JESK J Ergon Soc Korea 2020; 39(1): 73-86
http://dx.doi.org/10.5143/JESK.2020.39.1.73
http://jesk.or.kr eISSN:2093-8462

A Few Critical Design Parameters Affecting User's Satisfaction in Interaction with Voice User Interface of AI-Infused Systems

Jong Gyu Shin[1], In Gwon Jo[1], Wan Su Lim[2], Sang Ho Kim[1]

[1]Kumoh National Institute of Technology, School of Industrial Engineering, Gumi, 39177
[2]Kumoh National Institute of Technology, School of Electronic Engineering, Gumi, 39177

음성기반 지능형 시스템의 사용자 만족도에 영향을 미치는 인터랙션 설계변수들의 정의 및 분류방안

신종규[1], 조인권[1], 임완수[2], 김상호[1]

[1]금오공과대학교 산업공학부
[2]금오공과대학교 전자공학부

Corresponding Author
Sang Ho Kim
Kumoh National Institute of Technology, School of Industrial Engineering, Gumi, 39177
Email : kimsh@kumoh.ac.kr

Received : December 20, 2019
Revised : December 21, 2019
Accepted : January 04, 2020

Copyright@2020 by Ergonomics Society of Korea. All right reserved.

This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0/), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Objective: The aim of this study is to identify a few critical design parameters for enhancing user's satisfaction while interacting with the AI-infused intelligent systems through voice user interface (VUI).

Background: The interaction between the user and the AI-infused system is called as Human-AI Interaction (HAII) and supposed to have different features with respect to the human-computer interaction (HCI). It is therefore necessary to establish new criteria for designing and evaluating HAII in the point of user's satisfaction.

Method: This study identified 31 user requirements regarding with HAII from previous studies and organized them into 9 secondary and 3 tertiary level user requirement categories. It was investigated and selected 9 design parameters of VUI that might make differences in user's satisfaction. The priority of each design parameter was calculated using quality function deployment (QFD) technique.

Results: The amount of information, error control, and length of answer were found as the top three critical design parameters among others. They accounted for 51% of the total criticality score. It implies the reliability of information that the AI-infused systems provide during interaction is the most important factor for enhancing user's satisfaction.

Conclusion: This study suggested theoretically nine critical interaction parameters and their priority in designing VUI embedded in AI-infused systems.

Application: The result of the study can be used to derive various experimental research models and hypothesis in HAII.

Keywords: Human-AI interaction, Interaction design parameters, User satisfaction, Voice user interface

1. Introduction

1.1 Background

최근 들어 자율주행 자동차나 인공지능 스피커와 같이 진보된 인공지능 기술을 활용한 새로운 시스템들이 출시되고 있다. 이러한 시스템들은 헬스, 생활, 금융, 교육 등 다양한 분야에서 활용되고 있으며(NIA, 2017), 지능형 시스템(Intelligence System) (Kim et al., 2019) 또는 인공지능이 탑재된 시스템(AI-Infused System) (Amershi et al., 2019)이라 불리고 있다. Human-Computer Interaction (HCI) 분야에서는 지능형 시스템과 사용자의 인터랙션을 Advanced HCI 또는 Human-AI Interaction (HAII)이라 정의하며 HCI의 한 분야로 인식하고 이해하려는 경향을 보인다(Johnson, 1994). 하지만 Shin et al. (2018)의 연구에 따르면 HAII 상황에서 사용자가 인터랙션에 따른 만족을 얻기 위해서는 기존 HCI에 비해 더 다양한 요구사항을 충족시켜야 하는 것으로 나타났다. 또한 기존의 만족도 평가는 인터랙션에 따른 객관적 작업수행도와 인터랙션 후 설문방식을 통해 측정한 주관적 평가 결과를 중시하는 경향을 보여왔다(Nielsen and Levy, 1994). 하지만, Han (2001)의 연구에 따르면 사용자의 만족감 향상을 위해서는 작업수행도 뿐만 아니라 사용자의 요구사항을 체계적으로 반영한 사용 편의성 및 사용의 즐거움과 같은 감성적 특성까지 확보되어야 한다. 따라서 HAII의 평가를 위해서는 기존과 같은 기능적 측면에서의 평가를 뛰어넘어 사용자 요구사항을 기반으로 하는 감성적 만족도를 함께 평가할 필요가 있다. 또한 평가실험을 진행하기에 앞서 지능형 시스템의 어떠한 설계변수들이 사용자의 감성적 만족도에 더 큰 영향을 미칠 수 있는지에 관한 연구모형을 설정하여야 한다. 본 연구에서는 이미 상용화된 지능형 시스템과의 인터랙션 상황에서 사용자 만족도에 영향을 미칠 수 있는 주요 설계변수들을 우선으로 찾아내는 것이 중요하다고 판단하였다. 따라서 사전연구를 통해 분류한 지능형 시스템들의 유형 중에서 가장 큰 비중을 차지하고 있는(Shin et al., 2019) 빅스비(Bixby)나 시리(Siri)와 같은 음성기반 지능형 시스템을 중심으로 사용자 요구사항과 높은 연관관계가 있는 인터랙션 설계변수들을 정의해보고자 하였다.

1.2 Related works

음성기반의 HAII에 대한 평가와 관련하여 Yi and Kun (2017)은 인간 대 인간과 인간 대 AI의 인터랙션 비교를 통해 인간과는 다른 성격을 지닌 AI 인터랙션의 속성을 도출하였고, 사람들은 AI와의 인터랙션 시에 더 호의적이며 자신을 잘 드러내는 경향이 있음을 확인하였다. Marita et al. (2019)의 연구에서는 상용화된 챗봇과의 인터랙션 평가를 통해 챗봇 설계 시 의사소통의 투명성, 사용자 경험에 대한 맞춤형 서비스, 주제에 대한 이해력과 같은 사용자 경험 및 요구사항에 따라 만족도에 영향을 미칠 수 있는 요인을 도출하였다. 더불어 Georgios and Dimitris (2009)의 연구에서도 음성기반 인터페이스와의 인터랙션에 대한 사용성 평가방법과 고려사항이 도출된 바 있다. 이처럼 기술개발과 더불어 많은 연구자들이 음성기반의 인터페이스와 사용자 간의 더욱 만족스러운 인터랙션을 확보하기 위해 노력하고 있으며 이들 연구의 최종적인 목적은 '사용자 맞춤형(User-adaptive) 음성 인터페이스'의 개발에 있는 것으로 볼 수 있다. 하지만 지금까지의 인터랙션 설계 및 평가관련 연구들은 완성된 지능형 제품을 대상으로 사후평가 방식을 통해 사용자의 피드백을 받아 수정사항을 도출하는 방식이 주류를 이루어 왔다. 이러한 방식은 사후 피드백에 따른 단편적인 요구사항만 도출할 수 있어 사용자 만족도와 관련된 다양한 설계요인들을 세부적으로 확인할 수 없다는 한계점을 지니고 있다. 본 연구에서는 이런 문제점을 보완하기 위해 지능형 제품의 다양한 인터랙션 설계변수들을 도출하고 이들이 사용자의 요구사항과 어떻게 연관되어 있으며, 사용자 만족도에 더 큰 영향을 미치는 변수들은 무엇인지를 확인해보고자 한다.

2. Method

2.1 Research procedure

본 연구는 향후 진행할 평가실험을 통해 최적의 HAII를 설계하기 위한 기초연구의 성격을 띠고 있으며 기존연구 결과를 바탕으로 HAII 상황에서 음성기반 지능형 시스템에 대한 다양한 사용자 요구사항을 수집하였다. 수집한 요구사항들에 대해 Analytic Hierarchy Process (AHP) 분석을 진행하고 House of Quality (HoQ)를 구축함으로써 사용자 요구사항과의 연관성 정도에 따라 설계변수들의 중요도 우선순위를 도출하고자 하였다. 이를 위해 본 연구에서는 아래 Figure 1에 정리한 절차에 따라 연구를 진행하였다.

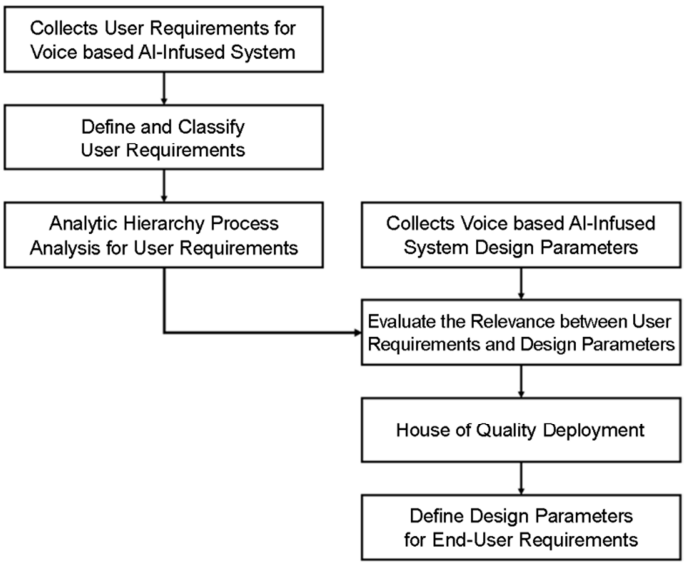


Figure 1. Research procedure of this study

2.2 Gathering and classification of user requirements

음성기반 지능형 시스템에 대한 사용자 요구사항을 도출하기 위해 2008년부터 현재까지 약 10년간 진행된 선행연구들 중 AI와의 인터랙션 및 대화형 인터페이스와 관련된 문헌을 조사하였다. '대화형 인터페이스', 'VUI', 'AI User Requirements' 등의 키워드를 이용하여 총 37건의 선행연구들을 수집하였다. 수집한 선행연구 중 인공지능 시스템의 기능개발을 위해 진행된 연구들은 배제하였고 최종적으로 사용자에게 영향을 미치는 요인에 대한 고려사항 및 정량적 설계변수들과 관련된 연구 8편을 중심으로 총 31개의 사용자 요구사항을 수집하였다.

2.3 Quality Function Deployment (QFD) process

QFD는 고객의 요구사항을 설계 단계에서부터 반영할 수 있도록 개발된 방법으로 사용자의 중요도와 우선순위에 따라 품질개선 과제를 선택하여 실행함으로써 이용자 지향적인 서비스 품질을 만드는데 용이한 품질관리 기법이다(Lee et al., 2018). 이 기법은 AHP를 통해 요구사항의 우선순위를 분석한 후 품질의 집(HoQ)을 구축하여 사용자 요구사항에 따른 설계변수와의 관계성을 도출하는데 용이하다. HoQ는 QFD 과정에서 주로 쓰이는 기법으로 Figure 2와 같이 총 5단계로 구성되어 있으며, 사용자가 요구하는 진정한 품질을 언

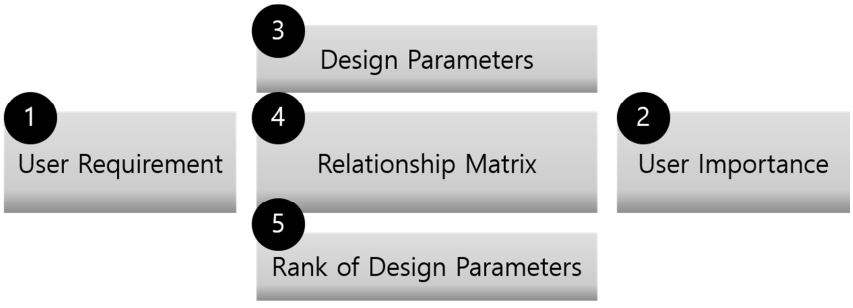


Figure 2. QFD - House of Quality (HoQ) configuration

어적 표현으로 체계화하여 설계변수의 기술특성과의 관련성을 파악하고 사용자의 요구를 정량적 특성으로 변환해 품질 설계를 하는 도구이다(Lee and Lee, 2003). 먼저 1) 사용자 요구사항 결정 단계와 2) 사용자 중요도 분석 단계는, 선행연구에서 도출한 사용자 요구사항에 대하여 중요도를 정량적으로 평가하고자 AHP를 활용하였다. AHP 분석은 의사 결정의 계층 구조를 구성하고 있는 요인 간의 쌍대비교를 통해 우선순위를 도출하기 위해 사용되는 방식이다(Park, 2014). 본 연구에서는 AHP 분석을 위해 수집한 요구사항 중 의미가 비슷하거나 중복되는 요구사항들을 통합 및 대표 요구사항을 정의함으로써 총 3단계의 계층으로 분류하였다. 계층은 3st 계층: 수집한 사용자 요구사항, 2nd 계층: 통합한 주요 요구사항, 1rd 계층: 대표 요구사항의 3단계로 각 계층 요소 간의 가중치를 상호 비교하여 우선순위를 도출할 수 있다. 이 기법을 통해 분석된 데이터를 사용하여 1)에는 계층화된 주요 및 최종 요구사항을, 2)에는 기법을 통해 도출된 우위 관계에 따른 가중치를 삽입하였다.

AHP 기법을 이용해 사용자 중요도를 평가하고자 시리나 빅스비 등 음성기반 지능형 시스템의 사용 경험이 있는 50명(남녀 구분 없이 20~30대)을 대상으로 설문조사를 실시하였다. 요구사항에 대한 쌍대비교 방식의 설문지를 작성하였고 중요도를 ±9점 척도로 수집하였다. 수집한 쌍대비교 데이터는 AHP 전용 분석 프로그램인 DRESS Ver. 1.5를 활용하여 각 사용자 요구사항 간에 중요도를 수치적으로 도출하였다. 분석의 정확도를 높이기 위해 참여자가 작성한 설문의 답변이 논리적으로 타당하며 일관성을 가졌는지 확인하였다. 일관성은 개인의 특정한 기준이 있을 경우 생각이나 사물들의 관계 강약을 그 기준에 따라 논리적인 방법으로 구성하는 것을 나타낸 성질로, 일관성 지수(Consistency Ratio, C.R.)를 통해 확인이 가능하며 C.R.이 0.1 이하인 경우 일관성이 있다고 판단한다(Saaty, 1980). 본 연구는 온라인으로 진행되어 적절한 피드백이 이루어질 수 없었기 때문에 C.R.이 0.1 이하인 데이터를 제거하고 분석을 진행하였다. 3) 설계변수 결정 단계에서는 음성기반 지능형 시스템의 사용자 요구사항을 토대로 기술적 측면에서 소비자의 인식에 영향을 줄 수 있는 항목들을 전문가 인터뷰를 통해 선정하고 정의하였다. 인터뷰는 인터랙션 디자인 전문가 4명과 인공지능 개발 전문가 2명을 대상으로 진행하였다. 기존의 HoQ 모형에서는 설계변수 간의 상호연관 관계를 나타내는 '지붕' 부분이 있지만 AHP 분석을 통해 각 요구사항 간의 계층화를 통해 우선사항 등의 관계를 이미 도출한 상태로 설계변수를 정의하였기에 이 부분은 고려하지 않았다. 4) 상관관계 기반 품질표 작성 또한 동일한 전문가들의 의견을 기반으로 사용자 요구사항에 대한 설계변수와의 관련 여부를 나타냈으며, 요구사항 행과 설계변수 열의 교차관계로 표시하여 높은 관계('◎', 9점), 보통 관계('○', 3점), 약한 관계('△', 1점), 그리고 관계없음(' ', 0점)으로 정의하였다. 이를 통해 사용자 요구사항과 설계변수 사이의 관계도를 도출하였고, 품질표에 나타난 관계 점수와 앞서 분석한 사용자 중요도 점수를 곱하여 5) 설계변수의 최종 중요도를 도출하였다.

3. Results

3.1 Classification of user requirements

3.1.1 3rd Level user requirements

앞 절에서 언급한 바와 같이 8개의 논문에서 총 31가지의 요구사항을 도출하였고, 이를 정리하면 아래의 Table 1과 같다.

Table 1. Gathered user requirements for voice-based AI-infused system

Papers	3rd Level user requirements	Papers	3rd Level user requirements
Amershi et al. (2019)	1. Make clear what the system can do	Farinazzo et al. (2010)	19. Output voice quality
	2. Time services based on context		20. Appropriate dialog start out
	3. Show contextually relevant information		21. Error prevention
	4. Encourage granular feedback		22. Handling errors
	5. Convey the consequences of user actions	Kim et al. (2017)	23. Sound levels
	6. Notify users about changes		24. Amount of information

Table 1. Gathered user requirements for voice-based AI-infused system (Continued)

Papers	3rd Level user requirements	Papers	3rd Level user requirements
Strohmann et al. (2019)	7. Navigation within the conversation 8. Minimizing cognitive load 9. Providing help 10. Feedback and system familiarity 11. Error handling 12. Emotions and empathy 13. Proactivity 14. Context awareness	Kang (2015)	25. Length of voice 26. Recognition wait time 27. Customized features based on the user's response type
		Hong et al. (2017)	28. Anthropomorphism difference
Dybkjær and Minker (2008)	15. Output voice quality 16. Feedback adequacy 17. Adequacy of dialogue initiative relative to the task 18. Sufficiency of interaction guidance	Kim et al. (2018)	29. Change commands during conversation
		Lee et al. (2018)	30. Prior understanding of error 31. Apologize for Error

수집한 31가지의 요구사항에 AHP 분석을 적용하고자 요구사항 중 의미가 비슷하거나 중복되는 항목들을 그룹화하여 9가지 주요 요구사항을 정의하고 이를 2nd 계층(주요 요구사항)으로 정의 및 최종적으로 Figure 3과 같이 3개의 계층으로 이루어진 사용자 요구사항

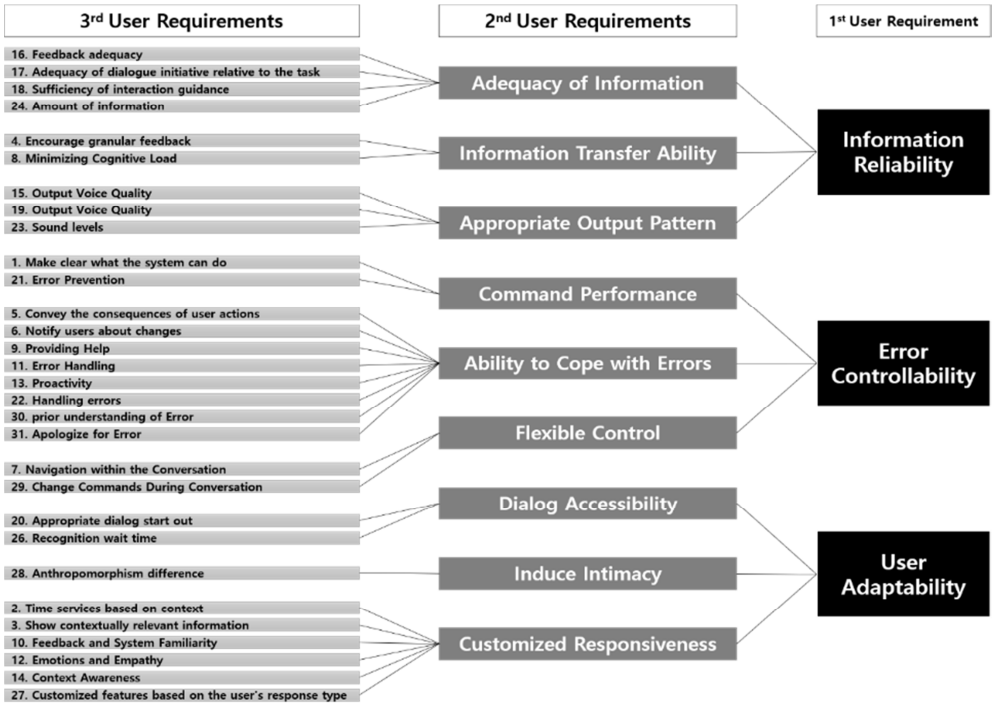


Figure 3. Hierarchies and layers of user requirements

네트워크를 작성하였다. 이를 바탕으로 2nd 계층(주요 사용자 요구사항) 및 1st 계층(대표 요구사항)을 설문에 활용하여 사용자 중요도 데이터를 수집하였다.

3.1.2 2nd Level user requirements

3.1.1절에서 정리한 31개 요구사항을 통합하여 결정된 2nd 계층 주요 사용자 요구사항에 대한 정의와 결정기준을 정리하면 다음과 같다.

Adequacy of Information

Adequacy of Information(정보에 대한 적합성)은 인터랙션 과정을 통해 제공되는 정보의 신뢰성과 관련된 사용자 요구사항을 반영한 것이다. 따라서 시스템으로부터 제공되는 피드백의 진실성 및 정확성과 관련된 요구사항들을 정보에 대한 적합성으로 분류하였다.

Information Transfer Ability

Information Transfer Ability(정보 전달 능력)은 명확한 정보를 전달받기 원하는 사용자 요구사항들의 그룹으로 사용자의 한계를 고려함과 동시에 의미 파악이 원활하도록 정보를 전달받고자 하는 요구사항을 반영하여 그룹핑하였다.

Appropriate Output Pattern

Appropriate Output Pattern(적정 출력 패턴)은 단순히 음질과 관련된 것이 아니라 사용자가 정확히 정보를 인지할 수 있도록 적절한 속도로 정보를 제시 받기를 원하는 요구사항들을 분류하였다.

Command Performance

Command Performance(명령 수행 능력)은 시스템이 사용자의 의도에 따라 원하는 결과를 명확하게 처리하여 제공해 주기를 바라는 사용자 요구사항과 관련된 그룹으로 분류하였다.

Ability to Cope with Errors

Ability to Cope with Errors(오류 대처 능력)은 시스템이 정보처리 도중 발생한 오류를 스스로 처리하는 등 사람과 유사한 업무처리를 할 수 있는 능력을 기대하는 요구사항과 관련된 항목을 포함하였다.

Flexible Control

Flexible Control(유연한 제어)는 사용자가 지능형 시스템과 인터랙션 시 실제로 사람과 대화하는 것처럼 반응하고 융통성 있게 행동하기를 기대하는 요구사항과 관련된 것이다. Walter (2011)의 'Hierarchy of User Needs'에 따르면, 사용자의 요구사항은 기능성 확보를 시작으로 신뢰성, 사용성을 거쳐 최고 단계인 즐거움의 방향으로 변화하고 있다. 이에 따라 사용자가 원하는 방향으로 컨트롤 할 수 있는 사용자 중심의 사용성 확보에 대한 요구사항들을 그룹핑하였다.

Dialog Accessibility

Dialog Accessibility(대화의 적절성)은 사람과 대화할 때 상대방에게 말을 건넸으면 그에 대한 대답이 적절한 시간 내에 되돌아오기를 기대하듯이, 시스템과 인터랙션 할 때 적절한 대화를 주고받는 시점에 대한 요구사항들을 토대로 주요 요구사항의 그룹으로 사람 사이에 대화를 할 때 상대방이 생각이 정리되지 않아 늦게 대답하거나, 생각없이 바로 대답하는 모습을 보이면 신뢰성이 떨어지듯이, 대

화형 시스템도 시스템 답변이 지연되거나 그 반대의 상황이 발생하면 만족도가 낮아지는 점을 고려하였다.

Induce Intimacy

Induce Intimacy(친밀감 유도)는 사람이 상대방과의 대화를 통해 감정을 주고받으면서 소통 관계를 형성하기를 원하는 것처럼 시스템에서도 친밀한 감정을 느끼고자 하는 요구사항을 반영하였다. 이는 자연어 처리가 선결되어야 하기 때문에 이 사항에 대해 고려한 연구가 많지는 않았지만 불쾌한 골짜기(Uncanny Valley) 이론을 고려해 볼 때, 향후 필수적으로 갖추어야 할 부분이라 판단되어 별도의 그룹으로 정의하였다.

Customized Responsiveness

Customized Responsiveness(맞춤형 대응력)는 인터랙션 과정에서 사용자의 심리상태 및 특성에 따라 대응방식을 달리할 수 있도록 진화하기를 바라는 요구사항에 관한 그룹이다. 사전 연구된 HAII의 요구사항 분류(Shin et al., 2018)를 기반으로 관련된 요구사항을 분류하였다.

이렇게 도출된 9가지의 주요 요구사항 중 선행연구에서 가장 높은 빈도로 고려된 항목은 Ability to Cope with Errors로 기존의 많은 연구에서 공통적으로 고려되고 있으며 사용자가 시스템의 도움을 받더라도 자신이 제어권을 가지고 오류를 해결하고 싶은 욕망이 강함을 대변하는 결과라고 볼 수 있다. Customized Responsiveness 또한 '개인화', '자기만의'와 같이 온전히 자신만을 위한 지능형 시스템을 원하는 사용자의 요구사항이 늘어남에 따라 다수의 연구에서 요구사항으로 포함한 것으로 판단된다. Information Transfer Ability, Command Performance, Flexible Control, Dialog Accessibility, Induce Intimacy는 각 그룹에 속한 개별 항목들을 고루 고려한 연구는 부족하지만, 최근에 향상된 기술을 통해 이에 관련된 요구사항을 충족시킬 수 있는 음성기반 지능형 시스템이 등장하고 있어 시스템을 사용할 때 사용자의 만족도를 더욱 향상시킬 수 있는 주요 요구사항이라 판단된다.

3.1.3 1st Level user requirements

아래는 3.1.2절에서 정리한 2nd 계층 주요 요구사항을 통합하여 정의한 1st 계층 대표 요구사항과 그 내용을 정리한 것이다.

Information Reliability

Information Reliability(정보 신뢰성)은 적절한 시간 내에 명확한 발음으로 정보를 전달받기를 원하는 요구와 관련된 항목이다. 사용자가 시스템과 인터랙션을 하는 목적은 필요한 정보를 얻는 것이며, 이 때 제공되는 정보는 알아듣기 쉬운 내용과 형식을 갖추어야 한다. 즉, 정보가 제시되는 속도 및 정보량, 정보 전달력과 관련해 정보를 이해할 수 있고 그에 대해 신뢰감을 가질 수 있기를 기대한다.

Error Controllability

Error Controllability(오류 제어력)은 명령 수행에 있어서의 한계를 명확하게 밝혀주고 오류 발생 시 오류가 나타난 이유 및 해결방안에 대해 도움을 받고 싶어하는 요구와 관련된 항목이다. 사용자는 시스템을 사용할 때 오류가 발생하지 않기를 기대하며, 발생한다면 이를 즉시 해결하기 위해 그 원인과 해결방법에 대해 시스템이 도움을 주기를 기대한다.

User Adaptability

User Adaptability(사용자 적응력)은 사용자가 음성기반 지능형 시스템과의 인터랙션에 있어서, 사람과 대화하는 것처럼 친근한 느낌과 함께 융통성을 원하는 요구와 관련된 항목이다. 사용자는 시스템이 적절한 맥락에 등장하는 것과 딱딱한 기계음보다는 자연스러운 음성으로 대화하며 사용자의 상태나 기분에 따라 다른 방식으로 대응해주기를 기대한다. 즉, 사용자는 사용자의 특성과 상황에 맞도록 맞춤식으로 대응할 수 있는 적응형 시스템이 되기를 기대한다.

3.2 QFD results

3.2.1 Relative importance of user requirements

앞서 도출한 지능형 시스템에 대한 3단계 계층 구조(3rd - 31가지 사용자 요구사항, 2nd - 9가지 주요 사용자 요구사항, 1st - 3가지 대표 요구사항)를 지닌 사용자 요구사항의 상대적 중요도를 설정하기 위하여 AHP 분석을 실시하였다. 온라인 설문을 통해 데이터를 수집하고 대표 요구사항 간 쌍대비교와 대표 요구사항에 속한 주요 요구사항 간 쌍대비교 데이터를 기반으로 아래 Table 2와 같은 사용자 요구사항에 대한 상대적 중요도를 도출하였다.

Table 2. Results of analytic hierarchy process (AHP) applied on user requirements

User requirements		Regional weight	Global weight
Representative requirements (Importance)	Key requirements		
Information reliability (0.523)	Adequacy of information	0.476	0.250
	Dialog accessibility	0.225	0.118
	Appropriate output pattern	0.299	0.156
Error controllability (0.347)	Command performance	0.524	0.182
	Ability to cope with errors	0.239	0.083
	Flexible control	0.237	0.082
User adaptability (0.130)	Information transfer ability	0.511	0.066
	Induce intimacy	0.110	0.014
	Customized responsiveness	0.379	0.049
Total weight			1.000

대표 요구사항 및 그룹 내 중요도를 도출하였고 그 두 값을 곱하여 주요 요구사항 9개에 대한 Global Weight(최종 중요도)를 도출하였다. 대표 요구사항들 중 상대적 중요도는 Information Reliability가 가장 높은 것으로 확인되었고 Error Controllability, User Adaptability 순으로 중요도가 감소하는 것으로 나타났다. 또한 주요 요구사항들을 비교하였을 때 Information Reliability 항목에서는 Adequacy of Information, Error Controllability 항목에서는 Command Performance, User Adaptability 항목에서는 Information Transfer Ability가 가장 높았으며, 전체적으로 Adequacy of Information 이 최우선 순위, 그 다음으로 Command Performance 및 Appropriate Output Pattern이 중요한 것으로 나타났다. 중요도 결과를 볼 때, 지능형 시스템은 제시되는 정보에 대해 신뢰성이 확보될 수 있도록 설계하는 것이 가장 중요하며 현재 시스템의 상태를 사용자가 잘 이해할 수 있도록 해주어야 한다. Adaptability 그룹의 요구사항들은 상대적으로 낮은 중요도를 보였는데 이를 통해 정보에 대한 정확성, 질(Quality), 접근성 등과 같은 기본요건이 우선적으로 확보되어야 하며 이를 충족시키고 나면 사용자의 특성에 감응하는 적응형 인터랙션에 대한 요구가 더 커질 것임을 예상해볼 수 있다.

3.2.2 Design parameters of voice user interface in AI-infused systems

전문가와의 인터뷰를 통해 사용자 만족도에 영향을 줄 수 있는 음성기반 지능형 시스템의 주요 설계변수들을 선정하였다. 앞서 도출한 사용자 요구사항을 기반으로 시스템의 기능적 측면에서 필요한 9가지 설계변수들을 선별하였고, 이를 정리하면 아래 Table 3과 같다.

Table 3. Design parameters and their definitions of voice user interface in intelligent systems

Design parameters	Definition
Amount of information in answer	The amount of information in the response presented by the intelligent system after the user requests the information
Length of answer	The sentence length of the answer that the user can recognize at once
Speed of answer	Speed of answers that users can easily recognize
Error control	The number of attempts taken for the user to get the correct answer from the intelligent system
Error feedback	Method of feedback provided when an intelligent system does not understand user's question or is disturbed by external noise
Stop working	The degree to which user can stop an existing running task to ask another question again if when their intentions differ so that they can effectively lead the conversation
Delay time	The time it takes for an intelligent system to answer user's questions
Way of speaking	The speech type of intelligent system answers that allow users to create intimacy
Intonation	The tone or timbre of the answer that can be tailored to the user's sensitivity

Amount of Information in Answer

Amount of Information in Answer(답변 정보의 양)은 사용자가 시스템에게 정보를 요청했을 때 지능형 시스템이 제시한 답변의 정보량이며, 사용자는 지나치게 단순하거나 복잡한 형태의 정보보다는 주어진 상황을 해결하기에 적합한 양의 정보를 제시 받기를 원하므로 이를 주요 설계변수로 지정하였다.

Length of Answer

Length of Answer(답변의 길이)는 시스템이 답변으로 사용하는 문장의 길이며, 사용자는 답변의 길이가 길어질수록 주의가 분산되고 혼란스러워하는 경향이 있으므로 이를 주요 설계변수로 지정하였다

Speed of Answer

Speed of Answer(답변의 속도)는 사용자가 쉽게 인식할 수 있는 답변의 속도로, 사용자는 이해하기 힘든 빠른 속도로 정보를 전달받거나 지나치게 느린 답답한 속도로 대화하기를 원치 않으므로 이를 주요 설계변수로 지정하였다.

Error Control

Error Control(오류 제어)는 시스템에 필요로 하는 내용을 전달하고 원하는 응답을 얻는 데까지 걸린 입력시도 횟수이다. 시스템이 사용자 의도를 제대로 이해하지 못해 명령을 여러 차례 반복해야 한다면 사용자가 부정적인 정서를 가지게 될 것이므로 이를 최소화할 수 있도록 주요 설계변수로 지정하였다.

Error Feedback

Error Feedback(오류에 대한 피드백)은 시스템이 사용자의 질문을 이해하지 못했거나 외부 소음으로 인해 방해받는 경우 이에 대한 피드백을 어떤 방식으로 제시하는 지와 관련된 설계변수이다. 사용자는 시스템을 이용하면서 정상적으로 작동하는 기능을 원하므로 오류가 발생했을 때 매우 답답해하며 해결할 수 있는 방법을 알고자 하는 경향이 있으므로 이를 설계변수로 지정하였다.

Stop Working

Stop Working(작업 중지)은 시스템이 사용자의 의도를 잘못 이해한 경우 사용자가 기존에 진행중인 작업을 중단시킬 수 있는지 여부와 관련된 설계 요소이다. 사용자는 시스템과의 인터랙션 시 자신의 의도가 훼손되지 않고 명확하게 드러내길 원하며 이에 반할 경우 하던 의도와는 다른 기능을 종료하고 빠르게 재시도하거나 올바른 기능을 수행하길 원하므로 이를 설계변수로 지정하였다.

Delay Time

Delay Time(답변 대기 시간)은 지능형 시스템이 사용자의 명령에 피드백을 할 때까지 걸리는 시간으로, 사용자는 명령 후 시스템의 답변이 지연되면 불쾌감이 생기므로 이를 설계변수로 지정하였다.

Way of Speaking

Way of Speaking(답변의 어투)는 시스템이 답변에서 사용하는 대화표현의 양식으로서 답변의 어투가 무례하면 듣는 사람이 불쾌감을 느낄 수 있고 사용자가 시스템에 대해 친밀감을 형성하는 정도가 달라질 수 있으므로 이를 설계변수로 지정하였다.

Intonation

Intonation(답변의 억양)은 사용자의 감성에 따라 맞춤형으로 제시할 수 있는 답변의 억양 또는 음색이다. 남성의 경우 젊은 여성의 음색으로 대답을 받고 싶거나, 방언을 쓰는 지역의 사용자는 답변을 방언으로 듣길 원하는 경우가 있으므로 이를 설계변수로 지정하였다.

3.2.3 Correlation between user requirements and design parameters of voice user interface

앞 절에서 제시한 전문가 인터뷰 결과를 이용해 사용자 요구사항과 시스템 설계변수 사이의 상관관계를 분석하였다. 사용자가 생각하는 요구사항의 중요도와 전문가 인터뷰를 통해 도출한 설계변수의 상관관계를 정리하면 Figure 4에 나타낸 HoQ 테이블과 같으며 이를 활용하여 각 설계변수의 중요도 우선순위를 도출하였다.

		Design Parameters									
User requirements		Amount of Information in Answer	Length of Answer	Speed of Answer	Error Control	Error Feedback	Stop Working	Delay Time	Way of Speaking	Intonation	◎ = 9 points ○ = 3 points △ = 1 point
Representative requirements	Key Requirements										User importance
Information Reliability	Adequacy of Information	◎	○				△			△	0.250
	Dialog accessibility	○	◎						○	△	0.118
	Appropriate Output Pattern		△	◎				△			0.156
Error controllability	Command Performance	△			◎	○					0.182
	Ability to Cope with Errors				○	◎					0.083
	Flexible control				△	△	◎			△	0.082
User Friendly	Information Transfer Ability			△			○	◎	△	△	0.066
	Induce Intimacy								◎	○	0.014
	Customized Responsiveness							○		◎	0.049
Importance		2.786	1.968	1.470	1.969	1.543	1.186	0.897	0.546	0.999	
Importance ratio(%)		20.847	14.726	11.000	14.734	11.546	8.875	6.712	4.086	7.475	
Priority		1	3	5	2	4	6	8	9	7	

Figure 4. House of quality (HoQ) for voice user interface embedded in AI-infused systems

사용자 중요도를 곱하여 HoQ를 구축하고, 설계변수의 최종 중요도 순위를 도출하였다. 설계변수의 중요도는 사용자의 중요도 * 상관관계 점수로 계산한다. 예를 들어 Amount of Information in Answer의 경우 가장 상관관계가 있다고 판단했던 Adequacy of Information에서 9점, 그 다음 Dialog Accessibility에서 3점, Command Performance에서 1점을 주어 사용자 중요도와 상관관계 점수를 곱하고 항목을 더하여 도출하였다.

산출된 상관관계 점수를 기준으로 사용자 요구사항에 부합하는 음성기반 지능형 시스템의 설계를 위해 우선적으로 고려해야 할 주요 변수들을 파악하려 파레토 분석을 실시하였다. 그 결과를 정리하면 Figure 5에 나타낸 그래프와 같다.

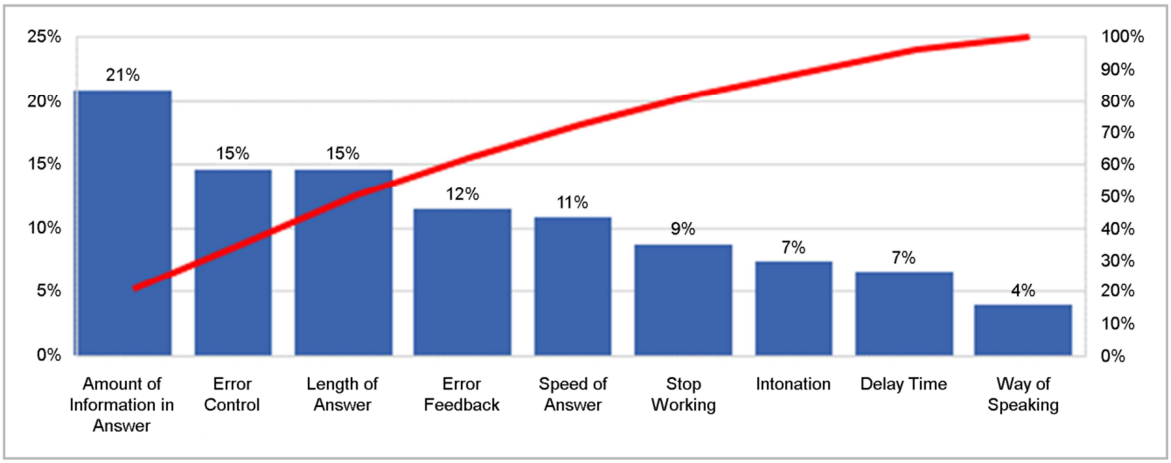


Figure 5. A few critical design parameters and their weights

파레토 분석을 통해 정보 신뢰성에 대한 요구사항과 관계가 깊었던 Amount of Information (21%)이 가장 중요한 설계변수임을 확인하였다. Error Control (15%)과 Length of Answer (15%)가 그 뒤를 따르고 있으며 상위 3순위 설계변수의 중요도 합이 전체의 51%로 절반 이상을 차지하였다. 사용자의 요구와 상대적으로 연관성이 낮은 하위 3순위 설계변수는 Way of Speaking (4%), Delay Time (7%), Intonation (7%)로 나타났으며 이들의 중요도 합은 전체 21% 수준이었다. Ability to Cope with Errors, Customized Responsiveness와의 가장 큰 연관을 보였던 Error Feedback, Intonation 항목은 중간수준의 중요도를 지닌 것으로 나타났다.

4. Discussion

본 연구에서는 점차 확산되고 있는 인공지능 기반 지능형 시스템의 개발 과정에서 사용자가 만족할 수 있는 인터랙션을 제공할 수 있도록 시스템을 설계하는 방안을 찾고자 하였다. 이를 위해 사용자 요구사항과 밀접한 연관이 있는 주요 설계변수들을 도출하였다. 우선 선행연구들을 통해 확인된 사용자 요구사항들을 계층적으로 그룹화한 후 이들의 상대적 중요도를 결정하였다. 이들 요구사항별 중요도와 지능형 시스템의 설계변수들 간의 연관성 정도에 따라 사용자 만족도에 더 큰 영향을 미칠 수 있는 설계변수들의 중요도 우선순위를 도출하였다. 본 연구에서 고려한 총 9개의 설계변수 중 '답변의 정보량'이 사용자 만족도 측면에서 가장 중요한 설계변수로 선정되었다. 답변의 정보량과 사용자 요구사항 간의 상관관계 설정에 있어서 주된 요구사항은 정보량이 적절하게 설정되지 않으면 시스템의 답변에 대한 신뢰성이 확보되지 않을 수 있기에 Adequacy of Information과 강한 연관성을 지닌 것으로 판단된다. 또한, 주어지는 정보량이 많아질수록 사용자의 주의와 정확한 명령에 영향을 주기 때문에 Dialog Accessibility와 Command Performance에 일부 연관관계가 존재하는 것으로 평가되어 3개의 주요 요구사항과 관련된 것으로 나타났다. 상관관계 점수와 AHP 분석을 통해 나타난 사용자 중요도를 종합한 결과로 볼 때, 지능형 시스템에서 제시되는 답변의 정보량이 사용자가 가장 중요히 요구하는 시스템의 신뢰성에 가장 큰 영향을 미치는 요인임을 확인할 수 있다. 상위 중요도의 설계변수를 기반으로 볼 때, 사용자가 신뢰할 수 있는 정보를 이해하기 쉽게 제시하며 오류를 해결할 수 있는 기능에 충실한 것을 가장 기본적인 특성으로 여겨 다른 항목보다 우선순위가 높게 나타난 것

으로 보인다. 그 밖의 설계변수들 중에서는 오류 제어와 답변의 길이가 상위권의 중요도를 가진 것으로 나타났다. 이는 답변이 조금 늦어지더라도(답변 대기 시간 8순위) 정확한 정보를 명확하게 응답해주기를 기대하고 있음을 의미한다. 이러한 분석 결과를 통해 지능형 시스템의 만족도에 있어 시스템이 제공하는 응답 정보에 대한 사용자의 신뢰도가 가장 큰 결정 요소임을 확인하여 이 부분에 대해 사용자의 만족도 확보를 위한 설계가 우선시되어야 한다고 판단된다. 한편 답변의 어조나 어투와 같이 답변을 통해 사용자와 친밀감을 형성할 수 있는 음성 언어적인 요소들은 상대적으로 중요도가 낮은 설계변수로 나타났다. 실제로 지능형 시스템 답변의 성별은 사람의 감정에 크게 영향을 미치지 못한다는 기존 연구 결과(Kim et al., 2019)로 보아 이들 설계요인들은 현재 시스템개발 수준에서는 부수적인 요소로 평가되며, 이는 인간이 아닌 시스템에게 가지고 있는 새로운 기술에 대해 거부감과 익숙하지 않은 경험(uncanny valley)을 가지고 있기 때문인 것으로 해석된다. 다만 사용자의 특성에 맞춤형으로 대응하는 새로운 기술들이 등장하고 있기 때문에 기술의 확산에 따라 더 많은 사용자들이 이러한 기능을 접하며 인식이 개선되면 점차적으로 관련 요소에 대한 수용도가 높아질 것으로 예상해 볼 수 있다. 한편 2nd 사용자 요구사항으로 도출된 항목들과 다수의 연관성이 있는 것으로 파악된 Error Feedback과 Intonation은 종합점수에서는 낮은 우선순위를 나타냈다. 이로부터 사용자가 시스템을 사용하는 과정에서 이들 설계변수의 특성에 따라 다소 불만족스러움을 느끼더라도 시스템이 제공하는 정보의 유용성과 신뢰성이 확보된다면 이를 감수하는 경향을 보일 것임을 예상할 수 있다.

5. Conclusion

인공지능 기술의 발전에 따라 기존의 HCI와는 차별화된 특성을 지닌 HAII가 등장함으로써 사용자는 새로운 인터랙션 상황에 직면하였고, 그에 따라 사용자 만족도를 향상시키기 위한 새로운 설계지침 및 방안을 수립할 필요가 생겼다. 본 연구에서는 음성기반 지능형 시스템과의 인터랙션 상황에서 중요하게 고려해야 할 설계변수들을 정의함으로써 향후 설계변수들을 통해 사용자들의 만족도에 미치는 영향을 분석할 수 있는 연구모형의 토대를 제공하였다. 특히 기존 사용자의 요구사항과 관련 항목 설문조사에 대한 우선순위 결과를 바탕으로 연구를 진행했으므로 사용자의 의견을 최대한 반영했다는 점에서 의의가 있다. 최근 Microsoft 연구팀이 발표한 연구는 학술자료 및 제품 가이드라인에서 수집한 150개 이상의 AI 관련 권장 사항을 18가지의 AI 설계 지침으로 코드화하는 연구를 진행하여 HAII 가이드라인을 제시한 바 있으며(Amershi et al., 2019) HAII에 대해 전체적으로 포괄할 수 있는 설계 지침을 도출한 바 있다. 본 연구는 기존의 HAII에 대한 연구를 토대로 단순한 가이드라인이 아니라 지능형 시스템의 인터랙션 특성을 변화시킬 수 있는 설계변수를 도출했다는 점에서 타 연구와 차별성을 가진 것으로 판단된다. 따라서 HAII 평가를 위한 향후 연구방향을 설정하는데 도움을 줄 수 있을 것으로 예상되며 새롭게 개발되는 지능형 시스템과의 인터랙션에 대한 통합적 평가의 기초자료로 활용될 수 있을 것으로 기대된다. 다만 본 연구는 한국인을 대상으로 진행되었으며 Jang et al. (2018)에 따르면 영어가 주언어인 사용자가 일본어, 한국어가 주언어인 사용자에 비해 AI에 대한 Uncanny Valley 현상이 더 발생하기 때문에 본 연구 결과를 타 언어 사용자에게도 일반화할 수 있을지에 대해서는 추가적인 고려가 필요하다. 또한 본 연구는 실험기반이 아닌 설문형식의 사용자 중요도와 전문가의 의견을 중심으로 도출해 낸 결과이기 때문에 차후 실험을 통해 중요한 설계변수들이 사용자의 만족도에 실제로 어느 정도의 영향을 주는지 확인할 필요가 있다. 후속 연구를 통해 본 연구에서 정의된 설계변수들을 그 우선순위에 따라 음성기반 지능형 시스템의 사용자 만족도에 영향을 미치는 독립변수로 설정한 평가실험을 시행한 후 그 영향정도를 평가할 계획이다.

Acknowledgement

This research was supported by the Brain Korea 21 Plus Project (School of Industrial Engineering, Kumoh National Institute of Technology).

References

Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., Suh, J.N., Iqbal, S., Bennett, P.N., Inkpen, K., Teevan, J., Kikin-Gil, R. and Horvitz, R., *Guidelines for Human-AI Interaction, Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, No.3, 1-13, 2019.

Dybkjær, L. and Minker, W., Recent Trends in Discourse and Dialogue, *Springer Science & Business Media*, 39, 2008.

Farinazzo, V., Salvador, M., Kawamoto, A.L.S. and Oliveira Neto, J.S., An Empirical Approach for the Evaluation of Voice User Interfaces, *User Interfaces*, IntechOpen, 2010.

Georgios, K. and Dimitris, S., Usability Methodologies for Real-Life Voice User Interfaces, *International Journal of Information Technology and Web Engineering*, 4(4), 78-94, 2009.

Han, S.H., Models for evaluating and predicting the user satisfaction of product designs, *Journal of the Ergonomics of Korea*, 20(1), 87-113, 2001.

Hong, E.J., Jo, G.S. and Choi, J.H., Effects of Anthropomorphic Conversational Interface for Smart Home : An Experimental Study on the Voice and Chatting Interactions, *Journal of the HCI Society of Korea*, 12(1), 15-23, 2017.

Jang, P.S., Jung W.H. and Hyun, J.S., Analysis of Differences in Uncanny Valley Phenomenon by Languages Using Social Media Data, *Journal of the Ergonomics Society of Korea*, 37(2), 1229-1684, 2018.

Johnson, H., Relationship between user models in HCI and AI, *in IEE Proceedings - Computers and Digital Techniques*, 141(2), 99-103, 1994.

Kang, K.D., A Study of Vehicle Voice Interface Design Guide: Focus on the Interface factor, *Ajou University Master Graduate Paper*, 2015.

Kim, A.Y., Cho, M.H, Ahn, J.Y. and Sung, Y.J., Effects of Gender and Relationship Type on the Response to Artificial Intelligence, *Cyberpsychology, Behavior, and Social Networking*, 22(4), 249-253, 2019.

Kim, J.B., Kim, S.H. and Lee, H.S., Pattern Recognition and Classifier Design of Bio-Signals based Interface in Human - Artificial Intelligence Interaction(HAII) Framework for Real Time Evaluation of Emotions, *Journal of The Korean Institute of Intelligent Systems*, 29(3), 242-249, 2019.

Kim, T.S., Kim, J.H. and Choi, J.H., The Effect of Barge-in Function of In-Vehicle Voice Conversational Interface on Driving Experience - Focus on Car Navigation and Music Services -, *Design Convergence Study*, 17(1), 17-28, 2018.

Kim, Y.J., Ahn, S.J. and Lee, T.E., UX Guidelines for Designing Audio Guidance of Multimedia Contents for Low-Vision - With a Focus on Sound Level Difference and Amount of Auditory Information, *Archives of Design Research*, 30(1), 131-143, 2017.

Lee, H.K. and Lee, S.B., A Study of HOQ Complexity Reduction by Quantification Method of Type III, *Journal of the Korean Society for Quality Management*, 31(2), 131-142, 2003.

Lee, S.J., Lee, S.H., Son, E.J. and Oh, H.J., A Study on the Improvement for Archives Service Quality of Central Government Agencies Using QFD, *The Korean Journal of Archival Studies*, No.58, 263-297, 2018.

Marita, S., Ida, M.H., Asbjørn, F. and Petter, B.B., Help! Is My Chatbot Falling into the Uncanny Valley? An Empirical Study of User Experience in Human-Chatbot Interaction, *Human Technology*, 15(1), 30-54, 2019.

Nation Information Society Agency (NIA), Korea A.I. Company Status Survey Report A.I. Plus Series Ver 1.0, *Nation Information Society Agency*, 2017.

Nielsen, J. and Levy, J., Measuring Usability: Preference vs. performance, *Communications of the ACM*, 37(4), 66-75, 1994.

Park, D.J., A Study on Priorities of the Key Competence of Port Logistics Enterprise using AHP Method, *Journal of Korea Port Economic Association*, 30(1), 159-173, 2014.

Saaty, T.L., The analytic hierarchy process: Planning, Priority Setting, Resource Allocation (2nd ed.), *McGraw-Hill*, 1980.

Shin, J.G., Jo, I.G. and Kim, S.H., A Study on the Difference between Human-AI Interaction and Human-Computer Interaction, *Proceedings of Ergonomics Society of Korea Fall Conference*, 2018.

Shin, J.G., Jo, I.G., Heo, J.H. and Kim, S.H., A Classification and Image Map of the Intelligent Systems based on their Features of Interaction, *Proceedings of Ergonomics Society of Korea Fall Conference*, 2019.

Strohmann, T., Höper, L. and Robra-Bissantz, S., Design Guidelines for Creating a Convincing User Experience with Virtual In-vehicle Assistants, *Proceedings of the 52nd Hawaii International Conference on System Sciences*, 4813-4822, 2019.

Walter, A., Designing for Emotion, *New York: A book apart*, 2011.

Yi, M. and Kun, X., The media inequality: Comparing the initial human-human and human-AI social interactions, *Computers in Human Behavior*, 72, 432-440, 2017.

Author listings

Jong Gyu Shin: shinjg@kumoh.ac.kr
Highest degree: MASc, School of Industrial Engineering, Kumoh National Institute of Technology
Position title: PhD Candidate, School of Industrial Engineering, Kumoh National Institute of Technology
Areas of interest: Human-AI Interaction, Usability Evaluation, Interaction Design

In Gwon Jo: bee9639@kumoh.ac.kr
Highest degree: Bachelor of Engineering, School of Industrial Engineering, Kumoh National Institute of Technology
Position title: MASc Candidate, School of Industrial Engineering, Kumoh National Institute of Technology
Areas of interest: Human-AI Interaction, Kansei Engineering, Usability

Wan Su Lim: wansu.lim@kumoh.ac.kr
Highest degree: PhD, School of Electrical Engineering and Computer Science, GIST
Position title: Assistant Professor, School of Electronic Engineering, Kumoh National Institute of Technology
Areas of interest: Machine learning, Statistical analysis, Optimization

Sang Ho Kim: kimsh@kumoh.ac.kr
Highest degree: PhD, Department of Industrial Engineering, POSTECH
Position title: Professor, School of Industrial Engineering, Kumoh National Institute of Technology
Areas of interest: Human-AI Interaction, Usability Evaluation, Interaction Design, Kansei Engineering, UI/UX Design