

# 智能技术赋能教育评价： 内涵、总体框架与实践路径\*

刘邦奇<sup>1,2</sup>, 袁婷婷<sup>1</sup>, 纪玉超<sup>1</sup>, 刘碧莹<sup>1</sup>, 李 岭<sup>1</sup>

(1.讯飞教育技术研究院, 安徽 合肥 230088; 2.西北师范大学 教育技术学院, 甘肃 兰州 730071)

**摘要:** 智能技术赋能为教育评价变革发展提供了重要机遇, 促进智能技术与教育评价融合创新是深化新时代教育评价改革的必然选择。据此, 对智能技术赋能教育评价进行了系统研究, 其历史渊源可追溯到管理工程领域的智能综合评价和教育评价领域计算机技术的应用。本质是基于智能技术对传统教育评价的突破与创新, 通过解构、重构形成新的教育评价模式, 具有科学化、多元化、立体化、最优化、精准化等主要特征。助力教育评价的智能技术及应用由“5+1”的总体框架(基础层、技术层、平台层、应用层、用户层5个层次结构和1个保障体系)构成。通过智能技术赋能“四个评价”的具体场景应用, 采取理念引领、标准规范、主体关照、数据驱动、专业支持等整体推进策略, 打造智能化教育评价生态体系。

**关键词:** 智能技术; 赋能; 教育评价; 总体框架; 实践路径

**中图分类号:** G434      **文献标识码:** A

## 一、引言

智能技术赋能教育评价提出的缘由, 可以追溯到两种历史渊源。一种是在管理工程领域, 为解决复杂系统问题, 将智能化方法用于综合评价, 如引入遗传算法、人工神经网络、蒙特卡罗模拟方法、人工免疫算法和蚁群智能算法等, 提高了综合评价的智能特性。由此出现了智能化综合评价<sup>[1]</sup>。现代教育评价具有多元性、复杂性、系统性等特征, 是多种学科专业、多种技术方法在教育评价领域的综合应用, 管理工程领域形成的智能化综合评价为创新教育评价模式提供了参考路径。有学者梳理了国际上“智能化教育评价”研究的演进路径, 析出“智能化教育评价”演进过程中的关键节点文献<sup>[2]</sup>, 其中3篇“智能化教育评价”文献综合了教育学、心理学、统计学、神经计算科学、信息科学等学科和技术应用。

另一种是在教育评价领域, 起源于计算机技术在教育评价中的运用。近二十多年来, 美国先后发布了六个国家教育技术计划(National Education

Technology Plan, 简称NETP), NETP1996在学生学习进度评价中提出应用多媒体技术和计算机技术, 形成学生作业的电子档案和进行计算机自适应测试<sup>[3]</sup>。NETP2000指出随着互联网的发展, 改善传统教育评价的机会出现了, 要开发新的学生评价工具, 传统的纸笔测评不能再准确反映21世纪学生的学习情况, 教育评价可以使用在线技术以更有效的方式进行<sup>[4]</sup>。NETP2004提出要建设集成数据系统, 指出集成、可互操作的数据系统是在线学生表现测评的关键, 测评结果可驱动日常教学决策和设计教学干预, 帮助教师进行差异化教学<sup>[5]</sup>。NETP2010、NETP2016、NETP2017均将教育评价作为教育技术应用的重要领域进行规划, 强调各级教育系统要利用技术去衡量什么最重要, 并利用评价数据来持续改进和创新教育<sup>[6-8]</sup>。在我国, 计算机技术最早运用于考试评价、教育测量分析等, 典型应用包括计算机化考试、网上评卷、网上录取等。同时, 现代信息技术也支持和发展了教学评价, 如电子档案、电子表决器、教学评价支持系统等技术应用, 有效提升了教育教学评价的信息化、智能化水平。

\* 本文系认知智能国家重点实验室2020年度智能教育开放课题重点课题“智能技术支持下的因材施教与教育治理”(项目编号: iED2020-Z003)阶段性研究成果。

智能技术赋能教育评价的正式提出,是近几年随着人工智能、大数据等智能技术的迅速发展及其在教育领域广泛应用,催生了智能教育新模式、新形态,出现了包括智能考试、智能评价等新样式<sup>[9]</sup>。随着智能时代的到来,智能技术为教育评价变革带来了机遇,深度融合智能技术与教育评价是必然趋势。早在2016年,杨现民提出大数据的发展使教育评价走向客观性评价、伴随性评价、综合性评价和智能化评价,教育大数据驱动教育评价体系重构<sup>[10]</sup>。董奇在2020年“人工智能与教育大数据峰会”上指出,科技赋能教育评价改革时机已至<sup>[11]</sup>。总之,教育评价的改革发展需要先进和前沿的技术支撑,智能技术赋能为此提供了重要的方向和路径。加强教育评价领域关键技术的研发力度,促进智能技术与教育评价的深度融合,是深化教育评价改革面临的重要议题。

## 二、智能技术赋能教育评价的内涵

智能时代技术为教育评价变革提供了支撑引领,智能技术赋能教育评价正走向教育评价改革的主战场,但目前业界对智能技术赋能教育评价的内涵并没有清晰的界定。有学者从技术视角提出如何赋能教育评价,Zhang Qingchen等人指出,在智能化教育评价中通过传感器网络和通信技术等实现大数据的收集,通过深度学习解决在图像分析和语音识别等方面的问题<sup>[12]</sup>。Wiley和Hastings等人指出可通过机器学习和自然语言理解技术评价学生的想法和概念的运用,评价结果具有较高的准确度<sup>[13]</sup>。也有学者从技术驱动教育评价手段变革的视角给出解释:通过利用技术的手段和方法,系统、科学、全面地采集、处理、分析各类教育数据,对教育活动做出客观判断的过程,实现了从经验主义走向数据主义,从模糊走向精准<sup>[14]</sup>;利用技术手段,有效采集和整合学生的各项数据,将教师评价、自我评价等多种评价数据融合,从而对学生进行多维、全面、深入的评价,将碎片化评价转化为系统化评价<sup>[15]</sup>。还有学者从传统教育评价向技术驱动教育评价观念转变的视角给出解释:技术使教育评价从“关注结果”转为“关注过程”,从“单一的考试评价”转为“多维度的综合素质评价”<sup>[16]</sup>;以技术为支撑的教育评价将“传统教育评价”转变为“智能化的教育评价”,从而真正走向“智慧评价”<sup>[17]</sup>。

综合国内外学者对智能技术运用于教育评价的理解,可以看到智能技术赋能教育评价将是未来发展的必然趋势。根据以上分析,可以认为“智能技

术赋能教育评价”是在走向智能时代背景下,充分利用人工智能、大数据等智能技术的优势,对传统教育评价进行革新与发展,改进教育评价过程与方法,提升教育评价的数字化、智能化水平,实现科学、客观和高效的评价与反馈,促进教育事业的改革与发展。

智能技术赋能教育评价本质是对传统教育评价的突破和创新,通过解构、重构形成新的有别于传统的教育评价新模式。而其实施的关键在于用智能技术助力教育评价数据的采集、处理、分析和运用,实现教育评价的数据化、智能化,由此而引发教育评价形态的变革。具体来说,充分利用人工智能、大数据、云计算、区块链等先进技术,实现对全过程、全方位教育评价数据的采集,进行深度挖掘分析和反馈应用,对教育教学过程、结果进行多元综合评价,为教育教学改进提供全面、有效的决策依据,实现了教育评价的现代化、专业化,同时由于智能化评价工具和手段的应用,教育评价过程也更加智能、高效。

## 三、智能技术赋能教育评价的主要特征

### (一)评价模型科学化

科学的评价模型是教育评价的核心。借助智能技术手段,通过教育评价人员、信息技术专业人员、教育教学人员等协作,针对不同评价对象和评价内容构建科学的指标体系、指标权重、评价模型。评价指标的设计是评价模型构建的前提条件,是保证评价模型科学化的重要支撑,指标构建方法有质化方法、量化方法和复合性方法等,其中指标权重设计的合理性既能够反映出决策的主观价值,又可以获得客观准确的测量结果。通过计算机软件系统、人工智能中的推理技术等构建评价模型,实现模拟评价,具体方法有专家系统、人工神经网络、机器学习等。

### (二)主体参与多元化

参与评价的主体由教师、家长、同伴、自我、评价专家等共同参与,形成“评价共同体”,使评价过程呈现民主化和人性化,评价结果也更具有真实性和可靠性。教师参与能够给出更加具有专业性、实效性的评价信息;家长参与使评价结果不仅聚焦于在校情况,还包含家庭表现情况,使评价结果更加全面;同伴参与可有效调动学习者的积极性,加强沟通交流;自我评价使被评价者的主体地位得到充分发挥,有效提升参与意识,主动反思发现自身的不足;评价专家为评价对象提供更精准、客观、全面的评价结果。

### (三)数据获取立体化

通过物联感知技术、可穿戴设备技术、视频监控技术、网评网阅技术等对评价数据进行全过程、全方位、多维度的采集,改变过去人工采集记录的方式,实现评价数据立体化获取。全过程是指依托数据采集平台和设备自动记录评价对象在整个活动中产生的各项数据,由过去的“间断性记录”转变为“全过程记录”;全方位强调数据的获取打破时空界限,不仅局限于传统教室,还包含线上学习数据的获取、户外教学活动数据的获取等;多维度是指采集的数据种类会更多样、更全面,包括行为数据、情感数据、体质数据、管理数据等。

(四)诊断分析最优化

通过采用数据融合、数据分析等技术,对多模态数据进行诊断分析,实现了多维、全局数据处理和分析的最优化,从而达成精准评估和测评。例如,在海量的多模态数据挖掘中,基于不同模态数据融合,可通过模态数据间的互补学习提取出复杂数据中的有效特征,从而提升了决策结果的准确性;使用机器学习等算法对不同种类的数据进行分析,包括文本分析、语音分析、图像分析、视频分析等,可准确表征评价对象的特征要素。

(五)评价反馈精准化

评价反馈是实现教育评价的应用价值体现,也是评价的重要组成部分。通过高度个性化定制、智能推荐引擎等技术,将评价结果以交互式可视化的形式及时精准地推送给用户,有效提升评价对象对自我的认知,使评价对象即时调整学习策略、教学目标等,进而有效地促进管理、教学、学习等。评价反馈贯穿整个教育活动的始终,有过程性的即时反馈,如课前的预习测评与反馈、课堂的实时检测反馈等,能够精准反馈评价结果即时调整学习策略;有根据评价结果精准推送教学和学习资源等,实现教师精准的教和学生个性化的学。

四、助力教育评价的智能技术及应用框架

智能技术赋能教育评价不同于传统的教育评价,它需要利用智能技术构建一个多层次、多用户、立体开放的教育评价生态体系。推进智能技术赋能教育评价的建设与应用,需要从整体上把握智能技术背景下技术、业务、应用和用户间的关系以及对评价产生的影响。因此,借助信息化顶层设计的方法论,在笔者团队提出的智能教育关键技术平台参考框架的基础上<sup>[18]</sup>,结合智能技术赋能教育评价活动的实现过程,可以提出和构建助力教育评价的智能技术及应用总体框架,如图1所示。

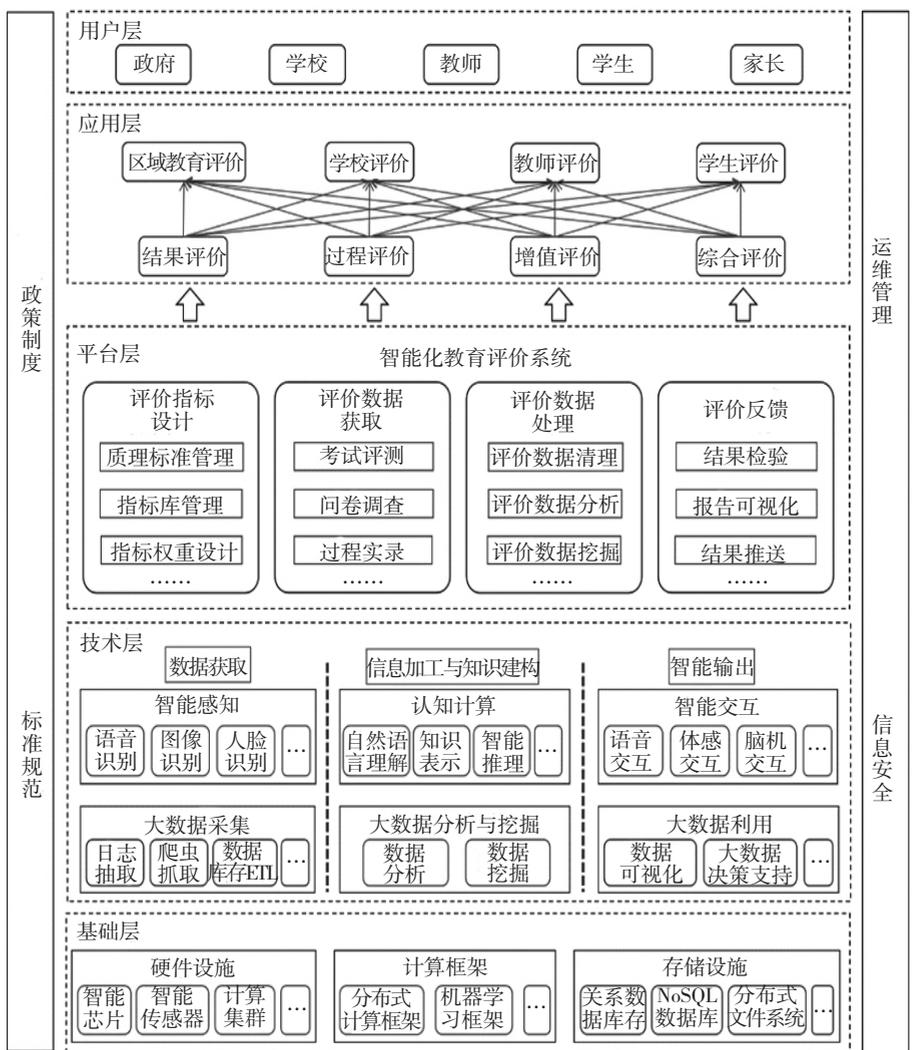


图1 助力教育评价的智能技术及应用总体框架

助力教育评价的智能技术及应用由“5+1”的总体框架构成,形成了科学化、多元化、立体化、最优化、精准化的智能教育评价生态体系,具体包

括基础层、技术层、业务层、应用层、用户层5个层次结构和1个保障体系。

#### (一)基础层：技术赋能教育评价的基础支撑

基础层由硬件设施、计算框架、存储设施等构成，为智能技术赋能教育评价提供了必备的基础设施条件，是智能化教育评价体系的基础保证。(1)硬件设施为整个教育评价过程提供硬件支撑环境，在数据采集集中提供智能传感器、智能穿戴设备等；在数据处理和分析中提供智能芯片、计算集群等；在评价结果反馈中提供具有交互功能的智能终端设备，如智能手机、智能PC等。(2)计算框架为系统提供统一的计算服务、统一的机器特征提取和算法模型建模服务，具体包含分布式计算、机器学习等。(3)存储设施为整个评价过程提供统一的数据存储、读写和管理等数据服务，保障数据的安全等，具体有关系数据库、NoSQL数据库、分布式文件系统等。

#### (二)技术层：实现技术赋能教育评价的核心

技术层是实现智能技术赋能教育评价的核心，通过利用各种智能技术实现数据的智能化采集、加工分析和可视化输出等。(1)在数据获取中，通过语音识别、图像识别、日志抽取、数据库ETL等技术，解决“听、说、看”等感知层面信息的转化，实现对原始数据源中数据的抽取。(2)在信息加工与知识构建中，通过自然语言处理、知识表示、数据分析、数据挖掘等技术实现对获取数据的统一处理和分析，挖掘数据潜在规律和价值。(3)在智能输出中，通过语音交互、体感交互、数据可视化等技术，将结果以可视化的形式反馈给用户。

#### (三)平台层：技术赋能教育评价的关键支撑

平台层是智能技术与教育评价深度融合的关键，其技术支撑体现在智能化教育评价系统的开发与运用，主要是通过人工智能、大数据等技术构建智能化的评价系统，为各类应用提供支撑平台。主要包括四类服务：(1)评价指标设计的设置是整个教育评价体系的关键，具有系统性、独立性、目标性等多重特点，提供质量标准管理、指标库管理、指标权重设计等服务。(2)评价数据获取主要是有效获取整个活动的各项数据，是数据处理分析的基础，具体方法有考试评测、问卷调查、过程实录等。(3)评价数据处理包含了对评价数据的清理、分析、挖掘等，能够有效处理缺失、错误、重复的数据，利用统计分析方法进行数据分析，用机器学习算法和计算机编程等实现数据的深度挖掘。(4)评价反馈是通过多种形式将评价结果反馈给不同用户，是有效教与学发生的必备条件，根据评价反馈结果个性化

推送相关资源，有效促进教与学等活动。

#### (四)应用层：技术赋能教育评价的价值体现

智能化教育评价支持系统的应用是各类教育评价场景应用智能技术的体现，为区域、学校、教师、学生等提供结果评价、过程评价、增值评价、综合评价。结果评价：利用智能技术优化智能阅卷与评测、智能考试分析等，注重对教学结果、学习成绩或成就等进行评定；过程评价：基于智能技术进行伴随式采集数据，对管理、教学、学习等行为进行密切跟踪，通过多模态诊断分析发现过程中存在的问题，即时反馈与调控；增值评价：有效促进教育的可持续发展，从发展性角度关注学生的学习进步程度、教师的教学进步程度、学校发展程度等；综合评价：利用智能技术实现综合评价模型的构建与优化、数据的深度挖掘分析等，实现对区域的教育质量综合评价、学生的综合素质评价等。

#### (五)用户层：技术赋能教育评价的目的归宿

用户层是用智能技术赋能教育评价，以实现评价的目的，是教育评价的归宿所在。具体来说，政府、学校、教师、学生、家长等用户通过智能终端设备实现一站式登录智能化教育评价系统，为不同用户提供不同的评价反馈，其评价反馈结果对用户具有导向、诊断、鉴定、调控、改进等功能。对政府而言可宏观了解区域教育发展情况，有助于促进区域教育优质均衡发展；对学校而言可有助于鉴定学校水平、评定学校优劣、促进学校发展等；对教师而言可以整体了解教学情况，进一步提高教师教学水平；对学生而言有助于促进学生的全面发展；对家长而言有助于即时了解学校的发展状况、学生的在校表现等。

#### (六)保障体系：技术赋能教育评价的重要保证

保障层是智能技术赋能教育评价有序实施的重要保证，使整个教育评价生态体系有效运转，包括政策制度、标准规范、运维管理、信息安全等保障。(1)以政策制度作为教育评价发展的根本导向，从建设内容、建设举措等方面制定支持教育评价的相关政策，并加强各地的督导落实。(2)建设工作要遵循技术标准规范，充分了解智能技术及其在教育评价中应用的特点，制定智能技术赋能教育评价的技术标准规范，包括管理规范、质量标准、技术标准等方面。(3)构建政府、学校、企业等多方协同的运维管理保障工作机制，为智能化教育评价支持系统提供技术支撑与服务保障，确保系统正常运行。(4)建立智能化教育评价的安全保障制度，明确信息安全责任主体，构建相关应急处置机制。

### 五、智能技术赋能教育评价的主要途径

《深化新时代教育评价改革总体方案》(下称《总体方案》)明确提出“坚持科学有效,改进结果评价,强化过程评价,探索增值评价,健全综合评价,充分利用信息技术,提高教育评价的科学性、专业性、客观性”。“四个评价”各有特点和优势,也不同程度存在着有待改进的方面,亟需采用新的技术手段加以改革创新。智能时代用技术赋能教育评价,促进评价改革创新的主要途径表现在智能技术赋能的结果评价、过程评价、增值评价和综合评价等方面,如图2所示。

进行预测。基于题库标注、神经网络训练、试卷结构生成、OCR文字识别、图像定位和自然语言处理等技术,依据教学需求检索出需要的题目自动生成试卷,实现快速、高效的智能化组卷。二是智能考场管理。基于人工智能技术开发智能化的网络电子巡查系统、考生身份验证系统、作弊防控系统等。比如通过在考场安装高清摄像头,利用图像识别技术、云技术、情感识别技术,对考场的动态视频数据进行实时监测和智能识别,自动分析考生的行为状态,及时识别考场的异常行为,实现考试过程中的智能监考。三是智能阅卷与评测。利用手写识别、图像识别、自然语言理解等技术,对选择题、

	评价目的	特点	存在问题	改革重点	技术赋能举例
结果评价	判别教育目标的实现程度	1.目标性 2.规范性 3.客观性 4.易操作性	1.命题的科学性有待提高 2.阅卷评分的区分度和科学性需要提升 3.测评方式以纸笔考试为主	“改进”结果评价,在当前重“量”的结果评价盛行的情况下,强调“结果”的“质”,关键要探索由“量”的评价向“质”的评价的转变	1.智能命题与组卷 2.智能考场管理 3.智能阅卷与评测 4.智能考试分析
过程评价	真实反映学生和教师在教学过程中的学习与发展状态	1.全过程 2.诊断性 3.即时性	1.缺乏客观性 2.缺乏整体性 3.难以量化、不易操作	“强化”过程评价,将师生在教与学活动过程中的全部信息纳入评价范围,展现学习与发展的结果成效、教与学方式的变化、师生成长与发展变化	1.全过程伴随式数据采集 2.多模态诊断分析 3.实时反馈与调控
增值评价	以学生进步幅度来衡量学校努力程度(看起点、比进步)	1.成长性 2.阶段性 3.激励性	1.研究内容片面 2.数据采集完整性难以保证 3.最佳模型需要进一步研究	“探索”增值评价,做好对学生的起点评估,要将所得到的各项学生发展指数作为增值情况分析的重要依据	1.追踪数据库链接与整合 2.增值数据分析模型构建 3.发展水平可视化报告 引导学生德智体美劳全面发展
综合评价	引导学生德智体美劳全面发展	1.系统性 2.多元性 3.复杂性	1.标准难以客观量化 2.评价数据复杂性增强 3.评价结果难以客观公正	“健全”综合评价,引入了人为因素,难点在于保障评价的客观性、公正性与公信力,要加强诚信文化建设,建立配套惩罚措施等	1.综合评价模型构建与优化 2.全域评价数据采集汇聚 3.数据的深度挖掘分析 4.数字画像建立与分享

图2 智能技术赋能教育评价主要途径分类

#### (一)智能技术赋能结果评价

结果评价通常是在教育活动完成后对最终结果做出价值判断,即结果与目标的一致程度的评价。结果评价具有目标性、规范性、客观性和易操作性等特点。当前我国的结果评价主要应用于教育考试,在科学命题、阅卷方式、测评方式等方面还存在问题,亟待通过技术手段进一步优化。通过智能技术赋能结果评价,将在命题与组卷、考场管理、组卷与评测、考试分析等方面推进评价工作的信息化和智能化水平提升,探索结果评价由“量”的评价向“质”的评价转变,实现智能化考评<sup>[19]</sup>。

智能技术赋能结果评价的主要应用场景包括以下方面:一是智能命题与组卷。智能技术在题库建设中的关键运用体现在试题难度预测方面,教师可以通过深度学习和人工特征两种方案对试题难度

填空题和是否判断题等客观题进行智能化阅卷;通过自动分析及训练构建分数预测模型对中英文作文、翻译、作答等主观题进行智能评分;利用多媒体、AR/VR等技术,构建人机交互、模拟操作、场景再现的智能化考试环境,有效实现英语听说、音乐、美术等方面的自动化考试。四是智能考试分析。利用数据挖掘、学习分析、数据预测模型等技术,对试卷的命题质量、教师教学质量、试卷讲评等提供客观有“证据”的数据分析,为改进教与学提供强大的量化支持。

#### (二)智能技术赋能过程评价

过程评价是对教师的教学过程和学生的学习过程作出即时的评判,聚焦于关注学生成长过程,注重学习者习得过程的评估与测量,从而实现以评促学、以评促教。过程评价具有全过程、诊断性和即

时性等特点。传统的过程评价存在诸多问题,如缺乏客观性、整体性,难以量化、不易操作等。在智能技术的赋能下过程评价发生了诸多改变,数据获取实现了从小规模的简单间断性数据向大规模的复杂连续性数据转变,数据分析实现了从数据结果的粗线条分析向多模态诊断分析转变,数据反馈实现了由单一文本的事后反馈向融合文本、语音、图片等多种方式的交互式实时反馈转变。

智能技术赋能过程评价的应用场景主要包括:一是全过程伴随式数据采集。依托可穿戴设备、传感器、视频录制技术、非接触式感知技术等数据采集工具,实现全过程数据的实时采集和及时生成,通过智能技术采集的数据更加具有动态性、真实性、连续性。二是多模态数据诊断分析。用数据挖掘算法、内容分析、预测性分析、系统建模等技术对行为数据、学习体征数据、学习资源数据、人机交互数据等进行多模态诊断分析,数据分析实现了从“不可量化”到“可量化”、从“碎片”到“集约”的转变,通过智能技术分析的数据更加具有客观性、准确性。三是实时反馈与调控。分析结果以可视化的形式即时反馈给管理者、教师、学生等不同对象,根据评价结果调控和改进管理、教学、学习等进程,评价结果涉及多个维度,如对学生的评价有学习动机态度、过程和效果等方面的反馈,通过数据可视化方法反馈数据更具有即时性、全面性。

### (三)智能技术赋能增值评价

增值评价是以学生进步幅度来衡量学校努力程度的一种新型评价方式,重视学生的学习起点,关注学习过程,依据一段时间内学生的学业水平、综合能力和素质等方面的表现,开展纵向比较,引导学生专注于自身的综合表现,并获得学校对学生的“净增值”。增值评价具有成长性、阶段性和激励性等特点。近年来我国的增值评价研究取得了一定的进展,但是在研究内容、数据采集、增值模型构建等方面依然存在着问题。在智能技术赋能下进行学习评价,能够更加精准、全面地获得增值评价的数据,更加精确地剔除学生自身和家庭、教师和学校等因素对学生成绩的影响,更加客观地获取到学生学业水平的“净增长”,也能够更加公正地对学生、教师和学校等进行评价,提升增值评价的科学性、客观性、有效性。

智能技术赋能增值评价主要应用场景包括如下:一是追踪数据库链接与整合。在进行增值评价时,建立大规模的追踪数据库是实施增值评价的前提条件。利用智能采集技术获取学生学业

水平成绩、品德、学习能力、艺术审美等纵向数据;利用数据库链接技术实现学生数据库与学生、教师、学校等数据库之间的链接,达到数据之间的有效整合<sup>[20]</sup>;通过SQL Server事件探查器的跟踪技术对数据进行跟踪,并在追踪数据库逐渐庞大时,利用智能技术完成对数据库的维护。二是增值评价模型的构建。开展增值评价,需要利用增值评价模型对收集的数据进行处理和分析,常用的主要有增分模型、成长达标模型、分类模型、残差模型、回归模型、学生成长百分等级模型、田纳西增值评估系统7种<sup>[21]</sup>。目前尚不存在适用于所有数据的模型,需要利用人工智能技术、借助统计分析手段和算法构建出能对学生进行全面、综合性评价的增值模型。三是发展水平可视化报告。增值评价主要关注学生的纵向发展水平,通过可视化技术能够将学生的学业情况、个人发展指数等信息直观形象、简单易懂、全面清晰的呈现出来,便于学生、家长和教师等查看和理解。

### (四)智能技术赋能综合评价

综合评价是从全面、多维的角度对教育活动作出系统的价值判断,其重点在于对评价对象进行分析的基础上,把多方面的评价结果整合起来,形成对评价对象清晰准确的总体认识。综合评价具有系统性、多元性、复杂性和全面性等特点。综合评价正在成为教育评价的趋势和导向,但也存在标准难以客观量化、数据复杂性增强、客观公正不足等问题,亟需通过智能技术助力解决这些问题。杨宗凯认为,要推进多种主体评价等甚至是机器评价的融合发展,采用基于智能技术的不同方法进行综合评价<sup>[22]</sup>。人工智能、大数据等作为健全综合评价的关键技术支撑,在评价模型构建、全域评价数据采集、数据深度挖掘分析、数字画像等方面极大地提高操作效率、评价效果和评价效能。

智能技术赋能综合评价应用场景主要包括以下方面:一是通过层次分析法、专家系统、机器学习、神经网络等算法,在对数据的有效分析中提取评价规则,实现综合评价模型的科学构建与不断优化,有效解决综合评价标准科学量化的问题。二是基于大数据和区块链等智能技术建立统一的数据平台,通过各类采集终端对评价对象进行全方位、全过程、全纳、全员和非干预的自然状态采集,汇聚全样本、混杂、海量的数据,解决综合评价中信息量小、失真和结构缺失的问题<sup>[23]</sup>。三是利用数据挖掘和学习分析技术,充分挖掘评价对象的特征和关系,全面分析评价对象的综合发展状态,深刻揭示评价结果的共性与个性,精准预测评价对

象的发展潜力与未来趋势。四是基于大数据分析技术和深度学习等算法,将最终的综合评价结果进行可视化呈现,生成个体或群体数字画像,能够让学生、教师及教育管理者清晰直观、形象生动地理解运用及分享综合评价结果。

### 六、智能技术赋能教育评价的推进策略

智能技术赋能教育评价是一项复杂的系统性、整体性、协同性工程,应以《总体方案》为依据,以人工智能、大数据等技术为依托,围绕服务各级各类教育主体,聚焦关键环节,采取理念引领、标准规范、主体关照、数据驱动、专业支持等实施策略(如图3所示),科学谋划、整体推进、分类实施。

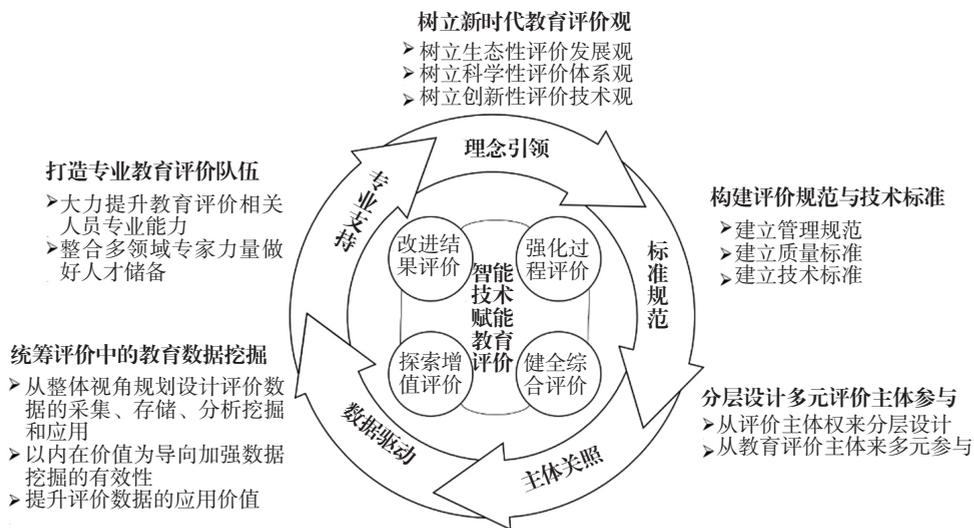


图3 智能技术赋能教育评价推进策略

#### (一)理念引领：树立新时代教育评价观

推进智能技术赋能教育评价,首要问题就是转变思想理念,树立新时代教育评价观,扭转不科学的教育评价导向。一是坚持立德树人,树立生态性评价发展观。教育评价的最终目标是保障和促进学生的全面发展和教育质量的全面提升,在实践中引导各类主体形成正确的政绩观、育人观和质量观,适应新时代教育发展的生态性要求。二是落实“四个评价”,树立科学性评价体系观。坚持以破“五唯”为导向,针对不同主体和不同学段、不同类型教育特点,通过立体、多元的评价方式,结果评价、过程评价、增值评价和综合评价相结合,形成科学的教育评价体系。三是聚焦智能转型,树立创新性评价技术观。人工智能、大数据等智能技术可以赋能教育评价的各个环节,形成基于数据的客观性思维、动态变化的持续性思维和价值呈现的发展性思维等创新性技术思维,树立教育评价实现智能

化转型。

#### (二)标准规范：构建评价规范与技术标准

评价规范与技术标准是科学实施教育评价的前提,通过建立管理规范、质量标准、技术标准“三位一体标准规范”,保障教育评价的专业、稳定、可持续发展。(1)管理规范是从顶层设计视角制定系列配套政策和制度,统筹规划教育评价改革。如制订实施方案、行动计划、实施指南等文件,明确目标任务、推进策略和工作机制,统筹实施教育评价的进度和具体安排。(2)质量标准是在贯彻国家及教育部相关标准文件要求基础上,结合区域实际构建教育质量评价指标体系,是对“评价什么”的具体描述和说明,也是评价工具的研发依据。《教育部2021年工作要点》

明确指出要推动出台中学生综合素质评价实施指南、义务教育质量评价指南、普通高中办学质量评价指南、普通高等教育质量评价指南等,为区域实践中教育质量标准的制订提供了指导和具体参考。(3)技术标准是推动智能技术赋能教育评价的重要技术准则,教育评价系统或平台可通过遵循国家技术标准、借鉴国际技术标准和制订地方

技术标准三种形式构建教育评价技术标准体系<sup>[24]</sup>。

#### (三)主体关照：分层设计多元评价主体参与

教育评价主体涉及党委和政府、学校、教师、学生和社会等多种类型,各个主体处于纵向教育系统的不同层级之中,每个层级主体的权益不同,因而要分层设计,全面关照每个主体,保证被评价者在评价过程中有充分的参与权、选择权、解释权,以体现教育评价的平等性和公正性。一方面,从评价主体权的利益诉求来看,可根据主体职责分工进行分层设计和管理主体权限赋予,确保各个主体按照各自的评价任务与责任深度参与评价过程,充分保障主体评价权,从而促进评价目标的有效达成。另一方面,从教育评价主体的多元化来看,要培育平等、公正、包容的评价文化,保障被评价者能参与评价标准的构建,并推进第三方评价、利益相关者评价和社会评价,使不同主体获得归属感和认同感,从而实现教育评价的教育性和发展性价值,促



进教育评价生态优化。

#### (四)数据驱动：统筹评价中的教育数据挖掘

数据是教育评价实践中最重要的“资源”，数据驱动教育评价是当前智能技术赋能教育评价的核心应用模式。通过数据挖掘找准规律、发现问题和预测趋势，从而提高教育评价的应用效能和服务水平。(1)从整体视角规划设计评价数据的采集、存储、分析挖掘和应用。包括统筹数据采集流程，做好数据选择、数据清洗和数据转换等数据处理；完善数据验证，保证数据的一致性和相关性，以便从整体和宏观上能够发现教育本质特征和潜在规律。(2)以内在价值为导向加强数据挖掘的有效性。面向学习者，要以促进个性化学习为导向，通过学习过程数据的深度挖掘，为学习者提供个性化学习建议。面向教师，要以差异化教学为导向进行学习数据挖掘分析，支持教师及时诊断学习者学习情况。面向管理者，要提高管理效率和质量为导向进行问题发现和趋势预测分析，以帮助管理者合理配置资源和优化教育决策。(3)提升评价数据的应用价值。要提高数据挖掘分析的易用性，重视和加强数据的可视化，让各类主体深度参与，强化大数据技术与教育评价深度融合的机制与能力，综合发挥智能化教育评价的导向、鉴定、诊断、调控和改进作用。

#### (五)专业支持：打造专业化教育评价队伍

专业化的评价需要专业化的评价队伍，面对智能时代的挑战，亟需培养一支精通教育评价业务且具备智能素养的专业化教育评价队伍。一方面，要大力提升教育评价相关人员专业能力。加强教育评价各个主体智能素养培养，通过开展人工智能、大数据等智能技术的培训和实践应用，提升信息素养和智能技术应用能力。加强一线校长、学科教师等主体的教育测量、教育评价等专业知识培训，提升实施教育教学评价的能力。同时，加强与国内外各种教育评价专业机构的交流与合作，及时跟踪教育评价的发展动态和技术革新。另一方面，整合多领域专家力量做好人才储备。《总体方案》明确提出支持有条件的高校设立教育评价、教育测量等相关学科专业，培养教育评价专门人才。在教育评价实践中，需要学科教学、心理学、统计学、计算机、教育测量、教育管理等领域的专家共同协作，各领域专家发挥各自的优势和特长，从而推进教育评价的创新研究<sup>[25]</sup>。同时要充分发挥心理学、统计学、计算机等基础学科的力量培养教育评价、教育测量专门人才，增强其跨领域综合应用和协同创新能力，促进智能技术赋能的教育评价不断改革创新，

构建智能化教育评价新体系。

#### 参考文献：

- [1] 杨勇. 智能化综合评价理论与方法研究[D]. 杭州: 浙江工商大学, 2014.
- [2] 张殿尉. 智能化教育评价领域演进路径、研究热点与前沿的可视化分析[J]. 工业技术与职业教育, 2020, 18(1): 110-114.
- [3] U.S. DEPARTMENT OF EDUCATION. The U.S. Department of education. National Education Technology Plan 1996 [EB/OL]. <https://files.eric.ed.gov/fulltext/ED398899.pdf>, 2021-02-19.
- [4] U.S. DEPARTMENT OF EDUCATION. The U.S. Department of education. National Education Technology Plan 2000 [EB/OL]. <https://files.eric.ed.gov/fulltext/ED444604.pdf>, 2021-02-19.
- [5] U.S. DEPARTMENT OF EDUCATION. The U.S. Department of education. National Education Technology Plan 2004 [EB/OL]. <https://files.eric.ed.gov/fulltext/ED484046.pdf>, 2021-02-19.
- [6] U.S. DEPARTMENT OF EDUCATION. The U.S. Department of education. National Education Technology Plan 2010 [EB/OL]. <https://files.eric.ed.gov/fulltext/ED512681.pdf>, 2021-02-19.
- [7] U.S. DEPARTMENT OF EDUCATION. The U.S. Department of education. National Education Technology Plan 2016 [EB/OL]. <https://tech.ed.gov/files/2015/12/NETP16.pdf>, 2021-02-19.
- [8] U.S. DEPARTMENT OF EDUCATION. The U.S. Department of education. National Education Technology Plan 2017 [EB/OL]. <https://tech.ed.gov/files/2017/01/NETP17.pdf>, 2021-02-19.
- [9] 刘邦奇, 吴晓如. 中国智能教育发展报告[M]. 北京: 人民教育出版社, 2020. 19-21.
- [10] 杨现民, 田雪松. 互联网+教育: 中国基础教育大数据[M]. 北京: 电子工业出版社, 2016. 147.
- [11] 董奇, 赵广立. 科技赋能教育评价改革时机已至[N]. 中国科学报, 2021-01-12(07).
- [12] Zhang Qingchen, Yang Laurence T., et al. A survey on deep learning for big data [J]. Inform Fusion, 2018, (42): 146-157.
- [13] Wiley J, Hastings P, et al. Different approaches to assessing the quality of explanations following a multiple-document inquiry activity in science [J]. International Journal of Artificial Intelligence in Education, 2017, 27(4): 758-790.
- [14] 杨现民, 顾佳妮等. “互联网+”时代数据驱动的教育评价体系构架与实践进展[J]. 浙江师范大学学报(社会科学版), 2019, 44(4): 16-26.
- [15] 郑燕林, 柳海民. 大数据在美国教育评价中的应用路径分析[J]. 中国电化教育, 2015, (7): 25-31.
- [16] 曹培杰. 以新思维新技术破解教育评价痛点[EB/OL]. [https://epaper.gmw.cn/gmrb/html/2019-12/10/nw.D110000gmrb\\_20191210\\_4-13.htm](https://epaper.gmw.cn/gmrb/html/2019-12/10/nw.D110000gmrb_20191210_4-13.htm), 2021-02-05.
- [17] 朱成晨, 闫广芬. 现代化与专业化: 大数据时代教育评价的新技术推进逻辑[J]. 清华大学教育研究, 2018, 39(5): 75-80.
- [18] 刘邦奇, 王亚飞. 智能教育: 体系框架、核心技术平台构建与实施策略[J]. 中国电化教育, 2019, (10): 23-31.
- [19] 范国睿. 教育评价改革需要新方向[N]. 郑州日报, 2020-07-24(08).
- [20] 边玉芳, 王焯晖. 增值评价: 学校办学质量评估的一种有效途径[J]. 教育学报, 2013, 9(1): 43-48.
- [21] 王晓平, 齐森等. 美国学校“成长测量”的7种主要方法[J]. 中国考试, 2018, (6): 21-27.
- [22] 杨宗凯. 利用信息技术促进教育教学评价改革创新[J]. 人民教

育,2020,(21):30-32.

[23] 张治,戚业国.基于大数据的多源多维综合素质评价模型的构建[J].中国电化教育,2017,(9):69-77+97.

[24] 刘邦奇,张振超等.区域教育大数据发展参考框架[J].现代教育技术,2018,28(4):5-12.

[25] 张娜.从对教育的评价到促进教育的评价——教育评价国际研究进展综述[J].基础教育,2017,(4):81-88.

#### 作者简介:

刘邦奇:教授,院长,认知智能国家重点实验室智能教育研究中心主任,中国教育技术协会人工智能专业委员

会常务理事,研究方向为人工智能教育应用、智慧校园、智慧课堂、区域智慧教育规划。

袁婷婷:硕士,研究方向为智能教育、智慧校园、智慧课堂。

纪玉超:硕士,研究方向为人工智能教育应用、智慧课堂。

刘碧莹:硕士,研究方向为智能教育、教育管理。

李岭:硕士,研究方向为智慧课堂、智能教育。

## Intelligent Technology Enabling Education Evaluation: Connotation, Overall Framework and Practice Path

Liu Bangqi<sup>1,2</sup>, Yuan Tingting<sup>1</sup>, Ji Yuchao<sup>1</sup>, Liu Biying<sup>1</sup>, Li Ling<sup>1</sup>

(1.iFLYTEK Educational Technology Institute, Hefei 230088, Anhui; 2.College of Educational Technology, Northwest Normal University, Lanzhou 730071, Gansu)

**Abstract:** The enablement of intelligent technology provides an important opportunity for the reform and development of education evaluation, and promoting the integration and innovation of intelligent technology and educational evaluation is the inevitable choice to deepen the reform of education evaluation in the new era. Accordingly, the intelligent technology enabled evaluation education is studied systematically, and its historical origin can be traced back to the intelligent comprehensive evaluation in the field of management engineering and the application of computer technology in the field of education evaluation. Its essence is the breakthrough and innovation of traditional education evaluation based on intelligent technology, and the formation of a new education evaluation model through deconstruction and reconstruction, which has the main characteristics of scientificization, diversification, three-dimensional, optimization and precision. The intelligent technology and its application to assist education evaluation consists of the overall framework of “5+1” (five hierarchical structures of basic layer, technology layer, platform layer, application layer and user layer and one guarantee system). Through the specific scene application of “four evaluations” enabled by intelligent technology, the overall promotion strategies of concept guidance, standard norms, subject care, data-driven and professional support are adopted to build an ecological system of intelligent education evaluation.

**Keywords:** intelligent technology; enablement; education evaluation; overall framework; practice path

收稿日期: 2021年3月18日

责任编辑: 李雅瑄