人工智能教育治理进展研究报告

（2026 年）



中国人民大学人工智能治理研究院

人工智能教育治理研究团队

摘 要

本报告系统研究了 2025 年至 2026 年 4 月期间，人工智能技术快速演进背景下全球教育体系正在经历的结构性变革。与此前教育信息化、在线教育、智慧教育等阶段相比，本轮人工智能浪潮的根本区别在于，AI 已不再仅仅是教育工具，而是正在成为教育活动中的“参与者”、“协作者”乃至“组织者”。报告从全球发展趋势、三大危机、各国治理模式、中国实践、治理原则及对策建议六个维度，对这一时代命题进行了系统分析。

全球人工智能教育发展正进入一个具有结构性意义的新阶段。人工智能正从辅助工具转向教育基础设施，重构教育系统的底层运行逻辑。生成式人工智能、大模型及智能体技术的快速发展，使 AI 开始承担教育协同与系统组织功能，推动教育从经验驱动转向数据驱动、从周期性管理转向实时性管理。据中国教育发展战略学会调研，人工智能已实现对中小学备课、授课、学练、考试、评价、管理六大核心教学环节的全流程覆盖，60.7%的教师基于提升效率目的、56.7%因减轻负担目的使用 AI 工具。教育系统对 AI 的依赖正从效率工具依赖转向系统能力依赖，课程内容开始呈现动态生成特征，传统统一教材、统一教学、统一评价的标准化模式正被统一目标、差异路径、动态反馈的智能化模式所替代。这一转型正在重塑全球教育竞争格局——教育竞争不再仅是课程与师资的竞争，更是教育智能系统能力与教育治理能力之间的综合竞争。据 HolonIQ 预测，全球 AI+智慧教育市场规模预计 2025 年达 850 亿美元，2030 年突破 2400 亿美元。

教育大模型成为核心竞争方向。OpenAI 推出 ChatGPT Edu，Google 推出 LearnLM，微软将 Copilot 深度整合教育场景；中国垂类教育大模型快速发展，科大讯飞、网易有道、学而思等围绕题库、课程、学习行为数据构建专业化模型体系，教育 AI 竞争正从通用能力竞争走向教育理解能力竞争。与此同时，学校私有化模型部署与教育主权意识上升，清华大学、北京大学等高校陆续推动校内 AI 平台建设，教育数据已成为核心战略资源，正在从结果记录转向学习过程建模，其战略价值已超越技术范畴，开始影响国家教育竞争力。联合国教科文组织在 2023 年发布的《生成式人工智能教育与研究应用指南》中明确提出，教育数据不应完全由商业平台控制，欧盟《人工智能法案》也将涉及未成年人的教育数据列为高敏感数据。

然而，人工智能深度嵌入教育也引发了三大危机。技术中心主义危机表现为三重异化：评价标准异化——将是否使用 AI 而非是否解决教育问题作为评价标准，技术从手段异化为目的，学校被迫为满足技术指标而表演性使用 AI；数据迷恋——过度采集注视时长、点击率、答题正确率等可量化行为数据而忽视顿悟、批判性质疑与价值判断等深层学习过程，可测量性僭越了教育性；复杂系统简化——将涉及认知、情感、社会关系、价值观念等多重维度的复杂教育问题简化为技术优化问题，人本要素被系统性排除。

主体性危机体现为三个层面：学生认知外包导致批判性思维、创造力等高阶认知能力面临萎缩风险，提问—获得答案的即时反馈循环取代了从探索到试错再到反思的完整认知过程；教师角色认

同危机——英国教育部调查显示 64%的教师担忧 AI 削减工作岗位，48%认为自身角色可能被取代，教师从知识传授者被弱化为技术操作者，教学决策空间被算法持续压缩；权力结构失衡——知识主权、教学决策权与评价权三重让渡于算法，学生被持续监测、分类和预期管理，人类主体沦为技术规训的产物。

信任危机表现为三重信任同时受损。学术诚信危机——英国谢菲尔德大学 AI 相关学术不端案例一年内从 6 起飙升至 92 起，增长 15 倍，哈佛、牛津、斯坦福等高校陆续发布生成式 AI 教学使用指南，强调“AI 辅助不等于 AI 替代”。算法与隐私信任危机——72%的人接受使用 AI，但 54%认为 AI 带来网络安全、隐私泄露及错误信息等显著风险，形成“用而不信”的矛盾心态。学历信号价值稀释——当社会无法确定一份优秀成绩单究竟反映学生真实能力还是其使用 AI 工具的熟练程度时，教育体系的社会公信力随之动摇，正规教育体系最核心的社会功能——能力认证面临失效风险。

在全球治理方面，报告梳理了七种代表性模式。欧盟以《人工智能法案》建立风险治理导向的监管框架，将教育 AI 列为高风险系统，强调透明性、可解释性与可追责性。美国采取原则导向与地方自治相结合的模式，联邦政府提供指导，州和学区自主实施，强调“以人为中心”。英国强调风险规制，将监管扩展至产品开发环节，2025 年发布《Generative AI Product Safety Expectations》。芬兰坚持教育价值与伦理导向，秉持“教育优先于技术”理念。新加坡构建国家统筹型治理，通过统一“学生学习空间”平台和 AIEd 伦理框架实现集中监管，提出主体性、包容性、公平性、安全性四项核心原则。日本采取审慎创新导向，以试点探索、风险分级和场景治理为特征。韩国推行技术驱动与国家推进相结合的模式，以 AI 数字教材改革为代表。这些模式反映了不同国家在技术创新、风险控制、教育价值和制度能力之间的差异化选择。

中国人工智能教育发展具有鲜明的政策驱动特征。2025 年教育部等九部门印发《关于加快推进教育数字化的意见》，2026 年印发《人工智能+教育》行动计划，形成从战略到专项行动的完整政策框架。国家智慧教育平台用户已突破 1.78 亿，覆盖 200 多个国家和地区，汇聚 13 万余条中小学资源、1.25 万余门职业教育课程和 14.5 万门高等教育课程。中国拥有全球规模最大、层次最完整、区域差异最显著的教育体系，构成了 AI 教育应用的独特场景优势。治理方面，顶层设计持续完善，治理方式逐步走向分学段、分场景的分类治理，初步形成政府引导、学校落实、企业参与、家庭协同的多元共治格局，但治理能力仍处于持续完善阶段，技术快速迭代带来新的制度适配压力。

基于上述分析，报告提出人工智能教育治理的六大原则：坚持教育价值引领，技术服务于人的全面发展；坚持以人为本，在人机协同中确立师生主体地位；以教育公平为底线，缩小数字素养鸿沟；打造适配技术迭代的监管体系；全面提升师生数智素养；推动多元共治，凝聚政府、学校、企业、家庭的治理合力。在此基础上，报告从五个方面提出具体对策：强化顶层设计，完善法律法规与政策体系；抬高准入门槛，建立以教育适配性为核心的产品标准，实施分类治理；统筹数据治理，建立贯穿教育 AI 应用全过程的数据安全体系，坚持数据最小必要原则；重塑教育评价体系，回归育人本质，建立适应 AI 时代的综合评价机制，提高高阶思维、创新能力等综合素养的评价权重；持续赋能教师队伍，加强 AI 应用与伦理培训，夯实人机协同的专业基础。

综上所述，2025—2026 年是人工智能与教育深度融合的重要转折阶段。AI 正在推动教育从工业化时代的大规模标准化体系向智能化时代的大规模个性化与人机协同体系转型。这一转型不仅是技术工具的升级，更意味着教育理念、学习方式、人才培养模式及教育治理逻辑的系统性重构。人工智能教育治理的根本命题是：在技术创新与风险防控、效率提升与公平保障、工具理性与价值引领之间建立动态平衡，使 AI 真正服务于人的全面发展和教育公共利益的实现。

关键词：人工智能教育　教育治理　大模型　教育数字化　人机协同

目 录

摘 要 1
目 录 4
前 言 6
1. 全球人工智能教育发展的主要趋势 7
1.1 AI 从辅助工具转向教育基础设施 7
1.2 教育大模型成为核心竞争方向 12
1.3 教育数据成为核心战略资源 15
2. 人工智能教育应用的三大危机 19
2.1 技术中心主义危机 19
2.2 主体性危机 22
2.3 信任危机 26
3. 全球人工智能教育治理的典型案例 31
3.1 欧盟《人工智能法案》：以风险治理为核心的教育 AI 监管框架 31
3.2 美国：原则导向与地方自治相结合的教育人工智能治理模式 32
3.3 英国：风险规制导向的教育人工智能治理模式 33
3.4 芬兰：教育价值与伦理导向的人工智能治理模式 34
3.5 新加坡：国家统筹型治理模式 35
3.6 日本：审慎创新导向的教育人工智能治理模式 36
3.7 韩国：技术驱动与国家推进相结合的教育人工智能治理模式 38
4. 中国人工智能教育发展与治理的总体特征 40
4.1 中国人工智能教育应用现状 40
4.2 中国人工智能教育治理现状 41
5. 人工智能教育应用与治理的六大原则 43
5.1 坚持教育价值引领，技术服务于人的全面发展 43
5.2 坚守人本内核，明确人机协同定位 44
5.3 以教育公平为底线，缩小数字素养鸿沟 46
5.4 打造适配技术迭代的监管体系 46

5.5 全面提升师生数智素养，培养理性使用技术的能力 48
5.6 推动多元共治，凝聚政府、学校、企业、家庭的治理合力 49
6. 对策与建议 52
结语 55

前 言

2025 年至 2026 年 4 月，人工智能技术特别是生成式人工智能与大模型技术的快速演进，正在推动全球教育体系进入新一轮结构性变革阶段。与此前教育信息化、在线教育、智慧教育等阶段相比，本轮人工智能浪潮最大的不同在于：AI 不再只是教育工具，而正在成为教育活动中的“参与者”、“协作者”甚至“组织者”。

自 2022 年底 ChatGPT 引发全球生成式 AI 热潮以来，人工智能技术在教育领域的应用经历了从“辅助工具”向“教育基础设施”的转变。进入 2025 年后，多模态大模型、推理模型、智能体系统以及长上下文模型逐渐成熟，AI 开始深度嵌入教学、学习、评价、管理与科研全过程。一方面，人工智能显著提升了教育资源供给效率。AI 能够低成本生成课程内容、自动化完成作业批改、提供个性化学习建议，并支持跨语言、跨文化教育传播。这使得教育资源在一定程度上突破了时间与空间限制。另一方面，AI 也对传统教育制度形成了深刻冲击。例如，生成式 AI 对作业、论文和考试体系提出挑战，教师权威结构发生变化，学生学习方式出现明显转型，教育公平问题也呈现新的复杂性。

在全球范围内，美国科技企业仍然在教育 AI 核心技术领域占据主导地位，OpenAI、Google、Microsoft、Anthropic 等公司不断推进教育场景布局；欧洲则更加强调 AI 治理与伦理规范建设；中国在教育数字化基础设施、教育大模型应用和智慧教育平台建设方面快速推进，形成了具有中国特色的发展路径。总体来看，2025—2026 年是人工智能与教育深度融合的重要转折阶段。这一时期不仅体现为技术工具的升级，更意味着教育理念、学习方式、人才培养模式以及教育治理逻辑的系统性重构。

1. 全球人工智能教育发展的主要趋势

1.1 AI 从辅助工具转向教育基础设施

1.1.1 人工智能正在重构教育系统的底层运行逻辑

全球人工智能教育发展正在进入一个具有结构性意义的新阶段。与此前人工智能主要作为教学辅助工具存在不同，生成式人工智能、大模型以及智能体技术的快速发展，使 AI 开始逐渐嵌入教育体系运行的底层结构之中。人工智能在教育领域的角色，正在从工具层走向系统层，从局部应用走向整体嵌入，并逐渐演变为支撑教育运行的重要基础设施。

在早期教育信息化阶段，人工智能更多承担的是功能性角色。例如，智能题库用于提高练习效率，自动批改系统用于降低教师工作量，自适应学习平台用于提升学生学习效率。这一时期的 AI 虽然提高了教育活动的数字化水平，但并未真正改变教育体系本身的组织结构与运行方式。学校仍然是教育活动的中心，教师仍然是知识传递的核心主体，课程体系与教学流程也基本保持工业化时代形成的标准化结构。AI 本质上只是对既有教育体系的局部优化。

然而，随着生成式人工智能能力的突破，AI 开始具备内容生成、逻辑推理、多轮交互、情境理解以及自主协同能力，其在教育领域中的作用开始发生根本变化。AI 已经不再局限于单一教学环节，而是逐渐覆盖课程设计、教学实施、学习支持、评价反馈、教师教研、校园管理以及教育治理等多个层面。由中国教育发展战略学会区域教育专业委员会联合全国 30 余家教育单位共同完成的《2025 年中小学教师人群应用人工智能情况调研报告》显示，人工智能已实现对中小学备课、授课、学练、考试、评价、管理六大核心教学环节全流程覆盖。中小学教师常用人工智能工具矩阵涵盖了 14 大类、90 多个具体工具；相关 APP 月活用户数量呈爆发式增长；中小学教师应用群体中，31-40 岁教师占比近 30%而成为主力，语文、数学等主科及初高中教师占比相对更高。 60.7%的教师基于提升效率目的、56.7%的教师因减轻负担目的使用人工智能工具。应用场景中，以备课为目的应用最为集中，写教案占比 50%、制作 PPT 占比 42%、检索教学资料占比 37.7%，课中、课后的评价与管理环节应用较低。

教育系统对于 AI 的依赖，也开始从效率工具依赖转向系统能力依赖。这一变化首先体现在课程与教学体系的重构上。过去课程内容具有相对稳定性，教材、教学进度以及教学方案通常按照统一标准进行设计，教师则负责在既定框架内完成知识讲授。然而，在 AI 深度参与教育之后，课程内容开始逐渐呈现动态生成特征。AI 能够依据学生能力水平、学习兴趣、知识掌握状态以及学习目标，自动生成差异化学习内容与练习方案。传统意义上统一教材、统一教学、统一评价的标准化模式，正在逐渐被统一目标、差异路径、动态反馈的智能化模式所替代。例如，美国可汗学院（Khan Academy）推出的 AI 学习助手“Khanmigo”，已经能够在数学、科学

和写作教学中提供实时互动辅导。系统不仅能够回答学生问题，还能够依据学生错误类型进行苏格拉底引导式教学，而不是直接提供答案。根据 Khan Academy 2025 年公布的数据，参与 AI 辅助学习的学生在数学学习持续性和课堂参与度方面均出现明显提升。与此同时，美国部分中学开始尝试利用生成式 AI 自动生成个性化阅读材料，使同一课程能够面向不同阅读水平学生提供差异化学习内容。在中国，国家智慧教育平台以及多地智慧校园建设，也正在推动 AI 深度嵌入课堂教学过程。部分地区已经实现 AI 辅助备课、课堂学情分析以及作业自动诊断等功能。AI 系统能够依据学生课堂行为数据和作业表现，自动生成知识掌握画像，并为教师提供教学调整建议。教育活动因此开始从经验驱动逐渐转向数据驱动。

其次，AI 正在改变教育系统的组织逻辑。工业化时代形成的学校体系，本质上是一种标准化、大规模、统一化的教育组织模式。班级授课制、固定课程结构以及统一教学节奏，是其核心特征。这种模式能够在有限资源条件下实现大规模教育普及，但也导致教育难以真正实现个性化。而生成式 AI 与智能体系统的发展，则正在使教育组织方式出现新的可能性。AI 不仅能够为学生提供个性化学习支持，还能够动态协调学习任务、自动安排学习节奏，并形成学习行为数据驱动的反馈闭环。部分 AI 系统已经具备一定程度的自主协同能力，能够依据学生学习状态自动生成学习计划，并在发现学生学习困难时主动干预。这一趋势意味着，AI 开始承担部分传统由教师和学校管理系统完成的组织功能。教育系统因此不再只是使用 AI 工具，而是逐渐形成 AI 原生教育环境。在这种环境中，AI 已经成为学校运行的重要基础能力，其地位越来越类似于互联网、云计算或数据平台。新加坡、韩国、阿联酋等国家甚至已经开始将 AI 学习平台视为国家教育基础设施的一部分。

与此同时，AI 基础设施化还意味着教育体系对于算力、模型和数据的依赖程度显著上升。未来教育系统的运行，不仅需要教师与课程资源，更需要稳定的大模型能力、数据系统以及智能协同平台。教育数字化因此正在进入“AI 原生阶段”。这一阶段与此前“互联网+教育”阶段的重要区别在于，AI 不再只是连接教育资源，而是开始直接参与教育过程本身。这种变化进一步推动学校从数字校园向智能校园转型。传统数字校园主要实现资源在线化和管理信息化，而智能校园则强调 AI 对教育全过程的动态参与。例如，AI 能够自动识别学生课堂状态、生成课堂反馈报告、分析学习风险、动态推荐学习内容，并对长期学习行为进行预测。学校运行逻辑因此越来越依赖实时数据分析与智能决策。从更深层次来看，AI 基础设施化意味着教育体系正在经历一次新的技术范式转型。过去两百年，现代学校制度主要建立在工业化逻辑基础之上，其核心目标是以标准化方式完成大规模人才培养。而 AI 时代的教育体系，则正在逐渐向智能化、个性化以及人机协同方向转变。教育活动不再只是知识传递，而越来越强调能力培养、认知支持以及持续学习能力建设。

因此，AI 从教育工具向教育基础设施的转变，并不仅仅是技术应用范围扩大，而是教育体系底层运行逻辑正在发生深刻变化。这种变化不仅影响教学方式，更影响未来学校组织结构、教师角色、课程体系以及教育治理模式。

1.1.2 AI 开始承担教育协同与系统组织功能

AI 基础设施化的一个重要标志，在于其开始承担传统教育体系中的部分组织与协同功能。过去，学校教育中的教学安排、学习管理、资源协调以及教学反馈，主要依赖教师与管理人员完成。教师不仅承担知识讲授任务，还负责学习进度安排、课堂组织、作业布置、学情分析以及学生沟通等多项工作。学校则通过行政管理体系维持整体教学秩序。然而，随着大模型与智能体技术的发展，AI 已经开始逐渐进入教育组织层，并在部分场景中承担类似教育协同系统的角色。

近年来，多智能体（Multi-Agent）系统在教育领域的发展速度明显加快。与传统 AI 工具只能完成单一任务不同，智能体系统能够进行任务分解、目标规划、自主调用资源以及多任务协同。教育领域中的智能体系统，已经开始具备课程安排、学习任务管理、学习进度跟踪以及协同学习组织等能力。例如，Google DeepMind 与部分教育机构合作开展的 AI 学习助手实验，已经能够依据学生学习目标自动生成学习路径，并根据学生学习表现动态调整内容难度。微软在教育版 Copilot 中，也开始尝试将 AI 嵌入教师工作流，使 AI 不仅能够生成教学内容，还能够帮助教师完成课堂管理与教学反馈。在高等教育领域，AI 开始更多承担科研与学习协同功能。部分高校已经利用 AI 系统完成课程资源管理、论文写作支持以及科研助手服务。例如，美国亚利桑那州立大学（Arizona State University）与 OpenAI 合作开展生成式 AI 校园应用实验，通过 AI 支持课程开发、学生辅导与教学管理。据公开资料显示，该校已有数百门课程接入 AI 辅助系统，用于教学内容生成和学习支持。在中国，教育智能体系统的发展同样十分迅速。部分智慧校园平台已经开始利用 AI 自动生成教学计划、进行课堂行为分析以及开展学情预测。部分地区还尝试利用 AI 系统进行区域教育资源调度，实现跨校课程协同与在线教研。

AI 承担组织功能的意义在于，其正在改变教育系统的运行结构。过去教育体系中的信息流动效率较低，教师很难持续跟踪每个学生的学习状态，学校管理也更多依赖周期性评价。然而 AI 能够实现对学习全过程的实时分析，并形成即时反馈机制。这意味着教育活动开始从周期性管理走向实时性管理。例如，在传统课堂中，教师通常只能通过考试或作业了解学生学习情况，而 AI 系统则能够实时识别学生在学习中的停顿点、错误模式以及知识漏洞。系统不仅能够自动生成学习建议，还能够预测学生学习风险，并在学生出现学习困难之前进行干预。这一变化实际上意味着教育系统开始形成数据驱动的闭环结构。AI 不仅负责内容生成，还能够持续跟踪学习过程、分析学习行为并调整后续教学方案。教学因此不再是单向知识传递，而越来越接近动态协同系统。

与此同时，AI 还正在改变教师之间以及学校之间的协作模式。过去教师教研活动往往具有较强封闭性，不同学校之间课程资源和教学经验共享程度有限。而 AI 平台则能够实现教学资源自动整合与共享，使优质课程、教学案例以及教学策略能够在更大范围内传播。例如，中国国家智慧教育平台目前已经汇聚大量课程资源，并逐渐接入 AI 能力，实现教学资源智能推荐与课堂分析功能。教育资源平台开始从资源仓库转向智能学习平台。

值得关注的是，开源教育 AI 生态的发展进一步强化了 AI 的基础设施属性。其中，OpenClaw 等开源智能体项目的出现，被视为 AI 从封闭工具体系走向开放协同生态的重要标志。与传统商业教育平台不同，OpenClaw 强调模块化、可扩展以及自主协同能力。学校和教师能够依据自身需求构建本地化教育 Agent 系统，并根据具体教学场景进行功能调整。这意味着未来教育 AI 的发展路径，可能并不完全依赖大型商业平台，而是逐渐形成开放式、多元化的教育智能体生态。OpenClaw 的重要意义，并不只是提供某种具体教学功能，而在于其展示了一种新的教育 AI 组织逻辑。AI 不再只是等待用户调用，而是能够主动规划学习流程、组织学习任务并协调多个系统共同完成复杂教育活动。这意味着教育 AI 开始从工具调用模式进入自主协同模式。这一变化可能对未来学校组织结构产生深远影响。在传统教育体系中，学校主要承担课程组织、教学管理以及资源配置功能。而在 AI 深度参与之后，部分组织功能可能逐渐由智能系统承担。未来学校的核心价值，或许将更多体现在价值引导、社会化培养以及创新能力培养等方面。

但与此同时，AI 承担教育组织功能也带来了新的问题。首先是教师角色变化问题。当 AI 能够完成部分教学组织工作之后，教师如何重新定位自身角色，已经成为全球教育领域的重要议题。越来越多研究认为，未来教师的核心价值将从知识讲授转向学习设计、情感支持以及价值引导。其次是平台依赖问题。随着教育活动越来越依赖 AI 平台，学校对于大型科技企业的依赖程度也显著上升。如果教育系统过度集中于少数平台，可能导致教育公共性受到影响。再次是算法治理问题。AI 系统在学习任务安排、学习评价以及内容推荐过程中，实际上已经开始影响学生学习路径与认知结构。如果算法逻辑缺乏透明性，可能进一步加剧教育偏见与教育分层。

因此，AI 承担教育协同与组织功能，虽然显著提升了教育系统运行效率，但也意味着教育治理体系需要同步转型。未来教育竞争，不再只是课程与师资竞争，更是教育智能系统能力与教育治理能力之间的综合竞争。

1.1.3 AI 基础设施化推动全球教育竞争格局重构

人工智能从教育工具演变为教育基础设施，不仅改变了学校内部运行逻辑，也正在重塑全球教育竞争格局。过去，国家教育竞争力更多体现在高等教育水平、科研能力以及人才培养质量等方面，而在 AI 时代，教育体系对于模型能力、数据能力、算力资源以及平台生态的掌控程度，正在逐渐成为新的核心竞争因素。

首先，教育竞争正在从教育资源竞争转向教育智能系统竞争。在传统教育体系中，优质教育资源主要表现为优秀教师、重点学校以及课程体系。而在 AI 时代，教育资源的核心形态开始发生变化。模型能力、数据资源、智能平台以及学习分析能力，正在成为新的关键资源。例如，美国 OpenAI、Google、微软等科技公司，正在持续推进教育版 AI 产品与教育生态建设。OpenAI 推出 ChatGPT Edu，Google 推出 Gemini 教育版与 LearnLM 模型，微软则将 Copilot 深度整合进入教育场景。这些平台不仅提供 AI 工具，更试图构建覆盖学习、教学、评价以及管

理的整体教育生态。根据 HolonIQ 2025 年全球教育科技报告，全球 AI+智慧教育市场规模预计将在 2025 年达到 850 亿美元,并于 2030 年突破 2400 亿美元, 2025 至 2030 年期间的年均复合增长率约为 23%。与此同时，各国政府也开始将教育 AI 基础设施纳入国家战略。韩国政府提出建设“AI 数字教科书”体系，计划逐步实现教材智能化；阿联酋则推动全国学校接入 AI 学习平台，并将 AI 素养纳入基础教育课程。这种趋势意味着，未来教育竞争不仅是学校之间的竞争，更是国家之间教育智能体系能力的竞争。

其次，教育平台生态竞争正在加剧。过去教育科技企业主要围绕课程资源和在线教学展开竞争，而 AI 时代的竞争核心则逐渐转向平台生态能力。大型科技公司希望通过教育 AI 平台掌握学习入口、数据入口以及内容入口，从而形成长期生态优势。例如，OpenAI 正在推动 ChatGPT 与教育插件体系结合，Google 则依托 Workspace 和 Gemini 形成校园生态，微软则通过 Office 与 Copilot 构建学习协同体系。中国企业也在加速布局 AI 学习平台，通过智能学习系统、学习机以及教育大模型构建本土教育生态。教育平台化竞争的重要变化在于，教育产品不再只是单点工具，而逐渐演变为学习操作系统。学生、教师以及学校越来越深地嵌入平台生态之中。平台不仅提供学习功能，还开始掌握学习数据、评价体系以及学习行为分析能力。这种变化带来了明显的数据集中趋势。由于 AI 模型能力高度依赖学习数据，拥有大规模用户的平台能够持续积累学习行为数据，从而进一步提升模型能力，并形成“数据—模型—平台”的正反馈结构。因此，教育数据逐渐成为新的战略资源。谁掌握教育数据，谁就可能在未来教育 AI 生态中占据主导地位。这也是近年来“教育数字主权”问题受到广泛关注的重要原因。

所谓教育数字主权，不仅涉及数据安全，更涉及国家对于人才培养体系与教育价值体系的控制能力。教育数据不仅记录学习行为，还反映学生认知结构、能力特征以及价值偏好。如果教育系统长期依赖外部平台与模型，可能导致教育体系在价值导向和知识结构方面受到深层影响。因此，越来越多国家开始强调本土教育模型与本地化教育平台建设。例如，欧盟近年来持续推动“可信 AI”与教育数据治理体系建设，强调教育数据不能完全由商业平台控制。中国也正在推动国家智慧教育平台以及教育数据体系建设，以提升教育数字基础设施自主能力。与此同时，AI 基础设施化还正在推动教育公共性问题重新受到关注。传统教育体系具有较强公共属性，而 AI 平台则往往具有明显商业属性。当教育活动越来越依赖平台生态之后，教育是否会被平台逻辑主导，已经成为全球教育治理的重要问题。例如，一些研究者担忧，如果未来学生长期依赖商业 AI 学习平台，平台算法可能通过学习推荐机制影响学生认知结构与学习路径，从而改变教育本身的公共价值导向。

此外，AI 基础设施化还可能进一步扩大教育不平等。理论上，AI 能够降低优质教育资源获取门槛，实现大规模个性化教育。然而现实中，高质量模型、算力资源以及教育平台仍然集中于少数国家和大型科技企业。部分发展中国家以及教育资源薄弱地区，可能在 AI 时代面临新的数字鸿沟。联合国教科文组织 2025 年报告指出，目前全球低收入国家中，能够系统性接入生成式 AI 教育平台的学校比例仍然较低。教育 AI 基础设施的不均衡，可能进一步加剧全球教

育差距。因此，AI 从教育工具转向教育基础设施，并不仅仅意味着技术升级，而意味着全球教育体系正在进入新的竞争阶段。这一阶段的核心特征，是教育系统越来越依赖模型、数据、平台与智能协同能力。未来教育竞争的关键，可能不再只是学校数量与师资规模，而是教育智能体系的整体能力。

总体来看，人工智能正在推动全球教育体系发生基础性转型。AI 不再只是服务教育的辅助工具，而正在成为教育运行的重要底层能力。这种变化不仅重构教学方式，也重构学校组织结构、教育资源配置方式以及全球教育竞争格局。未来教育体系，很可能将从工业化时代形成的大规模标准化体系，逐渐走向智能化时代的大规模个性化与人机协同体系。

1.2 教育大模型成为核心竞争方向

1.2.1 通用大模型加速进入教育场景

教育大模型逐渐成为全球人工智能教育竞争的核心方向之一。所谓“教育大模型”，并不仅仅是将通用大模型简单应用于教育场景，而是指围绕教育目标、教育逻辑与学习规律进行专门训练和优化的大模型体系。与传统教育软件相比，教育大模型的核心特点在于其具备跨任务生成、复杂语言理解、多轮交互、知识组织以及学习行为分析能力，其目标已经不仅是提供教学工具，而是形成适用于教育全过程的智能能力体系。

当前，全球教育大模型发展首先呈现出通用模型教育化的明显趋势。OpenAI、Google、Microsoft、Anthropic 等国际科技公司，正在通过教育版产品、教育插件体系以及教育知识库适配，加速推动通用大模型进入学校教育体系。教育因此成为全球大模型竞争中最重要的应用场景之一。其中，OpenAI 在教育领域的布局尤为典型。，OpenAI 持续推动 ChatGPT Edu 在高校中的应用，并与美国、英国、澳大利亚等多个国家高校开展合作。与普通版本相比，ChatGPT Edu 更加重视校园数据安全、学术管理以及教学协同能力。其功能已经从简单问答扩展至课程辅助、科研支持、论文写作、编程训练以及知识组织等多个领域。例如，美国亚利桑那州立大学已在课程开发、科研辅助以及学生学习支持中系统接入 OpenAI 模型。学校公开资料显示，大量教师开始利用 AI 辅助生成教学案例、课堂练习以及课程讨论方案。2024 年 Google 提出“LearnLM”教育模型概念，其目标并不仅是增强知识生成能力，而是强调模型对学习过程与认知逻辑的理解能力。Google 在相关研究中提出，教育 AI 需要具备“引导式教学”、“认知支架”和“过程反馈”等教育特征，而不仅是直接输出答案。这意味着全球科技企业已经开始意识到，教育 AI 与一般内容生成 AI 存在本质差异。微软则通过 Copilot 深度进入学校工作流系统。由于 Office 长期构成全球教育办公体系的重要基础设施，因此微软在教育 AI 领域的重要优势并不只是模型能力，而是其能够直接嵌入教师日常工作流程。Copilot 已经开始在课程设计、课堂材料生成、教学反馈以及学生管理等环节中广泛使用。部分欧美学校甚至开始尝试利用 AI 自动生成教学总结与学生成长档案。通用大模型进入教育领域的重要意义，在于其推动教育 AI 从专用功能系

统走向通用认知系统。传统教育软件通常只能完成某一单一任务，例如批改作业、生成题目或提供在线课程。而大模型则能够同时承担问答、推理、写作、翻译、编程、知识组织以及学习辅导等多重任务。

与此同时，多模态能力的发展进一步加速了教育大模型渗透。2025 年以来，文本、语音、图像以及视频理解能力开始逐渐融合，AI 能够同时识别学生语言表达、课堂互动以及视觉行为。例如，一些 AI Tutor 已经能够通过语音互动实时判断学生学习状态，并依据学生表情与语气调整教学方式。这种变化意味着 AI 开始从文本工具逐渐演变为沉浸式学习伙伴。

然而，通用大模型教育化也带来新的问题。首先是教育适配问题。通用模型虽然具备强大语言生成能力，但其对于课程结构、认知规律以及学习心理的理解仍然有限。例如，一些模型容易直接给出答案，而非引导学生进行思考；部分模型还可能出现知识错误、逻辑混乱或虚构内容等问题。其次是教育价值问题。教育活动并不仅是知识获取过程，更涉及价值引导、人格培养以及社会化发展。而通用大模型本质上仍然主要基于互联网语料训练，其价值体系与教育目标之间并不天然一致。因此，越来越多国家开始强调教育 AI 必须接受教育价值约束。此外，通用模型教育化还加剧了全球教育平台竞争。由于大模型具有明显规模效应，少数头部科技公司正在形成全球教育 AI 生态主导能力。未来教育体系是否会过度依赖少数商业平台，也已经成为国际教育治理的重要议题。

1.2.2 垂类教育大模型快速发展

在通用模型教育化快速推进的同时，另一条重要发展路径则是垂类教育大模型的崛起。与通用模型不同，垂类教育模型更加聚焦特定教育场景，其核心目标并不是形成万能 AI，而是围绕具体教育任务形成专业化能力。这一趋势在中国表现尤为明显。由于中国拥有全球规模最大的教育体系，以及高度结构化的考试体系和丰富的教育数据资源，因此大量企业开始围绕题库、课程、教研以及学习行为数据训练专用教育模型。

近年来，中国教育科技行业的竞争逻辑正在发生明显变化。过去教育企业的核心竞争力主要体现在课程内容、教师资源以及流量获取能力上，而随着生成式 AI 发展，竞争重点开始逐渐转向模型能力、学习行为数据以及学习系统能力。例如，科大讯飞推出“星火教育大模型”，重点布局课堂教学、作文批改、口语训练以及学习分析；网易有道推出“子曰教育大模型”，重点围绕学习辅导与语言教育进行优化；学而思则围绕数学学习、高考训练以及智能学习系统构建垂类模型体系。这些模型的核心特征，在于其更加重视教育知识结构与考试逻辑。例如，在数学学习场景中，教育模型不仅需要给出答案，更需要理解学生错误产生原因，并依据学生认知水平提供分步骤引导。在作文评价中，模型不仅需要完成语言生成，还需要理解评分标准、逻辑结构以及语言表达能力。因此，垂类教育模型的发展逻辑，本质上是“教育经验模型化”。大量长期积累的教学经验、考试逻辑以及教研体系，正在逐渐被编码进入 AI 模型之中。未来

教育 AI 竞争的核心，不再只是算法能力，而是教育知识体系与教学经验能否被有效转化为模型能力。

中国教育体系中的高频学习场景也为教育大模型快速落地提供了条件。例如，作业批改、考试训练、英语口语、数学辅导以及知识点诊断等场景，具有高度标准化和高频使用特点，因此非常适合 AI 规模化部署。艾瑞咨询公司测算 2025 至 2029 年中国 AI 产业将保持 32.1%的年均复合增长率 ，在 2029 年突破 1 万亿的市场规模。生成式 AI 凭借强语言交互特性，精准契合了教育作为沟通密集型行业的本质，使得教育行业迅速跃升为生成式 AI 技术快速渗透并落地的主要战场之一。AI 学习机、AI Tutor 以及智能学习平台，已经成为当前中国教育科技行业增长最快领域之一。

除了企业层面，教育垂类模型还开始向学校系统内部延伸。部分重点中学已经开始尝试利用本地模型构建校本题库、校本教研系统以及学情分析系统。AI 因此开始从外部教育服务逐渐进入学校内部教学体系。值得注意的是，垂类教育模型的发展，也正在推动教育评价体系变化。传统教育评价往往依赖考试结果，而 AI 则能够分析学生学习过程数据，包括思考路径、错误模式以及知识迁移能力。这意味着未来教育评价可能从结果评价逐渐转向过程评价。

此外，AI Tutor 的发展进一步强化了垂类模型的重要性。相比早期智能学习系统，新一代 AI Tutor 已经不只是答题工具，而是能够进行连续互动、启发式提问以及动态学习规划。部分 AI Tutor 甚至已经开始具备情境化教学与长期学习陪伴能力。例如，美国可汗学院推出的 Khanmigo，就强调 AI 作为学习引导者而非答案提供者。系统会通过连续追问帮助学生建立思考过程，而不是直接输出结果。

但与此同时，垂类教育模型的发展也带来新的风险。首先是应试化 AI 问题。由于部分模型高度围绕考试训练优化，可能进一步强化教育功利化倾向，使 AI 更多服务于分数提升，而非能力培养。其次是学习依赖问题。随着 AI Tutor 逐渐具备长期陪伴能力，部分学生可能过度依赖 AI 完成学习任务，从而削弱独立思考能力。再次是教育数据风险。垂类教育模型高度依赖学习行为数据，而这些数据往往涉及未成年人信息、学习轨迹以及认知特征，其敏感程度远高于一般互联网数据。因此，教育模型发展也越来越依赖教育数据治理能力。

总体来看，垂类教育模型的发展，意味着 AI 教育竞争已经从通用能力竞争进一步走向教育理解能力竞争。真正有效的教育模型，不仅需要语言生成能力，更需要理解教育规律、学习心理以及教学目标。

1.2.3 学校私有化模型部署与教育主权意识上升

除了通用模型教育化与垂类教育模型发展之外，近年来另一个重要趋势，则是学校和教育机构开始推动本地化、私有化教育模型部署。这一趋势出现的核心原因，在于教育数据安全、教育治理以及教育主权问题的重要性正在快速上升。随着 AI 深度进入教育体系，教育数据已经不再只是简单的信息记录，而逐渐成为影响人才培养体系与教育价值体系的重要战略资源。

教育数据具有高度敏感性。一方面，其涉及未成年人个人信息、学习行为轨迹以及长期认知发展数据；另一方面，其还涉及学校教学体系、课程资源以及评价机制。一旦教育系统长期依赖外部商业平台，可能导致教育数据控制权逐渐外移。因此，越来越多高校与重点学校开始探索本地化教育模型部署。特别是在科研型大学和重点中学中，学校希望通过本地模型实现教学支持、科研辅助以及校园治理能力建设，同时避免核心数据外流。例如，美国部分高校已经开始部署校内私有化大模型，用于学术写作支持、教学资源管理以及校园知识库建设。欧洲一些高校则更加重视教育数据合规与隐私保护，强调 AI 系统必须符合 GDPR 等数据治理要求。中国高校近年来也开始加速本地模型建设。清华大学、北京大学、浙江大学等高校陆续推动校内 AI 平台建设，部分高校已经实现本地知识库问答、教学辅助以及科研协同功能。

学校私有化模型部署的重要意义，在于教育系统开始重新掌握 AI 主动权。过去学校更多只是被动使用商业教育平台，而本地模型则意味着学校能够根据自身课程体系、教学逻辑以及管理需求进行定制化训练。这一趋势也反映出教育数字主权意识正在快速增强。在 AI 时代，模型不仅影响学习效率，更可能影响学生认知方式与知识结构。如果教育体系长期依赖外部模型，其知识生成逻辑、内容推荐机制以及价值导向，可能对本国教育体系产生长期影响。因此，越来越多国家开始强调教育 AI 本地化能力建设。。

然而，学校私有化模型部署同样面临现实挑战。首先是技术门槛问题。模型训练、算力支持以及系统维护，需要较高技术能力和资金投入，并非所有学校都具备独立建设能力。其次是资源不均衡问题。重点高校和头部学校更容易获得算力与技术资源，而普通学校则可能进一步依赖商业平台。这意味着 AI 时代教育资源不平等问题可能进一步加剧。再次是制度与伦理问题。学校在拥有本地模型之后，也需要面对数据治理、算法透明以及学术规范等问题。例如，AI 生成内容如何界定学术原创性，AI 辅助学习如何纳入教学评价体系，均需要新的制度回应。

总体来看，教育大模型已经成为全球人工智能教育竞争的核心方向。未来教育 AI 竞争的关键，不再只是模型参数规模，而是教育数据、教育经验、课程体系以及教育治理能力之间的综合竞争。谁能够真正理解教育规律，并建立兼顾技术能力与教育价值的 AI 体系，谁就可能在下一阶段全球教育竞争中占据主动地位。

1.3 教育数据成为核心战略资源

1.3.1 教育数据正在从结果记录转向学习过程建模

人工智能深度进入教育体系之后，教育数据的重要性正在被重新定义。传统教育体系中的数据，长期以来主要服务于教育评价与教育管理，其核心内容通常包括考试成绩、课堂表现、作业完成情况、升学结果等。这类数据具有明显的结果导向特征，本质上属于对学习结果的静态记录。无论是学校管理系统还是教育评价体系，过去更关注学生学得怎么样，而较少关注学生是如何学习的。

生成式人工智能和教育大模型的出现，正在改变这一逻辑。教育 AI 系统真正需要的，并不仅仅是最终答案，而是学习过程本身。学生在学习中的思考路径、停顿时间、错误模式、知识调用顺序、反复修改行为以及问题解决方式，开始成为比考试分数更具价值的数据资源。教育数据因此从结果数据逐渐演化为过程数据，其核心价值也从评价功能转向认知建模功能。这种变化意味着，教育系统开始第一次大规模、连续性地记录学生的学习过程。过去，一名学生在课堂上的理解障碍、注意力变化以及知识结构缺陷，很难被完整识别，教师通常只能通过考试成绩进行间接判断。但在 AI 环境下，学生几乎所有学习行为都能够被数字化记录。例如，学生在完成一道数学题时停留在哪一步时间最长、在哪类题型中最容易出现概念混淆、在阅读文本时哪些段落反复查看、在 AI 对话中如何表达自己的困惑，这些都成为可被分析的数据。

教育数据因此开始具有认知轨迹特征。它不再只是对结果的描述，而是在一定程度上反映学生的思维结构与学习机制。对于教育大模型而言，这类数据的价值远高于传统成绩数据。因为 AI 真正需要学习的，不是标准答案，而是人类学习过程中呈现出的认知规律。2025 年以来，大量 AI Tutor 系统已经开始围绕学习过程建模展开优化。系统并不直接给出答案，而是持续追踪学生的思考步骤，并根据学生错误类型动态调整提示方式。OpenAI 与部分高校合作开展的教育实验，也开始强调“推理过程数据”对于模型优化的重要性。这意味着未来教育 AI 竞争的重要方向，已经从谁拥有更多题目转向谁能够更准确理解学习过程。

中国教育 AI 的发展同样体现出这一趋势。过去国内教育数字化积累的大量题库、作业和考试数据，主要属于结果性数据，而 2025 年以后，越来越多教育企业开始建设全过程学习行为数据库。部分平台已经能够持续记录学生的学习节奏、知识掌握变化、错误迁移情况以及学习习惯特征，并据此形成动态学习画像。这种变化意味着教育数据的颗粒度正在不断细化。

教育数据结构的变化，也推动教育评价体系开始发生转型。传统评价主要基于阶段性结果，而 AI 系统更关注学习过程中的能力变化与认知演化。因此，越来越多国家开始讨论过程性评价与持续性评价的制度建设。例如，新加坡教育部已经开始在部分学校探索 AI 辅助学习分析系统，用于长期追踪学生能力成长过程，而不仅依赖统一考试结果。欧盟部分教育研究机构则提出，未来教育评价体系可能逐步从结果认证转向能力轨迹分析。这一变化的深层意义在于，教育数据正在成为理解“人类如何学习”的关键基础设施。过去教育学研究长期依赖小规模实验和有限样本，而 AI 时代的大规模学习行为数据，正在使教育研究进入数据驱动认知科学阶段。教育数据的战略价值，因此已经不仅属于教育技术范畴，而是开始影响未来教育理论、人才培养模式以及国家教育竞争力。

1.3.2 教育数据竞争正在重塑全球人工智能教育格局

教育数据战略价值的提升，正在改变全球人工智能教育发展的竞争结构。过去教育竞争更多表现为师资、课程、资金与制度之间的竞争，而在 AI 时代，数据资源开始成为决定教育模

型能力的重要基础。谁能够掌握更大规模、更高质量、更持续更新的教育数据，谁就更可能在教育 AI 生态中占据主导地位。

教育数据与普通互联网数据存在明显区别。互联网平台中的用户行为数据，更多反映消费偏好、社交关系与内容兴趣，而教育数据则具有高度结构化和因果化特征。例如，学生解决数学问题时的错误顺序、概念迁移能力以及推理方式，本质上反映的是认知结构。相比一般文本数据，这类数据更具深层认知价值。因此，教育数据不仅是 AI 训练资源，更是未来认知模型的重要基础。这一特点使拥有大规模教育体系的国家开始获得新的战略优势。中国、印度、美国等人口规模较大的国家，在长期教育实践中积累了海量学习行为数据。这些数据覆盖基础教育、高等教育、职业教育以及在线学习平台，形成了庞大的教育数据资源池。尤其是在中国，高频考试体系和在线教育普及，使结构化学习数据积累速度明显快于许多国家。

教育数据战略价值的提升，还推动国家层面开始重新审视教育数字主权问题。教育数据不同于普通商业数据，其背后涉及人才培养规律、国家知识体系以及文化价值结构。一旦核心教育数据长期被跨国平台掌握，教育体系可能在知识组织与能力评价层面形成外部依赖。因此，越来越多国家开始强调教育数据本地化、教育模型自主可控以及公共教育平台建设。联合国教科文组织在 2023 年发布的《生成式人工智能教育与研究应用指南》中提出，教育数据不应完全由商业平台控制。欧盟在《人工智能法案》中，也将涉及未成年人的教育数据列为高敏感数据。中国近年来持续推进国家智慧教育平台建设，其背后同样体现出对教育数字基础设施自主性的重视。教育数据因此开始超越技术问题，逐渐成为国家教育战略的重要组成部分。

1.3.3 教育数据治理成为人工智能教育发展的核心议题

教育数据战略价值的上升，也使数据治理问题迅速突出。人工智能教育的发展越深入，教育系统对数据的依赖程度就越高，而数据治理能力则越来越成为决定教育 AI 能否健康发展的关键条件。

教育数据具有极强敏感性。它不仅包含学生身份信息，还涉及学习行为、心理状态、认知特征以及成长轨迹。与普通互联网数据相比，教育数据更加接近个体长期认知结构。一旦数据泄露或被滥用，其影响远超一般商业隐私问题。尤其是在未成年人教育场景中，数据治理问题更加复杂。AI 系统为了实现精准学习分析，需要持续收集学生学习行为数据，包括课堂互动、语音交流、学习节奏、注意力变化等信息。这意味着教育系统正在形成高度数据化环境。虽然这种环境能够显著提高个性化教育水平，但也可能使学生长期处于“被监测”状态。部分国家已经开始关注数据化教育带来的伦理风险。美国一些教育研究机构提出，过度学习监测可能影响学生自主性与心理安全感。欧洲部分学者则担忧，AI 系统长期记录学生行为，可能形成算法标签化问题，即学生早期学习特征被长期固化，进而影响后续教育机会。

教育数据治理还涉及数据解释权问题。AI 系统会基于数据分析对学生能力进行判断，但这种判断是否真正科学、是否存在算法偏见、是否会强化教育分层，已经成为新的教育治理问题。

例如，如果 AI 系统长期依据历史数据判断某类学生学习能力较弱，那么系统可能不断降低其学习难度，从而形成低预期循环。这一问题已经在部分 AI 推荐系统中出现。研究发现，当算法过度依赖历史行为时，学生容易被固定在特定学习路径之中，个体发展可能受到限制。因此，教育 AI 不仅需要技术准确性，更需要教育价值导向。

教育数据治理还面临平台垄断风险。大型教育科技平台通过长期积累数据形成规模优势，中小学校和教育机构则逐渐失去数据自主权。一旦教育系统过度依赖少数平台，未来教育评价标准、知识组织方式乃至学习路径设计，都可能受到平台逻辑影响。因此，越来越多国家开始推动教育数据公共治理机制建设。欧盟强调建立“可信教育 AI”框架，日本提出“教育数据公共利用原则”，中国则持续推进教育数字化治理规范建设。全球范围内，一个基本共识正在形成：教育数据不能完全按照商业逻辑运行，而需要兼顾教育公平、公共利益与未成年人保护。

未来教育 AI 的发展，将越来越依赖制度能力而不仅是技术能力。教育数据的真正价值，并不只是支撑模型训练，而在于是否能够在保障教育公共性的前提下，推动教育体系更加公平、高效与个性化发展。如何在数据开放与数据保护之间建立平衡，如何在技术效率与教育伦理之间形成协调，已经成为全球人工智能教育治理的核心问题。

2. 人工智能教育应用的三大危机

2.1 技术中心主义危机

技术中心主义（technocentrism）是指将技术本身视为目的、以技术逻辑凌驾于教育逻辑之上的思维与实践倾向。人工智能进入教育体系的初始阶段，通常被赋予 解决教育难题"的技术想象。然而，当技术逻辑逐渐取代教育逻辑成为主导时，危机便由此产生。技术中心主义危机集中表现为三个相互关联的维度：评价标准异化、数据迷恋与复杂系统简化。

2.1.1. 评价标准异化

评价标准异化的直接表现，是将是否使用 AI 而非是否解决教育问题作为评价标准，导致为用技术而用技术。例如强制要求课程融入智能技术，把技术应用率本身当作教育改革的成果指标。要理解这一异化如何发生，必须回到人工智能改变知识传播结构这一深层背景。

现代学校教育体系建立在知识稀缺基础之上。在工业化时代，知识传播渠道相对有限，教师、教材和课堂构成了学生获取系统知识的核心途径。学校之所以能够成为现代社会最重要的人才培养机构，很大程度上是因为其掌握了知识组织与知识传播的主导权。教师不仅承担知识讲授功能，也在一定程度上承担知识筛选、知识解释和知识权威确认功能。长期以来，学生的学习过程本质上是围绕获取知识展开的。

生成式人工智能的发展正在改变这一长期稳定的知识传播结构。大模型具备强大的自然语言生成与知识整合能力，学生能够通过自然语言交互即时获得概念解释、案例分析、学习建议、论文总结、语言翻译以及问题解答。与传统搜索引擎不同，生成式 AI 不仅能够提供信息检索结果，还能够根据学生认知水平与问题背景生成适配性的表达方式。这意味着知识获取成本被大幅降低，知识传播方式开始从有限供给转向即时生成。这一变化对教育体系的影响是结构性的。

过去，学校课堂之所以重要，一个重要原因在于课堂是学生获取系统知识的主要场所。而在生成式 AI 广泛普及之后，学生越来越可以在课堂之外获得大量学习支持。无论是数学公式推导、语言学习、程序编写还是历史背景分析，AI 都能够在极短时间内提供较为完整的解释与辅助。这意味着课堂不再天然拥有知识供给优势。知识传播权威结构因此开始发生变化。长期以来，教师和教材构成知识权威中心，学生主要通过接受式学习获取标准化知识内容。然而，在 AI 时代，学生获取信息的渠道明显扩展，知识来源也更加多元。学生不仅能够接触课堂知识，还能够通过 AI 快速接触跨学科内容、国际案例和实时信息。这种变化使传统课堂中的单向知识传递模式受到冲击。

这种深刻的结构性变化，正是评价标准异化的温床。当一个技术系统被证明能够如此彻底地改变知识获取方式时，教育体系在应激反应中，极易将是否引入人工智能、是否使用智能工具误认为教育改革是否成功的标志。于是，评价的参照系从教育问题是否得到解决悄然滑向技术是否被使用。这种异化具有多重表现形式：在课程层面，智能技术融入度被列为课程评价的硬性指标，教师被迫在课程中嵌入 AI 元素，至于这些技术是否真正提升了教学质量反而退居其次；在学校层面，智慧课堂覆盖率、AI 应用频次成为办学水平的展示性指标；在教师评价层面，是否使用 AI 备课、是否在课堂中引入智能工具成为考核观测点。技术从解决问题的手段，异化为教育评价的对象本身。

这种评价标准的异化，其深层危害在于它颠倒了目的与手段的关系。教育的目的是人的发展，是学生认知能力、价值观念与社会能力的养成；技术只是服务于这一目的的手段。当是否使用 AI 成为评价标准，教育实践便开始围绕技术展示而非教育实效组织，技术应用本身被建构为教育进步的象征性符号。其后果是双重的：一方面，形式化的技术融入挤占了真正有效的教学设计空间，教师为了满足技术指标而表演性地使用 AI；另一方面，那些确实存在教育问题、但尚未找到合适技术方案的领域，反而因为无法展示技术成果而被边缘化。评价标准异化由此构成了技术中心主义危机的第一个入口，它让教育系统在尚未想清楚技术为何而用之前，就已经陷入了必须使用技术的压力之中。

2.1.2 数据迷恋

数据迷恋是指教育体系过度采集可量化行为数据，如注视时长、点击率、停留时间、答题正确率，却忽视无法量化的深层学习过程，如顿悟、批判性质疑与价值判断的形成。数据迷恋并非凭空产生，它根植于学习过程数据化、个性化的技术演进之中。

现代教育体系形成于工业化时代，其核心逻辑之一，是通过标准化方式实现大规模教育普及。统一课程、统一教材、统一教学进度和统一评价体系，构成了现代学校运行的基本框架。这种模式在工业化时期具有重要历史意义，它使国家能够以较低成本实现大规模人才培养，并推动基础教育快速普及。然而，标准化教育体系长期存在一个难以解决的问题，即难以充分适应学生个体差异。学生在认知能力、学习兴趣、知识基础、学习节奏以及思维方式方面存在明显差异，但传统课堂通常只能按照统一节奏推进教学。部分学生可能因为教学进度过快而难以理解知识，另一部分学生则可能因为内容重复而缺乏学习动力。工业化教育体系虽然提高了教育效率，却难以真正实现因材施教。

人工智能的发展，使大规模个性化学习开始具备现实可能。AI 系统能够持续分析学生学习行为数据，包括知识掌握情况、错误模式、学习节奏、注意力变化以及学习偏好等信息，并基于这些数据动态调整学习内容与学习路径。学习过程因此逐渐从固定化、线性化结构转向动态化、实时化结构。与传统在线学习平台相比，新一代 AI 学习系统最大的变化，在于其具备持续适应能力。过去的教育技术更多是资源分发工具，而 AI 系统则能够根据学生状态不断调

整教学策略。例如，当系统发现学生在某一知识点反复出错时，可以自动降低难度、补充基础知识或调整讲解方式；当学生表现出较高掌握能力时，则可以自动提高学习难度或扩展学习内容。这种动态适应能力，使学习过程开始摆脱传统统一进度模式。未来教育可能形成统一培养目标 与个性化学习路径并行的新结构。也就是说，学生仍然需要完成共同核心课程，但完成路径、学习节奏以及学习方式可能逐渐差异化。

这一变化意味着教育体系正在从以班级为中心转向以学习者为中心。传统学校组织方式强调集体教学，而 AI 支持下的学习系统则越来越强调个体学习体验。学生不再只是被动接受统一教学内容，而能够依据自身能力与兴趣形成差异化学习过程。值得注意的是，个性化学习不仅意味着学习节奏差异，更意味着学习结构本身的变化。AI 系统能够帮助学生形成更加灵活的学习路径。例如，对于语言能力较强但数学基础较弱的学生，系统可以自动强化数学训练；对于实践能力突出但理论理解较弱的学生，则能够增加案例化和情境化学习内容。学习组织方式因此开始从课程驱动逐渐转向数据驱动。过去教育体系主要依据统一课程结构组织教学，而 AI 时代的学习过程越来越依赖学习数据实时反馈。这意味着教育系统正在形成新的运行逻辑：教学不再只是依据固定教学计划推进，而是依据学生学习状态动态调整。这一趋势已经在全球范围内逐渐显现。美国部分高校开始使用 AI 学习分析系统持续跟踪学生学习状态，并依据学生学习风险自动生成干预方案。新加坡、日本等国家也在推动 AI 支持下的个性化学习实验，希望通过智能系统提高学生学习适应性与自主学习能力。中国教育领域同样出现明显变化，大量 AI 学习平台已经开始提供动态练习推荐、智能错题分析和个性化学习路径功能，一些学校正在探索 AI 辅助分层教学，通过 AI 系统实现学生学习状态实时分析，从而提高课堂差异化教学能力。

个性化学习的实现高度依赖行为数据的采集与计算，但可量化数据与真实学习之间存在着根本性的落差。教育 AI 系统能够精确记录的是学生的外在行为痕迹，比如注视时长、停留时间、答题正确率、错误模式、点击路径、学习时长。这些数据之所以被采集，是因为它们可以被测量、被计算、被反馈。然而，真正构成教育核心价值的深层学习过程，恰恰在数据可及的范围之外：学生是否在某个瞬间产生 顿悟式的理解跃迁，是否对既有结论提出批判性质疑，是否形成独立的价值观判断，是否在学习过程中产生深刻的情感体验与意义感。这些都无法被量化，也难以被分解为可计算的行为特征。

数据迷恋的实质，是可测量性对教育性的僭越。当教育评价、教学决策与学习反馈越来越依赖数据时，教育系统会不自觉地倾向于关注那些可被数据捕捉的行为，而忽略那些不可见但更重要的学习形态。教师被系统反馈引导去关注学生错了几道题、停留了多久，而不是去追问学生是否真正理解了、是否产生了自己的疑问。学生也被暗示：被系统记录的才是被认可的，未被记录的认知过程则无关紧要。学习的丰富性、复杂性与个体性，在数据化的过程中被还原为若干可计算的维度。值得警惕的是，教育评价向过程导向的转型本意是对结果导向的纠偏。越来越强调学生在学习过程中的思考、修正与合作，但若过程导向评价同样陷入数据化的逻辑，

以更精细的行为记录替代对深层认知过程的观察，那么它非但不能克服数据迷恋，反而可能以更隐蔽、更彻底的方式将学习全面数据化。

2.1.3 复杂系统简化

复杂系统简化，是指将复杂教育问题简化为技术优化问题，忽略情感、价值观等人本要素。教育是人类社会最复杂的系统之一，它同时涉及认知发展、情感体验、社会关系、价值观念、文化传统与制度环境等多重维度，各要素之间相互交织、动态演化。而技术中心主义的典型思维，是将这一复杂系统化约为信息传递效率或算法可优化路径问题。

这种简化首先体现在对教学过程的还原上。技术中心主义倾向于将教学理解为信息从教师（或教材、AI）向学生的高效传递，将学习理解为可分解、可度量、可优化的行为序列。在这一还原逻辑下，教学质量的提升被等同于信息传递效率的提升，学习效果的改善被等同于算法推荐路径的最优化。个性化学习系统正是这一简化逻辑的集中体现：系统将学习目标分解为知识点，将学生学习状态还原为可计算的特征向量，再通过算法不断调整难度、推荐内容、优化路径。这套技术逻辑在解决知识掌握类问题时确实显示出效率优势，但教育的复杂性远非知识传递所能涵盖。

当教育问题被化约为技术优化问题时，那些无法被算法表征的要素便系统性地被排除在视野之外。情感支持与心理关怀，例如学生面对挫折时的情绪波动、学习动机的起落、师生之间的信任与依恋，难以进入技术优化的计算框架。价值观念的引导，包括学生如何看待知识、如何理解自我与他人、如何形成社会责任意识，更无法被路径算法所捕获。技术优化逻辑越是精密，人本要素在教育中的位置就越是被边缘化。更严重的是，这种简化还可能主动扭曲教育行为本身：教育实践开始围绕算法能够优化的指标来组织，而不再围绕人之为人所需要的教育来组织。

复杂系统简化还加剧了教育分层风险。若 AI 依据历史数据不断降低某些学生的学习难度，可能形成低预期循环，影响学生长期发展；算法推荐可能倾向于强化符合主流评价标准的学习路径，而忽视不同文化背景与学习风格差异。技术优化逻辑表面上追求最优解，但由于其对教育复杂性的简化，其所谓的最优解往往只是单一维度上的最优，即知识掌握效率的最优，而非完整人的发展意义上的最优。当这一简化后的逻辑被上升为教育治理的总体逻辑时，技术中心主义便完成了从方法到价值观的全面渗透。

2.2 主体性危机

技术中心主义危机的深层后果，是教育中人的主体性危机。当技术逻辑成为教育的主导逻辑，学生与教师作为教育活动主体的地位便开始动摇。主体性危机表现为三个层面：学生认知外包导致高阶认知能力流失，教师角色认同危机导致专业身份焦虑，权力结构失衡导致人类主体沦为技术规训的产物。

2.2.1 学生认知外包

学生认知外包，是指过度依赖 AI 而丧失自主思考能力，导致批判性思维、创造力等高阶认知能力下降。认知外包（cognitive offloading）本是认知科学中的一个中性概念，指个体将记忆、计算等认知负荷转移给外部工具。然而，当 AI 能够即时生成解释、答案、方案与文本时，认知外包从偶发性的工具使用，演变为结构性的学习方式，其风险也随之凸显。

知识获取方式的即时化，为认知外包提供了技术前提。在生成式 AI 广泛普及的背景下，学生能够通过自然语言交互即时获得概念解释、案例分析、学习建议、论文总结、语言翻译以及问题解答，且 AI 能够根据学生认知水平生成适配性的表达方式。过去，由于知识获取成本较高，教育体系更加重视知识记忆与知识积累，学生需要通过长期训练掌握大量基础知识。而在 AI 能够即时提供信息支持的背景下，单纯依赖记忆的学习模式正在弱化，学生能够快速获得知识解释与参考内容。真正重要的问题开始变成：学生是否能够理解知识之间的关系，是否能够判断信息的真实性，是否能够在复杂情境中运用知识解决问题。

当 AI 几乎无条件地提供即时答案时，学生面临的最省力的选择便是跳过思考、直接取用结果。提问—获得答案的即时反馈循环，逐渐取代了从探索到试错再到反思的完整认知过程。论证的过程、推理的环节、反复修正的体验，恰恰是高阶认知能力得以养成的土壤；而认知外包恰恰将这些环节整体移交给了 AI。长期依赖 AI 生成答案，是否会削弱独立思考能力？如果 AI 持续替代阅读、写作与推理活动，是否会影响深层认知能力的形成？这些问题目前虽仍缺乏明确结论，但已经引发全球教育界的广泛讨论。

人工智能的发展更深层地改变了教育体系对于能力的理解。工业化时代的人才培养逻辑，长期建立在知识积累与标准化能力训练基础之上，学校教育强调知识掌握、标准答案、重复训练以及统一评价。而 AI 已经能够快速完成大量基于规则和标准化逻辑的任务，包括文本写作、代码生成、资料整理、信息总结以及部分分析工作。这意味着，传统依赖重复训练形成的部分能力，其教育价值正在下降；单纯依赖知识复现与标准化输出的能力，其竞争优势正在减弱。未来教育越来越强调高阶认知能力，批判性思维、复杂问题解决能力、创造力、跨学科整合能力、人机协同能力以及社会情感能力的重要性正在持续上升。

但必须清醒地看到这一转型的另一面：高阶认知能力的重要性上升，并不意味着这些能力会自动获得。恰恰相反，高阶认知能力从来不是天赋的自然显现，而是长期训练与深度参与的结果。批判性思维需要大量的论证实践，创造力需要大量的试错与发散，复杂问题解决能力需要真实情境中的反复磨砺。如果学生的认知过程被大规模外包给 AI，那么这些能力便失去了养成的机会，不是不再被需要"，而是 "得不到训练而萎缩"。知识能力结构转型为教育提出了更高的能力目标，而认知外包却在根本上瓦解着达成这些目标的训练基础，这正是当前教育面临的深层矛盾之一。未来教育需要培养能够理解知识逻辑、跨领域整合知识、在人机协同中有效使用知识的人，而这一切的前提，是学生首先保有独立、主动、深度的思考习惯。

2.2.2 教师角色认同危机

教师角色认同危机，是指教师的角色从知识传授者被弱化为技术操作者，教师群体由此产生职业焦虑与专业身份的不确定感。要理解这一危机，必须回到教师劳动结构的系统性变化。

人工智能进入教育领域之后，最先受到冲击的并不是课程内容本身，而是教师的劳动结构。长期以来，教师职业建立在知识讲授、课堂组织、作业批改、考试评价的线性工作链条之上，其中大量工作具有重复性、程序性与事务性特征。工业化时代的大规模学校教育，本质上是一种依赖教师高强度劳动维持运行的组织体系。教师不仅承担知识教学任务，还需要完成备课、批改、统计、记录、管理、沟通等大量非教学性工作。随着生成式人工智能、大模型以及智能教育系统的发展，这一长期稳定的教师劳动结构正在发生系统性变化。当前，人工智能已经能够在多个教学环节中替代传统重复性劳动。例如，在作业批改领域，AI 系统能够快速完成客观题批改、作文初步评分、语法识别以及知识点错误分析；在教学准备阶段，AI 可以依据课程目标自动生成教学大纲、课件、案例材料与课堂练习；在课堂管理环节，AI 能够实时记录课堂互动数据、生成课堂摘要并分析学生参与情况；在学情分析方面，AI 系统还能够根据学生作业、测试与课堂表现生成学习画像与能力评估报告。

这种变化首先带来了教学效率的显著提升。传统教学中，大量教师时间被消耗在事务性劳动之中，而 AI 介入之后，这些工作开始逐渐自动化。频繁使用 AI 的教师群体工作满意度提升 34%（EdWeek Research Center, 2024）；盖洛普调查显示，每周使用 AI 工具的教师平均节省 5.9 小时，相当于每年释放超过 200 小时（按 36 周教学计算）。这意味着教师能够将更多精力投入课堂互动、学生指导以及课程设计等更具创造性的工作之中。但人工智能对教师体系的真正影响，并不仅仅是减轻工作量或提高效率，更深层的变化在于教师职业功能正在发生重新定义。

过去，教师在教育体系中的核心权威，很大程度上建立在知识占有优势基础之上。在知识稀缺时代，教师与教材是学生获取系统知识的主要来源，课堂讲授构成学校教育的核心结构。然而，生成式 AI 的发展正在改变这一基础条件。学生越来越能够通过 AI 系统即时获取知识解释、案例分析、学习建议甚至完整解题过程。知识获取不再高度依赖教师讲授，教师作为知识传递者的传统功能因此开始弱化。与此同时，课堂权威结构也在发生变化。在传统课堂中，教师通常是知识的唯一权威，而 AI 时代的信息来源更加多元化，学生可以即时通过 AI 获取不同解释与观点，教师不再是学生唯一依赖的知识来源。

正是在这一背景下，教师角色认同危机得以生成。危机首先表现为 "去专业化" 的风险。当教师的核心功能 —— 知识传递 —— 被 AI 大规模承接之后，如果教师无法在 "组织学习" 的新定位中及时建立新的专业身份，那么其角色便可能被窄化为技术操作者：操作 AI 系统、解读学习数据、执行算法推荐的教学方案。在这一窄化中，教师失去了对教学过程的自主判断与专业决策，沦为智能教学流水线上的执行环节。其次，危机表现为专业自主性的压缩。当教

学决策越来越依赖数据与算法 —— 学什么、按什么节奏学、给什么反馈 —— 教师在教学中的决策空间不断收窄，其作为专业人员的判断力被系统性地边缘化。教师从教学的设计者与决策者，退化为技术系统的辅助者与执行者。

教师群体由此产生普遍的职业焦虑。英国教育部调查显示，64% 的教师担忧人工智能技术会削减教育领域现有工作岗位，48% 的教师认为自身教育角色未来可能被人工智能取代。教师群体产生职业焦虑，这并非简单的 "饭碗之忧"，而是一种深刻的角色认同危机：当教师的传统身份根基被动摇，而新的专业身份又尚未建立时，教师群体陷入了 "我是谁、我的价值何在" 的认同困境。值得注意的是，AI 并不会削弱教师的重要性，反而可能进一步凸显教师在人类教育中的不可替代性 —— 教育不仅是知识传递过程，更是人格塑造、价值引导与社会关系建构过程，AI 能够生成知识内容，却难以真正替代人与人之间的情感交流、价值认同与社会化互动。然而，这一判断的成立，恰恰取决于教师能否在 "知识传递者" 的角色被消解之后，顺利实现向 "学习组织者"" 价值引导者 " 的转型。转型过程中的身份悬置与专业焦虑，正是当前教师角色认同危机的真实形态。

2.2.3 权力结构失衡

权力结构失衡，是指师生将知识主权、教学决策权与评价权让渡于 AI 算法，人类主体沦为技术规训的产物。这是主体性危机中最深层的形态，它关涉教育权力在教育系统内部的重新配置。

学校组织形态的变革为权力让渡提供了结构前提。人工智能对教育系统的影响，并不仅仅局限于课堂层面，更正在深刻改变学校整体运行逻辑。现代学校制度形成于工业化时代，其基本结构建立在固定时间、固定空间和统一流程基础之上。学生按照年龄进入班级，在统一时间进入教室，依据统一课程体系接受教育。这种制度设计的核心目标，是实现大规模、标准化的人才培养。然而，人工智能的发展正在逐渐削弱传统学校结构形成的基础条件。因为当知识获取越来越数字化、学习支持越来越智能化之后，学习活动开始突破学校原有的时间与空间边界。首先，学校作为唯一知识中心的地位正在被削弱。过去，学校之所以在教育体系中占据绝对核心地位，很大程度上是因为其拥有知识资源、教师资源与教学组织资源。但 AI 时代的知识获取渠道正在高度多元化，学生能够通过在线平台、AI Tutor、数字学习社区以及全球开放课程获取学习资源，学校不再垄断知识供给。学校运行方式也开始逐渐平台化，AI 系统能够承担部分传统学校管理职能，例如学习任务安排、学习进度跟踪、资源推荐、学情分析以及教学协调。学校管理正在从传统行政管理模式，逐渐转向数据驱动与智能协同模式。班级制度同样可能受到冲击，AI 时代的个性化学习趋势，使 "同一班级、同一进度" 的组织逻辑逐渐受到挑战。

这些结构性变化，为权力向算法的转移铺平了道路。权力结构的失衡表现为三重让渡：

其一，知识主权的让渡。在传统教育体系中，知识的选择、筛选、组织与解释权掌握在教师、教材与学校手中，这种知识主权构成了教育机构权威的根基。而在 AI 时代，学生获得的

知识内容越来越由算法推荐决定：学习什么、接触什么解释、看到什么案例，取决于模型的训练数据与推荐逻辑。知识的筛选与解释权，正悄然从教育者转移至算法系统。

其二，教学决策权的让渡。教学不再只是依据固定教学计划推进，而是依据学生学习状态动态调整；学习内容、学习节奏、学习路径越来越多地由系统依据数据自动决定。教师与学生作为教学活动的共同决策者，其决策空间被算法持续压缩。谁决定学习推荐逻辑，谁就在事实上拥有了教学决策权。

其三，评价权的让渡。学生的评价越来越多地依据系统生成的学情分析、能力评估与数据画像，评价标准由算法预设，评价结果由系统生成。教师与学校在评价中的裁量权，被系统的数据输出所取代。

三重让渡的最终后果，是人类主体沦为技术规训的产物。学生在教育过程中被持续监测 —— 学习记录、行为习惯乃至情绪变化都被纳入数据系统；被持续分类 —— 依据算法评估被标记为不同能力水平、不同风险等级；被持续预期管理 —— 低预期循环使学生被锁定在算法预判的发展轨道上。这种持续的数据化监视、分类与预期塑造，构成了一种新型的规训机制：它不再通过外在的纪律与惩罚，而是通过看似中性的数据反馈与个性化推荐，将学生塑造为算法所预期的学习者。教师同样处于这一规训结构之中，其教学行为被数据反馈持续校准，其专业判断被系统建议持续规训。更深层的问题由此浮现：谁定义模型训练内容？谁决定学习推荐逻辑？谁掌握教育数据解释权？当这些权力从教育共同体让渡给算法及其背后的技术主体时，教育的 "人" 便不再是自己教育生活的主人。

主体性危机的实质，是教育中 "人" 的主体地位被技术主体所替代。学生丧失的是独立思考与自我发展的主体性，教师丧失的是专业判断与职业认同的主体性，师生共同丧失的，是对教育活动本身的支配权。

2.3 信任危机

技术中心主义与主体性危机的叠加，最终侵蚀了教育体系的信任基础。信任是现代教育体系运行的社会前提，它至少包含三个维度：师生之间的信任、公众对技术系统的信任、社会对正规教育体系的信任。当人工智能深度介入教育，这三重信任正在同时受损，形成系统性的信任危机。

2.3.1 学术诚信

学术诚信危机，首先表现为全球范围内与 AI 生成内容相关的学术不端行为激增。人工智能特别是生成式人工智能的发展，正在对现代教育评价体系形成深层冲击。长期以来，学校教育的评价逻辑建立在一个默认前提之上，即学生提交的作业、论文、考试答案以及研究成果，

能够在较大程度上反映学生自身真实能力。因此，无论是日常作业、课堂测试，还是标准化考试与学位论文，其核心功能都在于通过结果判断学生知识掌握程度与能力水平。

然而，生成式 AI 的快速发展正在削弱这一评价前提。当前的大模型已经能够高质量完成作文写作、论文生成、代码开发、阅读总结、语言翻译以及部分研究性任务，部分 AI 系统甚至已经能够根据不同年龄阶段与写作风格生成高度拟真的学习成果。在这种背景下，教育系统越来越难以仅凭最终成果判断学生真实能力。学生提交的文本、代码、报告乃至研究方案，可能既包含个人思考，也包含 AI 生成内容，两者之间的边界正在变得模糊。这种变化首先冲击了传统作业体系。过去，家庭作业和课后论文的重要功能之一，是帮助教师了解学生知识掌握情况与独立思考能力。但在 AI 广泛介入学习过程之后，许多传统作业已经难以准确反映学生真实水平。美国多所高校在 2023 年至 2025 年间已经开始公开讨论作业失效问题。一些教师发现，学生提交作业的整体语言质量与结构完整性明显提高，但课堂讨论与现场表达能力却并未同步提升。

学术不端行为的激增，正在以触目惊心的数据呈现出来。以英国谢菲尔德大学为例，其校内与人工智能相关的学术不端案例，在短短一年内从 6 起飙升至 92 起，整体增长 15 倍，增长趋势触目惊心。这一案例并非孤例，而是全球范围内 AI 相关学术不端蔓延的缩影。学术诚信体系因此面临新的挑战。传统教育中的作弊通常指直接抄袭、代写或考试违规行为，其判断边界相对明确。但 AI 时代的情况更加复杂：学生使用 AI 辅助润色语言是否属于合理使用？使用 AI 生成论文框架是否构成学术不端？AI 参与研究设计与数据整理应如何界定署名权？这些问题正在成为全球高校广泛争论的新议题。目前，全球范围内大量高校已经开始调整 AI 使用规范。例如，部分大学允许学生在论文写作中使用 AI 进行语言优化，但要求明确标注 AI 参与情况；部分学校则要求学生提交写作过程记录，以证明个人思考过程的真实性。2023 年以后，哈佛大学、牛津大学、斯坦福大学等高校陆续发布生成式 AI 教学使用指南，强调 "AI 辅助不等于 AI 替代"，并要求学生在学术活动中保持独立判断能力。

更深层的危机，发生在师生之间的信任关系之中。学术不端行为的激增，使师生之间经历严重的信任危机：教师无法确定学生提交的成果究竟反映其真实能力还是 AI 的生成能力，因而倾向于以怀疑的目光审视每一份作业；学生则感到自身被系统性不信任，即便其成果确为独立完成，也难以摆脱 "是否用了 AI" 的追问。信任一旦动摇，教学关系便从 "共同求真" 的合作关系，退化为 "相互防范" 的监督关系。教师不得不在教学之外投入大量精力用于 AI 检测与真伪核验，学生则在自我证明的压力下失去求知的从容。这种师生信任的损伤，是 AI 对教育关系最深刻的侵蚀之一。传统考试制度同样受到冲击。过去，标准化考试被视为相对公平和客观的人才评价机制，但 AI 的发展使大量基于知识复现和标准答案的考试价值下降。部分国家已经开始重新讨论考试制度改革问题，一些高校增加现场口头答辩、开放性问题分析以及项目展示等评价方式，以降低 AI 直接替代的可能性。更深层的问题在于，"原创性" 的概念正在

发生变化，教育评价开始从成果真实性问题进一步上升为能力真实性问题。教育系统必须重新建立新的能力判断机制，以适应 AI 深度参与知识生产的新环境。

2.3.2 算法与隐私信任

算法与隐私信任危机，是指算法的不透明和数据隐私泄露风险引发公众对技术系统的信任担忧。教育 AI 系统的运行，建立在大规模数据采集的基础之上，而这一基础本身正在制造新的信任风险。

教育数据治理问题在 AI 深度进入教育体系之后变得空前突出。教育 AI 系统通常需要持续收集学生学习行为数据，包括学习记录、课堂互动、行为习惯甚至情绪变化数据。这些数据高度敏感，一旦泄露或滥用，可能带来严重风险。近年来，欧洲与美国已经开始加强教育数据监管，欧盟《人工智能法案》以及《通用数据保护条例》（GDPR）均对未成年人数据使用提出严格要求。未成年人数据安全，已成为 AI 教育应用中最突出的治理议题之一 —— 未成年人的学习轨迹、行为习惯乃至情绪状态数据，既是最具分析价值的数据，也是最需要保护的数据。

算法的不透明性，构成信任危机的另一重来源。教育 AI 系统的评价与推荐逻辑，对于师生与家长而言往往是难以理解的黑箱：系统为何推荐这些内容、为何给出这样的评估、为何将学生判定为某一水平，其决策依据隐藏于复杂的模型参数之中，难以被审查与质疑。这种不透明性，使教育过程中的关键判断失去了可解释性与可问责性。当教师无法向学生和家长解释评价的由来，当家长无法理解孩子为何被系统如此 "画像"，对技术系统的信任便难以建立。算法偏见问题也正在受到广泛关注。AI 系统的运行依赖训练数据，而数据本身可能包含社会偏见。如果教育模型训练数据结构不合理，AI 可能在评价、推荐与学习分析过程中强化既有不平等。例如，部分算法可能更倾向于推荐符合主流评价标准的学习路径，而忽视不同文化背景与学习风格差异。

公众对 AI 的矛盾心态，进一步揭示了技术信任危机的深度。"用而不信" 现象显著：虽然 72% 的人接受使用 AI，但超过半数（54%）认为 AI 带来了网络安全、隐私泄露及错误信息等显著风险，导致 "用而不信" 的矛盾心态。这一数据表明，公众并非拒绝使用技术，而是对技术系统的安全性与可靠性缺乏信任。教育领域的算法与隐私信任危机，正是这一普遍心态在教育场景中的具体化：人们愿意借助 AI 改善学习，却担心自己的数据被滥用、担心算法偏见影响孩子的发展、担心教育评价被不可解释的黑箱所操控。当这种不信任累积到一定程度，教育技术系统的正当性便会受到根本性质疑。

2.3.3 学历信号价值稀释

正规教育体系的价值下滑，是指当社会无法确定一份优秀成绩单究竟反映学生的真实能力还是其使用 AI 工具的熟练程度时，学历的 "信号价值" 被稀释，教育体系的社会公信力随之动摇。

学历制度的功能，本质上是一种社会信号机制。在信息不对称的劳动力市场中，雇主无法直接观察求职者的真实能力，只能依赖学历、成绩单等教育信号进行判断。而这一信号之所以有效，前提是社会对教育评价体系存在基本信任 —— 相信成绩单上的分数与证书，确实反映了持有者的真实能力。人工智能的介入，正在从根本上侵蚀这一信任前提。当学生提交的作业、论文与考试成果可能部分或全部由 AI 生成时，社会无法再确信 "优秀" 意味着 "能力优秀" 而非 "AI 运用熟练"。学历的 "信号价值" 由此被稀释：一份成绩单所承载的信息含量下降，其作为能力证明的可信度降低。当教育评价无法有效区分真实能力与工具使用熟练度时，正规教育体系最核心的社会功能 —— 能力认证 —— 便面临失效风险。

正规教育体系的价值下滑，还体现在教育公平层面的信任危机。人工智能既可能扩大优质教育资源覆盖范围，也可能形成新的结构性差距。AI 确实为教育资源普及提供了新的可能性 —— 学生可以通过智能学习系统获得即时辅导、学习建议与个性化支持，小规模学校也能够共享高质量课程内容，部分发展中国家已经开始利用 AI 改善基础教育资源不足问题。但另一方面，AI 资源本身并不是均匀分布的，高质量模型、算力资源、数据平台与先进教育系统，往往集中于大型科技企业、发达国家与优质学校。不同地区之间仍然存在明显数字基础设施差异，部分地区网络条件不足、数字设备匮乏，学生难以稳定使用 AI 学习系统，全球仍有大量低收入地区学生缺乏稳定互联网接入条件。不同家庭之间的 AI 使用能力差距正在扩大，高收入家庭更容易接触先进 AI 工具与数字素养培训。学校之间的差距同样可能被进一步放大，优质学校能够率先部署 AI 系统、建设智能校园并开展教师培训，而资源有限学校则可能难以跟进。这意味着，人工智能可能形成 "强者更强" 的马太效应，使教育公平这一社会信任的基石面临新的侵蚀。

与此同时，教育体系内部的信心也在动摇。教师职业焦虑 ——64% 的教师担忧 AI 削减工作岗位、48% 的教师认为自身角色可能被 AI 取代 —— 不仅是个体层面的职业忧患，更是教育体系自我信任弱化的表征。当教育者自身都对教育体系的未来价值产生怀疑时，社会对正规教育体系的信任自然随之衰减。教师群体的职业焦虑与学历信号价值的稀释相互叠加，共同构成了正规教育体系价值下滑的双重推力。

更深层的问题在于，AI 开始影响教育价值体系本身。因为 AI 不仅影响学习方式，也可能影响知识选择、评价标准与价值判断。谁定义模型训练内容？谁决定学习推荐逻辑？谁掌握教育数据解释权？这些问题正在逐渐上升为教育治理中的核心问题。传统教育治理主要围绕考试作弊与抄袭展开，而 AI 时代的 "合理辅助" 与 "过度依赖" 之间边界越来越模糊，教育制度需要重新建立适应 AI 时代的新规范体系。未来教育竞争不仅是技术能力竞争，更是治理能力竞争。人工智能的发展正在迫使教育体系建立新的制度框架，包括数据治理制度、AI 伦理规范、教育公平保障机制以及算法透明制度。只有在技术创新与制度建设之间形成平衡，AI 才能真正服务于教育公共利益，而不是加剧新的结构性不平等。

信任危机的实质，是教育体系的信任基础在三个维度同时受损：师生信任被学术不端与真伪存疑所侵蚀，技术信任被算法黑箱与数据风险所削弱，制度信任被学历信号贬值与价值动摇所瓦解。当三重信任同时动摇，教育体系不仅面临技术适应问题，更面临自身社会正当性的重建问题。

3. 全球人工智能教育治理的典型案例

3.1 欧盟《人工智能法案》：以风险治理为核心的教育 AI 监管框架

欧盟《人工智能法案》（EU Artificial Intelligence Act）是全球范围内最早尝试系统性规范人工智能发展的综合性法律框架之一，也是当前国际社会在人工智能治理领域最具代表性的制度实践。该法案并不仅仅针对教育领域，但由于教育系统涉及未成年人保护、能力评价、人才选拔与社会流动等高度敏感问题，因此教育 AI 被明确纳入重点监管范围。欧盟的治理思路，并不是单纯鼓励技术扩张，而是在推动人工智能应用的同时，通过制度设计防止技术对社会公平、个人权利与公共价值造成结构性冲击。从全球教育 AI 治理趋势来看，欧盟模式具有明显的“风险治理导向”特征，其核心目标是在技术创新与公共安全之间建立制度平衡。

长期以来，欧洲国家对于教育公平、个人隐私与公共治理具有较强制度传统。在人工智能快速进入教育体系后，欧盟首先关注的并不是 AI 能否提升教学效率，而是 AI 是否可能对教育公平、学生权利与社会结构造成潜在影响。因此，《人工智能法案》对教育 AI 采取了较为严格的分类监管方式。根据法案规定，涉及学生录取、学习评价、考试评分、能力评估以及教育资源分配等功能的人工智能系统，被列为“高风险 AI 系统”。这意味着教育 AI 不再被简单视为普通技术工具，而是被视为可能影响个体发展机会与社会公平的重要公共系统。

欧盟之所以将教育 AI 纳入高风险领域，一个重要原因在于教育本身具有高度社会影响力。教育评价不仅关系学生当前学习结果，更可能影响升学机会、职业发展以及未来社会流动路径。如果 AI 系统在评价过程中存在偏见、误判或不透明问题，可能进一步强化既有社会不平等结构。例如，如果某类算法长期基于特定群体数据训练，就可能在能力预测、学习推荐与风险识别中形成隐性歧视，从而影响部分学生的发展机会。因此，欧盟强调教育 AI 必须具备透明性、可解释性与可追责性。

根据《人工智能法案》的要求，所有高风险教育 AI 系统在投入使用之前，都需要完成严格的风险评估程序。开发者必须明确说明模型训练数据来源、算法逻辑、风险控制措施以及人工干预机制。教育机构在使用 AI 系统时，也需要向学生和教师说明 AI 参与范围，避免出现黑箱决策问题。这种制度设计体现了欧盟在教育 AI 治理中的核心原则，即人工智能可以辅助教育，但不能完全替代人的教育判断。

在数据治理方面，欧盟同样采取了高度谨慎的态度。欧洲长期强调个人隐私保护，《通用数据保护条例》（GDPR）已经成为全球最具影响力的数据保护制度之一。在教育领域，由于涉及大量未成年人学习行为数据，欧盟更加强调数据最小化原则与数据使用边界。教育平台和 AI 系统不能无限制收集学生数据，也不能在未经授权情况下将学习行为数据用于商业模型训练。

学生与家长对于个人数据拥有更高知情权与删除权。这种制度设计，本质上是在防止教育平台形成对学生数据的长期垄断。

值得注意的是，欧盟并不反对人工智能进入教育体系。相反，欧盟认为 AI 在个性化学习、学习支持与教育资源共享方面具有重要潜力。例如，AI 能够帮助教师减轻事务性负担，能够为学习困难学生提供个性化支持，也能够帮助教育系统提升资源配置效率。因此，欧盟的治理逻辑并不是限制 AI 发展，而是在规则框架下推动 AI 发展。

在实践层面，欧洲多国已经开始依据《人工智能法案》调整本国教育 AI 政策。例如，法国、德国与荷兰等国家已经要求部分教育平台提高算法透明度，并强化学生数据使用规范。一些欧洲高校在使用生成式 AI 时，也逐渐形成有限开放、过程监管的政策模式。学校通常允许学生在资料整理、语言优化与辅助学习中使用 AI，但要求学生明确标注 AI 使用情况，并强调最终学习成果仍需体现个人思考过程。

欧盟模式对于全球教育 AI 治理具有重要影响。首先，它推动国际社会开始认识到，教育 AI 不仅是技术工具，更是影响社会公平与人才流动的重要制度力量。其次，欧盟强调“高风险 AI”概念，使教育系统成为人工智能治理中的核心领域之一。再次，欧盟推动形成了“算法透明”、“数据保护”、“人工监督”等治理原则，为全球教育 AI 监管提供了制度参考。不过，欧盟模式也面临一定争议。部分科技企业认为，过于严格的监管可能降低 AI 创新速度，提高企业研发成本。一些教育机构则担忧，复杂的合规程序可能影响学校快速引入新技术。此外，由于生成式 AI 技术迭代速度极快，传统法律框架能否长期适应 AI 发展变化，也仍然存在不确定性。

总体来看，欧盟《人工智能法案》代表了一种强调公共治理与风险控制的教育 AI 发展路径。其核心特点并不是技术优先，而是强调技术发展必须服从教育公平、学生权益与社会公共价值。这种治理模式表明，在人工智能深度进入教育体系之后，未来教育竞争不仅是模型能力竞争，也将是制度治理能力竞争。

3.2 美国：原则导向与地方自治相结合的教育人工智能治理模式

美国在教育人工智能治理方面并未形成统一的国家监管体系，而是呈现出联邦政府提供原则性指导、州政府和地方学区自主实施的治理格局。这种治理模式与美国长期形成的教育管理传统密切相关。美国宪法并未赋予联邦政府直接管理基础教育的权力，因此教育治理长期以州和地方学区为核心主体。在人工智能进入教育领域后，这一制度特征同样影响了教育人工智能治理框架的形成。

2023 年，美国教育部（U.S. Department of Education）发布《Artificial Intelligence and the Future of Teaching and Learning》报告[1]，这是美国联邦层面首个系统讨论教育人工智能问题的重要政

[1] U.S. Department of Education. *Artificial Intelligence and the Future of Teaching and Learning* (2023).

策文件。报告认为，人工智能可能深刻改变教学、学习和教育管理方式，但同时也带来公平性、隐私保护、算法偏见和责任归属等问题。因此，人工智能的应用必须坚持“以人为中心”（Human-Centered）的基本原则，确保技术服务于教育目标而非取代教育主体。美国教育部特别强调教师在教育过程中的核心作用，指出人工智能应被定位为支持教师专业实践的辅助工具，而不应替代教师的专业判断。

在治理实践层面，美国教育人工智能治理主要围绕数据保护、算法公平、教育公平和学术诚信四个方面展开[2]。首先，在数据治理方面，美国学校普遍受到《家庭教育权利和隐私法案》（Family Educational Rights and Privacy Act，FERPA）和《儿童在线隐私保护法》（Children's Online Privacy Protection Act，COPPA）等法律约束。教育机构在引入人工智能工具时，需要确保学生数据收集、存储和使用符合相关法律规定。其次，在算法治理方面，美国教育部多次强调应警惕人工智能系统可能存在的种族、性别和社会经济地位偏见，要求教育机构在采购和使用人工智能产品时进行公平性评估。再次，在教育公平方面，政策制定者担忧人工智能可能扩大不同地区、不同群体之间的数字鸿沟，因此强调保障弱势群体平等获取人工智能教育资源的机会。最后，在学术诚信领域，美国高校和基础教育机构陆续制定生成式人工智能使用规范，对学生作业、考试和科研活动中的人工智能使用边界进行界定。

值得注意的是，美国教育人工智能治理呈现出明显的地方创新特征。例如，纽约曾在 2023 年率先限制学校网络访问 ChatGPT，随后又调整政策并推动负责任使用；加州、华盛顿州等地区则制定了面向教师和学生的生成式人工智能应用指南。这些实践反映出美国教育治理体系中实验性治理的特点，即通过地方层面的探索逐步形成治理经验。

总体而言，美国案例代表了一种典型的原则导向型治理模式。其优势在于能够为技术创新保留较大空间，避免过早形成僵化监管框架，有利于人工智能教育应用的快速发展。然而，由于缺乏全国统一标准，不同州和学区之间的治理能力和资源配置水平存在明显差异，可能导致教育人工智能应用质量参差不齐。此外，随着大型生成式人工智能系统逐渐成为教育基础设施的一部分，如何在保持创新活力的同时建立更加系统化的风险治理机制，已经成为美国教育人工智能治理面临的重要课题。

3.3 英国：风险规制导向的教育人工智能治理模式

与美国强调创新优先的治理路径不同，英国在教育人工智能治理中更加强调风险管理和制度规范。尤其是在生成式人工智能快速普及之后，英国政府较早建立了面向学校和教育技术企业的专项治理框架，逐步形成了覆盖教育机构、技术开发者和监管部门的多层次治理体系。

[2] U.S. Department of Education. *AI Guidance for Schools and Educators* (2024–2025).

英国教育人工智能治理的政策基础来源于政府整体人工智能治理战略以及教育数字化转型政策。2023 年以来，英国教育部（Department for Education）持续更新《Generative Artificial Intelligence in Education》指导文件[3]，对学校、教师和教育管理者使用生成式人工智能提出明确要求。政府认为，人工智能有助于减轻教师行政负担、提高个性化教学水平和优化学习支持服务，但其潜在风险同样不容忽视，包括错误信息传播、隐私泄露、算法歧视以及学生过度依赖人工智能等问题。

在治理框架设计上，英国特别强调“安全、透明和责任”的治理原则。学校在引入人工智能工具时，应确保相关产品符合数据保护法规要求，特别是《英国通用数据保护条例》（UK GDPR）和《数据保护法案》（Data Protection Act 2018）的规定。同时，学校需要向教师和学生明确说明人工智能系统的功能边界和潜在风险，避免将人工智能输出结果视为绝对可靠的信息来源。

英国教育人工智能治理的一个重要特点是将监管范围扩展至教育技术产品开发环节。2025 年，英国教育部进一步发布《Generative AI Product Safety Expectations》[4]，明确要求教育人工智能产品开发商在产品设计过程中开展风险评估、内容安全测试和儿童保护审查。这意味着教育人工智能治理不再仅关注学校如何使用人工智能，而是开始关注技术产品本身是否具备进入教育场景的资格和条件。这种治理思路体现出源头治理的特点。

在教育实践层面，英国高度重视学术诚信问题。随着生成式人工智能能够完成论文写作、文本生成和代码编写等任务，学校和考试机构开始重新审视传统评价体系。英国资格认证与考试监管办公室（Ofqual）明确指出，人工智能生成内容不能替代学生真实学习成果，并要求学校完善作业设计和考试管理机制，以维护评价公平性。

从学术视角看，英国案例代表了一种风险规制型治理模式。其突出特点是通过制度规范和产品监管降低人工智能应用风险，在一定程度上实现了创新与安全之间的平衡。与此同时，英国经验也反映出一个重要趋势，即教育人工智能治理正在从“使用者治理”逐渐转向“技术生态治理”，监管对象不仅包括教师和学生，也包括人工智能开发企业和教育技术平台。这一治理思路对于未来全球教育人工智能监管体系建设具有重要参考价值。

3.4 芬兰：教育价值与伦理导向的人工智能治理模式

芬兰长期被视为全球教育改革和教育公平的重要代表国家。在教育人工智能治理领域，芬兰同样展现出鲜明的价值导向特征。与强调监管规则或技术创新的国家相比，芬兰更关注人工智能对教育价值、教师专业自主权和学习者发展的影响，因此形成了一种以教育伦理为核心的治理模式。

[3] UK Department for Education. *Generative Artificial Intelligence in Education* (2025).
[4] UK Department for Education. *Generative AI Product Safety Expectations* (2025).

芬兰教育与文化部以及芬兰国家教育署（Finnish National Agency for Education）认为，人工智能治理的根本目标不是控制技术本身，而是确保技术应用符合教育目的。因此，芬兰在推进人工智能教育应用过程中始终坚持“教育优先于技术”的理念。这一理念与芬兰长期坚持的以学生发展为中心的教育哲学高度一致。

在制度建设方面，芬兰国家教育署发布了面向学校和教育机构的人工智能应用指导文件（AI Guidelines）[5]，系统阐述人工智能在教学、学习和教育管理中的合理应用原则。文件强调，人工智能不应替代教师在评价、反馈和教育决策中的专业职责；教师应具备理解人工智能工作机制和识别算法局限性的能力；学生应理解人工智能生成内容的局限性和潜在偏见。

此外，芬兰积极推动人工智能伦理教育[6]。政府认为，未来公民不仅需要掌握人工智能使用技能，更需要具备理解人工智能社会影响的能力。因此，人工智能教育不仅包括技术知识，还包括算法透明性、数据隐私、公平正义和民主参与等内容。人工智能治理因此被纳入公民教育和数字素养教育体系之中。

综上，芬兰案例代表了一种典型的伦理导向型治理模式，核心特征在于将教育价值置于技术逻辑之上，通过强化教师专业自主性和学生数字公民能力来实现治理目标。相比以监管为中心的治理路径，芬兰模式更加关注教育本质问题，强调技术与教育价值之间的协调发展。这一经验表明，教育人工智能治理不仅是风险管理问题，更是教育哲学和教育价值选择问题。

3.5 新加坡：国家统筹型治理模式

作为全球数字化教育发展水平较高的国家之一，新加坡较早将人工智能纳入国家教育现代化战略，并逐步形成了由国家主导、多部门协同、伦理规范与技术应用并行推进的教育人工智能治理体系。与欧美国家较多依赖指导性文件和学校自主实践不同，新加坡的教育人工智能治理具有明显的国家统筹特征，其治理逻辑建立在国家人工智能战略（Singapore National AI Strategy 2.0）与教育科技总体规划（EdTech Masterplan 2030）的双重框架之上[7]。新加坡教育部（Ministry of Education, MOE）将人工智能视为促进个性化学习和提升教师专业能力的重要工具，但同时强调人工智能必须服务于教育目标，而非替代教师在教育过程中的核心地位。

在治理体系设计上，新加坡最具代表性的创新是构建了专门面向教育领域的“人工智能教育伦理框架”（AI-in-Education Ethics Framework，简称 AIEd Framework）[8]。该框架针对教育场景中的特殊风险进行了制度化设计。框架明确提出四项核心原则，即主体性（Agency）、包容性（Inclusivity）、公平性（Fairness）和安全性（Safety）。其中，主体性原则强调教师和学生应

[5] Finnish National Agency for Education. *AI Guidelines for Education* (2025).
[6] European Commission. *Ethical Guidelines on the Use of Artificial Intelligence and Data in Teaching and Learning for Educators* (2022).
[7] Singapore Ministry of Education. *Artificial Intelligence in Education*. 2025.
[8] Singapore Ministry of Education. *AI-in-Education Ethics Framework*. 2025.

始终保有重要教育决策的控制权；包容性原则要求人工智能系统能够兼顾不同背景和能力水平学生的学习需求；公平性原则关注算法偏见和教育机会平等；安全性原则则要求防范错误信息传播、隐私泄露和不当内容生成等风险。通过这一框架，新加坡实际上将人工智能治理从单纯的技术监管扩展到教育价值和教育伦理层面。

在具体治理实践方面，新加坡采取了平台化治理路径。教育部并未鼓励学校自由采购和部署各类生成式人工智能产品，而是依托国家统一建设的“学生学习空间”（Student Learning Space, SLS）平台集中提供人工智能服务。通过统一平台部署，政府能够对算法模型、数据使用、内容安全以及用户权限进行统一监管，降低教育机构分散采购带来的风险。目前，SLS 平台已集成多种人工智能功能，包括教师备课助手（Authoring Copilot）、数据分析助手（Data Assistant）、数学反馈助手（Feedback Assistant-Mathematics）、短答题反馈助手（SAFA）以及个性化学习系统（Adaptive Learning System）等。这些工具主要承担辅助教学、学习分析和个性化反馈功能，而最终的教学决策仍由教师完成。平台化治理模式不仅提高了教育资源配置效率，也增强了人工智能应用的可控性和透明度。

与此同时，新加坡高度重视人工智能素养教育。教育部提出学生应实现“学习人工智能（Learn about AI）、使用人工智能（Learn to use AI）、借助人工智能学习（Learn with AI）以及超越人工智能学习（Learn beyond AI）”四个层面的能力培养。这种理念体现出新加坡对未来人才培养的系统思考：一方面帮助学生掌握人工智能工具，另一方面培养批判性思维、创造力、价值判断和社会责任等难以被人工智能替代的能力。人工智能治理因此不仅被视为风险控制问题，也被视为未来教育转型的重要组成部分。

从学术角度看，新加坡案例代表了一种典型的国家统筹型治理模式。其优势在于能够通过统一平台和统一伦理框架降低学校层面的治理成本，实现技术创新与风险防控之间的平衡。同时，国家主导机制也有助于缩小学校之间的数字鸿沟，提高教育资源分配的公平性。然而，高度集中的治理结构可能在一定程度上限制基层学校的自主创新空间，并使人工智能应用路径更多依赖政府主导的技术生态。因此，新加坡经验表明，教育人工智能治理不仅是技术问题，更是国家能力、教育价值和治理体系现代化水平的综合体现。对于正在推进教育数字化转型的国家而言，新加坡提供了一个从国家战略、伦理规范到平台治理相互衔接的完整治理范式。

3.6 日本：审慎创新导向的教育人工智能治理模式

与美国强调市场创新、英国强调风险规制不同，日本在教育人工智能治理中体现出明显的“审慎创新（Responsible Innovation）”特征，即在鼓励技术创新和教育数字化转型的同时，通过试点探索、伦理规范和教师主导机制控制潜在风险。这种治理思路既反映了日本长期以来渐进式政策改革的特点，也体现了其在人工智能国家战略中强调“以人为中心”（Human-Centered）

的治理理念。日本教育人工智能治理并非孤立存在，而是嵌入国家数字化转型战略和人工智能治理框架之中，与“Society 5.0”建设目标密切相关。

日本教育人工智能治理的重要转折点出现在生成式人工智能迅速普及之后。2023 年，面对 ChatGPT 等大型语言模型进入学校场景所带来的机遇与挑战，日本文部科学省（Ministry of Education, Culture, Sports, Science and Technology, MEXT）发布《关于学校使用生成式人工智能的暂行指引》（暂定的なガイドライン）[9]，这是日本首个专门针对教育领域生成式人工智能应用的官方指导文件。该文件明确指出，生成式人工智能既可能成为促进学习创新的重要工具，也可能带来错误信息传播、学术诚信受损、隐私泄露以及学生思维能力弱化等问题。因此，学校应坚持谨慎试点原则，在充分认识风险的基础上探索合理应用路径。

在治理框架设计方面，日本采取了较为典型的风险分级与场景治理策略[[10],[11]]。文部科学省并未全面禁止学校使用生成式人工智能，而是根据不同教育场景设定差异化管理要求。例如，在资料检索、创意启发、语言学习和教学辅助等场景中，学校被允许探索人工智能应用；而在考试评价、独立论文写作和涉及学生个人信息处理等高风险场景中，则要求严格限制或加强监督。这种治理方式体现出日本政策制定者试图在教育创新与风险防控之间寻求平衡的思路。

数据安全和隐私保护是日本教育人工智能治理的重要内容[12]。日本政府特别强调学生和教师不得随意向公开生成式人工智能系统输入个人信息、未公开教育数据以及学校内部管理信息。学校在引入人工智能工具时，应审查服务提供商的数据处理机制，确保符合日本《个人信息保护法》（Act on the Protection of Personal Information）的相关要求。由于生成式人工智能往往依赖云端服务运行，日本教育主管部门尤其关注跨境数据流动可能带来的隐私风险，并建议学校优先采用经过安全审查的服务平台。

与此同时，日本高度重视教师在人工智能治理中的主体地位。文部科学省明确提出，人工智能生成内容不能直接替代教师的教学判断和评价职责。教师不仅需要掌握人工智能工具的基本使用方法，还应具备识别错误信息、分析算法偏见以及指导学生合理使用人工智能的能力。因此，日本近年来持续推进教师数字能力建设，将人工智能素养纳入教师专业发展体系之中。政府认为，教师不仅是人工智能工具的使用者，更是人工智能教育治理的重要实施主体。

在学生培养方面，日本将人工智能治理与数字公民教育紧密结合。文部科学省强调，未来教育不仅要培养学生使用人工智能的能力，更要培养其理解人工智能局限性、辨别信息真伪以

[9] Ministry of Education, Culture, Sports, Science and Technology (MEXT). *Guidelines on the Use of Generative AI in Schools*

[10] Government of Japan. *AI Strategy*（2019）

[11] G7 Hiroshima AI Process Comprehensive Policy Framework（2023）

[12] Ministry of Education, Culture, Sports, Science and Technology (MEXT). *Guidelines on the Use of Generative AI in Schools*

及进行独立思考的能力。生成式人工智能的普及使传统的信息素养教育面临新的挑战，因此学校需要帮助学生形成批判性思维和伦理判断能力，以避免对人工智能产生盲目信任或过度依赖。

整体上日本案例体现出一种典型的审慎创新型治理模式。其特点在于不将人工智能简单视为风险对象或技术工具，而是通过试点实践、动态评估和渐进式政策调整不断优化治理体系。这种模式有助于降低技术快速扩散带来的不确定性，同时避免因过度监管而抑制教育创新。

3.7 韩国：技术驱动与国家推进相结合的教育人工智能治理模式

韩国是全球推动教育人工智能应用最积极的国家之一。近年来，韩国政府将人工智能视为国家数字竞争力的重要组成部分，并将教育领域作为人工智能应用的重要试验场。韩国教育人工智能治理表现出鲜明的技术驱动特征，即通过国家主导的大规模教育数字化改革推动人工智能深度嵌入教学体系。在这一过程中，政府不仅关注人工智能治理问题，更试图通过人工智能重塑教育体系和人才培养模式。

韩国教育人工智能治理的发展与其国家人工智能战略密切相关。2019 年，韩国政府发布《国家人工智能战略》（National Strategy for Artificial Intelligence）[13]，提出建设“人工智能强国”的发展目标。教育被视为实现这一目标的重要基础领域。此后，韩国教育部陆续推出人工智能教育推广计划、数字教育创新方案以及人工智能数字教材（AI Digital Textbook，简称 AIDT）建设工程[14]，逐步形成覆盖课程体系、教材开发、教师培训和学习评价的教育人工智能治理体系。

韩国教育人工智能治理最具代表性的实践是 AI 数字教材项目[15]。该项目旨在利用人工智能技术构建个性化学习环境，通过学习数据分析、智能诊断和自适应学习路径设计，为不同学生提供差异化支持。根据韩国教育部规划，人工智能数字教材首先在数学、英语和信息技术等学科中推广，并逐步扩展至其他课程领域。人工智能系统能够根据学生学习过程中的表现数据动态调整学习内容和难度，从而提高教学效率和学习效果。

为了支撑人工智能数字教材的运行，韩国建立了较为完整的教育数据治理体系。学生在学习过程中的行为数据、作业完成情况和学习进度信息被汇总至统一平台，用于生成个性化学习分析报告。教师可以通过数据仪表盘实时掌握学生学习状态，并根据系统反馈调整教学策略。从治理角度看，这种模式体现出明显的数据驱动特征，即通过大规模教育数据收集和分析实现教育决策优化。

[13] Ministry of Education, Republic of Korea. *AI Digital Textbook Promotion Plan*（2023—2025）。

[14] Government of the Republic of Korea. *National Strategy for Artificial Intelligence*（2019）。

[15] Korea Education and Research Information Service (KERIS). *Digital Education Transformation Reports*（2023—2025）。

与此同时，韩国政府高度重视教师人工智能能力建设。教育部设立专项培训计划，推动教师掌握人工智能工具使用方法、数据分析技能以及数字化教学设计能力。政府认为，人工智能不能取代教师，而应帮助教师从重复性工作中解放出来，将更多精力投入学生指导和教育创新。因此，教师培训被视为人工智能教育改革成功实施的重要前提。

然而，韩国教育人工智能治理在实践过程中也面临较大争议。首先，部分教师和教育研究者担忧人工智能数字教材可能加重教师工作负担，因为教师需要学习新的技术系统并承担额外的数据管理任务。其次，家长群体对学生长期使用数字设备可能带来的注意力问题和身心健康影响表示担忧。再次，大规模教育数据采集引发了关于隐私保护和数据安全的讨论。此外，生成式人工智能在教育场景中的应用还涉及算法准确性、信息真实性以及教育公平等问题。围绕这些争议，韩国国会和教育界近年来持续开展政策讨论，并对人工智能数字教材的推广节奏和应用范围进行调整。

表 1 不同国家人工智能教育治理模式比较

国家	治理模式	核心特点	代表措施
美国	指导性治理	地方自主、原则监管	教育部 AI 指导框架
英国	规则治理	产品监管与学校治理并重	教育 AI 产品安全标准
芬兰	伦理治理	教师自主、教育价值优先	全国 AI 伦理指南
新加坡	国家统筹治理	统一平台、伦理框架	AIEd Ethics Framework
日本	审慎创新治理	试点先行、风险控制	学校生成式 AI 指引
韩国	国家推进治理	AI 数字教材改革	AI Digital Textbook

4. 中国人工智能教育发展与治理的总体特征

4.1 中国人工智能教育应用现状

4.1.1 政策驱动特征明显

中国人工智能教育的发展具有鲜明的政策驱动特征。与主要依赖企业创新、市场竞争和高校科研推动人工智能教育应用的部分国家不同，我国人工智能与教育的融合是在教育数字化和教育强国建设的整体战略框架下逐步展开的。人工智能并非作为孤立的技术产品进入教育，而是伴随着国家教育数字化战略持续推进，并逐渐从早期的信息化辅助工具发展为推动教育理念、教学模式和教育治理变革的重要技术力量。2025 年教育部等九部门印发的《关于加快推进教育数字化的意见》明确提出促进人工智能助力教育变革，并将人工智能与教育评价、师生数字素养、教育治理等领域的改革相结合。2026 年印发的《人工智能+教育》行动计划进一步提出推动智能技术与教育全要素融合、全过程贯通和全场景覆盖，表明人工智能已经从教育数字化的重要技术方向进一步上升为教育改革的重要驱动力。

这种政策驱动模式具有较强的系统推进能力。国家层面负责确定战略方向和建设公共数字基础设施，地方教育行政部门结合区域条件开展试点，学校和教育企业进一步将技术转化为具体的教学和管理应用，由此形成从国家战略到地方实践、再到学校场景的纵向推进机制。近年来，各地围绕人工智能课程、智能教学助手、智慧校园、学习分析等开展了大量探索，使人工智能教育逐步从局部试验走向规模化应用。与此同时，国家智慧教育平台持续升级，已经从单纯的数字资源平台向综合性的智能教育公共服务平台转变。2026 年教育数字化战略行动部署进一步提出推动人工智能融入教育全要素、全过程、全场景。

4.1.2 教育场景规模优势突出

中国拥有世界上规模庞大、层次完整且区域差异显著的教育体系，这构成了人工智能教育发展的重要场景优势。从学前教育、基础教育到高等教育、职业教育和继续教育，不同类型的教育场景能够为人工智能提供丰富而持续的应用需求。对于教育人工智能而言，这种优势并不仅仅意味着庞大的用户规模，更意味着技术可以在不同学段、不同学科和不同学习群体中持续接受真实场景检验，并据此不断调整和优化。

这一场景优势已经在国家智慧教育平台的规模化应用中得到体现。截至 2025 年底，国家智慧教育公共服务平台用户总量已经突破 1.78 亿，覆盖 200 多个国家和地区，并汇集 13 万余条中小学优质资源、1.25 万余门职业教育精品课程和 14.5 万门高等教育优质课程。 这种大规模数字教育基础不仅扩大了人工智能的应用空间，也为教育资源共享、智能学习支持和教育公共服务提供了较为坚实的基础。

更值得关注的是，中国教育场景具有明显的复杂性和多样性。不同地区在教育资源、数字基础设施、教师能力和学生学习环境方面存在差异，不同教育阶段的培养目标也并不相同。因此，人工智能在中国教育中的应用并不是简单的技术复制，而需要不断适应复杂多样的教育情境。这种大规模、多层次和差异化的应用环境，一方面为教育人工智能产品和模型提供了丰富的实践场景，另一方面也使技术的教育适配性、资源公平性和治理能力面临更高要求。

4.1.3 应用领域不断拓展

从目前的发展趋势来看，人工智能在教育中的应用已经不再局限于简单的内容生成和工具辅助，而是逐步向教学、学习、教研、评价、管理和公共服务等多个环节延伸。2025 年教育数字化工作部署已经提出从学生学习、教师教学、教育治理和科学研究等方面拓展人工智能应用场景；到 2025 年底，教育部介绍的人工智能应用已经涉及招生考试、教学评价、就业服务和校园治理等领域。

这意味着人工智能正在改变的已经不仅是教师如何备课、学生如何完成作业等具体环节，而是教育活动的组织方式和运行机制。一方面，人工智能能够降低教育资源生产和获取成本，提高教学辅助和学习反馈效率；另一方面，当人工智能逐渐参与学习分析、教育评价和学校治理之后，其影响也从教学工具层面进一步延伸至教育决策和学生发展。因此，人工智能教育应用越深入，技术与教育制度之间的关联就越紧密，对治理体系的要求也越高。

4.2 中国人工智能教育治理现状

4.2.1 顶层设计持续完善

中国人工智能教育治理目前最为突出的特点，是国家层面的制度供给持续加强。近年来，我国已经逐步形成从教育数字化战略到人工智能专项行动、从公共数字教育平台到应用规范的多层次政策体系。2025 年《关于加快推进教育数字化的意见》已经明确提出坚持应用导向、治理为基，同时强调健全适应数字化发展的制度体系和安全治理机制。2026 年《人工智能+教育》行动计划进一步提出建立智能应用能力评估体系，并推动形成更加开放协同的教育智能应用生态。

这说明我国人工智能教育治理正在从早期的数字基础设施管理逐步转向覆盖技术应用、数据治理、教育评价、人工智能素养和安全监管的综合治理。国家智慧教育平台本身也正在发生类似变化，从资源汇聚平台进一步向资源、数据、智能服务和教育治理相结合的平台发展。2025 年教育数字化相关部署已经提出建设平台智能版，并探索利用大数据和人工智能支持教育评价与科学决策。

4.2.2 治理方式逐步走向分层分类

人工智能在教育领域产生的风险具有明显的场景差异，因此治理不能简单采取统一标准。当前我国的治理思路已经开始强调分学段、分场景推进人工智能教育应用。例如，2025 年教育数字化相关部署明确提出分学段推进人工智能在教育领域的规模化应用。这种治理方式意味着，对于小学、初中、高中以及高等教育，不宜采取完全相同的人工智能使用规则，而应结合学生认知发展水平、教育目标和技术应用风险进行差异化安排。

与此同时，不同人工智能应用的风险程度也存在明显差异。一般性的资源检索、内容生成和教师备课辅助，与涉及学生画像、学习评价、招生考试和教育决策的应用，其潜在影响显然不同。因此，我国人工智能教育治理正在逐渐从对技术本身的笼统管理转向对具体教育应用场景的精细化治理。这种趋势也为后续建立风险分级、产品准入和动态监管机制提供了现实基础。

4.2.3. 治理体系逐渐形成政府引导、学校落实、多方参与的协同格局

人工智能教育涉及技术研发、产品供给、学校应用、教师教学、学生学习和家庭使用等多个环节，单一治理主体很难覆盖完整的责任链条。因此，我国人工智能教育治理正在形成政府引导、学校落实、企业参与以及社会多方协同的治理格局。政府主要承担战略规划、政策制定、公共资源供给和风险监管等职能；学校则承担人工智能技术选择、应用管理、教师培训和学生权益保护等具体责任；企业需要对产品安全、数据保护和技术服务承担相应责任；教师和家庭则直接参与人工智能在课堂和日常学习中的具体使用。

这一治理格局已经开始从原则性倡导走向制度化建设。2025 年的教育数字化政策明确提出强化部门协同、央地联动、区域协调并鼓励社会参与；2026 年的《人工智能+教育》行动计划则进一步提出推动高校、中小学协同研制人工智能课程和应用，并建设共创共享的智能应用体系。

4.2.4 治理能力仍处于持续完善阶段，技术快速迭代带来新的制度适配压力

总体而言，我国人工智能教育治理已经具备较好的政策基础和组织基础，但人工智能技术快速迭代所带来的治理挑战仍然十分突出。当前治理面临的问题已经不只是是否允许人工智能进入教育，而是如何判断具体产品是否适合教育、如何确定不同应用的风险等级、如何规范教育数据的采集和使用、如何保障学生权益，以及如何处理人工智能参与教育评价之后产生的责任问题。

与此同时，区域之间、学校之间的数字基础和教师人工智能素养仍存在差异。部分地区已经具备较好的人工智能应用基础，但一些基层学校在设备、网络、资源和教师能力方面仍然存在不足。因此，人工智能教育治理不仅需要解决技术风险，也需要关注技术应用可能产生新的教育不平等。换言之，中国人工智能教育治理已经从过去重点解决有没有技术、有没有平台的问题，逐步进入更加复杂的阶段，即需要进一步回答**技术如何进入教育、进入到什么程度、由谁负责以及如何确保其始终服务于育人目标**。

5. 人工智能教育应用与治理的六大原则

如前所述，人工智能技术的快速演进正在重构教育活动的技术基础、组织方式与运行机制。生成式人工智能、大模型、智能代理以及多模态人工智能逐步进入课程教学、教育评价、教师发展、教育管理和公共教育服务等场域，使教育数字化由资源数字化、业务数字化进一步转向智能化。在这一过程中，人工智能不仅表现为提高教育活动效率的技术工具，而且开始介入知识生产、学习过程、教学决策和教育评价，由此产生新的教育关系、权力关系与责任关系。

与传统教育信息化相比，人工智能具有较强的自主生成、模式识别、预测分析和交互能力，其教育影响具有更强的渗透性和不确定性。一方面，人工智能能够促进优质教育资源扩散，支持个性化学习，降低教师重复性工作负担，并为教育管理提供新的决策支持；另一方面，算法偏差、数据安全、技术依赖、认知惰性、评价失真以及教育机会分化等问题也随之出现。因此，人工智能教育治理不能局限于技术风险控制，而应当以教育基本规律和育人目标为依据，对技术应用的目标、边界、条件、责任和评价机制进行系统规范。基于此，人工智能教育治理需要形成以教育价值为根本遵循、以人的主体性为核心、以教育公平为基本底线、以适应性监管为制度保障、以师生数智素养为能力基础、以多元主体协同为治理机制的基本原则体系。

5.1 坚持教育价值引领，技术服务于人的全面发展

坚持教育价值引领，是人工智能进入教育领域必须遵循的首要原则。人工智能技术本身并不具有独立的教育价值，其教育意义取决于技术被置于何种教育目标、制度环境和实践场景之中。技术能力的快速增长并不意味着教育价值目标需要随之调整，相反，越是在技术快速发展的条件下，越需要以教育的根本目的对技术应用进行规范和校准。

从教育本质来看，教育不仅承担知识传授和能力培养职能，同时承担价值塑造、人格发展和社会化功能。学生的成长具有长期性、复杂性和情境性，人的认知能力、情感态度、价值观念、社会责任和创造能力不能完全通过技术指标加以量化。因此，人工智能教育治理必须避免将技术效率等同于教育质量，避免以人工智能应用程度替代对教育实际成效的判断。当前教育人工智能应用存在一定程度的技术先行现象，即先确定可以使用何种技术，再寻找相应的教育应用场景。这种路径容易导致技术能力对教育目标形成反向塑造，使原本服务于教育的技术逐渐成为教育活动的组织依据。

人工智能教育应用应当建立以教育目标为起点的决策逻辑。学校在引入人工智能系统或者开展人工智能教学应用之前，需要首先明确拟解决的教育问题以及预期达到的育人目标，再判断人工智能是否具有不可替代或者明显优于传统方式的功能优势。例如，在教师备课、教学资源整理、学习资料检索等领域，人工智能能够有效降低重复性劳动成本，其应用具有较强的合

理性；但在涉及价值判断、师生情感互动、学生人格发展以及复杂教育决策的场景中，则不宜以技术效率作为主要判断依据。由此，人工智能教育治理应当逐步形成教育目标优先的应用评估机制。技术是否先进并不是进入教育场景的充分条件，其必要性应当建立在教育价值和实际需求基础之上。

同时，人工智能能够显著提高教育活动的信息处理效率，但效率本身并不能充分代表教育质量。如果人工智能应用主要表现为提高任务完成速度、增加教学资源数量或者降低管理成本，而没有带来学生学习质量和发展水平的实质改善，那么这种技术应用的教育价值需要重新评估。尤其是在生成式人工智能广泛应用之后，教育活动中的部分任务已经可以由人工智能快速完成。例如，文本生成、资料整理、习题生成、语言润色等工作可以在较短时间内完成。这种变化一方面能够降低部分认知劳动成本，另一方面也使传统以作业结果为中心的教育活动面临重新调整的必要。如果学生可以通过人工智能直接获得完整的作业成果，而教育评价仍然主要依据最终成果进行判断，那么评价结果可能无法真实反映学生的知识掌握和能力发展状况。由此可见，人工智能对教育的影响并不仅仅是增加一种新型教学工具，而是要求教育重新审视学习过程、能力形成与教育评价之间的关系。因此，应当将人工智能的效率优势转化为促进高阶学习和深度学习的条件，而不是将效率本身作为教育改革的最终目标。

此外，人工智能教育治理需要改变以技术应用数量和平台使用规模为主要依据的评价倾向。学校是否部署人工智能平台、教师使用人工智能的频率、学生接触人工智能的比例等，可以作为描述教育智能化程度的指标，但不能作为判断人工智能教育成效的核心指标。更加重要的是考察人工智能是否促进了学生认知能力、实践能力、创新能力、合作能力和社会责任的发展，是否提高了教师专业工作的质量，是否扩大了优质教育资源的可及性，以及是否为不同类型学习者提供了更加适切的学习支持。这一原则意味着，人工智能教育治理需要建立技术指标与教育指标相结合的评价体系。其中，技术性能属于基础性指标，教育效果属于核心指标，育人成效则构成最终评价尺度。只有当人工智能技术的应用能够转化为真实的教育增益，其技术价值才能获得教育意义上的正当性。

5.2 坚守人本内核，明确人机协同定位

坚持人本内核，是人工智能教育治理区别于一般人工智能治理的重要方面。教育活动的对象不是抽象的数据主体，而是具有主体意识、情感需求、发展差异和社会关系的具体的人。因此，人工智能进入教育领域之后，其功能定位必须建立在维护人的主体性基础之上。人工智能能够承担部分知识处理、内容生成和决策支持功能，但教育过程并非单纯的信息加工过程。教师与学生之间形成的信任关系、情感联系、价值交流以及长期互动，是教育发挥育人功能的重要基础。这些因素具有高度情境化和个体化特征，难以由人工智能完整复制。

因此，人工智能教育治理的基本方向不应是追求机器对教育活动的全面替代，而应是形成

更加合理的人机协同关系，使人工智能承担适合技术完成的工作，同时将涉及教育价值判断、复杂情境判断和人格发展的职责保留在人类主体手中。

首先，人工智能进入教育以后，教师的职业角色将发生一定变化，但这种变化并不意味着教师主体地位的弱化。相反，随着人工智能承担更多知识检索、资源生成、数据分析和事务性工作的任务，教师专业判断的重要性可能进一步提高。教师需要对人工智能提供的信息和建议进行甄别，对生成内容进行专业审核，并结合具体学生的学习基础、认知特点和发展需求进行教学决策。因此，人工智能提供的是决策支持，而不是教育决策本身。特别是在学生评价、学业诊断、心理状态判断、奖惩决定以及其他可能对学生产生重大影响的教育事项中，应当保留明确的人类审核和干预机制。人工智能可以提供分析结果，但不能在缺乏人工审查的情况下直接决定学生的教育机会和发展路径。

其次，人工智能对学生学习方式的改变尤其值得关注。生成式人工智能能够快速回答问题、生成文章、解释概念并提供解决方案，使学习活动中的信息获取成本显著降低。但是，信息获取成本降低并不必然意味着学习质量提高。学习的本质不仅是获得答案，还包括提出问题、形成假设、搜集证据、进行推理、修正观点以及形成个人理解的过程。如果人工智能直接替代这些过程，学生虽然能够获得更完整的学习成果，却可能减少真正参与认知活动的机会。因此，人工智能教育治理需要关注技术使用对学生认知过程的影响。学校在设计人工智能支持的教学活动时，应当更加重视学生的问题提出、过程记录、观点形成和成果反思，并通过过程性评价减少单纯依赖最终成果的评价方式。

同时，人工智能高度依赖数据，由此可能推动教育活动进一步数据化。学生的学习行为、作业完成情况、课堂表现、兴趣偏好以及学习轨迹都可能成为算法分析对象。教育数据化能够为个性化教学提供依据，但也存在将学生简化为数据集合的风险。学生的学习状态具有动态变化特征，同一行为在不同情境下可能具有不同含义。算法能够识别数据中的模式，却未必能够理解学生行为背后的家庭环境、情绪状态、人际关系和成长经历。因此，人工智能教育治理不能将算法生成的学生画像视为学生真实状态的完整表达。对于涉及学生评价和教育决策的算法结果，应当明确其辅助性质，并建立教师复核、学生申诉和必要人工干预机制。

此外，人工智能可以模拟自然语言交流，但这种交互能力不能简单等同于教育关系。教师的教育影响不仅来自其传授的知识，还来自其人格示范、情感支持、价值判断以及对学生长期成长的理解。因此，在人工智能应用不断扩大的背景下，学校更需要明确哪些教育活动必须保留充分的人际互动。例如，对学生重大成长问题的指导、价值观教育、心理支持、师生关系建设以及涉及复杂情境的教育决策，都需要以真实的人际交流为基础。从治理角度看，人本原则并不意味着拒绝人工智能参与教育，而是要求在技术应用过程中始终维护人的主体地位、判断能力和关系网络，使技术承担辅助功能而不侵蚀教育的人文属性。

5.3 以教育公平为底线，缩小数字素养鸿沟

教育公平是人工智能教育治理必须坚守的基本底线。人工智能具有显著的资源放大效应，其教育收益并不会自动均等地分配给所有学生。如果不同地区、学校和家庭在人工智能工具获取、基础设施条件、教师能力以及家庭支持等方面存在较大差异，技术进步反而可能进一步扩大既有教育差距。因此，人工智能教育治理所关注的公平问题，已经不能停留于传统数字鸿沟意义上的设备和网络接入问题，而需要进一步关注人工智能使用能力、使用质量以及技术收益分配问题。

传统教育数字化阶段，数字鸿沟主要体现为设备、网络和资源获取方面的差异。人工智能时代，新的差距正在形成。同样拥有人工智能工具的学生，其实际使用效果可能存在显著差异。部分学生能够利用人工智能进行资料分析、知识拓展和问题探究，而另一些学生可能只能将其作为答案生成工具。这意味着人工智能教育公平不仅取决于是否能够使用技术，还取决于是否具有有效使用技术的能力。因此，未来人工智能教育公平需要同时关注技术接入、资源质量、使用能力和教育收益四个层面。

如果优质人工智能教育资源主要通过商业平台提供，并且不同家庭依据经济能力购买不同水平的人工智能服务，那么人工智能可能成为新的家庭教育投资工具。这种趋势如果缺乏制度约束，可能进一步强化家庭资本对学生教育机会的影响。因此，公共教育体系需要承担更重要的人工智能资源供给责任，通过建设公共数字教育资源、开放优质人工智能教育服务、加强薄弱学校技术支持等方式，降低人工智能教育的使用门槛。人工智能教育的发展不能以市场化资源供给作为唯一机制。对于具有基础公共服务属性的人工智能教育资源，应当进一步强化公共供给和普惠机制。

人工智能教育治理尤其需要关注农村地区、欠发达地区、特殊教育学生以及数字资源和家庭支持相对不足的学生群体。对于这些群体而言，人工智能不仅可能构成新的教育风险，也可能成为改善教育机会的重要工具。例如，人工智能可以通过语音交互、辅助阅读、个性化反馈和多模态资源，为部分特殊教育学生提供传统教育条件下难以获得的学习支持。因此，教育公平并不意味着所有学生接受完全相同的人工智能服务，而是根据不同群体的实际需求提供相对适切的技术支持，使人工智能成为弥补教育资源差异的重要机制。

5.4 打造适配技术迭代的监管体系

人工智能技术的快速迭代使教育治理面临显著不同于传统教育技术治理的制度环境。以生成式人工智能、大模型和智能体为代表的新一代人工智能技术具有持续更新、能力扩展和应用边界动态变化等特征，模型能力可能随着版本更新发生较大变化，其输出结果也具有一定程度的不确定性。这意味着学校在技术采购阶段完成的风险判断，并不能保证人工智能系统在后续

使用过程中始终处于同一风险水平。与此同时，人工智能正在由教学辅助工具逐步进入学习评价、学生管理、教育决策等更加敏感的领域，技术风险与教育风险之间的关联程度不断提高。在这一背景下，传统以产品准入、使用规范和事后处置为主要内容的监管方式，需要进一步向持续性、适应性和全过程治理转变。

人工智能教育监管首先需要实现监管对象从技术产品向教育应用场景的拓展。人工智能技术本身并不能完全决定其教育风险，同一种模型应用于不同教育场景时，其可能产生的影响存在明显差异。例如，将生成式人工智能用于教师备课和教学资源整理，主要涉及生成内容准确性、知识产权和教师专业判断等问题；而将人工智能用于学生综合素质评价、学业风险预测或者招生推荐，则可能进一步涉及学生隐私、算法偏见、教育机会分配和程序公平等问题。因此，人工智能教育监管不能仅根据产品类型或者模型性能确定监管强度，而应当综合考虑具体使用场景、影响对象、数据类型、决策后果和风险可逆程度。尤其对于可能影响学生教育机会和基本权益的人工智能应用，应当建立更加严格的审查和人工复核机制，从而避免技术工具在缺乏充分教育论证的情况下直接参与重大教育决策。

适应人工智能技术迭代，还需要建立覆盖技术全生命周期的治理机制。人工智能进入学校之后，其治理责任不能随着采购完成而终止，而应当贯穿产品选择、系统部署、数据处理、实际使用、模型更新、效果评价和退出等环节。学校在引入人工智能系统时，不仅需要关注产品功能和技术性能，还需要了解数据采集范围、数据存储方式、第三方数据使用情况、模型更新机制以及供应商的责任边界。在实际使用过程中，应当持续关注人工智能输出的准确性、稳定性和公平性，并根据技术版本更新和应用范围变化及时开展重新评估。对于已经停止使用的人工智能系统，还需要妥善处理学生和教师数据的删除、迁移和保存问题。只有将人工智能治理从一次性审批转变为持续性的过程管理，才能降低技术快速变化所带来的制度滞后风险。

人工智能教育治理还需要建立与风险程度相匹配的分类监管机制。当前一些人工智能治理实践容易在全面禁止和全面开放之间进行选择，但这两种方式均难以适应教育场景的复杂性。对于一般性的教学辅助工具，可以通过学校备案、教师审核和使用规范进行管理，以保证教育创新具有必要的空间；对于涉及学生行为数据、学习画像和个性化推荐的系统，需要强化数据治理、算法透明和使用边界；对于直接影响学生评价、升学、招生以及其他重大教育权益的人工智能系统，则需要提高准入标准，并明确人工审核、责任追溯和申诉机制。分类治理的意义在于使监管资源与风险水平相匹配，将有限的监管能力优先配置到可能产生重大教育影响的技术应用之中。

与此同时，人工智能教育监管需要解决责任主体分散所带来的责任认定问题。人工智能教育应用通常涉及模型开发者、产品提供者、平台运营者、学校管理者和教师等多个主体。当人工智能生成错误内容、形成不合理推荐或者产生具有偏差的评价结果时，如果缺乏明确的责任划分，容易出现技术责任和教育责任相互推诿的情况。因此，需要依据人工智能教育应用的实际运行过程明确各主体的责任边界。企业应当对产品设计、数据处理、模型安全和产品风险提

示承担相应责任，学校应当承担人工智能产品采购、部署和使用管理责任，教师则应当对具体教学活动中的人工智能使用承担专业判断责任。对于人工智能参与学生评价和教育决策的情形，还需要明确人工复核和最终决策责任。通过建立与技术链条相对应的责任链条，才能使人工智能教育治理真正形成可执行、可追溯的制度机制。

从更长周期看，适配人工智能技术迭代的监管体系并不是建立一套固定不变的规则，而是形成教育治理体系持续识别新风险、回应新问题和调整新规则的制度能力。教育行政部门需要加强人工智能教育应用的常态化监测和政策评估，学校需要形成内部人工智能治理机制，并通过实践反馈不断修正使用规范。人工智能教育监管由此应当由静态规则控制逐步转向动态风险治理，在维护教育安全、学生权益和基本教育秩序的同时，为合理的技术创新保留必要空间。

5.5 全面提升师生数智素养，培养理性使用技术的能力

人工智能教育治理不仅是制度建设问题，也是教育主体能力建设问题。人工智能最终需要通过教师和学生进入具体教育活动，如果教育主体缺乏理解人工智能、评价人工智能和合理使用人工智能的能力，即使外部监管制度较为完善，也难以有效降低技术误用和过度依赖所产生的风险。尤其是生成式人工智能降低了内容生产和知识获取的门槛，使学生能够更加容易地获得答案、生成文本和完成部分学习任务，也使教师面临教学设计、学习评价和专业判断方式的调整。因此，提升师生数智素养应当成为人工智能教育治理的重要基础，而不应仅被视为人工智能课程建设或者教师培训的附属工作。

人工智能时代的数智素养首先表现为对人工智能技术及其能力边界的基本理解。与传统数字工具相比，生成式人工智能能够以自然语言生成具有较高完整度和可信度的内容，这种表现形式容易使使用者将语言流畅性误认为知识可靠性。事实上，人工智能生成内容可能存在事实错误、逻辑错误、信息遗漏以及价值偏差，其输出结果与具体任务、数据条件和模型能力之间存在复杂关系。因此，师生数智素养不能简单等同于工具操作能力，而应当包括对人工智能基本运行机制、应用范围和技术局限的理解。只有认识到人工智能并不是天然可靠的知识来源，使用者才可能形成合理的技术预期，并在具体学习和教学活动中保持必要的判断意识。

在此基础上，人工智能素养的核心应当进一步体现为对技术输出进行验证和评价的能力。人工智能的广泛应用降低了信息生产成本，却相应提高了信息甄别和事实核验的重要性。学生面对人工智能生成的学习材料，需要能够判断信息来源是否可靠、事实是否能够得到其他证据支持以及不同答案之间是否存在逻辑冲突。教师则需要结合学科知识和课程目标，对人工智能生成的教学内容进行专业审核，判断其是否准确、是否适合学生的认知水平以及是否符合课程教学要求。因此，人工智能教育不能只培养学生如何提出有效指令和获得更好的生成结果，还应当将信息验证、证据意识、逻辑推理和批判性思维纳入人工智能素养培养的重要内容。

人工智能的发展还使教师专业能力的结构发生变化。传统教学环境中，教师的重要职责之一是进行知识选择、组织和传递，而人工智能正在降低知识获取和内容生成的成本。在这种情况下，教师专业价值并不会因为人工智能能够生成教学内容而降低，而是更加集中于课程目标确定、教学情境创设、学生差异识别、学习过程指导和教育价值判断。教师需要能够判断何时使用人工智能能够改善教学，何时使用反而会削弱学生的认知活动，并能够根据学生具体情况调整人工智能参与程度。因此，教师人工智能素养建设不应停留在工具培训层面，而应与教师专业发展和教学能力提升结合起来，使教师形成面向人工智能环境的课程设计能力、学习评价能力和专业判断能力。

对于学生而言，人工智能素养培养尤其需要处理技术便利与主体能力发展之间的关系。人工智能可以减少资料查找、文本表达和问题解决的时间成本，但如果学生将人工智能直接用于替代阅读、思考和表达过程，则可能形成技术依赖，使学习结果与真实能力之间出现偏离。因此，学校需要逐步建立与人工智能环境相适应的学习规范，将人工智能使用纳入学术规范和学习过程管理。例如，对于允许使用人工智能的学习任务，应明确人工智能可以介入的环节以及学生需要独立完成的环节；对于强调基础知识掌握和独立思考的任务，则需要保留必要的自主完成空间。同时，应当逐步加强过程性评价，使评价能够反映学生的问题提出、资料分析、论证过程和反思能力，而不仅仅评价最终提交的文本成果。

进一步而言，师生数智素养的提升还应当包含技术伦理和责任意识。人工智能使用涉及个人信息保护、知识产权、学术诚信、算法偏见和信息安全等问题，教师和学生如果缺乏相应的伦理判断能力，即使能够熟练使用人工智能工具，也可能产生新的教育风险。因此，人工智能素养教育应当帮助师生理解技术使用所对应的责任边界，使技术能力、批判性能力和伦理意识形成统一。其最终目标并非提高人工智能工具的使用频率，而是使教育主体具备根据教育目标和实际需要自主决定是否使用、如何使用以及何时停止使用人工智能的能力。

从治理角度看，师生数智素养实际上构成了人工智能教育治理的内生性保障。制度能够规定技术使用的基本边界，但无法覆盖每一次课堂互动和学习行为。只有当教师和学生形成稳定的技术判断能力，治理要求才能真正进入日常教育实践。因此，师生数智素养建设应当从阶段性培训逐步转向持续性的能力发展，并与课程体系、教师专业发展、学校文化和教育评价改革结合起来，形成能够伴随人工智能技术持续演进而不断更新的教育能力体系。

5.6 推动多元共治，凝聚政府、学校、企业、家庭的治理合力

人工智能进入教育之后，教育治理的主体结构正在发生变化。传统教育治理主要依托教育行政部门、学校和教师构成相对清晰的责任体系，而人工智能教育应用进一步引入了模型开发企业、平台企业和技术服务商等主体，同时将家庭和社会环境更加直接地纳入人工智能使用过程。人工智能技术本身具有跨组织、跨场景和跨地域运行的特点，其数据、模型和应用之间存

在复杂的技术链条和责任链条，仅依靠教育行政部门或者学校进行单独治理，很难覆盖人工智能教育应用的全部环节。因此，需要在明确主体责任的基础上形成政府、学校、企业和家庭共同参与的协同治理结构。

多元共治并不意味着简单增加治理主体，而是要求不同主体根据自身掌握的资源、能力和责任承担相应的治理职能。政府掌握公共政策制定和教育资源配置权，应当承担制度建设、公共服务和风险监管责任；学校处于人工智能教育应用的具体场域，需要承担技术选择、应用管理和师生权益保护责任；企业掌握人工智能模型、平台和产品能力，应当承担相应的技术安全、数据保护和产品责任；家庭则直接参与学生校外人工智能使用，需要承担与未成年人数字生活相关的教育和监督责任。只有各主体之间形成明确的责任分工，多元参与才不会演变为责任分散。

政府在多元治理体系中的作用，需要从单纯的行政监管进一步转向制度供给和公共治理能力建设。人工智能教育涉及教育、公平、数据、安全、未成年人保护和科技创新等多个政策领域，单一部门难以覆盖全部治理事项。因此，有必要加强教育行政部门与网信、科技、工业和信息化、市场监管等相关部门之间的政策协调，在人工智能教育应用标准、数据治理、产品管理和风险处置等方面形成更加稳定的制度衔接。同时，政府还需要承担公共人工智能教育资源建设和基础能力供给责任，避免人工智能教育资源过度依赖商业平台，使技术获取能力成为不同学校和家庭之间新的教育差距来源。

学校是多元共治结构中连接不同主体的关键组织。人工智能教育治理最终需要落实到课程、课堂、评价和学生管理等具体教育活动，因此学校不能仅仅作为人工智能产品的采购和使用者，而应当逐步形成独立的技术治理能力。学校需要建立人工智能应用管理制度，对人工智能工具的引入、使用场景、数据权限和教师使用规范进行管理，同时建立风险发现和问题报告机制。对于涉及学生评价、学生画像和个人数据处理的人工智能应用，应当形成更加严格的校内审查程序。学校还需要通过教师培训和学生教育，将人工智能治理要求转化为具体的教育实践规范，从而避免政策要求与课堂实践之间出现脱节。

企业责任是人工智能教育治理中需要进一步强化的重要环节。人工智能教育产品具有较强的技术复杂性，学校和教师往往难以独立判断模型训练、数据处理和算法机制所产生的全部风险。因此，企业不能仅仅承担一般意义上的产品供应责任，还需要承担与人工智能技术特征相适应的产品治理责任。对于进入学校的人工智能产品，企业应当向教育机构充分说明产品功能、数据处理范围、适用边界和已知风险，并在模型升级、功能变化和重大安全事件发生时及时履行告知义务。对于涉及未成年人的人工智能应用，还应当提高数据保护和风险控制标准，避免通过商业逻辑过度收集和利用学生数据。

家庭是人工智能教育治理体系中容易被忽视的主体。学生使用生成式人工智能和其他智能工具的行为并不局限于课堂，而是延伸到家庭作业、自主学习、信息获取和日常生活之中。学

校如果只对校内人工智能使用进行规范，而忽视家庭环境中的技术使用，就难以形成完整的治理闭环。因此，学校需要加强与家庭之间的信息沟通，为家长提供必要的人工智能素养教育，使家庭能够理解人工智能对学习方式和学生发展的影响。家庭治理的重点也不应只是限制使用时间，而应更多关注学生是否形成合理的使用习惯、是否具备信息辨别能力以及是否能够承担相应的学习责任。

多元共治还意味着建立有效的利益协调和问题反馈机制。人工智能教育治理过程中不可避免地存在不同主体之间的利益差异。企业关注技术创新和产品应用，学校关注教育效果和管理成本，教师关注专业自主性，家庭关注学生发展和权益保护，政府则需要兼顾创新发展、教育公平与公共安全。如果缺乏稳定的协商机制，任何单一主体的治理要求都可能难以得到充分落实。因此，应当探索由教育行政部门、学校、教师、技术企业、家长以及相关专业机构共同参与的咨询、评估和反馈机制，使人工智能教育政策能够吸收不同主体的实践经验，并在技术发展和教育需求发生变化时及时进行调整。

从治理结构的角度看，多元共治的最终目标并不是形成更加复杂的管理体系，而是建立清晰、协同和可追责的责任结构。政府需要解决制度供给和公共利益保障问题，学校需要解决教育场景中的具体实施问题，企业需要解决技术产品和数据治理问题，家庭需要承担学生校外使用过程中的协同教育责任。通过责任明确、信息共享、风险共担和协同决策，可以降低单一主体治理的局限性，使人工智能教育治理由行政控制逐步转向制度治理、专业治理和社会协同相结合的综合治理模式。

6. 对策与建议

上述六项原则从教育价值、治理边界、发展公平、监管机制、主体能力和协同关系等方面，构成了人工智能教育治理的基本框架，也为回答人工智能如何进入教育、应当遵循什么边界以及如何实现规范发展提供了基本遵循。但面对人工智能技术加速迭代以及应用场景不断拓展的现实，治理理念还需要进一步落实到具体的制度和实践之中。尤其是在技术产品进入教育场景、教育数据持续流动、人工智能参与学习评价以及教师教学方式发生变化的过程中，新的治理问题正在不断出现。如何将教育价值和治理原则落实到这些具体环节，既需要完善制度规则，也需要建立与技术发展相适应的治理机制。基于此，本报告进一步从顶层设计、产品准入、数据治理、教育评价和教师发展五个方面提出具体对策，推动人工智能教育治理由原则层面的价值确立进一步走向实践层面的制度完善。

第一，强化顶层设计，完善法律法规与政策体系，构建人工智能教育治理的制度基础。人工智能深度进入教育领域后，其影响已经超越单一技术应用范畴，涉及教育权利保护、未成年人保护、个人信息安全、算法治理、知识产权、教育评价以及教师专业权等多个政策领域。现有教育信息化治理体系在一定程度上仍以数字平台、教育软件和信息系统为主要治理对象，对于生成式人工智能带来的内容生成、算法决策、人机交互以及数据持续反馈等新型风险缺乏足够的制度回应。因此，应从国家层面加强人工智能教育治理的顶层设计，推动教育政策、数据治理、网络安全、未成年人保护和人工智能监管等相关制度之间的衔接，逐步形成权责清晰、层级协调、标准统一的人工智能教育治理制度体系。在制度建设过程中，应进一步明确人工智能进入教育领域的基本法律边界和责任边界。对于涉及学生个人信息处理、学习画像生成、智能评价、教育资源推荐以及招生升学等具有较大教育权益影响的应用，应明确其使用条件、人工介入要求和责任承担机制，避免由于法律责任模糊导致技术责任向教师和学校单向转移。同时，可探索建立跨部门的人工智能教育治理协调机制，对人工智能教育应用涉及的教育政策、技术标准、数据安全和伦理风险进行统筹协调。在此基础上，逐步建立覆盖政策制定、标准研制、风险监测、事件处置和效果评估的治理体系，使人工智能教育治理从政策层面的原则性倡导进一步走向具有可执行性和可问责性的制度安排。

第二，抬高准入门槛，建立以教育适配性为核心的产品标准，实施人工智能教育应用分类治理。人工智能技术具有明显的通用性，但教育场景具有高度的专业性和特殊性。一个在商业场景中表现良好的人工智能产品，并不意味着其天然适用于学校教育。尤其是教育活动直接关系学生的发展权、受教育权和人格成长，如果仅以模型性能、功能丰富程度或者市场使用规模作为产品进入教育领域的主要依据，容易形成技术逻辑对教育逻辑的替代。因此，应将教育适配性确立为人工智能教育产品进入学校的重要评价标准，将产品是否符合教育规律、是否有利于学生发展、是否能够支持教师专业判断以及是否具有可解释、可干预和可追责特征纳入准入

评价。具体而言，应建立人工智能教育产品的分类分级准入制度，根据应用场景、服务对象、数据敏感程度以及可能产生的教育影响，对人工智能应用进行差异化管理。对于一般性的备课辅助、资源生成和学习支持工具，可以通过备案、技术检测和使用规范等方式实施相对灵活的管理；对于涉及学生学习画像、自动评价、个性化推荐等具有较强决策影响的应用，则应提高安全性、公平性、准确性和人工监督方面的准入要求；对于可能直接影响学生重大教育权益的应用，应建立更加严格的审查制度，并明确人工复核和最终决策机制。在此基础上，可以建立动态白名单制度，但白名单不应成为一次性认证，而应与产品版本更新、模型能力变化和实际应用效果相联系，形成持续评估、动态调整和退出机制。对于发现存在严重安全风险、明显教育偏差或者无法满足教育伦理要求的产品，应及时暂停或退出教育应用场景，从而实现由静态产品准入向动态风险治理转变。

第三，统筹数据治理，打破数据孤岛，建立贯穿人工智能教育应用全过程的数据安全体系。数据是人工智能进入教育领域的重要基础,但教育数据同时具有高度敏感性和较强的公共属性。学生的学习记录、行为数据、家庭信息、心理特征以及发展评价等数据一旦被过度收集或者不当使用，不仅可能造成个人隐私泄露，还可能通过算法分析形成持续性的学生画像，并进一步影响教育资源配置、学习机会和评价结果。因此，人工智能教育治理不能仅仅关注模型本身的安全性，还需要将数据治理置于更加基础的位置，建立覆盖数据采集、存储、传输、处理、共享和退出全过程的数据治理机制。首先，应推动不同教育系统之间的数据标准衔接和规范建设，在保障数据安全与合法合规的前提下解决长期存在的数据孤岛问题，使必要的数据能够在明确授权和责任边界的基础上实现合理流动。其次，应坚持数据最小必要原则，严格控制人工智能教育应用的数据采集范围，避免以提升模型效果为由无限扩大数据收集。对于涉及未成年人的敏感数据，应进一步强化访问权限、使用目的、保存期限和第三方共享等方面的管理，并建立数据脱敏、权限控制和安全审计机制。同时，应明确数据的使用边界，防止学校或者企业将教育数据用于与教育目的无关的商业分析和用户画像。在此基础上，还应建立数据安全事件的发现、报告和处置机制，使数据治理从事后补救转向事前预防、事中监测和事后追责相结合的全过程治理，从根本上降低人工智能教育应用的数据风险。

第四，重塑教育评价体系，回归育人本质，建立适应人工智能时代的综合评价机制。人工智能对教育评价带来的影响，并不仅仅是学生是否使用人工智能完成作业的问题，更深层次的问题在于，传统评价体系所依赖的知识获取、文本表达和标准化作答方式正在受到人工智能技术的重新塑造。如果教育评价仍然主要关注最终答案、文本成果和知识记忆，那么人工智能越是降低信息获取和内容生成的成本，传统评价的区分度就越可能下降。因此，人工智能时代的教育评价改革不能简单采取禁止或者限制学生使用人工智能的方式回应技术挑战，而应重新思考教育评价究竟要评价什么，并通过评价机制推动学生保持真实的学习过程和主体性的认知活动。应进一步提高对高阶思维、问题解决、创新能力、实践能力、合作能力和价值判断等综合素养的评价权重，将评价关注点从单一结果逐步拓展到学习过程和思维过程。对于允许使用人

工智能的学习任务，可以重点考察学生如何提出问题、如何判断人工智能生成内容的可靠性、如何选择和整合不同信息、如何发现技术生成结果中的错误，以及能否对最终成果作出具有独立性的解释和反思。与此同时，对于涉及基础知识建构、核心能力形成和关键认知过程的学习任务，也应保留必要的自主完成环节，避免人工智能替代学生必须经历的阅读、思考、表达和实践过程。由此，教育评价才能真正成为人工智能教育治理的重要政策工具，通过评价导向约束技术的过度使用，同时利用人工智能推动评价方式更加多元、过程更加完整、反馈更加及时，最终实现技术创新与育人目标之间的协调。

第五，持续赋能教师队伍，加强人工智能应用与伦理培训，夯实人机协同的专业基础。教师是人工智能进入课堂并真正转化为教育实践的关键主体，因此，人工智能教育治理最终能否落地，很大程度上取决于教师是否具备理解、判断和驾驭人工智能的专业能力。当前教师人工智能能力建设不应局限于工具操作培训。如果教师培训仅仅停留在如何使用某一人工智能平台、如何生成教学材料或者如何设计提示词等层面，教师很容易形成对特定工具的依赖，却难以形成面对技术快速变化时所需要的独立判断能力。因此，应将教师人工智能素养纳入教师专业发展的整体框架，将技术理解、教学设计、数据意识、伦理判断和风险识别作为教师专业能力的重要组成部分。更为重要的是，应借助人工智能重新审视教师职业角色的变化。人工智能可以承担部分信息检索、材料生成、作业辅助和常规分析工作，但教师的核心价值并不会因此消失，而是更加集中于课程目标设计、学习情境创设、学生差异识别、情感支持、价值引导以及教育决策等具有高度专业判断性质的工作。因此，教师培训应帮助教师形成清晰的人机协同边界，使其能够判断何时适合使用人工智能、何种任务可以交由技术辅助、何种教育活动必须由教师自主完成。同时，应加强人工智能伦理、数据保护、算法偏差、学术诚信和学生权益保护等方面的培训，使教师不仅成为人工智能工具的使用者，也成为教育场景中的技术判断者和风险把关者。与此同时，还应建立持续性的教师发展机制，将人工智能能力培养与教师职前培养、入职培训、继续教育和校本研修相结合，避免将人工智能培训变成一次性的专项活动，从而为人机协同教学形成稳定而持续的专业支撑。

结语

人工智能正在深刻重塑全球教育体系。这种重塑的深刻性，不仅在于技术能力的持续跃升，更在于它正在重新定义知识生产与传播的方式、重新配置教育权力的结构、重新厘清人与技术之间的边界。本报告从发展趋势、核心危机、治理模式、中国实践等多个维度对这一变革进行了系统梳理，从中可以清晰地看到几个贯穿始终的核心判断。

第一，人工智能对教育的影响是结构性的，而非工具性的。 AI 不再只是提高效率的辅助手段，而是正在进入教育体系的底层运行逻辑。它改变知识传播的权威结构，改变课堂的组织方式，改变教师的劳动形态，也改变学生的学习路径。这种结构性影响意味着教育体系不能仅在技术应用层面做出调整，而必须在教育理念、制度设计、评价体系和人才培养目标上进行系统性回应。

第二，三大危机揭示了技术逻辑与教育逻辑之间的深层张力。 技术中心主义使教育评价从“是否解决问题”滑向“是否使用技术”，数据迷恋使可测量性僭越了教育性，主体性危机使学生和教师的教育主体地位面临被算法替代的风险，信任危机则动摇了学术诚信、技术信任和学历制度的社会基础。这些危机的本质，是技术逻辑对教育逻辑的系统性替代。治理的关键不在于拒绝技术，而在于重建教育逻辑对技术应用的主导权。

第三，治理能力已成为教育竞争的核心维度。 从欧盟的风险监管、新加坡的平台治理到日本的审慎创新，不同国家的治理模式差异表明：谁能够建立与技术发展相适应的制度框架，谁就能够在享受技术红利的同时控制其风险。中国拥有全球最大的教育场景优势和快速的政策响应能力，但在分层分类治理、产品准入、数据安全和教师赋能等方面仍需持续完善。

第四，教育公平面临新的挑战与机遇。 AI 既可能通过资源普惠缩小差距，也可能通过"强者更强"的马太效应扩大不平等。公共教育体系必须承担 AI 资源供给责任，确保技术获取能力不成为新的教育差距来源。对于农村地区、欠发达地区和特殊教育群体，AI 应成为弥补差距的工具而非加剧分化的力量。

第五，师生数智素养是治理有效性的内生保障。 制度能够规定边界，但无法覆盖每一次课堂互动。只有当教师和学生形成稳定的技术判断能力——理解 AI 的能力边界、验证生成内容、识别算法偏见、坚守学术诚信——治理要求才能真正转化为教育实践。教师的核心价值将从知识传授转向学习设计、情感支持和价值引导，这一转型需要系统性的教师发展机制支撑。

面向未来，人工智能与教育的融合将更加深入。智能体系统可能进一步承担教育组织功能，多模态技术可能创造沉浸式学习体验，个性化学习系统可能重塑班级授课制的基本形态。但无论技术如何演进，教育的基本使命始终不变：培养具有独立思考能力、创新精神和社会责任感的完整的人。使 AI 服务于这一使命而非背离它，是人工智能时代教育治理的根本遵循。技术

越是强大，人类越是需要回归对教育目的和人的发展的根本思考——这既是本报告的基本立场，也是人工智能时代教育变革必须坚守的方向。