

# 教育大数据的应用模式与政策建议

杨现民<sup>1</sup>, 王榴卉<sup>1</sup>, 唐斯斯<sup>2</sup>

(1.江苏师范大学 江苏省教育信息化工程技术研究中心, 江苏 徐州 221116;

2.国家信息中心, 北京 100045)

[摘要] 大数据是推进教育创新发展的科学力量。教育大数据是整个教育活动过程中所产生的以及根据教育需要采集到的一切用于教育发展并可创造巨大潜在价值的数据集。与传统教育数据相比,教育大数据的采集具有更强的实时性、连贯性、全面性和自然性,分析处理更加复杂和多样,应用更加多元、深入。教育大数据的五层架构包括个体层数据、课程层数据、学校层数据、区域层数据以及国家层数据,自下而上汇聚各种教育数据。自20世纪70年代起,教育数据的应用大体经历了初始起步阶段、重点探索阶段和快速发展阶段,三个阶段的教育数据应用的规模和水平不断递进。教育大数据的应用主要体现在:驱动国家教育政策科学化;驱动区域教育均衡发展;驱动学校教育质量提升;驱动课程体系与教学效果的最优化;驱动个体的个性化发展。最后,针对当前我国教育大数据发展存在的问题和挑战,提出了六点政策建议:出台《教育大数据应用发展指导意见》,制定《教育大数据安全管理办法》,成立国家教育大数据研究机构,成立国家教育大数据治理机构,颁布“教育数据运营商”牌照,加快建设教育大数据产业基地。

[关键词] 教育大数据;教育数据;教育发展;应用模式;政策建议

[中图分类号] G434

[文献标志码] A

[作者简介] 杨现民(1982—),男,河北邢台人。副教授,主要从事智慧教育、移动与泛在学习研究。E-mail: yangxianmin8888@163.com。

技术变革教育的时代已经来临,以信息化带动教育现代化已上升为国家战略。<sup>[1][2]</sup>当前,我国教育发展面临诸多难题(减负、公平、质量提升、均衡发展等),而云计算、大数据、学习分析、物联网、移动通信等信息技术的快速发展为解决教育难题、促进教育领域综合改革与发展,提供了重要机遇和巨大的可能性。其中,大数据技术无疑是推进教育创新发展的科学力量。近年来,大数据不断对社会各个领域产生深刻影响,正在实现人类工作、生活与思维的大变革。同样,其“威力”也强烈地冲击着整个教育系统,正在成为推动教育系统创新与变革的颠覆性力量。

教育大数据是大数据的一个子集,特指教育领域的大数据,是整个教育活动过程中所产生的以及根据教育需要采集到的,一切用于教育发展并可创造巨大潜在价值的数据集。教育大数据之“大”并非只指数

量之大,而是更加强调“价值”之大,即能从繁杂的教育数据中发现相关关系、诊断现存问题、预测发展趋势,发挥教育大数据在提升教育质量、促进教育公平、实现个性化学习、优化教育资源配置、辅助教育科学决策等方面的重要作用。当前,大数据已经引起教育研究者、管理者、决策者以及实践者的关注。有学者分别从教育模式的转变、<sup>[3]</sup>教育可能的转向、<sup>[4]</sup>教育研究的新范式、<sup>[5]</sup>学习方式的变革<sup>[6]</sup>等方面探讨了大数据对教育发展带来的影响。还有一些学者从技术的视角探讨了学习分析与教育数据挖掘的方法与应用。<sup>[7][8]</sup>然而,关于教育大数据的数据范畴及其应用模式尚缺乏系统的梳理和分析。

基于此,本文旨在探讨四个方面的问题:教育大数据有何特征,包含哪些数据类型?教育数据的应用过程大体呈现出哪几个阶段?教育大数据在教育领域

基金项目:江苏高校优势学科建设工程资助项目“江苏师范大学教育学省优势学科建设”(苏政办发[2014]37号);国家级大学生实践创新训练计划项目“大数据支持的学习行为记录与发展性评价运用研究”(项目编号:201410320029Z)

的应用价值与模式是什么?教育大数据发展面临哪些问题,如何解决?

## 一、教育大数据的特征与结构

与传统教育数据相比,教育大数据的采集具有更强的实时性、连贯性、全面性和自然性,分析处理更加复杂和多样,应用更加多元、深入。传统教育数据的采集往往是阶段性的,多在用户知情的情况下(非自然状态)进行,分析的手段多采用简单的汇总统计和比较分析,关注的重点是受教育者的群体特征以及国家、区域、学校不同层面教育发展的整体状况。在大数据时代,移动通信、云计算、传感器、普适计算等新技术将逐步融入教育的全过程,可以在不影响师生教学活动的情况下实时、持续地采集更多微观的教与学的过程性数据,比如学生的学习轨迹、在每道作业题上逗留的时间、教师课堂提问与微笑的次数等。教育大数据的数据结构更加混杂,常规的结构化数据(如成绩、学籍、就业率、出勤记录等)依旧重要,但非结构化数据(如图片、视频、教案、教学软件、学习游戏等)将越来越占据主导地位。

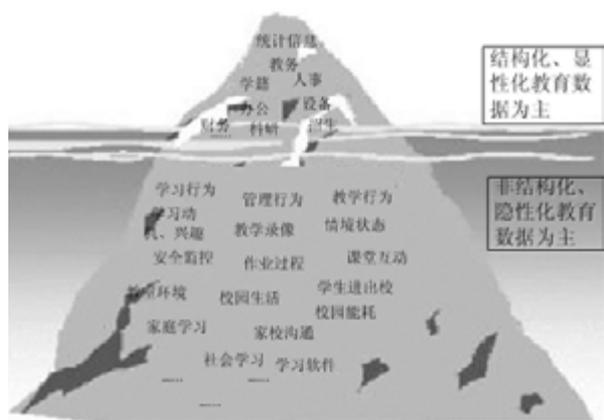


图1 教育大数据的“冰山模型”

教育数据每时每刻都在产生,然而教育领域究竟包含哪些数据?需要采集哪些数据?借鉴人力资源领域的人才素质“冰山模型”,可以构建教育大数据的“冰山模型”(如图1所示)。该模型将教育数据分为两大部分,分别是显露于冰面之上的数据和深藏于冰面之下的数据。多年来,国家采集的教育数据主要以管理类、结构化、结果性的数据为主,这些数据位于“冰面”以上,具有易测量、显性化等特点,重点关注宏观层面教育发展整体状况,在一定历史时期对我国制定教育政策、推动教育发展起到了积极的作用。然而,随着大数据时代的到来,国际社会对教育大数据战略资产的地位越来越认可和重视,教育

数据的全面化采集与深度挖掘分析就变得越来越重要。教育数据采集的重心将向非结构化的、过程性的数据转变,此类数据主要位于“冰面”以下,具有难测量、隐性化等特点。这些数据无论从数量上、增长速度上、还是潜在的价值上,都将远远超越传统的教育数据。

教育数据是客观的,其价值的发挥取决于操控和应用数据的人。因此,无论是冰面之上的数据还是冰面之下的数据都属于教育大数据的重要组成部分。只是从当前教育数据的采集与应用上来看,应当着重加强冰面之下部分教育数据的采集与深度挖掘,同时加强教育大数据与其他领域大数据(医疗、交通、经济、社保等)的融通和关联分析,进一步增强教育决策的科学性。为了更加清晰地认识教育大数据的概貌,这里根据教育数据的来源与范围,将其分成五层架构(如图2所示),从下向上汇聚各种教育数据。

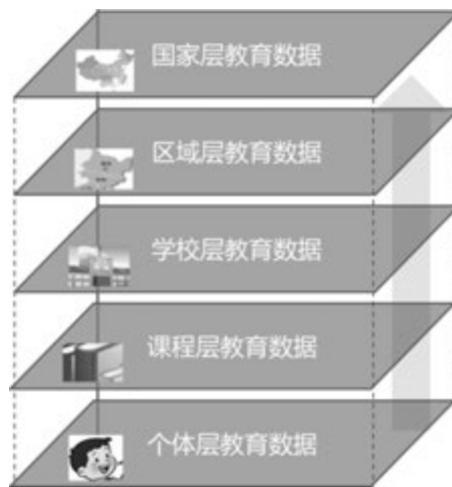


图2 教育大数据的分层架构

**个体层教育数据:**包括国家规定采集的教职工与学生的基础信息、用户各种行为数据(如学生随时随地的学习行为记录、管理人员的各种操作行为记录、教师的教学行为记录等)以及用户状态描述数据(如学习兴趣、动机、健康状况等)。

**课程层教育数据:**是围绕课程教学而产生的相关教育数据,包括课程基本信息、课程成员、课程资源、课程作业、师生交互行为、课程考核等数据,其中课程成员数据来自个体层,用于描述与学生课程学习相关的个人信息。

**学校层教育数据:**主要包括国家标准规定的各种学校管理数据(概况、学生管理、办公管理、科研管理、财务管理等)、课堂教学数据、教务数据、校园安全数据、设备使用与维护数据、教室实验室等使用数据、学

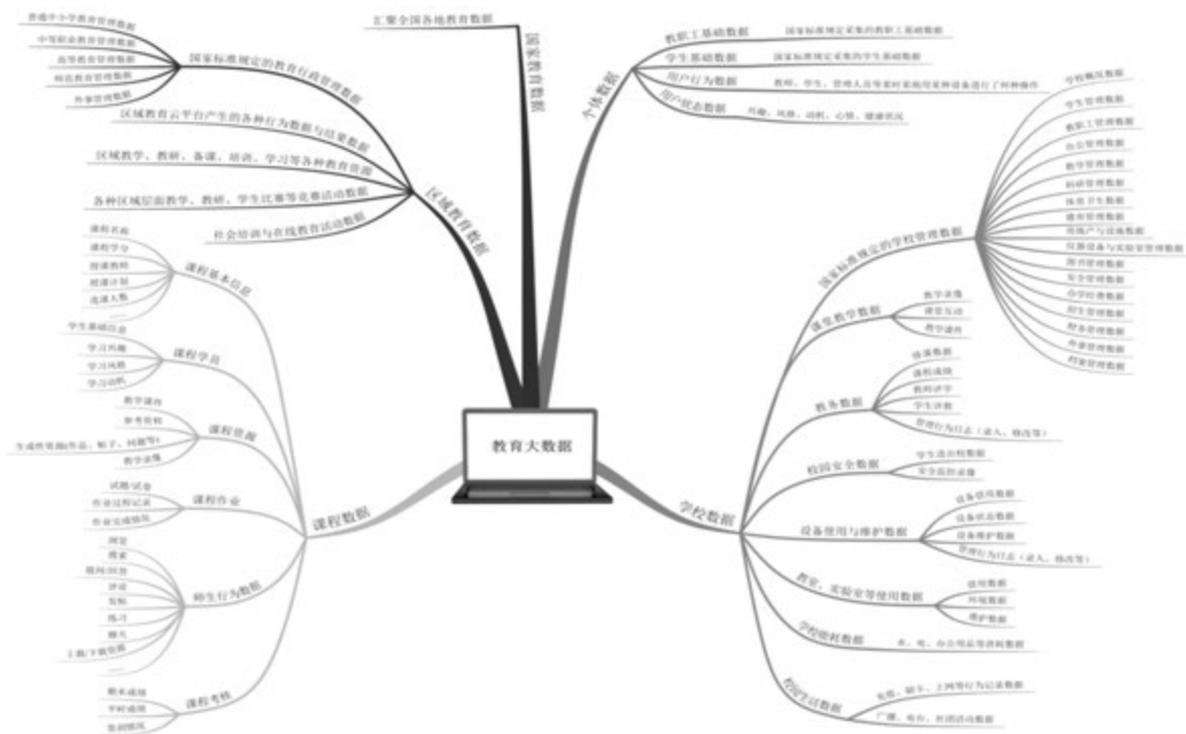


图3 教育大数据全景图

校能耗数据以及校园生活数据。

**区域层教育数据:**主要来自各学校以及社会培训与在线教育机构,主要包括国家标准规定的教育行政管理数据、区域教育云平台产生的各种行为与结果数据、区域教研训学等所需的各种教育资源、各种区域层面开展的教学教研与学生竞赛活动数据以及各种社会培训与在线教育活动数据。

**国家层教育数据:**主要汇聚来自各区域产生的各种教育数据。图3展示了各层教育数据包含的细分数据类型。

## 二、教育数据应用的发展阶段

随着人工智能、数据挖掘、机器学习等技术的发展,教育领域的数据处理与应用逐步走向成熟。从20世纪70年代算起,教育数据应用的发展阶段大体经历了初始起步阶段、重点探索阶段和快速发展阶段,这三个阶段的教育数据应用的规模和水平是不断递进的。

### (一) 初始起步阶段(1970—1997年)

人工智能产生于20世纪50年代中期,经过十几年的发展,到20世纪70年代开始进入教育领域。在AI技术的支持下,传统的CAI走向ICAI,随后发展为智能导师系统ITS。ITS主要包括专家模型、学生模型、指导模型和指导环境子系统,<sup>[9]</sup>通过对学习者的学习方法、学习习惯和学习过程等数据的收集与分析,

用以改善学习方法和提高学习效率。应该说,当时研究者对教育数据的应用水平有很高的期望,且已有了较清晰的应用思路。可惜的是,由于当时计算机使用成本较高、性能也较差,大大阻碍了该方法在教育领域的应用推广。

20世纪80年代中期,数据仓库技术开始兴起。由于其具有面向主题、集成性、时变性和非易失性特点,逐步成为数据分析和联机分析的重要平台。<sup>[10]</sup>数据仓库通过提供不同种类的应用系统的历史化数据,为各类教育机构进行全局范围的战略决策和长期趋势分析提供有效的支持。20世纪80年代末90年代初,数据挖掘和知识发现成为一个活跃的研究领域,各种专业性的国际会议陆续召开,掀起了数据挖掘与知识发现在各领域的研究与应用热潮。教育领域的研究者也开始关注数据挖掘与知识发现技术,进行了初步的尝试和探索,但由于网络、计算机等发展水平所限,对教育数据的挖掘利用也一直未有实质性的进展。

总的来说,该阶段计算机被初步引入到教育领域中,各种先进技术的应用范围以及程度都较低,采集到的教育数据也相对较少。该阶段的最大进展是实现了教育数据的数字化,从模拟走向数字为教育数据的永久存储、快速传播等提供了良好条件。其次,教育数据对教学的改善作用,开始引起了学术界的关注。但是,由于计算机普及性的不足以及数据库技术的滞

后性,教育数据在教育领域的应用价值未得到体现,应用水平处于初级阶段。

## (二)重点探索阶段(1998—2007年)

1998年,《科学》杂志上刊登了一篇介绍计算机软件HiQ的文章《大数据的处理程序》,第一次使用了大数据一词。但是,由于IT产业发展能力以及信息资源的产业利用在当时都还处于初级阶段,大数据的概念处于“萌芽期”,并未得到应有的重视。

2002年美国通过了《教育科学改革法》,明确了数据在教育决策中的决定性地位,“所有教育政策的制定都必须由实证数据进行支持”。经过多年的沉寂,2004年人工智能技术在高等教育领域发力,“智能导师系统”和“人工智能系统”再次开始流行,并掀起一股“教育数据挖掘”的研究热潮,同时也促进了学习分析技术的诞生。<sup>[11]</sup>自2005年起,在人工智能、人工智能教育应用及智能导师系统等国际会议上开展了多次“教育数据挖掘”主题研讨会,国际教育数据挖掘工作组以及国际教育数据挖掘协会也相继成立。教育数据挖掘主要研究数据挖掘技术在教育领域中的应用,从其研究领域的角度来看,教育数据挖掘研究包括“在教学研究中的应用”和“在教务管理中的应用”。<sup>[12]</sup>

这一阶段,计算机与网络技术飞速发展,在教育领域大范围普及应用。世界各国纷纷加大了对教育信息化的建设与推进力度,<sup>[13][14]</sup>出现了各种教育信息化应用系统(如学习管理系统、教务管理系统、人事管理系统、教学资源库等),各种教育业务的数字化、网络化促进了海量教育数据(学习者信息、学习过程数据、学习结果数据、教学管理数据等)的产生与存储。该阶段教育数据应用的最大进步是教育数据挖掘开始作为一个专门的研究领域得到来自计算机、信息管理、教育学等多学科研究者以及产业界的重视,发展迅速。此外,教育数据采集的数量和效率也大大提升。不足之处在于,由于对教育数据战略价值认识不足,加上教育数据挖掘技术仍处于发展时期,教育数据在教育领域应用价值的发挥仍未完全体现,大多数教育数据仍被用于简单的额统计分析以辅助教育和教学管理。

## (三)快速发展阶段(2008年至今)

2008年9月《自然》杂志版“Big Data”专刊,首次提出大数据一词。经过三年多的“发酵”,大数据在2012年引起了空前关注。全球数据量跨入了ZB时代,各种类型的数据呈指数增长,大数据时代已经到来。大数据技术不断对社会各个领域产生深刻影响,

正在实现人类工作、生活与思维的大变革。<sup>[15]</sup>在大数据时代背景下,世界各国纷纷加紧教育领域的大数据布局,出台相关政策和文件,数据驱动教育改革与发展已是大势所趋。美国教育部为推动“大数据”教育应用,于2012年10月发布了《通过教育数据挖掘和学习分析促进教与学》,指出要重点发展教育数据挖掘和学习分析技术,通过对教育大数据的挖掘与分析,促进美国高等院校及K-12学校教学系统的变革<sup>[16]</sup>。目前,我国教育部正在大力推进两级(国家、省)教育数据中心的建设,通过“两级建设、五级应用”实现对全国教育数据的统一规范管理。国内部分市、区也在依托智慧城市、智慧教育的建设与发展,大力发展区域教育大数据中心平台的建设与创新应用。

大数据时代的到来将教育数据的应用发展推向了“快车道”。云计算、物联网、大数据、移动通信等各种新型信息技术开始在教育的“沃土”上“生根发芽”。教育数据出现“爆炸式”增长,教育数据的采集更加实时、连贯、全面和自然,教育数据的分析处理更加复杂和多样,应用更加多元、深入。此外,教育数据作为重要资产的价值被逐渐认识和重视,教育数据挖掘和学习分析技术得到长足发展,正在被广泛应用到教学、管理、科研、评价、服务等教育的各个业务领域,规模效应正在凸显。教育数据的应用已经步入一个全新的历史时期,基于教育数据挖掘与学习分析技术,研发专用的教育数据分析决策模型、工具与算法,实现教育数据处理的高效能与数据应用价值的最大化,教育行业数据分析与应用体系的轮廓逐渐清晰。

## 三、教育大数据的应用模式

教育大数据是一种无形的资产,是一座可无限开采的“金矿”,充分的挖掘与应用是实现数据“资产”增值的唯一途径。从某种意义上来说,未来的国际教育竞争将是教育大数据舞台上的对抗。那么,教育大数据究竟如何发挥作用,有何应用模式?接下来,将依据上文提出的教育大数据分层架构,从不同层面论述大数据在教育中的应用模式。

### (一)教育大数据驱动国家教育政策科学化

数据在国家教育决策的制定方面起着关键性作用,而大数据时代的到来将使数据的收集和分析更加方便、快捷、全面、准确,教育政策的制定将更加依靠数据说话。

传统教育数据的采集渠道和数量是非常有限的,

仅能通过简单的统计分析反映某个时间段国家教育发展的局部状况,数据的预测价值难以发挥,无法为国家教育政策的制定提供科学支撑。大数据技术具有数据海量、途径多元化、挖掘深度化等优势,能够发现多种教育数据之间以及与其他社会行业数据之间的内在联系,有助于构建更加系统化的教育发展模型,推动国家教育政策制定与调整的科学化。此外,基于数据的教育决策能够增强广大民众对教育政策的理解和支持。<sup>[17]</sup>

美国是最先确定教育数据的战略地位并制定相关保障法案(《教育科学改革法》)的国家,早在2002年美国就以立法的形式明确指出,所有教育改革与决策必须有实证数据支持。决策科学化将使得教育整体成本呈下降趋势,同时实现教育质量和教育公平的大幅提升。目前,我国正在推行的学生“终身一人一号”电子学籍管理办法,为持续性记录每个学生的学业表现与全面发展情况提供了制度保障。如果能够由此建立全国联网的学生成长档案库,辅以家庭、教师、学校的多样化数据,这将为全国高考招生与录取政策、学生就业政策、资源分配政策、学生择校政策等当前众多教育政策的改善与优化,提供最宝贵的数据支持。

## (二)教育大数据驱动区域教育均衡发展

区域教育均衡发展是我国教育事业面临的重大现实问题。应用大数据技术可以准确把握区域教育发展动态和影响其均衡发展的关键因素,从教育环境均衡、教育资源均衡、教育机会均等、教育质量均衡等方面全面推进区域教育的均衡发展。<sup>[18]</sup>此外,不同区域有自己不同的教育现状,在大数据背景下不仅可以缩小区域间的教育差距,而且有助于不同区域根据自身环境条件、经济状况以及发展需要形成各具特色的区域教育发展路径。

实际上,目前一些国际组织已经开始将大数据技术应用到教育资源的分配上。联合国教科文组织发布的最新一份《全民教育全球监测报告》显示,全世界用于教育的宝贵经费正在被低质量教育所浪费,其损失金额高达每年1290亿美元。报告就此呼吁各国政府在投资教育时,要注重教师的数量与质量,确保把最好的教师配备给最需要的学生。<sup>[19]</sup>大数据可以帮助教育部门负责人和决策者了解资金政策的影响,美国弗吉尼亚州高等教育委员会政策研究和数据仓库的负责人Tod R. Massa说:“我们的目标是创造一种环境——所有讨论都是基于事实,而不是直觉或者猜想”。<sup>[20]</sup>

统一学籍信息管理制度为我国教育大数据的采集、管理与应用提供了重要保障。学生的入学、转学、休学、退学等教育管理数据,可以实现全面、实时的采集、监控、更新与分析处理。教育管理数据还可以与家庭收入、户籍、医疗、保险、交通等数据进行关联分析,有助于及早发现与预测学困生、择校生等需要进行教育帮助和干预的学生,进而提供针对性的教育支持服务,保障每位学生平等接受优质教育的机会。

此外,通过建立连续的制度化的区域教育发展数据采集机制,可以全面跟踪了解所有学生的在校学习情况以及毕业后的工作情况,进而更加客观的评价区域教育质量,根据评估结果动态调整区域教育体系,比如专业调整、课程计划调整、培养方式的调整等,实现教育与社会需求之间更加无缝的对接,帮助每一位学生获得成功。

## (三)教育大数据驱动学校教育质量提升

随着各种智慧教学与管理平台的不断涌现,大数据对学校教育的变革作用将不断凸显。大数据在提升学校管理质量和教学质量以及完善教育评价手段上具有独特的优势。

数字校园的建设大大推动了学校管理的数字化和网络化,<sup>[21]</sup>办公自动化系统、资产管理系统、教务管理系统、科研管理系统等各种应用系统为教育管理数据的实时采集和深度挖掘提供了条件。目前,国内已有一些高校率先开展基于大数据的教育管理服务。浙江大学对学校的设备资产数据进行了系统采集与整理,提供便捷的查询与分析服务,提升了实验室、教室、仪器、设备等资源的利用率和管理效率。江南大学通过物联网技术对学校用水、用电等数据进行全面监控和优化处理,实现了节能环保。华东师范大学利用学生的餐饮消费数据,对经济困难学生提供情感抚慰和助学金支持。此外,大数据还可以在教师招聘上发挥重要作用。通过对应聘者的个人信息进行分析和预测,从而将更有可能成功和更适合的教师招聘进来。美国一些学区开始与大数据公司合作,应用大数据工具辅助教师招聘。通过对教师的学位和专长以及信仰、人生观、态度、经验开放性等因素的分析,结合面试结果,综合决定教师是否被聘用。

除了学校管理服务质量的提高,通过应用大数据技术对海量教学数据的分析与预测,还将改变传统的千篇一律的教学模式,实现高质量、个性化的教学。大数据能够全面记录学习者的成长记录并进行科学分析,让学习者更了解自己,帮助教师预测学生成绩,并

为学习者提供科学的学习建议,帮助其提高学习成绩。加拿大的 Desire2Learn 科技公司面向高校研发了“学生成功系统”,该系统主要基于学生已有的学习成绩数据来预测并改善其在未来课程学习中的表现,并将分析结果详细地呈现给教师,以便教师进行个性化指导。

大数据技术还能为学生的学业成就评价提供全面的数据支持。学生的所有学习过程与结果数据将存储到学习档案袋,教师的所有教学过程与结果数据将存储到教学档案袋。基于档案袋数据可以建立科学的发展模型,定期评估学生和教师的发展情况,并提供相应发展建议。此外,大数据技术在学校科学研究活动中也大有用武之地。一方面,教育数据可以作为教育研究的数据源,通过深度数据挖掘,可以透过教育问题的表象发现本质,产出高质量的实证研究成果;另一方面,海量、多维数据的分析能够发挥“科研指南针”的作用,有助于研究者准确把握研究领域的前沿议题和走向。

#### (四)教育大数据驱动课程体系与教学效果的最优化

当前,国家正在积极推动部分本科高校向应用型大学转型,以顺应社会发展对高等教育提出的挑战。在转型过程中的一个核心问题便是开什么专业,如何设置课程体系。此时,便需要通过有效整合专业、行业、区域经济与社会发展等诸多数据,通过全面的数据采集和深入分析,准确把握市场对人才的需求,明确各应用型专业人才培养的目标,构建能力素质模型,配套适合的课程体系。

实际上,我国各级教育部门、各类学校已经通过学籍管理系统、教务管理系统、学位管理系统等,积累了大量的学生入学、毕业和课程设置数据,但这些数据基本处于“休眠”状态。对学校历届学生的课程成绩、授课方法、就业情况等诸多数据进行关联分析,可以找出影响课程成绩及就业的关键因素,为学校课程的调整和课程教学方法的优化提供可靠依据。<sup>[22]</sup>

通过大数据技术还可以持续跟踪教师授课历程,分析教师的教学特征及其优势,判断教师适合承担哪些课程的教学任务,采用何种方法可以达成最优化的教学效果。通过智能化的网络教学平台,教师可以提前准确判断学生的兴趣点、知识基础、学习偏好、学习难点等,从而进行针对性的教学。基于学生的在线学习数据可以构建多种预测模型,预测可能存在学习困难、有退学风险的学生,并进行及时的干预。美国普渡

大学的“课程信号”项目是国际知名的大数据教育应用典型案例之一。<sup>[23]</sup>该项目研发了课程学习预警平台,采集大量学生课程学习的过程性数据,通过一套预测算法分析学生课程学习成功的概率,根据预测结果教师可以给予针对性的帮助、指导和反馈,显著提高了课程学习的成功率。

此外,通过分析学生在课程资源上的点击、浏览、翻页、收藏、评价等操作轨迹,可以客观评价课程资源的受关注度、资源界面设计的合理性、资源导航的高效性以及对学生学习效果的影响,从而完善课程资源的结构与内容,实现大批优质课程资源的生成与进化。

#### (五)教育大数据驱动个体的个性化发展

当前的标准化人才培养模式已经不能适应信息社会的发展要求,大数据为人才培养模式的创新提供了条件和机遇。从大众化走向个性化已成为未来人才培养的重要趋势。学生和教师是教育领域人才培养的两大核心主体,大数据将真正推动学生和教师的个性化发展。

个性化发展的前提是个体首先要真正认识自己,知道自己的优势、不足、兴趣、偏好、风格、知识缺陷、能力缺陷、发展目标等;其次,需要提供最适合个体发展的环境、资源、活动、工具、服务等外部条件。而大数据的最大优势便是让学生和教师认识每个真实的“自我”,同时通过学习行为、教学行为数据的深度挖掘与分析,为每个真实的“自我”推送最合适的学习资源与学习路径。

网络学习虽然具有天然的“个性化”优势,然而缺少了大数据的支持,机器将无法真正了解每位学习者,也就无法实现个性化资源与服务的推送。如果说互联网促进了教育的民主化,那么大数据将实现教育的个性化。<sup>[24]</sup>在大数据的支持下,教师能够更有效地关注每个学生个体,记录每个学生的学习轨迹,分析每个学生的学习行为、预测其学习结果、诊断其学习需求与问题,开展真正的因材施教。教师逐步由教学者转变为帮助每位学生个性化学习与发展的指导者。传统的学习管理系统将升级为智慧学习平台,能够持续采集学习者的学习行为数据,并进行智能分析,依据学习者模型推送适合的学习资源,准确诊断、评价学习过程与结果,反馈给学习者最适合的学习建议,实现每位学生的个性化发展。

在大数据促进教师专业发展方面,有学者指出<sup>[25]</sup>:教育大数据可以提升作为在线学习者的教师的学习效率,激发其自主的专业发展意识;可以提高教

师作为在线教学者的教学效率,发展其在线教学实践智慧;可以提高教师作为研究者的研究绩效,提升其对学生在线学习的服务能力;可以提高教师作为管理者的管理效率,提升其在线教学领导力。实际上,教师的专业发展不仅仅体现在专业思想、专业知识与专业能力的提高上,更体现在教师成为一个独特的教学个体。通过大数据技术的支持,教师能够认识最真实的自我,彰显教学个性和智慧,开展灵活多样的个性化教学,最终实现个体的个性化专业发展。

#### 四、发展教育大数据的政策建议

大数据时代已经到来,随着政府、企业、教育机构、广大民众的数据意识逐步提高,可以预见大数据在教育领域发展应用的广阔前景。目前,教育部已经制定了《国家教育管理信息系统建设总体方案》,其首要任务就是要建立全国教育基础数据库,以支持教育事业更快、更好、更科学的发展,这标志着我国政府对教育大数据的管理与应用正在步入实质性阶段。此外,在线教育市场中的一些企业也开始着手教育数据产品和服务的研发与推广,爆炸式增长的“苗头”已经出现。

与商业、交通、环境、医疗等领域相比,教育领域具有更强的独特性和复杂性,大数据技术在教育领域的应用推广仍存在诸多难题,诸如教育环境的限制难以获取学习者线下的行为数据,教育数据的部门间共享仍存在制度壁垒,教育数据的规范化动态采集与实时更新的机制仍未建立,学习者隐私数据的泄露与不当应用的风险较大,等等。上述问题直接影响了教育大数据在教育教学领域应用推广的进程,导致大数据技术难以在短期内发挥实质性作用。此外,随着教育领域大数据产业的快速发展,教育数据的安全、治理、运营等诸多问题和挑战也将越来越凸显。

为了保障教育数据在采集、存储、分析、应用、管理等各环节的规范性,促进教育大数据在我国的良性发展,急需制定相关政策进行引导和监督。

##### (一) 出台《教育大数据应用发展指导意见》

为抢抓机遇,推进教育大数据事业与产业的健康快速发展,急需制定相应的应用发展指导意见。建议由相关中央部委牵头制定《教育大数据应用发展指导意见》。其核心内容包括:从国家层面加大对教育大数据应用、推广的支持力度,将教育大数据应用提升到更高的战略层面;加强教育数据采集的广度和深度,

规范采集流程,加大教育数据的挖掘与分析力度,为个性化教育的开展以及科学教育教学管理政策的制定提供基础;界定各级教育主管部门、学校、相关企业的职责,突出教育发展需求,注重教育数据的应用成效,杜绝浪费和面子工程;鼓励教育大数据技术产品开源,工信部电信研究院的《大数据白皮书》分析称“大数据技术发展及开源运动的结合是大数据技术创新中的一个鲜明特点”,教育大数据也应走向开源、开放;从资金、人才、政策等方面给予相应支持,引导教育大数据产业健康有序发展。

##### (二) 制定《教育大数据安全管理办法》

教育数据既是一笔宝贵的教育资产,同时也涉及到教育者和受教育者的隐私,应用保护不当则会带来严重的安全风险。国家应高度重视教育数据的隐私保护与安全管理,不断努力采取更先进、安全系数更高的措施来保障教育数据的安全,保护教育隐私数据不外泄、不被恶意使用。因此,我国需加快制定《教育大数据安全管理办法》,从体制、机制、技术、方法等多个层面制定管理细则,切实保障个体、机构、国家的教育数据安全。

《教育大数据安全管理办法》的核心内容应包括:建立健全教育大数据安全管理架构,包括数据生产部门、数据使用部门、数据管理部门等,并明确各部门的安全管理职责;建立教育数据资源的保密等级,按保密等级规定采取相关处理措施;设计高性能、高可靠、简化管理的教育大数据存储系统,为教育云平台提供充分的访问性能和数据安全保障;建立涵盖教育数据存储、传输、使用等多个环节的多种安全保障措施;采用数据隔离、数据加密、第三方实名认证、灵活转移、安全清除、完整备份、时限恢复、行为审计、外围防护等多种安全防护技术,解决教育大数据的云存储安全问题。

##### (三) 成立国家教育大数据研究机构

组织教育学、管理学、计算机科学、统计学等多学科的研究人员成立专门的教育大数据研究机构,集中优势力量破解教育大数据应用推广过程中存在的热点、难点问题,同时结合教育发展的战略需求,开展前瞻性研究,使其成为国家教育大数据发展的智库。

国家教育大数据研究机构的主要职责包括:深入钻研教育大数据技术,破解教育数据采集、存储、处理与应用过程中存在的关键性技术难题,着力研究教育大数据的应用模式与方法,更好地服务个体、学校、教育管理机构以及教育企业的发展;发挥好国家级智库

的创新与引领作用,研发教育大数据应用的相关技术规范、标准,规范教育大数据产业发展;结合国情构建特色化的教育大数据理论体系,涵盖理论、技术、模式、方法、机制、战略等多个层面,努力使我国在智慧教育与大数据研究方面从“随行者”变为“领跑者”;归纳、提炼典型的教育大数据应用案例与模式,总结教育大数据应用策略,为全国推广教育大数据应用经验提供指导。

#### (四)成立国家教育大数据治理机构

除了教育大数据研究机构,还应成立教育大数据治理机构。大数据时代,每时每刻都在产生海量的、各种来源的、各种类型的教育数据,如何协同多方力量进行高效的教育数据治理是摆在每位教育管理者面前的现实问题。教育数据治理的目的是提升教育数据质量,保护教育数据隐私安全,保障教育数据合理应用,促进教育数据合法共享。

国家教育大数据治理机构的主要职责包括:出台教育大数据治理的相关办法,建立完整的教育大数据治理模式,保护教育数据的隐私安全,指导教育数据的获取、归档、保存、互换以及重复利用;将教育数据库、教育服务平台、教育资源平台等产生的数据进行规范化采集与汇聚共享,构建综合性的一体化的教育大数据治理体系;制定教育基础数据采集标准,设置相关要求,形成清晰的数据治理机制和流程;制定教育数据质量标准和教育数据管理战略,确立教育数据集的归档和长期保存的机制与方法;建立教育大数据开放平台,发挥企业、个体、教育机构等主体的创造性,开发大量特色化的教育应用,让社会力量共同参与教育数据的治理与创新应用;监控教育大数据的应用状况,对非法应用、侵犯用户隐私与国家安全的单位和个人进行处罚,建立教育大数据合法、合理应用的良好环境。

#### (五)颁布“教育数据运营商”牌照

从国家层面来看,教育数据的安全性不亚于金融

数据。教育数据的开放程度、开放范围、开放对象都应进行深入的论证,以确保教育数据的使用是合法的,是有利于教育创新发展的,而非阻碍、破坏、威胁到国家安全。为此,政府部门应对教育数据加强监管,参考通信领域颁发运营商牌照的方式制定教育数据运营商准入标准,颁发运营牌照。建议由教育部、发改委、商务部等多个部委联合制定相关规定,对国内教育数据应用行业实施正式的监管,设置教育数据使用的标准和门槛以及相关申请资格与办法,颁布“教育数据运营商”牌照,非官方机构进行教育数据研究与使用,需要按条件满足相关规定并取得相关许可,才能获得教育数据的使用权利。

#### (六)加快建设教育大数据产业基地

以基于海量数据资源的挖掘和应用催生的大数据产业为基础的新经济是社会经济发展到一定阶段的必然选择。教育大数据产业作为大数据产业家族的一部分,同样是社会经济与教育发展的必然选择。当前,加快教育大数据产业基地建设,既是智慧教育事业发展的需要,也是大数据在教育领域发挥变革作用的需要。

为加快建设教育大数据产业基地,提出如下几点建议:多渠道增加教育大数据产业基地的投入;打造适应教育大数据产业发展的配套软硬件环境;加快成立教育大数据研究机构,产生一批引领型企业;打造教育大数据产业的基础配套设施、后台保障,力争打通全产业链条;领先企业进行前沿技术创新,创新成果通过开源得到不断完善并向全社会辐射,集众人之智实现教育大数据产品与服务的持续创新发展;成立教育大数据产业联盟和园区,综合投资机构、高校与科研机构、政府部门的力量,邀请国内外知名专家学者和产业人士定期研讨、指导产业基地发展,努力将基地建设成为国内外知名的教育大数据产业集群地和引领教育大数据产业发展的核心地带。

#### [参考文献]

- [1] 陈琳,陈耀华.以信息化带动教育现代化路径探析[J].教育研究,2013,(11):114~118.
- [2] [14] 陈琳.促进深层学习的网络学习资源建设研究[J].电化教育研究,2011,(12):69~75.
- [3] 杨淼淇,孙纳新,柴华.大数据时代教育模式的研究[J].计算机工程与科学,2014,36(A01):272~273.
- [4] 喻长志.大数据时代教育的可能转向[J].江淮论坛,2013,(4):34.
- [5] 祝智庭,沈德梅.基于大数据的教育技术研究新范式[J].电化教育研究,2013,(10):5~13.
- [6] 徐鹏,王以宁,刘艳华,张海.大数据视角分析学习变革——美国《通过教育数据挖掘和学习分析促进教与学》报告解读及启示[J].远程教育杂志,2013,(6):11~17.

(下转第69页)

## [参考文献]

- [1] 钟启泉. 教学活动理论的考察[J]. 教育研究, 2005, (5):36~42.
- [2] 夏家发. 教学活动设计[M]. 武汉:华中师范大学出版社, 2010:5.
- [3] 魏宁. 信息技术支持的教学分析方法——S-T 篇[J]. 信息技术教育, 2006, (1):55~57.
- [4] 魏宁. 信息技术支持的教学分析方法——FIAS 篇[J]. 信息技术教育, 2006, (2):60~62.
- [5] James, W., Stigler, P.G.T.K., Steffen Knoll, A.A.S.. The TIMSS Videotape Classroom Study[Z]. 2001.
- [6] 王晶莹. TIMSS 五国八年级科学教学录像研究述评[J]. 中小学电教, 2011, (Z1).
- [7] 顾小清, 王炜. 支持教师专业发展的课堂分析技术新探索[J]. 中国电化教育, 2004, (7):18~21.
- [8] 吴莉霞. 活动理论框架下的基于项目学习(PBL)的研究与设计[D]. 武汉:华中师范大学, 2006.
- [9] 陈瑶. 课堂观察指导(第1版)[M]. 北京:教育科学出版社, 2002:42~43.
- [10] An Interaction Analysis: A Teacher's Questions, Feedback, and Students' Production through Classroom Observation[DB/OL].[2015-02-15] <http://www.cels.bham.ac.uk/resources/essays/yamazaki1.pdf>.

## (上接第61页)

- [7] 顾小清, 张进良, 蔡慧英. 学习分析:正在浮现中的数据技术[J]. 远程教育杂志, 2012, (1):18~25.
- [8] 葛道凯, 张少刚, 魏顺平. 教育数据挖掘:方法与应用[M]. 北京:教育科学出版社, 2012.
- [9] 维基百科. Adaptive learning [EB/OL]. [2014-07-01]. [http://en.wikipedia.org/wiki/Adaptive\\_learning](http://en.wikipedia.org/wiki/Adaptive_learning).
- [10] 陶雪娇, 胡晓峰, 刘洋. 大数据研究综述[J]. 系统仿真学报, 2013, (1):142~146.
- [11] 魏顺平. 学习分析技术:挖掘大数据时代下教育数据的价值[J]. 现代教育技术, 2013, (2):5~11.
- [12] 李婷, 傅钢善. 国内外教育数据挖掘研究现状及趋势分析[J]. 现代教育技术, 2010, (10):21~25.
- [13] 陈琳. 中国高校教育信息化发展战略与路径选择[J]. 教育研究, 2012, (4):50~56.
- [15] Mayer-Schönberger, V., Cukier, K.. Big Data: A Revolution That will Transform How We Live, Work, and Think [M]. New York: Houghton Mifflin Harcourt, 2013.
- [16] 陆璟. 大数据及其在教育中的应用[J]. 上海教育科研, 2013, (9):5~8.
- [17] Schleicher, A.. Big Data and PISA [EB/OL]. [2014-04-25]. [http://www.huffingtonpost.com/andreas-schleicher/big-data-and-pisa\\_b\\_3633558.html](http://www.huffingtonpost.com/andreas-schleicher/big-data-and-pisa_b_3633558.html).
- [18] 刘雍潜, 杨现民. 大数据时代区域教育均衡发展新思路[J]. 电化教育研究, 2014, (5):11~14.
- [19] 通信世界网. 数据挖掘人工智能使教育定制化 [EB/OL]. [2014-05-01]. [http://www.cww.net.cn/tech/html/2014/3/10/20143101127294557\\_2.htm](http://www.cww.net.cn/tech/html/2014/3/10/20143101127294557_2.htm).
- [20] Gall Button. 大数据进校园[EB/OL]. [2014-05-01]. <http://weibo.com/2357403745/AFfeR4HCu?mod=weibotime>.
- [21] 王运武. 我国数字校园建设研究综述[J]. 现代远程教育研究, 2011, (4):39~50.
- [22] 杨永斌. 数据挖掘技术在教育中的应用研究[J]. 计算机科学, 2006, 33(12):284~286.
- [23] Pistilli, M.D., Arnold, K.E.. In Practice: Purdue Signals: Mining Real-Time Academic Data to Enhance Student Success[J]. About Campus, 2010, 15(3):22~24.
- [24] 翟博. 教育均衡发展:理论、指标及测算方法[J]. 教育研究, 2006, (3):16~28.
- [25] 张进良, 何高大. 学习分析:助推大数据时代高校教师在线专业发展[J]. 远程教育杂志, 2014, 32(1):56~62.