 78.63%, Cronbach's  $\alpha$  값 0.87으로 나타났다. 요인 3은 시스템으로 총 4문항이며 설명변량 76.16%, Cronbach's  $\alpha$  값 0.93으로 나타났고, 요인 4는 디자인으로 총 5문항이며 설명변량 75.62%, Cronbach's  $\alpha$  값 0.90으로 나타나 모든 요인에 대해 신뢰도가 검증되었다.

마지막으로 '척도 점검 및 확장'의 단계를 진행하기 위해 전문가 5인을 대상으로 내용타당성 검사를 실시한 결과 모든 전문가가 패션 챗봇의 품질을 평가하는 영역과 평가 문항의 적절성과 타당성을 인정하였다.

### 2. 패션 챗봇의 품질이 챗봇 신뢰 및 브랜드 신뢰에 작용하는 영향

#### 1) 종속변수의 신뢰도 검증

챗봇 신뢰와 브랜드 신뢰에 대한 요인 및 신뢰성 분석을 실시하였다. 그 결과는 <Table 2>와 같이 요인 1은 챗봇 신뢰로 총 4문항이며 설명변량 82.76%, Cronbach's  $\alpha$  값 0.93으로 나타났고, 요인 2는 브랜드 신뢰로 총 4문항이며 설명변량 99.89%, Cronbach's  $\alpha$  값 0.96으로 나타나 모든 요인에 대

<Table 1> Verification of Internal Conformity of Fashion Chatbot's Quality Evaluation Tool

Factor	Question	Factor Loading	Eigen Value	Variance (%)	Cronbach's $\alpha$
1 interaction	The chatbot conducted a coherent conversation.	.879	3.40	67.90	0.87
	The chatbot answered the question appropriately.	.845			
	The chatbot answered the question quickly.	.830			
	The chatbot used grammatically correct language.	.820			
	The chatbot accurately expressed.	.741			
2 information	The chatbot provided information that was easy to understand.	.922	3.93	78.63	0.87
	The chatbot provided accurate information.	.916			
	The chatbot provided highly complete information.	.898			
	The chatbot provided up-to-date information.	.849			
	The chatbot provided useful information.	.846			
3 system	Chatbot systems are easy to use.	.892	3.05	76.16	0.93
	The use of chatbot systems is easy to learn.	.887			
	Chatbot system can be used anywhere, anytime.	.871			
	Chatbot systems have features that help me do my job.	.839			
4 design	The chatbot designs are organized.	.916	3.78	75.62	0.90
	The chatbot's design is clean.	.888			
	The chatbot's design is entertaining.	.862			
	Chatbot's design is creative.	.846			
	The effect applied to chatbot designs is unique.	.833			

한 신뢰도가 검증되었다.

2) 가설 검증 결과

챗봇의 품질이 챗봇 신뢰와 브랜드 신뢰에 어

떤 영향력을 미치는가에 대한 연구 가설을 검증하기 위해 다음과 같은 분석을 실시하였다. 종속변수인 챗봇 신뢰, 브랜드 신뢰의 변화를 설명하기 위해 독립변수들을 투입하여 선형회귀분석을 실시

<Table 2> Verification of Internal Conformity of Chatbot trust and brand trust

Factor	Question	Factor Loading	Eigen Value	Variance (%)	Cronbach's $\alpha$
1 Chatbottrust	This chatbot is capable.	.924	3.31	82.76	0.93
	This chatbot is reliable.	.921			
	This chatbot works well without problems.	.902			
	This chatbot responds well to users' inquiries.	.892			
2 brand trust	I trust this brand.	.955	3.55	88.89	0.96
	I believe in this brand.	.951			
	I think this brand is honest.	.934			
	I can uwe this brand with confidence.	.931			

<Table 3> Correlation between chatbot quality factors

index	interaction	information	system	design
interaction	1			
information	0.79***	1		
system	0.75***	0.75***	1	
design	0.54***	0.57***	0.64***	1

\*\*\* $p < .001$  Verification on both sides

하고, 여러 독립변수를 순차적으로 투입한 뒤 독립변수들의 상대적인 효과를 알아보기 위해 위계적 다중회귀분석을 최종으로 실시하였다.

독립변수로 고려한 상호작용, 제공정보, 시스템, 디자인 요인 간의 상관관계는 <Table 3>과 같다. 그 결과 다중공선성이 의심되는 독립변수들이 없어 모두 회귀분석에 투입하였다.

사용자가 지각하는 패션 챗봇의 품질이 챗봇 신뢰에 긍정적인 영향을 미칠 것이라는 가설 1의 검정 결과 <Table 4>와 같이 나타났다. 상호작용이 챗봇 신뢰에 미치는 영향은  $t$ 값이 6.03,  $\beta$ 값이 0.25로 나타나 가설 1-1은 채택되었다. 제공정보의  $t$ 값은 9.43,  $\beta$ 값이 0.39로 나타나 가설 1-2가 채택되었다. 시스템의  $t$ 값은 4.91,  $\beta$ 값은 0.20으로 나타나 가설 1-3 역시 채택되었다. 디자인의  $t$ 값은 5.03,  $\beta$ 값은 0.16으로 가설 1-4도 채택되었다. 회귀모형은  $F$ 값이  $p = .000$ 에서 362.29의 수치를 보이고 있으며, 회귀식에 대한  $R^2 = .77$ 로 77%의 높

은 설명력을 보이고 있다.

패션 챗봇의 품질이 브랜드 신뢰에 긍정적인 영향을 미칠 것이라는 가설 2와 챗봇 신뢰가 브랜드 신뢰에 긍정적인 영향을 미칠 것이라는 가설 3에 대한 검정 결과 <Table 5>와 같이 나타났다. 상호작용이 브랜드 신뢰에 미치는 영향은  $t$ 값이 0.34로 가설 2-1은 기각되었다. 제공정보의  $t$ 값은 0.66으로 나타나 가설 2-2도 기각되었다. 시스템의  $t$ 값은 1.67로 가설 2-3 역시 기각되었다. 디자인의  $t$ 값은 4.88,  $\beta$ 값은 0.26으로 가설 2-4는 채택되었다.

이와 함께 챗봇 신뢰와 브랜드 신뢰의 관계를 보면  $t$ 값이 2.45,  $\beta$ 값은 0.15로 나타나 가설 3은 채택되었다. 회귀모형은  $F$ 값이  $p = .000$ 에서 43.51의 수치를 보이고 있으며, 회귀식에 대한  $R^2 = .34$ 로 34%의 설명력을 보이고 있다. 설명력이 다소 높지 않은 이유는 연구 대상 챗봇들이 도와주는 기능을 위주로 제공하거나 브랜드를 대표하는 소비자 접점 채널로 제공하는 경우로 나뉘기 때문인

<Table 4> Regression with chatbot trust by chatbot quality factors

index	standard error	$\beta$	$t$	Probability of significance
interaction	.05	0.25***	6.03	.000
information	.04	0.39***	9.43	.000
system	.04	0.20***	4.91	.000
design	.03	0.16***	5.03	.000
$R^2$	0.77			
$F$	362.29***			

\*\*\* $p < .001$  Verification on both sides



〈Table 5〉 Hierarchical analysis results for brand trust

구분		Model 1		Model 2		Model 3	
		$\beta$	<i>t</i>	$\beta$	<i>t</i>	$\beta$	<i>t</i>
process	interaction	0.21**	3.09	0.08	1.07	0.03	0.34
	information	0.31***	4.49	0.13	1.84	0.05	0.66
environment	system			0.16*	2.29	0.12	1.67
	design			0.30***	5.59	0.26***	4.88
Chatbot trust						0.20*	2.45
$R^2$		0.25		0.33		0.34	
$R^2$ Change		-		0.08		0.01	
$F$		69.95***		52.27***		43.51***	

\* $p < .05$  \*\* $p < .01$  \*\*\* $p < .001$  Verification on both sides

것으로 볼 수 있다. 도와주는 기능 위주로 제공하는 챗봇의 경우 사용자는 브랜드와 연관지어 인식할 확률이 낮기 때문이다.

이상의 결과를 바탕으로 연구가설의 검증 결과를 정리하면 〈Table 6〉과 같다. 패션 챗봇의 품질은 챗봇의 신뢰에 긍정적인 영향을 주며, 챗봇의 품질 중 디자인 품질만 브랜드 신뢰에 긍정적인 영향을 주었다. 또한 챗봇 신뢰는 브랜드 신뢰에 긍정적인 영향을 주는 것으로 나타났다.

### 3) 응답자의 일반적 특성

이 연구의 대상은 영어 능통자 56명, 일반대상자 102명으로 총 158명이 대답하였다. 응답자의 기본적 특성으로 〈Table 7〉과 같이 성별과 연령을 고려하였다. 성별은 남자가 45.6%(72명)이고, 여자는 54.4%(86명)이었다. 연령별로는 20대가 43.7%(69명), 30세 이상이 56.3%(89명)이었다.

〈Table 6〉 Hypothesis adoption status

	hypothesis	adoption status
<b>H1</b>	<b>Quality of chatbot will have a positive effect on chatbot trust.</b>	<b>adopt</b>
H1-1	The quality of interaction of chatbot will have a positive effect on chatbot trust.	adopt
H1-2	The quality of information of chatbot will have a positive effect on chatbot trust.	adopt
H1-3	The quality of system of chatbot will have a positive effect on chatbot trust.	adopt
H1-4	The quality of design of chatbot will have a positive effect on chatbot trust.	adopt
<b>H2</b>	<b>Quality of chatbot will have a positive effect on brand trust.</b>	<b>partial adopt</b>
H2-1	The quality of interaction of chatbot will have a positive effect on brand trust.	reject
H2-2	The quality of information of chatbot will have a positive effect on brand trust.	reject
H2-3	The quality of system of chatbot will have a positive effect on brand trust.	reject
H2-4	The quality of design of chatbot will have a positive effect on brand trust.	adopt
<b>H3</b>	<b>Chatbot trust will have a positive effect on brand trust.</b>	<b>adopt</b>

<Table 7> General characteristics of respondents

index		Number of persons	%
Gender	men	72	45.6
	women	86	54.4
	<b>total</b>	<b>158</b>	<b>100</b>
Age	20s	69	43.7
	30s	89	56.3
	<b>total</b>	<b>158</b>	<b>100</b>

<Table 8> Quality comparison of chatbot quality factors between chatbot

index	Talkjipsa (A)	Losa(B)	Shopbot(C)	Mode.ai(D)	F
interaction	5.01±1.00(B)	5.62±0.94(A)	5.35±1.21(AB)	5.23±1.04(AB)	9.62***
information	4.51±1.24(C)	5.35±1.10(A)	4.81±1.26(BC)	5.10±1.08(AB)	13.82***
system	4.75±1.20(B)	5.44±1.04(A)	4.91±1.34(B)	5.07±1.17(AB)	9.59***
design	4.46±1.14(B)	5.01±1.17(A)	4.10±1.30(B)	4.50±1.03(B)	10.77***

\*\*\*  $p < .001$

### 3. 브랜드 신뢰에 작용하는 패션 챗봇의 품질의 차이와 원인

#### 1) 브랜드 신뢰에 작용하는 패션 챗봇별 품질의 차이

연구 대상인 패션 챗봇 4개의 품질을 비교한 결과는 <Table 8>과 같다. 국내 챗봇 중 상호작용 품질은 로사가 특집사보다 우수한 것으로 나타났고 통계적으로 유의한 차이가 있었다. 국외 챗봇인 Shopbot과 Mode.ai도 상호작용 품질이 좋은 것으로 나타났다. 챗봇이 제공하는 정보의 품질은 로사, Shopbot, Mode.ai가 특집사보다 우수한 것으로 나타났다.

시스템 품질은 로사와 Mode.ai가 특집사, Shopbot

보다 높게 평가되었고, 통계적으로 유의한 차이가 있는 것으로 나타났다. 디자인 품질은 로사가 다른 3개의 챗봇보다 우수한 것으로 나타났고, 통계적으로 유의한 차이가 있었다.

챗봇 신뢰와 브랜드 신뢰에 대한 챗봇 간 비교 결과는 <Table 9>와 같다. 챗봇 신뢰는 로사, Shopbot, Mode.ai가 특집사보다 높았으며 통계적으로 유의한 차이가 있었다. 브랜드 신뢰는 로사가 다른 챗봇보다 높았으나, 집단 간의 차이는 크지 않았다.

패션 챗봇별로 품질이 브랜드 신뢰에 작용하는 영향의 차이를 규명하기 위해 다중회귀분석을 실시하였다. 먼저 챗봇 신뢰를 종속변수로 분석한

<Table 9> Quality comparison of chatbot trust and brand trust between chatbot

index	Talkjipsa (A)	Losa(B)	Shopbot(C)	Mode.ai(D)	F
Chatbot trust	4.34±1.29(B)	5.16±1.24(A)	4.94±1.18(A)	4.97±1.23(A)	12.02***
brand trust	4.75±1.14	4.81±1.17	4.36±1.18	4.45±1.26	2.94*

\* $p < .05$  \*\*\* $p < .001$

<Table 10> Regression with Chatbot trust by Fashion Chatbot

index	Talkjipsa		Losa		Shoptbot		Mode.ai	
	$\beta$	$t$	$\beta$	$t$	$\beta$	$t$	$\beta$	$t$
interaction	0.21**	3.24	0.29***	4.11	0.23	1.88	0.28*	2.15
information	0.29***	4.33	0.47***	6.85	0.37**	3.63	0.42**	2.75
system	0.33***	4.66	0.05	0.78	0.27*	2.28	0.21	1.64
design	0.17**	3.12	0.19***	3.64	0.19*	2.61	0.06	0.75
$R^2$	0.73		0.80		0.81		0.82	
$F$	104.85***		156.84***		54.42***		56.11***	

\* $p < .05$  \*\* $p < .01$  \*\*\* $p < .001$  Verification on both sides

결과 <Table 10>과 같이 나타났다. 특집사는 상호작용, 제공정보, 시스템, 디자인이 모두 챗봇 신뢰에 영향을 주는 것으로 나타났고, 로사는 상호작용, 제공정보, 디자인만 챗봇 신뢰에 영향을 주는 것으로 나타났다. Shoptbot은 제공정보, 시스템, 디자인이 챗봇 신뢰에 영향을 주었으며, Mode.ai는 상호작용과 제공정보만 챗봇 신뢰에 영향을 주는 것으로 나타났다.

브랜드 신뢰를 종속변수로 분석한 결과는 <Table 11>과 같다. Mode.ai의 경우 상호작용이 브랜드 신뢰에 부정적인 영향을 주는 것으로 나타났는데, 이는 영어를 잘하는 사용자라도 국외의 챗봇을 사용할 때 모국어가 아닌 언어로 상호작용하는 것에 대한 어느 정도 불편함이 있기 때문에 기인

한 결과로 볼 수 있다. 챗봇 신뢰와 브랜드 신뢰의 관계를 살펴보면 특집사와 Shoptbot의 경우 챗봇 신뢰가 브랜드 신뢰에 영향을 주지 않는 것으로 나타났다. 하지만 로사와 Mode.ai는 챗봇 신뢰가 브랜드 신뢰에 영향을 주는 것으로 나타났다.

## 2) 패션 챗봇별로 브랜드 신뢰에 작용하는 품질이 차이나는 원인

패션 챗봇별로 브랜드 신뢰에 영향을 미치는 품질과 챗봇 신뢰가 브랜드 신뢰에 영향을 미치는 여부가 다르게 나타나는 이유를 규명하기 위해 수정 델파이 조사를 진행하였다. 토론 주제는 다음과 같다.

<Table 11> Regression with brand trust by Fashion Chatbot

index	Talkjipsa		Losa		Shoptbot		Mode.ai	
	$\beta$	$t$	$\beta$	$t$	$\beta$	$t$	$\beta$	$t$
interaction	0.13	1.19	0.00	0.01	-0.32	-1.84	-0.47*	-2.70
information	0.11	0.89	-0.15	-1.09	0.37*	2.35	0.11	0.53
system	0.09	0.72	0.08	0.74	0.34*	2.02	0.17	1.02
design	0.19	1.92	0.24*	2.46	0.22*	2.06	0.35**	3.41
Chatbot trust	0.05	0.37	0.49**	3.34	0.27	1.40	0.70***	3.92
$R^2$	0.23		0.37		0.64		0.70	
$F$	9.13***		17.91***		17.78***		23.63***	

\* $p < .05$  \*\* $p < .01$  \*\*\* $p < .001$  Verification on both sides

첫째, 패션 챗봇에 따라 품질이 챗봇 신뢰에 주는 영향이 다른 이유는 무엇인가?

둘째, 패션 챗봇에 따라 품질이 브랜드 신뢰에 주는 영향이 다른 이유는 무엇인가?

모두가 한자리에 모여 논의한 결과 위와 같은 현상이 나타나는 이유에 대해 '패션 챗봇의 제공 목적'과 '패션 챗봇에 대한 사용자의 인식'에서 비롯된 것으로 의견을 모았다.

패션 챗봇의 제공 목적이란 챗봇 서비스를 제공하는 주체인 브랜드나 기업이 챗봇을 어떠한 관점에서 바라보고 어떠한 역할을 부여하는가를 의미한다. 전문가들은 특집사와 Shopbot의 경우 웹사이트를 보조해주는 기능을 수행하기 위한 목적으로 개발하였다고 보았다. 예를 들면 사용자가 원하는 상품을 빠르게 찾을 수 있도록 도와준다거나, 사용자의 불만을 실시간으로 처리해주는 것들을 들 수 있다. 로사와 Mode.ai는 웹사이트 등 기존의 채널과는 별개로 고객과 개별적으로 커뮤니케이션을 지속하기 위해 챗봇을 개발하였다고 보았다. 전문가들은 고객과 대면하기 위해 패션에 대한 정보를 숙지하고 제공하고 있다는 점과 원활한 상호작용을 위해 자연어 처리에 신경을 썼다는 점을 근거로 들었다.

사용자는 구매나 추천 등과 같은 특정 목적을 달성하기 위해 패션 챗봇에 접속하고 사용을 한다. 이 때 전문가들은 챗봇이 어떠한 방식으로 서비스를 제공하는가에 따라 사용자가 챗봇을 다르게 인식한다고 보았다. 챗봇이 서비스를 제공하는 방식은 챗봇의 제공 목적과 다시 연관된다. 전문가들은 특집사와 Shopbot은 공급자의 입장에서 수익을 높이기 위한 방편으로 챗봇을 개발한 것으로 보았고, 로사와 Mode.ai는 고객과의 지속적인 커뮤니케이션을 통해 충성도를 제고하기 위해 챗봇을 개발한 것으로 보았다. 그러므로 특집사와 Shopbot과 같이 쇼핑을 위한 정보 소비 행위를 촉진시키는 서비스를 주로 제공하는 경우 사용자는 챗봇을 '쇼핑과 관련된 기능을 제공하는 프로

그램'으로 인식할 것이다. 반면 로사처럼 브랜드나 기업이 원하는 점원의 이미지를 대변하거나 Mode.ai의 챗봇처럼 전문적이고 고품질의 패션 정보를 제공하는 경우 사용자는 챗봇을 '브랜드를 대표하는 전문가'로 인식할 것이다.

이것은 첫 번째 질문인 '챗봇에 따라 품질이 챗봇 신뢰에 다르게 영향을 미치는 이유'에 대한 답이 될 수 있다. '도와주는 기능을 가진 프로그램'로 인식되는 패션 챗봇은 사용자의 목적을 달성하는 것과 관련된 요인인 시스템과 제공정보의 영향력이 강하게 나타나고, '브랜드를 대표하는 전문가'로 인식되는 패션 챗봇은 제공정보와 상호작용의 영향력이 강하게 나타나게 된다. 이는 이 연구의 챗봇 신뢰에 대한 분석 결과 특집사와 Shopbot에서 시스템과 제공정보의 영향력이 1순위와 2순위로 나타났으며, 로사와 Mode.ai에서는 제공정보와 상호작용의 영향력이 1순위와 2순위로 나타난 것으로 뒷받침할 수 있다.

두 번째 질문인 챗봇에 따라 품질이 챗봇 신뢰와 브랜드 신뢰에 다르게 영향을 주는 이유는 다음과 같이 설명하였다. 브랜드는 종합적이고 총체적인 개념이기 때문에 '도와주는 기능을 가진 프로그램'으로서의 챗봇은 기능적인 성격이 강해서 브랜드의 신뢰에 유의한 영향을 미치기에는 한계가 있을 수 있다고 보았다. '브랜드를 대표하는 전문가'로서의 챗봇은 사용자가 이미 챗봇을 브랜드와 연관지어 생각하기 때문에 챗봇의 신뢰가 브랜드에 전이된다고 보았다. 이 때 사용자의 인식에 따라 시스템이나 상호작용이 브랜드 신뢰에 함께 영향을 미치게 된다고 보았다.

## V. 결론

이 연구는 패션 챗봇의 품질평가도구를 개발하고 브랜드 신뢰에 영향을 주는 품질과 그 원인을 규명하고자 하였다. 이 연구의 결과를 요약한 결론은 다음과 같다.

첫째, 패션 챗봇은 과정의 품질과 환경의 품질을 통해 평가할 수 있다. 과정의 품질은 상호작용과 제공정보의 품질로 구분되며, 환경의 품질은 시스템과 디자인의 품질로 구분된다. 패션 챗봇의 품질을 평가하기 위한 도구를 개발한 결과 상호작용 품질을 측정하는 5문항, 제공정보 품질을 측정하는 5문항, 시스템 품질을 측정하는 4문항, 디자인 품질을 측정하는 5문항으로 총 19문항을 개발하였다. 내적 일치도 검증 결과 Cronbach's  $\alpha$  값이 0.87에서 0.93으로 나타나 각 문항은 측정하고자 하는 개념을 충분히 설명하고 있었다.

둘째, 패션 챗봇의 품질이 브랜드 신뢰에 미치는 영향에 대한 연구가설을 검증한 결과 상호작용, 제공정보, 시스템, 디자인 품질은 챗봇 신뢰에 긍정적인 영향을 주며, 디자인 품질은 브랜드 신뢰에 긍정적인 영향을 주는 것으로 나타났다. 그리고 챗봇 신뢰는 브랜드 신뢰에 긍정적인 영향을 주는 것으로 나타났다.

셋째, 패션 챗봇에 따라 품질이 브랜드 신뢰에 영향을 줄 때의 차이점을 도출한 결과 특집사의 경우 상호작용, 제공정보, 시스템, 디자인 모두 챗봇 신뢰에 영향을 주었으며, 로사는 상호작용, 제공정보, 디자인이 챗봇 신뢰에 영향을 주었다. Shopbot은 제공정보, 시스템, 디자인이 챗봇 신뢰에 영향을 주었으며, Mode.ai는 상호작용과 제공정보가 챗봇 신뢰에 영향을 주는 것으로 나타났다.

패션 챗봇별로 품질이 브랜드 신뢰에 영향을 줄 때의 차이점을 도출한 결과 특집사는 브랜드 신뢰에 영향을 주는 품질이 없었으며, 로사는 디자인 품질이 영향을 주고 있었다. Shopbot은 제공정보와 시스템 품질이 브랜드 신뢰에 영향을 주고 있었으며, Mode.ai는 상호작용과 디자인이 브랜드 신뢰에 영향을 주고 있었다.

넷째, 이러한 차이가 나타나는 이유에 대해 수정 델파이 조사를 진행한 결과 '패션 챗봇의 제공목적'과 '패션 챗봇에 대한 사용자의 인식'을 원인으로 분석하였다.

이 연구는 챗봇에 대한 사용자의 인식이 명확히 정립되지 않고 챗봇에 대한 경험이 많지 않은 상황에서 연구를 진행했던 한계가 있다. 그러나 이 연구를 통해 패션 챗봇에 대한 학문적인 초석을 제시하였다는데 의의가 있다.

## References

- Bailey, J. E. & Pearson, S. W. (1983). Development of a Tool for Measuring and Analyzing Computer User Satisfaction, *Management Science*, 29(5).
- Bradach, J. L. & Eccles, R. C. (1989). Price, Authority and Trust: From Ideal Types to Plural Forms, *Annual Review of Sociology*, 15(2), 96-118.
- Brady, M. K. & J. J. Cronin (2001). Some New Thoughts on Conceptualizing Perceived Service Quality: a Hierarchical Approach, *Journal of Marketing*, 65(7), 34-49.
- Chaudhuri, A. & Holbrook, M. B. (2001). The Chain of Effects from Brand Trust and Brand Affect to Brand Performance: The Role of Brand Loyalty, *Journal of Marketing*, 65(2), 81-93.
- Cho, Y. J., Choi, Y. J., & Park, S. E. (2009). A qualitative study on analyzing the conversation of interacting with Embodied conversational agents on mobile environments, *Journal of Korea Design Knowledge*, 9, 63-71.
- Coniam, D. (2014). The Linguistic Accuracy of Chatbots: Usability from an ESL Perspective, *Text Talk*, 34(5), 545 - 567.
- DeLone, W. H. & McLean, E. R. (1992). Information System Success: The Quest for the Dependent Variable, *Information Systems Research*, 3(1), 60-95.
- DeLone, W. H. & McLean, E. R. (2003). The DeLone and McLean Model of Information Systems Success: A Ten-Year Update, *Journal of Management Information Systems*, 19(4), 9-30.
- Deshpande, A., Shahane, A., Gadre, D., Deshpande, M., & Joshi, P. M. (2017). A Survey of various chatbot implementation technique, *International Journal of Computer Engineering and Applications*, 6, ISSN(online): 2321-3469
- Doney, P. M. & Cannon, J. P. (1997). An Examination of the Nature of Trust in Buyer-Seller Relationships, *Journal of Marketing*, April, 35-51.
- Eteazadi-Amoli, J. & Farhoomand, A. F. (1996). A Structural Model of End User Computing Satisfaction and User Performance, *Information and Management*, 30(2), 65-73.
- Everard, A. & Galletta, D. F. (2006). How presentation

- flaws affect perceived site quality, trust, and intention to purchase from an online store. *Journal of Management Information Systems*, 22(3), 56-95.
- Gefen, D., Karahanna, E., & Straub, D. W. (2003). Trust and TAM in Online Shopping: An Integrated Model. *MIS Quarterly*, 27(1), 51-90.
- Goh, O. S., Ardil, C., Wong, W., & Fung, C. C. (2007). A Black-box Approach for Response Quality Evaluation of Conversational Agent Systems. *International Journal of Computational Intelligence*, 3(3), 195-203.
- Gronroos, C. (1984). A Service Quality Model and Its Marketing Implication. *European Journal of Marketing*, 14(4), 36-44.
- Hong, E. J. (2016). *A Study on the Anthropomorphic Effect depending on Interaction Type : Based on Conversational Agent in Smart Home Context* (master dissertation). Yonsei University, Republic of Korea.
- Jacoby, J., Jerry, C. O., & Rafael, A. H. (1971). Price, Brand Name and Product Composition Characteristics as Determinants of Perceived Quality. *Journal of Applied Psychology*, 55(6), 570-579.
- Koo, B. S., Baek, S. K., Kim, H. N., & Cho, K. S. (2016). Personalization of Conversational Agent Influences Users' Intimacy and Information Reliability. *EXTENDED ABSTRACTS OF HCI KOREA 2016*, 1, 456-458.
- Kosir, D. (2013). *Chatbot implementation and Testing* (Doctoral dissertation, University of Ljubljana, Slovenia. Retrieved from <http://eprints.fri.uni-lj.si/id/eprint/2041>
- Kuligowska, K. (2015). Commercial Chatbot: Performance Evaluation, Usability Metrics and Quality Standards of Embodied Conversational Agents. *Professionals Center for Business Research*, 2, 1-16.
- Lavie, T. & Tractinsky, N. (2004). Assessing Dimensions of Perceived Visual Aesthetics of Web Sites. *International Journal of Human-Computer Studies*, 60(3), 269-298.
- Lee, K. C., Kang, B. U., Suh, B. M., & Kim, J. U. (2005). Exploring the Influence of the Internet Shopping Mall's Quality Factors on User Trust and Acceptance. *Journal of the Korean Operations Research and Management Science Society*, 22, 27-46.
- Mayer, R. C., Davis, J. H., & Schroorman, F. D. (1995). An Intefrative Model of Organizational Trust. *Academy of Management Review*, 20, 709-734.
- Meira, M. O. & Canuto, A. M. P. (2015). Evaluation of Emotional Agents' Architectures: an Approach Based on Quality Metrics and the Influence of Emotions on Users. *In Proceedings of the World Congress on Engineering*, 1. Retrieved from <https://pdfs.semanticscholar.org/3494/f710bbd35afeef2577234828f8545c42302a.pdf>
- Moorman, C., Deshpande, P., & Zaltman, G. (1993). Factors affecting trust in Market research relationships. *Journal of Marketing*, 57(Jan), 81-101.
- Netemeyer, R. G., Bearden, W. O., & Sharma, S. (2003). *Scaling procedures: Issues and applications*. CA: SAGE Publications.
- Nicole, M. R. & Morgan, C. B. (2017). *Evaluating Quality of Chatbots and Intelligent Conversational Agents*. Retrieved from <https://arxiv.org/ftp/arxiv/papers/1704/1704.04579.pdf>
- Parasuraman, A., Valarie A. Zeithaml, & Leonard L. Berry (1988). SERVQUAL: A Multi-Item Scale for Measuring Consumer Perception of Service Quality. *Journal of Retailing*, 64(1), 12-40.
- Parker & Case (1991). *Management Information System* (2nd dc.). London, England: McGraw-Hill.
- Sirdeshmukh, D., J. Singh, & B. Sabol (2002). Consumer Trust, Value and Loyalty in Relational Exchanges. *Journal of Marketing*, 66(1), 15-37.
- Sirieix, L. & Dubois, P. L. (1999). Vers un modele qualite-satisfaction integrant la confiance? *Recherche et Applications en Marketing*, 14(3), 1-22.
- Srinivasan, A. (1985). Alternative Measures of System Effectiveness: Associations and Implications. *MIS Quarterly*, 9(3), 243-253.
- Stewart, K. J. (2003). Trust Transfer on the World Wide Web. *Organization Science*, 14(1). Retrieved from <https://doi.org/10.1287/orsc.14.1.5.12810>
- Swan, John, E., I. Fredric Trawick, & David, W. S. (1985). How Industrial Salespeople Gain Customer Trust. *Industrial Marketing Management*, 14(Fall), 203-211.
- Vetter, M. (2002). Quality Aspects of Bots(Eds.), *Software Quality and Software Testing in Internet Times*(pp. 165-184), Switzerland: Springer Berlin Heidelberg.
- Yoo, il, Na, K. Y., & Choi, H. R. (1999). Effects of Service Quality in Internet Shopping Mall on Electronic Commerce Performance. *The Journal of Society for e-Business Studies*, 4(3), 77-94.