MONTHLY EDTECH SOLUTION ANALYSIS REPORT

월간 에듀테크 솔루션 분석 보고서

Vol. 02

2026. 07

이번 호 주제 ISSUE TOPIC

적응형 학습 시스템

지식 추적에서 생성형 개인화까지, 적응형 학습 엔진의 구조와 고등교육 효과의 조건

대표 분석 솔루션 ALEKS (McGraw Hill) · Gemini Study Notebooks (Google)

키워드 적응형 학습 · 지식 추적 · 생성형 개인화 · 고등교육 · AI 디지털교과서

분석 프레임 4-QUADRANT FRAMEWORK

① 기술 구조 Technology	② 시장 동향 Market	③ 교육적 적용 Pedagogy	④ 한계·논쟁점 Limitations

한림대학교 AI 에듀테크 연구과제

글로컬대학 30 한림대학교 AI 에듀테크센터 연구과제「고등교육 AI 에듀테크의 통합적 발전 방안 연구」[연구 2] 산출물

발간 2026. 7. 31.

EXECUTIVE SUMMARY

핵심 요약

이번 호의 질문

지식 추적·문항반응이론 기반 전용 엔진과 범용 생성형 AI 의 진단 퀴즈 기반 개인화는 구조가 어떻게 다른가? 고등교육에서 효과가 확인된 맥락과 조건은 무엇이며 독립 근거는 충분한가? AI 디지털교과서의 지위 변화와 채택 감소는 대학의 적응형 학습 도입에 어떤 교훈을 주는가?

한 줄 정의

적응형 학습 시스템은 학습자의 응답 데이터로 지식 상태를 추정해 학습 경로·난이도·콘텐츠 순서를 자동으로 조정하는 엔진(코스웨어 포함)이다.

핵심 발견

1. 전용 적응형 엔진은 지식 공간·선수 관계 같은 도메인 모델을 미리 구축하고 학습자 응답으로 숙달 확률이나 능력을 추정해 경로를 정한다. 범용 LLM 1 종을 다룬 K-12 공개 데이터 비교 연구(Gemini Study Notebooks 대상 아님)에서 LLM 의 학습자 모델링은 예측 성능(AUC)이 DKT 보다 6% 낮고 방향이 틀린 갱신 등 시간적 일관성의 약점을 보여 전용 학습자 모델과 생성형 인터페이스를 짝짓는 결합형이 제안되었다. 제공사 문서상 Gemini Study Notebooks(학습 노트북)는 학습 목표를 100 개 이상의 세부 목표로 나누고 생성한 진단 퀴즈로 진도를 추적하며 사전 구조를 쓰는지와 숙련도 산출 기준은 공개되지 않았다.
2. 적응형 학습 메타분석의 g = 0.70 은 K-12 밖 연구가 7 개뿐인 K-12 중심 값이다. 대학 지능형 튜터링 시스템(ITS) 메타분석은 g = 0.32~0.37, 12 편을 종합한 국내 메타분석(이질성 높음)의 대학생 하위집단은 0.2306 으로 더 작았고 학부생 대상 연구 69 편 중 41%는 영향 없음을 보고했다. 14 개 고등교육기관의 적응형 코스웨어 평가에서는 분석 가능한 15 건 중 4 건만 평균 성적이 약간 올랐다. 제공사와 독립된 ALEKS 메타분석 2 편(K-12·고등교육 혼합, 고등교육 하위집단 추정치 없음)은 모두 전통 수업과 비슷한 성과를 보고했고(Sun 외 g = 0.05) 보충 사용의 효과는 엇갈렸다. 고등교육 단독 독립 근거는 ALEKS 도 관찰 연구 2 편과 준실험 비교 1 건(성과 혼재)에 그쳐 대표 제품 모두 충분하지 않다.
3. 두 대표 솔루션은 자료 기준일까지 공개된 문서로 평가했고 실사용 진단이 필요한 14 개 항목(게이트인 피드백·적응 포함)은 미측정으로 두었다. 필수 항목 13 개 가운데 두 제품 모두 5 개만 판정되어 도입 검토 의견은 모두 판정 유보(ALEKS 55.2 점, 학습 노트북 28.9 점)이며 점수는 순위가 아니다. ALEKS 는 학습목표 정렬과 교수자 통제가 문서로 확인되었으나 효과성 증거가 유망 수준이어서 측정을 마쳐도 조건부 검토가 상한이다. 학습 노트북은 출시 때 18 세 이상 개인 계정에 먼저 제공되고 교사 배정과 학교 계정 제공은 예정이어서 교수자·기관 개입 경로가 확인되지 않았고 독립 효과 연구가 없어 신중 검토가 상한이다.
4. AI 디지털교과서(K-12)는 2025 년 8 월 법 개정으로 교과서가 아닌 교육 자료가 되었다. 2025 년 1 학기에는 도입 대상 학교의 34.2%가 자율선정해 활용했으나 전환 뒤인 2 학기 채택률은 국정감사 질의에서 14.9%(대구를 빼면 10.4%)로 제시되었다(집계 기준 다름). 감사원은 시범 운영 생략과 현장적합성 검토 미흡, 시·도교육청과 협의 없는 구독료 예산 결정을 지적했고 선정 학교 학생 가운데 한 번도 활용하지 않은 비율은 학년·과목별 평균 60.0%(2025 년 3~5 월 접속기록)였다. 법적 지위와 재원에 기댄 정책 주도 도입은 조건이 바뀌면 지속되기 어려울 수 있다.

4 분면 요약

① 기술 구조

○ 적응형 엔진은 학습자 모델(지식 추적·문항반응이론)·도메인 모델(지식 공간·선수 관계)·교수 모델(숙달 기준·경로) 등이 맞물려 작동하며 생성형 AI 는 지식 요소 태깅·문항·힌트 저작을 돕지만 불일치와 품질 실패가 남아 전문가 검토와 측정학적 검증이 도입 요건이 될 것으로 보인다.

○ Google 의 학습 노트북은 목표를 세부 목표로 나누는 진단 퀴즈 기반 개인화로(산출 기준 미공개) 적응형 기능을 흡수했으나 학습자 모델의 일관성과 판정 기준의 투명성을 받칠 근거는 아직 부족하다. 결합형은 연구 제안(K-12 데이터)으로 나왔고 생성형 기능을 덧붙인 전용 엔진 제품(AI Reader)도 학습자 모델 연결은 확인되지 않았다.

② 시장 동향

○ 민간 조사 기관(IMARC Group)은 세계 적응형 학습 시장을 2025 년 58 억 달러로 추정했으나 추정 방법이 공개되지 않고 고등교육 몫도 드러나지 않아 참고 수치로만 읽어야 한다.

○ 교재 결합형 코스웨어의 기관 청구 구조 위에서 출판사의 기업공개·제품군 확장(McGraw Hill)과 전문 기업 간 합병 합의(Skillwell–Realizeit)가 이어졌다. Google 학습 노트북처럼 범용 LLM 이 기관판 무료 포함과 진단 퀴즈를 앞세워 적응형 기능을 흡수하는 사례도 나타났다.

③ 교육적 적용

○ 효과는 학교급·설계·측정 방식에 따라 달라지고 고등교육에서는 중간 이하에 결과도 엇갈린다(대학 ITS g = 0.32~0.37, 12 편을 종합한 국내 메타분석의 대학생 하위집단 0.2306, 이질성 높음).

○ 보충 사용(메타분석 간 엇갈림), 미시 수준 적응, 숙달 요건 조정, 교수자 통합이 효과와 관련된 조건으로 거론되지만 인과가 확인된 조건은 아니다. 제품 단위 독립 근거는 대표 제품 모두 충분하지 않고 생성형 AI 개인화의 대학 근거는 소표본·비무작위 설계 중심의 초기 단계다.

④ 한계·논쟁점

○ 콘텐츠 태깅 비용, 시스템 주도 경로와 학습자 자율성, 근거가 희소하거나 엇갈리는 알고리즘 편향, 배치 평가의 고부담 사용, 학습 이력의 2 차 이용, 제공사 효과 주장의 과대 추정 가능성이 쟁점이며 실험으로 인과가 확인된 위험은 드물다.

○ EU 는 고위험 AI 의무 적용일을 2027 년 12 월 2 일로 정했다. 집행위원회 초안 지침은 진행 중인 학습을 지원하는 적응형 시스템은 고위험에서 빼되 적응형 배치 도구와 성적 결정에 쓰이는 개인화 학습 보조 AI 는 고위험 예시로 든다(초안).

대표 솔루션 진단 요약

솔루션	사용성 (A)	교육적 특성(B)	AI 신뢰성 (C)	종합	공통 11 개 항목 기준	도입 검토 의견
ALEKS (McGraw Hill)	50.0	61.5	50.0	55.2	56.9	판정 유보
Gemini Study Notebooks (Google)	-	23.1	37.5	28.9	28.9	판정 유보

주: 두 제품 모두 자료 기준일까지 공개된 문서로 판정했고 실사용 진단이 필요한 14 개 항목(게이트 B3 포함)은 미측정이다. ALEKS 는 12 개 항목(필수 5 개)을 판정했고 축 A 는 1 개 항목 값이다. 학습 노트북은 11 개 항목(필수 5 개)을 판정했고 공개 만족도 조사가 없어 축 A 는 산출하지 않았다(종합은 축 B·C 로 계산). 공통 11 개 항목 기준은 두 제품을 모두 판정한 문서 판정 항목만의 점수다(축 A 없음). 미측정 항목을 모두 ●·모두 ○로 두면 ALEKS 는 26.7~79.5 점(B9 ◐라 조건부 검토가 상한), 학습 노트북은 12.9~68.3 점(B9 ○라 신중 검토가 상한)이다. 필수 측정 10 개 미만은 판정 유보, 게이트(B3·B9) 모두 ◐ 이상·종합 60 이상은 조건부 검토, 게이트 모두 ●·필수 ○ 없음·종합 80 이상은 적극 검토, 그 밖은 신중 검토다.

국내 적용 시사점

○ 고등교육 적응형 학습은 AI 디지털교과서의 법정 선정 기준이나 고영향 AI 규정의 명시 범위 밖으로 해석되므로 대학은 배치·진단 용도의 사람 검토와 이의 경로, 학습 이력의 가명처리와 2 차 이용 범위, 계약상 국외 처리 조건을 스스로 점검하는 기준을 고려해 볼 만하다(부록 4). 범용 LLM 진단 퀴즈는 학습 보조로 한정하고 출처를 검증하도록 안내할 수 있다.

○ 기초 교과 보완 학습에는 '진단 → 보완 → 재진단' 순환을 교수설계 지침과 조교·멘토 지원으로 뒷받침하고 숙달 요건과 학습 부담을 소규모 시범 운영으로 조정한 뒤 넓히는 방안을 검토할 수 있다. 이 방식의 효과는 확인되지 않았고 본교(한림대학교) AI-HiED 가 같은 방식이어서 이해관계가 있다.

○ 도입·실증 단계부터 분반 단위 무작위 배정(어려우면 비교 집단·지연 측정)을 넣어 근거 수준을 올리되 협약·운영 관계가 있는 제품은 관계없는 연구진의 분석이나 분석 계획 사전 공개로 독립성을 확보하고 지원이 끝난 뒤의 운영 재원도 미리 검토할 수 있다.

CONTENTS

목 차

Ⅰ. 배경 및 분석 개요 1

1. 선정 배경 1
2. 분석 범위와 분석 단위 2
3. 분석 대상 솔루션 2
4. 분석 방법 4

Ⅱ. 솔루션군 동향 분석: 4분면 6

1. 기술 구조 6
2. 시장 동향 10
3. 교육적 적용 13
4. 한계·논쟁점 17

Ⅲ. 대표 솔루션 심층 분석 22

1. 솔루션 프로필 22
2. 평가척도 적용 결과 23
3. 비교 종합 27

Ⅳ. 시사점 및 종합 29

1. 국내 적용 시사점 29
2. 본교·연구회 활용 제언 31
3. 종합 33

부록 35

부록 1. 참고 자료 목록 35
부록 2. 평가척도 채점표 42
부록 3. 용어 해설 52
부록 4. 기관 도입 점검표 53

Ⅰ. 배경 및 분석 개요

이 보고서는 「월간 에듀테크 솔루션 분석 보고서」의 둘째 호로 Vol.01 이 경계로 남긴 학습 경로·난이도 자동 조정, 곧 적응형 학습 시스템을 다룬다. 전용 엔진과 범용 대규모 언어 모델(LLM)의 진단 퀴즈 기반 개인화를 비교한다. 자료 기준일은 발간일과 같은 2026 년 7 월 31 일이며, 기준일 뒤에 발행된 자료와 기준일 뒤의 제품·정책 변화는 다루지 않는다.

1. 선정 배경

가. 정책 환경

2025 년 8 월 14 일 개정된 초·중등교육법은 지능정보기술을 활용한 학습지원 소프트웨어를 교과용 도서 외의 '교육 자료'로 규정했다(초·중등교육법, 2026). 개정 전인 2025 년 1 학기에는 도입 대상 1 만 1,985 개교 가운데 34.2%인 4,095 교가 AI 디지털교과서(AIDT)를 자율선정해 활용했다(감사원, 2025).

교육 자료 전환 직후인 2025 년 2 학기의 채택률(K-12)은 국정감사 질의에서 경남·제주를 뺀 집계로 14.9%로 제시되었고 대구를 빼고 계산하면 10.4%였다(국회사무처, 2025). 1 학기의 자율선정·활용 비율(34.2%)과 이 채택률은 집계 기준이 달라 두 수치를 직접 비교하는 데는 한계가 있다.

한국교육학술정보원(KERIS)은 본교(한림대학교)가 운영대학을 맡은 고등교육 에듀테크 소프트랩(Ⅰ장 3 절) 공고에서 에듀테크가 K-12 에 비해 고등교육에서는 아직 도입 초기 단계라고 진단했다(한국교육학술정보원, 2026).

집계 기준의 차이를 감안해도 초·중등에서는 교육 자료로 지위가 바뀐 뒤 채택이 크게 줄었고 고등교육은 아직 도입 초기로 진단되었다. 대학이 적응형 학습을 들이기 전에 초·중등의 경과에서 얻을 교훈을 셋째 핵심 질문(Q3)으로 삼았다.

나. 대학 현장

2020 년 스위스·남아공 두 대학의 전문가 델파이 연구는 고등교육 기관의 관심에 비해 강좌에서의 실제 사용이 낮다고 보았다(Mirata et al., 2020). 한국교육개발원(KEDI)의 2025 년 교원 조사(일반대학 80 개교 3,672 명, 자기 보고)에서도 AI 기반 적응형·맞춤형 학습시스템의 효과 인식은 4 점 척도로 2.62 점, 활용도는 2.03 점이었고 두 값의 응답자별 차이 평균은 0.58 점이었다(김지하 외, 2025).

국외 제품의 국내 대학 적용 사례로는 McGraw Hill 의 ALEKS(Assessment and Learning in Knowledge Spaces)를 2020 학년도 2 학기 간호학과 1 학년 학생 39 명을 대상으로 일반화학 수업에 운영한 사례연구가 있다(황은경·신종호, 2021).

이 보고서가 확인한 국내 대학의 공개 적용 사례는 이 일반화학 사례연구와 교양 수학 파일럿(Ⅱ장 3 절 바)처럼 기초 교과의 소규모 운영에 머문다. 사례 수가 적고 비교 집단을 둔 연구는 그 파일럿 1 편뿐이어서 효과는 Ⅱ장 3 절의 연구로 판단한다.

다. 연구 동향

2020~2024 년 SSCI Q1 상위 인용 논문 29 편을 검토한 체계적 리뷰에 따르면 고등교육 AI 개인화 학습 연구는 아시아, 특히 중국에 집중되어 있고 주로 교육학·컴퓨터과학 분야에서 이루어졌다(Peng & Li, 2025).

효과 근거는 K-12 에 치우쳐 있고 결과도 연구마다 다르며(Ⅱ 장 3 절 가·나) 구조 논의(Ⅱ 장 1 절 라)도 정리되지 않았다. Hooshyar 외(2026)는 K-12 에서 생성형 모델이 전통적 학습자 모델링을 대체할 수 있다는 오해가 생겼다고 지적한다.

이에 지식 추적(knowledge tracing)·문항반응이론(IRT) 기반 엔진과 생성형 AI 개인화를 중심으로 세 핵심 질문을 묻는다.

○ Q1. 지식 추적·문항반응이론 기반 엔진과 생성형 AI 개인화는 구조가 어떻게 다른가?

○ Q2. 고등교육에서 효과가 확인된 맥락과 조건은 무엇이며, 독립 근거는 충분한가?

○ Q3. AI 디지털교과서의 지위 변화와 채택 감소는 대학의 적응형 학습 도입에 어떤 교훈을 주는가?

2. 분석 범위와 분석 단위

미국 교육부 보고서(K-12 중심)는 개인화 에듀테크 다수가 제한적으로만 적응하며 난이도와 학습 순서 조정이 가장 흔한 두 방식이라고 보았다(U.S. Department of Education, Office of Educational Technology, 2023). KERIS 가이드라인(K-12)은 AI 디지털교과서를 개인별 학습 수준과 속도를 반영한 맞춤형 학습(Adaptive Learning)을 특성으로 하는 소프트웨어로 정의했다(한국교육학술정보원, 2023b).

이번 호는 학습자 데이터로 학습 경로·난이도·콘텐츠 순서를 자동 조정하는 엔진과 이를 강좌 단위로 묶은 코스웨어를 분석 단위로 삼으며 범용 LLM 개인화를 포함한 경계는 [표 Ⅰ-1]과 같다.

[표 Ⅰ-1] 분석 범위와 경계

구분	이번 호에서 다루는 것	제외·제한(넘기는 호)
사용 주체	학습자(시스템 주도 경로)	교수자 대시보드·위험 예측(Vol.04), 채점·피드백 생성(Vol.05)
핵심 기능	진단·숙달 추정에 따른 학습 경로·난이도·콘텐츠 순서 자동 조정	교과 특화 대화형 지도(Vol.03), 계획·점검·성찰 지원(Vol.06)
엔진 유형	지식 추적·IRT·지식공간이론 기반 전용 엔진, 범용 LLM 의 진단 퀴즈 기반 개인화(비교 항목)	범용 LLM 의 일반 대화·학습 모드(Vol.01)
제공 형태	교재 결합형 코스웨어, 기관 라이선스, 개인 구독, 국내 진단-보완 코스웨어(사례)	본교 자체 서비스는 사례로만(Ⅳ 장 2 절)
학교급	고등교육	K-12·시험 대비 제품은 경계·참고 사례('K-12' 표시)
근거 범위	독립 연구(효과), 제공사 문서(제품 사실), 동료심사 전 자료(표시)	제공사 수행 연구(표시), 마케팅 성과 수치는 효과 근거로 쓰지 않음(제공사 주장의 예로만 표시, Ⅱ 장 4 절 바)

자료: 연구진 정리

3. 분석 대상 솔루션

대표 솔루션은 국내외 균형, 시의성, 사업 연속성, 국내 도입 실적, 국내 이용 가능성, 실사용 진단 가능성의 6 개 기준으로 검토했다. ALEKS 는 국외 제품이며 독립 리뷰인 Leon(2026)은 이를 생성형이 아닌 지식공간이론 기반 적응형 엔진으로 분류했다. 시의성 면에서 McGraw Hill 은 2026 년 6 월 29 일

ALEKS 데이터에 기반한 머신러닝을 쓰는 ALEKS for Calculus II 를 발표했다(McGraw Hill, 2026e). 사업 연속성 면에서는 2025 년 7 월 23 일 기업공개(IPO) 가격을 확정했다(McGraw Hill, 2025). 국내 도입 실적으로는 대학 적용 사례가 있다(Ⅰ장 1 절 나). 국내 이용 가능성은 국제 이용약관으로 확인했으나 국내 기관 계약 조건은 공개 문서에서 확인되지 않았다(Ⅲ장 1 절).

Google 의 Gemini Study Notebooks(학습 노트북)도 국외 제품이다. 시의성 면에서 Google 은 2026 년 6 월 25 일 Gemini 앱에 학습 노트북을 출시하며 이를 적응형 학습 플랫폼이라고 소개했다(Kirtikar & Walport, 2026). 국내 이용 가능성 면에서 학습 노트북은 18 세 이상 개인 계정에 먼저 제공되었고(학교 계정은 제공 예정, Ⅲ장 1 절) 전 세계 모든 언어로 출시되고 있다고 안내되었다(Kirtikar & Walport, 2026; Walport, 2026). 사업 연속성을 해칠 발표는 확인되지 않았고 국내 도입 실적은 찾지 못했다. Gemini 기반 고등교육 연구는 대부분 소표본 비무작위 비교였고 학습 노트북을 시험하지 않았다(Ⅱ장 3 절 마).

실사용 진단 가능성은 선정 단계에서 함께 검토했으나 이번 호는 실사용 진단 결과를 싣지 않고 자료 기준일까지 공개된 문서로만 두 제품을 평가했다(Ⅰ장 4 절). 두 제품 모두 6 개 기준을 다 충족하지는 못했으나 Q1 의 유형 비교를 위해 전용 엔진과 범용 LLM 에서 하나씩 골랐다.

동향 분석 대상 가운데 Wiley 는 Knewton Alta 를 학생 수행에 따라 개인별 숙달 경로를 제공하는 숙달 기반 적응형 학습 코스웨어로 소개한다(Wiley, Retrieved 2026). Pearson 은 MyLab·Mastering·Revel·Pearson+ 안의 AI 학습 도구(AI Study Tools)를 연결된 학습 경험으로 소개하며 이 도구는 일부 MyLab·Mastering 강좌와 일부 Revel 강좌에서만 제공된다(Pearson, Retrieved 2026). Skillwell 은 2026 년 7 월 14 일 Blackboard 학습관리시스템(LMS)과 연동한다고 발표했다(Skillwell, 2026).

Area9 Lyceum 은 2026 년 4 월 27 일 기업 교육 대상 AI 기반 제품군의 첫 제품들을 공개했다(Area9 Lyceum, 2026). K-12 중심인 Squirrel AI(Squirrel Ai Learning)와 국내 시험 대비 서비스는 경계·참고 사례로만 적는다([표 Ⅰ-2]). 산타는 2025 년 9 월 30 일 뤼이드(Riiid)에서 사명을 바꾼 소크라 AI(Socra AI)가 운영한다(남혜현, 2025). 콴다 제공사 매스프레소는 2025 년 7 월 9 일 AI 학습 에이전트 '콴다 3.0'을 하반기에 출시한다고 밝혔다(박수빈, 2025).

이 보고서는 HAI-GPT 를 운영하는 본교 AI 에듀테크센터의 연구과제 산출물이며, 연구진에는 해당 부서·센터 구성원이 포함된다. 연구진 일부는 소프트랩 실증 교과를 맡았다(한림대학교 에듀테크소프트랩사업단, Retrieved 2026). 이 센터는 2025 년 12 월 프리윌린과 AI-HiED 플랫폼 구축 협약을 맺었고 플랫폼은 풀리캠퍼스와 본교 LMS SmartLEAD 를 학습 도구 상호운용성(LTI) 표준으로 연동한다(문화춘, 2025). 본교가 운영대학인 고등교육 에듀테크 소프트랩은 대학의 AI 코스웨어 실증에 드는 라이선스 사용료 일부를 지원한다(한국교육학술정보원, 2026). 이 관계 때문에 풀리캠퍼스·AI-HiED·소프트랩 AI 코스웨어는 채점에서 빼고 사례(Ⅱ장 3 절 바, Ⅳ장 2 절 가)로만 다루며 대표 솔루션이 모두 국외 제품이 되어 국내외 균형 기준은 충족하지 못했다.

자료 수집·선별·분석, 초안 작성, 채점 초안 작성에 생성형 AI(Anthropic Claude)를 보조로 활용했다. 그 제공사는 대표 솔루션 제공사인 Google, ALEKS AI Reader 의 기반 모델 제공사인 OpenAI 와 범용 LLM 시장에서 경쟁 관계에 있다. 한편 영국 경쟁시장청 결정문(2024)상 Google 은 Anthropic 의 투자자이며 Anthropic 은 2025 년 10 월 Google Cloud 의 AI 칩(TPU) 사용 확대를 발표했다(Anthropic, 2025; Competition and Markets Authority, 2024). 이번 호의 분석 대상 후보에는 Anthropic 제품이 없어 이 관계로 채점에서 뺀 제품은 없으며 두 대표 솔루션의 판정 근거는 부록 2 에 공개한다.

[표 Ⅰ-2] 분석 대상 솔루션 개요

구분	솔루션(제공사)	적응 방식	최근 변화·근거	이번 호 처리
대표	ALEKS(McGraw Hill)	지식공간이론 기반 숙달 경로(비생성형)	Calculus II 발표(2026. 6. 29.), IPO 가격 확정(2025. 7. 23.)	채점(공개 문서)
대표	Gemini Study Notebooks(Google)	진단 퀴즈 기반 생성형 개인화 레슨	출시(2026. 6. 25., 18 세 이상 개인 계정 우선), 학교 계정 제공 예정	채점(공개 문서)
동향	Knewton Alta(Wiley)	IRT 기반 숙련도 모델 (2018 년 제공사 설명)	숙달 기반 적응형 코스웨어로 소개(제공사 제품 페이지)	동향
동향	Pearson AI Study Tools	MyLab·Mastering·Revel·Pearson+ 안의 AI 학습 도구	일부 강좌에서만 제공(제공사 문서), 독립 리뷰 근거 신뢰도 '보통'(2025 년 중반)	동향
동향	Skillwell Adapt	학습 경로 실시간 조정	Realizeit 합병 합의(2025. 12. 9.), Blackboard LMS 연동 발표(2026. 7. 14.)	동향
동향	Area9 Rhapsode(Area9 Lyceum)	—	기업 교육용 AI 제품 공개 (2026. 4. 27.)	분류만
경계	Squirrel AI(Squirrel Ai Learning)	—	K-12 중심	분류만
참고	산타(소크라 AI, 구 뤼이드)·콴다(매스프레소)	시험 대비(산타)·문제 풀이 해설(콴다) 개인화 서비스	사명 변경(2025. 9. 30.), 콴다 3.0 출시 계획(2025. 7. 9.)	K-12·시험 대비
사례	풀리캠퍼스(프리윌린)	기초학력 진단·AI 맞춤형 보완 학습(AI-HiED 연동 설계)	소프트랩 실증 교과(Ⅱ장 3절 바)	이해관계로 채점 제외
사례	본교 AI-HiED·소프트랩 AI 코스웨어	풀리캠퍼스·SmartLEAD LTI 연동, 코스웨어 실증	협약(2025. 12.)	평가하지 않음(Ⅳ장 2절)

자료: 각 사 발표·문서, 독립 연구, 언론 보도로 연구진 정리

4. 분석 방법

분석은 [표 Ⅰ-3]의 다섯 단계로 진행했다. 효과성 증거(B9)는 평가척도 v1 이 따르는 미국 초중등교육법(ESSA) 증거 등급인 강함(무작위대조실험(RCT))·보통(준실험)·유망(통계적 통제를 둔 상관연구)·근거 없음으로 판정했다. KERIS 가 제안한 한국형 에듀테크 증거 기준(안)에서도 교육적 영향 연구(유형 3)의 연구방법은 무작위대조실험·인과관계 규명이다(한국교육학술정보원, 2023a). What Works Clearinghouse(WWC) 자료(K-12 중심) 분석에서는 개발사가 수행하거나 의뢰한 평가의 평균 효과크기가 독립 평가보다 상당히 컸다(Wolf et al., 2020). 이에 효과성 증거 항목(B9·B10)에는 제공사와 독립적인 연구만 반영했다. 또 대상 제품의 고등교육 연구로 한정하고 학교급 혼합 연구는 고등교육 하위집단 추정치가 없으면 참고로만 적었다.

이번 호의 평가는 자료 기준일까지 공개된 문서(제공사 도움말·정책·약관과 독립 연구)에 근거하며 실사용 진단 결과는 싣지 않았다. 평가척도는 항목마다 측정 방식을 AUTO(도구 실측)·HYBRID(실측 뒤

검토자 확인)·DOC(문서 판정)로 정해 두었다. 이에 AUTO·HYBRID 14 개 항목(게이트 B3 포함)은 공개 문서만으로 판정하지 않고 미측정으로 두었으며 DOC 12 개 항목 가운데 공개 근거가 있는 항목만 판정했다. 자료 수집·선별·분석, 초안 작성, 채점 초안 작성에는 생성형 AI 를 보조로 활용했고 판정은 연구진이 원문 자료와 대조해 확정했다.

평가척도 v1 은 사용성(A, 8 개), 교육적 특성(B, 10 개), AI 신뢰성·윤리·거버넌스(C, 8 개)의 26 개 항목이다. 항목 판정(● 2 점, ◐ 1 점, ○ 0 점)에 가중(필수 ×3, 권장 ×2, 보조 ×1)을 곱해 축마다 100 점으로 환산하고 측정하지 않은 항목(미측정)과 해당 없는 항목(N/A)은 계산에서 뺀다. 종합 점수에는 가중치 A:B:C = 25:45:30 을 적용하되 판정한 항목이 없는 축은 점수를 내지 않고 나머지 축의 가중치로 계산한다. 두 제품의 비교에는 공통 측정 항목(11 개)의 점수만 쓴다. 도입 검토 의견(적극·조건부·신중 검토, 판정 유보)은 게이트 항목 B3(피드백·적응)·B9, 종합 점수, 필수 항목 측정 범위로 정한다(기준은 부록 2 주). 고부담 게이트는 배치·진급 결정에 쓰일 때만 적용하며 이때는 C1(타당성·신뢰성과 환각 관리)·C3(공정성·편향 관리)·C4(프라이버시·데이터 거버넌스)·C5(인간 통제·교수자 역할)가 모두 ●여야 적극 검토가 될 수 있으며 일반 학습용으로 판정한 이번 호에는 적용하지 않았다.

[표 Ⅰ-3] 분석 방법 요약

단계	내용	자료·도구	산출
① 자료 수집·선별	학술·정책·제공사 문서 수집, 철회·정정 확인, 자료 기준일 뒤 발행 자료 제외	OpenAlex·Crossref	부록 1
② 4 분면 분석	기술 구조, 시장 동향, 교육적 적용, 한계·논쟁점	수집 자료	[표 Ⅱ-1]~[표 Ⅱ-4]
③ 효과 근거 등급화	설계·대상·효과크기·근거 수준 기록, 'K-12·동료심사 전·제공사 수행' 표시	메타분석·실험·준실험	[표 Ⅱ-3]
④ 평가척도 채점	26 개 항목 판정·가중, 게이트 B3·B9, AUTO·HYBRID 항목은 미측정	평가척도 v1, 제공사 문서·정책·독립 연구, A:B:C = 25:45:30	[표 Ⅲ-3], 부록 2
⑤ 국내 적용 해석	법령·감사·공공 조사 대조	수집 자료	[표 Ⅳ-1]·[표 Ⅳ-2], 부록 4

자료: 연구진 작성

Ⅱ. 솔루션군 동향 분석: 4 분면

네 분면 가운데 기술 구조(Ⅱ 장 1 절)는 핵심 질문 Q1 에 맞춰 전용 엔진과 범용 LLM 개인화의 구조를 비교하고 시장 동향(2 절)은 공급과 대학 도입의 흐름을 정리한다. 교육적 적용(3 절)은 Q2 에 맞춰 효과의 조건과 독립 근거를 점검하며 한계·논쟁점(4 절)은 남은 쟁점을 모은다. 근거의 성격(K-12, 제공사 자체 보고, 동료심사 전)은 괄호로 밝히고 1·3·4 절 소결에는 질문별 잠정 답을, 2 절 소결에는 공급·도입 흐름의 요지를 정리한다.

1. 기술 구조

가. 학습자 모델과 지식 추적

적응형 엔진의 출발점은 학습자가 지금 무엇을 아는지 추정하는 학습자 모델이다. Corbett 과 Anderson(1995)의 베이지안 지식 추적(BKT) 원전에서 지식 추적은 튜터가 학생이 각 규칙을 학습했을 확률을 계속 추정하는 과정이다. 튜터는 이 추정에 따라 학생이 각 규칙을 숙달할 때까지(숙달 기준은 지식 확률 0.95) 개인별 연습 문제 순서를 제시한다. 딥러닝 지식 추적(DKT)은 순환신경망(RNN)으로 학습을 모델링하며 사람이 도메인 지식을 명시적으로 부호화하지 않아도 된다. 칸아카데미의 K-12 데이터에서 다음 응답의 정오를 맞히는 예측 성능 지표 AUC(곡선 아래 면적)는 DKT 가 0.85 로 표준 BKT(0.68)보다 높았지만 이는 원전 보고치이며 학습 효과를 잰 값이 아니다(Piech et al., 2015).

Abdelrahman 외(2023)의 서베이에 따르면 BKT 등 전통적 방법은 심층 방법보다 계산 부담이 적고 학습된 전이 확률로 스스로 설명이 가능하다. 반면 심층 지식 추적 모델의 주요 과제는 지식 요소(KC) 간 선수 관계 같은 의존성을 찾아내는 것이다.

능력을 재는 계열에서는 컴퓨터 적응형 검사(CAT)가 개인 응답에 따라 문항을 동적으로 바꿔 능력을 효율적으로 추정한다. CAT 개발 첫 50 년 동안 IRT 가 지배적 측정 모델이었고 문항 모수는 사전에 보정되어 검사 중에는 고정된다(Zhuang et al., 2026). Knewton 의 2018 년 제공사 설명(현행 엔진 반영 여부 미확인)에 따르면 Knewton Alta 의 숙련도 모델은 IRT 를 토대로 하지만 IRT 는 지식이 고정된 적응형 검사를 위해 설계되어 지식이 계속 변하는 적응형 학습의 요구를 모두 충족하지 못한다. Knewton 은 이를 최근 응답에 더 큰 가중치를 두는 시간 모델로 보완했다고 밝혔다(Binger, 2018).

지식 추적은 연습하는 동안 변하는 지식을 따라가고 CAT·IRT 는 한 시점의 고정된 능력을 잰다는 점에서 두 계열은 학습자를 보는 가정이 다르다. 상용 엔진은 Knewton 처럼 둘을 섞기도 하므로 숙달 기준, 모수 추정 방식, 판정 근거의 설명 가능성을 엔진 검토 항목으로 삼아 봄 직하다.

나. 도메인 모델과 교수 모델

Essa(2016)의 정리에 따르면 적응형 교육 시스템은 서로 작용하는 다섯 모델(도메인·학습자·평가·교수·전이)에 의존한다. 이 가운데 도메인 모델은 지식 단위를 노드로, 선수 관계를 간선으로 하는 그래프로 나타낼 수 있고 전이 모델은 학습자가 다음에 배울 준비가 된 것을 정한다(게재 당시 McGraw-Hill Education 소속 저자의 방향 제시 논문).

McGraw Hill ALEKS 소속 연구팀(제공사 자체 보고)에 따르면 지식공간이론에 기반한 ALEKS 에서 항목은 교육과정의 한 단위를 묻는 문제 유형이며 같은 난도로 설계된 여러 예제로 구성된다. 평가는 이전 응답에 따라 다음 문항을 고르는 적응형 질의로 진행되며 지식 공간에 담긴 관계 정보 덕분에 대부분의 항목은 직접 묻지 않고도 숙달·미숙달로 분류된다. 연구팀은 RNN 을 이 평가와 결합하면

높은 정확도를 유지하면서 평균 평가 길이를 26% 넘게 줄일 수 있었다는 시뮬레이션 결과도 보고했다(Matayoshi et al., 2021).

다른 엔진도 선수 관계 구조를 미리 갖춘다. Realizeit 소속 공저자가 참여한 연구에 따르면 Realizeit 은 교육과정과 콘텐츠를 분리하는 것을 기본 원리로 하며 적응형 지능 엔진은 교수자가 만든 선수 네트워크에 베이지안 추정을 적용하는 확률 모델이다. 적응성의 주 원천은 엔진이지만 교수자와 학생도 개인화를 조정할 수 있다(Dziuban et al., 2018). KERIS 의 AI 디지털교과서 개발 가이드라인(K-12)도 교육과정 표준체계를 활용해 교과별 지식 맵(Knowledge map)을 설계하도록 했다(한국교육학술정보원, 2023b).

적응 단위도 제품마다 다르다. 미국 교육부의 정책 보고서(K-12 중심)에서는 일부 적응형 교실 제품이 최근 수학 문제 세 개의 정답 여부만 보는 제한된 추천 모델을 쓴다고 지적됐다(U.S. Department of Education, Office of Educational Technology, 2023). SRI International 이 독립 평가한 재단 지원 대학 적응형 코스웨어 사업(ALMAP, II 장 3 절 다)에서는 미시 수준(개별 학습 단위)에서 적응하는 제품이 거시 수준 제품보다 성과가 좋았다(Yarnall et al., 2016).

적응이 일어나는 단위와 추천에 쓰는 학습자 정보의 범위는 적응의 질과 관련될 수 있다. 이 해석은 ALMAP 의 제품 간 비교(비무작위)와 미국 교육부 보고서(K-12 중심)의 지적에 근거한다. 도입을 결정하려면 선수 관계 구조가 강좌 범위를 얼마나 덮는지, 적응이 문항 수준인지 단원 수준인지를 확인해야 한다.

다. 콘텐츠 태깅과 생성형 AI 의 결합

엔진이 응답을 해석하려면 문항마다 지식 요소가 연결되어 있어야 한다. 평가 문항에 지식 요소를 생성하고 연결하는 작업은 상당한 노력과 영역 전문지식을 요구한다. 한 연구에서는 고등교육 화학·e-러닝 객관식 문항에 OpenAI 의 GPT-4 로 지식 요소를 생성했다. 가장 효과적인 LLM 전략의 일치율은 화학 56%, e-러닝 35%였다. 상위 5 개 제안 중 어느 것과도 맞지 않는 문항도 화학 21%, e-러닝 38%였다(Moore et al., 2024). LLM 자동 문항 생성 연구 60 편을 정리한 리뷰는 LLM 이 여러 언어와 교과에서 다양한 유형의 문항을 유연하고 효과적으로 생성했다고 보면서도 생성 문항의 측정학적 속성과 교육적 타당성 평가를 필수 단계로 권고했다(Tan et al., 2025).

생성형 AI 는 노동 집약적인 힌트 저작에도 결합된다. 오픈소스 적응형 튜터 OATutor 에서는 OpenAI 의 ChatGPT 생성 힌트와 인간 튜터 힌트를 무작위 배정 온라인 실험(Mechanical Turk 성인 학습자, 대학생 표본 아님)으로 비교했다. ChatGPT 힌트 조건만 무도움 통제 대비 유의한 학습 향상을 보였다. 두 힌트 조건의 차이는 유의하지 않았으나 ChatGPT 생성 힌트는 문제의 32%에서 품질 점검에 실패했다. 자기일관성 기법을 적용하자 실패율은 대수 문제에서 거의 0%, 통계 문제에서 13%로 줄었다(Pardos & Bhandari, 2024).

뤼이드 연구팀(제공사 자체 보고)에 따르면 상용 시험 대비 서비스 산타의 문항 1 만 3,000 여 개와 강의 1,000 여 개는 293 개 기술 유형으로 태깅되어 있다(Choi et al., 2020). 제공사 문서에 따르면 McGraw Hill 의 AI Reader 는 eBook 에서 개념을 선택하면 생성형 AI 가 대안 설명·쉬운 설명·퀴즈 문항을 만들어 주는 기능이다. 자사 콘텐츠를 색인하는 검색 증강 생성(RAG) 방식으로 지정된 교재의 맥락만 모델에 제공하며 ALEKS 에서는 반 생성·반 설정 단계와 관리자의 기관 설정에서 사용 여부를 정할 수 있다(McGraw Hill, Retrieved 2026-a).

생성형 AI 는 적응형 엔진의 병목으로 꼽히는 태깅·문항·힌트 저작의 부담을 덜 수단으로 들어오고 있다. 다만 지식 요소 불일치와 힌트 품질 실패가 남아 있어 교과 전문가 검토와 측정학적 검증 절차가 도입 요건이 될 것으로 판단된다(태깅 비용은 Ⅱ장 4 절 가).

라. 범용 LLM 의 진단 퀴즈 기반 개인화

Google 의 학습 노트북은 학생이 프롬프트를 입력하거나 수업 자료를 올린 뒤 진단 퀴즈로 집중할 영역을 확인하면 퀴즈 결과와 새로 올린 자료에 따라 레슨을 자동으로 갱신한다(Kirtikar & Walport, 2026). 제공사 사용 안내에 따르면 Gemini 는 맞춤 진단 퀴즈로 학업 기준선을 정하고 레슨별 연습 퀴즈도 올린 자료로 만든다. 또 학습 목표를 100 개 이상의 세부 학습 목표로 나누고 주제별로 묶어 진도를 추적하며 이를 '강점'·'집중 영역'·'시작 전'으로 표시해 퀴즈를 마칠 때마다 자동으로 갱신한다. SAT 대비 노트북은 The Princeton Review 의 문항에 근거한다(Walport, 2026).

LLM 이 학습자 모델을 대신할 수 있는지 검토한 K-12 사례연구(학습 노트북 대상 아님)는 대규모 공개 데이터셋으로 DKT 와 널리 쓰이는 LLM(미세조정 전후)을 비교했다. 미세조정은 LLM 의 AUC 를 영샷(추가 학습 없음) 대비 약 8% 높였지만 여전히 DKT 보다 6% 낮았고 학습 초기 구간의 오류가 더 컸다. DKT 는 숙달 추정이 안정적으로 갱신된 반면 LLM 은 방향이 틀린 갱신 등 시간적 일관성의 약점을 보였다. 저자들은 LLM 을 학습자 모델링의 대체물이 아니라 전용 학습자 모델링 요소와 짝지은 교수 인터페이스·콘텐츠 생성기로 쓰는 것이 적절하다고 제안했다(Hooshyar et al., 2026).

전용 엔진 쪽에서는 반대 방향의 우려가 나왔다(자료와 한계는 Ⅱ장 3 절 마). McGraw Hill 직원 4 명이 참여한 ALEKS 분석(제공사 자체 보고, 동료심사 전)에서 저자들은 생성형 AI 로 인지적 관여 없이 정답을 만들면 플랫폼이 숙달 추정 보정에 쓰는 신호가 오도되고 이후의 적응 결정도 나빠진다고 해석했다(Rismanchian et al., 2026).

두 구조는 개인화의 출발점이 다르다. 전용 엔진이 미리 만든 도메인 모델에서 출발하는 것과 달리 공개 문서로 본 학습 노트북은 학습 목표를 100 개 이상의 세부 목표로 나누고 생성형 진단으로 출발점을 잡는다. 사전 구조를 쓰는지와 숙련도 산출 기준은 공개되지 않았다(SAT 노트북은 외부 문항에 근거). 주제 범위는 넓은 대신 학습자 모델의 일관성과 판정 기준의 투명성을 받칠 근거는 아직 부족하다. 전용 학습자 모델에 생성형 인터페이스를 짝짓자는 제안과 ALEKS 에 AI Reader 를 덧붙인 사례처럼 결합 방향도 나타나지만 전용 엔진의 신호 역시 생성형 AI 사용에 흔들릴 수 있다.

마. LMS 연동과 데이터 표준

LMS 연동의 대표 표준은 LTI 다. 제공사 문서에 따르면 McGraw Hill 은 LTI Advantage 를 지원하는 Blackboard·Brightspace·Canvas·Moodle 등 LMS 를 자사 플랫폼과 연동하며 연동 대상에는 ALEKS 도 포함된다. 이 연동으로 McGraw Hill 의 과제·평가 플랫폼 Connect 에는 단일 로그인·과제 단위 연결·성적부 동기화가 제공된다(McGraw Hill, Retrieved 2026-d).

학습 기술 표준 단체 1EdTech 의 인증 등록부에서 ALEKS 의 인증 이력을 보면 LTI 1.3 도구 인증은 2023 년 4 월 4 일 최초 인증, LTI Advantage 인증은 2024 년 6 월 25 일 인증이다. ALEKS 의 데이터 프라이버시 인증은 2025 년 7 월 1 일 인증으로 기록되어 있다(1EdTech Consortium, Retrieved 2026-a). Gemini 앱·NotebookLM 등 Gemini 기반 기능을 LMS 에 넣는 도구로 등록되어 있는 Gemini LTI 는 2026 년 3 월 23 일 LTI 1.3 도구 인증을 받았으나 학습 노트북이 LTI 로 제공되는지는 밝혀지지 않았다(1EdTech Consortium, Retrieved 2026-b).

교육 데이터 표준으로는 KERIS 가이드라인(K-12)이 교육과정 표준체계를 CASE 표준으로 표현하고 학습 데이터를 수집·저장할 때 Caliper Analytics 와 xAPI 를 참고하도록 권고했다(한국교육학술정보원, 2023b).

연동 표준 지원은 제공사 안내만으로 판단하지 말고 계약 시점에 제 3 자 인증의 유효 여부와 개인정보 심사 이력을 등록부로 대조해야 한다(계약 전 점검은 부록 4). Ⅱ 장 1 절의 구성 요소를 유형별로 정리하면 [표 Ⅱ-1]과 같다.

[표 Ⅱ-1] 적응형 엔진의 기술 구조 비교

구성 요소	전용 적응형 엔진(지식 추적·IRT·지식공간)	범용 LLM 개인화(학습 노트북)	결합형(전용 엔진 + 생성형 기능)
학습자 모델	규칙별 숙달 확률(BKT), 순환신경망(DKT), 능력 추정(IRT), 최근 응답 가중(제공사 설명)	진단 퀴즈 결과로 '강점'·'집중 영역' 표시, 산출 기준 미확인	전용 학습자 모델에 LLM 인터페이스·콘텐츠 생성기를 짝짓자는 제안(K-12 데이터)
도메인 모델	지식 공간·선수 관계 그래프 사전 구축	학습 목표를 100 개 이상 세부 목표로 분해(산출 기준 미공개)	LLM 지식 요소 생성(일치율 화학 56%·e-러닝 35%)
진단·교수	적응형 질의로 묻지 않은 항목까지 분류, 숙달까지 연습 순서 제시, RNN 결합 시 평가 길이 26% 넘게 단축(제공사 자체 보고)	수업 자료 기반 맞춤 진단 퀴즈, 퀴즈 결과·새 자료로 레슨 자동 갱신	생성 문항의 측정학적 검토 권고
콘텐츠	같은 난도로 설계한 예제 묶음, 태깅에 전문가 노력 필요	업로드 자료 기반 연습 퀴즈	AI Reader 설명·퀴즈 생성(제공사 문서), 생성 힌트(품질 점검 실패 32%, 대학생 표본 아님)
일관성·투명성	전통적 방법은 전이 확률로 자기 설명 가능, 학생의 생성형 AI 사용이 숙달 추정 신호를 오도할 수 있다는 해석(ALEKS, 제공사 자체 보고·동료심사 전)	숙련도 산출 기준 미공개, 범용 LLM 의 학습자 모델링은 DKT 보다 AUC 6% 낮음(K-12 데이터, 학습 노트북 대상 아님)	—
근거 제시·가드레일	해당 없음(생성형 출력 없음)	Gemini 앱 전체에 부정확한 정보 생성 가능성 고지(도움말), 학습 노트북 전용 고지·출처 제시 방식 미확인	지정 교재 맥락만 모델에 제공(AI Reader, 제공사 문서)
연동·인증	LTI Advantage 연동(제공사 안내), ALEKS 의 LTI 1.3·LTI Advantage·데이터 프라이버시 인증 이력(2023~2025 년), CASE·Caliper·xAPI 권고(AIDT 지침, K-12)	Gemini LTI 1.3 도구 인증(2026. 3. 23.), 학습 노트북 LTI 제공 미확인	—
대표 사례	ALEKS, Knewton Alta, Realizeit	Gemini Study Notebooks	McGraw Hill AI Reader(학습자 모델 연결 미확인), OATutor 생성 힌트 실험

자료: Ⅱ 장 1 절 인용 자료와 Ⅲ장 1·2 절의 제공사 문서를 바탕으로 연구진 정리

주: '미확인'은 연구진이 확인한 공개 문서에 내용이 없다는 뜻이며 기능이 없다는 뜻은 아니다. 인증은 인증일만 적었으며 유효 여부는 계약 시점에 등록부로 확인할 사항이다.

소결 전용 적응형 엔진은 지식 공간·선수 관계 같은 도메인 모델을 미리 구축하고 지식 추적이나 IRT로 숙달·능력을 명시적으로 추정해 학습 경로를 정한다. Q1에 비추어 보면 전용 엔진과 범용 LLM 개인화는 개인화의 근거를 공개된 사전 구조에 두는지, 공개되지 않은 생성형 분해에 두는지에서 갈린다. 둘을 잇는 결합형은 연구 제안(K-12 데이터)으로 나왔고 전용 엔진 제품에 생성형 기능을 덧붙인 사례(AI Reader)는 있으나 학습자 모델과 연결되는지는 공개 문서에서 확인되지 않았다. 콘텐츠 태깅·생성 품질과 학습자 모델 신호의 신뢰성은 유형을 가리지 않는 과제로 보인다. 제품별 문서 판정은 III장 2절, Q1의 최종 답은 IV장 3절에서 다룬다.

2. 시장 동향

가. 시장 규모 추정의 한계

IMARC Group(2026)은 세계 적응형 학습 시장을 2025년 58억 달러로 추정하고 2034년 297억 달러, 2026~2034년 연평균 19.30% 성장을 전망했다. 같은 요약은 2025년 북미 점유율을 37.5% 이상으로 보았으며 2026년 7월 7일에 갱신되었다.

민간 전망과 달리 대학의 제한된 예산은 오히려 성장의 걸림돌로 거론된다. Komljenovic 외(2024)는 영국 고등교육 에듀테크 스타트업 16곳의 인터뷰 24건과 문서 540건 등을 분석했다. 스타트업 참여자들은 수익 부족과 대학의 가용 자금 부족 때문에 성장 잠재력이 제한된다고 보고했다.

IMARC Group의 추정은 방법의 세부를 공개하지 않았고 고등교육 몫도 따로 제시하지 않아 이 보고서는 세계 시장의 참고 수치로만 적는다([표 II-2]). 대학 예산의 제약은 민간 성장 전망과 별개인 요인이어서 시장 수치로 대학의 도입 규모를 가늠하기는 어렵다.

나. 교재 결합형 코스웨어와 비용 구조

미국에는 교재 비용을 등록금에 묶어 청구하는 제도가 있다. 연방 규정 34 CFR 668.164(c)(2)는 요건을 갖춘 대학이 교재·학습자료 비용을 등록금·수수료의 일부로 포함해 청구할 수 있게 한다. 이 경로의 첫 요건은 대학이 출판사 등과 약정을 맺어 자료를 시장 경쟁 가격보다 낮게 학생에게 제공할 수 있는 것이다. 대학은 결제 기간 7일째까지 학생이 자료를 받을 방법을 제공해야 하고 학생이 빠질 수 있는 선택 해제(opt-out) 정책도 두어야 한다(Disbursing Funds, 2026).

출판사 플랫폼은 대학 적응형 학습 연구에서도 가장 많이 쓰인 축에 든다. 학부생 대상 개인화 적응형 학습 연구 69편을 검토한 주제범위 문헌고찰(II장 3절 나)에서 가장 많이 쓰인 플랫폼은 Moodle과 McGraw-Hill Connect LearnSmart(각 6편)였다(du Plooy et al., 2024).

학생이 실제로 부담하는 비용은 강좌 구성과 도입 시기에 따라 달랐다. LeMire(2025)는 자동 청구된 교재·코스웨어 가격을 다른 판매처의 같은 품목 가격과 비교했다(초록 기준). 숙제 사이트 같은 코스웨어를 쓰는 강좌에서는 자동 청구가 학생 비용을 크게 줄일 수 있었으나 코스웨어 없이 교재만 쓰는 강좌에서는 자동 청구된 학생이 더 많이 낼 수 있었다. SRI International이 대학의 적응형 코스웨어 조기 도입을 독립 평가한 보고서(2016년, 재단 의뢰)에 따르면 도입 첫 학기에는 대부분 강좌의 학생 1인당 비용이 올랐다. 그러나 2·3차 도입의 비용 자료가 있는 10건 중 7건에서는 적응형 코스웨어가 더 낮은 지속 비용과 연관되었다(Yarnall et al., 2016).

교재 결합형 코스웨어는 미국에서 교재비를 등록금과 함께 청구하는 기관 경로로도 공급되며 비용 효과는 강좌 구성과 도입 시기에 따라 달라지는 것으로 정리할 수 있다. 연구진이 확인한 범위에서는 국내에서 교재·코스웨어 비용을 등록금과 함께 청구하는 제도를 찾지 못했다. 국내 대학이 이런 제품을 들이려면 비용을 대학·학생·공공 지원 가운데 누가 부담할지부터 정할 필요가 있다(IV 장 1 절 라).

다. 사업자 재편과 제품군 확장

공급 측에서는 기업공개와 합병, 제품군 확장이 이어졌다. McGraw Hill 은 2025 년 7 월 23 일 보통주 2,439 만 주를 주당 17 달러에 공모하는 IPO 가격을 확정했다. 주식은 7 월 24 일 뉴욕증권거래소에서 MH 라는 종목 코드로 거래를 시작할 예정이었고 회사는 공모 순수익을 기간 대출 일부 상환에 쓰겠다고 밝혔다(McGraw Hill, 2025).

같은 회사는 ALEKS 의 적용 범위를 넓히고 있다. ALEKS for Calculus II 발표(I 장 3 절)에 이어 Calculus III 를 2027~28 학년도에 출시할 계획이라고 밝혔다(McGraw Hill, 2026e).

AI 학습 도우미를 내놓은 제공사는 출판사와 학습 플랫폼에 걸쳐 있다. Leon(2026)은 미국 학술 출판사와 이에 준하는 8 개 제공사(Cengage·Khan Academy·Macmillan Learning·McGraw Hill·Pearson·Wiley·Quizlet·Chegg)의 AI 도우미 11 종을 2025 년 중반 시점 기준으로 평가했다.

적응형 학습 전문 기업의 합병도 있었다. 몰입형 시뮬레이션 기업 Skillwell(구 ETU)은 2025 년 12 월 9 일 AI 기반 적응형 학습 기업 Realizeit 과 합병하기로 합의했다고 발표했다. 합병 회사는 Skillwell 브랜드를 쓰며 기업과 고등교육 기관을 대상으로 한다. Realizeit 의 운영 인력은 Skillwell 에 완전히 통합된다. 합병 뒤 Skillwell Simulate 고객은 학습자의 역할·성과·진도에 따라 학습 경로를 실시간 조정하는 적응형 학습 기능 'Skillwell Adapt'를 이용할 수 있게 된다(Skillwell, 2025).

데이터 자산이 인수의 동기로 지목되기도 한다. Komljenovic 외(2024)는 기업이 보유한 사용자 데이터 때문에 인수되는 경우를 '인수를 통한 데이터 통합' 전략으로 정의했다. 같은 연구의 참여자들은 스타트업이 디지털 제품을 만들고 대학 데이터로 개발한 뒤 나중에 이용료를 받는 방식을 흔히 쓴다고 관찰했다.

출판사의 기업공개·제품군 확장과 전문 기업 간 합병이 함께 나타나 적응형 학습의 공급 구조가 바뀌고 있다. 데이터 자산이 인수의 동기가 될 수 있다는 지적은 영국 에듀테크 전반을 다룬 연구의 결과여서 적응형 제품에 한정되지 않는다. 인수·합병이 일어나는 시장에서는 사업 연속성과 데이터 이전 조항이 계약 전 확인 사항이 된다(부록 4).

라. 시험 대비 서비스와 범용 LLM 의 흡수

시험 대비 시장에서는 국내 서비스가 대규모 학습 데이터를 쌓았다. 뤼이드 소속 연구팀인 Choi 외(2020)는 AI 튜터링 시스템을 갖춘 자습 서비스 산타에서 수집한 대규모 학습 상호작용 데이터 EdNet 을 공개했다. EdNet 은 2 년 넘게 78 만여 명에게서 모은 1 억 3,000 만여 건의 상호작용을 담았다(제공사 소속 연구팀, 2020 년 발표). 시험 대비 서비스인 만큼 이 보고서는 참고 사례로만 다룬다.

범용 LLM 쪽에서는 Google 이 학습 노트북으로 적응형 기능을 흡수한 사례가 나타났다(제공 범위는 I 장 3 절). SAT 대비 학습 노트북은 The Princeton Review 의 문항에 근거한다(Walport, 2026).

가격 구조도 교재 결합형 코스웨어와 다르다. Google 은 2025 년 6 월 30 일 Gemini for Education 을 발표하면서 이 서비스가 Workspace for Education 요금제에 무료로 포함된다고 밝혔다. 이 서비스는 프리미엄 AI 모델 접근을 기본 제공하며 소비자 무료판보다 높은 한도를 예고했다(Kirtikar, 2025).

Google 사례를 보면 범용 LLM 이 적응형 기능을 흡수하는 방식은 진단 퀴즈와 진도 추적(Ⅱ장 1절 라)에 시험 대비 문항을 더하는 쪽에 가깝다. 기관용 요금제에 무료로 포함하는 방식은 교재 결합형 코스웨어의 기관 청구 구조와 다르다. 무료 포함 범위와 개인 계정·학교 계정의 조건 차이는 범용 LLM 기능을 쓰는 대학이 계약·운영 단계에서 추적할 항목이다.

마. 대학의 도입 방식과 제약

Mirata 외(2020)가 스위스·남아공 두 대학의 전문가를 대상으로 수행한 델파이에서 도입 과제는 조직 문제가 가장 많았고 교수·학습 문제가 뒤를 이었다. 전문가들은 적응형 학습 사업에 인력·재정·시간 자원이 시급히 필요하다고 보았다. 저자들은 콘텐츠가 내장된 상용 플랫폼을 사는 방식이 자체 구축과는 다른 과제를 낳을 수 있다고 지적했다. 공급자 쪽에도 제약이 있다. 영국 연구에서 스타트업은 높은 비용, 고객이 비용을 내도록 설득하는 일, 대학에 대한 사용가치 입증 같은 어려움을 겪는다고 보고되었다(Komljenovic et al., 2024).

적응형 엔진이 대학에서 쓰이는 방식은 강좌 단위 보충 학습에 그치지 않는다. 미국 한 연구중심대학의 공학 지망 입학생 자료를 분석한 Ryan 외(2025)는 문헌 검토에서 ALEKS 평가가 원래 배치시험으로 설계되지 않았고 공학 분야에서 보충 튜터링 도구로 쓰여 왔다고 소개했다(Carpenter 와 Hanna(2006) 재인용). 같은 엔진은 신입생 배치에도 쓰인다. 관찰 연구인 Ayele 외(2023)는 미국 역사적 흑인 대학(HBCU)에서 5 년간 대학 대수 이하 과목에 배치된 1,659 명의 ALEKS PPL(배치 평가 모듈) 점수·SAT·첫 과목 성적을 분석했다(초록 기준).

대학의 도입은 강좌 단위 보충 과제, 신입생 배치 평가, 기관 계약의 세 갈래로 나눠 볼 수 있다. 두 대학 델파이와 영국 스타트업 연구에서는 조직(자원 포함)과 재정 문제가 주된 제약으로 꼽혔다. 배치 용도는 학생의 수강 경로를 정하는 고부담 결정이어서 Ⅱ장 4 절 라에서 따로 다룬다. 국내에서는 공공 실증 사업(본교 운영 사업 — 이해관계)의 사용료 지원이 비용 부담을 나누는 경로가 될 수 있다(Ⅰ장 3 절, Ⅳ장 2 절 가).

[표 Ⅱ-2] 시장 지표(수치·기준 시점·출처·정의)

지표	수치	기준 시점	출처	정의·범위	성격
적응형 학습 시장 규모(민간 추정)	2025 년 58 억 달러 → 2034 년 297 억 달러, 연평균 19.30%(2026~2034)	2026. 7. 7. 갱신	IMARC Group(2026)	세계 시장, 추정 방법 비공개, 고등교육 단독 수치 없음	민간 추정
지역 구성	북미 점유율 37.5% 이상	2025	IMARC Group(2026)	세계 시장 기준	민간 추정
사업자 재편(기업공개)	McGraw Hill IPO 가격 확정(보통주 2,439 만 주, 주당 17 달러)	2025. 7. 23.	McGraw Hill(2025)	순수익은 기간 대출 일부 상환에 사용	제공사 발표
사업자 재편(합병)	Skillwell–Realizeit 합병 합의 발표	2025. 12. 9.	Skillwell(2025)	합병 회사는 Skillwell 브랜드, 기업·고등교육 대상	제공사 발표
기관판 조건(범용 LLM)	Gemini for Education 은 Workspace for Education 요금제에 무료 포함	2025. 6. 30. 발표	Kirtikar(2025)	교육 기관용 요금제 기준	제공사 발표

지표	수치	기준 시점	출처	정의·범위	성격
대학 교원 활용도(국내)	AI 기반 적응형·맞춤형 학습시스템 활용도 2.03 점, 효과 인식 2.62 점(4 점 척도)	2025	김지하 외 (2025)	일반대학 80 개교 교원 3,672 명 자기 보고	공공 조사

자료: IMARC Group(2026), McGraw Hill(2025), Skillwell(2025), Kirtikar(2025), 김지하 외(2025)를 바탕으로 연구진 정리

주: 민간 시장 추정은 추정 방법이 공개되지 않아 대학 도입 규모의 근거로 쓰지 않는다.

소결 적응형 학습 시장의 규모 추정은 방법이 공개되지 않은 민간 수치이고 공개 요약에서 고등교육 몫도 드러나지 않아 참고 수치로만 읽어야 한다. 공급의 축은 교재 결합형 코스웨어이며 출판사의 기업공개·제품군 확장과 전문 기업 간 합병이 이어졌다. Google 학습 노트북처럼 범용 LLM 이 기관판 무료 포함과 진단 퀴즈를 앞세워 적응형 기능을 흡수하는 사례도 나타났다. 비용 효과는 강좌 구성과 도입 시기에 따라 달랐다. 두 대학 전문가 델파이와 영국 스타트업 연구에서는 조직과 재정이 주된 제약으로 꼽혔다. 신입생 배치 용도는 고부담 결정의 쟁점이 되므로 II장 4 절 라에서 검토한다. 국내 도입 모델과 비용 부담 주체는 IV장 1 절 라에서 다룬다.

3. 교육적 적용

가. 메타분석이 보여 준 효과의 범위

Wang 외(2024)의 메타분석은 2010~2022 년 문헌 45 편에서 독립 효과크기 47 개를 도출했고 종합 효과크기는 g = 0.70(중간~큰 효과)이었다. 포함 연구 중 K-12 밖의 연구는 7 개뿐이었다. 조절변수별로는 출판 유형 중 학술지 논문(g = 0.78)이 가장 컸고 설계 유형 중 실험 설계(g = 0.52), 국가 중 미국 연구(g = 0.33)가 가장 작았다.

학교급이 섞인 지능형 튜터링 시스템(ITS) 통제 평가 50 건의 효과 중앙값은 0.66 표준편차였으나 시험 유형에 크게 좌우되었다(Kulik & Fletcher, 2016). 효과크기 107 개의 메타분석에서 ITS 는 교사 주도 대집단 수업(g = 0.42)보다 성취가 높았으나 개별 인간 튜터링·소집단 수업과 ITS 사이에는 유의한 차이가 없었다(Ma et al., 2014).

적응형 학습 메타분석의 g = 0.70 은 K-12 중심 값이고 ITS 메타분석의 효과도 측정 도구와 비교 대상에 따라 크게 달라져 고등교육의 기대 효과로 옮겨 읽기 어렵다(고등교육 연구는 II장 3 절 나).

나. 고등교육 대상 종합 연구

대학 ITS 메타분석은 35 개 보고서에 실린 39 개 연구(ITS 22 종)를 종합했다. 이 분석에서 ITS 는 대학생 학업 학습에 중간 수준의 정적 효과(g = 0.32~0.37)를 보였다. ITS 는 인간 튜터링보다 덜 효과적이었지만 그 밖의 수업 방법·학습 활동보다는 효과가 컸다(Steenbergen-Hu & Cooper, 2014).

학부생 대상 개인화 적응형 학습 연구 69 편을 다룬 주제범위 문헌고찰에서는 41 편(59%)이 학업 성취 향상을, 28 편(41%)이 영향 없음을 보고했다(du Plooy et al., 2024). 고등교육 시뮬레이션 기반 학습의 개인화 전략을 다룬 연구 217 편의 메타분석에서는 컴퓨터가 결정하는 적응성이 비계 제공에 더 효과적이었고 학습자가 결정하는 적응 가능성은 과제 진행에 더 유익해 보였다(Chernikova et al., 2025).

국내 AI 기반 적응형 학습 플랫폼 효과 연구 12 편을 종합한 메타분석에서 처치/통제 연구의 전체 효과크기는 0.4909 였다(K-12·대학생 혼합). 대학생 대상 효과크기는 사전/사후 0.1348, 처치/통제 0.2306 으로 K-12(0.6516, 0.8270)보다 낮았고 포함 연구 간 이질성이 높았다(전용찬 외, 2024).

고등교육 대상 근거의 효과는 중간 이하이고 연구마다 결과가 갈리며 국내 메타분석 안에서도 대학생 효과크기는 K-12 보다 낮았다(12 편, 이질성 높음). 적응을 시스템에 맡길 부분(비계)과 학습자에게 맡길 부분(과제 진행)이 다르다는 근거도 있다(시뮬레이션 기반 학습 메타분석, 적응형 코스웨어 대상 아님). 대학은 기대 효과를 K-12 수치로 잡지 말고 적응의 주체를 기능별로 나누는 설계를 고려해 봄직하다.

다. 대학 현장 평가와 실험

SRI 는 14 개 고등교육기관이 참여한 ALMAP 의 비교 집단 준실험을 종합했다. 그 결과 분석 가능한 15 건 중 4 건만 평균 성적이 약간 올랐고 대부분은 식별 가능한 영향이 없었다. 이수 효과를 추정한 16 개 자료에서도 이수 확률에 영향이 없었다. 공통 학습평가 점수를 비교한 7 건의 평균 효과는 작지만 유의했다(Yarnall et al., 2016). WWC 는 대학 온라인 코스웨어 Open Learning Initiative(OLI) 연구 12 편 중 3 편만 기준을 충족한다고 판정했고 OLI 가 학점 취득·지속은 높일 수 있으나 학업 성취 효과는 일관되지 않다고 보았다. OLI 는 즉각 피드백형 코스웨어로 분석 단위의 경계 사례다(What Works Clearinghouse, 2020).

대학 경제학 강좌의 분반 단위 배정 실험에서는 3 학기 동안 9 개 분반 중 5 개가 처치, 4 개가 통제로 배정되었다(McGraw Hill Connect 의 적응형 기능, 성적 5%인 수업 전 과제에만 적용). 주 분석 646 명에서 처치 분반의 표준화 경제 이해도 시험(TUCE) 조정 향상도가 0.044 높았고($p = .040$) 적응형 학습은 성적이 높은 학생과 여학생에게 특히 이로울 수 있었으며 다른 학생의 성과를 떨어뜨리지 않았다. 학생은 적응형 여부를 모른 채 분반을 골랐고 무작위 배정 단위는 학생이 아닌 분반이었으며 처치 단위 수도 적었다(Eau et al., 2022).

일리노이대 대형 준비 물리 강좌의 숙달형 숙제 도입 첫해에는 학생들이 좌절하고 시스템을 효과적으로 쓰지 않았다. 2 년차에는 숙달 요건 완화·버전 간 전이 축소·학생과의 직접 논의를 도입했다. 코호트 비교에서 스스로 '좌절했다'고 답한 학생은 2014 년 60%에서 2015 년 30%로 줄고 평균 숙달률은 59%에서 69%로 올랐다(Gutmann et al., 2018).

실제 강좌의 효과는 작거나 혼재했고 숙달형 숙제 사례의 개선도 운영 조정 전후의 코호트 비교여서 인과는 확인되지 않았다. 도입 계획에는 운영을 조정하는 주기가 들어가야 하며 첫 학기 결과만으로 성패를 판단하기는 이르다.

라. 제품 단위 근거

2005~2015 년 연구 15 편(독립 표본 24 개)을 종합한 Fang 외(2019)의 메타분석에서 ALEKS 는 전통적 교실 수업만큼 효과적이었으나 더 낫지는 않았다. 효과크기는 중등과 고등교육, 주 수업과 보조 수업 사이에 유의한 차이가 없었다. 2000 년~2020 년 8 월의 K-12·고등교육 33 개 연구(학생 9,238 명)를 종합한 메타분석에서 ALEKS 의 학습 성과는 전통적 수업과 비슷했다($g = 0.05$, 95% CI –0.01~0.20). 전통 수업을 보충하는 용도에서는 $g = 0.43$(95% CI 0.02~0.83)으로 효과가 컸다(Sun et al., 2021).

K-12 근거로 WWC 는 2016 년 9 월 ALEKS 방과후 수학 프로그램을 다룬 6 학년 253 명 대상 RCT 를 검토해 통계적으로 유의한 긍정적 결과가 없다고 판정했다(What Works Clearinghouse, 2016). ESSA 기준의 Evidence for ESSA 는 ALEKS 를 'No Evidence' 등급으로 분류하지만 스스로 공식 지위가 없는

정보라고 밝힌다(Center for Research and Reform in Education, 2023). Leon(2026)의 독립 리뷰(Ⅱ장 2절 다)는 독립 동료심사 연구가 있는 ALEKS의 근거 신뢰도를 '높음'으로, 제품 문서와 제공사 보고 이용 실적에 기댄 Pearson AI Study Tools를 '보통'으로 매겼다(2025년 중반 기준).

제공사와 독립된 두 메타분석 모두 ALEKS의 효과가 전통 수업과 비슷하다고 보고했으나 보충 사용의 효과는 Sun 외(2021)에서만 나타나 사용 방식 간 차이가 없다는 Fang 외(2019)와 어긋난다. 두 메타분석은 수집 기간이 겹치고(2005~2015년, 2000년~2020년 8월) 학교급도 섞여 있어 독립된 두 확인으로 세지 않았다. 효과성 증거(B9·B10) 판정에서도 두 메타분석은 참고로만 두고 고등교육 단독 독립 연구를 썼다(Ⅲ장 2절 가). Knewton Alta와 Skillwell Adapt의 효과를 검증한 독립 연구는 이번 호 조사에서 확인되지 않았다. 리뷰의 '높음'은 근거 출처 유형에 매긴 등급이어서 효과의 크기를 나타내지 않는다.

마. 생성형 AI 개인화의 초기 근거

학습 노트북 자체의 효과를 검증한 연구는 공개 자료에서 찾지 못했다. Gemini를 다룬 고등교육 연구 세 편(비무작위 준실험 2편, 2집단 비교 실험 1편)은 모두 학습 노트북을 시험하지 않았다. 온라인 MBA 학생 64명을 등록 순서로 나눈 연구에서 실험집단 32명은 Gemini 플랫폼에 통합했다고 밝힌 연구용 알고리즘(이 연구용으로 따로 설계, 2023학년도 수행)을 썼다. 이 집단의 전체 효과성 점수는 사전 6.2에서 사후 7.5로 올랐고 통계적으로 유의했다($p < .05$). 그러나 통제집단의 점수도 6.1에서 6.5로 유의하게 올랐고 효과크기는 두 집단 행에 같은 $\eta^2 = 0.09$로 적혀 비교 대상이 불분명하다. 논리 구성 영역은 두 집단 간 차이가 없었다(Panwale & Vijayakumar, 2025).

학부 프로그래밍 수업의 비동등 집단 준실험에서는 문제기반학습(PBL) 안에서 Google Gemini를 쓴 실험집단 25명과 교수 영상을 활용한 통제집단 37명을 비교했다. 프로그래밍 기술 사후 점수에서 집단 효과가 컸으나(부분 $\eta^2 = 0.338$) 이 기술은 자기 보고 척도로 측정되었고 비판적 사고와 문제해결 기술은 집단 간 유의한 차이가 없었다(Omeh et al., 2025). 예비교사 80명의 사전·사후 2집단 실험(초록 기준)은 ChatGPT와 Gemini가 설명한 같은 개념의 학습을 비교했을 뿐 무처치 대조가 없어 Gemini 나 학습 노트북의 효과 근거가 아니다(Colaco & Antao, 2025).

국내 대학 작업치료학과 86명을 개인 단위로 무작위 배정한 실험은 LLaMA 3.1·Ko-PlatYi-6B 기반 AI 교과서로 자기주도 학습을 한 집단과 인쇄 교재 집단을 비교했다. 두 집단 사이에 학업 성취·학습 시간·만족도 모두 통계적으로 유의한 차이가 없었다. 원문은 같은 결과 절에 Cohen's $d = 0.86$도 적어 내부 불일치가 있다. 사전검사 평균은 인쇄 교재 98.73점, AI 교과서 98.14점이었다(Chun et al., 2025).

McGraw Hill 직원 4명이 참여한 동료심사 전 분석(제공사 수행, 저자 해석은 Ⅱ장 1절 라)에서는 ChatGPT 공개 이후 ALEKS를 쓰는 대학생의 AI 취약 문제(AI로 풀기 쉬운 문제) 학습 시간이 11개 분기 동안 26.9% 감소했다. 감독 조건에서는 차이가 사라졌고 AI 사용은 직접 관찰하지 않았다(Rismanchian et al., 2026).

생성형 개인화의 대학 근거는 소표본·자기 보고·비무작위 설계가 중심이고 결과도 혼재한다. 온라인 MBA 연구는 2023학년도 수행이어서 Gemini 공개(2023년 12월)와 시기가 맞는지 원문으로 확인되지 않는다. 국내 무작위 배정 실험은 사전검사 평균이 이미 98점대여서 집단 차이를 가려내기 어려운 측정 조건이었을 가능성도 있다.

바. 교수자 통합과 숙달 학습 설계

학부 문헌고찰은 적응형 알고리즘이 작동하더라도 학습 설계·콘텐츠 제작·모니터링·지원에서 교수자 역할이 실행과 효과의 중심으로 남는다고 결론지었다(du Plooy et al., 2024). 대학 ITS 메타분석도 ITS 기반 학습에서 교수자와 교수법이 중요하다는 근거가 일부 있다고 보았다(Steenbergen-Hu & Cooper, 2014). ALMAP 평가에서 교수자들이 가장 크게 우려한 점은 학생들이 적응형 코스웨어를 충분히 자주 쓰게 하는 것이었다(Yarnall et al., 2016).

신종호·황영아(2024)는 대학 기초교과목에서 AI 코스웨어 기반 적응형 학습과 학습자 참여 중심 교수법을 결합하는 교수설계 가이드라인을 개발했으나 학습 효과를 검증하지는 않았다. ALEKS 를 적용한 국내 대학 일반화학 수업에서는 숙달(논문 표기 '학습 완료') 단계까지 진행해야 하는 학습량 때문에 학생 피로도가 매우 높아 과제 난이도를 '학습인정' 단계로 조절했다(황은경·신종호, 2021).

국내 대학 교양 수학 강좌의 준실험 파일럿(AI 코스웨어 분반 34 명, 대면 수업 분반 25 명, 6 주)에서 코스웨어 분반의 기말성적 평균은 더 높았고 효과크기는 Cohen's d = 0.48 로 보고됐다. 독립표본 t-검정(p = .069)과 중간성적을 통제한 공분산 분석(p = .065)에서는 모두 통계적으로 유의하지 않았다(고미화 외, 2026). 논문은 제품명을 밝히지 않았으나 본교가 운영하는 소프트랩 사업단 누리집의 실증 교과 목록에 같은 대학·교수자의 같은 교과가 풀리캠퍼스(프리윌린) 실증 교과로 올라 있어 Ⅰ장 3 절에서 밝힌 이해관계가 있을 수 있다(한림대학교 에듀테크소프트랩사업단, Retrieved 2026).

효과는 수업 통합 방식과 숙달 요건 설정에 따라 달라질 수 있고 학생의 사용 빈도는 교수자가 가장 크게 우려한 운영 과제였다. 교수자 역할과 숙달 요건을 다룬 국내외 근거를 종합한 결과다. 효과 실증 연구의 설계·대상·효과크기는 [표 Ⅱ-3]에 정리했다.

[표 Ⅱ-3] 효과 실증 연구 요약

연구	설계	대상	표본	효과크기·주요 결과	근거 수준
X. Wang 외(2024)	메타분석	K-12 중심(K-12 밖 7 개)	45 편, 효과크기 47 개	g = 0.70, 실험 설계 g = 0.52, 미국 g = 0.33	메타(K-12 중심)
Kulik 과 Fletcher(2016)	메타분석(ITS)	학교급 혼합	통제 평가 50 건	효과 중앙값 0.66 SD, 시험 유형에 크게 좌우	메타
Steenbergen-Hu 와 Cooper(2014)	메타분석(ITS)	대학생	39 개 연구 (ITS 22 종)	g = 0.32~0.37, 인간 튜터링보다 낮음	메타(고등교육)
du Plooy 외(2024)	주제범위 문헌고찰	학부생	69 편	향상 41 편(59%), 영향 없음 28 편(41%)	메타(문헌고찰)
전용찬 외(2024)	국내 메타분석	국내 K-12·대학생	12 편	처치/통제 전체 0.4909, 대학생 0.2306	메타(국내)
Yarnall 외(2016)	다기관 준실험 종합(ALMAP)	14 개 고등교육기관	학생 1 만 9,500 명 이상	15 건 중 4 건 성적 소폭 상승, 이수 확률 영향 없음	하(독립 평가)
What Works Clearinghouse(2020)	공식 판정(OLI)	대학	12 편 중 기준 충족 3 편	학점 취득·지속 향상 가능, 학업 성취 비일관	공식 판정(경계 사례)
Eau 외(2022)	분반 단위 배정 실험	미국 대학 경제학	9 개 분반(주 분석 646 명)	TUCE 조정 향상도 0.044 높음(p = .040)	중

연구	설계	대상	표본	효과크기·주요 결과	근거 수준
Gutmann 외 (2018)	코호트 비교	미국 대학 준비 물리	2 개 학년도	숙달률 59% → 69%, 좌절 응답 60% → 30%	하
Fang 외(2019)	메타분석 (ALEKS)	중등·고등교육	15 편(표본 24 개)	전통 수업과 동등, 주·보조 수업 간 차이 없음	메타(제품)
Sun 외(2021)	메타분석 (ALEKS)	K-12·고등교육	33 개 연구, 9,238 명	전체 g = 0.05, 보충 사용 g = 0.43	메타(제품)
Panwale 와 Vijayakumar(2025)	비무작위 준실험	온라인 MBA(Gemini 통합 연구용 알고리즘)	64 명	집단 내 향상 η^2 = 0.09(통제집단도 유의한 향상, 보고 불일치)	하
Omeh 외(2025)	비동등 집단 준실험	학부 프로그래밍 (Gemini)	62 명	프로그래밍 기술(자기 보고) 부분 η^2 = 0.338	하
Chun 외(2025)	개인 무작위 배정	국내 대학 작업치료학과	86 명	학업 성취·학습 시간·만족도 유의차 없음	중(국내, 보고 불일치)
고미화 외(2026)	준실험 파일럿	국내 대학 교양 수학(제품명 미표기)	59 명	기말 평균 차이 d = 0.48 이나 비유의(p = .069), 중간성적 통제 시에도 비유의(p = .065)	하(국내, 이해관계 가능)
Rismanchian 외 (2026)	준실험(이중차분)	ALEKS 이용 대학생 등	상호작용 320 만 건	AI 취약 문제 학습 시간 26.9% 감소, 감독 시 차이 없음	참고(제공사 수행·동료심사 전)

자료: 각 연구 원문(일부 초록)을 바탕으로 연구진 정리

주: 근거 수준은 중 = 동료심사 무작위 배정 실험 가운데 분반 단위 배정이거나 표본 250 명 미만인 연구, 하 = 준실험·코호트 비교, 메타 = 메타분석·문헌고찰, 공식 판정 = 정부 클리어링하우스 판정, 참고 = 제공사 수행·동료심사 전이다. 이해관계 가능은 본교 운영 사업의 실증 교과로 추정되는 연구다.

소결 적응형 학습의 효과는 학교급·설계·측정 방식에 따라 달라지며 큰 평균 효과(g = 0.70)는 K-12 중심 메타분석의 값이고 고등교육에서는 중간 이하인 데다 결과도 일관되지 않다. 효과가 보고된 사례는 구조화된 기초 강좌와 일부 학습자 집단에 모여 있고 보충 사용(메타분석 간 엇갈림, K-12·고등교육 혼합), 미시 수준 적응(Ⅱ장 1절 나), 숙달 요건의 조정, 교수자 통합이 조건으로 거론된다. 제품 단위 독립 근거는 전통 수업과 비슷한 성과를 보고한 ALEKS 메타분석(K-12·고등교육 혼합)을 빼면 부족하다. 생성형 AI 개인화의 대학 근거는 소표본·비무작위 설계 중심의 초기 단계다. 핵심 질문 Q2 의 잠정 답은 '조건에 따라 달라지는 효과와 부족한 독립 근거'에 가깝다.

4. 한계·논쟁점

Ⅱ장 4 절은 쟁점의 성격과 대응 수단을 살피고 Ⅱ장 1~4 절의 쟁점 요지를 [표 Ⅱ-4]에 모은다. 효과의 지속성·전이를 검증한 고등교육 근거는 이번 조사에서 확인하지 못했다.

가. 콘텐츠 범위와 태깅 비용

게재 당시 McGraw-Hill Education 소속이던 저자의 논문은 ITS 설계에는 교수 콘텐츠 1 시간당 200~300 시간의 개발이 든다는 Pavlik 외(2013)의 추정을 재인용했다(Essa, 2016). 지식 요소를

생성하고 연결하는 작업(Ⅱ장 1절 다)도 부담이다. 정확한 지식 요소 모델이 없으면 숙달 평가와 적응형 학습 전략의 효과가 크게 저해될 수 있고 기존 태그를 고치는 일도 처음 만드는 것만큼 시간이 들 수 있다(Moore et al., 2024).

적응형 학습 시스템은 대개 선언적 지식과 절차적 유창성 습득용으로 설계되고 개념 학습은 거의 다루지 않는다는 지적이 있다(Nigon et al., 2024). LLM 자동 문항 생성 연구 60편을 정리한 리뷰에서 다수 연구가 생성 문항의 품질을 간과했고 교과 전문가나 측정 전문가가 생성 과정에 참여한 연구는 소수였다(Tan et al., 2025).

공급이 수학·화학 같은 구조화된 기초 과목에 치우치는 데에는 높은 비용도 작용했을 수 있다. 앞서 본 지식 유형 편중 지적과 구축 비용이 이 추정의 근거다. 국내 교과에 쓰려면 한국어 콘텐츠와 교과 범위부터 확인해야 한다(Ⅳ장 1절 라).

나. 시스템 주도 경로와 학습자 자율성

미국 한 대학 재료공학 학부 수업의 무작위 배정 준실험은 연구팀이 개발한 모듈 CALM의 개인화 적응 경로를 따르게 한 설계(80명)와 안내가 적은 자율 탐색을 허용한 설계(85명)를 비교했다. 두 설계 사이에도, 같은 학점 집단 안에서도 개념 이해 점수에 통계적으로 유의한 차이는 없었고 중위 학점 집단에서는 CALM이 개념 이해를 높인다는 근거가 일부 있었다. 하위 학점 집단은 보충 자료를 가장 많이 배정받았지만 스스로 시청한 영상은 가장 적었다(Nigon et al., 2024).

ALEKS를 쓴 미국 한 대학의 화학 입문 강좌에서는 학기 말 자기조절학습(SRL) 점수(M = 83.22)가 학기 초(M = 88.07)보다 유의하게 낮았고 학생들은 동기 부족, 시스템 복잡성, 가혹한 감점 등을 가능한 요인으로 들었다. 다만 이 연구는 통제집단 없는 단일집단 사전·사후 설계이고 두 검사에 모두 참여한 학생이 120명이어서 감소의 원인을 가릴 수 없다(Harati et al., 2021).

미국 교육부 보고서(K-12 중심)에는 어떤 학생이 보충학습에 끝없이 배정되는지가 점검 질문의 예로 제시됐다(U.S. Department of Education, Office of Educational Technology, 2023). 시뮬레이션 기반 학습 메타분석(Ⅱ장 3절 나)의 저자들은 적응성과 적응 가능성이 작동 기제는 다르지만 결합하면 선행지식 수준이 다른 학습자의 요구에 대응할 수 있다고 정리했다(Chernikova et al., 2025). 고등교육 AI 개인화 학습 리뷰는 AI가 고차 역량을 기를 잠재력이 있으나 무비판적으로 쓰면 학습자 자율성을 저해할 수 있다고 보았다(Peng & Li, 2025).

근거가 준실험 1건과 통제집단 없는 단일집단 연구뿐이어서 시스템 주도 경로가 학습을 해친다고 단정할 수는 없다. 그래도 배정과 실제 이용의 괴리는 점검 대상이 된다. 앞의 준실험에서 하위 학점 집단의 배정 대비 시청 부족이 나타났고 반복 배정을 점검하라는 정책 권고도 있다. 적응의 주체를 기능별로 나누는 결합 설계(Ⅱ장 3절 나, 적응형 코스웨어 대상 근거는 아님)도 따져 볼 대안이다.

다. 알고리즘 편향과 형평성

교육 알고리즘 편향 연구를 종합한 서술적 리뷰에 따르면 한 학습자 집단으로 훈련한 모델은 새로운 집단에서 성능이 떨어지는 경향이 있으나 근거는 매우 희소해 한 건 또는 소수 사례로만 보고된 편향이 많다. 교육 데이터마이닝 연구 대부분은 학습자 인구통계를 보고하지 않고 미국 밖 학습자 집단의 실증 부족이 가장 큰 공백으로 꼽혔다(Baker & Hawn, 2022).

시뮬레이션에서 지식 추적 알고리즘은 모든 학생에게 같은 연습량을 주는 방식보다 훨씬 공평했지만 부정확한 모델에 의존하면 불공평할 수 있었고 혼합 집단에 맞춘 BKT에서 기술을 익히지 못한 학생 비율은 느린 학습자 5.5%, 빠른 학습자 0.3%였다. 저자들은 작은 불공평도 선수지식 위에 쌓일수록

누적될 수 있다고 보았다(Doroudi & Brunskill, 2019). 반면 Carnegie Learning 의 MATHia 를 쓴 미국 중·고등학생 5,856 명의 관찰 자료(K-12, 제공사 플랫폼 자료, BKT 적합도를 높이도록 콘텐츠를 조정해 온 시스템)에서 BKT 의 인구통계 집단 간 예측 성능(AUC) 차이는 0.017 로 작았다. 다만 저자들은 이 결과가 문화·사회경제·언어 배경이 다른 지역으로 이전되는지는 불확실하다고 밝혔다(Zambrano et al., 2024).

편향이 있다거나 없다고 정리할 근거는 아직 부족하고 한국 학습자를 대상으로 한 검증은 이번 조사에서 찾지 못했다. 도입 기관이 학습자 집단별 진도와 성과를 직접 점검하면 이 공백을 기관 단위에서 일부 보완할 여지가 있다(IV 장 2 절 나).

라. 배치 평가의 고부담 사용

제공사 소속 저자들의 2021 년 연구(2012~2018 년 자료, 자체 보고)에 따르면 ALEKS PPL 은 기초 수학부터 미적분 준비 과정까지 314 개 항목을 최대 29 문항으로 평가해 수강할 수학 과목 배치를 추천한다(Matayoshi et al., 2021). 다른 연구는 ALEKS 배치시험을 쓰는 미국 한 연구중심대학의 공학 지망 입학생을 퍼지 회귀불연속으로 분석했다. 이 연구에서 미적분 준비 과목(Pre-Calculus) 수강은 기준점 부근 학생의 첫 학기 중도탈락 확률을 16.51%p 높인 것으로 추정되었다(국소 효과). 미적분 I 성적에는 유의한 영향이 없었고 저자들은 80% 기준점 자체는 적절하다고 보았다. 여학생은 학업 준비도 지표가 더 높았는데도 기준 점수를 넘을 가능성이 더 낮았고 기준 점수 이상 획득의 남녀 차이 가운데 94%는 학생 특성으로 설명되지 않았다. 자료가 단일 기관에 한정되지만 저자들은 과소대표 학생에게 의도치 않은 장벽이 된다며 고부담 수학 배치시험 사용의 재검토를 권고했다(Ryan et al., 2025).

HBCU 의 5 년 행정자료를 분석한 관찰 연구(II 장 2 절 마)는 ALEKS PPL 이 기초대수 성과 예측에서 SAT 보다 나았지만 교양 수학(Math for Liberal Arts)에서는 반대였고 이 과목을 빼면 대부분 과목에서 기준 점수가 잘 작동했다고 보고했다(Ayele et al., 2023). 2016 년 ALEKS 안내 문서도 기준점(cut score)은 기관마다 사정이 다르므로 권고값일 뿐이며 기관이 자기 과목 내용을 평가해 기준점을 정해야 한다고 안내한다(ALEKS Corporation, 2016).

배치 용도에서는 기관의 기준점 검증, 집단별 결과 점검, 사람의 검토와 이의 경로가 더 중요해지며 위 근거와 EU 초안 지침의 배치 도구 예시(II 장 4 절 사)가 함께 이를 뒷받침한다. 평가척도는 이 점을 고려해 고부담 게이트를 배치·진급 결정에 쓰일 때만 적용하도록 정했다(I 장 4 절, III 장 2 절 다).

마. 학습 데이터와 프라이버시

개인정보보호위원회는 초·중등(K-12) AI 디지털교과서 점검에서 개발사 자료를 모은 KERIS 통합포털 저장소의 '국가수준학습데이터셋'처럼 학생 개개인의 상세한 학습 행동 정보가 통합 DB 에 쌓이면 개인정보자기결정권 침해 우려가 있다고 판단했다(개인정보보호위원회, 2025). McGraw Hill 최종 사용자 고지에 따르면 비식별 집계 처리한 개인정보는 기존 제품 개선과 새 교육 제품 개발에 활용된다(McGraw Hill, 2026d). Google 학습 노트북의 데이터 조건은 III장 1 절·2 절 나에서 같은 기준으로 다룬다.

개정 EU AI Act 의 신설 제 4a 조는 고위험 AI 제공자의 민감정보 예외(종전 제 10 조 제 5 항)를 옮겼다. 이 조항은 고위험 AI 배포자와 그 밖의 AI 시스템·모델의 제공자·배포자도 편향 탐지·교정에 엄격히 필요한 범위에서 민감정보를 예외적으로 처리할 수 있게 했다(Regulation (EU) 2026/1744, 2026).

학습 이력의 축적, 제품 개발을 위한 2 차 이용(영국 스타트업 사례는 Ⅱ장 2 절 다), 편향 검사 사이에는 긴장이 생기기 쉽다. 편향을 점검하려면 집단 정보가 필요하지만 그 정보도 보호 대상이다. 기관 계약의 2 차 이용 조항과 가명처리 체계 확인이 대응 수단이 될 수 있다(Ⅳ 장 1 절 나).

바. 제공사 효과 주장의 검증

Ⅰ장 4 절에서 본 WWC 자료 분석(주로 K-12 프로그램)의 저자들은 개발사가 수행하거나 의뢰한 평가의 효과크기가 독립 평가보다 상당히 큰 것이 프로그램이 단순히 더 효과적이어서 생긴 차이만은 아니라는 근거를 제시했다(Wolf et al., 2020). 학술지 논문의 효과크기가 가장 컸던 적응형 학습 메타분석(K-12 중심)에서는 깔때기 그림이 비대칭이어서 저자들이 출판 편향 가능성을 위협 요인으로 보고했다(Wang et al., 2024). 학부생 대상 주제범위 문헌고찰도 포함 연구 대부분이 동료심사 학술지 게재 연구여서 긍정적 결과 쪽 출판 편향이 있을 수 있다고 밝혔다(du Plooy et al., 2024).

국내 메타분석(K-12·대학생 혼합)에서는 반대로 학위논문 연구의 효과크기가 학술지 연구보다 컸다(전용찬 외, 2024). 대학 ITS 메타분석에서는 초기 연구, ALEKS 메타분석(중등·고등교육)에서는 ALEKS 를 짧게 사용한 연구의 효과크기가 더 컸다(Fang et al., 2019; Steenbergen-Hu & Cooper, 2014).

McGraw Hill 은 2026 년 6 월 보도자료의 초등 4~5 학년용 제품 소개에서 ALEKS 사용 학생이 대개 90% 넘게 주제를 숙달한다고 했고(K-12·근거 미제시) PPL 페이지는 6 시간 공부하면 배치가 평균적으로 한 과목 이상 올라간다고 주장한다(McGraw Hill, Retrieved 2026-c, 2026e). Ⅱ장 4 절 라의 독립 관찰 연구에서는 PPL 이 배치를 최소 1 단계 높일 확률이 38%±8%로 추정되었다(Ayele et al., 2023).

Google 은 학습 노트북 출시 글에서 가이드 학습의 성과가 입증되었다고 주장했으나 연결된 근거는 시에라리온 중학교의 자사 실험(K-12, 노트북 대상 아님)이었다(Kirtikar & Walport, 2026).

두 제공사의 수치와 주장은 독립 검증 전에는 효과 근거로 쓰기 어렵다. 메타분석 수치도 조절 요인의 방향과 크기가 연구마다 달라 신중히 읽어야 한다. 이 점을 감안해 효과성 증거(B9·B10) 판정에는 제공사와 독립된 연구만 썼다(규칙은 Ⅰ장 4 절).

사. 규제 동향

EU AI Act 의 Digital Omnibus 개정으로 부속서 III 에 따라 고위험으로 분류되는 AI 시스템에 대한 의무의 적용일이 2027 년 12 월 2 일로 정해졌다(Regulation (EU) 2026/1744, 2026).

EU 집행위원회의 초안 지침(최종 채택 전 의견수렴용)은 학습 성과 평가 결과의 쓰임으로 성적·자격 부여와 함께 이후 교수 경로를 자동으로 조정·개인화하는 경우(보충 모듈 추천, 적응형 학습 자료 등)를 든다. 그러면서도 진행 중인 학습을 지원하는 AI 는 고위험으로 분류하지 않아야 한다고 해석한다. 학생 수행에 따라 교재 난이도를 조정하는 적응형 학습 시스템이 그런 예다. 초안은 총괄평가에 근거해 교육·직업 경로에 영향을 주는 최종 결정용 평가 AI 만 부속서 III 3(b)(학습 성과 평가)에 해당한다고 보면서도 출력이 교사의 학기말 성적 결정에 쓰이는 개인화 학습 보조 AI 는 형성적 피드백을 주더라도 3(b) 고위험 예시로 들었다. 입학생의 최적 수강 수준을 정하는 AI 적응형 배치 도구도 3(c)(교육 수준 판정)의 고위험 예시다(European Commission, 2026).

초안대로라면 일반 학습용 적응형 엔진은 고위험 밖에 놓이지만 평가 결과가 배치나 성적 반영으로 이어지는 용도는 고위험으로 분류될 가능성이 크다. 최종본에서 해석이 바뀔 여지가 있으므로 대학으로서는 같은 제품이라도 용도별로 나눠 관리하는 것이 한 가지 대비책이다.

[표 Ⅱ-4] 주요 쟁점 정리

쟁점	확인된 사실(요지)	근거의 성격	확실성(연구진 판단)	대학의 대응 수단
콘텐츠 범위·태깅 비용	콘텐츠 1 시간당 개발 200~300 시간 추정(재인용), LLM 지식 요소 일치율 화학 56%·e-러닝 35%	문헌 재인용·측정 연구	중간	콘텐츠 범위·한국어 지원 확인, 생성 콘텐츠 전문가 검토
시스템 주도 경로·자율성	개념 이해 유의차 없음(중위 학점 집단 일부 이득), 하위 집단의 배정 대비 이용 부족, SRL 감소(단일집단)	준실험·단일집단	낮음~중간	시스템 비계와 학습자 선택권 결합, 배정 대비 이용 점검
알고리즘 편향·형평성	근거 희소, 시뮬레이션 미숙달 5.5% 대 0.3%, 집단 간 AUC 차이 0.017(K-12)	리뷰·시뮬레이션·관찰	낮음(엇갈림)	집단별 결과 점검
배치 평가의 고부담 사용	기준 점수 남녀 차이의 94% 미설명(단일 기관), 기준점은 권고값, 초안 지침의 고위험 예시	준실험·관찰·제공사 문서·초안 지침	중간	기관 기준점 검증, 사람 검토, 이의 경로
학습 데이터·프라이버시	국가 단위 통합 DB 의 자기결정권 침해 우려(K-12), 비식별 집계 정보의 신제품 개발 활용	감독기구 점검·제공사 고지	중간(K-12 감독기구 판단·제공사 고지)	계약상 2 차 이용 제한, 가명처리
제공사 효과 주장	개발사 평가 효과크기가 독립 평가보다 상당히 큼(주로 K-12), 출판 편향 가능성, 초기 연구·짧은 사용에서 효과 큼	메타연구·메타분석	높음	독립 연구만 효과 근거로 사용
생성형 AI 사용과 숙달 신호	ChatGPT 공개 이후 대학생 AI 취약 문제 학습 시간 26.9% 감소, 감독 조건에서는 차이 사라짐	준실험(제공사 자체 보고·동료심사 전)	낮음	감독 조건 평가 병행, 학습 신호 점검
규제	EU 고위험 의무 2027. 12. 2. 적용, 초안상 일반 적응형 학습 제외·배치 도구와 성적 연계 보조 AI 는 고위험 예시	법령·초안 지침	중간(초안)	용도별 분류·관리

자료: Ⅱ 장 1~4 절에서 인용한 연구·정책 자료를 바탕으로 연구진 정리

주: 확실성은 연구 설계와 반복 확인 여부에 따른 연구진 판단이다. '생성형 AI 사용과 숙달 신호' 행은 Ⅱ장 1 절 라·3 절 마의 요약이다.

소결 [표 Ⅱ-4]의 쟁점 가운데 실험으로 인과가 확인된 위험은 드물어 위험의 크기는 근거의 성격과 함께 읽어야 한다. 학습자 자율성과 알고리즘 편향은 근거가 희소하거나 일관되지 않다. 제공사 평가의 과대 추정 가능성은 메타연구에서 거듭 지적되었으나 배치 용도의 고부담 위험은 단일 기관 준실험과 EU 초안 지침이 제기했고 다른 관찰 연구는 대부분 과목에서 기준 점수가 작동했다고 보고했다. 학습 데이터의 2 차 이용은 제공사 고지로 확인되는 사실이며 그 위험은 K-12 감독기구 점검과 영국 질적 연구에서 지적되었다. 연구진은 쟁점 다수를 용도 구분, 데이터 계약, 사람의 검토, 독립 근거 요구 같은 운영 설계로 관리할 성격으로 보고 이를 Q2 의 독립 근거 판단과 Ⅳ장의 시사점으로 이어 간다.

Ⅲ. 대표 솔루션 심층 분석

Ⅲ장의 판정은 보증이나 배척이 아닌 도입 검토 의견이다. 판정은 자료 기준일까지 공개된 문서에 근거하며 실사용 진단 결과는 싣지 않았다(Ⅰ장 4절).

1. 솔루션 프로필

McGraw Hill은 ALEKS를 적응형 학습과 교수자 주도 과제를 결합한 수학·화학용 AI 기반 온라인 학습 플랫폼으로 소개한다(McGraw Hill, Retrieved 2026-b). 제공사는 ALEKS를 전체 성적의 15~20%(Prep 과정은 10~15%) 이상으로 반영하도록 권고한다(McGraw Hill, 2023). 생성형 기능인 AI Reader는 지정된 교재의 맥락만 모델에 제공하는 검색 증강 생성 방식이며 교수자는 언제든 모든 학생의 AI Reader를 켜거나 끌 수 있다(McGraw Hill, Retrieved 2026-a).

최종 사용자 고지는 보관 기간을 필요한 최소 기간으로만 서술하고 실제 적용되는 기간은 개인정보 담당 부서에 문의하라고 안내한다(McGraw Hill, 2026d). 국제 이용약관은 McGraw Hill이 운영하는 모든 국제 웹사이트·앱에 적용되며 교육기관이 유사한 조건의 별도 서면 계약을 맺었다면 그 계약이 소속 이용자의 이용을 규율한다(McGraw Hill, 2026c). 프로필은 [표 Ⅲ-1]과 같다.

[표 Ⅲ-1] ALEKS 프로필

항목	내용
제품 유형·대상	수학·화학 적응형 학습 플랫폼(K-12~고등교육), 배치 평가 ALEKS PPL
적응 방식	지식공간이론 기반 비생성형 엔진, 이전 응답에 따른 적응형 질의
진단·진도 표시	초기 지식 점검(최대 25문항), ALEKS 파이, 주제당 5점이면 학습함(Learned), 숙달은 Knowledge Check로 판정
교수자·관리자 통제	모듈 주제·기한·숙달 비율, Knowledge Check 예약, 선수 주제 배정
생성형 기능	AI Reader(지정 교재 RAG, 반·기관 단위 사용 설정)
이용 형태	교수자가 만든 반 단위의 기관 이용, 기관이 별도 서면 계약을 맺으면 그 계약이 이용을 규율
제공사 현황	McGraw Hill(2025. 7. 23. IPO 가격 확정), ALEKS for Calculus II 발표(2026. 6. 29.)
데이터 처리	비식별 정보의 신제품 개발 활용, 보관 기간 미명시, EU 표준계약조항 기반 국외 이전, 개인·기관 조건 구분 미확인
연동·인증	LTI Advantage 연동(제공사 안내), 1EdTech LTI 1.3 도구 인증(2023. 4. 4. 최초)·LTI Advantage 인증(2024. 6. 25.)·데이터 프라이버시 인증(2025. 7. 1.)
독립 효과 근거	메타분석 2편(K-12·고등교육 혼합, 수집 기간 중복) 전통 수업과 비슷, 보충 사용 효과는 엇갈림 / 고등교육 단독 독립 연구는 관찰 연구 2편과 준실험 비교 1건(성과 혼재)
이번 호 평가	공개 문서로 12개 항목(필수 5개) 판정, 실사용 진단이 필요한 14개 항목(게이트 B3 포함)은 미측정

자료: McGraw Hill 공식 문서와 본문 Ⅱ장·Ⅲ장 2절의 인용 자료를 바탕으로 연구진 정리

학습 노트북은 개인 계정 18세 이상부터 출시되었고 학교 계정(18세 미만 포함)에는 수 주 안에 제공될 예정이었다. 교사가 Google Classroom에서 수업 자료에 근거한 학습 노트북을 배정하고

도움이 필요한 영역을 보는 기능도 예정으로 발표되었으며 학습 노트북은 NotebookLM 과 동기화된다(Kirtikar & Walport, 2026).

학교 계정에 쓰이는 Gemini for Education 은 Workspace 핵심 서비스로서 엔터프라이즈급 데이터 보호와 관리자 통제를 갖추며 Google 은 교육 사용자의 대화 데이터를 AI 모델 개선에 쓰지 않는다고 밝혔다(Kirtikar, 2025). 학교 계정에 적용되는 데이터 처리 부속서는 삭제한 고객 데이터를 법률상 보관 의무가 없으면 최대 180 일 안에 지우도록 정한다. 고객 데이터는 위치·이전 약정을 전제로 Google 이나 하위 처리자의 시설이 있는 어느 나라에서든 처리될 수 있다(Google, 2026d). 개인 계정에 적용되는 한국 거주자 대상 추가 정보는 수집한 개인정보가 국외에 있는 Google 데이터 센터로 전송되어 보관되며 Google 이 Gemini 앱 같은 AI 모델과의 상호작용을 모델 개발·학습·개선에 사용한다고 밝힌다(Google, 2026c).

1EdTech 등록부(II 장 1 절 마)에는 ALEKS 의 데이터 프라이버시 인증(2025 년 7 월 1 일)이 기록되어 있다(1EdTech Consortium, Retrieved 2026-a). Gemini 앱·NotebookLM 등을 LMS 에 넣는 도구인 Gemini LTI 도 2026 년 3 월 23 일 받은 LTI 1.3 도구 인증으로 등록되어 있다(1EdTech Consortium, Retrieved 2026-b). 학습 노트북 프로필은 [표 III-2]와 같다.

[표 III-2] Gemini Study Notebooks 프로필

항목	내용
제품 유형·대상	범용 LLM(Gemini 앱)의 목표별 개인화 레슨 기능('적응형 학습 플랫폼'으로 소개)
적응 방식	100 개 이상 세부 목표로 분해(산출 기준 미공개), 진단 퀴즈로 기준선 설정, 퀴즈 결과·새 자료로 레슨 자동 갱신
진단·진도 표시	진단 퀴즈, 진도 대시보드('강점'·'집중 영역'·'시작 전')
교수자·관리자 통제	Google Classroom 교사 배정·도움 영역 확인 기능은 예정(출시 발표)
생성형 기능	레슨·퀴즈 생성, NotebookLM 동기화
이용 형태	18 세 이상 개인 계정에 먼저 출시, 학교 계정(18 세 미만 포함)은 수 주 안 제공 예정, Gemini for Education 은 Workspace for Education 요금제에 무료 포함
제공사 현황	Google, 학습 노트북 출시(2026. 6. 25.)
데이터 처리	학교 계정은 고객 지시 범위 처리·삭제 데이터 최대 180 일 내 파기·처리 국가 제한 없음(약정 전제)·교육 사용자 대화의 모델 개선 미사용, 개인 계정은 국외 데이터 센터 보관·AI 상호작용의 모델 개발·학습·개선 활용
연동·인증	Gemini LTI 1.3 도구 인증(2026. 3. 23.), 학습 노트북의 LTI 제공 미확인
독립 효과 근거	학습 노트북 대상 연구 없음
이번 호 평가	공개 문서로 11 개 항목(필수 5 개) 판정, 실사용 진단이 필요한 14 개 항목과 A8(공개 만족도 조사 없음)은 미측정

자료: Google 공식 문서와 본문 II장·III 장 2 절의 인용 자료를 바탕으로 연구진 정리

2. 평가척도 적용 결과

두 제품 모두 자료 기준일까지 공개된 제공사 문서·정책과 독립 연구로 판정했다. 실사용 진단으로 판정하도록 정한 AUTO·HYBRID 14 개 항목(게이트 B3 포함)은 두 제품 모두 미측정이다. 문서로 판정하는 DOC 12 개 항목 가운데 ALEKS 는 12 개, 학습 노트북은 공개 만족도 조사가 없는 A8 을 뺀 11 개를 판정했으며 두 제품이 모두 판정한 11 개가 공통 항목이다.

가. ALEKS 문서 판정 결과

ALEKS 는 학습목표 정렬(B1)과 교수 통합·교수자 지원(B6)을 충족(●)했다. 학생은 교수자가 가장 중요하다고 지정한 내용을 바탕으로 한 최대 25 문항의 초기 지식 점검(Initial Knowledge Check)으로 시작한다. 교수자는 모듈마다 다룰 주제, 시작·마감 일시 또는 숙달 비율 요건, 모듈 종료 후 Knowledge Check 를 정할 수 있다(McGraw Hill, 2023). 교수자는 알고리즘으로 생성되는 문항으로 과제·퀴즈·시험을 배정하고 학습 진도·지식 격차·과제 수행 시간 데이터를 실시간으로 볼 수 있다(McGraw Hill, Retrieved 2026-b). McGraw Hill 은 Blackboard·Brightspace·Canvas·Moodle 등 LTI Advantage 호환 LMS 를 자사 플랫폼과 연동하며 연동 대상에는 ALEKS 가 포함된다(McGraw Hill, Retrieved 2026-d).

동기·참여(B4)와 주관적 만족도(A8)는 부분충족(◐)이다. 제공사 안내서에 따르면 ALEKS 파이는 영역별로 Knowledge Check 에서 숙달한 주제, 학습했지만 아직 숙달하지 않은 주제, 남은 주제를 서로 다른 색으로 보여 준다(McGraw Hill, 2023). 그러나 국내 대학 일반화학 수업의 2020 학년도 2 학기 수강생 39 명 대상 설문에서 시스템 전체 만족도는 5 점 만점에 3.21 점이었고 숙달 단계까지의 학습량 때문에 학생 피로도가 매우 높았다(황은경·신종호, 2021). 화학 입문 적응형 강좌의 학부생 연구에서도 동기 부족, 시스템 복잡성, 가혹한 감점이 자기조절학습 점수 하락의 가능한 요인으로 나타났다(Harati et al., 2021). 사용성 척도(SUS) 기준과 비교할 공개 점수는 없었다.

표준 준수·상호운용(B8)은 제공사의 LTI Advantage 연동 안내와 1EdTech 인증 이력(Ⅱ장 1절 마)을 함께 보아 부분충족(◐)으로 판정했다(근거는 부록 2). 투명성·설명가능성(C2)도 부분충족(◐)이다. ALEKS 평가가 이전 응답에 따라 다음 문항을 고르는 적응형 질의로 진행되고 대부분의 항목을 직접 묻지 않고도 숙달·미숙달로 분류한다는 점은 제공사 소속 연구팀의 학술지 논문으로 공개되어 있다(Matayoshi et al., 2021). 그러나 McGraw Hill 은 AI 기능별 투명성 문서(AI Model Overview)를 유지한다고 밝히면서도 그 목록에 적응형 엔진 자체의 문서는 없어 엔진의 작동·한계 고지는 확인되지 않았다(McGraw Hill, Retrieved 2026-e).

데이터와 통제 관련 항목(C4·C5·C7·C8)은 모두 부분충족(◐)이다. 최종 사용자 고지는 개인정보를 학습 솔루션의 제공·유지·개선에만 쓴다고 밝히지만 비식별 집계 처리한 개인정보는 새 교육 제품 개발에도 활용되고 보관 기간은 최소 필요 기간으로만 서술된다(McGraw Hill, 2026d). 구독 기관은 서면 요청으로 12 개월에 1 회 이내에서 데이터 보안 조치를 비공식 감사나 독립 감사인의 점검으로 검토할 수 있다(McGraw Hill, 2026a). 그러나 적응형 엔진의 판정 기록을 추적하는 감사 로그와 한국 이용자 대상 국내대리인은 확인되지 않았다. ALEKS PPL 의 배치 기준점은 기관마다 사정이 달라 권고값일 뿐이며 기관은 배치 결과 데이터와 다른 지표를 근거로 기준점을 미세 조정할 수 있다(ALEKS Corporation, 2016). 배치 결과에 대한 개별 이의 제기·사람의 검토 절차는 문서로 확인되지 않았고 학습자용 AI 리터러시 지원 자료도 찾지 못했다.

ALEKS 의 효과성 증거(B9)는 고등교육 단독 독립 연구인 대학 일반화학 관찰 연구 2 편(기말시험 점수 차를 다중회귀로 추정한 1 편, 성적 우위를 자기선택으로 해석한 1 편)과 다기관 평가 안의 준실험 비교 1 건으로 판정했다(Eichler & Peeples, 2013; Reid et al., 2021). 이 준실험(2 년제 대학, ALEKS 블렌디드 실습 대 같은 교수자의 강의식 분반)에서는 사후검사의 정적 효과가 나타났고 Pell 장학생의 효과는 유의했으나 학생 성과 요약은 혼재·비유의였다(Yarnall et al., 2016). 단일 현장의 혼재된 결과여서 보통 등급에 이르지 못해 B9 는 유망 수준의 부분충족(◐)이고 세 근거 모두 표준화 효과크기 수치가 없어 B10 은 미충족(○)이다.

ALEKS 의 실제 적응 동작과 화면 사용성(게이트 B3 포함 14 개 항목)은 체험·기관 계정에서 실사용 진단을 해야 판정할 수 있다(IV 장 2 절 나). 배치 용도에 필요한 타당성·환각 관리(C1)와 공정성·편향 관리(C3)도 미측정이며 제공사 외부 연구의 관련 결과는 II장 4 절 라에 정리했다.

나. Gemini Study Notebooks 문서 판정 결과

학습 노트북의 학습목표 정렬(B1)과 동기·참여(B4)는 부분충족(◐)이다. 제공사 사용 안내에 따르면 학생이 수업 자료를 올리면 Gemini 가 맞춤 진단 퀴즈를 생성해 학업 기준선을 정하고 학습 목표를 100 개 이상의 세부 학습 목표로 나누어 진도를 추적한다. 진도 대시보드는 '강점'·'집중 영역'·'시작 전'을 표시하고 후속 레슨의 퀴즈를 마칠 때마다 자동으로 갱신된다(Walport, 2026). 그러나 강좌 학습목표와의 정렬은 학생의 자료 업로드에 기대고 교사가 Google Classroom 에서 수업 자료에 근거한 학습 노트북을 배정하는 기능은 예정으로만 발표되었다(Kirtikar & Walport, 2026). 동기·참여 효과를 제공사와 독립적으로 확인한 연구도 없다.

교수 통합·교수자 지원(B6)과 인간 통제·교수자 역할(C5)은 미충족(○)으로 판정했다. 출시 발표에서 학습 노트북은 18 세 이상 개인 계정에 먼저 제공되었고 학교 계정 제공과 교사의 배정·학급별 도움 영역 확인 기능은 예정이었다(III 장 1 절). 이에 따라 자료 기준일까지 공개된 문서로는 학습 노트북에서 교수자·기관이 개입하는 경로와 교사용 지원이 확인되지 않았다. Gemini LTI 도 Gemini 앱·NotebookLM 등을 LMS 에 넣는 도구로 등록되어 있을 뿐 학습 노트북 제공 여부는 밝혀지지 않았다(1EdTech Consortium, Retrieved 2026-b). 학습 노트북 문서에 평가·배치 같은 고부담 결정 기능은 없어 고부담 게이트는 적용하지 않았다.

표준 준수·상호운용(B8)은 부분충족(◐)이다. Google 은 2024 년 12 월 9 일 타사 학습관리시스템에 통합되는 AI 도우미 Gemini LTI 를 발표했다(Google, 2024). 2026 년 5 월 12 일에는 Gemini LTI 의 지원 범위를 기존 Canvas·Schoology 에 더해 Moodle 로 넓혔다(Google, 2026b). 1EdTech 등록부에서 확인되는 Gemini LTI 의 인증은 LTI 1.3 도구 인증(2026 년 3 월 23 일)이며 LTI Advantage 인증과 학습 노트북의 Gemini LTI 제공 여부는 확인되지 않았다(II 장 1 절 마).

효과성 증거(B9)와 효과크기(B10)는 미충족(○)이다. 학습 노트북 대상 독립 연구가 없고 Gemini 를 쓴 고등교육 연구 세 편(II 장 3 절 마)은 연구용 알고리즘이나 Gemini 일반 기능을 다뤘거나 무처치 대조가 없어 제품 단위 근거로 세지 않았다. 그중 온라인 MBA 준실험 저자들은 실험집단 알고리즘을 이 연구용으로 따로 설계해 재현성에 한계가 있다고 밝혔다(Panwale & Vijayakumar, 2025). Google LearnLM 팀은 시에라리온 중학교 12 곳에서 사전등록된 2 집단 무작위 통제 실험을 수행했으나 K-12 이자 동료심사를 거치지 않은 보고서여서 반영하지 않았다(LearnLM Team & Fab AI, 2026). 이 B9 ○는 출시 약 1 개월 뒤인 자료 기준일에 독립 연구가 아직 없다는 근거 부재를 표시하며 효과 부재 판정과는 다르다.

투명성·설명가능성(C2)은 부분충족(◐)이다. Google 은 Gemini 가 할루시네이션으로 부정확한 정보를 사실처럼 제공하거나 자신의 작동 방식을 허위로 설명할 수 있다고 고지하며 Gemini 환경에 면책 조항을 둔다고 밝힌다(Google, Retrieved 2026-b). 학습 노트북 문서는 진도 표시와 레슨 갱신을 설명하지만(II 장 1 절 라) 숙련도 판정·추천의 근거와 응답 출처 제시 방식, 학습 노트북 전용 한계 고지는 확인되지 않았다.

프라이버시·데이터 거버넌스(C4)와 책무성·거버넌스(C7)도 부분충족(◐)이다. 자료 기준일에 확인되는 학습 노트북의 제공 범위는 18 세 이상 개인 계정이며 한국 거주자 대상 추가 정보는 국외 데이터 센터 보관과 AI 상호작용의 모델 개선 활용을 밝힌다(III 장 1 절). 같은 문서는 개인정보 보호법과

정보통신망법에 따른 국내대리인으로 구글코리아 유한회사를 고지한다(Google, 2026c). 학교 계정에는 교육 사용자 대화를 모델 개선에 쓰지 않는 방침과 데이터 처리 부속서가 적용되지만 학습 노트북의 학교 계정 제공은 예정 단계였고 학습 노트북 전용 개인정보 조항은 확인되지 않았다. 「인공지능 발전과 신뢰 기반 조성 등에 관한 기본법」(AI 기본법) 제 31 조는 생성형 인공지능을 이용한 제품·서비스를 제공하려는 인공지능사업자에게 그 사실을 이용자에게 사전에 고지하도록 규정한다(인공지능 발전과 신뢰 기반 조성 등에 관한 기본법, 2026). 학습 노트북이 이 규정에 맞춘 사전 고지·결과물 표시 대응을 했는지는 문서로 확인되지 않았으며 ALEKS 의 AI Reader 도 같다(부록 2).

AI 리터러시·연령 적합성(C8)도 부분충족(◐)이다. 한국에서 자신의 Google 계정을 직접 관리할 수 있는 최소 연령은 14 세다(Google, Retrieved 2026-a). Gemini 는 18 세 미만 사용자에게 더 엄격한 콘텐츠 정책을 적용한다(Kirtikar, 2025). 그러나 대학생(성인) 학습자를 위한 학습 노트북의 AI 리터러시 지원은 확인되지 않았다.

실제 진단 퀴즈의 피드백·적응(게이트 B3)과 콘텐츠 품질·정확성(B2), 환각 관리(C1)를 포함한 14 개 항목은 기관 계정과 자료 업로드 조건에서 실사용 진단을 해야 판정할 수 있다(Ⅳ 장 2 절 나). 학습 노트북을 대상으로 한 공개 사용성·만족도 조사(A8)는 찾지 못해 사용성 축(A)은 점수를 내지 않았다.

다. 점수와 판정

판정 규칙은 Ⅰ장 4 절과 같다. 축 점수는 소수 첫째 자리까지 반올림하고 종합 점수는 반올림한 축 점수로 계산했으며 결과는 [표 Ⅲ-3]과 같다.

[표 Ⅲ-3] 평가 결과 요약

구분	ALEKS(McGraw Hill)	Gemini Study Notebooks(Google)
판정 근거	자료 기준일까지 공개된 제공사 문서·정책·독립 연구	자료 기준일까지 공개된 제공사 문서·정책·독립 연구
측정 범위	12/26 항목(A 1·B 6·C 5), 필수 5/13	11/26 항목(A 0·B 6·C 5), 필수 5/13
축 A 사용성(가중치 25)	50.0(A8 1 개 항목)	미산출(판정 항목 없음)
축 B 교육적 특성(가중치 45)	61.5	23.1
축 C AI 신뢰성·윤리·거버넌스(가중치 30)	50.0	37.5
종합 점수	55.2	28.9(축 B·C 로 계산)
공통 11 개 항목 기준	56.9	28.9
게이트 항목(B3·B9)	B3 미측정·B9 ◐	B3 미측정·B9 ○
고부담 게이트	배치·진급(ALEKS PPL)에만 적용(C1·C3·C4·C5 모두 ● 필요), 이번 판정 미적용	문서상 배치·진급 기능 없음, 미적용
도입 검토 의견	판정 유보	판정 유보
핵심 보완점	체험·기관 계정 실사용 진단, 고등교육 효과크기, 배치 결과 이의 제기 절차	기관 계정·자료 업로드 조건 실사용 진단, 교수자·기관 개입 경로, 독립 효과 연구

자료: 자료 기준일까지 공개된 공식 문서와 독립 연구의 판정을 바탕으로 연구진 정리

주: 미측정 항목은 점수에서 제외했다. 판정 항목이 없는 축은 점수를 내지 않고 종합 점수는 점수가 있는 축의 가중치로 계산했으며 공통 항목 기준 점수는 축 A 를 산출하지 않았다. 판정 기호와 항목별 근거는 부록 2 에 실었다.

두 제품 모두 필수 항목 13 개 가운데 5 개만 판정되어 최소 범위(10 개)에 못 미치므로 종합 점수와 관계없이 '판정 유보'이며 이는 결함이 아니라 측정 범위 부족을 뜻한다. ALEKS 는 미측정 14 개 항목의 판정에 따라 종합 점수가 26.7~79.5 점으로 달라질 수 있고 B9 가 ◐여서 측정을 마쳐도 조건부 검토를 넘을 수 없다. 학습 노트북은 미측정 15 개 항목에 따라 12.9~68.3 점이 될 수 있고 B9 가 ○여서 측정을 마쳐도 '신중 검토'를 넘을 수 없다. 공통 11 개 항목 기준 점수는 두 제품을 모두 판정한 문서 판정 항목만의 비교이며 학습 노트북은 판정 항목이 이 11 개와 같아 종합 점수와 같다. 이 점수들은 측정 범위가 다르고 공개 문서에 기댄 판정이므로 순위로 읽지 않는다.

고부담 게이트(배치·진급 용도)는 이번에 적용하지 않았다. 제공사 제품 안내(2021 년 연구는 최대 29 문항, Ⅱ 장 4 절 라)에 따르면 ALEKS PPL 은 25 문항 이내의 배치평가로 기초수학부터 미적분까지 과목 배치에 쓰인다(McGraw Hill, Retrieved 2026-c). EU 고위험 분류 지침 초안(Ⅱ 장 4 절 사)은 입학생의 최적 수강 수준을 정하는 AI 적응형 배치 도구를 3(c), 성적 결정에 쓰이는 개인화 학습 보조 AI 를 3(b)의 고위험 예시로 든다(European Commission, 2026). ALEKS 를 배치에 쓴다면 고부담 게이트에 필요한 C1·C3 가 미측정이고 C4·C5 는 ◐여서 요건(모두 ●)을 확인하지 못한 상태다. 제공사 권고대로 성적에 반영하면(Ⅲ 장 1 절) 학습 지원을 넘어 성적 결정 용도가 되지만 이번 호 고부담 게이트는 이 용도를 판정하지 않았다.

3. 비교 종합

두 제품은 같은 '적응형'이라는 이름 아래 구조가 다르다(Ⅰ 장 3 절, Ⅱ 장 1 절 라). ALEKS 에서 교수자는 모듈마다 다룰 주제, 숙달 비율 요건, 모듈 종료 후 Knowledge Check 를 정할 수 있다(McGraw Hill, 2023). 학습 노트북에서는 학생이 입력한 학습 목표를 Gemini 가 세부 목표로 나누며 산출 기준은 공개되지 않았다.

효과 근거에서는 ALEKS 메타분석(K-12·고등교육 혼합, Ⅱ 장 3 절 라)이 전통 수업과 비슷한 성과(g = 0.05)를 보고했다(Sun et al., 2021). 학습 노트북은 아직 독립 효과 연구가 없다.

문서에서 확인되지 않은 기능도 제품에는 있을 수 있으므로 대학이 두 유형을 견줄 기준은 우열이 아니라 [표 Ⅲ-4] 마지막 열의 같은 질문이다.

[표 Ⅲ-4] 대표 솔루션 비교

비교 차원	ALEKS	Gemini Study Notebooks	대학 도입 시 확인할 점
엔진 구조	지식공간이론 기반 비생성형 엔진	학습 목표 세부 분해(산출 기준 미공개)·생성형 진단	도메인 모델·숙달 기준의 공개 수준
진단·경로 조정 (문서)	최대 25 문항 초기 지식 점검, 숙달은 Knowledge Check 로 판정, 실제 적응 동작은 미측정	진단 퀴즈로 기준선 설정, 퀴즈 결과·새 자료로 레슨 자동 갱신 (제공사 설명), 실제 동작은 미측정	체험·기관 계정 시범 운영
근거 제시·한계 고지(문서)	엔진의 작동·한계 고지 문서 미확인	Gemini 앱 전체의 부정확한 정보 생성 가능성 고지, 학습 노트북 전용 고지·출처 제시 방식 미확인	출처 검증 절차
교수자·관리자 통제(문서)	모듈별 주제·숙달 비율, Knowledge Check 예약	교사 배정·학교 계정 제공은 예정(출시 발표)	수업 설계와 통제 수단

비교 차원	ALEKS	Gemini Study Notebooks	대학 도입 시 확인할 점
데이터 조건(문서)	보관 기간 미명시, 비식별 정보의 신제품 활용	학교 계정은 삭제 데이터 최대 180 일 내 파기, 개인 계정은 AI 상호작용의 모델 개발·학습·개선 활용	보관·국외 이전·2 차 이용 계약
효과 근거(독립)	메타분석 2 편(K-12·고등교육 혼합, 수집 기간 중복) 전통 수업과 비슷, 보충 사용 효과는 엇갈림 / 고등교육 단독 독립 연구는 관찰 연구 2 편과 준실험 비교 1 건(성과 혼재)	학습 노트북 대상 연구 없음	자체 효과 측정 계획
고부담 용도	ALEKS PPL 배치(기준점은 기관이 정함), 고부담 게이트 대상	문서상 해당 기능 없음	배치 용도 시 C1·C3·C4·C5 확인

자료: 본문 Ⅱ장과 Ⅲ장의 인용 자료를 바탕으로 연구진 정리

IV. 시사점 및 종합

1. 국내 적용 시사점

IV장 1 절의 법제·데이터 보호·대학 정책·기관 도입 모델 시사점은 [표 IV-1]에 요약했다.

[표 IV-1] 수준별 시사점

수준	확인된 사실(요지)	시사점	점검 질문
국가	AIDT(K-12) 선정 기준은 개인정보 필수·효과성 선택, 고영향 AI 교육 영역은 유아·초·중등 학생 평가	고등교육 기준이 없어 대학 자체 점검이 필요할 수 있음	배치·진단 용도에 사람의 검토 절차가 있는가
기관	고등교육 에듀테크는 도입 초기(KERIS 진단), 소프트랩 라이선스 사용료 일부 지원	공공 실증으로 비용 분담(소프트랩은 본교 운영 — 이해관계), 효과 측정은 실증 설계에 넣을 때만 근거가 됨	비용 부담 주체와 지원 종료 뒤 재원을 정했는가
교수자	교수자 통합은 효과 조건으로 거론(문헌고찰), AIDT 교사 85.5%가 미활용·중단(513 명 설문, K-12)	교수설계 지원과 수업 자율성 보장	강좌 목표와 엔진 콘텐츠가 정렬되는가
학습자	숙달 요구에 따른 피로, 영어 시스템의 어려움, 범용 LLM 의 부정확한 정보 생성 가능성(제공사 고지), 학습 이력 축적	학습 부담 조정, 범용 LLM 진단 퀴즈는 학습 보조로 한정하고 출처 검증, 데이터 권리 안내	학습 이력의 이용 범위를 안내했는가

자료: 본문 II·III·IV 장 인용 자료를 바탕으로 연구진 정리

가. 법제

개정 초·중등교육법에 따라 학교장이 학습지원 소프트웨어를 교육 자료로 선정하려면 교육부장관이 개인정보보호위원회와 협의해 정한 기준을 따르고 학교운영위원회 심의를 거쳐야 한다. 이 법은 초·중등교육에 관한 사항을 정하는 것을 목적으로 한다(초·중등교육법, 2026). 교육부 선정 기준에서 개인정보 최소수집·안전조치 등 개인정보보호법 준수 사항은 필수기준이고 교육적 효과성과 활용 적합성은 선택기준으로 다뤄진다(교육부, 2025).

「인공지능 발전과 신뢰 기반 조성 등에 관한 기본법」(AI 기본법)의 고영향 AI 교육 영역은 유아·초등·중등교육의 학생 평가다. 과학기술정보통신부 가이드라인은 초·중등 기초학력 진단검사에 AI 를 쓰면서 교사 등 교육 전문가의 검토가 없으면 고영향 AI 에 해당한다고 보고 맞춤형 콘텐츠·난이도 추천은 교육 분야 활용 사례로 들었다(과학기술정보통신부·한국지능정보사회진흥원, 2026).

대학의 적응형 학습·배치 도구는 법정 선정 기준과 고영향 규정의 명시 범위 밖으로 해석된다. 목적 조항과 고영향 교육 영역의 규정을 근거로 한 판단이다. 그렇다고 위험이 사라지지는 않으며 EU 초안 지침은 배치 도구와 성적 결정에 쓰이는 학습 보조 AI 를 고위험 예시로 든다(II 장 4 절 사). 교수자 검토를 넣는 대학 자체 기준은 따져 볼 가치가 있다.

나. 데이터 보호

개인정보보호위원회는 AIDT 사전 실태점검(K-12)에서 AIDT 가 명확한 근거 법령 없이 동의·계약이행 등을 근거로 개인정보를 처리하고 있고 학생별 이용내역인 '국가수준학습데이터셋'이 KERIS 통합포털에 저장되고 있었다고 확인했다(개인정보보호위원회, 2025).

교육부의 교육분야 가명·익명정보 처리 가이드라인(2024 년 8 월 기준)은 고등교육법상 대학에도 적용되며 가명정보는 통계작성·과학적 연구·공익적 기록보존 목적이면 정보주체 동의 없이 처리할 수 있다고 본다(교육부, 2024).

Google Cloud 개인정보처리방침은 한국 거주 사용자를 위한 추가 정보를 별도로 제공하고 서비스 데이터는 해외 데이터 센터로 전송되어 보관될 수 있다고 고지한다(Google, 2026a).

McGraw Hill 공개 문서에서는 한국 이용자 전용 조항을 찾지 못했고 기관이 별도 서면 계약을 맺으면 그 계약이 국제 이용약관에 우선한다(Ⅲ장 1 절). 학습 이력의 2 차 이용과 국외 처리 조건은 계약과 학내 가명처리 체계로 정해질 가능성이 커서 계약 단계 점검(부록 4)이 중요하다.

다. 대학 정책

KERIS 는 고등교육 소프트랩 실증 기업 모집 공고에서 에듀테크 실증을 교육 현장에서 에듀테크의 기능과 교육성·효과성을 점검하고 기업에 개선을 요청해 현장 수요를 반영하는 프로그램으로 정의했다. 소프트랩은 대학의 AI 코스웨어 실증에 드는 라이선스 사용료(구독 비용) 일부를 지원한다(한국교육학술정보원, 2026).

KERIS 는 2023 년 종합방안에서 에듀테크 소프트랩 실증 프로그램에 '교육 효과성 검증' 유형을 추가하는 방안을 제시했다(한국교육학술정보원, 2023a). 대학 교원 조사를 수행한 KEDI 연구팀도 신기술 교육방법이 실제 효과로 이어지려면 효과에 대한 실증적 검증이 먼저 필요하다고 보았다(김지하 외, 2025).

국내 대학에서 AI 기반 적응형 학습시스템의 활용도는 낮고(Ⅰ장 1 절 나) 고등교육 에듀테크는 도입 초기로 진단되어(Ⅰ장 1 절 가) 재원과 효과 근거를 함께 쌓아야 하는 단계로 보인다. 이런 여건에서는 공공 실증(소프트랩은 본교가 운영대학 — 이해관계, Ⅰ장 3 절)과 기관 단독 계약이 도입 경로로 고려될 수 있다. 어느 경로든 근거를 쌓으려면 제공사·운영기관과 독립된 효과 측정을 실증 설계에 넣는 것이 전제다(Ⅳ장 2 절 나).

라. 기관 도입 모델

감사원은 교육부가 시범 운영을 생략하고 현장적합성 검토를 제대로 하지 않은 채 AIDT(K-12) 도입을 강행했으며 구독료 예산 부담을 시·도교육청과 협의 없이 결정했다고 지적했다. 2025 년 3~5 월 접속기록에서 선정 학교 학생 중 AIDT 를 한 번도 활용하지 않은 비율은 학년·과목별 평균 60.0%였다. 수학·영어 교사 513 명 설문에서는 응답의 85.5%가 미활용·중단이었다(감사원, 2025).

교육 자료 전환 뒤인 2025 년 2 학기에도 대구의 AIDT 채택률은 81.6%로 경남·제주를 뺀 전체 채택률(14.9%, Ⅰ장 1 절 가)과 차이가 컸다(K-12). 같은 국정감사 질의에서는 대구교육청이 6 월에 AIDT 예산 51 억 원을 추가 편성했으나 채택률 감소로 49 억 원이 불용될 것이라는 보고가 언급되었다(국회사무처, 2025).

KERIS 질적 연구(K-12)는 스스로 '초기 효과' 연구라고 밝히면서 정책 측면에서 AI·디지털 교육 자료가 도입-확장-성숙 단계를 거치지 못한 '미숙의 열매'라고 평가했다(김현진 외, 2025). 교사 설문(K-12)에서는 교육적 효과에 대한 의문 등이 방해요인으로 꼽혔고 교사들은 수업 자율성 보장과 제도적 안전망을 핵심 조건으로 제시했다(고보경 외, 2026).

교육 자료로 지위가 바뀐 뒤 채택은 크게 줄었고 대구의 채택률은 대구를 뺀 집계(10.4%)보다 크게 높게 남았다. 시범 운영과 현장 검토, 재원 협의를 거치지 않은 정책 주도 도입은 법적 지위와 재원

조건이 바뀌면 지속되기 어려울 수 있음을 이 경과가 보여 준다. 시범 운영 생략과 낮은 활용 같은 약점은 교과서 지위였던 2025 년 1 학기부터 있었다([표 IV-2]).

국내 대학 일반화학 수업의 ALEKS 사례연구에서 학생들이 겪은 어려움은 영어로 된 시스템(45%), 예측할 수 없는 학습 시간(35%), 숙달을 요구하는 토픽 학습(20%) 순이었다(황은경·신종호, 2021).

국내에서는 코스웨어 비용을 등록금과 함께 청구하는 제도가 확인되지 않았다(Ⅱ장 2절 나). 학생 개인 계정에 맡기는 방식은 학생 부담과 데이터 조건의 차이를, 공공 실증의 사용료 지원(Ⅰ장 3절)에 기댄 기관 라이선스는 지원 종료 뒤 재원을 과제로 남긴다. 기관판 무료 포함 조건도 제공사 정책에 따라 바뀔 수 있다(Ⅱ장 2절 라). 비용 부담 주체와 지원 종료 뒤 재원, 한국어 화면과 국내 교육과정에 맞춘 지식 구조·문항 범위(Ⅱ장 4절 가)는 계약 전에 정해 둘 사항으로 보인다.

[표 IV-2] AIDT 경과에서 본 도입 유형 점검

점검 차원	AIDT(K-12)	출판사 코스웨어형(ALEKS 등)	범용 LLM 개인화형(학습 노트북)
도입·시범 운영	시범 운영 생략, 현장적합성 검토 미흡	강좌 단위 도입, 기관 계약	18세 이상 개인 계정에 먼저 출시, 학교 계정은 제공 예정
효과 근거	확인된 공식 연구는 질적 초기 연구	ALEKS: 독립 메타분석(K-12·고등교육 혼합) 전통 수업과 비슷·보충 사용 효과 엇갈림 / 다른 출판사·전문 기업 제품: 독립 효과 연구 미확인	대상 독립 연구 미확인
교수자 역할	교사 85.5% 미활용·중단(513명 설문)	교수자가 모듈·숙달 비율 설정	교사 배정 기능은 예정(출시 발표)
이용 지속	학생 60.0% 미활용(2025. 3.~5.)	성적 15~20% 이상(전체 강좌) 반영 권고(제공사, 성적 연계 시 고부담 점검)	학습자 개인 선택
재정 구조	구독료 예산 부담을 시·도교육청과 협의 없이 결정, 전환 뒤 채택률 14.9%(대구 81.6%)	교재비 자동 청구(미국)·기관 계약	기관판(Gemini for Education) 무료 포함
데이터 보호	근거 법령 없는 처리, 국가 단위 데이터 저장	비식별 정보 2차 이용, 보관 기간 미명시	학교 계정은 고객 지시 범위 처리, 개인 계정은 AI 상호작용의 모델 개선 활용

자료: 본문 Ⅱ·Ⅲ장과 Ⅳ장 1절 인용 자료를 바탕으로 연구진 정리

주: 범용 LLM 개인화형은 자료 기준일 시점의 제공 범위(18세 이상 개인 계정 우선)를 기준으로 적었다.

2. 본교·연구회 활용 제언

가. 본교 적용 사례(관계 공개)

이 보고서는 HAI-GPT를 운영하는 본교 AI에듀테크센터의 연구과제 산출물이며, 연구진에는 해당 부서·센터 구성원이 포함된다. 이 센터는 프리윌린과 AI-HiED 협약(2025년 12월)을 맺었고 본교는 소프트랩 사업의 운영대학이며 연구진 일부는 실증 교과를 맡았다(Ⅰ장 3절). 이 관계를 고려해 풀리캠퍼스·AI-HiED·소프트랩 AI 코스웨어는 채점하지 않는다. 제품명을 밝혀 학습 효과를 검증한 공개

연구는 찾지 못했고 소프트랩 실증 교과로 추정되는 교양 수학 파일럿 1 편(Ⅱ 장 3 절 바)은 통계적으로 유의한 차이를 보이지 않았다.

본교 AI-HiED 플랫폼은 풀리캠퍼스와 SmartLEAD 의 LTI 연동(Ⅰ 장 3 절)을 바탕으로 기초학력 진단·AI 맞춤형 학습·보완 콘텐츠·평가·수업 설계 기능을 하나의 흐름으로 묶고 2025 년 동계 계절학기부터 적용되도록 설계됐다(협약 발표 보도)(문화춘, 2025). 소프트랩 실증 기업의 최종 선정은 복수의 대학이나 교수자가 매칭을 원하는 기업으로 한정된다(한국교육학술정보원, 2026).

연구회 계획서에 따르면 본교 기초프로그래밍 학습 플랫폼은 상호작용 로그·과제 수행 기록·오답 데이터를 생성하고 4 단계 자동 오류 분석 파이프라인을 갖췄다(연구진 자기 기술). 연구회의 프로그래밍 오류 데이터셋 연구(연구 3)는 데이터셋 구축 연구이며 학습 효과나 인과를 검증하지 않는다(김슬기, 2026).

나. 제언

교무 부서와 교수학습개발센터는 기초 교과(수학·화학 등) 보완 학습에 적응형 코스웨어를 쓸 때 '진단 → 보완 → 재진단' 순환을 교수설계 지침(Ⅱ 장 3 절 바)과 조교·멘토 지원으로 뒷받침하고 숙달 요건과 학습 부담은 소규모 시범 운영으로 조정한 뒤 넓히는 방안을 고려할 수 있다. 황은경·신종호(2021)는 학생들이 적응형 학습활동을 유용성보다 과제·평가 부담으로 인식했다고 해석했다. 이 방식의 효과는 확인되지 않았고 본교 AI-HiED 가 같은 방식이어서 이해관계가 있다.

입학·교무 부서로서는 배치·진급에 쓰는 진단의 기준점을 기관 자료로 검증하고 집단별 결과 점검과 사람의 검토·이의 경로를 갖추는 것이 한 방안이 될 수 있다. 적응형 학습 결과를 성적에 크게 반영하는 강좌에도 같은 점검을 적용해 봄 직하다(Ⅱ 장 4 절 라). 2016 년 ALEKS PPL 안내 문서도 기관이 배치 결과 데이터와 다른 지표를 근거로 기준점을 미세 조정할 수 있다고 안내한다(ALEKS Corporation, 2016).

교무 부서는 도입·실증 단계부터 분반 단위 무작위 배정(어려우면 비교 집단·지연 측정)을 넣어 KERIS 한국형 에듀테크 증거 기준(안)의 무작위대조실험·인과관계 규명 유형(유형 3)에 가까운 근거를 쌓는 설계를 검토할 수 있다(한국교육학술정보원, 2023a). 협약·운영 관계가 있는 제품은 관계없는 연구진이 분석하거나 분석 계획을 미리 공개하고 결과 보고에 관계를 밝힐 만하다. 집단별 진도·성과와 보충 배정 쏠림 점검도 이 설계에 담아 둘 가치가 있다(Ⅱ 장 4 절 나·다).

AI 에듀테크센터와 개인정보 보호 부서는 학습 데이터를 가명처리·목적 제한 체계(Ⅳ 장 1 절 나)로 관리하며 연구에 잇고 외부 계약에서는 비식별 정보의 2 차 이용 범위(Ⅱ 장 4 절 마)를 확인할 수 있다. 연구회 계획서에 따르면 연구 3 데이터셋은 가명정보 특례와 교육 분야 비식별화 제도에 맞춘 IRB(기관생명윤리위원회) 기반 이중 비식별화로 공개 데이터셋을 구축하려 한다(김슬기, 2026).

Google Workspace for Education 을 쓰는 대학의 정보화 부서와 교수학습개발센터는 학교 계정에 학습 노트북이 제공될 때에 대비해 사용 범위를 정하고 진단 결과를 성적·배치에 쓰지 않으며 학생 안내에 출처와 서지를 검증하는 방법을 넣는 방안을 검토할 수 있다. 출시 발표에 따르면 교사가 Google Classroom 에서 수업 자료에 근거한 학습 노트북을 배정하고 도움이 필요한 영역을 보는 기능은 예정 단계였다(Kirtikar & Walport, 2026). Google 도 Gemini 가 할루시네이션으로 부정확한 정보를 사실처럼 제공할 수 있다고 밝힌다(Google, Retrieved 2026-b).

연구회는 연구책임자 승인과 소프트웨어활용비 집행 가능 여부 확인을 거쳐 ALEKS 체험 계정과 학습 노트북 기관 계정·자료 업로드 조건에서 표준 과업으로 실사용 진단을 하고 미측정 항목(게이트 B3

포함)의 판정을 채점표 보완판으로 정리한다. 연구 3 의 4 단계 오류 유형은 지식 요소 태그 체계로 정리해 적응형 피드백 실험에 쓸 수 있는 형식으로 공개하는 방안을 검토하고 교양 AI 교육 프레임워크(연구회 설계안)의 데이터·알고리즘 기초소양 영역과 잇는다. 문항 분석·기준점 검증 절차는 '심리측정'·'에듀테크데이터분석' 실습 과제로 연결해 정기회의에서 공유한다.

3. 종합

Q1 에 대해, 전용 엔진은 지식 공간·선수 관계 같은 도메인 모델을 미리 구축하고 학습자 응답으로 숙달 확률이나 능력을 추정해 경로를 정한다. 범용 LLM 1 종을 다룬 K-12 공개 데이터 비교 연구(학습 노트북 대상 아님)에서 LLM 의 학습자 모델링은 DKT 보다 예측 정확도와 시간적 일관성이 낮아 전용 학습자 모델과 생성형 인터페이스를 짝짓는 결합형이 제안되었다(Ⅱ 장 1 절 라). 제공사 문서상 학습 노트북은 학습 목표를 100 개 이상의 세부 목표로 나누고 생성한 진단 퀴즈로 진도를 추적하며 사전 구조를 쓰는지와 숙련도 산출 기준은 공개되지 않았다. 두 제품의 실제 적응 동작은 실사용 진단이 필요한 항목이어서 이번 호에서는 판정하지 않았다(Ⅲ 장 2 절).

Q2 의 경우 고등교육의 효과는 K-12 중심 메타분석(g = 0.70)보다 작게 보고되었다(대학 ITS g = 0.32~0.37, 12 편을 종합한 국내 메타분석의 대학생 하위집단 0.2306, 이질성 높음). 효과나 개선이 보고된 사례는 경제학·물리·화학·수학 같은 구조화된 기초 강좌에서 주로 나왔고 성적 상위·여학생(경제학 분반 실험)이나 중위 학점 집단(재료공학 준실험)에 이로울 수 있다는 결과가 있다. 보충 사용(메타분석 간 엇갈림), 미시 수준 적응(제품 간 비교), 숙달 요건 조정(코호트 비교, 좌절·숙달률 지표), 교수자 통합(문헌고찰)이 효과와 관련된 조건으로 거론되지만 인과가 확인된 조건은 아니다. 제품 단위 독립 근거는 ALEKS 메타분석(K-12·고등교육 혼합, 고등교육 하위집단 추정치 없음)에 모여 있고 고등교육 단독 근거는 ALEKS 도 관찰 연구 2 편과 다기관 평가 안의 준실험 비교 1 건(성과 혼재, Ⅲ 장 2 절 가)에 그쳐 대표 제품 모두 충분하지 않다. Knewton Alta·Skillwell Adapt 와 학습 노트북에 대해서는 학습 효과를 검증한 독립 연구를 찾지 못했다(Ⅱ 장 3 절 라·마).

Q3 에 대해, AIDT(K-12)는 2025 년 1 학기에 도입 대상 학교의 34.2%가 자율선정해 활용했으나 교육자료 전환 뒤인 2 학기 채택률은 국정감사 질의에서 14.9%로 제시되었다(집계 기준이 다름, Ⅰ 장 1 절 가). 감사원은 시범 운영 생략과 현장적합성 검토 미흡, 시·도교육청과 협의 없는 구독료 예산 결정, 학생 60.0% 미활용(2025 년 3~5 월 접속기록)을 지적했다. 전환 뒤에는 예산을 추가 편성했던 대구의 채택률이 대구를 뺀 집계보다 크게 높게 남았다(K-12, Ⅳ 장 1 절 라). 법정 선정 기준이 없는 대학이 스스로 점검할 항목으로는 효과 근거 확인, 소규모 시범 운영 뒤 확대, 교수자의 수업 자율성 보장, 지원 종료 뒤 재원, 학습 이력의 가명처리를 들 수 있다(부록 4).

이번 호의 근거로 보면 적응형 학습의 효과는 적응 단위, 콘텐츠와 진단의 타당성, 교수자의 수업 통합 방식에 따라 달라질 수 있고 정책 주도로 도입된 AIDT(K-12)는 법적 지위가 바뀐 뒤 채택이 크게 줄었다. 대표 솔루션은 공개 문서만으로 필수 항목을 충분히 판정할 수 없어 두 제품 모두 판정 유보이며 점수는 순위가 아니다. 효과성 증거(B9)로 보면 측정을 마쳐도 ALEKS 는 조건부 검토, 학습 노트북은 신중 검토가 상한이다. 대학은 학습 지원과 배치·성적 반영 같은 고부담 결정을 구분하고 학내 효과 측정을 전제로 소규모 시범 운영에서 단계적으로 넓혀 갈 수 있다.

다음 호 예고

Vol.03 AI 튜터링 플랫폼(2026 년 8 월분) — 이번 호에서 경계로 남긴 대화형 지도를 이어받아 대화형 AI 튜터의 교수법 설계와 실증 근거를 다룬다

부록

부록 1. 참고 자료 목록

가. 국내 자료

Google. (Retrieved 2026-b). Gemini 앱의 대답에 관해 자세히 알아보기. Gemini 앱 고객센터. https://support.google.com/gemini/answer/16279220?hl=ko

Google. (2026a, 4 월 8 일). Google Cloud 개인정보처리방침. Google Cloud. https://cloud.google.com/terms/cloud-privacy-notice?hl=ko

Google. (2026c, 5 월 13 일). 한국 거주자를 위한 개인정보 관련 추가 정보. https://policies.google.com/privacy/additional?hl=ko&gl=KR

감사원. (2025). 감사보고서: AI 디지털교과서 도입 관련(국회감사요구). https://www.bai.go.kr/bai/result/branch/detail?srno=3279

개인정보보호위원회. (2025, 5 월 15 일). 개인정보위, AI 디지털교과서 사전 실태점검 결과 발표 [보도자료]. https://www.pipc.go.kr/np/cop/bbs/selectBoardArticle.do?bbsId=BS074&mCode=C020010000&nttId=11198

고미화, 김진아, 김윤경. (2026). 대학 수업에서의 AI 코스웨어 적용 및 효과를 탐색한 파일럿 연구. 교양교육연구, 20(2), 223–239. https://doi.org/10.46392/kjge.2026.20.2.223

고보경, 김유민, 정재현, 이채린, 곽헌동. (2026). AI 디지털교과서(AIDT) 도입의 방해요인과 해결방안: 현장 교사의 인식과 윤리적 쟁점을 중심으로. 교육공학연구, 42(2), 545–581. https://doi.org/10.17232/KSET.42.2.545

과학기술정보통신부, 한국지능정보사회진흥원. (2026). 고영향 인공지능 판단 가이드라인 (2026. 4. 29. 수정본). https://www.nia.or.kr/common/board/Download.do?bcIdx=28987&cbIdx=99835&fileNo=27

교육부. (2024). 교육분야 가명·익명정보 처리 가이드라인 (2024 년 8 월 개정). https://privacy.moe.go.kr/board_download.do?file_uid=758

교육부. (2025, 12 월 30 일). 학교의 안전하고 효과적인 학습지원 소프트웨어 선정을 위한 기준 안내 [보도자료]. https://www.moe.go.kr/boardCnts/viewRenew.do?boardID=294&boardSeq=105007&lev=0&searchType=null&statusYN=W&page=1&s=moe&m=020402&opType=N

국회사무처. (2025, 10 월 22 일). 2025 년도 국정감사 교육위원회회의록 제 2 반: 대구광역시교육청·경상북도교육청·강원특별자치도교육청. https://record.assembly.go.kr/assembly/viewer/minutes/download/pdf.do?id=55525

김슬기. (2026). Python 프로그래밍 학습에서 문제 유형별 학습자 오류 분류 체계 수립 및 벤치마크용 스키마를 갖춘 공개 데이터셋 구축 [미간행 연구계획서]. 한림대학교 AI 에듀테크 연구회.

김지하, 김정아, 김은영, 백승주, 손윤희, 조옥경, 김소형. (2025). 데이터 기반 대학 교육 혁신 모니터링 연구(Ⅲ) (연구보고 RR 2025-35). 한국교육개발원. https://www.kedi.re.kr/khome/main/research/selectPubForm.do?plNum0=16265

김현진, 이광호, 이상기, 이재진, 차현진, 김현영. (2025). 학습자 특성에 따른 AI·디지털 교육자료 활용 수업 교육효과 질적 연구(요약본) (연구보고 KR 2025-09). 한국교육학술정보원. https://www.keris.or.kr/main/ad/pblcte/selectPblcteRRInfo.do?mi=1138&pblcteSeq=13972

남혜현. (2025, 9 월 30 일). 뤼이드, '소크라 AI'로 사명 변경. 바이라인네트워크. https://byline.network/2025/09/30-443/

문화춘. (2025, 12 월 9 일). 한림대–프리윌린, 맞춤형 교육 실현을 위한 'AI-HiED 플랫폼' 구축 업무 협약 체결. 한국대학신문. https://news.unn.net/news/articleView.html?idxno=587279

박수빈. (2025, 7 월 9 일). 매스프레소, AI 학습 에이전트 '콴다 3.0' 출시. AI 타임스. https://www.aitimes.com/news/articleView.html?idxno=200475

신종호, 황영아. (2024). 대학 기초교육에서 AI 코스웨어 활용 수업을 위한 교수설계 가이드라인 개발 및 타당화. 교육정보미디어연구, 30(2), 571–594. https://doi.org/10.15833/kafeiam.30.2.571

인공지능 발전과 신뢰 기반 조성 등에 관한 기본법, 법률 제 21311 호 (2026). https://www.law.go.kr/LSW/lsInfoP.do?lsiSeq=282791&chrClsCd=010202&urlMode=lsInfoP&efYd=20260721&ancYnChk=0

전용찬, 김재현, 염제명, 이유경, 송지훈. (2024). 국내 AI 기반 적응형 학습 플랫폼 활용 교육의 교육적 효과에 대한 메타분석. 교육정보미디어연구, 30(6), 1763–1791. https://doi.org/10.15833/kafeiam.30.6.1763

초·중등교육법, 법률 제 21424 호 (2026) (제 29 조의 2 신설: 법률 제 21013 호, 2025. 8. 14. 일부개정). https://www.law.go.kr/LSW/lsInfoP.do?lsiSeq=283903&efYd=20260911&chrClsCd=010202&ancYnChk=0

한국교육학술정보원. (2023a). 디지털 기반 교육혁신을 위한 에듀테크 활성화 종합방안 (연구자료 RM 2023-12). https://www.keris.or.kr/main/ad/pblcte/selectPblcteRMInfo.do?mi=1139&pblcteSeq=13724

한국교육학술정보원. (2023b). AI 디지털교과서 개발 가이드라인 (GM 2023-11). https://webst.edunet.net/AIDT/AI%20%EB%94%94%EC%A7%80%ED%84%B8%EA%B5%90%EA%B3%BC%EC%84%9C%20%EA%B0%9C%EB%B0%9C%20%EA%B0%80%EC%9D%B4%EB%93%9C%EB%9D%BC%EC%9D%B8.pdf

한국교육학술정보원. (2026, 3 월 16 일). 2026 년 고등교육 에듀테크 소프트랩 실증 희망 기업 모집 공고. 기업마당. https://www.bizinfo.go.kr/sii/siia/selectSIIA200Detail.do?pblancId=PBLN_000000000119902

한림대학교 에듀테크소프트랩사업단. (Retrieved 2026). AI 코스웨어 활용 교육과정 운영. https://www.hiedsoftlab.com/ko/model/operation

황은경, 신종호. (2021). 대학 기초 과학 수업에서의 인공지능 기반 적응형 학습 시스템의 도입과 적용에 관한 탐색적 고찰: 일반화학 수업 사례 중심으로. 교양교육연구, 15(6), 71–86. https://doi.org/10.46392/kjge.2021.15.6.71

나. 국외 자료

1EdTech Consortium. (Retrieved 2026-a). ALEKS [Product certification listing]. TrustEd Apps Directory. https://site.imsglobal.org/certifications/mcgraw-hill/aleks

1EdTech Consortium. (Retrieved 2026-b). Google Gemini [Product certification listing]. TrustEd Apps Directory. https://site.imsglobal.org/certifications/google-llc/google-gemini

Abdelrahman, G., Wang, Q., & Nunes, B. (2023). Knowledge tracing: A survey. ACM Computing Surveys, 55(11), Article 224. https://doi.org/10.1145/3569576

ALEKS Corporation. (2016). Cut score recommendations. McGraw-Hill Education. https://cmn.edu/wp-content/uploads/alekstestinginformation.pdf

Anthropic. (2025, October 23). Expanding our use of Google Cloud TPUs and services. https://www.anthropic.com/news/expanding-our-use-of-google-cloud-tpus-and-services

Area9 Lyceum. (2026, April 27). Closing the gap in the learning ecosystem: Area9 Lyceum presents AI-powered content revolution at LEARNTEC 2026 [Press release]. https://area9lyceum.com/closing-the-gap-in-the-learning-ecosystem-area9-lyceum-presents-ai-powered-content-revolution-at-learntec-2026/

Ayele, A. D., Carson, Z., & Tameze, C. (2023). An efficacy study of ALEKS-based placement in entry-level college math courses. PRIMUS, 33(4), 414–430. https://doi.org/10.1080/10511970.2022.2073623

Baker, R. S., & Hawn, A. (2022). Algorithmic bias in education. International Journal of Artificial Intelligence in Education, 32(4), 1052–1092. https://doi.org/10.1007/s40593-021-00285-9

Binger, M. (2018, April 2). How does Knewton's Proficiency Model estimate student knowledge in Alta? Wiley. https://www.wiley.com/en-us/insights/how-does-knewton-proficiency-model-estimate

Center for Research and Reform in Education. (2023, June 16). ALEKS [Program entry]. Evidence for ESSA. https://evidenceforessa.org/program/aleks/

Chernikova, O., Sommerhoff, D., Stadler, M., Holzberger, D., Nickl, M., Seidel, T., Kasneci, E., Küchemann, S., Kuhn, J., Fischer, F., & Heitzmann, N. (2025). Personalization through adaptivity or adaptability? A meta-analysis on simulation-based learning in higher education. Educational Research Review, 46, Article 100662. https://doi.org/10.1016/j.edurev.2024.100662

Choi, Y., Lee, Y., Shin, D., Cho, J., Park, S., Lee, S., Baek, J., Bae, C., Kim, B., & Heo, J. (2020). EdNet: A large-scale hierarchical dataset in education. In I. I. Bittencourt, M. Cukurova, K. Muldner, R. Luckin, & E. Millán (Eds.), Artificial intelligence in education: 21st International Conference, AIED 2020, Ifrane, Morocco, July 6–10, 2020, proceedings, part II (pp. 69–73). Springer. https://doi.org/10.1007/978-3-030-52240-7_13

Chun, J., Kim, J., Kim, H., Lee, G., Cho, S., Kim, C., Chung, Y., & Heo, S. (2025). A comparative analysis of on-device AI-driven, self-regulated learning and traditional pedagogy in university health sciences education. Applied Sciences, 15(4), Article 1815. https://doi.org/10.3390/app15041815

Colaco, D. M., & Antao, D. (2025). Exploring self-learning efficacy: A comparative analysis of ChatGPT and Gemini. Interactive Technology and Smart Education, 22(4), 688–698. https://doi.org/10.1108/ITSE-12-2023-0251

Competition and Markets Authority. (2024). Alphabet Inc.'s partnership with Anthropic PBC: Decision on relevant merger situation (ME/7108/24). https://assets.publishing.service.gov.uk/media/676959bae6ff7c8a1fde9d33/Full_text_decision__.pdf

Corbett, A. T., & Anderson, J. R. (1995). Knowledge tracing: Modeling the acquisition of procedural knowledge. User Modeling and User-Adapted Interaction, 4(4), 253–278. https://doi.org/10.1007/BF01099821

Disbursing Funds, 34 C.F.R. § 668.164 (2026). https://www.ecfr.gov/current/title-34/subtitle-B/chapter-VI/part-668/subpart-K/section-668.164

Doroudi, S., & Brunskill, E. (2019). Fairer but not fair enough on the equitability of knowledge tracing. In Proceedings of the 9th International Conference on Learning Analytics & Knowledge (pp. 335–339). Association for Computing Machinery. https://doi.org/10.1145/3303772.3303838

du Plooy, E., Casteleijn, D., & Franzsen, D. (2024). Personalized adaptive learning in higher education: A scoping review of key characteristics and impact on academic performance and engagement. Heliyon, 10(21), Article e39630. https://doi.org/10.1016/j.heliyon.2024.e39630

Dziuban, C., Howlin, C., Moskal, P., Johnson, C., Parker, L., & Campbell, M. (2018). Adaptive learning: A stabilizing influence across disciplines and universities. Online Learning, 22(3), 7–39. https://doi.org/10.24059/olj.v22i3.1465

Eau, G., Hoodin, D., & Musaddiq, T. (2022). Testing the effects of adaptive learning courseware on student performance: An experimental approach. Southern Economic Journal, 88(3), 1086–1118. https://doi.org/10.1002/soej.12547

Eichler, J. F., & Peeples, J. (2013). Online homework put to the test: A report on the impact of two online learning systems on student performance in general chemistry. Journal of Chemical Education, 90(9), 1137–1143. https://doi.org/10.1021/ed3006264

Essa, A. (2016). A possible future for next generation adaptive learning systems. Smart Learning Environments, 3(1), Article 16. https://doi.org/10.1186/s40561-016-0038-y

European Commission. (2026). Draft Commission guidelines on the classification of high-risk AI systems under Article 6 of Regulation (EU) 2024/1689 (AI Act) for stakeholder consultation. https://digital-strategy.ec.europa.eu/en/library/draft-commission-guidelines-classification-high-risk-ai-systems

Fang, Y., Ren, Z., Hu, X., & Graesser, A. C. (2019). A meta-analysis of the effectiveness of ALEKS on learning. Educational Psychology, 39(10), 1278–1292. https://doi.org/10.1080/01443410.2018.1495829

Google. (Retrieved 2026-a). Age requirements on Google Accounts. Google Account Help. https://support.google.com/accounts/answer/1350409?hl=en

Google. (2024, December 9). Introducing Gemini Learning Tools Interoperability (LTI™). Google Workspace Updates. https://workspaceupdates.googleblog.com/2024/12/Gemini-Learning-Tools-Interoperability.html

Google. (2026b, May 12). Google Workspace Assignments LTI™ and Gemini LTI™ are now available for Moodle. Google Workspace Updates. https://workspaceupdates.googleblog.com/2026/05/workspace-assignments-lti-and-gemini-lti-for-moodle.html

Google. (2026d, June 8). Cloud data processing addendum (customers). Google Cloud. https://cloud.google.com/terms/data-processing-addendum

Gutmann, B., Gladding, G., Lundsgaard, M., & Stelzer, T. (2018). Mastery-style homework exercises in introductory physics courses: Implementation matters. Physical Review Physics Education Research, 14(1), Article 010128. https://doi.org/10.1103/PhysRevPhysEducRes.14.010128

Harati, H., Sujo-Montes, L., Tu, C.-H., Armfield, S., & Yen, C.-J. (2021). Assessment and Learning in Knowledge Spaces (ALEKS) adaptive system impact on students' perception and self-regulated learning skills. Education Sciences, 11(10), Article 603. https://doi.org/10.3390/educsci11100603

Hooshyar, D., Yang, Y., Šír, G., Kärkkäinen, T., Hämäläinen, R., Cukurova, M., & Azevedo, R. (2026). Why large language models alone fall short for responsible learner modeling in K–12 tutoring: A case study. IEEE Access, 14, 91267–91284. https://doi.org/10.1109/ACCESS.2026.3701047

IMARC Group. (2026, July 7). Adaptive learning market size, share, trends and forecast by component, application, deployment, end use, and region, 2026–2034. https://www.imarcgroup.com/adaptive-learning-market

Kirtikar, A. (2025, June 30). New Gemini tools for students and educators. Google. https://blog.google/products-and-platforms/products/education/gemini-iste-2025/

Kirtikar, A., & Walport, C. (2026, June 25). Supporting students with connected AI tools for more personalized learning. Google. https://blog.google/products-and-platforms/products/education/iste-students-2026/

Komljenovic, J., Birch, K., & Sellar, S. (2024). Monetising digital data in higher education: Analysing the strategies and struggles of EdTech startups. Postdigital Science and Education, 6(4), 1196–1215. https://doi.org/10.1007/s42438-024-00505-0

Kulik, J. A., & Fletcher, J. D. (2016). Effectiveness of intelligent tutoring systems: A meta-analytic review. Review of Educational Research, 86(1), 42–78. https://doi.org/10.3102/0034654315581420

LearnLM Team & Fab AI. (2026). Teaching with Gemini: Measuring the impact of Guided Learning on student mathematics progress in Sierra Leone [Technical report]. Google. https://storage.googleapis.com/deepmind-media/LearnLM/learnLM_sierraleone_may26.pdf

LeMire, S. (2025). Automatic textbook billing cost savings: Does the reality match the hype? Technical Services Quarterly, 42(3–4), 280–296. https://doi.org/10.1080/07317131.2025.2512283

Leon, M. (2026). Publisher-built generative AI assistants in U.S. higher education: A critical review and a reproducible TRIAD–JTBD evaluation framework. Algorithms, 19(6), Article 492. https://doi.org/10.3390/a19060492

Ma, W., Adesope, O. O., Nesbit, J. C., & Liu, Q. (2014). Intelligent tutoring systems and learning outcomes: A meta-analysis. Journal of Educational Psychology, 106(4), 901–918. https://doi.org/10.1037/a0037123

Matayoshi, J., Cosyn, E., & Uzun, H. (2021). Are we there yet? Evaluating the effectiveness of a recurrent neural network-based stopping algorithm for an adaptive assessment. International Journal of Artificial Intelligence in Education, 31(2), 304–336. https://doi.org/10.1007/s40593-021-00240-8

McGraw Hill. (Retrieved 2026-a). AI Reader technical information and disclosures. https://www.mheducation.com/highered/digital-products/ai/disclosures.html

McGraw Hill. (Retrieved 2026-b). ALEKS. https://www.mheducation.com/highered/digital-products/aleks.html

McGraw Hill. (Retrieved 2026-c). ALEKS Placement, Preparation and Learning. https://www.mheducation.com/highered/digital-products/aleksppl.html

McGraw Hill. (Retrieved 2026-d). LMS: Integrated tools for easy access. https://www.mheducation.com/highered/services/learning-management-system-integration.html

McGraw Hill. (Retrieved 2026-e). Our AI approach. https://www.mheducation.com/about-us/trust-center/our-ai-approach.html

McGraw Hill. (2023). ALEKS reference guide: Your go-to guide for navigating ALEKS. https://www.mheducation.com/content/dam/mhe/support/support-pdfs/migrated/article-25222.pdf

McGraw Hill. (2025, July 23). McGraw Hill announces pricing of its initial public offering [Press release]. https://investors.mheducation.com/news/news-details/2025/McGraw-Hill-Announces-Pricing-of-its-Initial-Public-Offering/default.aspx

McGraw Hill. (2026a, January 30). Data privacy and security guidelines. https://www.mheducation.com/about-us/trust-center/data-security-policy.html

McGraw Hill. (2026b, March 24). ALEKS frequently asked questions. McGraw Hill PreK–12 Support. https://www.mheducation.com/prek-12/support/knowledge/aleks-general-tech-questions.html

McGraw Hill. (2026c, April 24). Terms of use for international platforms. https://www.mheducation.com/about-us/policy-center/terms-use/international.html

McGraw Hill. (2026d, May 22). End user data privacy notice. McGraw Hill Privacy Center. https://www.mheducation.com/about-us/trust-center/privacy.html

McGraw Hill. (2026e, June 29). McGraw Hill expands its AI-powered ALEKS product portfolio [Press release]. https://www.mheducation.com/about-us/news-insights/press-releases/mcgraw-hill-expands-its-ai-powered-aleks-product-portfolio.html

Mirata, V., Hirt, F., Bergamin, P., & van der Westhuizen, C. (2020). Challenges and contexts in establishing adaptive learning in higher education: Findings from a Delphi study. International Journal of Educational Technology in Higher Education, 17(1), Article 32. https://doi.org/10.1186/s41239-020-00209-y

Moore, S., Schmucker, R., Mitchell, T., & Stamper, J. (2024). Automated generation and tagging of knowledge components from multiple-choice questions. In Proceedings of the Eleventh ACM Conference on Learning @ Scale (pp. 122–133). Association for Computing Machinery. https://doi.org/10.1145/3657604.3662030

Nigon, N., Tucker, J. D., Ekstedt, T. W., Jeong, B. C., Simionescu, D. C., & Koretsky, M. D. (2024). Adaptivity or agency? Educational technology design for conceptual learning of materials science. Computer Applications in Engineering Education, 32(6), Article e22790. https://doi.org/10.1002/cae.22790

Omeh, C. B., Ayanwale, M. A., Mnguni, L. E., & Olelewe, C. J. (2025). Fostering programming skill and critical thinking through AI-assisted PBL integration. Journal of New Approaches in Educational Research, 14(1), Article 22. https://doi.org/10.1007/s44322-025-00041-0

Panwale, S. B., & Vijayakumar, S. (2025). Evaluating AI-personalized learning interventions in distance education. The International Review of Research in Open and Distributed Learning, 26(1), 157–174. https://doi.org/10.19173/irrodl.v26i1.7813

Pardos, Z. A., & Bhandari, S. (2024). ChatGPT-generated help produces learning gains equivalent to human tutor-authored help on mathematics skills. PLOS ONE, 19(5), Article e0304013. https://doi.org/10.1371/journal.pone.0304013

Pearson. (Retrieved 2026). AI-powered tools for higher education. https://www.pearson.com/en-us/higher-education/products-services/aistudytool.html

Peng, J., & Li, Y. (2025). Frontiers of artificial intelligence for personalized learning in higher education: A systematic review of leading articles. Applied Sciences, 15(18), Article 10096. https://doi.org/10.3390/app151810096

Piech, C., Bassen, J., Huang, J., Ganguli, S., Sahami, M., Guibas, L., & Sohl-Dickstein, J. (2015). Deep knowledge tracing. In C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, & R. Garnett (Eds.), Advances in Neural Information Processing Systems 28 (pp. 505–513). Curran Associates. https://papers.nips.cc/paper_files/paper/2015/hash/bac9162b47c56fc8a4d2a519803d51b3-Abstract.html

Regulation (EU) 2026/1744 of the European Parliament and of the Council of 8 July 2026 amending Regulations (EU) 2024/1689, (EU) 2018/1139 and (EU) 2023/1230 as regards the simplification of the implementation of harmonised rules on artificial intelligence (Digital Omnibus on AI). (2026). Official Journal of the European Union, L, 2026/1744. https://eur-lex.europa.eu/eli/reg/2026/1744/oj

Reid, S. A., MacBride, L., Nobile, L., Fiedler, A. T., & Gardinier, J. R. (2021). Implementation and evaluation of an adaptive online summer preparatory course for general chemistry: Whom does it benefit? Chemistry Education Research and Practice, 22(2), 303–311. https://doi.org/10.1039/d0rp00283f

Rismanchian, S., Uzun, H., Matayoshi, J., Cosyn, E., & Kurd-Misto, E. (2026). Faster completion, less learning: Generative AI reduced study time on math problems and the knowledge they build (arXiv:2605.21629v3). arXiv. https://doi.org/10.48550/arXiv.2605.21629

Ryan, O., Sajadi, S., Barrera, S., & Tavakoli Jaghargh, R. (2025). Understanding the effects of a math placement exam on Calculus 1 enrollment and engineering persistence. Education Sciences, 15(2), Article 154. https://doi.org/10.3390/educsci15020154

Skillwell. (2025, December 9). Skillwell merges with Realizeit to supercharge personalized skill development [Press release]. PR Newswire. https://www.prnewswire.com/news-releases/skillwell-merges-with-realizeit-to-supercharge-personalized-skill-development-302634768.html

Skillwell. (2026, July 14). Skillwell announces personalized, skills-based learning directly inside Blackboard LMS [Press release]. PR Newswire. https://www.prnewswire.com/news-releases/skillwell-announces-personalized-skills-based-learning-directly-inside-blackboard-lms-302824565.html

Steenbergen-Hu, S., & Cooper, H. (2014). A meta-analysis of the effectiveness of intelligent tutoring systems on college students' academic learning. Journal of Educational Psychology, 106(2), 331–347. https://doi.org/10.1037/a0034752

Sun, S., Else-Quest, N. M., Hodges, L. C., French, A. M., & Dowling, R. (2021). The effects of ALEKS on mathematics learning in K-12 and higher education: A meta-analysis. Investigations in Mathematics Learning, 13(3), 182–196. https://doi.org/10.1080/19477503.2021.1926194

Tan, B., Armoush, N., Mazzullo, E., Bulut, O., & Gierl, M. J. (2025). A review of automatic item generation techniques leveraging large language models. International Journal of Assessment Tools in Education, 12(2), 317–340. https://doi.org/10.21449/ijate.1602294

U.S. Department of Education, Office of Educational Technology. (2023). Artificial intelligence and the future of teaching and learning: Insights and recommendations (ED631097). ERIC. https://files.eric.ed.gov/fulltext/ED631097.pdf

Walport, C. (2026, June 25). 5 ways to learn with study notebooks in the Gemini app. Google. https://blog.google/innovation-and-ai/products/gemini-app/gemini-study-notebooks/

Wang, X., Huang, R., Sommer, M., Pei, B., Shidfar, P., Rehman, M. S., Ritzhaupt, A. D., & Martin, F. (2024). The efficacy of artificial intelligence-enabled adaptive learning systems from 2010 to 2022 on learner outcomes: A meta-analysis. Journal of Educational Computing Research, 62(6), 1348–1383. https://doi.org/10.1177/07356331241240459

What Works Clearinghouse. (2016, September). WWC review of this study [Review of the study The impact of a technology-based mathematics after-school program using ALEKS on student's knowledge and behaviors, by S. D. Craig, X. Hu, A. C. Graesser, A. E. Bargagliotti, A. Sterbinsky, K. R. Cheney, & T. Okwumabua]. U.S. Department of Education, Institute of Education Sciences. https://ies.ed.gov/ncee/wwc/Study/82204

What Works Clearinghouse. (2020). Supporting Postsecondary Success intervention report: Open Learning Initiative (OLI). U.S. Department of Education, Institute of Education Sciences. https://ies.ed.gov/ncee/WWC/Docs/InterventionReports/WWC_InterventionReport_OLI_508.pdf

Wiley. (Retrieved 2026). Features: Knewton Alta. https://www.wiley.com/en-us/learning/teaching-resources/courseware/knewton-alta/features

Wolf, R., Morrison, J., Inns, A., Slavin, R., & Risman, K. (2020). Average effect sizes in developer-commissioned and independent evaluations. Journal of Research on Educational Effectiveness, 13(2), 428–447. https://doi.org/10.1080/19345747.2020.1726537

Yarnall, L., Means, B., & Wetzel, T. (2016). Lessons learned from early implementations of adaptive courseware. SRI Education. https://www.sri.com/wp-content/uploads/2021/12/almap_final_report.pdf

Zambrano, A. F., Zhang, J., & Baker, R. S. (2024). Investigating algorithmic bias on Bayesian knowledge tracing and carelessness detectors. In Proceedings of the 14th Learning Analytics and Knowledge Conference (pp. 349–359). Association for Computing Machinery. https://doi.org/10.1145/3636555.3636890

Zhuang, Y., Liu, Q., Bi, H., Huang, Z., Huang, W., Li, J., Yu, J., Liu, Z., Hu, Z., Hong, Y., Pardos, Z. A., Ma, H., Zhu, M., Wang, S., & Chen, E. (2026). Survey of computerized adaptive testing: A machine learning

perspective. IEEE Transactions on Pattern Analysis and Machine Intelligence, 48(8), 8744–8763. https://doi.org/10.1109/TPAMI.2026.3672850

부록 2. 평가척도 채점표

두 대표 솔루션의 항목별 판정과 근거는 다음과 같다. 판정은 자료 기준일까지 공개된 문서에 근거했다(Ⅰ장 4절).

대상 솔루션 ALEKS (McGraw Hill) **평가자** 연구진(문서 판정, 채점 초안 작성에 생성형 AI 보조) **접근 방식** 자료 기준일까지 공개된 제공사 문서·정책·연구로 판정했다(실사용 진단 결과는 이번 호에 싣지 않음).

ID	항목	측정	가중	판정	근거·출처
A. 사용성 (가중치 25%)					
A1	과업 완료 성과	AUTO	필수	미측정	실사용 진단이 필요한 항목이라 공개 문서만으로는 판정하지 않았다(실사용 진단 결과는 이번 호에 싣지 않음).
A2	효율성·반응성	AUTO	권장	미측정	실사용 진단이 필요한 항목이라 공개 문서만으로는 판정하지 않았다(실사용 진단 결과는 이번 호에 싣지 않음).
A3	시스템 상태 가시성	HYBRID	권장	미측정	실사용 진단이 필요한 항목이라 공개 문서만으로는 판정하지 않았다(실사용 진단 결과는 이번 호에 싣지 않음).
A4	오류 예방·복구	HYBRID	필수	미측정	실사용 진단이 필요한 항목이라 공개 문서만으로는 판정하지 않았다(실사용 진단 결과는 이번 호에 싣지 않음).
A5	일관성·표준 준수	HYBRID	보조	미측정	실사용 진단이 필요한 항목이라 공개 문서만으로는 판정하지 않았다(실사용 진단 결과는 이번 호에 싣지 않음).
A6	접근성 — 자동 검출 영역	AUTO	필수	미측정	실사용 진단이 필요한 항목이라 공개 문서만으로는 판정하지 않았다(실사용 진단 결과는 이번 호에 싣지 않음).
A7	접근성 — 수동 검증 영역	HYBRID	필수	미측정	실사용 진단이 필요한 항목이라 공개 문서만으로는 판정하지 않았다(실사용 진단 결과는 이번 호에 싣지 않음).
A8	주관적 만족도(SUS 등)	DOC	권장	◐	제공사 외부의 국내 대학 일반화학 수업 사례연구에서 2020 학년도 2 학기 수강생 39명의 시스템 전체 만족도는 5점 만점에 3.21점이었다. 같은 설문에서 꼽힌 어려움은 영어로 된 시스템(45%)·예측할 수 없는 학습 시간(35%)·숙달을 요구하는 토픽 학습(20%)이었다. 화학 입문 적응형 강좌 학부생 연구도 동기 부족·시스템 복잡성·가혹한 감점 등을 자기조절학습 점수 하락의 가능한 요인으로 들었으며, SUS 기준과 비교할 점수는 없어 부분 충족으로 판정했다. (출처: 황은경·신종호, 2021; Harati et al., 2021)
B. 교육적 특성 (가중치 45%)					

ID	항목	측정	가중	판정	근거·출처
B1	학습목표 정렬	DOC	필수	●	학생은 교수자가 가장 중요하다고 지정한 내용을 바탕으로 한 최대 25 문항의 초기 지식 점검으로 시작한다. 교수자는 모듈마다 다룰 주제·시작·마감 일시 또는 숙달 비율 요건·모듈 종료 후 Knowledge Check 를 정해 목표 주제와 평가가 정렬된다. 반 단위 ALEKS Pie Report 는 숙달·미숙달·학습 준비·시도 주제를 보여 주고, 기본 주제 구성은 내용 전문가 팀이 대부분의 강좌에서 흔히 다루는 주제를 조사해 정한다. (출처: McGraw Hill, 2023; McGraw Hill, 2026b)
B2	콘텐츠 품질·정확성	HYBRID	필수	미측정	실사용 진단이 필요한 항목이라 공개 문서만으로는 판정하지 않았다(실사용 진단 결과는 이번 호에 싣지 않음).
B3	피드백·적응	HYBRID	필수*	미측정	(게이트 항목) 실사용 진단이 필요한 항목이라 공개 문서만으로는 판정하지 않았다(실사용 진단 결과는 이번 호에 싣지 않음).
B4	동기·참여	DOC	권장	◐	ALEKS 파이는 영역별로 숙달·학습·남은 주제를 색으로 구분해 보여 주고, 한 주제에서 5 점을 얻으면 진도 표시가 채워지며 축하 화면이 뜬다. 그러나 화학 입문 적응형 강좌 연구는 동기 부족·가혹한 감점·사회적 실재감 부족을 자기조절학습 점수 하락의 가능한 요인으로 들었다. 국내 일반화학 수업 사례에서는 숙달 단계까지의 학습량으로 피로도가 매우 높았고, 연구진은 학생들이 적응형 학습활동을 유용성보다 새 과제·평가 부담으로 인식했다고 해석했다. (출처: McGraw Hill, 2023; McGraw Hill, 2026b; Harati et al., 2021; 황은경·신종호, 2021)
B5	교수설계·인지부하	HYBRID	권장	미측정	실사용 진단이 필요한 항목이라 공개 문서만으로는 판정하지 않았다(실사용 진단 결과는 이번 호에 싣지 않음).
B6	교수 통합·교수자 지원	DOC	권장	●	교수자는 알고리즘 생성 문항으로 과제·퀴즈·시험을 배정하고 학습 진도·지식 격차·과제 수행 시간 데이터를 실시간으로 보며, 반 단위 ALEKS Pie Report 로 수업 방향을 정하고 준비도별로 학생을 묶을 수 있다. 모듈 편집·Knowledge Check 예약·권장 선수 주제 배정 기능과 Blackboard·Brightspace·Canvas·Moodle 등 LTI Advantage 호환 LMS 연동이 공식 문서로 확인된다. (출처: McGraw Hill, Retrieved 2026-b; McGraw Hill, 2023; McGraw Hill, Retrieved 2026-d)
B7	상호작용 사용성·접근성(교육 맥락)	HYBRID	보조	미측정	실사용 진단이 필요한 항목이라 공개 문서만으로는 판정하지 않았다(실사용 진단 결과는 이번 호에 싣지 않음).

ID	항목	측정	가중	판정	근거·출처
B8	표준 준수·상호운용	DOC	보조	◐	McGraw Hill 은 Blackboard·Brightspace·Canvas·Moodle 등 LTI Advantage 호환 LMS 를 자사 플랫폼과 연동하며 연동 대상에 ALEKS 를 포함한다고 밝힌다. 1EdTech 등록부에는 ALEKS 의 LTI 1.3 도구 인증(2023. 4. 4. 최초)과 LTI Advantage 인증(2024. 6. 25.) 기록이 있으나 기준일 시점의 인증 유효 여부는 확인하지 않아 부분 충족으로 판정했다. (출처: McGraw Hill, Retrieved 2026-d; 1EdTech Consortium, Retrieved 2026-a)
B9	효과성 증거 등급	DOC	필수*	◐	(게이트 항목) 제공사 독립·고등교육 단독·제품 단위 조건을 충족하는 연구는 대학 일반화학 강좌의 관찰 연구 2 편과 다기관 평가(ALMAP) 안의 2 년제 대학 수학 준실험 비교 1 건이다. 학업·사회경제 배경을 통제한 다중회귀 연구에서 강좌 전 ALEKS 과제를 마친 학생은 비참여 학생보다 기말시험 점수가 평균 13 점 넘게 높을 것으로 추정되었다. 반면 예비 모듈 연구는 모듈을 절반 넘게 마친 학생의 성적 우위를 자기선택으로 해석했다. 준실험 비교(ALEKS 블렌디드 실습 대 같은 교수자의 강의식 분반)는 사후검사에서 정적 효과(Pell 장학생은 유의)를 보였으나 보고서 요약표는 학생 성과를 혼재·비유의로 정리했다. 단일 현장이어서 보통 등급(준실험의 유의한 정적 효과, 2 개 이상 현장·350 명 이상)에 이르지 못해 통계적 통제를 둔 상관 연구 수준(유망)의 부분 충족으로 판정했다. 참고: K-12·고등교육 혼합 메타분석 2 편은 고등교육 하위집단 추정치 없이 모두 전통 수업과 비슷한 효과를 보고했다(Sun 외 Hedges g = 0.05). Sun 외는 보충 용도에서 효과가 컸으나 Fang 외에서는 사용 방식에 따른 차이가 없었다. 독립 리뷰의 근거 신뢰도 '높음'은 근거 출처 유형의 등급이며, WWC 가 검토한 6 학년 무작위 통제 실험(유의한 긍정적 결과 없음)과 Evidence for ESSA 의 'No Evidence' 분류는 K-12 근거다. (출처: Eichler & Peeples, 2013; Reid et al., 2021; Yarnall et al., 2016; Fang et al., 2019; Sun et al., 2021; Leon, 2026; What Works Clearinghouse, 2016; Center for Research and Reform in Education, 2023)

ID	항목	측정	가중	판정	근거·출처
B10	효과크기 정량	DOC	권장	○	제공사 독립·고등교육 단독·제품 단위 조건을 충족하는 관찰 연구 2 편은 초록 기준으로 표준화 효과크기(d·g)를 보고하지 않았다. 다기관 평가(ALMAP)의 준실험 비교도 효과크기를 그림으로만 제시해 수치가 없어 미충족으로 판정했다. 참고: 다중회귀 연구의 차이는 기말시험 원점수 13 점 이상이었다. K-12·고등교육 혼합 메타분석 가운데 Sun 외는 전체 Hedges g = 0.05(95% CI –0.01~0.20)와 전통 수업 보충 용도 g = 0.43(95% CI 0.02~0.83)을 보고했으나 고등교육 하위집단 값은 없다. 대학생 대상 ITS 메타분석의 g = 0.32~0.37 은 ALEKS 를 포함한 ITS 전체 값이라 ALEKS 효과크기로 쓰지 않았다. (출처: Eichler & Peeples, 2013; Reid et al., 2021; Yarnall et al., 2016; Sun et al., 2021; Steenbergen-Hu & Cooper, 2014)
C. AI 신뢰성·윤리·거버넌스 (가중치 30%)					
C1	타당성·신뢰성과 환각 관리	HYBRID	필수	미측정	실사용 진단이 필요한 항목이라 공개 문서만으로는 판정하지 않았다(실사용 진단 결과는 이번 호에 싣지 않음). 참고: 제공사 외부 대학 연구는 ALEKS PPL 이 기초대수 성과 예측에서 SAT 보다 나았지만 교양 수학(Math for Liberal Arts)에서는 반대였다고 보고했다. 배치 용도라면 고부담 게이트에 필요한 C1 ●를 확인하지 못한 상태다. (출처: Ayele et al., 2023)
C2	투명성·설명가능성	DOC	필수	◐	제공사는 ALEKS 를 AI 기반 학습 플랫폼으로 소개하고 숙달 상태를 파이의 색 구분과 점수 규칙(한 주제에서 5 점이면 학습함(Learned)으로 표시되고 숙달은 Knowledge Check 로 판정)으로 안내한다. 이전 응답으로 다음 문항을 고르고 대부분의 항목은 직접 묻지 않고 분류하는 평가 방식은 제공사 소속 연구팀의 학술지 논문으로 공개되어 있다. 그러나 McGraw Hill 이 AI 기능별로 유지하는 투명성 문서(AI Model Overview) 목록에 적응형 엔진 자체의 문서는 없어 엔진의 작동·한계 고지는 확인되지 않았다. (출처: McGraw Hill, Retrieved 2026-b; McGraw Hill, 2023; McGraw Hill, 2026b; Matayoshi et al., 2021; McGraw Hill, Retrieved 2026-e)

ID	항목	측정	가중	판정	근거·출처
C3	공정성·편향 관리	HYBRID	필수	미측정	실사용 진단이 필요한 항목이라 공개 문서만으로는 판정하지 않았다(실사용 진단 결과는 이번 호에 싣지 않음). 참고: 제공사 외부의 단일 기관 연구에서 여학생은 학업 준비도 지표가 더 높았는데도 ALEKS 수학 배치시험의 기준 점수를 넘을 가능성이 더 낮았다. 남녀 차이의 94%가 학생 특성으로 설명되지 않아 저자들은 고부담 수학 배치시험 사용의 재검토를 권고했다. (출처: Ryan et al., 2025)
C4	프라이버시·데이터 거버넌스	DOC	필수	◐	최종 사용자 고지는 개인정보를 솔루션 제공·유지·개선에만 쓰고 판매하지 않으며 국외 이전은 EU 표준계약조항 등에 근거한다고 밝힌다. 데이터는 256 비트 이상으로 암호화되고 ALEKS 는 1EdTech 데이터 프라이버시 인증(2025. 7. 1.)을 받았다. 생성형 AI 기능(AI Reader)은 모델 학습용 데이터를 공급사로 보내지 않고 입력·출력 기록은 제품 개선에 쓰고 1 년 뒤 폐기한다. 그러나 보관 기간은 최소 필요 기간으로만 적고 실제 기간은 개인정보 담당 부서에 문의하게 하며, 비식별 집계 정보를 새 교육 제품 개발에 쓰고, 한국 이용자 대상 조항·국내대리인은 확인되지 않았다. (출처: McGraw Hill, 2026d; McGraw Hill, 2026a; 1EdTech Consortium, Retrieved 2026-a; McGraw Hill, Retrieved 2026-a)
C5	인간 통제·교수자 역할	DOC	필수	◐	교수자는 모듈별 주제·숙달 비율 요건을 정하고 Knowledge Check 를 예약하거나 예약 평가 전 최대 7 일 동안 자동 점검을 막을 수 있다. 학생은 Knowledge Check 를 직접 요청할 수 없고 교수자는 AI Reader 를 언제든 켜고 끌 수 있다. ALEKS PPL 의 배치 기준점은 기관마다 사정이 달라 권고값일 뿐이고 기관이 과목별 D/F/W·합격률로 평가해 미세 조정하며(2016 년 제공사 안내 문서), PPL 은 최대 5 회 응시를 제공한다. 그러나 PPL 배치 결과에 대한 개별 이의 제기·사람의 검토·결정 수정 절차는 문서로 확인되지 않아 부분 충족으로 판정했다(재응시는 재측정 기회로 보았다). 배치 결정에 쓰면 고부담 게이트가 적용되어 C1·C3·C4·C5 가 모두 ●여야 하나 이번 판정에서 C1·C3 는 미측정, C4·C5 는 ◐이다. 참고: 성적에 반영하는 기관은 II장 4 절 사의 용도 구분에 따라 따로 점검한다. (출처: McGraw Hill, 2023; McGraw Hill, 2026b; McGraw Hill, Retrieved 2026-a; ALEKS Corporation, 2016; McGraw Hill, Retrieved 2026-c)
C6	안전·견고성·보안	HYBRID	권장	미측정	실사용 진단이 필요한 항목이라 공개 문서만으로는 판정하지 않았다(실사용 진단 결과는 이번 호에 싣지 않음).

ID	항목	측정	가중	판정	근거·출처
C7	책무성·거버넌스·정책 정합	DOC	권장	◐	보관 기간 등은 개인정보 담당 부서(Privacy Office)에 문의하도록 안내되고, 교육기관과의 별도 서면 계약이 이용을 규율하며, 구독 기관은 12 개월에 1 회 이내에서 데이터 보안 조치를 감사할 수 있다. 그러나 적응형 엔진의 판정 기록을 추적하는 감사 로그, 한국 이용자 대상 국내대리인, ALEKS 에 통합된 생성형 AI 기능(AI Reader)의 AI 기본법 제 31 조 사전 고지·결과물 표시 대응은 확인되지 않았다. EU 집행위원회 초안 지침은 입학생의 수강 수준을 정하는 AI 적응형 배치 도구를 고위험 예시로 들지만, 과목 배치에 쓰이는 ALEKS PPL 의 고위험 적합성 대응 문서도 확인되지 않았다. (출처: McGraw Hill, 2026d; McGraw Hill, 2026c; McGraw Hill, 2026a; McGraw Hill, Retrieved 2026-e; 인공지능 발전과 신뢰 기반 조성 등에 관한 기본법, 2026; European Commission, 2026; McGraw Hill, Retrieved 2026-c)
C8	AI 리터러시·연령 적합성	DOC	보조	◐	최종 사용자 고지는 18 세 미만 이용자에게 보호자와 함께 읽도록 권하고 상업용 웹사이트는 13 세 미만 이용을 상정하지 않는다고 밝힌다. 국제 이용약관은 18 세 미만이어도 고등교육기관 재학생이면 본인이 약관에 동의할 수 있게 한다. 학습자를 위한 AI 리터러시 지원 자료는 확인되지 않았다. (출처: McGraw Hill, 2026d; McGraw Hill, 2026c)

축 점수 A 50.0 · B 61.5 · C 50.0 → 종합 55.2 | 고부담 게이트 조건부 — 배치·진급 결정에 쓰일 때만 적용 (일반 학습용 판정에는 미적용) | 도입 검토 의견 판정 유보

26 개 항목 중 자료 기준일까지 공개된 문서로 판정한 12 개(필수 13 개 중 5 개)만 점수에 반영했다. 실사용 진단이 필요한 14 개 항목(게이트 B3 포함)은 실사용 진단 결과를 이번 호에 싣지 않아 미측정으로 두고 점수에서 제외했다. 필수 항목 측정이 75%에 미치지 못해 종합 55.2 점과 관계없이 도입 검토 의견은 '판정 유보'이고, 축 A 점수는 A8 한 항목만으로 계산한 값이다. 미측정 14 개 항목을 모두 ●로 두면 종합 79.5 점, 모두 ○로 두면 26.7 점이며, 게이트 항목 B9 가 ◐여서 측정을 마쳐도 의견은 조건부 검토를 넘을 수 없다. 효과 근거(B9·B10)는 제공사 독립·고등교육 단독·제품 단위 연구로 판정했고 K-12·고등교육 혼합 메타분석은 참고로만 적었다. 공통 항목 기준 56.9 점은 두 제품을 모두 판정한 문서 판정 11 개 항목만으로 다시 계산한 비교용 점수다(축 A 미산출). 축 점수는 소수 첫째 자리까지 반올림하고 종합 점수는 반올림한 축 점수로 계산했다.

대상 솔루션 Gemini Study Notebooks (Google) **평가자** 연구진(문서 판정, 채점 초안 작성에 생성형 AI 보조) **접근 방식** 자료 기준일까지 공개된 제공사 문서·정책·연구로 판정했다(실사용 진단 결과는 이번 호에 싣지 않음).

ID	항목	측정	가중	판정	근거·출처
A. 사용성 (가중치 25%)					
A1	과업 완료 성과	AUTO	필수	미측정	실사용 진단이 필요한 항목이라 공개 문서만으로는 판정하지 않았다(실사용 진단 결과는 이번 호에 싣지 않음).
A2	효율성·반응성	AUTO	권장	미측정	실사용 진단이 필요한 항목이라 공개 문서만으로는 판정하지 않았다(실사용 진단 결과는 이번 호에 싣지 않음).

ID	항목	측정	가중	판정	근거·출처
A3	시스템 상태 가시성	HYBRID	권장	미측정	실사용 진단이 필요한 항목이라 공개 문서만으로는 판정하지 않았다(실사용 진단 결과는 이번 호에 싣지 않음).
A4	오류 예방·복구	HYBRID	필수	미측정	실사용 진단이 필요한 항목이라 공개 문서만으로는 판정하지 않았다(실사용 진단 결과는 이번 호에 싣지 않음).
A5	일관성·표준 준수	HYBRID	보조	미측정	실사용 진단이 필요한 항목이라 공개 문서만으로는 판정하지 않았다(실사용 진단 결과는 이번 호에 싣지 않음).
A6	접근성 — 자동 검출 영역	AUTO	필수	미측정	실사용 진단이 필요한 항목이라 공개 문서만으로는 판정하지 않았다(실사용 진단 결과는 이번 호에 싣지 않음).
A7	접근성 — 수동 검증 영역	HYBRID	필수	미측정	실사용 진단이 필요한 항목이라 공개 문서만으로는 판정하지 않았다(실사용 진단 결과는 이번 호에 싣지 않음).
A8	주관적 만족도(SUS 등)	DOC	권장	미측정	학습 노트북을 대상으로 한 SUS 등 공개 사용성·만족도 조사를 찾지 못했고 소규모 사용자 평정도 하지 않아 판정하지 않았다.
B. 교육적 특성 (가중치 45%)					
B1	학습목표 정렬	DOC	필수	◐	제공사 문서는 학생이 프롬프트를 입력하거나 수업 자료를 올리면 진단 퀴즈로 집중할 영역을 확인한다고 설명한다. 또 Gemini 가 학습 목표를 100 개 이상의 세부 학습 목표로 나눠 주제별로 진도를 추적하며 레슨별 연습 퀴즈를 올린 수업 자료로 만든다고 설명한다. 그러나 강좌 학습목표와의 정렬은 학생의 수업 자료 업로드에 의존하고 교사가 Classroom 에서 수업 자료 기반 학습 노트북을 배정하는 기능은 예정으로만 발표되었다. (출처: Kirtikar & Walport, 2026; Walport, 2026)
B2	콘텐츠 품질·정확성	HYBRID	필수	미측정	실사용 진단이 필요한 항목이라 공개 문서만으로는 판정하지 않았다(실사용 진단 결과는 이번 호에 싣지 않음).
B3	피드백·적응	HYBRID	필수*	미측정	(게이트 항목) 실사용 진단이 필요한 항목이라 공개 문서만으로는 판정하지 않았다(실사용 진단 결과는 이번 호에 싣지 않음).
B4	동기·참여	DOC	권장	◐	제공사는 진도 대시보드가 '강점'·'집중 영역'·'시작 전'을 표시하고 후속 레슨 퀴즈를 마칠 때마다 자동 갱신되며 레슨이 퀴즈 결과에 따라 바뀐다고 설명한다. 그러나 동기·참여 효과를 제공사와 독립적으로 확인한 연구는 없어 부분 충족으로 판정했다. (출처: Walport, 2026; Kirtikar & Walport, 2026)
B5	교수설계·인지부하	HYBRID	권장	미측정	실사용 진단이 필요한 항목이라 공개 문서만으로는 판정하지 않았다(실사용 진단 결과는 이번 호에 싣지 않음).

ID	항목	측정	가중	판정	근거·출처
B6	교수 통합·교수자 지원	DOC	권장	○	교사가 Google Classroom 에서 수업 자료에 근거한 학습 노트북을 배정하고 학급·학생별로 도움이 필요한 영역을 보는 기능은 출시 발표에서 예정으로만 제시되었다. 출시 때 학습 노트북은 18 세 이상 개인 계정에 먼저 제공되고 학교 계정에는 수 주 안에 제공될 예정이었으며, Gemini LTI 등록 항목은 Gemini 앱 ·NotebookLM 등을 LMS 에 넣는 도구로만 적혀 있다. 자료 기준일까지 공개된 문서로는 학습 노트북에서 쓸 수 있는 교사 자료·교실 통합 방안·교사용 차별화 지원이 확인되지 않아 미충족으로 판정했다(학습자별 개인화는 B3 항목에서 다룸). (출처: Kirtikar & Walport, 2026; 1EdTech Consortium, Retrieved 2026-b)
B7	상호작용 사용성·접근성(교육 맥락)	HYBRID	보조	미측정	실사용 진단이 필요한 항목이라 공개 문서만으로는 판정하지 않았다(실사용 진단 결과는 이번 호에 싣지 않음).
B8	표준 준수·상호운용	DOC	보조	◐	학습 노트북은 Gemini 앱 안의 기능이고, Gemini 앱·NotebookLM 등을 LMS 에 넣는 Gemini LTI 는 Canvas·Schoology 에 이어 2026 년 5 월 Moodle 로 연동이 넓어졌다(Moodle 은 LTI Dynamic Registration 규격으로 등록). 1EdTech 등록부에서 확인되는 Gemini LTI 의 인증은 LTI 1.3 도구 인증(2026. 3. 23.)이고, LTI Advantage 인증과 학습 노트북의 Gemini LTI 제공 여부는 확인되지 않아 부분 충족으로 판정했다. (출처: Kirtikar & Walport, 2026; 1EdTech Consortium, Retrieved 2026-b; Google, 2024; Google, 2026b)
B9	효과성 증거 등급	DOC	필수*	○	(게이트 항목) 제공사 독립·고등교육 단독·제품 단위 조건을 충족하는 연구는 없어(학습 노트북 대상 연구 없음) 독립 효과 근거 없음으로 판정했다. 고등교육에서 Gemini 를 쓴 독립 연구 3 편은 온라인 MBA 32 명 대 32 명 비무작위 비교, 프로그래밍 PBL 25 명 대 37 명 비무작위 비교, 예비교사 80 명의 ChatGPT 대비 실험이다. 이들은 연구용 알고리즘이나 Gemini 일반 기능을 다뤘거나 무처치 대조가 없어 제품 단위 근거로 세지 않았다. 참고: Google 연구팀이 수행한 시에라리온 중학교 무작위 통제 실험(가이드 학습, 동료심사 전)은 반영하지 않았다. (출처: Panwale & Vijayakumar, 2025; Omeh et al., 2025; Colaco & Antao, 2025; LearnLM Team & Fab AI, 2026)

ID	항목	측정	가중	판정	근거·출처
B10	효과크기 정량	DOC	권장	○	제공사 독립·고등교육 단독·제품 단위 조건을 충족하는 연구가 없어 학습 노트북의 효과크기는 보고되지 않았다. 고등교육에서 Gemini 를 쓴 독립 연구의 값으로는 온라인 MBA 전체 효과성 η² = 0.09(실험·통제 행에 같은 값으로 보고)와 프로그래밍 PBL 의 자기 보고 프로그래밍 기술 부분 η² = 0.338 이 있다. 두 값은 학습 노트북이 아닌 연구용 알고리즘·일반 기능의 결과라 반영하지 않았다. 참고: Google 연구팀의 시에라리온 중학교 수학 실험은 가이드 학습 사용 학급에서 의도처치 기준 +0.258 표준편차를 보고했으나 K-12·동료심사 전 결과다. (출처: Panwale & Vijayakumar, 2025; Omeh et al., 2025; LearnLM Team & Fab AI, 2026)
C. AI 신뢰성·윤리·거버넌스 (가중치 30%)					
C1	타당성·신뢰성과 환각 관리	HYBRID	필수	미측정	실사용 진단이 필요한 항목이라 공개 문서만으로는 판정하지 않았다(실사용 진단 결과는 이번 호에 싣지 않음).
C2	투명성·설명가능성	DOC	필수	◐	Google 도움말은 Gemini 가 부정확한 정보를 생성하고 환각으로 사실처럼 제공하거나 자신의 작동 방식을 허위로 설명할 수 있다고 고지하고, Gemini 환경 안에 면책 조항을 둔다고 밝힌다. 학습 노트북 사용 안내는 학습 목표를 100 개 이상의 세부 목표로 나눠 진도를 '강점'·'집중 영역'·'시작 전'으로 표시하고 후속 레슨의 퀴즈를 마칠 때마다 자동으로 갱신한다고 설명한다. 그러나 학습 노트북 전용 한계 고지와 진도 판정의 근거, 응답 출처를 제시하는 방식은 확인되지 않아 부분 충족으로 판정했다. (출처: Google, Retrieved 2026-b; Walport, 2026)
C3	공정성·편향 관리	HYBRID	필수	미측정	실사용 진단이 필요한 항목이라 공개 문서만으로는 판정하지 않았다(실사용 진단 결과는 이번 호에 싣지 않음).

ID	항목	측정	가중	판정	근거·출처
C4	프라이버시·데이터 거버넌스	DOC	필수	◐	출시 때 학습 노트북은 18 세 이상 개인 계정에 먼저 제공되었다(학교 계정은 제공 예정). 한국 거주자 대상 추가 정보는 수집한 개인정보를 국외 Google 데이터 센터로 옮겨 보관하고 Gemini 앱 같은 AI 모델과의 상호작용을 모델 개발·학습·개선에 쓴다고 밝힌다. 학교 계정용 Gemini for Education 은 교육 사용자의 대화를 AI 모델 개선에 쓰지 않는다고 밝히고, 데이터 처리 부속서는 고객 지시 범위의 처리와 삭제 데이터의 최대 180 일 내 파기(법적 보관 예외)를 정한다. 그러나 학교 계정의 고객 데이터는 Google·하위 처리자 시설이 있는 어느 나라에서든 처리될 수 있고 한국어 개인정보처리방침은 서비스 데이터만 다루며, 학습 노트북 전용 개인정보 조항은 확인되지 않았다. (출처: Kirtikar & Walport, 2026; Google, 2026c; Kirtikar, 2025; Google, 2026d; Google, 2026a)
C5	인간 통제·교수자 역할	DOC	필수	○	출시 발표에서 교사가 Classroom 에서 학습 노트북을 배정하고 도움이 필요한 영역을 보는 기능과 학교 계정 제공은 예정으로만 제시되었고, 학습 노트북은 18 세 이상 개인 계정에 먼저 제공되었다. 이에 따라 자료 기준일까지 공개된 문서로 확인되는 교수자·기관의 개입 경로가 없어 미충족으로 판정했다. 학습 노트북 문서에 평가·배치 같은 고부담 결정 기능은 없어 고부담 게이트는 적용하지 않았다. (출처: Kirtikar & Walport, 2026; Walport, 2026)
C6	안전·견고성·보안	HYBRID	권장	미측정	실사용 진단이 필요한 항목이라 공개 문서만으로는 판정하지 않았다(실사용 진단 결과는 이번 호에 싣지 않음).
C7	책무성·거버넌스·정책 정합	DOC	권장	◐	Google 은 개인정보 보호법·정보통신망법상 국내대리인으로 구글코리아 유한회사를 고지하고, 학교 계정용 데이터 처리 부속서는 고객 데이터를 고객 지시 범위에서만 처리하도록 정한다. 학교 계정용 Gemini 앱에는 관리자 접근 관리·Vault 대화 검색·사용 보고서가 있지만 학습 노트북의 학교 계정 제공은 출시 때 예정으로만 발표되었다. 또한 AI 기본법 제 31 조의 생성형 AI 사전 고지·결과물 표시에 맞춘 대응은 확인되지 않았다(도움말은 Gemini 환경 안에 면책 조항을 둔다고 밝힘). (출처: Google, 2026c; Google, 2026d; Kirtikar, 2025; Kirtikar & Walport, 2026; 인공지능 발전과 신뢰 기반 조성 등에 관한 기본법, 2026; Google, Retrieved 2026-b)

ID	항목	측정	가중	판정	근거·출처
C8	AI 리터러시·연령 적합성	DOC	보조	◐	학습 노트북은 18 세 이상 개인 계정에 먼저 출시되었고 학교 계정(18 세 미만 포함)에는 수 주 안에 제공될 예정이었으며, 한국에서 개인 Google 계정을 직접 관리할 수 있는 최소 연령은 14 세다. Gemini 는 18 세 미만에게 더 엄격한 콘텐츠 정책과 AI 리터러시 자료를 담은 청소년 온보딩을 제공하고, Gemini LTI 로 안전하고 생산적인 AI 활용 교육 자료를 쓸 수 있다고 밝힌다. 그러나 대학생(성인) 학습자를 위한 학습 노트북의 AI 리터러시 지원은 확인되지 않았다. (출처: Kirtikar & Walport, 2026; Google, Retrieved 2026-a; Kirtikar, 2025; Google, 2026b)

축 점수 A - · B 23.1 · C 37.5 → 종합 28.9 | 고부담 게이트 조건부 — 배치·진급 결정에 쓰일 때만 적용(일반 학습용 판정에는 미적용) | 도입 검토 의견 판정 유보

26 개 항목 중 자료 기준일까지 공개된 문서로 판정한 11 개(필수 13 개 중 5 개)만 점수에 반영했다. 실사용 진단이 필요한 14 개 항목(게이트 B3 포함)은 실사용 진단 결과를 이번 호에 싣지 않아 미측정으로 두었고, 공개 사용성·만족도 조사가 없는 A8 도 미측정이어서 축 A 점수는 산출하지 않았다. 필수 항목 측정이 75%에 미치지 못해 종합 28.9 점과 관계없이 도입 검토 의견은 '판정 유보'다. 미측정 15 개 항목을 모두 ●로 두면 종합 68.3 점, 모두 ○로 두면 12.9 점이며, 게이트 항목 B9 가 ○여서 측정을 마쳐도 의견은 '신중 검토'를 넘을 수 없다. 효과 근거(B9·B10)는 ALEKS 와 같은 기준(제공사 독립·고등교육 단독·제품 단위 연구)으로 판정했다. 학습 노트북은 출시일(2026. 6. 25.)에 18 세 이상 개인 계정에 먼저 제공되고 학교 계정 제공은 예정이었으므로, 교수자·기관 관련 항목(B6·C5)은 기준일 전에 공개된 출시 발표의 제공 범위로 판정했다. 공통 항목 기준 점수는 두 제품을 모두 판정한 11 개 항목으로 계산하며, 학습 노트북은 판정한 항목이 이 11 개와 같아 종합 점수와 같다(축 A 미산출). 축 점수는 소수 첫째 자리까지 반올림하고 종합 점수는 반올림한 축 점수로 계산했다.

주: ●(충족)·◐(부분충족)·○(미충족)·미측정으로 판정하고 미측정은 점수에서 제외했다. 측정 열의 AUTO(도구 실측)·HYBRID(실측 뒤 검토자 확인)는 실사용 진단으로 판정하는 항목으로 이번 호에는 실사용 진단 결과를 싣지 않아 모두 미측정이며 DOC 는 두 제품 모두 문서로만 판정한 항목이다. 가중 열의 별표는 게이트 항목(B3·B9)이다. B9 는 Ⅰ장 4 절의 증거 등급에 따라 강함·보통이면 ●, 유망(통계적 통제를 둔 상관 연구 수준)이면 ◐, 근거 없음이면 ○로 판정했다. 고부담 게이트는 배치·진급 결정에 쓰일 때만 적용하며 이때는 C1·C3·C4·C5 가 모두 ●여야 한다. 도입 검토 의견은 적극 검토(게이트 B3·B9 모두 ●, 필수 항목에 ○ 없음, 종합 80 점 이상), 조건부 검토(게이트 모두 ◐ 이상, 종합 60 점 이상), 그 밖은 신중 검토로 정하며 필수 항목 13 개 가운데 측정한 항목이 10 개 미만이면 판정 유보로 적는다.

부록 3. 용어 해설

용어	설명
적응형 학습 시스템	학습자의 응답 데이터로 지식 상태를 추정해 학습 경로·난이도·콘텐츠 순서를 자동으로 조정하는 엔진. 이 보고서의 분석 단위다
지능형 튜터링 시스템(ITS)	학습자의 응답을 분석해 개별 지도·힌트·피드백을 주는 컴퓨터 학습 시스템으로 대학 ITS 메타분석의 대상이다(Ⅱ장 3 절 나)
지식 추적(BKT·DKT)	학습자가 각 기술을 익혔을 확률을 시간에 따라 추정하는 방법. BKT 는 베이지안 확률 모형을, DKT 는 순환신경망을 쓴다
AUC	다음 응답의 정오를 맞히는 예측 성능 지표(곡선 아래 면적)로 학습 효과를 잰 값이 아니다(Ⅱ장 1 절 가·라)
문항반응이론(IRT)	문항의 난이도·변별도 모수로 학습자 능력을 추정하는 측정 이론. 응답에 따라 다음 문항을 바꾸는 컴퓨터 적응형 검사의 기반이다
지식공간이론	영역 지식을 선수 관계로 묶인 지식 상태의 집합으로 표현하는 이론. ALEKS 의 기반이다

용어	설명
지식 요소(KC)와 콘텐츠 태깅	지식 요소는 문항 풀이에 필요한 개념·기술 단위이고 태깅은 문항마다 지식 요소를 연결해 엔진이 응답을 해석하게 하는 작업이다(Ⅱ장 1절 다, Ⅱ장 4절 가)
숙달 학습	정해진 숙달 기준에 이르러야 다음 내용으로 넘어가게 하는 학습 설계로 숙달 요건이 높으면 학습 부담이 커질 수 있다(Ⅱ장 3절 다·바)
효과크기(g·d·η^2)	집단 간 평균 차이를 표준편차 단위로 나타낸 값(Hedges g·Cohen's d)이나 처치가 설명하는 변동의 비율(η^2)로 연구 간 효과 비교에 쓴다(Ⅱ장 3절)
배치 평가와 기준점(cut score)	입학생의 수강 수준을 정하는 평가와 그 판정 경계 점수. 과목 배치를 좌우하므로 고부담 용도에 해당한다
LTI와 1EdTech 인증	학습 도구 상호운용성(LTI)은 LMS와 외부 학습 도구를 잇는 표준이고 이를 관리하는 1EdTech는 LTI·데이터 프라이버시 인증의 유효 여부를 등록부에 공개한다(Ⅱ장 1절 마)
고위험 AI(EU)·고영향 AI(국내)	기본권에 중대한 영향을 줄 수 있어 추가 의무가 붙는 AI 분류. 교육 분야의 범위는 두 법제가 다르다(Ⅱ장 4절 사, Ⅳ장 1절 가)

부록 4. 기관 도입 점검표

본문의 점검 항목을 단계별로 모았으며 담당은 예시다.

단계	점검 항목	담당(예시)	관련 절
계약 전	독립 효과 연구가 있고 대상·설계가 학내 강좌와 비슷한가	교무·교수학습개발센터	Ⅱ장 3절 라, Ⅱ장 4절 바
계약 전	학습 이력 보관 기간, 비식별 정보 2차 이용, 국외 처리 조건을 계약 문서로 확인하고 학내 가명처리 체계를 갖췄는가	개인정보 보호·정보화	Ⅱ장 4절 마, Ⅳ장 1절 나, Ⅳ장 2절 나
계약 전	콘텐츠 범위·한국어 지원과 LTI·프라이버시 인증의 유효성을 확인했는가	교수학습개발센터·정보화	Ⅱ장 1절 마, Ⅱ장 4절 가, Ⅳ장 1절 라
계약 전	인수·합병·서비스 종료 때의 데이터 이전·삭제와 사업 연속성 조항을 계약에 넣었는가	정보화·개인정보 보호	Ⅱ장 1절 마, Ⅱ장 2절 다
도입 설정	숙달 요건·성적 반영 비율·학습 시간 기대치를 강좌 목표에 맞췄는가	교수자·교수학습개발센터	Ⅱ장 3절 다·바, Ⅱ장 4절 사, Ⅲ장 1절
도입 설정	전면 도입 전에 소규모 시범 운영(대상 강좌·기간·판단 기준)으로 숙달 요건·학습 부담을 조정하는가	교무·교수학습개발센터	Ⅱ장 3절 다, Ⅳ장 1절 라, Ⅳ장 2절 나
도입 설정	교수자가 강좌 목표에 맞춰 숙달 요건·진도·사용 범위를 정할 수업 자율성을 보장했는가	교무	Ⅳ장 1절 라
도입 설정	교수자·관리자 통제 범위를 확인하고 개인 계정의 데이터 조건(국외 보관·모델 개선 활용)을 안내했는가	교수학습개발센터·개인정보 보호	Ⅲ장 1·3절
수업 운영	보완 학습 부담을 점검하고 조교·멘토 지원을 두었는가	교수자·학생지원	Ⅱ장 3절 바, Ⅳ장 2절 나
수업 운영	학습자의 과제 선택 여지를 두었는가	교수자	Ⅱ장 4절 나

단계	점검 항목	담당(예시)	관련 절
수업 운영	일반 학습 용도에서도 집단별 진도와 보충 배정 쏠림을 점검하는가	교수자·교수학습개발센터	Ⅱ장 4절 나·다
수업 운영	범용 LLM 진단 퀴즈를 학습 보조로 한정하고 출처·서지 검증을 안내했는가	교수자·교수학습개발센터	Ⅲ장 2절 나, Ⅳ장 2절 나
고부담 용도	배치·진급 기준점과 성적 비중이 큰 강좌의 판정을 기관 자료로 검증하고 집단별 결과 점검과 사람의 검토·이의 경로를 두었는가	입학·교무	Ⅱ장 4절 다·라·사, Ⅳ장 1절 가, Ⅳ장 2절 나
점검·평가	학내 효과 측정(무작위 배정, 어려우면 비교 집단·지연 측정)을 설계하고 평가 주체의 독립성과 분석 계획 사전 공개, 개인 식별 없는 이용률 점검, 비용 부담 주체와 지원 종료 뒤 재원을 정했는가	교무·AI 에듀테크센터	Ⅳ장 1절 다·라, Ⅳ장 2절 나

PUBLICATION INFORMATION

발간 정보

항목	내용
보고서명	월간 에듀테크 솔루션 분석 보고서 Vol.02 적응형 학습 시스템
대상 월	2026 년 7 월분 (전 8 호 중 제 2 호)
발간일	2026. 7. 31.
자료 기준일	2026. 7. 31. — 제품의 기능·가격·정책 정보는 이 날짜 현재 기준임
발행처	한림대학교 AI 에듀테크 연구과제
과제	글로컬대학 30 한림대학교 AI 에듀테크센터 연구과제「고등교육 AI 에듀테크의 통합적 발전 방안 연구 — 전공 운영 모델·연구 방향성 정립·기초교양 AI 리터러시 교육」[연구 2] 월간 에듀테크 솔루션 분석
연구책임자	김슬기 (한림대학교 일송자유교양대학)
집필진	김슬기(연구책임자)·고영웅·김진환·김병정·김성우·김용수·서동기·육청민·이정근·정혜선·진서연·한민정·한수미·홍수현·황현석 (이상 한림대학교)
문의	skkim@hallym.ac.kr

인용 표기

김슬기·고영웅·김진환·김병정·김성우·김용수·서동기·육청민·이정근·정혜선·진서연·한민정·한수미·홍수현·황현석(2026). 월간 에듀테크 솔루션 분석 보고서 Vol.02: 적응형 학습 시스템. 한림대학교 AI 에듀테크 연구과제.

이용 안내

- ○ 이 보고서의 저작권은 한림대학교에 있습니다. 출처를 밝히면 비영리 목적으로 인용·활용할 수 있습니다.
- ○ 이 보고서에서 다룬 제품·서비스는 분석 목적으로 선정한 것이며 특정 제품의 홍보나 보증을 뜻하지 않습니다.
- ○ 제품의 기능·가격·정책은 자료 기준일(2026. 7. 31.) 이후 바뀔 수 있습니다.
- ○ 이 보고서의 내용은 연구진의 연구 결과이며 한림대학교 또는 AI 에듀테크센터의 공식 견해와 다를 수 있습니다.
- ○ 이 보고서는 HAI-GPT 를 운영하는 한림대학교 AI 에듀테크센터의 연구과제 산출물로, 한림대 AI-HiED 와 고등교육 에듀테크 소프트랩 AI 코스웨어, 풀리캠퍼스는 평가하지 않고 사례로만 다루었습니다. 이 센터는 2025 년 12 월 풀리캠퍼스 제공사인 프리윌린과 AI-HiED 협약을 맺었고 한림대학교는 고등교육 에듀테크 소프트랩 사업의 운영대학이며, 연구진에는 해당 부서·센터 구성원이 포함되고 일부는 소프트랩 실증 교과를 맡았습니다. 채점하지 않은 제품에 결함 근거가 드러나지 않은 것을 상대적 우위로 해석해서는 안 됩니다. 소프트랩 실증 교과로 추정되는 교양 수학 파일럿 연구 1 편은 관계를 밝혀 II장 3 절에서 인용했습니다.

사사

본 과제(결과물)는 2026 년도 교육부 및 강원특별자치도의 재원으로 강원 RISE 센터의 지원을 받아 수행된 지역혁신중심 대학지원체계(RISE) 글로컬대학 30 의 결과입니다.(2026-RISE-10-009)

This research was supported by the Regional Innovation System & Education(RISE) Glocal University 30 Project program through the Gangwon RISE Center, funded by the Ministry of Education(MOE) and the Gangwon State(G.S.), Republic of Korea.(2026-RISE-10-009)

MONTHLY EDTECH SOLUTION ANALYSIS REPORT

월간 에듀테크 솔루션 분석 보고서 시리즈

호	대상 월	주제
Vol.01	2026 년 6 월분	LLM 기반 학습 도구
Vol.02	2026 년 7 월분	**적응형 학습 시스템**
Vol.03	2026 년 8 월분	AI 튜터링 플랫폼
Vol.04	2026 년 9 월분	학습분석 도구
Vol.05	2026 년 10 월분	AI 평가·피드백 도구
Vol.06	2026 년 11 월분	인간-AI 공동조절 학습 지원 도구
Vol.07	2026 년 12 월분	교수자용 생성형 AI 도구
Vol.08	2027 년 1 월분	AI 거버넌스·학업 진실성 솔루션
종합	2027 년 1 월	종합 동향 리포트

이번 호 소개

「월간 에듀테크 솔루션 분석 보고서」 Vol.02(2026 년 7 월분)는 적응형 학습 시스템을 기술 구조·시장·교육적 적용·한계의 네 측면에서 분석했다. 전용 엔진은 미리 만든 도메인 모델로 숙달을 추정하고 Gemini 학습 노트북은 목표를 세부 목표로 나눠 진단 퀴즈로 출발점을 잡는다(산출 기준 미공개). 고등교육의 효과는 K-12 보다 작게 보고되었고 제품 단위 독립 근거는 부족했다. ALEKS 와 학습 노트북은 공개 문서로 판정한 필수 항목이 적어 모두 판정 유보였다. 대학은 학습 지원과 배치·성적 반영을 구분하고 시범 운영과 효과 측정을 거쳐 넓힐 수 있다.

키워드 적응형 학습 · 지식 추적 · 생성형 개인화 · 고등교육 · AI 디지털교과서

한림대학교 AI 에듀테크 연구과제

글로컬대학 30 한림대학교 AI 에듀테크센터 연구과제「고등교육 AI 에듀테크의 통합적 발전 방안 연구」[연구 2] 산출물