

ChatGPT의 교육적 활용 관련 연구동향 및 주제 분석

장혜지* · 소효정**†

* 이화여자대학교 교육공학과 통합과정 (E-mail: hyejijang@ewhain.net)

** 이화여자대학교 교육공학과 교수 (†교신저자, E-mail: hyojeongso@ewha.ac.kr)

(논문 접수 2023. 07. 25.; 수정 후 제출 2023. 08. 25.; 게재 승인 2023. 08. 28.)

The Analysis of Research Trends and Topics about the Educational Use of ChatGPT

Jang, Hyeji* · So, Hyo-Jeong**†

* Integrated Doctoral Student, Dept. of Educational Technology, Ewha Womans University (E-mail: hyejijang@ewhain.net)

** Professor, Dept. of Educational Technology, Ewha Womans University

(†Corresponding author, E-mail: hyojeongso@ewha.ac.kr)

(Received July 25, 2023; Received in revised form August 25, 2023; Accepted August 28, 2023)

요 약. 본 연구는 생성형 인공지능 ChatGPT의 교육적 활용 관련 연구동향 및 주제를 파악하고 시사점을 도출하기 위한 목적으로 실시되었다. 2022년 12월부터 2023년 5월까지 발표된 관련 문헌 72편(국내: 14, 국외: 58)을 수집하여 체계적 문헌 분석과 텍스트 마이닝 기법을 활용한 토픽 분석을 실시하였다. 체계적 문헌분석 결과, 문헌의 수가 3월에 급격히 증가하였으며 개념연구와 실증연구가 비슷한 비중으로 많이 나타났다. 연구 대상 학교급은 유아, 초등, 중등보다 고등교육 맥락에서 연구가 많이 이루어졌으며, 연구된 교육분야는 의학·간호와 언어교육에서 많이 수행된 것으로 나타났다. LDA(Latent Dirichlet Allocation) 기반 토픽 모델링을 분석한 결과 연구동향의 주제는 (1) 의학·간호교육에서 ChatGPT로 인한 영향과 대비, (2) 교과 내 ChatGPT 실제적 적용의 영향 탐색 및 프레임워크 제안, (3) 교과실제 및 연구에서 ChatGPT 활용으로 인한 미래 고찰, (4) 평가에서의 ChatGPT 성능 분석, (5) 교과 학습 활동 및 과제에서 ChatGPT 활용 방안 탐색, (6) 고등교육에서 ChatGPT 활용 시 고려사항, (7) 교과에서 전문가, 교사, 학생의 ChatGPT 활용 경험 기반 시사점, (8) 교과에서 ChatGPT 활용에 대한 가능성과 도전의 총 여덟 개로 나타났다. 도출된 결과를 토대로 ChatGPT의 교과교육적 활용 및 향후 관련 연구에 대한 시사점을 제시하였다.

주제어: 챗GPT, 인공지능, 교육적 활용, 연구동향, 키워드 빈도 분석, LDA 토픽 모델링

ABSTRACT. The aim of this study was to identify research trends and topics related to the educational use of ChatGPT, a generative artificial intelligence, and to draw implications. The study analyzed 72 related literatures (domestic: 14, foreign: 58) published from December 2022 to May 2023, and topics were analyzed through the systematic literature review and text mining techniques. Regarding the result of the systematic literature review, the number of related literatures increased rapidly in March, and conceptual studies and empirical studies appeared in a similar proportion. Regarding the LDA-based topic modeling analysis, the topics of research trends are (1) impact and preparation of ChatGPT in medical and nursing education, (2) exploration of the impact of ChatGPT in practices and proposing frameworks, (3) future consideration about the use of ChatGPT in practices and research, (4) evaluating ChatGPT performance, (5) exploring ways to use ChatGPT in subject-specific learning activities and tasks, (6) considerations for using ChatGPT in higher education, (7) implications drawn from the experience of using ChatGPT by experts, teachers, and students, and (8) possibilities and challenges of using ChatGPT in education. Based on the results, we present ways to use ChatGPT for education and implications for future related studies.

Key words: ChatGPT, Artificial Intelligence, Educational Use, Research Trends, Term Frequency Analysis, LDA Topic Modeling

I. 서 론

인간의 지능과 같이 사고하고 행동하는 기능을 가진 컴퓨터 시스템인 인공지능(Artificial Intelligence, AI)은 현재 인간 삶의 많은 영역에 영향을 주고 있는 신기술이다. 인

공지능이라는 용어는 John McCarthy가 1956년 다트머스 회의에서 ‘기계를 인간 행동의 지식에서와 같이 행동하게 만드는 것’으로 정의하며 탄생하였다. 교육 영역에서 AI의 활용은 AIED(Artificial Intelligence in Education)이라고 불리며, Holmes, Bialik, & Fadel(2019)이 AIED의 개념을

크게 ‘Learning about AI’와 ‘Learning with AI’의 두 가지로 나누어 제시하였다. 첫째, Learning about AI는 AI 기술 자체에 대한 학습으로 대표적으로 코딩과 같은 소프트웨어 교육 및 AI 리터러시 교육을 의미한다. 둘째, ‘Learning with AI’는 AI를 활용한 교육으로 다양한 교과영역의 교수 및 학습활동에 AI 기반 플랫폼 및 도구를 활용하는 것이다. 최근 AI 기술을 탑재한 교육용 애플리케이션이 증가하고 있고 다양한 교과영역에 적용이 가능하다. Lee, So, & Jin(2020)은 교육용 AI 기반 애플리케이션 분석 연구를 통해 어학 영역의 앱에서는 주로 음성인식 기술이, 수리 영역에서는 학습 데이터 분석과 적응형 학습 시스템이 적용되고 있음을 제시하였다.

최근 AI의 교육적 활용은 생성형 AI인 ChatGPT(Generative Pre-trained Transformer)의 등장으로 새로운 국면을 맞이하게 되었다. 2022년 11월 30일 OpenAI에서 발표한 ChatGPT는 전 세계가 생성형 AI의 가능성에 주목하도록 하였으며, ChatGPT는 출시된 지 두 달 만에 사용자가 약 1억 명을 돌파하는 등 과거의 신기술과 비교되는 급속한 파급력을 보였다. 대화기반 생성형 AI인 ChatGPT는 에세이 작성, 자료 수집 및 요약 등 다양한 방법으로 교수·학습 활동에 활용이 가능하다(van Dis et al., 2023). 하지만 동시에 ChatGPT를 활용하여 작성한 과제를 대학 강사가 0점 처리하여 논란이 된 경우와 같이(Choi, 2023) 이 신기술의 교육적 활용을 어디까지 허용할 것인가에 대한 합의가 현재는 부족한 상태이다.

최근 ChatGPT의 교육적 활용에 대한 학술적 논의가 시작되고 있으나 아직까지 연구결과는 단편적이며 초기 단계에 머물고 있다. 더불어 ChatGPT를 실제로 교육현장에 활용하여 시사점을 도출한 연구는 양적으로 충분하지 않은 상태이다. 따라서 본 연구는 현재 ChatGPT와 같은 생성형 AI의 교육적 활용에 대한 논의가 다각도로 펼쳐지고 있는 가운데 학술적으로 어떠한 주제와 관련된 연구가 실시되었으며 교육적 활용과 관련된 시사점은 무엇인지 문헌분석을 통해 살펴보고자 하였다. 이를 위해 ChatGPT가 출시된 시점인 2022년 12월부터 2023년 5월까지 약 6개월간 ChatGPT의 교육적 활용과 관련하여 국내·외에서 출판된 문헌을 대상으로 전반적인 연구 동향을 살펴본 후 텍스트 마이닝 기법인 키워드 빈도 분석과 LDA 기반 토픽 모델링을 실시하였다. 토픽 기반 분석을 통해 문헌에서 가장 빈번하게 도출된 키워드의 특징을 살펴보고 논의되고 있는 주요 토픽이 무엇인지 전반적으로 파악하고자 하였다. 본 연구에서 설정한 연구 문제는 다음과 같다.

1. ChatGPT 교육적 활용 관련 국내외 문헌의 전반적인 연구동향은 어떠한가?

2. ChatGPT 교육적 활용 관련 국내외 문헌에서 주요하게 언급된 키워드는 무엇인가?
3. ChatGPT 교육적 활용 관련 국내외 문헌에서 도출된 주요 연구주제는 무엇인가?
4. ChatGPT 교육적 활용 관련 국내외 문헌분석을 기반으로 도출된 시사점은 무엇인가?

II. 이론적 배경

1. 생성형 AI와 교육

생성형 AI란 통계, 확률 등을 통해 인공물을 생성하는 비지도 학습 또는 부분 지도 학습 기반의 기계 학습 프레임워크를 의미한다(Baidoo-Anu & Ansah, 2023). 생성형 AI는 딥러닝 모델에 영향을 주어 복잡하고 다양한 프롬프트(prompt)에 대응하여 이미지, 텍스트 등의 인공적 콘텐츠를 생성하는 기술로도 정의된다(Lim et al., 2023). 즉, 생성형 AI는 대규모 데이터를 통해 학습하며 창의적인 산출물을 생성하는 기술로 볼 수 있다.

생성형 AI의 대표적인 모델로는 생성적 대립 신경망(GAN: Generative Adversarial Network)과 사전 학습된 트랜스포머(GPT: Generative Pre-trained Transformer)가 있다(Baidoo-Anu & Ansah, 2023). GAN이란 주로 사용되는 생성형 AI 기법으로 생성자(Generator)와 판별자(Discriminator)라는 두 종류의 신경망으로 구성된다(Goodfellow et al., 2020). 생성자 신경망은 인공적인 데이터를 생성하며 판별자 신경망은 콘텐츠 진위 여부를 판별하는 역할을 한다. 이 두 종류의 신경망은 서로 적대적인 관계를 통한 훈련으로 새로운 콘텐츠를 생성하는 것이 특징이다. 다음으로 GPT란 OpenAI 기업이 개발한 모델로 트랜스포머와 비지도 학습을 결합한 기술이다. GPT는 주어진 텍스트 다음에 오는 단어를 예측하는 훈련을 하는 자연어 생성 모델로 인간과 유사하게 텍스트를 생성하는 기능을 가진다(Jovanović & Campbell, 2022).

생성형 AI는 기존에도 지속적으로 개발되고 있었지만 지난 2022년 11월 말에 출시된 ChatGPT의 등장으로 대중들이 쉽게 접하고 활용 가능한 기술이 되었다. 다양한 영역에서 ChatGPT 활용에 대한 영향과 관련된 논의가 지속적으로 이루어지고 있으며 교육 분야도 예외는 아니다. 그 예로서, Farrokhnia et al.(2023)은 SWOT 분석을 통해 ChatGPT의 강점(strengths)과 약점(weaknesses), 교육에 대한 기회(opportunities)와 위협(threats)을 탐색하였다. SWOT 분석 결과에 따르면 ChatGPT의 교육적 적용에 대한 강점은 정교한 자연어 모델을 사용하여 타당한 답변을 생성하고, 스스로 개선하는 능력, 맞춤형 및 실시간 응답을 제공하는 것으로 나타났다. 반면, ChatGPT의 약점은 깊은

이해 부족, 응답 품질 평가의 어려움, 편견 및 차별 위험, 그리고 학생들의 고차원적 사고 능력 저해 등으로 나타났다. 다음으로 기회로는 정보의 접근성을 높이고, 개별화 학습과 복잡한 학습을 촉진하며, 교수 부담을 감소시켜주는 것이 제시되었다. Kasneci et al.(2023)과 Rudolph, Tan, & Tan(2023)도 역시 ChatGPT가 교수자의 작업 부담을 감소시켜 중요한 프로세스 및 업무를 더 효율적으로 처리하게 도와줄 수 있음을 제시하였다. 마지막으로 교육에 대한 위협으로는 맥락에 대한 이해 부족, 학업 진실성 위협, 교육에서의 차별 유지, 표절 미화, 고차원적인 인지 능력 감소 등이 나타났다. 이러한 논의들은 현재 ChatGPT의 교육적 활용 가능성에 대해 잠재력과 제한점을 충분히 고려하여 적용해야 하며, 강점 또는 약점에 치우치기 보다는 균형된 관점을 취할 필요가 있음을 알 수 있다.

2. 관련 선행연구

현재는 ChatGPT 도입의 초기 단계로 볼 수 있기에 이를 교육 맥락에서 실제적으로 활용한 연구는 아직 충분하지 않으며, 대부분의 학술적 담화는 활용방안에 대한 개념적 논의 중심으로 이루어지고 있다. 소수이지만 선행연구들을 바탕으로 현재까지 논의된 교육 분야에서 ChatGPT 연구의 주제를 살펴보면 크게 (1) 각 교과에서 ChatGPT 성능평가, (2) 교수자 수업 준비 및 평가, (3) 학습자 학습 및 평가활동 지원으로 나뉘볼 수 있다.

첫째, 다양한 연구들에서 시험 또는 인터뷰를 통해 ChatGPT의 성능을 평가하였으며 교과에 따라 ChatGPT의 성능이 상이한 것으로 나타났다. 예를 들어, 거시경제학 분야(Geerling et al., 2023)와 프로그래밍(Surameery & Shakor, 2023)에서는 ChatGPT의 성능이 만족할만한 수준이었으나, 수학에서는 문제를 이해하더라도 정답의 오류가 빈번하여 성능이 가장 만족스럽지 않은 것으로 나타났다(Frieder et al., 2023; Wardat et al., 2023).

둘째, 교수·학습 상황에서 ChatGPT의 활용 방안에 대한 논의를 살펴보면 교수자를 위해서는 주로 수업 준비와 평가에서 ChatGPT의 활용 가능성이 제시되었다. 수업 준비로는 교육과정 및 강의계획안 설계(AI-Worafi et al., 2023; Kasneci et al., 2023), 외국어 학습을 위한 대화 주제 생성(Topsakal & Topsakal, 2022) 등이 제시되었다. 평가 측면으로는 ChatGPT가 수업 활동, 퀴즈, 시험을 준비할 때 보조도구로 사용될 수 있으며(AI-Worafi et al., 2023; Cooper, 2023; Khan et al., 2023), 학습자들에게 피드백을 즉각적이거나 개별적으로 제공할 수 있음이 제안되었다(Kasneci et al., 2023; Mollick & Mollick, 2023).

셋째, 학습자를 위한 활용 측면에서는 ChatGPT가 가상 튜터로서 크게 학습과 평가 활동을 지원할 수 있다. 먼저,

학습 측면에서는 학습자들의 협력학습을 위해 ChatGPT가 각기 다른 시나리오를 제공하고(Rudolph et al., 2023), 토론, 실시간 피드백, 개별화된 안내를 제공할 수 있는 것으로 나타났다(Kasneci et al., 2023). 또한, ChatGPT를 통해 다양한 설명 또는 예시를 제공하여 복잡한 교육적 내용에 대한 학습자들의 이해도를 높일 수 있다(Mollick & Mollick, 2023). 다음으로 평가 측면에서는 ChatGPT가 스캐폴딩 역할을 하여 학습자가 답변 초안을 수정할 때 도움을 줄 수 있으며, ChatGPT의 초기 답변을 토대로 학습자들이 질문하고 자신의 지식과 추론 능력을 적용하는 방안도 고려할 수 있다(Kasneci et al., 2023).

이와 같이 ChatGPT의 다양한 교육적 활용 가능성이 제시되고 있음에도 불구하고 동시에 여러 가지 우려점 및 제한점에 대한 논의도 활발하다. ChatGPT가 사실이 아니거나 거짓인 정보를 생성하고 표절 탐지기를 우회하는 문제점(C. Lo, 2023; van Dis et al., 2023), 학습 콘텐츠 제공의 부정확성, 평가의 적절성 여부, 감정의 부재 등에 대한 문제가 제기되었다(Tlili et al., 2023). 이와 같이 ChatGPT가 등장한지 약 8개월밖에 되지 않았음에도 불구하고 ChatGPT를 교육적으로 활용하는 방안에 대해 다양한 논의가 이루어지고 있는 것을 확인할 수 있다. 하지만, 실제 교육 현장에서 학생과 교사를 대상으로 ChatGPT를 적용한 연구는 충분히 이루어지고 있지 않는 이 시점에서 ChatGPT의 교육적 활용과 관련하여 전반적으로 어떠한 논의를 펼치고 있는지 다각도로 살펴보고 시사점을 도출할 필요가 있다.

III. 연구 방법

1. 분석대상 및 자료수집

생성형 AI 기술이 어떻게 교육적으로 활용 가능한지 탐색하기 위해 본 연구는 체계적 문헌분석을 실시한 후 텍스트 마이닝 기법인 키워드 빈도(TF: Term Frequency) 분석과 LDA(Latent Dirichlet Allocation) 기반 토픽 모델링을 실시하여 전반적 연구 주제를 파악하고자 하였다. 첫 번째 단계로 체계적 문헌분석을 통해 ChatGPT가 등장한 후 초기 기간의 전반적인 연구 동향을 파악하고자 하였다. 다음으로는 TF 분석을 통해 ChatGPT의 교육적 활용과 관련된 연구에서 빈번하게 출현되는 키워드가 무엇인지 파악하고자 하였으며, 이후 토픽 모델링 기반 주제별 분석을 실시하여 구체적인 연구동향을 분석하였다.

본 연구의 자료 수집을 위해 국내 문헌은 KCI와 RISS 데이터베이스를, 국외 문헌은 Web of Science와 Scopus를 선정하였다. 선정된 학술 데이터베이스에서 지난 2022년 12월부터 2023년 5월까지 6개월간 발간된 국내외 학술

문헌을 대상으로 본 연구에서 선정한 키워드를 통해 1차 검색을 실시하였다. 문헌 검색 시 포함 및 제외 조건으로 고려한 사항은 다음과 같다. 첫째, 생성형 AI인 ChatGPT가 2022년 11월 말에 처음 등장하였다는 점을 고려하여 2022년 12월부터 약 6개월로 한정하여 문헌을 검색하였다. 둘째, 포함조건으로 학술지와 학술대회에서 발표된 논문을 대상으로 검색하였다. 더불어 2023년 들어 많은 연구가 발표되고 있기 때문에 국외논문의 경우 ‘얼리 액세스’(early access), 즉 최종 출판 전 온라인으로 먼저 발간된 논문도 포함하여 검색을 실시하였다. 셋째, 국외 문헌 자료 수집을 위한 검색식으로는 ‘ChatGPT’를 메인 키워드로, 결합 키워드는 ‘education OR learning’으로 설정하였다. 국내 문헌 자료 수집을 위한 검색식으로는 ‘ChatGPT OR 챗GPT’를 메인 키워드로, 결합 키워드는 ‘교육 OR 학습’으로 설정하였다.

설정한 검색 키워드를 통해 검색된 논문들을 PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 절차에 따라 선별하여 Figure 1과 같이 최종 분석 대상 문헌을 결정하였다(Page et al., 2021). 문헌 선정 단계에서 본 연구진이 여러 차례 검토하여 최종 문헌을 선정하였다. 첫째, 검색(Identification) 단계에서 총 219편이 검색되었고, 이 중 중복된 문헌인 69편을 제외하여 총 150편이 선정되었다. 둘째, 선별(Screening) 단계에서는 검색 단계에서 선정된 문헌의 제목과 초록을 검토한 후 연구주제에 해당되지 않거나 관련성이 낮은 문헌

들 78편을 제외하여 총 72편이 남게 되었다. 마지막으로 포함(Included) 단계에서 국내 14편, 국외 58편으로 총 72편의 논문이 최종 분석 대상에 포함되었다¹⁾.

2. 분석방법

가. 체계적 문헌분석

ChatGPT의 교육적 활용에 대한 초기 연구들이 전반적으로 어떻게 이루어졌는지 최종 선정된 72편의 국내와 국외 논문을 대상으로 체계적 문헌분석을 실시하여 파악하고자 하였다. 본 연구에서 사용한 분석범주와 분석기준은 Table 1과 같다.

자세한 분석범주와 분석기준을 살펴보면 첫 번째로 문헌발표 시기로 게재년월은 최종 선정된 문헌들에 기재되어있는 게재된 연도와 월로 분류하였다. ChatGPT의 교육적 활용 연구는 ChatGPT가 2022년 11월 말에 공개된 것을 고려할 때에 대부분의 연구가 2023년에 이루어지고 게재된 것을 확인할 수 있었다. 따라서 연도만 파악하기 보다는 게재된 월까지 조사하여 구체적인 게재시기를 파악하고자 하였다.

두 번째, 연구의 유형은 실증적 데이터 유무 및 성격에 따라 (1) 개념연구, (2) 인간 대상 실증 연구, (3) ChatGPT 대상 실증 연구, (4) 인간과 ChatGPT 대상 실증 연구, (5) 문헌연구로 분류하여 분석하였다. 먼저, 개념 연구는 ChatGPT 활용 경험을 기반으로 교육적 활용에 대한 논의를 하거나 프레임워크를 제안하는 등 실증적 데이터가 없이 논의가 이루어진 연구를 의미한다. 다음으로 문헌 연구는 체계적 문헌고찰과 같이 다양한 연구들을 수집하여 종합적으로 논의한 연구를 의미한다. 데이터를 포함하고 있는 실증적 연구는 데이터의 성격(인간 대상 vs. 기계 대상)에 따라 분류하였다. 인간 대상 실증 연구란 사용자(교사 또는 학습자)가 직접 ChatGPT를 교육적으로 활용한 후 연구자가 양적 또는 질적 데이터를 수집하거나, ChatGPT의 경험이 있는 사용자를 대상으로 양적 또는 질적 데이터를 수집하여 분석한 연구를 뜻한다. 반면, ChatGPT 대상 실증 연구란 대표적으로 특정 교과에 대한 ChatGPT의 성능을 평가하는 연구와 같이, 인간 사용자 대상 데이터 없이 ChatGPT만의 성능을 실증적 데이터를 기반으로 논의한 연구를 의미한다. 마지막으로 인간과 ChatGPT 대상 실증 연구는 앞서 설명한 인간 사용자 및 ChatGPT 대상 데이터를 모두 포함하는 연구를 의미한다.

세 번째, 대상 학교급은 수준에 따라 특정되지 않음, 유아, 초등, 중등, 고등교육으로 분류하였다. 네 번째, 연구

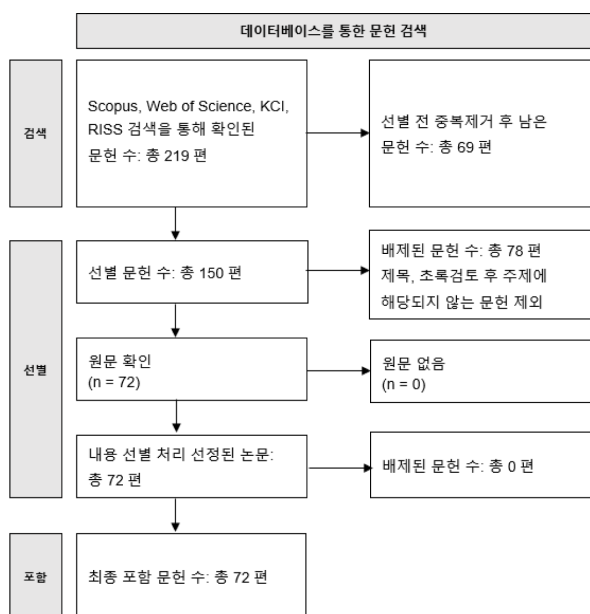


Figure 1. PRISMA of research trends analysis about the educational use of ChatGPT

1) 최종 분석 대상 문헌 전체 목록: <https://sites.google.com/view/chatgptineducation>

Table 1. Literature analysis categories and criteria

분석범주	분석기준
발표 시기	논문 게재 연도 및 월
연구 유형	1. 개념연구 2. 문헌연구 3. 인간 대상 실증연구 4. ChatGPT 대상 실증연구 5. 인간과 ChatGPT 대상 실증연구
대상 학교급	1. 특정되지 않음 2. 유아교육 3. 초등교육 4. 중등교육 5. 고등교육
연구된 분야	1. 특정되지 않음 2. 의학·간호교육 3. 언어교육 4. 과학교육 5. 관광교육 6. 공학교육 7. 미디어/영상교육 8. 수학교육 9. 기타

된 교과학습 분야로는 특정되지 않음, 의학·간호, 언어, 과학, 관광, 공학, 미디어/영상, 수학, 기타로 분류하여 분석을 실시하였다.

나. 데이터 전처리

PRISMA 절차를 통해 선정된 72편의 국내·외 논문들의 영문 초록을 추출하여 R 4.2.1과 RStudio 2023.03.0 +386 버전을 사용하여 데이터 전처리를 실시하였다. 분석 대상 국문 논문에서 영문 초록이 존재하지 않는 경우 국문 초록을 번역하여 데이터에 포함하였다. SnowballC 패키지를 활용하여 영어 어근을 추출하였고 이를 token으로 구성하였다. 다음으로 불용어를 설정하고 반복적으로 텍스트를 정제하여 정확한 분석을 실시하고자 하였다. 본 연구에서는 tidytext 패키지에서 제공하는 stop_words 외에도 ‘ChatGPT’, ‘OpenAI’, ‘AI’, ‘artificial’, ‘intelligence’, ‘LLM’, ‘NLP’, ‘model’과 같이 최종 분석 대상 문헌의 초록에서 많이 등장한 단어는 추가적으로 불용어로 설정하여 제외하였다.

다. 텍스트마이닝 기법

본 연구 분석을 위해 선정된 72편의 논문을 대상으로 ChatGPT의 교육적 활용과 관련된 문헌들에서 가장 많이 언급한 키워드들의 빈도 분석을 실시하였다. 키워드 빈도 분석은 문헌 안에서 저자가 강조한 부분이 무엇인지 파악하기 위한 방법으로 많이 활용되고 있다(Kim, 2021; Lee, Yoo, & Kim, 2020). 분석 후 빈도가 가장 높은 상위 30개의 키워드를 판별하고 이를 시각화하여 한눈에 파악할 수 있도록 하였다.

다음으로 LDA 기반 토픽 모델링을 통해 분석 대상 논문들의 텍스트 데이터를 기반으로 추출된 토픽으로 연구동향을 분석하고자 하였다. 이를 위해 전처리가 된 데이터를 기반으로 문서-단어 행렬(DTM: Document Term Matrix)을 생성하였으며, 하이퍼파라미터 튜닝(hyperparameter tuning)을 적용하여 토픽 모델링의 최적의 토픽 수를 결정하였다. 여러 모델의 성능 지표와 모델 복잡도는 Griffiths & Steyvers(2004)가 제시한 Griffiths2004로 확인이 가능하다(Kim, 2021). 본 연구는 최적의 토픽 수를 선

정하기 위해 토픽 모델을 2개에서 20개까지로 설정하여 비교 분석을 하였으며, 성능 지표를 기반으로 최적의 토픽 수를 선정하였다. 선정된 토픽 수를 기준으로 topic-models 패키지를 사용하여 LDA 기반 토픽 모델링을 실시하였다(Grün & Hornik, 2011). 토픽 모델링 후 Lee et al. (2022)의 연구와 같이 각 토픽에 포함된 문헌들을 분류하고 분류된 문헌들의 초록을 검토하여 정확도 높은 제목을 각 토픽에 부여하고자 하였다.

IV. 연구 결과

1. 체계적 문헌분석

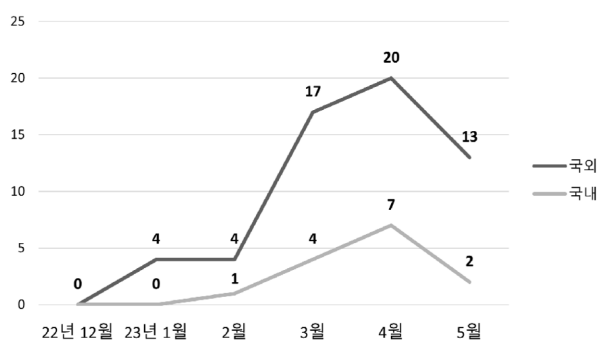
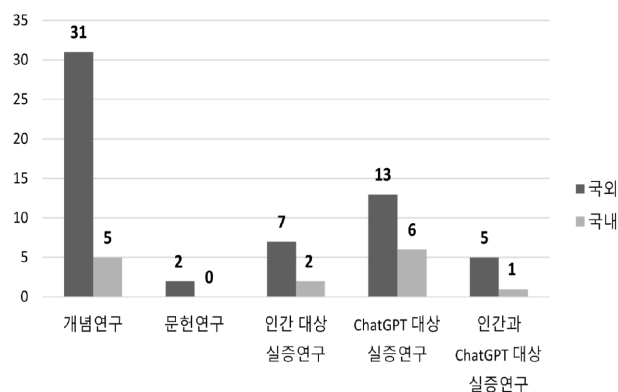
ChatGPT의 교육적 활용 연구동향을 체계적 문헌분석을 통해 파악하고자 (1) 문헌발표 시기(게재년월), (2) 연구 유형, (3) 대상 학교급, (4) 연구된 학술분야를 분석 범주로 설정하고 분석을 실시하였다. 각 분석기준에 따른 결과를 전체적으로 제시하면 Table 2와 같다.

첫 번째로 문헌발표시기를 게재년월로 살펴보면 Figure 2와 같이 국외에서는 2023년 1월을 시작으로 2월까지 각각 4편(5.6%)의 논문이 게재되었으며 3월부터 17편(23.6%)으로 그 수가 급증하고 4월도 역시 20편(27.8%)으로 증가하는 모습을 보였다. 5월에는 13편(18.1%)으로 문헌 수가 다소 감소된 것이 확인되었다. 국내에서는 2월에 1편(1.4%)을 시작으로 3월에 4편(5.6%), 4월에 7편(9.7%)의 논문이 게재되어 증가 추세를 보였으며 5월에 2편(2.8%)으로 감소된 것이 확인되었다. 전반적으로 5월에는 논문수가 감소되었으나 전반적으로 ChatGPT의 교육적 활용에 대한 관심이 급증하여 3월과 4월에 많은 연구들이 발표가 된 경향성을 알 수 있다.

두 번째로 연구유형을 국외 문헌부터 살펴보면 Figure 3과 같이 개념연구가 31편(43.1%)으로 가장 많이 이루어졌으며, 그 다음으로 ChatGPT를 통한 교과 지식, 시험 등에 대한 성능을 평가한 실증연구가 13편(18.1%)으로 많이 이루어진 것을 확인할 수 있다. 다음으로는 인간 대상 실증연구가 7편(9.7%), 인간과 ChatGPT 대상 실증 연구는 5편(6.9%), 문헌연구가 2편(2.8%)의 순으로 나타났다. 한편, 국내에서 발표된 연구들을 살펴보면 ChatGPT 대상

Table 2. The result of systematic literature review

분석범주	분석기준	논문 수(%)	분석범주	분석기준	논문 수(%)
발표 시기	2022년 12월	0 (0%)	대상 학교급	특정되지 않음	22 (30.6%)
	2023년 1월	4 (5.6%)		유아교육	1 (1.4%)
	2023년 2월	5 (6.9%)		초등교육	3 (4.2%)
	2023년 3월	21 (29.2%)		중등교육	4 (5.6%)
	2023년 4월	29 (37.5%)		고등교육	42 (58.3%)
	2023년 5월	15 (20.8%)		특정되지 않음	18 (25.0%)
	개념연구	36 (50.0%)	연구된 분야	의학·간호교육	16 (22.2%)
연구 유형	문헌연구	2 (2.8%)		언어교육	13 (18.1%)
	인간 대상 실증연구	9 (12.5%)		과학교육	8 (11.1%)
	ChatGPT 대상 실증연구	19 (26.4%)		관광교육	3 (4.2%)
	인간/ChatGPT 대상 실증연구	6 (8.3%)		공학교육	3 (4.2%)
				미디어/영상교육	2 (2.8%)
				수학교육	2 (2.8%)
				기타	7 (9.7%)

**Figure 2.** Publication year and month**Figure 3.** Research types

실증연구가 6편(8.3%)으로 가장 많이 이루어졌으며 그 다음으로 개념연구가 5편으로(6.9%) 나타났다. 그 외에 인간 대상 실증연구가 2편(2.8%), 인간과 ChatGPT 대상 실증연구가 1편(1.4%)이 실시되었다. 추가로 게재시기와 연구유형을 맵핑한 결과, Table 3에서 보는 것과 같이 3월에 게재된 21편의 문헌 중 6편(8.3%), 4월에 27편 중 14편 (19.5%), 5월에 15편 중 10편(14.0%)이 실증 연구인 것으

로 확인되었다. 즉, 게재된 문헌 중 실증 연구 문헌의 비율이 점차 증가하는 것을 알 수 있다.

세 번째로 대상 학교급에 대해 살펴보면 Figure 4와 같이 국외에서는 고등교육 맥락에서 실시된 문헌들이 42편 (58.3%)으로 압도적으로 많았다. 그 다음 맥락이 특정되지 않은 대상에 대한 연구가 13편(18.1%)으로 많이 이루

Table 3. The mapping result of publication period and research type

연구유형 게재년월	개념연구	문헌연구	인간 대상 실증연구	ChatGPT 대상 실증연구	인간/ChatGPT 대상 실증연구
2023년 1월	3 (4.2%)	0	0	0	1 (1.4%)
2023년 2월	2 (2.8%)	0	1 (1.4%)	2 (2.8%)	0
2023년 3월	14 (19.4%)	1 (1.4%)	0	6 (8.3%)	0
2023년 4월	12 (16.7%)	1 (1.4%)	4 (5.6%)	7 (9.7%)	3 (4.2%)
2023년 5월	5 (6.9%)	0	4 (5.6%)	4 (5.6%)	2 (2.8%)

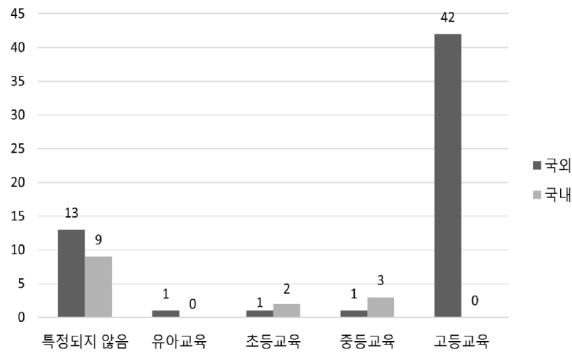


Figure 4. Education levels

어졌으며 유아, 초등, 중등교육에 대한 연구가 각각 1편(1.4%)으로 동일하게 이루어졌다. 국내 문헌을 살펴보면 대상이 특정되지 않은 논문이 9편(12.5%)으로 가장 많았다. 그 외의 연구들을 살펴보면 중등교육 3편(4.2%), 초등교육 2편(2.8%) 순으로 나타났다. 이와 같이 국외에서는 고등교육에서의 ChatGPT 활용에 대한 논의가 가장 많이 이루어졌고 국내에서는 유아와 고등교육을 대상으로 한 ChatGPT 활용 연구는 아직 이루어지지 않은 것을 알 수 있다. 구체적으로 유아, 초등, 중등교육 연구를 진행한 문헌들을 살펴보면, 유아교육 대상으로는 6명의 교수들을 대상으로 ChatGPT의 역할, 도전, 성장에 대한 논의를 실시한 연구가 확인되었다(Luo et al., 2023). 초등교육 연구로는 수학(Son, 2023), 영어(Jeon & Lee, 2023), 미술교육(Lee & Song, 2023)에서의 ChatGPT 활용에 대한 연구가 실시되었다. 중등교육을 대상으로는 영어교육 3편(R. Kim, 2023; T. Kim, 2023; Shin, 2023)과 화학교육 1편(Pawlak, 2023)의 연구가 수행된 것을 볼 수 있었다.

네 번째로 연구된 학문분야를 국외 문헌부터 살펴보면 Figure 5에서 보는 것과 같이 교과를 특정하지 않은 연구가 18편(25%)으로 가장 많은 것을 확인할 수 있다. 두 번째로 많이 연구가 실시된 분야는 의학·간호교육으로 16편(22.2%)이 나타났으며, 그 다음으로는 과학교육 6편(8.3%), 언어교육과 기타 각각 5편(6.9%), 관광교육 3편(4.2%), 공학과 미디어/영상교육 각각 2편(2.8%), 수학교육 1편(1.4%) 순으로 연구가 수행되었다. 국내 문헌을 살펴보면 언어교육이 8편(11.1%)으로 가장 많이 이루어졌으며, 과학교육과 기타 각각 2편(2.8%), 공학과 수학교육이 각각 1편(1.4%)으로 나타났다. 해당 분석 기준에 포함된 기타에는 경영학, 경제, 도서관, 미술, 안전, 정보리터러시, 회계교육 분야에는 각각 1편의 연구가 포함되었다.

2. 키워드 빈도 분석

2022년 12월부터 2023년 5월까지 게재된 ChatGPT의 교육적 활용 관련 국내외 문헌 72편의 초록을 대상으로 주요 출현 키워드의 빈도를 확인하고자 하였으며, 총 4,882개의 어근이 수집되었다. 어근의 출현 빈도를 분석하였으며 결과를 제시하기 위해 이를 명사 또는 형용사 중 데이터 내 가장 많이 출현된 품사의 형태로 변경하여 제시하였다. 예를 들어, 빈도 순위 1위의 'gener'는 'generative'로 기재하였다. 우선 총 72편 문헌의 초록에서 가장 많이 출현된 30개의 키워드들을 정리하면 Table 4와 같다. 결과를 살펴보면 'generative'가 125회로 가장 많이 등장하였으며, 'student'가 122회, 'learning'이 89회, 'research'가 81회, 'question'이 69회 순으로 많이 출현된 것이 확인되었다.

이를 통해 ChatGPT는 사용자인 교사, 학생 등이 기대하는 산출물을 생성하는 역할을 하여 ChatGPT의 교육적

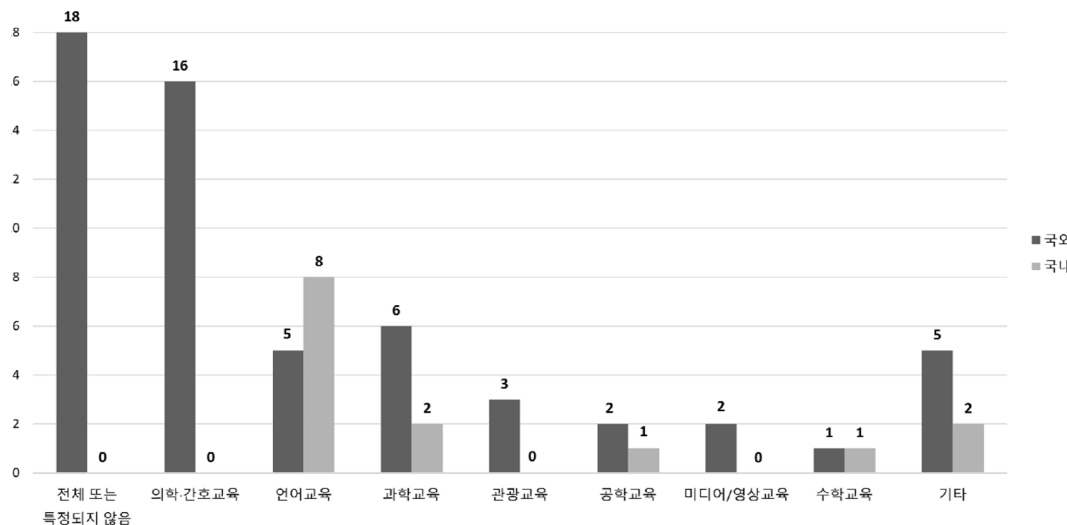


Figure 5. Disciplinary fields

Table 4. Top 30 term frequency result

순위	키워드	빈도	순위	키워드	빈도
1	generative	125	16	technology	42
2	student	122	17	academic	40
3	learning	89	18	evaluation	39
4	research	81	19	included	38
5	question	69	20	writing	38
6	language	67	21	challenge	37
7	study	63	22	process	36
8	tool	62	23	information	34
9	potential	61	24	limitation	34
10	base	60	25	integrity	33
11	response	56	26	human	32
12	answer	44	27	performance	32
13	teaching	43	28	providing	32
14	development	42	29	application	31
15	medical	42	30	chatbot	31

* 빈도수가 같을 경우 키워드의 알파벳순으로 순위 나열

활용에 대한 논의가 많이 이루어지고 있는 것으로 유추할 수 있다. 또한, 전반적으로 ChatGPT를 학습과 연구에서 활용할 수 있는 방안에 대한 논의가 많이 이루어지는 것으로 볼 수 있다. 그 다음으로 출현 빈도가 높은 단어들의 성격을 살펴보면 ‘language’, ‘medical’, ‘writing’, ‘potential’, ‘challenge’, ‘limitation’ 등 연구들에서 특정 과목영역을 대상으로 ChatGPT의 교육적 활용 방법 및 영향(도전 및 제한점)을 논의한 것으로 유추할 수 있다.

3. LDA 기반 토픽 모델링 분석

가. DTM 구성 및 최적 토픽수 선정

본 연구에서 최종적으로 선정한 72편의 문헌 초록을 수집하여 토픽 모델링 분석을 실시하였다. 하이퍼파라미터 튜닝을 통해 2-20개의 토픽 모델을 만든 후 각 모델의 성능을 Griffiths2004의 성능 지표로 파악하여 최적의 토픽수를 결정하였다. 토픽 모델의 성능이 1에 가까우면 가장 성능이 좋은 것으로 해석 가능하며, 일반적으로 토픽수가 증가하나 성능 지표가 두드러지게 향상되지 않고 등락이 반복되는 현상이 보여지는 시점에서 토픽의 수를 결정한다(Kim, 2021). 본 연구에서 도출된 토픽의 개수가 12개인 시점부터 등락이 반복되는 현상을 보이나 본 논문의 분석 대상이 72편인 것을 고려하였을 때에 토픽의 수가 너무 많으면 의미 있는 해석이 어려울 것으로 판단하였다. 따라서 Lee et al.(2022)의 연구와 유사하게 성능 지표의 기울기가 상대적으로 낮아지기 전인 8을 최종 토픽의 수로 결정하였다.

나. 토픽 분석 결과

여덟 개의 토픽을 기반으로 LDA 모델을 만들어 각 토픽에 포함된 문헌 초록과 도출된 주요 단어를 고려하여 토픽 주제명을 결정하였다. 토픽의 키워드는 어근으로 도출되었기 때문에 명사 또는 형용사 형태로 변형하여 제시하였다. 각 토픽마다 문헌이 6-11개로 분포되어 포함된 문헌 수가 상대적으로 적은 편으로 판단되었기 때문에 모든 문헌의 초록을 검토하여 주제명을 도출하였다. 각 토픽별 주제명과 자세한 정보는 Table 5와 같다.

먼저, 토픽 1의 등장 확률이 높은 키워드는 ‘health’, ‘limitation’, ‘potential’, ‘improvement’, ‘practice’, ‘benefit’, ‘risk’, ‘concern’, ‘skill’, ‘included’로 나타났다. 해당 토픽에서 나타난 연구의 주제로는 주로 의학 및 간호분야(예: 의학, 치의학, 간호학, 약학, 보건학 등) 교육에서의 ChatGPT의 장점과 제한점에 대한 연구(Eggmann et al., 2023; Lee, 2023; Sallam, 2023; Sallam et al., 2023)가 이루어졌다. 그 외에도 치의학에서 ChatGPT의 영향으로 인해 교육과정 검토를 다룬 연구(Thurzo et al., 2023), 간호 교육에서 교수진의 역할을 논의한 연구(Sun & Hoelscher, 2023)도 실시되었다. 전반적으로 토픽 1에서는 의학 또는 간호교육에서의 ChatGPT의 가능성과 제한점을 논의하고 전반적인 교육과정과 교수진들의 역할을 탐색하는 연구들이 주를 이루고 있는 것이 나타났다. 따라서 토픽 1의 주제명을 ‘의학·간호교육에서 ChatGPT로 인한 영향과 대비’로 선정하였다.

토픽 2에서 도출된 키워드는 ‘process’, ‘effect’, ‘support’,

Table 5. The result of LDA topic modeling analysis

토픽 번호	주제명	주요 키워드	문헌 수(%)	대표 문헌
1	의학·간호교육에서 ChatGPT로 인한 영향과 대비	health, limitation, potential, improvement, practice, benefit, risk, concern, skill, included	7 (10%)	Sallam et al. (2023) Thurzo et al. (2023)
2	교과 내 ChatGPT 실제적 적용의 영향 탐색 및 프레임워크 제안	process, effect, support, feedback, learner, design, task, activity, recent, base	6 (9%)	Bauer et al. (2023) Uddin et al. (2023)
3	교육실제 및 연구에서 ChatGPT 활용으로 인한 미래 고찰	research, paper, potential, critical, issue, discussion, impact, technology, teaching, application	11 (15%)	Ivanov & Soliman (2023) Hwang & Chen (2023)
4	평가에서의 ChatGPT 성능 분석	question, answer, medical, test, performance, response, score, knowledge, training, required	8 (11%)	Clark (2023) Das et al. (2023)
5	교과 학습 활동 및 과제에서 ChatGPT 활용 방안 탐색	language, development, study, writing, exploration, chatbot, conversation, examination, advancement, technological	10 (14%)	Kang (2023) Yu (2023)
6	고등교육에서 ChatGPT 활용 시 고려사항	student, tool, learning, academic, integrity, assessment, detection, creation, plagiarism, base	8 (11%)	Cotton et al. (2023) Crawford et al. (2023)
7	교과에서 전문가, 교사, 학생의 ChatGPT 활용 경험 기반 시사점	evaluation, challenge, study, data, teacher, teaching, method, material, offer, providing	11 (15%)	Kasneci et al. (2023) Lee & Song (2023)
8	교과에서 ChatGPT 활용에 가능성과 도전	generative, learning, human, response, capability, article, media, text, reflection, essay	11 (15%)	Bender (2023) Farrokhnia et al. (2023)

‘feedback’, ‘learner’, ‘design’, ‘task’, ‘activity’, ‘recent’, ‘base’로 나타났다. 토픽 2에 포함된 연구들을 살펴보면 안전 교육 및 훈련에서 ChatGPT를 실제로 활용하여 학생들의 위험 인식에 대한 영향 탐색(Uddin et al., 2023), 한국의 대학수학능력시험 영어 영역에서 ChatGPT가 학생들의 인지 처리 과정에 주는 영향 탐색(R. Kim, 2023) 등이 있었다. 또한, 동료 피드백 지원을 위해 ChatGPT와 같은 거대 언어 모델의 역할에 대한 학제간 프레임워크 연구(Bauer et al., 2023), 최적화된 ChatGPT와의 상호작용을 위해 프롬프트 엔지니어링 프레임워크 제안(L. Lo, 2023) 등 개념적 프레임워크를 탐색 및 제안한 연구들도 토픽 2에서 나타났다. 전반적으로 해당 토픽에서는 ChatGPT를 교육에 실제적으로 적용하여 영향을 탐색하거나 ChatGPT를 교육 현장에 적용하기 위한 프레임워크를 제안하는 연구들이 실시되었음을 알 수 있다. 따라서 이를 토대로 본 토픽 2의 주제명을 ‘교과 내 ChatGPT 실제적 적용의 영향 탐색 및 프레임워크 제안’으로 정하였다.

토픽 3에서 도출된 주요 키워드로는 ‘research’, ‘paper’, ‘potential’, ‘critical’, ‘issue’, ‘discussion’, ‘impact’, ‘technology’, ‘teaching’, ‘application’이 있었다. 해당 토픽에는 ChatGPT와 같은 생성형 AI 등장으로 인한 교육과 연구에서의 잠재력 논의(Hwang & Chen, 2023), 관광 교육과 연구에서의 ChatGPT 활용에 대한 잠재력 또는 윤리적 영향 관련 연구(Ivanov & Soliman, 2023; Skavronskaya, Hadinejad, & Cotterell, 2023), 화학 교육과 연구에서의 ChatGPT의 잠재적 영향을 탐색한 연구(Emenike & Emenike, 2023)가 포

함되었다. 그 외에도 교과별로 ChatGPT의 영향을 탐색하며 의사 결정, 연구 산출물, 윤리적인 영향과 같은 앞으로 연구 수행에 있어서의 논의를 이룬 연구 등을 살펴볼 수 있었다(Cooper, 2023; Corsello & Santangelo, 2023; Eyesenbach, 2023). 이러한 연구들은 미래를 바라보며 교육과 연구에 있어서 ChatGPT의 활용에 대해 깊은 논의를 하고 있는 것으로 볼 수 있어, 토픽 3의 주제명은 ‘교육실제 및 연구에서 ChatGPT 활용으로 인한 미래 고찰’로 선정하였다.

토픽 4에서 등장 확률이 높았던 키워드는 ‘question’, ‘answer’, ‘medical’, ‘test’, ‘performance’, ‘response’, ‘score’, ‘knowledge’, ‘training’, ‘required’로 나타났다. 토픽 4에 포함된 연구로는 의학 시험에서의 ChatGPT 성능 평가(Das et al., 2023; Ghosh & Bir, 2023; Gilson et al., 2023; Huh, 2023), 화학 시험에서의 성능 평가(Clark, 2023; Fergus, Botha, & Ostovar, 2023), 교사의 ChatGPT를 활용한 영어 시험 출제에 대한 성능 평가(Shin, 2023) 등이 있었다. 주로 해당 토픽에서 나타난 연구들은 의학과 화학 시험에서의 ChatGPT 성능을 평가하고 교사가 ChatGPT를 통해 시험을 출제하는 것에 대한 평가를 다루었다. 따라서 이를 토대로 토픽 4의 주제명은 ‘평가에서의 ChatGPT 성능 분석’으로 선정하였다.

토픽 5에서 도출된 키워드로는 ‘language’, ‘development’, ‘study’, ‘writing’, ‘exploration’, ‘chatbot’, ‘conversation’, ‘examination’, ‘advancement’, ‘technological’이 있다. 본 토픽에 포함된 연구는 ChatGPT의 출현으로 인한 국어 글쓰기 교육이 겪게 될 영향 및 대응(Kang, 2023), ChatGPT의

작문 능력 및 작문 교육이 마주하게 될 변화와 대응(Jang, 2023)이 있었다. 또한, 외국어 영작(Yan, 2023), 중국어 교육(Ha, 2023), 지리 교육(Yu, 2023)과 회계 교육(Yoon, 2023)에서의 ChatGPT의 역할을 탐구하고 이의 활용 방안들을 탐색하는 연구가 포함되었음이 확인되었다. 이와 같이 언어 교육과 보고서 등의 글쓰기 과제에서 ChatGPT가 가지게 될 역할을 논의하는 연구가 주를 이루고 특정 과목에서의 ChatGPT 활용에 대한 논의를 이루는 연구도 일부 나타났다. 따라서 토픽 5의 주제명은 ‘교과 학습 활동 및 과제에서 ChatGPT 활용 방안 탐색’으로 정하였다.

토픽 6에서 나타난 키워드는 ‘student’, ‘tool’, ‘learning’, ‘academic’, ‘integrity’, ‘assessment’, ‘detection’, ‘creation’, ‘plagiarism’, ‘base’가 포함되었다. 토픽 6에서는 고등교육 맥락에서 학생들이 ChatGPT를 활용함으로써 고려해야 할 학업 성실성(Chaudhry et al., 2023; Cotton, Cotton, & Shipway, 2023; Perkins, 2023)과 같이 ChatGPT를 윤리적으로 사용하기 위한 논의(Crawford, Cowling, & Allen, 2023)에 관한 연구들이 주로 발견되었다. 이렇듯 ChatGPT 등장으로 인해 학생들의 학업 성실성과 윤리적 이슈와 관련하여 ChatGPT를 교육적으로 사용해야 할 때의 고려사항들을 논의하는 연구들이 본 토픽에 포함됨이 확인되었다. 따라서 토픽 6의 주제명은 ‘고등교육에서 ChatGPT 활용 시 고려사항’으로 정하였다.

토픽 7에서 등장한 확률이 높은 키워드는 ‘evaluation’, ‘challenge’, ‘study’, ‘data’, ‘teacher’, ‘teaching’, ‘method’, ‘material’, ‘offer’, ‘providing’인 것으로 확인되었다. 해당 토픽에 포함된 연구로는 교수 및 학습을 위한 ChatGPT와 Stable Diffusion 기반 그림 자료 추천의 가능성 및 시사점 논의(Lee & Song, 2023), ChatGPT와의 경험에 기반한 신경외과 교육(Sevgi et al., 2023)과 영어 교수 자료 개발(T. Kim, 2023)에서의 활용 가능성을 탐색한 후 시사점을 제시한 연구가 있었다. 이와 더불어 Kasneci et al.(2023), Luo et al.(2023), Son(2023), Su & Yang(2023)과 같이 ChatGPT의 교육적 활용에 대해 저자 또는 전문가들의 관점을 기반으로 시사점을 제시하는 연구들도 나타났다. 이와 같이 해당 토픽에서는 ChatGPT를 실제로 교육 현장에 적용한 실증 연구 보다는 다양한 교과에서 전문가, 교사, 또는 학생들의 ChatGPT 활용 경험을 기반으로 ChatGPT의 영향을 논의한 연구들을 주로 찾아볼 수 있었다. 따라서 토픽 7 주제명은 ‘교과에서 전문가, 교사, 학생의 ChatGPT 활용 경험 기반 시사점’으로 선정하였다.

마지막으로 토픽 8에서 도출된 키워드는 ‘generative’, ‘learning’, ‘human’, ‘response’, ‘capability’, ‘article’, ‘media’, ‘text’, ‘reflection’, ‘essay’로 나타났다. 토픽 8에 포함된 연구로는 ChatGPT 활용을 위한 프롬프트 기반 글쓰기 가능

성(Li et al., 2023)과 미디어 교육에서의 ChatGPT와 그 외 생성형 AI의 장점, 가능성, 제한점, 또는 위협(Bender, 2023; Pavlik, 2023)을 논의한 연구들을 확인할 수 있었다. 비판적인 분석과 역설 이론에 기반한 경영학 교육에서의 생성형 AI 교육적 활용에 대한 영향을 탐색한 연구(Lim et al., 2023) 및 교육에서 ChatGPT를 활용할 때의 장점, 약점, 기회, 위협에 대해 논의한 연구(Farrokhnia et al., 2023) 등도 본 토픽에 포함되었다. 이를 토대로 토픽 8의 주제명을 ‘교과에서 ChatGPT 활용에 대한 가능성과 도전’으로 정하였다.

V. 결론 및 제언

본 연구는 생성형 AI의 한 종류인 ChatGPT 관련 최근 연구 문헌의 동향을 분석하고 ChatGPT 교육적 활용 가능성에 대하여 어떠한 학술적 담화가 이루어지고 있는지 토픽별로 살펴보고자 하였다. 아직까지 ChatGPT의 교육적 활용 연구가 초기 단계인 점을 고려하여 전반적 연구 동향 및 주제 분석에 초점을 두었으며, 이를 통해 향후 ChatGPT 교육적 활용의 실제 및 연구에 대한 시사점을 도출하고자 하였다. 본 연구를 위해 PRISMA 방법으로 국내외 문헌을 총 72편 선정하였으며, 체계적 문헌분석과 키워드 빈도 분석, LDA 기반 토픽 모델링 방법을 통해 전반적인 연구동향을 파악하고 연구동향 주제어 여덟 개를 도출하였다. 연구문제 별로 주요 연구 결과와 시사점을 논의하면 다음과 같다.

첫째, 전반적인 연구동향을 파악하기 위해 체계적 문헌 분석으로 ChatGPT의 교육적 활용 관련 연구동향을 살펴본 결과, 국내외에서 실시된 실증연구가 총 34편으로 전체의 47.2%의 비중을 차지하여 높은 것으로 나타났다. 이는 ChatGPT가 발표된 시점을 고려하여 최종 분석 대상 문헌을 2022년 12월부터 2023년 5월까지로 한정하여 선정한 것에 비해 실증연구의 수가 절반에 가깝게 분포되어 있음을 의미한다. ChatGPT 출시 2개월 후인 2023년 1월부터 인간과 ChatGPT 대상 실증연구(Huh, 2023)가 발표되기 시작하였으며, 이와 같이 실증연구가 빠르게 수행된 것은 ChatGPT가 발표와 동시에 사용자가 직접 해당 기술을 경험할 수 있는 환경을 제공하였기 때문인 것으로 해석할 수 있다. 또한, ChatGPT가 인간과 대화기반으로 상호작용하는 특성을 보유하고 있기 때문에, ChatGPT 대상 실증연구라는 새로운 연구유형이 가능하다는 점도 주목할 만하다. 본 연구의 분석 결과에 따르면 실증연구 중 ChatGPT 대상 실증연구가 국외 13편(18.1%)과 국내 6편(8.3%)으로 가장 많이 이루어진 것으로 나타났다. ChatGPT 대상 실증연구의 예로는 Clark(2023)과 Gilson et al.(2023)

과 같이 시험 또는 교과 관련 인지적인 질문들을 하고 양적 또는 질적 데이터로 성능을 분석한 연구들의 비중이 많았다. 따라서 향후 교육에서의 신기술에 대한 성능 평가 연구, ChatGPT와 같이 기계와 인간이 상호작용을 하는 연구 등 신기술 활용 연구 유형이 점차 다양화가 될 가능성을 예측할 수 있다. 다만 생성형 AI는 버전과 시기에 따라 입력한 프롬프트에 대한 결과물이 상이할 가능성이 있다는 점을 고려하여 연구를 실시할 필요가 있다.

둘째, ChatGPT의 교육적 활용에 대한 연구가 이루어진 대상 학교급을 살펴보면 고등교육이 42편(58.3%)으로 압도적으로 많은 것이 확인되었다. 본 연구의 LDA 기반 토픽 모델링 결과에서도 토픽 6이 ‘고등교육에서의 ChatGPT 활용 시 고려사항’이라는 독립적 주제로 나타났다. 반면, 유아, 초등, 중등교육 분야 대상 논문은 8편에 불과하여 향후 K-12 교육맥락의 실증연구가 더욱 많이 실시될 필요성이 있음을 알 수 있다. 연구된 학문분야를 살펴보면 최종 분석 대상 문헌 72편 중 의학·간호교육이 16편(22.2%)으로 많이 도출되었다. LDA 기반 토픽 모델링 결과에서도 토픽 1이 ‘의학·간호교육에서 ChatGPT로 인한 영향과 대비’로 의학교육이 주요 주제 중 하나로 도출되었으며 토픽 4인 ‘평가에서의 ChatGPT 성능 분석’에서도 절반 정도의 문헌이 의학교육 관련 연구인 것이 확인되었다. 이는 C. Lo(2023)에서 의학교육에 대한 ChatGPT의 성능 평가 연구가 다양한 국가에서 이루어진 것을 발견한 것과 유사한 결과이다. 의학교육에서 시험에 대한 ChatGPT의 성능을 평가한 연구가 많이 이루어진 것은 ChatGPT의 신뢰성을 확보한 후 의학교육에서 ChatGPT를 시뮬레이션 기반 학습 등과 같이 활용하기 위한 것으로 보여진다(Gilson et al., 2023). 또한, 선행연구를 살펴보면 ChatGPT를 퀴즈, 시험 준비 보조도구 등과 같이 평가에 활용하기 위한 방법이 다양하게 제시된 것을 확인할 수 있다(Al-Worafi et al., 2023; Cooper, 2023; Kasneci et al., 2023). 따라서 의학분야 외에도 다양한 교과 분야에서 평가목적의 ChatGPT 활용 가능성을 탐색할 필요가 있다.

다음으로 많은 연구가 이루어진 교과는 언어 13편(18.1%)과 과학교육 8편(11.1%)으로 과목영역에서 다소 편중된 경향성을 보였다. LDA 기반 토픽 모델링 결과에서도 토픽 5인 ‘교과 학습 활동 및 과제에서 ChatGPT 활용 방안 탐색’이 하나의 주제로 도출되었으며 본 토픽에 언어교육 관련 연구들이 많이 분포된 것으로 나타났다. 이러한 결과는 ChatGPT가 사용자와 대화를 기반으로 상호작용할 수 있는 기술적 특성을 지니고 있어 언어학습 및 대화기반 문제해결학습에서 쉽게 적용이 가능하기 때문인 것으로 보인다. 향후 연구에서는 보다 다양한 교과 영역에서 ChatGPT의 활용 방안을 탐색하여 ChatGPT를

각 교과영역에 활용했을 때의 강점 또는 제한점을 파악하여 관련 지식을 축적할 필요가 있다. 또한 교육 현장에서 많은 연구가 이루어진 교과들을 시작으로 ChatGPT 활용을 시도할 수 있으나, 학생들의 고차원적 인지 능력의 감소, 학업 진실성에 대한 위협, 표절 문제 등 잠재적 부작용도 균형 있게 고려해서 활용해야 할 것이다(Farrokhnia et al., 2023). 이는 학생과 교수자 모두에게 ChatGPT와 같은 생성형 AI를 효과적이며 윤리적으로 활용할 수 있는 AI 리터러시 교육이 필요함을 시사한다.

셋째, ChatGPT의 교육적 활용 관련 문헌의 키워드 빈도 분석을 실시한 결과, 상위 30개 키워드에서는 ChatGPT의 교육적 활용에 대한 잠재력과 도전과 같은 영향을 탐색하고 이를 어떻게 교육적으로 활용할 수 있을지에 대한 방안을 논의하는 키워드들이 주로 도출된 것을 확인할 수 있다. 대표적으로 ‘potential’(61회), ‘challenge’(37회), ‘limitation’(34회), ‘capability’(30회), ‘effect’(29회)와 같이 각각적으로 ChatGPT의 교육적 활용에 대한 영향을 탐색하는 연구들이 많은 것으로 유추할 수 있다. 이는 LDA 토픽 모델링 분석 결과에서 토픽 1인 ‘의학·간호교육에서 ChatGPT로 인한 영향과 대비’, 토픽 3인 ‘교육실제 및 연구에서 ChatGPT 활용으로 인한 미래 고찰’, 토픽 7인 ‘교과에서 전문가, 교사, 학생의 ChatGPT 활용 경험 기반 시사점’, 토픽 8인 ‘교과에서 ChatGPT 활용에 가능성과 도전’과 같이 ChatGPT의 교육적 영향을 다양한 차원에서 살펴본 연구주제들이 많았다는 점과도 일치하는 것이다. 이러한 결과는 체계적 문헌분석 결과에서도 개념 연구(36편, 50.0%)와 ChatGPT 대상 실증연구(19편, 26.4%)가 많은 비중을 차지하여 ChatGPT의 교육적 활용의 장점과 제한점 등 영향을 살펴본 연구가 상당 부분 이루어진 것으로 볼 수 있다.

넷째, 키워드 빈도 분석 결과 다양한 교과 영역에서 ChatGPT의 활용과 관련된 키워드가 많이 도출된 것을 확인할 수 있었으며, 대표적으로 ‘language’(67회), ‘tool’(62회), ‘medical’(42회), ‘evaluation’(39회), ‘writing’(38회) 등의 키워드가 출현하였다. 따라서 언어와 의학 교육 등에서 ChatGPT를 교육적으로 활용할 수 있는 방안을 탐색하거나 적용한 연구들이 다수 실시되었음을 알 수 있다. 이를 토픽 모델링 결과를 토대로 살펴보면 토픽 2에서 ‘교과 내 ChatGPT 실제적 적용’과 토픽 5에서 ‘교과 학습 활동 및 과제에서 ChatGPT 활용 탐색’을 다룬 것이 확인되었다. 토픽 2에서 나타난 연구들 중 Uddin et al.(2023)은 대학의 건축 교육 프로그램에서 ChatGPT를 활용한 후 학생들의 위험 인식 수준이 향상한 것을 확인하였으며, R. Kim(2023)은 한국 고등학생을 대상으로 ChatGPT를 활용하여 영어 수능문제에 대한 교수를 진행할 때에 학생들

의 인지적 과정의 영향을 탐색하였으며 사실 정보 파악과 대의파악 유형 문제에 ChatGPT의 교수 개입이 효과적인 것이 나타났다. 토픽 5에서는 언어, 지리, 회계 등 다양한 교과에서 학습 활동과 과제를 수행할 때 ChatGPT를 활용하기 위한 방안들을 탐색하였다. 특히 Yan(2023)은 학생들에게 ChatGPT를 활용하여 제2 외국어 작문을 하도록 하였으며, 학생들은 이러한 도구를 통한 제2 외국어 작문 활동이 효율성을 극대화할 수는 있으나 복잡한 작문 수정 과정이 단순화됨으로 인해 교육의 불공평을 야기하고 ChatGPT로 생성된 결과물이 낮은 표절률을 보임에 따라 나타나는 학업 성실성에 대해 우려를 표하는 것을 발견하였다. 이와 같이 키워드 빈도 분석 결과에서도 ChatGPT의 교육적 활용과 관련된 키워드가 많이 분포되어 있는 것을 통해 다양한 교과영역에서 연구가 실시된 것을 알 수 있다. 하지만, ChatGPT를 실제적으로 교육 현장에 적용한 연구가 포함된 것은 토픽 2, 즉 ‘교과 내 ChatGPT 실제적 적용의 영향 탐색 및 프레임워크 제안’에만 주로 분포된 것으로 나타났다. 이는 C. Lo(2023)에서 ChatGPT를 교육현장에 활용한 연구들이 더욱 필요함을 제시한 것과 유사한 결과라고 볼 수 있다.

다섯째, 본 연구에서 도출한 여덟 개의 토픽 중 토픽 3, 7, 8이 각각 11편(15%)의 문헌으로 가장 많이 분포된 것을 확인할 수 있다. 토픽 3은 ‘교육실제 및 연구에서 ChatGPT 활용으로 인한 미래 고찰’로 ChatGPT를 교과나 연구 수행에서 활용하는 가능성을 주로 다루는 연구가 포함되었으며, 토픽 7은 ‘교과에서 전문가, 교사, 학생의 ChatGPT 활용 경험 기반 시사점’으로 미술, 신경외과, 영어 교육 분야 등에서 전문가, 교사, 학생이 ChatGPT를 활용한 경험을 기반으로 이에 대한 시사점들을 탐색한 연구들이 많았다. 토픽 8은 ‘교과에서 ChatGPT 활용에 가능성과 도전’으로 전반적인 교육이나 특정 교과에서 ChatGPT의 교육적 영향을 강점 또는 제한점, 가능성, 위협으로 심도 있게 살펴본 연구들을 확인할 수 있었다. 반면, 토픽 모델링 분석 결과에서 문헌 비중이 가장 적은 토픽은 6편(9%)의 문헌이 포함된 토픽 2로 확인되었다. 토픽 2는 ‘교과 내 ChatGPT 실제적 적용의 영향 탐색 및 프레임워크 제안’이며 안전과 영어 교육에서 ChatGPT를 교육 현장에 활용하여 이에 따른 영향을 탐색한 연구들을 일부 살펴볼 수 있었다. 하지만, 토픽 2에서 Uddin et al.(2023)과 R. Kim(2023)만 ChatGPT를 실제로 활용한 연구이었으며, 대부분 실제로 활용할 때의 프레임워크를 제안하는 등의 개념적 연구인 것으로 나타났다. 이러한 결과를 통해 향후 개념적 논의를 넘어서 실제로 ChatGPT를 교육 현장에서 활용한 후 학생들에게 어떠한 영향을 끼치는지 다양한 변인을 통해 탐색하는 실증적 연구가 실시될 필요성

을 도출할 수 있다. 앞으로도 ChatGPT와 같은 생성형 AI는 지속적으로 발전 및 활용될 기술로 예측됨에 따라 교육 현장에서 ChatGPT를 직접 활용해보고 학생 또는 교사들에게 효과적으로 활용될 수 있는 방안이 무엇인지 다각적인 현장기반 연구가 이루어져야 할 것이다.

본 연구의 제한점과 후속연구를 제언하면 다음과 같다. 첫째, 본 연구는 ChatGPT의 교육적 활용에 대한 초기 연구들을 수집하여 연구 동향을 분석하여 분석된 문헌의 수가 많지 않다는 제한점이 있다. ChatGPT는 2022년 11월 30일에 등장한 것을 고려하여 본 연구는 2022년 12월부터 2023년 5월까지 발표된 72편의 문헌을 대상으로 분석을 실시하여 기존의 연구 동향 분석 기간 범위보다 상대적으로 좁은 것을 확인할 수 있다. 하지만, PRISMA 절차에 따라 최종 선정된 72편의 문헌이 ChatGPT의 교육적 활용과 관련하여 약 6개월 동안 출판된 문헌인 것을 고려했을 때에 기간에 비해 상대적으로 많은 문헌이 발견된 것을 확인할 수 있다. 본 연구는 신기술 도입 초기 시기의 연구경향을 파악하기 위해 실시되었으며, 앞으로 ChatGPT의 교육적 활용에 대한 연구가 충분히 축적된 후 추가 연구를 실시할 필요가 있다. 둘째, 본 연구는 머신 기반 방식인 키워드 빈도 분석과 토픽 모델링을 활용하여 분석을 실시하였기 때문에 글의 맥락을 고려한 분류 과정에 제한이 있다. 글의 맥락을 정확히 파악하고 세부적인 부분을 파악하기 위해서는 추후에 질적 분석 방법을 추가 적용할 필요성이 있다. 셋째, 본 연구는 국내외 문헌을 동시에 분석하기 위해 국내문헌의 경우 국문을 영문으로 번역하여 분석에 포함하였다는 제한점이 있다. 향후 국내에서 ChatGPT의 교육적 활용 연구의 수가 증가하면 국내 문헌만을 대상으로 한 연구수행이 가능할 것이다. 이러한 제한점에도 불구하고 본 연구는 ChatGPT 도입 초기에 발간된 국내외 문헌들을 대상으로 연구동향 및 주제를 다각적으로 파악했다는 점에서 의의가 있으며, 본 연구의 결과는 향후 ChatGPT의 교육적 활용의 실제 및 관련 연구에 대한 기초 자료로 활용될 수 있을 것이다.

참고문헌

- Al-Worafi, Y. M., Hermansyah, A., Goh, K. W., & Ming, L. C. (2023). Artificial intelligence use in university: Should we ban ChatGPT?. *Preprints*, 2023020400. <https://doi.org/10.20944/preprints202302.0400.v1>
- Baidoo-Anu, D., & Owusu Ansah, L. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *SSRN* 4337484. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4337484

- Bauer, E., Greisel, M., Kuznetsov, I., Berndt, M., Kollar, I., Dresel, M., Fischer, M. R., & Fischer, F. (2023). Using natural language processing to support peer-feedback in the age of artificial intelligence: A cross-disciplinary framework and a research agenda. *British Journal of Educational Technology*.
- Bender, S. M. (2023). Coexistence and creativity: Screen media education in the age of artificial intelligence content generators. *Media Practice and Education*, 1-16.
- Chaudhry, I. S., Sarwary, S. A. M., El Refae, G. A., & Chabchoub, H. (2023). Time to revisit existing student's performance evaluation approach in higher education sector in a new era of ChatGPT—A case study. *Cogent Education*, 10(1), 2210461.
- Choi, S. (2023, May 9). US university instructor “Students’ ChatGPT for homework...0 points” controversy. *Asia Economics*. <https://view.asiae.co.kr/article/2023052009344421383>
- Clark, T. M. (2023). Investigating the use of an artificial intelligence chatbot with general chemistry exam questions. *Journal of Chemical Education*.
- Cooper, G. (2023). Examining science education in ChatGPT: An exploratory study of generative artificial intelligence. *Journal of Science Education and Technology*, 32, 444-452.
- Corsello, A., & Santangelo, A. (2023). May artificial intelligence influence future pediatric research?—The case of ChatGPT. *Children*, 10(4), 757.
- Cotton, D. R., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 1-12.
- Crawford, J., Cowling, M., & Allen, K. A. (2023). Leadership is needed for ethical ChatGPT: Character, assessment, and learning using artificial intelligence (AI). *Journal of University Teaching & Learning Practice*, 20(3), 1-19.
- Das, D., Kumar, N., Longjam, L. A., Sinha, R., Roy, A. D., Mondal, H., & Gupta, P. (2023). Assessing the capability of ChatGPT in answering first-and second-order knowledge questions on microbiology as per competency-based medical education curriculum. *Cureus*, 15(3), e36034.
- Eggmann, F., Weiger, R., Zitzmann, N. U., & Blatz, M. B. (2023). Implications of large language models such as ChatGPT for dental medicine. *Journal of Esthetic and Restorative Dentistry*, 1-5.
- Emenike, M. E., & Emenike, B. U. (2023). Was this title generated by ChatGPT? Considerations for artificial intelligence text-generation software programs for chemists and chemistry educators. *Journal of Chemical Education*, 100(4), 1413-1418.
- Eysenbach, G. (2023). The role of ChatGPT, generative language models, and artificial intelligence in medical education: A conversation with ChatGPT and a call for papers. *JMIR Medical Education*, 9(1), e46885.
- Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2023). A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innovations in Education and Teaching International*, 1-15.
- Fergus, S., Botha, M., & Ostovar, M. (2023). Evaluating academic answers generated using chatgpt. *Journal of Chemical Education*, 100(4), 1672-1675.
- Frieder, S., Pinchetti, L., Griffiths, R. R., Salvatori, T., Lukasiewicz, T., Petersen, P. C., Chevalier, A., & Berner, J. (2023). Mathematical capabilities of chatgpt. *arXiv preprint arXiv:2301.13867*.
- Gabashvili, I. S. (2023). The impact and applications of ChatGPT: A systematic review of literature reviews. *arXiv preprint arXiv:2305.18086*.
- Geerling, W., Mateer, G. D., Wooten, J., & Damodaran, N. (2023). ChatGPT has aced the test of understanding in college economics: Now what?. *The American Economist*, 05694345231169654.
- Ghosh, A., & Bir, A. (2023). Evaluating ChatGPT's ability to solve higher-order questions on the competency-based medical education curriculum in medical biochemistry. *Cureus*, 15(4), e37023.
- Gilson, A., Safranek, C. W., Huang, T., Socrates, V., Chi, L., Taylor, R. A., & Chartash, D. (2023). How does CHATGPT perform on the United States Medical Licensing Examination? The implications of large language models for medical education and knowledge assessment. *JMIR Medical Education*, 9(1), e45312.
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2020). Generative adversarial networks. *Communications of the ACM*, 63(11), 139-144.
- Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. *PNAS*, 101(suppl.1), 5228-5235.
- Grün, B., & Hornik, K. (2011). Topicsmodels: An R package for fitting topic models. *Journal of Statistical Software*, 40(13), 1-30.
- Ha, D. J. (2023). Generative-based AI and the possibility of Chinese language education revolution. *Journal of Chinese Humanities*, 83, 411-432.
- Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promise and implications for teaching and learning*. Boston: Center for Curriculum Redesign.
- Huh, S. (2023). Are ChatGPT's knowledge and interpretation ability comparable to those of medical students in Korea for taking a parasitology examination?: A descriptive study. *Journal of Educational Evaluation for Health Professions*, 20(1), 1-6.
- Hwang, G. J., & Chen, N. S. (2023). Editorial position paper: Exploring the potential of generative artificial intelligence in education: Applications, challenges, and future research directions. *Educational Technology & Society*, 26(2).
- Ivanov, S., & Soliman, M. (2023). Game of algorithms: ChatGPT implications for the future of tourism education and research. *Journal of Tourism Futures*, 9(2), 214-221.

- Jang, S. (2023). ChatGPT has changed the future of writing education-Focusing on the response of writing education in the era of artificial intelligence -. *Writing Research*, 56, 7-34.
- Jeon, J., & Lee, S. (2023). Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT. *Education and Information Technologies*, 1-20.
- Jovanovic, M., & Campbell, M. (2022). Generative Artificial Intelligence: Trends and Prospects. *Computer*, 55(10), 107-112.
- Kang, D.-H. (2023). The advent of ChatGPT and the response of Korean language education. *Korean Language and Literature*, 82(82), 469-496.
- Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Grasser, U., Groh, G., Günemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274.
- Khan, R. A., Jawaid, M., Khan, A. R., & Sajjad, M. (2023). ChatGPT-Reshaping medical education and clinical management. *Pakistan Journal of Medical Sciences*, 39(2), 605.
- Kim, R. (2023). Effects of ChatGPT on the cognitive processing of K-CSAT English reading tasks by Korean high school learners: A preliminary study. *Secondary English Education*, 16(2), 179-205.
- Kim, T. (2023). A field study on the development of teaching materials for secondary English and its use of ChatGPT. *Secondary English Education*, 16(2), 207-218.
- Kim, Y. (2021). *Do it! R Text Mining*. Seoul: EasysPublishing.
- Lee, G. Y., So, H.-J., Lee, C., Jang, H., & Kim, Y. (2022). The use of blockchain technology in lifelong learning: An analysis of research trends and topics. *The Journal of Educational Information and Media*, 28(4), 1041-1067.
- Lee, H. (2023). The rise of ChatGPT: Exploring its potential in medical education. *Anatomical Sciences Education*.
- Lee, H., So, H.-J., & Jin, L. (2020). Analysis of the characteristics and design level of educational artificial intelligence applications. *The Journal of Educational Information and Media*, 26(3), 647-670.
- Lee, S.-H., & Song, K.-S. (2023). Exploring the possibility of using ChatGPT and Stable Diffusion as a tool to recommend picture materials for teaching and learning. *Journal of The Korea Society of Computer and Information*, 28(4), 209-216.
- Lee, S.-S., Yoo, I., & Kim, J. (2020). An analysis of public perception on artificial intelligence(AI) education using big data: Based on news articles and Twitter. *Journal of Digital Convergence*. 18(6), 9-16.
- Li, Y., Sha, L., Yan, L., Lin, J., Raković, M., Galbraith, K., Lyons, K., Gašević, D., & Chen, G. (2023). Can large language models write reflectively. *Computers and Education: Artificial Intelligence*, 100140.
- Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I., & Pechenkina, E. (2023). Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. *The International Journal of Management Education*, 21(2), 100790.
- Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences*, 13(4), 410.
- Lo, L. S. (2023). The CLEAR path: A framework for enhancing information literacy through prompt engineering. *The Journal of Academic Librarianship*, 49(4), 102720.
- Luo, W., He, H., Liu, J., Berson, I. R., Berson, M. J., Zhou, Y., & Li, H. (2023). Aladdin's Genie or Pandora's Box for early childhood education? Experts chat on the roles, challenges, and developments of ChatGPT. *Early Education and Development*, 1-18.
- Mollick, E. R., & Mollick, L. (2023). Using AI to implement effective teaching strategies in classrooms: Five strategies, including prompts. Available at <http://dx.doi.org/10.2139/ssrn.4391243>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *Systematic Reviews*, 10(1), 1-11.
- Pavlik, J. V. (2023). Collaborating with ChatGPT: Considering the implications of generative artificial intelligence for journalism and media education. *Journalism & Mass Communication Educator*, 10776958221149577.
- Pawlak, F. (2023). ChatGPT—a revolution for teaching and learning in chemistry education?!. *CHEMKON*. <https://doi.org/10.1002/ckon.202300010>
- Perkins, M. (2023). Academic integrity considerations of AI large language models in the post-pandemic era: ChatGPT and beyond. *Journal of University Teaching & Learning Practice*, 20(2), 07.
- Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning & Teaching*, 6(1), 1-22.
- Sallam, M. (2023). ChatGPT utility in healthcare education, research, and practice: Systematic review on the promising perspectives and valid concerns. *Healthcare*, 11(6), 887.
- Sallam, M., Salim, N., Barakat, M., & Al-Tammemi, A. (2023). ChatGPT applications in medical, dental, pharmacy, and public health education: A descriptive study highlighting the advantages and limitations. *Narra J*, 3(1), e103-e103.

- Sevgi, U. T., Erol, G., Doğruel, Y., Sönmez, O. F., Tubbs, R. S., & Güngör, A. (2023). The role of an open artificial intelligence platform in modern neurosurgical education: A preliminary study. *Neurosurgical Review*, 46(1), 86.
- Shin, D. (2023). A case study on English test item development training for secondary school teachers using AI tools: Focusing on ChatGPT. *Language Research*, 59(1), 21-42.
- Skavronskaya, L., Hadinejad, A., & Cotterell, D. (2023). Reversing the threat of artificial intelligence to opportunity: A discussion of ChatGPT in tourism education. *Journal of Teaching in Travel & Tourism*, 23(2), 253-258.
- Son, T. (2023). Exploring the possibility of using ChatGPT in mathematics education: Focusing on student product and pre-service teachers' discourse related to fraction problems. *Education of Primary School Mathematics*, 26(2), 99-113.
- Su, J., & Yang, W. (2023). Unlocking the power of ChatGPT: A framework for applying generative AI in education. *ECNU Review of Education*, 6(3), 355-366.
- Sun, G. H., & Hoelscher, S. H. (2023). The ChatGPT storm and what faculty can do. *Nurse Educator*, 48(3), 119-124.
- Surameery, N. M. S., & Shakor, M. Y. (2023). Use chat gpt to solve programming bugs. *International Journal of Information Technology & Computer Engineering (IJITC)* ISSN: 2455-5290, 3(1), 17-22.
- Thurzo, A., Strunga, M., Urban, R., Surovková, J., & Afrashtehfar, K. I. (2023). Impact of artificial intelligence on dental education: A review and guide for curriculum update. *Education Sciences*, 13(2), 150.
- Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, 10(15), 1-24.
- Topsakal, O., & Topsakal, E. (2022). Framework for a foreign language teaching software for children utilizing AR, voicebots and ChatGPT (Large Language Models). *The Journal of Cognitive Systems*, 7(2), 33-38.
- Uddin, S. J., Albert, A., Ovid, A., & Alsharif, A. (2023). Leveraging ChatGPT to aid construction hazard recognition and support safety education and training. *Sustainability*, 15(9), 7121.
- van Dis, E. A. M., Bollen, J., van Rooij, R., Zuldema, W., & Bockting, C. L. (2023). ChatGPT: Five priorities for research. *Nature*, 614(7947), 224-226.
- Wardat, Y., Tashtoush, M. A., AlAli, R., & Jarrah, A. M. (2023). ChatGPT: A revolutionary tool for teaching and learning mathematics. *Eurasia Journal of Mathematics, Science and Technology Education*, 19(7), em2286.
- Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Education and Information Technologies*, 1-25.
- Yoon, Y. (2023). Artificial intelligence and accounting education: Focusing on ChatGPT and its applications. *Korean Computers and Accounting Review*, 21(1), 1-29.
- Yu, J.-J. (2023). Application of artificial intelligence for geography education-Focusing on question answering on ChatGPT-. *Journal of the Association of Korean Photographers*, 33(1), 162-173.