科大讯飞
iFLYTEK



2024
智能教育发展蓝皮书

生成式人工智能教育应用

（精华版）

讯飞教育技术研究院
认知智能全国重点实验室智能教育研究中心

科大讯飞 iFLYTEK



2024 智能教育发展蓝皮书

生成式人工智能教育应用

（精华版）

讯飞教育技术研究院
认知智能全国重点实验室智能教育研究中心

版权声明

版权归科大讯飞股份有限公司所有。保留一切权利。

非经科大讯飞股份有限公司书面同意，任何单位和个人不得擅自摘抄、复制本蓝皮书内容的部分或全部，并不得以任何形式传播。

前　言

随着 ChatGPT 的推出，生成式人工智能（Generative Artificial Intelligence，GenAI）迅速崛起吸引了全球范围内的广泛关注。GenAI 作为人工智能技术的最新突破以及新质生产力的典型代表，不仅是引领新一代产业变革的关键力量，也为经济社会发展注入了强劲动能。GenAI 的自然语言理解、内容生成和逻辑推理等强大能力，可以与相关行业融合，赋能千行百业。2023 年 7 月，国家网信办出台的《生成式人工智能服务管理暂行办法》强调，要坚持发展和安全并重、促进创新和依法治理相结合的原则，采取有效措施鼓励生成式人工智能创新发展。2024 年 3 月，政府工作报告明确提出开展“人工智能+”行动，打造具有国际竞争力的数字产业集群。围绕 GenAI 的产业生态正在蓬勃发展。

在教育领域，国家高度重视人工智能对教育的深刻影响，积极推动人工智能和教育的深度融合与创新。早在 2017 年国务院发布的《新一代人工智能发展规划》中就提出，开展智能教育，利用智能技术加快推动人才培养和教学方法改革，构建新型教育体系。2024 年 1 月，怀进鹏部长在 2024 世界数字教育大会上提出，将实施人工智能赋能行动，促进智能技术与教育教学（AI for education）的深度融合，推动以智助学、以智助教、以智助管、以智助研。2024 年 3 月，教育部正式启动了人工智能赋能教育行动。2024 年 7 月，怀部长再次提出要打造中国版人工智能教育大模型，探索大规模因材施教、创新性与个性化教学，更好满足群众“上好学”的需要。GenAI 的出现将推动人类文明进入人机协同的新范式和新阶段，也为教育带来颠覆性变革[①]。但是从当前行业发展及学术研究现状来看，GenAI 教育应用仍处于起步阶段，虽然讨论度高但实际潜能尚未充分发挥。因此，迫切需要厘清技术应用推广的关键问题，提高教育主体对 GenAI 的技术认知和应用技能，明确 GenAI 教育应用的多元场景，同时关注 GenAI 技术伦理风险应对，加快推进 GenAI 在教育中的合理运用。

基于上述背景和思路，我们结合国内外 GenAI 相关政策、技术发展、学界研讨、行业分析和初步实践，对 GenAI 教育应用进行了专题研究和探讨，撰写了《2024 智能教育发展蓝皮书——生成式人工智能教育应用》。本报告是“智能教育发展蓝皮书”系列研究之 2024 年度报告。报告共分 6 章，包括发展现状、挑战与机遇，GenAI 技术框架，教育大模型的构建，GenAI 在教育中的应用场景，GenAI 教育应用伦理风险应对，以及教育领域 GenAI 发展与应用展望。同时，

① 北师大陈丽教授：人工智能对于教育，不是赋能，而是颠覆[EB/OL].(2024-07-03)[2024-07-30].https://mp.weixin.qq.com/s/kYISfhdRX84KHH0bmCKcrA.

报告从教、学、评、管、研视角对全国各地、各级各类学校最新 GenAI 教育应用实践案例进行了梳理和总结。希望通过本研究，为广大一线教育工作者开展 GenAI 教育应用实践提供借鉴和参考，激发更多关于如何利用 GenAI 推动教育创新的思考和讨论，共同迎接“人工智能+教育”的新时代。

本报告的研究和撰写是在科大讯飞智能教育专家委员会指导下，由讯飞教育技术研究院/认知智能全国重点实验室智能教育研究中心负责组织和实施。在报告研究撰写过程中得到了许多行业专家的指导和帮助，公司多个部门领导和专家给予了技术支持，众多一线教育工作者提供了大量实践案例，在此一并表示衷心感谢！在蓝皮书的整体润色中，讯飞星火认知大模型也发挥了重要的作用。由于撰写时间仓促，掌握资料不够，加上作者水平所限，本报告一定存在许多不足之处，敬请批评指正！

《2024 智能教育发展蓝皮书》编写组

2024 年 9 月

目　录

第 1 章 发展现状、挑战与机遇 1

1.1 GenAI 的兴起与影响 1

1.2 GenAI 国产化发展道路 1

1.3 推动 GenAI 教育应用的若干关键问题 2

第 2 章 GenAI 技术框架 4

2.1 GenAI 的提出与内涵 4

2.2 GenAI 的特点 6

2.3 GenAI 应用参考框架 8

第 3 章 教育大模型的构建 9

3.1 教育大模型的提出及发展现状 9

3.2 教育大模型的内涵与特征 10

3.3 教育大模型的体系架构 11

第 4 章 GenAI 在教育中的应用场景 13

4.1 应用场景概述 13

4.2 教学应用场景 13

4.3 学习应用场景 15

4.4 评价应用场景 18

4.5 管理服务应用场景 21

4.6 教育研究应用场景 23

第 5 章 GenAI 教育应用伦理风险应对 25

5.1 GenAI 教育应用伦理风险概述 25

5.2 GenAI 教育应用伦理风险应对策略 26

第 6 章 发展趋势与应用展望 29

6.1 教育大模型产业发展趋势 29

6.2 GenAI 教育应用趋势 30

蓝皮书编写工作组 33

第1章 发展现状、挑战与机遇

生成式人工智能（Generative Artificial Intelligence，GenAI），在自然语言理解与内容生成方面展现出高水平的认知智能，其涌现性、通用性等特性也使得应用范围不断扩张，为教育变革带来了新机遇和新挑战。

1.1 GenAI 的兴起与影响

GenAI 的快速兴起。2022 年 11 月 30 日，美国人工智能研究公司 OpenAI 上线了 GenAI 产品 ChatGPT，其强大的能力和所蕴含的潜力得到了国际社会各界的广泛关注，标志着 GenAI 技术发生了变革性突破，被国际顶级期刊《Science》评为 2022 年度十大科学突破之一[①]。OpenAI 公司自 2018 年起开始陆续推出 GPT 系列模型及产品，不仅性能逐步提升，交互方式也从单模态的文本交互升级为可以支持文本、图像、音频和视频等多种信息的多模态交互。随着以 ChatGPT 为代表的 GenAI 快速兴起和不断进步，全球范围内 GenAI 迎来新一轮蓬勃发展。

GenAI 为经济社会发展注入新动能。GenAI 作为先进科技生产力，凭借其高速性、全自动化、高度智能化等特点大大缩短了社会必要劳动时间[②]，渗透在生产力各要素中并综合作用于生产劳动过程[③]，是新质生产力发展的杰出代表。GenAI 在各行各业、各个领域的应用也不断普及和深化，将极大地提升社会生产效率，加速行业智能化升级发展，为经济社会创造巨大价值。有研究估计，预计在 2023 年到 2040 年间平均每年 GenAI 可使劳动生产率提升 0.1%~0.6%[④]，如果充分应用可能将为全球经济带来每年 2.6 万亿至 4.4 万亿美元的增长[⑤]。

1.2 GenAI 国产化发展道路

从全球创新到国产化落地。随着全球 GenAI 技术加速发展，国际竞争日趋激烈。当前，国外 GenAI 产品存在不对中国用户开放、用户使用数据被传输至境外、数据源难以保证中国语境下的内容质量等问题，可解释性、公平性、价值

① Science. The biggest science breakthroughs in 2022[EB/OL]. (2022-12-15)[2024-04-25]. https://www.science.org/content/article/breakthrough-2022#section_ai.
② 杜博士.科技创新背景下对劳动价值论的再认识:以 ChatGPT 为例[J].现代经济探讨,2024(01):28-34.
③ 张夏恒.新一代人工智能技术（ChatGPT）可能引发的社会变革[J].产业经济评论,2023,(03):22-30.
④ 张彦坤,王雪梅,汪卫国.生成式人工智能对经济社会的影响[J].通信世界,2023(16):43-44.
⑤ McKinsey & Company. The economic potential of generative AI[R/OL]. (2023-06-14)[2024-04-15]. https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier#introduction.

观等多方面均存在风险[①]。习近平总书记在中共中央政治局第二次集体学习时强调要“增强产业链供应链的竞争力和安全性，以自主可控、高质量的供给适应满足现有需求，创造引领新的需求”[②]，为加快国产化 GenAI 建设指明了方向，迫切需要大力提高 GenAI 产业链上关键基础设施的自主可控程度。当前我国在算法、算力和数据三大 GenAI 技术底座上已经有了一定的积累，为 GenAI 的国产化落地提供了较为坚实的基础保障。

国产 GenAI 正加快布局发展。政策引领国产 GenAI 发展多层级布局。在国家层面，推出了《生成式人工智能服务管理暂行办法》等多项相关政策和行动；在地方层面，北京、上海、广东、安徽等省市也均发布了 GenAI 相关政策，各地错位发展态势明显，且都围绕当地的经济特点展开[③]。同时，中国 GenAI 产业也正快速成长。从产业规模来看，2023 年我国 GenAI 的企业采用率已达 15%，市场规模约为 14.4 万亿元，预计 2035 年我国 GenAI 有望贡献超 30 万亿元的经济价值[④]。从产业结构来看，当前北京、长三角和珠三角区域的集聚效应明显，发布主体多以企业为主，包括互联网企业、人工智能企业等。

1.3 推动 GenAI 教育应用的若干关键问题

教育主体对 GenAI 的技术认知亟待加强。推动 GenAI 应用的前提是充分理解技术的效用边际与价值局限，但目前教育主体对于 GenAI 技术的理解和认知仍存在不到位的情况。一方面，存在盲目信任和依赖 GenAI 导致的技术万能论，使得技术在教学过程中脱离实际无限制使用，教学成为“炫技”的场所[⑤]；另一方面，也存在对 GenAI 技术本能的不信任，使得教师产生抵触情绪并对自身技术应用能力产生怀疑[⑥]。在人机融合转型过程中，为了更好发挥技术潜能，推动技术创新扩散，确保应用合规合理，教育主体认知也应转型升级[⑦]，加强 GenAI 教育应用培训，理解 GenAI 运作原理和应用局限，掌握正确使用技巧，建立教育主体自信。

提升 GenAI 教育适用性需要专用大模型加持。由于教育的复杂性和特殊性，

① 车万翔,窦志成,冯岩松,等.大模型时代的自然语言处理:挑战、机遇与发展[J].中国科学:信息科学,2023,53(09):1645-1687.
② 新华社.习近平主持中共中央政治局第二次集体学习并发表重要讲话[EB/OL].(2023-02-01)[2024-06-19].https://www.gov.cn/xinwen/2023-02/01/content_5739555.htm.
③ 人民网.强观察丨人工智能赛道“发令枪”已响，各地有哪些政策“助跑”？[EB/OL].(2023-06-13)[2024-06-28].http://www.people.com.cn/n1/2023/0613/c32306-40012416.html
④ 中国发改委报社.从业界新变化看战略性新兴产业的 2023 年[EB/OL].(2024-01-16)[2024-04-29].https://www.ndrc.gov.cn/wsdwhfz/202401/t20240116_1363298.html.
⑤ 崔晓慧,朱轩.转智成慧：智慧教育时代高职院校智慧教学的困局与突破[J].职业技术,2024,23(06):55-63.
⑥ 魏仙.公立小学教师的数字化技术压力及其管理对策研究[D].华东师范大学,2023.
⑦ 蔡连玉,金明飞,周跃良.教育数字化转型的本质：从技术整合到人机融合[J].华东师范大学学报(教育科学版),2023,41(03):36-44.

通用大模型存在难以适应教育场景的情况，需加强教育领域的专用大模型建设。通用大模型由于缺乏专业的数据训练，教育深度知识不够，回答准确性不足，在实际应用过程中往往难以灵活处理复杂多变的专业性教育任务[①]，不能较好地满足严肃教育教学场景的需求[②]，极大影响师生应用体验。当前业界已经开始开发教育大模型，需进一步加快关键技术的突破创新，充分利用教育领域多模态、长周期的海量数据，构建更具适用性的教育大模型[③]。

GenAI 多元教育应用场景及实效有待探索。为了适应更多元的教育情境特征，解决教育细分场景中的复杂综合性问题和需求[④]，需推进应用场景多元化覆盖与优质应用典型示范打造，并基于科学实证验证技术应用实效。当前 GenAI 教育应用仍处于起步阶段，应用场景较为单一，技术创新应用不充分、应用实践缺乏典型示范等问题依然严重。与此同时，国内研究领域当前更倾向于哲学思辨层面的理论探讨，从实践层面探讨 GenAI 对教育教学影响效果的实证研究还比较少[⑤]，对应用过程存在的问题缺少系统的分析以及针对性的改进意见，亟需在应用场景探索的基础上开展实证调查和案例研究，帮助教师明确 GenAI 的应用范围和能力边界，有效形成最佳实践。

保障 GenAI 合理应用亟需完善伦理规范。教育领域应高度重视 GenAI 教育应用的伦理风险应对，完善 GenAI 教育应用的伦理规范，构建风险应对策略，作为推进 GenAI 教育应用的重要保障。但当前理论层面存在顶层设计缺乏、监管框架缺位、政策标准缺失等问题，实践层面缺少科学完善的应用指南、监管认证工具、平台建设标准、创新实施方法和行动措施，削弱了 GenAI 技术全方位赋能教育的能力[⑥]。因此，教育领域需要高度关注 GenAI 的潜在安全与伦理风险，针对教育领域的应用场景和教育对象的特殊性，从多维度、多视角制定应对策略，构建多方位、多主体[⑦]监管模式，推进 GenAI 在教育领域的合规合理应用与可持续发展。

---

① 中国教育科学研究院数字教育研究所&之江实验室智能教育研究中心.重构教育图景：教育大模型研究报告(简版)[EB/OL].(2023-12-8)[2024-4-15].https://mp.weixin.qq.com/s/kJy-8JcA9Q9cRRTkCAOq7g.

② 曹培杰,谢阳斌,武卉紫等.教育大模型的发展现状、创新架构及应用展望[J].现代教育技术,2024,34(02):5-12.

③ 中国教育科学研究院数字教育研究所&之江实验室智能教育研究中心.重构教育图景：教育大模型研究报告(简版)[EB/OL].(2023-12-8)[2024-4-15].https://mp.weixin.qq.com/s/kJy-8JcA9Q9cRRTkCAOq7g.

④ 尹西明,陈劲,王冠.场景驱动：面向新质生产力的数据要素市场化配置新机制[J/OL].社会科学辑刊,1-11[2024-04-25].http://kns.cnki.net/kcms/detail/21.1012.C.20240419.1539.006.html.

⑤ 秦渝超,刘革平,许颖.生成式人工智能如何重塑教学活动——基于活动理论的模型构建与应用[J].中国远程教育,2023,43(12):34-45.

⑥ 兰国帅,杜水莲,宋帆,等.生成式人工智能教育:关键争议、促进方法与未来议题——UNESCO《生成式人工智能教育和研究应用指南》报告要点与思考[J].开放教育研究,2023,29(06):15-26.

⑦ 方海光,舒丽丽,王显闯.人工智能教育应用伦理的核心议题与应对之策[J].中小学数字化教学,2023(07):30-33.

第 2 章 GenAI 技术框架

GenAI 技术正引领着教育领域的变革创新，正确认识和有效利用 GenAI 技术，需要准确把握其提出背景、概念内涵、技术特点和应用框架。

2.1 GenAI 的提出与内涵

（1）GenAI 的提出

2022 年 11 月，由深度神经网络算法模型 Transformer 所驱动的 ChatGPT 在自由对话、内容生成等语言类任务上取得了重大突破，在将其应用到聊天应用场景后获得了用户的广泛认同和社会的广泛关注，一时间风靡全球，激发了人们关于人工智能新浪潮的讨论。这次浪潮的标志性起点可以追溯至 2014 年 Ian Goodfellow 等研究者提出新的神经网络架构——生成对抗网络（Generative Adversarial Network, GAN）。GAN 采用了两个模型相互对抗、交替改进的训练方式以生成图像，并在图像合成、艺术创作、数据增强等应用中获得成功。这种用人工智能算法模型从无到有地生成图像的方法成为了一种新范式，不仅推动了图像生成技术的发展，也为后续研究者在探索新的生成模型和应用场景方面带来了重要的启发。

Transformer 作为一种重要的基础性的新型神经网络结构，在本次浪潮中大放异彩。OpenAI 以 Transformer 为基础组件提出了“GPT（Generative Pre-trained Transformer）”模型，通过采用“仅解码器（Decoder-Only）”结构降低了模型的计算复杂度，通过采用规模庞大的“预训练（Pre-trained）”结构，使模型能够从海量文本数据中学习到丰富的语言知识和模式。GPT 专注于解决“预测文本序列中的下一个词以生成连贯且有意义的文本”这项任务，为各种语言生成类任务的解决提供了通用的方法，也为实现其他的生成类任务提供了全新视角和技术示范。2018 年，OpenAI 发布了首个 GPT 架构的模型 GPT-1，其参数量达 1.17 亿，可初步实现连贯文本，如文章、故事、对话等的生成[①]。2020 年，OpenAI 发布的 GPT-3 参数量跃升到 1750 亿，显著提升了文本生成的真实性和连贯性[②]。在 GPT-3 的基础上，ChatGPT 通过引入“基于人类反馈的强化学习（Reinforcement Learning from Human Feedback，RLHF）”技术，进一步实现了模型与人类意图的对齐，提升了对话结果的安全性和可靠性。

① 陈晓红,杨柠屹,周艳菊,等.数字经济时代 AIGC 技术影响教育与就业市场的研究综述——以 ChatGPT 为例[J].系统工程理论与实践,2024,44(01):260-271.

② 徐月梅,胡玲,赵佳艺,等.大语言模型的技术应用前景与风险挑战[J/OL].计算机应用,(2023-09-12)[2024-04-25].http://kns.cnki.net/kcms/detail/51.1307.TP.20230911.1048.006.html.

ChatGPT 在技术上的先进性和商业上的颠覆性，引发了全球，特别是中、美人工智能企业的急速跟进。谷歌采取正面竞争的方式，不断发布对标甚至超越 OpenAI GPT 最新版本的大模型 Bard 和 Gemini；Meta 通过不断更新可用于二次开发和调优的开源模型 LLMA，试图用开源生态颠覆闭源商业体系；而微软采取与 OpenAI 深度合作的方式，不断地将 GPT 最新技术引入到自己的产品生态之中。科大讯飞作为中国的人工智能龙头企业，基于其在智能语音、自然语言处理等领域长期积累的技术研究与场景应用优势，在国内首批发布了具有自主知识产权的大语言模型——“星火认知大模型”。2023 年 5 月，在 1.0 版本中形成了文本生成、语言理解、知识问答、逻辑推理、数学能力、代码能力、多模态等七大核心能力[①]，并发布教育、办公、汽车、数字员工四大行业应用成果；在 2024 年 6 月发布的 4.0 版本，七大核心能力全面升级，全面对标 GPT-4 Turbo，并在文本生成、语言理解、知识问答、逻辑推理、数学能力等方面的实现超越[②]。

总之，采用 GPT 结构的算法模型在语言类问题上展现出惊人的能力天赋，不仅使各种语言生成类任务有了系统、全面解决的可能性，而且为其他各种生成类任务提供了技术示范。由此，以实现生成类任务问题解决，以及内容生成与创造为特点的“生成式人工智能”技术，逐渐走向成熟。同时，由于其广泛的场景适用性和巨大的问题解决实用性，使得 GenAI 应用迅猛普及、蔚然成风。

（2）GenAI 的内涵

GenAI 是对解决生成类任务的一类人工智能技术的统称，与此相对的是解决判别类问题的“判别式人工智能”。判别式人工智能通过从带有人工标记的数据中学习到识别相关模式、判定不同类别之间边界的方法，在图像识别、语音处理、医疗诊断等领域取得了成功。但判别式人工智能只能对既有的数据做辨别判定，而对于处理新数据或未知分布的数据等方面的能力明显不足。而 GenAI 可以实现内容创造和新数据生成，有助于联结既有知识，激发新的想法，增强人类对世界的认知与理解，提高人类的创新力与创造力。

GenAI 作为一个快速发展和演变中的概念，其内涵特性也是在不断发展和丰富。考虑到抽象的 GenAI 概念是从具象的 GAN、VAE（变分自编码器）、GPT 等技术应用中总结和归纳形成的，而 GPT 已成为 GenAI 的典型代表和主要形态，由此研究者可以以 GPT 大语言模型为主要研究对象，以 GPT 的技术、应用及价值意义为观察视角，实现对 GenAI 重要特性的诠释。

---

① 汪永安.“大模型”赋能千行百业[N].安徽日报,2024-08-18:4.

② 讯飞星火 V4.0 正式发布，全面对标 GPT-4 Turbo，8 个国际主流测试集测评第一[EB/OL].(2024-6-27)[2024-08-27].https://mp.weixin.qq.com/s/6TnHz4xsLpJzHFmm0_I7aw.

①从技术原理视角来看，GenAI 应具备数据的自监督学习[①]和自动生成的机制。这表明 GenAI 允许模型在缺乏外部标注的情况下，通过内部预测机制进行自我学习，使得复杂的神经网络结构（如 Transformer 或 GAN）能够捕捉并学习数据的分布。例如，复旦大学肖仰华教授认为，GenAI 应能在处理无标注数据时，采用基于遮蔽内容还原的自监督学习任务来训练模型，从而引导模型生成与上下文语境相符合的数据内容[②]。

②从功能应用视角来看，GenAI 应具备通过人机交互明确任务，并根据任务创造性地生成多样化内容的能力。例如，联合国教科文组织官员苗逢春认为，GenAI 是根据人类语言表达的提示词（Prompts）自动生成内容的人工智能技术，既可以通过文字、语音、图像、视频或软件代码示例等方式向模型输入提示词，同样地可以通过文字、语音、图像、视频或软件代码等方式输出结果[③]。

③从价值视角来看，GenAI 应具有强大的创新与创造能力，使其在辅助人类决策、发展人类高阶思维等方面发挥重要价值。例如，联合国教科文组织强调，GenAI 超越了对现有信息整理的范畴，具备了创造性地生成新内容的能力[④]。国际知名咨询公司德勤指出，GenAI 能够生成连贯文本和超逼真图像，使其能够以以往必须通过人类努力地思考与创造才能实现的方式，生成新的数据[⑤]。

2.2 GenAI 的特点

（1）具有能力“涌现”特性

随着模型规模的增大，GenAI 表现出了能力“涌现”特性，即模型性能在其规模达到某个关键阈值之前接近随机水平，一旦超过该阈值，则提升至远超随机水平。例如，当 GPT-3 预训练语言模型接收到特定任务的提示（如自然语言指令）时，无需进一步的训练或参数梯度更新，就能完成新任务的响应，这表明其具有能力“涌现”特性。GenAI 模型的能力“涌现”特性与其规模、技术架构、数据质量和训练过程等息息相关。例如，在少样本（few-shot）学习模式下，当参数规模为 130 亿时，GPT-3 13B 版本的两位数加减法运算准确度低于 60%；当参数规模达到 1750 亿时，GPT-3 175B 版本的两位数加减法运算准确度则提高至

① 注：自监督学习是一种基于机器学习的训练格式和一系列鼓励模型从未标记数据进行训练的方法。
② 肖仰华.生成式语言模型与通用人工智能：内涵、路径与启示[J].人民论坛·学术前沿,2023,(14):49-57.
③ 苗逢春.生成式人工智能技术原理及其教育适用性考证[J].现代教育技术,2023,33(11):5-18.
④ Holmes W, Miao F. Guidance for generative AI in education and research[M]. UNESCO Publishing, 2023.
⑤ 德勤.亚太地区生成式人工智能应用与监管[R/OL].(2024-04-05)[2024-02-01].https://www2.deloitte.com/content/dam/Deloitte/cn/Documents/financial-services/deloitte-cn-fsi-acrs-gai-application-and-regulation-in-apac-zh-240131.pdf.

98%以上[①]。但在另一项研究中，人们在测试 GRE 阅读理解、语言学谜题等 14 个 BIG-Bench 任务后发现，参数规模较小的 PaLM 表现出了优于 LaMDA 和 GPT-3 的性能水平；研究者推测，PaLM 的能力“涌现”可能与训练数据质量和技术架构有关[②]。

（2）具备强大的泛化和迁移能力

GenAI 表现出持续适应新数据、新任务的特性，能够根据环境反馈不断地进行自我改进，具备强大的泛化和迁移能力。GenAI 通过上下文学习从提示词提供的示例中确定待解决任务的自然语言指令（instruct），在不需要额外训练的情况下生成符合预期的文本内容[③]；也能将已经学习到的知识和技能迁移到新的任务或领域中，或将从已有样本数据中学习到的一般性规律泛化到新类型数据的处理任务上面[④]。GenAI 模型卓越的泛化与迁移能力源于深度学习算法的应用、预训练机制的实施及技术架构的持续改进等多重因素。①通过深度学习算法，能够从大量数据中自动学习和提取特征，使得 GenAI 模型在面对新任务时，可以利用已学习的特征进行识别和理解，正确地响应新的任务。②通过预训练和微调，能够将学习到的知识应用到其他相关但不相同的任务上，实现知识迁移[⑤]。③通过技术架构的改进，如引入自注意力机制，能够进一步增强其语言理解和知识学习性能，实现泛化和迁移能力的提升[⑥]。

（3）存在认知“幻觉”现象

GenAI 有时会生成虚假信息或对知识盲区进行貌似合理的错误回答，特别是当使用者具有一定误导性的指令来引导其生成内容时，有很大可能会产生认知“幻觉”现象。Guo 等[⑦]通过构建人类专家和 ChatGPT 的回答对比数据集 HC3，发现 ChatGPT 有时会捏造事实来回答问题。人们将 GenAI 在处理输入任务、维持输出语境连贯性以及与现实世界事实保持一致性时存在的偏差或错误称之为“幻觉”[⑧]。GenAI 幻觉主要有 3 种表现形式：①生成内容与用户输入的任务不相关；②生成内容的逻辑前后矛盾；③生成内容与现实世界中的既定事实不一致。

---

① Brown T B, Mann B, Ryder N, et al. Language Models are Few-Shot Learners[J]. arXiv:2005.14165v4, 2020.
② Wei J, Tay Y, Bommasani R, et al. Emergent abilities of large language models[J]. arXiv preprint arXiv:2206.07682, 2022.
③ 桑基韬,于剑.从 ChatGPT 看 AI 未来趋势和挑战[J].计算机研究与发展,2023,60(06):1191-1201.
④ 张民,李俊涛.生成式预训练模型[J].中国科学基金,2021,35(03):403-406.
⑤ 罗锦钊,孙玉龙,钱增志,等.人工智能大模型综述及展望[J].无线电工程,2023,53(11):2461-2472.
⑥ 张华平,李林翰,李春锦.ChatGPT 中文性能测评与风险应对[J].数据分析与知识发现,2023,7(03):16-25.
⑦ Guo B, Zhang X, Wang Z, et al. How close is chatgpt to human experts? comparison corpus, evaluation, and detection[J]. arXiv preprint arXiv:2301.07597, 2023.
⑧ 赵月,何锦雯,朱申辰,等.大语言模型安全现状与挑战[J].计算机科学,2024,51(01):68-71.

2.3 GenAI 应用参考框架

为帮助教育研究者或实践者更精准地找到 GenAI 技术应用的切入点，并能提出创造性与可行性兼备的需求，参考 DIKW 模型理论框架，给出了一个从 GenAI 技术工作原理到工程应用实践的参考框架。使得研究者或实践者能够从不同的技术层次理解 GenAI，并在分析与应用大模型时能够找准技术脉络和组成模块。参考框架如图 2-1 所示，基础层描述了支撑大模型实现的软硬件基础环境；技术层阐释了从数据获取到信息加工与知识建构、再到智能输出各阶段所涉及的大模型实现所需的各项智能技术；能力层展示了大模型的七项核心能力；应用层描述了对大模型能力的四种常用应用方法；而行业层描述了大模型在各个行业或领域落地应用过程中，需要针对场景做的能力增强与模型优化。

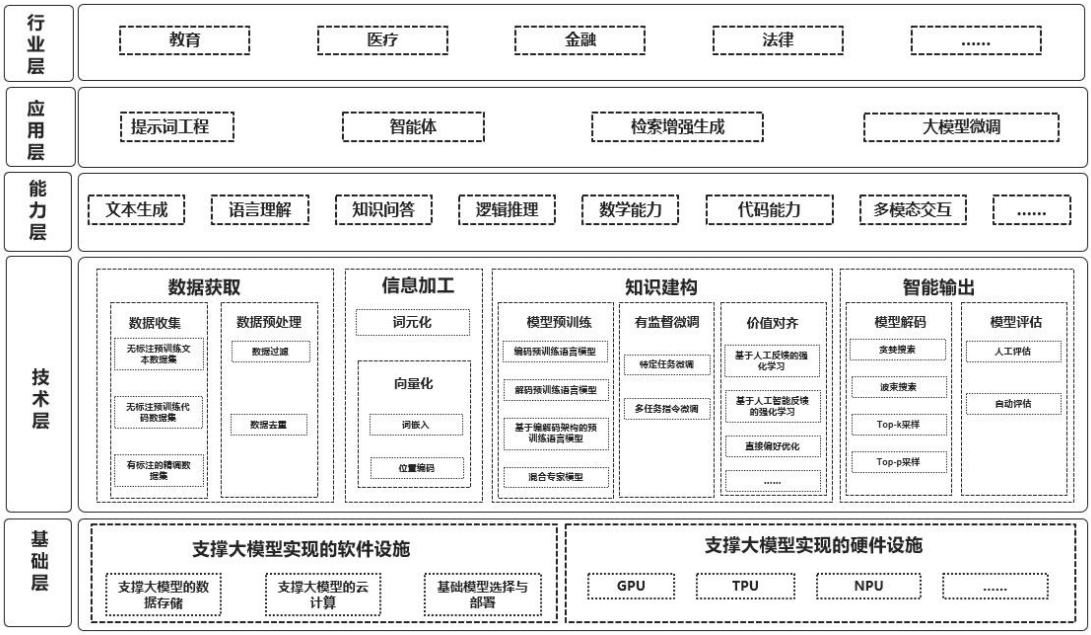


图 2-1 GenAI 应用参考框架

第 3 章 教育大模型的构建

大模型以其强大的泛化性能和跨领域的适应性，能够在多类场景中展现出强大能力，但若仅依靠海量数据训练带来的比较优势，往往很难满足教育场景下更高的知识准确度、价值观与安全性要求。为此，能够满足教育特定场景任务需要的教育大模型应运而生。

3.1 教育大模型的提出及发展现状

（1）教育大模型的提出

面对教育领域的多类迫切需求，需要能够深入理解教学资源、教学对象与教学过程三个教育要素的教育大模型[①]。产业界基于该背景正积极推动适用于教育领域的专用大模型设计及研发，试图通过各种最新技术手段提升模型性能，并探索其在教育全流程中的深化应用。学术界的研究焦点也逐渐从通用大模型在教育中的应用逐步转为对教育大模型的理论深化与价值探讨。推动教育大模型的建设与应用，将既是大模型在教育领域走向专业、科学和准确可靠的深度赋能，也是当前教育行业对科技发展与应用的现实需要。

（2）当前教育大模型的主要功能

为深入分析教育大模型的功能特点，对截至 2024 年 5 月收集到的国内教育大模型信息进行整理及可视化处理，形成图 3-1 所示的教育大模型功能统计。

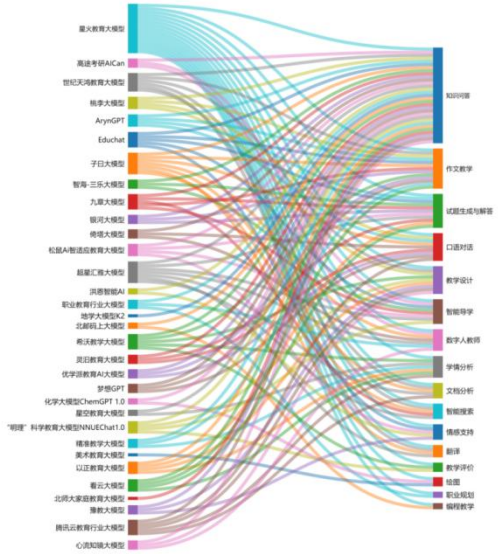


图 3-1 国内教育大模型功能统计

① 卢宇,余京蕾,陈鹏鹤,等.多模态大模型的教育应用研究与展望[J].电化教育研究,2023,44(06):38-44.

基于统计结果，“知识问答”“作文教学”“试题生成与解答”“口语对话”“教学设计”5 类为国内教育大模型的高频功能：①知识问答。基于大模型能力对海量专业信息进行筛选和检索，能够自动回答使用者提出的问题，并利用有效的交互对话界面面向使用者提供专业能力的支持[①]。②作文教学。大模型能够结合海量的作文教学数据，综合所学单元的具体要求，自动生成符合使用者需要的资源；通过交互式对话引导学生依据特定主题或指令帮助学生逐步构建思路等。③试题生成与解答。利用大模型不仅能够提升试题生成的情境适切性与内容的个性化，还能在试题答案解析中给予更为精准的启发辅导。④口语对话。能够为练习者创设真实、自然、完整的多轮口语对话练习环境，让练习者在各种模拟的、贴近真实情境的日常生活、学术或职业场景下练习，以此提高沟通技巧、加深语言技能的掌握和应用。⑤教学设计。能够获取教师输入的教学设计要求，通过调用模型能力并结合单元教学设计要求，了解教师教学意图；从知识库中检索与教学意图相关的知识信息，提供资源应用的来源参考；依据教学设计要求与教学意图相关的知识信息，生成教学设计内容。

3.2 教育大模型的内涵与特征

（1）教育大模型的内涵

目前不同的学术论文、研究报告和媒体报道对其各不相同，包括但不限于“教育行业大模型”“教育大模型”“教育专属大模型”“教育专用大模型”“教育领域大模型”“EdGPT”“教育垂直模型”等。结合当前行业大模型灵活建设实践的综合整理和分析，从广义层面上对教育大模型进行内涵界定：教育大模型是指利用大模型相关技术，针对教育特定数据和任务进行训练或优化，形成具备教育专用知识与能力的大模型及应用，旨在进一步赋能教育领域，满足教育环境中多样化的应用需求，为师生带来更具个性、高效和智能化的学习体验，从而加速推动教育数字化转型和智能化升级，助力教育的创新变革与高质量发展。

（2）教育大模型的能力特征分析

教育大模型的能力特征主要体现在：①场景应用的专业性。精准识别教育场景的差异有助于提升教育效果，因此在教育大模型的构建上也需要依据不同教育场景的需求进行优化和定制，并结合不同学科的特点和差异化的教学目标进行适配。②内容生成的合规性。在教育领域，教育大模型更需要严格遵守相关法规和

① 张春红,杜龙飞,朱新宁,等.基于大语言模型的教育问答系统研究[J].北京邮电大学学报(社会科学版),2023,25(06):79-88.

伦理标准，以确保输出内容符合相关学科的课程标准或课程大纲的要求，保护学生和教育工作者的数据安全，防止未授权访问、泄露或滥用[①]。③知识更新的持续性。教育大模型需要能够不断从最新的教育数据和专业知识中学习，以适应教育内容和教学方法的变化。④技术应用的集成性。除了在教育应用过程中不断提升模型本身的核心能力外，还需要针对各类复杂教育场景提供诸如插件库、知识库等基础构件支持，提供统一的开放接口服务，以此实现大模型技术应用、知识空间与外部资源工具等的集成应用。

3.3 教育大模型的体系架构

在教育大模型的构建过程中，可采取“夯实基础支撑、突出模型能力、健全模型评测”的基本思路，形成教育大模型体系架构，如图 3-1 所示。



图 3-1 教育大模型体系架构

在上述体系架构中，由基础支撑、模型能力、智能体平台、场景应用和模型评测五要素构成。①基础支撑包括硬件设施、算法模型、行业通用数据、教育专属数据训练集等方面。其中，教育专属数据训练集是专门为教育领域设计和收集的数据集，这些数据集包含了已标记的高质量教育数据样本。②在模型能力建设上，需要联动部署教育专有能力、AI 引擎以及多样化的公共基础构件和公共服务 API 等支持多类型通用场景的各类基础能力，以期实现高效协同、性能卓越的资源调度与能力发挥，推动各类资源的共享与高效利用，为用户提供更加智能、便捷的教育服务体验。③智能体平台允许用户创建、配置和部署个性化智能体，以此满足各类教育场景下的不同需求。在教育大模型智能体平台的构成上，往往

① 吴永和,颜欢,马晓玲.教育通用人工智能大模型标准体系框架研制[J].现代教育技术,2024,(4):28-36.

多集成支持教育智能体创建与管理的各类工具，包括教育智能体创建与编排、教育插件库、教育知识库等。④当前教育大模型的实践应用已覆盖教学、学习、评价、科研、管理等多类教育教学领域。聚焦到具体的技术应用场景层面，又可大致分为智能检索场景、咨询答疑场景、创意增强场景、个性化支持与反馈场景和情感互动场景等。⑤基于“可控、可信；安全、绿色；好用、高效”的教育大模型教学应用评测三项原则[①]，在结合教育实际应用需要的基础上，可从“专业能力+安全能力”两方面构建教育大模型的适应性评测框架，以推动教育大模型输出更符合教学需求的精准内容和对话响应，不断提升教育大模型的教育场景适应性能。

科大讯飞 iFLYTEK 讯飞教育技术研究院 iFLYTEK EDUCATIONAL TECHNOLOGY INSTITUTE

① 吴永和,颜欢,马晓玲.教育通用人工智能大模型标准体系框架研制[J].现代教育技术,2024,(4):28-36.

第 4 章 GenAI 在教育中的应用场景

GenAI 在教育领域中的应用，为实现个性化教学、共享教育资源、促进科学决策、提供精准服务、培养创新人才等多个方面创造了新的可能，主要包括教学、学习、评价、管理、科研等典型应用场景。

4.1 应用场景概述

近年来，无论是在具体实践还是在学者研究中，越来越多的区域、学校、教师和学生尝试在教育教学中应用 GenAI，或辅助教学设计，或进行个性化学习，或赋能教育评价和管理，新兴实践经验和案例层出不穷。随着大模型、知识库、检索增强、智能体、虚拟数字人等 GenAI 技术的不断发展，其教育应用也在迭代更新，适用教育领域的细分技术和场景应用工具逐步落地。GenAI 在教学、学习、评价、管理和科研等多个方面创造了众多典型的应用场景，不仅辅助教师优化了教学方法，丰富了师生学习体验，还有效推动智能技术赋能教育评价，提升了教育管理效率，并在科研工作中发挥了重要作用。本报告结合 GenAI 的教育教学实践应用，就各场景的典型应用进行分析，具体如图 4-1 所示。



图 4-1 GenAI 的教育应用场景

4.2 教学应用场景

4.2.1 教学设计智能生成

（1）教学设计方案生成。在提示词的支持下，大模型利用其专业能力，在

教育知识库中检索相关的信息和数据，基于设定生成框架，进而生成具体的教学设计方案，包括教学目标、教学重难点、教学过程和学习评价方法等关键内容。

（2）教学设计优化建议生成。大模型可以基于长文本能力，识别并提取教师上传的教学方案中的关键信息，根据设定的评估标准和模型对教案的各个组成部分进行评估，生成具体的优化建议，如教学目标的调整、教学方法的改进、评价方法的优化等，帮助教师提升教学设计的整体质量。

4.2.2 教学资源协同共创

（1）高效资源检索。资源高效检索。在语言交互界面（Language User Interface，LUI）的支持下，大模型能够理解教师通过自然语言输入的资源需求，从现有的教材、教学材料、在线课程等资源中提取关键信息，快速输出对应的资源。通过 RAG 技术的引入，大模型与外部数据库资源相结合，实现互补效果，大幅提升回答结果的准确性和相关性[①]。教师还可以利用大模型对知识点进行组织和联想，更好地把握知识点的内涵和外延，提升生成资源的丰富性和延伸性[②]。

（2）跨模态资源生成。跨模态资源生成主要包括文本到图像、文本到音频、图像到文本、图像到视频等。大模型在接到教师输入的文本信息后，抽取其中的关键信息，利用跨模态资源能力在不同类型媒体内容之间建立语义关联，合成新的教学资源。通过以上过程，大模型可以将某一形式的教学资源转换成其他多种模态，使得教学资源更加生动和灵活，从而适合不同学习风格和偏好的学生，并为存在视力障碍或其他特殊情况的学生提供帮助。

（3）课程知识图谱搭建。基于课程日常教学中教材、教学音视频、教学课件等材料，大模型提取和分析其中的关键信息并进行相互比对，在元知识库等辅助下形成课程知识图谱，清晰呈现图谱内不同知识点间的关联关系、知识点属性等，为课程资源建设奠定坚实的基础。

（4）课堂实录资源结构化。大模型能对不同课型的授课内容进行理解分析和提炼总结，生成结构清晰的课堂实录视频和文本，以图文并茂的形式展示课堂脉络，作为学生课后复习巩固、教师教学反思和教育教学研究的重要资源。

4.2.3 人机协同课堂互动

（1）创设互动情境。大模型支持教师创设个性化的互动情境，通过虚拟数字人技术实现与学生的角色扮演和场景模拟互动，使得互动活动的形式和内容更

① Lewis P, Perez E, Piktuss A, et al. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks[EB/OL]. (2020-05-22)[2024-04-18]. http://arxiv.org/abs/2005.11401v1.
② 陈静远,胡丽雅,吴飞.ChatGPT/生成式人工智能促进以知识点为核心的教学模式变革研究[J].华东师范大学学报(教育科学版),2023,41(07):177-186.

加生动有趣。在创设互动情境时，教师可以选择使用现有的数字人或自行创建数字人，从形象、声音、动作与表情等方面进行定制，增强数字人的真实感和互动性，使其更加符合课堂情境。

（2）AI 助教或学伴互动。大模型可以在课堂上实现与师生的实时互动，在课堂教学中充当助教或学伴角色，为师生提供教学过程的交互式支持，如解答学生的疑问、提供即时的反馈等[①]。

4.2.4 科学实验智能助教

（1）实验设计方案生成。科学教育大模型能够高效识别和理解教材内容与教师输入的相关信息，在科学教育知识库中检索相关内容，链接与之对应的实验设计方案，有效辅助教师进行实验设计和准备。实验设计方案包含了实验目的、实验准备、实验步骤、实验预期结果等，以及实验注意事项、实验教学引导、实验观察引导等内容。

（2）实验伴随指导。科学教育大模型可以在实验过程中作为智能助教，理解并实时回答学生提出的问题，提供操作指导和原理解释等帮助，精准推荐相关学习资源，促进实验的顺利开展，提升学习体验。推荐的学习资源可能包含实验原理解释和实验演示视频等。

（3）实验评估建议。科学教育大模型能够理解学生的实验报告，通过分析学生实验数据中的模式和趋势，判断其实验结果的合理性和准确性，诊断学生的实验操作情况，并据此输出评价意见和针对性的建议，促进学生实验技能的提升。

4.3 学习应用场景

4.3.1 个性化答疑辅学

（1）个性化问题答疑。问题答疑由学生发起对话。学生在自主学习过程中产生疑问时向大模型提问，大模型首先提取问题的关键词，再基于 RAG 技术，先从一个庞大知识库检索高质量的相关信息，最后基于高质量的相关信息生成更相关和更有意义的回答内容[②]。基于大模型能力的答疑可以满足学生的个性化学习需求，从而提高学生的学习效率和质量。

（2）启发式练习辅导。练习辅导可由大模型发起对话。针对学生练习过程中出现的错误，大模型首先基于学科专业能力和学科试题库数据，针对学生错题快速确定答题步骤，再利用文本生成能力逐步向学生进行启发式提问，最后根据

① 张锋,梁伯聪,范桂佳.“生成式人工智能+讲师”双师互动培训模式构建与应用[J].教育与装备研究,2024,40(02):93-96.
② 顾小清,刘桐.大模型时代的智适应学习研究：进展、实例与展望[J].中国教育信息化,2024,30(05):55-66.

学生的回答内容进行多轮引导，帮助学生完成错题的纠错和改正。基于大模型能力的启发式辅导可以启发学生自主思考解决练习错误，提高学生的思考能力和解决问题能力。

4.3.2 启发式阅读助手

（1）启发式引导。即通过提问的方式来引导学生对阅读文本进行思考。大模型首先可以利用其问题生成能力生成基于阅读目标的问题，对学生进行多轮启发式提问，然后利用其教育专业能力对学生的回答进行评估和分析，并提供个性化反馈，最终引导学生完成正确的回答。基于大模型的启发式引导可以促进学生进行深度阅读，提升学生的阅读学习质量。

（2）过程式伴读。即塑造开放自由式的问答环境，在阅读过程中对学生产生的问题和困惑进行解答，大模型基于 RAG 技术，结合接入的字词和阅读等专业知识库，对学生提出的字词和理解等问题进行识别和检索，然后利用其文本生成能力生成回答。通过以上过程，大模型可以帮助学生便捷地解决阅读中的问题，提升阅读效率和阅读体验。

4.3.3 情境化口语语伴

（1）情境对话。当学生进行口语对话练习时，大模型凭借其对话和翻译能力，结合语音识别、语音生成和虚拟数字人等多模态技术，为学生提供场景化的交互对话环境，根据学生的口语信息，提供个性化实时口语反馈，与学生完成多轮口语对话。基于大模型的情境对话可以增加对话环境的真实感，提高学生口语练习的沉浸感。

（2）对话纠错。学生在进行口语对话过程中，大模型利用其语法分析和检错能力，对学生回复的英语口语内容进行语法分析和错误检查，并实时反馈错误点和修改建议，帮助学生针对性提高口语语法薄弱点，提升语言学习效率。

（3）对话翻译。学生在口语对话过程中，当不理解大模型回复的英文句子时，或不知如何用英文表达句子时，可以利用大模型进行翻译。相较于传统机器翻译，基于大模型的翻译技术可以有更高的翻译效率[①]，并理解文本的上下文语境和语义[②]，还可以进行专业领域的翻译，提高了翻译结果的准确性、流畅性和通用性。

（4）对话练习报告。学生在完成口语对话练习后，结合口语评价标准，大

① 晁忠涛,叶传奇,韩雪磊,等.基于 Transformer 的中英机器翻译系统的研究与开发[J].电脑知识与技术,2022,18(27):16-17+20.
② 张文煜,赵璧.生成式人工智能开创机器翻译的新纪元了吗？——一项质量对比研究及对翻译教育的思考[J].北京第二外国语学院学报,2024,46(01):83-98.

模型可以从语法准确性与多样性、词汇丰富性、发音、流利性和连贯性等多维度生成评语，并给出个性化提升建议，还能根据本次对话主题推荐相关口语学习资源，形成个性化的口语练习报告。

4.3.4 自助式编程学习

（1）代码编写。大模型在学生编写代码的过程中，能够提供丰富的代码知识问答服务，如自动生成代码片段、代码补全，以及提供代码参考示例，从而提高学生的学习效率和编程技能。

（2）代码调试。基于大模型的代码调试不仅包括自动纠错和智能修改，还涵盖了生成测试用例和自动化测试等方面。一方面，大模型能够自动检测代码中的潜在问题，如语法和逻辑错误，并提供相应的提示和建议。另一方面，大模型可以根据学生提供的场景描述和预期性能要求，自动生成相应的性能测试用例，提升测试效率。此外，大模型还可以自动生成相应的测试脚本代码，以提高自动化测试的效率和准确性。

（3）代码解释。基于大模型的理解和推理能力，能够提供深入的代码解释功能，从而帮助学生更好地理解代码逻辑和功能。常见的代码解释有两种方式：第一，即时解释复杂代码，阐明每行代码作用及相互关系，适用于单片段及多文件项目。第二，自动添加注释，嵌入简明注释，提供概览性信息，便于快速理解代码的整体功能。

（4）代码翻译。大模型具备代码翻译能力，将代码从一种编程语言翻译成另一种其他语言，帮助学生在不同的平台或环境中复用代码。大模型进行代码翻译时，能够处理代码中的语法和语义细节，确保在转换过程中尽可能少地丢失信息。

4.3.5 陪伴式心理疏导

（1）心理状态评估。心理大模型可以通过视频、语音、文字等多模态方式的交互式对话，收集学生的必要信息数据，包括基本情况、表情、语速语调和交流内容等。经过对这些数据的深入分析，心理大模型可以准确地评估学生的心理状态。

（2）情感对话生成。情绪感知对话生成模型（Emotion-aware Chat Machine, EACM）能够识别和理解用户的情绪，并据此生成带有特定情感的对话回复[①]，这使得心理大模型在对话中更加贴近学生情感需求，有效提升心理疏导的效果。同时，根据学生的回答和反应，心理大模型能够动态调整对话策略。

① 刘家邑.基于深度学习的情感对话生成关键技术研究[D].华中科技大学,2021.

（3）心理指导与建议。心理大模型具有丰富的心理学知识，可以实现简单的心理咨询，并能够根据用户的个人背景、心理状况和需求，结合专业知识库提供个性化的心理指导与建议。如果心理大模型检测到学生出现严重情绪问题或危机，会及时转介给专业心理辅导师或心理健康机构进行进一步干预。

4.4 评价应用场景

4.4.1 多样化试题生成

（1）试题素材生成。试题素材是试题组成的重要部分。在多选题干扰项生成场景中，大模型能够根据多选题的题干背景文本、题目问题以及正确选项，有效创建在语法、语言流畅性、与问题的相关性以及干扰正确选项的能力方面均表现出较高质量的干扰项候选集[①]。在阅读理解问题生成场景中，大模型可以有效提取和理解大量阅读文本数据中的语义信息，生成句法流畅、符合阅读文章主题或情景、有正确答案的交互式阅读理解问题[②]。在主观题情景材料生成场景中，经过微调包含课本、报纸、论文和各种参考素材语料的教育大模型能够帮助教师生成形式与内容主题上创新，且具有教育意义的主观题情景材料。

（2）自动试题生成。基于大模型的试题助手可以根据选定试题学科、类型、考察知识点、难度等级等信息来批量生成试题；或者根据解析教师上传样例试题的文字内容、公式、图片等来生成相似解题技巧、知识点内容的新试题。此外，大模型能够根据试题模板批量生成试题，从而提高试题命制效率。

4.4.2 智能化作业批改

（1）智能评分。基于预设的评分标准，特别是在语言类作业的批阅中，批改系统能够保持评估的一致与客观性，减少由于人为因素导致的评分偏差，确保每位学生的作业都能得到标准化评分。例如，智能批改系统可以从作文中提取词汇丰富度、句子结构复杂度、语法错误率等特征，通过已评分作文数据训练的大模型得到评分结果。

（2）高阶反馈。基于大量同题作业大数据、高分作答以及作业评分标准数据训练的大模型，可以结合学生错误或者不准确的作答内容特征生成高阶反馈。以语言类作业为例，高阶反馈从篇章结构、文章立意和构思、内容连贯性等给出

① Bitew S K, Deleu J, Develder C, et al. Distractor generation for multiple-choice questions with predictive prompting and large language models[J]. arXiv preprint arXiv:2307.16338, 2023.
② Dijkstra R, Genç Z, Kayal S, et al. Reading Comprehension Quiz Generation using Generative Pre-trained Transformers[C]//iTextbooks@ AIED. 2022: 4-17.

针对性建议[①]。同时，在班级层面生成学情报告，包括班级学生在此次作业中展现的学科知识图谱以及核心素养能力特征的可视化图表。

（3）智能资源推荐。通过教育专用的调优大模型解析作业，识别学生在特定知识点上的掌握程度以及可能存在的学习障碍，从学生的学习进度、偏好以及认知风格等特征，推荐个性化学习资源，包括习题练习、阅读材料或模拟测试等，帮助学生进行薄弱知识点专项练习，素材积累提升对应能力[②]。针对班级共性问题，智能批改系统给教师推荐知识讲解和能力拓展习题，巩固提升班级学生的知识点掌握和能力素养，并提供相应的综合练习讲评讲义。

4.4.3 拟真化口语测评

（1）虚拟数字人模考。在实时对话中，虚拟数字人考官能够结合口语测评技术根据考试真题改编情景对话、话题讨论等贴近日常交流的题型，模拟还原考试全流程，帮助考生熟悉口语评测环节。除此之外，虚拟数字人考官能够呈现出丰富的表情、神态、手势等非语言信息，模拟真实机考场景中与外籍考官的交互，增强与考生的互动性体验。

（2）模考报告生成。大模型可以根据考生的口语对话文本记录，生成从考生回答的发音、语法多样性与准确性、词汇丰富性、流利性和连贯性四个维度包括评分、作答解析以及改进建议的模考报告，科学全面地对考生的口语能力进行评估。

4.4.4 伴随式学生核心素养评价

（1）评价方案设计。教育专用大模型可以辅助教师依据《义务教育课程方案和课程标准（2022 年版）》中的课程要求以及核心素养的评价意见，结合教师设定的具体学段以及评价目标信息，设计包括学生知识、思维能力、学习品质、社会责任等方面的素养评价标准、指标、工具和计划，推荐学生活动群以及点评项目等评价工具，支持写实计分、主观评价、表现性评价等，从而建立更具科学性的核心素养评价方案。

（2）伴随式评价数据采集。在学生核心素养评价实施过程中，特别是动态、复杂性高的学习活动中，智能评价系统可以高效获取评价文本、语音数据、学生作品图片的数据，完成采集和数据汇聚。评价形式包括电子徽章评价、写实评分评价和作品智能评语，伴随式记录学生学习表现，从而有效推动全面的评价数据

① 汪靖,米尔外提·卡马勒江,杨玉芹. 人机共生的复合脑：基于生成式人工智能辅助写作教学的应用发展及模式创新[J]. 远程教育杂志, 2023,41(04):37-44.

② Naseer F, Khalid M U, Ayub N, et al. Automated Assessment and Feedback in Higher Education Using Generative AI[M]//Transforming Education With Generative AI: Prompt Engineering and Synthetic Content Creation. IGI Global, 2024: 433-461.

采集[①]。在此基础上，智能评价系统汇聚整合实时数据并对接各平台，确保学生素养过程性评价数据的完整性、实时性以及交互同步[②]。

（3）评价结果生成。智能评价系统可以根据积累的徽章评价、计分评价以及作品评价等活动数据，依据教师设定的标签数据或者文本、语音交互编辑生成智能评语。大模型对不同种类数据，从文本、语音、图像等方面分析学习者特征要素的准确表征，结合学生素养标准从知识基础掌握与基本技能运用、思维能力和情感态度价值观等层面，自动生成智能评语。除此之外，还可以根据教师设定的人设、数据、字数等基础信息标签，生成定制的评语评价，并支持教师根据提示框架，从情景问题和目标、重要性描述、模型角色定位和输出样式四个方面与系统进行高质量提示交互，生成个性化评语[③]。

4.4.5 智能化课堂教学分析

（1）多模态数据识别。基于 GenAI 的课堂教学分析系统能够集成采集教师、学生的多模态数据，通过语音识别、姿态识别、表情识别和脑数据识别等技术，采集包括师生行为、师生言语、教学设计和教学特征等数据，将其转化为结构化数据以便进行智能分析，为教师改进教学方法、学生优化学习提供可靠依据[④⑤]。同时在教室可安装实时反馈评价数据与结果的电子显示屏或客户端，以便教师在课上实时获得系统的反馈数据，及时调整教学内容与方法[⑥]。

（2）课堂智能分析。多模态数据分析大模型负责对输入的多模态数据进行分类和识别，建立起针对教师、学生、课程三大主体映射关系的训练模型，并训练大量的数据集和测试集，同时引入人工标注进行矫正，提升模型的精度和可靠性，实现对课堂教学的智能解析。如专家评课垂直大模型集成众多教育专家的评课经验和标准，通过“大规模预训练+下游任务微调”方式训练专家评课语料库，实现模拟专家评课内容的生成应用[⑦]。

（3）AI 分析助手。在大模型已预设的分析维度和框架之外，教师或管理者可以根据自身的需求输入具体的指令，让 AI 分析助手提供关于课堂教学内容、

---

① Kusam V A. Generative-AI Assisted Feedback Provisioning for Project-based Learning in CS Education[D]. University of Michigan Dearborn, 2024.
② 程希来. 多数据库系统数据整合平台设计[D].电子科技大学,2009.
③ 罗恒,廖小芳,茹琦琦,等.生成式人工智能支持的教师评语研究：基于初中数学课堂的实践探索[J].电化教育研究,2024,45(05):58-66.
④ 方海光,洪心,舒丽丽,等. 基于课堂智能分析大模型的教师教学能力分析框架及其应用研究 [J]. 现代教育技术, 2024, 34 (02): 43-52.
⑤ 王运武,李雪婷,王藤藤,等. 人工智能助推教师队伍建设：新诉求、潜能与应用场域 [J]. 数字教育, 2022, 8 (06): 13-20.
⑥ 吴立宝,曹雅楠,曹一鸣. 人工智能赋能课堂教学评价改革与技术实现的框架构建 [J]. 中国电化教育, 2021, (05): 94-101.
⑦ 方海光,洪心,舒丽丽,等. 基于课堂智能分析大模型的教师教学能力分析框架及其应用研究 [J]. 现代教育技术, 2024, 34 (02): 43-52.

教师教学、学生学习情况的针对性分析和数据统计。基于 GenAI 的 AI 分析助手可以提供数据挖掘和预测分析，通过建模和算法优化，帮助教师或管理者发现隐藏在数据中的模式、趋势和关联性，为教学决策提供预测性支持。并且支持实时指令输入和分析，能够快速生成反馈和建议。

4.5 管理服务应用场景

4.5.1 管理应用对话生成

（1）应用创建。当管理者或教师需要创建管理应用时，可以与搭载在低代码平台上的智能体展开自然语言交流，完成目标应用的搭建。这种简单直接的操作过程简化了管理应用开发过程，进一步降低了应用开发的门槛，让更多了解区校实际业务需求的管理者或教师参与到应用创建过程中[①]。

（2）应用编辑。当管理者或教师想要对创建的管理应用进行修改和微调时，可以直接对智能体提出修改或优化需求，完成应用的编辑优化。这种操作方式降低了应用维护和更新的成本和难度，使得低代码开发具有更快的迭代速度和更高的灵活性。具体来说，智能体可以参考意图清单识别编辑需求中的编辑意图，然后基于编辑意图对目标应用进行编辑，得到编辑后的目标应用。

4.5.2 问答咨询智能高效

（1）业务咨询。大模型可以快速提供准确的信息和个性化的咨询服务，高效解答用户疑问，提升用户咨询体验。具体来说，大模型能够准确把握用户意图，提高搜索结果和回答的匹配度和针对性。大模型还能够提供文字、文档、链接、音频等多模态答复，确保用户可以直接查看原始信息，提高信息的准确性和可信度。此外，数字人技术与大模型的结合还能使用户感受到与真实人类交流一样的自然感，有效提升用户满意度和咨询质量。

（2）业务指引。当用户遇到不熟悉或操作复杂的业务流程时，大模型可以提供清晰的步骤引导，帮助用户高效完成业务办理。具体来说，大模型可以结合用户的搜索历史和已有提问，预测用户可能感兴趣的问题，并引导用户提问，从而提供更加贴近需求的回答。此外，它还能智能唤醒对应的业务应用或推送业务应用入口，引导用户完成相应的线上业务办理。

（3）业务办理。大模型能够根据用户提供的信息自动生成表单，从而提高业务办理效率。具体而言，大模型可以从与用户的对话中提取关键信息，或借助

① 裴太强,夏伟,张晓晓.基于低代码平台的学生综合素质评价系统建设与应用研究[J].中国信息技术教育,2022,(01):90-93.

OCR 技术从用户上传的图片或文档中提取文本，并将其转化为结构化的表单信息。最终，通过预定义的表单模板，大模型将提取的信息自动填入相应的表单字段，从而简化操作流程，自动引导用户完成业务办理。

4.5.3 行政办公效率提升

（1）公文撰写。基于大模型的公文助手能够通过处理相关人群提供或生成的数据或信息，自动生成内容规范、质量高的各类公文[①]，包括通知、决定、报告和批复等。此外，它还支持对内容的润色、扩写和续写，自动提取历史模板或根据预设的格式模板进行格式校验，提高公文撰写的规范性。公文助手还具备智能化的文档审核功能，能够自动检查文档中的错误和不规范之处，识别敏感词汇，减少人工审核的时间和精力。

（2）智能会议。会议策划阶段，大模型可以根据主题和提示词生成完整策划方案，包括会议环节、时间安排等。会议进行时，基于大模型的会议助手可提供同声传译服务，确保与会人员能够跨语言交流。此外，会议助手还能生成结构清晰、要点明确的会议记录，支持全文摘要、章节速览、说话人总结、多语种翻译和关键信息提取等功能，有效提升会后信息回顾和决策效率[②]。

（3）数字职工。融合机器人流程自动化（Robotic Process Automation，RPA）、低代码和大模型技术打造的数字职工具有处理复杂工作任务的能力，适用于招生、教务排课和财务等各个环节。它能自动化处理教育管理中的大量重复性任务，显著提升工作效率。同时，它能实现教育管理系统间的数据联动与集成，减少手工数据输入错误。此外，它还支持通过自然语言生成流程，降低流程创建难度和维护成本。

4.5.4 管理数据智能分析

（1）数据采集与汇聚。搭载在校园管理平台上的数据分析助手可以自动完成数据的汇聚和整理，减少人工操作的时间和错误，为后续的数据分析和决策提供可靠的数据基础。具体来说，数据分析助手可以自动提取智能设备采集的数据或者管理者上传的表格、视频、音频、图片、文本等材料，然后通过智能识别和解析技术，将其中的非结构化数据进一步转化为结构化数据，便于进一步处理和使用。

（2）数据分析与挖掘。数据分析助手可以根据管理者输入的数据分析请求，

① 魏顺平,范学健,王向旭,等.高等教育应用 ChatGPT 的潜能与风险——来自美国高校的经验与启示[J/OL].现代远距离教育,1-15[2024-07-03].https://doi.org/10.13927/j.cnki.yuan.20240628.003.
② 蔺欣.生成式人工智能的应用与治理[J].数字经济,2024,(05):20-24.DOI:10.19609/j.cnki.cn10-1255/f.2024.05.003.

采用合适的数据分析方法，生成相应的数据分析结果。此外，它还能够挖掘数据背后的深层信息，识别数据间的潜在模式和复杂关联，揭示传统方法难以觉察的信息，为管理者提供更为精准的数据洞察和决策支持。

（3）数据可视化与解释应用。数据分析助手可以通过调用各类可视化工具，将数据分析结果生成易于理解的图表，并将提取的特征和分析结果转化为解读报告，帮助管理者迅速获取有价值的信息。此外，管理者也可以通过自然语言对话方式获取个性化、有针对性的解读和建议，辅助管理决策。

4.6 教育研究应用场景

4.6.1 专业知识智能问答

（1）专业知识答疑。基于对特定领域知识库的深度学习，大模型能够精准理解该领域的专业术语、概念框架以及理论体系，从而提供更为专业的解答，使得研究者只需灵活运用对话指令，就能随时从生成的内容中获取新知[①]。

（2）专业知识推荐。大模型可以快速筛选并推荐最相关的资料，帮助研究者高效获取特定专业领域的信息。基于专业知识库，大模型能够根据主题和关键词精准推荐研究者所需的专业知识信息。同时，它还能处理视频、文本、图像、网站链接等非结构化学科知识，确保信息来源的丰富性和专业性。此外，它还可以智能推荐研究者可能感兴趣的高相关问题或关键词，加速信息检索进程。

4.6.2 学术文献高效研读

（1）文献解读。大模型或搭载了大模型的　科研助手能够批量处理文献，自动生成文献综述、文献主旨、论点分类以及对比分析等信息[②]，辅助文献归纳、整理和分析，帮助研究者深入探索专业领域[③]。具体来说，大模型或科研助手首先将非结构化的文献转换成结构化文本，解析出摘要、方法等模块；然后利用预训练的多语言模型将文本转换为向量表示，捕捉语义和上下文信息；最后从单篇或多篇文献中提取研究主题、结论和论点，基于预创建的模板生成文献解读报告或综述，分析文献相似点和不同点，生成对比分析报告。

（2）文献问答。通过回答针对特定文献内容提出的问题，大模型或科研助

① 刘宝存,苟鸣瀚.ChatGPT 等新一代人工智能工具对教育科研的影响及对策[J].苏州大学学报(教育科学版),2023,11(03):54-62.
② 荆洲,杨启光.生成式人工智能赋能教育研究范式变革：机理、风险与对策[J].中国电化教育,2024,(03):68-75.
③ 王美璇,王冠,王超.媒体融合背景下 AIGC 为学术期刊带来的机遇、风险及应对措施——以 ChatGPT 为例[J].传播与版权,2023,(18):64-66.

手能够帮助研究者快速理解文献，提升理解深度和准确性[①]。具体来说，大模型或科研助手可以结合研究者的问题和选择的引文内容来生成回答。为了提高检索增强生成的效果，避免内容“幻觉”，大模型还可以引入思维链的方法。此外，通过深度学习和自然语言处理技术，大模型可以更好地翻译专业文献中的复杂概念，确保专业术语翻译的准确性。

4.6.3 智能写作全面辅助

（1）研究思路生成。大模型或科研助手可以提供思路启发、假设生成、方法建议和框架拟定等方面的辅助和参考，帮助研究者克服瓶颈，理清思路。首先，它可以分析专业知识库中大量的文献和数据，帮助研究者快速了解前沿研究。其次，它能够根据输入的关键词或问题，生成初步的研究假设和问题框架，启发研究者思考，从而提出原创性想法。此外，通过分析研究主题和已有文献，它还能够生成逻辑性强、结构合理的框架，协助拟定写作大纲，提供写作要点和思路[②]，帮助研究者明确文章的主线和支持论点[③]。

（2）文章编辑与校对。大模型或科研助手可以辅助修改和润色学术文章，提升文章的质量和专业性。具体来说，它可以根据研究者的提示词进行定制化的内容生成，满足特定学科领域、篇幅和语言风格的需求[④]。它还可以优化篇章布局和语法句式，识别文本逻辑漏洞[⑤]，提高文章的逻辑严谨性和专业性。此外，它还能够对文本进行多语言翻译，帮助研究者纠正语法、拼写和标点错误，提供专业词汇和表达建议，以满足各领域期刊的征稿要求[⑥⑦]。

① 政光景,吕鹏.生成式人工智能与哲学社会科学新范式的涌现[J].江海学刊,2023,(04):132-142+256.
② 郝东杰.生成式人工智能时代学术论文写作规范之重申[J].出版参考,2023,(11):45-50.
③ 蒋雪颖,刘欣.生成式人工智能技术下的学术生产与出版：变革、失范与路径[J].数字图书馆论坛,2023,19(05):64-71.
④ 王树义,张庆薇.ChatGPT 给科研工作者带来的机遇与挑战[J].图书馆论坛,2023,43(03):109-118.
⑤ 程梦瑶,彭雪.ChatGPT 赋能学术生产：功能、伦理风险及应对策略[J].新闻前哨,2023,(20):31-33.
⑥ 纪小乐,魏建.生成式 AI 与学术期刊出版：冲击及应对[J/OL].沈阳师范大学学报(社会科学版),1-9[2024-03-14].https://doi.org/10.19496/j.cnki.ssxb.20240029.001.
⑦ 骆飞,马雨璇.人工智能生成内容对学术生态的影响与应对——基于 ChatGPT 的讨论与分析[J].现代教育技术,2023,33(06):15-25.

第 5 章 GenAI 教育应用伦理风险应对

随着 GenAI 技术在教育中越来越多的实践应用，其在应用中存在的隐私泄漏、机器依赖、偏见歧视、学术诚信、学习精神退化等伦理问题也引起了政、产、学、研、用各方的关注[①]。因此，在 GenAI 教育应用中应加强对于潜在伦理风险及其应对策略的研究。

5.1 GenAI 教育应用伦理风险概述

各国政府、高等院校、研究机构、科技企业等均高度关注 GenAI 教育应用的伦理风险，从不同视角展开了探析。本研究选取了 2024 年 3 月 10 日之前发布的与 GenAI 教育应用伦理风险相关的政策文件、行业报告和期刊论文作为研究样本，共计 104 篇。基于三级编码，对研究样本进行文本内容分析发现，当前 GenAI 教育应用伦理风险研究的关注重点包括主要伦理风险、造成风险的可能原因及该风险可能对教育产生的潜在影响三个方面，具体如图 5-1 所示。

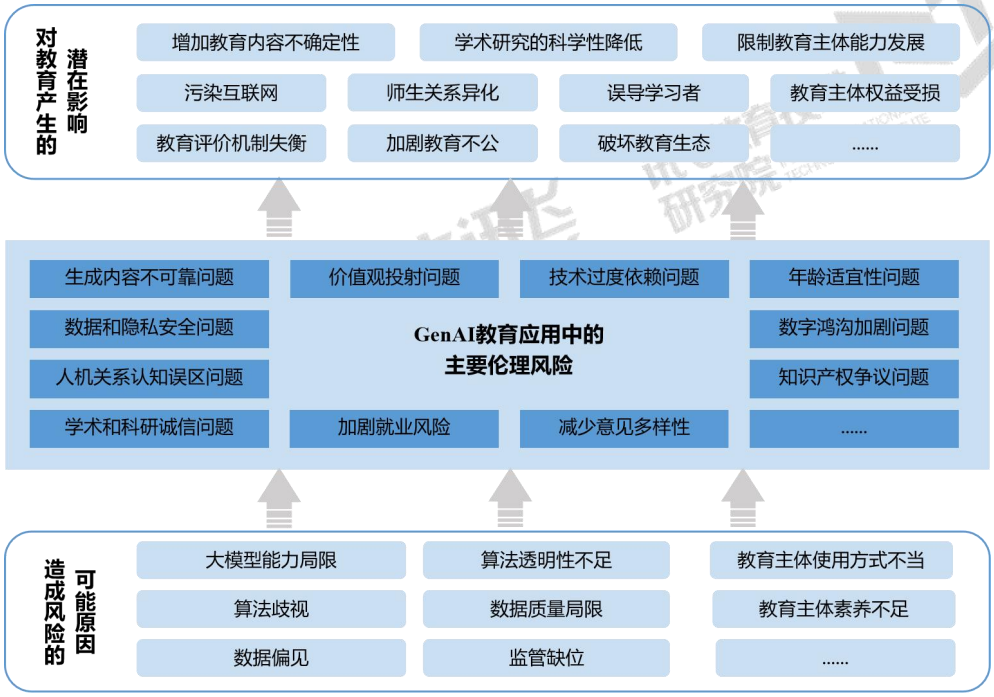


图 5-1 GenAI 教育应用伦理风险研究的关注重点

从主要伦理风险来看，当前主要关注生成内容、数据和隐私、价值观投射、用户认知和使用、知识产权争议等方面的伦理风险。政策文件更加关注价值观投

① 尚智丛,闫禹宏.ChatGPT 教育应用及其带来的变革与伦理挑战[J].东北师大学报(哲学社会科学版),2023,(05):44-54.

射问题，行业报告对数据和隐私安全、生成内容可靠性、知识产权等问题关注更多，期刊论文对各类风险均有关注。从造成风险的可能原因来看，大体可以归纳为三类：一是当前技术发展存在的客观原因，二是教育主体自身的主观原因，三是监管机制方面的原因。从对教育产生的潜在影响来看，当前主要关注 GenAI 教育应用对教育主体、教育整体发展、学术科研、互联网信息环境等多个方面产生的影响。其中，对于教育主体的潜在影响关注度最高，对教育整体发展的影响也受到了广泛讨论。

5.2 GenAI 教育应用伦理风险应对策略

5.2.1 规则视角：推进 GenAI 教育应用伦理风险治理规范化

从规则视角来看，需要以法规为准则，推进 GenAI 教育应用伦理风险治理规范化，保护教育主体权益，确保 GenAI 技术在教育领域的健康发展，最大限度地发挥其潜在价值。具体应对策略包括：①健全 GenAI 教育应用的监管法规。我国可以借鉴其他国家和地区的相关立法措施，加快推进《人工智能法》出台，修订完善现有相关法律法规，进一步健全 GenAI 教育应用的监管法规。②完善 GenAI 教育应用的监管审查制度。通过推进 GenAI 算法和服务备案，强化对教育应用中数据采集的监管[①]，设立专门的监管审查机构对 GenAI 教育应用进行审查、评估和监督应用[②]。③研制 GenAI 教育应用的技术标准。当前我国的 GenAI 技术标准尚不完善，针对 GenAI 教育应用有关的专项技术标准较少，应推动多部门协同，吸纳高等院校、教育研究机构等教育领域专家、科技企业技术专家共同参与，聚焦 GenAI 教育应用存在的数据和隐私安全、学业和科研诚信问题、价值偏见和歧视等具体伦理风险，推进 GenAI 教育应用相关技术标准的研制[③]。

5.2.2 技术视角：推动技术创新与伦理风险治理和谐发展

从技术视角来看，应坚持技术向善，正确处理 GenAI 技术创新与伦理风险治理的关系，提升 GenAI 技术的安全性、稳定性、隐私性。具体应对策略包括：①加强我国 GenAI 自主研发和技术突破，发展安全可控的人工智能。我国应推动基础性原创性技术突破，加大对 GenAI 技术自主研发的持续投入，尤其是技术安全伦理方面的经费投入，并促进开源生态与协同创新，强化国产大模型向开

① 田阳.风险社会视域下生成式人工智能教育应用悖论：生成逻辑、存在限度与风险治理[J/OL].阅江学刊,1-8[2024-06-03].https://doi.org/10.13878/j.cnki.yjxk.20240408.001.
② 兰国帅,杜水莲,宋帆,等.生成式人工智能教育的工作机理及监管框架构建研究[J/OL].西北工业大学学报(社会科学版),1-8[2024-04-19].http://kns.cnki.net/kcms/detail/61.1352.C.20231222.1504.006.html.
③ 苗逢春.生成式人工智能及其教育应用的基本争议和对策[J].开放教育研究,2024,30(01):4-15.

发者开放[①]，构建自主可控的大模型产业生态体系。②推动 GenAI 与人类价值对齐，发展有道德、有温度的人工智能。在技术研发过程中，使用包含正向价值观的数据集对 GenAI 进行训练，充分考虑 GenAI 教育应用中可能出现的伦理问题，提前或及时地嵌入 GenAI 技术运转的机制、程序、代码之中，并加强对于 GenAI 人类价值契合性验证。③研发更具适切性的教育大模型，发展懂教育的人工智能。通过构建高质量的教育领域专用数据集，完善教育大模型标准规范，定期对大模型的性能进行第三方评估，并推动科技企业开发面向教育特定场景的模型集合[②]，加强教师对于定制化智能体的自主开发与应用[③]，以提升教育大模型的教育适切性，为学校教学赋能。

5.2.3 人的视角：提升教育主体的 GenAI 应用意识与能力

从人的视角出发，应坚持以人为本，推动教师和学生提升人工智能素养，合理、道德地使用 GenAI 赋能教与学提质增效。具体应对策略包括：①引导教育主体形成正确的 GenAI 应用观念和意识。通过加强 GenAI 教育应用指南的宣传培训，推动高等院校、研究机构、科技企业和行业协会协同开展 GenAI 教育应用案例分析，组织开展各类 GenAI 教育应用实践活动，加强人机协同教学模式探索与实践等方式，让教师和学生能够在教学实践中正确认识人机关系[④]，调整对于自身教育主体地位的认知，主动转变观念。②培养教育主体掌握 GenAI 技术应用能力。参考联合国教科文组织提出的“针对学校学生和教师的人工智能能力框架草案”（AI Competency Frameworks for Teachers and for School Students），GenAI 技术应用能力可以包括学会使用和创造性地使用两个层面。通过完善学校人工智能相关课程，为师生提供 GenAI 相关设备和工具，将人工智能素养纳入师生评价体系等方式，提升师生的 GenAI 技术应用能力。③提升教育主体践行 GenAI 技术伦理的社会责任。通过定期开展法律法规、伦理规范宣传学习，完善 GenAI 应用的监督反馈机制[⑤]，充分发挥教师的引导作用，推动教育主体在 GenAI 教育应用中承担 GenAI 技术伦理的社会责任，共同维护 GenAI 教育应用的伦理边界。

---

① 科大讯飞.第 22 次参加全国两会，刘庆峰带来 8 份建议[EB/OL].(2024-03-06)[2024-04-17].https://mp.weixin.qq.com/s/-FByCkpi5mPRbBuUXs6Agw.
② 黎加厚.生成式 AI 大模型教育应用的四大秘诀[J].中小学信息技术教育,2024,(05):8-11.
③ 教育信息化 100 人.教育部信息科技课标组组长熊璋：AI Agent 是一种新质生产力[EB/OL].(2024-06-07)[2024-06-14].https://mp.weixin.qq.com/s/L2Kx2x6oONv3DiETZnPzbg.
④ 何昌旺,熊和平.ChatGPT 类生成式人工智能教育伦理危机及其应对[J].中国教育信息化,2024,30(02):81-90.
⑤ 姚志伟,李卓霖.生成式人工智能内容风险的法律规制[J].西安交通大学学报(社会科学版),2023,43(05):147-160.

5.2.4 应用视角：加强 GenAI 教育应用的指导、监测与研究

从应用视角来看，为有效防范 GenAI 教育应用的伦理风险，必须坚持应用为王，强化对 GenAI 教育的指导、监管和研究工作。具体应对策略包括：①制定 GenAI 教育应用指南。教育主管部门牵头制定面向教育领域的通用 GenAI 教育应用指南，并结合不同教育应用场景、面向不同年龄阶段/不同身份角色的使用者制定针对性的 GenAI 教育应用指南。②加强 GenAI 教育应用的全过程监测。通过加强 GenAI 工具和服务的质量与合规性检测，研发 GenAI 教育应用伦理风险预警系统，开发内容过滤和审查机制，为使用者建立良好反馈机制等方式，实现对 GenAI 教育应用全过程的监测。③持续开展 GenAI 教育应用伦理研究。主要包括开展 GenAI 伦理风险发生机理的研究，实施师生应用影响的长期跟踪研究，加强 GenAI 教育应用伦理的跨领域、跨学科的协同研究，全方位、多维度地持续推进 GenAI 教育应用伦理风险研究工作。

第 6 章 发展趋势与应用展望

GenAI 作为新质生产力要素将推动人类文明进入人机协同的新范式和新阶段[①]，将推动教育理论与实践的创新发展和教育教学形态的变革与重塑，促进教育高质量发展。

6.1 教育大模型产业发展趋势

6.1.1 教育大模型产业定位与分工日趋明晰

大中型企业大多占据教育大模型产业链上游，以 MaaS（Model as a Service，模型即服务）作为主要商业模式，更倾向于开发自研模型，通过打造大模型平台等方式为其他企业提供大模型能力服务。小微企业大多占据产业链中游，以 SaaS（Software as a Service，软件即服务）作为主要商业模式，主要使用上游提供的成熟模型，针对某一具体场景开发相应教育服务，以满足各类用户的多样化需求。

6.1.2 集成化一体化将成为大模型落地的重要支点

多种功能集成的教育大模型产品能更好满足教育实践需要。在初期应用探索阶段，出现了很多功能较为单一的教育大模型应用工具，且分散在不同网站或 APP 内，教师需要在多个系统平台间跳转应用，成本较高。随着技术的进步和应用探索的深入，为满足教育场景应用的多样化需求，教育大模型产品开始呈现出多种功能集成的特征。搭载教育大模型的软硬件一体化产品将在企业营收中占据重要地位。从模式识别、神经网络到大模型，教育智能硬件一直是 AI 前沿技术的试验田，同时“AI 竞逐”已成为教育硬件市场的趋势[②]，搭载教育大模型的教育硬件产品已在市场上获得广泛认可。

6.1.3 多模态教育大模型逐渐成为主流形态

具备多模态内容生成能力的教育大模型更能满足教育应用需要。图片、音频、视频等多模态教育资源有利于对宏观和微观事物、抽象事物进行生动直观表达、对复杂过程进行简化呈现等，可以促进学生理解和激发学生兴趣。当前，很多学校和教师已经开始使用大模型生成多模态的教育资源来辅助教学，兼具多种模态

① 北师大陈丽教授：人工智能对于教育，不是赋能，而是颠覆[EB/OL].(2024-07-03)[2024-07-31].https://mp.weixin.qq.com/s/kYISfhdRX84KHH0bmCKcrA.
② 21 经济网.教育市场创新趋势报告丨教育智能硬件：巨头博弈的下一个赛场[EB/OL].(2024-04-24)[2024-06-20].https://www.21jingji.com/article/20240424/herald/5c1efd0589220bfef940b465646f93b5.html.

教育资源生成能力的大模型能更好满足教师实际需要。多模态数据处理将成为教育大模型的重要应用方向。随着人工智能技术与教育教学融合的日益深入，交互式学习、探究式学习等学习理念不断普及，教学方式趋向于更加丰富和多样化，课堂实录、视频作业等多模态数据在教育数据中的占比将不断上升，大模型对多模态数据的处理能力也将随之不断提升。

6.1.4 教育大模型的专业性和通用性将日益增强

教育大模型的专业性将越来越强。教育领域涉及知识讲授与传承，其对于知识的准确性和专业性要求高，需要教育大模型使用高质量的专业教育数据进行训练和优化，提升教育专业性，为师生提供更加准确合理的知识问答和内容生成等服务。教育大模型的通用性将越来越强。教育领域涉及“教、学、评、管、研”等多类场景，经常涉及错综复杂的跨场景问题，需要教育大模型支持更多的功能，从而能够快速处理解决这些复杂问题，满足教育教学多场景应用的需要。

6.1.5 教育大模型研发与应用需要学校积极参与

高校将积极参与教育大模型的研究和开发。高校作为教育主体和科研主体，在教育大模型的研发上具有独特优势。当前，很多高校已经参与教育大模型研发，并取得优秀成果[①]。未来，高校将进一步参与大模型研究和开发，并产出更多高质量成果。中小学校将积极探索基于教育大模型的教学应用创新。当前，作为教育大模型的主要应用主体，国内许多中小学校都开展了大模型教育教学实践探索，积累了宝贵经验，并为教育大模型研发和优化提供了相应反馈和数据支持[②]。未来，随着大模型赋能的智能教育产品和技术服务越来越丰富和完善，大模型技术在学校教育教学中的应用探索将越来越广泛和深入。

6.2 GenAI 教育应用趋势

6.2.1 人才目标转向更加重视创新能力培养

高阶思维能力成为未来创新型人才培养的重要内容。GenAI 将引发社会生产生活方式和职业岗位的变革，能够批判性分析、理解和评价知识，并能创造性运用多领域多学科知识解决问题的创新型人才是未来社会发展的需要。因此，培养学生高阶思维能力将成为重要教学内容和人才培养目标。人机协作能力将成为未

① 浙江大学上海高等研究院.“智海-三乐”教育大模型正式发布，赋能 101 计划《人工智能引论》教与学[EB/OL].(2023-08-22)[2024-06-20].https://mp.weixin.qq.com/s/NYSRZnRU8TqWC_jnbXHFVA.
② 广州市天河区龙洞小学.AI 时代的语文启蒙之旅：基于大概念和 AI 技术支持的小学语文低段识字教学实践探索[EB/OL].(2023-12-06)[2024-06-20].https://mp.weixin.qq.com/s/tuOWhu6ymyjZfbdq16YWPQ.

来人才的必备关键能力。AI 技术将通过硬件、软件等多种形式充斥未来的工作、生活和学习，人机互动将成为人们常态化活动方式，机器的角色将从工具变成伙伴，人机协作能力将成为未来人才的关键能力之一，能够灵活运用智能技术、通过人机协同解决问题的创新应用能力将变得至关重要，不能利用 GenAI 技术有效赋能工作的人将逐渐落伍直至被淘汰。

6.2.2 教育环境转向智能升级和虚实融合

基于大模型研发出来的各类智能体将以独立软件系统或功能模块等方式加载到智能黑板、学习机等硬件以及教学平台、管理系统等软件平台中，形成各类 GenAI 助手，有效提升教学硬件和系统平台的智能化水平。而通过以大模型为基础的各类助手工具可以实现相关软硬件、系统平台之间数据的互联互通，并将促进教学环境一体化。同时，通过生成虚拟场景、建筑、景观、人物角色和物品等，和创设意义情境、故事情节，大模型可以为虚拟学习空间构建提供强大支持。通过制作虚拟数字人，创造个性化和引人入胜的体验，将打破虚拟空间和现实空间的边界，促进虚实空间更自然融合。

6.2.3 教学方式转向人机协同教学

在大模型赋能下，“机器教师”能够拓展和增强人类教师能力，为学生提供个性化指导，助力实施规模化因材施教。并能够在教学中不断提高和师生的契合度，根据师生的选择不断调整反馈的内容，及时为师生提供所需要的教学服务，提高教学质量和效果。“机器教师”可以独立完成绝大部分课堂练习批改、辅导答疑等简单重复性工作，以及多模态学习数据采集和处理分析等工作量巨大的复杂工作，人类教师将有更多时间和精力用于学生核心素养、道德情感等方面的培育及相关课堂活动组织[①]，并基于这些数据准确识辨学生学习水平和存在的问题、及时调整教学策略和教学活动[②]，不断推动教学方法、策略和模式等方面的改革创新。

6.2.4 学习方式转向生成式个性化学习

基于大模型的智能学习助手可以针对学生实际情况生成与之匹配的学习内容，并根据学生的即时性学习表现进行智能纠偏和学习路径调整，从而有效提高学习效果和学习效率，还可以智能生成图片、视频、课件等数字化学习资源，为

① 方海光,孔新梅,刘慧薇,等.基于共生理论的人机协同教育主体合作博弈及其优化策略研究[J].电化教育研究,2024,45(01):21-27.
② 宛平,顾小清.生成式人工智能支持的人机协同评价:实践模式与解释案例[J].现代远距离教育,2024,(02):33-41.

教与学提供支持。同时，智能学习助手可以引导学生开展思维链式的对话，通过一系列由表及里、由浅入深的连续提问和迭代追问，帮助学生在持续探索中了解知识内容、解除问题疑惑、加深知识理解和掌握，并支持学生在人机对话中开展对弈式训练培养高阶思维能力。

6.2.5 教育评价转向数据驱动的发展性评价

基于大模型的智能学习助手可以在持续的人机交互对话中完整记录学生的学习过程和成长轨迹，并综合分析学生的知识学习情况和表现，推动教学评价从依靠教师或专家经验走向追求数据实证。同时还可以基于所采集的人脸、语音、表情等多模态数据，对学生学科能力、情感态度、思维品质等进行综合分析，全方位刻画学生的问题解决能力、批判性思维能力、跨学科能力、创造力等高阶思维能力[①]。基于大模型对过程性学习数据的记录、分析、挖掘和呈现，智能学习助手能够从而全面、形象地展示学习者的学习状态和学习偏好，为学习路径规划、学习方式选择等提供实时数据支撑和建议，并可以通过实时反馈和调控，实现对每一个学习者学习状态的关注和及时指导，保障每一个学习者的学习体验和学习效果。

6.2.6 教育治理转向全球化智能化治理

以 GenAI 为代表的人工智能发展日新月异，在助力教育治理效能提升的同时，也带来许多新的诸如数字鸿沟、算法歧视等需要全世界共同解决的风险与挑战。2024 年世界人工智能大会发布《人工智能全球治理上海宣言》，指出只有在全球范围内的合作与努力下，才能充分发挥人工智能的潜力，为人类带来更大的福祉。依托大模型强化数据要素价值挖掘、分析、应用与管理，可以帮助教育管理者预测与感知教育治理风险，并为精准解决教育治理难题提供具体方案，促进教育治理效能提升[②]。同时，将大模型嵌入教育管理与服务平台中，作为业务处理助手，有助于提供个性化教育服务，提高业务办理效率和质量。

① 郑永和,王一岩,杨淑豪.人工智能赋能教育评价：价值、挑战与路径[J].开放教育研究,2024,30(04):4-10.
② 周洪宇,李宇阳.生成式人工智能技术 ChatGPT 与教育治理现代化——兼论数字化时代的教育治理转型[J].华东师范大学学报（教育科学版）,2023(7):36-46.

蓝皮书编写工作组

专家委员会

主任　陈　丽（北京师范大学教授、博导　中国教育技术协会副会长）

副主任　聂小林（科大讯飞股份有限公司高级副总裁）
郭绍青（西北师范大学智能教育研究院院长、教授、博导）

委员　（按姓氏笔画排序）

方海光（首都师范大学教育技术系数字化学习实验室主任、教授、博导）
孙　波（北京师范大学未来教育学院副院长、教授、博导）
刘邦奇（科大讯飞教育技术研究院院长、教授、博导）
刘革平（西南大学教育学部教育技术学院院长、教授、博导）
陈明选（江南大学江苏“互联网+教育”研究基地主任、教授、博导）
杨现民（江苏省教育信息化工程技术研究中心常务副主任、教授、博导）
吴　砥（华中师范大学国家数字化学习工程技术研究中心副主任、教授、博导）
汪　琼（北京大学数字化学习研究中心主任、教授、博导）
张　海（东北师范大学传媒科学学院副院长、教授、博导）
胡小勇（华南师范大学教育人工智能研究院常务副院长、教授、博导）
钟晓流（清华大学信息技术中心副总工程师　中国教育技术协会副会长）
贾积有（北京大学教育学院教育技术系主任、教授、博导）
顾小清（华东师范大学教育信息技术学系主任、教授、博导）
郭　炯（西北师范大学教育技术学院院长、教授、博导）
熊　璋（对外经济贸易大学信息学院院长、教授、博导）
黎加厚（上海师范大学教育技术学系教授）

科大讯飞专家　周佳峰、章继东、王士进

编写组（按姓氏笔画排序）

编写者　王亚飞、王　涛、邓　倩、代怿鑫、朱广袤、朱洪军、刘邦奇、刘梦蝶、刘碧莹、许容克、张国强、张金霞、赵子琪、胡婷婷、姚兰婷、袁立方、袁婷婷、耿　婧、凌婧婧、黄倩倩、康美玲、喻彦琨

技术支持　刘国飞、刘俊生、李晓枫、卓　晗、胡雯璟、查洪钰、贺　胜、郭红杰、黄国庆、常雪亮、韩　萌

案例支持　马长俊、王　颖、王雨蒙、石　楠、朱礼稳、刘　君、刘晓娟、许　磊、吴　溪、张　睿、张运豪、陈　龙、陈俊鹏、金　晶、孟　勇、索长安、钱立勇、殷　乐、姬润果、黄　倩、黄　涛、程　霞



了解更多关注
“讯飞 AIEd”公众号