

教育虚拟社区:教育大数据的必然回归

张洪孟 胡凡刚

(曲阜师范大学 传媒学院,山东日照 276826)

[摘要] 当前,基于教育大数据的实时决策、数据可视化和个性化学习服务受到人们的热捧,但很少有人从伦理规范和质量保证等视角对基于教育大数据的研究和应用进行反思。本文首先厘定了教育大数据的内涵、特征,然后分析了教育大数据与教育虚拟社区的契合以及教育大数据当前研究和应用中在技术器物、运维机制、使用主体三个层面所凸显出的问题,认为以学习者为中心、具有生态交往文化特性和一定伦理规范的教育虚拟社区可以规避教育大数据当前所面临的问题。基于上述论述,文章提出教育大数据回归教育虚拟社区将是必然,并阐述了实现教育大数据向教育虚拟社区回归的思路或方法。

[关键词] 教育大数据;教育虚拟社区;联通主义;自组织;超循环

[中图分类号] G434

[文献标识码] A

[文章编号] 1007-2179(2015)01-0044-09

近两年来,教育大数据的研究和应用得到人们的广泛关注。在线教育机构纷纷以教育大数据标榜自己,动辄以“教育大数据可视化”“个性化学习服务”宣传自己。教育大数据的作用被绝对化,教育大数据的宣传、研究和应用在一定程度上被异化和神化,一些机构疯狂涌入教育大数据领域。然而,很少有人对其深层问题进行探讨。比如,教育大数据生成的原动力是什么?其发展有无规律可循?研究和应用中伴有什么样的伦理问题?教育大数据的终极目的是什么?本文在对教育大数据内涵进行深刻解读的基础上,剖析了当前教育大数据在发展和应用中存在的问题,阐述了教育大数据回归教育虚拟社区的必然,以及如何实现教育大数据向教育虚拟社区的回归。

一、何为教育大数据

学生学习过程中会产生大量的学习数据,尤其是云计算、移动互联网等技术和泛在学习、移动学习等普及后,教育大数据自然而然成为教育领域的新生事物。教育事业发展的内在逻辑性,教育现象、内容和规律之间的级联性、高相关性被要求运用大数

据的思想和方法进行深度的分析和挖掘。

(一)对大数据的认知

已有大数据的定义大多从其特征出发进行描述性定义(Laney,2012;Barwick,2011;IBM;Department of Education Office of Educational Technology,2010)。大数据产生于网络生产过程中,伴随物联网、社交网络、云计算等网络服务的发展,传统的数据分析方法和分析工具很难再应对网络生产过程中的海量数据。大数据的产生主要来源于两点——在线和交互。“大数据”在内涵和外延上已超越传统意义上的“数据”概念。其实,“大数据”一词并不能准确概括当今大数据研究领域的全部。大数据的重点并不在于“大”,而在于其蕴涵的多元、有序、共享、生成等思想,在于人们在技术哲学指导下应对大数据挑战过程中所生成的技术、方法和思维。基于大数据的思维和技术在实际应用中应是稳定、可重复、可解释、可实施的。

(二)教育大数据的涵义

国内关于教育大数据的定义最早见于徐鹏的《大数据视角分析学习变革》一文。他从产生教育大数据的主体出发,将教育大数据区分为广狭二义:

[收稿日期] 2014-12-11

[修回日期] 2014-12-31

[DOI编码] 10.13966/j.cnki.kfjyyj.2015.01.005

[基金项目] 国家社会科学基金教育学一般项目“教育虚拟社区伦理的作用机制及评价研究”(BEA130026)。

[作者简介] 张洪孟,硕士研究生,曲阜师范大学传媒学院,研究方向:网络教育应用;胡凡刚,博士后,教授,硕士生导师,曲阜师范大学传媒学院(通讯作者),研究方向:教育虚拟社区理论与实践、网络教育应用、影视文化传播(hufangang5@163.com)。

广义的教育大数据泛指来源于日常教育活动中人类的行为数据;狭义的教育大数据指的是学习者的行为数据(徐鹏等,2013)。章怡等从教育大数据如何产生的角度将其界定为:信息技术支持下教与学各个环节所产生的符合教育大数据特征的数据集,是信息技术环境下教与学行为轨迹的产物(章怡等,2014)。2013年10月,美国哥伦比亚大学瑞安·贝克(Ryan Baker)在Coursera平台开设“教育中的大数据”(Big Data in Education)课程(Coursera,2013),向学习者介绍常用的数据挖掘方法,以及如何将这些方法应用到教育数据挖掘和学习分析中。美国Knewton教育公司将教育大数据分为两类:一类是有关学生基本信息的数据(例如,身份识别数据);另一类是基于学生学习活动用以提升学习效果的数据,包括学习交互数据、推断的内容数据、系统范围数据、推断的学生数据(Ferreira,2013)。根据已有研究并结合教育虚拟社区实践,本文将教育大数据界定为:学习者以学习为目的,在互联网学习环境中借助一定媒介,与外界交互过程中所产生的数据,主要包含学习行为数据、学习内容数据、虚拟社会网络关系数据以及学习管理数据。

——学习行为数据即记录学习者网上学习行为的数据。例如,学习者访问网站的分布数据,包括访问时间分布数据、停留时间、回访次数、回访相隔天数等。对学习者在在线学习行为数据的分析和研究有助于了解学习者的学习类型、学习风格以及所需学习服务类型等。

——学习内容数据指学习者在网络学习过程中形成的关于某一问题的观点、意见,这种数据多以文本、图像、声音和视频等形式沉淀下来。对学习内容的挖掘和分析,有助于描绘学习者的知识图谱,明确学习者的知识掌握情况。

——虚拟社会网络关系数据主要记录学习者在各虚拟学习社群中的交往属性和交往动态,以及在社群中的位置和作用。掌握学习者的虚拟社会网络关系数据有助于了解知识在虚拟社区中如何传递,进而更好地促进知识在虚拟社区中的传播。

——学习管理数据即对学生学习进行管理和评价中所记录、生成的数据。学习管理数据的管理和应用水平体现了一个组织或机构学习管理的能力。

教育大数据与其他行业领域(医疗、交通、气

象、金融等)大数据相比,其特点为教育性、交互性、生成性。教育大数据是以“教学”和“学习”为中心所产生的数据,主要生成于以“学习”为中心的在线交互过程中,且在形成和发展过程中呈现一定的自组织性。教育大数据的重点并不在于其存储、生成和计算上的“大”,更重要的是,它体现了一种方法论和技术论上的变革,一种对于教育和学习问题研究的新思路。

二、教育大数据回归教育虚拟社区的必然

当前,教育大数据的研究和发展面临一系列难题。世界银行前资深教育专家埃米利奥·门(Emilio Porta)认为,当前真正有用的教育大数据仅来自于数据网络中的关键节点,且这些教育大数据在制定国家层面教育政策时没有发挥很大的作用(Porta,2013)。教育大数据在技术器物、运维机制和应用主体方面存在的问题,使教育大数据在学习分析和个性化学习建构方面还未真正发挥作用。通过对比分析我们发现,教育大数据与教育虚拟社区在文化内涵、人本主义追求、联通虚拟教育与现实教育等方面存在一定的契合。而且,教育虚拟社区在教育、人文和社会价值上的功能和追求,很好地规避了教育大数据的缺点。结合当前的教育实践,我们认为,教育大数据与教育虚拟社区相结合,即教育大数据发展和应用的社区化将成为必然。

(一)教育大数据和教育虚拟社区的契合

教育虚拟社区基于跨时空、开放、自由的网络虚拟环境,社区成员之间进行专题研修、交互协作、资源共享,从而相互影响、相互促进,最终形成具有共同社区文化心理、生态式的社会关系共同体(胡凡刚,2005)。开放、自由、平等、共享等是教育虚拟社区的显著特征。《美国2010国家教育技术计划》强调虚拟社区和教育大数据在“技术支撑的学习模型”和“基于联结的教学来提升教师教学能力和专业知识模型”中的重要作用(The Enterprise Strategy Group,2010)。徐鹏在探讨美国教育大数据对我国教育大数据研究和应用的启示时指出,教育大数据为教育虚拟社区的建设和运维提供了机遇和挑战。教育大数据和教育虚拟社区都是在网络技术背景下,促进在线教育发展过程中涌现的。因此,目标的一致性和教育文化的相似性为两者的契合奠定了

基础。

1. 教育大数据和教育虚拟社区蕴含相同的文化内涵。教育虚拟社区和教育大数据本质上皆呈现开放、自由、平等、多元、自主生成的文化特性。教育虚拟社区是学习者在平等、自由、自主的文化氛围下自发聚集而成的。教育虚拟社区生态交往的文化特性是从线性走向网络,从依附走向平等,从他塑走向自塑(胡凡刚,2011)。在教育虚拟社区交往过程中,学生不用顾忌课堂教学中的等级、优劣、专家和教授的权威。来自优等生的优越感在教育虚拟社区完全被颠覆。学习者相互间平等交流和表达观点,不存在学习上的歧视和不平等现象。教育虚拟社区开放、自由、平等、多元、自主生成的特性能激发学习者表达自身学习需求、开展社区交往的愿望。

教育大数据突破了传统数据绝对集权的观点。每位学习者都可借助互联网和媒体设备生成学习数据,都是教育数据的中心。数据在网络世界没有高低贵贱之分,都借助数字信号在网络中传输。每位学习者都可表达自己真实的想法,各人在学习过程中生成的数据也各不同。同时,随着学习的不断深入,新数据又会在已有数据上不断生成。这些数据的不断积累,最终实现对学习者的准确“画像”。

2. 教育大数据和教育虚拟社区具有共同的人本主义追求。以往对学习者的学习数据的处理多为记录和储存,且数据多是静态、孤立的结构化数据。各教育部门、教育机构、网络学习社区之间的数据相互割裂,得不到有效联结,在数据库建设、应用上存在重复性。基于云计算的教育大数据,实现了以学习者为中心的学习数据的共建、共享,实现了不同系统间的耦合,即多种学习情境中产生的数据可有效联结。例如,由海峡两岸五所交通大学2013年8月发起并建立的ewant育网开放教育平台,创新性地实施微学程学习模式。微学程是一组精心设计、有系统关系的课程,每个微学程包含三门以上的课程,课程间有连贯性、逻辑性,每门微学程课程会在1-2年间陆续开设(ewant,2013)。这种基于教育大数据、跨地域的微学程模式使学习者对某一领域知识的学习历程更加系统、完整,充分体现了以人为本的服务理念。

教育虚拟社区的形成建立在交互的基础上,这种交互存在于学生与教师、学生与学生、学生与内

容、内容与内容之间,学习者在其中共同协作、相互学习、完成知识建构,并体验学习的归属感、成就感和存在感。在交往日益文明的教育虚拟社区中,学习者自我价值的实现将会成为最大、最现实的需求。在这种自我需求的驱动力下,学习者会积极发出自己的声音、共享自己的观点,进而增强其促进社区智慧和个体智慧的发展。

3. 教育大数据和教育虚拟社区——虚拟教育与现实教育有效联通的路径。联通主义学习观强调知识分布在网络中,这一网络有三层含义:个体的知识网络、小规模的学习社群网络(包括现实和虚拟两种形态)、通过社交媒体和网络等将各种学习社群连接起来形成的跨社群虚拟学习网络。联通主义学习观认为知识蕴含在这些网络的节点以及节点之间的联结中,并通过节点间的不断交互,在网络中不断生成新的知识,促进各个层次的网络在组织、内容和结构上更加有序、自组织。学习的关键是形成网络内外的有效联结,在不同层次的有效联结中丰富和完善个体知识结构(George,2005)。这种联结和交互通过不同层次的网络间和网络内的数据和信息的流动而实现。

学习者在虚拟教育和现实教育中的学习行为,展现的是学习者在不同空间的不同学习需求,都是学习者心理活动外化的表现。基于教育大数据的教育虚拟社区跨越时空阻隔将虚拟教育与现实教育有效联通,消除了二者的界限。这种联通的根本驱动力是学习者学习和交往的需求,联通的形式就是教育虚拟社区中数据和信息的流动。

(二)教育大数据存在的问题

海德格尔在评价技术的消极影响时指出,单纯的技术并不可怕,可怕的是受技术影响而产生的技术思维(许良,2005a)。当前,对教育大数据的研究和应用多停留在希冀的美好愿景上——如果教育大数据按照如何如何的顶层设计,那么将会产生如何如何的效果。这就暴露出人们对教育大数据的认识、研究和应用还存在一定问题。虽然教育大数据能够记录学习者的学习过程,但它不能解释这一过程为什么是这样的,而且学习过程还存在数据无法表征的因素,例如学习动机和情感因素等(祝智庭等,2013)问题。基于大数据的机器学习还无法构建类似人脑这样的智慧系统,主要技术瓶颈在于机

器无法像人一样去感知周围事物(Jordan, 2014; 赵勇, 2014)。此外,教育大数据还存在如何在价值观、性格塑造和公民参与等方面发挥作用(Tan, 2013),如何保卫学习者的自由意志、道德意志,如何保卫学习者的自主选择权、学习的能动性(Cukier, 2014)等问题。

德国哲学家 E·卡西尔(Emst Cassirer)等人主张把技术看作是文化的组成部分,技术和文化在交互过程中相互影响,形成技术文化。对技术文化分析研究可从三个层面进行,从外到内依次是:技术器物层、技术制度或技术体制层、技术意识形态层(许良, 2005b)。依据教育大数据的特性、当前研究和应用现状,借鉴卡西尔等人的思想,教育大数据凸显的问题可以划归为三个层面:技术器物层面、运维机制层面、使用主体层面。

1. 技术器物层面

1)教育大数据对学习者的数据化解构与重构。教育大数据的技术思维取向认为,教育大数据产生的过程即对学习者的数据化解构与重构的过程,每位学习者都是一个数据源,都可以逐步加以解构,直到每一部分、每一属性都以数据化形式表示,并可以一定的编码进行映像存储。然而,这种映像与学习者之间不能划等号,这种数据化解构与重构,忽视了学习者的自我意识、自我能动性、信仰和道德情感。学习者不再被视为有思维和能动意识的人,而被替代为一串串的符号记录。学习者之间的关系变成了干巴巴的数据关系,失去了之所以为学习者的根本属性。此外,学习者不仅有知识的需求,还有社会交往、情感和价值观方面的需求。学习者这些需求的程度,以及这些需求在学习过程中的作用机制也无法通过数据衡量。互联网进化论的提出者刘锋认为,互联网的发展和最终应用将以人为本,而不是以数据互联为中心(刘锋, 2009)。学习的思维过程是不可还原、不可重复的,每位学习者思维加工的经验 and 知识基础各有不同,不能机械地将已有学习者的成功经验移植到其他学习者身上。

2)教育大数据研究和应用过程中的单向性。教育大数据的真正力量在于将基于数据的形成性反馈准确地指向学习者(Clark, 2013)。当前,教育大数据的存储和聚集比较常见,且容易实现。然而,真正基于教育大数据的教育决策和个性化学习建构却

很少。教育数据只是单向地流向数据中心,很少形成基于教育大数据的意义反馈,学习者与教育大数据之间缺乏深层交互。数据被无意义地收集和存储,潜在意义得不到有效的开发和利用。即使存在反馈,那也是低层次的,很少涉及智慧层面。

3)教育大数据的产生多以网络终端设备为中心,并非以问题为中心。在处理学习者的某一问题时,研究者需要跨越多设备聚合数据,然后再对数据进行分析和处理。因此,数据准备阶段耗费时间和资源较多。

2. 运维机制层面

1)在线教育产业链上企业或机构间的利益博弈阻碍了教育大数据的联通。学校、教育培训机构等拥有教育大数据却没有处理数据的能力,企业拥有处理、分析教育大数据的能力,却得不到教育大数据。在线教育机构关于教育大数据的研究和应用未能实现良好的沟通和共享,多为竞争关系。目前,国内各大网络公司纷纷依托自身优势打造垂直在线教育,最令人关注的是 BAT(百度、阿里巴巴、腾讯)的动向(见表一)。三家互联网产业公司,分别依托自身数据、流量和产业优势,发展在线教育,相互间缺乏数据的沟通和交互。这种博弈导致数据断联并形成一个个“数据孤岛”,阻碍了他们通过相互合作达到整体的最大利益。近期,“学堂在线”正与阿里巴巴磋商,合作一旦成功,天猫盒子用户即可通过电视学习“学堂在线”提供的高清 MOOC 课程。这种将传统教育电视与新型 MOOC 在线教育相结合的创新模式能实现将互联网运营商、教育大数据、教育电视和 MOOC 等多种技术手段和优势资源的结合(辛巴, 2014)。

2)教育大数据算法限窄学习者信息来源。维克多·迈耶(Viktor Mayer-Schönberger)和肯尼斯·库克耶(Kenneth Cukier)反思教育大数据负面影响时指出,基于教育大数据的课程推荐一定程度限窄了学生的选课范围,挫伤了学生向困难挑战的积极性,使学生转而选学容易通过的课程(Mayer-Schönberger et al., 2014)。埃里帕泽(Eli Pariser)研究发现,基于大数据的媒体和网络过滤机制一定程度上阻碍了学习者信息视野的拓展(Pariser, 2011)。在教育虚拟社区中,相同或相反观点的学习者自发地聚集网络,形成强连接。然而,真正给社群带来新

表一 BAT 基于云计算的在线教育对比

公司	云平台	数据资源	技术优势	研发机构	教育社区	是否收费
百度	百度开放云	用户检索数据 公共 WEB 数据	数据聚合 语义分析 深度学习	深度学习研究院 百度数据研究中心	百度学堂 百度教育	免费 +有偿
阿里巴巴	阿里云 ("飞天"计划)	用户在线交易数据 部分社交数据和移动数据	底层系统 并发处理	阿里研究院 阿里集团数据委员会	淘宝同学	有偿
腾讯	腾讯开放平台 腾讯云平台	用户社交数据 关系数据	分布式数据存储 社会网络分析	腾讯数据平台部 腾讯数据中心	腾讯课堂 腾讯精品课 Q 群教育	免费 +有偿

观点、新知识的,往往是弱连接上的某些节点。刘锋认为忽略互联网上人脑的参与,是当前互联网发展面临的主要困境。基于教育大数据的算法和人工智能能否知道学习者的真实想法,提供的学习服务能否真正满足学习者的需求,有待进一步商榷。

3. 使用主体层面

1)教育大数据应用过程中的伦理失范。首先,教育大数据研究和应用中的隐私问题。受家长、数据隐私保护者和教育工作者对学术数据隐私的质疑,2014年4月美国知名教育大数据新兴公司 inBloom 宣布停止运营。inBloom 公司的宗旨是将经过存储、清洗和聚合分析后的教育大数据提供给政府批准的第三方机构开发教育产品,以更好地促进课堂教学。虽然 inBloom 公司的初衷值得称赞,但在具体操作中将学生的敏感数据提供给利益集团用作数据挖掘,没有充分保障学生的隐私权(Heröld,2014)。其次,教育大数据所有权应当归谁?是学生、教师,还是教育机构?最后,教育大数据研究和应用到底是公益还是商业?教育大数据的应用和研究迄今还未形成一套切实可行的伦理规范。教育大数据的伦理失范主要表现在对学生和教育相关从业人员的隐私侵犯,在实际研究和应用中目的与手段的不统一、研究方法的不规范等。

2)教育大数据的研究和应用缺乏一定的质量保证体系。随着大数据的“蹿红”,一些利益组织或个人在不具备条件的情况下疯狂涌入教育大数据领域,各大教育公司或组织摇身一变成为基于教育大数据的教育产业公司。然而,在线教育盈利周期较长,许多教育机构受时间制约和利益诱惑,急于从数据的浅层分析和表象中得出结论,并将其应用于教育服务。这种现象折射出的是教育大数据研究和应用质量标准鉴定体系的不完善,由此建议应建立一套相对成熟的教育大数据应用、评价体系,对各行

业、领域的教育大数据研究和应用的质量进行评定,以此提高涉足教育大数据的门槛。

3)专业教育大数据分析师的缺乏。当前制约教育大数据研究和应用发展的关键因素并非技术,而是人才——教育大数据分析师。教育大数据分析师应是兼具教育学、心理学、社会学等多学科知识,同时具备大数据思维,能从学生角度发现问题、思考问题和解决问题的复合型人才。美国求职网站 Career Cast 发布的 2014 年全美十大最佳职业中,与大数据相关的职业分别排第一、三和八位(Perman, 2014)。如今,电子商务、天文研究、社交网络分析、医疗健康、气象研究等领域已开始招聘大数据分析研究人员。微软全国广播公司 2010 年就开始招聘精通大数据相关知识的记者。BBC 新闻在线在早期项目研究中就已在由多领域专家合作的项目中培养具备多领域知识、精通大数据的记者。

不可否认,教育大数据的应用会对教育的发展产生重大影响,但我们不希望把一堆有关学习的数据扔给机器,认为有关学习的问题都可借此得以解决。这种思考问题的方式似乎改变了问题的本质。教育是关于人的复杂性系统,教育大数据仅仅是整个系统的一环。

三、教育大数据回归教育虚拟社区理路

以学习者为中心、具有一定生态交往文化特性和价值诉求的教育虚拟社区规避了教育大数据的缺点,为教育大数据的研究和应用提供了指导。结合当前教育实践,我们认为,教育大数据的发展和应用与教育虚拟社区相结合,即教育大数据的社区化将成为发展的必然。

(一)自组织教育虚拟社区避免了教育大数据的无意义增长

戴维·阿斯平(Aspin, D. N.)从教育哲学的高

度指出,由于人是具有自主性的个体,人类作为一个共同体具有相同的福祉和道德追求,因此在教育现象中必然会呈现自治性和群体性(戴维·阿斯平等,2013)。基于自组织的教育虚拟社区以超循环理论为指导,无论是学习个体,还是整个社群都具有认知意识。基于自组织的教育虚拟社区的发展能够避免教育大数据自发、无意义的聚集,引导教育大数据向合理、有序的方向生成。从整体生成论的角度看基于自组织的教育虚拟社区主要经过自选择、自组织、自稳定三个阶段(见图1)。个体水平的循环过程中,学习者借助已有认知结构参与学习交往,伴随以同化和顺应,促进学习者个体认知结构的不断完善。符合社区认知规律、发展需求和具有共同情感追求的学习者的社区交往过程被保存下来,成为社区发展的关键节点,同时这些关键节点在社区内交往过程中改变着社区的结构和功能,不断促进社区水平的认知循环。新认知水平上的教育虚拟社区作为一个单元进入更高层次(相关社区耦合为更大的教育虚拟社区)的进化循环,形成社区间的影响。这样,底层循环催化、耦合为高层循环,同时高层循环对底层循环有负反馈作用,通过自选择和自组织使整个生成系统趋于稳定,形成超循环形态的教育虚拟社区。

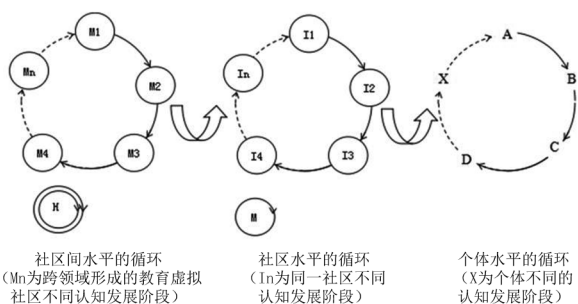


图1 自组织教育虚拟社区的认知发展

在这种超循环机制中,不同层次间能够耦合的基础有两点:一是学习者学习交往的本性;二是学习者追求自我归属感、自我发展的内驱力。耦合的具体表现是数据承载的信息在不同层次间交换和传递。自组织教育虚拟社区将教育大数据和学习者统筹考虑,同一问题的关联数据自发聚集并被收集,增强了数据之间的互通性、关联性,避免了教育大数据的无意义增长和教育大数据研究和应用中的单向性。学习者之间的关系不再是干巴巴的数据关联,而充满了人情味。

(二)构建基于教育大数据的教育虚拟社区学习者个性化学习服务

教育大数据的真正力量在于将基于数据的形成性反馈准确地指向学习者。共享内容对象参考模型 SCORM(Share Content Object Reference Model)注重教育资源的可重用性、可获得性、耐久性、互通性。然而,其思考问题的出发点是如何更好地促进教师的“教”。学习管理系统强调的是对学生学习事务的在线管理和处理、信息拥有者对信息的综合处理和应用,但其信息处理的对象多为某个学习群体或个体,很少研究同一学习者的跨群体表现。在教育大数据时代,在线学习正在从以内容为基础的弱互动向以关系为基础的强互动转变。基于教育大数据的教育虚拟社区学习者个性化学习服务将与学习者的多种教育云相整合,通过对教育云中教育大数据的分析,识别学习者,进而为学习者提供个性化学习服务(见图2)。

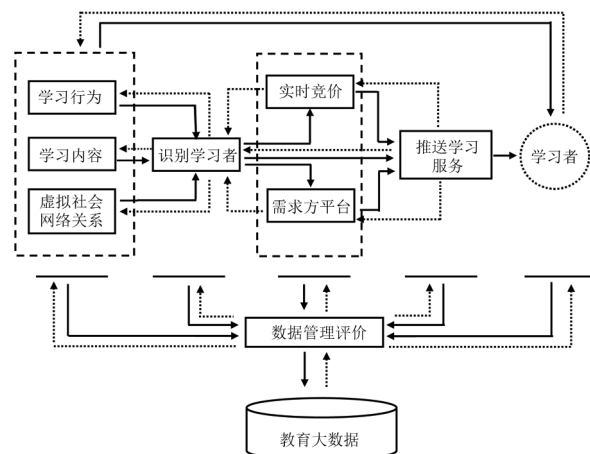


图2 基于教育大数据的教育虚拟社区学习者个性化学习服务

基于教育大数据的教育虚拟社区学习者个性化学习服务构建大体分为以下三个阶段:识别学习者、确立最优个性化学习服务、学习服务推送。首先,通过对学习行为数据、学习内容数据和虚拟网络社区关系数据的分析,结合学习者已有的管理评价数据,识别学习者,实现学习者基本属性和学习需求的准确定位;其次,学习服务提供方依据学习需求制定学习服务,第三方评价机构依据学习服务标准在相互“竞价”的学习服务中选择最优学习服务;最后,教育虚拟社区选择合理、有效、人性化的方式向学习者推送学习服务。

基于教育大数据的教育虚拟社区学习者个性化

学习服务的构建中,生成的每一步数据都是对上一步的反馈,同时这些数据也会以多种形式加以管理和评价,并存储于数据库中。基于教育大数据的教育虚拟社区学习者个性化学习服务,是一种遵从学习者学习需求,跨媒体、跨社群、跨终端、多情境、最大限度地整合与学习者相关教育大数据的学习服务,能实现教育大数据产业链上合作博弈的最优。

(三)以教育虚拟社区伦理建设规范教育大数据应用

康德说“你的行为要始终把人的本质作为目的,而绝不要仅仅作为手段对待,无论这本质是你自己,还是他人”(Kant, 1964)。也就是说,人不是通往任何目标的工具,而是目标本身。在线教育组织或机构将学习者作为教育大数据采集和实验的对象,并将获取的数据用以商业目的,在某种意义上,这种做法会导致意外的“目的的善”,但学习者在这个过程中对自己的数据和信息被利用是否知情,是否允许,又是否自愿?教育大数据的社区化有助于规范教育大数据的研究和应用,通过教育虚拟社区伦理的带动和示范,形成教育大数据研究和应用的伦理规范。

首先,保障教育虚拟社区学习者的信息知情权。在广播媒体时代,信息源(例如,媒体和报纸刊物的编辑等)根据一定的职业伦理和价值判断有甄别地向用户传递信息。但在自媒体时代,人人都是信息源,信息甄别被计算机算法所取代,然而这种机器算法是否考虑了人类思维的伦理道德、价值判断和情感因素?因此,教育虚拟社区应赋予学习者一定的信息甄别权、知情权和自主权,也就是学习者应知道教育虚拟社区通过什么样的机制使用户获取了哪些信息、丢失了哪些信息,学习者可以凭借自身需求定制、取消各类信息服务。

其次,保障教育虚拟社区中学习者的个人隐私。教育大数据应用过程中,学习者最为关注的是个人隐私。教育大数据的隐私和安全问题,不仅要诉诸法律、制度和技术,最根本的应加强人们在个人隐私和数据使用上的伦理规范教育。美国“大数据、伦理与社会理事会”(the Council for Big data, Ethics and Society)定期召集各领域的专家研究大数据的安全、隐私、平等和使用问题,并通过发布评论、白皮书,开发大数据标准等使大数据研究人员和公众及

时了解与大数据相关的社会、伦理、法律和政策问题(Boyd et al. 2014)。

最后,建立健全教育大数据应用的相关法律法规,制定教育大数据研究和应用的相关行业标准。某组织获取学习者的数据后,为何其他组织使用这些数据需要向其付费,即使这种付费是合理的,付费标准又如何制定?这些基于教育大数据的交易原则的合理性、可行性以及交易中的具体问题,应通过相关法律和行业标准规范教育大数据的操作流程,提高教育大数据服务领域的准入机制。

四、结语

教育大数据回归教育虚拟社区的过程应符合伦理规范,具有一定质量保证体系,以提升学习服务为目的。教育大数据回归教育虚拟社区不是把教育大数据的特性加以刨除,更不是两种技术或思想的简单叠加,而是以教育虚拟社区的特性规避教育大数据在思想、方法和实际应用中的弊端,回归意味着教育大数据在教育虚拟社区中是一种内在的存在和超越。教育大数据向教育虚拟社区的回归,是一种价值引领,其实质是一定意义上向学习者的回归,是对学习者的自我意识、自我价值和自我追求的肯定和彰显,目的是把学习者应有的权利归还学习者。

未来研究中,我们将会继续研究如何实现教育大数据在符合伦理规范要求下向教育虚拟社区的回归,如何使教育大数据的研究和应用更加符合教育虚拟社区生态交往的运行机制,从而变革现有教育虚拟社区的学习和交往模式。

[参考文献]

- [1] Barwick, H. (2011). IIIS: The 'four Vs' of Big Data_Storage vice-president says volume, variety, velocity and value key [EB/OL]. Retrieved from http://www.computerworld.com.au/article/396198/iiis_four_vs_big_data/.
- [2] Boyd, D., Bowker, G., Crawford, K., & Nissenbaum, H. (2014). Council for Big Data, Ethics, and Society [EB/OL]. Retrieved from <http://www.datasociety.net/initiatives/council-for-big-data-ethics-and-society/>.
- [3] Coursera (2013). 教育中的大数据 [EB/OL]. Retrieved from <https://www.coursera.org/course/bigdata-edu>.
- [4] Clark, D. (2013). WISE 2013 debate: What data for what purpose? [R]. wise summit. Retrieved from <https://www.youtube.com/watch?v=jkdBOOLjp98>.

- [5] Cukier, K. (2014). Big data is better data [R]. TED. Retrieved from https://www.ted.com/talks/kenneth_cukier_big_data_is_better_data#t-762372.
- [6] 戴维·阿斯平,朱迪斯·查普曼(2013). 终身学习的哲学思考[J]. 开放教育研究,(5):45-62.
- [7] Ewant(2013). 关于育网[EB/OL]. Ewant 育网平台. Retrieved from <http://www.ewant.org/aboutus.php>.
- [8] Ferreira, J. (2013). Big data in education; the 5 types that matter [Blog post] [EB/OL]. Retrieved from <http://www.knewton.com/blog/ceo-jose-ferreira/big-data-in-education/>.
- [9] Herold, B. (2014). inBloom to shut down amid growing data-privacy concerns [EB/OL]. Education Week. Retrieved from http://blogs.edweek.org/edweek/DigitalEducation/2014/04/inbloom_to_shut_down_amid_growing_data_privacy_concerns.html.
- [10] 胡凡刚(2005). 简论教育虚拟社区[J]. 电化教育研究,(9):42-46.
- [11] 胡凡刚(2011). 教育虚拟社区生态交往的理论与实践[M]. 青岛:中国海洋大学出版社:128-131.
- [12] IBM. What is big data? The FOUR V'S of big data [EB/OL]. IBM. Retrieved from <http://www-01.ibm.com/software/data/bigdata/what-is-big-data.html>.
- [13] Jordan, M. (2014). Machine-learning maestro Michael Jordan on the delusions of big data and other huge engineering efforts—Big-data boondoggles and brain-inspired chips are just two of the things we're really getting wrong[EB/OL]. IEEE SPECTRUM. Retrieved from <http://spectrum.ieee.org/robotics/artificial-intelligence/machinelearning-maestro-michael-jordan-on-the-delusions-of-big-data-and-other-huge-engineering-efforts>.
- [14] Kant, I. (1964), Groundwork of the metaphysic of moral [M]. translated H.J. Paton, New York: Harper&Row.
- [15] Laney, D. (2012). 3D Data management: Controlling data volume, velocity, and variety [Blog post] [EB/OL]. Retrieved from <http://blogs.gartner.com/doug-laney/deja-vvvue-others-claiming-gartners-volume-velocity-variety-construct-for-big-data/>.
- [16] 刘锋(2009). 互联网进化论[EB/OL]. Retrieved from <http://www.doc88.com/p-11561856601.html>.
- [17] Mayer-Schönberger, V., & Cukier, K. (2014). How big data will haunt you forever: your high school transcript [EB/OL]. Retrieved from <http://qz.com/185252/how-big-data-will-haunt-you-forever-your-high-school-transcript>. excerpted from Learning with Big Data: The Future of Education[M]. Viktor M. S. & Kenneth C. Houghton Mifflin Harcourt Publishing Company, 2014.
- [18] U. S. Department of Education Office of Educational Technology (2010). National Education Technology Plan: Transforming America education learning powered by technology [EB/OL]. Retrieved from <http://tech.ed.gov/wp-content/uploads/2013/10/netp2010.pdf>.
- [19] Pariser, E. (2011). Beware online “filter bubbles” [R]. TED. Retrieved from http://www.ted.com/talks/eli_pariser_beware_online_filter_bubbles?language=en,2011-03.
- [20] Perman, C. (2014). The best (and worst) jobs for 2014 [EB/OL]. Retrieved from <http://www.cnbc.com/id/101582641>.
- [21] Porta, E. (2013). WISE 2013 debate: what data for what purpose? [R]. Retrieved from <http://www.wise-qatar.org/emilio-porta>.
- [22] Siemens, G. (2005). Connectivism: a learning theory for the digital age [J]. Instructional technology and distance learning, (1): 3-10.
- [23] Tan, J. P-L. (2013). WISE 2013 Debate: What Data for What Purpose? [R]. wise summit. Retrieved from <http://www.wise-qatar.org/jennifer-pei-ling-tan>.
- [24] The Enterprise Strategy Group (2012). The 6 Vs: the BI/Analytics Game Changes so Microsoft Changes Excel [EB/OL]. Retrieved from <http://www.esg-global.com/blogs/the-6-vs-the-bianalytics-game-changes-so-microsoft-changes-excel/>.
- [25] 辛巴(2014). 天猫盒子 x 学堂在线——客厅里的 MOOC? [EB/OL]. Retrieved from <http://www.jmdu.com/news/detail/1007#0-qzone-1-5504-d020d2d2a4e8d1a374a433f596ad1440,2014-8-29>.
- [26] 许良(2005a). 技术哲学[M], 上海: 复旦大学出版社: 170-172.
- [27] 许良(2005b). 技术哲学[M], 上海: 复旦大学出版社: 138.
- [28] 徐鹏,王以宁等(2013). 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,(6):11-17.
- [29] 章怡,牟智佳.(2014). 电子书包中的教育大数据及其应用[J]. 科技与出版,(5):117-120.
- [30] 赵勇(2014). 在驾车上计算机比人安全几十倍[R]. Retrieved from <http://tech.ifeng.com/internet/special/chuanqixingkongjian/>.
- [31] 祝智庭,沈德梅(2013). 基于大数据的教育技术研究新范式[J]. 电化教育研究,(10):5-13.

(编辑:徐辉富)

Big Data and Learning Analytics for Educational Virtual Community

ZHANG Hongmeng & HU Fangang

(School of Communication, Qufu Normal University, Rizhao 276826, China)

Abstract: *In recent years, the education real-time decisions, data visualization and personalized learning services based on the educational big data have received scholars' interest, but few people have reflected on the researches and applications of the educational big data on the perspectives of ethical standards and quality assurance. In order to demonstrate the necessity that the development of the educational big data should be combined with the educational virtual community, this article consists of three parts. Firstly, the paper defines educational big data and describes its features and meaning. It also analyzes the commonalities between the educational big data and the educational virtual community on the perspectives of cultural characters, humanism and the contributions to the connection between online learning and traditional education.*

Secondly, the article reviews the internal and external development of the educational big data and sums up its existing problems in the current. According to Ernst Cassirer, we divide the problems into three aspects—technology, operation mechanism, and the users. In the aspect of technology, the digitally deconstructs and reconstructs for the learner, the collection and storage of the data do not take the students in center and form little constructive feedback for the learners. In the aspect of operation mechanism, the interest game between the educational institutions leads to the “data island” and hampers the data interchange. At the same time, the algorithms based on the educational big data could not perceive the circumstances around the learner and could not understand what the learner is really thinking. In the aspect of the users, it lacks of ethics, the quality assurance system and the educational big data analysts. All three aspects restrict the development of the educational big data.

On the basis of existing researches, we believe that the characters and the development of the educational virtual community could overcome the drawbacks of the educational big data. Thirdly, this paper states how to realize the educational big data regresses to the educational virtual community. We believe that the characteristics of learner centered of the educational virtual community could avoid meaningless growth of the big data; the ecological interaction between teachers and students could overcome the mechanical reduction for the learner; the community-based educational big data could connect the learners and form the constructive feedback for the learners.

So the educational big data should be combined with the educational virtual community. Integrating the educational big data and the virtual community can help highlight the learner's self-awareness, self-value and self-seeking, and value the learners' rights during the education process.

Key words: *educational big data; educational virtual community; connectivism; self-organizing theory; hypercycle theory*