

数智赋能高等教育科学评价与人才质量提升： 困境、挑战与突围路径

吴建红

(杭州师范大学, 浙江 杭州 311121)

摘要: 教育评价是保障人才质量、引导高校发展的核心指挥棒, 其科学化转型关乎国家创新体系建设。当前, 我国高等教育评价体系存在评价导向失衡、维度单一和功能异化三大困境, 难以适应新发展格局下对拔尖创新人才的迫切需求。数智技术为构建多维度、全过程、增值性的科学评价体系提供了可能, 然其应用在教学落地、伦理规制与制度协同层面仍存挑战。本文旨在系统探讨上述困境与挑战, 确立“以人为本、技术向善”的核心价值观, 构建“融通、智能、安全”的技术体系, 并通过政策引导、能力提升与多元协同完善评价实施生态, 最终推动高等教育评价回归育人本质, 实现人才质量的有效提升。

关键词: 数智技术; 高等教育评价; 科学评价; 人才质量提升

基金项目: 杭州师范大学 2025 年研究生科研创新推进项目-Deepseek 赋能小学语文个性化教学设计的路径研究 (2025HSDYJSKY096)

DOI: doi.org/10.70693/rwsk.v2i4.303

一、前言

“评价体系和评价标准是中国特色、世界一流大学建设的指挥棒”^[1]。习近平总书记指出, 教育评价事关教育发展方向, 事关教育强国成败。^[2]这深刻揭示了教育评价在国家发展中的战略地位。高等教育作为科技第一生产力、人才第一资源、创新第一动力的重要结合点, 其评价体系的科学性, 直接关系到拔尖创新人才的培养质量, 进而影响国家在全球竞争中的核心竞争力。

然而, 我国传统高等教育评价体系在取得历史性成就的同时, 也逐渐暴露出其与新时代发展要求不相适应的深层矛盾。“唯论文、唯帽子、唯职称、唯学历、唯奖项”的“五唯”倾向, 成为制约高等教育内涵式发展的顽瘴痼疾^[3]。这种评价体系难以全面、客观、动态地反映高校的办学成效、学科的真实水平、教师的教学贡献和学生的综合素养, 导致人才培养与国家战略需求之间存在一定程度的错位。

数智技术的迅猛发展, 为破解这一难题提供了历史性机遇。它使得对学习过程的全程记录、对能力素养的多维刻画、对成长轨迹的增值评价成为可能, 为实现从“知识本位”向“能力本位”和“素养本位”的评价转型提供了强大技术支撑。但是, 技术赋能并非一蹴而就的坦途。其应用过程本身也伴随着数据孤岛、算法偏见、伦理失范、制度惯性等复杂挑战, 是一场涉及价值重塑、技术攻坚与制度创新的系统性工程。

为实现真正意义上的评价转型, 本文遵循“困境剖析-挑战识别-路径构建”的逻辑脉络, 为确保数智技术真正服务于高等教育评价的科学化转型和人才质量的全面提升的, 必须确立正确的核心价值观、构建完善的技术体系, 并通过多方协同完善实施生态, 方能履行数智技术赋能高等教育现代化、助力人才强国使命。

二、传统高等教育评价体系的时代困境

传统高等教育评价体系所面临的深层困境, 并非简单的技术或方法问题, 而是深植于其价值导向、设计逻辑与系统功能之中的结构性矛盾。这一矛盾并非孤立存在, 而是呈现为一个环环相扣的因果链条。因此, 系统剖析这一矛盾综合体, 是构建新时代科学评价体系不可逾越的逻辑起点。

(一) 评价导向失衡: 知识灌输与创新能力培养的断裂

传统教育评价体系使得教育方式和人才培养模式侧重于对知识灌输。该体系在很大程度上仍沿袭工业化时代

作者简介: 吴建红(2001—), 女, 硕士研究生在读, 研究方向为基础教育。

通讯作者: 吴建红

的逻辑，侧重于对显性、确定性知识传授效率和易于量化的学术成果产出的考核。这种导向在高校中具体表现为“重科研轻教学”、“重知识灌输轻思维激发”的普遍倾向。然而，在创新驱动发展的数智时代，社会对高等教育的人才培养需求已从“知识继承者”转向“创新开拓者”。在蕴含“知识灌输”意味的教育评价体系之下，高校教师的教学内容易于聚焦于知识点的灌输，学生的学术训练则可能蜕变为对论文发表技巧的追求，逐渐忽视批判性思维与复杂问题解决能力的培养，这会大大削弱我国创新型人才储备队伍力量。

受到“分数”与绩效长期挂钩的教学评价体系影响，其结果是，高等教育在某种程度上培养了大量擅长应试和完成规定动作的“优秀考生”或“标准学者”，却难以培育出能够面对不确定未来、勇于进行原始创新的“创新型人才”。具体而言，这种断裂表现在：在认知维度上，评价侧重于对已有知识的掌握和再现，普遍忽视了对学科前沿的洞察、跨学科整合以及创新构想等高阶思维的评估与引导。在能力维度上，它难以衡量在真实科研与产业实践中至关重要的协作攻关能力、实验设计能力及创业精神。这一断裂的本质，是评价的“指挥棒”效应所强化的目标异化。教师的教学内容聚焦于考点，学生的学习活动蜕变为对知识点的记忆和答题技巧的反复演练。最终，这种失衡的导向导致高等教育产出与国家战略需求错位，构成了传统评价体系在价值逻辑上的核心困境。

（二）评价维度单一：静态分数与动态综合素养的脱节

习近平总书记提出，教育评价需要破“五唯”，“坚决克服唯分数、唯升学、唯文凭、唯论文、唯帽子的顽瘴痼疾，从根本上解决教育评价指挥棒问题”^[4]。在高等教育领域，“五唯”集中表现为过度依赖一系列静态、易量化的终结性指标。例如，用GPA（平均学分绩点）衡量学生整体学习水平，用SCI/SSCI论文数量、影响因子代表教师的学术水平，用国家级人才头衔、科研项目经费来评判学科实力。这种评价方式的根本缺陷在于，它将复杂、多维、动态演进的教育过程和人的发展，简化、压缩为少数几个冰冷的数字。这种评价方式与当代高等教育所倡导的“促进学生综合素养全面发展”和“引导高校特色发展”这一动态、多元的目标之间，产生了显著脱节。

对大学生而言，单一的分数和论文指标无法反映其品德修养、团队协作、社会实践、心理健康以及职业适应能力等关键素养。对高校教师而言，唯论文、唯项目的评价难以全面衡量其教学水平、学术指导贡献以及对校园文化的积极影响。这种“脱节”所带来的后果是严重的。它首先导致高等教育实践的窄化，形成“评什么就教什么、研什么”的功利主义倾向，使得那些无法被简单量化的通识教育、课程创新、社会服务等价值被边缘化。其次，它对师生个体价值的评判变得片面且不公，抑制了其个性化、多样化发展的空间与内在动力。最终，教育评价本应承担的“发展性功能”被其“管理性功能”所淹没，体系自身也因此陷入了无法真实反映高校育人成效和人才真实质量的困境之中。

（三）评价功能异化：选拔甄别与促进学生发展的背离

传统教育评价体系最深层的时代困境，在于其核心功能的系统性异化。从应然层面看，教育评价的根本宗旨应服务于学生发展，其理想功能是“以评促学、以评促教”，即通过持续的诊断、反馈与激励，优化教学进程并激发学生的内在学习动力。然而在实然层面，该体系在高压力的社会竞争语境下，已异化为一个以分等、筛选和淘汰为核心的功利性选拔工具，导致了手段与目的的严重背离，评价功能发生严重异化，进而对教育生态和个体成长产生了深刻的负面影响。

这种功能异化的发生机制源于高校评价结果与稀缺社会资源分配的强关联。当分数成为升学、就业等关键人生通道的几乎唯一“通行证”时，评价便不再仅仅是教学过程中的一个环节，而是演变成了教育竞争的终极目标本身。这一转变彻底颠倒了目的与手段的关系：学习不再是为了人的全面发展，而是为了在评价中获胜。

于学生个体而言，这种评价功能异化压抑了其学习的内在动机。对外部奖励和避免惩罚的过度追逐，侵蚀了由好奇心、探索欲和求知本身带来的愉悦感，导致学习焦虑普遍滋生，个性化发展空间被统一的应试标准所挤压。于教育整体生态而言，它引发了一系列连锁反应：课程内容趋于狭窄和僵化，教学方式强调机械训练而非理解创造，师生关系亦被简化的“教—考”绩效契约所主导。最终，教育评价从其促进发展的“仆人”，异化为支配一切教育活动的“主人”，背离了其以人为本的初衷，构成了传统体系难以自我超越的结构性困境。

三、数智技术赋能高等教育科学评价的现实挑战

数智技术赋能高等教育评价的转型，绝非简单的技术嫁接，而是一场触及深层学术治理结构的范式革命。但其赋能过程远非技术叠加的线性进程，而是在与学术治理结构深度互嵌中引发了一系列复杂挑战。这些技术、伦理与制度层面的挑战相互交织，共同构成了数智赋能需要突破的多重障碍。

（一）技术层面的应用挑战

尽管数智技术为教育评价转型描绘了宏伟蓝图，但其在实际应用层面仍面临诸多严峻的技术性挑战。这些挑战若不能得到有效破解，将成为阻碍评价改革落地的现实壁垒，甚至可能引发新的教育风险。

首先，是“数据孤岛”问题。技术是工具而非目的，需要“科学规范技术应用在教学评价中的适用边界，合

理赋值教学评价指标”^[5]理想中的全景式评价依赖于对学生学习全过程数据的全面采集与融合。然而现实是，数据通常散落、封闭于不同的系统和平台中，众多高校内部的教务系统、科研管理系统往往一分为二，而这就会导致数据之间出现割裂。这些数据格式不一、标准各异，形成了严重的“数据孤岛”。

其次，是算法本身的局限性。主要体现在模型的透明性、可解释性与可能存在的偏见。人工智能算法，特别是复杂的深度学习模型，常被视为“黑箱”，其内部决策机制难以被常人理解。当AI用于评价诸如研究生招生筛选、奖学金评定算法等复杂素养时，其给出的评分或诊断结果若无法被教师和学生所理解和信任，其教育价值将大打折扣。更甚者可能会放大教育不公，形成技术赋能下的“新歧视”，这与教育公平的初衷背道而驰。

最后，是普遍存在的技术门槛。这包括基础设施不均衡与师生数字素养不足。一方面，不同层次学校之间的信息化基础设施存在显著差距，部分双一流高校与地方应用型高校中之间可能会由于地区、资源的差异而形成天然的“数字鸿沟”。另一方面，师生的数字素养是决定技术能否“用好”的关键。高校是否开设数字素养课程的训练、对数字课程的重视程度也会导致师生数字素养存在不平衡的结果。若缺乏相应的培训与准备，再先进的技术也无法转化为有效的教育生产力，甚至可能因使用不当而产生反效果。

（二）伦理与价值层面的潜在风险

在处理教育大数据与人的相互关系时应遵循“有用”“无害”的伦理诉求^[6]，教育评价涉及到对人进行价值判断，这无疑会带来一定的伦理问题，若缺乏审慎的规制与人文关怀，技术理性可能异化为一种压迫性力量，侵蚀教育的本质。

数智评价和人之间的首要风险在于数据隐私的边界模糊。为实现全过程、多维度的评价，技术系统需持续采集海量学生数据，从学习行为、社交互动到甚至生物信息，几乎构建了一个数字化的“全景监狱”^[7]。这就由技术引发了严峻的隐私安全与数据所有权之问，若缺乏严格的法律法规与伦理审查，这种全面的数据监控不仅可能侵犯学生隐私，更可能使学生从“教育主体”异化为被观测、被利用的“数据客体”。

最深层的风险在于评价的异化，即技术理性对教育人文价值的侵蚀。当技术的使用逐渐忽视人伦理性，一切教育现象都被量化为数据时，教育中那些难以量化的珍贵品质——如好奇心、同理心、批判精神、道德勇气——也将面临被忽视和边缘化的命运。作为教育评价的核心评价主体，教师评价素养对于评价结果的质量取向、评价改革的进程深入有着深远的影响^[8]。教师的教育直觉、艺术性和情感关怀被冰冷的算法评分所替代，教学不再是充满创造性的师生互动，过度依赖数字技术也会使教师产生“教学惰性”，一切教学以数字技术给定的内容为准，教学过程进而丧失其最根本的育人温度与人文精神。

总之，伦理与价值风险是数智化评价转型中最深刻、最致命的挑战。它警示我们，必须将伦理规制置于技术应用之上，确保技术始终服务于人的全面发展，而非让人屈从于技术的逻辑。

（三）实践与制度层面的实施困境

数智技术赋能高等教育评价的最终落地，深受实践环境与制度安排的制约。当前，在从传统向现代转型的进程中，实践与制度层面存在着深层次的实施困境，存在“最后一公里”难题。

首要困境是强大的路径依赖与学术治理惯性的阻力。数十年来形成的世界一流大学以“五唯”指标为核心的评价体系，已深度嵌入高校管理的肌理。许多管理者习惯于依赖简洁的量化指标进行决策；教师担忧数智化评价的复杂性和额外工作量，且对其能否公正评价教学学术等软性贡献心存疑虑。这种强大的惯性使得新评价体系难以有效推行。

其次，科学、统一的标准与规范的普遍缺失。目前，对于如何利用学习分析进行增值评价、相关教育数据的产权归属与使用边界等关键问题，尚无权威的国家标准或行业共识。这导致各高校、各技术企业的探索“各自为战”，可能催生出一批缺乏教育测量学验证的“黑箱”产品，损害评价的公正性与权威性。

最后，多元主体间的协同不足与治理机制缺位。大学的信息化建设不是一个封闭系统，其创新发展与国家、政府的制度环境息息相关^[8]。同样的，基于信息化的数智化评价也是一项需政府、高校、技术企业、学术专业组织多元共治的系统工程。然而目前，政府在顶层设计、标准制定与监管问责上的角色有待强化，高校作为需求方和应用主体能力不均，企业在商业利益与教育公益之间需找平衡点。各方权责利模糊，难以形成推进科学评价的合力。

四、数智技术赋能高等教育评价转型的突围路径

面对传统评价体系的深层困境与数智赋能过程中的复杂挑战，高等教育评价改革必须超越局部优化的思路，转向价值重塑、技术攻坚与生态重构协同推进的系统性路径，才能最终实现高等教育评价回归育人本质的根本目标。

（一）确立“以人为本、技术向善”的评价价值观

理念是行动的先导。面对数智技术带来的机遇与挑战，首要的突围路径在于进行价值重构与理念引领，必须

旗帜鲜明地确立“以人为本、技术向善”的评价核心价值观。这一理念与联合国教科文组织强调的“以人民为中心”的人工智能发展原则高度契合^[9]，即教育领域的人工智能必须服务于增进人类能力和权益的根本目标。此核心价值观必须深刻融入高校治理理念和学术文化之中，确保技术发展始终服务于师生全面发展与高等教育质量提升这一根本目的。

技术的应用必须被约束在伦理的框架之内。为确保数据安全与算法公正，高校应率先建立校级“数智评价伦理审查委员会”。委员会成员应包括教育专家、伦理学者、技术专家、法律工作者、教师及学生代表，对重大数据应用项目进行前置性伦理评估，严格遵循“目的明确、最小够用、知情同意、安全可控”的原则。其次，必须建立常态化的算法审计与影响评估机制。要求对核心评价算法的设计逻辑、训练数据构成及决策结果进行定期审查与评估，检测并消除其中可能存在的偏见，确保算法的公平性与正义性。^[10]

技术工具的设计理念直接决定了其应用的导向。必须大力倡导并推动包容性设计，开发公平、透明、可解释的智能评价系统。“公平”要求工具的设计充分考虑不同学科、不同类型高校、不同群体师生=的适用性。“透明”与“可解释性”则要求评价系统能提供清晰的解释性反馈，这能将技术从“评判者”转变为“赋能者”，服务于师生的发展。

（二）构建“融通、智能、安全”的评价技术体系

先进的理念与价值观需要依托坚实的技术基础方能落地。推动数智技术赋能教育评价转型，当务之急是攻克核心技术难题，着力构建一个集“融通性、智能性、安全性^[11]于一体的新一代数字教育评价体系。数据是数智化评价的基石，推动高校建设校级“教育数据中台”或“人才培养数据中心”，制定校内统一的数据标准，强制要求各类业务系统依标对接。实现多源异构数据的融合治理。这是实现全景式、过程性评价的底层基础

马克思指出，“吸引人的劳动，成为个人的自我实现”。^[12]在教育数字化飞速发展的当下，数据分析、人工智能等技术因其可提高评价过程的效率与准确性，使评价结果更加客观、公正，并能为教育决策提供强有力的数据支持^[13]，有研究者就从学校特色理念和方向的適切度、特色课程和管理的一致度、师生和学校发展的优质度三个维度，构建了学校特色发展的测评体系。

算力与算法是数智化评价的核心引擎。未来的技术攻坚方向必须从单纯的知识点评估转向对核心素养的精准刻画。应汇聚教育测量专家、学科专家与人工智能科学家，共同研发基于多模态数据的下一代评价模型，从而将评价内容真正引导至素质教育所倡导的核心素养轨道上来。

安全与信任是技术体系得以可持续运行的保障。在数据融通与应用的同时，必须将隐私保护技术深度嵌入系统架构的底层。如差分隐私、联邦学习等，在保障数据安全和个人隐私的前提下，最大化地发挥数据价值，建立师生对数智化评价的信任。为了进一步提升协作学习的成效，研究者还提出了多元化的协作策略，如同伴对话反馈策略^[14]、在线学习支架策略等^[15]。通过这些技术手段，在最大化发挥数据价值的同时，最小化隐私泄露风险，筑牢教育数据安全的技术防线。

（三）完善“协同、渐进、规范”的评价实施生态

数智技术赋能教育评价的最终成功，不仅依赖于技术体系的构建，更在于其能否在一个健康、可持续的生态中落地生根。这要求从政策、人与治理三个维度协同发力，构建一个“协同、渐进、规范”的实施环境。

强有力的顶层设计与政策引导是避免市场混乱、确保改革方向的关键。国家及教育主管部门应牵头制定数智化评价的课程标准、数据标准、质量标准和伦理标准，明确“评价什么、怎么评价、为了什么评价”。数据如何管理以及算法需遵循的伦理底线，为各类实践提供清晰、统一的规范框架。同时，应采取“试点先行、典型引路”的渐进式推广策略。选择有条件的地区和学校开展示范项目，鼓励其在保障安全的前提下探索多种模式，最终实现由点及面、稳步推进的健康发展，有效降低全面铺开可能带来的系统性风险。

高校教师与管理人员的评价素养与数据素养是决定改革成败的核心，其专业能力是决定技术能否“用好”的核心，因此，我们必须对其进行系统的数据素养与评价素养培训。评价素养培训则需帮助教师掌握数智化评价的新理念、新工具与新方法，理解其优势与局限，学会在人机协同中发挥主导作用——技术负责处理海量数据和常规分析，教师则专注于技术难以衡量的情感关怀、创造力激发和价值观引领等高级育人工作。

数智化评价跨越教育、技术、法律等多个领域，绝非任何单一主体能够独立承担，必须构建一个多元共治的新格局。政府应承担起宏观规划，统筹管理的“元治理”角色；高校作为主阵地，负责具体实施、日常管理以及营造健康的评价文化；技术企业则需恪守伦理底线，提供安全、可靠、透明的技术与产品，并建立良好的售后服务和迭代升级机制；学术专业组织应参与标准制定与质量认证。通过建立常态化的沟通协商机制，明确各方的权责利边界，方能形成推动评价改革持续深化的强大合力。确保这场数智时代高等教育评价的深刻变革能够真正落地见效，惠及每一个高校学生。

参考文献:

- [1]周光礼.破“五唯”立新标:建构中国特色哲学社会科学评价体系[J].中国人民大学学报,2022,36(03):15-18.
- [2]习近平:扎实推动教育强国建设[EB/OL].(2023-09-15).https://www.gov.cn/yaowen/liebiao/202309/content_6904156.htm.
- [3]张炜.高校人才评价改革如何突破“五唯”困局[J].高等工程教育研究,2020(3):5-11.
- [4]习近平在全国教育大会上强调:坚持中国特色社会主义教育发展道路培养德智体美劳全面发展的社会主义建设者和接班人[N].人民日报,2018-09-11(1)
- [5]于孟任,何晓芳,牛姝颖.器道之间:论大学教学新技术应用的工具理性与价值理性[J].当代教育论坛,2024(6):68-75.
- [6]戚万学,等.教育大数据的伦理诉求及其实现[J].教育研究,2019,(7)
- [7]韩炳哲.精神政治学[M].关玉红,译.北京:中信出版社,2019:49-53
- [8]刘永贵,郑旭东.专业协会在大学创新发展中的独特作用——美国高等教育信息化的成功之道探秘[J].复旦教育论坛,2009,(01):21-24+29.
- [9]UNESCO. (2021). AI and education: Guidance for policy-makers. Paris: UNESCO.
- [10]Selwyn, N. (2021). Digital Ethics. Bloomsbury Academic.
- [11]范涌峰,宋乃庆.学校特色发展测评模型构建研究[J].华东师范大学学报(教育科学版),2018(2):68-78.
- [12]马克思,恩格斯.马克思恩格斯全集:第三十卷[M].北京:人民出版社,1995:616.
- [13]倪娟.高质量基础教育评价体系的内涵、重点及实现策略[J].人民教育2024(10):10-14.
- [14]张缨斌,陈孝然,胡小勇.协作学习中同伴对话反馈策略的异质效应研究[J].电化教育研究,2025,(5):86-93、102.
- [15]李梅,葛文双.基于项目的在线协作学习支架策略探究[J].现代远距离教育,2021(1):40-47.

Empowering Higher Education with Digital Intelligence for Scientific Evaluation and Talent Quality Improvement: Dilemmas, Challenges, and Breakthrough

Paths

JianHong Wu

(Jing Hengyi school of Education, HangZhou Normal University name, HangZhou, China)

Abstract: Educational evaluation is the core guide to ensuring talent quality and directing the development of higher education institutions, and its scientific transformation is crucial to the construction of a national innovation system. Currently, China's higher education evaluation system faces three major challenges: imbalanced evaluation orientation, single-dimensional criteria, and functional distortion, making it difficult to meet the urgent demand for top-notch innovative talents under the new development framework. Digital intelligence technologies make it possible to build a scientific evaluation system that is multidimensional, full-process, and value-added; however, their application still faces challenges in terms of technological implementation, ethical regulation, and institutional coordination. This paper aims to systematically explore the above challenges, establish the core values of people-oriented, technology-for-good, construct a technological system that is integrative, intelligent, and secure, and improve the evaluation implementation ecosystem through policy guidance, capacity building, and multi-stakeholder coordination, ultimately promoting a return to the educational essence of higher education evaluation and effectively enhancing talent quality.

Keywords: Digital Intelligence Technology; Higher Education Evaluation; Scientific evaluation; Improvement of Talent Quality