

教育大数据应用:挑战和利用

——兼论教育小数据的大作用

方柏林

(艾比林基督教大学,艾比林,美国)

[摘要] 高等教育机构试图利用庞大、复杂、集成的大数据,更好地为学生服务,但大数据在教育领域应用存在误区,比如对边缘群体更有效,数据不能和课堂教学衔接,教师对数据缺乏兴趣,也缺乏利用数据的相应支持。不过,在数据时代,教育界可以利用一些学习分析数据和相关工具,给老师和学生提供优质反馈。本文介绍了美国一所私立大学使用教师评估和课程管理软件的案例,证明一维的“小数据”和“微数据”也可以有效地帮助教师改进教学和学习过程。本文还介绍了学生可以使用哪些数据、工具、方法,提高学习结果。

[关键词] 小数据;学习分析;教师评估;课程管理系统;学习工具

[中图分类号] G434

[文献标识码] A

[文章编号] 1007-2179(2016)04-0018-07

一、引言

大数据指那种庞大、复杂,需要一系列工具收集、分析、展示的数据。《福布斯》杂志撰稿人普莱斯曾追踪这个概念的来龙去脉:从历史沿革看,大数据的发展是典型的“灰姑娘”故事。它一开始是个贬义词,指数据过于复杂和庞大,对处理它的计算机内存空间、处理时间都提出了挑战。《牛津英语字典》对它的定义也是负面的,宣称大数据“构成了重大的处理上的挑战”(Press, 2014)。后来,政府和商业机构发现,这种多渠道来源的数据,可以帮助他们更好地认识人、问题、地域以及其他社会现象,得出一些平时使用一维数据无从提供的结论。大数据这时在人们头脑中的印象才得以改观。2012年和2014年,涂子沛、舒恩伯格相继出版著作(涂子沛, 2012; Mayer-Schönberger, 2014),并得到各界人士的大力推荐,大数据顿时成了炙手可热的富矿。大众传媒报道的传奇故事,也给大数据蒙上了神秘色彩。2012年,《纽约时报》发表了一篇大数据深度报道(Duhigg, 2012),介绍塔吉特(Target)公司通过大数据分析,比一名高中女生的父亲更早知晓该女孩怀孕。在统计分析师安德鲁·波尔(Andrew Pole)

的带领下,该公司分析人员利用大数据,对孕期妇女常购买的产品进行归类,分别给予预测值权重,然后利用相关模型,预测妇女是否怀孕,甚至什么时候生产。塔吉特公司由此可以提前发放折扣券,争取客户。

大数据的商业应用让人惊叹。每次在网上搜索物品,脸书网(Facebook)框边就“贴心”地出现相关广告,列出我们要找的物品。其他商家如亚马逊更是登峰造极。国内的淘宝网和各种社交媒体也积极采集并使用多维度数据,使得这个概念广为人知。

个人化学习概念的走红,也和大数据有关。脸书网创办人马克·扎克伯格(Mark Zuckerberg)在女儿出生后,宣布将捐出99%的财产用于社会公益。他想到的公益事业之一,是个人化学习。以脸书网的统计和分析实力,他自然有底气认为可以采集、分析、预测人的学习行为,让教育得到卓有成效的改进。教育领域会不会出现塔吉特公司波尔那样的神奇人物?可不可以像商界那样,能根据一系列变量、数据,帮助老师和学校准确认识学生学习过程,预测学习结果?我想问题并不这么简单。

在美国高校,大数据做得好的领域包括校友管理、新生宣传等。例如,很多学校需要校友捐款,因

[收稿日期] 2016-04-30

[修回日期] 2016-06-25

[DOI编码] 10.13966/j.cnki.kfjyyj.2016.04.003

[作者简介] 方柏林,美国德克萨斯州艾比林基督教大学课程设计总监(berlin.fang@acu.edu)。

此校友管理部门需要了解校友的行为习惯、捐赠行为、捐赠时机等,然后和校友建立长效联系。大数据能够有效地做出预测,使得劝捐有的放矢。自主招生的体制,使得各个高校都在寻找一流的生源。而未来学生的数据因一时无法获得,高校只能从高中和其他机构(如参展收集的资料),多角度地了解哪些学生可能进入本校,并展开相应宣传。由于本文重点是学习型数据,这方面暂不详述。

在课堂教学领域,大数据之光似乎还没有照进。老师对此兴趣索然。笔者所在的艾比林基督教大学亚当斯教学中心曾组织两次大数据培训,几乎没有老师报名听讲。“大数据”成了人人都曾听说,但很少有人真去探究的“熟悉的陌生人”。下面我们将分析这到底是什么原因所致。

二、教育大数据的挑战

关于大数据,教育界存在不少挑战:

一是大数据出现神秘化倾向。各种关于大数据的传说,让人们见证了大数据的神奇。但有讽刺性意味的是,大数据一定程度上被神秘化了,人们有可能把大数据变成新时代的看手相和查星座。没有数据的决策是不科学的,但是有数据的决策也未必靠得住。“数据”本身带有“科学决策”的光环。“大数据”给人的印象是有大量数据佐证的结论。但事实上,数据可以被操纵,同样的数据甚至可以得出完全不同的结论。人们可以借“大数据”的大旗,为自己的不合理结论或决策辩护。

教育界有时误以为问题无法得到解决,乃是因为数据不足,数据工具落后。由此就期待:一旦数据充分了,一连串问题就能解决。数据的完整、复杂、庞大,可以让人们接近真相。这会为教育机构服务学生和其他群体上的不作为提供借口。教育机构可能只关注数据的收集、分析和利用,但是舍本逐末,缺乏对人的关心。比如,上文说的校友大数据为例,建设校友网络,挖掘校友数据,这些操作都可以实现,但离开了关系的长远经营,离开了对现有学生人生的投资和关怀,就是照搬哈佛大学校友会的具体做法,恐怕也是事倍功半。

二是大数据管理能力欠缺。塔吉特公司能够设计出孕期预测模型,开展精准营销。教育上这么做恐怕不现实。尤其是规模不大的高校,研发部门往

往难有跨国公司的实力,建立仅针对本校的分析和预测模型。同时,悖论的现象是,有能力采集、分析、应用大数据的高校,如美国常青藤大学,中国 985、211 大学,有足够的实力吸引员工、学生、捐款人,反而不需要大数据帮忙。需要大数据的小学校,濒临生存危机的学校,大数据可以帮其精准营销,但学校声誉、工资待遇等现实因素,使得它们无法挽留开发和数据管理高端人才。

三是大规模上大数据对高校未必需要。顾客和商家是双向选择的买卖关系。他们依靠收集的大数据建立关系,是更为合理的选择。学校与学生、老师和学生关系直接、相互了解,是被鼓励的结果。而把学生当成甲乙丙丁,缺乏个体关注,反而是不被赞赏的做法。这种关系和环境,使得认识 and 了解可以且必须更为直接。能直接得到的知识,却绕弯通过大数据去采集,是舍本逐末,是为了大数据而大数据,无端浪费社会资源。

四是大数据悖论地让大多数学生成了被忽略的群落。大数据分析有分类(classify)、归总(cluster)(Sin & Muthu, 2015),然后呢?下文在哪里?被贴了标签的个体学生,怎样得到帮助?学生怎样才能吸引学校关注,让相关资源来到自己身边?这些都是问题。大数据对学生数据的收集,是多渠道的,是集成而不是分散的。它提供了一个出色的预警系统,使得困难甚至危险的苗头及时被识别,并有机会设计干预方案。但大数据分析的雷达往往只能探测处在常模分布比较边缘的人。学生状况变得特别,学校才会发现,而对大多数学生的学习过程是不清楚的,除了成绩外,知之甚少。学生不同的课有不同的老师,这些老师未必像大数据那样会将各自的认知汇总。老师可能对某些学生学习知之甚少,这些学生慢慢被边缘化。最终,当他们出现在大数据雷达上时,或许已回天乏术。消防部门常说,事前防火胜于临时救火。教育领域如何“防火”颇为有趣。这里我们可以看到教育数据如何影响教育。

五是文字陷阱导致方法陷阱。中文语境中,我们常把数据(data)和数字混为一谈,误以为必须进行复杂的统计和分析,利用高端而复杂的工具,才能分析人。数据不应是一堆统计数字,也应包括定性分析的工具。有时,即便商家收集的数据也应包括开放、定性的内容。定性内容多了,从中能看出规

律。文字一样可以打败数字,成为帮助我们形成洞察的工具。我们对于人的洞察、同理心和理解,未必都要通过各种渠道采集。有时候,和学生吃一次饭得到的洞察,会多过很多部门搜集的数据。

三、教育者如何利用学习数据

教育界说的大数据,多属名词滥用。学习的数据,往往不能和政府人口统计数据那样,涉及成千上万人和多个采集维度。更准确地说,学习数据属分析数据(learning analytics),或者说是“小数据”,甚至是数量不大、范围狭窄的“微数据”。数据是“大”还是“小”,其实关系不大,关键是要给学生带来学习过程和结果改进的“实惠”。下文举例说明不是严格意义上的大数据如何实实在在地改变教与学的过程和结果。

(一)预警失学风险

高风险学生(at-risk students)指濒临辍学、转学、退学风险的学生。美国高等教育市场化程度高,学校自主招生,学制灵活。学生选择多,从常青藤名校、巨型州立大学、普通州立大学、半公立大学、人文学院,到社区学院、两年制副学士学位学校等,种类繁多。大部分人只要想上大学,基本都可以上。由于宽进严出,很多人报读后发现跟不上,或觉得学校或专业选择错误,或源于其他原因,会选择中途放弃或转入其他学校。学生频繁“跳槽”会影响学校财源、排名和声誉。部分学校尤其是小学校,挽留学生成了大问题。应对的办法是及时识别学生辍学转学的动向,利用适当的干预手段,影响学生决定。

有些数据采集和分析公司专门帮高校收集各部门的数据,了解学生状况,方便学校及时开展工作,防止学生辍学或转学。笔者所在的艾比林基督教大学设有学生应急干预部门(Student Opportunities, Advocacy and Resources,简称SOAR),老师和员工可以根据学生情况,及时发出预警。预警系统收集的信息包括:

- 学生学习情况,如成绩滑坡、出勤、作业、考试焦虑、学习压力过大等;
- 身体状况,如疾病、缺少睡眠、遭遇虐待或无法完成相关任务;
- 精神问题,如有自杀倾向、情绪低落、失恋、抑郁、压力、焦虑等;

- 生活问题,如家里出现了危机、家人去世、朋友去世、经济困难等;
- 专业和职业选择问题,如无法确定专业、不适合目前专业等;
- 行为问题,如行为不当、行为对他人造成干扰、语言骚扰、肢体骚扰;
- 退学风险,如考虑转学,因为经济困难无法继续;
- 其他问题。

这些数据如果束之高阁,则浪费所有人的心血和时间。不过,这些信息也不是校方想公布就公布的。信息的披露和分享,严格受制于《家庭教育权利和隐私法》(Family Educational Rights and Privacy Act,简称FERPA)规定。学校一般要求发布预警的老师、员工告知自己是谁,甚至告诉学生自己通过SOAR系统求助,尽量不匿名发布预警。这样学校可以把预警问题和提出预警的人员联系起来,把发布预警的人当成解决问题的合作伙伴。大部分情况下,工作人员发布预警时也告知学生本人。这也体现了对学生的关心和爱。

发布预警的人,包括老师,未必有兴趣、精力、时间和专才,解决濒危学生面临的各种问题。学校相关部门可以和学生对接,比如学生面临压力、焦虑、家庭危机、自杀倾向等问题,心理咨询部门可以提供帮助;如果面临财务困难,学校可和学生重新商定学习计划,比如推迟毕业,或是帮助其寻找奖学金和助学金等;如果遇到写作、公开演讲、调查研究等“横向能力”方面的困难,学校可推荐学生向写作中心、演讲中心、图书馆研究服务台求助。

(二)深化教师评估

教育机构收集数据的手段之一,是开展问卷调查,尤其是对期末的课程评估。大数据多采集用户外围数据,形成特定认识,而直接向学生发放问卷,则可以让认识更直接。这种评价,几乎所有高校都在做,可惜数据质量差,不能帮助老师改进教学,反而还制造不满和怨言。究其原因,这种评估是给老师下判断,而不是帮助他们改进工作。

教学中,我们常提到形成性评估和终结性评估。对两者的区别,美国协和大学威斯康辛分校教育学教授伯纳德·布尔(Bernard D. Bull)形象地比喻为:“形成性评估如同每年去医院体检,终结性评估

如同验尸。”大部分学生调查是终结性评估,目的是区分得到学生好评的老师和被学生差评的老师,与评教授网(Ratemyprofessor)和中国“评师网”无异。它涉及教师绩效,事关升迁甚至工作安全,有些老师对这种评估比较警惕,若缺乏自信,或是不认为公平合理,则会在数据收集整理上不予配合。

我们使用一家叫 Idea Center 机构的调查问卷,目的是帮助老师改进教学。问卷问题包括:

- 获得事实性知识(术语、分类、方法、趋势);
- 学习基本原理、概括或者理论;
- 学习应用课程材料(提高思维能力、解决问题和决策能力等);
- 培养课程相关行业所需的具体技能、能力和观点;
- 培养团队协作技能;
- 培养创造能力(写作、发明、设计、表演、音乐、戏剧等);
- 拓展基本知识面和对高智力活动(如音乐、科学、文学等)的欣赏能力;
- 培养口头或文字表达技能;
- 学习如何利用资源回答问题,解决问题;
- 培养个人价值观,深化认识,坚定信念;
- 学习如何分析和审辩地评估观点和思路;
- 提出问题,寻求答案,以拓展进一步了解的興趣。

老师在发放问卷前,会评估这些教学目标哪些必不可少,哪些重要,哪些不重要。不是所有的教学目标都同等重要,把这种评估权交给老师,是现实的举措,也是对老师专业才能的尊重。

根据这些教育目标,问卷设计了一些问题,让学生逐条评分。下面以针对老师的问题为例说明:

- 关心学生和学习;
- 设法帮学生自行寻求问题答案;
- 课程所需功课(课堂活动、考试、项目)的时间部署合理,学生能跟得上;
- 展现出学科内容的重要性;
- 创建“团队”和“讨论组”促进学习;
- 明确说明讨论话题和课程的相关性;
- 评价学生的学业表现,能给出明确原因;
- 激发学生在课程基本要求之外深入思考;
- 鼓励学生使用多种资源(如数据库、图书馆

资料、外部专家)以强化了解;

- 将课程内容和现实生活结合;
- 明确、简要地介绍课程材料;
- 考试和作业项目等覆盖课程要点;
- 介绍和内容相关的有趣观点;
- 让学生参与研究、案例研究和“现实任务”等比较应用型的教学活动;
- 激发学生设置挑战性的目标并努力实现;
- 让学生和背景或观点不同的同学分享自己的观点和体验;
- 在考试、报告、项目等方面提供及时、经常性的反馈,帮助学生改进;
- 让学生互相帮助,以增进对观点和概念的了解;
- 布置的项目、考试和作业,能激发学生开展新颖而有创造性的思考;
- 鼓励学生和老师课外互动(如办公室辅导、电话、邮件等)

问卷还包括学习进度、课程内容等方面的问题。这些问题基本吸纳了教学和学习研究中已被证明的结论,比如学生相互帮忙,可以促进学习;课程内容要有挑战性、有激励效果等。

老师收集这些问卷后,分析工具会根据学生的打分、老师先前设置的权重,计算老师总分,同时以整个数据库的结果为基准,让老师知道自己在全行业做得怎样。有些课程可能其他学校评分也低,就需要依据整个数据库的基准,将老师的评分重新调整;依据老师先前设定的目标重要性,将原始评分(raw score),换算成调整分(adjusted score),这有助于提高评估结果的准确性和可靠性,而不会因为极端情况,导致评分失真。

分析结果出来后,老师可能会发现,自己的学习内容设置比其他学校容易,他们由此可以主动改变课程内容设置。如果老师们都存在类似问题,这种结果可以反馈到教师发展部门,以提供相应的教师培训和发展方案。这种数据的收集,过程并不复杂,也不需要大规模采集,只需依据某一班级的反馈,但是它的形成性评估效果非常好,这就是典型的利用小数据办好事的范例。

(三) 学习数据收集和应用

另一种数据收集、分析和应用的软件为课程管

理系统。它能很好地帮助老师及时认识学生学习过程和结果。美国绝大部分学校都在使用课程管理软件,包括 Canvas、Moodle、Blackboard 等。这些课程管理软件可以放置大量教学内容,如课程纲要、讲义、视频、播客,也可以安排教学活动,如收集、批改作业,发布成绩,开展网络测试,开展在线讨论,甚至举办网络会议。我曾使用过 Moodle、Blackboard 等,现在使用 Canvas。三种软件都有较好的数据采集功能。下面以 Canvas 为例,说明老师可以从课程软件中采集什么数据,这些数据分别有什么用途。

1. 班级活动汇总数据

Canvas 软件的班级活动汇总能显示学生浏览量、活动量分布,从中了解哪些天学生的浏览或者活动比较丰富。对于网络课程来说,该数据可以让老师了解、预测学生行为习惯。这有利于老师合理发布教学内容,使学生不会一时负担过重,一时又无所事事。另外,学生提交作业的汇总数据,能表明哪些学生按时提交,哪些学生延迟,哪些学生没交,从而帮助老师了解学生在完成作业可能遇到哪些困难,比如可能有的作业学生找不到。老师还能看到学生成绩的大致分布,比如是否分布均匀,还是过于容易,过于困难等。

2. 学生活动汇总

该数据还可以显示学生的网页浏览量、参与频率、提交数量(包括按时、迟到、未交)以及分数统计。老师可以对这些数据进行排序,然后将学生活动分类,必要时将属于学习过程的活动和代表学习结果的成绩进行关联。数据也能让老师迅速将学习高风险的学生汇集起来,通过邮件提醒他们完成功课或是告知学习中存在的问题。

如果老师点击学生名字,还能查看学生个人活动统计图,从中判断学生是否积极完成课程任务,如有拖延或者耽误,一般出现在什么时间(比如是不是期中考试期间),方便教师采取相应干预行动。

3. 测试内容分析

测试分析可以让老师查看任何一门网络测试课程的成绩分布,包括最高分、最低分、标准差和平均完成时间,让老师掌握测试内容的难度和学生对学习内容的总体掌握程度。

4. 测试问题分析

对测试中的问题,Canvas 可以提供答案分布,

帮助老师了解哪道题哪个答案学生容易出错。更宝贵的是,它能够提供单道考题的区分度(discrimination index),即某道考题和学生的总成绩关联度如何。关联度越高,说明该考题越能区分学生的水准。比如,某题的区分度为 0.87,说明该考题区分度高,即这道题答对,在很大程度上能判断出学生总成绩好坏。如果区分度为 0,说明这道题没有太大差别,答对机会均等。如果区分度为负值,如 -0.19,则说明这道题好学生可能失误,也说明题目的设计可能出了问题,老师可以采取相应行动,包括调整分数(系统有自动重新计算成绩功能)、修改考题等。

5. 作弊风险分析

Canvas 还有一种功能,亦即记录学生参与测试过程的数据,比如哪道题花了多少时间,来回修改次数,或是遇到某道题时离开了页面(可能上网搜索了)。这能给老师提供关于学生考试的很多有用信息,尤其是存在作弊风险时。比如,老师可以从测评日志中看出,学生可能在做第一题时遇到了困难,当时离开了页面。第二题学生改了一次答案。这些细微的数据足迹,能帮助老师了解学生对某项内容的熟悉程度,以及学生有无作弊的可能性。顺便说一句,网络课程中测试作弊风险的工具还有其他软件。比如,在写作上,老师可以利用 Turnitin 软件,将学生作文和网络中的海量文章比对,查看作文抄袭的可能性。

课程管理软件的这些数据,对老师非常有用,很多数据可以马上促进行动,比如可以对教学活动(例如讨论)进行排序,让未发帖的同学在页面上显示,老师可以统一公布提醒。试题的分析,让老师可以对题库进行修订,测验中发现某道题不合理,可以及时删除或者修改。这些数据公布给学生,可以帮助学生了解、督促或激励自己的学习。

四、大数据时代学生的选择

学生有无办法利用数据改进学习呢?学生如何防范大数据风险,以及如何采集、利用“小数据”“微数据”,提高学习效果?

(一)警惕大数据气泡

首先,要警惕大数据风险。隐私的暴露,自然是风险之一。另外一大风险,是大数据可能形成气泡和蚕茧,让使用者感觉贴心而舒适,但这么做恰恰有

害于未来发展。在线电影网站 Netflix 在了解客户兴趣方面做得很好。观众每看过一部电影,它就会推荐类似电影。有段时间我纳闷为什么老看野外生存电影,原来是 Netflix 根据我的观影记录推荐名单,并显著地放置在首页上。不知不觉地,大数据就给我们制造了相关蚕茧。或许我应该看别的电影,让自己的兴趣有所拓展。除娱乐活动外,商家会根据大数据识别出形形色色根据人们行为习惯定制的内容引导我们消费。如果学校在大数据方面做得好,号称学习可以定制,建议大家要小心。有时学习者需要打破那些“贴心”的诱惑,不要待在自己的舒适区不可自拔。学生需要所在领域的老师和专家加以指引,适当走出舒适区,让技能得到丰富或延伸,不能仅根据大数据推送的信息,而应该让自己接触尚且未知、但了解之后有益的内容。

(二)蓄意训练学习习惯

多份研究显示,人的行为和决策近 50% 来自习惯,而不是蓄意的思索(Neal et al, 2006; Wood et al, 2002)。人做事有时是顺其自然、随波逐流的,根据方便而不是理性进行决策行事。这么做自然有好处,比如好的习惯,使得我们的行为更有效,心理更放松,也将思维空间释放出来从事其他活动。但是这种自动化过程也代价沉重,比如人们可能在情况发生微妙变化时,仍按照原有方式做事,这就失去了学习的机会。另外,不良的习惯,比如顺手刷微信,也造成了注意力的分散和时间的浪费。

如何打破不良习惯,使得思想更为专注? 一个好办法是把提醒我们从事某种习惯的提示消灭掉,或者藏起来。习惯行为一般追随“提示-行为-奖赏”(cue-routine-reward)。如果蓄意改变提示,比如去阅览室时把手机装进书包,放到一边,专注地看两小时书,然后给自己奖赏,比如出去用一下手机,这会改变行为规律。

(三)过“量化自我”的生活

《连线》(Wired)杂志记者盖瑞·沃尔夫(Gary Wolf)在题为《量化的自我》演讲中,提出了人类如何从智能腕带到智能耳塞,追踪身体的各种信息,包括摄入热量、睡眠时间、走路步数等。换言之,凡能通过机器搜集的身体信息,几乎都可搜集到,并可以借此规划生活。

但关于学习的信息,却零碎、分散、抽象,没法准

确收集到,因为我们尚缺乏好的手段整理学习过程方面的活动。大家可以自己做些登记,比如用时间日志的土方法登记时间。我用一种叫抢救时间(Rescuetime)的工具,帮助了解时间到底都去了哪里? 该工具会对用户一天时间归类,让用户知道有多少时间浪费,有多少时间有效使用。这种数据能给用户警醒和鼓励,让他们更好地利用时间。

(四)主动使用技术工具帮助学习

比如,学生可以利用 Quizlet 测试工具帮助学习。很多时候,学生把大量时间花在重复练习上,不知道自己到底掌握了没有。Quizlet 能够帮助学生以自测或游戏的方式,在反馈中重复练习,并生成有关数据。这种数据极简单,比如哪些题常出错,完成了多少,纯属微数据,但学习效果很好。由于制作这些练习耗费时间,学生可以组建成学习小组,各承担自建复习题库的部分任务。

还有一种叫 Brainscape 的工具,能够让人学习基本概念,同时也生成并掌握学习者对内容的精熟程度。学生可以一边练习,一边标记自己是否熟悉,如果标记为“非常熟悉”,同样的内容就不会出现,系统如此继续循环,直到学习者熟悉所有内容。

(五)熟悉如何在信息过剩时代找到最优质的学习资源

网络学习资源十分丰富。比如,国内教育资源包括网易公开课、果壳网的 MOOC 学堂等。各种微信公众号也向人们推送非正式学习资源,比如罗辑思维、知乎、澎湃新闻等。这些资源在非正式学习上战功赫赫,但是在正式学习上,我们不能寄希望于现有的比较流行的资源。这些市场上颇受追捧的资源,包括畅销书,不足之处在于往往为了贴近大众口味,而变得过于通俗,或是通过名人咀嚼后让学习者消化。貌似对学习者来说省时省力,但是知道一大堆东西拿出去炫耀的时代已经过去了。我们未来需要的知识结构和能力,要求我们目前就要把根扎深。我们还得去看业内的论文,养成深度阅读的习惯,知道自己所在领域的来龙去脉,这会给我们带来更为长久的乐趣和更大的改进。

五、结语

在人人都谈大数据的时代,教育界未必能像塔吉特公司那样,有雄厚实力利用多维数据,批量了解

特定的细分客户,继而开展精准营销。在学生方面,大数据识别边缘学生的状况能起很大作用,但是往往不能为教师和学生所用,为其教和学服务。教育界多非营利机构,采集、分析、利用大数据的力量有限。但我寄希望于大家通过各种工具和技术,打通教与学之间的毛细血管,形成新的对接关系,让老师与学生、学生与自己、学生与学习内容能有更有效地互动。如能这样,我们可以借助小数据和各种数字技术平台,给学习帮大忙。

[参考文献]

- [1] Bernard Bull, E. (2014). 10 assessment design tips for increasing online student retention, satisfaction and learning. Faculty Focus. Retrieved 2 April 2016, from <http://www.facultyfocus.com/articles/online-education/10-assessment-design-tips-increasing-retention-satisfaction-student-learning-online-courses/campaign=FF140203article>.
- [2] Duhigg, C. (2012). How companies learn your secrets[DB/OL]. Nytimes.com. Retrieved 31 March 2016, from <http://www.nytimes.com/2012/02/19/magazine/shopping-habits.html>.
- [3] Mayer-Schönberger, V. , & Cukier, K. (2014). Big data: A

revolution that will transform how we live, work, and think (First Mariner Books edition)[M]. Boston: Mariner Books, Houghton Mifflin Harcourt.

- [4] Neal, D. T. , Wood, W. , & Quinn, J. M. (2006). Habits: A repeat performance[J]. Current Directions in Psychological Science, 15(4): 198-202.
- [5] Press, G. (2014). A Very Short History Of Big Data[DB/OL]. Retrieved 25 June 2016, from <http://www.forbes.com/sites/gilpress/2013/05/09/a-very-short-history-of-big-data/#35e769d855da>.
- [6] Sin, K. , & Muthu, L. (2015). Application of big data in education data mining and learning analytics: A literature review[J]. IC-TACT Journal on Soft Computing, 5(4): 1035-1049.
- [7] 涂子沛(2012). 大数据:正在到来的数据革命[M], 南宁:广西师范大学出版社.
- [8] Wood, W. , Quinn, J. M. , & Kashy, D. A. (2002). Habits in everyday life: Thought, emotion, and action[J]. Journal of Personality and Social Psychology, 83(6): 1281-1297.
- [9] Wolf, G. (2010). The quantified self. TED Talk, Cannes [DB/OL]. Retrieved from https://www.ted.com/talks/gary_wolf_the_quantified_self.

(编辑:徐辉富)

Challenges and Opportunities of Using Big Data in Education, and the Use of Small Data for Educational Impact

FANG Bolin

(Abilene Christain University, USA)

Abstract: While colleges and universities seek to utilize big, complex and aggregated data to generate insights to support students, big data is limited in its use for the design of interventions due to a few limits it inherently has, including the overt focus on quantitative data, disconnect between university services and classroom teaching, and lack of support for faculty who may not be interested in big data. However, “small data”, one-dimensional teaching and learning analytics, provide quality feedback to teachers and students. In this paper, we will discuss how a small private university uses IDEA evaluations to gather feedback to improve teaching practices, and data from the use of a learning management system (Canvas) to identify ways to improve student learning by providing formative evaluation for faculty to improve their practices. Towards the end of the paper we will discuss how students can use learning analytics to improve their own learning through self diagnosis, “quantified self” behavior data collection and reflection on learning practices.

Key words: small data; learning analytics; teacher evaluation; Learning Management Systems; learning tools