

인공지능 기반 교육 시스템의 사회적 영향 평가 프레임워크 연구*
**A Framework for Social Impact Assessment of AI-powered Systems in
Education**

방준성¹, 이상민²
Junseong Bang, Sangmin-Michelle Lee

Bang, Junseong & Lee, Sangmin-Michelle. (2025). A framework for social impact assessment of AI-powered systems in education. *Multimedia-Assisted Language Learning*, 28(1), 1-21.

The integration of artificial intelligence (AI) into education has significantly changed the dynamics of the classroom by enabling personalized and tailored learning through AI-powered learning analytics that address individual student needs. While the pedagogical benefits of personalized learning have long been recognized, implementation has been challenging in traditional classroom settings where teachers manage multiple students simultaneously. AI offers a potential solution to this challenge and could promote educational equity. However, the adoption of AI in education raises significant concerns regarding fairness, inclusivity, transparency, algorithmic bias, technology dependence, trustworthiness, and reduced social interaction-factors that could potentially lead to ineffective learning outcomes and educational inequality. Therefore, social impact assessment (SIA) becomes crucial to analyze and predict the outcomes of AI implementation in education, with the goal of minimizing negative impacts while maximizing positive societal contributions. This paper proposes a comprehensive framework for conducting SIA of AI in education, presenting core assessment items and procedures. The framework provides systematic guidelines for future SIA implementations in educational contexts, contributing to the responsible integration of AI in education.

Key words: AI-powered education system, social impact assessment, customized learning, AI digital textbook
Applicable level: all levels

* This study was supported by Institute of Information & Communications Technology Planning & Evaluation (IITP) under the metaverse support program to nurture the best talents (IITP-2024-RS-2024-00425383) grant funded by the Korea government (MSIT).

* This study was supported by NRF-2023S1A5A2A21083590.

¹ Junseong Bang (CEO, 1st author, Ymatics Corp., 19 Nambusunhwan-ro 351-gil, Gangnam District, Seoul, Korea; E-mail: hjbang21pp@gmail.com)

² Sangmin-Michelle Lee (Professor, Corresponding author, Kyung Hee University, 1732, Deogyong-daero, Yongin, Seoul, Korea; E-mail: sangminlee@khu.ac.kr)

Received: January 15, 2025; Reviewed: January 31, 2025; Accepted: February 15, 2025

I. 서론

사회적 영향 평가(Social Impact Assessment, SIA)는 정책, 법안, 기술, 혁신, 프로젝트 등이 사회에 미치는 영향을 체계적으로 분석, 예측, 평가하고 관리하는 과정이다. SIA의 주요 목적은 이러한 분석을 통해 긍정적 영향(사회적, 경제적, 문화적 영향)은 극대화하고 부정적 영향은 최소화하는 방안을 모색하는 데 있다(Vanclay, 2003). SIA는 다양한 이해관계자들의 의견을 수렴하여 특정 집단이 소외되지 않도록 하며, 잠재적 갈등 요인을 초기 단계에서 파악하여 해결할 수 있게 한다. 또한 경제적, 사회적으로 지속가능한 발전을 지원하는 데 기여한다. 이러한 과정을 통해 SIA는 해당 정책, 기술, 프로젝트의 사회적 수용성을 높여 장기적 성공을 도모하는 한편, 공동체의 복지와 권리 증진에도 이바지한다(Vanclay, 2003; Vanclay et al., 2019).

교육 분야에서도 SIA는 핵심적인 역할을 수행한다. 교육에서의 SIA는 교육 정책, 프로그램, 기술 도입이 학생, 교사, 지역사회와 전반적 사회에 미치는 영향을 예측하고 평가하여 공정하고 효과적인 교육 시스템 설계에 기여한다(Nuary et al., 2022). 이는 교육 자원의 공정한 분배를 통한 교육 형평성을 보장하고 이해관계자들의 의견 수렴을 통해 지역사회의 신뢰를 확보하는 한편, 장기적으로 지속가능하도록 해준다. 인공지능(Artificial Intelligence, AI)은 도입 초기 교육계에 엄청난 충격으로 다가왔으며, 상당한 거부감과 반발을 불러일으켰으나, 현재는 빠른 속도로 교육 현장에 적용되고 있다(Park et al., 2024; Shin, 2023). 이는 교육 분야 AI 활용의 잠재적 효과 때문인데, Bloom(1984)에 의하면, 코칭과 같은 개인별 맞춤형 학습이 가장 효과적인 교수방법으로, 전통적 교실 수업과 비교해 학습자의 평균 성취도를 2-시그마 정도 향상시키는 것으로 나타났다. 그러나 교사 1인이 다수의 학생을 지도해야 하는 현행 교실 환경에서는 사실상 이러한 맞춤형 학습을 실현하기 어려웠다. 이러한 상황에서 AI는 학습분석(learning analytics)을 기반으로 개별 학습자의 학습과정과 결과를 실시간으로 분석하여, 학습자의 수준과 요구에 맞는 학습자료와 학습경로를 제공함으로써 개인별 맞춤형 학습을 실현 가능하게 하고 있다(Do et al., 2022; Hong et al, 2024; Khan, 2024). 이러한 맥락에서 현재 진행 중인 교육계의 급격한 AI의 도입과 활용에 대해서 그 사회적 영향과 파장을 고려하여 시급하게 SIA를 시행해야 할 것으로 보인다.

특히, 우리나라 교육부는 교육에서의 디지털 전환(digital transformation)과 교육 효과성 제고, AI 시대에 대비하고자 초·중·고등학교 교실 현장에 AI 디지털교과서(AI digital textbook, AIDT) 도입을 결정하였다. 2025년 첫 도입 시에는 영어, 수학, 정보 교과를 우선 시행하고, 이후 사회와 과학 교과로 확대 시행할 예정이다(Ministry of Education of Korea, 2023). 공교육에서 AIDT의 도입은 세계 최초의 시도으로써, 기대와 우려가 공존하고 있다. AIDT는 모든 학습자에게 공정한 학습 기회를 제공하고 개별적 필요에 따른 학습을 가능하게 한다는 점에서, 교육의 평등성을 넘어 교육의 형평성 실현에 한 걸음 더 나아갈 수 있는 토대를 마련할 것이라는 기대를 받고 있다(OECD, 2024). 그러나 AIDT를 비롯한 AI 기반 교육 시스템은 교육 접근성 향상과 개인화된 맞춤형 학습을 통해 교육의 형평성과 학습 효과성을 높일 수 있으나, 이러한 시스템이 가져올 수 있는 잠재적 부작용에

대한 우려도 크다. AI 기반 교육 시스템은 데이터와 알고리즘의 편향성, 개인정보 보호 문제, AI 디지털 격차, 교사-학습자 및 학습자 간 상호작용 감소, 창의성과 비판적 사고력 같은 상위인지 능력 발달 저해 등 다양한 윤리적, 교육적, 사회적 문제를 초래할 수 있으며(Hong et al., 2024; Lee et al., 2023; Park et al., 2024), 이러한 문제들은 학습자의 학습 동기와 학습 효과에 부정적 영향을 심각하게 미칠 수 있다(Bang & Lee, 2024). AIDT가 올해부터 교실 현장에 도입될 예정인 만큼, 이에 대한 SIA의 수행이 시급한 과제로 대두될 수 있다. 교육에서의 AI 활용 관련 기존 연구는 특정 AI 툴을 활용하는 사례와 소규모 집단을 대상으로 한 효과성 연구에 주로 초점이 맞춰져 있으며, 국내에서 SIA 연구는 아직까지 찾아볼 수 없다. 이에 본 연구에서는 교육 분야 SIA를 위한 평가 지표를 설정하고, 구체적인 평가 프로세스와 그 예시를 제안하고자 한다. 본 연구는 국내외에서 발간된 자료를 바탕으로 교육에서의 AI 사용 시 고려해야 할 AI 윤리 지침과 가이드라인, SIA 관련 선행연구를 분석하고, 이를 바탕으로 AI 기반/활용 교육 시스템의 사회적 영향을 평가할 수 있는 핵심 지표와 프레임워크를 개발하고자 한다. 본 연구에서 개발된 교육에서의 AI의 SIA 프레임워크는 AIDT 도입에 따른 잠재적 AI 리스크를 사전에 진단하고, 교실 현장에서 발생할 수 있는 문제점을 체계적으로 분석할 수 있는 가이드라인을 제시할 수 있을 것이다. 본 연구의 연구질문은 다음과 같다.

1. AI 기반/활용 교육 시스템의 사회적 영향 평가를 하기 위한 핵심 지표는 무엇인가?
2. AI 기반/활용 교육 시스템의 사회적 영향을 체계적으로 분석하기 위해 평가 프로세스와 방법론은 어떻게 구성되어야 하는가?

II. 선행연구

1. 사회적 영향 평가의 정의 및 방법

사회적 영향 평가(SIA)는 프로젝트 기획과 정책 개발에 있어 시행 전후로 사회에 미치는 긍정적, 부정적 영향을 분석하고, 계획된 개입으로 인한 사회적 결과를 분석, 모니터링, 관리하기 위한 체계적인 프레임워크 역할을 하는 중요한 도구이다(Vanclay et al., 2019). 구체적으로, SIA는 정책, 프로그램, 계획, 프로젝트와 같은 계획된 개입과 그로 인해 발생하는 사회 변화 과정의 의도된 결과와 의도되지 않은 사회적 결과를 분석하여, 잠재적 영향에 대해 예측하고 부정적 영향을 해결하기 위한 완화 전략 개발을 도출하는 데 그 목적이 있다(Vanclay et al., 2015). Esteves 외(2012)에 따르면, SIA는 그 자체로 연구, 시행, 담화, 또는 패러다임이라고 할 수 있으며, 현재까지 다양한 분야(예: 환경, 경제, 정책 등)에서 SIA를 수행하는데 필요한 지식, 연구방법, 연구도구, 실질적 경험, 사례연구 등이 상당히 축적되어 있는 상태이다. SIA는 특정 정책이나 프로그램, 혁신 등이 사회와 구성원들에게 미치는 영향에 대해 광범위하게 분석·조사하기 때문에 통상 사회학, 심리학,

인구통계학, 사회연구 방법론 등 다양한 분야의 전문가들이 모여서 학제 간 또는 초학제 간 연구로 진행된다(Esteves et al., 2012).

SIA를 효과적이고 투명하게 진행하기 위해서 다양한 지침들이 개발되어 활용되고 있다. 대표적으로 국제영향평가협회(International Association for Impact Assessment, IAIA)(2013)는 SIA 수행을 위한 5가지 주요 지침을 제시하였다. 첫째, 영향평가를 통해 인간사회에서 발생할 수 있는 불평등과 갈등 같은 주요 과제를 해결해야 한다. 둘째, 부정적 영향을 줄이고 구성원들의 혜택과 복지를 증진시키는 긍정적인 기여를 해야 한다. 셋째, 투명성, 참여, 대중을 포함한 모든 관련 이해관계자의 참여를 촉진하고 의사결정에 반영한다. 넷째, 영향평가의 과정과 결과에 대한 양질의 정보를 보장하여 영향평가의 효율성을 높인다. 다섯째, 영향평가 이후의 주요 결정에 대해서도 의사결정을 도모한다. SIA 분야의 대가인 Vanclay(2003)는 SIA 시 고려해야 할 제반사항으로 형평성, 계획된 개입의 사회적 영향 예측 가능성, 부정적 사회적 영향을 감소시키고 긍정적 사회적 영향을 강화시킬 것, 지속가능한 개발에 중점을 둘 것, 경제적 이익과 사회적 비용 사이의 중재자 역할을 담당할 것, 그리고 민주적 절차를 지킬 것 등을 제시하였다.

SOPACT(2024) 또한 SIA 가이드라인을 제시하였는데, 이 가이드라인에서는 국제표준을 준수할 것, 이해관계자를 포함할 것, 투명성을 보장할 것, 문화적 민감성 고려할 것 등을 포함한다. 이와 함께 SOPACT(2024)는 SIA 과정의 절차로서 다음과 같은 5단계를 제안하였다. 첫째, 스크리닝(screening)과 스코핑(scoping) 단계에서는 1)프로젝트나 정책에 대한 SIA 필요성을 판단한 후 2)평가 범위와 규모를 결정하고, 3)평가할 주요 사회적 이슈를 정의한 뒤 4)정책이나 프로그램의 영향을 받게 될 이해관계자를 식별하고, 5)평가의 지리적, 사회적, 시간적 경계를 설정한다. 둘째, 이해관계자 분석(stakeholder analysis)단계에서는 영향을 받는 개인, 그룹 또는 커뮤니티와 소통하여 그들의 우려, 기대, 삶에 미칠 수 있는 잠재적 영향을 파악한다. 셋째, 기초연구(baseline studies) 단계에서는 영향을 받을 인구의 현재 사회적 조건에 대한 정성적 및 정량적 기초 데이터를 수집하고 연구의 기준점을 설정한다. 넷째, 영향 예측(impact prediction) 단계에서는 SIA 대상의 잠재적인 긍정적, 부정적 사회적 변화를 예측한다. 다섯째, 완화 계획(mitigation planning)단계에서는 긍정적인 결과를 강화하는 동시에 부정적인 영향을 제거하거나 최소화하는 전략을 개발한다. SIA 측정방법으로는 1)분야별 특성을 반영한 지표 개발, 2)이해관계자들을 대상으로 설문과 인터뷰 실시, 3)유사 사례 연구, 4)프로젝트 실행 중과 실행 후에 사회적 영향을 평가하기 위한 지속적인 추적 프로세스 구현을 제안하였다(SOPACT, 2024). 종합해서 SIA의 지침이나 권고를 살펴보면, SIA의 방법론과 실행방식은 다양하나, 그 궁극적 목적이 지속가능하고 공정하며 사회적으로 책임 있는 의사결정을 촉진하는 데 있음을 알 수 있다.

교육 분야에서도 다양한 정책, 프로그램, 기술 등에 대하여 SIA가 시행되어 왔다. 교육에서의 SIA의 목표는 새로운 교육 정책, 프로그램, 또는 기술의 도입이 학생, 교사, 학부모, 그리고 지역사회에 미치는 영향을 분석하여 새로운 정책과 기술의 도입이 교육에 미치는 영향을 사전에 예측하고, 시행 후의 영향을 평가하여 정책을 개선하는 한편, 궁극적으로는 구성원들의 삶의 질 향상에 도움을 주는데 있다(Bornmann, 2013). 교육 분야에서의 성공적인 SIA는 학생과 교사의

성과를 극대화하고 지역사회의 불만이나 부정적 영향을 최소화하며, 해당 정책이나 프로그램, 기술에 대한 신뢰를 향상시키는 데 기여한다. 예를 들어, Gillani와 Eynon(2023)의 연구에서는 학습에서의 디지털 플랫폼 사용이 교육의 사회적, 문화적, 경제적, 정치적 측면에서 갖는 의의를 조사하고 이를 통해 공평한 교육과 삶에서의 공평한 결과를 촉진하고자 하였다.

외국어 교육 분야의 SIA 사례를 살펴보면, OECD(2020)는 외국어 학습에서의 사회경제적 격차에 대한 연구를 수행하고 사회경제적 배경에 따라 외국어 교육의 접근성과 학습 성과에 차이가 있음을 밝혔다. 즉, 이 연구에서는 고소득 가정의 학생들이 더 많은 외국어 학습 기회를 가지며 더 높은 외국어 능력을 보이는 것으로 나타났다. OECD는 이 결과를 바탕으로 외국어 교육의 접근성을 높이는 정책의 필요성을 부각시켰다. Guerard(2021)는 미국 유학생들의 영어 능력, 사회적 유대, 차별 경험이 학업 성과에 미치는 영향을 조사하였는데, 영어 능력이 높고 사회적 유대가 강한 유학생들은 더 좋은 학업 성과를 얻은 반면, 차별 경험은 부정적 영향을 미치는 것으로 나타났다. Guerard(2021) 또한 연구 결과를 바탕으로, 부정적 효과를 감소시키기 위한 정책적 방안 마련을 제안하였다. 국내 사례로는 Park 외(2024)이 영어 AI 코스웨어 활용 현황 분석을 근거로 하여 영어 AIDT 영향평가 설계 및 실행방안 연구를 수행하였다. 이 연구는 국내 상황에서의 SIA에 대한 가이드라인과 방법론을 제시하는 시도를 하였으나, 이는 AIDT 시행 전에 수행된 것으로 AI에 대한 기술 도입에 따른 사회변화를 다루는 데에는 한계가 있었다. 현재까지 국내에서는 교육에서의 AI와 관련된 SIA 연구를 찾아보기 어려운 실정이다.

2. AI 기반 교육의 장점 및 잠재적 위험 요인

AI 기술을 활용한 에듀테크 시장규모는 현재 빠른 속도로 팽창하고 있으며, 2030년에는 약 8,000억에 이를 것으로 전망되고 있다(Samil PwC Business Research, 2024). 교육 분야에서 AI는 AI가 갖고 있는 사용성, 편리성, 교육적 장점 등으로 인하여 학습자, 교수자, 관리자 측면에서 다양한 목적으로 사용되고 있다. 학습자 측면에서는 지능화된 콘텐츠(smart content)를 제공하고 지능형 튜터링 시스템(intelligent tutoring system)에 기반한 맞춤형 피드백을 통하여 개인화된 학습을 지원한다(Ayeni et al., 2024; Park et al., 2024; Shin, 2023). 따라서 일대다 방식의 수업이 진행되는 전통적인 교실상황에서는 거의 불가능하였던 학습자 맞춤형 학습이 가능하게 됨에 따라 교육의 공정성과 형평성을 향상시키고 교육 격차를 감소시키는 효과를 기대할 수 있다(Bang & Lee, 2024; Khan, 2024). 교수자 측면에서는 수행평가의 자동화, 학습자와 학습에 대한 분석 데이터 제공과 수업 콘텐츠의 효과적인 큐레이션(curation) 지원을 통해 교수 업무를 경감시키고 교수 효율성을 증대시킬 수 있다(Hong et al., 2024; Kim et al., 2020). 관리자 측면에서는 행정 업무의 자동화를 통한 업무 효율성 향상과 학습자들에 대한 과학적 관리를 가능하게 한다(Do et al., 2022; Tahiru, 2021).

그러나 다른 한편으로 교육에서의 AI 활용은 여러 잠재적인 위험 요인을 내포하고 있다. 기술적 측면에서는 AI 기반 교육시스템에서 데이터와 알고리즘의 편향성 문제가 발생할 수 있다. Suresh와 Gutttag(2021)에 의하면, AI에서는 머신러닝 생애주기 전반에 걸쳐 다양한 편향성이 발생하는데, 이는

데이터 수집, 처리, 배포, 운영, 모델 개발 및 학습 단계 중 어디서나 발생할 수 있으며, 이런 단계를 거치는 동안 편향성은 누적되어 나타난다. 더욱이, 데이터 편향은 알고리즘 편향과 결합하여 전체적인 편향성을 증폭시키는 결과를 초래한다. 교육 환경에서의 데이터 편향은 특정 그룹에 불공정한 결과를 부여함으로써 교육의 형평성을 저해하고, 학습 성과를 떨어뜨리며, AI에 대한 사회적 신뢰와 윤리적 가치를 훼손하게 된다(Bang & Lee, 2024; Byun et al., 2023). 교육적 측면에서는 AI의 활용이 AI에 대한 지나친 의존도를 초래할 수 있고, 학습자의 상위인지 발달을 저하시키며, 교사-학습자 간, 학습자-학습자 간의 상호작용을 감소시켜 학습자의 사회정서적 영역에서 부정적인 결과를 가져올 수 있다는 우려 또한 제기되고 있다(Cho et al., 2024; Lim et al., 2024; Park et al., 2024).

윤리적 측면에서도 교육에서의 AI 활용은 여러 잠재적 위험을 내포하고 있다. 우선, 데이터와 알고리즘의 투명성(transparency, AI 시스템 작동 방식, 의사 결정 과정 및 결과에 대한 정보를 명확하게 공개하여 이를 신뢰할 수 있도록 함)과 설명가능성(explainability, AI 활용 상황에서 관련 데이터, 설계, 절차가 이해/설명 가능해야 함)이 보장되어야 하는데, 앞서 지적한 기술적 측면에서의 위험성인 AI 알고리즘의 불투명성은 의사결정 과정에서 구성원들에게 사회적 신뢰를 저하시킬 수 있다(Suresh & Guttag, 2021). 이는 공정성(fairness, AI 시스템이 특정 집단에게 편향적이지 않고 모든 사용자에게 평등하고 공정한 결과를 제공하도록 설계되고 운용되어야 함)과 교육의 형평성을 해치게 되어 결과적으로 교육에서 의도하지 않은 악영향을 초래할 수 있다(Holmes, 2023; Holstein & Doroudi, 2021; OECD, 2024; Suresh & Guttag, 2021). 또한 편향성 문제가 발생했을 때 책임소재를 명확하게 가리기 어려운 책임성(responsibility, AI 시스템의 개발, 배치 및 사용의 궁극적 책임은 인간에게 있으며, 문제 발생 시 책임 주체를 명확히 규정해야 함)의 문제도 존재한다(Stahl, 2023). 학습자의 사회경제적 배경, 신체적 장애, AI 디지털 리터러시 등에 따라 접근성(accessibility, 사용자의 사회경제적 배경, 신체적 장애, 기술적 환경 등에 관계없이 모든 사용자가 AI 기술과 서비스를 동등하게 이용할 수 있는 정도)과 포용성(inclusiveness, AI 시스템이 다양한 사용자 그룹의 특성과 요구사항을 고려하여 설계되어, 특정 집단이 소외되거나 차별 받지 않고 모든 사용자가 포함될 수 있도록 하는 특성)에서도 차이가 발생할 수 있으며, AI 사용 시 개인정보 보호의 문제도 상존한다(OECD, 2024).

교육에서의 AI 활용이 이러한 잠재적 위험 요소를 내포함에 따라 국내외에서는 AI 사용에 대한 권고사항과 가이드라인을 제시하고 있다. OECD(2019)는 AI 원칙으로 데이터 투명성과 공정성, 책임성, 포용적 성장, 안정성 및 보안을 포함한 5대 기본 원칙을 제시하였다. EU(2024)가 제시한 AI Act도 투명성, 안전성, 차별 방지 등을 필수 요건으로 포함하고 있으며, UNESCO(2021)의 AI 윤리 권고는 AI의 4대 가치로 인권·자유·인간 존중, 환경 및 생태계 번영, 다양성 및 포용성 보장, 평화롭고 정의로우며 상호 연결된 삶을 권고하고 있다. 국내에서는 KERIS가 AI 학습용 데이터 윤리 가이드라인(2023)과 AIDT 자체기술검증 가이드라인(2024)을 통해 교육 분야 AI 윤리 대원칙으로 '사람의 성장을 지원하는 AI'를 제시하고, AI 학습용 데이터 윤리 프레임워크의 6개 범주로 동의, 투명성, 보안, 공정성, 합목적성, 거버넌스를 포함하였다. AI의 일반적인 사용에 대한 법령, 가이드라인, 규제, 개인정보보호법 등이 이미 상당히 배포가 되어 있으나, 아직은 교육에서의 AI

Digital Learning(2024)										
Ministry of Education of Korea(2022)	O	O	O	O	O	O	O	O	O	O
KERIS(2023, 2024)	O	O	O	O	O	O	O	O	O	O

III. AI 기반 교육의 SIA 지표 설정

SIA를 위해 지표 설정은 필수적인 단계이다. 지표는 특정 프로젝트나 정책이 교육 환경과 사회 전반에 미치는 영향을 체계적으로 측정하고 분석하는 데 있어 객관적이고 일관된 기준을 제공하게 된다. 앞에서 설명한 것처럼, AI는 맞춤형 학습, 교사 지원, 행정 효율화와 같은 긍정적 영향을 줄 수 있지만, 동시에 디지털 격차, 알고리즘 편향, 개인정보 보호 문제와 같은 부정적 영향을 초래할 가능성을 갖는다(Unite. AI, 2023). 지표 설정은 AI 기술이 교육에서 어떤 방식으로 형평성과 포용성 등과 같은 교육의 가치를 촉진하거나 저해하는지를 구체적으로 파악하는데 도움이 되며(Stahl et al., 2023), 이는 정책 수립 및 실행 과정에서 중요한 데이터로 활용될 수 있다(Shin, 2022).

SIA를 위한 지표를 설정할 때 정량적(예: 학업 성취도, AI 도구 사용 빈도) 및 정성적(예: 사용자 경험, 윤리적 수용성) 요소를 모두 포함해야 AI에 의한 복잡한 사회적 영향을 충분히 반영할 수 있게 된다(SOPACT, 2024). 이를 위해 현장의 이해관계자(학생, 교사, 학부모 등)와 전문가 의견을 반영하여 다층적이고 협력적인 방식으로 지표를 도출하는 것이 중요하다. 지표는 반복적으로 검토 및 조정 가능해야 하며, 한 번 설정한 뒤에 고정하는 것이 아니라 기술 발전과 교육 환경 변화에 따라 문화적, 사회적, 경제적 차이를 고려하여 동적으로 적용할 수 있어야 한다.

AI 기술이 도입되는 상황에서 교육 분야가 주목해야 하는 가치에 대한 논의가 있어왔다(Bura & Myakala, 2024; OECD, 2024). 교육적 가치와 교육 분야에서 추구하는 일반적인 목표를 참고하여, 형평성(equity), 포용성(inclusiveness), 접근성(accessibility), 교육 격차 감소(reduced educational inequality), 데이터 편향(data bias), 개인정보 보호(privacy protection), 기술 의존성(technology dependency), 사회적 신뢰도(social trustworthiness), 학습 성과(learning outcomes), 사회적 상호작용(social interaction)을 SIA를 위한 지표로 설정해 볼 수 있다(<Table 2>). 본 연구에서 제시된 지표는 AI 활용 시 지켜야 할 AI 윤리에 대한 국내외 가이드라인 및 법령 등을 근거로 설정되었는데(<Table 1> 참조), 이는 추후 교육 분야 이해관계자들의 심층 논의와 실제 현장 적용을 통한 검증 절차를 거쳐 확정하는 단계를 필요로 한다.

Table 2
Examples of Indicators for SIA of AI Use/AI Systems in Education

Indicators	Description	Question(example)
Equity	The extent to which AI services provide equal opportunities for diverse users, including gender, region, economic background, and minority groups.	<ul style="list-style-type: none"> - Did AI have a negative impact on certain groups (e.g., economically disadvantaged students, students with disabilities)? - Did it reduce gender or regional differences in learning outcomes?
Inclusiveness	The extent to which AI services are accessible to and useful for marginalized groups, ensuring their active participation and engagement in learning: it concerns the ease of access and participation in learning for marginalized groups, such as students with disabilities, students from multicultural backgrounds, and students with learning difficulties.	<ul style="list-style-type: none"> - Does the service include accessibility features for students with visual and hearing impairments? - Does the AI system work effectively in multicultural and multilingual environments?
Accessibility	The extent to which AI services are accessible without technical, economic or geographical barriers.	<ul style="list-style-type: none"> - Can low-income families easily use the AI services? - Are alternatives provided in regions with limited access to digital devices or the Internet?
Equality	The extent to which academic achievement gaps are reduced across different regions, schools, and socioeconomic backgrounds.	<ul style="list-style-type: none"> - Has the academic achievement gap between rural and urban areas narrowed after the introduction of AI? - Has the learning tool contributed equitably to improving academic performance?
Data bias	Whether the data and algorithms used by AI are biased against particular groups.	<ul style="list-style-type: none"> - Does the training data reflect different social and cultural backgrounds? - Did the learning recommendation disadvantage certain groups (e.g., certain genders or regions)?
Privacy protection	The extent to which students' personal data is securely protected and whether data management policies are transparent.	<ul style="list-style-type: none"> - Is the collection, storage, and deletion of learning data ethical and legal? - Has any data been shared with third parties without students' consent?
Technology dependency	How AI services respect the autonomy of human teachers and learners in educational settings	<ul style="list-style-type: none"> - Has a heavy reliance on AI reduced the creative role of teachers or students? - Are alternative learning methods available in the event of technical failure?
Social trustworthiness	The extent to which students, teachers, and parents perceive AI	<ul style="list-style-type: none"> - Can the AI explain why it produced certain learning outcomes?

	services to be fair and reliable	- Are users satisfied with the AI service?
Learning outcomes	Whether AI services have actually improved student motivation and achievement	- Has the AI-based learning recommendation system improved students' academic performance? - Did the provision of personalized learning paths increase learner satisfaction and engagement?
Social interaction	The impact of AI adoption on the interactions between teachers and students, as well as among students	- Has AI technology increased or decreased interaction-based learning? - Has the lack of face-to-face interaction led to a decline in emotional empathy?

SIA를 위한 지표가 설정되었으면, 그 다음으로 각 지표에 대한 다각적인 분석을 위해 데이터 수집을 정량적 및 정성적으로 진행하여 결합할 필요가 있다. 정량적 데이터는 AI의 영향을 객관적으로 측정할 수 있는 근거를 제공하며, 정성적 데이터는 숫자로 표현되지 않는 미묘한 사회적 변화와 경험적 차이를 포착한다. 정량적 방법으로는 설문조사와 통계 데이터 활용이 있을 수 있다. 설문조사는 학습자, 교수자, 학부모를 대상으로 형평성, 접근성, 학습 성과 등에 대해 점수화된 응답을 수집하며, 결과는 특정 지표에 대한 인식을 수치화하여 비교할 수 있다(Saseanu et al., 2024). 통계 데이터는 학업 성취도, 교육 참여율, 학습 결과 격차와 같은 객관적 지표를 기반으로 AI 도입 전후의 변화를 분석하는 데 활용될 수 있다(Han et al., 2024). 정성적 데이터 수집 방법으로는 심층 인터뷰와 사례 연구가 있다. 심층 인터뷰는 학습자와 교수자의 경험과 인식을 깊이 탐구하여, AI 도입이 실제로 어떻게 교육 환경과 사회적 상호작용에 영향을 미쳤는지에 대한 질적 통찰을 제공하게 된다(SOPACT, 2024). 예를 들어, 전문가와의 논의를 통해 학습자가 느낀 포용성이나 교수자가 경험한 AI 도구의 편리성 및 한계점을 분석할 수 있다. 사례 연구는 특정 학교나 지역을 대상으로 AI 서비스의 효과를 평가하며, 이 과정에서 구체적인 사례를 통해 교육 격차 감소와 같은 지표에 대해 입증 가능한 데이터를 제공받을 수 있게 된다.

AIDT는 학습자 중심으로 설계된 도구로 개인화된 학습 경험과 다문화적 이해를 촉진하며 교육적 격차를 해소할 가능성을 가지고 있다(Hong et al., 2024; OECD, 2024). SIA는 도입 초기부터 학습자와 교사, 지역사회의 요구를 파악하고, 디지털 격차, 알고리즘 편향, 문화적 다양성 부족 등 부정적 영향을 사전에 완화하는 데 기여할 수 있다. 특히 정책 결정자와 교육자들에게 기술의 윤리적, 사회적 책임을 강조하며 데이터 수집 및 AI 활용 방식을 투명하게 유지할 수 있는 프레임워크를 제공한다. 이를 통해 AIDT의 긍정적 효과를 극대화하고 부정적 효과를 최소화하는 방안을 마련할 수 있다.

교육에서의 AI 활용에 대한 SIA 수행 시, 각 교과목의 특성에 따라 SIA의 초점은 달라질 수 있다. 예를 들어, 수학 교과목에서는 문제 풀이 형식과 데이터 편향 여부가 주요 과제로, 특정 학습 유형에 치우치지 않는 공정한 설계가 중요하다. 정보 교과목에서는 최신 기술 트렌드를 반영하는 교육 자료와 디지털 리터러시를 지원하는 실습 환경이 초점이 될 수 있다. 이는 기술 격차를 해소하고 미래 지향적인 학습 기회를 제공하는 데 기여한다. 영어 교과목은 읽기, 쓰기, 듣기, 말하기 역량을

훈련하며, 발음 교정과 문화적 이해를 포함하는 상호작용형 학습 자료가 중요한 특징이다(Han et al., 2024). 이러한 과정은 학습자의 사회적 상호작용과 다문화적 포용성을 직접적으로 지원하는 것과도 관련이 있다. 특히 영어 교육에서 사회적 영향 평가를 수행하는 것이 중요한데, 이는 언어와 문화적 포용성이 학습자의 사회적 관계와 글로벌 역량 형성에 미치는 영향이 크기 때문이다. AI가 학습자의 발음과 억양을 공정하고 정확하게 인식하고, 다양한 억양과 문화적 맥락을 반영한 콘텐츠를 제공하도록 설계되었는지 확인하는 것이 필요하다(Han et al., 2024; Holmes, 2023). 또한 AI 피드백은 학습자가 문화적으로 민감하고 포용적인 태도를 기를 수 있도록 돕는 것이 중요하다(Baker & Hawn, 2022).

IV. AI 기반 교육의 사회적 영향 평가 프로세스

본 연구에서는 AI 기반 교육의 SIA 프로세스를 다음과 같이 5단계로 제안하고자 한다(<Figure 1>). 1단계는 사회적 영향(Social Impact) 지표(Indicator) 설정이다. 형평성, 포용성, 접근성, 교육 격차, 데이터 편향 등 주요 평가 지표를 정의한다. 2단계는 지표별 평가 질문 설계이다. 각 지표의 영향을 측정할 수 있는 질문을 작성하며, 예컨대 형평성은 "성별, 지역, 계층 간 학습 자원이 공정하게 배분되었는가?"로 설정할 수 있다. 3단계는 긍정/부정 영향의 점수화이다. 각 질문에 대해 특정 범위 내에서 점수를 부여하며(예: 0~10), 이를 통해 긍정적·부정적 영향을 계량화한다. 4단계는 영향 점수 계산 및 시각화이다. 평가자들의 점수를 평균하여 긍정적·부정적 영향을 비교하고, 이를 그래프로 시각화하여 개선이 필요한 영역을 도출할 수 있다. 마지막으로 5단계는 정책 개선 방안의 제안이다. 디지털 격차 해소, 데이터 편향 제거, 포용성 확대 등 지표별 부정적 영향을 줄이기 위한 구체적인 개선 전략을 제안할 수 있다. 이 프로세스는 AI 기반 교육 도입의 사회적 영향을 체계적으로 분석하고, 정책 설계와 실행의 방향성을 제시하는 데 유용하다.

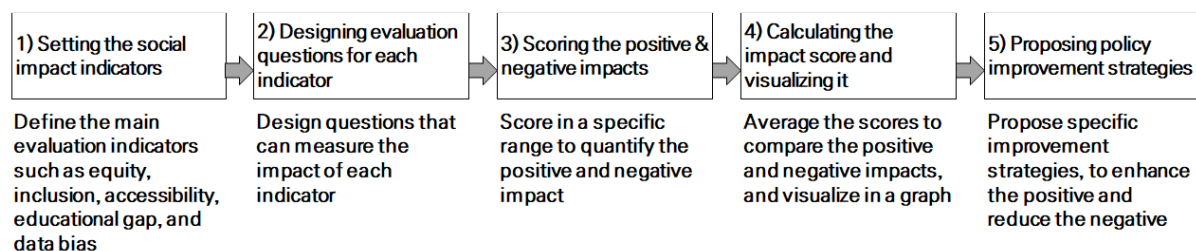


Figure 1

Framework for Social Impact Assessment

1단계인 사회적 영향 지표 설정과 2단계인 지표별 평가 질문 설계는 III장에서 다루었다. 각 지표에 대한 질문은 다수일 수 있으며 그 지표의 최종 점수 산정을 위해 평균값을 취할 수 있다.

3단계에서는 2단계의 지표별 질문에 대해 긍정/부정 영향을 평가하여 점수화한다. 점수화 방법은 각 지표에 대해 0~10 또는 -10~10 범위로 할 수 있다. 본 논문에서는 긍정적 영향과 부정적 영향 부분에 대해 각각 0~10으로 점수화하였다. <Table 3>에서 각 지표별 질문, 긍정/부정 평가, 관련 사례를 예를 들어 설명하였다.

Table 3
Example of SIA of AI Use/AI-powered Systems in Education

Indicator	Item(question)	Score (example)	Evaluation (example)	Sample case (example)
Equity	Are learning resources and outcomes distributed equitably across gender, economic class, and geography?	6/10	Positive	Students from the lowest 30% of income households receive the same level of learning materials and opportunities in AI digital textbooks.
Inclusiveness	Are underrepresented groups, such as students with disabilities, multicultural students, and low-income students, supported to participate in learning?	7/10	Positive	Increase learning engagement for students from multicultural backgrounds (e.g., voice assistants) and visually impaired students (e.g., multilingual support.)
Accessibility	Are AI digital textbooks easily accessible to all students and teachers?	8/10	Positive	Limited access to AI systems due to low rate of device penetration in remote areas.
Achievement gap	Has it helped reduce achievement gaps between regions or schools?	5/10	Positive	Achievement gap between urban and rural schools narrowed after use of AI.
Data bias	Has the AI system not treated certain groups unfairly (e.g., gender, income level)?	4/10	Negative	AI makes better recommendations for students from higher-income families based on training data, identifying a data-biased problem.

4단계에서는 긍정/부정 영향 점수를 계산하고 시각화하는데, 이때 두 가지의 대표적 방식이 가능하다. 첫번째는 각 지표마다 사람들의 평균 점수를 계산하여 확정하는 방식이다. 예를 들어, 여러 사람이 형평성 지표에 6점, 9점, 7점, 9점, 9점을 부여하여 $40/5 = 8$ 점이 되는 경우이다. 이 때 각

지표별 가중치를 설정하여 중점 지표를 선택적으로 지정할 수 있다. 두번째는 사람들 각자가 개별로 지표 전체에 대해 점수를 기입하여 전체 평균 점수를 긍정과 부정으로 나누어 계산하는 방식이다. 단, 이 경우에는 지표별 가중치를 설정하여 계산하기 어렵다는 단점이 있다.

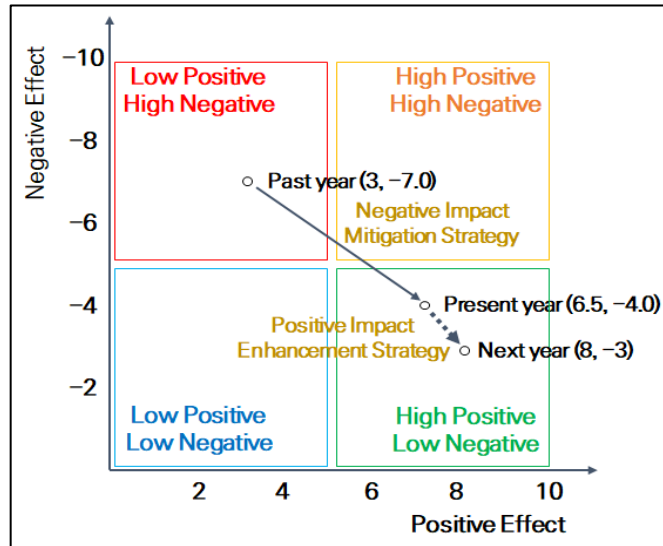


Figure 2

Example of Visual Representation of Positive/Negative Impact Scores

이를 바탕으로 SIA 결과를 시각화하여 표현할 수 있으며, 이는 직관적으로 의사결정을 하는데 도움이 된다(<Figure 2>). 이는 교육 분야 정책 및 전략 수립 뿐 아니라, AIDT의 도입 효과를 객관적으로 평가하고 개선이 필요한 영역을 시각적으로 표현하거나, 정책 입안자와 개발자 등이 효과적인 활용 및 운영 방안을 설계하는 데 활용이 가능하다. 시각화 단계에서는 평가된 긍정적 영향(X축)과 부정적 영향(Y축)을 기준으로 그래프를 작성하여, 각 사례를 네 가지 사분면 중 하나에 배치한다. 이 과정은 AI의 도입 가능성과 개선 필요성을 명확히 판단하는 역할을 수행한다. 그래프의 사분면은 다음과 같이 해석된다.

- 긍정적 영향과 부정적 영향이 모두 강한 경우(High Positive, High Negative): 부정적 영향을 줄이는 개선 전략 필요.
- 긍정적 영향은 강하고 부정적 영향은 약한 경우(High Positive, Low Negative): AI 도입이 적합한 상황으로 판단.
- 긍정적 영향은 약하고 부정적 영향이 강한 경우(Low Positive, High Negative): 도입 여부를 재검토하거나 재설계 필요.
- 긍정적 영향과 부정적 영향이 모두 약한 경우(Low Positive, Low Negative): 긍정적 효과를 강화하기 위한 방안 모색 필요.

위의 시각화 예시에서 X축의 긍정적 영향 점수가 6.5로 접근성과 포용성에서 높은 점수를

받았고 Y축의 부정적 영향 점수는 -4.0으로 데이터 편향 문제가 주요 부정적 요소로 가정되었다. 이러한 위치는 긍정적 영향이 부정적 영향보다 우세하며, 데이터 편향 문제를 개선할 경우 더 큰 사회적 효과를 기대할 수 있음을 짐작할 수 있다.

필요에 따라 3차원 그래프를 활용하여 Z축을 추가하면 더욱 정교한 분석이 가능하다. 예를 들어, Z축에 학습 성과, 사회적 신뢰도, 교수자와 학습자 간 상호작용과 같은 지표를 포함하여 부정적 영향이 있음에도 학습 성과가 개선되는지 확인할 수 있다.

5단계인 정책 개선 방안 제안에서는 AI 기반 교육의 사회적 영향을 분석한 결과를 바탕으로 긍정적 영향을 극대화하고 부정적 영향을 최소화하기 위한 구체적인 실행 전략을 수립하는 과정이다. 이 단계에서는 분석된 문제 영역별로 적절한 개선 방안을 도출한다. 예를 들어, 접근성 개선을 위해 농어촌 지역에 디지털 기기와 인터넷 보급을 확대하거나, 교육 격차 해소를 위해 교육 취약 지역에 AIDT를 활용한 보충 학습 지원을 강화하는 정책을 제안할 수 있다. 또한, 데이터 편향 문제를 해결하기 위해 다양한 배경의 학습 데이터를 추가 확보하고 AI 알고리즘의 공정성을 평가할 수 있는 도구를 도입하며, 포용성을 높이기 위해 장애 학생과 다문화 가정 학생을 위한 사용자 친화적 기능을 지속적으로 개발하는 방안을 제안할 수 있다. 이러한 정책적 개선 방안은 사회적 영향을 체계적으로 관리하고 AI 도입의 효과를 극대화하며, 교육의 공정성과 포용성을 강화하는 데 기여한다.

영어 교과 AIDT 도입 시 사회적 영향 평가는 서비스 특징과 도입 대상의 맥락에서 긍정적 및 부정적 영향을 정량적으로 분석하고, 이를 기반으로 교육적 효과를 극대화하고 한계를 개선하기 위한 근거를 만든다. 본 연구에서는 제안된 SIA 프레임워크와 평가 지표의 실제적 적용을 예시하고자, 영어 교과에서의 AIDT 도입을 가상 시나리오로 설정하여 구체적인 평가 사례를 제시하였다(<Table 4>). 이 가상 시나리오에서는 중학교 영어 교과에 AIDT를 도입하여 읽기, 쓰기, 듣기, 말하기 네 가지 역량을 고르게 강화하고, 학생 개개인의 학습 요구를 충족시키는 것을 목표로 설정했다. 주요 서비스 특징으로는 맞춤형 학습 경로, 실시간 발음 교정, 다문화 콘텐츠 제공, 그리고 즉각적인 AI 피드백이 포함되며, 도입 대상은 도시와 농촌 지역의 전국 중학교와 그 학생 및 교사로 가정한다. 긍정적 영향 측면에서, 학생마다 수준별 콘텐츠를 제공하여 개인화된 학습이 가능하며, 음성 인식 기술을 통해 발음과 억양을 교정하여 발화 능력을 향상시킬 수 있다(Jeon et al., 2021; Khan, 2024). 또한, 다양한 문화와 억양을 반영한 콘텐츠는 글로벌 감각을 강화하고 다문화 학생의 학습 참여를 확대할 수 있다. 교사에게는 반복적인 피드백 작업을 AI가 대체함으로써 심화된 교수 활동에 집중할 시간을 제공할 수 있다. 이 환경에서의 부정적 영향으로는 농어촌 지역의 인터넷 품질 저하 및 디지털 기기 부족이 학습 참여를 제한하거나, 기술 의존성이 증가하여 독립 학습 능력을 약화시킬 수 있는 점을 포함하였다(Han et al., 2024; Park et al., 2024). 또한, 외국어 학습에서의 대표적인 AI 문제인 AI 학습 데이터가 특정 억양에 편향되거나 학생들 간 대면 상호작용이 감소할 가능성도 함께 포함하였다(Holmes, 2023).

Table 4
Example of SIA of English AI Digital Textbook

Category	Item(example)	Method
Accessibility	<ul style="list-style-type: none"> - AI can deliver content at each student's level so they can learn at their own pace. - AI can provide the same learning resources to everyone. 	<ul style="list-style-type: none"> - Analyzing statistics on access to learning materials (urban vs. rural) - Investigation of device and network accessibility
Digital divide	<ul style="list-style-type: none"> - Students in remote areas may not have access to real-time learning opportunities due to poor internet quality. - Lack of digital devices at home can limit engagement in learning. 	<ul style="list-style-type: none"> - Investigation of internet connectivity - Investigation of digital device penetration
Learning outcomes	(Examples of pronunciation correction) <ul style="list-style-type: none"> - Speech recognition technology instantly corrects students' pronunciation and intonation to improve their speaking skills. - AI analyzes training data to reduce pronunciation errors. 	<ul style="list-style-type: none"> - Examining student learning outcomes - Student satisfaction surveys, student feedback
Cultural inclusiveness	<ul style="list-style-type: none"> - AI can help improve students' global awareness by providing materials that reflect different accents and cultural backgrounds. - AI can increase engagement for minority language speakers and students from multicultural backgrounds. 	<ul style="list-style-type: none"> - Analyzing cultural diversity of content. - Measuring multicultural student satisfaction
Effective teaching environment (teacher)	<ul style="list-style-type: none"> - AI can save teachers' time by providing feedback to students (e.g., grammar, pronunciation correction). - AI can allow teachers to focus on more creative and interactive activities. 	<ul style="list-style-type: none"> - Measuring teachers' time savings - Teacher interviews and surveys
Technology dependency	<ul style="list-style-type: none"> - There is a potential risk that students will become overly dependent on AI, undermining their ability to learn independently. - Technical errors negatively impact student learning. 	<ul style="list-style-type: none"> - Technology dependency surveys - Analysis of learning experiences during technical failures
Data bias	<ul style="list-style-type: none"> - AI training data may be biased toward certain accents and may not be representative of a wide variety of accents. 	<ul style="list-style-type: none"> - Analyzing diversity of accents in AI training data - Comparing student learning outcomes by accent
Social interaction	<ul style="list-style-type: none"> - As students rely on AI, there is the potential for less interaction with teachers and less interaction between learners. - There is a potential risk that students' face-to-face communication skills will be diminished by the use of AI. 	<ul style="list-style-type: none"> - Examining frequency of teacher-student interactions - Examining student engagement in collaborative learning

<Table 4>에 있는 지표들에 점수를 임의로 가정하여 부여하고 해석한 결과는 다음과 같다. X축은 긍정적 영향과 관련된 지표들로부터 점수(0~10점)를 산정한다. 접근성 = 8점, 발음 교정 = 9점, 문화적 포용성 = 7점, 교사 업무의 효율성 향상 = 8점이라고 가정할 때, 평균 점수는 8.0점이 된다.

Y축은 부정적 영향과 관련된 지표들로부터 점수(-10~0점)를 산정한다. 디지털 격차 = -6점, 기술 의존성 = -4점, 데이터 편향 = -3점, 사회적 상호작용 감소 = -5점이라고 할 때, 평균 점수는 -4.5점이 된다. 이러한 가정에서 긍정적 효과가 부정적 효과를 상회하며, AIDT의 영어 서비스는 효과적일 수 있다고 해석된다. 또한, 부정적 영향인 기술 의존성과 디지털 격차 문제를 개선하면 더 높은 사회적 수용성을 기대해볼 수 있다고 해석된다. 이러한 상황에서 개선 방안을 제안할 수 있다. 기술 의존성 감소 방안으로 AI와 대면 수업을 혼합한 학습 모델을 적용하거나, 데이터 다양성 강화 방안으로 다양한 억양과 문화적 맥락을 반영한 학습 데이터를 추가 확보하거나, 디지털 격차 완화 방안으로 농어촌 지역에 인터넷 인프라와 디지털 기기 보급을 확대해 볼 수 있다.

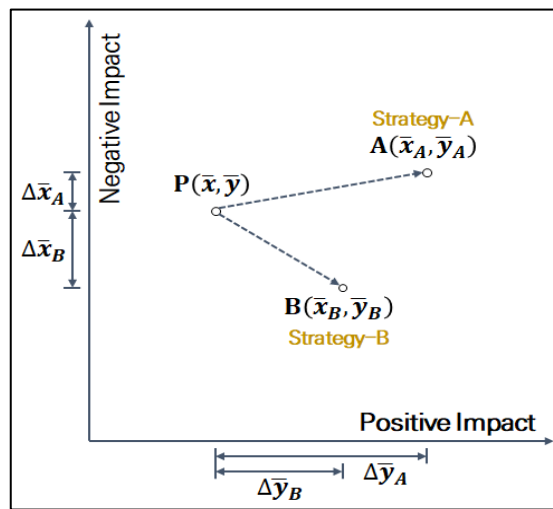


Figure 3

Example of Evaluating the Choice Between Strategy A and Strategy B

SIA를 시행한 이후에 결과를 바탕으로 전략적 정책 시행이 가능하다. <Figure 3>에서 전략A는 긍정적인 영향을 많이 높이지만 부정적인 영향도 일부 높아질 수 밖에 없는 전략이며, 전략B는 긍정적인 영향을 약간 높이고 부정적인 영향을 약간 낮출 수 있는 전략이다. 현재 SIA에 대한 좌표 위치를 (\bar{x}, \bar{y}) , $\bar{x} = (x_1 + x_2 + \dots + x_M)/M$, $\bar{y} = (y_1 + y_2 + \dots + y_N)/N$ 일 때, \bar{x} 는 x_1, x_2, \dots, x_M 에 대한 M 개의 긍정 지표들의 각 평가값을 나타낸다. \bar{y} 는 y_1, y_2, \dots, y_N 에 대한 N 개의 부정 지표들의 각 평가값을 나타낸다. 전략A와 전략B의 좌표 위치를 각각 (\bar{x}_A, \bar{y}_A) 와 (\bar{x}_B, \bar{y}_B) 인 경우도 마찬가지로, $\bar{x}_A = (x_{A,1} + x_{A,2} + \dots + x_{A,M})/M$, $\bar{y}_A = (y_{A,1} + y_{A,2} + \dots + y_{A,N})/N$, $\bar{x}_B = (x_{B,1} + x_{B,2} + \dots + x_{B,M})/M$, $\bar{y}_B = (y_{B,1} + y_{B,2} + \dots + y_{B,N})/N$ 로 설정한다.

여기서 $\bar{x}_A = \bar{x} + \Delta\bar{x}_A$, $\bar{y}_A = \bar{y} + \Delta\bar{y}_A$, $\bar{x}_B = \bar{x} + \Delta\bar{x}_B$, 그리고 $\bar{y}_B = \bar{y} + \Delta\bar{y}_B$ 이다. 긍정 지표들의 수와 부정 지표들의 수가 동일하다고 가정하자($M = N$). 만약, 특정 지표에 대해 선택적 통제를 하는 상황이고 지표 x_1, x_2, \dots, x_M 의 값들이 x_k 를 제외하고 고정되어 있다면, $\frac{(x_{A,1} + x_{A,2} + \dots + x_{A,k} + \dots + x_{A,N})}{N} = \frac{(x_1 + x_2 + \dots + x_k + \dots + x_N)}{N} + \Delta\bar{x}_A$ 에서 $x_{A,k} = x_k + N \cdot \Delta\bar{x}_A$ 로 표현될 수 있다. 만약 각 지표에 가중치가 있다면,

앞의 표현식은 $w_k \cdot x_{A,k} = w_k \cdot x_k + N \cdot \Delta \bar{x}_A$ 이고 이는 다시 $x_{A,k} = x_k + N \cdot \Delta \bar{x}_A / w_k$ 로 표현할 수 있는데, 이는 2개 이상의 특정 지표에 대해 선택적으로 통제를 해야 하는 상황에서 유용하게 활용할 수 있다. 이와 유사하게 $y_{A,k}$, $x_{B,k}$, $y_{B,k}$ 도 계산할 수 있다. 이러한 값이 계산되면 전략A와 전략B의 기대되는 사회적 영향을 고려하여 전략 수립이 가능해진다.

V. 결론

본 연구에서는 선행연구와 국내외 AI 윤리 지침 및 가이드라인을 분석하여 AI가 교육 분야에서 미칠 사회적 영향을 평가하기 위한 프레임워크와 이에 필요한 지표, 절차, 방법을 제안하였다. AI 기반 교육 시스템의 SIA는 기술 도입이 교육 및 사회에 미치는 긍정적, 부정적 영향을 체계적으로 분석하여 공정하고 포용적인 교육 환경을 조성하는 데 핵심적 역할을 한다(Vanclay, 2003; Vanclay et al., 2019). 교육 분야에서의 SIA는 교사, 학습자, 관리자, 개발자 등 다양한 이해관계자의 관점을 포괄적으로 수용하고, AI가 각 구성원에게 미치는 영향을 체계적으로 분석·예측해야 한다. 이를 통해 AI의 긍정적 효과는 극대화하고 부정적 영향은 최소화할 수 있는 실행 가능한 구체적 방안을 도출하는 것이 필요하다(Stahl, 2023). AIDT 도입 이후에는 각 평가 지표에 대한 체계적 분석을 위해 AIDT를 비롯한 교육용 디지털 플랫폼에서의 데이터 수집 방법이 구체적으로 논의되어야 하며, 이를 위한 교육 분야 SIA 인프라가 조속히 구축될 필요가 있다.

AI 기술이 교육 현장에 급속도로 확산되고 있으며 이해관계자들에게 미칠 영향력이 지대함을 고려할 때, 교육에서의 AI 활용에 대한 사회적 영향 평가는 시급히 수행되어야 할 중요 과제이다. 그럼에도 불구하고 국내에서는 교육 분야 AI 도입과 활용에 대한 SIA 수행 사례를 찾아보기 어려우며, 이에 대한 체계적인 프레임워크나 가이드라인도 부재한 실정이다. 이러한 맥락에서 본 연구는 교육에서의 AI를 평가하는 프레임워크를 제안하였다. 본 연구에서 제안한 프레임워크는 형평성, 포용성, 접근성, 교육 격차, 데이터 편향, 개인정보 보호, 기술 의존성, 사회적 신뢰도, 학습 성과, 사회적 상호작용을 주요 평가 지표로 설정하고, 5단계의 평가 절차(지표 설정, 평가 질문 설계, 영향 점수화, 결과 시각화, 정책 개선 방안 제안)를 통해 체계적인 영향 평가가 가능하도록 하였다. 특히 정량적, 정성적 평가 방법을 결합하여 AI 도입의 복합적 영향을 종합적으로 분석할 수 있도록 설계하였으며, 시각화 도구를 통해 의사결정자들이 직관적으로 결과를 이해하고 개선 방향을 도출할 수 있도록 하였다. 본 연구에서 제안하는 프레임워크는 형평성, 포용성, 데이터 편향, 접근성 등 다양한 평가 지표를 중심으로 구체적인 평가 프로세스와 방법론을 제시함으로써, 정책 입안자와 교육자들이 AI 기술의 사회적 영향을 예측하고 실효성 있는 개선 방안을 모색하는 데 실질적인 도움을 줄 수 있을 것이다. 본 연구는 이를 통해 AI 도입 초기 단계에서부터 윤리적, 사회적 문제를 선제적으로 파악하고 체계적으로 관리함으로써 지속 가능하고 책임 있는 기술 활용을 지원하는 데 기여하고자 하였다. 본 연구에서 개발된 프레임워크는 국내외 관련 문서의 포괄적 분석을 바탕으로 하여, AIDT를 포함한 모든 교육용 AI 시스템의 사회적 영향 평가에 활용될 수 있을 것이다.

본 연구는 이론과 문헌연구를 기반으로 프레임워크를 제시하였다는 점에서 실제 교육 현장에서의 검증이 필요하다는 한계를 지닌다. 후속 연구에서는 본 프레임워크를 실제 교육 환경, 특히 2025년 도입 예정인 AIDT와 같은 구체적 사례에 적용하여 그 실효성을 검증하고, 다양한 이해관계자의 의견을 수렴하여 본 연구의 프레임워크를 보완할 필요가 있다. 또한 교과별 특성을 반영한 세부 평가 지표 개발과 실제 데이터에 기반한 장기적 영향 분석 연구도 수행되어야 할 것이다. 이러한 지속적인 연구와 실천을 통해 교육 분야의 AI 도입이 형평성과 효과성을 제고하고, 사회적 신뢰를 증진하며, 모든 구성원을 위한 포용적 학습 환경 조성에 기여할 수 있을 것으로 기대된다.

REFERENCES

- Australian Government Department of Education. (2023, November). *Australian framework for generative artificial intelligence (AI) in Schools*. Australian Government Department of Education. <https://www.education.gov.au/schooling/resources/australian-framework-generative-artificial-intelligence-ai-schools>
- Ayeni, O., Hamad, N., Chisom, C., Osawaru, B., & Adewusi, O. (2024). AI in education: A review of personalized learning and educational technology. *GSC Advanced Research and Reviews*, 18(2), 261-271. <https://doi.org/10.30574/gscarr.2024.18.2.0062>
- Baker, R. S., & Hawn, A. (2022). Algorithmic bias in education. *International Journal of Artificial Intelligence in Education*, 32, 1052-1092. <https://doi.org/10.1007/s40593-021-00285-9>
- Bang, J., & Lee, S-M. (2024). Addressing bias for equity in AI-driven customized learning systems. *Multimedia Assisted Language Learning*, 27(4), 70-86. <https://doi.org/10.15702/mall.2024.27.4.70>
- Becker, H. A. (2001). Social impact assessment. *European Journal of Operational Research*, 128(2), 311-321. [https://doi.org/10.1016/S0377-2217\(00\)00074-6](https://doi.org/10.1016/S0377-2217(00)00074-6)
- Bloom, B. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, 13(6), 4-16. <http://dx.doi.org/10.3102/0013189X013006004>
- Bornmann, L. (2013). What is societal impact of research and how can it be assessed? a literature survey. *Journal of the American Society for Information Science and Technology*, 64(2), 217-233. <https://doi.org/10.1002/asi.22803>
- Bura, C., & Myakala, P. K. (2024). Advancing transformative education: Generative AI as a catalyst for equity and innovation. <https://doi.org/10.48550/arXiv.2411.15971>
- Byun, S., Song, S., Lee, C., Kim, J., Bang, E., Jang, E., & Jang, H. (2023). *The research on ethical guidelines of learning data for artificial intelligence (RR-2023-03)*. KERIS. <https://www.keris.or.kr/main/ad/pblcte/selectPblcteRRInfo.do?mi=1138&pblcteSeq=13683>
- Center for Innovation, Design, and Digital Learning. (2024). *Inclusive intelligence: The Impact of AI on education for all learners*. <https://ciddl.org/inclusive-intelligence-the-impact-of-ai-on-education-for-all-learners/>
- Cho, Y., Xie, Y., & Kim, H. (2024). Exploring factors influencing interaction between learners and AI. *Journal of*

- Korean Association for Educational Information and Media*, 30(6), 1819-1845.
<http://dx.doi.org/10.15833/KAFEIAM.30.6.1819>
- Do, J., Eur, J., Na, Y. J., & Kim, S. (2022). A study of teachers' use and perception of learning analytics based dashboard for customized education. *The Journal of Korean Teacher Education*, 39(4), 261-289.
<https://doi.org/10.24211/tjkte.2022.39.4.261>
- Esteves, A., Franks, D., & Vanclay, F. (2012). Social impact assessment: the state of the art. *Impact Assessment and Project Appraisal*, 30(1), 34-42. <https://doi.org/10.1080/14615517.2012.660356>
- European Commission. (2024). *AI Act*. <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>
- European Parliament. (2021). *Report on artificial intelligence in education, culture and the audiovisual sector*. (A9-0127/2021). Committee on Culture and Education. https://www.europarl.europa.eu/doceo/document/A-9-2021-0127_EN.pdf
- Gillani, N., & Eynon, R. (2023). Understanding and improving social factors in education: A computational social science approach. In T. Yasseri (Ed.), *Handbook of computational social science* (pp. 1-23). Edward Elgar Publishing.
- Guerard, J. (2021). *The impact of English language proficiency, social connectedness and perceived discrimination on academic performance and self efficacy amongst international students* [Doctoral dissertation, Florida State University]. FSU Research Repository. https://purl.lib.fsu.edu/diginole/2021_Summer_Guerard_fsu_0071E_16642
- Han, J., Yoo, H., Myung, J., Kim, M., Lee, T. Y., Ahn, S.-Y., & Oh, A. (2024). *RECIPE4U: Student-chatGPT interaction dataset in EFL writing education*. *ArXiv, abs/2403.08272*. <https://doi.org/10.48550/arXiv.2403.08272>
- Holmes, W. (2023). *The unintended consequences of artificial intelligence and education*. *Education International*. <https://www.ei-ie.org/en/item/28115:the-unintended-consequences-of-artificial-intelligence-and-education>
- Holstein, K., & Doroudi, S. (2021). Equity and artificial intelligence in education: Will "AIEd" amplify or alleviate inequities in education? *ArXiv, abs/2104.12920*. <https://doi.org/10.48550/arXiv.2104.12920>
- Hong, S., Hwang, Y., Park, Y., & Lee, S. (2024). Expectations and concerns about adopting AI digital textbooks: Based on investigation of teachers' use of AI and digital tools. *The Journal of Studies in Language*, 40(1), 7-20. <https://doi.org/10.18627/jslg.40.1.202405.7>
- IAIA. (2009). Social impact assessment key citations [online]. Fargo, ND: International Association for Impact Assessment. Retrieved January 10, 2025, from <http://www.iaia.org/resources-networking/key-citations.aspx>
- Jeon., H., Lee, S-M., & Park, I. (2021). A systematic review of AI technology use in English education. *Multimedia Assisted Language Learning*, 24(1), 87-103. <https://doi.org/10.15702/mall.2021.24.1.87>
- KERIS. (2023). *AI digital textbook development guidelines* (GM 2023-11). Ministry of Education, KERIS. <https://www.keris.or.kr/main/ad/pblcte/selectPblcteETCInfo.do?mi=1142&pblcteSeq=13722>
- KERIS. (2024). *Guidelines for self-evaluation of technologies in AI digital textbook* (GM 2024-04 AI-PM-03). Ministry of Education, Korea Institute for Curriculum and Evaluation, KERIS, Korea Foundation for the Advancement of Science & Creativity. <https://www.keris.or.kr/main/ad/pblcte/selectPblcteETCInfo.do?m>

- i=1142&pblctSeq=13791
- Khan, S. (2024). *Brave new words: How AI will revolutionize education (and Why That's a Good Thing)*. Viking.
- Kim, H., Park, J. H., Hong, S., Park, Y., Kim, E., Choi, J., & Kim, Y. (2020). Teachers' perceptions of AI in school education. *Journal of Educational Technology*, 36(3), 905-930. <https://doi.org/10.17232/KSET.36.3.905>
- Lee, E.-B. (2024). *The study on legislative approach for AI impact assessment (24-20-①)*. Korea Legislation Research Institute. <https://www.klri.re.kr/kor/publication/2261/view.do>
- Lee, J. H., & Ahn, S. (2023). Case analysis for developing key functions of AI digital textbooks. *Journal of Creativity and Information Culture (JCIC)*, 9(4), 379-387. <https://doi.org/10.32823/jcic.9.4.202311.379>
- Lim, E., Hong, S., & Lim, J. Y. (2024). Effects and limitations of participatory design for teaching with digital textbooks. *Journal of Educational Information and Media*, 30(2), 303-323. <https://doi.org/10.15833/KAFEIAM.25.4.767>
- Ministry of Education of Korea. (2022, August 11). *Principles for AI ethics to support human growth* [Press release]. <https://www.moe.go.kr/boardCnts/viewRenew.do?boardID=294&lev=0&statusYN=W&s=moe&m=020402&opType=N&boardSeq=92297>
- Ministry of Education of Korea. (2023, February 23). *Digital-driven education reform plan announced* [Press release]. <https://english.moe.go.kr/boardCnts/viewRenewal.do?boardID=265&boardSeq=94073&lev=0&searchType=null&statusYN=W&page=1&s=english&m=0201&opType=N>
- Nuany, M., Judijanto, L., Nurliyah E., Muriyanto, M., & El-Farra, S. (2022). Impact of AI in education and social development through individual empowerment. *Journal of Artificial Intelligence and Development*, 1(2), 89-97. Retrieved from <https://edujavare.com/index.php/JAI/article/view/301>
- OECD. (2019). *Scoping principles to foster trust in and adoption of AI – Proposal by the expert group on artificial intelligence at the OECD*. OECD. <https://doi.org/10.1787/d62f618a-en>
- OECD. (2020). *The socio-economic gap in foreign language learning*. OECD. https://www.oecd.org/en/publications/the-socio-economic-gap-in-foreign-language-learning_953199e1-en.html
- OECD. (2024). *The potential impact of artificial intelligence on equity and inclusion in education*. OECD. https://www.oecd.org/en/publications/the-potential-impact-of-artificial-intelligence-on-equity-and-inclusion-in-education_15df715b-en.html
- Office of Educational Technology. (2024). *Designing for education with artificial intelligence: An essential guide for developers*. <https://tech.ed.gov/files/2024/07/Designing-for-Education-with-Artificial-Intelligence-An-Essential-Guide-for-Developers.pdf>
- Park, H., Kim, H., Maeng, E., & Lee, S.-M. (2024). *Design and implementation strategies for impact evaluation of AI digital English textbook (CRR2024-06)*. Korean Educational Development Institute. <https://www.kedi.re.kr/khome/main/research/selectPubForm.do>
- Samil PwC Business Research. (2024, February). *The revolution of hyper-personalized learning begins: EduTech*. Samil Insight. <http://pwc.com/kr/ko/insights/samil-insight.html>
- Saseanu, A. S., Gogonea, R.-M., & Ghita, S. I. (2024). The social impact of using artificial intelligence in education. *Amfiteatru Economic*, 26(65), 89-105. <https://doi.org/10.24818/EA/2024/65/89>

- Shin, D. (2023). Case study on English test item development training for secondary school teachers using AI tools: Focusing on ChatGPT. *Language Research*, 59(1), 21-42. <https://doi.org/10.30961/lr.2023.59.1.21>
- Shin, S. (2022). A study on implementation plan for AI service impact assessment. *The Journal of the Institute of Internet, Broadcasting and Communication*, 22(5), 147-157. <https://doi.org/10.7236/JIIBC.2022.22.5.147>
- SOPACT. (2024). *Social impact assessment guide*. https://www.sopact.com/guides/social-impact-assessment?utm_source=chatgpt.com
- Stahl, B. C. (2023). Embedding responsibility in intelligent systems: from AI ethics to responsible AI ecosystems. *Scientific Reports*, 13, 7586 (2023). <https://doi.org/10.1038/s41598-023-34622-w>
- Stahl, B. C., Antoniou, J., Bhalla, N., Brooks, L., Jansen, P., Lindqvist, B., Kirichenko, A., Marchal, S., Rodrigues, R., Santiago, N., Warso, Z., & Wright, D. (2023). A systematic review of artificial intelligence impact assessments. *Artificial Intelligence Review*, 56, 12799-12831. <https://doi.org/10.1007/s10462-023-10420-8>
- Suresh, H., & Gutttag, J. (2021). A framework for understanding sources of harm throughout the machine learning life cycle. *Equity and Access in Algorithms, Mechanisms, and Optimization*, 17, 1-9. <https://doi.org/10.1145/3465416.3483305>
- Tahiru, F. (2021). AI in education: A systematic literature review. *Journal of Cases on Information Technology (JCIT)*, 23(1), 1-20. <https://doi.org/10.4018/JCIT.2021010101>
- TeachAI. (2023). *AI guidance for schools toolkit*. Code.org, CoSN, Digital Promise, European EdTech Alliance, Larimore, J., PACE. <https://www.teachai.org/toolkit>
- The Institute for Ethical AI in Education. (2021). *The ethical framework for AI in education*. <https://www.buckingham.ac.uk/wp-content/uploads/2021/03/The-Institute-for-Ethical-AI-in-Education-The-Ethical-Framework-for-AI-in-Education.pdf>
- UNESCO. (2021). *Recommendation on the ethics of artificial intelligence*. <https://unesdoc.unesco.org/ark:/48223/pf0000381137>
- UNICEF. (2021). *Policy guidance on AI for children*. UNICEF, Ministry of Foreign Affairs of Finland. <https://www.unicef.org/innocenti/reports/policy-guidance-ai-children>
- Unite. AI. (2023). *Thought leaders AI and educational equity: A blueprint for closing the gap*. Retrieved January, 9, 2025, from <https://www.unite.ai/ai-and-educational-equity-a-blueprint-for-closing-the-gap/>
- Vanclay, F. (2003). International principles for social impact assessment. *Impact Assessment and Project Appraisal*, 21(1), 5-12. <https://doi.org/10.3152/147154603781766491>
- Vanclay, F., Baines, J. T., & Taylor, L. (2019). Reflections on social impact assessment in the 21st century. *Impact Assessment and Project Appraisal*, 37(1), 1-11. <https://doi.org/10.1080/14615517.2019.1685807>
- Vanclay, F., Esteves, A. M., Aucamp, I., & Franks, D. M. (2015). *Social impact assessment: Guidance for assessing and managing the social impacts of projects*. International Association for Impact Assessment. https://www.iaia.org/uploads/pdf/SIA_Guidance_Document_IAIA.pdf
- Washington Office of Superintendent of Public Instruction. (2024). *Human-centered AI guidance for K-12 public schools*. <https://ospi.k12.wa.us/sites/default/files/2024-06/comprehensive-ai-guidance.pdf>