

# 생성형 AI 기반 지능형 학생 중심 과정 평가 시스템 개발 연구

(Development of an Intelligent, Student-Centered Process Evaluation System Based on Generative AI)

이해인\*

(Haein Lee)

## 요약

본 연구는 2022 개정 교육과정에서 강조하는 과정 중심 평가의 실질적 운영을 지원하기 위해 생성형 AI 기반 학생 중심 과정 평가 시스템(SCPES)을 개발하는 것을 목적으로 한다. 본 시스템은 과목-단원-성취기준-성취수준 선택 구조를 기반으로 학생 정보, 학습진단수준, 교사 진단노트를 입력하면 학습 역량·태도 분석, 개념 정의, 학습 전략, 탐구 활동, 서술형 문항 및 해설을 자동 생성한다. 알고리즘은 성취기준 의미 분석, 인지적 요구 분석, 학생 수준 기반 난이도 조절, 자연어 생성 및 품질 검증 단계로 설계되었다. 실제 교육과정 사례 적용 결과, 동일 성취기준이라도 학생 수준에 따라 전략, 문항, 탐구 활동이 효과적으로 차별화되어 생성됨을 확인하였다. 이는 SCPES가 개인화된 과정 평가와 교사 업무 경감에 기여할 수 있음을 시사한다.

■ 중심어 : 생성형 AI ; 과정 중심 평가 ; 성취기준 ; 학생 맞춤형 평가 ; 교육용 AI 시스템

## Abstract

This study aims to develop SCPES, a generative AI-based student-centered process evaluation system designed to support the practical implementation of the 2022 revised Korean curriculum. SCPES automatically generates learning competency analysis, learning attitude descriptions, concept explanations, learning strategies, inquiry tasks, and constructed-response items with explanations based on selected subjects, units, achievement standards, achievement levels, and student information. The system's core algorithms include achievement-standard semantic analysis, cognitive-demand analysis, student-level alignment, content generation, and quality verification. Application of SCPES to curriculum-based scenarios demonstrated that learning strategies, task difficulty, and inquiry activities were effectively differentiated according to student proficiency levels. These results suggest that SCPES can enhance personalized process evaluation and reduce teacher workload.

■ keywords : Generative AI ; Process Evaluation ; Achievement Standards ; Personalized Learning ; Educational AI System

## I. 서론

### 1. 연구 배경과 필요성

2022 개정 교육과정은 학습자의 변화를 중심으로 평가를 설계하도록 강조한다[1]. 이는 단순

한 결과 중심 평가에서 벗어나 학습 과정에서 나타나는 사고의 흐름, 문제 해결을 위한 전략적 선택, 협력 과정에서의 상호작용, 학습 태도의 안정성 등을 포함한 다면적 평가를 의미한다. 그러나 실제 교육 현장에서 교사가 학습자의 과정을 면밀히 관찰하고 그 결과를 성취기준에 맞추어

\* 정회원, 충남도립대학교 컴퓨터공학과

이 논문은 2011년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No. 2000-000000).

정리하며 학생별 맞춤 피드백까지 제공하는 것은 현실적으로 큰 부담이 된다[2]. 교사가 수업과 행정 업무를 병행하면서 동시에 평가 문항 개발, 과정 기록 작성, 성취기준 해석 등의 작업을 수행해야 하는 상황에서 평가의 질을 유지하기 어렵다는 문제도 지속적으로 제기되고 있다. 또한 성취기준 해석 방식이 교사마다 달라 평가 결과의 일관성이 확보되지 않는다는 점 역시 중요한 한계로 지적된다. 이러한 문제를 해결하기 위해서는 교사의 전문적 판단을 보완하면서도 일관된 기준을 유지할 수 있는 기술적 지원 도구가 필수적이며, 최근 급속히 발전한 생성형 인공지능 기술은 이러한 요구를 충족할 수 있는 적절한 대안으로 주목받고 있다[3],[4].

서론의 마지막에는 본 연구에서 다룬 과정 중심 평가 업무의 흐름을 시각적으로 정리한 그림을 제시할 수 있다. 그림 1은 교육과정 해석, 수업 관찰, 기록 작성, 피드백 제공, 평가 조정이라는 과정 중심 평가의 전체 흐름을 나타내며, 이러한 모든 단계가 교사에게 큰 업무량으로 다가온다는 점을 명확히 보여준다. 연구의 목적은 바로 이 복잡한 과정을 자동화하여 교사가 본질적인 교육 활동에 더 집중할 수 있도록 지원하는데 있다.



그림 1. 과정 중심 평가 전체 프로세스

## 2. 연구 목적

본 연구는 생성형 인공지능 기반 자연어 생성 기술을 활용하여 학생 중심 성취기준 기반 평가

를 자동화하는 시스템을 개발하는 것을 목표로 한다. 구체적으로는 학생의 학습 수준과 성향, 교사 진단 정보, 교육과정의 성취기준을 입력하면 AI가 학습자의 수준에 적합한 평가 문장을 자동으로 생성하도록 설계하였다. 이를 통해 교사의 기록 업무를 대폭 줄이고, 성취기준 정합성을 유지하며, 동시에 학생별로 차별화된 피드백을 제공하는 평가 생태계를 구축하는 데 목적이 있다.

## II. 관련 연구

### 1. 과정 중심 평가와 2022 개정 교육과정

과정 중심 평가는 학습자의 수행 과정 전반을 평가하여 그 과정에서 나타나는 강점과 약점을 파악하고, 이후 학습 방향을 설정하는 데 활용하는 평가 방식이다[2]. 이러한 평가는 학습자의 다양한 특성을 드러내기 때문에 교육적으로 큰 의의가 있지만, 교사가 지속적으로 관찰 자료를 수집하고 이를 해석해야 하는 부담이 존재한다. 2022 개정 교육과정은 이러한 과정을 근거 기반으로 운영할 것을 강조하며, 성취기준 기반 평가, 성장 중심 평가, 수업 - 평가 - 기록의 일체화를 통해 평가의 실효성을 높이도록 요구하고 있다 [1], [6]. 이러한 변화는 교사에게 더 많은 정보 처리와 판단을 요구하므로 기술적 지원이 필요하다.

### 2. AI 기반 평가 자동화 연구 동향

대규모 언어모델을 활용한 자동화 연구는 주로 문제 생성, 피드백 문장 작성, 학습 전략 제시 등 자연어 처리 기능을 활용한 교육적 응용 사례로 확장되고 있다[3],[4]. 기존 연구에서는 학습자의 능력 수준을 예측하거나, 기본적인 문장 템플릿에 기반한 피드백을 자동으로 생성하는 방식이 널리 사용되었다. 그러나 템플릿 기반 출력은 표현의 다양성이 낮고, 학생의 실제 수행과 성취기

준의 의미를 정확히 반영하지 못하는 경우가 많았다. 또한 교육과정의 구조적 특성과 연계성을 반영한 AI 모델 개발이 부족하다는 한계도 존재했다. 본 연구는 이러한 한계점을 보완하기 위해 성취기준의 의미를 정교하게 분석하고 학생 수준에 따라 난이도를 조절할 수 있는 구조적 알고리즘을 설계했다는 점에서 차별성을 갖는다.

관련 연구 비교는 표 1을 통해 제시될 수 있다.

표 1. 기존 연구 vs SCPES 비교표

구분	기존 연구	SCPES
성취기준 연계성	템플릿 중심, 낮은 연계성	성취기준 의미 기반 자동 분석
개인화 수준	제한적 개인화	학습 수준 기반 난이도 조절
피드백 생성 방식	고정 문장 중심	전략·문항·탐구 자동 생성
정합성 검증	부재 또는 약함	규칙 기반 정합성 검증 포함
데이터 기반성	단일 유형 데이터	대규모 사전 + 학생부 데이터 기반

표 1은 기존 연구들이 활용한 AI 기법, 평가 자동화 범위, 개인화 수준, 교육과정 연계성 유무와 같은 요소를 비교하여 SCPES가 기존 연구 대비 구조적 완성도와 개인화 능력이 높다는 점을 시각적으로 보여준다.

### III. 학생 중심 과정 평가 시스템(SCPES)설계

#### 1. 시스템 개요 및 요구 분석

SCPES(Student-Centered Process Evaluation System)는 학생의 학습 수준, 성향 정보(MBTI 등), 교사 진단 메모 그리고 성취기준을 입력받아 학습자의 수준에 적합한 분석 문장과 평가 요소를 자동으로 산출하도록 개발된 시스템이다. 이 시스템은 과정 중심 평가를 실제 현장에서 운영할 때 발생하는 교사의 과중한 기록 업무와 성취기준 해석의 어려움을 해결하기 위한 목적으

로 설계되었으며, 학습자 한 명 한 명의 특성을 반영하는 개인화된 평가를 보다 구조적이고 일관된 방식으로 제공하기 위해 고안되었다.

요구 분석 과정에서 다수의 교사들은 성취기준을 직접 탐색하고 해석하는 작업에 상당한 시간이 소요될 뿐 아니라, 해석 과정에서 교사마다 관점과 표현 방식이 달라 평가 문서의 일관성이 유지되기 어렵다는 문제를 반복적으로 언급하였다. 특히 과정 중심 평가는 학생의 학습 과정에서 나타나는 비교적 세밀한 행동 변화, 사고 전략, 개념 적용 방식 등을 모두 기록해야 하기 때문에 정량 평가보다 훨씬 더 높은 기록 노동을 요구한다. 이러한 기록은 단순히 문장을 적는 행위를 넘어, 학생의 수행을 해석하고 성취기준과 연계하여 의미를 구성하는 전문적 과정이기 때문에 교사에게 상당한 인지적 부담을 초래한다.

성취기준 기반 평가를 정확히 수행하기 위해서는 교사가 해당 기준을 의미 단위로 분해하여 핵심 개념과 요구 기능을 명확히 이해해야 한다. 그러나 많은 교사들은 단위 내 다수의 성취기준을 빠르게 파악하기 어렵고, 기준마다 요구하는 사고 수준이나 개념 요소가 다르기 때문에 이를 수업과 평가에 일관되게 적용하는 데 어려움을 겪고 있다. 또한 평가 문장을 작성할 때 교사마다 사용하는 문장 표현과 구조가 크게 다르기 때문에 동일한 수행 상황에 대한 평가 문장도 교사에 따라 이질적으로 작성되는 경우가 많다. 이런 비표준성은 학교 내 평가의 신뢰도와 공정성 확보에 위협 요소로 작용한다.

교사들은 학생별 정성적 기록을 개별적으로 작성하는 과정이 가장 큰 부담이라고 응답하였다. 예를 들어 학습 과정에서 드러난 이해도, 전략의 적절성, 태도의 특징 등을 일관성 있게 기록하려면 상당한 시간이 필요하며, 학급당 학생 수가 많을 경우 이러한 기록 작성은 현실적으로 지속되기 어렵다. 특히 고교학점제가 확대되면서 다양한 선택 과목을 동시에 운영해야 하는 교사들은 평가 문항 개발, 피드백 작성, 성취수준 판단

까지 모두 수행해야 하므로 업무량은 더욱 증가하는 상황이다.

이러한 여러 문제점을 해결하기 위해 SCPES는 자동화된 성취기준 로딩 기능, 자연어 기반 평가 문장 생성 기능, 성취기준과의 정합성 검증 기능, 학생 수준을 기반으로 한 차별화된 문장 생성 기능을 필수 요소로 포함하였다. 시스템의 전체 구조는 그림 2의 흐름으로 나타낼 수 있다.

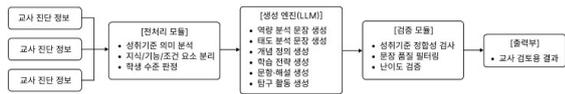


그림 2. SCPES 전체 시스템 구조도

입력 단계에서는 학생 정보와 교사 진단 정보가 수집되며, 전처리 단계에서는 성취기준 의미 분석과 학생 수준 분류가 이루어진다. 생성 단계에서는 대규모 언어모델을 기반으로 한 자연어 생성 기술이 작동하여 전략, 문항, 개념 정의, 탐구 활동 등 다양한 평가 요소가 자동 산출된다. 이후 검증 단계에서 생성된 문장이 성취기준과 의미적으로 일치하는지, 표현 오류는 없는지 확인한다. 마지막 출력 단계에서는 교사가 쉽게 검토하고 수정할 수 있도록 자연스러운 문장 형태로 결과가 제시된다.

이러한 일련의 과정은 교사의 전문성이 반영된 입력 정보를 기반으로 하되, 시간과 노력이 많이 드는 반복적 기록 업무는 AI가 처리하도록 설계된 구조이며, 교사는 최종 판단과 결과 조정 역할에 집중할 수 있게 된다. 이는 교사가 학생의 학습 성장에 보다 본질적으로 기여할 수 있도록 돕는다는 점에서 교육적 의의가 크다.

## 2. 시스템 구성과 기능 설명

SCPES는 입력 계층, 성취기준 분석 계층, 생성 계층, 검증 계층, 출력 계층의 다섯 개 계층으로 구성되며, 각각의 계층은 독립적인 기능을 수행

하면서도 서로 유기적으로 연결되어 하나의 평가 자동화 체계를 형성한다. 이 계층 구조는 교사가 수행하는 평가 과정을 단계별로 모사한 것으로, 과정 중심 평가의 실제 업무 흐름을 분석한 결과를 기술적으로 구현한 형태라고 할 수 있다.

첫 번째 계층인 입력 계층에서는 학생 정보와 교사 진단 정보가 입력된다. 학생의 학습 수준, 성향 정보, 교사의 관찰 기반 진단 메모 등이 입력되며, 이러한 정보는 이후 생성 과정에서 문장 내용의 난이도, 전략의 구체성, 탐구 활동의 복잡도 등을 조절하는 데 활용된다. 과목과 단원을 선택하면 해당 단원에 속하는 성취기준이 자동으로 로딩되며, 교사는 목록에서 원하는 성취기준을 선택하기만 하면 된다. 이 과정은 기존에 교사가 교육과정 문서를 직접 탐색해야 했던 시간을 크게 절약한다.

이러한 입력 계층의 실제 인터페이스 구성은 아래 그림 3에서 확인할 수 있다.



그림 3. SCPES 시스템의 학생 정보 및 성취기준 입력 화면

두 번째 계층인 성취기준 분석 계층에서는 선택된 성취기준 문장이 자동 분해되어 의미적 구조가 분석된다. 성취기준 문장은 일반적으로 지식 요소, 기능 요소, 상황 요소로 이루어져 있는데, 시스템은 이를 형태소 분석과 의미 단위 분리를 통해 구조화한다. 또한 성취기준에 포함된 행동 동사를 기반으로 요구되는 인지 수준을 판단하고, 이전 단원이나 선행 개념과의 연계성을 분석함으로써 생성 단계에서 제공해야 할 설명이나 전략의 깊이를 결정한다.

세 번째 계층인 생성 계층은 SCPEs의 핵심 기능이 작동하는 부분이다. 생성형 AI는 앞선 단계에서 분석된 성취기준 의미 구조와 학생 정보를 결합하여 학습 역량 분석 문장, 태도 분석 문장, 개념 정의, 학습 전략, 탐구 활동, 서술형 문항 및 해설 등을 자동 생성한다. 이 과정에서 모델은 학생 수준에 따라 문장의 복잡성을 조절하고, 교사 진단 메모에서 나타난 특징을 반영하여 실제 교실 상황과 유사한 평가 문장을 생성한다. 예컨대, 문제 해결 과정에서 조급함이 있는 학생이라면 전략 문장에 '접근 방법을 신중히 점검할 필요가 있다'와 같은 서술이 자연스럽게 포함된다.

네 번째 계층인 검증 계층에서는 생성된 문장의 성취기준 정합성과 표현 품질을 점검한다. AI 언어모델은 높은 자연어 생성 능력을 갖추고 있지만, 때때로 성취기준 핵심 개념을 누락하거나 지나치게 복잡한 문장을 생성하는 경우가 있다. 이러한 문제를 방지하기 위해 시스템은 성취기준의 필수 요소가 포함되었는지, 인지적 요구 수준이 적절한지, 문장 간 논리적 연결이 자연스러운지 확인하는 검증 절차를 거친다. 오류가 발견되면 문장을 자동으로 재생성하거나 수정하며, 이를 통해 교사가 별도의 검토 과정을 거치지 않아도 일정 수준 이상의 품질이 확보된 기록을 제공받을 수 있다.

마지막 계층인 출력 계층에서는 검증된 문장을 교사가 검토하기 쉬운 형태로 제시한다. 출력된 문장은 즉시 활용 가능한 자연어 서술 형태로 제공되며, 교사가 필요에 따라 일부 표현을 조정하거나 학생의 실제 수행 상황에 맞게 보완할 수 있다. 이렇게 생성된 결과는 학교생활기록부 및 과정 중심 평가 기록 문서 작성 시 바로 적용할 수 있는 수준을 목표로 한다.

최종적으로 출력되는 평가 화면의 예시는 아래 그림 4와 같다.

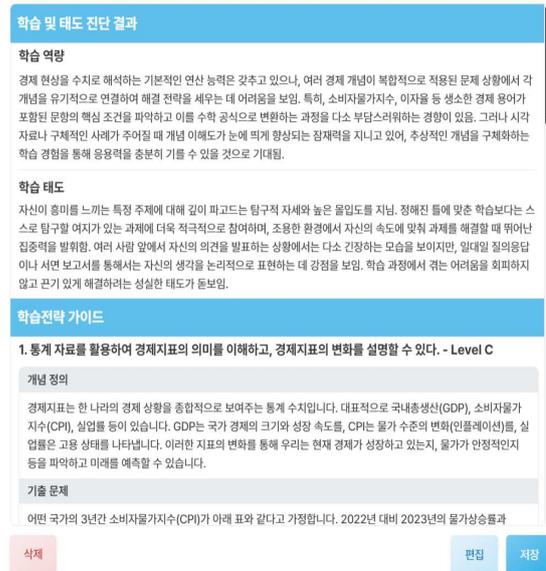


그림 4. SCPEs 시스템의 최종 과정 평가 기록 출력 화면

이와 같은 다층 구조는 과정 중심 평가의 복잡한 흐름을 효율적으로 자동화하면서도 교사의 전문성을 유지·보완하는 방향으로 설계되었다. 시스템은 단순히 평가 문장을 생성하는 도구가 아닌, 성취기준 기반의 평가 패러다임을 실제 학교 현장에서 실천할 수 있도록 돕는 종합적 지원 도구로 기능한다.

## IV. 알고리즘 설계

### 1. 성취기준 분석 및 개인화 알고리즘

성취기준 분석 과정은 성취기준 문장을 구성하는 의미 단위, 즉 지식 요소, 기능 요소, 조건 요소를 분해하여 그 의미적 구조를 파악하는 것으로 시작된다. 알고리즘은 이러한 요소들을 자동으로 추출하고, 문장 내에서 이들이 어떤 관계를 형성하고 어떤 학습 요구를 제시하는지 의미적으로 해석한다. 이후 Bloom의 인지 수준 분류 체계를 기반으로 성취기준에 포함된 사고 활동을 분석하여 해당 기준이 요구하는 인지적 깊이를 판별한다[7]. 이러한 인지 수준 분석은 단순히 행

동 동사를 분류하는 데 그치지 않고, 성취기준이 담고 있는 전체적 사고 요구를 파악하여 학습자의 상황에 맞게 난이도를 조절할 수 있도록 돕는다.

학생 수준은 교사가 사전에 관찰한 정보, 학생의 학습 패턴, 수행 기록 등 다양한 정성적·정량적 자료를 종합적으로 고려하여 판단된다. 이 과정에서 학생의 이해 정도, 과제 수행 방식, 개념 간 연결 능력, 탐구 과정에서의 태도와 같은 요소가 함께 분석된다. 알고리즘은 이러한 정보를 기반으로 학생의 현재 학습 상황을 다면적으로 해석하며, 특정 범주나 일률적인 등급으로 단순 분류하기 보다는 학생의 학습 양상과 특성을 세밀하게 반영하여 문장의 수준과 복잡성을 조정하도록 설계되었다.

따라서 동일한 성취기준이라 하더라도 학생의 현재 학습 맥락과 특성에 따라 생성되는 문장의 깊이는 자연스럽게 달라진다. 예를 들어 사고의 폭이 넓고 자료 해석 능력이 잘 드러나는 학생에게는 상대적으로 분석적이고 추론적 사고를 요구하는 문장이 생성되며, 개념 이해 과정에서 어려움을 드러내는 학생에게는 핵심 개념을 명확히 확인하고 점진적으로 사고를 확장할 수 있도록 돕는 문장이 제시된다. 이 과정은 학생의 수준을 단순히 구분하는 방식이 아니라, 학생이 현재 어떤 학습적 흐름에 놓여 있는지를 파악하여 생성 문장의 난이도를 자연스럽게 조정하는 데 중점을 두고 있다.

이러한 성취기준 분석 및 개인화 알고리즘은 교육과정 기반 평가의 본질을 AI가 이해하도록 설계된 핵심 요소로, 단순한 자동 생성이 아니라 학습자의 실제 상황을 세밀하게 반영한 교육적 타당성을 확보하는 데 중요한 역할을 한다.

## 2. 전략·문항·탐구 활동 생성 과정

전략·문항·탐구 활동 생성 알고리즘은 성취기준 분석 결과와 학생의 학습 정보가 결합되어 작

동하는 핵심 과정이다. 전략 생성 단계에서는 성취기준이 요구하는 기능 요소를 기반으로 학습자가 목표 성취에 도달하기 위해 어떤 사고 절차를 밟아야 하는지 체계적으로 구성한다. 이 과정에서 성취기준에 포함된 동사는 중요한 기준으로 작동하며, 예를 들어 ‘분석하다’, ‘해석하다’, ‘비교하다’와 같은 동사가 포함되어 있다면 전략 문장도 자연스럽게 자료의 구조를 파악하고 관계를 해석하는 과정 중심으로 생성된다.

문항 생성 알고리즘은 성취기준의 핵심 개념이 반드시 반영되도록 설계되며, 문항의 복잡성과 사고 요구 수준은 학생의 학습 양상과 특성에 따라 조정된다. 문항은 단일 사고 원칙을 바탕으로 구성되어 학습자가 한 번에 하나의 사고 단위를 명확하게 수행할 수 있도록 하고, 필요 이상으로 복잡한 요소가 중첩되거나 개념이 과도하게 확장되는 것을 방지한다. 또한 성취기준과 문항이 요구하는 사고 수준이 일치하는지 자동 점검하여 교육적 타당성을 확보한다[5].

탐구 활동 생성 과정은 교육과정이 강조하는 탐구 구조를 반영하고, 동시에 학생이 실제로 수행 가능한 수준에서 활동을 구성하는 것을 목표로 한다. 활동을 단순히 제시하는 데 그치지 않고 활동의 목적, 수행 절차, 필요한 자료, 기대되는 결과 등이 자연스럽게 연결되도록 구성하여 수업과 평가가 유기적으로 이어지도록 설계하였다. 학습 과정에서 자신감을 보이거나 자료 해석에 강점을 가진 학생에게는 다층적 자료 분석이나 실제 경제·사회 상황을 적용한 심화 탐구가 제시되며, 개념 이해 과정에서 어려움을 보이는 경우에는 기초 개념 중심의 단순 자료 해석 활동처럼 실천 가능한 범위의 탐구가 제시된다.

이 전체 과정은 입력 정보가 의미 분석을 거쳐 평가 요소 생성 모델로 전달되고, 이후 생성된 문장이 검증을 통해 정제된 후 최종 평가 요소로 제공되는 일련의 구조 속에서 이루어진다. 생성 과정과 검증 과정은 서로 보완적인 역할을 하며, 생성된 문장이 성취기준과 일치하는지, 표현이

자연스러운지, 교육적 요구와 부합하는지를 반복적으로 점검하여 교사가 사용할 수 있을 만큼 안정적인 수준으로 정리된다. 이러한 구조는 AI가 단순히 문장을 만들어내는 수준을 넘어, 교육 과정 기반 평가 체계에 부합하는 결과를 생성하도록 돕는 데 목적이 있다.

## V. 적용 사례 분석

### 1. 사례 적용 개요

본 연구에서는 SCPES의 개인화 기능을 검증하기 위해 경제 수학 과목의 「수와 경제」 단원에 포함되는 경제지표와 단리·복리 성취기준을 중심으로 두 명의 학생을 선정하여 시스템을 적용하였다. 학생 A는 교사의 사전 진단에 따라 중상 수준 학습자로 분류되었으며, 수학적 개념 이해가 비교적 안정적일 뿐만 아니라 통계 자료를 해석하고 이를 바탕으로 문제를 추론하는 능력이 우수한 것으로 나타났다. 반면 학생 B는 중하 수준 학습자로 판단되었으며, 경제 용어의 의미를 연결하는 과정에서 난점을 보였고, 기초 개념을 실제 상황에 적용하는 데 어려움을 겪는 경향이 있었다. 이러한 두 학생의 상이한 학습 특성은 개인화된 전략과 문항 생성의 적절성을 평가하는 데 중요한 기준이 되었다. SCPES는 성취기준을 동일하게 설정하더라도 학생 개인의 학습 수준과 진단 정보를 입력값으로 반영하여 서로 다른 분석과 추천 결과를 생성할 수 있으므로, 두 학생 간 차별화 수준을 직접적으로 비교함으로써 시스템의 개인화 기능을 검증할 수 있었다.

### 2. 학생 수준에 따른 생성 결과 비교

SCPES는 학생 A와 B에 대해 동일한 성취기준을 입력받더라도 각 학생의 학습 수준, 이해도, 교

사 진단 메모에 따라 서로 다른 전략, 문항, 탐구 활동을 생성하였다. 학생 A에게 제시된 문항은 통계 자료를 활용하여 추세를 해석하거나 여러 경제 지표 간의 관계를 분석하는 고차적 사고를 요구하는 형태로 나타났다. 예를 들어 물가상승률 변화 그래프를 해석하는 문제나 실질GDP와 명목GDP의 차이를 비교·설명하는 문항 등이 포함되었다. 탐구 활동 역시 한국은행 ECOS 통계자료를 활용하여 직접 데이터를 수집하고, 해당 데이터를 기반으로 경제 상황을 해석해보는 심화 활동이 중심을 이루었다. 이러한 생성 결과는 학생 A가 높은 수준의 자료 해석 능력과 추론 능력을 갖추고 있다는 교사 진단과 일치하였다.

반면 학생 B에게 제공된 생성 결과는 보다 기초적이고 적용 중심의 문제로 구성되었다. 예를 들어 소득이 일정 비율로 증가했을 때 소비지출에 어떤 변화가 생기는지 계산하는 문제, 단리·복리 개념을 단순한 수치 계산으로 확인하는 문항 등이 포함되었다. 탐구 활동 또한 생활 속 물가 변동 사례를 찾아 그 원인을 단순 비교하는 활동 등 일상 맥락 중심의 분석을 강조하는 형태로 제시되었다. 이는 학생 B가 경제 개념을 실제 상황과 연결하는데 어려움을 보인다는 진단 결과를 반영하여, 기초 개념의 이해와 적용을 돕기 위한 구조로 설계된 출력임을 의미한다.

두 학생의 결과를 비교하면, SCPES는 성취기준이라는 동일한 출발점을 유지하면서도 학생의 학습 수준에 따라 전략의 깊이, 문항 복잡성, 탐구 활동의 구조와 범위가 명확히 차별화된 형태로 생성됨을 확인할 수 있었다. 특히 고난도 문항과 심화 활동은 학생 A에게만 제시되었고, 기초 계산 중심 문항과 개념 확인 활동은 학생 B에게만 제공되었다는 점에서 SCPES의 개인화 알고리즘이 효과적으로 작동했음을 보여준다. 이와 같은 차별화는 단순한 난이도 조절을 넘어 학습자의 사고 과정과 이해 수준을 고려하여 추론 중심, 적용 중심, 개념 확인 중심 등 다양한 유형의 평가 요소를 생성하는 시스템의 구조적 강점을 반영한다.

표 2. 학생 A/B 생성 결과 비교표

항목	기존 연구	SCPES
역량 분석	자료 해석·추론 강점	개념 연결·용어 이해 어려움
전략	고난도 추론 기반 전략	기초 개념 중심 전략
문항 난이도	추론·분석 중심 고난도	연산·기초 개념 중심
탐구 활동	ECOS 기반 데이터 분석	생활 맥락 중심 경제 탐구

표 2는 학생 A와 B의 생성 결과를 역량 분석, 전략, 문항 난이도, 탐구 활동의 네 가지 항목으로 분류하여 정리한 것으로, 두 학생 간 생성 결과의 차이를 명확히 보여준다. 이 비교를 통해 SCPES는 성취기준을 기반으로 하되 학생 맞춤형 평가 요소를 체계적으로 제시할 수 있는 구조적 완성도를 갖추고 있으며, 교육 현장에서 교사의 판단을 지원하는 도구로서의 가능성을 충분히 지니고 있음을 확인하였다.

## VI. 결론 및 향후 연구

본 연구는 생성형 인공지능을 기반으로 성취기준 중심 평가 요소를 자동 생성하는 지능형 학생 중심 과정 평가 시스템(SCPES)을 설계·개발하였다. 시스템은 성취기준 의미 분석, 학생 수준 기반 난이도 조절, 자연어 생성 기술을 통합하여 교사의 평가 업무 부담을 줄이고 학생 개별 특성에 맞는 피드백 제공을 가능하게 한다. 사례 분석을 통해 시스템이 실제 교육과정 기반 상황에서 효과적으로 작동하며 학생별 차별화를 구현할 수 있음을 확인하였다.

향후 연구에서는 실제 학교 단위에서의 대규모 실증 연구를 통해 알고리즘의 안정성과 신뢰성을 검증할 필요가 있다. 또한 학생 행동 데이터, 수업 참여 로그 등 다양한 데이터를 활용하여 개인화 모델을 고도화하고, 생성형 AI의 편향을 최소화하는 안전한 알고리즘 구조를 개발해야 한다. 마지막으로 교사와 AI가 협력하여 평가를 수행하는 하이브리드 평가 체계를 구축하는 연구도 필요하다.

## REFERENCES

- [1] 교육부, 2022 개정 교육과정 총론 주요사항. 2021년
- [2] 김정민, “과정중심평가의 개념과 교육적 의의 탐색,” *학습자중심교과교육연구*, 제18권, 제20호, 839-859쪽, 2018년
- [3] 한옥영, “생성형 AI기반 학습자 맞춤형 교육 시스템 설계를 위한 구성요소,” *컴퓨터교육학회지*, 제26권, 제6호, 127-141쪽, 2023년
- [4] 김정아·강두식·고용철, “생성형 AI의 교육적 활용 방안 연구- ChatGPT 활용을 중심으로,” *정보교육학회논문지*, 제27권, 제6호, 691-703쪽, 2023년
- [5] 조용상, “생성 AI 기반의 평가 체계 및 지식 추적 방법 설계 연구,” *한국컴퓨터교육학회논문지*, 제27권, 제1호, 143-156쪽, 2024년
- [6] Hye-Bin Kim, Eunseo Pyo, Anna Shin and Il-Hyun Jo. (2023). “A Systematic Literature Review of Learning Analytics and Data Utilization in a Digital Teaching and Learning Environment in Korea,” *Journal of Educational Technology*, vol. 39, no. 2, pp. 595-638.
- [7] Anderson, L. W., & Krathwohl, D. R. (2001). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom’s Taxonomy of Educational Objectives*.

### 저자 소개



이해인(정회원)

2006년 한밭대학교 산업경영학과 학사 졸업.

2013년 공주대학교 컴퓨터교육학과 석사 졸업.

2018년 공주대학교 컴퓨터교육학과 박사 수료.

<주관심분야 : 생성형 교육프로그램, 인공지능, AI, ICT, 빅데이터 등>