



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方的广泛关注，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 发展多层次布局。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，为了更好地发挥技术潜能，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|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，教育专业知识不够，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，亟需在教育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，由深度学习算法模型 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—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。

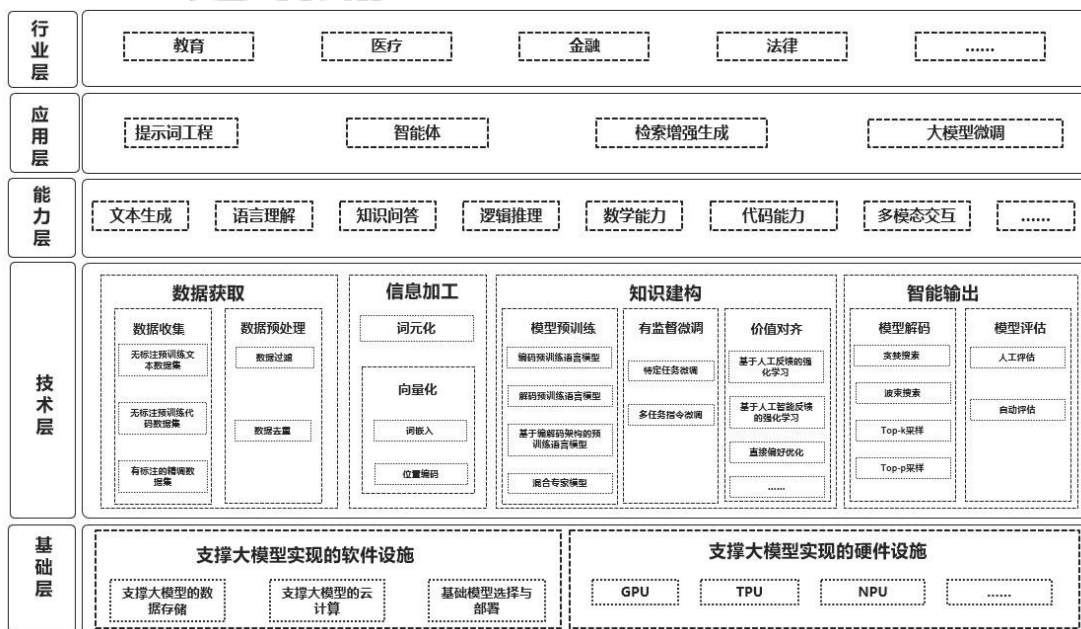


图 2-1 GenAI 应用参考框架

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。

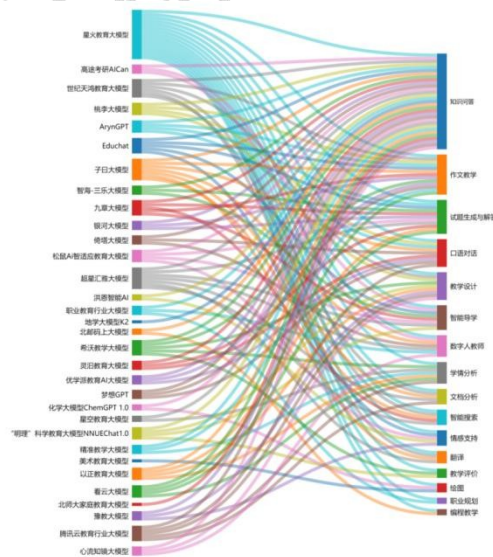


图 3-1 国内教育大模型功能统计

^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適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和 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

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性能。



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

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/438017057042006135>