

打造智慧考试 服务智慧教育

于涵

(教育部教育考试院,北京 100084)

摘要: 教育考试数字化是立足新发展阶段、促进教育考试高质量发展与实现教育考试现代化的必然要求、必要选择和必由之路。本文聚焦国家教育数字化战略行动总体要求,对标智慧教育总体架构,以数字化思维重构考试业务流程,提出以智慧考试为核心概念的全新考试业务系统,描绘我国大规模教育考试数字化发展的愿景和技术路线。智慧考试按照“3133”的整体架构进行总体布局,以海量题库、个性试卷、科学评价为核心特征,以试题创生为龙头,并与考试投送、编码评分和智能评价共同构成核心圈层,以数据确权、数据安全和数据标准作为技术支撑。当前,通用人工智能技术的涌现正在掀起全面智能化发展浪潮,智慧考试的整体架构和研究成果对于探索适应智能时代的教育考试新模型、新工具、新应用具有重要的理论价值和实践意义。

关键词: 数字化;智慧考试;智慧教育;智能算法;大数据;原型机

【中图分类号】G405 【文献标识码】A 【文章编号】1005-8427(2023)05-0001-10

DOI: 10.19360/j.cnki.11-3303/g4.2023.05.001

当今时代,数字技术作为世界科技革命和产业变革的先导力量,日益融入经济社会发展的各领域、全过程,深刻改变着生产方式、生活方式和社会治理方式^[1]。人工智能技术持续变革,特别是通用人工智能技术的涌现,更是加速推动社会智能化水平迈入新的阶段。进入新时代,以习近平同志为核心的党中央高度重视数字化发展,明确提出建设数字中国^[2-3]。“推进教育数字化”写入党的二十大报告,为新时代新征程进一步发展教育数字化指明了前进方向,提供了根本遵循^[4]。2023年2月,国务院印发的《数字中国建设整体布局规划》提出,到2025年,基本形成横向打通、纵向贯通、协调有力的一体化推进格局,数字中国建设取得重要进展^[5]。《“十四五”数字经济发展规划》明确提出深入推进智慧教育^[6]。教育部以“应用为王、服务至上、简洁高效、安全运

行”为总要求^[7],全面推进国家教育数字化战略行动。2023年2月,以“数字变革与教育未来”为主题的世界数字教育大会召开,围绕数字化转型、数字学习资源开发与应用、教育数字治理和教育发展评价等专题展开深入讨论,获得国际社会广泛关注^[8]。国家智慧教育平台上线以来,在应对疫情、服务“双减”、促进优质资源共享等方面发挥了重要作用^[9]。综上,推进教育数字化转型是贯彻落实科教兴国战略、人才强国战略、创新驱动发展战略的重要举措,教育数字化的变革性实践和标志性成果正在成为教育领域的重要驱动力。

作为国家教育制度的组成部分,我国大规模教育考试通过人才甄别和选拔的组织实施,在保障教育公平公正、服务拔尖创新人才培养、促进人力资源优化、助力经济社会发展等方面发挥着

收稿日期: 2023-03-29

作者简介: 于涵(1978—),男,教育部教育考试院党委副书记、副院长,研究员。

积极作用。同时,考试又同教育教学紧密衔接,既是教育过程中的关键环节,又通过反馈和评价机制深刻影响教育教学,与之形成循环和互馈。随着国家教育数字化战略行动深入推进,利用现代数字技术和理念助力教育考试内容、形式、评价及治理改革,建设与智慧教育相适应、为智慧教育服务的数字化考试体系(以下简称智慧考试),是我国教育考试现代化进程中一项重要且急迫的任务。

1 背景和意义

首先,数字化发展是教育考试进入新发展阶段的必然要求。在加快建设数字中国、推进教育数字化的总体要求下,国家教育数字化战略行动是立足新发展阶段、引领教育变革、推进教育现代化的重大举措。聚焦到教育考试领域的改革发展历史进程,实现教育考试现代化是一以贯之的主题和矢志不渝的目标。20世纪八九十年代的考试标准化改革^[10]和进入新世纪以来的考试信息化建设^[11]有力促进了我国教育考试现代化发展,为进一步推动教育考试现代化建设奠定了坚实基础。在教育考试领域,深入推进教育考试数字化,实现数字化变革和转型,是落实国家教育高质量发展的根本要求,对实现教育考试现代化具有十分重要的意义。

其次,数字化发展是教育考试贯彻新发展理念的必要选择。新发展理念倡导创新、协调、绿色、开放、共享。数字化所依托的数据科学被称为继实验科学、理论科学、计算科学之后的第四科学范式,直接影响并促进各个领域的数字化发展和转型。新范式具有与新发展理念相一致的特征:在创新方面,数字化能够带来产业变革,考试数字化将促进教育考试创新发展和业态重塑;在协调方面,数字化能够实现交流和互通,考试

数字化有助于加强考试各环节之间的信息共享和协调联动,保障考试的质量和公平;在绿色方面,数字化能够优化服务、提高效率,考试数字化可以营造简洁、高效、低碳的考试服务环境;在开放和共享方面,数字化的基础就是通过数据流通打破信息孤岛,考试数字化有利于促进各类教育资源和服务的共建共享,构建更加开放、融通的教育生态。

最后,数字化发展还是教育考试实现高质量发展的必由之路。实现高质量发展是事关教育考试事业发展全局的系统性、深层次变革,是立足当前、着眼长远的战略谋划。《深化新时代教育评价改革总体方案》是教育考试改革的纲领性文件,明确了改革发展的总体目标,包括落实立德树人、坚持“五育并举”和实现“四个评价”等^[12]。对照总体目标,教育考试必须面对包括如何破“五唯”、如何尊重教育规律以更好发挥考试正向反拨作用、如何引导和鼓励创新、如何在超大规模考试中考出能力和素养、如何实现因材施教而非抹杀个性等一系列重大问题。通过教育考试数字化发展寻求突破、化解难题并取得成效,是贯彻落实国家教育数字化战略行动、实现教育考试高质量发展的必由之路。对教育考试而言,数字化发展既是走进新时代、迈上新征程的战略机遇,也是推进我国从考试大国迈向考试强国的关键着力点。

立足教育考试数字化发展,对标国家智慧教育总体架构,坚持理念创新和技术实践相结合,面向服务科教兴国和人才强国大局,聚焦国家考试重要功能,本文提出智慧考试的核心理念和目标框架,助力考试业务数字化转型和智能化升级,构建以数据驱动为核心的发展模式,探索适应时代要求的教育考试新模型、新应用,打造可为育人选才全过程服务的考试新业态。

2 智慧考试整体框架

智慧考试按照“3133”的框架进行总体布局,架构示意图见图1。在整体架构中,海量题库、个性试卷和科学评价是智慧考试的3个核心特征。其中,海量题库汇集大量优质试题,承载数据资源,确保试题对教育教学目标及选才需求的有效覆盖;个性试卷使试卷难度和考生群体能力水平更加匹配,提高考试信效度,减轻大量刷题带来的学习负担和偏题怪题带来的心理压力;科学评价提供合理有效的结果报告及使用方法,既直接弱化“唯分数”导向,又间接服务教育管理改进和教育质量提升。3个核心特征聚焦实现“有考无类”“因材施教”和破除“五唯”,目标是促进教育公平、提升教育质量、服务创新人才选拔和助力教育评价改革。

在智慧考试总体架构中,试题创生是龙头,拥有以人工智能为基础的智能出题、难度预估、自动组卷和全网查重等创新技术,支持多媒体动态题型,由此衍生的国家题库系统发挥着数据中枢和超级节点作用。考试投送、编码评分^[13]和

能评价三线赋能,与试题创生一起共同构成智慧考试的核心圈层。其中,考试投送支持多种模式的考生试卷交互,记录和获取作答数据及其他无感采集信息;编码评分的过程实现了对考生主观题作答从单维评分到多维编码的转变,力求充分理解作答思路,发现闪光点,保护个性,鼓励创新;智能评价包含数据分析、结果报告、数据仓库和监控预测等功能,智能生成评价报告,从“按分排队”转变为“立体塑像”。数据确权、数据安全和数据标准构成3大支撑。智慧考试架构的一个突出特点是数据驱动,图1中的箭头表示数据流动的方向,从而形成考试全链条的数字化联动。

3 智慧考试的关键要素

智慧考试核心圈层由试题创生、考试投送、编码评分和智能评价4部分共同构成。它们是智慧考试的实体,也是重要功能单元,直接体现数字化思维对考试业务流程的重塑和对考试整体生态的重构。

3.1 试题创生

试题创生是命题人员进行试题创作和生产

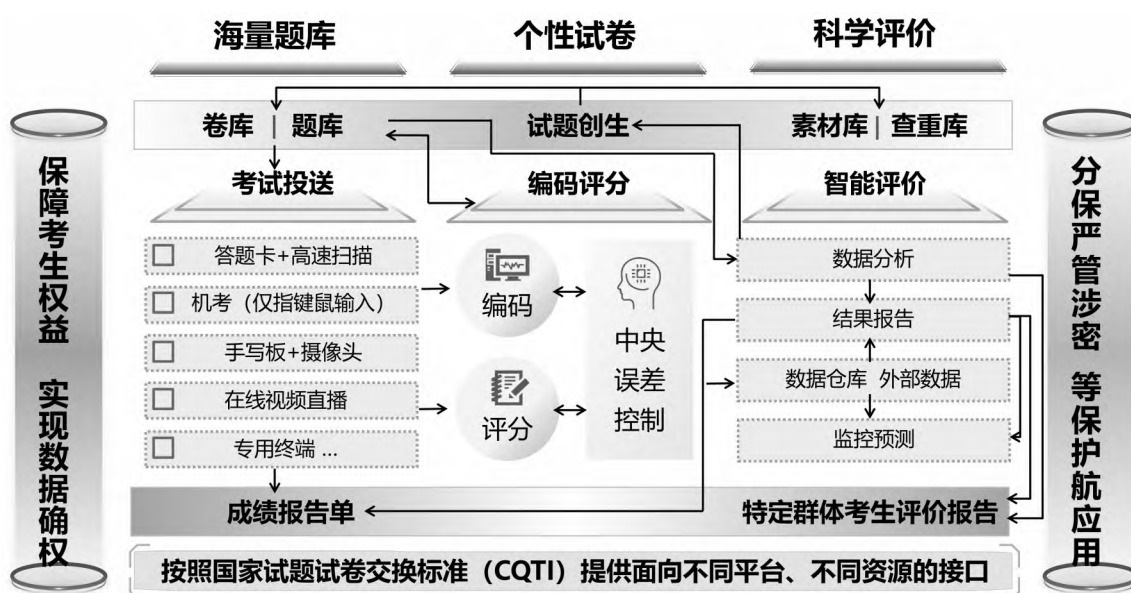


图1 智慧考试架构示意图

的平台,既为试题构思设计提供智能化的技术支持,也为试题精雕细琢提供全方面的工具配置。在试题创生所搭建的环境中,命题人员可以通过自动化、智能化的方式批量加工试题,还能够像塑造工艺品一样逐个研磨试题。通过人工智能和大数据技术赋能,以智能出题、难度预估、自动组卷和全网查重为代表的创新技术已经成为试题创生的重要功能,在命题的相关环节发挥着关键作用。智能出题聚焦试题命制,助力生产端;难度预估和全网查重着眼试题质量,守好出口端;自动组卷面向试题管理,服务考试端。这些技术在很大程度上提高了试题创生的交互性、精准化和个性化程度,在提质、增效、保稳和控险等多个方面体现出了数字化的特点和优势,使试题创生成为智慧考试的核心单元。具体来说,试题创生具有一个重要特征和两项主要内容。

试题创生的一个重要特征是支持创新试题类型。试题创生能够提供更加丰富多样的试题类型,除了面向已有考试项目的传统命题需求提供通用功能组件外,更为未来考试命题的创新需求提供个性化配置服务,目的是通过丰富的交互方式引导和激发命题人员的创作灵感,突破传统试题类型的限制,设计出更多且能够体现出数字化特点和优势的试题。研究表明,技术增强题型更能激励学生的作答兴趣,其互动元素可以帮助教育者更多地了解学生如何思考问题以及如何根据提示作出反应,鼓励学生对知识应用进行实践探索,让学生对于自己的学习有更强的责任感^[14-15]。与传统计算机化题型相比,通过个性化配置服务设计出来的技术增强题型凭借技术优势和数据采集多样性,更符合数字化发展的趋势和教育教学的要求,也更有利于服务拔尖创新人才的培养与选拔。

试题创生包括两项内容:一是充分借助人工

智能技术在试题创生中创作方面的研发助手;二是在试题创生中生成方面的设计工具。试题创作指的是以传统的专家命题创作为主导形式、以人工智能技术为辅助形式的人机协同模式;这种形式基本可以服务所有考试项目,也包括一些高关注度、高利害的大规模考试项目。试题生成指的是以智能生成试题为主导形式,以专家审校修改为辅助形式的人机协同模式。这种形式由于机器参与程度更高、比重更大,初始阶段可以在一些分散性、小规模和非高利害的考试项目中使用。

试题创生在智慧考试中所发挥的龙头作用可以从3个角度得到体现。从系统功能而言,试题创生是智慧考试运转的“发动机”。试题的创作和生产方式直接关系到考试的呈现方式、格式标准和运行效率。在试题创生的牵引下,智慧考试的其他环节与之深度对接,实现联动。从节点位置而言,试题创生是智慧考试数据的“原生地”。试题创生是智慧考试数据流的最初起点、关键节点和重要支点。试题创生伊始,试题数据即进入题库或卷库,进行统一存储、标注和管理。试题进入智慧考试后续流程,所产生的数据也会与试题创生发生关联。从核心贡献而言,试题创生是智慧考试的“主产区”。试题创生提供各类丰富题型的创作和生产环境,除满足当前考试项目和科目的命题需求之外,还顺应数字化发展的趋势为技术增强题型等的创造生产提供服务支撑。

3.2 考试投送

考试投送是把试题创生输出的试题试卷通过不同形式呈现给考生,同时记录考生作答数据、日志信息和其他无感采集信息。考试投送是连接考试和考生的桥梁,是从试题创生到考生作答的枢纽,关系到命题人员的试题创作意图能否

得到准确体现,关系到考试能否安全有效组织实施,还关系到考生的真实作答过程及结果能否得到完整记录。考试投送主要包括如下5种方式。

一是纸笔形式(答题卡+高速扫描)。纸笔考试是最为常见也是历史最为悠久的考试形式,在当前一些重要的大规模考试中还占据主导地位。近些年,随着信息技术逐步应用于大规模考试,纸笔形式也在不断优化完善。例如,考生答题卡设计更加科学,考生信息收集和管理更加规范,考生作答能够通过扫描快速转换成电子文档上传至阅卷平台。纸笔形式对考场软硬件要求都不高,组织管理成熟,能够完整记录考生在答题纸上的书写笔迹,这对以主观题为主、考试和评卷统一组织的大规模高利害教育考试而言,几乎是不可或缺甚至难以取代的。

二是机考形式(仅键盘输入)。以键盘和鼠标作为输入设备的机考形式已在众多考试项目中应用和推广。计算机屏幕作为重要输出端,能够呈现比纸笔形式更为丰富多样的试题类型。对命题人员而言,试题创作形式更加灵活,一些受限于纸笔形式的命题理念能够在机考形式下得以体现。例如,增加声音、动画和视频等多媒体内容,设计人机交互的试题类型等。正如计算机已经改变了人类生活的方方面面一样,从长远发展来看,机考形式更能顺应信息化时代要求,基于计算机和网络的考试将是未来很长一段时间内考试研究和技术发展的主战场。

三是手写板+摄像头。这种方式同时具备机考和纸笔考试的特点,既可以呈现丰富的题型,又可以支持考生手写笔迹的记录,特别是通过摄像头还能全程捕捉考试作答过程中的视频画面信息。因此,这种方式能够无感采集考生在考试场景下的所有作答信息和操作信息,为多维分析评价考生能力提供重要数据基础。

四是在线视频直播。在线视频直播考试已在一些考试项目中,尤其是在面试环节里得到使用,其主要形式是监考人员通过音视频在线发出指令,实时指导并监控考生在居家等非标准化考试场地进行考试。当考试受到某些突发事件影响时,这种形式能够在一定程度上模拟并还原考试场景,观察和了解考生现场作答,因此正成为当下在线考查和评估考生的重要途径。在线下考试组织管理工作受到客观因素制约的情况下,在线视频直播考试能有效缓解组考压力,完成小型考试任务。

五是专用考试终端。专用终端立足未来考试发展,整合科技创新元素,从软硬件方面同时进行架构搭建,涵盖芯片设计、系统开发和兼容扩展等内容。作为专用考试设备,这些终端将来能够进行定制化的考试服务,有助于增强考试体验、激发考生兴趣,同时对于推动教育考试新基建和提升教育考试专业化水平也具有促进作用。例如,基于虚拟现实(VR)和增强现实(AR)技术的终端设备能够呈现出更加逼真、更加贴近现实的试题情景,提供身临其境的考试体验,增强试题的实操性、探究性和开放性,不但能够全面考查考生能力,更有利于寓教于考,促进教与考的衔接和融合。

当然,考试投送方式不限于以上5种类型,而且在某个特定考试项目上也并不限于使用一种投送方式。数字化考试投送也并不意味着把考试直接照搬上计算机,其本质上还是为考试目标和考查要求来服务的,同时需要与不同地区的教育基础设施和信息化发展水平相匹配。对于低利害度的较大规模考试而言,考试投送方式可以作出更加灵活的尝试,而不拘泥于某一特定形式;对于高利害大规模考试而言,考试投送要更加注重安全和平稳;对于一些规模较小的考试而言,

则要勇于创新 and 试验不同考试投送方式,为将来在更大范围推广和应用探索路径,积累经验。随着更多科技融入考试投送环节,考试安全也能得到更全面、更有效的保障。

3.3 编码评分

在大规模考试中,主观题评分问题是世界性难题。传统上,评卷老师根据经验“踩点”给分来判定考生得分,容易产生两方面的问题:一是主观判断带来分数误差和公平问题;二是在一定程度上限制考生个性发挥,抑制创新。我国大规模教育考试多采用以双人评分、全程质控、实时预警、统计校验为核心的网上评卷模式,以尽量减少主观评分的不确定性,这在世界考试行业居于领先地位。

在数字化发展的新阶段,主观题评分将开启从传统模式到编码评分的新探索。编码评分引入大数据和人工智能技术(AI),大体分成4个步骤进行。第一步,分类归集。AI对所有考生作答进行快速筛选和分类,总结提炼出有代表性的考生答案类别的集合,对每一个类别赋予一个编码值(code),每个编码值代表一类答题思路或解法。第二步,标定编码。编码员根据与不同答案类别的近似度对所有考生作答标定相应的编码值,在此过程中,编码员会及时发现不同考生作答的创新点,必要时提交专家组研判,产生新的答案类别和编码值。第三步,专家赋分。专家组综合评判分类的合理性和编码结果,对每个编码值或类别进行赋分(score)。第四步,自动给分。系统根据每个考生作答的编码自动给出相应分数。

编码评分模式实现智能技术与人工评判的相互配合,最大限度减少误差、保障公平、保护个性、鼓励创新。编码评分收集的信息更加丰富,有助于更好了解考生能力特点,实现更加精准的评价,并有利于命题评价反馈。

3.4 智能评价

科学产出和使用考试结果是考试工作流程中的最后一个重要环节。在智慧考试架构中,这个环节被定义为智能评价。智能评价是各环节数据的“汇集地”和“加工厂”,是使用和呈现的关键“最后一公里”。它接收来自编码评分的评分数据、来自试题创生的信息数据、历史数据以及来自外部的其他相关数据(如考生群体数据、教育发展数据等),既通过传统的教育测量模型和统计方法对这些数据进行常规加工处理,又能够利用人工智能和大数据技术进行综合分析,从而形成多维度、全方位、立体化的评价结果。

评价结果有3个应用场景。第一,生成结果报告,服务于考生或者特定群体。传统的考试结果报告通常只包括分数或等级,智能评价则可以提供个体、群体、局域及全域的考生能力的立体画像,实现考试结果报告的科学化、精准化和多元化。第二,运用教育测量学中的统计方法进行考试的信度和效度分析,生成试题性能的相关评价指标,流转回试题创生环节,为改进命题、提高考试服务专业化水平提供信息反馈。第三,部分数据通过数字大屏或者仪表盘等形式服务行政单位和相关部门的实时监控和分析预测。学校、区/县、市、省、全国5级考试评价组织和管理机构在系统中可以即时、便捷了解到本级相关结果,改善传统模式下的信息割裂、信息滞后的情况,为数字化决策提供可靠依据,对防治“唯分数论”、改进教育管理发挥积极作用。

3.5 三“数”支撑

在智慧考试总体架构中,数据确权、数据安全和数据标准是确保考试数字化健康、平稳和持续发展的3个重要方面。

实现数据确权关系到考试的健康有序发展。数据已经成为一种与土地、劳动力、资本和技术

等生产要素并列的新型生产要素,数据确权是在数字化时代构建数据综合治理体系的重要一环。考试数据范围很广,包括与试题试卷相关的数据、与考生相关的数据以及在考试过程中产生的伴随数据等多种类型。只有明确这些数据的权属、责任和权利,才能够更好保障考试战线高效管理数据、合理利用数据,更好保障考生权益。

保障数据安全有利于考试的安全平稳发展。考试是对数据安全要求极高的行业,安全在考试发展的任何时期都是非常关键的因素。考试数字化发展依托的是考试各信息系统的支撑和服务,信息系统的数据安全是考试数字化发展的重要前提。针对不同系统的功能、性质和场景,加强信息安全等级保护或涉密信息系统分级保护,能够从不同的层面和维度切实维护考试信息系统安全。

规范数据标准有助于考试的持续融合发展。数字化发展必然要求数据的互通和共享。在教育考试领域已经有相关机构制定了关于考试数据存储、管理和共享的标准,提高了考试数据的规范性和交互性,有利于数据在各个平台、环节和系统之间高效流转和交换^[6]。在考试数字化发展过程中,研究和制定符合我国教育考试特点的数据交换标准,特别是针对试题和试卷等的关键数据制定新标准(China Question and Test Interpretability, CQTI),对构建整个智慧考试生态的底层数据基础具有重要意义。

4 未来展望

随着优质数据的持续积累、工程技术的升级迭代、算力水平的不断提升,由量变引起质变的新产品和新应用将不断出现在数据科学领域。当前,大语言模型的出现掀起新一轮技术变革^[7],也标志着“智慧涌现”的通用人工智能时代正加

速到来。目前,通用人工智能技术正在以迅猛的态势与行业领域和业务工作相融合,并已经产生更加智能的教育平台、更加便捷的办公系统以及更加高效的辅助工具。在不久的将来,通用人工智能技术将推动社会的全面智能化。从前文的论述中可以看到,智慧考试强调优质试题、考试相关数据资源的收集和标准规范的建立,强调应用需求导向和持续迭代升级,强调技术与业务的深度融合,其理念和架构与数据科学范式相统一,与数字化发展规律和发展方向相适配。在未来实践中,智慧考试将坚持守正创新,坚持基础性、开放性和灵活性相统一的原则,并将从以下两个方面付诸实施。

第一,基于总体架构进行主体推进,切实保障理念落地见效。按照体系构建、原型机研究、系统中试和应用迭代四步走的方案,进行规划和实施。智慧考试的体系已经初步构建,设计智慧考试原型机以及后续中试和逐步迭代就成为确保智慧考试从理念变为现实的关键举措。目前,结合国内外考试发展趋势,立足中国式教育考试的基本特点,围绕智慧考试总体架构,在主体推进方面已经梳理出6个主要实践方向,即试题呈现动态化、试题情景真实化、解题答题实操化、信息采集无感化、编码评分开放化和考试评价智能化。同时,稳步推进集试题创生、考试投送、编码评分、智能评价等核心功能于一体的原型机研发,初步勾勒出在可见的未来关于智慧考试的总体样貌,有关这些实践方向的科学性和可行性的相关研究和论证也在积极进行。目前,智慧考试原型机已经开展了两轮小范围中试,创新题型、编码评分等创新设计引起相关领域的强烈反响,得到教师、考生和专家的一致认可。

第二,基于创新技术进行局部突破,着力探索新模型新应用。人工智能和大数据技术已经

成为这个时代重要的科技引擎,及时跟踪了解、分析研判新技术的特点和优势,进而寻找切入点并实现深度融合是智慧考试在实践过程中保持活力和与时俱进的关键因素。大语言模型作为当前最为活跃的科技产品之一,正在成为新的创新驱动力和产业助推剂,其赋能传统行业已经产生了多种类型的应用助手和定制化产品^[18-19]。这种行业数据、应用场景、大语言模型三者融合的典型模式正在成为一种颇具代表性的创新模式。早在大语言模型发展初期,在医学领域已有专家学者进行了这种模式的探索。研究表明,在大语言模型中加入高质量的行业数据进行微调,能够在诸如试题自动生成这样的应用过程中产生特定的文件草案,人类只需要进行适当调整就可以得到预期效果^[20]。随着大语言模型日渐成熟,智慧考试可以借鉴利用这种模式,进而开发适用于教育考试的专用模型或定制化应用,服务于试题创生和智能评价等关键环节。这将为智慧考试带来新的活力,也将为教育考试智能化发展提供新的动力。

历史不断向前,时代不断进步。在科技变革的浪潮之中,数字化发展是建立数字化时代新优势的战略选择,是打造高质量发展新驱动的现实需要。智慧考试的本质属性和必然逻辑都决定了其自身要始终面对供与需的衔接和转换,既作为供给方为办好人民满意的教育考试提供更多的支撑和服务,又作为需求方不断吸收和融合更多创新技术;始终兼顾矛与盾的对立和统一,既需要通过先进数字技术来提高教育考试的现代化水平,又需要防范新技术对教育考试带来伦理及安全方面的影响和破坏;始终平衡历史与未来的过渡和碰撞,既以自信开放的胸怀充分吸收先进技术、勇于面对变革挑战,又要以客观谨慎的态度审视和判断技术创新的风险和隐患。可以预见,在今后很长一段时间,数字化变革带来的

机遇与挑战相互交织,相互转化。教育考试将持续面对新技术、新应用的不断冲击,特别是人工智能和大数据技术带来的深远影响。智慧考试作为教育考试进入数字化新阶段的重要理论和实践创新,更要把握历史主动性,抓住战略机遇期,积极识变、应变、求变,在服务拔尖创新人才选拔和助力人才综合评价等方面肩负起教育考试的历史责任。

5 结束语

智慧考试的初衷是通过考试数字化的发展推动我国从考试大国迈向考试强国。在国家教育数字化战略行动的大背景下,打造智慧考试的核心目标更是被提升定位在服务智慧教育上。总体而言,智慧考试将从3个层面实现对智慧教育的全方位服务。第一,汇聚考试资源,包括面向考试全环节的海量题库、命题素材、无感信息和考生作答等,对教育而言,这些都是含金量极高的数据资源;第二,强化考试服务,力求在教学端服务学生学习和学校教学,在教育端服务教育评价、服务教育改革;第三,创新考试生态,通过可及性更好的考试促进教育公平,通过准确性更高的评价保障教育质量,通过适配性更强的考试与智慧教育总体架构实现深度嵌套、全面融合。当然,这仅是基于现阶段数字化的发展情况、基于当前对数字化的理解和认识,也基于我国教育考试的自身特点和发展现状形成的智慧考试初步构想。随着时代的发展和进步,智慧考试将不断迈入新的阶段。

当今世界正在经历百年未有之大变局,3年来世纪疫情给国家教育考试的举办和考生参考造成很大困难。然而,危机之中总是孕育先机和转机,正如恩格斯所说:“没有哪一次巨大的历史灾难不是以历史的进步作为补偿的。”^[21]按照党的二十大作出的实施科教兴国战略、人才强国战

略、创新驱动发展战略的全面部署,在广大教育工作者的共同努力下,我国教育必将在开辟发展新领域、新赛道和塑造发展新动能、新优势中迎来一个引领世界潮流、实现古代先贤有教无类和因材施教千年梦想的数字化发展新时代。作为教育考试工作者,通过打造智慧考试、服务智慧教育,必将在我国迈向教育强国的新征程上作出新的更大贡献。

参考文献:

- [1] 习近平向 2022 年世界互联网大会乌镇峰会致贺信[EB/OL]. (2022-11-09) [2022-11-11]. http://www.qstheory.cn/yaowen/2022-11/09/c_1129113635.htm.
- [2] 习近平. 决胜全面建成小康社会 夺取新时代中国特色社会主义伟大胜利: 在中国共产党第十九次全国代表大会上的报告[EB/OL]. (2017-10-28)[2022-11-11]. <http://cpc.people.com.cn/n1/2017/1028/c64094-29613660.html>.
- [3] 习近平. 实施国家大数据战略 加快建设数字中国[EB/OL]. (2017-12-09) [2022-11-11]. <http://jhsjk.people.cn/article/29696290>.
- [4] 习近平. 高举中国特色社会主义伟大旗帜 为全面建设社会主义现代化国家而团结奋斗: 在中国共产党第二十次全国代表大会上的报告[EB/OL]. (2022-10-26) [2022-11-11]. <http://cpc.people.com.cn/20th/n1/2022/1026/c448334-32551867.html>.
- [5] 中共中央 国务院印发《数字中国建设整体布局规划》[EB/OL]. (2023-02-28) [2023-03-15]. <http://politics.people.com.cn/n1/2023/0228/c1001-32632549.html>.
- [6] 国务院. “十四五”数字经济发展规划[EB/OL]. (2022-01-12) [2022-11-11]. http://www.gov.cn/zhengce/content/2022-01/12/content_5667817.htm.
- [7] 教育部举行党组理论学习中心组集体学习暨教育信息化首场辅导报告会[EB/OL]. (2022-02-21)[2022-11-11]. http://www.moe.gov.cn/jyb_xwfb/gzdt_gzdt/moe_1485/202202/t20220221_600942.html.
- [8] 怀进鹏. 数字变革与教育未来: 在世界数字教育大会上的主旨演讲[EB/OL]. (2023-02-13)[2023-03-23] http://www.moe.gov.cn/jyb_xwfb/moe_176/202302/t20230213_1044377.html.
- [9] 以教育数字化战略引领未来: 教育部举行国家智慧教育平台启动仪式[EB/OL]. (2022-03-28) [2022-11-11]. http://www.moe.gov.cn/jyb_zzjg/huodong/202203/t20220328_611461.html.
- [10] 中共中央关于教育体制改革的决定(一九八五年五月二十七日)[J]. 江苏教育, 1985(增刊2): 6-12.
- [11] 于涵, 王蕾, 佟威. 国家题库建设: 从1.0到2.0: 新高考国家题库创新研究[J]. 清华大学教育研究, 2020, 41(1): 94-100.
- [12] 中共中央 国务院印发《深化新时代教育评价改革总体方案》[A/OL]. (2020-10-13)[2023-01-18]. http://www.gov.cn/zhengce/2020-10/13/content_5551032.htm.
- [13] 王蕾, 佟威. 从PISA编码看高考网上阅卷[J]. 教育科学研究, 2016(4): 33-38.
- [14] VELDKAMP B P, SLUIJTER C. Theoretical and practical advances in computer-based educational measurement[M]. Cham: Springer, 2019: 14-31.
- [15] GIFFORD C. Technology enhanced items in assessment [M]. A Pass Educational Group, LLC. 2017.
- [16] IEDTECH. IMS Global learning consortium[EB/OL]. [2023-03-20]. <http://www.imsglobal.org/home>.
- [17] 赵朝阳, 朱贵波, 王金桥. ChatGPT给语言大模型带来的启示和多模态大模型新的发展思路[J/OL]. 数据分析与知识发现: 1-12 [2023-03-10]. <http://kns.cnki.net/kcms/detail/10.1478.G2.20230320.1508.004.html>.
- [18] KHAN ACADEMY. Introducing... Khanmigo! [EB/OL]. [2023-03-15]. <https://support.khanacademy.org/hc/ens/community/posts/13992414612877-Introducing-Khanmigo->.
- [19] Introducing Microsoft 365 copilot[EB/OL]. (2023-03-16) [2023-03-20]. <https://www.microsoft.com/en-us/microsoft-365/blog/2023/03/16/introducing-microsoft-365-copilot-a-whole-new-way-to-work/>.
- [20] DAVIER M. Training optimus prime, M.D.: generating medical certification items by fine-tuning openAI's GPT2 transformer model[EB/OL]. (2019-08-21) [2023-03-15]. <https://arxiv.org/pdf/1908.08594.pdf>.
- [21] 恩格斯. 致尼古拉·弗兰策维奇·丹尼尔逊[M]//马克思恩格斯全集(第39卷). 北京: 人民出版社, 2016: 49.

Smarter Assessment for Smarter Education

YU Han

(National Education Examinations Authority, Beijing 100084, China)

Abstract: Digitalization is an essential step in the modernization of China's national education assessment system and is key to the further enhancement of education assessment quality. This paper focuses on the national strategy of digitalization in education and proposes a vision and a roadmap under the smarter education framework for developing a "smarter assessment" system, which reshapes the entire workflow and redefines the profession from a data flow perspective. Rooted in rich education assessment practice experience and relying on cutting-edge artificial intelligence technologies, the smarter assessment system provides data-driven functions and aims to provide all stakeholders with comprehensive services. The smart assessment system is designed with an overall layout based on the "3133" framework. Its core features include a massive item bank, adaptive test, and scientific evaluation. Its driving force is the item creation powerhouse, and the three core components are the digitalized assessment delivery sub-system, the intelligent coding and rating sub-system, and the intelligent evaluation sub-system. The emergence of AGI is causing a wave of comprehensive intelligence development in society. The overall framework and research results of the "smarter assessment" system have important theoretical value and practical significance for exploring new models, tools, and educational assessment applications suitable for the new intelligent era.

Keywords: digitalization; smarter assessment; smarter education; artificial intelligence algorithm; big data; system prototype

(责任编辑:王海东)