https://doi.org/10.33778/kcsa.2024.24.3.117

LLM 기반의 생성형 AI 응답 데이터 품질이 업무 활용 만족도에 미치는 영향에 관한 연구

이 승 환*, 현 지 은**, 김 광 용***

요 약

2017년 새로운 형태의 아키텍처인 트랜스포머(Transformer)가 발표되면서 언어모델에도 많은 변화가 있었다. 특히 대형 언어 모델인 LLM(Large language model)의 발전으로 검색이나 챗봇(Chatbot)과 같은 생성형 AI 서비스가 다양한 업무 영역에 활용되고 있다. 하지만 개인정보 유출과 같은 보안 이슈나 거짓 정보를 생성하는 할루시네이션(Hallucination)과 같은 신뢰성 문제가 발생하면서 이러한 서비스의 실효성에 대한 우려의 목소리도 커지고 있다. 이에 본 연구에서는 이러한 우려에도 불구하고 생성형 AI를 업무 영역에 활용하고 있는 빈도가 점점 증가하고 있는 요인에 대해서 분석하고자 하였다. 이를 위해 LLM 기반의 생성형 AI 응답 데이터 품질에 영향을 미치는 8가지 요인을 도출하고 유효 표본 195개를 대상으로 이러한 요인들이 업무 활용 만족도에 미치는 영향을 실증 분석하였다. 분석결과 전문성, 접근성, 다양성, 편리성이 지속적 사용의도에 유의한 영향을, 보안성, 안정성, 신뢰성 등이 부분적으로 유의한 영향을, 완전성이 부정적 영향을 미치는 요인으로 나타났다. 본 연구에서는 응답 데이터 품질에 대한 수요자의 인식이 업무 활용 만족도에 어떠한 영향을 미치는지 학문적으로 규명하고, 이러한 서비스에 대한 수요자 중심의 의미 있는 실무적 시사점을 제시하는데 그 목적이 있다.

A Study of how LLM-based generative AI response data quality affects impact on job satisfaction

Lee Seung Hwan*, Hyun Ji Eun**, Gim Gwang Yong***

ABSTRACT

With the announcement of Transformer, a new type of architecture, in 2017, there have been many changes in languag e models. In particular, the development of LLM (Large language model) has enabled generative AI services such as searc h and chatbot to be utilized in various business areas. However, security issues such as personal information leakage and reliability issues such as hallucination, which generates false information, have raised concerns about the effectiveness of t hese services. In this study, we aimed to analyze the factors that are increasing the frequency of using generative AI in t he workplace despite these concerns. To this end, we derived eight factors that affect the quality of LLM-based generativ e AI response data and empirically analyzed the impact of these factors on job satisfaction using a valid sample of 195 res pondents. The results showed that expertise, accessibility, diversity, and convenience had a significant impact on intention to continue using, security, stability, and reliability had a partially significant impact, and completeness had a negative imp act. The purpose of this study is to academically investigate how customer perception of response data quality affects busi ness utilization satisfaction and to provide meaningful practical implications for customer-centered services.

Key words : (LLM, Generative AI, System Quality, Information Quality, Service Quality, TAM, IS Success)

접수일(2024년 06월 28일), 게재확정일(2024년 07월 09일)

* 숭실대학교 IT정책경영학과(주저자)
** 숭실대학교 IT정책경영학과(공동저자)
*** 숭실대학교 경영학부(교신저자)

1. 서 론

BERT 및 GPT 등과 같은 대형 언어 모델(LLM)의 등장은 인공지능 기술의 전개를 바꿔놓았다[1]. 특히 LLM 기반의 생성형 AI는 최근 몇 년 동안 괄목할 만한 성장을 이루면서 다양한 영역의 애플리케이션 서비스로 이어졌다[2]. 하지만 이러한 생성형 AI는 사람이 지각하지 못한 사이에 존재하지 않는 사실을 조작해서 환각에 빠뜨리는 할루시네이션과 같은 현상을 일으킬 수 있어 이러한 응답 데이터를 완전히 신뢰하기는 힘든 실정이다[3]. 또한 개인정보 유출과 같은 취약점을 드러내며 보안 관련 커뮤니티에서도 주목받았다[4]. 그럼에도 불구하고 지난 수년간 다양한 생성형 AI 서비스가 출시되고 있다. 구체적으로, 이러한 서비스는 사용자의 일반적인 질문에 대해 답변을 제공하거나 예술적인 이미지를 자동으로 생성하는 등의 작업을 수행할 수 있다[5]. 특히 인간의 고유 영역으로 분류되던 문학, 음악, 영화, 미술과 같은 예술 분야에서도 활용할 수 있다. 기업에서는 인사, 영업, 제품 개발, 마케팅, 고객 관리와 같은 다양한 비즈니스 프로세스 영역에서의 활용을 통해 직원의 업무수행 방식을 변화시킬 것이다[6]. 이렇듯 생성형 AI는 보안성, 신뢰성과 같은 이슈에도 불구하고 다양한 업무 영역에서 활용하는 빈도가 점점 높아지고 있다. 하지만, 이러한 생성형 AI의 응답 데이터 품질이 업무 활용 만족도에 미치는 영향에 대한 연구는 부족한 실정이다. 이에 본 연구에서는 LLM 기반의 생성형 AI 응답 데이터 품질이 업무 활용 만족도에 미치는 영향에 대해 분석해 보고자 한다. 이를 위해 다양한 선행연구를 바탕으로 응답 데이터 품질에 영향을 미치는 8가지 요인을 도출하였다. 이렇게 도출한 요인과 업무 활용 만족도 및 지속적 사용의도 간에 관계를 실증 분석하여 업무에 지속해서 사용하고자 하는 수요자 중심의 학문적, 실무적 시사점을 제시하고자 한다. 나아가 LLM 기반의 생성형 AI 응답 데이터를 업무에 활용하고자 하는 잠재적인 수요에 대한 기대와 함께 서비스 측면에서의 품질 향상에도 기여하고자 한다.

2. 이론적 배경

2.1 LLM(Large language model)

일반적으로 LLM은 BERT, GPT, LLaMA 및 PaLM과 같은 방대한 텍스트 데이터에 대해 훈련된 수천억 또는 그 이상의 매개변수를 포함하는 트랜스포머(Transformer)를 기반으로 하는 언어 모델을 의미한다[7]. 트랜스포머는 기존의 RNN(Recurrent neural network)이나 CNN(Convolutional neural network)과 같은 알고리즘에서 사용하던 반복 순환 연산 방식을 배제하고 어텐션 메커니즘(Attention mechanism)[8]을 기반으로 언어를 처리하는 네트워크 아키텍처이다[9]. 트랜스포머의 등장으로 LLM을 기반으로 하는 검색, 챗봇과 같은 형태의 생성형 AI 서비스가 발전하는 계기가 되었다.

2.2 생성형 AI(Generative AI)

생성형 AI는 훈련 데이터로부터 텍스트, 이미지 또는 오디오와 같은 새롭고 의미 있는 콘텐츠를 생성할 수 있는 컴퓨팅 기술을 의미한다. 새로운 텍스트나 새로운 이미지를 만드는 목적으로 사용될 수 있을 뿐만 아니라 지능형 질의응답 시스템을 통해 인간과의 소통을 지원할 수 있다[10]. 골드만삭스 리서치(Goldman Sachs Research)에 따르면 900개가 넘는 직업에 대한 데이터베이스를 분석한 결과, 직업의 약 3분의 2가 생성형 AI에 의해 자동화 하는 것이 가능하다고 추정하였다[11]. Avishek Choudhury & Hamid Shamszare(2023)는 사용자 신뢰가 ChatGPT와 같은 생성형 AI 채택 및 사용에 미치는 영향에 대해 연구[12] 하였고, 이동영(2024)은 AI 리터러시 조절효과를 중심으로 ChatGPT 지속사용의도에 관해 연구[13] 하였으며 신재형(2024)은 인공지능 학습데이터의 수집 및 이용에 대한 규제[14]에 대해 연구하였다. 이렇듯 국내외 생성형 AI에 관한 연구는 다양한 분야에서 이루어져 왔지만 주로 ChatGPT를 중심으로 하는 수용의도나 알고리즘 개선을 위한 기술 자체에 대한 연구가 많아 수요자 중심의 생성형 AI 응답 데이터 품질에 영향을 미치는 요인에 관해서는 연구가 드문 실정이다.

2.3 기술수용모델(TAM)

기술수용모델(TAM, Technology Acceptance Model)[15]은 조직에서 컴퓨터 시스템을 사용하는데 영향을 미치는 요인에 대해 체계적으로 정리한 이론으로 1989년 Davis에 의해 제안되었다. Davis는 이 논문에서 지각된 유용성과 지각된 사용용이성이 컴퓨터 시스템 사용의도와 실제 사용행위에 영향을 준다고 가정하였다. 이후 2001년도에 Bhattacherjee에 의해 후기수용모델(PAM, Post Acceptance Model)로 발전 되었는데 기술수용모델이 정보시스템 사용에 대한 모델이라면 후기수용모델은 정보시스템의 사용에 대한 만족도와 지속적인 사용의도에 초점을 두고 제안된 이론이다[16]. 본 논문에서는 기술수용모델의 선행연구를 바탕으로 지각된 유용성과 지각된 사용용이성이 만족도 및 지속적 사용의도에 미치는 영향을 분석하였다.

2.4 정보시스템 성공모델(IS Success Model)

정보시스템 성공모델(ISSM, Information System Success Model)[17]은 정보시스템에 대한 사용자 만족에 영향을 미치는 요인에 대한 이론으로 1992년 DeLone & McLean에 의해 제안되었다. DeLone & McLean은 1981년부터~1987년 사이에 발표된 정보시스템 성공 요인과 관련된 논문 약 180편에 대한 연구를 바탕으로 정보시스템 성공모델을 구성하였다. 시스템 품질(System Quality), 정보 품질(Information Quality), 사용(Use), 사용자 만족(User Satisfaction), 개인적 영향(Individual Impact), 조직적 영향(Organizational Impact)의 총 6가지 요인으로 구성되어 있으며 정보시스템 품질을 시스템 품질과 정보 품질로 구분하고 이 두 가지의 품질에 따라 사용자 만족도가 달라진다고 가정하였다. 이후 Pitt et al.(1995)에 의해 서비스 품질을 추가한 수정된 정보시스템 성공모델이 제시되었다[18]. 본 논문에서는 정보시스템 성공모델의 선행연구를 바탕으로 LLM 기반의 생성형 AI 응답 데이터 품질에 영향을 미치는 요인을 시스템 품질, 정보 품질, 서비스 품질 관점에서 도출하였다.

2.4.1 시스템 품질

시스템 품질은 정보시스템 자체에 대한 품질을 의미한다[19]. 유은정(2020)은 시스템 품질의 편리성, 보안성이 지각된 유용성, 지각된 사용용이성에 미치는 영향에 대해 검증하였다[20]. 문은경(2022)은 시스템 품질의 안정성이 지각된 유용성에 미치는 영향에 대해 연구하였다[21]. 본 연구에서는 이러한 선행 연구를 바탕으로 지각된 유용성과 지각된 사용용이성에 영향을 미치는 시스템 품질의 요인으로 보안성, 편리성, 안정성을 선택하였다.

2.4.2 정보 품질

정보 품질은 시스템 출력의 특성으로 이해하기 쉽고 정확해야 하며, 필요한 모든 정보가 다양하고 완전하게 포함되어야 한다[22]. 박성준(2012)은 정보 품질의 신뢰성이 지각된 유용성에 미치는 영향에 대해 검증하였다[23]. 이승육(2015)은 정보 품질의 완전성이 지각된 유용성과 지각된 사용용이성에 미치는 영향에 대해 실증하였다[24]. 김광철(2023)은 정보품질의 다양성이 만족도에 미치는 영향에 대해 연구하였다[25]. 본 연구에서는 이러한 문헌적 연구를 바탕으로 정보 품질의 요인으로 완전성, 다양성, 신뢰성을 본 논문에 맞게 보완 하였다.

2.4.3 서비스 품질

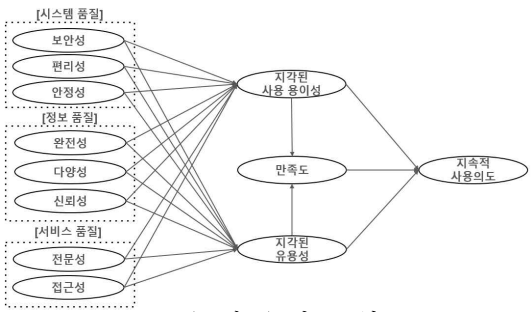
서비스 품질은 서비스에 대한 사용자 기대치와 서비스 경험 후의 사용자 인식 간의 차이로 정의할 수 있다. 서비스 품질이 높아질수록 만족도와 재사용 의향이 높아진다[26]. 임나윤(2023)은 서비스 품질이 지각된 유용성, 지각된 사용용이성에 미치는 영향을 검증하였다[27]. 최주원(2021)은 서비스 품질의 전문성, 반응성이 지각된 유용성, 지각된 사용용이성에 미치는 영향에 대해 연구하였다[28]. 박명준(2022)은 접근용이성이 지각된 유용성에 미치는 영향에 대해 실증하였다[29]. 이러한 연구 결과를 바탕으로 서비스 품질의 요인으로 전문성, 접근성을 선정 하였다.

3. 연구모형 및 가설설정

3.1 연구모형

LLM 기반의 생성형 AI 응답 데이터 품질이 업무

활용 만족도에 미치는 영향을 알아보기 위하여 이론적 배경에서 제시한 기술수용모델과 정보시스템 성공모델의 선행연구를 기반으로 아래 (그림 1)과 같은 연구모형을 설계하였다.



(그림 1) 연구모형

3.2 변수의 조작적 정의

본 연구에서는 LLM 기반의 생성형 AI 응답 데이터 품질이 업무 활용 만족도에 미치는 영향을 분석하기 위해 독립변수로 시스템 품질(보안성, 편리성, 안정성), 정보 품질(완전성, 다양성, 신뢰성), 서비스 품질(전문성, 접근성)의 관점에서 도출하였다. 매개변수와 종속변수로는 지각된 유용성, 만족도, 지각된 사용용이성, 지속적 사용의도로 하였으며 각 변수에 대한 조작적 정의 및 측정 항목은 다음의 <표 1>과 같다.

<표 1> 조작적 정의 및 측정 항목

측정 항목		내용	참고
시스템 품질	보안성	● LLM 기반의 생성형 AI 시스템은 위변조 가능성이 없고, 사용자의 개인정보가 안전하게 보호된다고 인식하는 정도	[19] [20] [21]
		1. 외부 위협으로부터 보호된다 2. 안전하게 전송된다 3. 개인정보는 안전하게 보호된다 4. 위조 및 변조 위협으로부터 안전하다 5. 개인정보는 다른 사람이 볼 수 없다	
	편리성	● LLM 기반의 생성형 AI 응답 데이터를 이용하면 원하는 정보를 쉽게 조회하고 이용할 수 있다고 인식하는 정도	
		1. 자료를 요약해주기 때문에 편리하다 2. 정보를 한꺼번에 찾아주기 때문에 편리하다 3. 체계적으로 정보가 분류되어 있어 편리하다 4. 정보를 복사해서 가져올 수 있어 편리하다 5. 여러 정보를 취합해주기 때문에 편리하다	
	안정성	● LLM 기반의 생성형 AI 응답 데이터는 오류나 중단, 끊김 없이 일정한 속도로 제공된다고 인식하는 정도	
		1. 끊김 없이 전송된다 2. 전송 시 오류 발생 빈도가 낮다 3. 전송이 지연되지 않는다 4. 반응 속도가 빠르다 5. 응답 속도는 느리지 않다	
정보 품질	완전성	● LLM 기반의 생성형 AI 응답 데이터는 요구한 범위의 정보를 누락 없이 제공하고 있다고 인식하는 정도	[22] [23] [24] [25]
		1. 요구한 범위의 정보를 제공한다 2. 나의 요구를 충족시킨다 3. 모든 범위를 포함한다 4. 요구한 범위를 충족시킨다 5. 필요한 정보를 모두 제공한다	
	다양성	● LLM 기반의 생성형 AI 응답 데이터는 여러 가지 주제에 대해서 풍부한 정보가 제공된다고 인식하는 정도	
		1. 다양한 분야의 정보를 제공한다 2. 여러 주제의 정보를 제공한다 3. 여러 분야의 정보를 제공한다 4. 제공되는 정보는 다양하다 5. 다양한 주제의 정보를 제공한다	
	신뢰성	● LLM 기반의 생성형 AI 응답 데이터는 믿을 수 있는 출처로부터 사실과 일치하는 정보를 제공한다고 인식하는 정도	
		1. 학습데이터를 신뢰한다 2. 출처는 믿을 수 있다 3. 정보제공 프로세스는 믿을 수 있다 4. 정보는 믿을 수 있다 5. 믿을 수 있는 학습 데이터를 사용한다	
서	전	● LLM 기반의 생성형 AI는 전문지식	[26]

비스품질	문성	과 능력을 보유하고 있다고 인식하는 정도	
		1. 전문 지식 데이터를 가지고 있다 2. 전문 지식을 분석해서 알려준다 3. 문제 해결을 위한 정보제공 능력이 있다 4. 전문적인 응답을 제공한다 5. 데이터 학습에 의해 전문성이 높다	〔27〕 〔28〕 〔29〕
	접근성	● LLM 기반의 생성형 AI 서비스는 다양한 디바이스에서 쉽게 접속할 수 있다고 인식하는 정도	
		1. 필요할 때 언제나 접속할 수 있다 2. 언제 어디에서나 쉽게 접속할 수 있다 3. 스마트폰으로도 접속할 수 있다 4. 다양한 디바이스에서 접속할 수 있다 5. 시간과 장소에 관계없이 접속할 수 있다	
지각된 사용용이성		● LLM 기반의 생성형 AI 서비스는 적은 노력으로도 쉽게 사용할 수 있다고 인식하는 정도	
		1. 사용하기 용이하다 2. 사용법을 배우기 쉽다 3. 사용하기 용이한 형태로 제공된다 4. 적은 노력으로 사용법을 배울 수 있다 5. 사용하기 쉽다고 생각한다	
지각된 유용성		● LLM 기반의 생성형 AI 응답 데이터는 업무에 도움이 된다고 믿는 정도	
		1. 업무 처리에 도움이 된다 2. 업무를 해결하는데 도움이 된다 3. 업무 시간을 줄이는데 도움이 된다 4. 업무를 개선하는데 도움이 된다 5. 업무 지원에 도움이 된다	〔15〕 〔16〕 〔17〕 〔18〕
만족도		● LLM 기반의 생성형 AI 응답 데이터로부터 얻는 정보에 대해 만족하는 정도	
		1. 데이터 품질에 만족한다 2. 데이터 활용에 만족한다 3. 활용 경험에 만족한다 4. 정보에 만족한다 5. 응답 내용에 만족한다	
지속적 사용의도		● LLM 기반의 생성형 AI 서비스는 가치가 있다고 생각하여 계속하여 사용하고자 하는 의지의 정도	
		1. 계속해서 업무에 사용할 것이다 2. 지속적으로 사용할 것이다 3. 업무에 자주 사용할 것이다 4. 계속해서 사용 빈도를 늘릴 것이다 5. 업무에 지속적으로 사용할 것이다	

3.2 가설설정

본 연구에서는 LLM 기반의 생성형 AI 응답 데이터 품질이 업무 활용 만족도 및 지속적 사용의도에 미치는 영향을 확인하기 위해 다음과 같은 21개의 가설을 설정하였다.

3.2.1 시스템 품질에 대한 가설

H1-1. 응답 데이터의 보안성은 지각된 유용성에 정(+)의 영향을 미칠 것이다.

H1-2. 응답 데이터의 편리성은 지각된 유용성에 정(+)의 영향을 미칠 것이다.

H1-3. 응답 데이터의 안정성은 지각된 유용성에 정(+)의 영향을 미칠 것이다.

H1-4. 응답 데이터의 보안성은 지각된 사용용이성에 정(+)의 영향을 미칠 것이다.

H1-5. 응답 데이터의 편리성은 지각된 사용용이성에 정(+)의 영향을 미칠 것이다.

H1-6. 응답 데이터의 안정성은 지각된 사용용이성에 정(+)의 영향을 미칠 것이다.

3.2.2 정보 품질에 대한 가설

H2-1. 응답 데이터의 완전성은 지각된 유용성에 정(+)의 영향을 미칠 것이다.

H2-2. 응답 데이터의 다양성은 지각된 유용성에 정(+)의 영향을 미칠 것이다.

H2-3. 응답 데이터의 신뢰성은 지각된 유용성에 정(+)의 영향을 미칠 것이다.

H2-4. 응답 데이터의 완전성은 지각된 사용용이성에 정(+)의 영향을 미칠 것이다.

H2-5. 응답 데이터의 다양성은 지각된 사용용이

성에 정(+)의 영향을 미칠 것이다.

H2-6. 응답 데이터의 신뢰성은 지각된 사용용이성에 정(+)의 영향을 미칠 것이다.

3.2.3 서비스 품질에 대한 가설

H3-1. 응답 데이터의 전문성은 지각된 유용성에 정(+)의 영향을 미칠 것이다.

H3-2. 응답 데이터의 접근성은 지각된 유용성에 정(+)의 영향을 미칠 것이다.

H3-3. 응답 데이터의 전문성은 지각된 사용용이성에 정(+)의 영향을 미칠 것이다.

H3-4. 응답 데이터의 접근성은 지각된 사용용이성에 정(+)의 영향을 미칠 것이다.

3.2.4 만족도에 관한 가설

H4-1. 응답 데이터의 지각된 유용성은 만족도에 정(+)의 영향을 미칠 것이다.

H4-2. 응답 데이터의 지각된 사용용이성은 만족도에 정(+)의 영향을 미칠 것이다.

3.2.5 지속적 사용의도에 관한 가설

H5-1. 응답 데이터의 지각된 유용성은 지속적 사용의도에 정(+)의 영향을 미칠 것이다.

H5-2. 응답 데이터의 만족도는 지속적 사용의도에 정(+)의 영향을 미칠 것이다.

H5-3. 응답 데이터의 지각된 사용용이성은 지속적 사용의도에 정(+)의 영향을 미칠 것이다.

4. 실증분석 및 결과

4.1 분석방법

본 연구는 LLM 기반의 생성형 AI 응답 데이터 품질이 업무 활용 만족도에 미치는 영향에 대하여 조사하기 위해 리커트 7점 척도의 설문지를 사용하였다. 온라인 설문을 통해 203부의 표본을 수집하였고, 그 중 유효 표본 195부의 설문 응답을 실증분석에 이용하였다. 회수된 설문 분석을 위해 SPSS 22.0 프로그램을 이용하여 응답자 현황을 인구통계학적 특성으로 분석 하였으며, 탐색적 요인분석과 내적 일관성 분석을 통해 변수의 타당성과 신뢰성을 검증하였다.

4.2 인구통계학적 특성

빈도 분석을 통해 성별, 나이, 직업에 대한 인구통계학적 특성을 조사하였으며 결과는 <표 2>와 같다. 최종 분석에 사용한 응답자는 195명으로 남자 128명(65.6%), 여자 67명(34.4%)이었으며 40대가 74명(37.9%)으로 가장 많았다. 학력으로는 대학교 졸업이 96명(49.2%)으로 가장 많았고 대학원 졸업 이상이 79명(40.5%)으로 그다음 순이었다. 직업으로는 IT(정보보안 포함)가 92명(47.2%)으로 가장 많았다. 업무에 주로 사용하는 목적은 정보검색 91명(46.7%), 아이디어 도출 31명(15.9%), 문서작성 30명(15.4%) 순이었다. 이것은 사용자의 대다수가 생성형 AI를 정보검색이나 아이디어 도출 등을 위한 업무 보조적인 목적으로 활용하고 있다는 것을 보여준다.

<표 2> 인구통계학적 특성

구분		명(%)
성별	남자	128(65.6%)
	여자	67(34.4%)
나이	20~29세	25(12.8%)
	30~39세	52(26.7%)
	40~49세	74(37.9%)
	50~59세	36(18.5%)
	60세 이상~	8(4.1%)
학력	고졸 이하	12(6.2%)
	대학교졸	96(49.2%)
	대학원졸(재학)	79(40.5%)
	전문대졸	8(4.1%)
직업	IT(정보보안 포함)	92(47.2%)
	경영/사무	34(17.4%)
	공무원	10(5.1%)
	교육/연구	10(5.1%)

	금융/보험	8(4.1%)
	기타	13(6.7%)
	보건/의료	5(2.6%)
	서비스/영업	10(5.1%)
	전문직	13(6.7%)
사용 용도	프로그램 코딩	10(5.1%)
	데이터 분석	24(12.3%)
	문서 번역	8(4.1%)
	문서 작성	30(15.4%)
	아이디어 도출	31(15.9%)
	정보 검색	91(46.7%)
	기타	1(0.5%)
총		195(100%)

4.3 측정 도구의 검증

본 연구에서는 SPSS 22.0 프로그램을 사용하여 측정 도구 변수들의 타당성과 신뢰성을 검증하였다. 주성분 분석(PCA, Principle component analysis), 베리맥스(Varimax) 방법을 활용한 탐색적 요인분석을 실시하였으며 신뢰도 분석을 위해 크론바흐 알파(Cronbach's alpha) 계수를 사용하였다. 탐색적 요인 분석은 변수 간의 관계를 탐색적으로 찾아내는 방법으로 요인 적재량이 0.4 이상이면 유의한 변수로, 0.5 이상이면 중요한 변수로 판단할 수 있다. 신뢰성 분석은 측정 요인에 대한 결과 값이 일관성 있는 패턴을 보이는지를 분석하는 방법으로 크론바흐 알파 계수값이 0.6 이상이면 수용 가능, 0.7 이상이면 바람직한 신뢰도, 0.8 이상이면 아주 높은 신뢰도로 판단 할 수 있다[30]. 본 연구에서는 시스템 품질, 정보 품질, 서비스 품질을 구성하는 하위 요인들을 외생변수로, 나머지를 내생변수로 구분하여 탐색적 요인분석을 수행하였으며 타당성 및 신뢰성 검증 결과 값은 각각 <표 3-1>, <표 3-2>와 같다. 분석 결과 각각의 요인에 대한 요인적재량이 모두 0.5 이상으로 측정되었으며, 크론바흐 알파 계수도 모두 0.8 이상으로 측정되어 타당성, 신뢰성이 모두 확보된 것으로 나타났다. 측정변수와 측정데이터 수의 적합성을 보기 위한 Kaiser-Meyer-Olkin(KMO) 표본적합도 값이 모두 0.8 이상으로 적합한 수준으로 나타났으며, Bartlett의 구형성 검증 결과 p=.000으로 요인분석이 통계적으로 유의하여 적합한 것으로 나타났다.

<표 3-1> 측정도구의 탐색적 요인분석 결과 1

변수	요인적재치								Cronbach's α
	1	2	3	4	5	6	7	8	
보안성1_3	**.876**	.080	.059	.168	.204	.135	.026	.151	.947
보안성1_4	**.860**	.055	.046	.146	.194	.170	.066	.119	
보안성1_1	**.827**	.092	.034	.107	.203	.192	.130	.161	
보안성1_2	**.816**	.082	.106	.125	.115	.245	.141	.149	
보안성1_5	**.798**	.086	.080	.128	.224	.167	.027	.135	
접근성1_2	.171	**.840**	.179	.039	.067	.032	.138	.087	.924
접근성1_3	.005	**.830**	.143	.093	-.081	.058	.252	.167	
접근성1_1	.156	**.828**	.155	.023	.082	.123	.237	.062	
접근성1_4	-.004	**.815**	.221	.091	.053	.170	.151	.038	
접근성1_5	.064	**.760**	.241	.128	.032	.113	.205	-.002	
다양성1_2	.123	.167	**.845**	.225	.094	.223	.134	.140	.944
다양성1_3	.054	.204	**.845**	.165	.108	.180	.196	.093	
다양성1_5	.017	.288	**.798**	.228	.127	.204	.047	.137	
다양성1_1	.140	.268	**.789**	.100	.097	.087	.217	.201	
다양성1_4	.017	.287	**.655**	.248	.095	.200	.144	.244	
전문성1_4	.186	.059	.169	**.782**	.219	.166	.162	.194	.913
전문성1_3	.106	.123	.173	**.780**	.224	.198	.043	.056	
전문성1_2	.157	.088	.126	**.774**	.268	.107	.186	.174	
전문성1_1	.164	.089	.230	**.748**	.300	.014	.115	.157	
전문성1_5	.150	.094	.250	**.624**	.145	.168	.262	.144	

신뢰성1_3	.201	.100	.158	.141	**.780**	.216	.092	.180	
신뢰성1_2	.254	-.028	.109	.280	**.764**	.166	.049	.245	
신뢰성1_5	.226	.049	.086	.344	**.740**	.145	.127	.126	.923
신뢰성1_4	.258	.033	.065	.347	**.739**	.205	.081	.199	
신뢰성1_1	.269	.020	.112	.247	**.640**	.074	.207	.276	
안정성1_5	.154	.108	.150	.251	.095	**.821**	.152	.027	
안정성1_3	.208	.099	.171	.078	.184	**.803**	-.008	.140	
안정성1_4	.220	.095	.143	.174	.056	**.784**	.117	.122	.893
안정성1_2	.128	.114	.236	-.001	.222	**.667**	.117	.119	
안전성1_1	.290	.131	.099	.121	.159	**.643**	.132	.238	
편리성1_3	.100	.105	.027	.198	.126	.139	**.766**	.168	
편리성1_1	.138	.302	.213	.075	.088	.072	**.762**	-.010	
편리성1_2	.133	.261	.155	.127	.066	.008	**.746**	.080	.873
편리성1_5	.027	.251	.258	.127	.085	.115	**.722**	.206	
편리성1_4	-.047	.319	.043	.156	.119	.261	**.575**	.272	
완전성1_1	.207	.142	.062	.269	.198	.226	.198	**.679**	
완전성1_4	.189	.128	.280	.147	.285	.163	.209	**.674**	
완전성1_5	.314	.036	.303	.134	.314	.159	.168	**.645**	.906
완전성1_2	.263	.100	.228	.282	.258	.105	.272	**.626**	
완전성1_3	.249	.125	.305	.152	.308	.241	.052	**.609**	
고유값	4.576	4.178	4.170	3.926	3.787	3.724	3.427	2.983	
분산설명(%)	11.441	10.444	10.425	9.814	9.466	9.309	8.567	7.459	
Kaiser-Meyer-Olkin 표본적합도	0.921								
Bartlett의 구형성검정(p)	.000								

<표 3-2> 측정도구의 탐색적 요인분석 결과 2

변수	요인적재치				Cronbach's α
	1	2	3	4	
지속적사용의도1_3	**.805**	.130	.313	.265	
지속적사용의도1_2	**.802**	.222	.265	.261	
지속적사용의도1_4	**.798**	.116	.217	.290	.939
지속적사용의도1_1	**.744**	.272	.266	.322	
지속적사용의도1_5	**.731**	.213	.410	.216	
지각된사용용이성1_5	.128	**.860**	.073	.202	
지각된사용용이성1_2	.197	**.852**	.144	.093	
지각된사용용이성1_4	.163	**.824**	.045	.113	.919
지각된사용용이성1_1	.124	**.808**	.204	.249	
지각된사용용이성1_3	.117	**.785**	.281	.156	
만족도1_1	.240	.136	**.794**	.200	
만족도1_5	.335	.160	**.784**	.202	
만족도1_4	.420	.116	**.701**	.267	.897
만족도1_3	.214	.285	**.689**	.257	
만족도1_2	.422	.248	**.504**	.427	
지각된유용성1_2	.370	.186	.216	**.777**	
지각된유용성1_1	.248	.328	.279	**.744**	.912
지각된유용성1_3	.326	.298	.274	**.704**	
지각된유용성1_5	.489	.190	.310	**.592**	

지각된유용성1_4	.414	.117	.456	**.543**	
고유값	4.412	4.094	3.571	3.221	
분산설명(%)	22.062	20.472	17.856	16.105	
Kaiser-Meyer-Olkin 표본적합도	0.94				
Bartlett의 구형성검정(p)	.000				

4.3 가설검증

본 연구는 LLM 기반의 생성형 AI 응답 데이터 품질에 대해 사용자가 만족도를 느끼고 업무에 지속적으로 사용하게 되는 과정을 규명하는데 초점을 두었다. 이를 위해 연구모형에서 제시된 각 변수 간의 영향관계인 가설 검정을 위한 다중회귀분석(Multiple regression analysis)을 실시하였다. 잠재변수에 대한 결정계수(R2) 값이 지각된 사용용이성과 독립변수 간의 관계를 제외하고는 모두 0.5 이상으로 중간 정도 이상의 설명력을 보이고 있으며, Durbin-Waston 지수가 모두 2에 가까운 값을 가지므로 본 데이터는 회귀분석에 적합한 것으로 나타났다. 먼저 독립변수와 지각된 사용용이성과의 관련성을 분석한 결과, <표 4>과 같이 접근성(β=.443), 다양성(β=.205), 전문성(β=.184), 안정성(β=.177), 편리성(β=.396)은 모두 지각된 사용용이성에 정(+)의 영향을 미치지만, 보안성(β=.051, p>.05), 신뢰성(β=.093, p>.05), 완전성(β=.053, p>.05)은 지각된 사용용이성에 유의한 영향을 미치지 않는 것으로 나타나 가설 H1-4, H2-4, H2-6은 기각되고 H1-5, H1-6, H2-5, H3-3, H3-4는 채택되었다.

<표 4> 지각된 사용용이성과 독립변수간의 관계 검증 결과

독립변수	비표준화계수		표준화계수	t값	p값	결과
	B	표준오차	β			
보안성	.216	.226	.051	.953	.342	기각
접근성	1.884	.226	.443	8.329	.000***	채택
다양성	.872	.226	.205	3.855	.000***	채택
전문성	.781	.226	.184	3.454	.001**	채택
신뢰성	.398	.226	.093	1.759	.080	기각
안정성	.755	.226	.177	3.337	.001**	채택
편리성	1.686	.226	.396	7.452	.000***	채택
완전성	.226	.226	.053	.999	.319	기각
R=.689, R2=.474, F=20.980, P=.000, Durbin-Watson=1.894						

**p<.01, ***p<.001

독립변수와 지각된 유용성과의 관련성을 분석한 결과, <표 5>과 같이 보안성(β=.173), 접근성(β=.262), 다양성(β=.193), 전문성(β=.391), 신뢰성(β=.278), 편리성(β=.387)은 모두 지각된 유용성에 정(+)의 영향을 미치지만, 안정성(β=.068, p>.05), 완전성(β=.085, p>.05)은 지각된 유용성에 유의한 영향을 미치지 않는 것으로 나타나 가설 H1-3, H2-1은 기각되고 H1-1, H1-2, H2-2, H2-3, H3-1, H3-2는 채택되었다.

<표 5> 지각된 유용성과 독립변수간의 관계 검증 결과

독립변수	비표준화계수		표준화계수	t값	p값	결과
	B	표준오차	β			
보안성	.725	.212	.173	3.422	.001**	채택
접근성	1.100	.212	.262	5.191	.000***	채택
다양성	.810	.212	.193	3.819	.000***	채택
전문성	1.645	.212	.391	7.761	.000***	채택
신뢰성	1.171	.212	.278	5.523	.000***	채택
안정성	.284	.212	.068	1.341	.182	기각
편리성	1.627	.212	.387	7.674	.000***	채택
완전성	.358	.212	.085	1.688	.093	기각
R=.726, R2=.527, F=25.941, P=.000, Durbin-Watson=2.004						

**p<.01, ***p<.001

매개변수와 종속변수와의 관련성을 분석한 결과, <표 6>, <표 7>과 같이 지각된 유용성(β=.710), 지각된 사용용이성(β=.289)은 모두 만족도에 정(+)의 영향을 미치고 지각된 유용성(β=.633), 지각된 사용용이성(β=.246), 만족도(β=.246)는 모두 지속적 사용의도에 정(+)의 영향을 미치는 것으로 나타나 가설 H4-1, H4-2, H5-1, H5-2, H5-3은 모두 채택되었다.

<표 6> 만족도와 독립변수간의 관계 검증 결과

독립변수	비표준화계수		표준화계수	t값	p값	결과
	B	표준오차	β			
지각된유용성	2.972	.194	.710	15.297	.000***	채택
지각된사용용이성	1.210	.194	.289	6.225	.000***	채택
R=.766, R2=.587, F=136.376, P=.000, Durbin-Watson=1.758						

***p<.001

<표 7> 지속적 사용의도와 독립변수간의 관계 검증 결과

독립변수	비표준화계수		표준화계수	t값	p값	결과
	B	표준오차	β			
지각된유용성	3.110	.228	.633	13.667	.000***	채택
지각된사용용이성	1.207	.209	.246	5.787	.000***	채택
만족도	1.208	.228	.246	5.295	.000***	채택
R=.811, R2=.658, F=122.270, P=.000, Durbin-Watson=1.693						

***p<.001

5. 결론 및 향후과제

5.1 결론

본 연구는 국내외적으로 실증적 연구가 부족한 LLM 기반의 생성형 AI 응답 데이터 품질이 업무 활용 만족도에 미치는 영향에 관한 연구를 수행하였고, 다음과 같은 결론을 도출하였다.

첫째, 지각된 사용용이성에 편리성, 안정성, 다양성, 전문성, 접근성은 정(+)의 영향을 미치는 것을 확인하였다. 반면, 보안성, 완전성, 신뢰성은 부정(-)의 영향을 미치는 것을 확인하였다. 이것은 지각된 사용용이성에 유의한 영향을 미치기 위해서는 사용자들로 하여금 응답 데이터가 사용하기 용이하다는 지각을 증대시키고, 보안성, 신뢰성, 완전성 등의 요인을 개선해야 한다는 것을 보여준다. 마찬가지로 업무에 활용하고, 성과를 높임에 있어 당장 필요한 정보를 검색하고, 아이디어를 얻는데

보조적인 수단으로 활용하기 때문에 보안성, 신뢰성, 완전성이 지각된 사용용이성을 인정할 정도에는 미치지 못한다는 뜻으로 해석된다.

둘째, 지각된 유용성에 보안성, 편리성, 다양성, 신뢰성, 전문성, 접근성은 정(+)의 영향을 미치는 것을 확인하였다. 반면, 안정성, 완전성은 부정(-)의 영향을 미치는 것을 확인하였는데, 이것은 지각된 유용성에 유의한 영향을 미치기 위해서는 안정성, 완전성을 혁신해야 한다는 것을 보여준다. 사용자들은 응답 데이터가 전문성이 있다고 인식하고 있으며 접근성과 다양성을 확보하고 있기 때문에 업무 성과를 내는데 유용하게 활용하고 있다고 보여 진다. 하지만 업무 활용에 유용성과 만족도를 증대시키기 위해서는 요청한 내용에 대해 정보가 누락 없이 일정한 속도로 제공되는 것이 필요하다고 해석된다.

셋째, 지각된 사용용이성과 지각된 유용성은 만족도에 정(+)의 영향을 주었고, 지각된 사용용이성, 지각된 유용성, 만족도 모두 지속적 사용의도에 정(+)의 영향을 주었다. 이것은 응답 데이터가 전문적이고 접근하기 편리해서 별다른 노력 없이도 필요한 정보를 쉽게 얻을 수 있으며, 응답 데이터 활용에 만족하기 때문에 업무 성과를 내는데 지속적으로 사용되고 있음을 보여준다.

5.2 시사점

본 연구는 LLM 기반의 생성형 AI 응답 데이터를 실제로 업무에 활용해본 직장인을 대상으로 설문조사를 수행하였으며 195개의 유효 표본으로 실증 연구를 수행하였다. 기존 생성형 AI와 관련된 연구가 대부분 ChatGPT 같은 서비스의 수용의도나 사용의도, 기술 자체에 관한 연구에 집중된 반면 본 연구는 응답 데이터의 품질을 중심으로 연구를 수행하였다. 이것은 응답 데이터를 실제 업무에 활용해본 직장인을 대상으로 만족도 및 지속적 사용의도에 영향을 미치는 요인에 대한 실증적 연구를 수행하였다는 점에서 의의가 있다. 또한 최근 쟁점이 되는 보안성, 신뢰성 이슈에도 불구하고 생성형 AI를 업무에 활용하려는 사용자가 증가하고 있는 요인에 대해서 학문적 측면에서 설명할 수 있는 계기를 제공하였다. 서비스 제공자 관점에서는 사용자 만족도 제고를 위한 방향성을 제시해 주었으며 나아가 잠재적인 수요에 대한 예측은 물론 생성형 AI 응답 데이터 품질 향상에도 기여했다는 점에서도 가치가 크다.

5.2 향후과제

연구 결과를 토대로 발견하게 된 한계점을 바탕으로 앞으로의 연구 방향을 제시하면 다음과 같다.

첫째, 본 연구에서는 사용자의 지속적 사용의도를 분석하기 위해 전통적인 TAM 및 IS Success 모델을 적용하였지만 최근 인공지능 기술수용과 관련된 모델을 활용하여 보다 설명력을 높일 필요가 있다.

둘째, 본 연구의 결과로 응답 데이터 품질에 영향을 미치는 요인을 통해 지속적 사용의도에 대한 영향력을 확인하였으니, 향후에는 본 연구에서 부분적으로 기각된 보안성, 신뢰성을 중심으로 보다 심도 있는 연구를 수행해 볼 필요가 있다.

셋째, 인과 관계를 확인하기 위하여 다중 회귀분석을 사용하였으나, 향후 연구에서는 구조방정식과 같은 다양한 연구 방법을 적용할 필요가 있다.

참고문헌

[1] Shanahan,M., "Talking about large language models", 2022.

[2] Gozalo-Brizuela,R., E. C. Garrido-Merchán, "A survey of generative AI applications", 2023.

[3] J. Yao, K. Ning, Z. Liu, M. Ning, and L. Yuan, "Llm lies: Hallucinations are not bugs, but features as adversarial examples", arXiv preprint arXiv:2310.01469, 2023.

[4] Yao,Y., J. Duan, K. Xu, Y. Cai, Z. Sun, and Y. Zhang, "A survey on large language model (LLM) security and privacy: The good, the bad, and the ugly", 2023.

[5] Gozalo-Brizuela,R., E. Garrido-Merchan, "C

hatGPT is not all you need. A state of the art review of large generative AI models", 2023.

[6] "Generative AI: Perspectives from stanford HAI", Standford University, 2023.

[7] Zhao,W. X., K. Zhou, J. Li, T. Tang, X. Wang, Y. Hou, Y. Min, B. Zhang, J. Zhang, Z. Dong, Y. Du, C. Yang, Y. Chen, Z. Chen, J. Jiang, R. Ren, Y. Li, X. Tang, Z. Liu, P. Liu, J. Nie, and J. Wen, "A survey of large language models", 2023.

[8] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate", arXiv preprint arXiv:1409.0473, 2014.

[9] Vaswani,A., G. Brain, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, and Ł Kaiser, "Attention is all you need", 2017.

[10] S. Feuerriegel, J. Hartmann, C. Janiesch, and P. Zschech, "Generative ai", Business & Information Systems Engineering, vol.66, no.1, pp.111-126, 2024.

[11] "Generative AI could raise global GDP by 7%", Goldman Sachs, 2024.05.15. https://www.goldmansachs.com/intelligence/pages/generative-ai-could-raise-global-gdp-by-7-percent.html

[12] A. Choudhury, H. Shamszare, "Investigating the impact of user trust on the adoption and use of ChatGPT: Survey analysis", Journal of Medical Internet Research, vol.25, pp.e47184, 2023.

[13] 이동영. (2024). "ChatGPT 지속사용의도에 관한 연구", 「국내박사학위논문」, 동의대학교 대학원.

[14] 신재형. (2024). "인공지능(AI) 학습데이터의 수집 및 이용에 대한 규제", 「국내박사학위논문」, 서울대학교 대학원.

[15] F. D. Davis, R. P. Bagozzi, and P. R. Warshaw, "User acceptance of computer technology: A comparison of two theoretical models", Management Science, vol.35, no.8, pp.982-1003, 1989.

[16] A. Bhattacherjee, "Understanding information systems continuance: An expectation-confirmation model", MIS Quarterly, vol.25, no.3, pp.351-370, 2001.

[17] W. H. DeLone, E. R. McLean, "Information systems success: The quest for the dependent variable", Information Systems Research, vol.3, no.1, pp.60-95, 1992.

[18] L. F. Pitt, R. T. Watson, and C. B. Kavan, "Service quality: A measure of information systems effectiveness", MIS Quarterly, vol.19, no.2, pp.173-187, 1995.

[19] 박인근. (2016). "기업의 영업지원 소프트웨어 사용 성과에 영향을 미치는 요인에 관한 실증적 연구 : 정보시스템 성공 모델 중심으로", 「국내박사학위논문」, 숭실대학교 대학원.

[20] 유은정. (2020). "공유숙박 플랫폼의 품질특성이 한·중 이용자의 이용의도에 미치는 영향", 「국내박사학위논문」, 숭실대학교 대학원.

[21] 문은경. (2022). "대학 원격수업 유형에 따른 지속사용의도에 대한 영향 요인 비교 연구", 「국내박사학위논문」, 고려대학교 대학원.

[22] Y. H. Al-Mamary, A. Shamsuddin, and N. Aziati, "The relationship between system quality, information quality, and organizational performance", International Journal of Knowledge and Research in Management & E-Commerce, vol.4, no.3, pp.7-10, 2014.

[23] 박성준. (2012). "모바일 여행콘텐츠 정보품질이 지각된 유용성과 행동의도에 미치는 영향", 「국내박사학위논문」, 安養大學校.

[24] 이승육. (2016). "농.수산물 이력추적 정보품질이 지속사용의도에 미치는 영향", 「국내박사학위논문」, 서울벤처대학원대학교.

[25] 김광철. (2023). "플랫폼 품질속성이 고객만족과 재이용의도에 미치는 영향", 「국내박사학위논문」, 호서대학교 벤처대학원.

[26] P. Asubonteng, K. J. McCleary, and J. E. Swan, "SERVQUAL revisited: A critical review of service quality", Journal of Services Marketing, vol.10, no.6, pp.62-81, 1996.
[27] 임나윤. (2023). "기술수용모델을 적용한 전시관 가이드 품질이 지속이용의도에 미치는 영향", 「국내박사학위논문」, 공주대학교 대학원.
[28] 최주원. (2021). "인공지능기반 스마트양식시스템의 수용의도에 영향을 미치는 요인에 관한 연구", 「국내박사학위논문」, 숭실대학교 대학원.
[29] 박명준. (2022). "대학의 비대면 실시간 및 비실시간 교육서비스품질과 지속사용의도에 관한 연구", 「국내박사학위논문」, 서강대학교 경영전문대학원.
[30] J. F. Hair, "A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)", Sage, 2017.

〔저자소개〕

이 승 환 (Seung-hwan Lee)
2022년 건국대학교 공학석사
2023년 숭실대학교 IT정책경영학과 박사과정
email : whaani@hanmail.net

현 지 은 (Ji-eun Hyun)
2022년 건국대학교 공학석사
2023년 숭실대학교 IT정책경영학과 박사과정
email : j.morethanwords@gmail.com

김 광 용 (Gwang-yong Gim)
1984년 고려대학교 학사
1991년 조지아 주립대학 석사
1995년 조지아 주립대학 박사
1996년 숭실대학교 경영학부 교수
email : gygim@ssu.ac.kr