

컴퓨터교육학회 논문지 2026년 제29권 제8호
<https://doi.org/10.32431/kace.2026.29.8.009>



AI 역량 측정 도구의 구성 특성과 타당화 동향에 대한 체계적 문헌 고찰

A Systematic Review of the Structural Characteristics and Validation Trends of AI Competency Measurement Tools

조민국[†] · 김귀훈^{††}
 Minkuk Cho[†] · Kwihoon Kim^{††}

요약

본 연구는 국내 AI 교육 연구에서 활용된 AI 역량 측정도구의 구성 특성과 타당화 동향을 종합적으로 파악하고, 향후 측정도구 개발 및 타당화 연구를 위한 시사점을 도출하는 것을 목적으로 한다. 이를 위해 2020년부터 2025년까지 발표된 국내 학술지 논문 32편을 대상으로 문헌 고찰을 수행하였다. 분석 결과, 관련 연구는 2023년 이후 증가하는 경향을 보였으며, 교사와 예비교사를 중심으로 이루어지고 있었다. 측정 도구는 기능 및 가치 영역의 비중이 높게 나타났으며, 타당도 검증에서는 내용타당도와 요인분석이 주로 활용되었다. 이를 바탕으로 학습자 중심의 측정도구 개발과 통합적 역량 평가, 체계적인 타당화 절차의 필요성이 도출되었다.

주제어 인공지능 교육, AI 역량, 역량 측정도구, 체계적 문헌 고찰, 타당도

ABSTRACT

This study aims to comprehensively examine the structural characteristics and validation trends of AI competency measurement tools used in domestic AI education research, and to derive implications for future measurement tool development and validation studies. To this end, a systematic literature review was conducted on 32 domestic journal articles published between 2020 and 2025. The results showed that related studies have increased since 2023, with a strong focus on teachers and pre-service teachers. The measurement tools emphasized skills and values domains, and content validity and factor analysis were the most commonly used validation methods. Based on these findings, the need for learner-centered measurement tool development, integrated competency assessment, and more systematic validation procedures was identified.

Keywords AI Education, AI Competency, Measurement Tools, Systematic Literature Review, Validity and Reliability

†정회원	한국교원대학교 일반대학원 초등컴퓨터 교육 전공 박사과정
††중신회원	한국교원대학교 인공지능융합교육전공 부교수(교신저자)
논문투고	2026년 05월 29일
심사완료	2026년 07월 07일
게재확정	2026년 07월 13일
발행일자	2026년 08월 27일

1. 서론

최근 AI 기술의 발달로 인해 교육 분야에서도 인공지능 활용과 AI 리터러시의 중요성이 빠르게 커지고 있다. 이러한 변화는 사회 전반의 구조적 변화를 가져오고 있으며, 미래 사회를 살아갈 학생들에게 요구되는 역량 역시 변화하고 있다. 특히 인공지능 기술은 다양한 분야에서 활용되며 사회와 산업의 패러다임을 변화시키고, 이에 따라 인공지능을 이해하고 활용할 수 있는 역량은 미래 사회의 핵심 역량으로 강조되고 있다.

이와 같은 변화에 대응하기 위해 여러 국가에서는 인공지능 교육을 학교 교육과정에 도입하고 있으며 국내에서도 초·중등 교육과정에서 소프트웨어 교육과 인공지능 교육을 강화하고 있다. 2015 개정 교육과정에서는 초등학교에서 SW 교육을 의무화하며 정규 교육과정 내에서 17차시의 SW 교육을 실시하였다[1]. 나아가 2022 개정 교육과정에서는 디지털 및 인공지능 소양을 체계적으로 기를 수 있도록 실과와 정보 영역 시수와 학교 자율시간 등을 활용하여 34차시 이상 편성·운영하도록 하였다[2].

이에 따라 인공지능 교육과 관련된 다양한 연구가 활발하게 이루어지고 있다. 기존 연구들은 인공지능 교육 프로그램 개발, 교수·학습 방법 탐색, 교육 효과 분석 등 다양한 측면에서 진행되고 있으며 학습자의 컴퓨팅 사고력, 문제 해결력, AI 리터러시 등의 역량 향상을 검증하는 연구가 증가하고 있다. 또한 인공지능 교육을 다양한 교과와 연계하여 적용하거나 프로젝트 기반 학습, 문제 기반 학습 등 다양한 교수·학습 방법을 적용한 연구들도 꾸준히 수행되고 있다.

이러한 연구들의 교육 효과를 검증하기 위해서는 신뢰성과 타당성을 갖춘 측정 도구가 필수적으로 요구된다. 측정 도구는 학습자의 역량 변화와 교육 효과를 객관적으로 분석하는 데 중요한 역할을 하며, 연구 결과의 신뢰성을 결정하는 핵심 요소이다. 특히 인공지능 교육 연구에서는 학습자의 컴퓨팅 사고력, 문제 해결력, AI 이해도, 태도, 흥미 등 다양한 변인을 측정하기 위한 도구가 활용되고 있다[3, 4].

하지만 현재 인공지능 교육 연구에서 활용되는 측정 도구는 연구마다 다양한 형태로 사용되고 있으며, 동일한 역량을 측정하더라도 서로 다른 측정 도구가 사용되는 경우가 많다. 또한 측정 도구의 구성 요소나 측정 영역에 대한 체계적인 분석이 부족하여 인공지능 교육 연구에서 활용되는 평가 도구의 특성을 종합적으로 이해하기 어려운 한계가 있다. 이러한 상황은 향후 인공지능 교육 연구에서 적절한 측정 도구를 선택하거나 새로운 평가 도구를 개발하는 데 어려움을 줄 수 있다.

기존 연구들은 AI의 교육적 활용 사례를 분석하거나 AI 리터러시 개념 및 연구 동향을 살펴보는 데 주로 초점을 두었다. 이러한 연구들은 개별 측정 도구의 개발과 타당화, 혹은 AI 리터러시나 AI 윤리 등 특정 영역에 한정되어 있어, 국내에서 활용된 AI 역량 측정 도구 전반의 구성 특성과 타당화 절차를 종합적으로 검토한 연구는 찾아보기 어렵다.

측정 변인과 하위 요인이 어떻게 구성되어 있는지, 내용타당도와 구인타당도, 수렴 및 변별타당도, 신뢰도는 어느 정도 수준으로 검증되어 왔는지를 비교한 경우도 찾아보기 힘들다. 이에 본 연구는 개별 도구의 개발 연구를 넘어 AI 역량 측정 도구와 역량 요소를 종합적으로 분석하고, 측정 도구의 구조적 특징과 타당도·신뢰도 검증 방식까지 함께 검토함으로써 기존 연구와 차별성을 가진다. 이러한 맥락에서 국내 AI 교육 연구에서 활용된 측정 도구를 종합적으로 분석하여 그 특성과 검증 수준을 체계적으로 파악할 필요가 있다.

따라서 본 연구는 국내 인공지능 교육 연구에서 활용된 측정 도구를 체계적으로 분석하여 그 유형과 특징을 파악하고자 한다. 이를 위해 국내 학술지에 게재된 인공지능 교육 관련 연구를 대상으로 측정 도구의 종류, 측정 영역, 문항 구성 및 연구 동향을 분석하였다. 이를 통해 인공지능 교육 연구에서 활용되는 측정 도구의 현황을 파악하고, 향후 평가 도구의 개발 및 활용 방향에 대한 시사점을 제시하고자 한다.

2. 이론적 배경

2.1 인공지능 교육

인공지능 교육은 2015 개정 교육과정에서 도입된 소프트웨어 교육을 기반으로 발전한 교육 영역으로, 학생들이 인공지능의 개념과 작동 원리를 이해하고 이를 활용하여 다양한 문제를 해결할 수 있는 역량을 기르는 것을 목적으로 한다. 초등학교 교육과정에서는 인공지능 교육의 내용을 인공지능의 이해, 인공지능의 원리와 활용, 인공지능의 사회적 영향의 세 영역으로 구분하여 제시하고 있다[5].

인공지능 기술은 소프트웨어와 하드웨어 기술의 결합을 기반으로 작동하며, 이러한 기술은 일상생활의 다양한 문제를 해결하는 과정과 밀접하게 관련되어 있다. 초등학교 실과 교육과정에서는 ‘디지털 사회와 인공지능’ 영역을 통해 컴퓨터의 개념, 문제 해결 절차, 프로그래밍, 데이터의 표현, 생활 속 인공지능 등을 다루며, 컴퓨팅 사고를 바탕으로 실생활 문제를 해결하는 역량을 기르는 것을 목표로 한다[2].

인공지능은 특정 교과의 지식에 국한되지 않고 사회, 경제, 문화 등 다양한 영역에서 활용되고 있다는 특성을 가진다. 이러한 특성을 고려할 때 특정 교과 중심으로 인공지능 교육을 운영하는 것에는 한계가 있을 수 있으며, 다양한 교과와 연계하여 융합적으로 접근하는 교육이 필요하다[6]. 따라서 초등학교 AI 교육은 실과 교과를 중심으로 삼아 프로그래밍, 데이터 활용, 생활 속 AI 사례 탐색 등의 학습 경험을 제공하되, 이에 그치지 않고 다양한 교과와의 융합을 통해 보다 넓은 맥락에서 AI 교육을 확장해 나가는 것이 필요하다.

2.2 AI 역량

인공지능 교육 연구에서 다루어지는 AI 역량은 단순히 AI 기술을 다루는 능력을 넘어, 지식, 기능, 가치, 태도가 함

게 포함된 종합적인 개념으로 볼 수 있다. 이러한 역량은 인공지능의 원리와 작동 방식을 이해하고, 이를 다양한 상황에서 적절하게 활용하는 능력을 포함한다. 또한 인공지능이 사회와 인간의 삶에 미치는 영향을 이해하고, 이를 책임감 있게 활용하려는 태도 역시 중요한 요소로 제시된다. 따라서 AI 역량은 인공지능에 대한 이해와 활용뿐 아니라, 기술을 받아들이고 실제로 적용하는 능력까지 포함하는 포괄적인 역량으로 이해될 수 있다[7, 8].

학습자의 AI 역량은 인공지능 이해, 사고, 활용, 윤리, 사회·정서 등 여러 영역으로 구성될 수 있다. 특히 인공지능의 개념과 원리, 데이터의 구조와 활용 방식을 이해하는 지식과, 문제 해결 과정에서 알고리즘을 설계하고 데이터를 활용하는 사고 및 활용 능력이 중요하게 강조된다. 또한 인공지능이 사회에 미치는 영향과 윤리적 문제를 이해하고, 책임 있게 활용하려는 태도와 가치 역시 중요한 요소로 포함된다. 이러한 요소들은 지식, 기능, 태도가 함께 어우러진 형태로 나타나며, 학습자가 인공지능을 이해하고 활용하면서 사회적 맥락 속에서 비판적으로 판단할 수 있는 능력을 의미한다[9, 10].

AI 역량은 학습자뿐만 아니라 수업을 설계하고 실행하는 교사에게도 중요한 전문 역량이다. 교사의 AI 역량은 인공지능 기술을 이해하고, 이를 교수·학습 과정에 효과적으로 적용할 수 있는 능력을 의미한다. 특히 인공지능을 활용한 수업을 설계하고, 교육자료를 만들며, 수업을 운영하고 평가하는 전 과정에서 인공지능을 적절하게 활용하는 역량이 중요하게 강조된다. 따라서 인공지능 교육을 효과적으로 실행하기 위해서는 교사의 인공지능 이해와 수업 실행 역량을 함께 고려할 필요가 있다[11]. AI 역량은 지식, 기능, 태도, 가치 등 여러 요소가 복합적으로 작용하는 개념이므로, 그 구성 요소를 보다 체계적으로 살펴볼 필요가 있으며, OECD Learning Compass 2030이 제시하는 지식, 기능, 태도, 가치의 네 범주와 맞닿아 있다. AI의 개념과 원리, 데이터 구조에 대한 이해는 지식 영역에 해당하고, AI 도구를 활용한 문제 해결이나 교수·학습 설계 및 실행 능력은 기능 영역에 속한다. AI에 대한 정서적 반응이나 자기효능감은 태도 영역으로, AI 윤리와 사회적 책임에 대한 인식은 가치 영역으로 볼 수 있다. 이에 본 연구에서는 이 네 범주를 분석틀로 삼아, 분석 대상 문헌에서 도출된 측정 변인과 하위 요인을 재분류함으로써 AI 역량 측정도구의 구조적 특징을 살펴보았다.

2.3 AI 역량 측정 도구의 필요성

인공지능 기술의 발전과 함께 교육 현장에서 인공지능 교육의 중요성이 점차 증가하고 있으며, 이에 따라 학습자와 교사의 AI 역량을 객관적으로 측정할 필요성이 지속적으로 제기되고 있다[12, 13]. 특히 역량 기반 교육에서는 교육 프로그램의 효과를 체계적으로 검증하고 학습자의 역량 수준과 변화를 파악하기 위해 신뢰성과 타당성을 갖춘 측정 도구가 필수적으로 요구된다[7]. 실제로 인공지능 교육 연구에서는 학습자나 교사의 AI 역량을 보다 체계적으로 파악하

기 위해 다양한 역량 측정 척도와 평가 도구를 개발하고 이를 검증하는 연구들이 활발하게 수행되고 있다[14]. 이러한 측정도구는 AI 역량 수준을 진단하고 교육 요구도를 분석하며, 교육 프로그램의 효과와 학습 성과를 평가하는 데 중요한 역할을 수행한다[9]. 따라서 인공지능 교육 연구에서는 학습자와 교사의 AI 역량을 보다 정확하고 체계적으로 측정할 수 있는 평가 도구의 개발과 활용이 중요한 연구 과제로 강조되고 있다[7].

2.4 측정도구의 타당도 및 신뢰도

측정도구의 질적 수준을 평가하기 위해서는 타당도와 신뢰도에 대한 체계적인 검증이 필요하다. 내용타당도는 측정도구의 문항이 측정하고자 하는 구성개념의 내용을 얼마나 적절히 대표하는지를 의미하며, 주로 전문가의 검토나 델파이 조사를 통해 확보되고, 내용타당도비율(CVR)이나 항목별 내용타당도지수(I-CVI) 등의 지수로 정량화된다[15]. 구인타당도는 측정도구가 연구자가 의도한 이론적 요인 구조를 실제로 반영하는지를 확인하는 것으로, 탐색적 요인분석(EFA)을 통해 초기 요인 구조를 도출하고, 확인적 요인분석(CFA)을 통해 모형의 적합도(CFI, TLI, RMSEA, SRMR 등)를 검증하는 절차로 이루어진다[16]. 수렴타당도는 동일한 구성개념을 측정하는 문항들이 서로 높은 상관을 보이는지를 의미하며, 평균분산추출(AVE)이나 개념신뢰도(CR) 등을 통해 확인한다[17]. 변별타당도는 서로 다른 구성개념을 측정하는 요인 간에 충분한 차이가 있는지를 의미하며, AVE와 요인 간 상관계수의 비교 또는 HTMT 지수 등을 통해 검증한다[17]. 신뢰도는 측정도구의 내적 일관성을 의미하며, 일반적으로 Cronbach's α 계수를 통해 확인한다. 이러한 일련의 검증 절차는 측정도구가 의도한 구성개념을 정확하고 일관되게 측정하고 있는지를 판단하는 근거가 되며, 본 연구에서는 이러한 개념을 기준으로 분석 대상 문헌의 타당화 수준을 검토하였다.

3. 연구 방법

본 연구는 인공지능 교육 연구에서 활용된 AI 역량 측정도구와 역량 요소를 체계적으로 분석하고, 이를 바탕으로 향후 연구를 위한 시사점을 도출하는 것을 목적으로 한다. 이를 위해 체계적 문헌 고찰(SLR) 방법을 적용하여 관련 연구를 수집·선정하고, 측정도구의 유형, 구성 요소, 적용 대상, 신뢰도와 타당도를 중심으로 분석하였다[18]. 또한 문헌 검색과 선정 과정의 체계성과 투명성을 확보하기 위해 PRISMA 절차를 준용하였다.

SLR 절차는 일반적으로 계획(Planning), 수행(Conducting), 보고(Reporting)의 단계로 이루어지며, 본 연구에서는 이를 연구 문제 설정, 문헌 검색 및 선정, 문헌 분석 기준 수립, 자료 분석 및 결과 종합의 순서로 구체화하여 진행하였다.



Figure 1. PRISMA flow diagram of the literature selection procedure

3.1 연구 문제

본 연구는 인공지능 교육 연구에서 활용된 AI 역량 측정 도구와 역량 요소의 유형과 특성을 체계적으로 분석하기 위해 다음과 같은 연구 문제를 설정하였다.

첫째, 국내 인공지능 교육 연구에서 활용된 AI 역량 측정 도구 및 역량 요소의 전반적인 활용 현황은 어떠한가?

둘째, AI 역량 측정 도구는 어떠한 측정 영역과 하위 구성 요소로 이루어져 있으며, 역량 요소는 어떻게 도출되고 구성되어 있는가?

셋째, AI 역량 측정 도구의 신뢰도와 타당도는 어떠한 방식으로 검증되고 제시되고 있는가?

3.2 문헌 검색 및 선정

본 연구에서는 문헌 검색과 선정의 체계성을 확보하기 위해 명확한 기준과 절차를 설정하였으며 문헌 검색은 국내 학술 데이터베이스인 KCI, RISS, DBpia를 중심으로 수행하였다. 검색어는 인공지능 교육에서 활용되는 역량 측정 도구와 타당화 연구를 포괄적으로 탐색하기 위해 선행 연구에서 반복적으로 사용된 핵심 용어를 중심으로 선정하였으며, “AI 역량”, “AI 리터러시”, “측정 도구”, “타당화” 등의 검색어를 조합하여 관련 문헌을 탐색하였다. 또한 인공지능 교육이 본격적으로 활성화된 시기를 고려하여 2020년 1월부터 2025년 12월까지 발표된 연구를 대상으로 하였다.

문헌 선정은 다음과 같은 절차로 이루어졌다. 먼저 데이터베이스 검색을 통해 총 811편의 문헌을 수집하였다. 이후 제목과 초록을 검토하여 본 연구의 주제인 AI 역량·리터러시 측정도구의 개발 및 타당화와 직접 관련이 없는 문헌 728편을 제외하고 83편을 1차로 선별하였다. 이 중 동일한 문헌 37편을 제거하여 원문 검토 대상으로 46편을 확정하였다. 46편에 대한 원문 검토를 실시하여 측정도구의 구체적인 문항이나 타당화 절차가 제시되지 않은 문헌, 학위논문과 중복 출판된 학술지 논문, 본 연구의 분석 기준에 부합하지 않는 문헌 등 14편을 추가로 제외함으로써 최종적으로 32편이 분석 대상으로 선정되었다.

3.3 문헌 분석 기준

연구의 문헌 분석 기준은 임수훈 외(2025)의 분류 체계를 참고하되, 본 연구의 목적과 대상에 적합하도록 일부 항목을 수정·보완하여 구성하였다[19]. 최종 선정된 문헌을 대

상으로 크게 ‘일반적 현황’, ‘측정도구 구성 특성’, ‘측정 도구 타당도 및 신뢰도’의 세 가지 범주로 구분하여 분석 틀을 구성하였다. 문헌 분석 과정의 신뢰성과 일관성을 확보하기 위해 분석 기준을 설정하고 단계적으로 문헌 분석을 수행하였다. 먼저 일부 문헌을 대상으로 예비 분석을 실시하여 분석 범주와 적용 기준의 적절성을 점검하였으며, 이후 전체 문헌에 대한 본 분석을 수행하였다. 문헌 분석 과정에서는 연구 목적과의 적합성을 지속적으로 검토하며 동일한 기준을 일관되게 적용하고자 하였다. 문헌 분석의 신뢰성을 확보하기 위해, 연구자 1인이 분석 대상 32편 전체에 대해 측정 변인, 하위 요인 분류, 타당도 및 신뢰도 검증 방식 등을 1차로 분석하였다. 이후 10년 이상의 교육 경력을 바탕으로 AI 융합교육 및 컴퓨터교육 분야에서 관련 연구를 수행해 온 석·박사 과정 전문가 10명을 대상으로 분석 결과 전반에 대해 검토를 하였다. 특히 OECD Learning Compass 2030 기반의 하위 요인 재분류 결과에 대해 의견 불일치가 있는 항목을 중심으로 논의하여 분류 기준을 수정·보완하였다. 이를 통해 문헌 분석 결과의 일관성과 객관성을 높이고자 하였다.

첫째, ‘일반적 현황’은 연구의 전반적인 흐름과 동향을 파악하기 위한 기준으로, 발행 연도, 연구 목적, 연구 대상의 유형 등을 포함하였다. 이를 통해 인공지능 교육에서 AI 역량 측정도구 및 역량 요소 관련 연구가 어떠한 시기에 집중적으로 이루어졌는지와 연구 대상의 전반적인 분포를 확인하였다.

둘째, ‘측정 도구 구성 특성’은 각 연구에서 개발 또는 활용된 측정 도구의 구조를 분석하기 위한 기준으로, 측정 변인, 하위 요인 구성을 포함하였다. 이를 통해 측정 도구의 구성 방식과 특징을 비교·분석하였다.

셋째, ‘측정 도구의 타당도 및 신뢰도’는 도구의 질적 수준을 분석하기 위한 기준으로, 탐색적 요인분석, 확인적 요인분석, 내용타당도 검증, 신뢰도 계수 등의 통계적 검증 여부와 방법을 포함하였다. 이를 통해 측정도구의 타당화 수준과 연구의 엄밀성을 확인하고자 하였다.

Table 1. Criteria and Detailed Framework for Literature Analysis of AI Competency Measurement Tools

Category	Sub-analysis Items	Analysis Content and Purpose
General Trends	-Publication year -research participants -research purpose	To examine the overall trends of studies on AI competency measurement tools and identify temporal patterns and the distribution of research participants
Characteristics of Measurement Tools	-Measurement variables, -sub-factor composition	To analyze the types of measurement variables (cognitive, performance, affective, value & ethics) and the structural characteristics of the tools

Category	Sub-analysis Items	Analysis Content and Purpose
Validity and Reliability of Measurement Tools	-Validation methods -reliability coefficients -analytical techniques (EFA, CFA)	To analyze how validity and reliability were established and to evaluate the quality and rigor of the measurement tools

4. 연구 결과

4.1 일반적 현황

4.1.1 연도별 발행 추이

발행 연도별 분포를 분석한 결과, 2020년 1편, 2021년 2편, 2022년 1편, 2023년 9편, 2024년 12편, 2025년 7편으로 나타났다. 초기인 2020년과 2021년에는 관련 연구가 소수에 그쳤으나, 2022년 이후부터 연구 수가 점차 증가하는 경향을 보였다. 특히 2023년과 2024년에 연구가 급격히 증가하여 가장 활발하게 이루어진 것으로 확인되었다. 이러한 경향은 2022 개정 교육과정에서 디지털 및 인공지능 교육이 강조된 점과 생성형 AI 확산, AI 디지털교과서 도입 논의 등의 영향과 관련이 있는 것으로 해석할 수 있다. 이에 따라 교사와 학생의 AI 역량을 측정하고자 하는 연구의 필요성 또한 증가한 것으로 보인다.

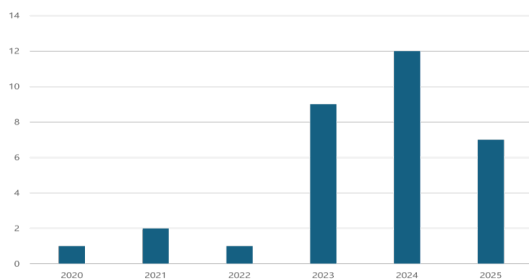


Figure 2. Number of Papers by Year

4.1.2 연구 대상별 분포

연구 대상별 분포를 분석한 결과, 본 연구에서 분석한 32편의 논문은 다양한 대상 집단을 포함하고 있었으며, 분석 대상 중 1편의 논문에서 교사와 예비교사를 동시에 연구 대상으로 삼아 총 연구 대상 수는 33개로 집계되었다. 연구 대상별로 살펴보면, 교사를 대상으로 한 연구가 10편으로 가장 높은 비중을 차지하였으며, 예비 교사를 대상으로 한 연구가 8편으로 그 뒤를 이었다. 이는 인공지능 교육 관련 측정 도구 연구에서 교사와 예비 교사가 주요 연구 대상으로 활용되고 있음을 보여준다. 한편, 측정 도구 개발 초기 단계에서 델파이 조사나 전문가 검토만으로 역량 요소를 도출하거나 내용 타당도를 확보한 연구가 6편으로 나타났는데, 이는 실제 조사 대상이 일반 응답자가 아닌 전문가 패널로 한

정되어 있었기 때문이다. 대학생을 대상으로 한 연구는 5편으로 나타났으며, 일반 성인을 대상으로 한 연구와 초·중등 학생을 포함한 일반 학생을 대상으로 한 연구는 각각 2편으로 나타나, 다른 대상에 비해 상대적으로 낮은 비율을 보였다. 전체적으로 연구 대상은 교사와 예비 교사에 집중되어 있으며, 전문가, 대학생, 성인, 학생 순으로 분포하는 경향이 나타났다. 교사와 예비교사를 대상으로 한 연구가 가장 많은 것은, 2022 개정 교육과정에서 교사의 AI·디지털 역량이 인공지능 교육 실행의 핵심 변수로 강조되면서, 교사 연구 및 양성 과정에서 활용할 진단도구의 수요가 먼저 발생했기 때문으로 해석할 수 있다. 반면 초·중등 학생을 대상으로 한 측정도구 연구가 상대적으로 부족한 것은, 학생용 도구 개발에 앞서 교사의 역량을 진단하고 교육 환경을 정비하려는 정책적 흐름이 우선적으로 반영된 결과로 볼 수 있다.

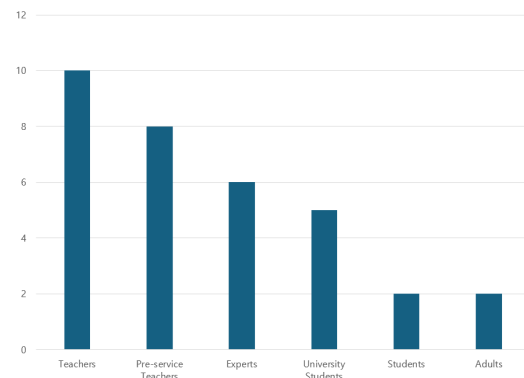


Figure 3. Distribution of Research Participants

4.1.3 연구 목적별 분석

연구 목적에 따라 분석한 결과 전체 32편의 논문 중 29편의 연구는 측정 도구의 개발 및 타당화를 주요 목적으로 수행된 것으로 나타났으며, 인공지능 리터러시, AI 역량, 태도, 윤리 의식 등 다양한 변인을 측정하기 위한 도구를 개발하고 도구의 타당도와 신뢰도를 검증하는데 초점을 두고 있었다.

반면, 3편의 연구는 델파이 조사를 기반으로 역량 요소를 도출하는 데 목적을 두었다. 이러한 연구는 측정 도구 개발 이전 단계에서 수행되는 기초연구의 성격을 가지며, 전문가의 합의를 통해 핵심 역량이나 구성 요소를 정립하고 내용타당도를 확보하는 데 중점을 두고 있다.

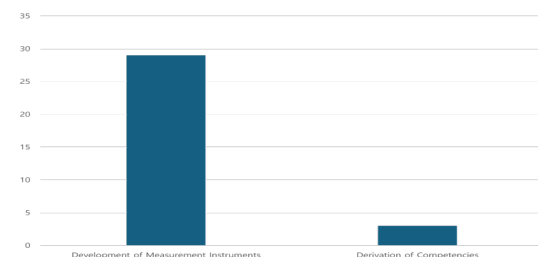


Figure 4. Distribution of Studies by Research Purpose

4.2 측정 도구 구성 특성

4.2.1 측정 변인 유형

측정 변인 유형을 분석한 결과, 인공지능 교육에서 활용된 AI 관련 측정 도구는 다양한 변인을 중심으로 개발되고 있는 것으로 나타났다. 이 중 AI 역량을 중심으로 한 연구가 15편으로 가장 높은 비중을 차지하였으며, 다음으로 AI 리터러시를 측정 변인으로 설정한 연구가 9편으로 나타났다. 또한 태도 및 인식을 측정하기 위한 도구 개발 연구는 4편이었으며, AI 윤리와 교수효능감을 측정하기 위한 도구 개발 연구는 각각 2편으로 확인되었다.

이처럼 측정 변인 중 AI 역량과 AI 리터러시가 가장 높은 비중을 차지하고 있으며, 두 개념이 유사하게 혼용되는 경우가 많아 구체적으로 구분할 필요가 있다. AI 역량은 인공지능의 원리와 작동 방식을 이해하고 이를 다양한 상황에서 적절하게 활용하며, 인공지능이 사회와 인간의 삶에 미치는 영향을 책임감 있게 고려하는 지식, 기능, 태도, 가치를 포괄하는 종합적 역량을 의미한다[7, 8]. 이에 비해 AI 리터러시는 인공지능의 개념과 원리, 사회적 영향에 대한 이해와 비판적 사고 능력에 초점을 두는 개념으로 구분할 수 있다[12, 20, 21]. 즉, AI 역량이 '할 수 있는가'에 초점을 둔다면, AI 리터러시는 '이해하고 판단할 수 있는가'에 초점을 두는 개념이라 할 수 있다.

이러한 결과는 인공지능 교육에서 단순한 이해 수준을 넘어 학습자의 실제 수행 능력과 문제 해결 능력을 평가하려는 연구가 중심으로 이루어지는 한편, 인공지능에 대한 이해와 비판적 사고 등 기초적인 역량을 측정하려는 연구도 함께 이루어지고 있음을 시사한다.

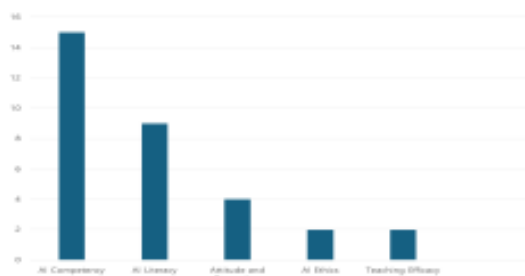


Figure 5. Distribution of Studies by Measurement Variables

4.2.2 하위 요인 및 문항 구성

본 연구에서는 OECD Learning Compass 2030에서 제시한 핵심 역량 체계를 참고하여 분석 기준을 재구성하였다. 구체적으로 AI 개념 및 원리 이해, 컴퓨팅 사고력 이해, 인공지능 교육 및 융합 이해, 디지털·AI 리터러시 이해는 지식 영역으로 분류하였다. 기술 활용 및 실행, 문제 해결 및 프로그래밍 수행, 교수·학습 설계 및 실행, 평가 및 피드백, 상호작용 및 참여는 기능 영역으로 구분하였다. 태도 영역에는 태도 및 정서, 자기효능감 및 기대, 사회적 관계 및 상

호작용, 자기개발 및 성장을 포함하였으며, 가치 영역에는 윤리 원칙 및 규범, 데이터 및 개인정보 보호, 사회적 영향 및 책임, 가치 및 공공성을 포함하였다.

각 하위요인은 OECD Learning Compass 2030의 개념 정의를 공통 기준으로 삼아 분류하였다. 즉, 인공지능에 대한 사실적, 개념적 이해를 다루는 하위요인은 지식 영역으로, 이를 실제 상황에 적용하고 수행하는 능력을 다루는 하위요인은 기능 영역으로, 인공지능에 대한 정서적 반응이나 동기, 자기효능감을 다루는 하위요인은 태도 영역으로, 인공지능 활용에 따른 윤리적 책임과 사회적 가치 판단을 다루는 하위요인은 가치 영역으로 분류하였다.

이러한 기준에 따라 AI 개념 및 원리 이해에는 인공지능의 정의와 작동 원리를 다룬 하위요인이 포함되었으며, AI 기술 자체에 대한 개념적 지식을 측정한다는 공통점을 근거로 지식 영역으로 분류되었다. 교수·학습 설계 및 실행에는 AI 수업 설계, AI 활용 교수·학습 실행 등의 하위요인이 포함되었는데, 이는 AI에 대한 이해를 실제 수업 행동으로 구현하는 수행적 성격을 지닌다는 점에서 기능 영역으로 분류되었다. 태도 영역의 자기효능감 및 기대에는 AI 활용에 대한 자신감과 성공 기대 등의 하위요인이 포함되었으며, 이는 인공지능 활용에 대한 정서적, 동기적 반응을 다룬다는 공통점을 근거로 태도 영역으로 분류되었다. 가치 영역의 윤리 원칙 및 규범에는 AI 윤리, 윤리적 판단 기준 등의 하위요인이 포함되었으며, AI 활용 시 요구되는 책임 의식과 윤리적 판단 기준을 다룬다는 공통 기준에 따라 분류되었다.

하위 요인을 재분류한 결과 기능 영역이 36개로 가장 높은 비중을 차지하였고 다음으로 가치 영역이 34개, 지식 영역이 25개, 태도 영역이 20개 순으로 나타났다.

하위 요인 분포에서 기능과 가치 영역의 비중이 지식·태도 영역보다 높게 나타난 것은, AI 역량이 단순한 개념적 이해를 넘어 실제 도구 활용 및 문제 해결 능력과, AI가 사회에 미치는 영향에 대한 책임 있는 활용을 핵심으로 요구하는 최근의 인공지능 교육 정책 방향과 맞물려 있는 것으로 해석된다. 이는 2022 개정 교육과정에서 인공지능의 사회적 영향과 윤리적 활용을 강조하고 있는 정책 방향과도 부합한다. 특히 가치 영역의 비중이 높게 나타난 것은, 생성형 AI의 확산과 함께 AI 윤리 및 데이터 보호에 대한 사회적 관심이 커지면서, 측정도구 개발 연구에서도 윤리·가치 영역을 별도의 핵심 구인으로 다루는 경향이 반영된 결과로 볼 수 있다. 실제로 최근 AI 윤리 역량이나 윤리인식을 측정하기 위한 독립적인 도구 개발 연구가 증가하고 있는 추세도 이러한 해석을 뒷받침한다.

Table 2. Distribution of Domains and Sub-competencies in AI Competency Measurement Tools

Competencies	Sub-competencies	N	Total
Knowledge	Understanding of AI concepts and principles	8	25 (21.7%)
	Understanding of computational thinking	5	
	Understanding of AI education and integration	5	
	Understanding of digital and AI literacy	7	
Skills	Technical use and implementation	6	36 (31.3%)
	Problem-solving and programming performance	8	
	Instructional design and implementation	9	
	Assessment and feedback	7	
	Interaction and participation	6	
Attitudes	Attitude and emotion	8	20 (17.4%)
	Self-efficacy and expectation	5	
	Social relationships and interaction	3	
	Self-development and growth	4	
Values	Ethical principles and norms	14	34 (29.6%)
	Data and privacy protection	3	
	Social impact and responsibility	9	
	Values and public interest	8	

1) 지식 영역

지식 영역 연구 25건 중 AI 개념 및 원리 이해가 8편으로 가장 높은 비중을 차지했으며, 디지털·AI 리터러시 이해 7편, 컴퓨팅 사고력 이해 5편, 인공지능 교육 및 융합 교육 이해 5편 순으로 나타났다.

AI 개념 및 원리 이해 영역에서는 인공지능의 기본 개념과 정의를 이해하고, 인공지능이 인간의 인지, 학습, 추론과 같은 지능적 기능을 구현하는 기술임을 파악하도록 한다. 또한 인공지능이 데이터와 알고리즘을 기반으로 작동한다는 원리를 이해하고, 문제 해결 과정의 구조를 설명할 수 있도록 한다[20-22].

컴퓨팅 사고력 이해 영역에서는 문제를 체계적으로 해결하기 위한 사고 과정을 이해하도록 한다. 이를 위해 문제를 분해하고 핵심 요소를 파악하며, 패턴을 인식하고 알고리즘을 설계하는 과정에 대한 개념적 이해를 포함한다. 또한 데이터를 수집·분석·해석하는 과정의 의미를 이해하고, 이를 바탕으로 문제 해결 과정의 구조를 이해하도록 한다[9, 23].

인공지능 교육 및 융합 이해 영역에서는 인공지능 기술이 교육 현장에서 어떻게 활용되는지를 이해하고, 이를 교과 및 교수·학습 과정과 연계하는 원리를 이해하도록 한다. 이를 위해 인공지능 기반 디지털 도구와 학습 환경의 특성을 이해하고, 학습자의 수준과 교육 목표에 맞게 수업이 설계되고 재구성되는 과정을 이해하는 것을 포함한다[24-27].

디지털·AI 리터러시 이해 영역에서는 인공지능과 디지털 기술이 일상생활과 다양한 맥락에서 어떻게 활용되는지

를 이해하는 데 중점을 둔다. 이를 위해 인공지능 기술과 서비스의 기능과 특징을 이해하고, 이를 비판적으로 해석하는 기초적인 이해를 포함한다. 또한 인공지능과의 상호작용과 소통 방식에 대한 이해와 함께, 다양한 상황에서 인공지능이 어떻게 활용되는지를 이해하는 것을 포함한다[22, 28-31].

2) 기능 영역

기술 활용 및 실행 영역에서는 인공지능 및 디지털 기술을 실제 상황에서 활용하는 능력을 기르는 것을 목표로 한다. 이를 위해 다양한 AI 도구와 플랫폼을 활용하여 정보를 탐색하고 과제를 수행하며, 문제 해결 과정에서 기술을 적절하게 적용하는 능력을 포함한다[32, 33].

문제 해결 및 프로그래밍 수행 영역에서는 문제를 해결하기 위한 절차적이고 논리적인 수행 능력을 기르는 것을 목표로 한다. 이를 위해 문제를 정리하고 해결 방법을 설계하며, 프로그래밍을 통해 실제로 구현하는 능력을 포함한다. 또한 다양한 문제를 해결하고 새로운 해결 방법을 만들어내는 능력을 포함한다[7, 26].

교수·학습 설계 및 실행 영역에서는 인공지능을 활용해 수업을 계획하고 운영하는 능력을 기르는 것을 목표로 한다. 이를 위해 학생의 수준과 특성을 고려해 수업을 설계하고, AI 도구를 활용해 수업을 진행하는 능력을 포함한다. 또한 학생의 참여를 높이기 위해 수업을 다시 구성하고 운영하는 능력을 포함한다[11, 12, 25, 26, 34, 35].

평가 및 피드백 영역에서는 학습 결과를 분석하고 이를 바탕으로 수업을 개선하는 능력을 기르는 것을 목표로 한다. 이를 위해 학습자의 성취를 평가하고 결과를 해석하여 적절한 피드백을 제공하며, 평가 결과를 바탕으로 교수·학습 과정을 성찰하고 개선하는 능력을 포함한다[12, 25, 34].

상호작용 및 참여 영역에서는 학습 과정에서 타인 및 인공지능과 협력하며 학습을 확장하는 수행 능력을 기르는 것을 목표로 한다. 이를 위해 다양한 상황에서 협업하고 의사소통하며, 인공지능과의 상호작용을 통해 문제를 해결하는 능력을 포함한다. 또한 학습 활동에 적극적으로 참여하고 협력을 이어가는 수행 능력을 포함한다[36, 37].

3) 태도 영역

태도 및 정서 영역에서는 인공지능에 대한 감정과 인식, 그리고 수용 태도를 형성하는 것을 목표로 한다. 이를 위해 인공지능에 대한 흥미, 호감, 불안과 같은 정서적 반응을 이해하고, 인공지능을 긍정적으로 받아들이려는 태도를 포함한다. 또한 인공지능 기술에 대한 인식과 가치관을 바탕으로 적절한 태도를 형성하는 것을 포함한다[10, 38-40].

자기효능감 및 기대 영역에서는 인공지능을 잘 활용할 수 있다는 믿음과 기대를 기르는 것을 목표로 한다. 이를 위해 인공지능을 사용할 때 자신감을 가지고, 좋은 결과가 나올 것이라고 기대하는 태도를 포함한다. 또한 어려움이 있어도 계속 도전하려는 태도를 포함한다[41, 42].

사회적 관계 및 상호작용 영역에서는 학습 과정에서 타인 및 인공지능과의 관계를 형성하고 상호작용하려는 태도를 기르는 것을 목표로 한다. 이를 위해 협업과 의사소통의 중요성을 이해하고, 타인 및 인공지능과 함께 학습하려는 태도를 포함한다. 또한 다양한 상황에서 적극적으로 참여하고 관계를 형성하려는 태도를 포함한다[36].

자기개발 및 성장 영역에서는 인공지능 시대에 필요한 역량을 지속적으로 발전시키려는 태도를 기르는 것을 목표로 한다. 이를 위해 변화하는 기술 환경에 적응하려는 태도와 인공지능 관련 지식과 기술을 꾸준히 학습하려는 태도를 포함한다. 또한 자신의 학습 과정을 돌아보고, 지속적으로 성장하려는 태도를 포함한다[12, 13].

4) 가치 영역

윤리 원칙 및 규범 영역에서는 인공지능 활용 과정에서 지켜야 할 윤리적 기준과 책임 의식을 이해하는 것을 목표로 한다. 이를 위해 인공지능을 공정하고 책임감 있게 사용하려는 태도와, 기술 활용 과정에서 발생할 수 있는 윤리적 문제를 인식하고 바르게 판단하려는 태도를 포함한다. 또한 인공지능이 개인과 사회에 미치는 영향을 고려하여 윤리적 기준에 따라 행동하려는 태도를 포함한다[37, 43, 44].

데이터 및 개인정보 보호 영역에서는 인공지능 활용 과정에서 데이터와 개인정보를 보호하기 위한 기준과 책임을 이해하는 것을 목표로 한다. 이를 위해 개인정보 보호의 중요성과 데이터 관리 원칙을 이해하고, 정보 활용 과정에서 발생할 수 있는 문제를 인식하는 것을 포함한다. 또한 데이터 활용과 관련된 윤리적 기준을 이해하는 것을 포함한다. 사회적 영향 및 책임 영역에서는 인공지능이 사회에 미치는 영향을 이해하고, 이에 따른 책임을 인식하는 것을 목표로 한다. 이를 위해 인공지능 기술이 사회 구조, 문화, 인간의 삶에 미치는 긍정적·부정적 영향을 이해하고, 기술 활용에 따른 사회적 책임의 의미를 파악하는 것을 포함한다. 또한 인공지능 활용이 공동체에 미치는 영향을 고려하는 관점을 포함한다[21, 33].

가치 및 공공성 영역에서는 인공지능 활용 과정에서 요구되는 사회적 가치와 공공성을 이해하는 것을 목표로 한다. 이를 위해 공정성, 투명성, 안전성, 다양성, 인간 존중과 같은 핵심 가치를 이해하고, 인공지능 활용에서 이러한 가치가 왜 중요한지 인식하는 것을 포함한다. 또한 개인의 이익을 넘어 사회 전체의 이익과 공공의 가치를 고려하는 관점을 포함한다[43, 44].

4.3 측정 도구의 타당도 및 신뢰도

본 연구에서는 분석 대상 논문에서 제시된 AI 역량 측정 도구의 타당도와 신뢰도 확보 방법을 분석하였다. 이를 통해 각 연구에서 측정 도구의 적절성과 신뢰성을 어떻게 검증하였는지 파악하고 전반적인 연구 동향을 살펴보고자 하였다.

4.3.1 내용 타당도 검증

측정도구 개발의 초기 단계에서 문항의 적합성을 확인하기 위한 내용 타당도는 전체 32편의 연구 중 31편에서 고려된 것으로 나타났다. 이 중 29편의 연구에서는 교육학, 컴퓨터공학 등 관련 분야의 전문가 집단을 구성하여 델파이 조사 또는 전문가 검토를 통해 문항의 타당성을 검토하였으며, 일부 연구에서는 내용타당도 비율(CVR)이나 I-CVI와 같은 지수를 산출하여 정량적으로 검증하였다. 반면, 나머지 2편의 연구에서는 전문가 검토나 델파이 절차를 수행하지 않고 선행연구 및 이론적 고찰을 바탕으로 문항을 구성하여 개념적 수준의 타당도를 확보한 것으로 나타났다. 또한 타당성 판단의 기준으로는 주로 Lawshe가 제안한 패널 수 대비 최소 임계값을 적용하였으며, 이를 충족하지 못하거나 전문가 간 합의도가 낮은 문항은 삭제하거나 수정하는 과정을 거쳐 최종 문항을 정제한 것으로 나타났다[15].

4.3.2 구인 타당도 검증

개발된 문항들이 연구자가 의도한 이론적 요인 구조를 적절히 반영하는지를 확인하기 위해 구인 타당도 검증이 수행되었다. 분석 결과, 다수의 연구에서 탐색적 요인분석(EFA)과 확인적 요인분석(CFA)을 순차적으로 실시한 것으로 나타났다.

전체 32편 중 23편의 연구에서는 초기 문항의 요인 구조를 탐색하기 위해 탐색적 요인분석(EFA)을 실시하였다. 이들 연구는 요인분석이 가능한 자료인지 확인하기 위해 KMO 지수와 Bartlett의 구형성 검정을 활용하였다. 또한 주성분 분석(PCA) 등을 적용하여, 고유값, 공통성, 요인부하량 등의 기준을 충족하는 문항을 선별하고 요인 구조를 도출하였다.

요인 구조의 적합성을 검증하기 위해 전체 32편 중 21편의 연구에서 구조방정식 모형을 활용한 확인적 요인분석(CFA)을 실시한 것으로 나타났다. 이들 연구는 χ^2 통계량의 한계를 보완하기 위해 CFI, TLI, RMSEA, SRMR 등의 다양한 적합도 지수를 함께 사용하여 모형의 적합성을 평가하였다.

수렴타당도와 변별타당도를 검증한 연구는 총 16편으로 나타났다. 이 중 CR과 AVE를 활용하여 수렴 및 변별타당도를 검증한 연구는 11편이었으며, 나머지 5편의 연구에서는 표준화 요인부하량, 요인 간 상관계수 또는 기존 척도와 상관분석을 활용하는 방식으로 수렴타당도와 변별타당도를 간접적으로 검증한 것으로 나타났다. 반면, 나머지 16편의 연구에서는 수렴타당도와 변별타당도에 대한 별도의 검증이 수행되지 않은 것으로 나타났다.

4.3.3 신뢰도 검증

측정도구의 내적 일관성을 평가하기 위한 신뢰도 검증은 전체 32편 중 25편의 연구에서 수행된 것으로 나타났다. 이들 연구는 문항 간 일관성을 확인하기 위해 Cronbach's α 계수를 산출하였으며, 대부분의 연구에서 높은 수준의 신뢰도 계

수를 보고하여 척도의 안정성을 확보한 것으로 나타났다.

5. 결론 및 제언

본 연구는 국내 인공지능 교육 연구에서 활용된 AI 역량 측정 도구의 유형과 특성을 체계적으로 분석하기 위해 2020년부터 2025년까지 발표된 32편의 학술지 논문을 대상으로 문헌 고찰을 하였다. 분석 결과, AI 역량 측정 도구의 활용 경향과 구성 특성, 적용 대상, 타당도 및 신뢰도 검증 방식의 주요 특징을 확인할 수 있었다. 이러한 결과를 구체적으로 살펴보면 다음과 같다.

첫째, AI 역량 측정 도구 관련 연구는 2023년 이후 크게 증가한 것으로 나타났으며, 특히 교사와 예비교사를 대상으로 한 연구가 가장 많은 비중을 차지하였다. 이는 인공지능 교육을 실제로 수행하는 교사의 역량이 중요하게 인식되고 있음을 보여준다. 반면, 초·중등 학생을 대상으로 한 연구는 상대적으로 적어, 앞으로 학생의 발달 수준과 학교급을 고려한 학습자 중심의 측정도구 개발 연구가 더 필요함을 시사한다.

둘째, 측정도구의 구성 측면에서는 기능 영역이 가장 높은 비중을 차지하였으며, 그 다음으로 가치, 지식, 태도 영역 순으로 나타났다. 이는 인공지능 교육에서 단순한 개념 이해뿐만 아니라 실제 문제 해결과 기술 활용 능력을 중요하게 다루고 있음을 보여준다. 또한 윤리와 사회적 책임과 같은 가치 영역도 높은 비중을 차지하고 있어, 인공지능 기술이 사회에 미치는 영향에 대한 교육적 중요성이 반영된 것으로 볼 수 있다.

셋째, 측정도구의 타당도와 신뢰도 검증에서는 대부분의 연구에서 내용타당도를 확인하였으며, 특히 전문가 검토나 델파이 조사를 활용한 경우가 많았다. 또한 구인 타당도 검증에서는 탐색적 요인분석과 확인적 요인분석을 함께 사용한 연구가 많았고, 일부 연구에서는 수렴타당도와 변별 타당도를 추가로 검증하여 도구의 타당성을 더욱 높였다. 그러나 모든 연구에서 이러한 절차가 이루어진 것은 아니며, 일부 연구에서는 제한된 검증 방법만 적용된 것으로 나타났다. 신뢰도 검증에서는 많은 연구에서 Cronbach's α 계수를 활용하여 문항 간 일관성을 확인하였다. 그러나 일부 연구에서는 신뢰도 검증이 이루어지지 않은 경우도 있었다. 이는 측정도구의 질을 높이기 위해 보다 체계적인 검증 절차가 필요함을 보여준다.

이상의 결과를 바탕으로 다음과 같은 시사점을 도출할 수 있다. 첫째, 초·중등 학생을 대상으로 한 AI 역량 측정도구 개발 연구가 필요하다. 본 연구의 분석 결과, 교사와 예비교사를 대상으로 한 연구는 많이 이루어졌지만, 초·중등 학생을 대상으로 한 측정도구 연구는 상대적으로 부족한 것으로 나타났다. 따라서 학생의 발달 수준과 학교급별 교육 내용을 고려한 학습자 중심의 AI 역량 측정 도구를 개발할 필요가 있다. 특히 프로그래밍, 피지컬 컴퓨팅, 인공지능을 개별

적으로 측정하는 데 그치지 않고, 이들을 통합적으로 활용하여 문제를 해결하는 역량을 반영한 측정도구 개발이 필요하다.

둘째, AI 역량 측정도구 개발 과정에서 보다 체계적인 타당도 및 신뢰도 검증 절차가 적용될 필요가 있다. 본 연구의 분석 결과, 32편 중 16편의 연구에서는 수렴타당도와 변별 타당도에 대한 별도의 검증이 이루어지지 않은 것으로 나타났다. 이는 측정도구가 의도한 구성개념을 실제로 측정하고 있는지, 그리고 이론적으로 구분되는 다른 구성개념과 변별되는지를 확인하지 않은 채 도구가 활용되고 있음을 의미한다. 따라서 향후 연구에서는 내용타당도, 구인타당도, 수렴타당도, 변별타당도, 신뢰도에 이르는 종합적인 검증 절차를 적용하여 측정도구의 신뢰성과 타당성을 더욱 강화할 필요가 있다.

셋째, AI 역량의 다양한 요소를 균형 있게 반영한 통합적 측정도구 개발이 필요하다. 기존 연구에서는 AI 도구 활용이나 문제 해결 능력과 같은 기능 영역이 높은 비중을 차지하는 경향이 있었다. 그러나 AI 역량은 인공지능 개념을 이해하는 지식, AI 도구를 활용하여 문제를 해결하는 기능, AI에 대한 태도, 윤리와 사회적 책임 의식을 함께 포함하는 개념이다. 따라서 향후 측정도구는 특정 영역에 치우치기보다 지식, 기능, 태도, 가치 영역을 균형 있게 반영하여 개발될 필요가 있다. 이는 최근 국내 교육과정이 지식 습득 중심에서 역량 중심 교육으로 전환되고 있는 흐름과도 맥을 같이 하며, 인지적 영역뿐 아니라 정의적·가치적 영역까지 포괄하는 통합적 평가 체계의 필요성을 뒷받침한다.

본 연구는 개별 측정도구 개발 연구를 넘어, 국내 AI 역량 측정도구 연구 전반을 대상으로 측정 영역, 하위 요소, 타당화 절차, 연구 대상별 특징 등을 종합적으로 분석하였다는 점에서 차별성을 가진다. 특히 OECD Learning Compass 2030 기반의 분석 틀을 적용하여 AI 역량 측정도구의 구조적 특징과 영역별 경향을 종합적으로 분석하고, 연구 대상에 따라 강조되는 역량 요소의 차이와 측정도구 연구의 한계를 함께 검토하였다는 점에서 학술적 의의가 있다. 하지만 본 연구는 국내 학술지 논문만을 대상으로 분석하였기 때문에 국제 연구와의 비교 분석에는 한계가 있다. 또한 KCI, RISS, DBpia 등 국내 학술 데이터베이스로 검색 범위를 한정하였고 분석 대상 기간을 2020년부터 2025년까지로 설정하였기 때문에, 이 기간 이전에 개발된 측정도구나 데이터베이스에 색인되지 않은 학위논문, 연구보고서 등은 분석에서 제외되었을 가능성이 있다. 또한 학술지에 게재된 연구를 중심으로 문헌을 수집함에 따라 출판 편향의 가능성을 배제할 수 없다. 향후 연구에서는 해외 AI 리터러시 및 AI 역량 프레임워크와의 비교를 통해 국내 측정도구 연구의 특징과 한계를 보다 심층적으로 분석하고 다양한 시사점을 도출할 필요가 있다.

참고문헌

- [1] Ministry of Education, (2015). *2015 Revised Curriculum Overview*. Sejong city. Ministry of Education
- [2] Ministry of Education, (2021). *2022 Revised Curriculum Overview*. Sejong city. Ministry of Education
- [3] Jang, E., & Oh, K. (2025). Analysis of the Impact of Generative AI Utilization on Computational Thinking and Problem-Solving Ability. *The Journal of Korean Association of Computer Education*, 28(7), 83-98. <https://doi.org/10.32431/kace.2025.28.7.008>
- [4] Han, J. (2020). Changes in attitudes and efficacy of AI learners according to the level of programming skill and project interest in AI project. *Journal of the Korean Association of Information Education*, 24(4), 391-400. <https://doi.org/10.14352/jkaie.2020.24.4.391>
- [5] Korea Foundation for the Advancement of Science and Creativity. (2020b). *Guide to artificial intelligence education*. Seoul: Korea Foundation for the Advancement of Science and Creativity.
- [6] Jang, H. (2020). The Domestic Research Trends of Artificial Intelligence Related in Elementary Practical Arts Education. *The Journal of Korea Elementary Education*, 31, 33-48
- [7] Lee, E., & Pyun, Y.(2024). Development of checklist questions to measure AI capabilities of elementary school students. *Journal of The Korea Internet of Things Society*, 10(3), 7-12. <https://doi.org/10.20465/kiots.2024.10.3.007>
- [8] Jung, J., Kim, H., & Joo, Y. (2023). An Analysis of Priorities for AI-Digital Competencies of Elementary and Secondary School Teachers in Korea. *The Korea Educational Review*. 29(3), 177-203. <https://doi.org/10.29318/KER.29.3.8>
- [9] Lee, E., & Han, J. (2024). Development of checklist questions to measure AI core competencies of middle school students. *Journal of The Korea Internet of Things Society*. 10(3), 49-55. <https://doi.org/10.20465/kiots.2024.10.3.049>
- [10] Kim, S., & Lee, J. (2020). Development of Test Tool of Attitude toward Artificial Intelligence for Middle School Students. *The Journal of Korean association of Computer Education*. 23(3), 17-30. <https://doi.org/10.32431/kace.2020.23.3.003>
- [11] Heo, H., & Kang, S. (2023). Teacher competencies for designing artificial intelligence-integrated education. *The Journal of Korean Association of Computer Education*. 26(2), 89-100. <https://doi.org/10.32431/kace.2023.26.2.008>
- [12] Heo, H., Seo, J., Lim, K., Lim, K., Lee, H., & Kang, S. (2024). Development and validation of an assessment instrument for teachers' AI-digital competencies. *Journal of Korean Association for Educational Information and Media*. 30(2), 497-521. <http://dx.doi.org/10.15833/KAFEIAM.30.2.497>
- [13] Im, C., Han, Y., Kim, S., Lee, C., Hong, S., & Jeong, J. (2025). Development of an AI and digital competency assessment instrument for in-service teachers using the behaviorally anchored rating scale (BARS). *The Journal of Educational Information and Media*, 31(3), 945-969. <http://dx.doi.org/10.15833/KAFEIAM.31.3.945>
- [14] Yu, M., & Lim, E. (2024). Development and Validation of a Career Guidance Competency Scale using AI for Pre-service Teachers. *Journal of the Future of Society*. 15(1), 242-263. <https://doi.org/10.22987/jifso.2024.15.1.242>
- [15] Lawshe, C. (1975). A quantitative approach to content validity. *Personnel Psychology*, 28(4), 563-575.
- [16] Choi, C., & You, Y. (2017). The Study on the comparative analysis of EFA and CFA, *Journal of Digital Convergence*, 15(10), 103-111. <https://doi.org/10.14400/JDC.2017.15.10.103>
- [17] Fornell, C., & Larcker, D. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39-50.
- [18] Kitchenham, B. (2004). Procedures for performing systematic reviews. Keele, UK, Keele University, 22(2004), 1-26.
- [19] Lim, S., Go, H., Kim, S., & Lee, Y. (2025). A Systematic Literature Review on the Use of Physical Computing in Primary and Secondary Artificial Intelligence Education. *The Journal of Korean Association of Computer Education*, 28(8), 1-14.
- [20] Hwang, H., Hwang, Y., Park, J., Shin, M., & Lee, H. (2024). A study on the development and validity verification of a measurement tool for AI literacy. *The Korean Journal of Literacy Research*, 15(2), 247-278. <http://dx.doi.org/10.37736/KJLR.2024.04.15.2.09>
- [21] Ryu, M., & Han, S. (2023). The study on test standard for measuring AI literacy. *Journal of the Korea Society of Computer and Information*, 28(7), 39-46. <https://doi.org/10.9708/jksci.2023.28.07.039>
- [22] Han, S., & Chung, H. (2025). Development and validation of an artificial intelligence literacy scale for elementary school teachers. *CNU Journal of Educational Studies*, 46(3), 265-294. <http://dx.doi.org/10.18612/cnujes.2025.46.3.265>
- [23] Jang, E., & Oh, K. (2023). Development of core competency assessment tools in basic SW-AI education. *Korean Journal of General Education*, 17(2), 275-287.
- [24] Kim, D., So, H., & Lim, J. (2023). Development of the measurement instrument of AI convergence education competency for pre-service teachers. *Journal of Educational Studies*, 54(3), 139-168.
- [25] Park, G., Hwang, S., & Lee, J. (2023). Development and validation of a teaching competence scale for teachers' artificial intelligence convergence education. *Journal of Educational Technology*, 39(1), 315-344. <http://dx.doi.org/10.17232/KSET.39.1.315>
- [26] Bae, Y., Park, G., Hong, J., & Lee, J. (2023). Development of the measurement instrument of artificial intelligence convergence teaching competency for pre-service teachers. *The Journal of Learner-Centered Curriculum and Instruction*, 23(23), 339-351.
- [27] Lee, S., & Kim, K. (2024). Development and Validation of an AI Convergence Education Competency Measurement Tool for In-Service Teachers. *The Journal*

- of Korean Association of Computer Education, 27(4), 159-171. <http://dx.doi.org/10.32431/kace.2024.27.4.012>
- [28] Kim, S., & Yu, K. (2024). Development and Validation of an AI Literacy Rating Scale for Early Childhood Teachers. *Early Childhood Education Research & Review*, 28(4). <https://doi.org/10.32349/ECERR.2024.8.28.4.31>
- [29] Lim, C., Cho, Y., & Jeong, D. (2024). A Study on Developing of AI Literacy Scale for College Students. *The Journal of Information Technology and Architecture*, 21(3), 285-296. <http://doi.org/10.22865/jita.2024.21.3.285>
- [30] Kim, I., Yi, S., & Kim, K. (2024). A Study on the Development and Validation of Generative AI Literacy Measurement Tool for Elementary and Secondary Teachers. *The Journal of Learner-Centered Curriculum and Instruction*, 24(22), 697-708. <http://doi.org/10.22251/jlcci.2024.24.22.697>
- [31] Noh, H., Kim, H., & Kim, M. (2024). Development and Validation of a Korean Generative AI Literacy Scale. *Knowledge Management Research*, 25(3), 145-171. <http://doi.org/10.15813/kmr.2024.25.3.007>
- [32] Lim, H., & Lee, M. (2023). Development and Validation of the Pre-secondary school Teacher's Artificial Intelligence Literacy Scale. *The Journal of Learner-Centered Curriculum and Instruction*, 23(12), 875-892. <http://dx.doi.org/10.22251/jlcci.2023.23.12.875>
- [33] Kim, S., Kim, S., Park, C., Hong, J., & Park, J. (2023). Development of an AI competency measurement tool for pre-service teachers to enhance expertise in digital education. *The Journal of Korean Association of Computer Education*, 26(4), 21-32. <http://dx.doi.org/10.32431/kace.2023.26.4.003>
- [34] Ko, E., Kim, D., Kang, D., & Lee, J. (2025). Development and validation of a diagnostic tool for preservice teachers' competence in implementing AI-based digital textbooks. *Journal of Educational Technology*, 41(3), 727-753. <http://dx.doi.org/10.17232/KSET.41.3.727>
- [35] Lee, H., & Park, J. (2024). Development of AIED competencies and achievement standards for pre-primary teachers in elementary school. *Journal of The Korean Association of Information Education*, 28(2), 181-189. <http://doi.org/10.14352/jkaie.2024.28.2.181>
- [36] Jang, Y., Yoon, I., & Kim, H. (2023). Deduction and validation of universal artificial intelligence core competencies in the AI-based intelligent information society. *Journal of the Korean Association of Information Education*, 27(5), 477-488.
- [37] Jeong, Y., & Joo, S. (2025). Development and validation of an AI literacy measurement tool for responsible learning and ethical AI use of university students. *Cultural Exchange and Multicultural Education*, 14(5), 475-503.
- [38] Lee, S., Kim, H., & Ahn, H. (2025). Validation of the Korean Version of the Attitudes Toward Artificial Intelligence Scale (ATTARI-12). *Korean Psychological Journal of Culture and Social Issues*, 31(2), 269-292. <http://dx.doi.org/10.20406/kjcs.2025.5.31.2.269>
- [39] Choi, B. (2025). Validation of an AI Transition Attitude Scale for University Students. *The Journal of Saramdaum Education*, 19(2). <http://dx.doi.org/10.18015/edumca.19.2.202505.83>
- [40] Kim, D., Park, Y., Ryu, S., Hwang, Y., Jang, Y., & Kim, J. (2024). The validation of preschool teachers' perception of the educational use of artificial intelligence. *The Journal of Korea Open Association for Early Childhood Education*, 29(1), 155-171. <http://dx.doi.org/10.20437/KOAECE29-1-07>
- [41] Yi, S., Kim, S., & Lee, Y. (2021). Development of Teaching Efficacy Belief Instrument about Artificial Intelligence for Pre-service Teachers. *The Journal of Korean Association of Computer Education*, 24(1), 47-61. <http://dx.doi.org/10.32431/kace.2021.24.1.006>
- [42] Park, H., Kim, H., Choi, J., & Jeon, Y. (2021). Development of a teaching efficacy instrument in informatics (software and AI) subject. *The Journal of Korean Association of Computer Education*, 24(4), 39-52. <http://dx.doi.org/10.32431/kace.2021.24.4.004>
- [43] Bae, J., Lee, J., Hong, M., & Cho, J. (2022). The Development of AI Ethical Competence Scale for Secondary School Students. *The Journal of Korean Association of Computer Education*, 25(6), 115-130. <http://dx.doi.org/10.32431/kace.2022.25.6.008>
- [44] Koo, M., Yoon, J., & Kim, J. (2025). Development of artificial intelligence ethical awareness measurement tool for early childhood teachers. *The Journal of Korea Open Association for Early Childhood Education*, 30(2), 139-163.



조민국

· 2010년 청주교육대학교 물리교육과
· 2023년 충남대학교 교육대학원 AI융합교육전공 (교육학석사)
· 2024년~현재 한국교원대학교 초등컴퓨터교육과 박사과정

✚ 관심분야 : 인공지능 융합교육, 컴퓨터교육, 정보교육

✉ lotem011@korea.kr



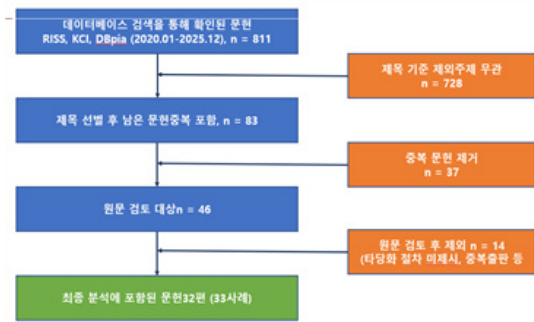
김귀훈

· 1998년 KAIST 전기및전자공학(공학사)
· 2000년 KAIST 전자전산학과(공학석사)
· 2019년 KAIST 전기및전자공학부(Ph.D.)
· 2000년~2005년 LG 데이콤 주임연구원
· 2005년~2020년 ETRI 실장, 책임연구원
· 2006년~현재 ITU-T SG11 Rapporteur, Editor
· 2022년~현재 한국AI융합교육연구소 소장
· 2024년~현재 교육정보원 원장
· 2020년~현재 한국교원대학교 인공지능융합교육 전공교수

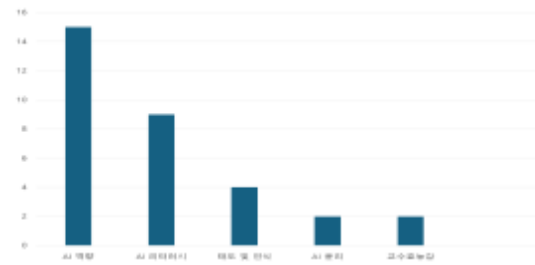
✚ 관심분야 : 인공지능융합교육, ICT디지털교과서, 지능형 에지컴퓨팅, 강화학습

✉ kimkh@knue.ac.kr

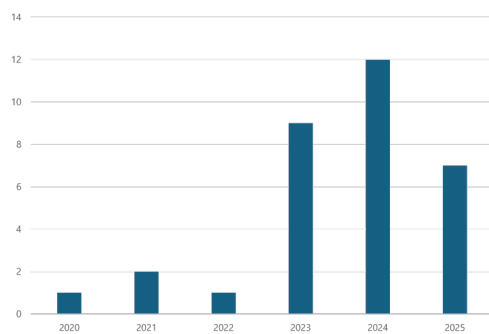
부록



〈그림 1〉 문헌 선정 과정에 대한 PRISMA 흐름



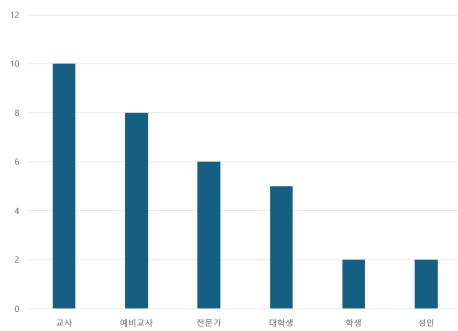
〈그림 5〉 측정 변인별 분포



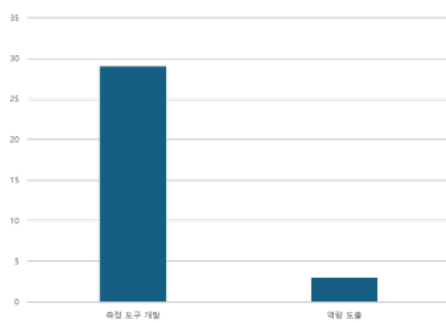
〈그림 2〉 연도별 논문 수

〈표 1〉 AI 역량 측정도구 관련 문헌 분석 기준과 세부 내용

대분류	세부 분석 항목	분석 내용 및 목적
일반적 현황	발행 연도, 연구 대상, 연구 목적	AI 역량 측정도구 관련 연구의 전반적인 흐름을 파악하고, 연구가 수행된 시기별 경향과 연구 대상의 분포를 확인
측정도구 구성 특성	측정 변인, 하위 요인 구성	인지, 수행, 정의, 가치·윤리 등 측정 변인의 유형과 도구의 구조적 특징을 분석
측정도구의 타당도 및 신뢰도	타당화 방법, 신뢰도 계수, 분석 기법 (탐색적·확인적 요인분석 등)	측정도구의 신뢰도와 타당도를 검증한 방법을 분석하여 도구의 질적 수준과 연구의 엄밀성 확인



〈그림 3〉 연구 대상별 분포



〈그림 4〉 연구 목적별 분포

〈표 2〉 AI 역량 측정도구의 영역 및 하위 역량 분포 분석 결과

역량	하위역량	빈도	합계
지식	AI 개념 및 원리 이해	8	25
	컴퓨팅 사고력 이해	5	
	인공지능 교육 및 융합 이해	5	
	디지털·AI 리터러시 이해	7	
기능	기술 활용 및 실행	6	36
	문제 해결 및 프로그래밍 수행	8	
	교수-학습 설계 및 실행	9	
	평가 및 피드백	7	
	상호작용 및 참여	6	
태도	태도 및 정서	8	20
	자기효능감 기대	5	
	사회적 관계 및 상호작용	3	
	자기개발 및 성장	4	
가치	윤리 원칙 및 규범	14	34
	데이터 및 개인정보 보호	3	
	사회적 영향 책임	9	
	가치 및 공공성	8	

〈표 3〉 체계적 문헌 고찰 분석 대상 논문 목록

저자(연도)	논문 제목	학술지	측정변인
황현정 외(2024)	AI 리터러시 측정 도구 개발과 타당화 연구	리터러시연구 15(2)	AI 리터러시
장은실, 오정선(2023)	SW·AI 기초교양에서의 핵심역량 성과 평가도구 개발	교양교육연구 17(2)	SW·AI 핵심역량
박가영 외(2023)	교사의 인공지능 융합교육 교수역량 척도 개발 및 타당화	교육공학연구 39(1)	AI융합교육 교수역량
허희옥 외(2024)	교원의 AI·디지털 역량 측정 도구 개발과 타당화	교육정보미디어연구 30(2)	AI·디지털 역량
최보라(2025)	대학생용 AI 전환 태도 검사도구 타당화	인격교육 19(2)	AI 전환 태도
임철홍 외(2024)	대학생을 위한 AI 리터러시 척도 개발	정보기술아키텍처연구 21(3)	AI 리터러시
정연재, 주소영(2025)	대학생의 책임 있는 학습과 윤리적 AI 사용을 위한 AI 리터러시 측정도구 개발 및 타당화 연구	문화교류와 다문화교육 14(5)	AI 리터러시
김성에 외(2023)	디지털 교육 전문성 강화를 위한 예비교사용 AI 역량 측정 도구 개발	컴퓨터교육학회논문지 26(4)	AI 역량
장운재 외(2023)	보편적 인공지능 교육을 위한 핵심 역량 도출 및 타당화	정보교육학회논문지 27(5)	AI 역량
임혜진, 이문수(2023)	예비 중등교사의 인공지능 리터러시 진단도구 개발 및 타당화	학습자중심교과교육연구 23(12)	AI 리터러시
김동심 외(2023)	예비교사 대상 AI융합교육 역량 척도의 개발 및 타당화	교육학연구 54(3)	AI융합교육 역량
배윤주 외(2023)	예비교사의 AI 융합교육 교수역량 측정 도구 개발	학습자중심교과교육연구 23(23)	AI융합교육 교수역량
유미옥, 임은미(2024)	예비교사의 AI 활용 진로지도 역량 척도 개발 및 타당화	미래사회 15(1)	AI 활용 진로지도 역량
이소울 외(2021)	예비교사의 인공지능 교수효능감 측정 도구(AI-TEBD) 개발	컴퓨터교육학회논문지 24(1)	AI 교수효능감
고은지 외(2025)	예비교원의 AI 디지털 교과서(AIDT) 활용 역량 진단도구 개발 및 타당화	교육공학연구 41(3)	AIDT 활용 역량
김선희, 유구중(2024)	유아교사 AI 리터러시 평정척도 개발 및 타당화	유아교육학논집 28(4)	AI 리터러시
구민경 외(2025)	유아교사의 인공지능 윤리의식 측정도구 개발	열린유아교육연구 30(2)	AI 윤리의식
류미영, 한선관(2023)	인공지능 리터러시 측정을 위한 검사 지표 개발	한국컴퓨터정보학회논문지 28(7)	AI 리터러시
허희옥, 강신현(2023)	인공지능융합교육 설계를 위한 교사 역량	컴퓨터교육학회논문지 26(2)	교사 AI융합교육 역량
김대웅 외(2024)	인공지능의 교육적 활용에 대한 유아교사 인식 척도의 타당화	열린유아교육연구 29(1)	AI 교육 활용 인식
박희정 외(2021)	정보(SW·AI) 교수효능감 측정도구 개발	컴퓨터교육학회논문지 24(4)	정보(SW·AI) 교수효능감
배진아 외(2022)	중등학생 대상 인공지능 윤리역량 평가척도 개발	컴퓨터교육학회논문지 25(6)	인공지능 윤리역량
김성원, 이영준(2020)	중학생의 인공지능에 대한 태도 검사 도구 개발	컴퓨터교육학회논문지 23(3)	인공지능에 대한 태도
이은철, 한정수(2024)	중학생의 AI 핵심역량 측정을 위한 체크리스트 문항 개발	사물인터넷융복합논문지 10(3)	AI 핵심역량
김인주 외(2024)	초·중등교사의 생성형 AI 리터러시 측정도구 개발 및 타당화 연구	학습자중심교과교육연구 24(22)	생성형 AI 리터러시
이희란, 박정호(2024)	초등 예비교원의 AIED 역량 및 성취기준 개발	정보교육학회논문지 28(2)	AIED 역량
한소미, 정혜원(2025)	초등교사 인공지능 리터러시 측정 도구 개발 및 타당화	교육연구논총 46(3)	AI 리터러시
이은철, 변영신(2024)	초등학생의 AI 역량 측정을 위한 체크리스트 문항 개발	한국사물인터넷학회 논문지 10(3)	AI 역량
이선영 외(2025)	한국판 인공지능에 대한 태도 척도(ATTARI-12)의 타당화	한국심리학회지: 문화 및 사회문제 31(2)	AI 태도
노환호 외(2024)	한국형 생성 인공지능 리터러시 척도 개발 및 타당화	지식경영 연구 25(3)	생성 AI 리터러시
임철일 외(2025)	행동기준 평정척도(BARS)를 활용한 현직 교원의 AI·디지털 역량 측정 도구 개발 및 타당화	한국교육정보미디어학회 31(3)	AI·디지털 역량
이소울, 김귀훈(2024)	현직 교사의 AI융합교육 역량 측정 도구 개발 및 타당화	컴퓨터교육학회논문지 27(4)	AI융합교육 역량