

基于教育大数据生态体系的高校智慧校园建设研究*

余 鹏¹, 李 艳²

(1.中南民族大学 现代教育技术中心,湖北 武汉 430074; 2.中国科学院 武汉文献情报中心,湖北 武汉 430071)

摘要:近些年,教育大数据的研究与应用为高校智慧校园的建设与落地开辟了新的思路。针对当前高校智慧校园发展中所面临的学科科研、管理服务及综合决策等方面的痛点问题,该文从制约高校教育大数据发展的六项主因入手,深入探究了智慧校园背景下教育大数据(生态)的特点及研究意义。从不同规模、类别、等次的高校样本出发,归纳了教育大数据生态体系的数据应用与服务模式,并探索了一种基于智慧校园良性发展的教育大数据生态体系架构;结合设计的教育大数据应用氛围视图,提出了一种教育大数据生态体系模型用以提升数据治理的精准化,管理服务的协同化,学科科研的智能化,决策应用的科学化;最后,基于智慧校园中长期的发展需求,进一步探究了教育大数据生态体系的建设目标。

关键词:教育大数据;智慧校园;教育大数据生态体系;大数据应用氛围视图;大数据生态体系模型

中图分类号: G434 **文献标识码:** A

一、引言

十三五以来,我国正处于由教育大国向教育强国迈进的攻坚期和转型期,不断地提升高校管理水平,促进高校实现由数字化校园向智慧校园的快速转型发展,是当前诸多高校建设的目标之一,也是高校办学水平、现代化进程的重要考评指标之一。智慧校园代表的是高校信息化发展的高级形态,它融合了虚拟化、云计算、大数据、人工智能、数据挖掘、知识管理与知识发现等新兴技术^[1],智慧校园代表的是以无感知服务为中心的发展理念,是数字化校园阶段以管理信息化向服务信息化的升华。智慧校园以更为科学化、精细化、维度化的建设模式,全面感知校园的物理环境,以提供知识化的服务需求、科学化的数据管理、智能化的决策分析为导向,实现校内外各类人、物、业务、流程的深度智能融合。

其中,教育大数据作为智慧校园建设的载体之一,为驱动智慧校园的发展,起到了至关重要的作用。大数据能够影响高等教育实践活动,以多元

维度数据为基础,感知和分析教育活动中人员的所有行为,以提供有效的数据流转优化业务流程为依据,为实现科学化的战略决策,使复杂化、非结构化数据转化为可以操作的信息提供保障^[2]。

二、高校智慧校园发展中的痛点现状分析

信息化的不断发展,催生了教育大数据的生根落地。我国高等教育正从基础建设发展阶段逐步向改革创新冲刺阶段迈进。“十三五”期间,高校的信息化进程已全方位铺开,各地高校在积极响应国务院颁布的《统筹推进世界一流大学和一流学科建设总体方案》的同时,期望抓住发展契机,建设一流师资队伍、培养拔尖创新人才、提升科学研究水平、传承创新优秀文化、推进成果转化^[3]。

笔者认为,高校综合实力的提升离不开业务水平及流程服务的升华,高校在不断建设业务系统、完善数据交换平台、优化管理制度、改进教学模式的同时,面对改革进程中集教学、管理服务、决策等方面一系列的痛点问题,结合教育大

* 本文系中南民族大学教研项目“基于‘大数据挖掘与决策分析+微服务’体系的高等院校云辅助教学研究(项目编号:GJYX16057)”、中央高校基本科研业务费人文社科青年项目专项资金项目“‘互联网+’背景下教育大数据挖掘与决策分析体系辅助智慧校园建设模式研究(项目编号:CSQ17011)”、湖北省教育厅2017年度人文社会科学研究指导性项目“基于大数据挖掘与决策分析体系的教育云辅助高校信息化发展模式研究(项目编号:17G040)”阶段性研究成果。

① 李艳为本文通讯作者。

数据的特点^[4],利用教育大数据作为介质,通过数据交换、数据治理、分析挖掘、决策分析辅助高校规范数据标准、肃清业务流程、健全管理制度、深挖发展痛点,从而正确定位自身中长期发展的绩效目标,逐步实现高校自身阶梯式上升是必要可行的。

以下从教学、管理服务、综合决策三个方面剖析高校在智慧校园发展中的痛点现状。

(一)教学方面

教与学是高校生存的生命线,老师和学生作为高校里重要的两类载体,几乎串联起高校所有的活动,教学过程中发展累积的结果化数据是教育大数据的重要来源之一。诸如课堂学习、测评互动、网上教学等一系列活动过程,构筑起教学活动的数据集。教学活动中的过程化数据贯穿了从数据产生、输出、信息展示,到知识决策整个过程。课程与实际应用脱节,理论与动手实践分离,教师所期望的和学生所需要的相差甚远,都是高校教学发展中的痛点。传统的以“教师为中心”的教学方式,已无法满足学生所想和学生所需。作为载体之一的高校教育管理者,尽管不断地改进教育管理流程、优化学科建设、出台激励政策,却无法精准的定位不同的学科建设目标、学科发展趋势以及学生培养模式。面对不同的学科专业、不同的学生群体、不同的知识背景,如何选择合适的点去改进,如何感知学生所想、探寻学生所需,是教育工作者永远摸不透的谜题。

(二)管理服务方面

当前,我国众多高校仍无法脱离“行政化”的管理方式。尽管各级部门不断优化业务流程、梳理服务目录、改进工作方式,但鉴于业务流形成闭环,数据孤岛林立,高校内部跨部门业务数据缺乏有效的共享与交互,仍无法解决行政事务管理流程繁琐、跨部门业务办理困难且效率低下的问题。高校师生作为服务的受体,本应享受学校改革和信息化建设成果带来的便捷,却被无形套上繁杂的流程枷锁。业务系统作为管理服务的支撑,业务流程扮演着系统主心骨的角色。随着高校的不断改革,业务部门管理者的思路在不断发生变化,主心骨的业务逻辑关系同样发生着改变,庞大而复杂的系统逐步无法完全满足业务部门的新需求,过去花大把力气、高额经费建设的业务系统面临着“大修”和“淘汰”。另外,很多功能相近、流程相仿的应用,管理部门重复建设,造成高校资源的过度浪费。管理流程的简化、服务体验的提升,尤其是用户个性化的需求,高校实现管理体系的“去行政化”迫在眉睫。大学改革的核心要素应是顺应高校

从“管理”走向“治理”的趋势,开展“管理服务治理”和“业务数据治理”,才能真正提升服务对象的用户体验。

(三)综合决策方面

综合决策涉及的对象是高校的领导者和管理者,“火车开得快,全靠车头带”准确的反映了决策的重要性。“拍脑袋式的决策”、“屁股决定脑袋”的表述充分反映了当前领导者和管理者在管理能力、管理水平上的不足。

综合决策需要深入掌握基础数据、准备把握发展不足、敏感抓牢政策动态,但鉴于高校基础数据不全、数据孤岛林立、数据质量低下、数据服务体系缺失,使得高校开展科学化、智能化的综合决策举步维艰,另加上领导者和管理者自身管理方法、管理水平和管理视野的差异,都严重的影响了综合决策工作的开展。诸如:招生计划决策、财务状况与政策决策、师资队伍发展状况与人才政策决策、科研发展状况与学科建设决策、教学评估状况与人才培养模式决策等一系列关乎高校生死存亡的决策项,缺少了校级层面的数据分析与发展趋势预测为依据,综合决策仅是“纸上谈兵”的空想。以数据为基础,实现由事实数据向评价指标地智能转换,预测趋势的发生轨迹,以拓新管理、服务质量,让数据说话才能为科学的决策提供依据^[5]。

三、智慧校园建设中面临困境的成因分析

当前,我国高校由数字化校园向智慧校园迈进的进程中,尽管积累了大量教学、科研、运行管理等方面的业务数据,构建了数据中心与业务服务平台,却依然面临着许多的困境。经归纳分析,制约高校教育大数据发展的因素主要表现在以下六个方面,以教育大数据生态体系链的形式呈现,如下页图1所示。

(一)高校信息化办公室(信息化领导小组)缺少全局性、统筹性的指导方案及信息化专业人员梯队

智慧校园的发展,离不开全局性、统筹性的顶层设计与规划决策。众多高校信息化的建设模式依然是按照“被推着向前”“看一看,模仿别人建”的发展模式,加上信息化专业人员梯队得不到有效补充,成为制约教育大数据发展的重要瓶颈。高校因为个体的差异,在管理模式、专业建设、经费保障、应用需求、自身定位等方面不尽相同,需要规划好适合自身发展的指导方案,构建强有力的专业技术团队,才能保障智慧校园的稳步推进。

(二)高校缺少对各类数据源的有效整合

数据的流转是业务生命周期的主心骨。数据源

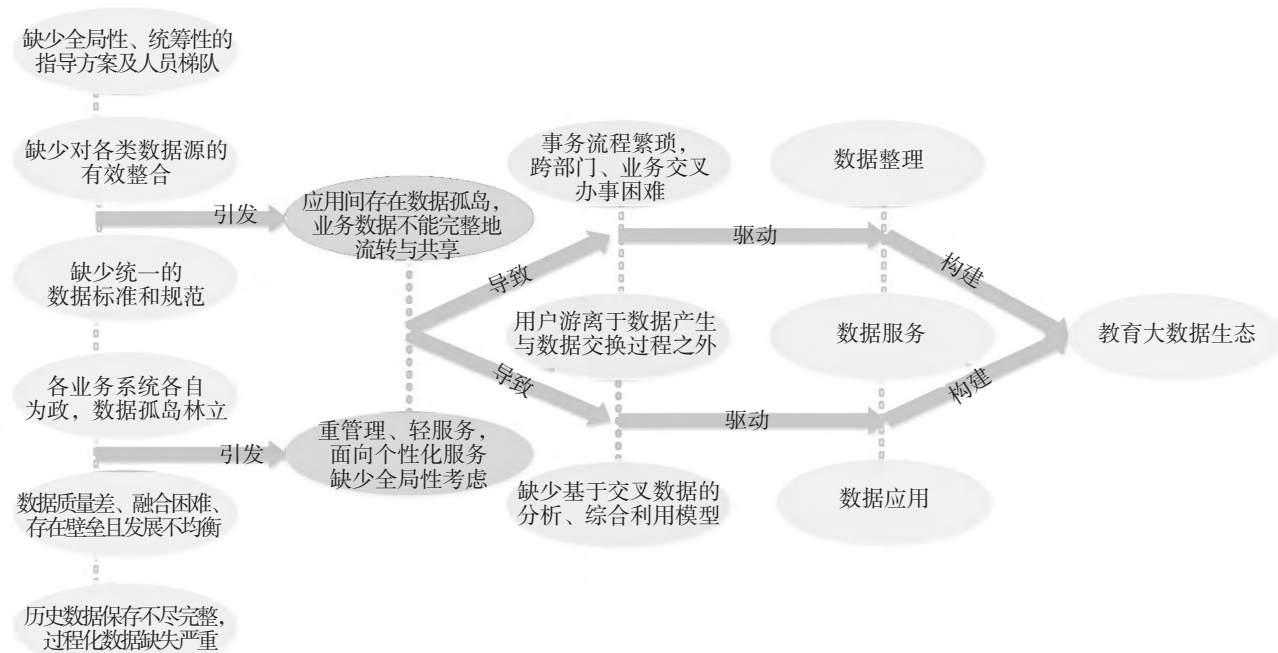


图1 教育大数据生态体系链

的产生，应由业务归口部门产生，依托数据交换平台向需要使用数据的应用进行推送，实现教育数据的实时交互。当前，由于高校业务及职能交叉管理的需要，使得数据来源的入口较多，缺乏有效、统一的整合，给业务带来的必将是数据的不一致，引起数据中心数据管理的混乱，给教育大数据的分析与挖掘埋下极大的隐患。

(三)信息化建设过程中缺少统一的数据标准与规范

数据标准与数据规范是顶层设计阶段的必要性步骤。在全球规划环节中，应充分考虑各类业务需求，制定并规范合理化的数据标准，从而为后期流程化的业务建设奠定坚实的基础。高校理应建立公共代码库，构建代码编码规则，统一取值、规范赋值；按照国标、部标、校标逐级筛选，依托数据共享平台，应用到各类业务系统；对公共代码库进行持续化的管理，关注其对全校范围、所有数据的影响^[6]。

(四)很多高校业务建设各自为政，数据孤岛林立，缺乏有效的数据管理与数据交互

智慧校园的大数据生态体系应以业务流转过程中产生的过程化数据和结果化数据为基础。业务的流转离不开业务系统间的数据共享与交换，构建标准的开放性业务接口(例如：JSON、WebService等接口)，利用中间件、前置机实现基于标准接口的多业务之间的数据转换与共享、同库共享、异库交换，有利于深度解决数据孤岛问题，业务系统的构建应保证其具备一定的开放性。

(五)数据质量差、数据融合困难，存在壁垒且发展不平衡

由于长期对数据的重视程度不高、对数据价值的认识不够，且受限于行政管理架构影响，跨域实施困难，加上数据分析模型及工具的匮乏，高校的数据利用与前期规划、设计、开发脱节，造成了积累数据与实际应用需求存在严重的偏差。数据质量差、数据融合困难，严重的制约了高校数据的整合与共享。构建智慧校园的大数据生态体系，第一步应开展数据治理，实现包含不同业务系统、不同数据库、不同表格、不同字段、不同数据项的数据整合；第二步进行数据监控，实现数据质量的持续性保障，构建数据监控规则，动态监测并报警提示数据异常^[7]。

(六)高校对不同业务产生的历史数据的保存不尽完整，过程化数据严重缺失

我国高校信息化发展了几十年，很多业务系统都进行了多轮的更新与重建。由于各个时期数据标准、数据规范不同，数据保护意识不够，造成了大量历史数据的流失。“推翻重来”“升级换代”无形加剧了数据的遗失。大数据生态体系的构建，更加强了除结果化数据外的过程化数据的重要性。结果化数据往往会以结构化数据的形式存入关系型数据库、或以非结构化数据的形式存入存储系统、文件系统、文本中进行保存。相反，过程化数据因其生命周期短、体量大、源头多、异构性强等“缺陷”，往往被用户忽略和丢弃(例如：各类机器在

应用过程中产生的日志数据、用户行为数据等)。过程化数据是大数据分析的重要数据源,其生成了多维度的数据集,依托大数据分析平台,可实现教育大数据的解析与关联,并利用类似于SparkSQL工具完成过程化数据的计算与检索,通过对解析和关联结果进行选择、过滤、统计等,筛选数据价值,并进行可视化的数据输出。

基于上述因素制约,在日积月累的过程中,信息化应用之间的数据孤岛也不断累加,业务数据无法较好地完成数据流转与共享,加上高校本身存在“重管理、轻服务,面向个性化服务缺少全局性考虑”的诸多“先天不足”,加剧了跨部门或跨业务交叉办事困难、用户游离于数据产生与数据交换过程之外的种种不利局面。数据治理、数据服务,甚至是更深层次的数据挖掘分析与决策应用工作在高校智慧校园建设环节中的必要性亟待进一步加强。构建教育大数据生态体系的重要目标,是实现信息化应用与教育数据治理的深入结合。综合运用信息化技术,将高校业务数据通过采集、处理、利用加以高效地组织和科学化管理,使其成果得以有效的应用到智慧校园建设的方方面面,从而提升高校的核心竞争力。

四、智慧校园背景下教育大数据(生态)的特点及研究意义

随着“互联网+”技术的不断革新,大数据应用服务在物联网及公共服务等领域遍地开花,产生出较好的社会效应和巨大的社会影响。近些年,基于大数据的研究快速延伸到象牙塔,对高校智慧校园的建设与发展提出了更深层次的要求。伴着高等教育改革的不断深入,高等院校由数字化校园向智慧校园演进已是社会发展的必然趋势。当前,全国众多高校正如火如荼开展“移动互联网”“大数据+教学/科研/管理/服务”“云计算”“软件定义网络(SDN)+网络功能虚拟化(NFV)/网络定义安全(SDS)”“虚拟化应用服务”等新兴技术的研究与应用,旨在为师生、行政管理人员提供合理化地管理、精细化地服务、智能化地应用及科学化的决策。

在持续化的信息化建设与应用中,高校已积累了丰富的教学、科研、管理、服务等方面的“小数据”^{[8][9]},小数据集为教育大数据的深层次研究提供了夯实的多维度数据基础。同时,传统流程化的业务模式沉淀出大量的结果化数据,加上过程化的机器日志数据,逐步衍化为价值巨大的优质数据资产。智慧校园的核心任务是实现教育大数据的反哺,使得高校由“以事物为核心驱动应用(封闭式

的命题)”逐步向“以数据为核心驱动应用(开放式的命题)”转变,让教育大数据成为创新精准服务的基石。

(一)高校教育大数据特征分析^[10]

1.数据源的多样化:数据源自学校各级机构,重点围绕师生、横跨教学科研、管理服务,所有的业务流程能够产生丰富的过程化数据,最终转化为结果化数据。过程化数据逐渐被用户重视,成为感知用户个性化的重要数据源。

2.数据异构维度高:教育数据体量一般,但异构性强、跨度大、维度高、含义丰富。维度数据源多,数据能够更加准确的定义个性化服务,极大的提升用户的无感知体验。

3.数据类型多样化:除结构化数据外,半结构化、非结构化数据规模庞大。教育大数据由结构化数据为主导逐渐向半结构化数据、非结构化数据为主导进行转变。

4.数据整体价值高:单一的“小数据”意义甚小,但整体数据可描绘全局。小数据作为大数据的支撑,能够产生个体性的过程化小数据,从而构建全局性的维度化大数据。

(二)教育大数据的核心作用及研究意义

中国高校信息化发展历程普遍经历了“底层设计”“基础建设”“应用推广”及“服务提升”的四元化阶段。教育信息化的实施亦是不断促进信息化全方位辅助高校教学水平、科研模式、管理手段、服务精度及决策能力的全方位提升。当前,教育大数据的应用水平和应用程度是标志高校向“创新发展型”“世界一流化”大学迈进的重要指标。开展对教育大数据的深度研究与思考,从本质上是为了实现全面的解放与提升,让智慧化的成果受惠于高校的主体对象,其研究意义主要体现在以下方面,如图2所示。

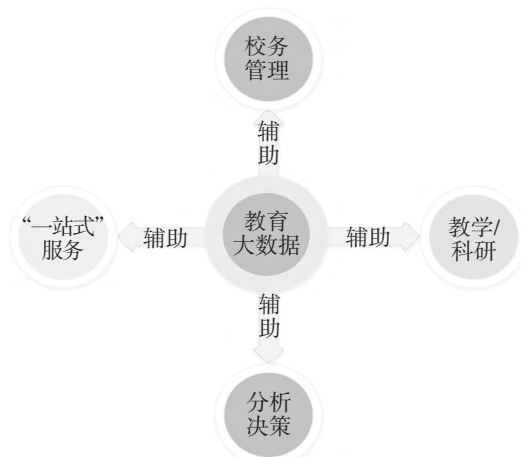


图2 教育大数据核心作用

续表1

服务: 查询展示	挖掘分析
发展: 事后分析	事前预测
目标: “全数据”	“大数据”

表2 样本高校教育大数据发展存在的共性问题

问题属性	特征体现	问题映射层面
缺乏整体观念	缺少健全、长远的全局指导方案	概念
数据来源众多	缺少真正有效的多源、多类数据整合	规范
技术标准不一	推广示范效应不理想	标准
数据孤岛林立	数据来源分散, 采集困难多	技术
数据融合欠佳	存在一定的数据壁垒、发展不均衡	应用
数据衔接脱节	不同教育领域数据脱节	政策

通过分析10所样本高校教育大数据发展现存的共性问题, 结合样本高校教育大数据的发展目标, 开展基于教育大数据生态体系的数据应用与服务模式的探讨与改进是智慧校园的发展的自然选择, 亦是互联网模式新信息化环境下应用需求催生的新趋势。经调研总结, 结合上述10所高校教育大数据(生态)发展现状, 围绕其智慧校园的发展模式、应用场景与发展思路, 将教育大数据(生态)体系应用与服务模式划分为水平化与垂直化两类, 如表3和表4所示^[20-22]。

表3 教育大数据生态体系的水平化应用与服务模式

支持“精准管理”的数据应用与服务模式	支持“科学决策”的数据应用与服务模式	支持“综合评估”的数据应用与服务模式
培养计划执行预警(学生) 毕业资格审查预警(学生) 项目经费执行预警(项目负责人) 学风学分预警(学校、学生) 就业发展状况预警(高校) 舆情监控(高校) 面向个体智能导学……	招生计划决策支持 财务状况与政策决策支持 师资队伍发展状况与人才政策决策支持 科研发展状况与学科建设决策支持 教学评估状况与人才培养模式决策支持 ……	学习能力与效果(学生) 教学质量与水平(教师) 综合素质评价(教师、学生) 发展能力评估(教师、学生) 学生行为分析与预测(学生画像) 教学绩效分析与评价(教师画像) ……

表4 教育大数据生态体系的垂直化应用与服务模式

支持“学校全局”的数据应用与服务模式	支持“学科发展”的数据应用与服务模式	支持“院系业务”的数据应用与服务模式	支持“大学事务”的数据应用与服务模式
高基表分析评估 全校资产分析评估 招生就业分析评估 高校舆情监控评估 师资水平分析评估 科研实力分析评估 学校发展决策分析……	本科生教学质量分析评估 研究生教学质量分析评估 留学生教学质量分析评估 发展规划部门的学科评估 教师学术表现分析评估 教学状况与人才培养分析评估 ……	学院教学工作分析评估 学院科研水平分析评估 奖学金、三好学生评定绩点分析评估 学科综合发展分析评估 专业学生学习能力分析评估 ……	一站式服务应用 “一表通”流程化服务 “大数据+微服务”应用 学生失联预警应用 贫困生资助应用 创新创业孵化应用 ……

表3中, 基于高校教育大数据生态体系的水平化应用与服务模式被划分为三类: 支持“精准管理”

1.解放高校“重管理、轻服务”的“以管理为主、用户感受为辅”管理模式, 实现管理体系中管理者的“去行政化、去权利化、去一元化”;

2.解放高校“重讲授、轻反馈、轻实践”的“课堂为主、课后为辅、实践能力弱”教学模式, 实现教学体系中角色角色的“个性化学习与自主学习”;

3.解放高校“屁股决定脑袋”的“拍脑袋式”决策模式, 实现决策体系中决策者的“数据化决策与智能化决策”;

4.解放高校“重流程、轻效率”的“跑断腿、折断腰、处处碰壁式”的服务模式, 实现服务体系中对服务对象的“精细化服务与一站式服务”。

当前, 智慧校园的建设仍在数据管理与应用等方面存在诸多问题, 构建大数据生态治理服务体系(本章节中简称“体系”)旨在利用虚拟化、云计算、人工智能、数据挖掘等技术, 将高校各类过程化、结果化的数据通过采集、清洗、分析、存储等流程, 加以高效的组织与科学化管理, 使其被最大化的挖掘与利用, 最大限度的释放数据红利^[11]。通过数据应用为导向开展数据治理可以有效降低数据治理的难度和风险, 以致于更好地提升高校在教学、科研、管理、服务、决策等方面的能力。

五、高校教育大数据生态体系的应用服务模式建设

(一)教育大数据生态体系的数据应用与服务模式

为了较为客观地了解我国高校教育大数据的建设及应用需求, 结合高校智慧校园建设发展需求和服务需要, 针对不同规模、不同类别、不同等次的高校大数据(生态)体系进行了调研, 选取了国内双一流A类高校清华大学^{[12][13]}、复旦大学^[14]、武汉大学; 双一流学科高校华中农业大学^[15]、南京理工大学^[16]、中国矿业大学^[17]、中国地质大学(武汉); 普通本科高校中南民族大学^[18]、湖南农业大学^[19]、常熟理工学院等三类高校为样本。

尽管上述10所高校在教学、科研、管理、服务水平方面存在较大的差异, 但从当前的教育大数据发展现状以及智慧校园中、长期的发展目标分析(如表1所示), 存在如下共性问题(如表2所示)。

表1 样本高校教育大数据的发展目标

目标属性	问题映射层面
特点: 抽样数据	全量数据
算法: 复杂算法	简单算法
应用: 零星点状	系统全面
方式: 授人以鱼	授人以渔

的数据应用与服务模式、支持“科学决策”的数据应用与服务模式、支持“综合评估”的数据应用与服务模式。水平化的划分方式，主要从校务职能的角度入手，围绕提升“教学、科研、管理、服务”四大类主体工作水平为目标。通过教育大数据(生态)体系在校务职能上的全方位应用为载体，创新发展模式、提升治理能力、强化管理意识、简化办事流程、开创一流服务，推进智慧校园建设的快速落地。

在上页表4中，基于高校教育大数据生态体系的垂直化应用与服务模式被划分为四类：支持“学校全局”的数据应用与服务模式、支持“学科发展”的数据应用与服务模式、支持“院系业务”的数据应用与服务模式、支持“大学事务”的数据应用与服务模式。垂直化的划分方式，主要从自上而下的逻辑层次关系入手，围绕高校的机构单元职能开展大数据应用与服务模式探索。通过教育大数据(生态)体系在层级式的机构单元职能方面的全面应用为切入点，不断推进智慧校园建设中释放数据红利为核心，辅助提升高校现代化的治理能力为契机，构建长久的教育大数据(生态)体系发展规划为任务，健全科学的决策管理为目标。

(二) 教育大数据生态体系架构

通过对上节中的高校教育大数据应用现状调查，智慧校园背景下的教育大数据生态体系的核心工作流程可归纳为：理清(数据)源头 → 规范(数据)标准 → 优化(数据)质量 → 促进(数据)共享 → 开展(数据)分析 → 挖掘(数据)价值 → 生成(数据)应用 → 反馈(数据)源头。教育大数据生态体系的构建，旨在加快教育大数据与高校各类应用之间的深度耦合，重构教育数据服务基础架构，建立层次化数据服务与管理体系，形成拥有良性循环机制的数据生态环境，进一步实现高校的信息化应用模式由“粗放式、大一统”向“精细化、个性化”的转变。

基于上述分析，教育大数据生态体系架构设计

为“一源(数据源)”+“一平台(大数据平台)”+“一端(用户端)”的三级架构模式(如图3所示)。

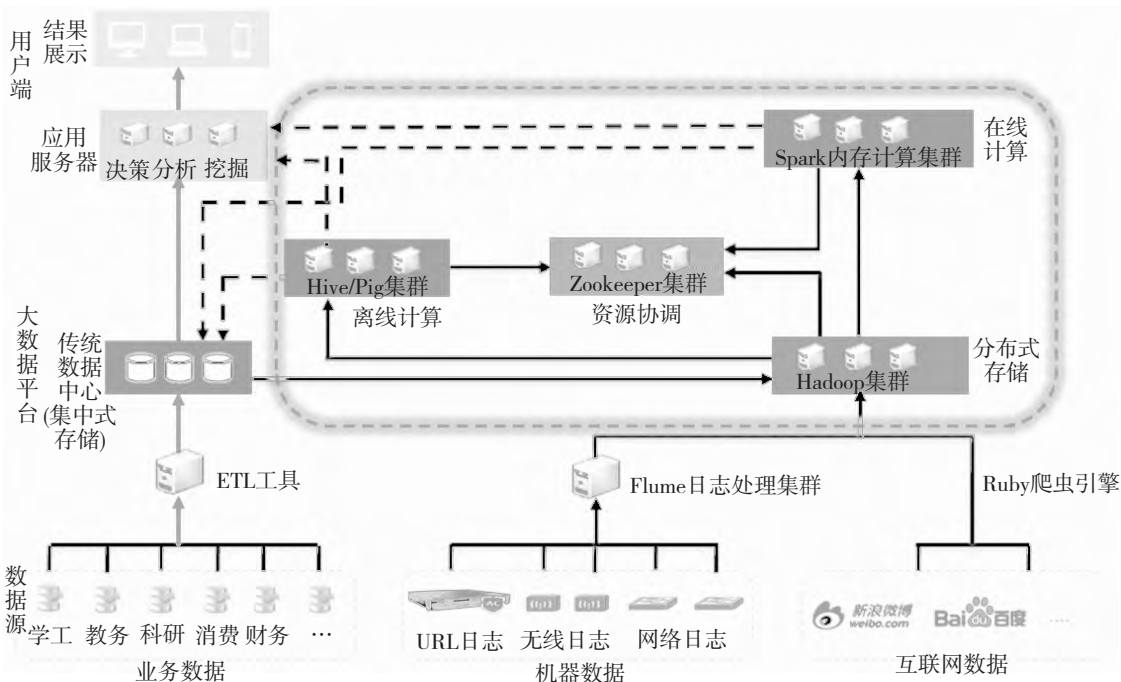


图3 教育大数据生态治理体系平台架构^[23]

底层的“数据源”是体系架构的基石。该层主要完成高校内部以业务数据为主的结构化数据，以文本、音视频、机器数据(日志数据)为主的半结构化数据与非结构化数据、外部互联网数据的采集工作，依靠ETL工具、Flume日志处理集群及Ruby爬虫引擎进行数据预处理，实现对各类数据有序且规范化地梳理、采集、清洗、标准化规范存储和应用，为中间层“大数据平台”所需的多维度的数据交换、计算与挖掘服务提供干净且可靠的数据源；

中间层“大数据平台”是教育大数据生态体系的核心层。该层基于目标应用开展对多源数据的样本分析、建模、多维度数据抽取及规则库的定义和递归优化，依托Hadoop数据仓库集群工具(Hive/Pig集群)、分布式应用程序协调服务工具(ZooKeeper集群)、大规模数据计算处理工具(Spark计算集群)、分布式存储工具(Hadoop集群)等完成对实时及离线结构化、非结构化数据集全面的存储、交换、计算、分析、挖掘，实现异构数据的归一化处理，为上层“用户端”的数据应用报表呈现、个性化的需求服务推送、智能化的分析评估、科学化的决策定位提供坚实的依据。

上层“用户端”是实现数据的可视化输出。根据用户的属性特点和个性化需求，为用户推送与其相关的知识服务信息。

(三) 教育大数据生态体系建设模型

信息化发展所面临的新挑战与新机遇，驱使高

校的体制改革向更高的要求推进,高校改革的核心要素主要体现在:

1.管理治理:在技术上实现管理体系的“去行政化”,以顺应高校从“管理”走向“治理”的良性发展趋势。

2.服务治理:从学校行政服务规划着手,梳理与整合业务流程,以完成流程的再造从而提供更为明晰化的服务。

3.数据治理:构建数据资产平台,提高针对全生命周期的数据治理能力,为后续教育大数据的应用创新提供基础。

针对改革的三项核心要素,教育大数据生态体系的主要工作表现为:

1.流程协同方面:进一步厘清高校各类管理流程及形式转变,采取高效便捷的工作流引擎,快速配置及部署实施。

2.服务模式方面:进一步构建业务与服务的“标准化”“耦合化”及“碎片化”,定义标准的服务接口,建立服务化的数据驱动模型,将服务端独立于后端,开发和推广采用统一、创新技术,以达到服务模式与需求应用的深度融合。

3.数据基础方面:进一步建立更为完备的数据管理和服务机制,完善各种数据采集、交换、分析和利用,以较低的经济成本和学习成本交付给业务应用主体(用户)。

教育大数据生态体系模型的构建过程,实际是充分利用虚拟化、云平台、云计算等信息化技术和统计、在线分析处理、情报检索、机器学习、专家系统和模式识别等计算机技术,将高校各类数据通过采集、处理、利用加以高效地组织和科学化管理,并在流程服务中得以有效应用于教学、科研、校务管理与决策的全过程。教育大数据生态体系不单是相关技术的集合,更是智慧校园发展过程中衍生出来的一种应用思路,是一种使用全量数据的意愿和能力。如图4所示,智慧校园中的教育大数据生态体系的核心环节可被理解为是一条全生命周期的“应用闭环流”,本文将之定义为“教育大数据应用氛围视图”。多维度的数据源经过原始数据采集→数据整合→数据分析→数据加工→数据展示→输入反

馈→回溯至数据源的迭代过程,逐步构建起良性的教育大数据生态环境。

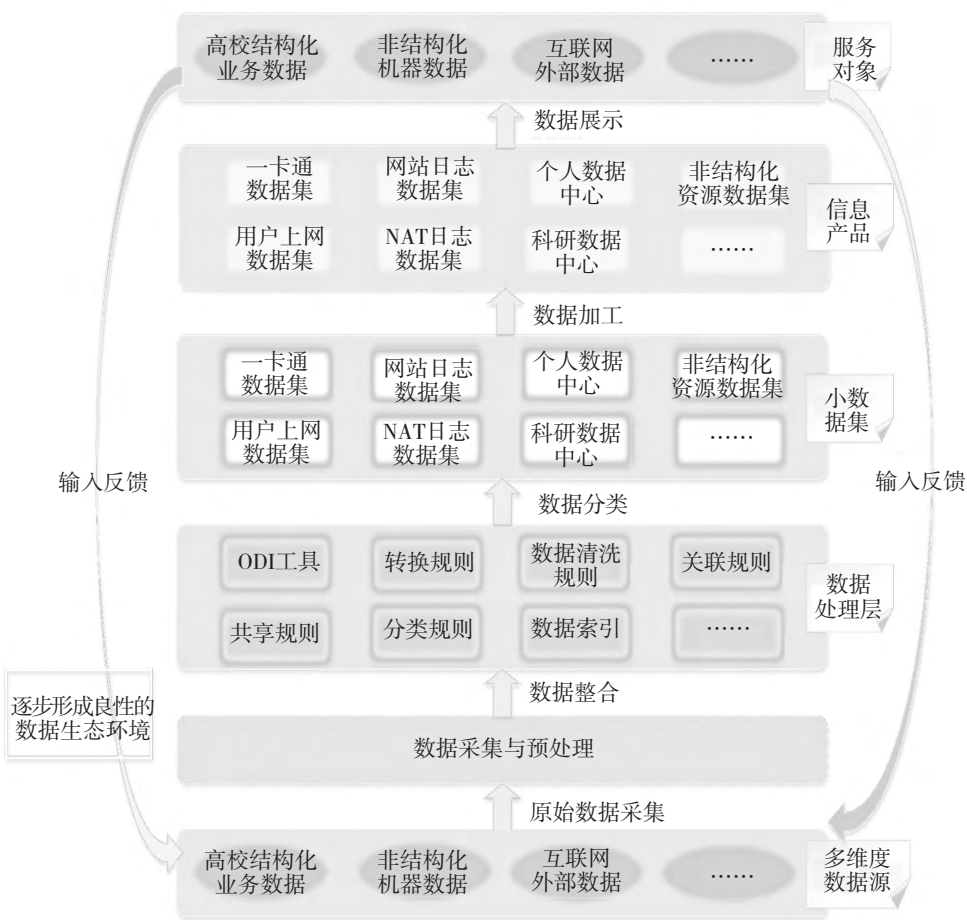


图4 教育大数据应用氛围视图

基于教育大数据应用氛围视图,智慧校园中的教育大数据体系模型可理解为是一条环回的“逻辑单元流”。如下页图5所示,教育大数据体系模型由以下6个单元构成。其中,单元1至单元6以逻辑承接的方式,构建起闭环式教育大数据生态体系模型。

单元1:“顶层设计”单元

顶层设计单元应建立在大数据生态体系链的顶端,是体系的纲领性部分。该单元主要完成对整个高校教育大数据发展规划的制定、建设目标的定位以及对数据标准(包括环境规划、管理策略、管理措施)的统筹。

单元2:“基础条件”单元

基础条件单元是高校构建体系的底层建筑。该单元包含建设资金的合理化调配、数据管理/治理/监控过程中专技人员队伍的构建、体系制度的建设及构建体系过程中一系列数据分析、应用工具的配备。

单元3:“数据环境”单元

数据环境单元是数据治理、数据分析挖掘与数据应用最为关键的基础环节。要做好数据治理,必须先厘清业务服务流程、梳理业务生命周期、规范数据字

典类型及数据标准。统筹协调业务部门,制定各类数据源的业务归口,建立数据交换与共享标准。其次,构建业务流程的数据服务模型和数据对应关系,建立数据服务的基础架构。最后,设计基于用户角色服务的“一站式”服务框架,为数据治理提供支撑。

数据全生命周期管理包括对数据的“发生—传递—分析—呈现”过程的全周期管理,通过构建合理的数据架构体系,在数据业务流推进的过程中,保证数据源可靠、数据传递畅通、数据分析精准、数据呈现可读^[24]。

数据质量的评判标准具有及时、完整、一致、有效、准确五大特性。围绕五大特性对接口数据、数据仓库层面的数据质量进行监控,建立业务数据监控规则,动态监测并报警体术数据异常。

单元5:“数据分析与挖掘”单元

业务不仅仅是独立存在,数据交换解决了数据孤岛的问题,但没有发掘数据价值的能力。数据分析与挖掘单元的主要工作是开展对不同业务系统产生的多维度过程化数据、结果性数据进行采集、抽取、清洗、关联、分析,实现应用与数据的全“解耦”,提升数据的“可复用”度。

单元6:“数据应用”单元

数据应用即是数据红利的释放,利用数据分析与挖掘部分的成果,不断提升高校管理者的科学决策水平,实现全面精准化的管理,辅助高校增强教学、科研、管理、服务能力,为建立师生个性化的综合服务体系提供支

撑。

(四)教育大数据生态体系建设目标^[25]

教育大数据生态体系的建设目标旨在实现“构建教育大数据的各个组件标准规范”“参与教育大数据的各个部门密切配合”“进入教育大数据的各个环节良性循环”“服务教育大数据的各个对象全面入微”。高校紧密围绕自身的发展需求,不断促进智慧校园创新应用,最大限度的挖掘数据价值,进而提升

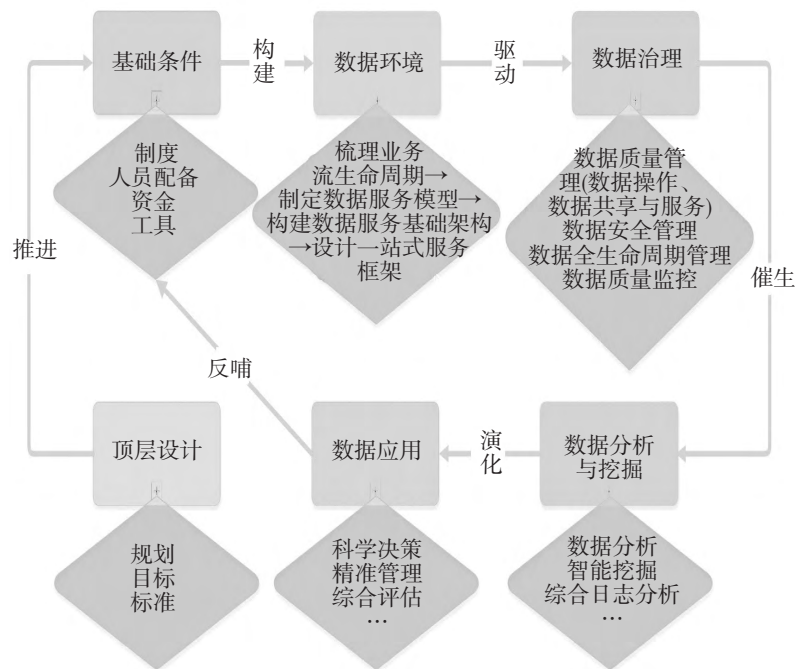


图5 教育大数据生态治理体系模型

单元4:“数据治理”单元

数据治理是体系生命周期“业务流”中最为核心的部分,包含对数据质量管理、数据安全、数据全生命周期管理及数据质量的监控。

数据质量管理是基于数据操作、数据共享与数据服务的统筹化治理,通过构建多维度元数据模型实现对数据共享关系的管理,实现对权威数据源的核查,对数据交换与共享过程中的脏数据、冗余数据进行清洗。

如图6所示,数据安全是在数据交换过程中,杜绝源业务数据原表进行写操作,数据交应基于共享数据平台通过数据集成工具(ODI、WebService等)以接口方式将信息系统需要的字段信息抽取到中间表或视图中,为大数据分析和数据服务管理提供可靠的数据支撑。在开放数据服务的同时,健全数据脱敏安全管理体系,实现对敏感数据和关键数据的分级与分类保护。

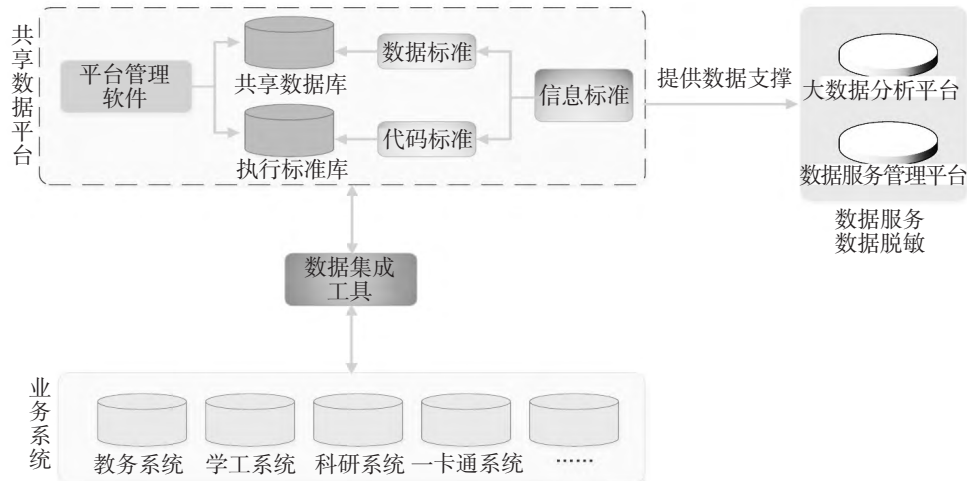


图6 数据安全框架



高校治理能力、推进高校治理体系的现代化发展。尽管高校教育大数据生态体系源于应用、管理、服务,但应在应用、管理和服务之上,是智能化、科学化、统筹化智慧校园的集中性体现,为高校全方位“精准管理、科学决策、综合评估”提供保障。

参考文献:

- [1] 蒋东兴,付小龙等.大数据背景下的高校智慧校园建设初探[J].华东师范大学学报(自然科学版),2015,(S1):119-125.
- [2] 李馨.高等教育大数据分析:机遇与挑战[J].开放教育研究,2016,22(4):50-56.
- [3] 陶坚.依法治校:高等教育的法治进行时[J].北京教育(高教版),2016,(1):44-46.
- [4] 杨现民,王榴卉等.教育大数据的应用模式与政策建议[J].电化教育研究,2015,(9):54-61.
- [5][11] 李艳,吕鹏等.基于大数据挖掘与决策分析体系的高校图书馆个性化服务研究[J].图书情报知识,2016,(2):60-68.
- [6][7] 比特网.数据全生命周期管理,大数据分析引领企业未来[EB/OL].
<http://bigdata.chinabyte.com/164/13066164.shtml>, 2014-09-02.
- [8] 王欣,张冬梅等.大数据环境下基于科研用户小数据的图书馆个性化科研服务研究[J].情报理论与实践,2017,40(10):85-90.
- [9] 李立睿,容仲华.面向科研用户小数据微知识服务研究(下)—基于系统动力学的视角[J].图书与情报,2016,(5):110-118.
- [10][14][21][25] 宓詠.教育大数据治理体系探索[R].南京:北京希嘉创

智教育科技有限公司,2017.

- [12][20] 袁芳.说说数据那点事儿[R].南京:北京希嘉创智教育科技有限公司,2017.
- [13][23] 袁芳.大数据时代高校数据管理与应用的探索[R].武汉:北京希嘉创智教育科技有限公司,2017.
- [15] 明廷庆.数据应用总体设想[R].武汉:北京希嘉创智教育科技有限公司,2017.
- [16] 涂庆华.“大势所趋,以数为据”——南京理工大学数据分析实践探索[R].南京:北京希嘉创智教育科技有限公司,2017.
- [17] 陈越.校园数据分析:问题与探索[R].武汉:北京希嘉创智教育科技有限公司,2017.
- [18] 魏大江.面向学生成长的大数据应用初探[R].南京:北京希嘉创智教育科技有限公司,2017.
- [19] 贺志勇.湖南农业大学数据治理工作汇报[R].武汉:北京希嘉创智教育科技有限公司,2017.
- [22][24] 余鹏,李艳等.高等院校大数据挖掘与决策分析体系的应用研究[J].现代教育技术,2016,26(8):102-108.

作者简介:

余鹏:硕士,高级工程师,研究方向为教育大数据应用、SDN网络(wildwolfyup@sina.com)。

李艳:硕士,馆员,研究方向为大数据应用、科技查新与检索(2328099300@qq.com)。

Research on Education Big Data Ecosystem in the Construction of Smart Campus

Yu Peng¹, Li Yan²

(1. Modern Educational Technology Center of South Central University for Nationalities, Wuhan Hubei 430074;

2. Wuhan Library, Chinese Academy of Sciences, Wuhan Hubei 430071)

Abstract: In recent years, the researches and applications of education big data have opened up new ideas for the construction and development of the smart campus. In view of the pain-point problems about teaching, scientific research, management, service and comprehensive decision-making, this paper is started from the six main causes in restricting the development of college big data, and delved into the characteristics and research significance of big data (ecology) in the background of smart campus. We selected the different scale, different category and different order as samples in order to summarize the data application and service mode of the education ecological system, as well as explore a kind of education big data of ecological architecture based on healthy development of smart campus. Combined with the design of education big data application atmosphere view, the paper is proposed an education big data ecosystem model to put forward to improve the precision of accurated data management, collaborative management services, intelligent teaching and research and scientific decision-making application. Finally, based on the long-term development demand of smart campus, the paper is further explored the construction target of education big data ecosystem.

Keywords: Education Big Data; Smart Campus; Education Big Data Ecosystem; Big Data Application Atmosphere View; Education Big Data Ecosystem Model

收稿日期: 2017年12月19日

责任编辑: 赵云建