Journal of Digital Art Engineering & Multimedia Vol.10, No.2, June (2023), pp.175-186
ISSN 2508-9099(p) 2672-1198(e) *http://dx.doi.org/10.29056/jdaem.2023.06.04*

AI 서비스의 사용성 향상을 위한 인터랙션 디자인 연구 : 헤이카카오, 네이버 클로바를 중심으로

A Study of Interaction Design for Improving the Usability of AI Service: Focused on HeyKakao and Naver Clova

권아름[1], 김승인[2]*

Arum Kwon[1], Seung In Kim[2]*

요 약

본 연구는 AI 서비스 플랫폼의 사용성 향상을 위한 사용자 경험에 관한 연구이다. AI 기술이 생활 속으로 파고들지만 자연스러운 대화를 이어가지 못하는 자연어 처리와 기술적 문제가 있다. 이에 사용자 경험을 고찰하고자 국내 AI 복합서비스를 제공하는 카카오의 헤이카카오와 네이버의 클로바 애플리케이션의 사용자 경험을 비교·분석하였다. 연구 방법으로는 VUI 사용에 익숙한 20-40대 남녀를 대상으로 1:1 대면 심층 설문조사를 진행했으며 사용성 평가 설계를 위해 VUI 사용성 범주 항목인 효율성, 정확성, 유연성, 일관성, 의미성과 AI 사용성 평가를 위한 항목 개인맞춤화, 다양성, 편의성, 전문성, 실시간성을 매칭하여 사용성 평가 목록을 작성한 후 사용자 경험을 분석하였다. 그 결과, 두 애플리케이션 모두 명령어 인식에 대해 이해도가 기능마다 상이했으며, 정보 및 콘텐츠 제공에 있어 VUI와 UI 사용에 대해서도 일관성이 떨어졌다. 따라서 사용성을 높이기 위해서는 명령어 인식에 대한 기술적 문제와 고차원적인 정보 제공, 일관적인 인터랙션 디자인이 고려되어야 할 것이다.

핵심어 : AI 서비스, 인공지능 서비스 플랫폼, AI 사용성 테스트, 인터랙션 디자인

Abstract

In this paper, study on user experience to improve the usability of AI service platforms. There are natural language processing and technical problems that AI technology penetrates into life but cannot continue natural conversations. Accordingly, in order to examine the user experience, the user experience of Kakao's Hey Kakao and Naver's Clova application, which provide domestic AI complex services, was compared and analyzed. As a research method, a 1:1 face-to-face in-depth survey was conducted on men and women in their 20s and 40s who are familiar with VUI use, and user experience was analyzed by matching VUI usability category items such as efficiency, accuracy, flexibility, consistency, meaning, and AI usability evaluation. As a result, both applications had a different understanding of command recognition

1 Department of Design Studies, IDAS, Hongik University, Seoul, Korea [Graduate Student]
e-mail: kwonarum707@gmail.com
2 Department of Digital Media Design, IDAS, Hongik University, Seoul, Korea [Professor]
e-mail: r2d2kim@naver.com (Corresponding author)

Received(March 15, 2023), Review Result(1st: April 9, 2023, 2nd: May 2, 2023), Accepted(June 12, 2023), Published(June 30, 2023)

© 2023 The Authors. Published by NCISS.
This is an open access article licensed under the Creative Commons Attribution-NonCommercial 4.0 International License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc/4.0/.

from feature to feature, and were also inconsistent with the use of VUI and UI in providing information and content. Therefore, in order to increase usability, technical problems with command recognition, high-level information provision, and consistent interaction design should be considered.

Keyword : AI service, AI service platform, AI Usability test, Interaction design

1. 서론

본 연구는 인공지능 서비스 애플리케이션의 사용성 향상을 위하여 인터랙션 디자인에 관한 사용성 테스트를 분석한 연구이다.

비대면 경제가 활성화되고, 데이터가 중요한 시대가 도래되면서 인공지능 기술은 4차 산업혁명 시대의 핵심 요소가 되었다. 디지털 혁신의 강력한 주역으로 부상해 시장의 가속화와 경쟁력, 고객의 새로운 경험과 기대에 직접적으로 영향을 미치는 기술로 진화하고 있다 [1]. 인공지능 기술은 사람들의 일과 생활의 활동을 대체할 수 있게 되었고, 이는 플랫폼 비즈니스에 융합되어 커다란 비즈니스 혁신의 가능성을 제시하고 있다 [2]. 그중 인공지능 플랫폼은 인공지능 서비스를 가능하게 하는 여러 기술 요소를 한 곳에 통합하여 다양하게 적용할 수 있도록 구성되어 있다. 사람의 음성을 인식해 텍스트로 변환하거나, 텍스트를 다시 음성으로 변환하기도 하며, 언어를 이해하여 질의응답과 대화를 처리하는 기술이 대표적이다. 전 세계 기업들이 개별적인 인공지능 애플리케이션보다는 총체적인 인공지능 서비스를 제공하는 인공지능 플랫폼 개발에 힘을 쏟고 있다 [3]. 실제로 AI 기반 지능형 개인 비서 서비스의 글로벌 시장 규모는 글로벌 시장조사업체인 리서치앤마켓에 따르면 연평균 약 46% 성장으로 2023년엔 252억 달러 규모를 예상할 만큼 높은 성장률을 보이고 있다 [4]. 그중 음성 기반 인공지능 서비스는 특성에 따라 스마트 스피커, 개인 가상비서, 디지털 어시스턴트, 챗봇 등으로 나뉠 수 있다. 대표적으로 글로벌 IT 기업을 중심으로 아이폰의 시리(Siri), 인공지능 스피커 아마존의 알렉사(Alexa), 구글 어시스턴트가 있다. AI 스피커의 등장이 증가하면서 국내에서도 SKT의 누구(NUGU), KT 기가지니(GiGA Genie), 네이버의 클로바, 카카오의 헤이 카카오 순으로 앞 다투어 출시되고 있다. 모두 음성인식 인터페이스를 통해 검색, 음악 스트리밍, 쇼핑, 날씨 등 사용자의 요구사항을 받아 서비스를 제공한다. 이러한 서비스의 사용이 가정, 자동차, 산업 기기 등에서 활용되며 새로운 플랫폼인 애플리케이션으로 확대되고 있다 [5]. 그러나 자연어 처리의 한계와 대화 맥락을 충분히 파악하지 못하는 기술 문제 등 보편적 활용에 여러 한계가 있어 실제 음성 인터랙션(이하 VUI)에 대한 사용자들의 불편함이 존재한다 [6].

따라서 본 연구는 이러한 사용자의 요구사항을 파악하여 이를 중심으로 서비스 구성과 개선사항을 제시하는 데 목적이 있다.

2. 이론적 배경

2.1 인공지능 서비스의 개념 및 특성

인공지능(이하 AI) 서비스는 인간의 지각, 이해, 인지, 학습, 판단 능력을 기기에 구현시킨 종합 기술을 통해 문제를 해결할 수 있도록 제공한다. 사람의 음성을 인식해 텍스트로 변환하거나, 텍스트를 다시 음성으로 변환하고, 언어를 이해하며 질의응답과 대화를 처리한다. AI 서비스에 관련된 선행연구를 통해 AI 서비스의 특성을 [표 1]과 같이 정리하였다. 선행연구에서 공통으로 도출한 주요 특성으로는 개인의 취향과 특성을 이해한 개인 최적의 맞춤 서비스를 제공하는 개인맞춤화, 요구사항을 명확하게 이행하고 상황에 맞게 높은 수준의 결과나 품질 서비스를 제공하는 전문성, 다양한 요구사항을 인지하여 복합적인 서비스를 처리하는 다양성, 상황이나 환경에 따라 빠르고 편리하게 이용할 수 있고 효율적인 서비스를 제공하는 편의성, 양방향 상호작용으로 실시간 서비스를 제공하는 실시간성, 소비자가 필요로 하는 서비스를 완벽하고 안정적으로 처리하며 개인정보를 안전하게 관리하는 신뢰성으로 정리할 수 있다 [3][7].

[표 1] 인공지능 서비스의 특성
[Table 1] The figure of AI service

연구자	인공지능 서비스 특성
백창화, 임성욱, 최재호	KJ법을 이용하여 AI 서비스와 관련된 산업계와 학계의 전문가들과 함께 개인맞춤화, 전문성, 다양성, 편의성, 시공간성, 실시간성, 신뢰성 7가지 주요 특성을 범주화하였다[8].
정용석, 이홍우, 서혜애	빅데이터 기반의 AI 서비스는 방대한 양의 데이터 소스로 다양성및 데이터의 정확성, 신뢰성, 데이터의 생성 속도 및 전송, 그리고 수익성을 가진다 [9].
백창화	델파이 기법을 이용해 AI 서비스 품질평가 항목을 도출했다. 선행연구를 바탕으로 예비 평가항목 50개를 도출했고 개인맞춤화, 전문성, 다양성, 편안함, 편의성, 시공간성, 실시간성, 신뢰성의 카테고리로 분류해 주요 키워드로 선정하였다 [10].
정의영, 조성주	음성 명령에 대한 정보의 신뢰성과 명확성이 AI 스피커에 기본적이고 중요한 기능으로써 사용성 평가에서도 큰 영향을 미치는 요소로 도출했다. 즉 효율성, 이해성, 용이성, 피드백, 정보전달속도, 신뢰성, 명확성, 호환성, 개인성으로 정의하였다 [11].

2.2 음성 기반 인터랙션(VUI)

VUI는 음성인식 기술을 기반으로 기계가 인간의 음성에서 의미를 추출하여 인간과 시스템 간 인터랙션을 가능하게 하는 대화형 인터페이스다. VUI는 휴대폰, 태블릿을 넘어 AI 서비스의 확대로 주요 인터페이스가 되고 있다. 구두의 언어로 소통하는 상호작용 방식을 지닌 응용프로그램으

로 작업 수행을 도울 때 주로 사용된다 [6][12].

음성으로 소통하는 인간에게 친숙한 정보 입출력 방법으로 비교적 학습과 훈련의 필요성이 낮고 손쉽게 사용할 수 있다. 또 손으로 입력하는 인터랙션 방식에 비해 비교적 빠른 특징으로 긴급한 정보 입출력 상황에서 중요한 인터랙션 방식으로 주목받고 있다 [12]. 음성 사용자 인터페이스의 사용성 평가에 관한 선행연구를 [표 2]와 같이 정리했다. 선행연구에서 공통으로 언급하고 있는 요소로 효율성, 정확성, 의미성, 유연성, 일관성, 신뢰성을 도출할 수 있다.

[표 2] VUI 사용성 평가

[Table 2] The usability test of VUI

연구자	VUI 사용성 평가
김민정, 한지은, 강효진, 권규현	질적 연구 분석과 전문가 인터뷰를 통해 VUI의 사용성 평가 기준에 있어 13가지 속성을 추리고 이 속성을 효율성, 정확성, 의미성, 유연성, 일관성 다섯 가지의 사용성 범주 항목으로 도출했다 [13].
Yao Song, Yanpu Yang, Pejyao Cheng	57개의 기사를 분석하고 VUI의 유용성을 평가하는 데 사용되는 9개의 설문지를 검토한 결과 모든 설문지에는 세 가지 가용성 차원인 효과성, 효율성, 만족성과 관련된 항목이 포함되어 있었다 [14].
천재민, 한성호, 조영석, 박원규, 김종서	VUI의 사용성 향상을 위해서는 사용자의 실제 발화에 근거한 시스템-사용자 대화 모델의 수립과 인터랙션 시 발생하는 오류 상황에 안정적으로 답변할 수 있는 유연한 대화 전략 수립이 필요함을 확인했다 [15].
송지성, 정다희	HCI의 3대 요소인 유용성, 사용성, 감성을 중심으로 인공지능 검색 서비스에 대한 만족도와 신뢰도를 확인하기 위한 사용성을 분석했다. 제이콥 닐슨 Jakob Nielsen의 휴리스틱 평가 요소 10가지 원칙 중 HCI 3대 요소와 관련된 인식성, 일관성, 정확성, 직관성, 효율성, 접근성으로 6가지 주요 항목을 도출했다 [16].
신지선, 이승헌, 김진우	전자들의 차량 내 VUI의 사용성에 대한 탐구에서 VUI의 낮은 사용률의 원인은 기술적 신뢰도의 문제로 파악하여 정확한 음성인식율을 높여 VUI의 수용성을 확대하는 일이 중요하다고 주장했다 [17].

3. 연구 방법

3.1 실험 애플리케이션

본 연구에서는 애플리케이션을 통해 AI 서비스, 라이프 어시스턴트 서비스를 제공하는 네이버 클로바와 카카오의 헤이카카오의 사용성 테스트를 하기 위하여 [표 3]과 같이 기능별로 유형화하여 비교하였다.

라이프 어시스턴트의 기능으로써 생활 속 필요한 정보 검색 및 엔터테인먼트 콘텐츠 제공을 기본으로 크게 생활, 편이, 미디어로 나눌 수 있다. 두 에플리케이션 모두 생활 영역에서 기본적인 서비스로 날씨, 일정, 교통, 브리핑, 스마트홈을 제공한다. 더 나아가 헤이카카오는 카카오톡 메시지 보내기, 전화 걸기, 카카오 홈, 카카오 선물하기 기능을 활용한 주문하기로 카카오 계열사에서

연계되는 서비스를 제공한다. 편의 부분에서도 두 애플리케이션의 공통적인 기능인 알람/타이머, 메모, 번역 외에도 헤이카카오는 택시 애플리케이션을 활용한 택시 서비스, 홈트레이닝 서비스, 음성 대화를 텍스트로 받아쓸 수 있는 음성 받아쓰기 기능으로 카카오톡으로 공유할 수 있으며, 운전 중 말로 간편하게 메시지를 보내고 노래를 들을 수 있는 드라이빙 모드가 탑재되어 있다. 미디어 기능으로는 팟캐스트, 오디오북, 뮤직, 뉴스, 검색, 게임, 키즈를 두 애플리케이션 모두 공통으로 제공하며 각 서비스별 제공되는 세부적인 콘텐츠의 다양성에 차이점이 있다. 예를 들면 헤이카카오는 콩순이, 아이를 칭찬할 수 있는 칭찬 서비스, 인터랙티브 동화, 들리는 동물원, 목표 달성 스탬프 등 교육 서비스를 제공한다. 네이버 클로바는 동요, 동화, 키즈 라디오, 퀴즈를 제공한다. 그리고 각각 다른 애플리케이션과 연결되어 정보를 제공하는 네이버 클로바의 스킬스토어 기능과 헤이카카오의 플러그인 서비스가 있다.

[표 3] 라이프 어시스턴트 기능

[Table 3] The function of life assistant

분류	네이버 클로바	헤이카카오
생활	날씨, 일정, 교통, 브리핑, 스마트홈	날씨, 일정, 교통/장소, 브리핑, 카카오톡, 전화걸기, 카카오 홈, 주문하기
편의	알람, 장소, 메모, 번역	알람/타이머, 메모, 번역, 음성 받아쓰기, 드라이빙 모드, 택시, 홈트레이닝, 계산기
미디어	오디오북, 팟캐스트, 뮤직, 뉴스, 검색, 키즈, 스킬스토어	라디오, 팟캐스트, 오디오북, 뮤직, 뉴스, 검색, 게임, 키즈, 교육, 플러그인 서비스

3.3 실험 대상자 선정

설문조사는 3일간 6명에게 사전검사를 실시한 결과 인공지능 서비스 애플리케이션에 대한 50-60대 참여자들이 이해도가 낮아 인공지능 검색 서비스의 개념에 대해 이해도가 높고 모바일 활용도가 높은 조사대상의 범위를 20-40대로 조정하였다.

제이콥 닐슨에 따르면 사용성 평가에 있어 5명의 실험자를 대상으로 연구한 결과가 사용성에 대한 문제를 파악하고, 그 결과 값이 유의하다고 판단했다. 본 연구는 유효한 값을 위해 두 배수인 10명을 대상으로 네이버 클로바와 헤이카카오의 서비스 사용성 평가를 시행하였다.

3.4 사용성 평가 설계

VUI 사용성 범주 5가지 항목인 효율성, 정확성, 유연성, 일관성, 의미성과 AI의 개인맞춤화, 다양성, 편의성, 전문성, 실시간성 등 다섯 가지 특성을 매칭하여 각 특성별 평가 항목을 구성하였다. AI 서비스의 특성 중 서비스를 이용하는 고객의 개인정보를 안전하게 관리하며 문제 발생 시 적절하게 대응하는 여부에 대한 신뢰성은 서비스를 제공하는 네이버와 카카오의 내부 관리에 대한 요

소로 사용성 테스트로 결과를 입증하기 어려운 부분으로 제외하였다.

본 연구에서는 먼저 상기의 VUI 사용성 범주 항목에 AI 서비스 특성들을 나눴다. 그리고 백창화 AI 서비스의 품질평가속성 연구에서 델파이 기법으로 제안된 24개의 세부 품질 평가항목을 차용하여 두 AI서비스 애플리케이션 사용성 설문 문항을 [표 4]와 같이 구성하였다 [8].

[표 4] 사용성 평가 설문 문항

[Table 4] Questionnaire of usability test

특성	세부 특성	평가항목
효율성	편의성	◦ 서비스 이용에 있어 손쉽게 접근할 수 있다. ◦ 서비스 대응방식에서 편안함을 준다.
정확성	전문성	◦ 사용자의 요청사항에 정확하게 대응한다. ◦ 가장 최신 정보를 반영한다. ◦ 사람보다 더 수준 높은 결과를 제공한다.
유연성	개인맞춤화	◦ 개인에 맞춘 정보를 제공한다. ◦ 개인의 특성을 정확히 분석하여 알고 있다.
일관성	실시간성	◦ 사용자의 요청에 신속하게 반응한다. ◦ 사용자와 양방향으로 원활하게 상호작용한다. ◦ 실시간으로 서비스를 제공한다. ◦ 때와 장소에 적합한 정보를 제공한다.
의미성	다양성	◦ 사용자의 다양한 요구사항을 명확하게 인지한다. ◦ 사용자가 요구하는 여러 가지 서비스를 연계한다. ◦ 복합서비스 처리가 가능하다.

네이버 클로바와 헤이카카오 두 AI 서비스 기능별로 각 평가항목을 설문하였으며, 설문지는 객관성과 정확성을 위해 2022년 5월 30일부터 2022년 6월 7일에 걸쳐 대면으로 시행하였다. 설문지는 리커트 5점 척도에 따라 매우 그렇지 않다(1점), 그렇지 않다(2점), 보통이다(3점), 그렇다(4점), 매우 그렇다(5점)로 측정하였다.

3.5 사용성 테스트 결과

설문지 분석 결과 [표 5]와 같은 수치가 나타났다. 카카오와 네이버의 각 애플리케이션의 3가지 항목별 분석 결과 생활 기능에서 헤이카카오와 네이버 클로바 두 애플리케이션 모두 효율성, 편의성 부분의 점수가 각각 4점과 3.4점으로 높지만, 헤이카카오는 유연성, 개인맞춤화 부분에서 2.5점, 네이버 클로바는 의미성과 다양성에서 2.9점으로 낮게 평가되었다. 편의 기능에서는 네이버 클로바는 일관성, 실시간성에서 3.7점으로 높은 점수가 매겨졌으며, 헤이카카오는 효율성과 편의성 부분에서 4점의 점수를 얻었다. 미디어 기능에서도 모든 항목에 있어 비슷한 수치들을 보였으나 헤이카카오와 네이버 클로바 두 애플리케이션 모두 일관성과 실시간성에서 3.64점, 3.44점으로 다른 특

성들에 비해 높게 평가되었다.

[표 5] 설문지 평균 분석

[Table 5] The average analysis of survey

구분		효율성 편의성	정확성 전문성	유연성 개인맞춤화	일관성 실시간성	의미성 다양성
생활	헤이카카오	4	3.08	2.5	3.79	3.06
	네이버 클로바	3.4	3.26	3	3.29	2.9
편의	헤이카카오	4	3.7	2.8	3.59	3.04
	네이버 클로바	3.4	3.3	3.1	3.7	2.5
미디어	헤이카카오	3.4	3.58	2.8	3.64	2.78
	네이버 클로바	3.5	3.24	3	3.44	2.44

3.6 사용성 평가 분석

생활 기능에서는 [표 6]과 두 애플리케이션 모두 효율성과 편의성 부분이 좋게 평가되었다. 이는 VUI 특성상 음성으로 원하는 서비스에 접근할 수 있는 부분과 더불어 시각적인 UI도 함께 제공하기 때문이다. 헤이카카오는 명령어가 끊김에도 불구하고 앞 명령어와 이어지며 4점으로 높게 평가되었다. 그러나 네이버 클로바는 애플리케이션 메인 화면에 제공되는 서비스에 관련된 질문들을 예시로 제안하지만, 이전 명령어와 문장 및 문맥 연결이 어려워 헤이카카오보다는 상대적으로 낮은 3.4점으로 나타났다. 두 애플리케이션 모두 가장 낮게 평가된 부분은 유연성과 개인맞춤화 특성이다. 헤이카카오는 내 명령어 서비스를 통해 명령어를 설정 후 원하는 기능을 설정할 수 있지만, 직접 설정이 아닌 정해진 명령어 옵션 중 선택해야 하므로 2.5점으로 낮은 점수로 평가되었다. 네이버 클로바는 명령어를 직접 정할 수 있어 3점으로 헤이카카오보다는 비교적 높게 평가되었지만, 명령어 인식에 대한 이해도가 낮고, 정보 및 콘텐츠를 선택할 수 있는 옵션이 적어 개인에게 필요한 정보 선택의 어려움이 있었다. 또 실시간성 평가에 대해 위치에 따라 실시간으로 정보를 제공하지만, 일차적인 정보 제공에 한계가 있으며, 다른 애플리케이션으로 연결되는 부분에서 일관성이 떨어지는 것으로 나타났다.

[표 6] 생활 기능 서비스 평가 분석

[Table 6] The analysis of life service function

		헤이카카오	네이버 클로바
생활	효율성	◦ 기본적인 서비스만 제공. 더 자세한 정보는 추가 액션이 필요하며 다른 애플리케이션으로 연결 ◦ 일정 등록 시 카카오톡으로 서비스가 연결되어 알람 기능 사용이 가능	◦ 기본적인 서비스만 제공. 더 자세한 정보는 추가 액션이 필요하며 다른 애플리케이션으로 연결. ◦ 가능한 서비스의 명령어 예시를 메인 화면에 제공

	편의성	◦ 카카오톡 메시지를 보낼 때 명령어의 끊김에도 불구하고 앞 명령어와 이어지며 메시지 전달이 수월함 ◦ 정보 제공 시 카카오 이모티콘과 관련 메시지를 함께 제시해 친근함	◦ 이전 명령어와의 문장 및 문맥 연결의 어려움
	정확성	◦ 소음이 있으면 명확하게 명령어를 인식 못함 ◦ 원하는 서비스의 관련 단어만 명령해도 정확하게 대응함 ◦ 전화 걸기 기능 시 등록된 번호의 이름에 따라 인식을 정확하게 못 하는 경우가 있음 (된소리 발음 인식이 어려움)	◦ 소음이 있으면 명확하게 명령어를 인식하지 못함 ◦ 원하는 서비스의 관련 단어만 명령해도 정확하게 대응함 ◦ 일정 등록 시 네이버 캘린더 외에 구글 캘린더와 함께 연동할 수 있어 다양한 옵션을 가짐
	전문성		
	유연성	◦ 내 명령어 서비스로 명령어를 설정 후 원하는 기능을 이용할 수 있으나, 정해진 명령어에서 선택해야 함 ◦ 교통 기능에서 명령어에 구체적인 이동 수단을 언급하지 않아도 다양한 옵션을 제공 ◦ 개인이 등록한 장소 명령 인식 오류	◦ 내 명령어 서비스로 직접 명령어를 설정 후 원하는 기능을 이용할 수 있음 ◦ 교통 기능에서 명령어에 교통수단과 장소로 구체적인 문장을 완성해야만 서비스 제공 ◦ 개인이 등록한 장소 명령 인식 가능
	개인맞춤화		
	일관성	◦ 위치에 따라 실시간 정보 제공 ◦ 교통 기능에서 길 찾기 외의 다른 교통 관련 정보를 제공함	◦ 위치에 따라 실시간 정보 제공 ◦ 길 찾기 외의 다른 교통 정보를 제공하지 않음
	실시간성		
	의미성	◦ 전반적인 서비스가 일차적이며 고도화가 필요 ◦ 서비스 제공에 관련된 다양한 정보를 아래에 제시함	◦ 전반적인 서비스가 일차적이며 고도화가 필요 ◦ 서비스 제공에 관련된 다양한 정보를 아래에 제시함
	다양성		

편의 기능에서는 [표 7]과 헤이 카카오는 효율성과 편의성 부분에서 4점으로 높게 평가되었다. 다른 애플리케이션과 연계되는 번거로움에도 불구하고, 피실험자가 모두 사용하고 있는 카카오톡 애플리케이션으로 이어진다. 피실험자들은 이를 편리하다고 여기며 효율성과 편의성에 큰 점수를 매겼다. 그러나 헤이카카오와 네이버 클로바 모두 타이머 기능, 알람 기능을 사용할 수 있는 문구가 있지만, 실제로 애플리케이션에서는 지원하지 않아 피실험자들에게 혼란을 가져다주었다. 이는 인공지능 스피커를 통해 지원되는 서비스로 기기별 구별이 필요해 보인다. 네이버 클로바는 애플리케이션 내에서 리마인더 기능과 메모를 확인할 수 있는 개인 화면을 제공함으로써 개인 맞춤화와 유연성 부분에서 3.1점으로 헤이카카오보다는 높은 점수로 평가되었다. 그러나 카카오 계열사와 연계되는 다양한 서비스를 제공하는 헤이카카오와 달리 제공되는 서비스가 적은 네이버 클로바가 상대적으로 의미성과 다양성 부분이 2.5점으로 낮게 평가되었다.

[표 7] 편의 기능 서비스 평가 분석

[Table 7] The analysis of convenience service function

		헤이카카오	네이버 클로바
생활	효율성	◦ 카카오 계열사의 연계 애플리케이션으로 다양한 서비스를 이용할 수 있음 ◦ 알람 기능은 지원이 안 됨 ◦ 번역, 음성 받아쓰기, 드라이빙 모드와 같은 주요 기능을 시각적인 UI와 함께 전달해 쉽게 접근 가능	◦ 알람, 타이머 기능은 애플리케이션에서 지원이 안 됨 ◦ 개인적인 메모와 알람 서비스를 볼 수 있는 탭이 별도로 마련되어 있음 ◦ 이전 명령어와의 문장 및 문맥 연결의 어려움
	편의성		
	정확성	◦ 19개의 언어로 번역이 가능 ◦ 음성 받아쓰기 서비스에서는 발음이 명확하지 않으면 문맥과 관련 없는 다른 단어로 입력이 됨 ◦ 명령어가 다를 경우 반응이 없음	◦ 4개의 언어로 번역이 가능 ◦ 번역 기능에서는 하나의 정보를 다양한 옵션으로 제공 ◦ 명령어가 다를 경우 네이버 검색으로 연결됨
	전문성		
	유연성	◦ 음성을 인식해 메모나 받아쓰기를 할 때 중간에 수정 되지 않음	◦ 애플리케이션 내에 리마인더, 메모 등을 확인할 수 있는 서비스가 탑재되어 있음
	개인맞춤화		
	일관성	◦ 위치와 상황에 따라 실시간으로 서비스를 제공	
	실시간성		
	의미성	◦ 홈트레이닝을 위한 타이머 기능은 있으나 콘텐츠를 제공하지 않는 일차적인 부분이 아쉬움	◦ 서비스가 한정적이고, 예시 명령어를 통해 명확한 서비스 파악이 어려움
	다양성		

미디어 기능에서는 [표 8]과 같이 두 애플리케이션 모두 일관성과 실시간성이 가장 높은 점수로 평가되었다. 헤이카카오는 실시간 정보 검색이 되지만 네이버 클로바는 정보 유형에 따라 가능 여부가 달라졌다. 예를 들면 스포츠 경기가 있던 날, 현재 점수를 알고 싶을 때 카카오는 정보를 실시간으로 명확하게 제공하는 반면 네이버 클로바는 정보를 제공하지 못하고 네이버 검색 결과로 연결되도록 나타났다. 이런 점에 헤이 카카오가 3.64점으로 3.44 네이버 클로바 점수보다 높게 평가되었다. 그러나 두 애플리케이션 모두 뉴스 서비스 중 각각 기본으로 연결되는 뉴스 방송사가 정해져 있으며 이를 제외하고 다른 채널의 뉴스를 제공하는 방식이 기존과 다른 점으로 일관성이 떨어졌다. 효율성과 편의성 부분에서는 네이버 클로바가 헤이카카오보다 0.1점 더 높은 점수를 받았다. 이는 시각적인 요소를 제공함으로써 다양한 콘텐츠를 선택할 수 있도록 한 점과 관련 서비스의 VUI 가이드를 함께 전달함으로써 더 다양한 서비스를 받을 수 있기 때문이다. 그리고 유연성과 개인맞춤화에서 연계되는 애플리케이션의 옵션뿐만 아니라 키즈 서비스 중 학교 정보도 등록할 수 있어 네이버 클로바가 3점으로 2.8점의 헤이카카오보다 더 높게 평가되었다.

[표 8] 미디어 기능 서비스 평가 분석

[Table 8] The analysis of media function

		헤이카카오	네이버 클로바
미디어	효율성	◦ 뉴스 서비스 중 한 기사만 제안함 ◦ 카카오톡 애플리케이션과 연계가 되어 주문하기 기능이 가능(음식에 한함)	◦ 오디오 클립 서비스 중 다른 채널 선택 및 중지 가이드를 함께 전달함 ◦ 뮤직 검색이 되지 않음 ◦ 뉴스 서비스 중 다양한 기사를 보고 고를 수 있도록 시각적으로 제안함 ◦ 오디오북 서비스 시 콘텐츠 리스트를 보여줌으로써 이용이 편리함.
	편의성		
	정확성	◦ 명령어가 긴 문장일지라도 문맥을 파악해 서비스를 제공	◦ 명령어가 긴 문장일 경우 문맥 파악이 어려움
	전문성		
	유연성	◦ 연결되는 음악 애플리케이션이 한 가지(멜론)로 이용권이 없는 사용자들은 불편함. ◦ 키즈 서비스 중 자녀 정보 등록 가능. ◦ 뉴스 방송사가 한정적임	◦ 연결되는 음악 애플리케이션이 세 가지로 선택할 수 있음(바이브, 지니, 벅스) ◦ 음악 재생 시 연결된 개인 뮤직 애플리케이션에 있는 플레이리스트가 재생이 안 됨(바이브는 가능) ◦ 키즈 서비스 중 자녀 정보 등록 및 학교 등록도 가능해 정보를 받을 수 있음 ◦ 뉴스 방송사가 다양함
	개인맞춤화		
	일관성	◦ 뉴스 서비스 중 기사 본문을 함께 보여주는 뉴스가 있는 반면 본문 없이 음성으로만 제공하는 뉴스가 있음 ◦ 실시간 정보 검색이 가능	◦ 음성으로 읽어주는 뉴스가 있지만, 시각적으로 기사로만 전달하여 네이버 애플리케이션으로 연결되는 기사가 있음 ◦ 실시간 정보 검색이 정보의 유형에 따라 가능 여부가 달라짐
	실시간성		
	의미성	◦ 영상 서비스를 지원하지 않음	◦ 유튜브 영상 검색 및 서비스 가능
	다양성		

4. 결론

본 연구는 향후 인공지능 서비스 플랫폼의 사용성 향상을 위한 도움에 가치를 두고자, 음성 인식 기반 지능형 개인 비서 서비스의 보편화에 활용방안이 될 수 있도록 하는 데 목적이 있다. 이에 국내에서 인공지능 스피커를 넘어 독자적인 애플리케이션으로 확대되어 사용할 수 있는 네이버의 클로바와 카카오의 헤이카카오 애플리케이션의 AI 비서 서비스에 대해 사용성 평가를 진행하였다. VUI 사용성 범주 항목인 효율성과 정확성, 유연성, 일관성, 의미성을 기본으로 AI 서비스 특성 편의성, 전문성, 개인맞춤화, 실시간성, 다양성을 매칭하여 14개의 세부 품질평가를 도출하였다.

분석 결과 카카오의 헤이카카오는 카카오 계열사에서 선보이고 있는 다양한 서비스를 연결해서

주문, 카카오톡 보내기, 택시 부르기, 드라이빙 모드, 전화 걸기, 홈트레이닝 등 다양한 서비스를 제공한다. 반면 네이버 클로바는 카카오 서비스와 비교했을 때 서비스의 다양성은 부족하지만 각 서비스당 연결되는 애플리케이션의 옵션을 여러 개 제안해 사용성을 높였다. 그리고 서비스당 제공하는 콘텐츠가 헤이카카오보다는 다양한 것을 알 수 있었다. 그러나 두 애플리케이션 모두 기능마다 명령어의 끊김을 이해하는 수준이 달라 일관성이 떨어졌다. 그리고 애플리케이션 내에서 모든 정보와 서비스를 제공하는 것이 아닌 다른 애플리케이션을 통해 복합처리가 되는 불편함이 있었다. 무엇보다 일차원적인 서비스 제공으로 고차원적인 서비스가 필요하며, 주변 소음을 구분하는 것에 대한 정확성을 높이는 것이 필요하다.

날씨와 일정, 교통 관련된 서비스를 제공하는 생활 영역에서 헤이카카오는 이모티콘과 함께 시각적 요소를 함께 제공함으로써 사용성 및 친근감을 높여주는 부분을 확인할 수 있었다. 네이버 클로바는 구체적인 명령어를 전달하지 않으면 서비스 제공에 어려우며 한정적인 정보 제공으로 유연성과 다양성 부분을 보완할 필요가 있다.

알람, 장소, 메모, 번역과 같은 편의 영역에서는 기능별로 플랫폼에 대한 구별이 필요해 보인다. 그리고 두 애플리케이션 모두 명령어가 다를 경우 헤이카카오는 반응이 없고, 네이버 클로바는 네이버 검색 서비스로 연결이 되어 명령어의 인식에 대한 폭을 넓힐 수 있는 기술적 훈련이 더 필요하며, 단어만으로도 문맥을 유추해 서비스를 제안할 필요가 있다.

뉴스, 검색, 뮤직, 키즈 등 다양한 미디어를 제공하는 서비스에서는 제공되는 콘텐츠가 부족하다. 헤이카카오는 정해진 뉴스채널 외에는 서비스 제공이 안 되며, 두 애플리케이션 모두 전달되는 방식이 VUI가 아닌 UI 또는 VUI로만 전달한다는 점들로 일관성 부분에서 떨어졌다. VUI와 UI를 함께 사용해 다양한 콘텐츠를 선택하고 제공받을 수 있는 편의성을 높여야 한다.

본 연구의 분석 결과를 바탕으로 인공지능 서비스 플랫폼 사용성이 최적화되고, 일상생활에서 효율적이고 편리한 개인 맞춤형 인공지능 비서 서비스 발전에 활용될 것으로 기대한다.

References

[1] H. J. Park, G. M. Choi, “The domestic artificial intelligence market is small! AI starts-ups are targeting the global market with technology”, Artificial Intelligence Times, aitimes.kr, http://www.aitimes.kr/news/articleView.html?idxno=21461, (accessed July 25, 2022).

[2] K. J. Lee, E. Y. Kim, “The Role and Effect of Artificial Intelligence (AI) on the Platform Service Innovation: The Case Study of Kakao in Korea”, The Knowledge Management Society of Korea, vol. 21, no. 1, January 2020, pp. 175-195, doi: 10.15813/kmr.2020.21.1.010.

[3] T. J. Yoo, “Pay attention to the AI platform where the AI service starts”, sti.kostat.go.kr, http://sti.kostat.go.kr/window/2017b/html/2017_win_8.html, (accessed July 25, 2022).

[4] S. M. Park, “Jarvis, help me!… AI secretary technology, how far have you come?”, sisaweek.com, http://www.sisaweek.com/news/articleView.html?idxno=142492, (accessed July 25, 2022).

[5] Y. J. Kang, “An Exploratory Study on the Activation of Speech Recognition AI Assistants through the Relation Analysis among Business Ecosystems’s Operators*”, The e-Business Studies, vol. 22, no. 5, October 2021, pp. 95-112, doi: 10.20462/tebs.2021.10.22.5.95.

[6] M. Porcheron, J. E. Fischer, S. Reeves, S. Sharples, “Voice Interfaces in Everyday Life”, 2018 CHI Conference on Human Factors in Computing Systems, April 21-26, 2018, Montréal, QC, Canada, pp. 1-12, doi: 10.1145/3173574.3174214.

[7] B. P. Kim, “If Google and Amazon dominate the artificial intelligence platform...”, www.joongang.co.kr, https://www.joongang.co.kr/article/23500429#home, (accessed July 25, 2022).

[8] C. H. Baek, S. U. Lim, J. H. Choe, “A Study on Major Characteristic Analysis and Quality Evaluation Attributes of Artificial Intelligence Servic”, Journal of Korean Society for Quality Management, vol. 47, no. 4, December 2019, pp. 837-846, doi: 10.7469/JKSQM.2019.47.4.837.

[9] Y. S. Jeon, H. W. Lee, H. A. Seo, “Development and Application of Teacher-Friendly AI Service Platform for Individual Customized Career Guidances”, The Korean Association Of Computer Education, vol. 25, no. 1, December 2022, pp. 59-71, doi: 10.32431/kace.2022.25.1.006.

[10] C. H. Baek, “A Study on Service Quality Evaluation in Artificial Intelligence Era Using Delphi Technique”, Journal of Korea Service Management Society, vol. 21, no. 3, September 2020, pp. 1-15, doi: 10.15706/jksms.2020.21.3.001.

[11] U. Y. Jeong, S. J. Cho, “A Study on the Usability Evaluation Items of AI Speakers through AHP Analysis”, Journal of the Ergonomics Society of Korea, vol. 41, no. 2, April 2022, pp. 53-61, doi: 10.5143/JESK.2022.41.2.53.

[12] J. U. Na, J. Y. Heo, Y. H. Pan, “Classification of customer Utility for Voice Interactive Function of AI Personal Assistant Services : Based on KANO Model”, Design Convergence Study, vol. 16, no. 4, August 2017, pp. 67-80.

[13] M. J. Kim, J. E. Han, H. J. Kang, G. H. Kwon, “Extracting usability dimensions of the voice user interface - Focusing on AI assistants-”, The HCI Society of Korea, vol. 15, no. 1, March 2020, pp. 53-64, doi: 10.17210/jhsk.2020.03.15.1.53.

[14] Y. Song, Y. Yang, P. Cheng, “The Investigation of Adoption of Voice-User Interface (VUI) in Smart Home Systems among Chinese Older Adults”, Sensors, vol. 22, no. 4, February 2022, pp. 1614, doi: 10.3390/s22041614.

[15] J. M. Cheon, S. H. Han, Y. S. Cho, W. G. Park, J. S. Kim, “An analysis of exceptional dialogue pattern for a conversational speech interface”, Ergonomics Society of Korea, October 26-28, 2006, Daejeon, SouthKorea, pp. 81-84.

[16] J. S. Song, D. H. Jung, “A Study on the Usability of AI-based Naver App Search Service”, Journal of the korean society design culture, vol. 27, no. 1, March 2021, pp. 197-207, doi: 10.18208/ksdc.2020.27.1.197.

[17] J. S. Shin, S. H. Lee, J. W. Kim, “Explore the Usability of the Driver's Voice Interface in the Vehicle”, The HCI Society of Korea, February 11-14, 2020, Hongcheon, SouthKorea, pp. 567-577.