

继续教育周刊

CONTINUING EDUCATION WEEKLY

总第109期



封面摄影：董彦



北京大学继续教育学院编

2017年4月24日



目录

Contents



主办： 北京大学

继续教育学院

承办： 综合办公室

编委会主任：

章政 李胜

编委会副主任：

杨虎 舒忠飞 屈兵

白彦

编委会委员（以姓氏笔画为序）：

马睿 刘宁 张孜孜

陈瑞 岳枫 曹建

常靖 廖来红

编辑部顾问：

李胜

主编：

刘宁

副主编：

文天骄 李丽

编辑：

董彦 门吉越

电子邮箱：

jxjyzk@163.com

【国际动态】

高校在线教育国际论坛举办 互联网+教育席卷全球 2

【国内动态】

教育部专家透析 VR 教育的三大困境 5

大数据驱动教育变革与创新 7

高职教育要“眼中有人” 9

【理论前沿】

教育大数据的核心技术、应用现状与发展趋势 10



【国际动态】

高校在线教育国际论坛举办 互联网+教育席卷全球

2017-04-23 来源：新浪教育¹

为拓展“互联网+”教育的国际视野，促进中美教育资源合作共享、融合互通，4月21日上午，由全国地方高校优课联盟主办，深圳市优课在线教育有限公司、美国 CMS Global、中关村互联网教育创新中心承办的“2017 高校在线教育国际论坛”在深圳大学举办。



本次“高校在线教育国际论坛”的主会场以中美在线教育的发展与展望为主题，在线教育行业专家齐聚一堂，围绕在线教育发展现状、教学模式、课程设计、平台建设、在线教学管理等高校教育和职业教育领域炙手可热的议题，展开了主题学术研讨和思想交锋，共同探讨在线教育的新趋势、新发展。



¹ 转载自新浪教育 <http://edu.sina.com.cn/l/2017-04-23/doc-ifyepsch2660509.shtml>



深圳大学校长李清泉在发言中表示，近几年来，我国数字化教育进展迅速，在线教育形式不断创新，成果显著，而“网络化、数字化、个性化、终身化”的教育体系将是未来在线教育发展的新趋势。未来中国在线教育应该集中各种学习方法的优点，线上线下多种教学方式相融合，尊重教育规律，坚持教育本质，服务学生学习。



美国科罗拉多州立大学校长 Becky Takeda-Tinker 博士介绍了目前美国在线高等教育的发展趋势和研究成果，分享了科罗拉多州立大学关于在线教育的线下服务、信息系统建设、学生完课率、学习质量、教学评估等多方面的实践经验。他认为，在线课程的完善和课堂教学的进步是一个互相促进的过程，这种促进对课程质量的提高有非常大的帮助。



加州大学伯克利分校网络学院运营总负责人 Dr. Scott Shireman 也在论坛上进行了演讲，他从伯克利在在线教育领域的实践经验、发展与推广计划、支持



与激励机制的设置、跨国教育联动等四个方面深入浅出地分享了加州大学伯克利分校网络学院的在线教育经验。



论坛上，全国地方高校UOOC（优课）联盟、深圳市优课在线教育有限公司、美国CMS Global公司、美国科罗拉多州立国际大学、美国荷晶大学、美国威尔明顿大学还共同成立中美高校在线教育联盟，该联盟将促进中美高校在线教育领域的合作发展、资源共享，推动中美高校在线课程的共建与共享，学分互认，为学生提供更好的国际教育机会。



圆桌论坛环节议题“中美在线教育的机遇与挑战”，深圳大学副校长王晖、中关村互联网教育创新中心总经理杨丹、亚利桑那州立大学国际合作处主任Rivadavia Alvarenga Neto、美国CMS Global公司CEO Charlie Nguyen、美国Wellsprings公司CEO Greg Schrade参加与讨论。



本次论坛还设置有四个分会场，与会专家学者将围绕在线教育科技、在线教学课程设置与开展、在线教育教师管理考核、学生在线管理与服务等主题，共同探“互联网+”时代的产教融合、创新教育新模式、在线教育促进高等教育变革、国际合作新格局等话题，深入解析了互联网智能化技术与高校教学深度融合下的教育生态大势。

【国内动态】

教育部专家透析 VR 教育的三大困境

2017-04-19 来源：中国教育装备采购网²

“很多 VR+教育没有涉及教育的核心问题，没有系统化、没有深入，没有引起教育行业的共鸣。”教育部虚拟现实应用工程研究中心主任周明全在接受采访时表示。

被称为虚拟现实元年的 2016，虚拟现实在全世界掀起了新的高潮。全国各地开始陆续建立了 VR 产业基地。越来越多的科技巨头以及创业团队在虚拟现实 VR 这块新领域大展身手。VR+教育也引发更多投资者的关注和教育人士的探讨。

周明全认为，虚拟现实是人类历史进步的重要一步，也是信息技术发展的必然。给人们提供如同现实的虚拟世界，输入的现实世界，输出的虚拟世界。虚拟现实是基于理论研究、计算研究的基础上，是认识世界的基本方法。

周明全坚信，虚拟现实改变我们的生活。也能颠覆传统教育。

² 转载自中国教育装备采购网 <http://www.caigou.com.cn/news/2017041925.shtml>



图片来自网络

VR+教育是未来教育发展方向？

VR同网络直播、人工智能一样受到投资界的关注。似乎没有这些“明星”技术，就跟不上互联网发展的潮流。教育作为特殊行业，潜力巨大，很多企业家纷纷打出教育牌，VR+教育、直播+教育、人工智能+教育都狠狠地火了一把，到底什么才是未来教育发展的方向？

在接受未来网记者采访时，周明全抛出一个问题，“现在信息技术进入好多领域，唯有教育领域进入很难。现在的孩子不用计算机和用计算机的教育效果差距到底在哪？这个问题需要去反思。”

周明全认为，信息化大潮将冲击教育领域，什么时候“大堤”冲破了，教育就开放了。

与网络直播、人工智能不同，周明全认为，虚拟现实将给教育带来彻底的变革。他表示，“中国古代讲求行先于知，由行致知。行就是要体验。比如地理课，可以通过虚拟画面看到沙漠变绿洲，绿洲变沙漠，历史课，可以通过虚拟看到恐龙如何进化等；带爆炸性的实验等这些实验都可以用虚拟现实来实现的。”

VR融入教育还有多远？

2016年9月，北京微视酷科技宣布获得华融天泽等知名投资公司千万级Pre-A轮投资，成为国内为数不多获得千万级投资的VR教育企业。

完美幻境、网龙华渔、黑晶科技等也纷纷开拓教育领域；泛美教育、新东方、巧克互动、邢帅教育、智课网等教育机构引入VR技术，百度、立思辰、川大智胜、天喻信息等巨头、上市公司布局VR教育。

但相比国外，国内VR教育一直是雷声大、雨点小。

“2016年在虚拟现实的资本投入乱砸钱，人才、技术跟不上。”周明全说。

虚拟现实在与教育融合方面，周明全认为主要有三个大的问题：第一个是概念的问题，教育是特殊的行业，需要的不光是技术，虚拟现实技术要与教育本身的发展规律结合起来。现在是搞教育的还是原来的模式，搞信息技术的在四处吆喝。而在之间是割裂的没有很好的融合；第二是技术上的问题，VR眼镜不太可取，对孩子眼镜伤害太大。要用全息技术等代替VR眼镜等。第三个问题是内容的问题，VR课程内容的制作，内容巨大且复杂，需要教师、政府、公司三方面结合，构成邮寄的教育改革系统工程。相比之前的D-Learning、E-Learning、M-Learning



三个阶段, V-Learning 的教学内容、方式、管理完全不一样, 对学生的测评也需要重新做规定。

除了这些, 在 VR 人才方面也存在很大的缺口。周明全认为, 预计到 2018 年中国有近 50 万 VR 人才缺口, 而这个数字也会根据时代发展不断变化。

大数据驱动教育变革与创新

2017-04-18 来源: 光明日报³

当前, 教育大数据研究和应用已经引起我国政府的高度重视。国务院《促进大数据发展行动纲要》提出“探索发挥大数据对变革教育方式、促进教育公平、提升教育质量的支撑作用”, 《教育信息化“十三五”规划》强调积极发挥教育大数据在教育管理平台建设和学习空间应用等方面的重要作用。“十三五”期间, 大数据与教育的深度融合已成为必然趋势。可以预见, 在今后一段时间我国教育大数据研究和应用将获得更快发展。

大数据给教育行业带来重大影响

大数据技术是 21 世纪最具时代标志的技术之一。国务院发布的《促进大数据发展行动纲要》中提出“大数据是以容量大、类型多、存取速度快、应用价值高为主要特征的数据集合”。

大数据给教育行业带来了重大影响。基于大数据的精确学情诊断、个性化学习分析和智能决策支持, 大大提升了教育品质, 对促进教育公平、提高教育质量、优化教育治理都具有重要作用, 已成为实现教育现代化必不可少的重要支撑。教育大数据的主要作用突出体现在:

第一, 有利于促进个性化学习。基于大数据, 可以精细刻画学生特点、洞察学生学习需求、引导学生学习过程、诊断学生学习结果。通过对学习者学习背景和过程相关的各种数据测量、收集和分析, 从海量学生相关的数据中归纳分析各自的学习风格和学习行为, 进而提供个性化的学习支持。例如美国亚利桑那州立大学运用 Knewton 在线教育服务系统来提高学生的数学水平, 系统通过数据分析区分出每个学生的优缺点并提供有针对性指导, 全校 2000 名学生使用该系统两学期之后, 毕业率从 64% 升高到 75%, 学生成绩也获得大幅增长。

第二, 有利于实现差异化教学。大数据可以在保障教育规模的情况下实现差异化, 一方面可以因材施教, 教师可以根据学生的不同需求推荐合适的学习资源, 另一方面可以达成更大的教育规模。比如, MOOC (慕课) 平台突破了传统教育中实体教室的限制, 课程受众面极广, 能同时满足数十万学习者学习需求。在教学过程中, MOOC 平台可依托大数据构建学习者体验模型对其线上课程进行评估, 进行线上课程的再设计、改变课程学习顺序、优化教学策略, 为每一个学习者提供不同的教学服务, 从而实现规模化下的多样化、个性化教学。

第三, 有利于实施精细化管理。传统教育环境下, 教育管理部门或决策制定者依据的数据是受限的, 一般是静态的、局部的、零散的、滞后的数据, 或是逐级申报、过滤加工后的数据。很多时候只能凭经验在做管理、决策。大数据根据社会各方面的综合数据来源, 可实现实时精确观察和分析, 对于推进教育管理从经验型、粗放型、封闭型向精细化、智能化、可视化转变具有重要意义。以学校课程设计为例, 美国加州马鞍山学院所开发的 SHERPA (高等教育个性化服务建议助理系统), 能根据学生的喜好为他们的课程、时段和可选节次作出推荐, 帮

³ 转载自新华网 http://news.xinhuanet.com/tech/2017-04/18/c_1120826948.htm



助学校课程设计咨询专家解决学生所面临的选课难题。此外，该系统还通过智能分析为教师和课程设计者提供反馈，使他们能有的放矢改进教材。

第四，有利于提供智能化服务。大数据可以采集分析管理者、家长、教师、学生的各方面行为记录，全面提升服务质量，为学习者、教师、家长等提供更好的服务。对教育大数据的全面收集、准确分析、合理利用，已成为学校提升服务能力，形成用数据说话、用数据决策、用数据管理，利用数据开展精准服务的驱动力。如在择校服务方面，运用大数据智能分析技术可助力破解教育择校感性化难题，推进理性择校。美国教育科学院推出的“高校导航(college navigator)”项目，该项目通过对全美7000多所高校各类资源指标（如所在地区、学费、奖学金资助、入学率和毕业率等）进行大数据分析并对所有大学进行排序和筛选，进而帮助家长和学生找到理想中的大学。

我国教育大数据开发利用面临难题

我国教育大数据开发利用已经具备一定基础，但还面临不少问题。自《教育信息化十年发展规划（2011-2020年）》颁布实施以来，通过建设教育资源公共服务平台和教育管理公共服务平台积累了教育教学和教育管理两方面大量数据；同时，我国在基于大规模在线开放课程的数据汇聚方面也已初具规模，这都为开展教育大数据研究和应用提供了一定基础。然而，还存在以下不足：

第一，数据类型比较单一。不是实时采集，很多是结果性数据，而并不是及时的、过程性数据。

第二，数据规模还很不够，目前通过“两平台”和第三方服务机构积累的数据量相比于我国庞大教育规模而言还较为有限，难以满足深度挖掘分析的需求。

第三，缺乏对教育数据的深度分析、挖掘、利用。对海量教育数据的挖掘分析具有很强的专业性，缺乏有效的技术、工具、产品、服务。

第四，跨界数据的整合不够。教育行业的很多应用需求不是光靠教育系统内部的数据能分析的，还需进行跨界数据整合，比如人口数据、地理数据等对于提升教育治理水平就具有重要价值。

第五，对大数据研究和开发的支持力度不足、队伍不强。我国虽已发布大数据的国家发展计划，但对于教育大数据的各项研究还需要更多关注，不仅需要尽快考虑在教育大数据研究专项等方面进行布局，还需要一批既懂技术又懂教育的专业机构和人员，目前我们在力量准备上有所不足。

第六，教育大数据标准与安全问题面临挑战。对教育大数据的准确、规范、统一使用和管理缺乏相应法律法规和技术标准支持，在教育大数据共享、开放、交换、交易、安全等方面的有效监管和规范还不太到位，制约教育大数据的健康发展。

如何积极稳妥发展我国教育大数据

为积极稳妥发展我国教育大数据，笔者认为应注意以下问题：

第一，制定大数据教育应用的宏观规划和行动计划。基于国家大数据驱动战略，尽快编制我国教育行业大数据研究、开发、应用的宏观规划和行动计划方案，指导各级教育管理部门、教育机构推进教育大数据开发和应用。

第二，加强我国教育大数据基础建设。依托现有国家教育资源公共服务平台、中小学生学籍系统等基础数据库，建立覆盖全国的教育大数据基础平台和管理体系，一方面有利于充分挖掘利用，另一方面有利于合理监管。

第三，积极开展教育大数据的应用创新和示范。以教育机构、企事业单位为依托，在全国建立若干教育大数据应用示范区、示范点，在个性化学习服务、教



育管理决策等方面开展有针对性的应用创新示范，为后续推广积累经验。

第四，大幅提升我国教育大数据服务创新能力。成立教育大数据联盟，组建专业团队，启动专项工程，针对教育大数据基本理论、关键技术和应用模式开展研究，聚焦学生核心素养、高考改革等教育热点问题，开展应用示范。

第五，加快推进我国教育大数据标准规范和法律法规建设。积极制定教育大数据采集、交易等相关法律法规，尽快启动教育大数据相关标准规范研究制定工作，保障教育大数据的安全、规范获取和及时、有效利用。

高职教育要“眼中有人”

2017-04-23 来源：人民网-人民日报⁴

高等职业教育经过两轮评估和示范、骨干院校建设，发展重点从外延深入到内涵建设，利益相关方的认可度稳步提高，但学生的地位和作用没有得到突显。在供给侧改革视域下，高职院校是职业教育服务的提供者，学生是职业教育服务的体验者和消费者。因此，树立“以学生为中心”的人才培养观，通过提供高水平的职业教育服务，全面提升人才培养质量，就成为高职院校改革发展的重点。

“以学生为中心”的人才培养观，就是在充分尊重学生个体独立性基础上，从学生的知识发展需求出发，按照职业能力成长规律和人的可持续发展需要，构建服务和成就学生的育人体系，提高服务学生的能力和水平，实现成就学生的目标。这既符合人的全面发展理论，又反映了高职教育供给端转型升级的内在要求。

重塑学生自信，矫正学生品格，提升自我认同是提供“有效性”供给的基础和条件。《2015 中国高等职业教育质量年度报告》显示，有 91% 的 2014 届高职毕业生是家庭第一代大学生，52% 来自“农民与农民工”家庭，这些学生普遍存在着学习经历坎坷、自我认同度低、自信缺乏等问题。提高职业教育供给的有效性，一是搭建学生与校长、学生与教师、学生与学生之间的交流平台，使学校充分了解并尊重学生的需求；二是构建激励学生成长成才的荣誉体系，满足学生多样化需求，促进学生个性化发展；三是抓住新学期升旗仪式和重大节日等关键节点，强化仪式教育在学生成长成才中的作用。通过搭平台、建体系、抓节点，实现为学生做主向由学生做主的转变，提升学生自我认同。

厚植文化底蕴，精湛一技之长是提供“精准性”供给的路径和手段。应当为学生提供涵盖文化素养和职业技能的双重服务菜单。一方面，推进产业文化进校园，打造产教融合的专业文化，培养学生的职业理想和职业态度；推进地方文化进校园，传承优秀的传统文化，培养学生的爱国之心和责任关怀；推进境外文化与校园文化交流融合，营造多元文化氛围，培养学生的世界眼光和民族情怀；将文化的积淀、传承、发展融汇于高职教育，以文育人，以文化人，最终实现无需提醒的文化自觉。另一方面，利用高水平的双师型教师队伍、先进的实训设施、紧贴工作过程的教学项目等优质教育资源，提供高质量的服务性教育教学产品，开展个性化精细培养，使学生精通一门技艺，提高可持续发展能力，并在此基础上实现高质量就业，让学生有尊严地走向社会。

锻造“工匠精神”是提供“引领性”供给的目标和归宿。工匠精神是一种对职业的敬畏、对工作的执着、对产品和服务追求完美与极致的价值取向，是精益求精。虽然短短 3 年的高职教育无法培养出大国工匠，但可以而且应该培养学生

⁴ 转载自济宁大众网 http://jining.dzwww.com/jiaoyu/jiaoyuarc/201704/t20170423_15819162.htm 大众



的工匠精神，高职院校要牢固树立质量意识和品牌意识，将敬畏写入人生信条，将专注刻进理想信念，将执着融入师生血液，持续强化精益求精的工作态度和求，不断提升服务学生的能力和水平，将锻造学生工匠精神融入人才培养全过程。

教育的本质是为人的发展提供服务，校长心中有学生，学生才会有未来，教师心中有学生，学校才会有希望。高职教育应当树立“以学生为中心”的人才培养观，强化服务与成就学生理念，全面提升人才培养质量，为国家发展、民族进步提供技术技能人才支撑。

【理论前沿】

教育大数据的核心技术、应用现状与发展趋势⁵

作者简介：孙洪涛，博士，中央民族大学现代教育技术部高级工程师，研究方向：教育大数据，教育信息化规划。北京 100081；郑勤华，博士，北京师范大学教育学部副教授，研究方向：教育经济学，教育大数据。北京 100875

内容提要：大数据领域近年来蓬勃发展，作为大数据的细分领域，教育大数据具有推动教育变革的巨大潜力。大数据技术正在快速演进之中，这为大数据应用提供了新的可能。为了深入分析教育大数据发展，文章从大数据技术的最新进展入手，从基础设施、分析技术和领域应用方面阐述了大数据的发展趋势。进而通过教育领域大数据构成与特征的分析，对教育大数据的含义进行了解析。并结合国际范围内教育大数据典型实践，从适应性教学、教育规律发现和精准管理支持的角度，对教育大数据应用进行了探讨。最后，针对我国教育大数据的发展状况，对教育大数据发展所面临的挑战进行了分析，并提出了应对挑战的建议。

关键词：教育大数据 大数据技术 数据湖 云计算 人工智能 适应性学习 精准管理

标题注释：本文系北京师范大学自主科研基金项目“学习者在线学习状态分析与可视化工具研发”（SKZZB2015013）课题成果，并获得中央高校基本科研业务费专项资金资助。

[中图分类号]G434 [文献标识码]A [文章编号]1672-0008(2016)05-0041-09

大数据是近年来快速发展的技术领域。关于大数据的研究与应用与日俱增，并不断深入影响社会生活。购物推荐、路况分析乃至高考预测等与大众密切相关的应用，充分展现了大数据的力量。2016年3月，AlphaGo 与李世石的人机大战，让人们从更深层次上认识了大数据驱动下的人工智能对人类社会的深层影响。根据大数据版图(Big Data Landscape) 3.0 版本的描绘，大数据相关基础设施、分析工具和应用系统都在快速发展中[1]。这个逐年扩展的图景表明了大数据的疆域正在不断延展，领域应用不断深化，影响力与日俱增。

在教育领域中，大数据已经在多方面引起了研究者和实践者的关注。无论是从研究范式、技术应用，还是实践案例都在快速发展之中。教育大数据正在成为教育领域不可忽视的新型驱动力，在教育教学研究与实践中发挥着越来越重要的作用。

作为一个新兴领域，大数据技术仍在快速迭代之中，新方法、新工具和新模式不断涌现。在教育大数据这个细分领域之中，在契合大数据发展整体趋势的同

⁵ 浏览网址 http://ex.cssn.cn/jyx/jyx_jyjsx/201704/t20170421_3494998_7.shtml



时,具有自身的鲜明特性。在教育大数据日趋瞩目的今天,在研究大数据技术的基础上,分析教育大数据的定义内涵、实践范例、发展趋势与面临挑战,有助于我们把握教育大数据的整体图景,因应技术发展,推动教育的系统化变革。

一、大数据技术的发展趋势

大数据技术的缘起,可以回溯到2004年谷歌公司提出的MapReduce模型[2]。在十几年时间里,大数据技术从概念走向应用,形成了以Hadoop为代表的一整套技术。时至今日,大数据技术仍在快速发展之中,无论是基础框架、分析技术,还是应用系统都在不断演变和完善。据统计,2015年美国大数据初创企业获得的融资额达到了66.4亿美元,占整个技术领域总融资额的11%。这代表着大数据领域具有蓬勃的活力并受到市场的肯定。大数据技术的发展方向是技术与应用需求相互推进的结果,对大数据技术趋势的分析,有助于从更本质的层面理解这个领域的现状。

(一)基础架构

历经多年发展,大数据基础设施正在向着快速、便捷与整合的方向发展。Hadoop框架是大数据分析的重要基础框架。但它存在着计算速度慢、运维复杂等问题。基于Hadoop衍生出了如Spark、Pig等框架,正在不断提升计算性能和优化处理流程。与Hadoop相比,Spark的抽象层次更高,计算速度更快,编程更加简便。更重要的是,Spark提供了统一的数据平台,通过不同的模块支持了不同类型的数据应用。通过Spark Core支持批处理,通过Spark SQL支持数据交互,通过Spark Streaming支持流式存储,通过MLlib支持机器学习,通过GrphaX支持图计算[3]。

在大数据基础设施中,各种新技术不断产生,数据湖(Data Lake)和雾计算(Fog Computing)分别从数据的集中与分布的不同角度给出了解决方案。数据湖是大型的基于对象的存储库,数据以其原始格式存储。不需要对数据进行转换,就可以进行全面的监控和分析,并建立数据模型。与一般意义的数据库不同,数据湖不需要改变原始数据的结构,而是支持分析原始数据。这个方式消除了数据抽取、转换和加载ETL的成本。为了达到不改变数据结构直接存储和技术的目标,数据湖对元数据有很高的要求。目前,数据湖技术仍在起步阶段,还存在原始数据差别大、类型复杂、分析应用困难等问题。但它有助于企业完成更长远的数据规划,建立数据治理结构,并预先解决安全问题[4]。数据湖与一般大数据汇集方式的对比,如表1所示。

表1 数据湖与一般大数据汇集方式的对比

	汇集方式	元数据要求	数据结构	计算	安全性
一般大数据	需花大量时间转换格式	低	结构化、非结构化分立	移动数据	高
数据湖	原始格式汇集	高,需要通过元数据实现聚合	结构化、非结构化整合	不移动数据	低

与数据湖侧重数据的聚集不同,雾计算则提出了一种分布式解决方案。雾计算这一名词最早来自网络安全领域,后来由思科(Cisco)公司借用,并赋予了分布式计算的含义。思科将雾解释为“更贴近地面的云”,雾计算是云计算的延伸。与云计算不同,雾计算并非由性能强大的服务器组成,而是由性能较弱、更为分散的各类计算模块和智能网络设置组成,这些低延迟且有能力进行位置感知的模块可以融入各类基础设施,乃至生活用品[5]。



可以预见，随着物联网的不断发展，来自各类终端的数据量会激增。面对这一情况，云计算的瓶颈可能会凸显。在雾计算中，数据、分析和应用都集中在网络的终端节点，只在需要的时候汇集到云中。云计算与雾计算的对比，如表2所示。

	计算单元	计算能力	节点数量	分布情况	数据存储	延迟	位置感知
云计算	服务器	强	较少	物理集中与分散并存	全部存储于云端	较高延迟	无
雾计算	各类处于网络边缘设备	弱	较多	物理分散	存储于网络边缘设备	低延迟	有

雾计算将计算能力延伸到了网络的边缘的各类智能设备。在这种模式下，智能设备的管理与交互就变得非常重要。比如，比特币的底层技术“区块链”(Block Chain)形成了行动登记、权属确认和智能管理模式。这为通过网络实现各种智能终端和设备实现自我管理和智能交互，提供了新的技术支持[6]。

数据湖和雾计算着眼于大数据的源头和终端，从分布和集中两个角度提供了解决方案。诚然，这些方案需要通过实践进行检验。但总体而言，数据湖和雾计算代表着大数据分析基础设施的发展趋势，即采用更灵活的方式获取和处理终端数据，合理分布计算负载，对核心数据进行广泛汇集，通过定制标准实现数据治理。

(二) 分析技术

分析技术是基于大数据进行模型构建，并进行评价、推荐和预测等具体应用的基础。大数据分析技术在近年得到快速发展，智能化、实时化和易用性成为了分析技术的发展特征。

1. 智能化

在分析技术方面，大数据与机器学习相结合形成的新型人工智能，已经成为近年最引人瞩目的趋势。大数据与机器学习正让数据分析在统计分析的基础上，更快速地实现智能关系发现和预测，如图1所示。AlphaGo就是这一趋势的典型应用范例。在海量数据的基础上，以深度学习为代表的创新算法，通过大规模并行计算，不断迭代演化，最终形成了能够战胜人类的数据智能。

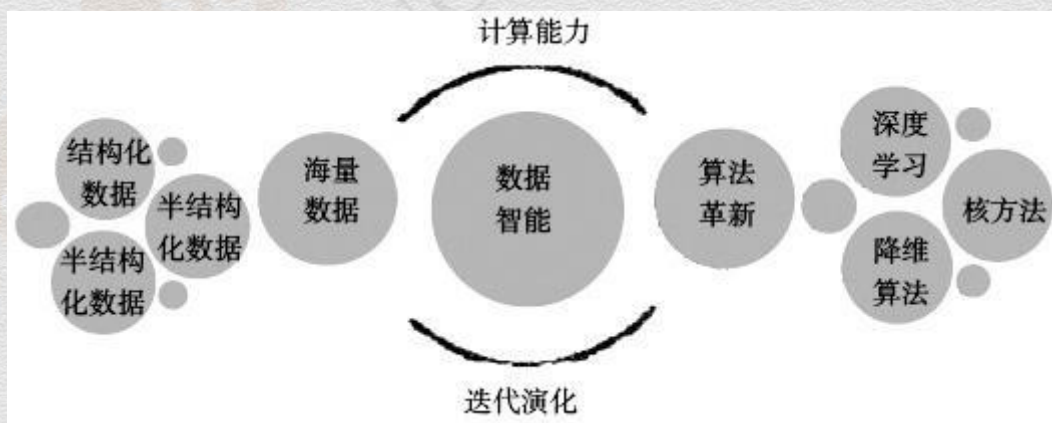


图1 数据与算法迭代演化形成数据智能

大数据与机器学习整合所实现的人工智能，其意义不限于特定的领域应用，而是实现了一般性人工智能技术的突破。这一突破将在医疗、交通、金融和教育等为代表的各个应用领域产生重大影响。从更为广阔的角度，以智慧城市为代表的智能化系统解决方案，预示着智能化大数据技术综合应用的未来前景。由各类设备和传感器获得的数据，可以成为智能化分析的数据来源。基于大数据的机器学习在完成海量数据汇集与分析的同时，不断演化、提高自身智能水平。数据分



析结果驱动智慧城市各个组成部分的智能化活动,基于数据智能的新型技术架构,为未来城市的智慧生活奠定了基础。

2. 实时化

实时分析是大数据技术的另一个发展方向。随着大数据技术的深入发展,各类应用对于数据的实时分析和处理的要求不断提高。与针对历史数据的聚合和分析不同,实时数据分析具有更强的时效性,也对数据存储、计算和呈现提出了更高要求。Hadoop 中的批处理框架在对实效性要求较高的分析,例如,实时用户行为分析、用户分类和推荐等应用场景中的局限日益凸显。Spark Streaming、Samza、Storm 等流式实时计算框架应运而生。以 Spark Streaming 为代表的实时分析框架具有优秀的调度机制,快速的分布式计算能力,在数据的汇聚和批处理之间通过关键参数建立平衡,提升了数据吞吐量和性能,对实时计算提供了有效支持[7]。实时性预示着大数据将更深度地融入人们的工作和生活之中,在交通、翻译等需要及时响应的领域中,大数据会体现出更强大的作用。

3. 易用性

近年来,随着技术的不断成熟,大数据应用的门槛不断降低。Google、微软等巨头不断推出大数据技术平台。我国互联网三巨头百度、阿里和腾讯分别推出了百度开放云、阿里数加和腾讯大数据平台,在应用技术方面提供了全面的支持。从数据汇集、模型构建到可视化应用方面都提供了高质量的解决方案。并且,这些分析框架中存在很多优秀的开源项目,如,Caffe、Torch 等[8]。Google 为 Tensor Flow 的开源分析工具提供了一个重要选择, Tensor Flow 的开发者来自 Google Brain 团队,它整合了 Google 在搜索引擎、电子邮件和翻译、图像识别等方面的分析成果。并且应用了数据图技术(Data Flow Graphic)将模型构建过程和产品开发紧密结合,在完成建模实验之后就可以直接将代码应用到产品中。易用性为大数据在垂直领域的应用铺平了道路。

(三) 领域应用

在基础框架和应用技术的支持之下,大数据在各个领域中的应用也在不断快速地深入发展,展现出了领域应用深化与融合、可视化应用广泛和产业生态链萌发的特征。

1. 领域深化与融合

大数据在方法论层面上影响着多个领域的研究与实践[9-11],作为新的研究范式影响着众多学科。在各个领域应用中,大数据作为基础方法与工具有着一定的普适性,也具有鲜明的领域特征与领域差异。数据不同于金融、交通、零售等领域有着较为明确的量化指标作为机器学习的依据。在教育等社会科学相关领域中,大数据分析模型建立过程中形成的类量化指标往往很难获得。这就使得教育领域的模型构建具有了一定的独特性。同时,教育教学自身的周期性和复杂性,也为模型构建提出了新的挑战。

随着大数据的发展,领域应用将逐步深入。在各个领域中需要借助领域知识,针对领域问题进行深层次研究与实践。在此过程中,以数据为桥梁,各个领域的融合将成为可能。例如,始于气象系统的 DMSP/OLS 夜间灯光数据,已经在遥感测绘、城市规划、人口估计、国民经济测算、能源消耗以及生态环境影响评估方面取得了令人瞩目的成果[12]。基于大数据,各个领域自身发生深刻变化的同时,领域之间的比较出现加速融合的趋势。大数据技术在领域内的深入发展,和领域间的融合发展将日趋重要。

2. 可视化应用



可视化是大数据应用的呈现层面,直接面向终端用户,并通过各类应用场景服务各类人群。数据可视化可以通过多种方式实现,从较为底层的R语言Ggplot扩展包、D3函数库,到SPSS Modeler、Tableau等数据分析和可视化工具。数据可视化的方法和工具种类繁多,近年来,可视化工具的应用门槛不断降低。SAP、Tableau等重量级数据分析企业都推出了移动端数据可视化工具。以SAP的Roambi为例,只需要导入数据集,选择模板,Roambi就能够马上完成精美的可视化图表并支持互动[13]。Tableau不仅推出了Tableau Mobile支持移动端数据分析,还通过Tableau Public和Desktop等工具,构建了包含桌面分析、在线发布和移动应用的整体可视化方案[14]。

在各类工具支持下,数据可视化的应用门槛大大降低,为更加广泛的应用奠定了基础。数据可视化作为大数据技术的表现层,是数据分析与洞察的“最后一英里”。随着这个环节的不断优化与人性化,数据分析的广泛应用指日可待。

3. 生态链萌发

2015年8月,国务院发布的《促进大数据发展行动纲要》,将大数据定位于推动经济转型发展的新动力,重塑国家竞争优势的新机遇以及提升政府治理能力的新途径[15]。《促进大数据发展行动纲要》成为了大数据产业发展的政策依据,必将对大数据产业发展起到催化作用。大数据产业的资金投入、基础设施、数据标准、应用平台、区域实践必将呈现加速发展趋势。同时,正如前文所述,大型互联网企业如百度、阿里和腾讯等,已经在大数据领域发力,并开始构建基础设施、制定标准、推广应用,在各个应用领域的大数据实践也在快速开展。

可见,在政策重点支持、工具平台日渐成熟、领域应用不断深入的合力之下,大数据产业链正逐步形成,生态体系正在孕育之中。生态链将催生一系列数据标准,形成多种整合型技术路线,打通原始数据到终端应用,将大数据应用推向新的层次。

二、教育大数据的含义

教育大数据的含义,需要从数据和技术两个层面进行解析。在引用较多的大数据定义中,维基百科定义[16]和麦肯锡(McKinsey)定义[17]都强调了大数据的量,无法用常见数据工具处理;而高德纳(Gartner)定义则着眼于数据的特性与价值[18]。为了解析教育大数据的真正意义,需要对教育大数据的构成和特性进行分析。

在教育大数据的构成方面,在线学习的数据首当其冲。可以说教育大数据的广受关注,与在线教与学的盛行有着密不可分的关系。在舍恩伯格的《与大数据同行——学习和教育的未来》一书中,第一个大数据教育应用案例就来自在线学习。随着在线教学的日益普及,在教与学过程中,由学习管理系统和各类移动设备所记录下来的各类海量数据,成为分析教学过程的重要来源。这些数据包括记录学习过程的行为数据,记录学习结果的评价数据,以及学习形成的社会网络关系数据等。由这些数据拓展开来,教育大数据还包含着各类学生个人信息数据、教学管理数据等。可见,教育大数据来自于教育教学的主体和过程。

依照不同层级的主体和教育教学活动的各项内容,教育大数据可以分为四个层次和六大类型。四个层次包括个体、学校、区域和国家;六大类型包括基础数据、教学数据、科研数据、管理数据、服务数据和舆情数据。其中,基础数据包括以人口学为代表的学习者基本信息数据;教学数据包括教学过程中涉及的过程、内容和结果数据;科研数据包括各类教育教学实验与科研项目当中所获得的数据;管理数据包括各类教育管理系统当中所记录下来的数据,如,学生的学籍数据、



档案数据和各类统计数据等等；服务数据包括各类与教育教学相关的服务系统当中记录的数据，如，各类师生生活服务、图书档案服务等；舆情数据包括各类公开媒体中与教育相关的数据，如，各类教育新闻数据、微博等社会网络系统中教育相关数据等。

从特征的角度看，大数据的特点往往被概括为4V，包括海量规模(Volume)、快速流转(Velocity)、多样构成(Variety)和巨大价值(Value)。教育大数据的特征与4V既有重合又有不同：首先，从规模上看，教育大数据的体量尚未达到零售业、电信业等领域的规模，但已经超出了传统数据工具的处理能力。其次，从流动速度的角度，教育大数据流转速度相对较慢，并不像交易数据、搜索数据或通讯数据具有快速流转的特性。相应地，教育教学的周期性决定了教育大数据具有典型的周期性。进而从数据构成方面看，教育大数据中非结构化数据，特别是音视频数据占很大比重。这些数据来自课堂录像、教学资源等，不同于传统数据库记录的数据，具有一定的分析复杂性。同时与电商等领域中步骤清晰、结果明确、周期较短的交易活动不同，教育教学活动具有更高的过程复杂性。通过教育大数据分析发现规律也就更为困难。可见教育大数据的特征可以概括为强周期性、高复杂性和巨大价值。

综上所述，我们可以把教育大数据定义为：服务教育主体和教育过程，具有强周期性和巨大教育价值的高复杂性数据集合，具体如图2所示。

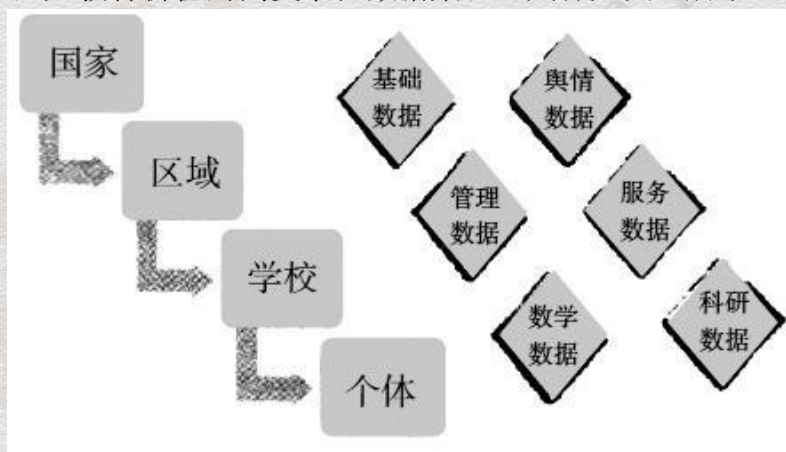


图2 教育大数据的构成

三、教育大数据的应用

对于教育大数据的应用，研究者从不同的角度提出了各自的思考。祝智庭教授从研究范式的角度，提出了大数据对教育技术研究方法的启示，并着重强调了数据支持下的自适应学习。郑燕林和柳海民认为，教育大数据的应用主要是对教育评价和教育教学决策的支持[19]。胡弼成和王祖霖将大数据应用总结为通过评价和预测促进教学有效性，基于变化的教育形式和复杂关系推动教育决策的科学性，完整、全面、动态的质量监控体系[20]。杨现民等研究者认为，教育大数据应用可以分为政策科学化、区域教育均衡、学校教育质量提升、课程体系与教学效果最优化、个体的个性化发展等层面[21]。

大数据对教育领域的冲击是全面性的。它能够改变个体学习者的学习状况、对教育规律的认识深度、教育政策的制定方式，乃至整个教育的结构。从需求的角度，教育大数据的应用可以概括为五个层次，即学习、教学、研究、管理与政策。学习层与教学层需求着眼于适应性学习；研究层需求着眼于发现教育教



学规律；管理层需求着眼于精细管理和科学决策；政策层需求来自获得机制设计依据，如图3所示。



图3 教育大数据的应用

针对不同层次的需求，教育大数据应用形成了各种产品和服务。从适应性教学到动态跟踪测评，从管理模型构建到数据共享门户，各种类型的应用勾勒出了大数据影响教育领域的整体图景。

我们不妨通过国际范围内典型的技术、产品和服务，从适应性教学、教育规律发现和精准管理支持三个方面，对有较大影响的教育大数据国际应用进行分析，以期对我国教育大数据发展应用提供借鉴。

(一) 适应性教学支持

适应性教与学是教学的最优化状态。适应性教学中的内容、方法和过程都可以根据学习者的状况来进行定制，让每个学习者都有可能获得适合自己的最大程度的发展。适应性教学的实现，需要基于学习者的个体特征和学习状况的全面分析。大数据为追踪和整合这些数据，并对学生进行个性化支持提供了可能，如图4所示。

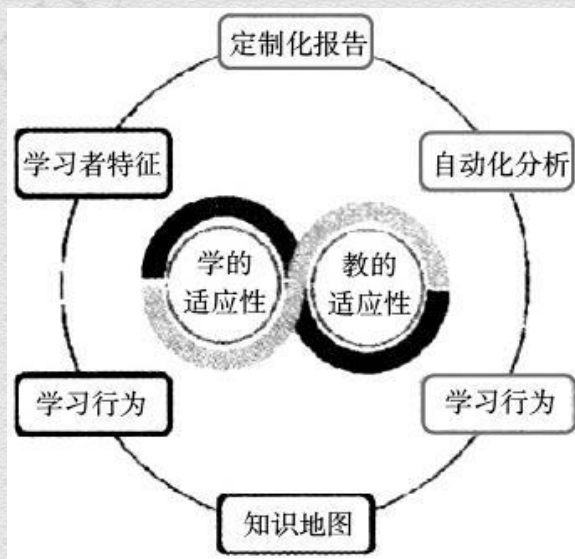


图4 适应性教学的构成



最为常见的适应性教学系统来自在线学习领域。在各种学习管理系统和在线学习平台中,学习者的学习过程能够得到完整记录。学习过程的记录结合人口学和学习风格等学习者特征数据,可以清晰地表征学习者的学习路径和学习者特征,在有效记录学习过程、综合评价学习状况的基础上,进行诊断和推荐,开展有针对性的教学。

适应性学习支持几乎已经成了在线学习的“标配”,在每一个商业在线学习平台中,都有不同程度的适应性。内容推荐是适应性的一种主要形式。然而,真正有效的适应性教学系统需要整合三个系统,即知识系统、行为系统和特征系统。通过知识系统来描绘知识体系;通过行为系统来记录学习、练习和反馈过程;通过特征系统去分析学生的个体特征和学习特质。

当前,最具代表意义的适应性学习系统当属 Knewton 和可汗学院(Kehan Academy)。此类适应性学习系统重点支持了学生的学。学习系统试图扮演教师的角色,对学生的学进行自动化记录、诊断和干预。适应性学习的另一个层面对教师教的支持。应当看到,目前的学习系统还存在诸多局限,在 MOOCs 发展的初期,教师是否会被在线课堂所替代曾经成为一个引人瞩目的话题。然而,教师作为教育过程中的关键角色,不可能在短时间消失,而是会借助技术实现专业水平提升和角色转变。大数据将成为教师教学的强大助手,帮助教师更好地发挥自身作用,更好地促进学生的学习。

在大数据技术支持下,教师可以根据自身的需求对学生的学习进行监测,并通过自己设定的标准,对学生进行自动化或半自动化的评价。在数据的支持下,教师可以结合自身的教学经验对学生进行诊断和干预。在教师的训练下,大数据工具将对教师的教提供更有力的支持。大数据工具将成为教师最好的帮手,而不是竞争者。

以 Masteryconnect 为例, Masteryconnect 对教师的教提供了全面的数据化支持,它从教师的日常工作出发,提供了数据采集、分析、呈现和基于数据的协作支持。教师可以在其支持下采集各种教学数据,包括课堂观察数据、答题卡数据、量表数据和在线测试数据。在采集数据之后, Masteryconnect 可以进行自动化分析和可视化呈现。分析的结果可以通过该系统分享给其他教师,教师可以在数据的基础上进行交流和协作。Masteryconnect 为各种形成性评价提供了全面的支持,教师可以自行建立教学内容的结构,并为各个模块和知识点设计问卷、练习、试卷等各种测评方式。测评可以发布到 PC 和移动设备,学生可以选择自己喜欢的方式完成测评。同时,测评结果可以形成定制报告,并发送给家长[22]。

对教与学的支持是大数据在适应性学习中应用的两个侧面。基于数据,学生的学习状态得以完整记录,学习系统可以推送定制化内容,教师可以开展更具针对性的教学。可见,数据正在改变着线上和线下的教学过程。

(二) 教育规律发现

教育研究是一项复杂的系统性研究。长期以来,小样本量、个案研究对教育规律探索起到了重要作用。而教育大数据的引入,大大拓展了教育规律探索的视角。图灵奖得主吉姆·格雷在《第四范式:数据密集型科学发现》一书中,提出了一种新型研究范式,即数据密集型研究,这将成为大数据时代教育研究的利器之一。

换言之,在大数据的驱动下,教育研究将出现不同的态势,通过挖掘、分析教育大数据,研究者可以量化学习过程,表征学习状态,发现影响因素,找到干预策略,从更深的层次揭示教育规律。诚然,规律发现并非易事。但可以肯定的



是，在多来源、大体量数据的基础上，通过技术手段进行数据汇集和共享，组织研究者进行群体协作，开展大量能够进行标准化，具有对比意义的研究，最终更易发现真实的教育规律，如图 5 所示。



图 5 通过大数据揭示教育规律

这里可以通过对美国的三个大数据教育应用的案例，即“预测分析报告项目 (Predictive Analytics Reporting, PAR)、数据商店 (Data Shop) 和 Data.gov”的介绍与分析，充分了解大数据标准化研究和数据门户对教育规律探索的意义。

美国的预测分析报告项目 (PAR) 对高校学生的学习状况进行了全面分析，在学生辍学等重要风险的预测方面进行了探索。从 2011 年开始，该项目与美国的高校合作，建立了标准化数据收集框架，通过对学生学习数据的收集和分析，发现影响因子并构建预测模型。该项目分析的原始数据包括学生人口学数据、教学管理数据、学习过程数据、成绩数据和学生财务信息数据等。通过分析，该项目建立了通用分析标准和风险预测模型，并为各个学校提供了定制化风险因素模型。迄今，PAR 已经服务 351 家院校，分析了超过 2000 万条课程数据。PAR 也为学校分析学生学业表现提供了综合视角，为防止学生辍学提供了有效手段[23]。

教育大数据研究需要大量汇集数据，数据汇集需要大量研究者的群策群力，也需要行之有效的组织机制。匹兹堡大学学习科学中心 (Pittsburgh Science of Learning Center) 的数据商店 (Data Shop)，是美国自然科学基金支持建立的学习科学数据库。经过十多年的建设，已经成为全球最大的学习数据分享社区[24]。数据商店具有数据存储和数据分析两大类功能：一方面，它为全球学习科学研究者提供安全的数据存储与共享工具；另一方面，它提供了数据分析工具和调用接口，便于分析应用。数据商店中的数据，分为公开数据和私有数据，研究者可以根据需要选择自己的数据是否要公开。数据商店中的数据，包括教学软件应用数据、在线课程数据、智能教学系统 (Intelligent Tutoring Systems)、虚拟实验室数据、协作学习系统数据等。在数据汇集和分享功能的基础上，数据商店提供了丰富的数据分析功能，支持探索性统计分析和数据挖掘，提供 Web Service 支持远程调用，以及 R 语言、Excel 等工具的接口。

美国教育部在教育研究与实践数据汇集方面开展了大量工作，汇集了大量数据。截至 2016 年 5 月，在其数据门户 Data.gov 中，汇聚的数据包含了 325 个大型数据集。数据包含范围很广，涉及人口统计、学习成绩、贷款情况、校园安全等情况。Data.gov 针对各类数据提供了多种数据格式，同时，提供了在线数据



分析功能，它可以实现灵活便捷的在线数据可视化。同时，Data.gov 还为每个数据集提供了 API，便于外部调用与分析。

由此可见，大数据分析对于教育研究及规律的发现意义重大。上述三个案例——预测分析报告项目、数据商店和 Data.gov，对于我们利用大数据研究教育现象、发现与探索教育规律，有着重要的借鉴意义。因为规律的发现，需要针对特定研究主题进行长期研究，需要大量标准化数据支持，需要对基于数据的研究与协作作广泛支持。对于重要的教育研究主题，设定数据标准，进行长期纵向跟踪和广泛横向比较研究，是探索教育规律的有效途径。同时，还需要建立更加通畅的数据分享渠道，通过开放，最大化实现数据的应用价值，以便为研究提供更加广泛的支持。

(三) 精准管理支持

在学校和教育机构中，管理者时常面对无法及时掌握教学与管理综合状况的困境。这也导致了教育管理常常是粗放的、由直觉驱动的现状。数据对于学校和教育机构的精准管理和科学决策，可以起到重要的支持、调节作用。基于数据的管理，需要通过汇集各类管理与教学数据，构建多维模型。以 Learnspout、Calarity、Altschool 等为代表的数字化管理应用，体现了数据建模对于精准管理的重要支撑作用，如图 6 所示。

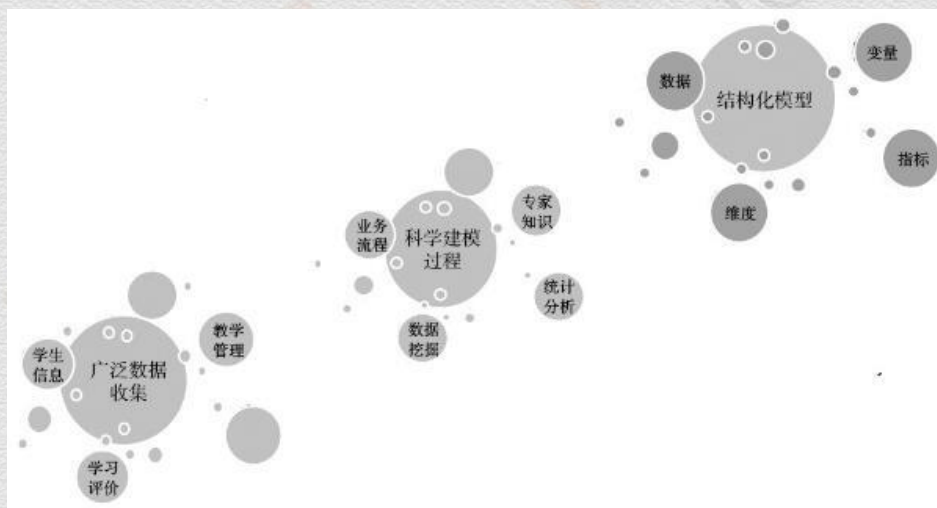


图 6 数据支持下管理模型构建

Learnspout 能够整合学生信息系统和学习管理系统的数据库，对学生日常表现进行监测，通过构建模型对学生综合状况进行分析，并对管理和教学提出建议。例如，Learnspout 能够对高中学生进入大学的准备情况等评价，对存在问题的学生进行早期预警，提出教学干预建议，并评估干预效果。Learnspout 不仅提供了自动化分析，还建立了有教育专家和数据专家构成的团队对数据进行深入挖掘，并形成分析报告。Learnspout 已经在美国 42 个州、200 多个学区的 2500 多所学校中得以广泛应用。该公司于 2016 年 1 月被苹果公司收购，这也从一个侧面说明了市场对其应用效果的认可。

Calarity 是通过构建模型提供管理支持的另一个典型案例。Calarity 旨在评估学校应用信息化技术促进教学的状况，并提出改进建议。它针对学校师生的技术应用情况，采集了 280 个数据点，建立 82 个变量，形成 21 个指标，进而聚合成了 4 个维度，分别为课堂学习、技术应用、信息化技能和信息化环境[25]。在模



型化数据分析的基础上, Calarity 对存在的问题进行了分析和诊断, 并提出了解决方案建议。

Altschool 是学校层次上进行整体性数据采集、分析与应用的代表。这所由前谷歌工程师 Max Ventilla 创办的学校, 将教育教学和工程化思维密切结合, 将提供个性化的教育作为办学宗旨。在教育教学中采用全方位数据采集和分析, 成为了这所学校提供适应性教与学的重要依据和支撑。Altschool 构建了独特的信息化基础设施进行数据采集, 它为教室设计的 Alt Video 系统, 通过各种传感器、摄像头和麦克风采集学生行为数据, 这些数据每天将被分析, 用来改进教学过程和教学系统。Altschool 还开发了 Stream 等教学系统和移动应用, 在支持教学过程的同时, 其基于数据开展教学研究, 并通过研究结果支持教学和管理。不同于传统教育研究, Altschool 的研究周期非常短, 教学设计更新以周为单位, 教学管理调整速度极快。快速迭代的工程化思维、系统化信息技术环境支持、完整的数据采集与分析, 构成了 Altschool 在精准教学管理方面的整体解决方案。

可见, 通过系统化数据采集, 采用科学方法建立的数据模型, 可以帮助教育管理者对学校和机构的管理状况进行持续动态监控和综合性评价。在数据支持下, 管理者能够更容易地发现管理和教学问题, 设计可能的解决方案, 并追踪问题解决的成效。

四、教育大数据所面临的挑战

教育大数据领域方兴未艾, 既具有巨大的发展潜力, 又面临着诸多的挑战。我们认为, 这些挑战包括数据标准、数据采集、模型构建、产品服务和开放共享和隐私保护等方面, 具体如图 7 所示。



图 7 教育大数据应用所面临的挑战

(一) 数据标准有待完善

大数据分析需要多来源、多类型数据的汇集, 数据汇集需要建立统一标准与规范。教育部于 2012 年发布了《教育管理信息教育管理基础代码》等七个教育信息化行业标准, 对教育管理、行政管理、教育统计、中小学、中职学校和高等学校管理的信息进行了规范。这一规范, 虽然对统一教育管理信息有着重要意义。然而, 教学环境、教学过程相关数据标准尚存大量空白。同时, 近年来在线教育



蓬勃发展，其数据标准也尚未建立。上述标准的缺失，成为了教育大数据有效应用的瓶颈。

（二）数据采集覆盖面窄

“十二五”期间，我国教育管理公共服务平台基本建成，教育管理数据的收集具备了较好的条件。但其他教育教学数据，特别是教学过程数据的采集，尚存在较大不足。现有在线学习平台在设计上，往往并未考虑数据分析的需要，对教与学过程的记录不够完整，对数据的分析应用造成了困难。近年来，各高校和部分中小学开始进行智慧校园建设，传感器、Wifi网络、移动设备等技术手段，为更全面的数据采集提供了一些条件。但在现有智慧校园建设中，对数据应用尚缺乏整体设计，对数据采集的支持还远远不够。

（三）模型构建专业性不足

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。从教学的适应性到管理决策支持，都需要通过科学的模型来支撑。但在现阶段模型构建过程中，教育大数据模型构建的专业化水平明显不足。一方面，教育研究成果没有得到很好地应用，教育领域专家知识应用明显不足；另一方面，前沿数据分析方法与技术的应用不足，未能有效借助通用大数据技术的力量。

（四）产品服务单一

在我国教育领域中，以网龙、科大讯飞、猿题库、优答、一起作业、学堂在线为代表的企业，都开始对教育数据的分析与应用展开探索。但总体而言，目前，我国教育大数据相关应用主要聚焦于适应性教学，题库类产品居多，缺乏管理类的应用，对于教学决策的支持不足。同时，教育数据分析应用在功能上较为单一，所采用的分析方法也有一定局限，统计分析仍占有很大的比重，缺乏高水平产品与服务。

（五）开放共享尚未形成

教育大数据的重要价值，首先来自于其数据的大规模和全面性，规模的形成需要广泛的数据共享与开放。当前，行之有效的数据共享开放和应用规则尚未建立，所以，还需要建立有效的资源共享机制，通过多种途径汇聚教学，研究和管理数据，扩大数据的规模，才能形成教育大数据的独特优势。

（六）隐私保护有待完善

教育大数据涉及庞大规模的受教育者与教育者群体。对于这些人群，特别是对于大量的未成年学生而言，隐私保护至关重要。所以，应当从法律上明确、规范公开数据与私有数据的边界，有效的保护隐私数据。在来源清晰、责权明确、应用有序的前提下，才能有效地开展教育大数据研究与应用。

我国教育大数据领域正处于起步阶段。教育大数据研究与应用具有鲜明的特点，其发展需要将大数据技术与教育领域进行深度融合。在“互联网+”时代，为了更好地应对教育大数据所面临的一系列挑战，我们认为，目前迫切需要在体制与机制上，多方协同，各尽其力，以形成一种合力，如图8所示。

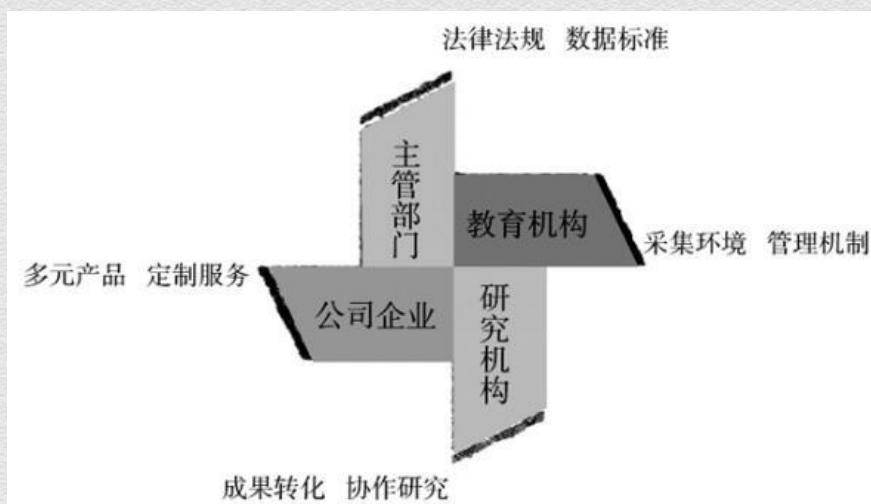


图8 教育大数据发展需要多方合力

具体来说,教育主管部门需要推动教育大数据方面的相关法律、法规的制定,划定边界,明晰责权,建立更加全面的教育数据标准,为国家层面的大规模数据共享和分析奠定基础。以学校为代表的教育机构,需要提升数据驱动教学与管理的意识,构建综合数据采集环境,并建立数据管理与应用机制。一些研究机构需要将教育科学与数据科学紧密结合,开展多学科协同研究,并注重研究成果的转化。公司、企业则需要从教学与管理的整体流程出发,设计与开发多元化产品,并根据教育需求提供灵活可扩展的定制化服务,从而共同促进教育大数据的健康发展,更好地服务于教育事业。

参考文献:

[1]Turck M. Is Big Data Still a Thing?(The 2016 Big Data Landscape) [EB/OL]. [2016-02-01]. <http://mattturck.com/2016/02/01/big-data-landscape/>.

[2]程学旗,靳小龙,王元卓,等. 大数据系统和分析技术综述[J]. 软件学报, 2014, (9): 1889-1908.

[3]与 Hadoop 对比,如何看待 Spark 技术?[EB/OL]. [2016-04-06]. <https://www.zhihu.com/question/26568496>.

[4]O'Brien J. Critical Factors for Data Lake Success[EB/OL]. [2015-09-01].

<http://www.teradatamagazine.com/v15n03/Tech2Tech/Critical-Factors-for-Data-Lake-Success/>.

[5]Cisco. Fog Computing and the Internet of Things:Extend the Cloud to Where the Things Are[EB/OL]. [2015-04-15].

http://www.cisco.com/c/dam/en_us/solutions/trends/iot/docs/computing-overview.pdf.

[6]Norton S. CIO Explainer:What Is Blockchain? The Wall Street Journal[EB/OL]. [2016-02-02].



<http://blogs.wsj.com/cio/2016/02/02/cio-explainer-what-is-hlockchain/>

[7]周明耀. Spark Streaming 指南 [EB/OL]. [2015-08-03]. <http://www.ibm.com/developerworks/cn/opensource/os-cn-spark-streaming/index.html>.

[8]Li F. Convolutional Neural Networks for Visual Recognition [EB/OL]. [2016-01-20]. <http://cs231n.stanford.edu>.

[9]刘红, 胡新和. 数据革命: 从数到大数据的历史考察 [J]. 自然辩证法通讯, 2013, (6): 33-39, 125-126.

[10]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.

[11]喻丰, 彭凯平, 郑先隽. 大数据背景下的心理学: 中国心理学的学科体系重构及特征 [J]. 科学通报, 2015, (5/6): 520-533.

[12]范长煜, 朱艳婷, 高雅静. 大数据在社会科学中的价值: 以 DMSP/OLS 夜间灯光数据为例 [J]. 华东理工大学学报: 社会科学版, 2016, 31(1): 1-9.

[13]SAP. Roambi Analytic Understand Your Numbers [EB/OL]. [2016-06-29]. <https://roambi.com/>.

[14]Tableau. 5 Steps to Self-Service Analytics that Scales [EB/OL]. [2016-06-29]. <http://www.tableau.com/>.

[15]国务院. 国务院关于印发促进大数据发展行动纲要的通知 [EB/OL]. [2015-09-05]. http://www.gov.cn/zhengce/content/2015-09/05/content_10137.htm.

[16]Big Data-Wikipedia [EB/OL]. [2016-06-29]. https://en.wikipedia.org/wiki/Big_data.

[17]McKinsey Global Institute. Big Data: The Next Frontier for Innovation, Competition, and Productivity [EB/OL]. [2016-06-15].

<http://www.mckinsey.com/business-functions/business-technology/our-insights/big-data-the-next-frontier-for-innovation>.

[18]Gartner. Big Data [EB/OL]. [2016-06-29]. <http://www.gartner.com/it-glossary/big-data/>.

[19]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.

[20]胡弼成, 王祖霖. “大数据”对教育的作用、挑战及教育变革趋势——大数据时代教育变革的最新研究进展综述 [J]. 现代大学教育, 2015, (4): 98-104.

[21]杨现民, 王榴卉, 唐斯斯. 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, 69.

[22]Edshelf. Masteryconnect Review [EB/OL]. [2016-06-29]. <https://edshelf.com/tool/masteryconnect/>.

[23]PAR. Predictive Analytics Reporting Framework [EB/OL]. [2016-06-29]. <https://public.datacookbook.com/public/institutions/par>.



[24]Kiesinger. Baker R. ,Cunningham K. ,et al. A Data Repository for the Learning Science Community:The PSLC DataShop[J]. Handbook of Educational Data Mining, 2010.

北京
继续教育
周刊
CONTINUING EDUCATION WEEKLY