豆包科研手册

——基于工作任务模式

北京航空航天大学 人文与社会科学高等研究院

清华大学 新闻与传播学院、人工智能学院 双聘教授 @新媒沈阳 团队

何静（副教授）

（提示词受数据样本、测试环境等因素影响，仅用于参考）

工作任务模式：Turbo vs Pro

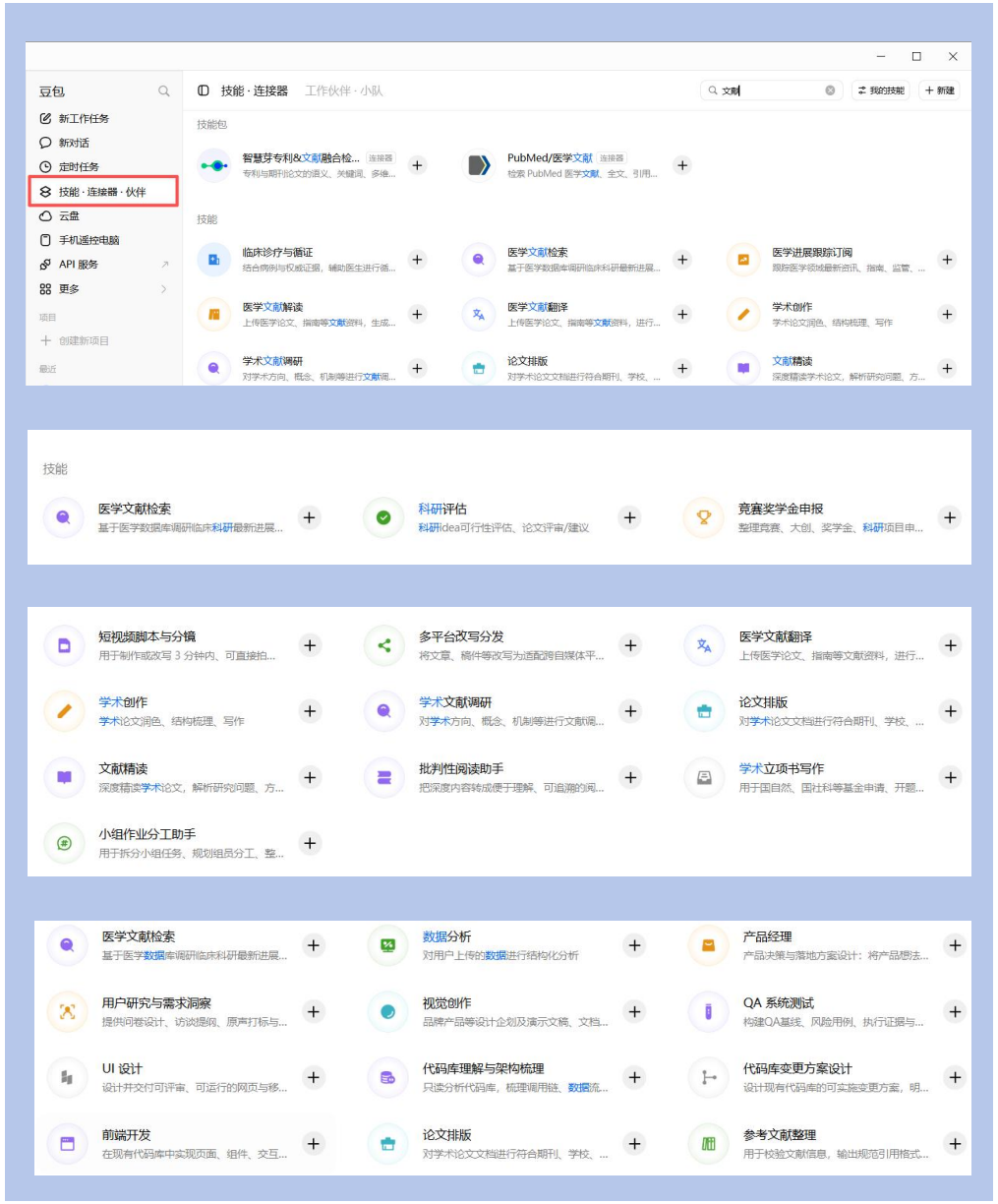
官方把 Seed2.1 定义为一个模型系列，包含 Pro 和 Turbo 两个不同规模的生产力模型。两者都具备 Agent 能力，都能完成多步骤任务、文件处理和工具调用；区别更多体现在复杂专业任务、长程执行和部分高难度基准表现上。

对比维度	Turbo	Pro
核心定位	面向日常、高频、常规任务	面向复杂、专业、高价值任务
任务复杂度	更适合目标明确、步骤较短的任务	更适合开放性强、多步骤、强依赖任务
自主规划	可完成基础任务拆解与执行	更擅长复杂任务规划、阶段推进与动态调整
长链路执行	适合短流程或有限环节连续执行	更适合跨多个阶段持续执行直到最终交付
多文件协同	可处理常规文件与材料	更适合多文件、多格式、多来源信息联合处理
上下文保持	适合一般长度任务	更适合长文档、长任务和跨阶段信息保持
工具调用	支持文件、浏览器、办公工具等调用	更适合多工具组合调用与复杂流程协同
额度消耗	相对更低，适合高频日常使用	相对更高，复杂任务通常消耗更多额度
效率倾向	更强调速度、频率与性价比	更强调深度、质量与复杂任务完成度

Turbo更适合高频、低消耗地完成常规任务；Pro更适合用更高额度换取复杂任务的深度执行与完整交付。

豆包科研技能全景

科研任务	核心功能	代表技能	适用科研场景
文献调研	围绕研究方向、概念与机制开展系统资料调研	学术文献调研	选题调研、研究现状梳理、研究空白识别
文献精读	深度解析研究问题、方法、结果与结论	文献精读、医学文献解读	单篇论文精读、方法学习、证据提取
文献翻译	对专业论文和学术材料进行中英文转换	医学文献翻译	外文论文阅读、国际文献使用
循证研究	结合病例与权威证据辅助循证判断	临床诊疗与循证	医学循证研究、临床证据整理
学术写作	论文润色、结构梳理与学术内容生成	学术创作	论文初稿、章节撰写、语言润色
科研评审	判断科研选题可行性并提供论文评审建议	科研评估	选题评价、论文自审、研究方案诊断
课题申报	生成科研项目立项与申请材料	学术立项书写作	基金申报、课题申请、开题材料
数据分析	对科研数据进行结构化处理与分析	数据分析	数据清洗、统计分析、结果提取
科研可视化	生成和编辑视觉内容、辅助科研结果表达	视觉创作	科研示意图、研究框架图、汇报视觉
论文规范化	处理论文格式、文献引用及真实性核查	论文排版、参考文献整理、参考文献审计	投稿排版、引用规范、参考文献核验
专利成果	辅助形成专利申请文件	专利文件申请	发明专利申请、科研成果转化



豆包工作任务模式可根据任务需求自动调用相关技能，完成多步骤科研任务。

开题报告——技能辅助规范生成

Prompt + 技能 | 学术立项书写作

学术立项书写作
用于国自然、国社科等基金申请、开题...

提示词公式

请围绕【研究主题】，调用“学术立项书写作”技能，生成一份完整的开题报告初稿。具体要求如下：

1. 明确研究背景、研究问题与研究意义；
2. 梳理国内外相关研究现状，并总结已有研究的主要进展与不足；
3. 在研究现状基础上凝练本研究的核心问题和研究目标；
4. 设计整体研究内容和研究框架，确保各部分之间逻辑衔接；
5. 根据研究问题选择合适的研究方法，并说明每种方法解决什么问题；
6. 明确研究重点、难点及可能的突破方向；
7. 提炼具有依据的创新点，避免使用“首次”“填补空白”等无法证明的表述；
8. 设计研究技术路线与实施步骤；
9. 制定阶段性研究计划和预期成果；
10. 对全文进行一致性检查，确保“研究问题—研究目标—研究内容—研究方法—创新点”相互对应；
11. 不得虚构已有成果、实验数据、项目基础或个人经历，缺失部分请标注“待补充”；
12. 最终输出一份结构完整、可继续修改的开题报告初稿。

技能把开题报告的常用结构、写作步骤和交付规范提前固化，Prompt只需补充研究主题和个性化要求。

明确硕士开题报告核心主题 >

我来帮你完成这份硕士开题报告。先加载学术立项书写作相关技能，并确认输出路径。

确认文件夹并准备读取资料 >

文件夹已确认存在。现在按技能流程读取开题报告场景所需的参考资料。

读取相关文档完善写作准备 >

继续读取剩余参考资料，并同步开展文献检索。

补充元认知与AI依赖文献 >

文献检索已获得丰富素材。再补充检索元认知调节与 AI 依赖测量相关文献。

整合文献素材筹备开题报告 >

文献素材已充分。现在创建工作目录并撰写完整开题报告。

启动硕士开题报告正文撰写 >

目录已就绪。现在撰写完整的开题报告正文。

一、选题背景与研究意义

（一）现实背景：生成式人工智能进入高校学习场景

生成式人工智能（Generative·Artificial·Intelligence，GenAI）自 2022 年

三、研究问题与研究目标

（一）核心概念界定

1. 生成式人工智能使用方式（GenAI·Usage·Patterns）

本研究中的 GenAI 使用方式指大学生在学习场景中使用生成式 AI 工具（如 ChatGPT、文心一言、通义千问、Kimi 等大语言模型及相关 AI 辅助学习工具）的具体模式和目的。基于文献梳理，本研究将其划分为两种典型类型：

- **策略性使用（Strategic·Use）**：将 AI 作为思维脚手架，用于提出问题、获取多元视角、对比观点、反思自身推理、拓展知识边界等，学生在使用过程中保持主动思考和批判性评估。
- **替代性使用（Substitutive·Use）**：将 AI 作为答案生成器，用于直接获取作业答案、代写论文、替代推理过程、跳过分析环节等，学生在使用过程中减少或放弃主动思考。

2. AI 学习依赖（AI·Learning·Dependency）

本研究中的 AI 学习依赖指大学生在学习过程中对 GenAI 工具形成的心理和行为上的过度依赖状态，表现为离开 AI 则难以完成学习任务、对 AI 输出不加批判地接受、将认知思考外包给 AI 等。借鉴 AIDep-22 量表的维度结构，本研究将 AI 学习依赖操作化为四个维度：

- **情感依赖**：对 AI 产生情感上的依恋和信任，离开 AI 感到焦虑或不安。
- **功能依赖**：学习任务的完成高度依赖 AI 提供的功能支持，无法独立完成。
- **认知依赖**：将分析、推理、判断等认知过程外包给 AI，减少自身认知投入。
- **失控感**：感到自己无法控制 AI 使用行为，使用频率和时长超出预期。

硕士学位论文开题报告

论文题目：生成式人工智能对大学生批判性思维的影响机制研究

——基于 AI 学习依赖的中介作用与元认知的调节效应

研究生姓名：【待补充：请填写研究生姓名】

学号：【待补充：请填写学号】

学院：【待补充：请填写所在学院】

专业：【待补充：请填写专业名称】

研究方向：【待补充：请填写研究方向】

指导教师：【待补充：请填写指导教师姓名及职称】

开题日期：2026 年【待补充：请填写具体月份】月

二、国内外研究现状述评

（一）生成式人工智能与高等教育学习

GenAI 进入高等教育领域后，相关研究呈现快速增长态势。国外研究方面，一项基于 89 篇同行评审论文（2024—2026 年）的系统综述指出，GenAI 对大学生高阶认知技能的影响呈现“放大器或替代者”的双重面貌：当学生将 AI 用于反思、提问和观点对比时，其批判性思维倾向得到增强；而当 AI 被用于直接获取答案、替代推理过程时，则可能导致认知技能萎缩

["https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.

（二）研究问题

基于上述背景和文献分析，本研究提出以下三个核心研究问题：

研究问题 1：大学生 GenAI 使用方式（策略性使用·思维的影响是否存在差异？具体而言，策略性使用是否i性使用是否负向预测批判性思维？

研究问题 2：AI 学习依赖在 GenAI 使用方式与批判·作用？即替代性使用是否通过增强 AI 学习依赖进而削弱是否通过降低 AI 学习依赖（或不依赖）进而促进批判性

研究问题 3：元认知在“AI 使用方式—AI 学习依赖-径中是否发挥调节作用？具体而言，元认知是否调节（a依赖的影响，以及/或者（b）AI 学习依赖对批判性思维否比低元认知个体更能抵御 AI 依赖对批判性思维的侵蚀？

（三）研究意义

1. 理论意义

第一，本研究构建“AI 使用方式—AI 学习依赖—批判性思维”的中介模型，并引入元认知作为调节变量，有助于揭示 GenAI 影响大学生批判性思维的内在心理机制，弥补现有研究多停留在相关性描述、缺乏机制解释的不足。第二，本研究将 AI 学习依赖从泛化的“技术依赖”概念中剥离，聚焦学习场景下的情感依赖、功能依赖、认知依赖和失控感四个维度，丰富了教育情境下 AI 依赖的理论内涵。第三，本研究将元认知理论引入 AI 学习研究领域，探讨元认知监控与调节能力如何缓冲 AI 依赖对批判性思维的侵蚀，为理解“AI 使用—认知发展”关系提供新的理论视角。

2. 实践意义

第一，研究结果可为高校制定 GenAI 使用规范和教育干预策略提供实证依据，帮助教育者区分“有益使用”与“有害依赖”，引导学生形成健康的 AI 学习习惯。第二，通过识别元认知的保护作用，本研究可为高校批判性思维培养课程和 AI 素养教育提供设计参考，将元认知训练融入 AI 时代的人才培养体系。第三，研究成果可为教师调整教学评价方式、设计 AI 融合的教学活动提供指导，促进 GenAI 从“替代思考的工具”向“促进思考的脚手架”转变。

（三）变量关系设计

本研究涉及四类变量，其操作化定义和测量方式如下：

变量类型	变量名称	维度	测量方式
自变量	GenAI 使用方式	策略性使用、替代性使用	自编/修订量表，Likert·5 点计分
中介变量	AI 学习依赖	情感依赖、功能依赖、认知依赖、失控感	AIDep-22 量表（学习场景适配版），Likert·5 点计分
结果变量	批判性思维	认知技能、情感特质	加利福尼亚批判性思维量表（CCTDI/CCTST）中文版或修订版
调节变量	元认知	元认知知识、元认知监控、元认知调节	元认知意识量表（MAI）中文版或学习过程元认知量表
控制变量	性别、年级、学科、院校层次、AI 使用频率、AI 使用时长	—	人口学问卷

科研计算——连接器辅助独立核验

Prompt + 连接器 | Wolfram

Wolfram 连接器
Wolfram 知识库查询、符号计算与精...

提示词公式

请对我提供的【论文/数据结果/公式】进行计算核验，并调用 Wolfram 连接器 对可以独立计算的内容进行重新验证。具体要求如下：

1. 识别文中所有可以通过公式或数值重新计算的关键结果；
2. 提取【原始数据/参数/公式/统计量】；
3. 调用 Wolfram 对关键计算步骤进行独立计算；
4. 核验【均值、标准差、比例、置信区间、效应量、统计量、概率、函数结果等】是否正确；
5. 如包含数学公式，检查公式推导、变量定义和计算逻辑是否一致；
6. 如包含统计模型，核验能够独立复算的指标，但不要凭空补充原文未提供的数据；
7. 将原文结果与重新计算结果逐项对比；
8. 对存在差异的项目说明差异大小及可能原因；
9. 区分【计算错误】【四舍五入差异】【参数口径差异】【信息不足无法核验】；
10. 最终输出【核验项目—原结果—重新计算结果—是否一致—问题说明】对照表。

借助 Wolfram 连接器，可对论文公式和关键统计结果进行独立复算与交叉核验。

读取论文分析结果文档 >

我来读取 Wolfram 技能说明和你上传的论文文档，然后进行统计核验。

读取文档与技能说明并规划核验步骤 >

文档内容已读取。现在获取 Wolfram 计算工具的接口定义，然后逐项核验。

获取Wolfram工具调用规范 >

工具接口已获取。现在用 Wolfram 一次性重新计算所有核验项目。

启动Welch t检验结果对比流程 >

Wolfram 重新计算全部完成。我发现了多处计算错误，现在整理完整的核验对照表。

统计核验结果总览

共核验 **6 大类、18 项** 统计指标，其中 **5 项存在错误**，13 项正确。以下是逐项对照表：

核验对照表

表格

发消息或创建任务... / 使用技能 @ 添加资料

工作任务 | 按需确认 | 技能 | 连接器

12	回归 t 值	−4.43	−4.432	✅ 一致	—
13	回归 95% CI	[−4.31, −2.25]	[−4.73, −1.83]（正态）/ [−4.75, −1.81]（t,df=120）	❌ 错误	CI 半宽仅 1.03，对应临界值≈1.39（约 83% 置信水平），非 95%
14	回归 p 值	0.021	<0.001（正态 p=9.3×10⁻⁶；t,df=120 时 p=2.1×10⁻⁵）	❌ 严重错误	t=−4.43 对应 p 应远小于 0.001；p=0.021 对应 t≈±2.3，两者矛盾
15	比例值	68.8%	68.8%	✅ 一致	—
16	比例 95% CI	[61.2%, 75.5%]	[60.7%, 76.9%]（Wald）/ [60.2%, 76.3%]（Wilson）	❌ 错误	区间偏窄，半宽约 7.2%，而标准误 4.14%×1.96=8.1%
17	ANOVA F 值	5.42	6.053	❌ 错误	210÷34.69=6.053，非 5.42
18	ANOVA η^2	0.076	0.0761	✅ 一致	—

错误 1：Cohen's d 计算错误（第三节，P8）

- 论文原文：`Cohen's d = (78.4 − 74.1) / 8.90 = 0.61`
- 实际计算：4.3 ÷ 8.90 = **0.483**
- 影响：论文将效应解释为 "中等效应"，d=0.48 按 Cohen 标准（0.2 小 / 0.5 中 / 0.8 大）应解释为**"小到中等效应"**，而非 "中等效应"
- 建议修改值：**d = 0.48，解释为 "小到中等效应"**

核验对照表

表格

序号	核验项目	论文原结果	Wolfram 重新计算	是否一致	差异原因
1	Welch t 均值差	4.30	4.30	✅ 一致	—
2	Welch t 标准误 SE	1.595	1.595	✅ 一致	—
3	Welch t 值	2.696	2.696	✅ 一致	—
4	Welch 自由度 df	121.38	121.38	✅ 一致	—
5	Welch 双侧 p 值	0.0081	0.0080	⚠ 基本一致	四舍五入精度差异（第 4 位小数）
6	Welch 95% CI	[1.14, 7.46]	[1.14, 7.46]	✅ 一致	—
7	**合并标准差 Sp**	**8.90**	**8.898**	✅ 一致	—
8	**Cohen's d**	**0.61**	**0.483**	❌ **错误**	算术错误：4.3÷8.90=0.483，非 0.61

错误 2：回归系数 95% 置信区间错误（第五节，Table 3）

- 论文原文：95% CI = [−4.31, −2.25]
- 实际计算（B=−3.28, SE=0.74）：
  - 正态近似：−3.28 ± 1.96×0.74 = **[−4.73, −1.83]**
  - t 分布（假设 df=120）：−3.28 ± 1.98×0.74 = **[−4.75, −1.81]**
- 论文 CI 半宽仅 1.03，对应临界值约 1.39，相当于约 83% 置信水平，不是 95%
- 建议修改值：**[−4.73, −1.83]**（或根据实际 df 用 t 分布）

错误 3：回归 p 值与 t 值严重矛盾（第五节，Table 3）

- 论文原文：t = −4.43，p = 0.021
- 实际计算：t = −4.43 时，无论正态近似（p = 9.3×10⁻⁶）还是 t 分布（df=120 时 p = 2.1×10⁻⁵），p 值均**远小于 0.001**
- p = 0.021 对应的 t 值约为 ±2.33（df=120 时），与报告的 t=−4.43 完全不匹配
- 建议修改值：**p < 0.001**（或更精确的 p = 0.00002，取决于实际 df）
- 注：论文未报告回归自由度，若实际样本量或预测变量数不同，df 会变化，但 t=−4.43 在任何合理 df 下 p 都不会是 0.021

交叉核对发现

1. **正文与表格一致性**：Welch t 检验部分正文（P5）表述 "显著高于" 与表格 p=0.008 一致，无矛盾。
2. **效应量解释一致性**：Cohen's d=0.61 被解释为 "中等效应"，但正确值 d=0.48 应解释为 "小到中等效应"——**解释需随数值同步修改。**
3. **ANOVA η^2 与 F 值**：η^2=SS 组间 / SS 总 = 420/5520=0.076 正确，但 F 值错误导致两者不匹配（F=6.05 对应 η^2=0.076 是自洽的，F=5.42 则不对应）。

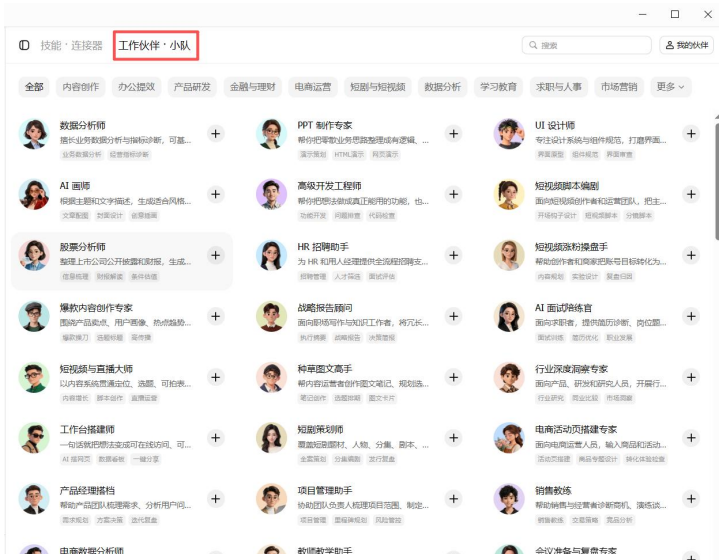
研究方案——工作伙伴协同设计

Prompt + 工作伙伴 | ① 论文写作助手 ② 数据分析师 ③ 一图总结大师 ④ 项目管理助手

提示词公式

请围绕【研究主题】，由我选择的多个工作伙伴协同完成一套完整研究方案。请充分发挥不同工作伙伴的专业能力，不要由一个伙伴包办全部任务。

1. **论文写作助手**负责梳理研究问题、理论逻辑、研究假设和整体框架；
2. **数据分析师**负责设计变量、数据获取方案和统计分析方法；
3. **一图总结大师**负责将最终研究逻辑转化为清晰的研究框架图或技术路线图；
4. **项目管理助手**负责制定研究实施步骤、阶段计划、关键节点和风险应对方案。
5. 各伙伴完成自己的部分后，请对不同方案之间的一致性进行检查，最终整合为一套完整、可执行的研究方案。











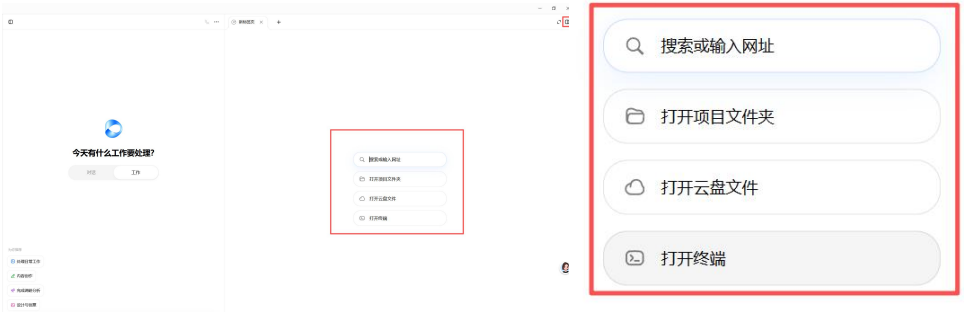


不是一个Agent包办，而是多个专业伙伴像科研团队一样分工协作。

科研协作——侧边工作台多源联动

提示词公式

请结合侧边工作台中当前打开的【对象1】和【对象2】，围绕【当前科研任务】进行联动分析和处理。具体要求如下：
1. 识别对象：分别读取【对象1】和【对象2】的核心内容，明确各自提供了哪些关键信息；
2. 提取重点：围绕【当前科研任务】提取最相关的内容、数据、要求或观点；
3. 交叉比较：对两个对象中的信息进行对应、比较和核对，识别一致点、差异点和潜在冲突；
4. 任务判断：根据比较结果，对【适配性/一致性/完整性/合理性/可行性】等目标作出判断；
5. 问题定位：如发现缺失、矛盾或需要进一步确认的内容，指出具体位置和原因；
6. 提出处理建议：给出针对性的修改、补充、调整或下一步操作建议；
7. 持续联动：如有必要，可继续查看侧边工作台中的相关网页、云文档或项目文件，不需要我重复复制或上传；
8. 最终输出：【对象概况】+【关键信息】+【比较结果】+【问题判断】+【处理建议】。



三、匹配度等级：中等

四、判断依据

1. 主题匹配：中等偏高 ✅

- **契合**：AIGC 治理是该刊近期高频选题。2023—2025 年已连续发表 ChatGPT 治理、生成式 AI 政府治理、算法治理、AI 治理体系构建、生成式 AI 个人信息规范等多篇文章，2025 年第 2 期专设 "人工智能时代下信息技术治理规则" 专题。论文主题在**大方向上与该刊热点高度吻合**。
- **偏差**：论文聚焦 "网络视听内容精品化" 这一**传播学 / 文化产业**细分领域，而该刊已发表的 AI 治理文章多从**政府治理、法学、公共政策**视角切入，尚无专门针对网络视听内容生产的选题。

2. 研究问题与方法：中等 ⚠️

- **契合**：论文采用政策文本分析与理论框架构建，与该刊已发表的《生成式 AI 嵌入政府治理》（"技术 — 制度 — 价值" 框架）、《中国 AI 治理体系构建》（历史制度分析）风格接近，均为**定性理论研究**。
- **偏差**：该刊部分治理类文章引入量化方法（如 DEA 模型、政策文本量化分析、社会网络分析），论文目前以纯理论论述为主，**缺乏实证数据支撑**，说服力可能偏弱。

3. 相近选题：存在但不重叠 ✅

表格

该刊已发表	与你的论文关系
《生成式 AI 嵌入政府治理的应用前景、潜在风险和防范机制》(2024)	同为 AIGC + 治理，框架相似产
《算法治理的整合型沟通架构》(2025)	同为算法 / AI 治理，关注制度
《中国 AI 治理体系构建的需求、经验与路径》(2024)	宏观治理体系，论文是细分领域
《生成式 AI 个人信息利用的规范体系》(2025)	法学视角，论文是传播学 / 治理

结论：AIGC 治理主题已被该刊多次接纳，但 "网络视听内容精品化 + 可信机制" 这一具体切入点**尚未见发表**，具备一定新颖性。

五、投稿建议

方案 A：调整后投该刊（匹配度可提升至中高）

题目调整建议（突出治理视角，弱化传播学色彩）：

- 《专项治理背景下 AIGC 内容生产的可信治理机制研究 —— 以网络视听领域为例》
- 《生成式 AI 内容规模化生产的质量治理：技术逻辑、供给结构与可信底线》

内容调整方向：

1. **强化治理理论对话**：将 "精品化" 重构为 "质量治理" 或 "内容治理目标"，增加与公共治理、平台治理理论的对话，而非仅停留在传播学层面
2. **深化政策文本分析**：系统梳理国家广电总局、网信办关于 AIGC 视听内容的专项治理政策（如 "AI 魔改" 治理、深度合成规定等），增强政策分析厚度
3. **突出理论贡献**：将 "可验证可信" 提炼为一种治理机制创新，与该刊已发表的 "技术 — 制度 — 价值" 框架形成对话和推进
4. **补充实证元素**：可增加典型平台案例的对比分析（如抖音 / 快手 / B 站的 AIGC 内容治理规则差异），或政策文本的编码统计，提升研究的实证性
5. **投稿栏目**：优先投 "公共政策与治理" 或 "跨学科研究热点" 栏目

方案 B：推荐更匹配的期刊（若保持现有传播学视角）

表格

期刊	匹配理由
《现代传播（中国传媒大学学报）》	CSSCI，传媒类顶刊，AIGC + 视听内容是核心选题
《新闻与写作》	CSSCI 扩展，关注新媒体内容生产与治理
《出版广角》	北大核心，关注数字内容产业与治理
《中国编辑》	CSSCI 扩展，关注内容质量与编辑规范
《电子政务》	CSSCI，关注数字治理与平台治理，若强化治理视角可投

总结

表格

项目	结论
期刊定位	工科高校综合性社科期刊，以公共政策 / 法学 / 治理为主，AI 治理为近期热点
论文定位	AIGC 网络视听内容治理，传播学 + 文化治理交叉
匹配度	**中等**
核心优势	AIGC 治理是该刊热点，理论框架风格契合，切入点有新颖性
核心短板	传播学视角与该刊主流的政治学 / 法学 / 公管视角有偏差，缺乏实证支撑
建议	若投该刊，需将题目和内容向 "治理" 视角调整，强化政策分析和理论贡献；若保持传播学视角，建议投传媒类期刊

网页、项目文件和云文档同屏联动，让科研任务不再依赖反复上传和复制粘贴。

选题生成

测试模式：工作任务 Turbo

豆包工作任务Turbo模式可将选题从“一次问答”**升级为“完整调研任务”**，**自主完成**多源文献检索、研究空白识别、候选方向比较与优先选题推荐，并**自动沉淀为结构化研究文档**，并支持多种格式下载。

提示词公式

我的研究方向是【研究方向】，希望围绕【具体关注领域/研究对象】寻找具有一定创新性、研究价值和可操作性的论文选题，目标成果为【硕士论文/博士论文/CSSCI论文/SCI论文/基金项目等】。请使用工作任务模式开展选题研究。你可以根据任务需要自行规划研究步骤、检索相关资料并进行分析，不需要局限于我预设的检索路径。

最终请给出【3—5】个值得进一步研究的候选选题，并分别说明：

1. 选题名称；2.该选题基于什么已有研究或现实问题提出；3.当前研究还存在哪个具体缺口；4.可以重点回答什么研究问题；5.可以采用哪些数据和研究方法；6.选题的创新性与实际可行性；7.可能存在的研究风险。

最后对候选选题进行综合比较，并推荐一个最值得优先开展的选题。所有涉及已有研究和研究趋势的判断均需提供可核验的信息来源，不确定的信息请明确说明，不要编造文献。



问问豆包 用户870097 › 生成式人工智能进入高校教学后的新问题：CSSCI期刊论文选题研究

最近修改：17 分钟前

任务已完成（40m41s）

已完成 "生成式人工智能进入高校教学后产生的新问题" 的 CSSCI 期刊论文选题调研，共纳入 42 篇经文件），最终给出 5 个候选选题并推荐 1 个优先方向。

核心发现

生成式 AI 进入高校教学后，研究界已从早期 "机遇 - 挑战" 二元框架转向对**制度性、认知性和劳动性**研究明显不足。学术诚信领域研究最为成熟，教师劳动和学生认知领域争议最突出，教育公平与高校治究稀缺。

5 个候选选题概览

表格

序号	选题方向	核心切口	方法取向
1	评估范式重构	从 "检测 - 惩罚" 到 "过程 - 表现" 评估的中国高校实践	多案例研
2	教师劳动重构	GenAI 背景下教学工作量隐性增加与劳动异化	混合方法谈）
3	**认知外包与深度学习**	GenAI 使用对批判性思维与元认知能力的影响机制	准实验 + 纵向追踪
4	新数字鸿沟	"双一流" 与地方高校间 AI 使用能力的分层效应	大规模问卷 + 结构方程模型
5	高校治理适配	GenAI 使用政策的制定逻辑、执行偏差与治理框架	政策文本分析 + 多案例比较

推荐优先方向

选题三：认知外包的代价 —— 生成式 AI 使用对大学生批判性思维与元认知能力的影响机制研究

推荐理由：该选题处于 "争议最大 × 证据最弱 × 中国情境空白" 的交叉点，现有研究多为横断面调查且集中于西方学生群体，缺乏长期追踪和学科差异分析；准实验 + 纵向追踪的设计可产出因果性证据，符合 CSSCI 教育技术类期刊对实证研究的偏好；同时与您的 AI 与高等教育研究方向高度契合，数据可及性和伦理门槛相对可控。

详细的 9 维度展开、综合比较表、文献多维地图、核心逻辑图和完整参考文献均已整理在飞书文档中。

生成式人工智能进入高校教学后...

创建时间：14:03

生成式人工智能进入高校教学后的新问题：CSSCI期刊论文选题研究

一、核心结论

二、研究范围与方法

十二、参考文献

生成式人工智能进入高校教学后的新问题：CSSCI期刊论文选题研究

用户870097 今天修改

部分内容由豆包生成

一、核心结论

总体判断：生成式人工智能进入高校教学后，研究界已从早期的"机遇与挑战"二元框架转向对制度性、认知性和劳动性新问题的深入考察，但中国情境的实证研究明显不足。现有研究在学术诚信领域最为成熟，在教师劳动和学生认知领域存在显著争议，在教育公平和高校治理领域的政策讨论较多但微观执行研究稀缺。这些缺口共同指向一个核心问题：在中国高等教育的特定制度环境下，GenAI引发的教学变革如何与现有的评估体系、绩效制度、分层结构和治理框架发生互动，以及这种互动产生了哪些具有理论价值和政策意义的新问题。

1. 学术诚信问题已从"检测-惩罚"的技术应对转向评估范式重构，全球顶尖高校普遍开始修订评估指南，但过程性评估在中国高校的落地机制尚未被系统研究（Moorhouse等，2024；Kolade等，2024）。

2. 教师劳动正在经历隐性重构：AI既可能减轻重复性劳动，也可能因课程 redesign、评估创新和AI素养培训增加隐性工作量，中国高校绩效制度下的劳动异化风险被低估（宋霍和林敏，2023；侯浩娟和王旦，2025）。

3. 生成式AI作为认知代理可能导致学生批判性思维和元认知能力的认知外包风险，现有研究多为横断面调查，缺乏长期追踪和学科差异分析（Roy and Jony，2024；Qian，2025）。

4. 教育公平呈现新形态：从硬件接入鸿沟转向AI使用能力鸿沟和算法偏见，双一流与地方高校间的分层效应尚未被实证检验（刘三女牙和郝晓鲍，2024；UNESCO，2023）。

5. 高校治理面临制度适配挑战：中国政策从试点示范走向全面推进，但高校层面的政策制定逻辑、执行偏差和治理框架建构缺乏系统研究（Chan，2023；教育部，2024）。

二、研究范围与方法

十二、参考文献

O'Dea, C. Generative AI: is it a paradigm shift for higher education? Studies in Higher Education, 2024. https://www.tandfonline.com/doi/pdf/10.1080/03075079.2024.2332944

Giannakos, M. et al. The promise and challenges of generative AI in education. Behaviour & Information Technology, 2024. https://oro.open.ac.uk/99236/9/99236EA.pdf

Jensen, E. et al. Generative AI and higher education: a review of claims from the first months of ChatGPT. Higher Education, 2024. https://link.springer.com/content/pdf/10.1007/s10734-024-01265-3.pdf

杨俊锋. 生成式人工智能与高等教育深度融合：场景、风险及建议. 中国高等教育, 2024. https://wap.cnki.net/touch/web/Journal/Article/ZGDJ202405029.html

李今春. ChatGPT的智慧生成特征及对高等教育的挑战. 江苏高教, 2023. ://m.qikan.cqvip.com/Article/ArticleDetail?id=7110216888

ide, O. et al. Is AI changing learning and assessment as we know it? Heliyon, 2024. s://www.cell.com/heliyon/pdf/S2405-8440(24)01984-4.pdf

rhouse, B. et al. Generative AI tools and assessment: Guidelines of the world's top-ranking versities. Computers and Education Open, 2024. s://scholars.cityu.edu.hk/files/272161619/268445269.pdf

nas, A. et al. The impact of generative AI on academic integrity of authentic assessments. sh Journal of Educational Technology, 2024. https://bera-nals.onlinelibrary.wiley.com/doi/pdfdirect/10.1111/bjet.13585

tzer, T. et al. Can you spot the bot? Identifying AI-generated writing in college essays. rnational Journal for Educational Integrity, 2024. s://edintegrity.biomedcentral.com/counter/pdf/10.1007/s40979-024-00158-3.pdf

ins, K. Academic Integrity considerations of AI Large Language Models in the post-pandemic CORE, 2023. https://core.ac.uk/download/554156096.pdf

on, D. et al. Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. ovations in Education and Teaching International, 2023. s://core.ac.uk/download/564450553.pdf

i, 林敏. ChatGPT/生成式人工智能时代下教师的工作变革：机遇、挑战与应对. 华东师范大学学报 育科学版），2023. https://xbjk.ecnu.edu.cn/CN/article/downloadArticleFile.do?chType=PDF&id=10976

文献检索--基于主题

测试模式：工作任务 Turbo

提示词公式

请围绕【研究主题/研究问题】开展系统性的文献检索，目标是帮助我快速了解该领域的研究现状，并建立可用于后续研究的核心文献池。
请根据研究主题自行判断检索思路、关键词组合、资料来源与筛选方式，并完成必要的多轮检索与交叉核验。
检索要求：时间范围限定为【起始年份—截止年份】；优先选择与研究主题高度相关的高质量【期刊论文/综述/会议论文等】，兼顾代表性研究与最新研究；所有文献必须真实可核验，作者、标题、期刊、年份、DOI等信息不得编造；对检索结果进行去重和相关性筛选，不要简单堆砌文献；
如某项文献信息无法可靠核验，请明确标注。
最终请输出：

1. 核心文献列表；
2. 文献的主要研究方向分类；
3. 各类代表性研究及核心发现；
4. 当前研究中的主要共识、争议与不足；
5. 最值得继续阅读的【数量】篇核心文献及推荐理由。
6. 请保留可追溯的文献来源和DOI/原文链接，方便后续核验。

■ 适用场景：从研究主题出发，系统建立文献池。

Plan literature review for CSSCI paper >

我先加载学术文献调研技能，然后按其流程开展系统性检索。

Identify research task complexity >

这是一个典型的学术深度综述任务，且用户明确要求 DOI 与可核验题录，属于**题录严格模式**。自行串行执行：检索 → 题录核验 → 综合 → 交付编译。先建立任务清单并读取子技能细则。

Acquire scholar_search tool >

子技能细则已读取。现在初始化工作目录并读取核心脚本（台账 schema 与题录核验脚本）。

生成式人工智能对大学生批判性思维与元认知能力的影响：系统性文献综述与核心文

> 检索时间范围：2022—2026 年 | 检索完成日期：2026 年 8 月 13 日
> 文献核验方式：Crossref API 题录核验 + 期刊官网 / 数据库网页交叉核验
> 核心文献池规模：29 篇（英文 24 篇 + 中文 5 篇），其中 22 篇经 Crossref 精确核验，7 篇经期刊

一、检索方法与文献筛选说明

1.1 检索策略

围绕研究主题，从五个互补视角构建检索式，在学术数据库（Google Scholar 语义检索、Crossref、W
及中文数据库（知网、维普、期刊官网）进行多轮检索：

表格

视角	核心检索词组合
正面效应	generative AI / ChatGPT + critical thinking + college students + empirical / expe
负面风险	cognitive offloading / dependency / metacognitive laziness + university students
机制调节	self-regulated learning / AI literacy / prompt engineering + mediating / moderating
方法综述	meta-analysis / systematic review + large language models + higher education
中文情境	生成式人工智能 + 大学生 + 批判性思维 / 元认知 + 实证研究 + CSSCI

1.2 纳入与排除标准

- **纳入**：2022—2026 年发表；大学生 / 高等教育情境；实证研究（定量 / 定性 / 混合）、系统综述 / 元分析、高质量理论框架；中英文文献。
- **排除**：纯技术实现论文、K12 为主的研究、无直接关联的泛 AI 教育讨论、会议摘要 / 短文。
- **特别处理**：Wang & Fan (2025) 发表于 *Humanities and Social Sciences Communications* 的元分析经 Crossref 核验确认已于 2026 年 4 月被撤稿（Retraction Note, DOI: 10.1057/s41599-026-07310-z），本综述不将其纳入核心文献池，其报告的效应量数据不予引用。



豆包可**自主拆解研究主题并制定多维检索策略**，开展中英文多来源文献检索，同时**完成相关性筛选、真实性核验与撤稿识别**，最终形成质量较高、可直接用于后续研究的核心文献池。

文献检索--基于段落

测试模式：工作任务 Turbo

■ 适用场景：从已有论文观点反向寻找能够支撑它的真实文献。

提示词公式

请依据我提供的【论文段落】，为其中需要文献支撑的核心论断检索和匹配真实参考文献。
请使用工作任务模式，先理解段落的论证逻辑，自行识别其中需要引用支持的关键论断，再针对不同论断分别开展文献检索，而不是仅根据整段文字的主题推荐一组泛相关文献。要求：

1. 将段落拆分为若干需要证据支持的核心论断；
2. 判断每个论断需要何种类型的文献证据；
3. 为每个核心论断寻找高度匹配的真实文献；
4. 优先选择【时间范围】发表的高质量研究，经典文献确有必要时可以保留；
5. 核验作者、标题、期刊、年份和DOI等基本信息；

不得为了给每句话配引用而推荐相关性不足的文献；
如果某一论断暂时找不到充分证据，请直接指出，不要强行匹配。
最终请按照“【原文论断—推荐文献—文献核心发现—支撑程度—建议引用位置】”形成对应关系，并提供DOI或可核验来源。

豆包能够先将段落拆解为具体论断，再针对每个论断匹配对应文献、核心证据与支撑强度，实现从“按主题找文献”到“按论点精准配证据”的检索。



文献检索--基于种子文献

测试模式：工作任务 Turbo

提示词公式

请以我提供的【种子文献/上传论文】为起点，对其相关研究进行扩展检索，帮助我建立围绕该论文的核心文献网络。
请使用工作任务模式，首先理解种子文献的研究问题、理论基础、方法与核心结论，再自行规划文献扩展路径，不要仅按照论文标题关键词寻找相似文章。文献扩展至少需要考虑：
前序研究：种子文献建立理论、概念或方法基础时依赖重要研究；
后续研究：发表后引用、延伸、验证或应用该研究的重要文献；
同类研究：研究问题、研究对象或方法高度相近的代表性研究；
争议研究：结论与种子文献存在差异、质疑或相反证据的研究；
最新进展：该研究方向近【时间范围】出现的新问题、新方法或新发现。所有推荐文献必须真实可核验，并核查作者、标题、期刊、年份、DOI以及与种子文献之间的实际关系，不得仅凭标题推测引用关系。
最终请形成一份以种子文献为中心的文献扩展结果，分别列出前序研究、后续研究、同类研究、争议研究和最新进展，并说明每篇文献为什么值得继续阅读。最后总结该种子文献在整个研究脉络中的位置，以及基于现有文献还可以继续研究哪些问题。

■ 适用场景：以核心文献为起点，扩展相关研究脉络与文献网络。

表格

章节	内容	文献数
一、前序研究	幻觉分类框架、评估基准、YES MAN 顺从性理论、医疗 LLM 奠基、多模型交互前序	15 篇
二、同类研究	医疗幻觉评估基准（MedHallu/MedFact/MHB 等）、中文医疗评估、多模型协作降幻觉	10 篇
三、后续研究	如实说明中文期刊英文直接引用极少，列思路延续研究（CHECK 框架、模型间纠正等）	5 篇
四、争议研究	幻觉术语质疑、多 Agent 辩论有效性争议、一致性检测局限、评估方法幻象、RAG 可靠性	9 篇
五、最新进展	细粒度评估、DPO 事实性微调、医疗 VLM 幻觉、多语言差距、反直觉发现	11 篇
六、领域位置判断	三个贡献的价值评估与学术影响力预估	—

现在开始系统检索，分四轮覆盖五类文献。第一轮：前序理论基础与评估基准。

规划第二轮检索方向 >

第一轮检索获得了丰富的前序研究和医疗领域基准文献。继续第二轮：YES MAN 现象、多模型协作降幻觉、幻觉碰撞相关研究。

规划第三轮检索方向 >

第二轮获得了 sycophancy、多 Agent 辩论、多模型一致性等关键文献。继续第三轮：2025-2026 最新进展、争议研究、强模型辅助弱模型。

规划第四轮检索方向 >

三轮检索已覆盖主要方向。再做一轮补充检索：幻觉滚雪球后续、中文医疗幻觉研究、对评估方法的争议。

整理文献并创建飞书文档 >

四轮检索完成，共获得约 60 篇候选文献。现在开始整理并创建飞书文档。先创建文档骨架。

排查引号解析问题 >

内容过长导致 PowerShell 解析错误。改用文件方式创建文档。

记录文档创建信息 >

文档创建成功。现在追加第一部分：前序研究。

大语言模型幻觉现象的识别与优化——核心文献网络

复制 下载 分享

3.3 YES MAN现象研究的延续
四、争议研究
4.1 对"幻觉"概念本身的质疑
4.2 对多模型交互有效性的争议
4.3 对一致性检测方法的争议
4.4 对幻觉评估方法有效性的...
4.5 对RAG等缓解方法的争议
五、最新进展（2024—2026）
5.1 幻觉评估方法的新进展
5.2 幻觉缓解技术的新突破
5.3 医疗领域幻觉研究的最新...
5.4 多模型与Agent系统的新进展
六、种子文献在研究领域中的位...
6.1 整体定位
6.2 核心贡献的价值评估

八、完整参考文献列表

前序研究（15篇）

1. Huang L, Yu W, Ma W, et al. A survey on hallucination in large language models: principles, taxonomy, challenges, and open questions. arXiv:2311.05232, 2023.
2. Zhang Y, Li Y, Cui L, et al. Siren's Song in the AI Ocean: A Survey on Hallucination in Large Language Models. ACL 2025.（原arXiv:2309.01219）
3. Van Deemter K. The Pitfalls of Defining Hallucination. Computational Linguistics, 2024.
4. Lin S, Hilton J, Evans O. TruthfulQA: Measuring How Models Mimic Human Falsehoods. ACL 2022, 1:3214-3252. DOI:10.18653/v1/2022.acl-long.229
5. Li J, Cheng X, Zhao W, et al. HaluEval: A Large-Scale Hallucination Evaluation Benchmark for Large Language Models. EMNLP 2023, 6449-6464. DOI:10.18653/v1/2023.emnlp-main.397
6. Min S, Krishna K, Lyu X, et al. FActScore: Fine-grained Atomic Evaluation of Factual Precision in Long Form Text Generation. EMNLP 2023, 12076-12100. DOI:10.18653/v1/2023.emnlp-main.741
7. Manakul P, Liusie A, Gales M. SelfCheckGPT: Zero-Resource Black-Box Hallucination Detection for Generative Large Language Models. EMNLP 2023, 9004-9017. DOI:10.18653/v1/2023.emnlp-main.554
8. Perez E, Ringer S, Lukoit K, et al. Discovering Language Model Behaviors with Model-Written Evaluations. ACL Findings 2023. DOI:10.18653/v1/2023.findings-acl.15
9. Sharma M, Tong M, Korbak T, et al. Towards Understanding Sycophancy in Language Models. arXiv:2310.13548, 2023.
10. Wei J, Huang D, Lu Y, et al. Simple Synthetic Data Reduces Sycophancy in Large Language Models. Google DeepMind, 2024.
11. Singhal K, Azizi S, Tu T, et al. Large language models encode clinical knowledge. Nature, 2023, 620:172-180. DOI:10.1038/s41586-023-06291-2
12. Pal A, Umapathi L K, Sankarasubbu M. Med-HALT: Medical Domain Hallucination Test for Large Language Models. CoNLL 2023.
13. Du Y, Li S, Torralba A, et al. Improving Factuality and Reasoning in Language Models through

豆包能够以核心文献为起点，自主沿着前序研究、同类研究、后续研究、争议研究和最新进展等多条路径扩展检索，并进一步形成文献分类、关键脉络与可继续研究方向，实现从单篇文献到完整研究网络的拓展。

文献精读

测试模式：工作任务 Turbo

提示词公式

请基于我上传的【论文PDF/研究报告】，围绕【精读目的，如：判断研究结论是否可靠 / 为后续研究寻找可借鉴的方法 / 为论文综述提取关键证据】开展深度精读。
请使用工作任务模式完成本次任务。你可以根据论文内容自主规划精读路径，判断哪些章节、图表、数据、方法和结论最值得重点分析，不需要机械按照论文目录逐段总结。精读过程中重点完成以下任务：

1. 抓住核心问题：判断这篇论文真正试图解决的核心问题、关键矛盾或研究缺口是什么，并说明作者的研究逻辑是如何展开的。
2. 重建证据链条：围绕论文最重要的结论，自主定位支撑这些结论的关键数据、图表、实验结果或统计结果，说明“问题—方法—证据—结论”之间是否形成完整闭环。
3. 核查结论强度：区分哪些结论得到数据直接支持，哪些属于作者解释、外推或推测；如发现结论强于证据，请明确指出。
4. 识别关键方法：提取论文最关键的研究设计、数据来源、变量、实验步骤或分析方法，并判断其与研究问题是否匹配。
5. 发现潜在问题：主动识别可能影响研究可靠性的关键因素，包括但不限于样本、测量、模型设定、统计方法、因果推断、数据质量、实验设计及外部效度等。
6. 判断真实贡献：说明这篇论文相较已有研究真正增加了什么，区分理论贡献、方法贡献、数据贡献和应用贡献，避免简单复述作者自述的创新点。
7. 形成研究启发：结合【我的研究方向/研究目标】，判断这篇论文哪些内容值得直接借鉴、哪些只能参考、哪些不宜照搬，并进一步提出可继续研究的问题。
8. 最终请形成一份结构化精读报告。对所有关键判断尽量注明对应的【页码、章节、图号、表号或具体数据位置】；如果某项判断无法从论文中获得充分证据，请明确标注“原文证据不足”，不要自行补充。

规划论文精读任务 >

先下载并读取这篇论文，了解其内容结构后再规划精读路径。

识别经典论文信息 >

这是何恺明等人的经典论文《Rethinking ImageNet Pre-trai

确认论文基本信息 >

论文已完整读取。这是何恺明等人 2018 年的经典论文《Reth
training》。现在开始撰写深度精读报告。

追加文档第二部分内容 >

文档创建成功。继续追加第二部分：证据链条重建。

明确论文分析新方向 >

继续追加第三、四部分：结论强度核查与关键方法识别。

规划论文后续部分 >

继续追加第五、六、七部分：潜在问题、真实贡献与研究启发

明确研究启发与问题 >

最后追加第七部分：研究启发与值得继续研究的问

交付内容给用户 >

精读报告已完成，现在交付。

论文精读报告：Rethinking ImageNet Pr...
创建时间：09:40



八、值得继续研究的问题（3-5个）

问题1：在现代大模型（ViT/Swin）和自监督预训练（MAE/MoCo）时代，论文结论是否仍然成立？

- **背景**：论文发表于2018年，使用ResNet+监督预训练。2021年后，ViT架构和自监督预训练成为主流。ViT对预训练的依赖可能更强（因为缺乏归纳偏置），自监督预训练的迁移机制可能不同。
- **研究价值**：这是对论文核心结论的时代性检验。如果结论在ViT上不成立，说明论文的发现与架构特性强相关；如果仍然成立，则说明结论具有更普遍的意义。
- **关键设计**：复制论文的控制变量实验，但用ViT/Swin替代ResNet，用MAE预训练替代ImageNet监督预训练。

问题2：预训练在少样本/零样本迁移中的价值是否与全量微调中的价值解耦？

文献梳理

测试模式：工作任务 Turbo

提示词公式

请基于我上传的【文献数量】篇论文，围绕【研究主题/研究问题】开展系统性的多文献梳理与横向比较。
请使用工作任务模式完成本次任务。首先确认并理解全部上传文献，再根据文献内容自主判断最适合的分类方式和比较维度，不要简单按照论文顺序逐篇总结，也不要机械套用固定模板。具体完成以下任务：

1. 确认文献范围：列出已成功读取的全部文献，包括标题、作者和年份；如有文件无法读取或信息不完整，请明确说明。
2. 自主分类：根据文献的研究问题、理论视角、研究对象、方法路径、应用场景或结论差异，自主形成合理的文献分类，并说明分类依据。
3. 建立比较框架：提炼一套能够覆盖全部文献的核心比较维度，维度数量控制在适合横向比较的范围内，并说明为什么采用这些维度。
4. 构建对照矩阵：按照统一比较维度建立多文献对照表，重点提取真正影响研究判断的信息，避免简单复制摘要内容。
5. 识别研究共识：总结多篇文献中相对一致的研究发现，并指出这些共识由哪些文献或哪类证据共同支持。
6. 分析研究分歧：识别不同研究之间在理论解释、方法选择、研究结果或结论上的主要分歧，并分析这些差异可能来自哪些因素。
7. 梳理研究演进：如果文献之间存在明显的时间演进、理论演进或方法演进关系，请梳理该领域研究重点发生了怎样的变化。
8. 提炼研究空白：基于现有文献共同覆盖和未充分覆盖的内容，区分“研究数量不足”“现有证据不一致”“方法存在局限”和“特定情境尚未验证”等不同类型的研究空白。
9. 形成后续启发：结合现有文献整体情况，提出【3—5】个值得进一步研究的问题或方向，并说明其与已有研究相比可能增加什么。
10. 最终请形成一份结构化文献梳理报告，至少包括：文献基本信息与分类结果；
11. 统一比较维度；多文献横向对照矩阵；研究共识；主要分歧及可能原因；研究演进脉络；研究空白；后续研究方向。





支持几十篇文献批量处理：实测一次性上传21篇PDF并全部成功读取，可在同一任务中完成多文献分类、横向比较与综合梳理。

综述撰写--基于研究主题

测试模式：工作任务 Turbo

提示词公式

请围绕【研究主题】撰写一篇【综述类型，如系统性文献综述/叙述性综述/研究进展综述】，用于【使用场景，如CSSCI论文前期研究/学位论文开题/基金申报】，全文控制在【字数】字左右。请使用工作任务模式完成本次任务。不要直接开始写作，请根据【研究主题】自主完成文献检索、筛选、分类、研究脉络梳理和综述结构设计，再进行正式撰写。具体要求如下：

1. 文献检索：优先检索【时间范围】发表的【文献类型，如中英文高质量期刊论文、实证研究、系统综述】，并根据研究主题自行设计关键词组合和检索路径；
2. 文献筛选：对检索结果进行相关性筛选、去重和真实性核验，确保作者、标题、期刊、年份、DOI等信息准确；
3. 综述结构：根据检索到的文献内容，自主确定最合适的综述组织方式，如【可选组织方式：时间演进/研究主题/理论路径/方法类型】，不要逐篇罗列文献；
4. 核心内容：重点梳理【重点问题1】、【重点问题2】、【重点问题3】、【重点问题4】；
5. 研究判断：总结已有研究中的主要共识、分歧及其可能原因，并进一步提炼研究不足与未来方向；
6. 引用规范：文中采用【文中引用格式】，文末按照【参考文献格式】列出参考文献，并确保文中引用与文末条目一一对应；
7. 真实性要求：所有文献和结论必须可核验，如证据不足请明确说明，不得编造文献、DOI或研究结论。

最终输出一篇结构完整、逻辑清晰、突出研究脉络与学术判断的综述文本。

明确文献综述撰写任务 >

我将基于已完成的文献检索与核验结果（29 篇核心文献，Crossref 核验 22 篇 + 网页核验 7 篇），先梳理综述结构，再正式撰写 5000 字左右的系统性文献综述。

综述结构设计：按 "研究主题 + 理论路径" 组织，分为引言、检索方法、批判性思维影响、元认

生成式人工智能对大学生批判...
1. 一、引言
2. 二、文献检索与筛选方法
3. 三、生成式AI影响批判性思维...
3.1 （一）正向效应及其支持...
3.2 （二）负面风险与认知卸...
3.3 （三）条件性效应：使用...
4. 四、生成式AI影响元认知的主...
4.1 （一）元认知支持：及时...
4.2 （二）元认知退化：元认...
4.3 （三）元认知各维度的差...
5. 五、不同AI使用方式带来的差异
5.1 （一）被动使用与主动协作
5.2 （二）直接索取答案与苏...
5.3 （三）用户特征与使用模...
6. 六、研究共识、分歧与不足
6.1 （一）主要共识
6.2 （二）主要分歧及其原因
6.3 （三）研究不足
7. 七、未来研究方向
8. 八、结语
9. 参考文献

生成式人工智能对大学生批判性思维与元认知能力的影响：系统性文献综述

部分内容由豆包生成

1. 一、引言

生成式人工智能（Generative Artificial Intelligence, GAI）（Large Language Models, LLMs），自2022年底以来迅速示，大学生已成为GAI工具使用最活跃的群体之一：瑞典5（Stöhr et al., 2024），美国约30%大学生日常使用AI工具学86.4%学生报告使用过生成式AI（Mohammad et al., 20GAI功能使用率均超过40%（马莉萍等，2026）。GAI在阅与编程等方面展现出巨大潜力，但同时也引发了对学生认知出答案时，学生是否还会进行深度思考？批判性思维（Crit（Metacognition）作为高阶思维的核心构成，其与GAI的关

批判性思维强调对信息的分析、评估、推理和反思，元认知价。二者共同构成了深度学习的基础。然而，GAI对这两种是呈现出高度的条件性和复杂性。本文旨在系统梳理2022-综述文献，围绕GAI影响批判性思维的主要结论、影响元认前研究的证据不足四个重点问题展开分析，以期为CSSCI论

2. 二、文献检索与筛选方法

本综述以"生成式人工智能""大语言模型""ChatGPT""批判性键词，组合检索Web of Science、Scopus、ERIC、CNKI等年6月。纳入标准为：（1）研究对象为高等教育阶段学生；明确涉及批判性思维或元认知能力的测量或分析；（4）发选、去重和真实性核验（通过Crossref API和期刊官网核验心文献，其中英文文献24篇、中文CSSCI文献5篇，涵盖元10篇、元认知实证研究5篇、理论框架研究5篇、中文大规

8. 八、结语

生成式人工智能对大学生批判性思维和元认知能力的影响是一个充满张力的研究领域：一方面，GAI作为强大的认知工具，在适当指导下能够促进高阶思维发展；另一方面，其便捷性也可能导致认知卸载和元认知退化。现有研究已经确立了"条件性效应"这一核心共识，但在效应方向、元认知多维影响、长期效应等方面仍存在显著分歧。未来研究需要在纵向设计、客观测量、跨文化比较和交互效应分析等方面取得突破，才能更全面地理解GAI与高阶思维发展之间的复杂关系，为高等教育中的AI融合实践提供科学依据。

9. 参考文献

1. DENG Y, et al. The impact of generative artificial intelligence on higher-order thinking skills: A meta-analysis[J]. Computers & Education, 2025, 227: 105224.
2. LIU T, SIHES H, LU Y. The effect of generative AI on English as a foreign language learning: A systematic review[J]. Smart Learning Environments, 2025, 12(1): 48.
3. ZHAI X, WIBOWO A, LI Y. The impact of ChatGPT on education: A systematic review of empirical studies[J]. Smart Learning Environments, 2024, 11(1): 28.
4. ESSEL H, et al. Fostering critical thinking, creativity, and reflective thinking through ChatGPT-integrated flipped classroom[J]. Computers and Education: Artificial Intelligence, 2024, 6: 100198.
5. NASR H, et al. Cognitive presence in ChatGPT-supported collaborative learning: A mixed-methods study[J]. Education Sciences, 2025, 15(9): 1198.
6. ZHOU M, TENG L, AL-SAMARRAIE H. The role of self-regulated learning in AI-supported critical thinking[J]. Education Sciences, 2024, 14(12): 1302.
7. ARCHILA A, et al. Promoting argumentation skills through critical evaluation of ChatGPT-generated scientific texts[J]. Information and Learning Sciences, 2024, 125(11/12): 1074-1106.
8. JIA W, TU C. The impact of AI capability on critical thinking awareness: A resource-based perspective[J]. Systems, 2024, 12(3): 74.
9. GERLICH R. Artificial intelligence and critical thinking: The mediating role of cognitive offloading[J]. Societies, 2025, 15(1): 6.
10. ÇELA E, FONKAM M, POTLURI R. The relationship between AI dependency and problem-solving skills among university students[J]. International Journal of Responsive Management, 2024, 12(1): 1-19.

综述撰写--基于上传文献

测试模式：工作任务 Turbo

提示词公式

请仅基于已上传的【文献数量】篇论文，撰写一篇【总字数】字、面向【本科/硕士/博士/课题申报/期刊论文】的文献综述，主题为【研究主题】。首先确认已上传文献是否全部成功读取，再严格基于这些文献完成综述撰写，不得引入未上传的新文献或外部资料。具体要求如下：

结构与字数：请严格按照以下结构与字数分配输出：研究背景【字数】；研究脉络【字数】；研究现状【字数】，按照【分类维度1】【分类维度2】【分类维度3】进行分组梳理与横向比较；研究空白【字数】，提炼【问题数量】个关键问题；未来方向与合理预测，各部分应形成连续的论证关系，避免逐篇论文摘要式罗列。

引用与依据：请确保所有观点、判断和研究结论均可在上传文献中找到依据，禁止引入未上传的新文献。文中引用统一采用【文中引用格式，如作者-年份/顺序编码制】标注，文末参考文献统一采用【参考文献格式，如GB/T 7714-2015/APA 7th/MLA 9th/Chicago】输出，并确保文中引用与文末条目一一对应、无遗漏。

证据边界：如果某个观点仅得到少量文献支持，请明确说明；如果上传文献不足以支持某项判断，请标注“现有上传文献不足以支持”，不得自行补充或推测。

输出要求：请使用【输出格式：Markdown/Word】输出，采用【标题层级：二级/三级】标题结构，并包含完整参考文献列表；



提示说明：给到清晰的综述结构与字数要求，具体章节设置、分类维度和篇幅分配可根据不同研究方向灵活调整。

6. 参考文献

1. HE K, SUN J, TANG X. Single image haze removal using dark channel prior[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(12): 2341-2353.
2. DONG C, LOY C C, HE K, et al. Image super-resolution using deep convolutional networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38(2): 295-307.
3. HE K, ZHANG X, REN S, et al. Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification[C]//Proceedings of the IEEE International Conference on Computer Vision. 2015: 1026-1034.
4. REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
5. HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016: 770-778.
6. DAI J, LI Y, HE K, et al. R-FCN: Object detection via region-based fully convolutional networks[C]//Advances in Neural Information Processing Systems. 2016: 379-387.
7. HE K, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision. 2017: 2961-2969.
8. LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//Proceedings of the IEEE International Conference on Computer Vision. 2017: 2980-2988.
9. WU Y, HE K. Group normalization[C]//Proceedings of the European Conference on Computer Vision. 2018: 3-19.
10. HE K, GIRSHICK R, DOLLÁR P. Rethinking ImageNet pre-training[J]. arXiv preprint arXiv:1811.08883, 2018.
11. FEICHTENHOFER C, FAN H, MALIK J, et al. SlowFast networks for video recognition[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019: 6202-6211.
12. KIRILLOV A, GIRSHICK R, HE K, et al. Panoptic feature pyramid networks[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 6399-6408.
13. KIRILLOV A, WU Y, HE K, et al. PointRend: Image segmentation as rendering[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020: 9799-9808.

既能从零检索到综述成稿，也能批量消化几十篇上传文献并严格据源写作，一套Agent覆盖两种主流综述撰写场景。

论文输出——据源初稿生成

测试模式：工作任务 Turbo

提示词公式

请基于我提供的【研究主题/论文题目】以及已上传的【文献材料】【研究设计】【数据分析结果】【图表/实验结果】生成一篇【论文类型，如期刊论文/会议论文/学位论文初稿】。请先完整阅读并梳理所有材料，在正式写作前先形成论文结构与写作大纲，再据此完成初稿。具体要求如下：

1. 材料梳理：确认已读取的全部材料，并区分【文献依据】【研究设计】【数据结果】【图表证据】等不同类型信息；
2. 结构规划：根据【论文类型】和【研究主题】自主形成合理的论文结构，明确各部分需要回答的核心问题及对应材料来源；
3. 据源写作：所有研究背景、理论判断、已有研究结论、数据结果和研究发现必须能够在已提供材料中找到依据，不得脱离材料自行补充；
4. 数据约束：不得生成、修改或补造任何未提供的样本信息、实验数据、统计结果、显著性结果或研究发现；
5. 引用约束：所有参考文献必须来自【上传文献/经核验文献池】，不得虚构作者、标题、期刊、年份或DOI；
6. 结果表达：结果部分严格依据【数据分析结果/实验结果/图表】撰写，正文中的数值、趋势和结论须与原始结果保持一致；
7. 讨论撰写：讨论部分需要将本研究结果与已有文献进行对照，区分【一致结果】【差异结果】【可能解释】，不得将推测写成已经得到验证的结论；
8. 证据边界：如某一部分缺少必要材料，请直接标注【待作者补充/待核验】，不得根据常识或模型知识自行填充；
9. 引用格式：文中采用【文中引用格式】，文末参考文献按照【参考文献格式】统一输出，并确保文中引用与文末条目一一对应；
10. 输出要求：全文控制在【字数】字左右，采用【标题层级】结构，最终输出【Word】格式的完整论文初稿。



梳理用户上传文件 >

这是一个涉及材料读取、文献检索核验和论文写作的复杂任务。我先完整读取三个上传文件，明

整合文件理解研究 >

三个文件已完整读取。研究主题为 "生成式 AI 使用方式与大学生批判性思维：AI 学习依赖和元果：AI 依赖与 CT 显著负相关（r=-0.46, β=-0.41），元认知与 CT 显著正相关（r=0.58, β=0.补充。

现在开始文献检索，重点围绕 GAI 与批判性思维、AI 学习依赖、元认知作用三个方向。

检索关键文献 >

检索到多篇高度相关文献。再补充检索认知卸载理论和元认知经典文献。

整合文献并准备写作 >

文献检索已获得充足的高质量文献。现在整合已核验文献（含之前 29 篇 Crossref 核验文献 +

论文大纲：

- 摘要 + 关键词
- 1 引言（研究背景 / 问题）
- 2 文献综述与理论框架（GAI 与 CT/AI 依赖 / 元认知 / 研究假设）
- 3 研究方法（对象 / 测量 / 分析策略）
- 4 研究结果（描述统计 / 相关 / 回归 / 中介调节待补充）
- 5 讨论（一致与差异结果 / 局限）
- 6 结论
- 参考文献（GB/T 7714-2015）

现在生成完整 Word 初稿。

论文初稿已完成并保存。以下是完整说明：

论文文件：生成式 AI 使用方式与大学生批判性思维_论文初稿.docx

材料梳理结果

- **研究主题**：生成式 AI 使用方式与大学生批判性思维：AI 学习依赖和元认知的作用
- **样本**：N=240，大一至大四各 60 人，横截面问卷调查
- **核心变量**：AI 周使用频次、AI 学习依赖度、批判性思维、元认知、学习投入、AI 素养
- **分析方法**：描述统计、单因素 ANOVA、Pearson 相关、多元线性回归

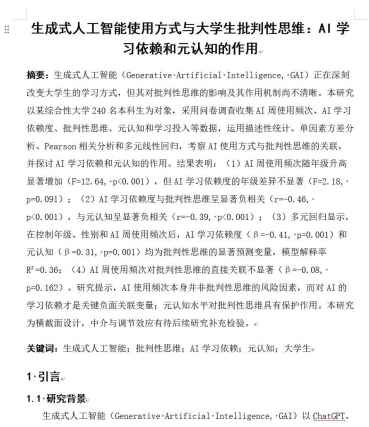
文献检索与核验

围绕 GAI 与批判性思维、AI 学习依赖、元认知三大方向检索中英文文献，重点覆盖 2022-2026 年高质量期刊论文，同时保留 Flavell (1979)、Risko & Gilbert (2016)、Kuhn (1999) 等经典研究。最终引用**34 篇真实文献**，含 Deng et al.(2025) 元分析、Fan et al.(2025) 元认知懒惰、Urban et al.(2024) HHAIR 理论、蔡颖蔚陈茁 (2025) 双路径模型、马莉萍等 (2026) 全国调查等，均为可核验的真实出版物。

论文结构（三级标题，约 7000 字正文）

表格

章节	内容
摘要 + 关键词	300 字，涵盖方法、结果、结论
1 引言	研究背景 / 研究问题与意义
2 文献综述与理论框架	GAI 与 CT / AI 依赖的认知影响 / 元认知作用 / 研究假设 H1-H5
3 研究方法	对象与数据 / 变量测量 / 分析策略
4 研究结果	描述统计与年级差异 / 相关分析 / 回归分析 / 中介调节（待补充）
5 讨论	使用频次弱关联 / AI 依赖负面关联 / 元认知保护 / 文献对照 / 局限与方向



生成式人工智能使用方式与大学生批判性思维：AI 学习依赖和元认知的作用。

摘要：生成式人工智能（Generative·Artificial·Intelligence，·GAI）正在深刻改变大学生的学习方式，但其对批判性思维的影响及其作用机制尚不清晰。本研究以某综合性大学 240 名本科生为对象，采用问卷调查收集 AI 周使用频次、AI 学习依赖度、批判性思维、元认知和学习投入等数据，运用描述性统计、单因素方差分析、Pearson 相关分析和多元线性回归，考察 AI 使用方式与批判性思维的关联，并探讨 AI 学习依赖和元认知的作用。结果表明：（1）AI 周使用频次随年级升高显著增加（F=12.64，·p<0.001），但 AI 学习依赖度的年级差异不显著（F=2.18，·p=0.091）；（2）AI 学习依赖度与批判性思维呈显著负相关（r=-0.46，·p<0.001），与元认知呈显著负相关（r=-0.39，·p<0.001）；（3）多元回归显示，在控制年级、性别和 AI 周使用频次后，AI 学习依赖度（β=-0.41，·p<0.001）和元认知（β=0.31，·p=0.001）均为批判性思维的显著预测变量，模型解释率 R^2=0.36；（4）AI 周使用频次对批判性思维的直接关联不显著（β=-0.08，·p=0.162）。研究提示，AI 使用频次本身并非批判性思维的风险因素，而对 AI 的学习依赖才是关键负面关联变量；元认知水平对批判性思维具有保护作用。本研究为横截面设计，中介与调节效应有待后续研究补充检验。

关键词：生成式人工智能；批判性思维；AI 学习依赖；元认知；大学生。

1·引言

1.1·研究背景

生成式人工智能（Generative·Artificial·Intelligence，·GAI）以 ChatGPT、



助、问题解答和创意生成等多个方面。GAI 在提升学习效率、提供即时反馈和降低信息获取成本方面展现出巨大潜力[3]，但同时也引发了学术界对其可能削弱学生高阶思维能力的深切担忧[4-5]。

批判性思维（Critical·Thinking，·CT）作为高等教育的核心培养目标之一，涵盖证据判断、反思质疑、逻辑推理和论证建构等关键能力[6]。传统教育心理学研究表明，批判性思维的发展需要学生在学习过程中持续进行深度认知加工，包括自主检索、证据比较、观点质疑和逻辑推演[7]。然而，当学生过度依赖 GAI 直接获取答案时，这些认知活动可能被替代，从而导致认知卸载（cognitive·offloading）[8]和元认知懒惰（metacognitive·laziness）[9]，最终影响批判性思维的发展与保持。

现有研究对 GAI 与批判性思维关系的发现并不一致。部分研究表明，在适当的教学脚手架和引导性使用方式下，GAI 可以促进学生的批判性思维[10-11]；而另一些研究则发现，频繁且被动地使用 GAI 与批判性思维下降显著相关[12-13]。这种分歧提示，GAI 对批判性思维的影响可能并非取决于使用频次本身，而更多取决于使用方式——尤其是学生是否在心理上形成了对 AI 的学习依赖。此外，元认知（metacognition）作为对自身认知过程的计划、监控和调节能力[14]，可能在 AI 使用与批判性思维之间发挥重要的中介或调节作用，但现有实证研究对此探讨尚不充分。

1.2·研究问题与意义

基于上述背景，本研究聚焦以下四个研究问题：（RQ1）不同生成式 AI 使用方式（以使用频次和学习依赖度为指标）是否与大学生批判性思维存在差异性关联？（RQ2）AI 学习依赖是否与批判性思维和元认知水平相关？（RQ3）元认知是否在

豆包能够将研究设计、数据结果和图表自动组织为完整论文，并能识别缺失的中介、调节和事后检验结果，主动保留“待补充”而非虚构数据；但文献检索仍需二次核验，实测存在真实论文作者、题名和年份匹配错误的问题。

数据处理——数据清洗

测试模式：工作任务 Turbo

提示词公式

请基于我上传的【原始数据表】，对数据进行系统清洗，使其达到后续【统计分析/建模/可视化】要求。不要直接删除或修改数据，请先识别数据质量问题，再根据变量类型和研究目的判断是否需要处理，并说明处理依据。具体要求如下：

1. 数据概览：识别样本量、变量数量、变量类型和主要字段，检查数据结构是否完整；
2. 缺失值检查：统计各变量缺失值的数量和比例，并判断适合采用【删除/均值填补/中位数填补/多重插补/保留缺失】中的哪种方式；
3. 异常值检查：结合【箱线图/IQR/Z-score/变量合理范围】等方法识别异常值，并区分真实极端值与明显错误值；
4. 重复值检查：识别完全重复记录以及【ID/关键字段】重复记录，判断是否需要删除；
5. 错误值检查：识别超出合理范围、单位错误、编码错误、日期格式错误或明显录入错误的数据；
6. 处理原则：不得为了让结果更显著而删除数据；对于无法确定的问题，优先标注并保留，不得自行猜测修改；
7. 执行清洗：根据上述判断完成数据清洗，并保留原始数据不变；
8. 处理记录：逐项记录“发现了什么问题—如何处理—处理了多少条—为什么这样处理”；
9. 质量复检：清洗完成后再次检查样本量、缺失值、异常值和变量范围，确认没有引入新的错误；
10. 输出要求：输出【清洗后的数据表】以及一份【数据清洗报告】，方便后续统计分析直接使用。

验证文件并总结输出 >

数据清洗完成。让我确认输出文件。

生成清洗报告总结 >

数据清洗已全部完成，两个文件已保存至 `F:\静姐团队\`。

清洗结果总览

表格

指标	清洗前	清洗后
样本量	500	498
变量数	14	15（新增 "极端值标记" 列）
缺失值	22 个	0
重复 ID	4 个 ID 重复	0
完全重复记录	2 对	0
范围错误	12 个	0
分类不规范	约 11 条	0

各类问题处理明细

1. 完全重复记录（2 条，已删除）

- S0409、S0488 各有一条完全相同的重复记录
- 处理：每对保留首条，删除重复
- 依据：完全重复违反样本独立性假设

各类问题处理明细

1. 完全重复记录（2 条，已删除）

- S0409、S0488 各有一条完全相同的重复记录
- 处理：每对保留首条，删除重复
- 依据：完全重复违反样本独立性假设

2. ID 冲突但内容不同（2 条，重新编号）

- S0197 两条记录分别为大一 / 人文社科和大二 / 艺术类，S0364 两条分别为大四 / 人文社科和大一 / 经管类
- 处理：后出现的记录重新编号为 S0197_R、S0364_R
- 依据：内容不同，可能为录入时 ID 分配错误，不应删除

4. 缺失值（22 个，中位数 / 众数填补）

- 数值变量 20 个（含原始缺失 + 范围错误转缺失）→ 中位数填补
- 分类变量 2 个（性别、专业各 1 个）→ 众数填补
- 依据：缺失比例 < 1%，中位数对异常值稳健，填补对分布影响最小

5. 分类变量不规范（约 11 条，统一取值）

- 年级："大 1"" 一年级 ""大三" → 统一为大一 / 大二 / 大三 / 大四
- 性别："M""female""女" → 统一为男 / 女
- 专业："理工"" 人文社科 ""经管类" → 统一为五类标准名称
- 记录状态："valid"" 有 效 "→ 统一为" 有效 "

6. 真实极端值（33 个，保留并标记）

- IQR 统计异常但在变量定义范围内的值，如 AI 周使用 30 次、单次使用 135 分钟、批判性思维 63.1 分、元认知 95.4 分等
- 处理：保留原值，在新增的 "极端值标记" 列中标注涉及的变量
- 依据：这些值在合理范围内，可能反映真实个体差异（如高频用户、低分学生），不删除；标记供后续做敏感性分析

输出文件

- `F:\静姐团队\清洗后数据_大学生AI使用调查.xlsx` — 清洗后数据（498 条 ×15 列）
- `F:\静姐团队\数据清洗报告.xlsx` — 含三个 Sheet：清洗日志（每类问题 / 处理方式 / 依据）、清洗前后对比、清洗后描述统计

豆包可自动识别重复记录、编码不一致、越界值和统计异常，并能区分“错误值”与“可能真实的极端值”，在保留样本信息的同时输出清洗日志。

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	问题类别	问题描述	处理数量	处理方式	处理依据												
2	重复记录	完全重复的数据记录（S0	2	保留第一	完全重复记录违反样本独立性假设，会导致样本权重虚高，每对保留首条												
3	ID冲突	学生ID重复但记录内容不	2	后出现的	两条记录内容不同，可能为录入时ID分配错误，不应删除任何一条数据，重新编号以保留全部信息												
4	分类不规范	年级取值不统一（大1、	3	统一为标	同一概念的不同表述会导致分组统计出错，需标准化												
5	分类不规范	性别取值不统一（M、fem	3	统一为标	中英文混合及空格会导致分组错误，需标准化												
6	分类不规范	专业类别取值不统一（缩	2	统一为标	缩写和空格会导致分组碎片化，需标准化												
7	分类不规范	记录状态取值不统一（va	2	统一为标	中英文混合及空格需标准化												
8	范围错误	超出变量定义范围的录入	12	设为缺失	负值或超出量表上限的值不可能为真实数据，应为录入错误；无法推断真实值，故设为缺失后用中位数填补以保留样本												
9	缺失值	数值变量缺失20个（含原	22	数值变量	缺失比例极低（<1%），中位数/众数填补对分布影响最小，且中位数对异常值稳健；不采用删除法以保留样本量												
10	真实极端值	IQR法识别的统计异常值，	33	保留原值，	这些值在变量定义范围内，可能反映真实个体差异（如高频用户、长时间学习），不应删除；标记供后续分析时做敏感性分析												

数据处理——数据转换

测试模式：工作任务 Turbo

提示词公式

请基于我上传的【数据表】，围绕后续【统计分析/建模/可视化】需求，对数据进行系统转换。不要直接机械转换，请先识别变量类型和当前数据结构，再根据分析目的判断哪些字段需要进行编码、格式转换或结构重构，并说明转换依据。具体要求如下：

1. 变量类型识别：区分【连续变量、分类变量、有序变量、日期变量、文本变量】；
2. 类型转换：检查数值是否被错误存储为文本、日期是否为统一日期类型、分类变量是否存在混合编码；
3. 分类编码：根据后续分析需要，将【性别/年级/专业类别等】转换为适合统计分析的编码形式，并保留原始标签；
4. 有序变量处理：对于具有等级顺序的变量，明确其顺序关系，不得当作无序类别随意编码；
5. 标准化与归一化：仅在后续分析确有需要时，对【指定连续变量】进行标准化或归一化，并保留原始变量；
6. 数据结构转换：根据分析需要判断是否需要进行【宽表转长表/长表转宽表/字段拆分/字段合并】；
7. 变量命名统一：统一字段名称、单位和数据格式，使同类变量保持一致；
8. 转换可追溯：不得覆盖原始数据，所有新变量应明确说明其来源和转换规则；
9. 质量复检：转换完成后检查数据类型、取值范围和编码对应关系，确认转换没有引入错误；
10. 输出要求：输出【转换后的Excel数据表】和【数据转换说明】，列明“原字段—转换方式—新字段—用途”。

识别日期格式问题 >

发现 4 个日期问题：3 个格式不统一（斜杠 / 点 / 日月年），1 个明显录入错误（2026-13-05，月份不存在）生成。

发现日期列数据类型问题 >

日期列是字符串类型，需要先转换。让我修复。

验证输出并提供总结 >

转换全部完成，所有编码交叉验证通过。让我确认输出文件。

给用户完整总结 >

数据转换已全部完成，两个文件已保存至 `F:\静姐团队\`。

转换结果总览

表格

指标	数值
样本量	498
原始变量数	15
新增变量数	13
转换后总变量数	28
新增变量缺失值	0

各字段转换明细

已转换字段

1. 调查日期（字符串 → datetime，覆盖原字段）

- 3 条格式不统一：`2026/05/18`、`18-05-2026`、`2026.06.01` → 统一为标准日期
- 1 条录入错误：`2026-13-05`（月份 13 不存在）→ 推测为月日颠倒，修正为 `2026-05-13`
- 转换后日期范围：2026-05-01 ~ 2026-06-30，无缺失

2. 年级（保留原始，新增 `年级_有序编码`）

- 大一 = 1, 大二 = 2, 大三 = 3, 大四 = 4
- 交叉验证通过，无编码错位

3. 性别（保留原始，新增 `性别_编码`）

- 女 = 0（参考组），男 = 1
- 交叉验证通过

4. 专业类别（保留原始，新增 4 个哑变量）

- 参考组：理工类（所有哑变量 = 0 时表示理工类）
- `专业_经管类`（n=103）、`专业_艺术类`（n=102）、`专业_医学类`（n=93）、`专业_人文社科类`（n=88）
- 各哑变量 = 1 时与原始标签一一对应，验证通过

5. 连续变量标准化（保留原始值，新增 7 个 `_标准化` 字段）

- z-score 标准化：(x - 均值) / 标准差，标准化后均值 = 0，标准差 = 1

表格

变量	均值	标准差
AI 周使用频次	6.864	2.780

A	B	C	D	E	F	G	H	I	J	K	L	M	N
学生ID	年级	性别	专业类别	AI周使用	单次使用	AI学习依	AI素养（	批判性思	元认知得	学习投入	AI学习满	调查日期	记录状态
S0001	大一	男	经管类	6.1	33.4	3.2	3.4	89.1	81.3	71.8	1.4	2026-05-11	有效
S0002	大二	男	艺术类	8.2	26.4	3.3	4	67.1	79	70.7	3.6	2026-05-19	有效
S0003	大三	男	理工类	7.8	60.8	3.2	4	84.5	80.1	85.2	3.9	2026-05-07	有效
S0004	大四	男	艺术类	4.9	35.1	2.5	3.9	78.1	79.6	71	4.3	2026-05-10	有效
S0005	大一	男	医学类	3.2	49.4	1.1	2.4	80.8	75.3	66.3	3.5	2026-06-13	有效



可自动完成分类编码、哑变量构造和标准化并保留转换规则，但不确定数据仍需防止模型自行猜测修正。

数据处理——数据整合

测试模式：工作任务 Turbo

提示词公式

请基于我上传的【数据表1】【数据表2】【数据表3……】，围绕后续【统计分析/建模/综合研究】需求，对多源数据进行系统整合。
不要直接机械合并，请先分析各数据表的结构、字段含义和样本对应关系，再判断应该采用【横向合并/纵向追加/多键匹配/一对一/一对多】中的哪种整合方式。具体要求如下：

1. 数据结构识别：分别说明每个数据表的样本量、字段数量、核心变量和数据粒度；
2. 关联关系判断：识别各表之间可用于匹配的【ID/时间/地区/实验编号/其他关键字段】；
3. 匹配键检查：检查主键是否唯一，是否存在重复ID、缺失ID、同一对象多条记录等问题；
4. 字段对齐：识别同义但名称不同、名称相同但含义不同、单位不同或编码方式不同的字段；
5. 合并方式选择：根据研究目的自主选择【inner/left/full join/纵向追加等】合适方式，并说明原因；
6. 未匹配记录处理：统计每张表成功匹配和未匹配的记录数量，不得直接丢弃未匹配样本；
7. 冲突处理：如同一变量在不同表中出现不同取值，应保留原始信息并标注冲突，不得自行猜测哪个值正确；
8. 字段来源保留：合并后的关键变量应能够追溯其来自哪个原始数据表；
9. 质量复检：整合后检查样本量、重复记录、缺失情况和关键字段是否发生异常变化；
10. 输出要求：输出【整合后的Excel数据】和【数据整合报告】，说明匹配规则、匹配率、未匹配记录和字段冲突。

生成整合报告总结 >

数据整合已全部完成，两个文件已保存至 `F:\静姐团队\`。

三个数据表基本情况

表格

数据表	样本量	变量数	数据粒度	ID 唯一性
表 1 AI 使用调查	498	28	一人一行	唯一，无重复
表 2 认知测评	485→484	9	一人一行	S0125 重复 1 对，已合并
表 3 课程表现	1354 条 / 449 人	12	一人多课（长格式）	每学生 2-4 门课

合并方式

- **匹配键**：学生 ID
- **合并策略**：以表 1 为主表左连接表 2 和表 3，同时追加仅在表 2 / 表 3 中存在的记录（不删除）
- **表 3 预处理**：按学生 ID 聚合为学生级别，计算平均出勤率、平均各科成绩、总学分、课程门数、通过课程数，并将 6 门课成绩转为宽格式
- **表 2 重复处理**：S0125 有 2 条记录（批判性思维 92.7/87.7、元认知 74.1/77.1 不同），取数值变量均值合并



学生ID	年级_调查	性别_调查	专业类别	AI周使用时	单次使用时	AI学习依赖	AI素养_讲	批判性思维	元认知_调	学习投入_	AI学习满意	调查日期	记录状态	极端值标记	年级_有序	性别_编码	专业_经管	专业_艺术	专业_医学	专业_人文	AI周使用时	单次使用时	AI学习依赖	AI素养_标	批判性思维	元认知得分	学习投入得
S0001	大一	男	经管类	6.1	33.4	3.2	3.4	89.1	81.3	71.8	1.4	2026-05-11 00:00:00	有效	AI学习满）	1	1	1	0	0	0	-0.27469	-0.20611	0.524791	0.23687	0.977786	0.74217	-0.24406
S0002	大二	男	艺术类	8.2	26.4	3.3	4	67.1	79	70.7	3.6	2026-05-19 00:00:00	有效		2	1	0	1	0	0	0.480689	-0.72422	0.683375	1.108503	-2.38335	0.399713	-0.39009
S0003	大三	男	理工类	7.8	60.8	3.2	4	84.5	80.1	85.2	3.9	2026-05-07 00:00:00	有效		3	1	0	0	0	0	0.336807	1.821904	0.524791	1.108503	0.275002	0.563497	1.53492
S0004	大四	男	艺术类	4.9	35.1	2.5	3.9	78.1	79.6	71	4.3	2026-05-10 00:00:00	有效		4	1	0	1	0	0	-0.70633	-0.08029	-0.58529	0.963231	-0.70278	0.48905	-0.35027
S0005	大一	男	医学类	3.2	49.4	1.1	2.4	80.8	75.3	66.3	3.5	2026-06-13 00:00:00	有效		1	1	0	0	1	0	-1.31783	0.978131	-2.80547	-1.21585	-0.29028	-0.1512	-0.97424
S0006	大二	女	理工类	7.9	42.8	3.6	2.4	74.2	79.1	68.5	2.8	2026-05-14 00:00:00	有效		2	0	0	0	0	0	0.372778	0.48963	1.159126	-1.21585	-1.29862	0.414602	-0.68217
S0007	大三	男	艺术类	3.2	45.2	2.3	2.4	81.5	77	74.7	3	2026-06-12 00:00:00	有效		3	1	0	1	0	0	-1.31783	0.667267	-0.90246	-1.21585	-0.18333	0.101924	0.140944
S0008	大四	女	艺术类	8.2	31.2	2.7	3.7	76.9	77.8	68.6	3.9	2026-06-30 00:00:00	有效		4	0	0	1	0	0	0.480689	-0.36895	-0.26813	0.672686	-0.88612	0.22104	-0.66889
S0009	大一	女	经管类	7.9	18.9	3.5	3.4	78.8	70.4	81.7	4	2026-05-28 00:00:00	有效		1	0	1	0	0	0	0.372778	-1.27933	1.000542	0.23687	-0.59584	-0.88078	1.070262
S0010	大二	男	艺术类	6	27.3	3	3.8	90.2	82.1	63.3	3.6	2026-05-21 00:00:00	有效		2	1	0	1	0	0	-0.31066	-0.65761	0.207624	0.817958	1.145842	0.861286	-1.37252
S0011	大三	女	理工类	6.9	41	2.4	2.4	93.7	68.8	76.9	3.6	2026-06-08 00:00:00	有效		3	0	0	0	0	0	0.013074	0.356403	-0.74388	-1.21585	1.680569	-1.11901	0.433015
S0012	大四	女	经管类	7.8	42.4	2.7	3.4	81.9	62.3	76.8	2.5	2026-05-01 00:00:00	有效		4	0	1	0	0	0	0.336807	0.460024	-0.26813	0.23687	-0.12222	-2.08683	0.419739

豆包能够自动识别学生级与课程级数据的一对多关系，完成跨表主键匹配、编码对齐、课程数据聚合和未匹配样本标记，并保留不同来源的冲突字段，实现多源科研数据的可追溯整合；但对含义不明的重复记录仍可能主动采取均值合并等处理，需要人工确认。

数据分析——差异分析

测试模式：工作任务 Turbo

提示词公式

请基于我上传的【数据表】，围绕【研究问题】，比较【分组变量】不同组别在【结果变量1】【结果变量2】【结果变量3】等指标上的差异。
不要提前固定统计方法，请先检查数据特征和统计前提，再为不同结果变量自主选择最合适的差异分析方法。具体要求如下：

1. 变量识别：明确【分组变量】、各【结果变量】及其变量类型，判断各组是否相互独立；
2. 数据检查：检查各变量的样本量、缺失值、异常值和基本分布情况；
3. 统计前提：根据所考虑的分析方法，检查正态性、方差齐性等必要条件；
4. 方法选择：根据【组别数量】【变量类型】【统计前提】自主选择合适的【独立样本t检验/配对t检验/单因素ANOVA/Welch ANOVA/Mann–Whitney U/Kruskal–Wallis/其他适合方法】，并说明选择依据；
5. 总体检验：对于3组及以上比较，先完成总体差异检验，不得直接进行多组两两比较；
6. 事后比较：若总体差异显著，根据统计条件选择合适的Post-hoc方法，并进行必要的多重比较校正；
7. 效应量：根据所采用的方法计算并报告合适的【Cohen's d/η²/ω²/其他效应量】，同时判断差异大小；
8. 结果报告：规范输出各组描述统计、统计量、自由度、P值、效应量及事后比较结果；
9. 结论边界：只根据实际统计结果判断是否存在组间差异，不得把差异直接解释为因果关系；
10. 输出要求：最终输出【统计方法选择说明】+【完整差异分析结果表】+【可直接用于论文结果部分的规范表述】。

明确差异分析需求 >

我先下载数据，然后进行完整的统计前提检验和差异分析。

整理结果并提供完整报告 >

差异分析全部完成。结果已保存至 `F:\静姐团队\年级差异分析结果.xlsx`（含 4 个 Sheet）。

一、统计方法选择说明

表格

变量	正态性 (Shapiro-Wilk)	方差齐性 (Leve
AI 周使用频次	不满足（大一 W=0.75,p<0.001；大二 W=0.82,p<0.001，呈右偏）	满足 (p=0.729)
AI 学习依赖度	满足 (各组 p>0.05)	满足 (p=0.828)
批判性思维得分	满足 (各组 p>0.05)	满足 (p=0.268)
元认知得分	满足 (各组 p>0.05)	满足 (p=0.910)

选择依据：AI 周使用频次在大一、大二年级显著偏离正态（存在 0 次和 3
非参数 Kruskal-Wallis 检验；其余三个变量均满足正态性和方差齐性，采

	A	B	C	D	E	F	G	H	I
	变量	年级	n	M	SD	Median	Min	Max	SE
	AI周使用	大一	126	5.91	3.1	5.8	0	30	0.276
	AI周使用	大二	124	6.47	3.1	6.3	0.6	28	0.279
	AI周使用	大三	123	7.34	2.22	7.2	1.3	12.5	0.2
	AI周使用	大四	125	7.76	2.18	7.8	3	14.9	0.195
	AI学习依	大一	126	2.87	0.64	2.95	1.1	4.2	0.057
	AI学习依	大二	124	2.87	0.65	2.9	1.2	4.8	0.059
	AI学习依	大三	123	2.85	0.62	2.8	1.5	4.4	0.056
	AI学习依	大四	125	2.89	0.61	2.9	1.2	4.3	0.055
	批判性思	大一	126	83.69	7.06	83.45	68	100	0.629
	批判性思	大二	124	81.8	6.54	81.65	63.1	100	0.587
	批判性思	大三	123	82.71	5.83	82.9	70.1	96.9	0.526
	批判性思	大四	125	82.59	6.62	82.1	67.2	100	0.592
	元认知得	大一	126	76.18	6.88	76.35	55.2	95.4	0.613
	元认知得	大二	124	75.22	6.92	75.25	56.7	92.3	0.622
	元认知得	大三	123	77.47	6.43	76.7	62	92.9	0.58
	元认知得	大四	125	76.41	6.5	77	59.8	93.7	0.581

三、总体检验结果

表格

变量	检验方法					
AI 周使用频次	Kruskal-Wallis					
AI 学习依赖度	ANOVA					
批判性思维得分	ANOVA					
元认知得分	ANOVA	F=2.36	3,494	0.071	η²=0.014	不显著 (边缘)

五、论文 "研究结果" 部分规范表述

以年级为分组变量，分别对 AI 周使用频次、AI 学习依赖度、批判性思维得分和元认知得分进行组间差异检验。Shapiro-Wilk 检验显示，AI 周使用频次在大一（W=0.75, p<0.001）和大二（W=0.82, p<0.001）组不满足正态性，故采用 Kruskal-Wallis H 检验；其余三个变量各组均满足正态性和方差齐性（Levene 检验 p 均> 0.05），采用单因素方差分析。

结果显示，AI 周使用频次在不同年级间存在显著差异（H (3)=55.71, p<0.001, ε²=0.107），呈随年级递增趋势（大一 5.91±3.10 次 / 周→大四 7.76±2.18 次 / 周）。Dunn 事后检验（Bonferroni 校正）表明，大一与大三（p<0.001）、大一与大四（p<0.001）、大二与大三（p=0.005）、大二与大四（p<0.001）差异显著，而大一与大二、大三与大四之间差异不显著。AI 学习依赖度在不同年级间无显著差异（F (3,494)=0.10, p=0.961, η²=0.001）。批判性思维得分的年级差异亦不显著（F (3,494)=1.77, p=0.152, η²=0.011）。元认知得分呈现边缘显著趋势（F (3,494)=2.36, p=0.071, η²=0.014），大三学生得分最高（77.47±6.43），但未达到 0.05 显著性水平。

豆包能够针对不同变量分别检查统计前提，并自主选择ANOVA或Kruskal-Wallis等合适方法，在总体显著后继续完成事后比较和多重校正，而不是机械套用同一种检验。

数据处理——相关分析

测试模式：工作任务 Turbo

提示词公式

请基于我上传的【数据表】，围绕【研究问题】，分析【变量1】【变量2】【变量3】等核心变量之间的相关关系。不要提前固定使用Pearson相关，请先检查变量类型、数据分布和异常值情况，再为不同变量组合选择最合适的相关分析方法。具体要求如下：

1. 变量识别：明确需要分析的【核心变量】，判断其属于连续变量、有序变量还是其他类型；
2. 数据检查：检查样本量、缺失值、异常值和基本分布情况；
3. 方法选择：根据变量类型、分布特征和线性关系条件，自主选择【Pearson相关/Spearman相关/Kendall相关/其他适合方法】，并说明选择依据；
4. 相关计算：计算变量之间的相关系数，并报告样本量、相关系数和P值；
5. 强度判断：根据实际相关系数判断关联方向和强弱，不得只用“显著/不显著”概括结果；
6. 多重比较：如果同时检验较多变量之间的相关关系，判断是否需要进行多重比较校正，并说明处理方式；
7. 异常影响：如存在明显异常值，判断其是否可能显著影响相关结果，必要时进行敏感性比较；
8. 结果筛选：从全部相关结果中提炼与【核心研究问题】最相关的关系，不需要机械逐项描述整个相关矩阵；
9. 结论边界：相关结果只能说明变量间关联，不得直接表述为“影响”“导致”或其他因果关系；
10. 输出要求：最终输出【方法选择说明】+【完整相关分析结果表】+【关键关系解读】+【可直接用于论文结果部分的规范表述】。

一、方法选择说明

表格

项目	说明
主要方法	**Spearman 秩相关**
辅助方法	Pearson 积差相关（用于对比验证
选择依据	AI 周使用频次（偏度 = 2.01，W= 用频次存在 6 个 IQR 极端值；Spe
多重比较校正	Bonferroni 校正，共 15 对检验，
校正前 / 后显著	校正前 10 对显著，校正后**6 对显**

所有变量无缺失值（n=498），异常值敏感性分析显示关

二、Spearman 相关系数矩阵

表格

	AI 周使用频次	AI 学习依赖度
AI 周使用频次	1.000	0.083
AI 学习依赖度	0.083	1.000
AI 素养	0.071	-0.102
批判性思维	0.011	**-0.348**
元认知	0.089	**-0.198**
学习投入	**0.194**	-0.112

加粗为 Bonferroni 校正后显著（p<0.0033）。

三、完整相关结果（校正后显著的 6 对）

表格

变量对	n	Spearman r	P 值	校正 P	方向	强弱
AI 学习依赖度 ↔ 批判性思维	498	-0.348	<0.001	<0.001	负	中等
AI 学习依赖度 ↔ 元认知	498	-0.198	<0.001	<0.001	负	弱
AI 素养 ↔ 批判性思维	498	0.203	<0.001	<0.001	正	弱
AI 素养 ↔ 元认知	498	0.154	0.0005	0.008	正	弱
元认知 ↔ 批判性思维	498	0.188	<0.001	<0.001	正	弱
AI 周使用频次 ↔ 学习投入	498	0.194	<0.001	<0.001	正	弱

其余 9 对校正后不显著（含 AI 周使用频次↔AI 学习依赖度 r=0.083,p=0.064；AI 周使用频次↔批判性思维 r=0.011,p=0.807 等）。

四、关键关系解读

1. AI 学习依赖度与批判性思维（r=-0.348，中等负相关）
这是本研究中最强的关联。AI 学习依赖度越高的学生，批判性思维得分倾向于越低。效应量达到中等水平（|r|>0.3），提示这一关联具有实际意义，而非仅统计显著。但横截面数据不能推断“AI 依赖导致 CT 下降”，也可能是批判性思维较弱的学生更倾向于依赖 AI。

2. AI 学习依赖度与元认知（r=-0.198，弱负相关）
AI 学习依赖度与元认知呈显著负相关，但效应量较弱。高 AI 依赖可能与较少的自我监控和反思活动相关，但关联强度有限。

3. AI 素养与批判性思维（r=0.203，弱正相关）
AI 素养较高的学生批判性思维得分也较高。这可能反映 AI 素养包含对 AI 生成内容的辨识、质疑能力，与批判性思维技能有重叠。

4. AI 素养与元认知（r=0.154，弱正相关）
AI 素养与元认知呈显著正相关，效应量较弱。较高的 AI 素养可能伴随更好的自我监控和学习调节能力。

A	B	C	D	E
	AI周使用频	AI学习依赖	AI素养	批判性思维
AI周使用频次	1	0.0831	0.071	0.011
AI学习依赖度	0.0831	1	-0.102	-0.3482
AI素养	0.071	-0.102	1	0.2032
批判性思维	0.011	-0.3482	0.2032	1
元认知	0.0891	-0.1981	0.1544	0.1876
学习投入	0.1939	-0.1119	0.0552	0.047

A	B	C	D	E	F	G
	AI周使用	AI学习依	AI素养	批判性思	元认知	学习投入
AI周使用频次	0	0.063814	0.113583	0.807117	0.046896	1.31E-05
AI学习依赖度	0.063814	0	0.022832	1.21E-15	8.39E-06	0.012439
AI素养	0.113583	0.022832	0	4.88E-06	0.000546	0.218723
批判性思维	0.807117	1.21E-15	4.88E-06	0	2.51E-05	0.295141
元认知	0.046896	8.39E-06	0.000546	2.51E-05	0	0.013813
学习投入	1.31E-05	0.012439	0.218723	0.295141	0.013813	0

不只计算相关系数，还能根据数据条件选方法、控制多重比较，并检验关键相关结果是否受异常值影响。

数据处理——回归分析

测试模式：工作任务 Turbo

提示词公式

请基于我上传的【数据表】，围绕【研究问题】，分析【核心自变量】与【因变量】之间的关系，并根据数据结构自主建立合适的回归模型。
不要提前固定使用哪一种回归方法，请先判断因变量类型、变量分布、共线性和模型假设，再选择最合适的分析方案。具体要求如下：

1. 变量识别：明确【因变量】【核心自变量】【控制变量】，判断变量属于连续、二分类、多分类或有序变量；
2. 模型选择：根据因变量类型和研究目的，自主选择【线性回归/Logistic回归/有序Logistic/其他适合模型】，并说明依据；
3. 数据检查：检查缺失值、异常值和变量分布，明确进入模型的有效样本量；
4. 共线性检查：检查核心自变量和控制变量之间的多重共线性，报告【VIF/容忍度等】指标；
5. 模型假设：根据所用模型检查必要条件，如线性关系、残差正态性、同方差性、独立性等；
6. 模型构建：根据研究问题合理安排变量进入模型的顺序，如【基础模型→加入核心变量→加入控制变量】；
7. 结果报告：规范输出回归系数、标准误、标准化系数或OR值、置信区间、P值和模型拟合指标；
8. 效应解释：解释核心变量的方向、大小和实际含义，不得只写“显著/不显著”；
9. 稳健性判断：如发现异常值、共线性或假设违反，说明采取了什么处理，并判断核心结果是否稳定；
10. 结论边界：如研究为横截面数据，不得把回归结果直接解释为因果效应；
11. 输出要求：最终输出【模型选择说明】+【模型诊断结果】+【完整回归结果表】+【关键结果解释】+【可直接用于论文结果部分的规范表述】。

明确回归分析核心变量设定 >

我先进行数据准备和模型诊断，再执行分层回归分析。

安装statsmodels统计分析库 >

需要先安装 statsmodels 库。

完成statsmodels安装并重启分析 >

statsmodels 安装完成，重新运行回归分析。

完成回归分析并整理报告 >

回归分析全部完成，结果已保存至 `F:\静姐团队\回归分析结果.xlsx`（含

一、模型选择说明

表格

项目	说明
因变量	批判性思维得分（连续变量，0-100 分）
模型类型	**普通最小二乘多元线性回归（OLS）**
选择依据	因变量为连续型，模型诊断显示残差满足正态性、
有效样本量	498（无缺失值）
分类变量编码	年级有序编码（大一 = 1... 大四 = 4）；性别女 =
控制变量纳入	全部纳入候选控制变量（年级、性别、专业、使用

二、模型诊断结果（模型 4）

表格

诊断项目	统计量	结果
残差正态性 (Shapiro-Wilk)	W=0.997, p=0.483	**满足正态性**
同方差性 (Breusch-Pagan)	LM=8.80, p=0.640	**满足同方差**
残差独立性 (Durbin-Watson)	DW=1.930	**无自相关**（接近 2）
多重共线性 (VIF)	最大 VIF=1.568	**无严重共线性**（<5）
强影响点 (Cook's d>4/n)	20 个，最大 d=0.019	无极端强影响点
残差均值	≈0	线性关系满足

稳健性检验：剔除 20 个强影响点后（n=478），AI 学习依赖度 B 从 - 3.52 变为 - 3.23（p<0.001），AI 素养 B 从 1.59 变为 1.52（p<0.001），方向和显著性均未改变，核心结果稳健。

三、分层回归模型汇总

表格

模型	预测变量数	R^2	调整 R^2	F(df)	P 值	ΔR^2	ΔR^2 显著性
模型 1：基本控制（年级 / 性别 / 专业）	6	0.014	0.002	1.19(6,491)	0.308	—	—
模型 2：+AI 周使用频次	7	0.015	0.001	1.06(7,490)	0.387	0.001	不显著 (p=0.598)
模型 3：+AI 学习依赖度、AI 素养	9	**0.185**	**0.170**	12.30(9,488)	<0.001	**0.170**	**显著 (p<0.001)**
模型 4：+ 元认知、学习投入	11	0.192	0.173	10.46(11,486)	<0.001	0.007	不显著 (p=0.140)

关键发现：模型 1 和模型 2 的解释力极低（R^2<0.02）且不显著；加入 AI 学习依赖度和 AI 素养后，R^2 从 0.015 跃升至 0.185（ΔR^2=0.170），模型解释力的提升几乎全部来自这两个核心自变量 ↓ 一步加入元认知和学习投入后，R^2 仅增加 0.007 且不显著。

	A	B	C	D	E	F	G	H
	模型	预测变量数	R^2	调整R^2	F值	F自由度	P值	Δ R^2
	模型1（基本控制）	6	0.0144	0.0023	1.1945	6, 491	0.307750	0.0144
	模型2（+使用频次）	7	0.0149	0.0009	1.062	7, 490	0.386918	0.0006
	模型3（+核心变量）	9	0.1849	0.1699	12.2996	9, 488	0.000000	0.17
	模型4（+元认知/投入）	11	0.1915	0.1732	10.4628	11, 486	0.000000	0.0066

豆包可自主完成分层回归建模，并同步检查多重共线性、残差假设和影响点，同时比较核心变量在不同模型中的系数变化，判断其独立关联是否稳定。

数据挖掘——聚类分析

测试模式：工作任务 Turbo

提示词公式

请基于我上传的【数据表】，围绕【研究目标】，对【研究对象】开展聚类分析，识别具有不同特征的潜在群体。不要提前固定聚类方法和类别数量，请先检查变量类型、量纲、缺失值和异常值，再自主选择适合的聚类方案。具体要求如下：

1. 变量选择：围绕【研究目标】筛选适合参与聚类的变量，排除ID、日期、结果标签等不适合直接用于聚类的字段；
2. 数据检查：检查参与聚类变量的缺失值、异常值、分布和量纲差异；
3. 预处理：根据需要完成标准化、归一化或其他必要处理，并说明原因；
4. 方法选择：根据样本量、变量类型和数据结构，自主选择【K-means/层次聚类/GMM/DBSCAN/其他方法】，并说明依据；
5. 聚类数判断：不要提前指定类别数量，可结合【肘部法、轮廓系数、Calinski-Harabasz指数、Davies-Bouldin指数等】判断合理聚类数；
6. 方案比较：如存在多个可行方案，应比较不同类别数或不同聚类方法的结果，说明最终方案选择依据；
7. 聚类结果：报告每一类样本数量及占比，并给出各核心变量的中心值或代表性统计特征；
8. 群体画像：根据各类真实数据特征为不同群体提炼简洁、具有研究含义的画像，不得仅按某一个变量高低命名；
9. 稳定性判断：检查聚类结果是否受随机初始化、异常值或变量选择明显影响，必要时进行重复聚类或敏感性比较；
10. 结果边界：聚类结果属于探索性模式发现，不得把人为命名的类别直接视为客观固定人群类型；
11. 输出要求：最终输出【方法选择说明】+【聚类方案比较】+【最终聚类结果表】+【各类群体画像】+【可直接用于研究结果部分的规范表述】。



三、聚类方案比较

表格

k	方法	轮廓系数	CH	DB	与 K-mea
2	K-means	0.1475	92.70	2.190	—
2	层次聚类 (Ward)	0.0997	61.51	2.537	0.438
2	GMM	0.2595	12.58	2.516	-0.0002
3	K-means	0.1252	73.73	2.125	—
3	层次聚类	0.1023	55.43	2.315	0.318
3	GMM	0.1111	37.50	2.486	0.192
4	K-means	0.1145	63.55	2.083	—
4	层次聚类	0.0864	49.00	2.356	0.327
4	GMM	0.0559	29.40	2.837	0.118

最终选择：K-means, k=2

- K-means 在所有 k 下轮廓系数和 CH 指数均优于层次聚类
- GMM 在 k=2 时轮廓系数虽高（0.260），但 CH 指数极低（12.58），且类结构不稳定或不适合本数据的非高斯分布
- K-means 在不同随机初始化下高度稳定（见下文）

七、论文 "研究结果" 部分规范表述

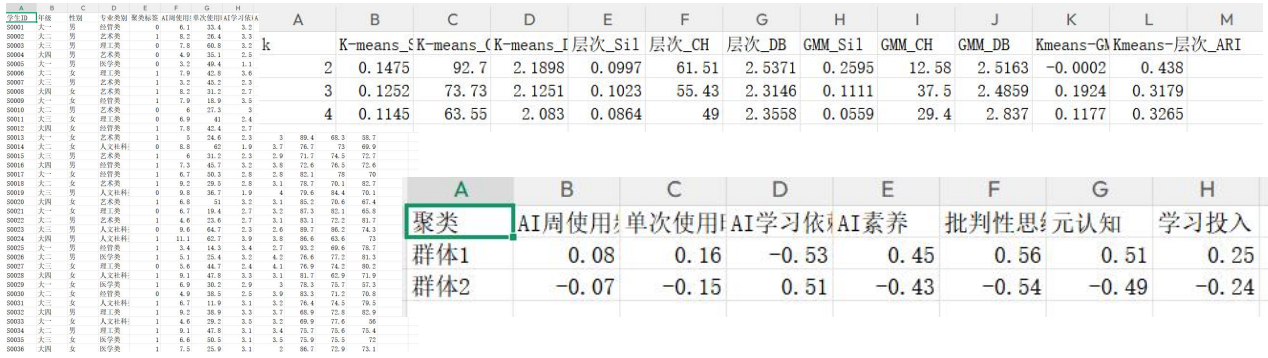
采用 K-means 聚类分析识别大学生 AI 学习使用与认知特征的潜在群体。聚类变量包括 AI 周使用频次、单次使用时长、AI 学习依赖度、AI 素养、批判性思维、元认知和学习投入共 7 个连续变量，均进行 z-score 标准化处理以消除量纲差异。有效样本量为 498，所有变量无缺失值。

通过肘部法、轮廓系数、Calinski-Harabasz（CH）指数和 Davies-Bouldin（DB）指数综合判断最优类别数。结果显示，当 k=2 时轮廓系数最高（0.148），CH 指数最高（92.70），肘部法亦显示 k=2 为明显拐点（k=2→3 时 SSE 下降 8.6%，之后下降率持续放缓），各项指标一致支持二类别方案。同时比较了 K-means、层次聚类（Ward 法）和高斯混合模型（GMM）三种方法，K-means 在各 k 值下的轮廓系数和 CH 指数均优于层次聚类；GMM 在 k=2 时虽轮廓系数较高，但 CH 指数极低且与 K-means 的调整兰德指数（ARI）接近 0，表明其聚类结构不稳定，故最终选择 K-means 方法。

稳定性检验显示，10 次不同随机初始化的聚类结果与基准标签的平均 ARI 为 0.998（最小 0.992），轮廓系数完全一致；剔除 22 个多变量异常值后（n=476），轮廓系数为 0.141，两个群体的大小比例（51.1% vs 48.9%）与原始结果（49.0% vs 51.0%）基本一致，表明聚类结果稳健。

最终识别出两个学生群体。群体 1（n=244, 49.0%）表现为较高的 AI 素养（M=3.55, +0.45σ）、较低的 AI 学习依赖度（M=2.53, -0.53σ）、较高的批判性思维（M=86.36, +0.56σ）和元认知（M=79.71, +0.51σ），可命名为 "高素养低依赖型"—— 该群体学生虽然 AI 使用频次和时长略高于总体，但不依赖 AI 直接获取答案，善于批判性和元认知性地使用 AI，认知表现整体优于总体。群体 2（n=254, 51.0%）表现为较高的 AI 学习依赖度（M=3.19, +0.51σ）、较低的 AI 素养（M=2.94, -0.43σ）、较低的批判性思维（M=79.18, -0.54σ）和元认知（M=73.05, -0.49σ），可命名为 "高依赖低素养型"—— 该群体学生 AI 使用频次和时长略低于总体，但更倾向于依赖 AI 直接获取答案，认知表现整体低于总体。值得注意的是，两个群体在 AI 使用频次（z=0.08 vs -0.07）和单次使用时长（z=0.16 vs -0.15）上差异很小，但在 AI 依赖度、AI 素养和认知表现上差异显著（|z|>0.4），提示 AI 使用方式而非使用量是区分学生群体的关键特征。

卡方检验显示，两个群体在年级（$\chi^2(3)=7.18$, p=0.066）、性别（$\chi^2(1)=2.36$, p=0.125）和专业类别（$\chi^2(4)=4.85$, p=0.303）构成上均无显著差异，说明聚类结果主要由 AI 使用方式和认知特征驱动，而非人口学背景差异。



k	K-means_S	K-means_C	K-means_I	层次_Sil	层次_CH	层次_DB	GMM_Sil	GMM_CH	GMM_DB	Kmeans-GM	Kmeans-层次_ARI
2	0.1475	92.7	2.1898	0.0997	61.51	2.5371	0.2595	12.58	2.5163	-0.0002	0.438
3	0.1252	73.73	2.1251	0.1023	55.43	2.3146	0.1111	37.5	2.4859	0.1924	0.3179
4	0.1145	63.55	2.083	0.0864	49	2.3558	0.0559	29.4	2.837	0.1177	0.3265

聚类	AI周使用	单次使用	AI学习依	AI素养	批判性思	元认知	学习投入
群体1	0.08	0.16	-0.53	0.45	0.56	0.51	0.25
群体2	-0.07	-0.15	0.51	-0.43	-0.54	-0.49	-0.24

豆包可自主比较不同聚类方法和类别数量，并结合多项聚类指标选择方案，再依据多维特征形成群体画像，实现从无标签数据中发现潜在样本结构。

数据挖掘——分类预测

测试模式：工作任务 Turbo

提示词公式

请基于我上传的【数据表】，围绕【预测目标】，利用【候选特征变量】建立分类预测模型。不要提前固定使用某一种算法，请先检查目标变量、特征变量、类别分布和数据质量，再自主选择适合的分类模型并进行比较。具体要求如下：

1. 目标识别：明确【目标变量】及其类别含义，判断属于二分类还是多分类任务；
2. 特征筛选：从【候选变量】中筛选适合建模的特征，排除ID、日期、结果泄漏变量等不应进入模型的字段；
3. 数据检查：检查缺失值、异常值、类别是否均衡，以及连续变量和分类变量的编码情况；
4. 数据划分：合理划分训练集和测试集，必要时采用分层抽样，避免训练集和测试集类别比例失衡；
5. 模型选择：根据任务特点自主比较【Logistic回归/决策树/随机森林/XGBoost/SVM/其他适合模型】，不得只运行单一模型；
6. 模型验证：采用交叉验证评估模型稳定性，避免只根据一次训练结果判断模型优劣；
7. 性能评价：根据任务特点报告【Accuracy、Precision、Recall、F1、ROC-AUC等】指标，类别不平衡时不得只看准确率；
8. 过拟合检查：比较训练集、验证集和测试集性能，判断是否存在明显过拟合；
   模型比较：综合预测性能、稳定性和模型复杂度选择最终模型，并说明依据；
   结果解释：分析哪些变量对预测最重要，但不得将预测贡献直接解释为因果关系；
9. 输出要求：最终输出【建模方案说明】+【模型性能比较表】+【最佳模型结果】+【关键特征解释】+【可直接用于研究结果部分的规范表述】。

一、建模方案说明

高依赖分组规则

表格

项目	说明
目标变量	是否属于较高 AI 学习依赖群体
分组阈值	**3.20 分（67% 分位数）**
高依赖定义	AI 学习依赖度 ≥ 3.20
非高依赖定义	AI 学习依赖度 < 3.20
高依赖样本	n=171（34.3%）
非高依赖样本	n=327（65.7%）

选择依据：AI 学习依赖度量表为 1-5 分，均值 2.87、中位数 2.90。"较高依赖"语义上对应分保证高依赖群体有足够区分度（非仅略高于量表中点），又使类别比例约为 1:2，避免过度不平表中点，区分度不足。

预测特征（10 个变量，13 个特征）

- **数值特征（7 个）**：AI 周使用频次、单次使用时长、AI 素养、批判性思维、元认知、学习投
- **分类特征（3 个）**：年级（有序编码）、性别（0/1）、专业类别（哑变量）
- **排除**：AI 学习依赖度本身及其直接衍生字段（避免数据泄漏）

数据划分

- 分层随机抽样 7:3 划分，训练集 n=348（高依赖 34.2%），测试集 n=150（高依赖 34.7%
- 类别比例在训练 / 测试集中保持一致

二、模型性能比较

表格

模型	测试 Acc	测试 Precision	测试 Recall	测试 F1	测试 AUC	5 折 CV AUC	过拟合度（训练 - 测试 AUC）
Logistic 回归	0.607	0.451	**0.615**	**0.520**	0.641	**0.696±0.026**	**0.103**
决策树	0.567	0.377	0.385	0.381	0.598	0.600±0.037	0.229
随机森林	0.613	0.446	0.481	0.463	**0.654**	0.659±0.028	0.325
梯度提升树	**0.667**	**0.528**	0.365	0.432	0.642	0.642±0.043	0.356

关键观察

1. **所有模型均存在过拟合**：训练集 AUC（0.74-0.998）远高于测试集 AUC（0.60-0.65），其中梯度提升树和随机森林过拟合最严重（>0.32），Logistic 回归过拟合最轻（0.103）。
2. **整体预测性能有限**：所有模型测试 AUC 在 0.60-0.65 之间，属于较低水平，说明仅用现有特征预测高 AI 学习依赖的区分能力有限。
3. **指标关注点**：由于类别不平衡（高依赖 34%），Accuracy 具有误导性（全预测非高依赖可达 65.7% 准确率），应重点关注**Recall、F1 和 AUC**。Logistic 回归在 Recall（0.615）和 F1（0.520）上表现最好，能识别更多真正的高依赖学生。

三、最佳模型选择

综合推荐：Logistic 回归

表格

选择维度	理由
泛化稳定性	5 折 CV AUC 最高（0.696），标准差最小（0.026）
过拟合控制	训练 - 测试 AUC 差异最小（0.103），远优于树模型（>0.23）
实际效用	Recall 最高（0.615），能识别更多高依赖学生，F1 最高（0.520）
模型复杂度	最简单，可解释性最强，可直接解读各特征的预测方向和大小
性能差距	测试 AUC（0.641）与随机森林（0.654）差距仅 0.013，可接受

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	模型	训练Acc	训练F1	训练AUC	测试Acc	测试Preci	测试Recal	测试F1	测试AUC	CV_AUC_me	CV_AUC_st	过拟合度(AUC差)	
2	Logistic	0.6839	0.5956	0.7442	0.6067	0.4507	0.6154	0.5203	0.6413	0.6961	0.0258	0.1029	
3	决策树	0.7557	0.6502	0.8265	0.5667	0.3774	0.3846	0.381	0.5978	0.6002	0.0366	0.2287	
4	随机森林	0.8966	0.8605	0.979	0.6133	0.4464	0.4808	0.463	0.654	0.6594	0.0281	0.325	
5	梯度提升	0.9655	0.9469	0.9977	0.6667	0.5278	0.3654	0.4318	0.6421	0.6418	0.0425	0.3556	

豆包可自主定义分类目标、避免明显数据泄漏，并比较多种机器学习模型，通过测试集、交叉验证和过拟合指标评估预测性能，同时给出关键特征解释。

科研示意图--图形摘要

测试模式：工作任务 Turbo

提示词公式

请基于我提供的【论文摘要/研究内容/上传论文】，生成一张适用于科研表达的图形摘要。
你需要先理解研究的核心问题、研究方法、关键结果和研究意义，再将最重要的信息提炼并组织成一张结构清晰、逻辑完整的图形摘要，不要简单堆砌元素，也不要把论文内容逐字转成图片文字。
具体要求如下：

1. 内容提炼：从研究内容中提取最关键的信息，重点包括【研究背景】、【研究方法】、【核心结果】和【研究意义/应用价值】；
2. 结构组织：采用【布局方式，如从左到右/从上到下/中心辐射式】组织画面，体现研究逻辑的推进关系；
3. 核心元素：图中应包含【关键对象/关键变量/关键实验对象/关键技术环节】等必要元素；
4. 关系表达：通过箭头、流程连接、模块分区或图示符号等方式呈现研究过程与结果之间的关系；
5. 视觉风格：整体采用【风格要求，如简洁科研插画风格/BioRender风格/扁平化学术风格】，配色克制，突出科研感与可读性；
6. 文字要求：图中文字尽量精炼，使用短语而非长句，避免文字过多影响图示表达；
7. 输出要求：输出【图片数量，如1张】图形摘要，图像比例为【比例，如4:3/16:9/1:1】，背景为【白色背景/浅色背景】。



豆包的突出优势在于：**既能把论文内容直接转化为图形摘要，又能基于生成结果继续做局部修改与迭代优化，形成从内容理解到成图细化的一站式科研绘图闭环。**



科研示意图--作用机制

测试模式：工作任务 Turbo

提示词公式

请基于我提供的【研究内容/论文摘要/上传论文】，生成一张适用于科研表达的作用机制图。你需要先理解研究中的核心问题、关键变量、作用对象、中间机制和最终结果，再将这些内容组织成一张逻辑清晰、关系明确的作用机制图，不要只罗列概念，也不要把论文原文逐字搬到图中。具体要求如下：

1. 内容提炼：从研究内容中提取最关键的机制信息，包括【输入因素/起始因素】、【关键作用对象】、【中介机制/调节机制】和【最终结果】；
2. 结构组织：采用【布局方式，如从左到右/从上到下/中心发散式】组织画面，清晰体现因果路径和机制层次；
3. 核心元素：图中应包含【核心变量】【关键主体/对象】【中间作用路径】【结果变量】等必要元素；
4. 关系表达：通过【实线箭头】表示直接作用关系，通过【虚线箭头】表示间接作用或推测性关系，通过【T形抑制线】表示抑制作用；如有中介或调节效应，请明确标示；
5. 视觉风格：整体采用【风格要求，如简洁科研插画风格/BioRender风格/扁平化学术风格】，配色克制，突出科研感与逻辑清晰度；
6. 文字要求：图中文字尽量简洁，使用概念词或短语，避免长段文字堆积；
7. 输出要求：输出【图片数量，如1张】作用机制图，图像比例为【比例，如4:3/16:9/1:1】，背景为【白色背景/浅色背景】。

核心通路与抑制节点梳理：

- 主通路（激活）：LPS → TLR4 → MyD88 → IKK 复合物 → IκBα 降解 → NF-κB p65 核转位 → TNF-α/IL-6/IL-1β 转录 → 炎症加重
- Compound X 抑制（T 形线）：TLR4 激活、MyD88 招募、IKK 活化、IκBα 降解、p65 核转位，共 5 个节点

检查细胞机制图比例 ›

细胞作用机制图已生成，4:3 比例，白色背景，BioRender 风格，清晰区分胞外 / 细胞膜 / 胞质 / 细胞核四个区域。

给用户交付图的说明 ›

细胞作用机制图已生成，核心结构如下：

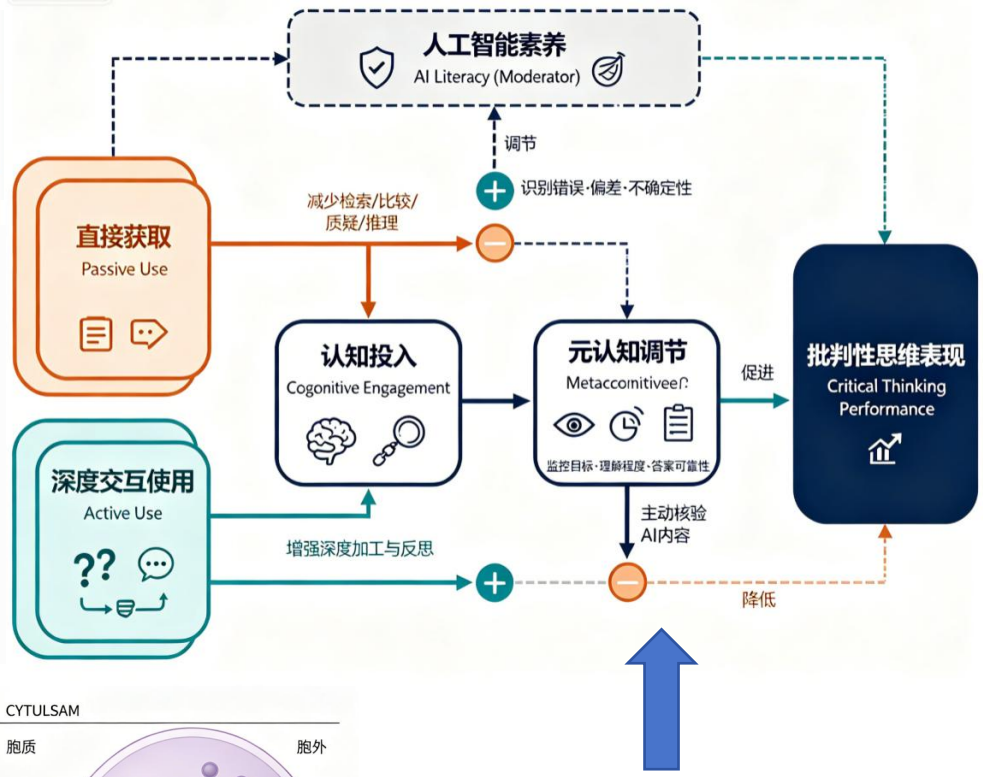
四个区域划分：胞外（顶部）→ 细胞膜（横向脂双层）→ 胞质（中部）→ 细胞核（右侧椭圆形核膜）

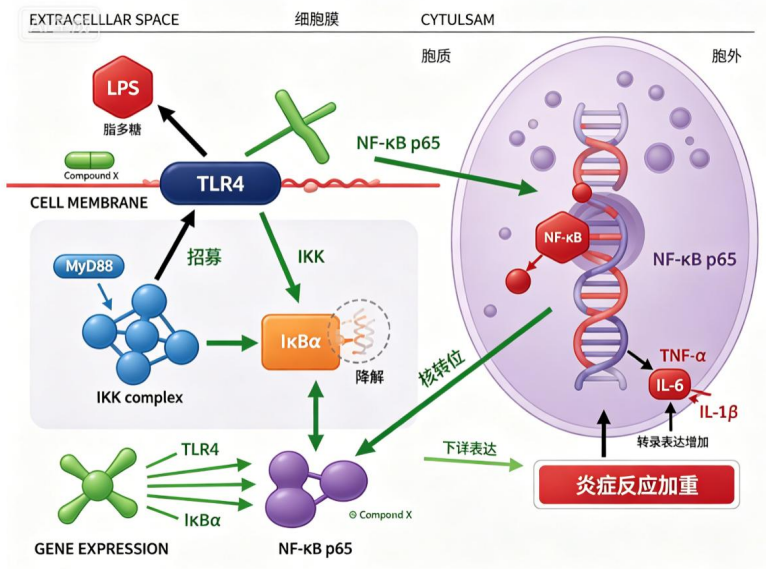
主激活通路（黑色实线箭头，从左到右）：
LPS（胞外）→ TLR4（跨膜受体）→ MyD88 招募 → IKK 复合物活化 → IκBα 磷酸化降解 → NF-κB p65 释放 → 核转位入核 → 结合 DNA → TNF-α / IL-6 / IL-1β 转录表达增加 → 炎症反应加重

Compound X 抑制作用（绿色 T 形抑制线，5 个关键节点）：

1. 抑制 TLR4 激活
2. 减少 MyD88 招募
3. 抑制 IKK 复合物活化
4. 阻止 IκBα 降解
5. 降低 NF-κB p65 核转位
   最终下调炎症因子表达，减轻炎症反应

配色克制：LPS 红色、受体 / 信号分子蓝色系、NF-κB 紫色、炎症因子红色、Compound X 绿色（抑制剂标准色）。如需调整分子图标样式、区域布局或补充细节，可以告诉我。





给用户说明交付的图 ›

作用机制图已生成，核心结构如下：

左侧（自变量）：两种 AI 使用方式上下并列

- 直接获取答案（橙色，被动使用）→ 实线箭头带减号 → 降低认知投入
- 深度交互使用（青色，主动使用）→ 实线箭头带加号 → 增强认知投入

中部（中介链）：认知投入 → 元认知调节（监控目标 / 理解程度 / 答案可靠性，含 "主动核验 AI 内容" 分支）→ 批判性思维表现

右侧（因变量）：批判性思维表现（深色结果框）

顶部（调节变量）：人工智能素养以虚线框 + 虚线箭头指向 "AI 使用方式→认知投入" 路径，标注 "识别错误 · 偏差 · 不确定性" 的调节作用

关系区分：实线箭头 = 主要作用路径，虚线箭头 = 调节作用；青色加号 = 促进，橙色减号 = 抑制。如需调整变量位置、箭头方向或配色，可以告诉我。

科研示意图--工作流程

测试模式：工作任务 Turbo

提示词公式

请根据我提供的【研究流程/实验方案/方法描述】，生成一张适用于科研表达的工作流程图（Workflow Diagram）。
你需要先理解整个研究过程，提炼关键阶段、核心步骤、输入信息、处理方法和输出结果，再将其组织成逻辑清晰、层级明确的科研流程图。具体要求如下：

1. 流程提炼：将研究过程划分为【流程阶段数量】个核心阶段，并提取每个阶段最关键的任务；
2. 结构组织：采用【横向流程/纵向流程/分区式流程】布局，按照研究实施顺序组织各阶段；
3. 模块内容：每个模块重点呈现【输入数据/处理方法/核心任务/输出结果】中的关键信息；
4. 关系表达：使用箭头清晰连接各步骤，如存在并行分析、分支任务或结果回流，应明确表示；
5. 重点突出：对【核心方法/关键步骤/创新环节】进行适当视觉强调；
6. 视觉风格：采用【简洁科研流程图风格/BioRender风格/学术信息图风格】，配色克制，层级清晰；
7. 文字要求：模块名称和说明采用简洁短语，避免大段文字；
8. 输出要求：生成【图片数量】张工作流程图，比例为【16:9/4:3等】，背景为【白色/浅色】。

我先梳理 6 个阶段的数据流转，再生成这张竖版工作流程图。

6 阶段数据流转梳理：
原始数据（采集）→ 清洗数据（预处理）→ 特征集（特征分析）→ 候选模型（模型训练，4 模型 + 交叉验证）→ 最优模型（模型评价，5 指标比较）→ 解释结果（SHAP 解释）

9:16 竖版画布采用两行蛇形布局，每行 3 个阶段。

检查工作流程图比例 ›

工作流程图已生成，9:16 竖版比例，白色背景，蛇形两行布局呈现 6 阶段完整流程。

多组学联合分析工作流程图已生成，核心结构如下：

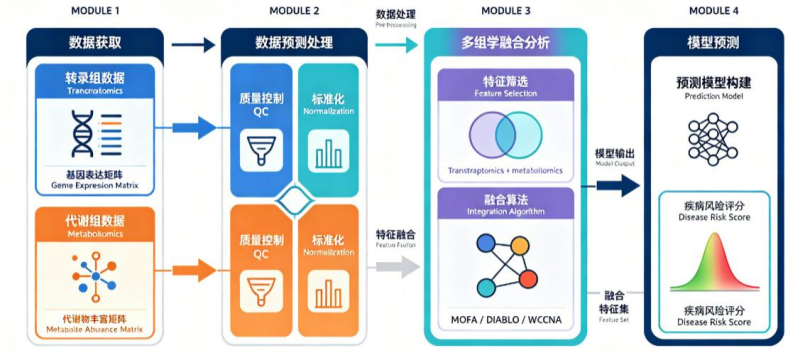
四模块横向布局（16:9 宽屏）：

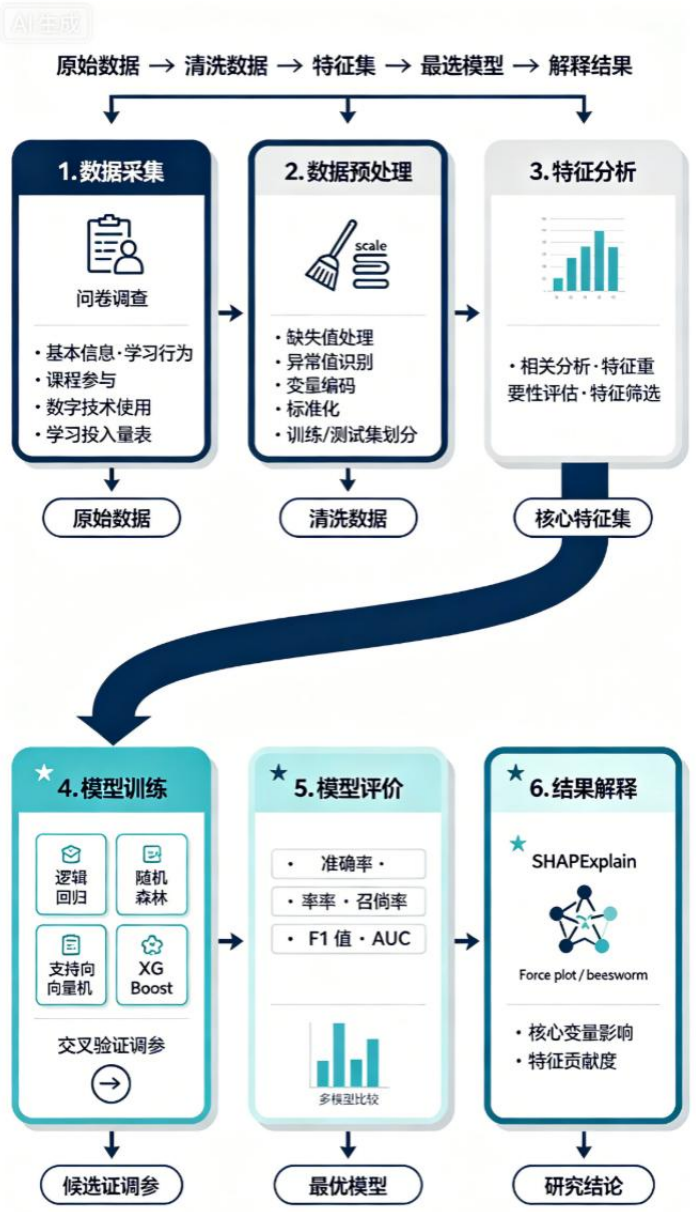
模块 1 数据获取（最左）：双轨并行输入

- 转录组数据（蓝色，DNA/RNA 图标，基因表达矩阵）
- 代谢组数据（橙色，分子 / 质谱图标，代谢物丰度矩阵）

模块 2 数据预处理（左中）：双轨并行处理后汇聚

- 转录组：QC 质量控制 + 标准化
- 代谢组：QC 质量控制 + 标准化
- 两轨在模块右侧合并为单一输出





科研示意图--技术路线图生成

测试模式：工作任务 Turbo

提示词公式

你是一位精通【研究领域】的科研绘图专家。请根据以下研究内容，为我设计一张期刊级别的技术路线图（Research Roadmap），用于清晰展示研究从【起始环节】到【最终目标】的整体研究路径。
研究内容：【用一段文字描述研究对象、数据来源、主要研究方法、分析步骤和最终产出】
具体要求如下：
1. 流程结构：按照“【阶段1】—【阶段2】—【阶段3】—【阶段4】—【最终输出】”组织整体技术路线。
2. 模块设置：将【不同数据/研究对象/实验内容】设置为并行或分支模块，并在【关键阶段】进行汇聚。
3. 信息关系：使用箭头清晰表示各研究阶段之间的先后关系、并行关系和汇聚关系。
4. 重点突出：重点展示【核心方法/关键技术/关键分析环节】，弱化非核心辅助步骤。
5. 文字表达：每个模块以简洁关键词呈现，避免大段文字。
6. 视觉风格：整体采用BioRender风格科研插图，配色克制、结构清晰、层级明确。
7. 输出规格：图像比例16:9，白色背景，适用于论文、基金申请书或科研汇报。

你是一位精通医学人工智能与多模态数据分析的科研绘图专家。请根据以下研究内容，为我设计一张期刊级别的技术路线图（Research Roadmap），用于清晰展示研究从多源数据获取到疾病风险预测与临床应用的整体研究路径。
研究内容：围绕人工智能辅助疾病风险预测开展研究。首先收集临床数据、医学影像数据和多组学数据，并分别进行数据清洗、标准化和特征提取；随后构建临床特征、影像特征和组学特征体系，通过多模态融合方法形成综合特征；在此基础上构建疾病风险预测模型，并通过内部验证和外部验证评价模型性能；最终形成疾病风险分层结果和临床辅助决策方案。
具体要求如下：
整体按照“数据获取 — 数据处理 — 特征构建 — 多模态融合 — 模型构建 — 模型验证 — 结果输出”的技术路线展开。
临床数据、医学影像数据和多组学数据设置为三个并行分支，分别完成预处理和特征提取后汇聚至多模态融合模块。
使用箭头清晰表示各研究阶段之间的推进关系，并突出“多源数据 — 特征融合 — 模型预测 — 临床应用”的主线。
模型验证阶段设置“内部验证”和“外部验证”两个并列模块。
每个模块采用简洁关键词和科研图标表达，避免大段文字。
整体采用 BioRender 风格科研插图，配色克制、结构清晰、层级明确。
图像比例 9：16，白色背景，适用于论文、基金申请书和科研汇报。

明确技术路线图需求 >

我先梳理 7 阶段技术路线与三并行分支结构，再生成这张 9:16 竖版技术路线图。

结构梳理：

- 第 1-3 阶段：临床 / 影像 / 组学三列并行（数据获取→数据处理→特征构建）
- 第 4 阶段：三分支汇聚至多模态融合
- 第 5 阶段：模型构建
- 第 6 阶段：内部验证 + 外部验证双列并列
- 第 7 阶段：疾病风险分层 + 临床辅助决策

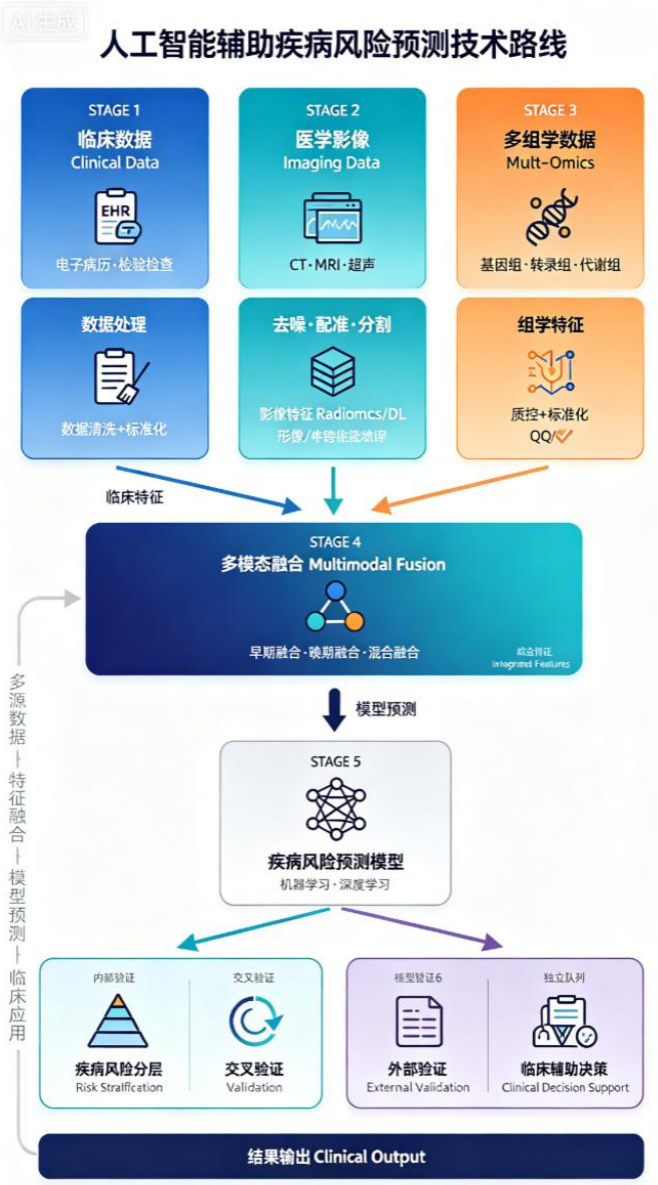
第 6 阶段 模型验证（双列并列）：

- 内部验证（交叉验证）
- 外部验证（独立队列）

第 7 阶段 结果输出（深色重点框）：

- 疾病风险分层（低 / 中 / 高）
- 临床辅助决策方案

左侧标注主线："多源数据→特征融合→模型预测→临床应用"。配色采用蓝 / 青 / 橙 / 深蓝的克制 BioRender 风格，适用于论文、基金申请书和科研汇报。如需调整具体模态、融合方法名称或补充模型类型，可以告诉我。

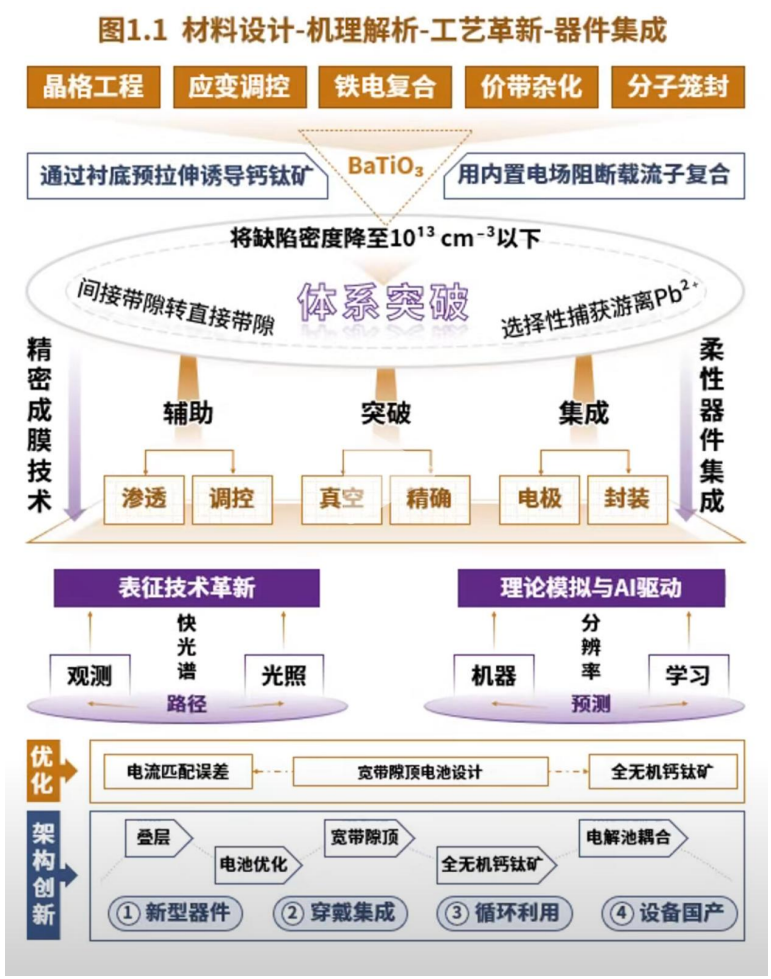


科研示意图--技术路线图仿绘

提示词公式

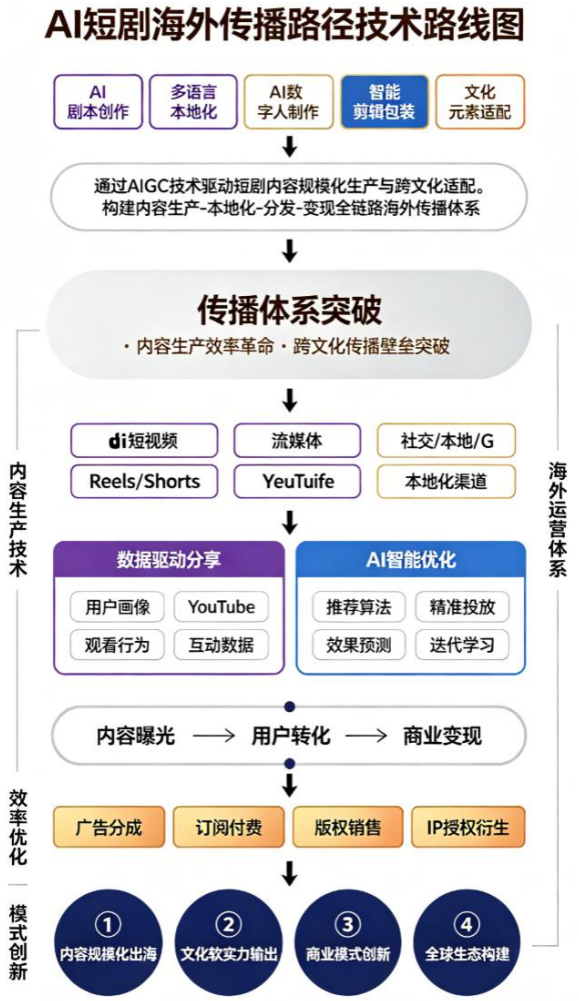
测试模式：工作任务 Turbo

上传【原图】，请仿照这张【技术路线图】，制作新的【技术路线图】，将内容转化成【主题/上传的研究内容】相关的



原图





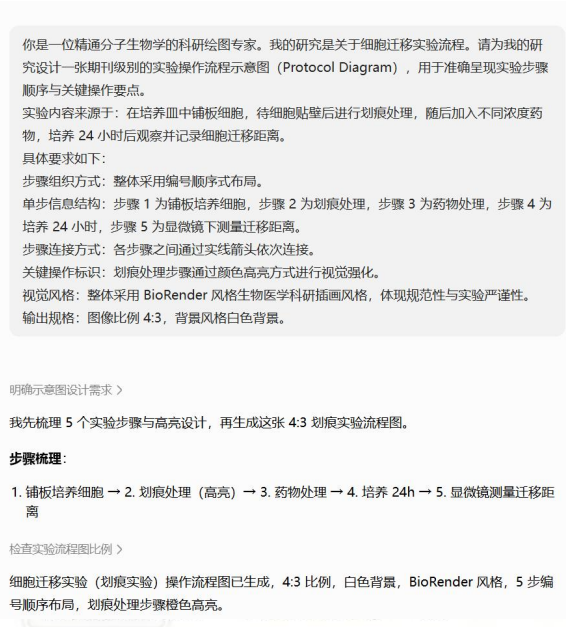
仿图

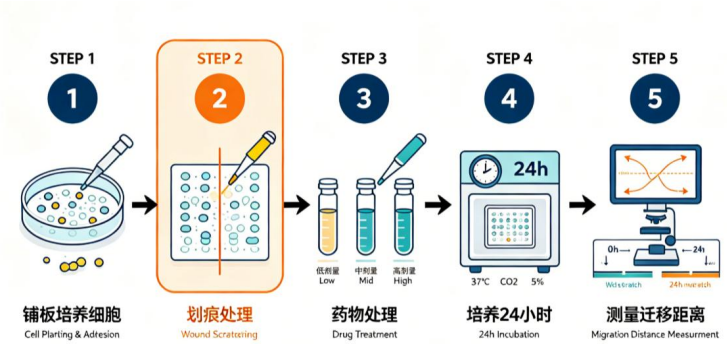
科研示意图--实验设计

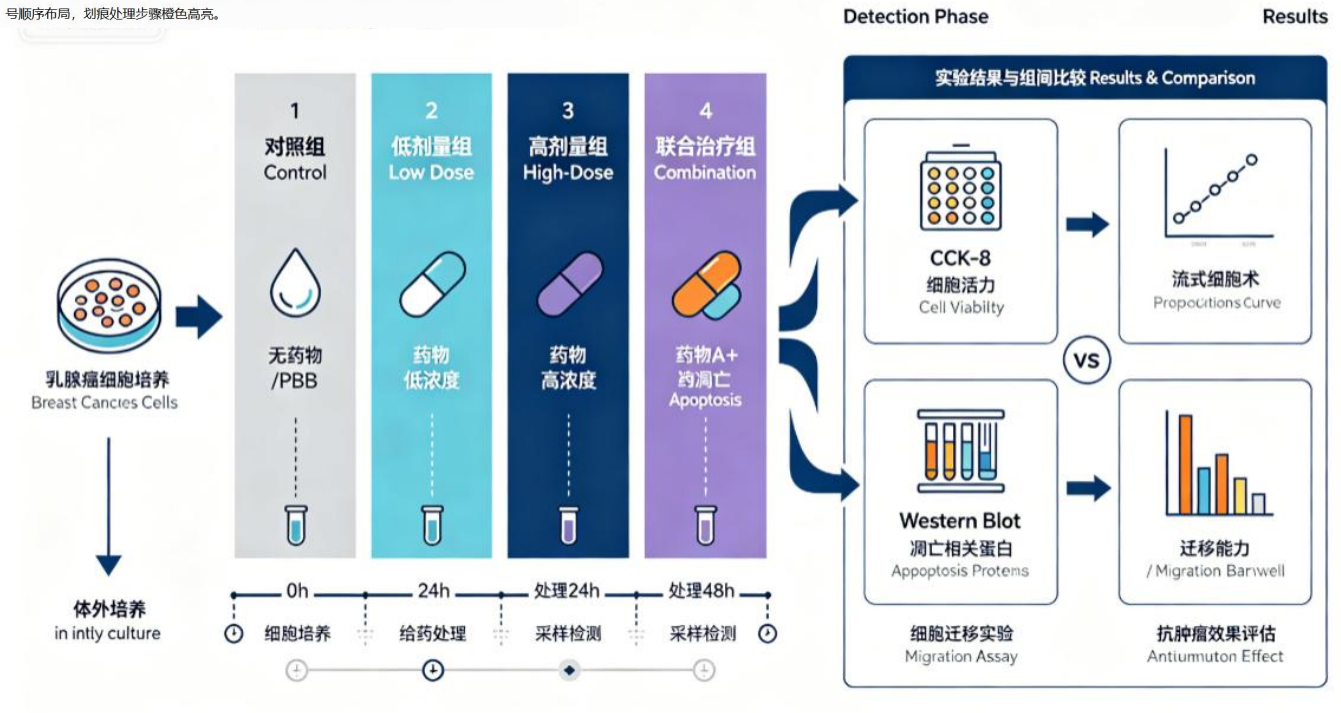
测试模式：工作任务 Turbo

提示词公式

你是一位精通【研究领域】的科研绘图专家。请根据以下实验方案，为我设计一张期刊级别的实验设计示意图，用于清晰展示实验对象、分组方式、处理条件、检测指标和实验流程。
实验内容：【用一段文字描述实验对象、样本数量、实验分组、处理方式、时间节点、检测方法和最终分析目标】
具体要求如下：
1. 实验对象：明确展示【细胞/动物/患者/样本】及其基本信息。
2. 分组设计：将实验划分为【对照组】、【实验组1】、【实验组2】等不同处理组，并清晰标注各组条件。
3. 实验流程：按照“【样本准备】—【分组处理】—【干预过程】—【样本采集】—【检测分析】”展示实验推进过程。
4. 时间节点：如涉及不同处理时间，请使用时间轴标注【时间点1】、【时间点2】、【时间点3】。
5. 检测指标：在实验末端展示【检测方法1】、【检测方法2】、【检测方法3】及对应分析指标。
6. 信息关系：通过箭头、分支和汇聚关系清晰表达不同实验组的处理路径及最终检测过程。
7. 视觉风格：整体采用BioRender风格科研插图，实验对象和实验操作使用简洁科研图标表示，配色克制、结构清晰。
8. 输出规格：图像比例16:9，白色背景，适用于论文、基金申请书和科研汇报。







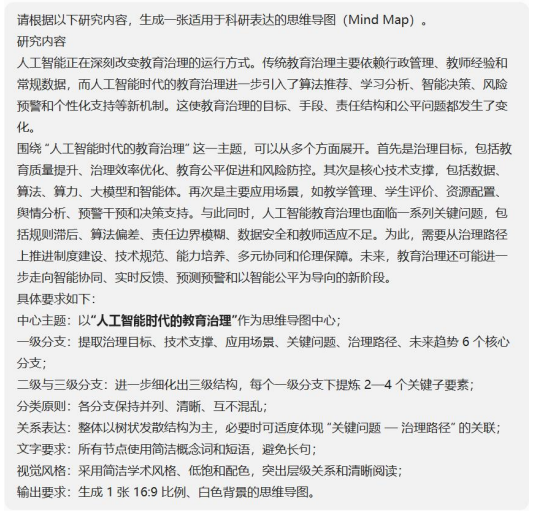
科研示意图--思维导图

测试模式：工作任务 Turbo

提示词公式

请根据我提供的【研究主题/研究内容/概念说明】，生成一张适用于科研表达的思维导图（Mind Map）。
你需要先理解研究主题的核心内容，再围绕中心主题提炼主要分支和子分支，将其组织成层级清晰、逻辑明确的思维导图，不要简单把原文切分成若干文本框。具体要求如下：

1. 中心主题：以【中心主题】作为思维导图中心；
2. 一级分支：围绕中心主题提取【一级分支数量】个核心分支；
3. 二级与三级分支：根据内容需要进一步细化出【2—4层】结构，突出主题展开逻辑；
4. 分类原则：各分支应具有清晰的分类依据，如【研究背景/核心概念/影响因素/研究路径/方法类型/应用场景/问题挑战/未来方向等】；
5. 关系表达：以树状发散结构为主，必要时可通过连接线适度体现不同分支之间的关联；
6. 文字要求：各节点使用简洁概念词或短语，避免完整长句，同一层级尽量保持表达方式一致；
7. 视觉风格：整体采用【简洁学术风格/低饱和科研信息图风格/清晰树状结构风格】，突出层级关系与可读性；
8. 输出要求：生成【数量】思维导图，比例为【16:9/4:3】，背景为【白色/浅色】。



给用户说明交付的图 ›

思维导图已生成，核心结构如下：

中心主题：人工智能时代的教育治理

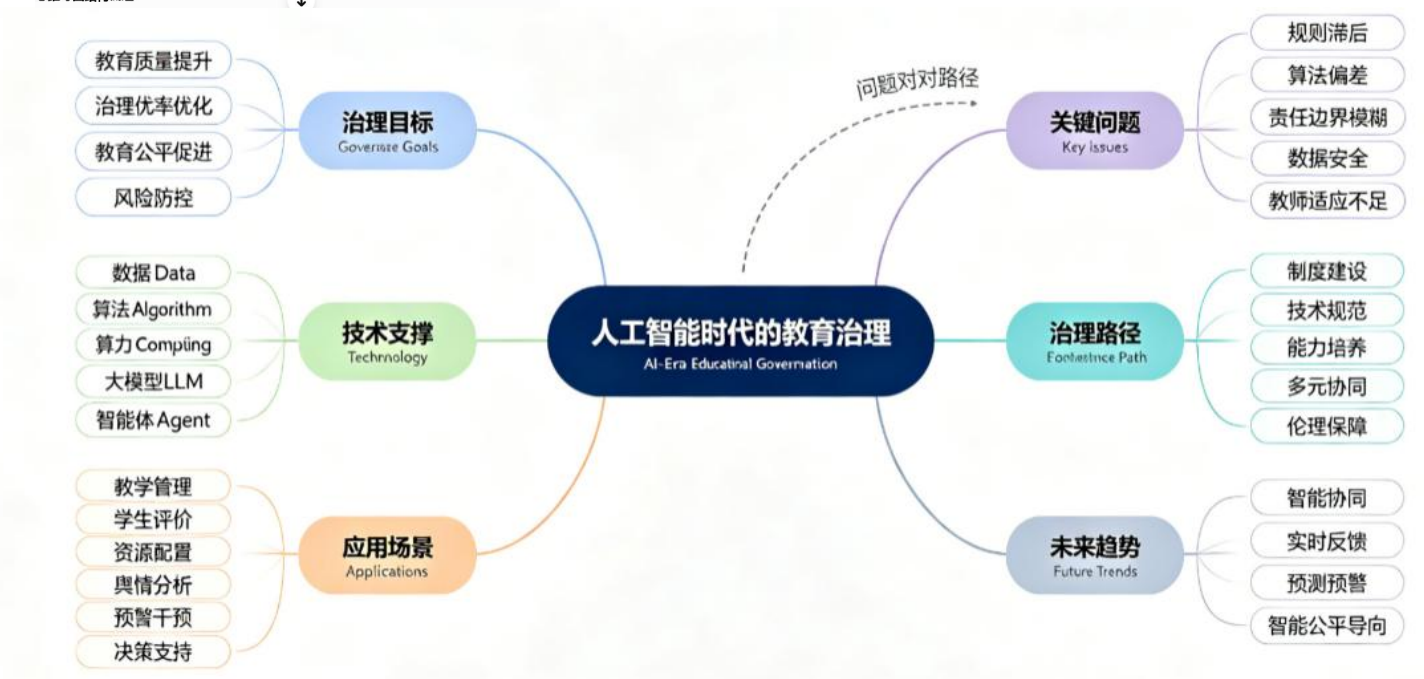
左侧 3 分支：

- **治理目标**（淡蓝）：教育质量提升 / 治理效率优化 / 教育公平促进 / 风险防控
- **技术支撑**（淡绿）：数据 / 算法 / 算力 / 大模型 / 智能体
- **应用场景**（淡橙）：教学管理 / 学生评价 / 资源配置 / 舆情分析 / 预警干预 / 决策支持

右侧 3 分支：

- **关键问题**（淡紫）：规则滞后 / 算法偏差 / 责任边界模糊 / 数据安全 / 教师适应不足
- **治理路径**（淡青）：制度建设 / 技术规范 / 能力培养 / 多元协同 / 伦理保障
- **未来趋势**（灰蓝）：智能协同 / 实时反馈 / 预测预警 / 智能公平导向

关联表达：关键问题与治理路径之间以虚线 "问题→路径" 连接，体现问题导向的治理回应关系。各分支采用低饱和差异化配色，层级清晰，所有节点均为简洁概念词。如需调整子项内容、分支配色或补充三级节点，可以告诉我。



科研数据图--描述型

测试模式：工作任务 Turbo

提示词公式

请基于我提供的【数据表/统计结果】，生成用于描述【研究对象/核心指标】整体特征的科研数据图。请先分析数据结构、变量类型和需要表达的信息，再自行选择最合适的【柱状图/折线图/箱线图/直方图/散点图/均值±误差图等】进行可视化，不要在未理解数据的情况下直接作图。具体要求如下：

1. 数据解析：识别【核心变量】、【分组变量】、【时间变量】及其数据类型，并检查是否存在缺失值、异常值或明显格式问题；
2. 图表选择：根据【表达目的，如分布特征/变化趋势/集中趋势/离散程度】选择最合适的图表形式，并简要说明选择理由；
3. 数据表达：重点突出【需要呈现的核心特征】，不得改变、补充或虚构原始数据；
4. 坐标与标注：清晰标注横轴、纵轴、单位、图例及必要的统计信息；
5. 统计规范：如呈现均值、标准差、标准误、置信区间等信息，应明确标注其具体含义；
6. 视觉风格：采用【简洁学术风格/期刊论文风格】，减少无关装饰，保证打印和论文阅读时清晰可辨；
7. 输出要求：生成【图片数量】张科研数据图，适用于【期刊论文/研究报告/学术汇报】使用。

先诊断数据，再决定怎么画：豆包可自主处理缺失与异常情况，并根据分布、趋势和个体差异选择互补科研图表。

4 张科研数据图已全部交付。以下是每张图的核心数据特征总结：

表格

图号	图表类型	主要表达的数据特征
图 1	箱线图	中位数随年级递增（5.0→6.0→7.0→9.0 次 / 周）；IQR 稳定在 3.0-4.0；大一、大四各 1 例高值异常点
图 2	小提琴图	大一大二分布集中（峰值 4-6 次 / 周）；大三大四分布变宽，大四右偏长尾延伸至 22 次 / 周；密度峰值整体右移
图 3	均值 ±SD 柱状图	频次 4.97→8.65 次 / 周、时长 23.5→38.4 分钟，双指标均单调递增；大四约为大一的 1.6-1.7 倍
图 4	抖动散点图	个体分布范围大一 1-12、大四 2-22 次 / 周；低值个体（0-2 次 / 周）随年级减少，高值个体增多

统计规范：所有图均标注样本量（n=59-60 / 组），图 3 柱顶标注均值 ± 标准差具体数值，图 4 标注各组中位数。单因素 ANOVA 显示年级间差异极显著（F=19.32, p<0.001）。所有图表 300dpi，无 3D / 渐变 / 阴影，严格基于原始数据未做平滑或修饰。

交付图片并说明理由 >

数据解析与图表选择说明

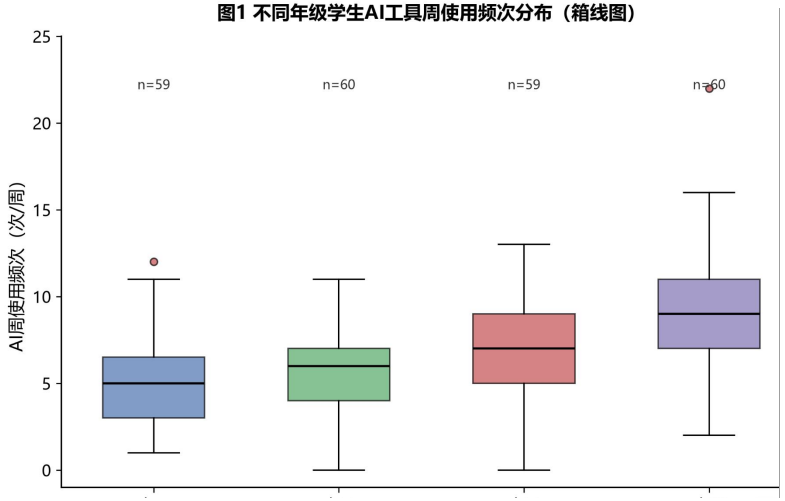
数据结构：240 名大学生，4 个年级各 60 人（均衡设计）。核心变量为 AI 周使用频次（次 / 周）和单次使用时长（分钟）

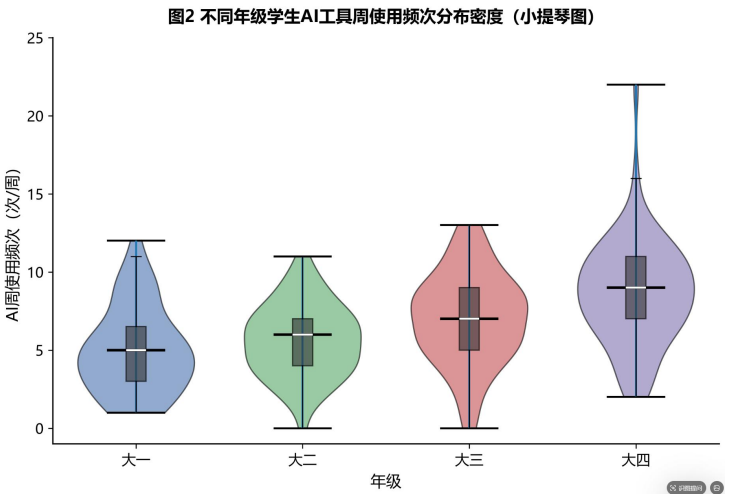
数据质量：

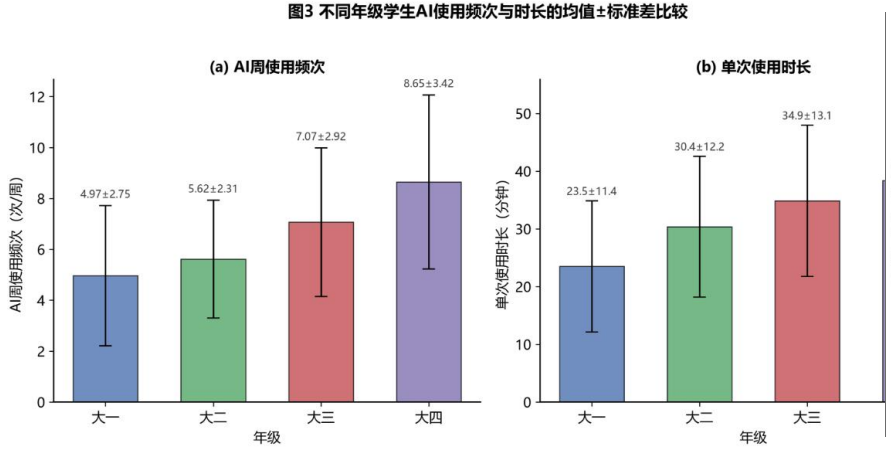
- 缺失值：AI 周使用频次 2 例、单次使用时长 3 例（已在绘图时剔除，不做插补）
- 异常值（IQR 法）：大一 1 例高值（12 次 / 周）、大四 1 例高值（22 次 / 周），均保留并在图中标识
- 年级趋势：AI 周使用频次随年级显著递增（大一 4.97→大四 8.65 次 / 周，单因素 ANOVA F=19.32, p=3.15×10⁻¹¹）

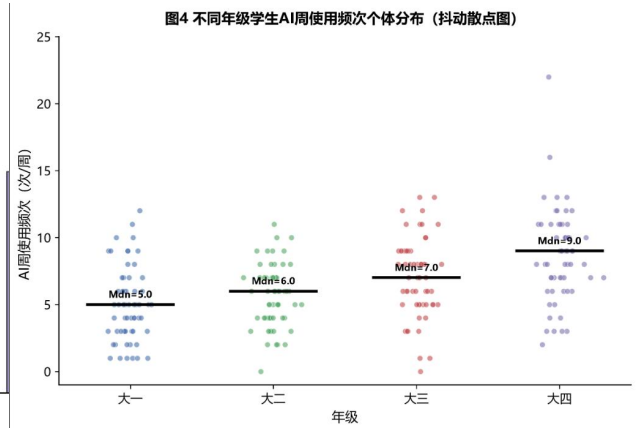
图表选择理由：

1. **箱线图**—— 比较多组分布的标准形式，同时呈现中位数、四分位距、极值和异常值
2. **小提琴图**—— 补充箱线图，展示分布密度形状与偏态特征
3. **均值 ± 标准差柱状图**—— 直观呈现集中趋势与离散程度的组间比较，双指标并列
4. **抖动散点图**—— 展示个体层面数据分布，避免汇总统计掩盖个体差异









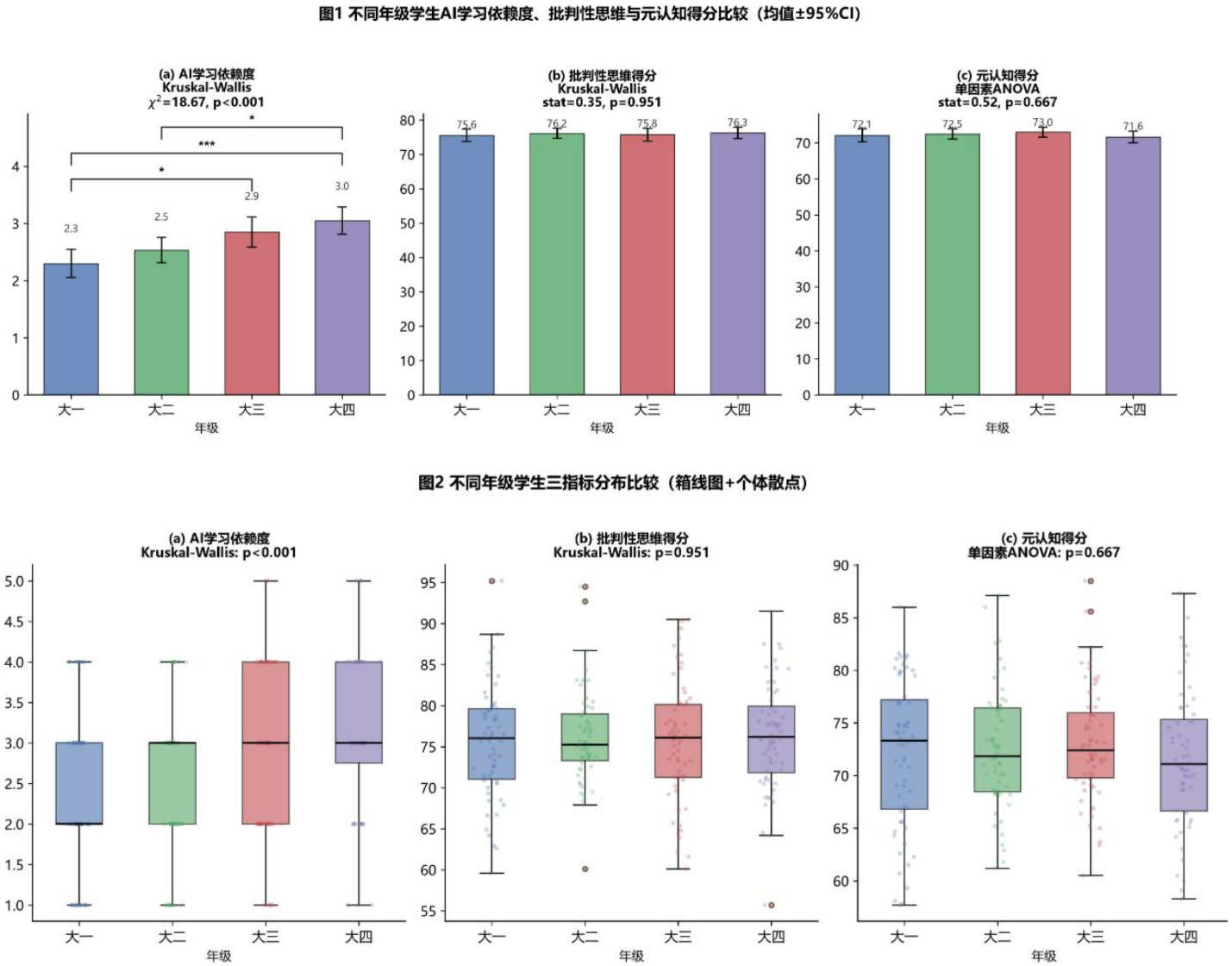
科研数据图--比较型

测试模式：工作任务 Turbo

提示词公式

请基于我上传的【数据表/统计结果】，围绕【研究问题】，生成用于比较【不同组别/不同实验条件/不同时间点】差异的科研数据图。
请先理解变量结构和比较目的，再根据数据特点自行选择最适合的【柱状图/箱线图/小提琴图/折线图/点图等】进行可视化。
具体要求如下：

1. 变量识别：明确【分组变量】、【结果变量】以及需要控制或辅助说明的其他变量；
2. 比较设计：根据【比较目的，如组间差异/干预前后变化/不同条件比较】选择合适的图表形式；
3. 统计分析：根据数据类型自行判断是否需要进行【t检验/方差分析/非参数检验/重复测量分析等】，并说明选择依据；
4. 差异表达：在图中准确呈现各组的【均值/中位数/分布】以及【标准差/标准误/置信区间】等信息；
5. 显著性标注：如存在统计显著差异，可使用规范方式标注【P值/显著性符号/置信区间】，不得虚构统计结果；
6. 数据真实性：所有数值必须严格来自原始数据或实际统计结果，不得自行修改、补全或制造显著性；
7. 视觉规范：采用【简洁期刊风格/学术论文风格】，避免3D、渐变和过度装饰；
8. 输出要求：生成【图片数量】张比较型科研数据图，并简要说明每张图所表达的主要研究发现。



科研数据图--关系型

测试模式：工作任务 Turbo

提示词公式

请基于我上传的【数据表】，围绕【研究问题】，分析【变量1】、【变量2】、【变量3】等核心变量之间的关系，并生成适合科研论文使用的关系型数据图。请先理解数据结构和变量含义，再根据数据类型和研究目的自主选择最合适的【散点图/相关矩阵/热力图/回归拟合图/气泡图等】进行可视化。具体要求如下：

1. 变量识别：明确需要分析的【核心变量】及其数据类型，判断哪些变量适合开展相关关系或趋势分析；
2. 数据检查：检查缺失值、异常值和变量分布，必要时说明相应处理方式；
3. 关系分析：根据变量类型和数据条件，自行选择合适的【Pearson相关/Spearman相关/回归分析等】方法，并说明选择依据；
4. 图表选择：根据不同变量关系选择最适合的图形，避免为了丰富形式而重复表达相同信息；
5. 统计标注：如涉及相关系数、回归系数、P值、置信区间或拟合结果，应准确计算并规范呈现；
6. 关系解释：区分正相关、负相关、弱相关和无明显关系，不得将相关关系直接表述为因果关系；
7. 视觉规范：采用【简洁学术风格/期刊论文风格】，坐标、单位、图例和统计标注完整；
8. 输出要求：生成【图片数量】张最有分析价值的关系型科研数据图，并说明每张图呈现的主要关系。

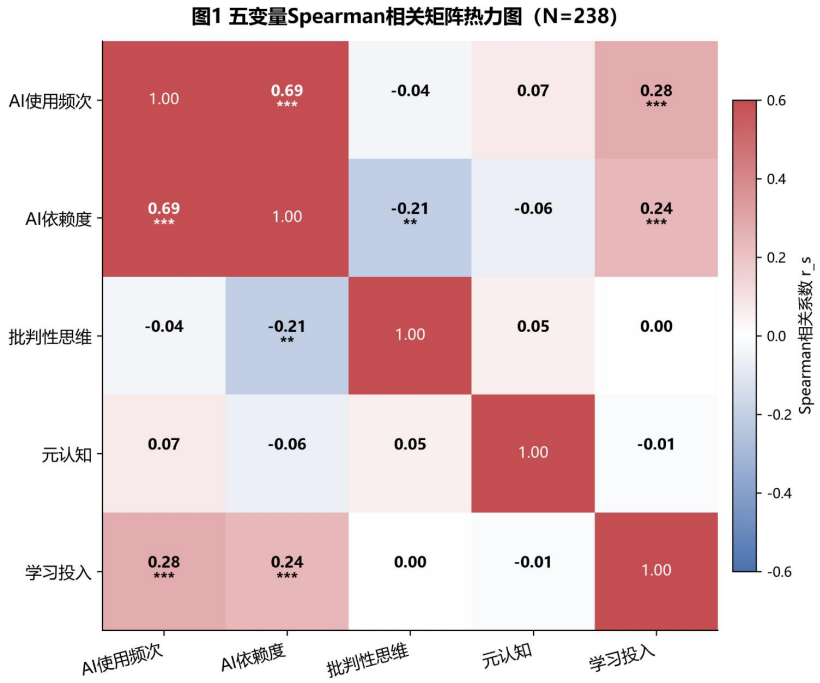
总结交付图的主要发现 >

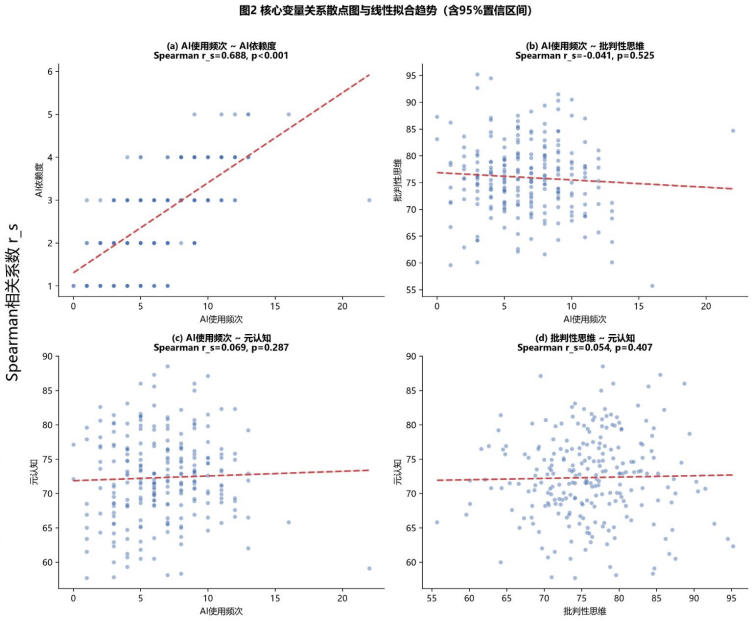
两张关系型数据图已全部交付。主要发现如下：

图 1（相关矩阵热力图）：五变量间最强的关联是 AI 使用频次与 AI 依赖度（r_s=0.688, p<0.001，强正相关）；AI 依赖度与批判性思维呈弱负相关（r_s=-0.205, p=0.001）；AI 使用频次和依赖度均与学习投入呈弱正相关（r_s=0.276 和 0.243, p<0.001）；而 AI 使用频次与批判性思维（r_s=-0.041）、元认知（r_s=0.069）以及批判性思维与元认知之间（r_s=0.054）均无统计学显著关联。

图 2（散点图矩阵）：(a) AI 使用频次与依赖度呈现清晰的正线性趋势，拟合线斜率明显；(b)(c) AI 使用频次与批判性思维、元认知的散点分布呈无规律弥散状，拟合线近乎水平，直观印证无相关；(d) 批判性思维与元认知的散点同样呈圆形弥散分布，二者不存在线性关联。

以上均为横截面数据的相关分析结果，仅反映变量间的关联程度，不构成因果推断。所有图表 300dpi，无 3D / 渐变 / 阴影，符合期刊论文规范。





课题申报——从研究主题到完整申报书

测试模式：工作任务Pro

提示词公式

请围绕【研究主题】，按照【课题类型/申报要求】，生成一份完整的课题申报书初稿。
不要等待我逐项提供大纲和材料，请先自主理解研究主题，制定完整任务计划，再连续完成文献检索、研究设计、科研绘图、申报书撰写和文档排版。具体要求如下：

1. 文献检索：围绕【研究主题】自主检索真实的中英文文献，中文重点覆盖【知网、维普等国内学术来源】，英文重点检索【国际期刊/数据库】，兼顾近年研究与代表性经典文献；
2. 文献核验：核对作者、题名、期刊、年份、DOI等书目信息，无法确认的内容不得自行补全；
3. 研究综述：梳理国内外研究的主要方向、核心观点、研究进展、争议与不足，不得简单堆砌文献；
4. 问题凝练：在文献综述基础上明确本课题拟解决的【核心问题】，说明现有研究为什么还不足以回答这些问题；
5. 框架设计：自主形成【研究目标、研究内容、重点难点、研究思路、研究方法和整体逻辑】；
6. 方法设计：根据不同研究内容选择适合的方法，并明确“哪一种方法解决哪一个问题”；
7. 创新提炼：从【理论视角、研究问题、研究内容、方法或应用】等方面提炼创新点，不得使用缺乏证据的“首次”“填补空白”等表述；
8. 科研绘图：根据研究内容自动生成至少【研究总体框架图】和【技术路线图】，必要时增加【机制图/研究模型图】；
9. 图文融合：将科研图片自动插入申报书对应章节，并在正文中加入“如图X所示”等引用，同时生成规范图题；
10. 完整撰写：按照【申报书结构】生成完整正文，包括但不限于【选题依据、研究现状、研究目标、研究内容、重点难点、研究思路与方法、技术路线、创新之处、研究计划、预期成果、参考文献】；
11. 证据边界：不得虚构申请人已有成果、前期项目、实验数据、论文发表情况等未提供的信息，缺失内容标注【待申请人补充】；
12. 一致性检查：检查【研究问题—研究目标—研究内容—研究方法—创新点—预期成果】之间是否一一对应，避免前后概念不一致或内容重复；
13. 最终交付：生成一份【Word格式】完整课题申报书初稿，完成标题层级、段落、表格、科研图片、图题和参考文献排版，确保图片直接嵌入正文而不是单独输出。


生成式人工智能赋能高校教育治理...
741KB

确认文件交付状态 >

课题申报书初稿已生成并交付，以下是完成情况说明。

文档概况

- 全文约 1.5 万字，A4 排版，宋体 / 黑体分级标题，1.5 倍行距
- 含封面、基本信息表、11 个正文部分、3 个数据表格、28 条参考文献

正文结构

1. 选题依据（研究背景、问题提出、理论与实践意义）
2. 国内外研究现状述评（按四个研究方向系统梳理，非逐篇罗列）
3. 研究目标（5 项具体目标）
4. 研究内容（5 个模块，含机制图）
5. 重点难点（3 个重点、4 个难点）
6. 研究思路与方法（5 种方法，各对应明确问题）
7. 技术路线（五阶段，含路线图）
8. 创新之处（4 点，未使用“首次”“填补空白”等表述）
9. 研究计划（表格）
10. 预期成果（5 项，表格）

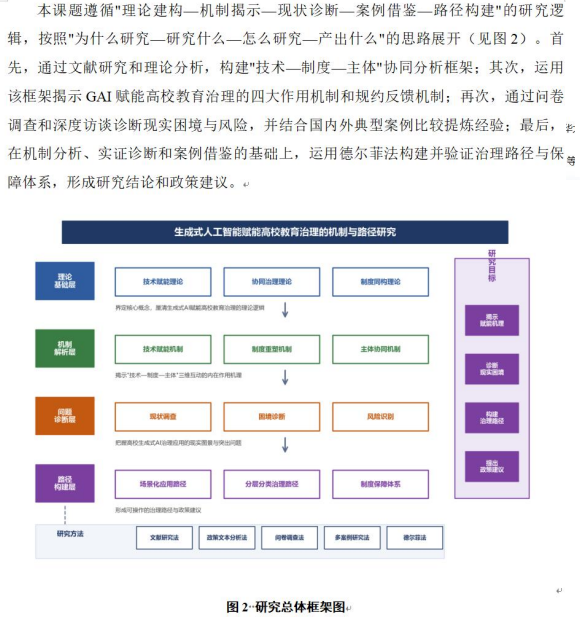

本课题遵循"理论建构—机制揭示—现状诊断—案例借鉴—路径构建"的研究逻辑，按照"为什么研究—研究什么—怎么研究—产出什么"的思路展开（见图 2）。首先，通过文献研究和理论分析，构建"技术—制度—主体"协同分析框架；其次，运用该框架揭示 GAI 赋能高校教育治理的四大作用机制和规约反馈机制；再次，通过问卷调查和深度访谈诊断现实困境与风险，并结合国内外典型案例比较提炼经验；最后，在机制分析、实证诊断和案例借鉴的基础上，运用德尔菲法构建并验证治理路径与保障体系，形成研究结论和政策建议。

图 2 研究总体框架图

测试结果：从一个研究主题直接生成21页图文一体课题申报书

课题申报书

（初稿）

课题名称：生成式人工智能赋能高校教育治理的机制与路径研究

申请人：待申请人补充

所在单位：待申请人补充

申请日期：2026 年 8 月

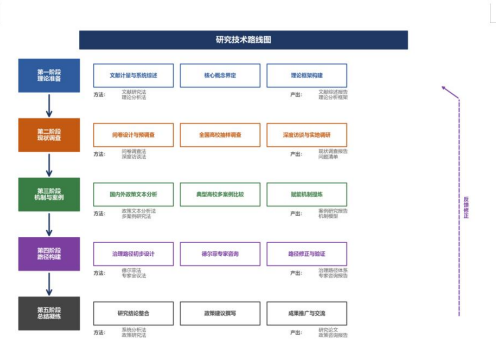


图 3 研究技术路线图

第一阶段（第 1—3 个月）：理论准备。完成文献检索与系统综述，界定核心概念，构建理论分析框架，形成文献综述报告和理论分析框架。

第二阶段（第 4—9 个月）：现状调查。设计问卷和访谈提纲，开展预调查并修订工具，实施全国高校抽样问卷调查和深度访谈，完成现状调查报告和问题清单。

第三阶段（第 10—15 个月）：机制分析与案例研究。开展国内外政策文本分析和典型高校多案例比较，结合调查数据提炼 GAI 赋能高校教育治理的作用机制，形成案例研究报告和机制模型。

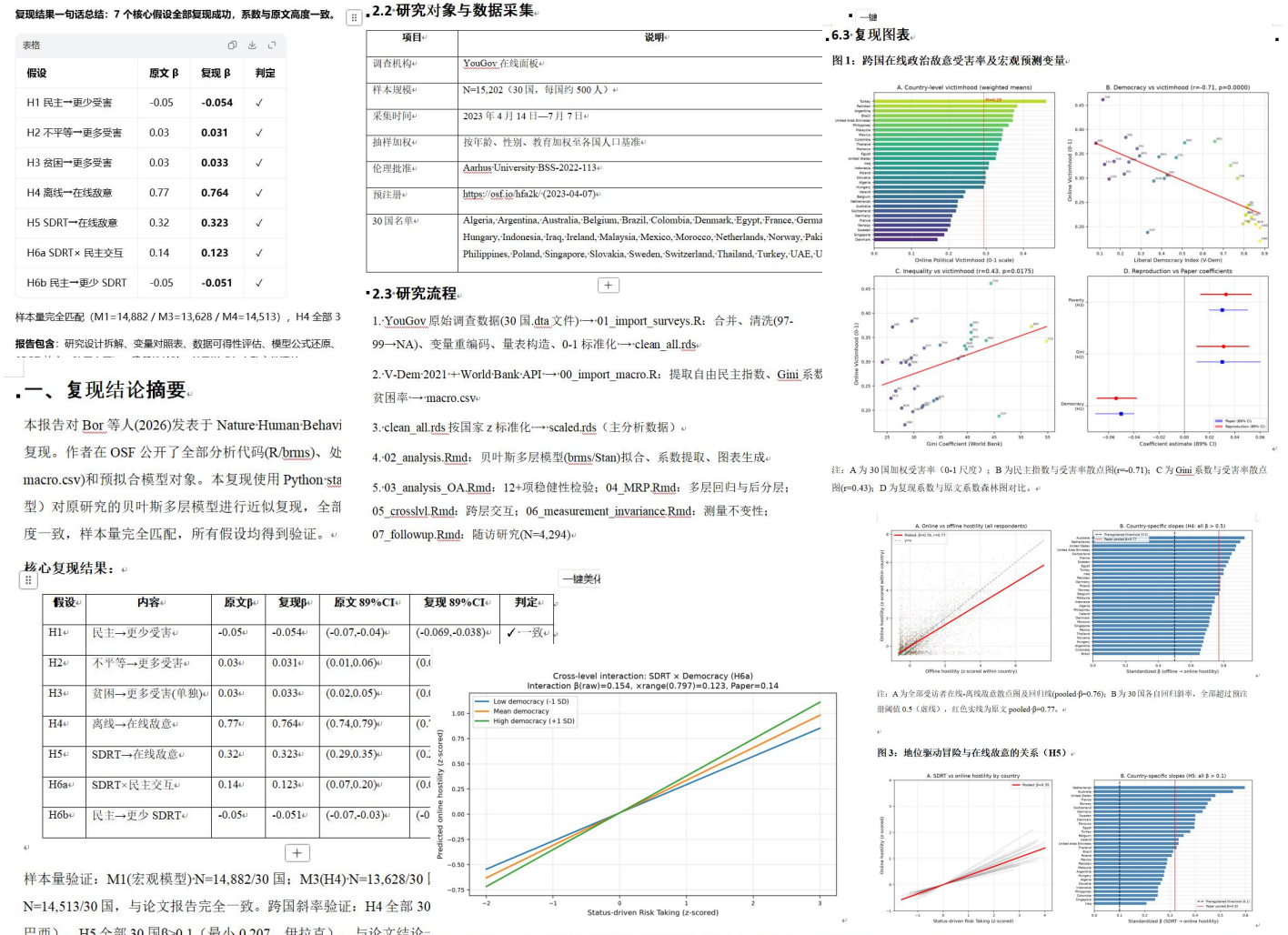
论文复现——顶刊研究全流程复刻

测试模式：工作任务Pro

提示词公式

请基于我上传的【目标论文】，尝试完整复现其研究过程。
不要只做论文总结，请将论文反向拆解为可执行的科研流程，并根据论文公开的信息、数据、代码和补充材料自主推进复现。具体要求如下：

1. 研究问题拆解：提取论文的研究问题、理论假设、核心变量、研究对象和主要创新；
2. 研究设计还原：识别论文采用的研究设计、样本来源、样本量、数据获取方式和研究流程；
3. 变量识别：整理因变量、自变量、控制变量、中介/调节变量及其测量或计算方式；
4. 数据定位：查找论文正文、附录、补充材料、数据仓库或作者公开页面中是否提供原始数据或可复现数据；
5. 代码定位：检查是否存在作者公开的分析代码、项目仓库或补充脚本，并优先使用原始代码进行复现；
6. 方法复刻：还原数据清洗、变量构造、统计分析、模型设定、参数配置和稳健性检验等关键步骤；
7. 代码重建：若作者未公开完整代码，根据论文方法部分重建可执行分析代码，并明确哪些步骤属于依据论文信息推断；
8. 结果复现：尽可能复现论文中的主要统计结果、模型结果和关键结论；
9. 图表复现：优先复现论文中的核心表格和科研图，尽量保持变量、指标和呈现逻辑一致；
10. 结果比对：将复现结果与原论文逐项比较，包括样本量、描述统计、系数、显著性、模型拟合、图表趋势等；
11. 差异诊断：若无法完全复现，分析差异可能来自【数据版本、预处理方法、软件版本、参数设定、随机种子或论文信息缺失】中的哪一部分；
12. 证据边界：不得虚构原论文未公开的数据、代码、参数或实验过程；无法确认的内容必须明确标注；
13. 最终交付：输出【论文复现流程说明】+【变量与方法对照表】+【可执行分析代码】+【复现结果表】+【复现图表】+【原论文与复现结果对比】+【复现报告】。



复现结果一句话总结：7 个核心假设全部复现成功，系数与原文高度一致。

表格

假设	原文 β	复现 β	判定
H1 民主→更少受害	-0.05	**-0.054**	✓
H2 不平等→更多受害	0.03	**0.031**	✓
H3 贫困→更多受害	0.03	**0.033**	✓
H4 离线→在线敌意	0.77	**0.764**	✓
H5 SDRT→在线敌意	0.32	**0.323**	✓
H6a SDRT× 民主交互	0.14	**0.123**	✓
H6b 民主→更少 SDRT	-0.05	**-0.051**	✓

样本量完全匹配（M1=14,882 / M3=13,628 / M4=14,513），H4 全部 3

报告包含：研究设计拆解、变量对照表、数据可得性评估、模型公式还原、

一、复现结论摘要

本报告对 Bor 等人(2026)发表于 Nature Human Behavi
复现。作者在 OSF 公开了全部分析代码(R/brms)、处
macro.csv)和预拟合模型对象。本复现使用 Python sta
型）对原研究的贝叶斯多层模型进行近似复现，全部
度一致，样本量完全匹配，所有假设均得到验证。

核心复现结果：

假设	内容	原文β	复现β	原文 89%CI	复现 89%CI	判定
H1	民主→更少受害	-0.05	-0.054	(-0.07,-0.04)	(-0.069,-0.038)	✓ 一致
H2	不平等→更多受害	0.03	0.031	(0.01,0.06)	(0.(
H3	贫困→更多受害(单独)	0.03	0.033	(0.02,0.05)	(0.(
H4	离线→在线敌意	0.77	0.764	(0.74,0.79)	(0.7
H5	SDRT→在线敌意	0.32	0.323	(0.29,0.35)	(0.2
H6a	SDRT×民主交互	0.14	0.123	(0.07,0.20)	(0.(
H6b	民主→更少 SDRT	-0.05	-0.051	(-0.07,-0.03)	(-0

样本量验证：M1(宏观模型)N=14,882/30 国；M3(H4)N=13,628/30 国
N=14,513/30 国，与论文报告完全一致。跨国斜率验证：H4 全部 30
巴西），H5 全部 30 国β>0.1（最小 0.207，伊拉克），与论文结论一

二、研究设计拆解

2.2 研究对象与数据采集

项目	说明
调查机构	YouGov 在线面板
样本规模	N=15,202（30 国，每国约 500 人）
采集时间	2023 年 4 月 14 日—7 月 7 日
抽样加权	按年龄、性别、教育加权至各国人口基准
伦理批准	Aarhus University BSS-2022-113
预注册	https://osf.io/hfa2k/ (2023-04-07)
30 国名单	Algeria, Argentina, Australia, Belgium, Brazil, Colombia, Denmark, Egypt, France, Germa Hungary, Indonesia, Iraq, Ireland, Malaysia, Mexico, Morocco, Netherlands, Norway, Paki Philippines, Poland, Singapore, Slovakia, Sweden, Switzerland, Thailand, Turkey, UAE, U

2.3 研究流程

1. YouGov 原始调查数据(30 国.dta 文件)→01_import_surveys.R：合并、清洗(97-99→NA)、变量重编码、量表构造、0-1 标准化→clean_all.rds
2. V-Dem 2021+World Bank API→00_import_macro.R：提取自由民主指数、Gini 系数
贫困率→macro.csv
3. clean_all.rds 按国家 z 标准化→scaled.rds（主分析数据）
4. 02_analysis.Rmd：贝叶斯多层模型(brms/Stan)拟合、系数提取、图表生成
5. 03_analysis_OA.Rmd：12+项稳健性检验；04_MRP.Rmd：多层回归与后分层；05_crosslvl.Rmd：跨层交互；06_measurement_invariance.Rmd：测量不变性；07_followup.Rmd：随访研究(N=4,294)

一键美化

Cross-level interaction: SDRT × Democracy (H6a)
Interaction β(raw)=0.154, ×range(0.797)=0.123, Paper=0.14
Low democracy (-1 SD)
Mean democracy
High democracy (+1 SD)
Predicted online hostility (z-scored)
Status-driven Risk Taking (z-scored)

注：在高民主国家，SDRT 对在线敌意的斜率更陡，即地位驱动者在更民主社会中更具敌意，交互效应×民主范围=0.123（原文 0.14）。

一键

6.3 复现图表

图 1：跨国在线政治敌意受害率及宏观预测变量

A. Country-level victimhood (weighted means)
B. Democracy vs victimhood (r=-0.71, p=0.0000)
C. Inequality vs victimhood (r=0.43, p=0.0175)
D. Reproduction vs Paper coefficients
Online Political Victimhood (0-1 scale)
Liberal Democracy Index (V-Dem)
Gini Coefficient (World Bank)
Coefficient estimate (89% CI)

注：A 为 30 国加权受害率（0-1 尺度）；B 为民主指数与受害率散点图(r=-0.71)；C 为 Gini 系数与受害率散点图(r=0.43)；D 为复现系数与原文系数森林图对比。

A. Online vs offline hostility (all respondents)
B. Country-specific slopes (H4: all β > 0.5)
Offline hostility (z-scored within country)
Standardized β (offline → online hostility)

注：A 为全部受访者在线-离线敌意散点图及回归线(pooled-β=0.76)；B 为 30 国各自回归斜率，全部超过预注册阈值 0.5（虚线），红色实线为原文 pooled-β=0.77。

图 3：地位驱动冒险与在线敌意的关系（H5）

A. SDRT vs online hostility by country
B. Country-specific slopes (H5: all β > 0.1)
Online hostility (z-scored)
Status-driven Risk Taking (z-scored)
Standardized β (SDRT → online hostility)

注：A 为各国 SDRT-敌意回归线（灰色）及 pooled 回归线（红色，β=0.32）；B 为 30 国各自斜率，全部超过预注册阈值 0.1（虚线），红色实线为原文 pooled-β=0.32。

上传一篇顶刊论文，豆包 Pro 可自主查找公开数据与代码、还原研究设计、重建分析流程并复现核心结果和图表，最终形成完整复现报告。

公开数据研究——从数据库到实证论文

测试模式：工作任务Pro

提示词公式

请围绕【研究问题】，基于适合的公开数据库完成一项完整的实证研究。
不要等待我提供数据文件或变量名称，请先根据研究问题自主判断适合使用的公开数据库，定位并获取数据，再完成变量识别、数据处理、统计分析、科研绘图和论文级结果输出。具体要求如下：

1. 数据库选择：根据【研究主题/研究对象】自主比较多个可能的数据源，选择最适合的【国内/国际公开数据库】，并说明选择依据；
2. 数据获取：定位数据库官方网站、公开下载页面或可信数据仓库，获取可用于研究的数据文件、数据字典和相关说明；
3. 变量匹配：围绕【因变量】【核心自变量】【控制变量】从数据字典中寻找最合适的变量，明确变量名称、编码、测量方式和所在数据文件；
4. 样本构建：根据研究问题确定分析样本范围、纳入排除标准，并说明最终样本是如何形成的；
5. 数据处理：完成缺失值、异常值、变量编码、反向题、合成指标、标准化或权重处理等必要步骤，不得随意改变原始变量含义；
6. 研究设计：根据数据结构和研究问题自主选择适合的分析方法，不要提前固定统计模型；
7. 模型分析：完成核心模型，并根据研究需要加入控制变量、分层模型、固定效应、多层模型或其他适合方法；
8. 复杂抽样处理：如果数据库包含抽样权重、分层、聚类设计或重复权重，应根据数据库规范进行处理，而不是把复杂调查数据当作普通随机样本；
9. 稳健性检验：根据研究设计自主选择替代变量、替代模型、样本调整、异常值处理或其他稳健性检验；
10. 异质性分析：如研究问题适合，可进一步分析不同【性别/年级/地区/学校类型/社会经济背景】群体之间的差异；
11. 科研绘图：根据核心结果自动生成适合论文使用的数据图，图中的数值必须来自实际分析结果；
12. 结果解释：区分统计关联、预测关系和因果效应，不得因为使用回归模型就直接作因果推断；
13. 可复现交付：保留数据处理和统计分析代码，并清楚说明使用的数据版本、变量和分析流程；
14. 最终输出：形成【数据来源说明】+【变量定义表】+【数据处理流程】+【分析代码】+【模型结果表】+【科研图】+【稳健性结果】+【论文级实证分析初稿】。



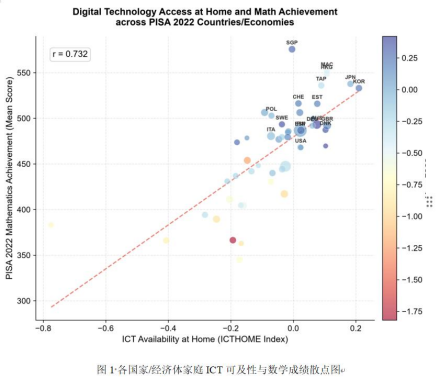


图 1·各国家/经济体家庭 ICT 可及性与数学成绩散点图。



图 2·分层模型系数图（含 95%置信区间）

图 3·稳健性检验与异质性分析

没有提供任何数据文件，豆包 Pro 自己找到PISA原始数据库、读取数据字典、提取所需变量、正确处理10个PV和80个BRR权重，再完成模型、稳健性、异质性、科研图和Word论文初稿。

实测亮点总结

从“单点AI工具”走向“科研任务执行器”

核心判断　豆包的优势不在于某个单点功能，而在于能把文献、数据、绘图和写作串成一条连续科研工作流。

实测关键词

- 国内外检索
- 综述闭环
- 局部重绘
- 自主分析
- 据源写作

实测边界：参考文献书目信息仍需复核；不确定数据避免模型自行推测修正；统计模型解释需人工把关。

01　国内外文献一站式检索

既能找国外期刊，也能覆盖知网、维普等中文学术来源；更适合中文科研场景。

02　从检索到综述连续完成

主题、段落、种子文献三种入口，继续完成筛选、分类、共识、分歧、空白与综述成稿。

03　科研绘图支持局部重绘

图形摘要、机制图、流程图生成后，可指定局部修改：哪里不满意改哪里，不必整图重做。

04　从“指定方法”到“自主分析”

清洗、转换、整合、差异、相关、回归、聚类、预测等环节，可根据数据条件自主选择方法。

05　科研任务连续执行

围绕同一主题串联检索、数据、绘图、论文输出与审稿修改，形成任务链而非单点问答。

扫码（小助理）入科研交流群



让科研像聊天一样简单

愿光阴所及，皆是暖阳；
愿人心所向，俱为温良。



AI科研小助手



扫一扫上面的二维码图案，加我为朋友。



让科研像聊天一样简单



扫一扫上面的二维码图案，加我为朋友。