

基于大数据的教学评价研究*

李葆萍 周颖



(北京师范大学 教育技术学院, 北京 100088)

摘要: 文章基于教育大数据的视角, 结合教育评价活动的转型, 对教育大数据的采集、可视化分析与呈现技术等开展了研究。文章认为, 教育大数据技术推动了教育评价中数据驱动决策的实现, 为多方参与教育评价、实现发展性学生评估提供了良好的支持。平板电脑、数码笔、可穿戴设备等能够实时地将不同类型的学习数据数字化, 实现了对学生学习全过程数据的采集, 为教育领域中实现基于数据分析与理性证据的教育评估与决策提供了数据基础。各类可视化分析工具能够在稀疏的教育大数据中过滤、挖掘各类隐含的教育信息和规律, 帮助我们理解学生个人知识体系的构建过程, 探索学生个体的社会学习网络的演化规律, 揭示教育事件在特定时空呈现的特征。

关键词: 教育大数据; 教育评价; 发展性评估; 数据采集; 可视化分析

【中图分类号】G40-057 【文献标识码】A 【论文编号】1009—8097(2016)06—0005—07 【DOI】10.3969/j.issn.1009-8097.2016.06.001

数据(Data)是按照一定规则排列组合的物理符号, 可以表现为符号、文字、数字、语音、图像、视频等形式, 是信息的表现形式和载体。2001年高德纳咨询公司在研究报告中指出, 人类社会数据爆炸从三个维度展开: 一是同一类型的数据量在快速增大, 二是数据增长的速度在加快, 三是数据的来源和新的数据种类在不断增加——上述事实清楚地说明人类真正进入到了大数据时代^[1]。

关于大数据的定义暂时还没有达成共识, 现存的定义主要从大数据具备的性质出发进行界定。目前, 研究界认可的是大数据应当满足3V特点, 即规模性(Volume)、多样性(Variety)和高速性(Velocity)^[2]; 随后, 不同的机构将其扩展成为4V, 如IBM提出的准确性(Veracity)以及IDC提出的价值性(Value)等^[3]。大数据立足于对大量数据的深度挖掘与科学分析, 寻求数据背后的隐含关系与价值, 使得人们可以从基于小样本数据的推测或基于感性的偏好性选择转向基于数据分析与理性证据的决策^[4]。当前, 大数据正不断地深入教育领域, 对一些“数据密集型”的教育业务如考试、学习分析等产生了强劲的冲击。

一 大数据与教学评价变革

教育评价指的是在系统、科学、全面地搜集、整理、处理和分析教育信息的基础上, 对教育的价值做出判断的过程。从个人层面来看, 教育评价的目的在于了解学生的发展情况, 对学生的学习情况进行客观总结、对教师的教学质量进行评估; 从宏观来看, 评价的目的更在于促进教育改革, 提高整个国家的教育质量^[5]。我国《国家中长期教育改革和发展规划纲要(2010-2020年)》指出: “要改进教育教学评价, 根据培养目标和人才理念, 建立科学、多样的评价标准。开展由政府、学校、家长及社会各方面参与的教育质量评价活动。做好学生成长记录, 完善综合素质评价, 探索促进学生发展的多种评价方式。”美国于2010年发布的《国家教育技术计划》(National Education Technology Plan)中强调各级各类教育系统要利用技术来测量、评价学习过程, 教育管理者应该利用技术来收集学习过程中的实时数据, 为持续改善学习效果提供依据^[6]。

中美两国的重要教育文件都揭示了教育评价在整个国家教育系统中的重要地位,并预示着教育评价将在理念、内容、方式等方面的转型取向,在现代教育价值趋于多元化的基础上,教育评价方式面临全面转变的现实需要^[7]。这种转变主要表现在如下方面: 做好学生成长记录、收集学生学习过程中的实时数据等评价措施的提出,体现了评价理念从以往“经验主义”向“数据主义”转变的趋势; 完善综合素质评价、探索促进学生发展的多种评价方式以及提供持续改善学习效果等评价目标的提出,体现了评价内容从以往注重认知水平的“总结性评价”向综合素养的“发展性评价”转变的趋势; 政府、学校、家长及社会各方面参与的评价主体的提出,体现了评价方式从“单一封闭”向“多元开放”转变的趋势。事实上,上述教育评价的转型并非全新的理念和思想。依据数据所提供的证据进行判定是实证主义的基本思想,发展性原则一直是教育评价本身应有之义,学生的自我评价、家长的校外评价^[8]也是很多学校采取的评价方式之一。然而近几年来,这些思想在国家教育纲要性文件中被显著地提及,说明其在过去落实的力度欠缺,或者说直至今日这些思想的落实方才具备可操作性的条件,这就是现代信息技术和大数据在教育领域中应用的用武之地。

教育评价是为了让我们更好地了解学生、审视我们的课堂和教学过程。在传统的教育环境下,了解学生的主要方法为问卷调查、课堂行为观察、考试、作业分析等。这些方法存在着耗时长、数据不准确、过程型数据遗漏或者无法采集等多种弊端,建立在这种不完整数据之上所获得的分析结果只能揭示某些特定的问题,缺乏综合性。此外,不同来源的数据之间难以整合,因为采集成本等原因,获得的数据缺少持续性,导致数据内部隐含的信息连接被割裂。比如学生作业水平和学生课堂学习行为之间的所存在关系的挖掘; 学生阅读能力对其数学学科表现的影响分析等都难以实现。因而教师往往只能根据经验来处理教学问题,这些都对于科学、精准地了解学生,做出教学决策甚至制定教育政策造成不利影响。教育大数据的应用则为克服现有教育评价中的不足提供了效果良好的解决方案。

1 大数据推动数据驱动的教学决策

数据驱动决策在教育中是指收集、分析、报告和使用数据用于教育教学改进的过程^[9]。比如美国普渡大学的“课程信号灯”(Course Signals)项目是国际知名的大数据诊断学生、提供教育决策的典型示例之一^[10]。如图1所示,“课程信号灯”系统主要以成果算法为基础,对学生课程表现、课程努力程度、前期学业历史、学习者特征等数据进行采集和计算,实现对课程的实时预测。预测结果将通过红、黄、绿三种颜色信号灯的形式,呈现在学生的学习页面以及教师的课程控制页面中——红灯表示课程学习中存在极大的失败可能性;黄灯表示在课程学习中存在一定的问题,有失败的可能;绿灯表示学习成功的几率很高。根据不同的信号显示,教师通过发送电子邮件、短信以及面谈等方式对学生学习进行适当的干预,还可以通过“课程信号灯”系统自带的推荐学习导师与学习资源模块,对学习者的学习提供适当的帮助,以促进其在课程学习中取得成功。

2 大数据促进了学生发展性评估

早在1940年,美国史密斯—泰勒报告中就指出教育评价不能只测量学生的某些能力和特征,而应该根据教育目标来评价学生发展及成长的进程和水平,这是发展性评价理念在现代教育评价中的确立。发展性评价是指通过系统地搜集评价信息和进行分析,对学生的教育活动进行价值判断,实现其发展目标的过程。发展性评估主要发挥评价诊断的功能,突出评价的过程,重

视学生的个性差异,因此,其往往要和学生的学习过程紧密结合,进行长期追踪。如北京、成都、深圳等地的中小学校,以发展性评价理念为指导,持续跟踪学生历次考试成绩,通过时间序列分析、聚类分析等手段,对学生的学习数据进行挖掘,构建学生的学科知识地图,进行学习风格和学习行为分析,最终完成对每个学生的学习力诊断^[11]。

2015年成立的北京市未来教育高精尖创新中心,正致力于利用学生学习全过程的大数据分析,帮助学生发现并提升优势学科、诊断和补救劣势学科,以适应新型的学生评价机制,实现个性化、差异化的学生发展目标。

3 大数据提供了多方参与评价的途径

过去的学生评价主要是针对学生的学业水平测试,评价主要由学校相关部门和教师完成,整个评价体系呈现出封闭性的特征。当前强调学生的发展性评价和综合素质评价,评价活动贯穿学生的整个学习过程,覆盖学生在校园内外的学习活动和行为表现。多种来源、结构不同的数据汇总将用来分析学生的综合素质,并通过数据的不断积累,使各类参数和模型得以确立,以提高分析的精确性。可见教育大数据直接产生于各种教育活动(包括教学活动、管理活动、科研活动、校园活动等),每个教育利益相关者既是教育数据的生产者也是教育数据的消费者^[12],基于开放性大数据的评价活动,为那些能够掌握和提供学生不同情境下学习数据的多方主体共同参与评价活动架设了桥梁。可以说,数据是驱动教育评估转型变革的核心因素。研究认为,从用户的视角来看,在教育评价活动中很容易从数据的入口和使用,即数据的采集和分析两个方面感受到大数据带来的变化。

二 教学评价数据的采集

1 教学评价数据的采集类型

2013年颁布的《教育部关于推进中小学教育质量综合评价改革的意见》强调要改革评价方式,将定量与定性评价相结合,注重全面客观地收集信息,根据数据和事实进行分析判断^[13]。该意见说明,教育质量评价不再单一地依托考试成绩,而是要将学生发展的所有信息收集、整理、分析并得出结论性的认识,也就是说要基于“数据”和“证据”对教育质量做出评价^[14]。在当前倡导的以学生为中心的学习环境中,学生学习“数据”和“证据”的主要产生途径如图1所示。其中,学生在学校内外的学习过程数据包括学习交互、学习行为、学习路径、各类过程性学习档案等,学习者个性数据则包括学生的生理、情感、认知状态数据等,以及各类以考试、作业、作品等形式展现的学习成果。

2 教学评价数据的采集技术

为了实现“数据”+“证据”为基础的教育评价,学校采取了很多评价数据采集措施,如考试、问卷、作业档案袋等。然而这类信息多是采集学生的学习结果信息或者状态信息(如学习风格),属于静态信息,而学生在学习过程中实时产生的诸多动态信息如学习路径、学习行为等未能实现有效采集。不完整的数据采集势必会影响评价结果的准确度和可信度,也会制约大数据下教育评价的实施,因而教育全过程数据的采集研究是大数据应用与教育评价的关键要素。

目前,智能学习环境以及具有数据采集能力的学习终端如平板电脑、智能手机、数码笔、可穿戴设备等的应用,为破解学生学习数据采集难题提供了技术方案。



图1 学生学习大数据来源

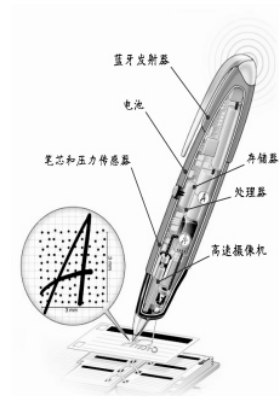


图2 数码笔结构图

平板电脑是一种便携的智能设备，很多学校将其配备给学生开展数字化学习。通过平板电脑中电子教材系统、数字资源系统、作业与考试系统和互动交流系统可以记录和生成学习行为大数据^[15]，主要包括海量的学习内容序列节点数据、海量的学习资源序列节点数据、海量的练习评测序列节点数据和海量的互动交流序列节点数据等。这些数据涵盖了学生在学习过程中所产生的所有学习行为数据，再结合平板电脑环境感知数据能够对学生进行多元深层次评价，从而发现学生真正的不足和潜能。

一些学校也尝试在学生最自然的纸笔环境下采集学习数据，而数码笔可以帮助实现这样的目标。如图2所示，数码笔与普通笔的外形和使用方法完全一致，所不同的是数码笔内置高速摄像机，书写时压力传感器可以把笔尖的信息传递到摄像机，从而实时记录和采集笔迹的轨迹。利用数码笔可以实时记录学生在纸笔环境下的学习行为，并对学生的学习状态和教师教学提供诊断信息。如数码笔采集的XXX同学在语文课堂学习文言文的记录和诊断信息如下¹：

XXX同学从45秒处画了个小人。1分50秒点断第一段结束，2分50秒开始点断第二段，估计是一分钟通读。3分30秒点断结束，同时有回看，将前边疑难断点点出，**说明其有回看意识和整合文段的意思**。4分49秒翻页，第一篇点断完毕。

5分15秒完成第二篇第一段点断，同时有回读。6分32秒大体完成第二篇点断，7分52秒回到第一页。此后一直没有动作，直到教师朗读指令发出。16分59秒开始画小人，18分50秒小人画完，同时还完善了前边画的小人，**说明这2分钟该生在走神**。

22分20秒开始翻译，翻译速度很快，“就之而不见所畏焉”一句不会翻译，主要是**不了解此话所代表的含意，缺少认知背景**。

数码笔也可以在教师批阅作业或试卷的同时，将每个学生的平时作业及考试信息以数字化的方式记录下来，结合课堂学习行为表现能够为学生和教师提供个性化的评价和学习分析诊断。

学生学习的全过程大数据除了我们熟知的课堂中的学习数据以外，还包括课堂外学习数据和学生的个人生活、心理等状态数据。尽管人们早已认识到这类数据的重要性，但现实中却难以精准采集这类数据。学习分析专家Siemens教授指出，学习分析所需的数据主要来源于学习

1 样例数据来源于拓思德公司数码笔实验课堂

管理系统、键盘数据、浏览器数据等学生操作电脑的数据，缺少对现实情境中学习过程数据的捕捉，从而导致分析者难以深入了解学习和教学过程^[16]。2007年凯文·凯利提出了量化自我的概念，就是用实时测量或记录的方法，测试、量化和记录个人的生命数据（如饮食、运动、睡眠、情绪等），并通过数据反馈进行自我调整^[17]。量化自我概念的提出和大数据密切相关，其中传感器网络、可穿戴技术、移动终端等对量化自我起关键作用。

教育领域也可使用RFID芯片、眼动仪以及可穿戴设备如各种手环等技术，真实地采集学生在教室内外的学习信息和学生的日常行为数据，供精确化学习分析和个性化教育评估与管理的使用。如眼动技术通过对眼动轨迹的记录，从中提取诸如注视点、注视时间和次数、眼跳距离、瞳孔大小等数据，从而研究个体的内在认知过程。有研究就采用眼动仪采集2~3岁婴幼儿对图画书页面区域的注视点个数、注视注视点时间长度、回视等行为数据和生理数理，用来评估其图画书阅读的注意力水平、阅读喜好、读图能力、阅读理解能力等^[18]。手环等设备则可以实时地记录学生的位置信息、运动数据和身体健康指标等，可以随时将这些数据与学生的学业表现关联，用来分析、预测不同类型学生的学习表现和发展。

三 教学评价数据的分析

采集到教育大数据以后，需要通过数据挖掘和分析技术来提取数据中蕴含的教学信息以便开展教学评价。除了常规的教育统计手段外，还可对大数据构成的数据立方体进行多维度的下钻或上卷操作以提炼出更深层次的知识需求，促使聚类分析和关联分析等数据挖掘技术越来越多地应用于学生的教学评价当中。

如根据历次试题答题情况对学生进行分类分析，可以确定学生的基本类型并帮助甄别表现异常的学生。若聚类结果发现部分学生多次解答难度高的题目能够保持正确，而难度低的题目错误率较高，可能这部分学生是因为粗心造成，可以对其答题注意力状态进行辅导；部分学生若长期聚类在同一知识点答题正确率起伏较大的区间，则可能这部分学生存在抄袭或者随意选择的情况，可以进一步关联其它数据对其进行分析和诊断。再如对学科不同知识点得分情况进行关联和回归分析，就能精准地进行教学归因分析，如B知识点得分低，是因为A知识点得分低，那么应当先解决A知识点的学习问题等。

上述这些数据分析很多是基于结构性数据的分析，它能够确定或者否定我们对教学问题的预期性判断。然而教育大数据具有非结构化、稀疏性等特征，难以像结构化数据的方式构建出其内部的正式关系，或者使用先验的模型来进行分析，这就需要在教学评价领域引入新的数据分析技术。在当前大数据环境下，信息可视化分析可以通过可视化图形呈现数据中隐含的信息和规律，建立起符合人类的认知规律的心理映像，成为人们分析复杂问题的强有力工具^[19]。

教育领域中的大数据主要包括师生基本信息数据、课业测试与作业数据、校园实录数据、课程资源数据^[20]、课外学习数据、学生身体健康数据、社会行为数据等。这些数据主要以文本、网络（图）、时空数据等信息类型呈现，基于大数据的学习评估可以引入相应的可视化分析工具。

使用文本可视化分析工具，能够将文本中蕴含的语义特征（如词频与重要度、逻辑结构、主题聚类、动态演化规律等）直观地展示出来。可以使用基于文本的语义结构可视化技术，以树形结构或者放射性层次圆环结构来展示文本结构、文本相似度统计、修辞结构以及相应的文本内容等。还可以结合文本的空间、时间属性，进一步展示主题的合并和分支关系以及演变，

并借助文本可视化分析工具，通过大数据对学生知识点和学科知识体系的评估，能够迅速了解学生的知识掌握状态以及学生个人知识体系的发展进化状态。

网络关联关系是大数据中最常见的关系，在联通主义的影响下，学生学习活动的社会性和连结性日益凸显，学生—学习资源—资源的提供者—使用者—学习硬件—学习场所等以前呈现出弱连接甚至无连接的信息和结点越来越密切地关联在一起、越来越频繁地发生互动。通过关联分析，可以将学习者个人信息数据、学习者学习内容数据、学习者社会性学习交互数据、学习设备数据等多维度数据关联在一起，计算之间的关联度，使用如H状树、圆锥树、气球图、放射图、双曲树等来表征这类信息。评估和揭示学生的学习网络关联，有助于辨识学生社会学习网络的构建和演化规律，有助于以学习者为中心的学习方式的实现。如图4所示，采集了社会化学习资源平台上与一个知识点关联的行为数据和用户数据，构建成该知识点的网络关联，借此可以评估不同学习者的学习参与度和贡献度等。

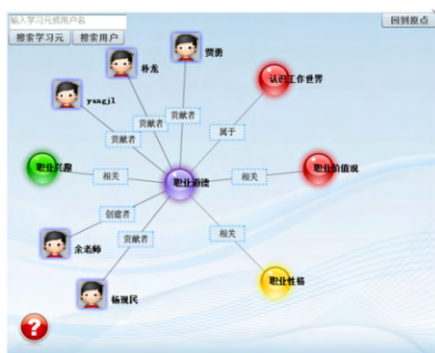


图4 基于知识的社会网络可视化分析

时空数据是指带有地理位置与时间标签的数据，传感器与移动终端的迅速普及使得时空数据成为典型的教育大数据。这类数据更多地反映了学习者在学校以外的学习行为，揭示和补充了教育教学这个复杂系统的真实的状态数据。将教育数据的时间流与地图进行融合，可以立体地展示教育事件在特定时空呈现的特征，预测未来的发展，为教育决策提供有力的支持。如在美国匹兹堡教育网站上，可以通过“动态地图”把复杂的数据转化为形象直观的图像，用不同颜色的图标表示不同的学校状态，绿色的图标表示这个学校超过平均水平，橙色的表示平均水平，而红色的就是预警学校了。通过图5展示的这种可视化分析方式，直观显示不同地区学校的教育质量，能有效支持科学的教育决策。

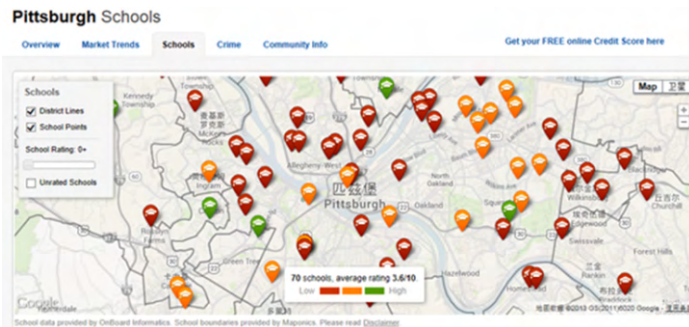


图5 教育时空数据的可视化分析

四 小结

教育是一个复杂系统,对教学问题进行科学的评价和精确的诊断干预是极具挑战性的任务。借助现代化的教育教学技术得以保留和采集的教育大数据,为完成这样的任务提供了很好的基础环境。尽管如此,应该认识到基于全学习过程的教育大数据,在数据类型上具有半结构化和非结构化的特征,在数据容量上具备海量和天量的特征,在数据处理上具有实时性的特征,过去常规的数据采集和分析技术难以满足大数据背景下的需求。这就要求在教学评估领域快速积累有关爆炸式的大数据采集、存储、清洗等方面的知识和技术,在教育数据采集的质和量之间进行有效权衡;同时,加强对教育大数据所揭示出的新的学习和教育规律的研究,将其吸纳于教学评价、学生学习模式的构建和验证之中,以提升教育大数据挖掘的效率和质量。此外,虽然教育大数据在对学生进行个性化、差异化、精确化教学诊断方面具有潜力,但也应建立数据采集、使用和公开规范,以保护学生、教师的隐私和安全等。

参考文献

- [1]涂子沛.大数据[M].桂林:广西师范大学出版社,2012:55-56.
- [2]Grobelenik M. Big data computing: Creating revolutionary breakthroughs in commerce, science and society[OL]. <http://videlectures.net/eswc2012_grobelenik_big_data>
- [3]孟小峰,慈祥.大数据管理:概念技术与挑战[J].计算机研究与发展,2013,(1):146-169.
- [4]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.
- [5][8]王道俊,王汉澜.教育学[M].北京:人民教育出版社,1999:290-293.
- [6]周由游,施建国.技术推动学习的新模式——美国国家教育技术计划的启示[J].中国电化教育,2011,(10):54-58.
- [7]戚业国,杜瑛.教育价值的多元与教育评价范式的转变[J].华东师范大学学报(教育科学版),2011,(2):11-18.
- [9]University of New Mexico College of Education. Data-based decision making for student success[OL]. <<http://coe.unm.edu/uploads/docs/ipd/issue3.pdf>>
- [10]Pistilli M D, Arnold K E. In practice: Purdue signals: Mining real-time academic data to enhance student success[J]. About Campus, 2010,(3):22-24.
- [11]台湾网奕.clouDAS 云端诊断分析服务[OL.] <<http://www.habook.com.tw>>
- [12][20]杨现民,唐斯斯,李冀红.教育大数据的技术体系框架与发展趋势——“教育大数据研究与实践专栏”之整体框架篇[J].现代教育技术,2016,(1):5-12.
- [13]教育部.关于推进中小学教育质量综合评价改革的意见[Z].北京:教育部,2013:2-10.
- [14]覃明勇.“数据+证据”:中小学教育质量综合评价的呈现与表达方法[J].广西教育学院学报,2014,(6):196-199.
- [15]章怡,牟智佳.电子书包中的教育大数据及其应用[J].科技与出版,2014,(5):117-119.
- [16]魏雪峰,宋灵青.学习分析:更好地理解学生个性化学习过程——访谈学习分析研究专家 George Siemens 教授[J].中国电化教育,2013,(9):1-4.
- [17]陈然,杨成.量化自我:大数据时代教育领域研究新机遇[J].现代教育技术,2014,(11):5-11.
- [18]金慧慧.2-3 岁婴幼儿图画书阅读眼动水平研究[D].上海:华东师范大学,2010:83.
- [19]任磊,杜一,马师,等.大数据可视分析综述[J].软件学报,2014,(9):1909-1936.

Research on Educational Evaluation based on Big Data

LI Bao-ping ZHOU Ying

(School of Educational Technology, Beijing Normal University, Beijing, China 100088)

Abstract: Based on the perspective of educational big data and the transformation of educational evaluation activity, the collection and visual analysis and presentation technique of big data in education area were discussed in this paper. The paper stated that big data promoted the realization of data-driven decision making in education evaluation, provided good support for multi-stakeholder participation in education evaluation and developmental student evaluation. Tablet computer, digital pen, and wearable devices can digitalize the different kinds of learning data in real-time, which made the data collection during the whole process of students' learning and provided data base for the data-driven and evidence-based education evaluation and decision. All kinds of visual analysis tools can filter and dig all kinds of implicit education information and rules in the sparse educational data, which helped us understand the construction process of the students' personal knowledge system, explore the evolution law of the students' individual social learning network, and reveal the characteristics of educational events under certain time-space conditions.

Keywords: educational big data; education evaluation; developmental assessment; data collection; visual analysis

*基金项目：本文受2014年度北京市教育科学十二五规划青年课题“北京市中小学师生对智慧教室环境感知研究”（项目编号：CJA14186）资助。

作者简介：李葆萍，博士，研究方向为智慧学习环境、平板电脑教学，邮箱为libp@bnu.edu.cn。

收稿日期：2016年3月12日

编辑：小西