Creative Education Studies 创新教育研究, 2026, 14(11), 672-686
Published Online November 2026 in Hans. https://www.hanspub.org/journal/ces
https://doi.org/10.12677/ces.2026.1411933

Hans汉斯

生成式人工智能赋能高等数学教学改革的路径演进与反思

郑成勇

五邑大学数学与计算科学学院，广东 江门

收稿日期：2026年10月16日；录用日期：2026年11月11日；发布日期：2026年11月26日

摘　要

目的：总结生成式人工智能技术在高等数学教学中的实践路径，厘清通用生成式人工智能工具与学科专属教学功能之间的能力边界。方法：采用行动研究路径，按“计划－行动－观察－反思”的循环推进：依托广东省教育科学规划课题(2024GXJK275)开展两年多的三阶段教学实践，各阶段同时采集量化运行数据与质性反馈，分析结论直接决定下一阶段的行动。第一阶段尝试基于无代码平台自主开发微信小程序；第二阶段转向“现有教学平台＋成熟生成式人工智能工具”的轻量化融合模式；第三阶段在课题经费支持下自主研发本地大模型驱动的“数学学习平台”。结果：三阶段实践表明：轻量化路径与自主研发并非替代关系，而是“工具应用保运转、自主研发保内核”的分工格局；平台的核心创新在于以同一模型、同一知识库配合不同的提示词设计，实现标准模式与苏格拉底模式两种问答模式的对照运行，将“AI作为答案提供者”的工具属性转化为“AI作为思维引导者”的教学属性；平台以三维知识图谱组织学科知识结构，将知识点与配套教学视频、结构化问答关联起来；学科内容与系统逻辑解耦的架构设计使平台由高等数学向线性代数的低成本迁移得到工程验证。结论：对资源有限的高校，“使用成熟工具”仍是最稳妥的起步路径，但引导式教学、学科结构化题库与学习过程数据沉淀三项超出通用工具的功能边界，需自建轻量内核。需要如实说明：平台当前数据为系统联调阶段的功能验证数据，证明的是数据链路已跑通与工程实现完整，教学成效尚待规范对照实验检验。

关键词

生成式人工智能，高等数学，教学改革，引导式教学，知识图谱，本地大模型，路径选择

Evolution and Reflection on the Pathways of Generative Artificial Intelligence-Empowered Reform in Higher Mathematics Teaching

Chengyong Zheng

School of Mathematics and Computational Science, Wuyi University, Jiangmen Guangdong

文章引用: 郑成勇. 生成式人工智能赋能高等数学教学改革的路径演进与反思[J]. 创新教育研究, 2026, 14(11): 672-686. DOI: 10.12677/ces.2026.1411933

Received: October 16, 2026; accepted: November 11, 2026; published: November 26, 2026

Abstract

Purpose: This study summarizes the practical path of applying generative artificial intelligence (GenAI) in advanced mathematics teaching and clarifies the capability boundary between general-purpose GenAI tools and discipline-specific pedagogical functions. Methods: Following an action-research approach organized as successive plan-act-observe-reflect cycles, a three-stage teaching practice lasting more than two years was carried out under a provincial education science planning project (No. 2024GXJK275). Quantitative operational data and qualitative feedback were collected at each stage, and the conclusions drawn from their analysis determined the actions of the following stage: in Stage 1, a WeChat mini-program was developed independently on a no-code platform; in Stage 2, the focus shifted to a lightweight integration model combining an existing learning management platform with mature GenAI tools; in Stage 3, a locally deployed large-model-driven mathematics learning platform was developed in-house. Results: The three stages show that the lightweight path and in-house development are complementary rather than substitutable, forming a layered pattern in which "tool adoption keeps routine teaching running while in-house development secures the pedagogical core". The platform's key innovation lies in running two answering modes, a standard mode and a Socratic mode, on the same model and the same knowledge base with different prompt designs, thereby transforming AI from an answer provider into a thinking guide. A three-dimensional knowledge graph presents the disciplinary knowledge structure and links each knowledge point to its teaching video and structured questions. A discipline-agnostic architecture, with content decoupled from system logic, was validated by the low-cost migration from advanced mathematics to linear algebra. Conclusion: For resource-constrained universities, adopting mature tools remains the safest starting point; however, guided instruction, structured discipline-specific question banks, and the accumulation of learning-process data exceed the capability boundary of general-purpose tools and call for a lightweight in-house core. It should be noted that the current platform data are functional verification data from the system integration-testing phase, demonstrating that the data pipeline works and the engineering implementation is complete; pedagogical effectiveness still awaits rigorous controlled experiments.

Keywords

Generative Artificial Intelligence, Advanced Mathematics, Teaching Reform, Guided Instruction, Knowledge Graph, Local Large Language Model, Path Selection

Copyright © 2026 by author(s) and Hans Publishers Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/
Open Access

1. 引言

1.1. 研究背景与问题提出

近年来，教育数字化转型成为高校教学改革的重要背景；2023 年以来，生成式人工智能(generative artificial intelligence, GenAI)技术的迅速普及，进一步改变了教学工具的可获得性与使用方式。高等数学是理工科专业的核心基础课程，具有学时长、受众广、抽象性强等特点，普遍采用大班授课模式，每学期单班 100~150 人，教学组织难度大。传统教学存在四类突出痛点：一是教学模式单一，大班授课难以

实现个性化教学；二是学生兴趣不足，抽象概念理解困难，学习积极性不高；三是教师负担沉重，答疑与作业批改占用大量时间；四是学用脱节，数学知识与应用场景分离，学生难以体会学科价值。

与此同时，通用生成式人工智能工具在高校迅速普及，为破解上述痛点提供了新的可能。但一线教师在拥抱技术的同时普遍面临一个实践困惑：通用生成式人工智能工具能替代多少教学工作？哪些教学功能必须自主研发？已有的相关讨论多停留于理念倡导或工具推介，缺乏基于完整实践过程的经验证据。本文基于一项持续两年多的省级教改课题实践，尝试回答这一问题。

1.2. 文献综述与研究现状

国内外学者围绕生成式人工智能在教育领域的应用已展开大量研究，总体上呈三个方向。其一，生成式人工智能赋能课程教学的混合式改革探索，涵盖高等数学、线性代数等数学类课程的课堂教学创新与混合式教学模式重构[1]-[9]。其二，生成式人工智能作为学习支持工具的价值与挑战分析，涉及生成式人工智能对学习、能力与评价的赋能机制[10]，以及对师生角色转变的影响[11]-[13]。其三，面向学科场景的 AI 教辅系统研发[1] [2] [5] [7]，如基于生成式人工智能的高等数学教辅系统开发、人工智能融入实践混合式教学等。

综观现有研究，仍存在两点不足：一是多数成果停留在理论探讨或系统开发层面，对“使用而非开发”这一更贴近一线教师实际的视角关注不足，缺乏完整的失败－调整－再选择的过程性证据；二是面向学科专属的引导式教学与学习过程数据采集研究较少，通用生成式人工智能工具在此类场景中的功能边界尚少被明确刻画。本文的实践在三阶段演进中同时覆盖了工具应用与自主研发两类路径，并对二者的能力边界进行了对照分析，可为上述两个缺口提供经验补充。

1.3. 研究意义

本研究的价值体现在三个方面。实践价值上，为高校教师提供一条“生成式人工智能工具 + 现有教学平台”的低门槛实施路径，降低改革启动成本；路径价值上，通过三阶段完整演进厘清“通用工具能做什么、学科专属功能必须自己做什么”的边界，为同类院校的路径选择提供决策参照；方法价值上，平台以同一模型、同一知识库实现两种问答模式的对照运行，使“引导是否比直接给答案更有效”从一句主张变成一个可用数据检验的问题。

2. 研究设计与实践历程

2.1. 研究方法与分析框架

本研究采用行动研究路径。行动研究的基本循环是“计划－行动－观察－反思”：针对真实教学情境中的问题制定行动计划，实施后系统收集证据，依据证据反思并据此修正下一阶段的计划。本课题的实践依托广东省教育科学规划课题(高等教育专项)“AIGC 赋能的数智化大学数学教学改革与实践”(2024GXJK275)开展，周期为 2024 年 10 月至 2026 年 10 月，覆盖高等数学、线性代数等课程的教学全过程。两年多的实践恰好构成三个连续的循环，每一循环的“反思”环节都直接生成了下一循环的“计划”，三个阶段因此不是三次并列的尝试，而是一条由证据驱动的决策链。

为使这条决策链可核查，本研究对每个循环作三点约定：其一，观察环节同时收集量化证据与质性证据；其二，分析环节不止于描述结果，还须判断结果与计划的偏离程度及其原因；其三，反思环节以书面形式明确“继续、调整或放弃”的结论，并将其作为下一循环的行动前提。

2.2. 数据收集与数据分析

本研究的证据包括三类。量化数据来自各阶段系统的自动记录：第一阶段的访问人次、日均活跃用

户、功能使用占比与公式识别准确率；第二阶段学习通后台的签到、作业提交、测验成绩等教学过程数据，以及智能批阅工具的处理记录；第三阶段统计看板的总问答数、活跃学生数、习题成功率与高频知识点 TOP15。此类指标由系统自动生成，可追溯、可复核。

质性数据包括教学观察记录、师生反馈与平台对话实录。教学观察聚焦学生在课堂与课后的实际使用行为；师生反馈以日常交流和使用中反映的问题为主，不作结构化问卷处理；第三阶段的对话实录按模式标记入库，可逐轮比对同一问题在两种模式下的应答差异。过程性凭证包括题库与功能建设的过程记录、系统本体与在线运行验证记录、模块运行截图等。

数据分析遵循三条路径：一是描述性统计，用于刻画使用规模与功能结构，如功能使用占比可直接反映学生的真实需求分布；二是分类归纳，将质性反馈按“功能覆盖－技术可行性－使用意愿”三个维度归类，与量化结果相互印证；三是阶段间纵向对照，检验上一阶段计划所依据的假设是否成立。本研究的分析不追求统计推断，其功能是回答行动研究所特有的问题——下一步该做什么，以及为什么。

需要事先说明数据性质：平台统计看板当前的问答日志、习题执行记录与知识点频次，系统联调阶段的功能验证数据，样本量较小。因此本文主张的是“采集－存储－可视化－按模式对照”这条数据链路已跑通、系统工程实现完整且核心创新已落地，而非教学成效已获证实。

2.3. 三阶段路径演进总览

课题实践经历了三个特征分明的阶段，如表 1 所示。三阶段的关系是一个边界逐步清晰的过程：第一阶段暴露了全面自主开发在时间投入与推广难度上的不可行；第二阶段证明轻量化路径足以解决教学主流程数字化问题，但触及其能力边界；第三阶段针对该边界做定向的轻量内核自建。各阶段的计划、观察证据与反思结论汇总于表 2，2.4~2.6 节依次展开。

Table 1. Comparison of the characteristics of the three stages of practice
表 1. 三阶段实践的特征对比

维度	第一阶段(2024.10~2025.03)	第二阶段(2025.04 至今)	第三阶段(2026.06~2026.10)
路径取向	自主开发	工具应用	自主研发(内核定向)
载体	“高数自修室”微信小程序	学习通 + 成熟生成式人工智能工具	“数学学习平台”
投入规模	6 个月开发 + 持续维护	1~2 周工具筛选 + 培训	课题经费支持，5 个月建成
核心目标	教学补充工具	教学主流程数字化	引导式教学与数据沉淀
主要结果	功能重叠、体验不佳，使用率低	全流程数字化，批改负担大降	五模块闭环，双模式对照运行

Table 2. The plan-act-observe-reflect process and stage decisions of the three stages
表 2. 三阶段“计划－行动－观察－反思”过程与阶段决策

阶段	计划与行动	观察证据(量化/质性)	反思与决策
第一阶段 (2024.10~2025.03)	自研课后自修工具；无代码平台开发微信小程序，集成四类功能，历时 6 个月	量化：访问约 500 人次，日均活跃 20~30 人，功能占比 45%/35%，公式识别 70%~80%。质性：公式显示与语音播报体验不佳，学生倾向“一站式”平台	判断：全面开发成本不成立，功能与学习通重叠。决策：停止扩张，转向轻量化融合

续表

第二阶段 (2025.04 至今)	以已采购平台承载主流程，成熟工具单点增强；六方面功能落地，引入智能批阅	量化：批改由数小时降至分钟级，过程数据全程留痕。质性：点对点提问增加；复杂推理题仍需人工复核	判断：主流程已数字化，引导式推导与思维数据仍缺。决策：定向自建教学内核
第三阶段 (2026.06~2026.10)	只做通用工具做不到的三件事；5 个月建成五模块平台，题库扩至 8 章 110 问并接入线性代数	量化：282 个结构化问题、52 道在线习题、50~60 个核心知识点。质性：双模式对话可逐轮比对；迁移仅需更换题库	判断：数据链路跑通、工程完整，成效待验证。决策：开展双模式对照实验

2.4. 第一阶段：全面自主开发的受挫与边界初显

第一阶段(2024 年 10 月~2025 年 3 月)的行动计划是开发一个覆盖高等数学课后自修全场景的教学补充工具。项目初期对生成式人工智能技术抱有较高预期，技术选型采用“扣子(Coze)无代码平台 + 大语言模型”的轻量化开发方案，历时 6 个月开发“高数自修室”微信小程序，集成知识问答、习题解答、出题练习、教案生成四类功能，上线后持续维护。具体成果见我们的论文[1]。

行动结果可由三类证据刻画。量化方面，系统累计访问约 500 人次，日均活跃用户仅 20~30 人，功能使用集中于知识问答(45%)与习题解答(35%)；数学公式识别准确率仅 70%~80%，复杂题目误判率高。质性方面，师生反馈集中于三点：微信端公式显示不友好，虚拟人语音播报数学公式存在难以突破的技术瓶颈，学生课后需要在多个入口之间切换。系统对照方面，将小程序与学校已采购的泛雅学习通逐项比对(表 3)可以看到，两者在课程管理、作业提交与批阅、测验与成绩统计、考勤与学情数据等主流程环节高度重叠，小程序仅在“大模型知识问答”一项上具有增量，而这一增量不足以支撑一个独立入口的存在。

Table 3. Functional comparison between the self-developed mini-App and Chaoxing platform
表 3. 第一阶段自研小程序与泛雅学习通的功能对比

对比维度	“高数自修室”小程序(第一阶段自研)	泛雅学习通(学校已采购)
课程与班级管理	不具备	具备
作业发布、提交与批阅留痕	不具备	具备(拍照上传、批阅留痕)
章节测验与成绩统计	仅离线“出题练习”，无班级成绩管理	具备(即时评分、成绩汇总)
考勤与学情数据	不具备	具备(签到、任务完成度)
大模型知识问答	具备(公式识别准确率 70%~80%，微信端公式显示不佳)	本课题使用范围内未用于实时问答
学科结构化题库	未建成	需教师自行建题，无伴随问题结构
学生入口与使用习惯	需新注册独立小程序	已采购、已形成使用习惯

注：表中判断基于本课题 2024~2025 年的实际使用范围；平台功能随时间更新，此处的比较仅就当时版本而言。

据此，第一阶段末形成的反思是：生成式人工智能赋能教学的核心价值不在于“开发新平台”，而在于“使用成熟工具提升现有教学效果”；对一线教师而言，时间是最稀缺的资源，全面自主开发的投入产出比不成立。相应决策是停止小程序的功能扩张，将资源转向“现有平台 + 成熟工具”的轻量化融合，即第二阶段的计划。需要说明的是，这一反思在第三阶段被修正而非推翻——受挫的实质是“全面

开发”不成立，而非“任何自建”都不成立，二者的边界条件见 2.6 节。

2.5. 第二阶段：轻量化融合模式的常态化运行

第二阶段的计划由此确定：以学校已采购的平台承载教学主流程，以成熟工具做单点增强。经过两年多实践，学习通已成为高等数学教学的核心支撑平台，实现六方面常态化功能。

其一，教学资料数字化分享，课件、教学大纲、课外学习资料统一上传，经典练习题与往年试卷集中管理，学生可随时随地获取学习资源；其二，作业发布与批改流程重构，作业直接在学习通发布、学生拍照上传，取消纸质作业本，采用“手工批改 + 学生互批(每人随机批改 2 份他人作业)”的混合批改方式，作业痕迹可追溯、批改反馈即时；其三，课堂签到与考勤管理自动化，考勤数据自动统计，为平时成绩评定提供依据；其四，章节测验即时反馈，每章结束后立即进行选择题、判断题测验，系统即时评分，教师即时掌握全班知识掌握情况；其五，平时成绩数据化管理，签到、作业、测验、互动等数据全程记录，平时成绩有据可查，避免人为因素干扰；其六，课后答疑与师生交流，学生通过点对点私下聊天功能大量提问，弥补课堂时间不足。

2025 年 10 月起，项目引入“数你最灵”(https://smartestu.cn)智能批阅系统，实现手写作业自动批改。该系统可识别学生拍照上传的手写作业，自动判断答案正误并给出批改结果，支持选择题、填空题、计算题等多种题型。使用效果体现在三方面：批改负担大幅降低，原本需要数小时的作业批改工作可在分钟级完成；批改一致性提升，标准统一，避免人工批改的主观性；反馈及时性增强，学生提交后即可获得反馈，学习闭环更快。其局限性同样明显：复杂推理题仍需人工复核；手写潦草时识别准确率下降；目前定位为“辅助工具”，未完全替代人工批改。

第二阶段的观察结果与第一阶段形成对照：教学主流程的数字化目标基本达成，签到、作业、测验、互动数据全程留痕，作业批改由数小时压缩到分钟级，学生对点对点提问的使用明显增加；同时暴露两类局限——复杂推理题仍需人工复核、手写潦草时识别率下降，工具定位仍为辅助。更关键的是，主流程数字化并未触及教学的核心环节：学生获得的仍是“对/错”与标准答案，而非推导过程的引导；系统沉淀的仍是“结果型”数据，而非“思维型”数据。这一判断构成第三阶段的行动前提，2.6 节对此展开。

2.6. 第三阶段的问题来源：通用工具的能力边界

第二阶段解决了教学主流程的数字化问题，但实践中逐渐暴露出两类无法用通用工具解决的诉求。其一，学生对“被直接告知答案”普遍不满，希望获得引导式的自主推导过程，而通用问答工具只能生成标准答案；其二，通用工具无法沉淀学科专属的结构化题库，也无法采集学生的思维过程数据。这两类诉求共同指向一个判断：引导式教学、学科结构化题库与学习过程数据沉淀三项超出通用工具的功能边界，教学内核必须定向自建。

这一判断可以从本课题深度使用的两类工具中得到具体印证。以泛雅学习通为代表的通用学习管理系统，其逻辑以“资源－作业－测验－讨论”的流程管理为核心：它能准确记录学生是否提交、得分多少、任务完成到哪一步，却无法介入“学生如何想到这一步”的过程；答疑以师生异步留言为主，教师无法对每名学生的每次追问给出逐轮差异化应答，“引导式教学”在大班条件下因而缺乏可执行的载体。以“数你最灵”为代表的学科专用批阅工具，其能力集中于手写答案的正误判定，输出的是批改结果而非“下一步该问什么”；批改记录留在工具侧，未与课程的过程数据打通，难以支撑完整的学情画像。

两类工具的共同短板可归纳为三点：应答形态以“结论”为终点，缺乏面向概念形成过程的引导机制；数据颗粒度以“结果”为单位，缺乏对思维过程的结构化留痕；内容组织依附于既有课程流程，缺乏可按知识点调取的伴随问题体系。这三点正是第三阶段自建内核的功能清单，也构成平台功能设计的出发点。

3. 平台核心功能的设计理念与实现

第三阶段的研发遵循一条原则：只做通用工具做不到的事。以下依次说明平台的总体架构、双模式问答的教学设计及其在提示词层面的实现、该设计对传统教学痛点的回应，以及五模块闭环、知识图谱与题库架构。

3.1. 系统架构与技术栈

平台采用三层架构：浏览器端、Streamlit 应用层与本地 GPU 推理层，五大模块形成闭环。前端以 Streamlit 结合 Three.js (知识图谱三维可视化)与 KaTeX (数学公式渲染)构建；后端采用 Python 3.12 与 SQLite；AI 推理由 vLLM 本地服务加载 Qwen3.8 27B 量级开源模型，学生提交的代码通过 NBClient 子进程隔离运行，并设定 CPU、内存与超时限制以保证安全。平台部署于校内算力服务器(8×RTX 4090)。平台以课题立项经费自主研发，采用本地大模型私有化运行，学习数据不出校园，不依赖任何商业大模型 API，知识产权归本课题所有。平台首页与五模块总览如图 1 所示。



Figure 1. Homepage and five module overview of the math learning platform
图 1. 数学学习平台首页与五模块总览

3.2. 双模式问答系统：教学设计

双模式问答是平台最具辨识度的设计，其出发点是回答一个具体的教学问题：同一个模型面对同一个提问，能否走出两条不同的认知路径。平台在同一模型、同一知识库的前提下，提供标准模式与苏格拉底模式两种问答模式，仅通过提示词的差别实现两种截然不同的认知路径。以偏导数概念的建立为例：标准模式下，AI 直接从一元导数类比出发，给出偏导数的完整定义与公式，学生被动接收，属于“灌输

式答疑”；苏格拉底模式下，AI 不直接给出定义，而是以具体函数 $f(x, y) = x^2y$ 反问“你会把 y 看成什么”“你觉得偏导数会怎么定义”，引导学生从已掌握的一元导数知识中自行“长出”偏导数概念，属于“引导式教学”。两种模式的对照设计如表 4 所示，典型对话实录见图 2、图 3。

为确保这一教学核心功能不被埋没，模式开关置于侧边栏顶部并高亮显示(绿色提示条“已启用引导式提问”)，学生打开问答页面第一眼即可见、一键切换。两种模式的完整对话均入库留存、可按模式对照分析，这使得“引导是否比给答案更有效”成为一个可被数据检验的问题。

Table 4. Contrastive design of the two modes of the platform’s AI Q&A
表 4. 平台 AI 问答两种模式的对照设计

对照维度	标准模式	苏格拉底模式
提示词策略	直接讲解，给出完整定义与公式	以问题引导，不直接给结论
典型交互	类比一元导数，输出偏导数定义	以 $f(x, y) = x^2y$ 反问 y 的角色与定义设想
认知路径	灌输式答疑	引导式教学
数据归属	完整对话入库，可对照分析	同左，按模式标记区分
检验方式	两模式对话可对照，统计学习效果差异	同左



Figure 2. A dialogue record of the Socratic mode
图 2. 苏格拉底模式对话实录



Figure 3. A dialogue record of the standard mode
图 3. 标准模式对话实录

3.3. 提示词工程：把教学理念转化为可复用的应答约束

两种模式的差异只在提示词，这既是教学设计问题，也是工程问题：模型、知识库与知识检索范围保持一致，唯有系统提示词不同，两种模式的应答差异才可归因于提示词所承载的教学意图，而不是模型能力或知识来源的差异。

从平台的实现看，两种模式的提示词约束可归纳如下。标准模式要求“直接、完整”——不作反问，依次给出定义、公式与求解过程。苏格拉底模式要求三条约束：其一，不直接给出定义与结论，以反问推进；其二，以学生已掌握的知识为设问起点，如由一元导数类比引出偏导数；其三，优先调取该知识点在问答库中的伴随问题，作为引导下一步的素材。第三条是两种模式在数据依赖上的实质差异：标准模式只需知识点内容本身，苏格拉底模式还需要“下一步该问什么”，而这类知识无法由模型凭常识生成，只能来自问答库的人工建设。

提示词层面的设计带来两个后果。其一，应答差异可归因：由于模型、知识库与检索范围保持一致，两条对话路径的分歧点可以定位到具体的提示词约束，而非归因于模型的随机性。其二，教学经验可配置：教师关于“这个概念该怎么问”的经验一旦固化为提示词约束，便对全部会话生效，其传递方式由口耳相授转为系统配置。当然，这一机制的有效性最终取决于约束设计是否恰当，仍需通过对话实录的逐轮比对与后续对照实验检验(图 2、图 3)。

3.4. 双模式设计对传统教学痛点的回应

引言所列四类痛点中，双模式设计主要回应前两类。

一是教学模式单一。大班授课(单班 100~150 人)难以针对个体差异调整讲授节奏，而模式切换把“讲多深、给不给结论”的选择权交给学生：需要快速核对结论的学生用标准模式，需要重建概念的学生用苏格拉底模式；多轮对话本身也提供了大班课堂难以实现的“一对一追问”环境。

二是抽象概念理解困难。苏格拉底模式的价值不在“更友好”，而在于把概念的形成过程还原为可参与的步骤：以 $f(x, y) = x^2y$ 为例，学生先判断 y 在求偏导时的角色，再由一元导数的经验自行给出定义，

概念不是被告知的，而是被构造的。这一路径与高等数学概念抽象、依赖已有知识生长的学科特征相契合。

后两类痛点不是双模式所能解决。教师负担主要经第二阶段的智能批阅工具缓解(见 2.5 节)，作业批改由数小时压缩到分钟级；学用脱节则由在线习题与知识图谱部分回应——前者以可运行的代码验证数学结论，后者把知识点置于结构之中，缓解“只见公式不见结构”的问题。但“学用脱节”的根治仍需课程设计层面的改造，超出本平台的功能范围。

3.5. 五模块闭环与学习数据链路

平台由五个模块构成闭环。其一，问答库：采用“N 个为什么”结构化知识体系，配章节导航与伴随问题入口，每个问题含核心问题、提示(hint)、参考答案、关键词、类型标签与伴随问题；构建问答库的初衷在于降低学生的提问门槛——**提出有价值的问题本身有知识和能力门槛**，许多学生(尤其是初学数学的新生)并不擅长提问，问答库把经过精心构建的结构化问题前置呈现，使学生得以从感兴趣的问题入手进入学习，并沿“核心问题 - 提示 - 参考答案 - 伴随问题”的路径逐步深入；伴随问题是苏格拉底式追问的数据基础，AI 可优先调取它作为引导下一步的素材。其二，AI 对话：多轮流式对话，思考过程可见，对话自动记入统计看板。其三，在线习题：学生读题、编写代码，系统以 nbclient 子进程隔离运行并返回结果，用于验证数学结论。其四，知识图谱：以三维引力星球可视化组织学科知识结构，节点与知识点详解、教学视频、配套问答相互关联，详见 3.6 节。其五，统计看板：实时聚合总问答数、活跃学生数、习题成功率、高频知识点 TOP15 与最近问答记录，数据来自 SQLite，形成“问答日志→存储→指标聚合→可视化”的完整链路；两种模式的对话分别入库，可对照分析。五模块及其数据链路如表 5 所示，问答库与学习数据看板分别如图 4、图 5 所示。



Figure 4. Question bank interface with AI follow-up questions
图 4. 问答库界面与 AI 伴随追问

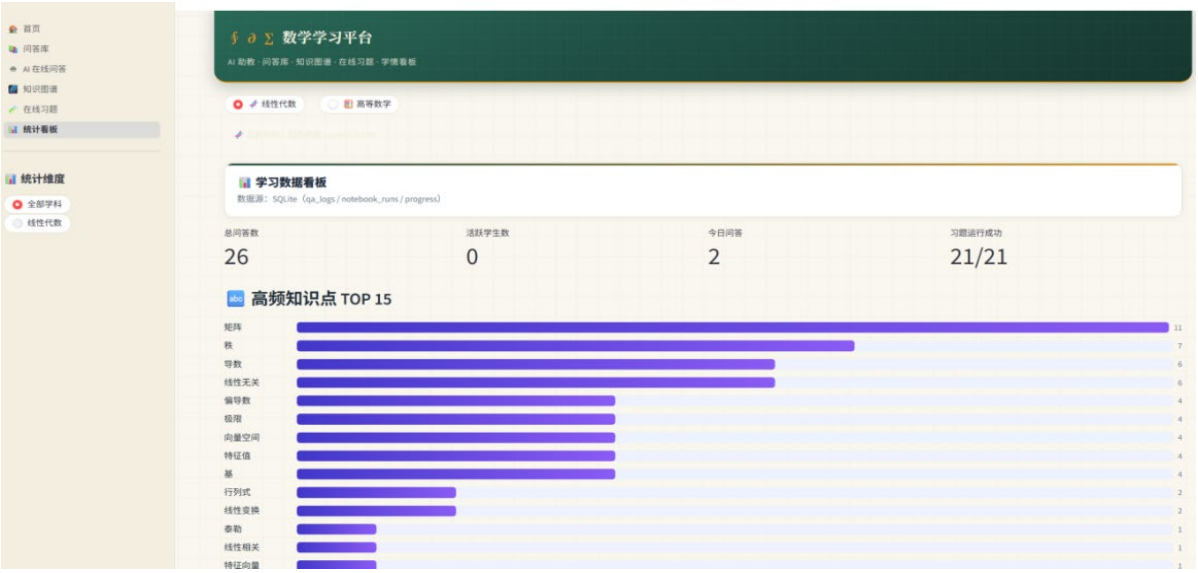

Figure 5. Learning data dashboard
图 5. 学习数据看板

Table 5. Functions of the platform's five modules and the learning data pathway
表 5. 平台五模块功能与学习数据链路

模块	核心功能	沉淀的学习过程数据
问答库	“N 个为什么”结构化知识体系，章节导航与伴随问题	问题访问与调取记录
AI 对话	多轮流式对话，思考过程可见	全量对话日志(按模式标记)
在线习题	代码编写与隔离运行，验证数学结论	习题提交与执行结果
知识图谱	Three.js 三维引力星球可视化，知识点引力关联	知识点浏览路径
统计看板	指标实时聚合与可视化	总问答数、活跃学生数、习题成功率、高频知识点

3.6. 知识图谱：学科知识结构的组织与关联

知识图谱承担两件事：把学科知识的结构呈现出来，把知识点与教学内容关联起来。平台采用 Three.js 渲染的“三维引力星球”可视化，以课程知识体系为中心、知识板块为聚类单元、知识点为节点，节点之间以引力连线表示彼此的关联，左侧配可检索、可展开的知识点列表，直观呈现学科知识的整体结构与关联关系(图 6)。

点击任一知识点节点即展开该知识点的详解面板，提供知识点精讲、LaTeX 渲染的数学公式与配套教学视频，使“知识点 - 讲解 - 视频”在同一次交互中完成(图 7)。

知识图谱是平台组织学科知识结构、支撑结构化提问的基础模块，其设计有三点考虑。其一，以可视化方式呈现知识之间的关联，帮助学生建立整体结构认知，缓解高等数学知识点多、跨度大、学习过程中易“见树不见林”的问题。其二，把知识点与教学视频、结构化问答打通：学生既可从图谱节点直接进入知识点精讲与配套教学视频，也可由此转向问答库中该知识点对应的问题条目与伴随问题，使苏格拉底式追问有明确的知识点锚点，这一关联也正是伴随问题能够按知识点调取的前提。其三，图谱的节

点点击与浏览行为构成学习过程数据的一部分，与问答日志、习题执行记录一并进入统计看板，为识别高频知识点与学习薄弱环节提供依据。受构建成本约束，平台当前的知识图谱采用轻量化方案，优先覆盖高等数学的核心知识点，节点规模与关联密度将随教学使用持续扩充(见 4.7 节)。



Figure 6. Overview of the knowledge graph of advanced mathematics
图 6. 高等数学知识图谱全貌



Figure 7. A knowledge point with its linked teaching video in the knowledge graph
图 7. 知识图谱中的知识点与配套教学视频

3.7. 题库建设与学科无关架构

平台当前接入线性代数(7 章 172 问)与高等数学(8 章 110 问)两门学科，合计 15 章 282 个结构化问题，并配套 52 道可一键运行的在线习题 Python 解答(线性代数 25 道、高等数学 27 道)。架构上，学科内

容与系统逻辑解耦，新增课程只需更换题库数据即可接入。该设计已通过从高等数学扩展到线性代数的实践验证，意味着平台可低成本迁移至其他数学类乃至其他理科课程的教辅场景。

4. 讨论

4.1. 路径选择的决策逻辑

三类路径在投入与产出上的系统对比如表 6 所示。对大多数教师而言，“使用生成式人工智能工具”比“开发生成式人工智能平台”更具可行性。但第三阶段的自主研发并不推翻上述判断，而是补充了一个更清晰的边界条件：轻量化路径解决的是“教学主流程数字化”，而引导式教学、学科结构化题库与学习过程数据沉淀三项超出通用工具的功能边界，必须自主研发才能实现。因此三阶段的关系是分工而非替代——工具应用保运转、自主研发保内核。

Table 6. Comparison of decision-making factors between self-development and tool-application modes
表 6. 自主开发与工具应用模式的决策要素对比

维度	自主开发模式	工具应用模式
时间投入	6 个月开发 + 持续维护	1~2 周工具筛选 + 培训
资金投入	服务器、API 调用、维护成本	购买成熟工具或免费使用
技术门槛	需掌握平台开发、API 对接	无需开发，直接使用
整合难度	与现有平台独立，数据不通	与学习通等自然融合
推广难度	需改变学生使用习惯	基于现有平台，易接受
可持续性	需持续更新维护	工具方负责更新

4.2. 技术选型：场景匹配优于技术先进

两年多的工具使用经验表明，技术选型的第一标准不是先进性，而是与教学场景的匹配度。不同教学场景对工具能力的要求存在结构性差异：知识讲解场景依赖长文本理解与逻辑推演能力，教案生成场景依赖结构化输出能力，习题识别场景依赖多模态识别能力，作业批改场景则要求对学科规范与解题步骤完整性的精细判断，此类学科专用工具通常比通用大模型更精准。工具选择的核心标准有三：准确性优先——教学场景对准确性要求高，宁可牺牲速度；集成便捷——能否与现有平台自然融合；成本可控——免费或低成本工具更适合长期推广。需要补充的一点方法论启示是：生成式人工智能工具生态迭代迅速，具体产品的性能位次随时间变化较快，相对稳定的是“按场景匹配能力、按标准筛选工具”这一选型方法本身，它比任何特定工具推荐都更易于迁移到新的工具环境。

4.3. 实施路径：轻量化优先于全面覆盖

实施策略上，轻量化优先于全面覆盖。项目初期曾尝试全覆盖(问答、批改、出题、教案同时推进)，结果资源分散；调整后聚焦核心痛点——作业批改耗时与个性化反馈不足，效果更集中，师生获得感更强。同时坚持与现有平台融合、不另起炉灶：学习通已承载教学主流程，生成式人工智能作为增强层叠加其上，避免让学生切换多个平台，降低使用门槛，并使数据在同一平台积累，形成完整学情画像。

4.4. 成本结构：时间成本往往大于金钱成本

两类路径的隐性成本结构差异显著。自主开发模式下，教师需投入平台开发技术学习(20~40 小时)、

持续调试与 Bug 修复，且与现有平台对接难以实现；工具应用模式下，成本主要是工具筛选与试用(1~2周)与师生培训(1~2 次讲座)，持续使用成本多为免费或低价。对一线教师而言，时间是最稀缺资源，工具应用模式的时间投入产出比更高；这也决定了“自主开发”必须是有明确边界的定向投入。本课题的数学学习平台只做三件通用工具做不到的事——引导式教学、学科结构化题库(282 个结构化问题)与学习过程数据采集，不追求功能大而全，这是自主研发在有限经费与人力下仍然可行的关键。

4.5. 推广策略：降低门槛比提升功能更重要

推广经验的核心是降低门槛而非堆砌功能。学生端，基于现有平台(学习通)无需重新注册，操作流程尽量简单，并与成绩挂钩(平时成绩包含平台互动)；教师端，无需技术背景、工具开箱即用，配套操作手册与案例库，并建立教师交流社区分享经验。激励机制上，将平台使用纳入平时成绩(签到、测验、作业)，设立“AI 学习之星”等荣誉称号，提供个性化学习报告作为学习证明。

4.6. 教学属性的转换：AI 从答案提供者到思维引导者

平台最深层的价值不在于功能集合，而在于完成了 AI 教学属性的转换。通用生成式人工智能工具的本质属性是“答案提供者”——高效、准确、即时应答，但其输出形态天然偏向结论；而数学概念学习(如偏导数)的关键难点恰在概念的形成过程。平台以同一模型、同一知识库承载两种问答模式，仅通过提示词差别切换“灌输式答疑”与“引导式教学”两条认知路径，其意义有三：一是把教学理念创新落实为可运行、可切换的系统功能，而非停留在口号；二是两种模式的对话全量入库并按模式标记，使两种模式的效果对比成为可量化研究的问题；三是苏格拉底模式所依赖的“伴随问题”体系，为引导式提问提供了结构化的学科知识基础，使“怎么问”本身成为可建设、可复用的教学资产。

4.7. 现实挑战与应对策略

实践中仍面临四类挑战。一是数学公式处理难题：虚拟人语音播报数学公式仍是未解难题，公式朗读规则复杂(∫f(x)dx 应读作“f(x)的不定积分”而非符号朗读)，现有 TTS 引擎对数学公式支持不足，拍照识别准确率仅 70%~80%；短期采取接受文字输出、放弃语音播报的策略，长期探索专用数学公式识别引擎与教育 TTS 厂商合作(成本高，需评估)。二是知识图谱构建成本高：完整知识图谱需数百上千节点，人工构建工作量大、商业产品价格昂贵；平台采用轻量化方案，梳理 50~60 个核心知识点(图 6、图 7)，基于大模型实现知识点动态关联，并利用用户行为数据持续优化。三是师生接受度问题：学生习惯传统学习方式、担心技术出错影响成绩，教师存在畏难情绪、担心增加工作负担；应对策略是渐进式推广(先小范围试点)、明确工具定位(辅助而非替代)、提供持续培训与技术支持。四是平台整合难题：学习通与“数你最灵”等工具数据独立，无法形成完整学情画像；改进方向包括探索与学习通 API 对接、建立数据中间层实现部分数据互通、等待平台方开放更多接口。

5. 结论与展望

5.1. 主要结论

第一，三阶段演进澄清了通用工具与自建内核的能力边界。从“第三方平台受挫 - 转向工具应用 - 自主研发内核”的演进表明：轻量化路径解决了教学主流程的数字化问题，自主研发解决了引导式教学与学习数据沉淀问题，二者构成“工具应用保运转、自主研发保内核”的分层格局。第二，轻量化路径仍具有可复制性。“学习通平台 + 成熟生成式人工智能工具”的融合模式具有时间投入少、技术门槛低、易于推广、可持续性强等特点，对资源有限的高校依然是最稳妥的起步选择，但需在其上针对三项能力

边界自建轻量内核。第三，自主研发的价值在于教学内核而非技术堆砌。在时间成本这一核心约束下，自主开发必须是定向投入；只做通用工具做不到的三件事，是其在有限经费与人力下仍然可行的关键。

5.2. 研究局限

本研究存在三点局限。其一，平台统计看板当前数据为系统联调阶段的功能验证数据，样本量较小，本文主张的是数据链路已跑通与工程实现完整，教学成效尚未获规范证实；其二，实践在单一高校的单一课程群中开展，结论的外推需谨慎；其三，标准模式与苏格拉底模式的效果对照目前只完成了系统设计与数据留痕，尚无对照实验结果。

5.3. 未来工作

后续将沿三个方向推进：短期(1~2 年)完善“学习通 + 生成式人工智能工具”融合模式，扩大应用规模积累实证数据，并基于已跑通的数据链路开展两种教学模式的对照实验，检验引导式教学相对灌输式答疑的学习效果差异；中期(3~5 年)推动生成式人工智能工具与主流教学平台的数据互通，形成生成式人工智能赋能教学的标准化实施路径，扩充知识图谱的知识点规模与跨课程关联，建设跨校共享的生成式人工智能教学资源库；长期(5 年以上)探索生成式人工智能技术与教学流程的深度融合、智能化个性化学习路径推荐，以及“人机协同”的新型教学生态。

基金项目

广东省教育科学规划课题(高等教育专项)“AIGC 赋能的数智化大学数学教学改革与实践”(项目编号：2024GXJK275)。

参考文献

[1] 郑成勇, 陈帅. 基于 AIGC 的高等数学教辅系统开发[J]. 创新教育研究, 2025, 13(8): 703-715.

[2] 范洪光, 周安然. 人工智能专业高等数学教学模式改革探究[J]. 创新教育研究, 2025, 13(9): 155-161.

[3] 李珊, 杨天愉. AI 赋能高等数学课程混合式教学改革研究与探索[J]. 教育进展, 2026, 16(1): 915-923.

[4] 贾晓彤, 王利岩, 赵晓丽, 等. 人工智能融入高等数学实践混合式教学模式的探究[J]. 教育进展, 2025, 15(10): 822-826.

[5] 刘仁龙. “AI+”赋能高等数学课程教学改革与实践[J]. 美术与设计, 2025(12): 72-75.

[6] 江桐. 人工智能背景下高等数学课堂教学创新实践[J]. 美术与设计, 2025(11): 87-90.

[7] 金迎迎. 人工智能在“高等数学”教学中的应用探究[J]. 科技风, 2024(1): 135-137.

[8] 李玲. 融合人工智能技术的高等数学课程混合式教学改革与实践[J]. 数字技术与应用, 2024, 42(1): 134-136.

[9] 魏重庆, 李安然. AI 赋能线性代数课程混合式教学研究与实践[J]. 信息系统工程, 2025(4): 169-172.

[10] 白雪梅, 郭日发. 生成式人工智能何以赋能学习、能力与评价? [J]. 现代教育技术, 2024, 34(1): 55-63.

[11] 周文辉, 赵金敏. ChatGPT 对研究生创新能力培养的价值与挑战[J]. 高校教育管理, 2024, 18(2): 42-52.

[12] 刘勇, 徐佳慧, 董跃武, 等. 高等教育中如何应用 ChatGPT 类生成式人工智能——联合国教科文组织《高等教育中的 ChatGPT 和人工智能: 快速入门指南》解析[J]. 中国教育信息化, 2024, 30(2): 71-80.

[13] 罗恒, 廖小芳, 茹琦琦, 等. 生成式人工智能支持的教师评语研究: 基于初中数学课堂的实践探索[J]. 电化教育研究, 2024, 45(5): 58-66.