



GPT Assistants API 기반 데이터 시각화 사고 촉진 튜터 시스템 설계 연구*

Design of a Data Visualization Thinking-Facilitation Tutor System Based on the GPT Assistants API

이정훈[†] · 양수경^{††} · 조정원^{†††}
 Jeonghun Lee[†] · Sugyeong Yang^{††} · Jungwon Cho^{†††}

요약

디지털 전환 시대에 따라 데이터의 수집과 활용이 일상화되면서, 데이터를 해석하고 시각화하여 의미를 도출하는 능력은 중등 교육에서 중요한 핵심 역량으로 부상하고 있다. 그러나 기존의 데이터 시각화 교육은 도구 조작 중심의 기술 습득에 치중되어 있어, 학습자의 탐색적 사고와 해석 중심의 학습을 충분히 지원하지 못하는 한계가 존재한다. 이에 본 연구는 중등 학습자의 데이터 기반 사고력과 해석 역량을 향상시키기 위한 방안으로, OpenAI의 GPT Assistants API를 활용한 대화형 데이터 시각화 튜터 시스템(DataViz GPT 튜터)을 설계하고 구현하는 것을 목적으로 하였다. 이 시스템은 학습자가 엑셀 또는 시각화 이미지를 업로드하고 자연어로 질문을 입력하면, GPT가 사고를 유도하는 질문을 통해 시각화 방향을 안내하고 해석을 도와주는 방식으로 작동한다. 시스템 설계에는 GPT의 역할을 ‘사고 촉진 파트너’로 정의한 System Prompt 구성, 분석 흐름에 맞춘 다단계 질문 설계, Tableau 기능 설명서 및 AI 윤리 가이드를 참조하는 Retrieval 기능이 통합되었다. 사용자는 직관적인 웹 인터페이스에서 데이터를 확인하고 질문을 입력하며, GPT는 정답을 제공하는 대신 학습자의 사고를 유도하는 방식으로 대화를 이어간다. 실제 구현된 화면에서는 학습자의 수준과 상황에 맞춰 시각화 대상, 분석 범주, 차트 유형, 시계열 조건 등을 탐색적으로 유도하며, 분석 목적의 명확화와 인사이트 도출을 돕는 구조가 구현되었다.

주제어 GPT, Assistants API, 데이터 시각화, 대화형 튜터, 데이터 리터러시, ITS

ABSTRACT

As data collection and utilization have become routine in the digital transformation era, the ability to interpret and visualize data to derive meaningful insights has emerged as a key competency in secondary education. However, traditional data visualization education has focused primarily on tool operation, often falling short in supporting students' exploratory thinking and interpretive learning. To address this limitation, this study aims to design and implement an interactive data visualization tutor system—DataViz GPT Tutor—based on OpenAI's GPT Assistants API, with the goal of enhancing students' data-driven reasoning and interpretation skills. In the proposed system, students upload Excel datasets or visualization images and ask natural language questions. The GPT model then guides the visualization process and interpretation by providing thinking-oriented questions, rather than direct answers. The system architecture includes a system prompt that defines the GPT's role as a “thinking facilitator,” a multi-step dialog flow aligned with analytical reasoning stages, and a Retrieval feature that links to a Tableau user guide and AI ethics resources. Through an intuitive web-based interface, users can review their data and engage in guided conversations, where GPT elicits reasoning rather than delivering fixed solutions. The implemented interface dynamically adjusts to the learner's context, supporting exploration of variables, chart types, and time-series conditions, ultimately helping clarify analytical goals and derive insights.

Keywords GPT, Assistants API, Data Visualization, Conversational Tutor, Data Literacy, Intelligent Tutoring System (ITS)

†정회원 한국교육과정평가원 부연구위원
 ††정회원 제주대학교 교육대학원 컴퓨터교육전공 석사과정
 †††중신회원 제주대학교 컴퓨터교육과 교수(교신저자)
 논문투고 2025년 06월 12일
 심사완료 2025년 07월 01일
 게재확정 2025년 07월 09일
 발행일자 2025년 08월 06일

* 이 논문은 2021년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임(NRF-2021S1A5C2A04088646)

1. 서론

디지털 전환이 가속화되면서 우리는 데이터를 중심으로 구성된 사회, 이른바 ‘데이터 중심 사회(data-centric society)’로 빠르게 이동하고 있다. 인공지능, 사물인터넷, 자율주행 등 첨단 기술의 대부분은 방대한 양의 데이터를 기반으로 작동하며, AI가 일상화된 지능정보사회에서 데이터를 읽고 해석하며 활용하는 역량은 개인의 정보주체성과 사회참여 능력을 가능하는 중요한 지표로 자리 잡고 있다. 이와 같은 흐름은 OECD의 Learning Compass 2030에서도 ‘데이터 문해력(data literacy)’을 미래핵심역량 중 하나로 강조하며, 초·중등 교육에서의 체계적 접근 필요성을 부각시킨 바 있다[1][2].

특히 최근에는 데이터의 분석이나 해석만큼이나, 이를 시각적으로 구조화하고 의미를 전달하는 능력, 즉 데이터 시각화(data visualization) 역량의 중요성이 강조되고 있다. 데이터 시각화는 단순한 그래픽 기술을 넘어서, 분석과정 전반에서 사고를 촉진하고 핵심 정보를 효과적으로 전달하는 탐색적 사고 도구로 기능한다. 이는 단순한 기술 숙련 차원을 넘어, 데이터를 기반으로 사고하고 판단하며 타인과 소통하는 역량을 포함한다는 점에서 교육적 가치가 높다[3].

그러나 실제 중등 교육 현장에서 데이터 시각화 교육은 여전히 ‘기능 중심의 도구 사용’에 국한되거나, 수동적인 그래프 생성 수준에서 멈추는 경우가 많다. 특히 비전공 학습자의 경우, 데이터 해석 경험이 부족하거나 시각화 도구 사용이 낯설어, 어떤 그래프를 선택하고 어떤 관점에서 해석해야 할지 판단하는 데 어려움을 겪는다[4]. 또한 교사는 학습자의 탐색 과정에 일일이 개입하기 어려워, 자율적 사고와 피드백의 순환이 제한되는 구조적 한계도 존재한다.

이러한 교육 현장의 제약을 보완하고 학습자의 사고 촉진을 지원하기 위한 방안으로, 최근에는 대화형 인공지능 기반 튜터링 시스템이 주목받고 있다. 특히 OpenAI의 GPT 기반 Assistants API는 목적에 따라 언어적 역할을 설정하고 학습자의 질문에 반응하는 멀티턴 대화 시나리오 설계가 가능하다는 점에서, 교육적 상호작용을 위한 유연한 도구로 활용될 수 있다[5][6]. 인공지능 기반 튜터링 시스템이 가지는 의미는 교육 현장에서 선생님을 대체하는 것이 아니라, 선생님을 보조하고, 교육의 질적 향상을 도모하는 것이다.

본 연구는 중등 학습자의 데이터 기반 사고력과 해석 역량, 시각화 전략 설계 역량을 향상시키기 위한 방안으로, GPT Assistants API를 활용한 데이터 시각화 튜터링 시스템을 설계하고자 한다. 본 시스템은 단순한 정답 제시가 아닌, 질문 중심의 사고 유도를 통해 학습자의 인지 과정을 지원하고, Tableau와 같은 시각화 도구와 연계함으로써 기술적 장벽을 낮추고 실천 가능성을 높이고자 한다. 이를 통해, 기존의 AI 튜터링 시스템이 정답 지향적 피드백에 머무는 한계를 넘어, 사고의 유도와 확장을 중심으로 구성된 튜터의 언어적 구조를 설계함으로써 교육적 상호작용의 깊이를 확장한다. 또한, 시각화 도구 활용에 어려움을 겪는 비전공

학습자에게 직관적인 인터페이스와 질문 기반 피드백을 제공함으로써 데이터 리터러시 교육의 포용성과 실용성을 높이는 기반을 마련하고자 한다.

2. 이론적 배경

2.1 데이터 리터러시

데이터 리터러시(data literacy)는 단순한 데이터 활용 능력을 넘어, 데이터를 바르게 해석하고 비판적으로 사고하며, 이를 통해 새로운 의미를 구성하고 타인과 소통할 수 있는 복합적 역량이다. Ridsdale 등(2015)은 이를 “데이터에 접근하고, 이해하며, 분석하고, 평가하고, 효과적으로 의사소통할 수 있는 능력”이라고 정의하며, 디지털 시민으로서의 필수 소양임을 강조하였다[7]. 특히 데이터 리터러시 역량은 오늘날 정보 과잉 시대에서 중요한 필터 역할을 하며, 데이터 기반의 합리적 의사결정, 정책 판단, 사회적 참여 등을 가능하게 한다.

OECD(2019) 또한 미래사회 핵심역량 중 하나로 데이터에 대한 윤리적 태도와 해석 능력, 그리고 복잡한 문제 해결을 위한 데이터 기반 접근을 강조하며[2], 데이터 리터러시를 단일 기술이 아닌 다차원적 역량군으로 규정하고 있다. 이처럼 데이터 리터러시는 이제 더 이상 특정 직군이나 전공자만을 위한 역량이 아니며, 초·중등 교육에서부터 모든 학습자에게 필수적으로 길러져야 할 ‘소양’으로 자리 잡고 있다.

데이터 리터러시의 중요성과 효과는 실제 교육 현장에서도 확인된다. 예를 들어, Witte et al.(2025)의 K-12 대상 프로젝트 기반 학습 연구에서는 실제 데이터를 활용한 수업이 학생들의 분석적 사고와 비판적 의사소통 능력을 크게 향상시켰다고 보고했다[8]. 또한 Hasim et al.(2024)는 중등 교과에 통계적 사고 중심의 프로젝트 기반 활동을 적용한 결과, 학습자의 데이터 해석 역량과 자기주도적 문제 해결 능력이 향상되었음을 실증적으로 제시하였다[9]. 송유경 외(2021)는 고등학생 대상의 데이터 기반 토론 수업 모형을 개발하여, 실제 적용 결과 학습자들의 데이터 리터러시 점수가 유의미하게 향상되었음을 제시하였다[10].

하지만 현행 교육과정에서는 데이터 리터러시가 단편적인 통계 교육이나 ICT 활용 교육의 하위 개념으로 다루어지는 경우가 많으며, 정량 정보의 해석이나 시각적 전달력보다는 도구 사용법이나 결과 수치의 산출에 초점을 두는 한계가 존재한다. 이는 학생들이 실제 사회문제를 탐색하거나 데이터 기반으로 의사결정을 내리는 데 필요한 사고 역량을 충분히 개발하기 어렵게 만든다. 따라서 학교 교육은 단순한 ‘데이터를 다루는 기술’에서 벗어나, 데이터를 통해 사고하고, 의미를 구성하며, 사회적 맥락에서 비판적으로 활용할 수 있는 역량을 함양할 수 있도록 교육 내용을 재편성해야 하며, 이를 지원하는 교수학습 도구 역시 이와 같은 방향성을 반영해야 한다.

2.2 데이터 시각화와 학습자 사고 촉진

데이터 시각화는 수치적 또는 비정형 데이터를 시각적으로 표현하여 의미를 도출하거나 전달하는 과정으로, 데이터 분석의 마지막 단계이자 가장 직관적인 정보 전달 수단이다. 하지만 교육적 관점에서 데이터 시각화는 단지 결과를 아름답게 표현하는 기술이 아니라, 학습자가 데이터를 탐색하고 구조화하며 인과 관계를 사고하는 일련의 과정을 촉진하는 중요한 학습 도구로 작용한다.

예컨대, 추세 파악을 위한 선그래프, 구성비를 위한 원그래프, 이상치 감지를 위한 그림 등은 각각의 목적과 해석 방식이 다르며, 학습자가 이를 선택하고 사용하는 과정 자체가 사고의 흐름을 반영하게 된다. 주애리·김미량(2023)은 초등학생을 대상으로 시각화 자료 활용 토론 수업을 설계·운영한 결과, 학생들의 데이터 해석력과 토론 능력이 눈에 띄게 향상되었다고 보고하였다[11]. Chang et al.(2024)의 STEM 교육 연구에서는 고등학생과 대학생이 대기 환경 데이터를 시각화하는 과정에서 ‘메타비주얼 전략(metavisual strategies)’을 활용하여 인지 과정을 조절하고 사고를 확장하는 양상이 확인되었다[12]. 이처럼 시각화는 학습자의 인지 과정과 직접적으로 연결되며, 그 선택과 해석을 통해 분석적·비판적 사고가 촉진될 수 있다.

그러나 국내외 다수의 연구들은 실제 교육 현장에서 데이터 시각화가 사고를 촉진하는 매개로 충분히 기능하지 못하고 있으며, 다음과 같은 문제가 반복된다고 지적한다. 첫째, 학습자들은 시각화 유형을 선택하는 기준에 대한 이해가 부족하고, 교과서나 틀에 제시된 예시를 기계적으로 따라 하는 경향이 강하다. 둘째, 시각화 결과에 대한 자기화 능력, 즉 “이 그래프가 무엇을 의미하는가?”에 대한 설명 능력이 부족하다. 셋째, 시각화 활동이 창의적 탐색이 아닌 정해진 절차 수행에 그치는 경우가 많다[3][4].

이러한 상황에서 데이터 시각화가 사고의 도구로 기능하기 위해서는, 학습자의 판단 과정에 개입할 수 있는 피드백 시스템이 필요하다. 특히 어떤 그래프를 선택하고, 무엇을 강조하며, 어떤 관점에서 해석할 것인가에 대한 판단은 교사나 튜터의 질문을 통해 촉진될 수 있으며, 질문을 통한 상호작용이 결여될 경우 시각화는 단지 기술적 작업으로 전략하게 된다. 본 연구가 시도하는 AI 튜터링 시스템은 단순한 지식 전달의 문제를 해소하고, 시각화 활동 자체를 사고의 촉진 도구로 재구성하고자 한다.

2.3 대화형 AI 튜터와 GPT Assistants API의 교육적 가능성

인공지능 기반 지능형 튜터링 시스템(Intelligent Tutoring System, ITS)은 1980년대부터 학생 개인에게 적절한 시점에 개입하여 안내하는 기능으로 교육 효과를 입증해 왔다. 대표적으로 CMU(2006)의 ‘Cognitive Tutor Authoring Tools (CTAT)’는 프로그래밍 지식이 없어도 시연 기반(example-tracing) 튜터를 설계할 수 있게 함으로

써, 교육 연구자들이 보다 신속하고 효율적으로 인공지능 기반 튜터를 개발할 수 있도록 지원하였으며, 실제 연구에서는 기존 방식 대비 1.4배에서 2배 빠른 개발 효율을 보여주는 등 교수학습 도구 개발의 접근성과 효과성을 크게 향상시킨 것으로 보고되었다[13]. 또한, VanLehn(2011)의 연구에 따르면, ITS는 인간 튜터와 거의 동등한 수준의 학습 효과를 보이는 것으로 나타났다[14]. 하지만 기존 ITS는 학습자가 입력한 문제를 자동 평가하고 단계별 힌트를 제공하는 등 정형화된 절차 중심 학습에 최적화되어 있어, 탐구 기반 학습이나 시각화와 같은 비정형적 과제에 개입하기에는 한계가 있다.

최근에는 GPT-3 및 GPT-4 같은 대규모 언어모델(LLM)이 ITS의 한계를 극복할 수 있는 가능성으로 부상하고 있다. 예를 들어, Baillifard et al.(2023)는 GPT-3 기반 개인 AI 튜터를 활용하여, 기존 강의 자료로부터 자동 생성된 microlearning 질문과 동적 학습 모델을 통해 학생 맞춤형 반복 학습을 제공하였고, AI 튜터의 활발한 사용이 심리학 학생들의 시험 성적을 평균 15 퍼센트 상승시키는 데 기여했다고 보고했다[15]. 특히 GPT-4를 활용한 튜터링은 열린 자연어 대화를 기반으로 사고를 유도하고, 학생의 반응 흐름에 유기적으로 개입하는 방식으로 발전하고 있다. 하지만, Bastani et al. (2024)의 무작위대조실험에 따르면, AI 튜터를 활용하면 연습 문제 수행은 향상되지만, 정작 시험에서는 AI 접근권이 없는 상황에서 복사/붙여넣기 중심의 수행으로 인해 성과가 저하된 것으로 나타났다. 이들은 결과적으로 AI를 단순 도구로 사용하기보다, 교사 개입과 설계 통합이 필수적이라고 강조하였다[16].

최근 인공지능 기술의 비약적인 발전은 학습자 맞춤형 피드백, 개별화된 학습 지원, 상호작용 기반의 문제 해결 등을 가능하게 하는 지능형 튜터링 시스템의 확산을 가속화하고 있다. 이 중에서도 OpenAI의 GPT Assistants API는 기존 GPT 모델을 교육적으로 재구성할 수 있는 기능을 제공하는 고급 인터페이스로, System Prompt를 통해 GPT의 역할을 명확히 설정하고, 특정 시나리오에 맞는 멀티턴 대화를 설계할 수 있도록 지원한다. GPT Assistants API는 일반적인 GPT 모델과 달리 System Prompt를 통해 튜터의 역할, 대화 흐름 설계, 질문·응답 스타일을 사전 정의할 수 있는 고급 인터페이스를 제공한다. 예를 들어, ‘데이터 시각화 튜터’로 설정할 시, GPT는 정답을 제시하기보다 “왜 이 시각화를 선택했는가?”, “이 결과에서 무엇을 해석할 수 있는가?” 등 사고 유도형 질문 중심으로 응답하도록 유도할 수 있다. Khan Academy는 ‘Khanmigo’라는 GPT-4 기반 Assistants API를 교육 현장에 적용하기도 하였다. 학생들은 ‘Khanmigo’에게 질문하고, ‘Khanmigo’는 단순히 답을 제시하는 것이 아니라 ‘힌트를 주며 학생 스스로 개념을 발견하도록 돕는 방식’으로 작동된다. 파일럿 테스트 결과, 학생들이 스스로 질문하는 횟수가 증가했고, 교사들은 수업 준비와 학생 분석에서 큰 도움을 받았다고 보고하였다 [17],[18].

이는 기존의 AI 튜터가 주로 정형화된 문항에 한정된 반면, 비정형적이고 개방형 탐구 활동에서도 사고를 유도하는 대화형 구조를 만들 수 있다는 점에서 교육적 활용 가능성이 매우 크다. 특히 Tableau와 같은 시각화 도구와 연계함으로써, 언어적 유도과 도구적 조작이 결합된 하이브리드 구조를 구성할 수 있으며, 이는 비전공 학습자에게도 직관적이고 실천 가능한 교육환경을 제공할 수 있다.

3. GPT 기반 데이터 시각화 튜터의 System Prompt 설계

3.1 GPT Assistants 기반 사고 촉진 튜터 설계 전략

본 연구의 설계는 단순히 기술적 구현에 그치지 않고, 학습자의 탐색적 사고와 자율적 문제 해결을 촉진하기 위한 교육적 원리를 토대로 이루어졌다. 특히 구성주의 학습 이론과 메타인지 촉진 전략은 GPT 기반 튜터 설계의 핵심 이론적 기반으로 작용하였다. 구성주의 관점에서는 학습자가 능동적으로 지식을 구성하고, 이를 위해 다양한 인지적 자극과 반응이 상호작용해야 한다고 본다[19]. 이에 따라 본 연구는 GPT 튜터를 정보 제공자가 아닌 사고 촉진자로 정의하고, 학습자의 사고 흐름을 자극하고 확장하는 대화 구조를 설계하였다. 또한 시각화 과정에서 자신의 판단을 성찰하고 조정할 수 있도록 돕는 메타인지적 개입 역시 학습 효과를 높이는 주요 설계 전략으로 반영하였다[20]. 이러한 이론적 토대는 다음의 세 가지 튜터 설계 원칙으로 구체화된다.

첫째, ‘질문 중심의 사고 유도’이다. 기존 시각화 교육은 정답을 제시하는 방식의 단방향 피드백에 치우쳐 있어, 학습자의 사고를 확장하고 주도적으로 분석 방향을 설정하는데 한계가 있다는 지적이 있어 왔다[3][4]. 이에 따라 본 연구는 GPT 튜터가 학습자에게 정답을 제공하기보다는 질문을 통해 스스로 사고를 전개할 수 있도록 유도하는 방식을 채택하였다.

둘째, ‘도구 사용 안내와 해석적 사고의 균형’이다. 시각화 교육이 도구 조작에만 집중될 경우, 결과에 대한 해석이나 의미 구성은 소홀해질 수 있다. 본 연구는 도구 사용법 안내와 더불어, 시각화 결과에 대한 학습자의 해석을 이끌어내는 인지적 개입이 병행되어야 함을 강조하며, 이를 튜터 설계의 핵심 원칙으로 설정하였다.

셋째, ‘개방형 탐색과 시각화 전략 구성’이다. 정해진 절차를 따라가는 방식보다는, 학습자가 분석 목적에 따라 변수와 조건을 스스로 탐색하고 적절한 시각화 전략을 수립하는 경험이 중요하다. 이를 통해 분석의 자율성과 문제 해결력을 함께 기를 수 있도록 설계하였다.

GPT Assistants API에서는 system 역할을 담당하는 Prompt를 통해 모델의 말투, 역할, 응답 구조 등을 사전에 정의할 수 있다. 본 연구에서는 ‘Table 1’과 같은 구성 요소로 System Prompt를 설계하였다.

Table 1. Key Components of the System Prompt Design for the DataViz GPT Tutor

Component	Design Details
Design Details Role Definition	“You are a data visualization tutor for middle and high school students. When asked how to visualize data, do not provide direct answers. Instead, guide the student with questions that promote thinking.”
Language Style	Friendly and clear tone. Use guided questions rather than statements. Complex terms should be explained in simple language.
Feedback Principles	<ul style="list-style-type: none"> - Do not suggest specific chart types directly - Use metacognitive questions such as “Why did you choose this variable?”, “What relationship are you looking for?” - For interpretation, ask “What does this mean?”, “Do you see any patterns?”
Tool Guidance	Focus on contextual feedback rather than technical commands. For example: “Using a dual axis lets you compare two variables—what would you like to compare?” Also, encourage reflection on data and AI ethics.
Dialog Flow Design	Multi-step guidance process: ① Identify data structure → ② Clarify analysis goals → ③ Suggest visualization direction → ④ Ask interpretation questions → ⑤ Prompt insight articulation
Few-shot Examples	Include 2–3 conversation samples in the prompt: User: “I don’t know what to do with this dataset.” GPT: “What columns do you have? What are you interested in exploring?” → This helps GPT follow the intended reasoning flow

본 연구에서는 Assistant의 정체성을 ‘중학생과 고등학생을 위한 데이터 시각화 튜터’로 명확히 정의함으로써, GPT가 단순한 질의응답 도우미가 아닌 학습자의 사고를 촉진하는 교육적 파트너로 기능할 수 있도록 역할을 고정하였다. 특히, 정답을 직접 제시하지 않고 질문을 통해 유도하는 방향을 명시함으로써 탐구 기반 학습에 적합한 피드백 구조를 갖추게 된다. 언어 스타일은 ‘친절함’과 ‘명료함’을 기반으로 하되, 단정적인 응답 대신 유도적인 질문을 중심으로 구성되도록 설계하였다. 복잡한 용어나 분석 개념은 학습자의 수준에 맞춰 쉽게 풀어 설명하도록 설정함으로써, 비전공 학습자도 접근 가능하도록 언어적 장벽을 최소화하였다.

GPT가 제시하는 피드백의 방향성을 명확히 하기 위해, 세 가지 원칙을 포함하였다. 첫째, 시각화 방법을 직접적으로 제시하지 않도록 제한하고, 둘째, “왜 이 변수를 선택했나요?”, “어떤 관계를 보고 싶나요?” 등의 메타인지적 질문을 통해 학습자가 스스로 분석 목적을 성찰할 수 있도록 유도한다. 셋째, 시각화 이후 결과 해석 단계에서는 “이 그래프는 어떤 의미인가요?”, “눈에 띄는 패턴이 있나요?”와 같은 질문을 통해 해석적 사고를 촉진한다.

Tool Guidance 측면에서는 Tableau와 같은 데이터 시각화 도구에 대한 안내는 기능 중심 설명보다는 학습 맥락에 맞는 안내를 우선시하였다. 예를 들어, 이중축 그래프를 단순히 ‘어떻게 만드는지’를 설명하는 대신, “이중축을 사용하

면 두 변수를 비교할 수 있어요. 지금 비교하고 싶은 항목이 있나요?”처럼 의도-기능 간 연결을 유도하는 방식이다. 또한 데이터 및 AI 윤리에 대한 질문이 들어올 경우, 관련 문서를 기반으로 사고를 촉진할 수 있도록 GPT가 응답하도록 설정하였다.

대화는 단일 질문-응답 구조가 아니라, 데이터의 구조 파악, 분석 목표 설정, 시각화 방향 탐색, 시각화 결과 해석, 인사이트 정리의 흐름으로 구성된다. 해당 구조 흐름은 학습자의 분석 사고 과정 전체를 언어적으로 지원하며, 일관된 대화 시나리오 구성을 가능하게 한다. 또한, 실제 시나리오 예시(Few-shot examples)를 System Prompt에 포함시켜, GPT가 추론 가능한 사고 유도 흐름을 학습하도록 하였다.

예를 들어 “이 데이터를 어떻게 시각화해야 할지 모르겠어요.”라는 질문에 대해, GPT는 “어떤 열이 있고, 관심 있는 항목은 무엇인가요?”와 같이 탐색을 유도하는 방식으로 반응한다. 이와 같은 예시는 GPT가 학습자의 수준에 맞춰 일관된 반응 패턴을 유지하는 데 중요한 기준이 된다. 즉, GPT가 무조건적인 정답을 제시하기보다는, 학습자의 질문에 따라 맥락 기반 사고를 유도하도록 돕는 방향으로 작동하게 만든다. 특히 데이터 분석 경험이 부족한 중·고등학생의 경우, 무엇을 시각화할 것인지, 어떻게 해석할 것인지에 대한 판단을 내리는 데 있어 적절한 안내가 매우 중요하다. 본 설계는 학습자의 탐색적 사고와 해석 능력 향상을 목적으로 한 GPT 튜터 설계의 핵심이라 할 수 있다. 또한 Tableau 기능 설명서와 AI 윤리 자료를 함께 활용하는 Retrieval 기능과 연계되어, 단순 질문응답을 넘어 문서 기반 피드백 제공도 가능하게 하였으며, 이는 향후 GPT 튜터링 시스템의 확장 가능성을 높이는 기반이 된다.

3.2 Assistants API를 사용한 DataViz GPT 튜터의 시스템 아키텍처

GPT 기반 튜터링 시스템이 교육적으로 의미 있는 상호

작용을 제공하기 위해서는 단순히 언어모델을 호출하는 수준을 넘어, 역할 기반 언어 설정, 문서 기반 정보 참조, 대화 흐름 설계, 학습자 질문의 맥락 이해 등이 종합적으로 작동하는 구조가 필요하다. 이러한 요구를 충족하기 위해 본 연구는 OpenAI에서 제공하는 Assistants API를 활용하여, DataViz GPT 튜터 시스템을 설계하였다.

Assistants API는 사용자가 GPT 모델을 단일 호출 방식으로 사용하는 기존 구조와 달리, 하나의 ‘Assistant’ 객체를 생성하여 해당 객체의 언어적 정체성, 반응 방식, 참조 문서, 사용 도구 등을 일관성 있게 정의할 수 있도록 지원하는 고도화된 인터페이스이다. 즉, 시스템적으로 ‘튜터’ 역할을 수행하는 GPT를 사전에 정의하고, 이를 중심으로 학습자의 질문 흐름을 다단계로 설계할 수 있다는 점에서, 기존 챗봇 시스템과는 본질적으로 다른 작동 방식을 지닌다.

해당 시스템의 아키텍처는 ‘Figure 1’과 같이 구성되어 있으며[21], Assistant는 총 다섯 가지 핵심 요소로, 모델(Model), 지시문(Instructions), 도구(Tools), 문서(Files), 이름 및 설명(Name/Description)으로 구성된다. 이 요소들은 학습자의 질문에 GPT가 어떻게 응답할지, 어떤 언어 스타일을 유지할지, 어떤 정보를 참조할 수 있는지, 어떤 맥락에서 기능 안내가 가능할지를 구조적으로 제어하는 역할을 한다. 우선 Assistant의 기반 모델로는 GPT-4 (gpt-4-1106-preview)가 사용되었으며, 이는 복잡한 대화 흐름을 이해하고 문맥을 장기적으로 반영할 수 있는 고도화된 언어 이해 능력을 기반으로 한다. Assistant는 학습자의 입력을 단발성 응답으로 처리하는 것이 아니라, 이를 Thread 단위로 저장하여 대화차 대화에서도 맥락을 유지하며 피드백을 이어갈 수 있도록 설계되었다. 예를 들어, 학습자가 처음에는 데이터의 열 구조에 대해 질문하고, 이어서 그중 일부를 시각화하는 방법을 묻는 경우, Assistant는 이전 대화 내용을 바탕으로 현재 질문을 연결지어 해석하고 응답할 수 있다.

Assistant의 시스템 프롬프트(System Prompt)와

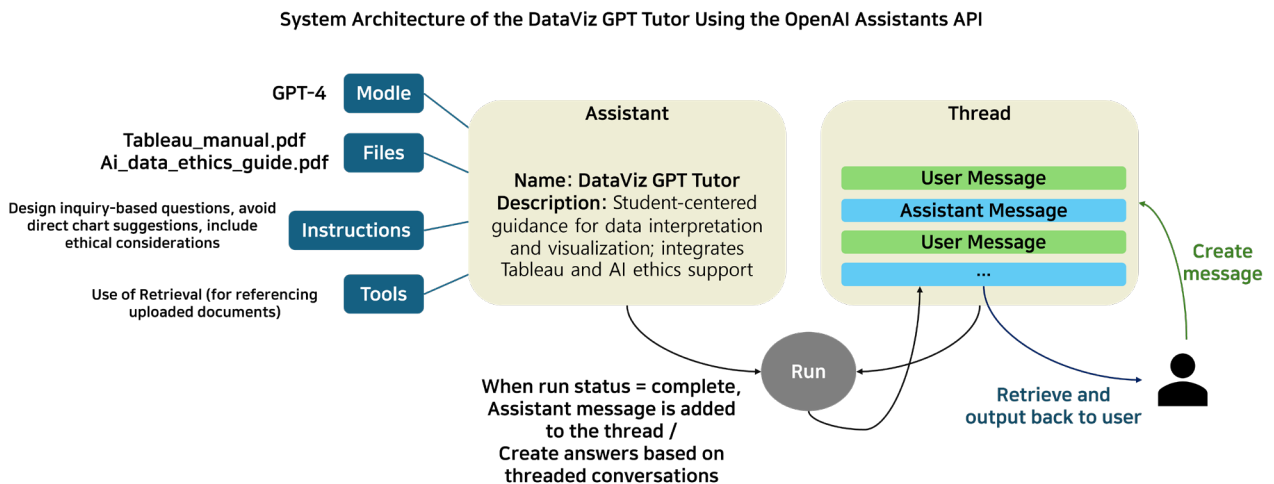


Figure 1. System Architecture of the DataViz GPT Tutor Using the OpenAI Assistants API

Instructions에는 GPT의 역할을 ‘중학생과 고등학생을 위한 데이터 시각화 튜터’로 정의하고, 응답의 언어적 특징과 피드백 방식이 구체적으로 명시되어 있다. 또한 본 시스템은 GPT의 응답이 단순히 사전 훈련된 지식에만 기반하지 않도록 하기 위해, Retrieval 도구를 활성화하고 학습 자료를 문서 형태로 업로드하여 응답의 정보적 신뢰도를 보완하였다. 구체적으로는 Tableau 사용법에 대한 상세 매뉴얼 (tableau_manual.pdf)과 인공지능 및 데이터 윤리에 대한 학습 가이드(ai_data_ethics_guide.pdf)를 파일로 업로드하였다. Retrieval 도구는 사용자의 질문에 따라 문서 내용을 실시간으로 검색하고, 그중 관련된 부분을 응답에 포함시킬 수 있도록 설계된 기능이다. 예를 들어 사용자가 “Tableau에서 이중축은 어떻게 설정하나요?”라고 질문할 경우, Assistant는 단순한 설명이 아닌, 해당 기능이 언제 효과적으로 사용되는지에 대한 교육적 맥락 중심의 안내를 제공하며, 필요 시 업로드된 설명서의 내용을 참조하여 정확한 안내를 수행한다. 마찬가지로, 데이터 수집의 윤리성이나 개인정보 활용에 대한 질문이 들어올 경우, GPT는 AI 윤리 가이드 문서에 기반한 사고 유도형 응답을 구성할 수 있다.

본 아키텍처는 학습자의 질문을 중심으로 대화형 피드백, 맥락 반영, 기능 안내, 윤리적 사고 촉진 등이 유기적으로 연결되도록 구성되어 있으며, 교육 현장에서 교사가 개별적으로 제공하기 어려운 피드백을 자동화하면서도 개별화된 방식으로 지원할 수 있다는 장점을 갖는다. 특히 단순 지식 전달을 넘어서 학습자의 탐구적 사고를 촉진하고, 데이터 분석에 대한 실질적 해석 역량을 함양할 수 있도록 돕는다는 점에서, 본 연구에서 제안한 시스템은 AI 기반 교육의 실천적 모델로서 의미를 갖는다.

3.3 실제 구성 모델 및 인터페이스 구현

DataViz GPT 튜터는 실제 사용자 환경에서 직관적인 상호작용이 가능하도록 웹 기반 인터페이스로 구현되었다. 학습자는 웹페이지 상에서 자신의 데이터를 업로드하고 질문을 입력하는 방식으로 GPT 튜터와 상호작용할 수 있으며, 시스템은 학습자의 입력에 따라 다단계 언어 피드백을 제공한다. 인터페이스는 기술적으로는 GPT Assistants API를 기반으로 작동하지만, 사용자 입장에서는 복잡한 설정 없이도 데이터를 탐색하고 해석하는 사고 과정을 자연스럽게 경험할 수 있도록 설계되었다.

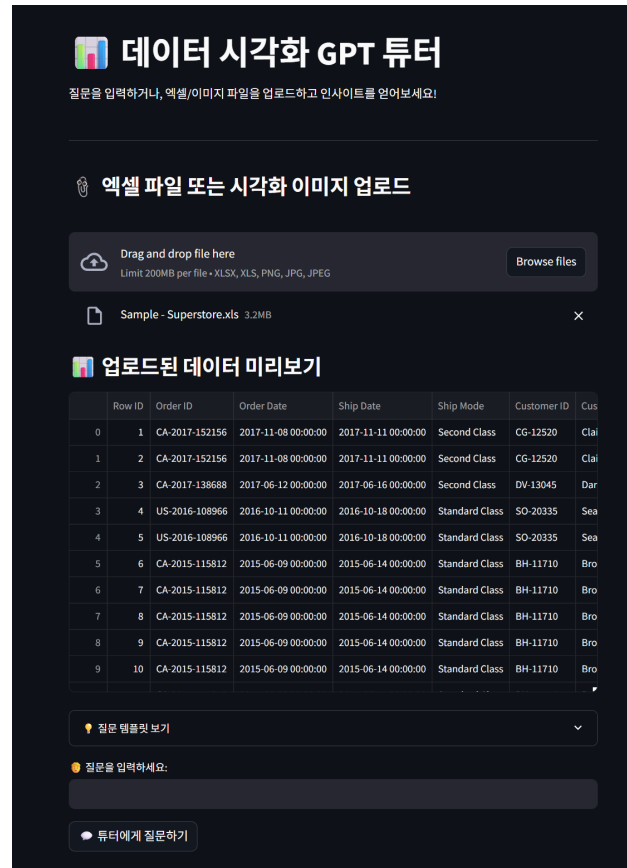


Figure 2. Data upload and input interface of the DataViz GPT Tutor; where learners can upload Excel or image files and preview their dataset in a structured table to support question formulation.

‘Figure 2’는 튜터의 초기 실행 화면으로, 학습자가 자신의 데이터 파일을 업로드하고 시각화 작업을 시작하기 위한 입력창을 제공하는 구조이다. 상단에는 ‘엑셀 파일 또는 시각화 이미지 업로드’ 영역이 배치되어 있으며, 이곳에 .xlsx, .xls, .png, .jpg 형식의 파일을 드래그 앤 드롭하거나 직접 선택하여 업로드할 수 있다. 파일 업로드가 완료되면, 해당 데이터는 자동으로 파싱되어 테이블 형태로 하단에 시각적으로 미리보기 된다. 이를 통해 학습자는 자신의 데이터가 어떤 열(column)로 구성되어 있는지, 어떤 값들이 포함되어 있는지를 직관적으로 파악할 수 있으며, 이는 이후 질문을 구성하는 데 중요한 기반 정보를 제공한다.

시각적 미리보기는 데이터 분석 경험이 부족한 학습자에게도 매우 중요한 기능으로 작동한다. 기존의 분석 도구에서는 코드를 입력하거나 복잡한 메뉴 구조를 통해 데이터를 불러와야 했던 반면, 이 시스템에서는 업로드만으로 데이터 구조를 한눈에 확인할 수 있도록 함으로써, 접근성과 분석 출발점의 진입 장벽을 크게 낮추는 효과를 거두었다. 또한 화면 하단에는 자연어로 질문을 입력할 수 있는 인터페이스가 마련되어 있으며, 사용자는 자유롭게 자신의 궁금한 점을 입력하거나, ‘질문 템플릿 보기’ 버튼을 눌러 제시된 질문 예시를 참고할 수도 있다.

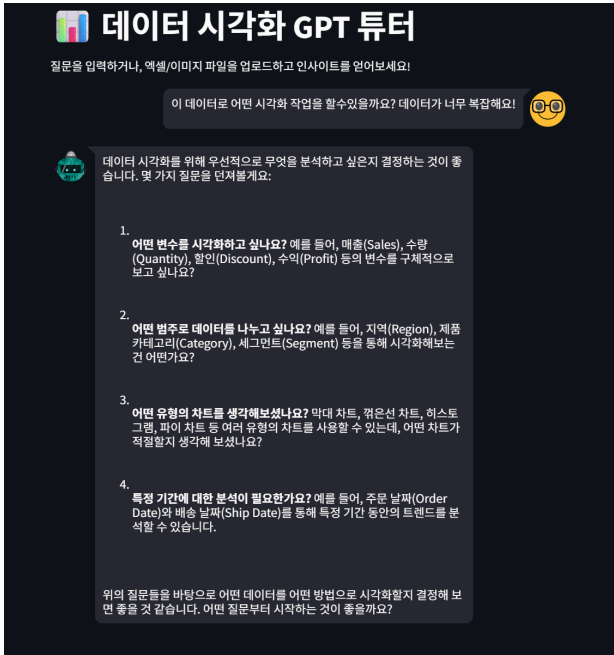


Figure 3. Example of guided questioning by the DataViz GPT Tutor
An example of the GPT tutor's guided response, showing how it prompts the learner with structured questions to support data visualization planning.

'Figure 3'은 학습자가 데이터를 업로드한 이후, "이 데이터로 어떤 시각화 작업을 할 수 있을까요? 데이터가 너무 복잡해요!"라는 질문을 GPT 튜터에게 입력했을 때의 응답 화면이다. 이 화면은 단순한 정보 제공이 아닌, 사고를 유도하기 위한 질문 중심 피드백이 어떻게 작동하는지를 보여주는 실제 예시로서, 시스템 설계 원칙과의 연결성을 명확히 드러낸다. 튜터는 먼저 분석하고 싶은 변수가 무엇인지, 데이터를 어떤 범주로 나누고 싶은지, 어떤 유형의 차트를 고려해본 적이 있는지, 특정 시간 범위에 대한 분석이 필요한지를 차례로 묻는다. 피드백은 정해진 정답이나 시각화 방식을 제시하지 않으며, 오히려 학습자의 의도를 끌어내고 이를 구체화하도록 돕는 방식으로 구성되어 있다.

해당 질문 구조는 단발성 피드백이 아니라, 학습자가 스스로 분석 목표를 명확히 하고 적절한 시각화 방향을 설정하도록 사고를 점진적으로 확장시키는 구조를 지닌다. 사용자는 GPT의 질문을 통해 자신이 주목하고 싶은 변수나 범주를 명확히 인식하게 되고, 이어지는 피드백을 바탕으로 어떤 시각화가 적절할지를 판단하게 된다. 이 과정은 단순히 기술적 도구 사용법을 익히는 차원을 넘어, 분석 목적의 명확화, 변수 선택, 시각화 전략 구성, 해석 질문, 인사이트 도출로 이어지는 사고 흐름 전개를 유도한다는 점에서 교육적으로 큰 의미를 가진다.

또한 해당 인터페이스는 비전공 학습자나 중등 수준의 사용자도 쉽게 접근할 수 있도록 설계되었다. 학습자가 데이터를 다루는 데 익숙하지 않거나, 시각화 도구의 조작에 대한 배경지식이 부족한 경우에도, 튜터가 제공하는 사고를 유도하는 질문을 통해 점차 분석 방향을 설정해 나갈 수 있

다. 실제로 사용자는 별도의 코드 작성 없이도 데이터 구조를 이해하고, 필요한 분석 목표를 자연어로 표현할 수 있으며, 튜터는 이를 바탕으로 도구 사용법 안내나 윤리적 고려 사항에 대한 안내까지 통합적으로 제공한다.

본 시스템은 GPT 언어모델의 생성 능력에 Retrieval 기능을 결합함으로써, 학습자가 시각화 기능에 대해 묻는 경우 업로드된 Tableau 설명서 문서를 기반으로 정확한 기술적 안내를 제공할 수 있다. 예를 들어, 사용자가 "이중축 그래프는 어떻게 만들어요?"라고 질문했을 때, 튜터는 단순히 조작법을 알려주는 데 그치지 않고 "두 개의 변수를 동시에 비교하고 싶은 경우 이중축을 사용할 수 있습니다. 비교하고 싶은 항목이 무엇인가요?"와 같은 사고 유도형 안내를 함께 제공한다. 이는 단순한 기능 안내에서 나아가, 기능의 개념적 의미와 학습자의 분석 목적을 연결하는 피드백 구조를 구현한 것이다.

마찬가지로, 데이터 수집, 활용, 프라이버시와 관련한 질문이 들어올 경우, 시스템은 사전에 업로드된 AI 윤리 가이드 문서를 검색하여 해당 내용을 기반으로 응답한다. 이는 데이터 시각화 과정에서도 윤리적 사고를 병행하도록 유도함으로써, 단순 분석 능력뿐 아니라 데이터 리터러시의 핵심 구성 요소인 윤리적 판단 능력까지 함께 강화하는 효과를 기대할 수 있게 한다.

결과적으로 본 시스템은 시각적 인터페이스, 자연어 입력 구조, 사고 유도형 피드백, 문서 기반 안내 기능을 통합하여, 학습자가 자신의 데이터를 탐색하고 해석하는 전 과정에서 능동적이고 반성적인 사고를 경험할 수 있도록 설계되었다. 이는 기존 분석 도구가 제공하지 못했던 교육적 기능을 구현하는 새로운 형태의 대화형 튜터 시스템으로서, 향후 실험 연구나 수업 현장 적용을 통해 효과성을 더욱 입증할 수 있을 것으로 기대된다.

4. 결론

본 연구는 데이터 시각화 교육에서 학습자의 탐색적 사고와 해석 역량을 효과적으로 지원하기 위한 방안으로, OpenAI의 GPT Assistants API를 기반으로 한 대화형 튜터 시스템인 DataViz GPT 튜터를 설계하고 구현하였다. 기존의 시각화 교육 도구들이 주로 기능 중심의 기술 습득이나 단방향 설명 방식에 머물렀던 반면, 본 연구는 학습자 스스로 분석 방향을 탐색하고 의미를 구성해나갈 수 있도록 사고 유도형 언어 인터페이스와 문서 기반 피드백 기능을 통합하여 새로운 학습 지원 방식을 제안하였다.

시스템 설계에 있어 핵심이 된 요소는 GPT 모델의 언어 생성을 단순한 응답 생성이 아닌, 의도된 교육적 목표에 따라 제어 가능한 형태로 구조화했다는 점이다. 특히 System Prompt와 Instructions를 통해 GPT의 역할을 '사고 촉진 파트너'로 정의하고, 정답 중심이 아닌 질문 중심 피드백 구조를 설계함으로써, 학습자가 데이터를 바라보는 관점과 해

석 사고를 점진적으로 확장할 수 있도록 유도하였다. 또한 Retrieval 도구를 활용하여 Tableau 설명서와 AI 윤리 문서를 참조할 수 있도록 함으로써, GPT의 응답 신뢰도와 교육적 깊이를 동시에 확보하고자 하였다.

실제 구현된 사용자 인터페이스에서는 데이터 업로드, 미리보기, 자연어 질문 입력, GPT 튜터 응답까지의 전 과정을 직관적이고 간편하게 구성하였으며, 사고 흐름을 다단계 언어 피드백을 통해 자연스럽게 유도하는 기능을 구현하였다. 이를 통해 학습자는 시각화 도구 사용에 대한 부담 없이 분석 목적을 명확히 설정하고, GPT와의 대화를 통해 스스로 인사이트를 구성해볼 수 있는 경험을 할 수 있다.

본 연구에서 개발한 대화형 튜터는 특히 GPT 모델의 언어 생성을 학습 목표에 맞게 제어하기 위한 System Prompt 설계를 통해, 정답 제공 중심의 응답이 아닌 질문 중심의 사고 촉진형 대화가 가능함을 확인하였다. 이는 AI가 학습자의 주도적 사고를 유도하는 교육적 파트너로 기능할 수 있다는 가능성을 보여준다. 다만 본 연구는 실제 수업 적용을 통한 학습 효과 분석이나 정량적 성과 측정은 향후 연구 과제로 남아 있다. 추후 연구에서는 본 시스템을 중등 교육 현장이나 실험 환경에 적용하고, 학습자의 대화 로그, 시각화 결과, 인사이트 도출 과정 등을 기반으로 GPT 튜터의 반응 품질과 사고 유도 효과를 정량적으로 검증할 필요가 있다.

참고문헌

- [1] Echtenbruck, M. M., Fühles-Ubach, S., Naujoks, B., & Kaliva, E. (2025). A data literacy competence model for higher education and research. *arXiv*. <https://doi.org/10.48550/arXiv.2504.15690>
- [2] OECD. (2019). *OECD Future of Education and Skills 2030: OECD Learning Compass 2030 – A series of concept notes*. OECD Publishing.
- [3] Kim, H.-N., & Kim, S.-H. (2023). Analysis of user preference of data visualization: Focusing on distribution visualization methods. *Journal of Digital Contents Society*, 24(1), 129–138. <https://doi.org/10.9728/dcs.2023.24.1.129>
- [4] Hyun, E.-R., Kim, J.-H., & Shin, T.-S. (2022). Development of an infographic-based STEAM program. *Journal of Engineering Education Research*, 25(5), 68–74. <https://doi.org/10.18108/jeer.2022.25.5.68>
- [5] Baillifard, A., Gabella, M., Lavenex, P. B., & Martarelli, C. S. (2024). Effective learning with a personal AI tutor: A case study. *Education and Information Technologies*, 30(1), 297–312. <https://doi.org/10.1007/s10639-024-12888-5>
- [6] OpenAI. (n.d.). *Assistants API overview*. OpenAI. <https://platform.openai.com/docs/assistants/overview>
- [7] Ridsdale, C., Rothwell, J., Smit, M., Bliemel, M., Irvine, D., Kelley, D., Matwin, S., Wuetherick, B., & Ali-Hassan, H. (2015). *Strategies and best practices for data literacy education: Knowledge synthesis report*. Dalhousie University. <https://doi.org/10.13140/RG.2.1.1922.5044>
- [8] Witte, V., Schwering, A., & Frischmeier, D. (2025). Strengthening data literacy in K–12 education: A scoping review. *Education Sciences*, 15(1), 25. <https://doi.org/10.3390/educsci15010025>
- [9] Hasim, S. M., Rosli, R., & Halim, L. (2024). A quantitative case study of secondary school students' level of statistical thinking. *Eurasia Journal of Mathematics, Science and Technology Education*, 20(4). <https://doi.org/10.29333/ejmste/14358>
- [10] Song, Y., Song, S., Kim, Y., & Lim, C. (2021). A developmental study of an instructional model and strategies for data-driven debate (DDD) to improve data literacy. *Journal of Educational Technology*, 37(4), 943–982. <https://doi.org/10.17232/KSET.37.4.943>
- [11] Joo, A., & Kim, M. R. (2024). Effects of discussion classes using data visualization materials on data literacy of elementary school students. *Journal of the Korean Association of Computer Education*, 27(2), 37–47. <https://doi.org/10.32431/kace.2024.27.2.004>
- [12] Chang, H.-Y., Chang, Y.-J., & Tsai, M.-J. (2024). Strategies and difficulties during students' construction of data visualizations. *International Journal of STEM Education*, 11(1), 11. <https://doi.org/10.1186/s40594-024-00463-w>
- [13] Aleven, V., McLaren, B. M., Sewall, J., & Koedinger, K. R. (2006). The Cognitive Tutor Authoring Tools (CTAT): Preliminary evaluation of efficiency gains. *Intelligent Tutoring Systems: 8th International Conference, ITS 2006*, Jhongli, Taiwan, 4053, 61–70. https://doi.org/10.1007/11774303_7

- [14] VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, 46(4), 197–221. <https://doi.org/10.1080/00461520.2011.611369>
- [15] Baillifard, A., Gabella, M., Banta Lavenex, P., & Martarelli, C. S. (2023). Implementing learning principles with a personal AI tutor: A case study. *arXiv*. <https://doi.org/10.48550/arXiv.2309.13060>
- [16] Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakci, Ö., & Mariman, R. (2024, July 15). *Generative AI can harm learning*. The Wharton School Research Paper. SSRN. <https://doi.org/10.2139/ssrn.4895486>
- [17] Khan Academy. (2023). *Why Khan Academy's Khanmigo is a promising new tool to drive student learning*. <https://annualreport.khanacademy.org/khanmigo>
- [18] Shetye, S. (2024). An evaluation of Khanmigo, a generative AI tool, as a computer-assisted language learning app. *Studies in Applied Linguistics and TESOL*, 24(1). <https://doi.org/10.52214/saltv24i1.12869>
- [19] Bruner, J. (1996). *The Culture of Education*. Harvard University Press. <https://doi.org/10.2307/j.ctv136c601>
- [20] Azevedo, R., & Cromley, J. G. (2004). Does Training on Self-Regulated Learning Facilitate Students' Learning With Hypermedia?. *Journal of Educational Psychology*, 96(3), 523–535. <https://doi.org/10.1037/0022-0663.96.3.523>
- [21] Horsey, J. (2023, November 13). *Learn how to use the OpenAI Assistants API to build ChatGPT AI agents*. Geeky Gadgets. <https://www.geeky-gadgets.com/learn-how-to-use-the-openai-assistants-api-to-build-chatgpt-ai-agents/>



이정훈

· 2018년 제주대학교 전자공학과(공학사)
· 2020년 제주대학교 컴퓨터교육전공 (교육학석사)
· 2025년 제주대학교 컴퓨터교육전공 (교육학박사)
· 2025년~현재 한국교육과정평가원 부연구위원

✚ 관심분야: 소프트웨어 교육, 인공지능 교육, 지능형 시스템

✉ 2ehdrks@gmail.com



양수경

· 2023년 서울사이버대학교 컴퓨터공학과(공학사)
· 2023년~현재 제주대학교 컴퓨터교육전공 석사과정
· 2023년~현재 제주대학교 지능소프트웨어교육 연구소 연구원

✚ 관심분야: 컴퓨터(정보, AI)교육, 지능정보윤리

✉ sugyeongyang1@gmail.com



조정원

· 2004년~현재 제주대학교 컴퓨터교육과 교수
· 2020년~현재 한국컴퓨터교육학회 부회장, 논문지 편집위원장
· 2012년~현재 한국정보과학회 부회장, 교육위원회 위원장
· 2018년~현재 제주대학교 지능소프트웨어교육 연구소 연구소장

✚ 관심분야: 컴퓨터(정보, AI)교육, 지능정보윤리, 지능형시스템, 멀티미디어

✉ jwcho@jejunu.ac.kr