



教育大数据生态圈构建：“3+3”模型的逻辑与实践*

□ 吴南中 黄治虎 曾 靓 谢青松 夏海鹰

【摘要】

《教育信息化2.0行动计划》将教育信息化作为教育系统性变革的内生变量，教育大数据的价值期待被无限扩大。教育大数据来源成为主要掣肘，探索积极的教育大数据生态体系建设，解决教育大数据来源和应用问题，提升教育大数据对教育教学支持程度显得尤其重要。文章通过开发研究与行动研究，归纳了教育大数据生态圈构建的基本特征：大数据生态圈的构建是有目标、有计划、有体系的整体构建；蕴含多样性、广泛性、全程性、持续性等特性；反映了人在环境实践改造中的能动性；对教育发展的价值引领性和自身具备一定的成长性。基于此，本研究尝试构建学校层面的“3+3”教育大数据生态圈模型，并开展实践。结果表明，“3+3”教育大数据生态圈的构建对学校理解和应用好教育大数据有促进作用。

【关键词】 大数据；“互联网+教育”；生态圈；建构；学习分析；模型；逻辑；实践；数据采集

【中图分类号】 G434

【文献标识码】 A

【文章编号】 1009-458x(2019)7-0077-09

DOI:10.13541/j.cnki.chinade.20190226.005

《教育信息化2.0行动计划》正式开启了信息化建设的第二阶段，提出了“信息技术与智能技术融入教育全过程，推进改进教学、优化管理、提升绩效”的发展目标。“加强教育信息化从研究到应用的系统部署、纵深推进”的战略任务，需要在教育的方方面面落实。在这样的目标体系下，教育大数据作为沟通物化的环境、符号化环境和人的反馈的载体，蕴含了更加理性地认识教育本质和功能、认识学习者学习过程和心理世界、改进教育研究的方法、提供人本化教育等认知价值和实践价值，备受学者重视。杨现民等(2016)认为教育大数据正在“成为推动教育系统创新和变革的颠覆性力量”。其他学者也提出基于教育大数据的教学模式转变(郑立海, 2015)、学习方式转变(徐鹏, 等, 2013)、研究范式转变(吴南中, 等, 2017)等。也有学者从理论上构建教育大数据的整合平台，如徐珊(2014)提出了基于XML的数据整合模型；孟玲玲等(2012)基于元数据标准设计了知识共享系统；常燕(2013)设计了基于本体技术的异构教育资源数据集成框架。“但鲜有学者对大数

据平台架构进行统一、完整的设计”(李振, 等, 2018)。有的从应用层面构建一些数据收集系统，起到了一定的数据收集效用(刘宏达, 等, 2018)，但很多文章对应用数据捕捉效果和应用有效程度语焉不详。综上所述，教育大数据促成了教育发展的新愿景，获取数据成为教育大数据发展和利用的掣肘，探讨形成大数据获取的环境构建，构建教育的数据生态圈，通过对教育大数据的捕捉、聚合、理解，促使其成为教育发展和变革的元素、动力、方法与支撑，进而形成可持续的信息化发展机制，积极支持国家信息化新战略，成为教育研究者关注的问题。

一、教育大数据应用掣肘：数据在哪里？

(一) 教育大数据内涵及其来源

1. 教育大数据内涵

互联网时代的经济运行发现，数据达到一定量之后，各种相关关系和关联关系隐匿在传统因果联系无法识别和解释的框架中。随后，有人总结，大数据体

* 本文系2017年度重庆市教委科学技术研究项目“在线教育大数据挖掘技术与理解模型建构研究”(编号：KJ1737457)和2017年重庆市社会科学规划项目“以信息化推进教育精准扶贫的机制与路径研究”(编号：2017YBWT09)的阶段性研究成果。

现了数据量大、输入和处理速度快、数据多样和精准性等内在特性 (Doug, 2017), “数据的交换、整合、分析, 不断发现新的知识、创造新的价值, 从而带来了大知识、大科技、大利润和大发展” (徐子沛, 2012, p. 12)。教育大数据是指在教育体系中的一切数据, 包括了发生在教育领域内各个在线应用中的数据, 也包括大教育场景中的其他数据 (如学校、社会、家庭等), 还包括教育管理和教学管理的相关数据等。整个校内校外数据, 构成了教育大数据形成环境, 是一个教育时间发展和空间拓展的概念。由于教育大数据要服务于特定目标的人群, 过于宽泛的内涵会限制教育大数据的应用落地, 需要对教育大数据进行分层: 宏观层面上, 教育大数据是指在广义的教育中发生的一切数据, 涵盖全员 (所有人)、全过程 (从出生到死亡)、全方位 (学习、生活、工作和闲暇); 中观层面上, 教育大数据指的是人在教育体系中发生的一切数据, 包括学前教育、初等教育、高等教育与职后教育; 微观层面指的是在学校体系所发生的数据, 包括学习数据、生活数据和相关管理数据。从现有的研究范畴来看, 后两者是研究的主要领域, 前者着重于决策, 后者着重于微观的应用。本研究着眼于学校数据生态, 在研究对象选取上指的是在一个学校中的所有数据和与外界的数据沟通, 包括教学过程产生的数据、教育管理采取的数据、科学研究与校园生活产生的数据, 也包含利用校外资源开展教育教学活动, 提升生活体验等数据, 是一个以校内数据为主的数据生态。从实践来看, 教育大数据具有三种特色: 第一, 数据来源复杂化。教育大数据既有各类学习者使用终端的数据, 也有主动的情景感知终端的数据; 既有蕴含过程各类结构化数据和非结构化数据, 也有展示结果的标准化数据和非标准化数据。第二, 数据应用对称化。教育大数据在理论上讲, 既能利用于教学管理, 也能利用于学生自主使用; 既能便利学生管理的数据挖掘, 也能适应学生学习的自我调适。可以说, 多种教育主体都能通过教育大数据找到其用处, 这与商业情境中的数据挖掘和利用仅围绕公司和企业的利益最大化有一定区别。第三, 数据应用多层次化。在商业数据中, 大数据主要用于预测关系和相关关系的挖掘; 在教育大数据中, 数据利用具有多层次化的特征: 在宏观层面, 可以用于教育政策制定; 在中观层面, 可以用于解释和预测;

在微观层面, 可以用于教学策略的适时调整和学习方式的即时转变。

2. 教育大数据来源

在社会经济领域, 大数据主要通过社会应用 (如借助移动终端) 来收集数据, 用于识别和预测日常生活、人际关系等 (Eagle & Pentland, 2006)。这种层面的数据收集忽视了环境信息和个性化信息的处理, 在教育应用层面的个体画像中存在缺陷, 也无法提供教育大数据核心的价值归旨——服务于个性化学习。发挥教育大数据的作用, 需要围绕学习者建立一系列的数据, 才能开展基于数据的挖掘和分析。研究者进行了大量的探索, 比如有学者基于在线教学平台, 通过事先“埋点”的布局, 以量化的方式获取数据, 通过数据的增益分析学习者的进步程度 (Seel, 2012, p.1176)。ADL 使用学习管理系统, 跟踪学习者与学习资源之间的交互 (顾小清, 等, 2014), 通过网络会话实现对学习经历数据的捕捉。传统的这些数据来源依然重要, 但在泛在学习的概念和空间中, 分析学习者学习进展, 提供有针对性的支持, 不能借助于简单的汇总统计和比较分析, 需要借助移动通信、云计算、情境感知设备等, 对伴随性的数据进行无感知、实时和持续的采集, 比如学习者学习轨迹、停留时间、游离时刻、主动倾向等。除了学习过程本身, 还包括学校的管理信息、状态信息 (设备、天气环境等) 和学校活动等。在某种程度上, 微观的有意义的教育大数据需要涵盖学校生活的方方面面, 实现实时的采集, 从而挖掘与个性化学习、发展性评价、学习路径设计、教育行为和政策预测等相关的数据, 服务于学校教学、科研、管理和服务。

(二) 教育大数据的应用掣肘

教育大数据展现了丰富的应用场景和无法想象的发展空间, 被许多学者认定为“金矿”。但从目前来讲, 几个核心的问题始终无法解决。第一, 数据的采集通道并不健全。在大数据视域下, 数据需要覆盖教育机构全过程的所有数据。从实践来看, 可以通过学校 Wi-Fi 和移动运营商捕捉到学习者的移动数据, 进而判断学习者的活动地点。但在各种情境中学习者的究竟在做什么? 投入程度如何? 情绪体验如何? 对于当前学校生态而言, 并不具备完整的采集能力。就研究者所在学校而言, 部分教室已经实现了学习者表情的捕捉, 但脱离教室之外的学习者情形很难捕



捉。第二,个人身份识别存在技术难点。对于学习过程而言,参与在线学习的识别始终存在问题。由于越来越多的学习行为通过翻转课堂、混合式学习转移到互联网,互联网的身份隐匿特征加大了个人身份识别的难度。同时,课堂教学和日常生活的个体识别也有技术门槛过高、推广需要一定的硬件和软件准备等难题,需要从顶层设计上做足文章,保持框架的包容性,然后逐步进行拓展。第三,各个系统中“孤岛”还大量存在。在信息化1.0时代,建设了大量的应用,产生了和正在产生着大量有价值的信息,比如学生管理系统、教育管理系统、教材管理系统、在线学习系统等,要整合各类数据,对于学校来说是一件浩大的工程。第四,数据标准缺乏,导致互操作问题。大数据的价值是多维度、多阶段数据的积累,由于独立开发的各应用并没有相应的标准,产生的数据难以被其他系统识别、整合和交换,无法进行互操作,限制了数据的有效来源。第五,数据采集的实时性、连贯性和伴随性,考验了学校的技术实力和经济实力。大数据的“大”不仅体现在“大”的价值上,还体现在“大”的量上。就物理课堂数据采集和智慧校园摄像头系统采集为例,按照可识别的技术操作,保存相应的视频材料,动辄产生TB计的数据量。实现整个学校完整的课堂数据采集,从数据量来看几乎是天文数字,这对教育机构而言,考验其云计算技术和经济实力。因此,在已有的教育大数据研究与实践中,真正使用大数据的典型案例不够多,更多停留在理论层面。第六,各类教育应用并不充分。对于教育而言,数据挖掘技术、学习分析技术、数据可视化技术等是教育大数据应用的主要领域,但整体而言,围绕这些方面的应用开发还停留在初期阶段,与成熟的应用体系、与教育大数据来源形成良好的互动关系,还相去甚远。

二、教育大数据生态圈建设: 解决数据来源的思路

从已有的研究视野来看,教育大数据的来源问题隐匿在数据挖掘和学习分析的浪潮里,似乎并不是问题。深究教育大数据的发展,尤其是“循数治理”(南旭光,2016)、“循数诊改”(胡娜,2017)、学习分析、智慧教育等研究的不断加深,带来的困惑和

难题,越来越多地来源于教育大数据本身的问题。当然,教育大数据并非是孤立的因素,而是在复杂的应用系统中多种因素的相互作用机制,有数据获取的问题,有应用的问题,也有意识思维的问题。只有围绕教育大数据产生、挖掘和利用形成一种内在统一性、外在平衡性和稳定性的体系构建,才能促进局域教育大数据的深度利用,实现“与教育主流业务的深度融合以及持续推动教育系统的智慧化变革”(杨现民,等,2016)。以此为驱动,课题组提出了教育大数据生态圈的概念,也就是围绕教育大数据,通过有计划、有目的、有系统的生态构建,形成以校内数据为主,沟通校外数据,努力协调数据产生、数据挖掘和数据利用之间的关系,构建一个具有内在一致性和持续发展能力的学校教育大数据生态环境,基础在于数据的获取,价值在于数据的挖掘和利用,其特征如下。

(一)教育大数据生态圈的系统性构架

与现有教育大数据的来源零散不同,教育大数据生态圈的建设是有目标、有计划、有体系的整体构建,体现了系统的特征:一是数据构成的多样性,就学校教育大数据生态而言,其构成有学习过程之中,也有学习过程之外。学习过程之中的数据指的是在线学习和实体学习数据。在线学习过程之中的数据主要渠道有门户学习数据、搜索学习数据和社交学习数据。门户学习数据指的是在学校学习门户以及与学校相关的学习平台的过程数据;搜索学习数据指的是学习者为了解决客观问题,通过搜索获取知识所产生的数据;社交学习数据指的是同伴和师长关系以及校外相关人员通过松散的或者正式的知识共享体系,共同解决问题所产生的数据。实体学习数据指的是在教室、图书馆等实体学习空间,通过情景感知、移动终端数据挖掘形成的数据。学习过程之外的数据指的是与学习者相关、影响学习过程的一切数据,包括管理系统数据、校园消费数据、图书借阅数据、校园活动数据等。二是数据之间存在多维的相互联系,形成相互影响和相互制约的关系。比如图书借阅数据与学习者学习参与存在一定关系,从校园消费数据可以挖掘出学习者的同伴关系等。除此之外,学校大数据生态的各类数据,与校园社会环境有一定的相关关系。总之,整个教育大数据生态圈通过数据收集的总体布局,形成系统性的生态设计。

（二）教育大数据生态圈的复杂性影响

教育大数据生态圈的复杂性主要体现在以下几个方面：第一，影响因素的多样性和广泛性。由于教育大数据生态圈关注的是人的发展，记录的是人的活动，人的需求的多样性和学校生活供给的多样性都会影响大数据生态的结构和内容。第二，影响时段的全程性和持续性。由于人参与学校生活受到环境（尤其是文化氛围）的影响是全程的和持续性的，这些都会促使学习者改变自身情况。比如计划参与在线学习的学习者由于网页新闻的影响，投入到了游戏之中；去图书馆的途中受到英语社区的影响，参与了英语角的活动，等等。第三，影响方式的多样性和影响效果的多重性。教育大数据的影响方式有很多，比如教师对学习者的学习成就的判断，学习者对自身感知的理解进而调整自适应学习路径。从影响效果来讲，有的是深层次的影响，有的是表浅的影响；有的是持久的影响，有的是暂时的影响。这就要求我们在设计感知与采集系统时，充分考虑各个因素之间的相互作用，形成内外数据与线上应用的作用关系，搭建持续的交付体系。

（三）教育大数据生态圈的建构性形成

与其他生态不同，教育大数据生态圈的构建更具有建构性，反映了人对环境的实践改造能动性的发挥。教育大数据生态圈的建构性体现在以下三点：第一，教育实践者的主动布局对生态圈的形成功具有关键作用。随着信息化的进展，教育大数据终究会实现数据支持教育教学的行为，会改造现有教学体系。但教育实践者的主动选择、顶层设计和平台打造，会缩短实效产生的时间。第二，由于使用大数据的意识和技术受限于社会发展，尤其是教师的意识对教育大数据的需求在某段时间内释放不够，会导致教育大数据的作用受限。第三，使用教育大数据还需要学习者的主动赋权。一方面，教育大数据的推动者需要清晰自身的伦理义务；另一方面，学习者需要将自身无关隐私的信息通过授权的方式共享出来，服务于教育大数据的完善和发展。同时，教育大数据的实践者还需要设计好相应的隐私保护机制，防止数据被滥用，让数据提供者有所信任。这些要素都凸显了主动性的作用和主动构建的价值。

（四）教育大数据生态圈的引领性

教育大数据的价值不断得到恰当的认识和评估，逐渐定位为一种变革教育的战略资本、科学研究支持

和智慧教育发展的机制。在资本层面，教育大数据生态将海量数据聚集在一起，教育决策者不再依赖经验和简单的统计结果做决定，而是依托数据实现科学决策，提升了决策的科学性，优化了资源配置效率，在消解教育公平等社会问题上有了可以依托的工具；在教学过程支持上，教育大数据有助于实现智慧教育所追求的个性化教育、自适应学习和发展性评价；在教育管理层面，教育大数据通过各类元素的相关关系挖掘，找准制约学习和生活的制度问题，优化教育管理模式，在安全管理、精准扶贫、治理体系构建等方面有极大的发展价值；在学习层面，通过教育大数据实现个性化路径设计，精准学习支持服务，支持教育服务的人性化；在教学科研方面，在传统的教育学研究范式中，教育作用机制难以让人信服，比如学习者的内在领悟、高阶智慧形成、内省感知等。依托生态圈所产生的数据——尤其是可穿戴设备、情境感知设备与脑科学等结合，各类原本隐匿的、无法测量的关系通过大数据有效展示出来，原本教育教学过程的“黑匣子”变成了“可见系统”（徐子沛, 2012, p.122），如同医学有了“核磁共振成像”技术（吴南中, 等, 2016），学习者的学习过程变成一种可以测量的东西。此外，教育大数据的发展可以引领教育产业发展，加深校企合作，通过校企在信息技术应用于教育教学等方面的合作，加强与企业的联系。

（五）教育大数据生态圈的成长性内在特征

教育大数据生态圈实际上形成与现存大数据产生技术、捕捉技术和利用技术相吻合的结构和形态，在不断成长中实现平衡。首先，教育大数据生态圈的构成要素在发展变化。随着数据产生布局的不断扩大、感知和采集技术的不断提高、教育大数据相关应用的不断拓展，与学习者相关的数据不断纳入生态圈的范畴，在大数据生态圈内流动，同时也提升了生态圈的层次。其次，教育大数据生态圈元素的相互“斗争”与“协同”，促使教育大数据生态圈的转型升级。生态环境动态发展的两种状态是平衡和失衡。在平衡状态中，相关要素体现了和谐共生状态；在失衡状态中，相关要素不断“斗争”，逐渐发生改变。在教育大数据生态圈中，各类大数据应用的使用，一方面提升了教育的水平，另一方面挖掘出新的问题，需要在不断的技术进步和数据使



用中寻求新的发展,通过技术的使用、挖掘的深度等打破原有的平衡状态。随着问题的逐渐解决,系统又进入到更高层次的平衡,实现了可持续的成长。这种成长同自然生态圈一样,没有顶点,只有过程。这就要求在教育大数据生态圈的构建中加强顶层设计,按照标准化数据、模块化思路、分阶段分步骤实施建设,通过标准一致、数据互访等方式,支持教育大数据生态圈不断升级。

三、“3+3”教育大数据生态圈： 一种基于实践的解决方案

从实践层面谈教育大数据生态圈的构建,基础是数据问题,构建则不限于数据,而是融入多种相关要素(如大数据应用),参照生态系统各个要素的相互关系,形成相互影响的体系(如大数据思维)。教育管理者、教育工作者和学习者对教育大数据的关注,影响了教育大数据生态的整个取向。依托学校转型升级的整体背景,本研究以“建设3个平台,构建3个体系”为抓手,以“数据支撑决策、数据优化教学、数据提升服务”为目标,积极探索学校教育大数据生态圈,提升学校治理质量,促进信息技术与教育教学深度融合,创造便捷舒适的学习体验。

(一)“3+3”教育大数据生态圈的建构思路

1. 以人为本,问题驱动和顶层设计相结合的整体思路。伴随着教育大数据应用的深入,势必引发传统教与学方式的变革;同时,学校管理模式、办学模式、人才培养模式等不同结构、空间和方式的解构与重构,不可避免地影响到不同主体的增量利益和存量利益,甚至形成各种利益诉求。举例而言,基于教育大数据的专业评价引导机制建成以来,不同层面的利益相关者对自身专业数据、专业地位、专业前景等的相关判断提出质疑,这就要求整个教育大数据生态圈的构建,在问题驱动和顶层设计中规划不同角色的职责和利益。教育大数据解决的是深层次的问题,挖掘的是教与学以及与教与学相关的问题,强调的是各个利益主体的整体价值。可以说,教育大数据生态圈的运行是围绕“人”在教育、科研、管理和生活过程中的实际问题。但解决教育生态圈运行的现实问题不是一蹴而就的,顶层设计需要考虑技术发展的持续接入

性、资金投入的连续性、人员保障机制和后续措施的立体化构建,逐渐形成“有序”的心理场域、“有效”的政策支持体系、理论与实践相互适配的体系,逐步解决教育大数据生态圈构建和持续发展的动力问题。在课题组所在的学校中,顶层设计主要体现了“互联网+”特色学校的目标定位、独立的大数据应用研究院的构建,以及在此基础上建立的系统化的制度,体现了发展的整体性和设计的全面性,既有解决当前问题的问题驱动意蕴,又有优化学校整体办学效益的顶层设计,如专业建设和课程建设的大数据评价等。

2. 结构优化,梳理不同相关数据来源的关系。普遍认为现有教育体系不是缺乏数据(与教育大数据的量有区别),而是数据混乱,也就是常说的“数据孤岛”现象(李香勇,等,2017)。教育大数据生态圈的构建,结构优化的目标是要解决不同相关数据来源的关系梳理问题。具体而言,在应用端,需要按照特定的数据标准,构建具有互操作能力的应用;在处理层,需要将劣构、复杂的数据转变为可执行的信息(Global,2012)。一般而言,来自各种终端的数据分为结构化数据、半结构化数据和非结构化数据。传统的结构化数据需要通过人工设计,教育大数据生态圈的建设将更多富有价值的半结构化数据和非结构化数据纳入到自身的视野,需要依托一定的数据收集模型,同时教育大数据的应用还需要考虑不同数据来源之间的数据共享,在数据安全和数据伦理的基础上,实现各个应用专注于自身的责任,在规范的约束下不断深入拓展,当用户需求改变时,不需要改变底层结构,只需要改变数据来源的接口即可。

3. 良性互动,建立完善的数据应用运作机制。在教学领域,数据挖掘和学习分析方兴未艾;在科研领域,大数据支持的相关应用逐步建立;在教育管理领域,依托教育大数据开展决策成为一种崭新的思考问题和决策行为方式。人人生产数据、人人共享数据、人人热爱数据、人人管理数据的大数据全民参与方式,正在改变教育领域数据使用的基础、内容和方式,促进了数据生产、加工和使用的互动机制的形成。首先,数据采集平台的开发,实现已有数据源的整合与利用,为教师、学生和管理者养成数据习惯、分享数据行为,建立了良好的意识;其次,

教师、学生和管理者意识的提升促进了更多数据源的需求,促进了更多应用的开发和使用,丰富了教学模式,提高了教学效果。如对基于在线学习开展混合式教学、智慧教室的建设等,实现了从排斥到喜欢、喜欢到依赖的过程,丰富了数据来源;最后,丰富的数据来源激发了管理和科研的积极性,围绕大数据的采集和应用的学习分析和教育决策方式大量开展起来,冲击了原有范式,同时也促使相关数据感知与采集平台的完善和拓展。

4. 技术能力提升,通过培训逐渐实现全员参与和使用。教育大数据为教育发展带来了新的机遇,然而教育大数据应用的基础首先是来源广泛的教育数据的获取;其次是大量应用的投入,与数据开展对话和互动;再次是社会对大数据产业的投入,包括建设大量的基础设施、数据标准、应用平台、推广应用等。可以预见的是,教育大数据生态圈在政策的重点支持下成长,已经不能单纯依靠信息技术部门来支持大数据的发展,全员参与、全过程参与、全方位参与是教育大数据生态圈不断升级和完善的基本条件。培训教师如何使用各类教育大数据应用,理解教育大数据思想,依托教育大数据设计自适应学习路径,开发不同层次和方式的资源,支持教育教学过程,这些都必不可少。在教师群体中,培养教师的数据模型建构能力,提升技术水平,是提升全员参与层次的有效方式。对于教育大数据的进一步应用而言,数据模型是对教学与管理进行有效监测、评价、诊断和预测的核心支持(孙洪涛,2016)。教育大数据的模型构建,不仅需要正确认识教育大数据采集与挖掘系统的技术水平,还需要具备教育教学知识。而现在的教育大数据工作者以技术人员和科研人员为主,一线教师参与较少,使得其模型建构和应用开发不能满足教与学对教育大数据深层次的应用。通过全员的理念、技术、应用使用等培训,配合激励、制约、引导机制的出台,促使全员使用教育大数据应用并开展科研和教学,尤其是参与到相应的模型构建中来,是实现教育大数据发展的重要抓手。

(二)“3+3”教育大数据生态圈模型及其内容阐述

为了更清晰地阐述“3+3”教育大数据生态理念,课题组设计了如图1所示的模型。

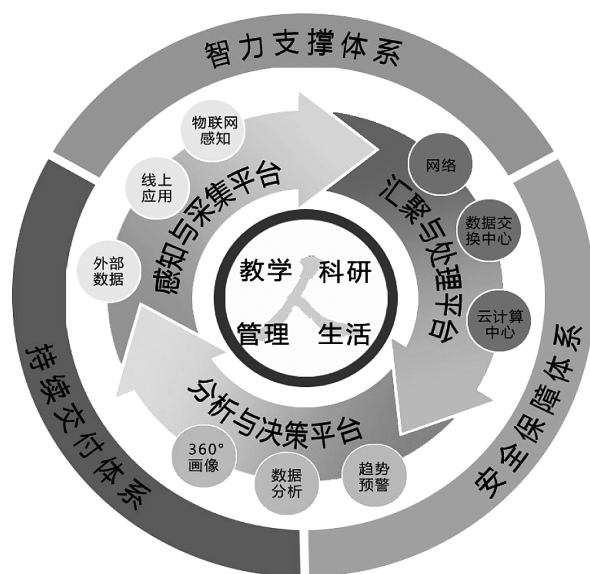


图1 “3+3”教育大数据生态圈

整个“3+3”教育大数据生态圈的核心是人,分内圈和外圈,其目标是人的教学、科研、管理和生活,是建立在系统性、复杂性、建构性、价值引领性和成长性的教育大数据整体构建思路上的。内圈的主要内容是数据感知与采集、数据汇聚与处理和数据分析与决策3大平台。数据感知与采集平台是通过多种情境感知装备的投入使用、物联网的进一步应用、线上应用的逐步开发和整合、外部数据的收集等,实施对校园设施设备、师生活动、环境互动的全面感知与采集。数据汇聚与处理平台主要是通过网络、数据交换中心和云计算中心,实现校园有线、无线、移动互联网和数据交换系统,实时汇聚和处理科学、科研、管理和生活服务全方位的数据,形成中心数据仓库,并利用云计算、大数据、人工智能等技术为数据的清洗、存储、挖掘和分析提供高效的基础平台。分析与决策平台是通过360°画像、数据分析、趋势预警等实现数据的价值。基本运作过程是基于中心数据仓库,运用聚类分析、多维分析、报表展示、KPI分析、驾驶舱等技术,进行多层次、多维度的数据分析,构建实体校园的360°数据画像,实时展现学校的运行状态,为教学、科研、管理和生活服务提供科学、全面的趋势预警和决策服务。

外圈主要是智力支撑、持续交付和安全保障3大体系。智力支撑体系是依托大数据智库、专家库建

设,通过提供决策类支持,实现办学事务、资源调度、治理结构优化的“数据化”。比如通过专业数据库的建设,形成专业评价引导的整合数据,实现专业评价和调整(吴南中,2016)。持续交付体系指的是通过各类标准化的应用开发,以数据互访问、系统模块化、出入口一致化为目标,支持数据的聚合利用,开发各类教育教学应用,缩短教学、科研、管理和生活服务等信息化应用的交付周期,提高应用的迭代速度,提升数据的共享程度,满足生态圈各类主题的应用需求。比如学校为跟踪学习者,建设了“学习轨道”应用,跟踪和整合学习者在学校生活的轨迹,判断学习者学业投入的程度,对偏轨学生进行引导性的提醒。安全保障体系指的是在利用教育大数据全过程中,要有信息意识、伦理意识和安全意识。坚持信息系统和网络安全“同步规划、同步建设、同步运行”的原则,从策略方针、管理规范、技术和服务等多个层面入手,完善防护、检测、响应和恢复网络安全体系。完善防火墙、实名认证等网络防护设施,搭建信息安全的第一道防线;完善病毒检测、漏洞扫描、入侵检测、等级保护等检测设施和机制,发现安全薄弱环节;完善应急响应、事件处理、调查取证等制度,确保网络安全事件的及时处理;完善备份、恢复等制度和设施,确保数据和系统的安全运行;坚持伦理为先,确保教育大数据的利用要强化对数据伦理的尊重,强化数据个人伦理观,体现为强调自主性、使用的善意、公平公正和可持续发展(吴南中,等,2017)。

（三）“3+3”教育大数据生态圈运行成效

1. 初步建成了学校层面的数据感知与采集平台

构建了从新生招录和报到开始的学习者生活数据采集系统,涵盖了校园生活的方方面面,如课程学习计划、课程学习成绩、其他学习成果、奖励申报与批准情况、课程满意度、图书借阅、在线学习情况、食堂超市等后勤体系、智慧长跑等常规性应用的数据。在安全领域,对学校校园内流动的人员开展了数据采集。在学生授权的前提下,还可以采集学生学校位置数据、上网数据等,部分教室实现了学生学习全过程数据采集,并实现了基于flanders和s-t两个模型分析学生投入状态,以表盘方式进行显示。教师在上课完毕之后,可以对学生表现进行基于整体的查阅,分析课堂结构,以优化教学。学校与腾讯云和中国移动等公司合作,通过采购云服务和信息化建设方

面的合作，开发了数据交换中心和云计算中心，加强了数据存储能力和计算能力。

2. 建设了一大批大数据应用

开发了短信系统（用于会议提醒、考试提醒、上课提醒、节日注意事项、放假提醒等）、迎新系统（新生入籍的相关环节）、决策支持系统、教学日志、交通车订票系统、资产管理系统、财务报账系统、办事大厅、大数据平台、学工系统、人事系统、办公自动化系统、open-X在线学习系统等一大批应用，并设计了相应的数据采集标准，各类数据之间可以相互访问（如图2所示）。如人事系统可以读取科研系统的相关信息，查阅教师的科研进展，用于判断教师职称晋升材料真实状况，分析教师的发展趋势等；食堂可以查阅交通车订票系统，用于准备教师工作餐的数量等；办事大厅将师生常用应用进行整合，登录上网账号之后，可以进入相关应用来开展工作，减少了签字和“跑腿”的次数。

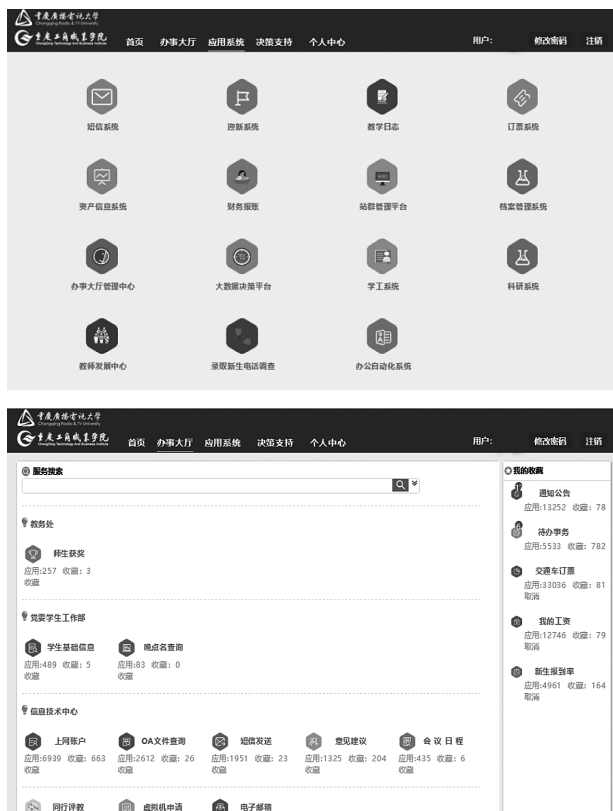


图2 大数据应用与办事大厅

(资料来源: <http://szxy.cqtbj.edu.cn/cqdddt/index.html>)

3. 打通了原有应用之间的数据屏障

在原有的框架下，学校建立了一些平台，如学分银行平台、教改平台、教材管理平台、质量保证控制

系统、教务管理系统等,保存了一些数据。学校遵循“不破而立”的战略方针,对具备改造条件的应用(如学分银行平台和OA管理系统)进行改造;不具备条件改造但能有效读取数据的进行数据读取,如教务系统、教材管理平台等;对于用处不大甚至没有用处的应用在对数据进行清洗、保存后废弃。经过新建、改造、换代等方式,建立了统一入口,规范了数据标准,实现了以单一身份登录,可以使用学校所有信息化体系应用,并对过程进行了数据采集。

4. 试点建设了智慧教室

在部分教室实现了课堂数据的全过程采集,通过身份识别(点名、学生座位分布等)、行为识别(教师活动轨迹、PPT采集、板书设计、学生行为采集)、情绪识别(学生和教师)等,自主分析上课视频数据。申请进入教室上课的教师,需要授权数据采集,可以帮助教师分析教学内容的教学效果,引发学生关注度提升的学习活动类型等相关信息,并根据学生的表现、教学设计与内容的匹配程度、教师的显著性教学行为等,客观形成教师课堂质量评价报告等。

5. 建成了分析与决策平台

根据已有应用、已开发应用以及学习平台相关数据的聚集,以及通过“爬虫”技术对相关网站信息的获取,形成了“分析与决策平台”,并逐步扩大应用范围和应用层次。现已运行的平台有实训室使用情况、科学研究计分情况、科研项目情况、人均借阅情况、专业发展情况等相关平台(如图3所示),并按照需求对不同层次的人开放查阅。管理人员和决策者可以根据数据进行决策,开展有针对性的管理工作。

四、结束语

文章从教育大数据应用的问题谈起,分析了教育大数据应用的掣肘——教育大数据本身的缺乏,通过



图3 大数据决策平台 (资料来源: <http://szxy.cqbtu.edu.cn/cqdddt/index.html>)

教育大数据生态圈的概念和特征分析,提出了学校层面的“3+3”教育大数据生态圈构建模型,并进行了3年多的实践,在3个平台和3大体系建设上都取得了一些进展,体现为数据收集和整合能力越来越强、师生使用的大数据应用越来越多、教与学过程的大数据支持越来越广泛和深入。我们正在开发和完善基于教育大数据校园贷预警、学习轨迹偏离预警、职业潜能与职业推荐、专业评价引导机制、质量保证体系、校园安全预警等应用。但客观来说,发展教育大数据、构建教育大数据局域生态圈还有较长的路要走。由于教育体系大数据应用人才本身不足,深入到学校层面的企业力量和社会力量不足,教育大数据本身的特性也在一定程度上束缚着教育大数据的发展。“3+3”教育大数据生态圈模型项目的实施,得到了许多同行专家的支持和赞许,但还未能达到预期的成效。我们深知,取得的成效只是初级阶段甚至是萌芽阶段。在国家整体产业转型升级的大背景下,如何进一步落实教育大数据的价值和功能,发挥教育大数据与教育教学结合的增值效应,仍然需要进一步思考和实践。

[参考文献]

- 常燕. 2013. 基于本体的教育资源异构数据集成的研究[D]. 呼和浩特: 内蒙古大学.
- 顾小清, 郑隆威, 简菁. 2014. 获取教育大数据: 基于 xAPI 规范对学习经历数据的获取与共享[J]. 现代远程教育研究(5): 13-23.
- 胡娜. 2017. “循数诊改”: 高职质量保证体系构建的理念与实践途径[J]. 重庆广播电视大学学报(6): 34-39.



- 李振,周东岱,刘娜,等. 2018. 教育大数据的平台构建与关键实现技术[J]. 现代教育技术(1):100-106.
- 李香勇,左明章,王志锋. 2017. 学习分析的研究现状与未来展望——2016年学习分析和知识国际会议述评[J]. 开放教育研究(2):46-55.
- 刘宏达,潘开艳. 2018. 融合创新:思想政治教育大数据研究的视角转换[J]. 学校党建与思想教育(2):52-56.
- 孟玲玲,顾君忠. 2012. 基于元数据管理的远程教育知识共享体系[J]. 现代远程教育(6):79-82.
- 南旭光. 2016. 大数据时代高等教育“循数治理”解析及实现路径[J]. 中国电化教育(8):20-26.
- 孙洪涛. 2016. 教育大数据的核心技术、应用现状与发展趋势[J]. 远程教育杂志(9):41-49.
- 吴南中. 2016. 高职专业评价引导的实践理性、运行逻辑与保障[J]. 教育与职业(16):15-19.
- 吴南中,夏海鹰. 2017. 教育大数据范式的基本理念与建构策略[J]. 电化教育研究(6):82-87.
- 徐鹏,王以宁,刘艳华,等. 2013. 大数据视角分析学习变革——美国《通过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(6):11-17.
- 徐珊. 2014. XML与WebServices数据整合模型及其在教育考试信息系统中的应用[D]. 南昌:南昌大学.
- 徐子沛. 2012. 大数据:正在到来的数据革命,以及它如何改变政府、商业与我们的生活[M]. 桂林:广西师范大学出版社.
- 杨现民,唐斯斯,李冀红. 2016. 发展教育大数据:内涵、价值和挑战[J]. 现代远程教育研究(1):50-61.
- 郑立海. 2015. 大数据时代的教育管理模式变革刍议[J]. 中国电化教育(10):5-13.
- Doug, L. (2012). 3-D Data Management: Controlling Data Volume, Velocity and Variety [EB/OL]. [2017-04-10]. <http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf>
- Eagle, N., & Pentland, A. (2006). Reality Mining: Sensing Complex Social Systems. *Personal and Ubiquitous Computing*, (10): 255-268.
- Global, P. (2012). Big Data for Development: Challenges & Opportunities[EB/OL]. [2018-5-11]. <http://www.unglobalpulse.org/sites/default/files/BigDataforDevelopment-UNGlobaPulseJune2012.pdf>
- Seel, N. M. (2012). *Encyclopedia of the Sciences of Learning* (pp.1176-1179). NY: Springer.

收稿日期:2018-05-29

定稿日期:2018-08-27

作者简介:吴南中,博士研究生,副教授;黄治虎,博士,教授;曾靛,硕士,讲师;谢青松,博士研究生,副教授。重庆工商职业学院(401520)。

夏海鹰(本文通讯作者),博士,教授,博士生导师,西南大学教育学部(400716)。

责任编辑 韩世梅

(上接第65页)

- Breneman, D. W., (2003). The University of Phoenix: Poster Child of For-Profit higher education [EB/OL]. Retrieved October 5,2015, from http://www.ncspe.org/publications_files/breneman.pdf
- Siemens, G., (2005). Connectivism: Learning as Network-Creation [EB/OL]. [2017-8-23]. <http://www.elearnspace.org/Articles/networks.htm>
- Pirani, J. A., (2003). Supporting E-Learning at the University of Phoenix [EB/OL]. [2013-10-05]. <https://net.educause.edu/ir/library/pdf/ers0303/cs/ecs0304.pdf>
- PayScale. (2018). Average Salary for University of Phoenix (UOP) [EB/OL]. [2018-03-07]. Employees [https://www.payscale.com/research/US/Employer=University_of_Phoenix_\(UOP\)/Salary](https://www.payscale.com/research/US/Employer=University_of_Phoenix_(UOP)/Salary)
- UOPX. (2007). Faculty-Handbook 2007 [EB/OL]. [2018-01-09]. <https://ecampus.phoenix.edu/content/FacultyAcknowledgement/Sources/2007FacultyHandbook.pdf>
- UOPX. (2013). 2013-2014 faculty payroll & benefits schedule [EB/OL]. [2017-10-05]. <http://webmkt.s3.amazonaws.com/faculty/Email/2013-2014%20Faculty%20Payroll%20and%20Benefits%20Schedule%20v1.pdf>
- UOPX. (2014). 2014-2015 Faculty-Handbook [EB/OL]. [2014-10-05]. <http://www.ebooksde.org/download/Vis/2014-2015-faculty-handbook-university-of-phoenix.pdf>
- UOPX. (2016). recruiting system [EB/OL]. [2016-04-27]. <https://faculty-apollo.icims.com/jobs/intro?hashed=-435711714>
- UOPX. (2018). Campus Locations [EB/OL]. [2018-03-07]. <http://www.phoenix.edu/campus-locations.htm>

收稿日期:2018-03-28

定稿日期:2018-07-31

作者简介:袁松鹤,副研究员,博士,国家开放大学党政办公室(100039)。

责任编辑 郝丹