

生成式人工智能支持下的多模态作品创作课实践探索 ——以“跨界艺术家”一课为例

上海民办南模中学 史今金

摘要：

在人工智能融入教育体系的背景下，课堂教学正经历由“工具辅助”向“智能共创”的变革。本文以上海市人工智能课程“跨界艺术家”一课为实践案例，以初二学生为教学对象，依托扣子平台开展智能体绘本创作实践，从创作任务设计、人机协同创作及作品迭代等方面，分析人工智能赋能课堂教学的实施路径。实践发现，将智能体引入绘本创作过程，能够为学生提供开放的创作空间和即时的智能反馈，促进学生运用文本、图像、声音等多模态技术进行表达，并在问题发现、方案调整与作品优化过程中，提升跨界思维、问题解决与创新表达能力。实践探索发现，基于智能体的项目化创作能够推动人工智能由课堂中的“辅助工具”转变为学生学习与创作的“协作伙伴”，为构建以学生为主体、人工智能深度参与的人机协同课堂提供实践参考。

关键词： 人工智能教育；项目式学习；智能体；人机协同；

一、背景分析

在 AI 快速发展的背景下，人工智能、虚拟现实等新兴技术不断应用于教育教学，推动教育模式由传统“教师知识传授”逐步向“学生能力建构”转型，学习者数字素养的内涵也由单一的技术素养逐渐拓展至综合跨学科素养。^[1] 生成式人工智能的发展在文本、图像、语音等方面展现出的多模态生成能力，为课堂教学创新和学生综合表达提供了新的实践空间。然而，从当前人工智能融入课堂的实践来看，仍存在一些亟待解决的问题。首先，学习活动呈现一定的碎片化特征，生成式人工智能往往以单一任务或局部环节介入课堂，缺乏系统性的任务设计和持续性的学习过程，难以促进学生形成结构化认知与深度学习。其次，人工智能应用存在一定的表层工具化倾向，部分课堂主要将 AI 用于资源获取、信息检索或内容生成，学生更多关注“使用工具”，而缺乏对生成机制、技术逻辑及其应用边界的深入理解。学生容易直接依赖 AI 生成结果，弱化自主思考、批判性判断以及主动迭代的过程。此外，人工智能与其他学科的

融合程度仍显不足，多模态生成技术在语言、艺术、信息科技等领域的综合应用尚未得到充分发挥，人工智能赋能学生跨学科创作与综合表达的教育价值仍有较大的探索空间。因此，如何从单纯的“AI 工具使用”走向“AI 理解与协同创作”，如何通过系统化、项目化的学习活动促进学生深度参与、持续迭代与跨学科表达，成为当前人工智能课程实践中值得进一步研究的问题。

基于上述背景，本文以“跨界艺术家”一课为实践案例，开展智能体绘本创作项目，探索生成式人工智能支持下的人机协同学习路径及其对学生技术认知、能力提升、情感表达的促进作用，为人工智能与基础教育深度融合提供实践参考。

二、设计思路：从生成到调试的五个阶段

基于相关研究，本课将绘本创作任务分解为“创意构思—提示词生成—多模态生成—反馈调试—成果展示”五个阶段，形成完整的生成式学习路径。该设计思路融合生成式学习与深度学习机制，强调学生在“构思—生成—反馈—优化—展示”的循环过程中实现知识建构与能力提升。

（一）理论框架支撑

本课程构建了“人机协同多模态生成学习框架”。该框架为课程设计与实践教学提供理论支撑，其核心包括以下五点：一是项目式学习结构，以绘本创作为载体实现综合能力培养；二是建构主义学习观，以真实项目情境为基础促进知识建构；三是人机协同模式，基于 AI 平台构建动态互动的学习环境；四是多模态学习路径，整合语言、视觉与声音表达；五是深度学习机制，通过迭代优化项目实践结果促进学生对技术的深入理解。

（二）课程导入

教师首先回顾上节课的学习内容：学生通过输入关键词生成故事，对故事进行分镜，并借助提示词优化插件将故事内容转化为可视化的文字描述。在此基础上，教师提出问题：“如果要將故事进一步制作成完整绘本，还需要哪些要素？”引导学生思考如何为故事添加图像与声音，自然导入本课学习。本环节通过“回顾旧知—提出问题—引入新知”建立知识衔接，激活学生已有经验，为后续多模态绘本创作做好铺垫。

（三）介绍插件

教师以“小工具”的类比介绍智能体插件，能使智能体具备更多功能（API），一个插件内可包含一个或多个工具。通过问题引导学生思考不同插件能够为智能体增加哪些能力，并进一步明确插件在绘本创作中的作用。该环节旨在通过插件介绍，让学生了解其在绘本创作过程中的重要作用。

（四）绘本图像生成

在绘本图像生成环节，教师首先引发学生思考能否根据故事内容让 AI 协助绘画。接着开展活动探索，介绍图像生成插件，以“日记里的牛奶店”为例，带领学生探索图像生成插件的使用过程，完成“添加插件—调用工具—生成预览”的基本操作。在此基础上，教师组织学生对比不同图像生成插件的生成效果，从画面风格、色彩表现和内容契合度等维度进行分析，讨论不同工具与创作需求之间的匹配关系。本环节通过插件选择与生成效果对比，帮助学生理解文生图的基本逻辑，形成“根据创作目标选择合适工具”的意识，深入了解文生图插件的作用。

（五）绘本语音生成

针对绘本语音生成，教师先引发学生思考能否根据故事内容让 AI 协助配音。随后开展活动探索，介绍语音合成插件，演示从无到有（添加插件-调用插件-预览）生成语音的过程，引导学生了解如何根据绘本风格主题进行插件参数调整。最后组织学生参与语音参数调节的试听活动，体验 AI 语音对故事表达的作用。此环节通过模拟活动让学生直观体验语音合成插件的使用及学会根据故事内容进行参数调节的方法。

（六）调试智能体

在调试智能体环节，教师引导学生发现问题，思考图片上的人物形象、画面风格是否统一，图像与文字内容的排版是否合理，接着探讨问题解决的方法。针对生成结果与预期存在偏差的情况，教师引导学生采用优化提示词、修改生成要求等方式进行迭代。对于输出格式不稳定的问题，则通过设置固定输出格式提高结果的一致性。学生根据反馈持续调整智能体，比较优化前后的生成效果。本环节将创作过程由“生成”进一步延伸至“分析—调整—再生成”，突出调试在创作中的重要作用，培养学生的问题发现、分析解决和迭代优化能力。

三、典型活动剖析

（一）活动一：“日记里的牛奶店”

活动目的：

本活动以“牛奶店的奇遇”故事为情境，旨在探究图像生成插件的使用方法和功能，培养学生 AI 工具的选择能力和视觉审美一致性的把控能力。活动要求学生围绕故事段落文字生成对应图片，并在相同提示词（Prompt）的基础上，分别调用不同底层模型（如豆包图像生成大模型、通义万相、Midjourney、DALL-E 3、Stable Diffusion 等）的插件进行生成测试，这一过程不仅涉及技术的操作，更是一次关于风格迁移与审美匹配的深度学习挑战。

活动实施：

学生首先完成了添加与调用图像生成插件的基础操作，随后进入核心对比探究环节。学生需对不同的文生图插件进行多维度评估，具体考察指标包括来源平台、风格特点（如卡通风、油画风等）、色彩表现力以及与故事内容的契合程度。例如，在对比测试中，学生发现“豆包图像生成大模型”倾向于生成色彩鲜艳、符合儿童绘本特征的卡通风格，而“通义万相”则可能在油画质感和写实风格上表现更为突出。通过预览并记录不同模型的生成结果，学生能够直观地观察到同一提示词在不同算法模型下的差异化呈现。

活动成效：

学生不再被动接受 AI 生成的结果，而是通过比较不同 AI 图像生成模型插件，分析不同插件在风格特点、色彩表现等方面的优劣，形成“工具匹配目标”的思维。

（二）活动二：语音参数调节实验

活动目的：

本活动旨在引导学生深入探究语音合成（TTS）插件的功能特性与使用方法，重点探索参数设计对听觉体验的具体影响。通过实践操作，使学生理解技术参数不仅是控制输出的数值，更是传递情感温度、增强叙事感染力的关键手段，从而培养其在多模态创作中对声音元素的审美把控能力。

活动实施：

活动依托已生成的绘本故事文本展开，学生首先完成语音合成插件的添加

与调用，并通过语音链接、音频文件或二维码等形式进行输出。在此基础上，活动设置了核心的参数调节与对比实验环节：

（1）音色情感匹配实验：学生对比测试“奶气萌娃”与“京腔侃爷”等差异化音色，分析不同声音特质所承载的叙事情感。例如，探究“奶气萌娃”音色契合儿童角色的对话场景，而“京腔侃爷”音色更适配旁白叙述或成人角色的演绎，从而实现音色与故事角色的精准匹配。

（2）语速节奏调控实验：学生将语速参数从 0.5 倍速的迟缓状态调整至 2.0 倍速的急促状态，观察语速变化对叙事节奏的影响。通过反复试听，寻找最符合故事内在节奏的语速区间（如儿童故事通常适配 0.8-1.0 倍速的稍慢语速），以营造适宜的聆听氛围。

（3）情境嵌入与验证：学生将调节后的音频嵌入故事情节中，进行整体听感测试。通过对比不同参数组合下的听觉效果，评估哪组参数最能契合文本内容并引发情感共鸣，最终确定最优的声音设计方案。

活动成效：

学生通过实验可以直观感受语音语速在特定语境下产生的感情，技术参数不仅是简单的数值，更是传递温度、增强叙事感染力的“指挥棒”。

四、教学实践成效与反思改进

通过持续一个月的教学实践，结合课堂观察、学生作品及学习表现，从认知、能力与情感三个维度对教学成效进行分析，并针对实践中发现的问题提出相应的改进策略，形成“实践—评价—反思—优化”的教学闭环。

（一）认知层面：从工具操作到技术理解

教学初始阶段，学生对生成式人工智能理解仅用于简单对话，对生成式人工智能的内在机制与运行逻辑缺乏系统性认知。随着课程深入化学习，学生逐步掌握调用 AI 模型插件操作方法：引入插件观察生成结果—对比不同插件生成结果—选取适合故事内容的插件。

具体而言，全部学生能够认识到文字、图像与语音之间的转换并非简单映射，而是基于 AI 模型的语义理解，同样提示词生成结果并不一样，结果具有随机性。在多次实践中，大部分学生意识到提示词的结构化表达对生成内容结构具有关键作用，能够从角色设定、故事合理分段等方面优化提示词设定。部分

学生开始主动探究不同工具之间的差异，初步形成了工具比较与技术选择意识。

然而，实践也发现，部分学生倾向于直接接受 AI 生成的内容，对生成结果背后的逻辑缺乏进一步思考。针对这一问题，后续教学应由单纯关注“生成了什么”进一步转向关注“为什么这样生成”，通过要求学生在获取 AI 输出后，反向标注出哪些关键词决定了风格，哪些指令导致了偏差，强调“人主导”的地位，建立“构思—初步生成—人工干预—微调优化”的迭代流。

（二）能力层面：从完成任务到问题解决与迭代优化

本课程以实现双语绘本为目标，引导学生在真实任务情境中开展学习，有效提升学生解决问题的思维能力。在故事生成与图像生成过程中，学生能够主动识别生成内容中的逻辑矛盾、角色不一致或画面风格偏差等问题，并尝试用多种方式来解决问题。例如，在图像生成阶段，部分学生发现图像连续性不足，进而想到通过在提示词中设定图片风格来解决。另一方面，学生逐步形成调试与迭代思维，课堂观察表明，学生的学习方式逐渐由最初的“尝试一次—接受结果”，转变为“尝试—反馈—分析—修正—再生成”的循环过程，初步形成了调试与迭代优化思维，体现出过程导向的学习特征。

由于学生的计算机素养与艺术感知力存在天然差异，一刀切的教学目标易导致“优生吃不饱，差生跟不上”，后续教学可进一步实施动态分层任务设计，构建“基础支撑—进阶探究”的弹性学习空间。基础层学生可以借助预设提示词模板和插件完成双语绘本创作，重点理解 AI 生成的基本逻辑；进阶层学生则可以进一步探索参数调节、不同 AI 模型及工具的特点，尝试通过多种工具实现多模态内容生成；对于能力较强的学生，可以设置技术探究任务，并通过“技术互助小组”让其承担技术顾问角色，在帮助同伴解决问题的过程中实现知识的内化与迁移。

（三）情感层面：从被动参与到主动创作

与传统创作任务相比，学生能够快速看到自己的想法转化为文字、图像和声音，为学生提供了较强的互动体验。在课程实施过程中，通过作品展示、同伴评价和成果分享，为学生提供了多样化的表达与交流平台。学生不仅关注自身作品的完成情况，也开始主动观察和评价其他同学的作品，并从同伴的创意

和修改方式中获得新的思路。课堂中逐渐形成了较为积极的学习氛围，学生的合作意识和表达意愿得到增强。学生普遍认为 AI 绘本创作任务“具有挑战性但充满乐趣”，并愿意对生成作品进行多轮修改和完善。这表明，生成式人工智能在降低创作门槛的同时，也能够通过即时反馈和成果可视化增强学生的学习动机。

同时，教师需要警惕 AI 带来的依赖风险，部分学生容易满足于直接获得生成结果，弱化自身的构思、判断和创造。因此，课程后续优化应进一步构建“人机协同”的学习模式，将 AI 定位为学生的创意伙伴和思维助手，引导学生在使用 AI 的过程中保持主体性、批判性和创造性，从而真正实现技术赋能学习，而非技术替代思考。

五、结论与展望

“跨界艺术家”一课实践表明，人工智能技术为课堂教学提供了新的结构性支持，学生在创作过程中不断生成想法、优化表达，并完成多模态作品的设计与呈现。一方面，生成式工具降低了创作门槛，使更多学生能够参与复杂任务。另一方面，多模态生成过程强化学生在文本表述、视觉设计与声音表达之间的整合能力，有助于推动学习从单一知识学习向综合素养发展转变。这一过程体现了从“技术辅助教学”向“技术重构学习方式”的转型趋势。

基于本研究的实践经验，未来课程可在以下方面进一步拓展与深化：

多模态生成的深度整合。 在现有文本、图像与语音生成基础上，进一步探索多模态内容的协同设计与动态交互，构建更加完整的创作生态。

情感计算与个性化支持。 引入情感识别与情绪反馈机制，关注学生在创作过程中的情绪体验与学习投入，提升学习支持的精准性与适切性。

跨学科协同教学模式。 推动信息技术、艺术与语言学科的深度融合，探索多学科教师共同参与的课程设计与实施路径，促进学生综合素养的发展。

参考文献:

- [1] 吴砥,朱莎,王美倩.学生数字素养培育体系的一体化建构:挑战、原则与路径
[J].中国电化教育,2022(7):43-49,63.

作者单位: 民办南模中学

姓名: 史今金

通讯地址: 上海市徐汇区龙华西路 323 号

邮编: 200030

手机: 18707158669

电子信箱: 1249716584@qq.com