

智能教育领域监管沙盒的国际经验及启示

刘鑫怡

(中国科学技术信息研究所, 北京 100038)

摘要:当前,智能教育应用蓬勃发展,为建设人人皆学、处处能学、时时可学的学习型社会提供了支撑。然而,技术快速迭代和广泛应用也伴生出内容生成、个人信息权益保护和人机关系等风险,传统治理范式日益显现出化解风险的局限性。探讨了监管沙盒在智能教育领域的应用价值,作为平衡技术创新与风险防控的新型治理工具,其通过构建可控试验环境,破解数据算法监管薄弱、风险外溢、人机权责边界模糊等关键难题。在梳理典型国家的智能教育监管沙盒实施路径的基础上,结合中国现行法律法规体系,建议以三类高风险场景作为智能教育监管沙盒适用对象,构建安全规制与激励措施并重的监管沙盒规则,为智能教育领域的监管创新提供可操作的制度设计。

关键词:智能教育; 监管沙盒; 人机关系

中图分类号: G311 **文献标识码:** A **DOI:** 10.3772/j.issn.1009-8623.2025.11-12.010

当前,人工智能技术快速发展,深刻推动学习方式转型和教育范式变革。“智能教育”利用智能技术加快推动人才培养模式和教学方法改革,构建包含智能学习、交互式学习的新型教育体系^[1],为提升教学质量、形成个性化学习体验、减轻教师负担和促进教育公平提供了新工具,广泛应用于教师教学、学业辅导、教育评价等场景。2025年联合国教科文组织的调查显示,在参与联合国教科文组织教席和姊妹大学网络的高等教育机构中,有近2/3的机构正在使用或开发人工智能^[2]。2025年8月,《国务院关于深入实施“人工智能+”行动的意见》指出,把人工智能融入教育教学全要素、全过程,创新智能学伴、智能教师等人机协同教育教学新模式,推动育人从知识传授为重向能力提升为本转变,加快实现大规模因材施教,提高教育质量,促进教育公平^[3]。然而,“学生课堂监测”等智能教育应用也带来了数据安全、隐私保护和人机伦理关系等问题,人工智能赋能教育面临风险防控和安全

监管的新挑战。《教育强国建设规划纲要(2024—2035年)》强调,要建立基于大数据和人工智能的教育评价和科学决策制度。加强网络安全保障,强化数据安全、人工智能算法安全和伦理安全。如何在提升教育质量的同时,形成系统性、前瞻性的风险防控机制,以及建立科学合理的智能教育监管体系,是亟待解决的问题。

1 智能教育应用快速发展引发的风险挑战

智能教育主要应用于学生学习、精准教学、画像评价等场景,在相关应用的研发、上线服务等环节,可能存在数据处理不合规、算法偏见、模型测试不到位和人工审核缺失等问题,导致在实践中产生较多风险隐患,对学生和教师的个人权益及教学环境产生不利影响。

1.1 生成式人工智能的教育内容输出风险

教育领域强调知识输出的严谨性、可解释性和无偏性。生成式人工智能广泛应用于智能教育服务,

作者简介:刘鑫怡(1994—),女,硕士,助理研究员,主要研究方向为科技政策、人工智能治理。

项目来源:中国科学技术信息研究所创新研究基金“法律与技术协同视角下人工智能监管沙盒制度研究”项目(QN2025-03)。

收稿日期:2025-09-10

然而因其技术的不确定性和不透明等内在特征，可能输出不准确的内容，甚至会加剧偏见和歧视，易给认识力、判断力尚不成熟的学生带来不良影响。首先，智能教育应用可能输出错误事实和不当内容，影响学生日常学习和教师教学。有些错误易于察觉，如对数学题解答错误、生成的地图和国家形象标识错误等，可由教师及时纠正；但有些文字或图像表达隐晦，审核易疏漏，可能误导学生，甚至扩大错误的传播范围。例如，一些儿童智能手表产品出现错误回答和不良导向内容，导致产品下架整改^[4]。其次，知识的输出过程需要可验证的推理过程作为支撑，但是生成式人工智能算法具有“黑箱”属性，透明度问题在技术层面和监管政策上仍未完全解决，导致其输出结果的可信度存疑。最后，人工智能的输出内容还可能加剧既有的偏见和歧视。大模型在研发过程中的训练数据可能不完全符合国家法律法规、教育规范、学科标准，算法模型可能隐含了研发者的主观意志等，这些情况均会导致输出结果带有一定的偏见和歧视。有研究显示，当ChatGPT在输出与职业相关的文本时，男性职业类型的文本表现出更高的多样性^[5]。某人工智能检测器在评估学生写作时，有61.3%的概率将非英语母语学生的文本错误地标记为人工智能生成，而对于英语母语学生的写作评价，误报率不超过10%，这可能会加剧对非英语母语学生的偏见^[6]。

1.2 智能教育监测评价可能侵害学生的个人信息权益

目前，多家企业推出了面向课堂管理和学生画像的智能教育产品，如某产品可以通过人工智能全自动采集课堂观测点数据，包括学生的人脸识别、身体姿态识别和情感分类等，以分析学生的活跃度和专注度；也有关于学生成长画像的智能教育产品，能够通过对学生行为进行分析和情感计算等，生成学生能力报告，为家长和教师提供个性化教育建议。虽然此类智能教育产品的初衷是更客观、更精准地对学生表现进行评价，但其应用涉及学生个人隐私等信息，可能存在使用缺乏合法性、合理性的情形。第一，在智能教育应用过程中，未经当事人同意就处理其个人信息。对于人脸信息这类敏感个人信息，《中华人民共和国个人信息保护法》规定，应当取得个人的单独同意，并告知处理敏感个人信息的必

要性以及对个人权益的影响，特别是对于不满14周岁的未成年人，还应征得其监护人的同意。但是实践中出现了学校未经家长同意，强制学生“刷脸”进行校园消费的事件，未能有效保障个人信息所有者的知情权和选择权。第二，相关应用的必要性存疑。《人脸识别技术应用安全管理办法》规定，应用人脸识别技术处理人脸信息，应当具有特定的目的和充分的必要性，采取对个人权益影响最小的方式，并实施严格保护措施。“充分的必要性”意味着该项个人信息的收集对于实现特定处理目的是不可或缺的，但对于学生课堂表现和学习能力评价这一场景，即使不进行人脸识别、情绪识别，也能通过学生回答内容、协作贡献程度、作业完成质量等指标实现评价，因此对学生个人信息的收集不是必须的。使用这一智能技术会伴随多重弊病，既造成学校、教师对学生个人信息和个人自由的过度干预，又提高了处理个人信息的合规成本。第三，相关技术应用的信息保护措施可能存在不足。公安部在“护网—2025”专项工作中发现，某主营业务为对外提供人工智能模型训练基础数据（算力）的科技有限公司，在处理人脸等生物识别类敏感个人信息前，未按《个人信息保护法》有关规定进行个人信息保护影响评估^[7]。涉及人的情绪、心理以及升学评价的智能教育产品，在使用过程中处理大量个人敏感信息，一旦被滥用或误用，可能对学生造成评价偏见等不利影响，这对信息保护提出了更高的监管和技术要求。

1.3 智能教育应用引发人机关系重塑风险

人类中心主义认为，人类比非人类具有更高的内在道德价值^[8]。然而，人工智能的应用正在打破这一传统观点。目前，学生、研究人员在使用智能教育产品时具有较高的自由度，而人工智能在回答问题、数据分析和文献整理等环节的高效性，可能导致学生过度依赖技术，忽视对问题的深入思考和原创性探索。长此以往，学生的主体性逐渐让位于机器的“意志”。除了学生对智能技术的依赖，教师和学校管理者同样面临着如何更好地应用智能技术的困惑。“数字人教师”、人工智能教科书等智能教育技术快速发展，引发了教师对其工作被替代的深刻担忧——教师可能从教育的主导者转变为技术的辅助者，致使传统的“教书育人”面临向“算

法育人”异化的风险。智能教育系统通过数据采集和分析形成的数字规训机制，决定了教学资源的分配，并且潜移默化地影响师生的行为模式。此外，学校智能化管理也涉及自主决策问题，如英国科茨莫尔学校聘用人工智能机器人作为校长以提高决策能力，该事件引发了巨大争议。有观点认为，严重依赖人工智能分析考试成绩和出勤记录数据会削弱信任^[9]。为了防止人工智能影响人类主体地位，新加坡教育部发布了《人工智能教育伦理指南》，强调人工智能系统应保留学习者对重要决策的选择和控制权，避免学生和教育工作者的核心能力被削弱^[10]。

由于智能教育应用可能带来的各类风险，欧盟发布《人工智能法》，将用于教育和职业培训的部分人工智能系统列入高风险清单，对其实施专门监管。其他国家主要采取“软法”指导方式，如美国教育部发布《利用人工智能设计教育：开发人员必备指南》，要求警惕数据集偏见问题等。然而，“软法”不具备强制约束力，难以有效应对智能教育应用中涉及的未成年人个人信息权益保护、机器决策对人的异化等高风险问题，因此，在规范智能教育产品准入、平衡技术创新与安全治理的总体导向下，监管沙盒作为一种试验性监管方法展现出重要的实践价值。该方法通过容错试错机制，有效协调了智能教育规模化应用的需求和安全合规的保障，加速了高风险、高敏感类智能教育产品的培育和开放，推动了教育新场景、新模式的大规模应用。

2 智能教育领域监管沙盒的理论基础与国际实践

2.1 智能教育监管沙盒的理论溯源

传统观念认为，规制体系是包含规则或规范制定、信息反馈及行为纠正机制三元素的控制体系^[11]，而新兴技术的快速迭代，使政府部门在行使传统规制职能时，其所掌握的知识与信息滞后，无法对预期行为做出准确判断。这一问题映射到智能教育领域，导致了部分场景监管缺位、监管措施准确性不足等问题。传统“命令—控制”类的治理模式在信息匮乏和不确定性的状态下可能会陷入治理困境^[12]。为了将被规制对象的自我组织能力与

相关部门的规制能力相关联，以形成对新兴技术的有效规制体系，试验性监管理念应运而生，其中监管沙盒是一种典型的监管方法。监管沙盒指规制机关设定的与外界隔绝的安全空间，在此安全空间内，市场主体可以测试科技产品、商业模式和营销方式，并且通常承担较轻的规制负担，享有部分的法律责任豁免^[13]。

监管沙盒制度最初适用于金融科技领域。英国金融行为监管局于2015年率先推出监管沙盒，通过客户数量、资金规模等方面的限制，构建了一个内嵌于真实市场的“安全空间”，通过利益相互冲突的各方平等协商，对性状不明、风险不定、无法满足完全监管要求的金融创新进行试验，并根据试验结果决定是否将其推向市场^[14]。2019年，经济合作与发展组织（Organisation for Economic Co-operation and Development, OECD）发布《人工智能原则》，首次提出“人工智能监管沙盒”这一概念，建议政府“考虑通过创建一个受控的环境，以便在其中进行适当的测试和人工智能系统探索”^[15]。目前，法国、挪威等已将其应用于智能教育领域，在整体监管框架下，针对智能教育中的不同应用开展监管沙盒实践。

2.2 智能教育监管沙盒的应用价值

部分智能教育产品或服务存在不确定性，违背了教育领域对准确性、严谨性的要求，这些产品或服务在与学生心理、发展成长相关的场景应用时，可能产生更多的法律、伦理问题，对未成年人的社会影响更深远。同时，目前对人工智能服务提供者、学校、学生及家长之间的责任分配没有明确规定，导致相关科技企业在布局智能教育市场时面临较大的不确定性风险。在此背景下，监管沙盒模式凭借其容错性高和多元主体协作等特点，为智能教育创新提供了一种兼具灵活性与规范性的混合治理范式，有助于推动既有监管规则针对特定技术创新情境的动态优化。一是该模式可针对目前监管的薄弱环节进行强化，明确人工智能教育应用的内容输出标准，确立其在知识准确性、伦理规范等方面的内容合规底线和质量高线。相关企业可从训练数据源头入手，在沙盒内探索数据存储、加工等方面的技术安全保障措施，降低数据安全隐患，并进一步筛选适合开展教育活动的优质训练数据。二是在特

定范围和期限内，该模式为尚未进入市场的产品或服务创建了受控的环境，避免了对个人合法权益的侵害。对于学生评价画像、课堂监测等仍存在合法性争议的应用，可通过监管沙盒将风险控制一定范围内。同时，在监管沙盒内构建减责免责规则，既能为企业提供容错试错空间，也能使相关部门通过与沙盒测试对象的交互深入地理解技术，并对其监管。三是该模式能够在可控沙盒内测试具有高度自主性或人机融合类的智能教育产品，如具身智能等，探索人类与人工智能协作的边界。通过构建受控环境，监管者可以系统性地评估人工智能在教育决策、师生互动中的角色定位，监测可能出现的算法偏差与伦理风险，同时通过实时数据反馈，不断优化人机协同模式，最终形成兼顾教育质量与个人信息安全的智能教学范式。

2.3 智能教育监管沙盒的国际实践

欧盟、美国、英国等国家和地区陆续在智能教育领域开展监管沙盒制度实践，以推进人工智能技术创新发展。监管沙盒最重要的价值是对传统监管方式的突破和创新，根据监管限制程度与责任豁免范围的不同，各个国家和地区在实践中形成了不同的智能教育监管沙盒实施路径。

欧盟以《人工智能法》为基础构建了“有限豁免”的智能教育监管沙盒制度。该法案指出，监管沙盒旨在指导、帮助企业适应既有的欧盟人工智能监管规则，并以沙盒的测试结果作为验证、改进监管制度的依据。为落实欧盟的这项法案，西班牙、法国和挪威等国家陆续启动教育领域监管沙盒项目，聚焦智能教育应用带来的数据安全、个人隐私等问题，帮助企业遵守现有法规，且均强调不允许在沙盒中突破现行法律限制，不能给予民事责任豁免，仅可在一定条件下免除沙盒测试对象的部分行政责任。挪威数据保护机构为了防控个性化教育中的学生隐私风险，以智能教育辅导和学习分析系统作为沙盒监管对象，在沙盒中开展了隐私风险的识别分析，进而提出监管建议，如对不同学生群体采用分层披露的方式以增强透明度，明确处理学生个人数据的法律依据等^[16]。2022年，法国国家信息与自由委员会支持了4个智能教育监管沙盒项目，入盒应用包括技能评估智能系统、家校沟通管理系统、学习轨迹分析系统等，监管部门主

要针对智能教育产品部署的合法性、相关主体对数据的控制权、学生数据共享的限制条件等提出建议^[17]。

不同于欧盟“有限豁免”的制度设计，美国 and 英国采取了“临时调整”的监管沙盒模式，在沙盒的有限时空内，突破既有监管规则的束缚，为人工智能创新营造宽松的环境。英国科学、创新与技术部对沙盒环境下既有监管规则的调整范围进行了明确规定：为了加速人工智能发展，除了4类必须遵守的法律规则（包括消费者保护和条款、基本权利条款、工人保护条款和知识产权条款^[18]）外，可在沙盒测试环境中临时暂停或调整其他监管规则，以便企业和开发者能在真实条件下测试新的智能教育产品。这一机制在保障公众健康、人身安全和知识产权等核心权益的前提下，为人工智能创新开辟了制度空间。特别是在智能教育领域，英国教育部表示，将对人工智能监管采取比欧盟要求标准更低的规范性立场，仅依托现有的学校数据保护政策和知识产权框架指南进行规范，没有采取欧盟《人工智能法》的强制性规定^[19]。此举为智能教育企业提供了更具适应性的测试与部署环境，在确保学生隐私与教育公平的同时，促进了智能教育产品的快速迭代与应用。美国人工智能监管政策的核心目标是维持美国人工智能产业在全球范围内的领先地位^[20]。尽管美国尚未在智能教育领域正式推行监管沙盒制度，但相关立法进程已显现出政策方向：2025年9月，美国参议院商务委员会主席特德·克鲁兹提出了《沙盒法案》^[21]，规定沙盒参与者可申请在一定时间内免于遵守可能阻碍其发展的法律法规，以快速推进人工智能技术应用。

3 对中国构建智能教育监管沙盒的启示与政策建议

中国的人工智能监管沙盒尚处于起步阶段，应用实践以自动驾驶和大模型领域的监管为主。北京、上海和深圳等多地建立了自动驾驶监管试点制度，北京依据《北京市自动驾驶汽车条例》，建立了“道路测试—示范应用—道路应用试点”的自动驾驶汽车上路三级试验机制；深圳依据《深圳经济特区智能网联汽车管理条例》，实行道路测试和示范应用申报管理制度等。这些实践均限制在一定的地域范

围内,依申请授权相关智能网联汽车进行测试。有学者分析,上述现象的根源在于中国的监管沙盒模式有待完善,在前端的许可、后端的出盒激励以及终端的责任制度等方面,可探索对智能网联汽车管制放宽的规定^[22]。为了平衡智能教育领域的监管效能与创新活力,综合智能教育风险、域外监管沙盒经验及中国监管需求,可在智能教育领域分场景、分阶段采取监管沙盒模式。

3.1 明确智能教育监管沙盒的适用对象

智能教育产品或服务并不必然产生不可控风险,因此并非所有的智能教育提供者都需要进行入盒测试,测试的适用对象应聚焦于与个人隐私、数据安全高度相关的智能教育产品。结合中国的人工智能伦理高风险清单和敏感个人信息相关法律规定,中国智能教育监管沙盒适用于以下3类对象:一是涉及学生人脸、声纹等生物特征信息的智能系统,如课堂监测、校园消费等智能产品;二是对学生的主观行为、心理情绪和生命健康等具有较强影响的智能系统,如学生个性化评测、心理健康识别等智能产品;三是与具身智能等物理空间相结合,具有高度自主性、可能颠覆人机关系,或者存在安全、人身风险的场景。尽管第三类场景尚未形成规模化应用,但技术能力快速迭代,可能导致人工智能发展出非预期的有害行为^[23],因此应纳入监管沙盒范围内进行前瞻考量。对于监管沙盒的适用时间,为了确保智能教育产品不会产生严重的不利影响,应将通过监管沙盒测试作为相关产品或服务上线的条件。但目前法律并未对智能教育产品或服务上线做出禁止性或限制性规定,因此沙盒主要适用于已上线的智能教育产品。其为企业提供自愿性的测试工具,帮助企业发现应用中的风险隐患并进行动态调整,增强应用的可信度。这一举措也将与地方政策相衔接,如《北京市教育领域人工智能应用工作方案》规定,要定期对智能教育产品进入学校后的应用流量、数据保护、大模型计算、敏感内容等进行动态监测,而监管沙盒可以为这一政策落实提供支持。

3.2 实现监管沙盒与现行法规体系的动态协同

目前,国家和地方相继出台法规政策,推动教育智能化发展,为智能教育监管沙盒提供了制度支持。一方面,完善监管沙盒与教育相关法律法规

的衔接机制。《个人信息保护法》《未成年人网络保护条例》《生成式人工智能服务管理暂行办法》等对智能教育相关的数据处理行为、服务提供者的权利义务做出了规定。监管沙盒的运作既要严格遵守法律法规,也要在受控环境中探索和验证法律法规的具体落实方式,如在使用大模型时应注重保障未成年人的个人信息安全等。对于部分无法确定是否符合监管制度的创新性较强的智能教育产品或服务,在全面上线应用之前,可通过监管沙盒提前识别风险。另一方面,完善监管沙盒与地方智能教育试点示范经验的衔接机制。越来越多的地区通过设立人工智能教育试验区、建设教育领域国家智能社会治理实验基地等方式探索人工智能在课堂教育、师生发展等方面的大规模应用,以提前识别技术应用可能带来的数据安全、技术适应、教育伦理等各类问题^[24]。例如,《浙江省推进“人工智能+教育”行动方案(2025—2029年)》明确提出,搭建大模型预训练和微调的“沙盒”。应充分发挥沙盒的技术创新测试和监管工具测试的双重作用,保障教育大模型输出可控、合规、适切。

3.3 形成多主体参与的监管沙盒组织体系

监管沙盒制度强调利益相关者的共参共治,监管部门应当与入盒企业沟通交流,共同制定监管规则。世界经济论坛与印度政府首席科学顾问办公室联合发布的《塑造智能时代的人工智能沙盒生态系统》报告指出,在教育沙盒中,教育科技公司、相关国家部门和国家教育研究与培训委员会的加入有利于保证解决方案与基层教育需求和政策路径保持一致^[25]。监管沙盒应为企业提供申请入盒、完备性审核、方案评估、训练验证、出盒、成果转化与宣传推广等全流程服务。为了弥合监管部门与企业监管认知与市场理解之间的信息差,建议引入大数据交易所、成果转化服务机构等第三方机构,为入盒企业提供人工智能训练数据合规治理、算法评估优化方面的专业服务。鉴于人工智能技术在教育领域的应用专业性较强、认知机制复杂,校方管理人员及教育学、心理学、社会学等跨学科专家也应积极参与其中,为智能教育产品或服务提供精准指导,形成“监管部门—校方—企业—第三方机构—外部专家”协同合作模式。以学生画像类智能教育产品为例,当产品提供者和学校不能确定该产品的

潜在风险时，产品提供者可申请入盒，向监管部门提交合规方案，如建立学生数据收集和保护机制、阐明产品运作机制、形成生成内容审核机制等，由监管部门研究审定通过后，在一定的时空范围内进行试点，并由第三方机构和外部专家对潜在风险进行评估。若产品通过评估，将由监管部门为其提供认证，证明其合规能力符合特定标准，以增强该产品进入市场的可信度。

3.4 构建规范和激励并重的监管沙盒规则

科学完善的监管规则体系是智能教育监管沙盒实现价值的核心保障，其设计水平的高低从根本上决定了沙盒机制能否有效平衡创新激励与风险防控。

在规范方面，监管沙盒的主要任务是明确法律法规的“底线”，并对企业能否达到“高线”伦理标准进行测试，确保测试过程可控、风险可承受、问题可追溯。以新加坡为例，人工智能验证基金会和资讯通信媒体发展局推出了全球人工智能保障沙盒，其以测试人工智能真实应用和优化测试方法为核心，确保企业能够在多个地区实施一致的测试框架和标准^[26]。中国的智能教育监管沙盒应与算法备案、大模型安全评估等制度相协调，在监管沙盒中不对前述制度的要求进行重复操作，主要在此基础上对智能教育相关算法的可解释性、智能教育领域的语料多样性、生成内容适当性等作出进一步规定。具体而言，要求企业在保护自身商业秘密的基础上，最大限度地公开所使用的语料情况、语料内容过滤方式、对个人信息的处理情况、对输出内容的审核及使用情况等，并对其来源、用途、目的、采取的技术方法等进行解释。另外，在沙盒中应支持开展智能教育产品的伦理审查，从实践角度划定智能教育服务中高风险活动的判定标准，明确相关智能教育应用可能对学生、教师产生的伦理风险并提出解决方案，利用真实场景数据持续优化伦理审查标准并调整高风险清单。

作为促进发展的创新监管工具，监管沙盒应强化激励机制，如鼓励企业参与试验性监管、采取高于合规要求的自律方式，从而将包容审慎的监管理念落到实处，为智能教育技术的迭代提供包容、弹性的空间。一是构建责任豁免规则。在企业已经充

分满足智能教育应用监管要求的情况下，如果发生不利后果，可根据后果严重程度、企业合规完备程度等因素，合理限定责任承担范围。一般仅豁免监管部门权限内的行政责任，企业仍应根据法律法规承担相应的民事责任和刑事责任。二是构建出盒后的认证规则。以企业出盒后监管部门的认证为企业合规能力提升提供背书，进而为智能教育领域规范化发展提供可复制的标杆。三是以监管沙盒测试结果作为下一步调整监管手段的依据。例如，韩国个人信息保护委员会在个人信息保护监管沙盒运行过程中发现，“要求送货机器人取得个人事先同意”不具备可操作性，于是对相关条款做出扩张解释，降低了相关公司的合规成本。将监管沙盒的成果融入国家监管政策和监管标准中，有助于形成有效的正反馈机制，切实解决人工智能发展的痛点问题。

监管沙盒作为一项新的监管工具，应用于智能教育领域，有利于促进智能教育应用创新，保障信息安全合规，但在实际应用中也面临一系列挑战，如企业自证合规的负担增加、沙盒的准入资质不明晰、监管部门需要对可能的损害事件和救济方案进行大量预设工作、监管效果不达预期等。未来，随着人工智能与教育领域融合程度的不断深化，监管沙盒将依托“实践—反馈—优化”的闭环机制，持续提升监管的适用性和前瞻性，促进智能技术潜力充分释放，助力教育质量大幅跃升。■

参考文献：

- [1] 中国政府网. 国务院关于印发新一代人工智能发展规划的通知 [EB/OL]. [2025-10-15]. https://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm.
- [2] UNESCO. UNESCO survey: Two-thirds of higher education institutions have or are developing guidance on AI use [EB/OL]. [2025-10-15]. <https://www.unesco.org/en/articles/unesco-survey-two-thirds-higher-education-institutions-have-or-are-developing-guidance-ai-use>.
- [3] 中国政府网. 国务院关于深入实施“人工智能+”行动的意见 [EB/OL]. [2025-10-15]. https://www.gov.cn/gongbao/2025/issue_12266/202509/content_7039598.html.
- [4] 法治日报. 儿童手表为何频现“毒答案”?[EB/OL]. [2025-10-16]. <http://epaper.legaldaily.com.cn/fzrb/>

- PDF/20240905/04.pdf.
- [5] 张旭, 郭梦清, 朱述承, 等. 大语言模型开放性生成文本中的职业性别偏见研究[J]. 数字人文, 2024(3): 20-37.
- [6] LIANG W, YUKSEKGONUL M, MAO Y, et al. GPT detectors are biased against non-native English writers[J]. Patterns, 2023, 4(7): 1-4.
- [7] 央视网. 公安网安部门依法对某人工智能服务科技有限公司予以行政处罚[EB/OL]. [2025-10-15]. <https://news.cctv.cn/2025/09/15/ARTIWIqvhOOVoLSlYcZGX1li250915.shtml>.
- [8] RIGLEY E, CHAPMAN A, EVERS C, et al. Anthropocentrism and environmental wellbeing in AI ethics standards: A scoping review and discussion [J]. AI, 2023, 4(4): 844-874.
- [9] ADAMS D, THOMPSON P. Transforming school leadership with artificial intelligence: applications, implications, and future directions[J]. Leadership and policy in schools, 2025, 24(1): 77-89.
- [10] Ministry of Education of Singapore. A guide to AIEd ethics [EB/OL]. [2025-09-16]. <https://file.go.gov.sg/aied-ethics-guide.pdf>.
- [11] 斯科特. 规制、治理与法律: 前沿问题研究[M]. 安永康, 译. 北京: 清华大学出版社, 2018: 116-120.
- [12] 杨丰一. 人工智能“监管沙盒”: 理论溯源、实践考察与制度构建[J]. 中国行政管理, 2024, 40(12): 25-35, 56.
- [13] 闫海, 王天依. 论颠覆性技术对规制的挑战及其应对[J]. 中国科技论坛, 2023(1): 58-68.
- [14] 刘盛. 监管沙盒的法理逻辑与制度展开[J]. 现代法学, 2021, 43(1): 115-127.
- [15] OECD. AI principles[EB/OL]. [2025-10-15]. <https://www.oecd.org/en/topics/ai-principles.html>.
- [16] Datatilsynet. The AVT project: Activity data for assessment and adaptation[EB/OL]. [2025-10-15]. <https://www.datatilsynet.no/en/regulations-and-tools/reports-on-specific-subjects/reports/the-avt-project-activity-data-for-assessment-and-adaptation/>.
- [17] CNIL. Bac à sable EdTech Les recommandations de la CNIL aux lauréats[EB/OL]. [2025-10-15]. <https://www.cnil.fr/en/digital-health-and-edtech-cnil-publishes-results-its-first-sandboxes>.
- [18] Government of the United Kingdom. New blueprint for AI regulation could speed up planning approvals, slash NHS waiting times, and drive growth and public trust[EB/OL]. [2025-10-15]. <https://www.gov.uk/government/news/new-blueprint-for-ai-regulation-could-speed-up-planning-approvals-slash-nhs-waiting-times-and-drive-growth-and-public-trust>.
- [19] Department for Education of the United Kingdom. Generative artificial intelligence (AI) in education[EB/OL]. [2025-10-15]. <https://www.gov.uk/government/publications/generative-artificial-intelligence-in-education>.
- [20] 周恒. 主要经济体人工智能全球治理诉求分化研究[J]. 全球科技经济瞭望, 2025, 40(7): 68-76.
- [21] U.S. Senate Committee on Commerce, Science, & Transportation. Sen. Cruz Unveils AI Policy Framework to Strengthen American AI Leadership[EB/OL]. [2025-10-15]. <https://www.commerce.senate.gov/2025/9/sen-cruz-unveils-ai-policy-framework-to-strengthen-american-ai-leadership>.
- [22] 邓文娟, 高圣平. 论智能网联汽车监管的沙盒模式: 兼评中国相关地方立法[J]. 江汉论坛, 2023(4): 125-128.
- [23] SHAH R, IRPAN A, TURNER A M, et al. An approach to technical AGI safety and security[EB/OL]. [2025-10-15]. https://storage.googleapis.com/deepmind-media/DeepMind.com/Blog/evaluating-potential-cybersecurity-threats-of-advanced-ai/An_Approach_to_Technical_AGI_Safety_Apr_2025.pdf.
- [24] 鲍婷婷, 柯清超, 马秀芳. 人工智能教育社会实验的理论基础与实践框架[J]. 电化教育研究, 2023, 44(1): 54-60.
- [25] World Economic Forum. Shaping the AI sandbox ecosystem for the intelligent age[R/OL]. [2025-10-15]. <https://www.weforum.org/publications/shaping-the-ai-sandbox-ecosystem-for-the-intelligent-age/>.
- [26] IMDA. Singapore launches new tools to help businesses protect data and deploy AI in a trusted ecosystem [EB/OL]. [2025-10-15]. <https://www.imda.gov.sg/resources/press-releases-factsheets-and-speeches/press-releases/2025/singapore-launches-new-tools-to-help-businesses-protect-data-and-deploy-ai-in-a-trusted-ecosystem>.

International Experience in Regulatory Sandboxes for Intelligent Education and Its Implications for China

LIU Xinyi

(Institute of Scientific and Technical Information of China, Beijing 100038)

Abstract: Currently, the rapid development of intelligent education applications provides robust support for building a learning society where everyone can learn, everywhere is a place for learning, and learning is accessible at any time. However, the rapid iteration and widespread adoption of these technologies have also given rise to risks concerning content generation, the protection of personal information rights, and human-computer relationships, exposing the limitations of traditional governance paradigms in mitigating these risks. This paper explores the application value of introducing the regulatory sandbox into the field of intelligent education. As a novel governance tool designed to balance technological innovation with risk prevention, the regulatory sandbox helps address critical challenges—such as weak supervision of data algorithms, risk spillovers, and ambiguous boundaries of rights and responsibilities between humans and machines—by establishing a controllable testing environment. Based on an analysis of the implementation pathways of intelligent education regulatory sandboxes in typical countries and in conjunction with China’s current legal and regulatory framework, this paper proposes focusing on three high-risk scenarios as the target for the application of the regulatory sandbox. It further suggests constructing sandbox rules that emphasize both safety regulation and incentive measures, thereby providing an operational institutional design for regulatory innovation in the field of intelligent education.

Keywords: intelligent education; regulatory sandbox; human-machine relationship

(上接第52页)

Analyzing Germany’s Practices for SMEs Development from a “Technology-Industry-Finance” Cycle Perspective and Their Implications for China

LIU Jing, GUO Wen

(China Center for Information Industry Development, Beijing 100081)

Abstract: Establishing an ecosystem for technology, industry, and finance is crucial to boosting economic vitality, driving innovation, and supporting high-quality development. Countries worldwide have developed distinct models tailored to their national characteristics during the process of building such ecosystems. This paper proposes the concept of the “Technology-Industry-Finance (TIF) Cycle”, clarifying the government’s important role in constructing this cycle. Meanwhile, taking Germany as an example and adopting methods such as literature review and empirical research, this paper analyzes the policies of the German government in building the TIF Cycle ecosystem, and summarizes the typical practices of Germany in constructing an application-oriented TIF Cycle ecosystem. Based on Germany’s development experience, this study further puts forward policy implications for promoting China’s TIF cycle and achieving leapfrog development of enterprises.

Keywords: industrial technological innovation; technology-industry-finance cycle; development of small and medium-sized enterprises