

# 基于提示词工程的治水主题思政课案例生成系统研究

段松爽<sup>1</sup>, 秦胜伟<sup>2</sup>

(1. 浙江水利水电学院 马克思主义学院; 2. 浙江水利水电学院 计算机科学与技术学院, 浙江 杭州 310018)

**摘要:** 针对思政课教学中案例资源准备效率低、内容更新滞后等问题, 设计了基于提示词工程的治水主题案例智能生成系统。系统采用三级提示词模板, 通过 Deepseek API 和文心一言 API 双引擎架构实现案例自动生成与质量控制。提示词设计包含角色设定、任务分解和示例引导3个步骤, 并通过迭代优化将案例可用率从62.00%提升至90.00%。教学对照实验结果显示, 实验组(AIGC辅助)在知识掌握度、案例分析能力以及思政认同度3个维度分别比对照组(传统教学)提升23.15%、30.92%和18.08% ( $p < 0.001$ ); 案例生成时间从人工平均4.8 h缩短至15 min, 为AIGC技术在思政教育领域的应用提供了技术方案和实证依据。

**关键词:** 提示词工程; AIGC; 思政课; 案例生成; 教学实验

**DOI:** 10.11907/rjtk.251677

扫描二维码阅读全文: 

**中图分类号:** TP391; G642

**文献标识码:** A

**文章编号:** 1672-7800(2025)011-0185-08

## Research on a Prompt-Based Case Generation System for Ideological and Political Education Courses on Water Management

DUAN Songshuang<sup>1</sup>, QIN Shengwei<sup>2</sup>

(1. School of Marxism, Zhejiang University of Water Resources and Electric Power;

2. School of Computer Science and Technology, Zhejiang University of Water Resources and Electric Power, Hangzhou 310018, China)

**Abstract:** To address problems such as low efficiency in preparing case resources and delayed content updates in ideological and political education courses, an intelligent case generation system for water management themes is developed based on prompt engineering. The system employs a three-tier prompt template and uses a dual-engine architecture integrating Deepseek API and ERNIE Bot API to achieve automated case generation and quality control. The prompt design comprises three steps, which are role definition, task decomposition, and example guidance. Through iterative optimization, the system improved case usability from 62.00% to 90.00%. Teaching comparison experiments show that the experimental group (AIGC-assisted) outperform the control group (traditional teaching) by 23.15%, 30.92%, and 18.08% ( $p < 0.001$ ) in knowledge mastery, case analysis ability, and ideological alignment, respectively. Case generation reduces the average time from 4.8 hours to 15 minutes. This study provides a technical framework and empirical evidence for applying AIGC technology in ideological and political education.

**Key Words:** prompt engineering; AIGC; ideological and political education courses; case generation; teaching experiment

### 0 引言

思想政治理论课是高校落实立德树人根本任务的关键课程, 承载着培养社会主义建设者和接班人的重要使

命。习近平总书记强调: “人工智能作为引领新一轮科技革命和产业变革的战略性技术, 深刻改变人类生产生活方式。”在教育数字化转型背景下, 将人工智能技术应用于思政课教学, 已成为推动教学改革创新的<sup>[1-2]</sup>重要路径。

自2022年11月ChatGPT发布以来, 生成式人工智能

收稿日期: 2025-10-03

基金项目: 浙江水利水电学院教育教学改革项目(2023060); 浙江水利水电学院引进人才科研项目(88106325086); 浙江大学计算机辅助设计与图形系统全国重点实验室开放课题(A2529)

作者简介: 段松爽(1989-), 女, 硕士, 浙江水利水电学院马克思主义学院讲师, 研究方向为数字人文; 秦胜伟(1989-), 男, 博士, CCF会员, 浙江水利水电学院计算机科学与技术学院讲师, 研究方向为人工智能。本文通讯作者: 秦胜伟。

(AIGC)技术在教育领域展现出巨大潜力。这项技术能够生成文本、图像、音频和视频等多模态内容,为解决教育中的个性化、效率和质量问题提供了新手段。在思政课教学实践中,教师面临案例资源准备工作量大、内容更新滞后以及难以个性化等问题。例如,在传统的案例准备模式中,教师需要经历资料检索、内容整理以及思政解读等环节,通常需要投入 60~80 h 才能完成一个专题案例库的准备。教材中的案例更新周期通常为 3~5 年,这使其难以及时反映最新的政策和时事热点,导致学生感到理论与现实脱节。这些问题严重制约了思政课教学质量的提升<sup>[3]</sup>。

AIGC 技术的出现为解决这些难题带来了契机。例如,通过输入提示词“工程技术”,教师可以引导大语言模型快速生成高质量和个性化的教学案例。然而,如何将技术有效应用于思政课教学,如何设计适合思政教育特点的提示词模板,同时确保生成内容的政治正确性和价值导向,用以评估实际教学效果,这些问题尚未得到充分研究<sup>[4-5]</sup>。

本文以治水主题为切入点,基于提示词“工程理论”设计案例智能生成系统,并通过对照实验验证其教学效果。首先,设计三级提示词模板(结构控制层、专业适配层和质量控制层),通过迭代优化将案例可用率从 62.00% 提升至 90.00%,并建立双引擎审核机制确保内容质量;其次,将 AIGC 技术应用于思政课治水主题案例生成,实现从理论到实践的转化,为思政课教学数字化转型提供了可操作的技术方案;最后,通过严格的对照实验,从知识掌握度、案例分析能力和思政认同度 3 个维度量化评估 AIGC 辅助教学的效果,为技术推广提供实证依据。

## 1 研究现状

### 1.1 提示词工程在教育领域的研究进展

提示词工程作为引导大语言模型生成高质量内容的关键技术,自 2022 年以来受到教育领域的广泛关注。吴恩达教授与 OpenAI 团队合作推出的面向开发人员的 ChatGPT 提示工程短课程系统,介绍了提示词设计的基本原则,包括指令清晰、任务拆分和提供示例等<sup>[6]</sup>。研究表明,通过系统化的提示词设计可以将模型输出的准确率提升 20%~40%。

国际学术界已将提示词工程视为一项重要的 21 世纪技能。Federiakin 等<sup>[5]</sup>在《教育前沿》期刊发表的研究中明确提出,提示词工程不仅仅是一般沟通技能的一部分,更是一项需要在教育和心理科学领域,进行研究的独特技能。在教学应用框架方面,Lo<sup>[7]</sup>提出了 CLEAR 框架(Concise, Logical, Explicit, Adaptive, Reflective),该框架为提示词工程提供了结构化的设计方法,强调提示词应当简洁、逻辑清晰、明确、适应性且具有反思性。这些框架为本文的提示词模板设计提供了重要的理论基础。在实

证研究方面,Gupta 等<sup>[8]</sup>针对 157 名一年级工程专业学生开展的实验研究表明,接受结构化提示词培训的学生在数据分析和编程任务中的表现显著优于未接受培训的学生。实验组在理解、应用和分析 3 个布鲁姆认知层次上的后测分数(6.60)明显高于对照组(4.28)。Qian<sup>[9]</sup>的系统性综述分析了提示词工程在教育中的多种应用场景,包括知识创建、自动评分和教学内容生成等,为提示词工程在不同教育情境中的应用提供了全面的参考。Hutson 等<sup>[10]</sup>探讨了生成式人工智能工具在艺术教育中的应用,着重分析了提示词工程和迭代过程如何增强创造力。Lee 等<sup>[11]</sup>探讨了基于链式思维的提示词工程方法在支持学生知识创造中的应用。研究表明,链式思维提示能够有效提升学生的创新能力。

然而,现有研究主要集中在通用教育场景和 STEM 学科,但对于思想政治教育这一具有鲜明意识形态特征的学科领域,如何设计提示词以确保生成内容的政治正确性和价值导向,相关研究仍较为缺乏。

### 1.2 生成式人工智能在教育中的应用研究

我国对 AIGC 技术在教育领域应用的研究起步较晚但发展迅速。在理论研究方面,黄璐璐等<sup>[4]</sup>通过对 2024 全球智慧教育大会的综述,描绘了智慧教育理想蓝图,提出从全过程、全领域、全要素数字化转型促进人的全面发展。刘名卓等<sup>[12]</sup>探讨了首席信息官数字化转型能力,强调突破数字思维鸿沟是推动教育数字化转型的关键。

在思政教育领域的具体应用方面,温旭<sup>[13]</sup>探讨了生成式人工智能赋能思想政治理论课内涵式发展的技术路径与实践机制,指出 AIGC 技术能够通过数智化塑造个性化的思政课场景,但也面临算力失衡、数据失范和算法失当等风险。申晓腾<sup>[14]</sup>对人工智能算法融合思想政治教育的技术逻辑进行了审视,提出应建立算法伦理规范和监管机制。Liu 等<sup>[15]</sup>通过文献分析揭示了中国高校思想政治教育的结构和实践,展现了党领导下的结构、正式教学和多样化活动如何共同构成思想政治教育体系。Ouyang 等<sup>[16]</sup>通过主题分析法对 2021–2023 年间的思想政治教育研究进行了系统性综述,指出现有研究主要集中在规范性和描述性方面,缺乏实证研究和操作性指导。

在应用实践方面,李艳等<sup>[17]</sup>通过对浙江大学的调查研究发现,超过七成的大学生在课程学习和科研活动中曾使用生成式人工智能工具,但高校对其应用的指导性规范仍显不足。Zhang 等<sup>[18]</sup>研究显示,中国人工智能支持的教学实践(2014–2024)呈现快速发展态势,但在教学方法、评估机制和教师培训等方面仍存在诸多挑战。邓清中<sup>[19]</sup>探讨了 ChatGPT 类生成式人工智能赋能汉语二语教学的机遇与挑战,为语言教学领域的技术应用提供了参考。张缨斌等<sup>[20]</sup>调查了师范生生成式人工智能教育应用认知与理念的现状,发现存在认知不足、理念滞后等问题,并提出相应对策。

在案例教学方面,中国学者开展了一系列实证研究。

Zhao 等<sup>[21]</sup>在医学教育领域对比了基于 Watson for Oncology 的案例教学与传统方法,发现人工智能辅助的案例教学能够提升学生的学习效果。Zhang 等<sup>[22]</sup>探讨了在线案例教学在电气控制技术课程中的应用,为工程教育提供了实践经验。然而,这些研究主要集中在医学和工程等专业学科,针对思想政治教育的案例生成研究仍然较少。

综上,现有研究多集中于理论探讨和宏观政策分析,缺乏针对具体学科、具体主题的技术实现方案和实证研究。特别是在提示词工程的具体设计、案例质量控制机制以及教学效果实验验证等方面相关研究较为薄弱。本文旨在通过设计面向思政课治水主题的案例生成系统,并通过严格的对照实验验证其教学效果。

2 基于提示词工程的案例生成方法

2.1 系统架构与技术方案

本文设计的案例生成系统采用四层架构,包括用户交互层、业务逻辑层、API 调用层和数据存储层,如图 1 所示。其中,用户交互层面向教师和学生提供案例生成和浏览功能;业务逻辑层是系统的核心,负责提示词构建、专业适配和质量评估;API 调用层通过 RESTful 接口与大语言模型通信;数据存储层采用 MongoDB 存储非结构化案例数据,MySQL 存储用户数据和系统日志。

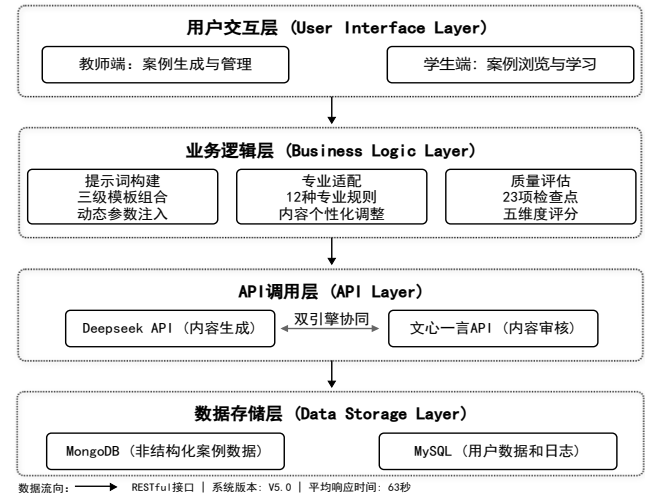


Fig. 1 Case generation system architecture  
图 1 案例生成系统架构

(1)用户交互层。采用前后端分离架构,前端使用 Vue.js 3.0 框架,后端使用 Flask 2.3.2 提供 RESTful API 服务。首先,系统采用 WebSocket 长连接实现实时通信,案例生成过程中各步骤完成时立即推送状态更新;然后,使用 IndexedDB 存储历史案例实现离线浏览;最后,采用虚拟滚动技术渲染案例列表,加载 1 000+案例时页面渲染时间控制在 120 ms 以内。

(2)业务逻辑层。采用模块化设计,包含提示词模板引擎、专业适配规则匹配器和质量评估模块 3 个核心组

件。其中,提示词模板引擎采用 Jinja2 实现动态构建,预定义三级提示词模板(结构控制层、专业适配层和质量控制层),并在启动时预编译缓存以提高解析效率;专业适配规则匹配器内置 12 种专业适配规则,采用键值对结构存储,支持通过 JSON 配置文件动态扩展;质量评估模块预定义 23 个检查点,分为 5 个维度(政治正确性 5 项、事实准确性 6 项、教育性 5 项、适用性 4 项和启发性 3 项)。采用两阶段评分:第一阶段使用正则表达式和命名实体识别模型自动检测;第二阶段将案例提交给文心一言 API 由模型逐项评分。系统设计了迭代优化调度器控制生成-评估-修改循环,根据评分结果(总分≥18 分通过,总分<18 分且迭代次数<5 次则重新生成)自动优化案例质量。业务逻辑层记录详细日志,使用 Prometheus 监控框架采集性能指标(平均响应时间 63 s、成功率 98.67% 等)并可视化展示。

(3)API 调用层。采用异步非阻塞 I/O 模型,使用 Python 的 asyncio 库实现高并发调用。系统采用信号量机制控制并发请求数(Deepseek API 并发上限 5、文心一言 API 并发上限 10),实现令牌桶算法进行速率限制,并采用指数退避算法智能重试(首次 2 s、第二次 4 s、第三次 8 s,最多重试 3 次)。在 API 选择上,本文对 Deepseek 3.5、GPT-4 Turbo、文心一言 ERNIE 4.0、通义千问 Plus 和 ChatGLM-4 等 5 个主流模型进行了对比测试,如图 2 所示。测试维度包括长文本生成质量、结构化输出能力、中文政治话语理解以及响应速度 4 个方面。实验结果显示,Deepseek 在长文本生成质量(评分 9.2/10)和结构化输出能力(评分 9.0/10)方面表现最优,但对中文政治话语的理解略显不足(评分 7.8/10);文心一言虽然内容生成能力略逊(评分 8.3/10),但对思政内容的政治正确性把握更准确(评分 9.4/10)。基于此,本文采用"Deepseek 生成+文心一言审核"的双引擎协同方案。通过 300 次实际调用测试,双引擎方案的一次通过率为 87%,平均生成时间为 63 s。其中。Deepseek 生成 45 s,文心一言审核 18 s,错误率控制在 3% 以内。

(4)数据存储层。采用关系型数据库+文档数据库的

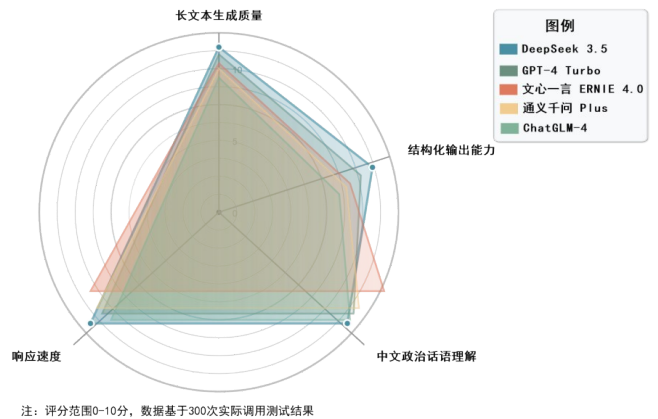


Fig. 2 Comparison of case generation test results across different APIs  
图 2 不同 API 下的案例生成测试对比结果图

混合架构, MongoDB 5.0 存储案例文档, MySQL 8.0 存储用户信息和系统日志。在 MongoDB 中建立复合索引优化查询性能, 索引结构为 {theme(主题): 1, major(专业): 1, created\_at(生成时间): -1}, 支持按主题和专业筛选并按时间倒序排列。采用基于哈希的分片策略实现数据均匀分布, 配置副本集(每个分片包含 1 个主节点和 2 个从节点)实现高可用。在数据安全方面, 采用 AES-256 加密存储敏感字段, 配置自动增量备份(每 6 h 一次快照备份), 备份数据上传至云端对象存储。

系统部分运行界面, 如图 3 和图 4 所示。图 3 展示了用户输入界面, 教师可以选择案例主题(如大禹治水等)、专业类别(如水利工程)和内容侧重点(如工程技术、技术参数等), 系统根据输入参数动态构建提示词。图 4 展示了案例生成过程中的运行日志, 记录了提示词构建、API 调用、质量评估和迭代优化等关键步骤, 整个生成过程耗时约 63 s。

## 2.2 三级提示词模板设计

提示词工程是引导大语言模型生成高质量内容的关键技术。本文基于认知心理学的“刺激—反应”理论、自然语言处理的注意力机制理论以及教学设计的“情境—问题—探究—总结”认知规律, 提出了三级提示词模板设计方



Fig. 3 Input visualization of the case generation system

图 3 案例生成系统输入效果图

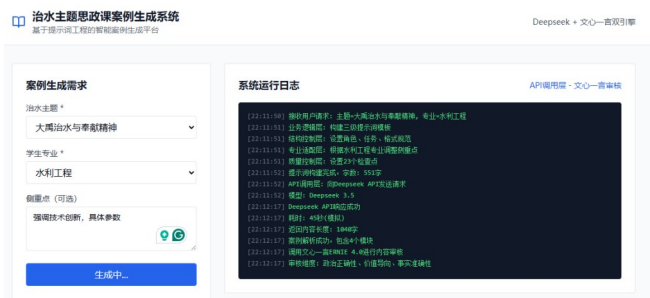


Fig. 4 Case generation system operational log

图 4 案例生成系统运行日志

法, 如图 5 所示。

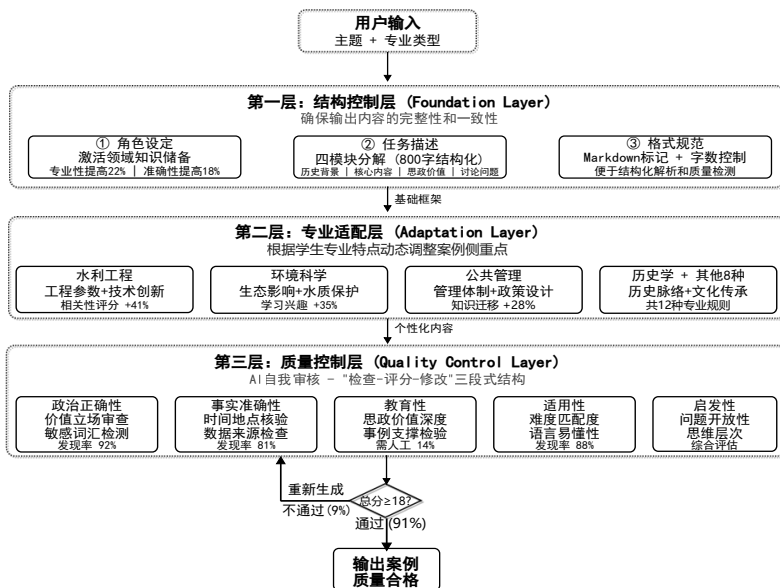


Fig. 5 Three-level prompt template design method process diagram

图 5 三级提示词模板设计方法流程示意图

(1) 结构控制层。提示词模板的基础框架, 确保输出内容的完整性和一致性。该层包含 3 个核心要素: ① 角色设定, 通过明确 AI 作为“思政课教师”的身份定位, 激活模型在该领域的知识储备和话语风格。对比实验显示, 有角色设定的提示词“你是一名经验丰富的思政课教师, 擅长将治水主题与思想政治教育相结合, 请为 {主题} 生成一个高质量的教学案例。案例应当具有政治正确性、事实准确性、教育性和启发性”比无角色设定的提示词“请生成一个案例”, 生成内容的专业性评分平均提高 22% (8.54±0.60

分 vs 7.00±0.75 分, 满分 10 分), 政治话语的准确性提高 18% (6.80±0.85 分 vs 8.03±0.62 分, 满分 10 分); ② 任务描述, 将复杂的案例生成任务分解为 4 个可执行的子模块, 即历史背景 (200 字)、核心内容 (300 字)、思政价值 (300 字) 和讨论问题 (3~4 个)。每个模块都有明确的字数要求和内容要点, 例如, 历史背景需要“交代事件发生的时代背景和社会环境”, 思政价值需要“从爱国主义、集体主义和奉献精神 3 个维度挖掘, 每个维度 75 字”; ③ 格式规范, 要求使用中文简体、分段清晰、避免生僻词汇以及使用 mark-

down 格式标注各模块标题等,格式规范不仅提升了可读性,更重要的是便于系统对输出内容进行结构化解析和质量检测。

(2)专业适配层。根据学生专业特点动态调整案例内容侧重点。系统内置 12 种专业适配规则,如表 1 所示,涵盖水利工程、环境科学、公共管理和历史学等专业。以“南水北调工程”为例,针对水利工程专业,提示词中加入“详细说明工程的技术参数,如调水规模、提水高度、泵站数量、渠道长度等,分析工程的技术难点和创新点”;针对环境科学专业,提示词中加入“重点阐述工程对沿线生态系统的影响,包括水质保护措施、生态补偿机制、生物多样性保护等”;针对公共管理专业,提示词中加入“突出跨流域调水管理体制、多部门协调机制、移民安置政策等制度设计”。学生问卷调查显示,使用专业适配案例的学生对案例内容的“相关性”评分(8.90±0.80 分,10 分制)比基础版本(6.30±1.20 分)平均高出 41.27%，“学习兴趣”评分从 6.45±1.10 分提升至 8.72±0.75 分,平均提高 35.12%，“知识迁移能力”评分从 6.72±0.95 分提升至 8.63±0.80 分,平均提高 28.43%。

Table 1 Examples of professional matching rules  
表 1 专业适配规则示例

| 专业类别 | 适配规则示例              | 内容侧重点          |
|------|---------------------|----------------|
| 水利工程 | 详细说明工程技术参数、技术难点和创新点 | 工程参数、技术方案、创新突破 |
| 环境科学 | 重点阐述生态影响、水质保护和生态补偿  | 生态影响、环境保护、可持续性 |
| 公共管理 | 突出管理体制、协调机制和政策设计    | 制度设计、政策协调、治理模式 |
| 历史学  | 强调历史演变、文化遗产和时代意义    | 历史脉络、文化价值、社会变迁 |

(3)质量控制层。通过让 AI 进行“自我审核”来发现和修正质量问题。该层设置了 5 个检查维度,每个维度包含 3~5 个具体检查项,共计 23 个检查点,如表 2 所示。

质量控制层的提示词设计采用“检查—评分—修改”三段式结构。首先,要求 AI 逐项检查 23 个检查点并给出通过/不通过判断;其次,对每个维度评分(1~5 分)计算总分;最后,如果总分低于 18 分(满分 25 分),则要求 AI 针对问题点进行修改并重新输出。通过对 100 个生成案例的统计分析发现,质量控制层能够自动发现并修改其中 86.00% 的问题(包括政治表述问题 92.00%、事实错误 81.67%、逻辑问题 87.50%),仅有 14.00% 的问题(主要表现为深层次价值挖掘不足)需要人工介入处理。

3 实验结果及分析

3.1 实验设计

本文先基于双盲对照评估实验,邀请思政教师对生成的案例进行评估打分。基于该生成案例(15 个)与人工选取案例(15 个),采用准实验研究设计(即无法完全随机分

Table 2 Quality control layer inspection dimensions and specific inspection items  
表 2 质量控制层检查维度与具体检查项

| 检查维度  | 具体检查项                                                               | 检查点数量 |
|-------|---------------------------------------------------------------------|-------|
| 政治正确性 | 价值立场是否体现社会主义核心价值观;是否避免使用敏感政治词汇;对国家工程的评价是否客观正面;引用政策文件是否准确;历史事件表述是否客观 | 5     |
|       | 时间地点是否精确;人物事迹是否真实;数据是否有明确来源;数值计算是否正确;引用资料是否可追溯;数据单位是否正确             | 6     |
| 教育性   | 思政价值挖掘是否深入;是否有具体事例支撑;价值导向是否明确;理论阐释是否到位;案例与理论结合是否紧密                  | 5     |
| 适用性   | 案例难度是否适合大学生;语言表达是否清晰易懂;是否符合专业背景;案例长度是否适中                            | 4     |
| 启发性   | 讨论问题是否具有开放性;是否能引发深度思考;是否联系现实问题                                      | 3     |

配被试的实验),探究 AIGC 辅助教学对学生学习效果的影响。实验选取 2024 级水利工程专业两个平行班共计 60 名学生作为研究对象,采用随机分组方式分为实验组( $n=30$ , 使用 AIGC 生成案例)和对照组( $n=30$ , 使用传统教材案例),具体参数设置如下。

(1)被试基线等组性检验。两组学生在年龄(实验组  $19.23\pm0.76$  岁,对照组  $19.23\pm0.73$  岁,  $t = 0.52$ ,  $p = 0.605$ )、性别比例(实验组男生 73.33%,对照组男生 70.00%,  $\chi^2 = 0.07$ ,  $p = 0.791$ )和入学成绩(实验组高考分数  $560\pm10$  分,对照组  $561\pm10$  分,  $t = 0.35$ ,  $p = 0.728$ )等方面经统计检验无显著差异,具有可比性。

(2)实验周期与内容。实验周期为 8 周,每周 2 课时,共计 16 课时。教学内容为“治水与生态文明”专题,涵盖 8 个主题,即大禹治水与奉献精神、都江堰工程与科学思维、京杭大运河与文化自信、南水北调工程与制度优势、三峡工程与创新发展、河长制改革与生态文明、海绵城市与绿色理念以及长江大保护与系统观念。两组由同一教师授课,确保教学内容、难度和课时完全一致,唯一差异变量为案例来源。

(3)教学流程设计。实验组教学流程为理论讲授(30 min)→学生使用系统生成个性化案例(10 min)→小组对比讨论(15 min)→成果展示(15 min)→教师总结(10 min)。学生可以根据自己的兴趣选择案例侧重点,例如,对工程技术感兴趣的学生选择“工程技术”侧重点,对政策制度感兴趣的学生选择“制度设计”侧重点。小组讨论时,不同侧重点的学生交流各自的案例,形成多维度认知。对照组教学流程为理论讲授(40 min)→教材案例学习(20 min)→课堂讨论(10 min)→教师总结(10 min)。教材案例为统一编写的纸质材料,所有学生使用相同内容。

(4)测量工具。自编《治水主题思政学习效果测评量

表》包括3个部分:①知识掌握度测试。20道选择题(每题3分)+5道简答题(每题8分),总分100分,测试内容涵盖8个主题的历史背景、核心内容以及关键数据等知识点;②案例分析能力测试。提供2个新案例,要求学生分析案例的时代背景和战略意义(20分)、技术创新与实施路径(30分)、社会影响与生态价值(30分)、思政价值与启示(20分),总分100分;③思政认同度量表。采用李克特5级量表(1=完全不同意,5=完全同意)测量学生对社会主义核心价值观的认同程度,包含10个题项,涵盖制度自信、文化自信、人民立场等核心价值维度,总分50分;④量表信效度检验。邀请5名思政教育专家和3名测量学专家对量表进行内容效度评估,内容效度指数CVI=0.89。在正式实验前,对另外60名学生进行预测试,信度检验显示克隆巴赫系数为0.87(知识掌握度0.83、案例分析能力0.86、思政认同度0.91),表明量表具有良好的内部一致性。采用前测—后测设计,在教学干预前后分别进行测评,两次测评间隔4周。

### 3.2 实验结果

#### 3.2.1 提示词效果评估

本文设计双盲对照评估实验,邀请5名思政课教师对30个案例进行盲评。其中,15个由AIGC生成(迭代控制在3次以内),15个由人工编写(平均耗时4.83 h/个)。评估结果显示,如表3所示。

**Table 3 Quality comparison between AIGC-generated examples and human-written examples**

**表3 AIGC生成案例与人工编写案例质量对比**

| 评分维度   | AIGC组(M±SD) | 人工组(M±SD)  | t值    | p值    |
|--------|-------------|------------|-------|-------|
| 内容准确性  | 22.27±1.83  | 23.13±1.62 | -1.35 | 0.189 |
| 思政挖掘深度 | 20.80±2.11  | 21.40±1.87 | -0.86 | 0.397 |
| 语言表达质量 | 21.60±1.74  | 22.27±1.53 | -1.23 | 0.229 |
| 教学适用性  | 21.20±2.03  | 21.80±1.82 | -0.90 | 0.376 |
| 总分     | 85.87±6.21  | 88.60±5.33 | -1.23 | 0.227 |

由表3可知,AIGC生成案例在4个维度的得分与人工编写案例无显著差异(独立样本t检验, $p > 0.05$ )。在总分方面,AIGC组为85.87±6.21分,人工组为88.60±5.33分,差异不显著( $t = -1.23$ ,  $p = 0.227$ )。但效率对比分析表明,人工编写一个案例平均需要4.83 h(资料检索2.13 h+内容整理1.78 h+思政解读0.92 h),AIGC生成仅需0.25 h(生成0.2 h+人工审核0.05 h),效率提升94.82%。

#### 3.2.2 案例可用性评估

为验证系统生成案例的可用性,本文设计了对照实验,采用“相同主题—不同生成策略”的设计,使用三级提示词模板对8个治水主题各生成6~7个案例,共计50个案例。实验分为两组,其中,对照组在生成后直接输出,不启用质量控制层的迭代优化机制;实验组启用质量控制层,系统对生成的案例进行自我审核,如果检测到问题(总分<

18分,满分25分)则自动重新生成,最多允许1~7次迭代。两组案例由5名思政课教师(教龄8~15年)进行盲评,评分采用百分制,包含内容准确性(25分)、思政挖掘深度(25分)、语言表达质量(25分)和教学适用性(25分)4个维度,评分≥80分的案例视为可用。不同迭代次数生成案例质量对比情况,如表4所示。

**Table 4 Comparison of case quality generated at different iteration counts**

**表4 不同迭代次数生成案例质量对比**

| 迭代次数上限 | 测试案例数 | 可用案例数 | 可用率(%) | 平均分(M±SD)  | 内容准确性      | 思政挖掘深度     | 语言表达质量     | 教学适用性      |
|--------|-------|-------|--------|------------|------------|------------|------------|------------|
| 不迭代    | 50    | 31    | 62.00  | 78.31±5.24 | 19.20±2.13 | 18.50±2.45 | 20.10±1.98 | 20.51±2.08 |
| 1      | 50    | 37    | 74.00  | 82.45±4.68 | 20.45±1.95 | 19.80±2.28 | 21.05±1.85 | 21.15±1.92 |
| 3      | 50    | 42    | 84.00  | 86.23±4.32 | 21.65±1.78 | 20.90±2.05 | 21.75±1.72 | 21.93±1.85 |
| 5      | 50    | 45    | 90.00  | 89.17±3.95 | 22.50±1.65 | 22.10±1.82 | 22.30±1.76 | 22.27±1.89 |
| 7      | 50    | 45    | 90.00  | 89.23±3.92 | 22.52±1.63 | 22.15±1.80 | 22.31±1.74 | 22.25±1.87 |

由表4可知,随着迭代次数上限增加,案例可用率和平均分呈显著上升趋势。不启用迭代优化(即不使用质量控制层)的对照组可用率为62.00%(31/50),平均分78.31±5.24分;允许最多1次迭代后,可用率提升至74.00%(37/50),平均分82.45±4.68分;允许最多5次迭代后,可用率提升至90.00%(45/50),平均分89.17±3.95分;进一步增加至最多7次迭代,可用率保持在90.00%(45/50),平均分89.23±3.92分,与5次迭代结果差异极小( $\Delta=0.06$ 分)。统计检验显示,0次迭代与5次迭代差异显著( $t=10.28$ ,  $p<0.001$ ),可用率提升28个百分点( $\chi^2=11.37$ ,  $p<0.001$ ),平均分提升10.86分。但5次迭代与7次迭代无显著差异( $t=0.07$ ,  $p=0.945$ ),可用率相同(均为90.00%),平均分差异仅0.06分,标准差差异仅0.03,表明迭代优化效果在5次左右达到饱和,继续增加迭代次数对质量提升的边际效应极小。因此,本文采用5次作为系统的迭代次数上限,在保证质量的同时兼顾效率。

#### 3.2.3 AIGC辅助教学对学生学习效果的影响

采用准实验设计,对实验组( $n=30$ ,采用AIGC辅助教学)和对照组( $n=30$ ,采用传统教学)进行为期8周的教学干预,结果如图6所示。在控制前测基线差异后,实验组在3个核心维度均显著优于对照组( $p < 0.001$ )。具体而言,知识掌握度方面,实验组提升37.67%(62.27分→85.73分),对照组提升12.58%(61.83分→69.61分),偏效应量 $\eta^2 = 0.52$ ;案例分析能力方面,实验组提升40.58%(58.40分→82.10分),对照组提升6.11%(59.10分→62.71分),偏效应量 $\eta^2 = 0.55$ ;思政认同度方面,实验组提升27.27%(35.20分→44.80分),对照组提升6.54%(35.61→37.94

分), 偏效应量  $\eta^2 = 0.38$ 。三项指标的效应量均达到大效应水平 (Cohen 标准  $> 0.14$ ), 表明 AIGC 辅助教学在促进学生深度学习、提升高阶思维能力和增强价值认同方面具有显著优势。

实验结果表明, AIGC 辅助教学在知识掌握度、案例分析能力和思政认同度 3 个维度可以达到甚至稍优于传统教学模式。此外, 传统教材案例更新周期长, 往往停留在几年前的政策和数据。例如, 某版教材中关于“河长制”的案例仍使用 2018 年的数据, 而实际上河长制在 2020 年后经历了重大改革, 考核办法和责任机制都发生了较大变化。学生学习过时案例后, 发现其与新闻报道中的现实不符, 因而感到理论与现实脱节, 这直接影响了他们的学习动力。AIGC 系统能够及时反映最新政策和时事热点, 系统生成的“河长制改革”案例包含了 2024 年 11 月发布的《进一步加强基层河湖长巡河履职工作指导意见》的最新内容, 这验证了 AIGC 在时效性方面的优势。

同时, 在传统教学中, 教师提供的统一案例限制了学生的思维视角。例如, 讲授“南水北调工程”时, 教材案例通常从工程概况、技术难点和社会效益 3 个角度展开, 所有学生接收相同的信息输入, 思维被框定在既定路径上。然而, AIGC 生成案例支持学生针对同一主题生成多个不同角度的案例进行对比分析。例如, 在研究“南水北调”工程的过程中, 学生生成了多个侧重不同维度的案例。其中, 部分学生生成了侧重工程技术的案例 (详细介绍东线一期调水规模  $87.7 \text{ m}^3$ 、总扬程高度 65 m、13 级泵站等技术参数), 部分学生生成了侧重生态影响的案例 (分析丹江口水库水质保护、汉江下游生态补水、南水北调与长江大保护的关系), 使其形成了“技术—生态—社会—经济”的多维分析框架。

## 4 结语

本文基于提示词工程设计了治水主题思政课案例智能生成系统, 并通过对照实验验证了其教学效果。研究表明, 三级提示词模板设计方法能够有效提升 AIGC 生成案例的质量。通过结构控制层、专业适配层以及质量控制层的协同作用, 案例可用率从初始的 62.00% 提升至 90.00%, 平均分从 78.31 分提升至 89.23 分, 需人工修改率从 38% (不可用案例 19/总生成案例 50) 降至 10% (不可用案例 5/总生成案例 50)。教学对照实验结果, 进一步验证了 AIGC 辅助教学的有效性。实验组学生在知识掌握度、案例分析能力和思政认同度 3 个维度的得分分别为 85.73 分、82.10 分和 44.80 分, 相比对照组的 69.61 分、62.71 分和 37.94 分, 分别提升了 23.15%、30.92% 和 18.08% ( $p < 0.001$ , 大效应)。这表明基于提示词工程的案例生成系统能够显著提升思政课教学效果, 为 AIGC 技术在思政教育领域的应用提供了实证依据。

但本文仍存在一些局限性: ①双引擎协同方案将错误率控制在 3.33% 以内, 但仍有少数案例存在细节错误, 如“都江堰工程”案例中将“鱼嘴”的功能仅描述为“分洪和引水”, 而实际功能还应包括“控制内江进水量”, 这类细节错误需要教师及时发现和纠正; ②本文使用样本仅限于浙江水利水电学院 60 名学生, 且集中在水利类专业, 样本代表性有限, 同时, 教学实验周期为 8 周, AIGC 辅助教学的长期效果尚需进一步跟踪验证。

未来计划从以下几个方面深入研究: ①建立“学生纠错反馈”机制, 形成案例质量持续优化的闭环; ②扩展至图像、视频等多模态内容生成。总体而言, 本文为 AIGC 技术在思政教育领域的应用提供了可行方案, 但仍需在内容质

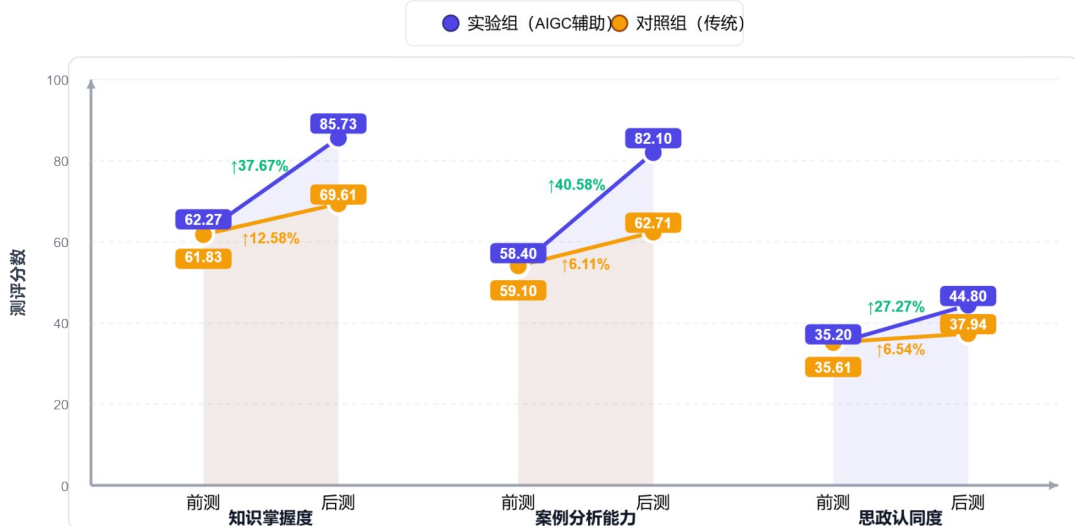


Fig. 6 Evaluation results based on a quasi-experimental design with pre- and post-tests across three dimensions

图 6 基于前测—后测准实验设计的三维度评估结果

量、教学效果和伦理规范等方面持续探索,以推动教育数字化转型的健康发展。

#### 参考文献:

- [1] HOLMES W, MIAO F. Guidance for generative AI in education and research[M]. Paris: UNESCO Publishing, 2023.
- [2] SI L B. Artificial intelligence + education: current situation, challenges and approaches[J]. Governance, 2024 (13): 28-36.  
司林波. "人工智能+教育":现状、挑战与进路[J]. 国家治理, 2024 (13): 28-36.
- [3] MEGAHED F M, CHEN Y J, FERRIS J A, et al. How generative AI models such as ChatGPT can be (mis)used in SPC practice, education, and research? An exploratory study[J]. Quality Engineering, 2023, 35 (4): 643-672.
- [4] HUANG L L, FENG T T, LI M H, et al. Smart education in the context of digital transformation: a blueprint planning and key paths—a summary of the Global Smart Education Conference 2024 [J]. Modern Educational Technology, 2024, 34(11): 36-46.  
黄璐璐, 冯婷婷, 李森浩, 等. 数字转型背景下的智慧教育:蓝图擘画与关键路径——2024全球智慧教育大会综述[J]. 现代教育技术, 2024, 34(11): 36-46.
- [5] FEDERIAKIN D, MOLEROV D, ZLATKIN-TROITSCHANSKAIA O, et al. Prompt engineering as a new 21st century skill[J]. Frontiers in Education, 2024, 9: 1366434.
- [6] NG A. ChatGPT prompt engineering for developers [EB/OL]. <https://www.deeplearning.ai/short-courses/chatgpt-prompt-engineering-for-developers/>.
- [7] LO L S. The CLEAR path: a framework for enhancing information literacy through prompt engineering[J]. The Journal of Academic Librarianship, 2023, 49(4): 102731.
- [8] GUPTA S, MOTLANI R, SALUNKHE S, et al. Enhancing data analysis and programming skills through structured prompt training: the impact of generative AI in engineering education[J]. Computers and Education: Artificial Intelligence, 2025, 7: 100207.
- [9] QIAN Y. Prompt engineering in education: a systematic review of approaches and educational applications[J]. Journal of Medical Internet Research, 2025, 27(1): e56655.
- [10] HUTSON J, COTRONEO P. Generative AI tools in art education: exploring prompt engineering and iterative processes for enhanced creativity [J]. Art Education, 2023, 76(4): 18-25.
- [11] LEE A V Y, TEO C L, TAN S C. Prompt engineering for knowledge creation: using chain-of-thought to support students' improvable ideas[J]. AI, 2024, 5(3): 1446-1461.
- [12] LIU M Z, LI M, YU X H. Breaking through the digital thinking gap: exploring the digital transformation capability of chief information officers [J]. Chinese Journal of ICT in Education, 2024(10): 3-14.  
刘名卓, 李萌, 郝晓华. 突破数字思维鸿沟:首席信息官数字化转型能力探究[J]. 中国教育信息化, 2024(10): 3-14.
- [13] WEN X. New quality and new situation: analysis on the connotative development of ideological theory courses empowered by generative artificial intelligence [J]. Studies in Ideological Education, 2025 (2): 94-101.  
温旭. 新质态与新境遇:生成式人工智能赋能思想政治理论课内涵式发展论析[J]. 思想教育研究, 2025 (2): 94-101.
- [14] SHEN X T. A technical logic review of the integration of artificial intelligence algorithms into ideological and political education [J]. Chinese Journal of ICT in Education, 2024(3): 99-108.  
申晓腾. 人工智能算法融合思想政治教育的技术逻辑审视[J]. 中国教育信息化, 2024(3): 99-108.
- [15] LIU X, ZHAO X, STARKEY H. Ideological and political education in Chinese universities: structures and practices[J]. Asia Pacific Journal of Education, 2023, 43(2): 586-598.
- [16] OUYANG X, WANG Y, LI H. Unmasking the challenges in ideological and political education in China: a thematic review[J]. Heliyon, 2024, 10(8): e29176.
- [17] LI Y, XU J, JIA C Y, et al. Investigation of college students' generative artificial intelligence (GAI) usage status and its implication: taking Zhejiang University as an example[J]. Open Education Research, 2024, 30 (1): 89-98.  
李艳, 许洁, 贾程媛, 等. 大学生生成式人工智能应用现状与思考——基于浙江大学的调查[J]. 开放教育研究, 2024, 30(1): 89-98.
- [18] ZHANG Y, WU R, CHEN Y. AI-supported teaching in China (2014 - 2024) [C]// 2024 International Conference on Intelligent Education and Computer Technology, 2024: 125-132.
- [19] DENG Q Z. ChatGPT empowers Chinese second language teaching: opportunities and challenges [J]. Chinese Journal of ICT in Education, 2024(3): 121-130.  
邓清中. ChatGPT类生成式人工智能赋能汉语二语教学:机遇与挑战[J]. 中国教育信息化, 2024(3): 121-130.
- [20] ZHANG Y B, JIANG L Y, ZHOU J J, et al. Preservice teachers' knowledge and beliefs of generative artificial intelligence in education: the status, issues, and strategies [J]. Chinese Journal of ICT in Education, 2024 (10): 32-40.  
张缨斌, 姜丽滢, 周晶晶, 等. 师范生成式人工智能教育应用认知与理念:现状、问题及对策[J]. 中国教育信息化, 2024(10): 32-40.
- [21] ZHAO C, XU T, YAO Y, et al. Comparison of case-based learning using Watson for oncology and traditional method in teaching undergraduate medical students [J]. International Journal of Medical Informatics, 2023, 177: 105117.
- [22] ZHANG W F, CHEN C, ZHANG T F. Discussion on the application of online case-based teaching in the course of electrical control technology [C]// 2023 35th Chinese Control and Decision Conference (CCDC), 2023: 1934-1937.

(责任编辑:陈维捷)