

教育数据可视化:发展历史、研究热点及典型应用

王一凡^{1,2}

- (1. 华中师范大学 教育学院,湖北 武汉 430079;
2. 武汉市教育科学研究院,湖北 武汉 430021)

[摘 要] 随着大数据兴起,各领域掀起了大数据分析的浪潮,教育领域也不例外。近年 E-learning 的兴起形成了教育数据的海量存储,构成教育大数据。教育数据可视化即对教育大数据进行可视化分析,为学生、教师、家长等不同角色人群带来极大便利,在教育领域中存在着巨大潜在价值。与国际相比,我国的教育数据可视化起步较晚,研究与应用探索尚待进一步加强。文章基于 citespace 文献分析法,追溯了数据可视化从古至今的历史渊源,分析了数据可视化在教育领域的发展趋势和热点主题研究,重点探讨了数据可视应用载体的典型代表——学习仪表盘在教育领域的应用,指出在线学习系统可视化过程中存在的问题,并提出应在未来应用中注重隐私保护、关注学习风格、加强用户反馈,引入动态交互、融合线下数据等相关举措,借此为教育大数据可视化的发展研究提供可行性参考和借鉴。

[关键词] 大数据; 可视化; 学习分析; 在线学习系统

[中图分类号] G434

[文献标识码] A

[文章编号] 1671-6973(2022)05-0076-13

一、引言

可视化来源于英语的^[1] visualization,原意是形象化显示(显像化或图示化)。数字技术中的可视化最先是指科学计算可视化,即运用计算机图形学的原理方法,将科学、工程等计算产生的大规模数据转换为图形、图像来直观显示,之后又衍生出数据可视化、信息可视化、知识可视化等概念。数据可视化研究数据视觉表征形式,以一种概要形式抽取信息(包括信息的属性和变量等),给人以视觉冲击。而视觉是人类最重要的信息获取通道,人类的大脑有一半以上的功能用于视觉感知^[2],当数据被图形取而代之展现在人的眼球下时,可以降低人类的脑部认知负荷,从而使人们快速地吸收理解数据。数据可视化借助图形化手段融入美学来传递信息,数据可视化的概念^[3-5]一直处于不断演变之中,其边界也在不断扩大,目前学界已存在多种数据可视化的定义,从广义的角度来讲数据可视化包括了比较成熟的科学可视化和较年轻的信息可视化以及知识可视化。从狭义的角度来说,数据可视化是指运用计算机图形学知识和图像处理技术将大型数据集中的数据以可交互式

[收稿日期] 2022-06-21

[基金项目] 本文受教育部重点课题“大数据智能采集终端赋能区域教育质量评测实践研究”(DHA210338)资助。

[作者简介] 王一凡(1973-),男,湖北广水人,博士,武汉市教育科学研究院副研究员,研究方向为教育质量评测大数据,学校德育。

图形图像的方式来展现的理论、方法和技术。

当前,信息技术发展推动教育领域的数据以指数级增长,形成教育数据的海量存储,教育领域进入大数据时代。所谓教育大数据是大数据的一个子集,是在教育教学和校园活动中产生,在教育管理和科学研究活动中采集,对推动教育发展有巨大潜在价值的数据集合^[6]。对在线学习系统数据进行可视化研究,进而发掘数据背后的价值,能够帮助学生高效学习,同时可以惠及教师、家长、教育管理者等,在提升教育质量、促进教育公平、提升教育服务等方面存在巨大研究价值。

二、发展历史

数据可视化就是让信息以更容易理解、记住、辨识或能够快速捕获重点信息的方式呈现出来。其发展历史可以分为三个阶段。早期数据可视化可追溯到17世纪,最初可视化用于表现星星和其他天体位置的图表,以及导航地图,可视化形式是基本的地图与图表,一个著名的例子是托勒密的世界地图,这也应用到了后来的地理科目教学当中。17世纪由于分析几何的兴起、测量误差理论和概率论的诞生,人口统计等的展开,数据可视化在理论和实践表现巨大的增长,对当前数据可视化的研究奠定了初步基础。18世纪末,人们开始进行地质、经济和医学数据专题制图尝试,出现类似柱状图的线图表述方式,引入抽象图和函数图。此时可视化开始应用到教学内容,对课本知识结构进行可视化展现以帮助学习者整体把握学习内容。此阶段的可视化方式已经接近当代科学使用形式^[7]。可视化在渐渐走向教育教学当中。

19世纪初到20世纪上半叶,是可视化应用于教育教学的蓬勃发展时期。立足前期设计和技术创新基础,这个时期统计图形和专题制图的运用成爆炸性增长,其速度是以前所不能比的,出现了统计图形中所有现代形式的数据显示形式:条形图和饼图、直方图、折线图和时序图、等高线图等。由于认识到数字信息在社会规划、工业化、商业和交通方面日益增长的重要性,欧洲建立了官方国家统计数据库,同时,高斯和拉普拉斯发起统计理论,并由Guerry^[8]和Quetelet^[9]扩展到社会领域^[10],提供了理解大量数据的手段,这也为后面的大数据可视化奠定了基础。

20世纪上半叶,数据可视化处于稳定发展时期,研究者认为这是一个必要休眠期,是应用和普及时期而不是创新时代。在此期间,统计图形成为主流,图形方法大量进入教科书^[11-13]、课程^[14]、政府标准^[15]、商业^[16]和科学,在天文学、物理学、生物学和其他学科中首次使用图形方法表达新见解、新发现和新理论。此时,静态图形的可视化技术大量应用到课堂教学内容中。同时开始对各种图形形式的可视化效果进行实验比较,例如研究者开发了许多实用的图形辅助工具^[17]。

从20世纪70年代开始,教育数据可视化发展进入了质变阶段。1975年至20世纪末,高维数据可视化开始发展。这一时期数据可视化的发展变化多端、发展迅速,涉及更广泛的学科,很难简要进行概述。布鲁斯·麦考梅克等所编写的美国国家科学基金会报告《Visualization in Scientific Computing》对于可视化领域产生了重大冲击,这份报告中强调计算机可视化技术的必要性^[18]。20世纪90年代初期,出现了信息可视化研究领域,分析抽象的异质数据集。21世纪初,由于动态图形的发展,兴起动态交互式的可视化,推动可视化领域的进一步发展。大数据时代的到来,数据可视化越来越受到研究者重视,大数据的爆发对数据可视化领域的发展有着冲击性作用。这一时期,针对电子设备存储的学习行为数据进行可视化分析发展迅速,大多数是借助于在线学习平台存储的数据进行可视化分析。如何合理有效地利用学习者产生的海量学习行为数据成为了这一领域的研

究重点,而可视化能帮助人们获取到更多数据背后隐藏的或不明显的信息或发展趋势,教育数据可视化的应用是在线学习平台和可视化技术的结合,应用载体的典型代表为学习仪表盘,如 LOCO-Analyst、Course Signal、快乐学等,还有高校和社会机构开发的在线教育平台,如清华大学雨课堂、华中师范大学师大云课堂、网易云课堂、希沃教学等。近些年来也涌现出了更多样的可视化工具,如数据可视化库类:Echarts、Antv、Leaflet、Polymaps 等,以及报表类的,如百度图说、Tableau,可视化大屏类阿里 dataV、FineReport,还有专业编程类如 R 语言和 python,均可用于教育数据可视化。

三、热点分析

(一)数据与方法

为了把握教育领域数据可视化不同时期的特点,本文采用文献计量学方法,使用 citespace 软件对数据可视化领域内 web of science 核心期刊进行文献计量分析,以“educational Data Visualization、data visualization of education、educational visualization”等为关键词进行检索,分别检索出 240、576、653 条记录,时间为 1985 年至 2021 年。对这些文献进行共被引分析,发现其中起关键作用的文献,呈现出教育领域可视化不同时期研究主题演变。经过浏览文章各级标题和摘要,剔除相关度不大的文献,由于文献大部分集中在近十年,所以本文选取了 2010 到 2021 十一年间 653 篇 WOS 核心文献,导入到 citespace 中进行分析,以一年为时间切片,选取每个切片的前 80% 的文献进行聚类分析。

(二)结果与分析

1. 研究热点

教育领域可视化研究集中于 E-learning,比如学习仪表盘(learning dashboard)、报告系统(reporting system)等,数据可视化前提是大数据,借助于电子设备和数据库可以方便收集用户数据。2012 年维克托·迈尔-舍恩伯格的《大数据时代》前瞻性地指出,大数据带来的信息风暴正在变革我们的生活、工作和思维,大数据开启了一次重大的时代转型^[19]。2013 年,数据可视化的典型应用——学习仪表盘蓬勃发展,对于教育领域的海量数据,利用数据挖掘和数据可视化技术,给学生的学习、教师的教学以及教育管理者的工作带来极大便利。

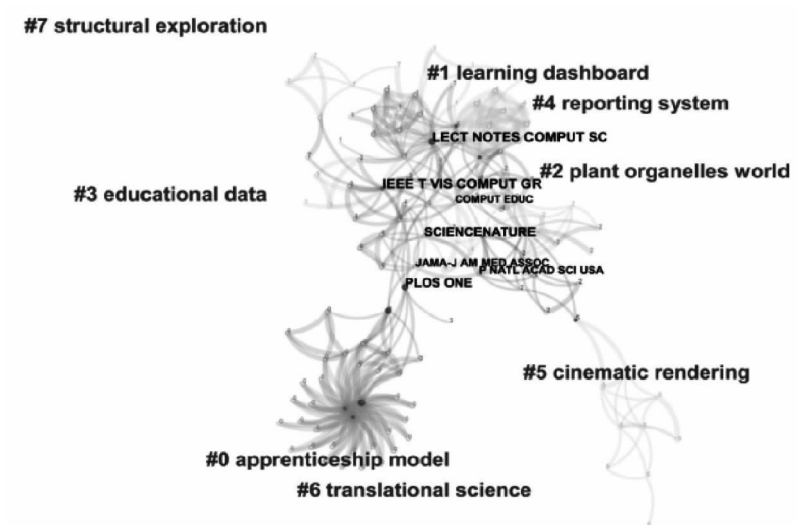


图 1 热点聚类图

2. 相关热点趋势

用 citespace 对教育领域内有关数据可视化文献进行关键词共现分析,文章选取 2010 年到 2021 年的文献,以 1 年为时间切片。分析发现,这段时间教育领域内数据可视化研究的热点词汇有可视化(visualization)、学习分析(learning analytics)、教育(education)、信息可视化(information visualization)、医学教育(medical education)和教育数据挖掘(educational data mining),其次还有模型(model)、学生(student)、3D 可视化、工具(tool)等。并且,用于教育领域的可视化技术主要是进行学习分析,对信息进行可视化等。

表 1 关键词频率

Keywords	Year	centrality	count
visualization	2010	0.39	15
education	2016	0.05	8
design	2017	0.27	6
medical education	2012	0.21	6
model	2010	0.02	6
information visualization	2017	0.01	6
learning analytics	2015	0.19	6

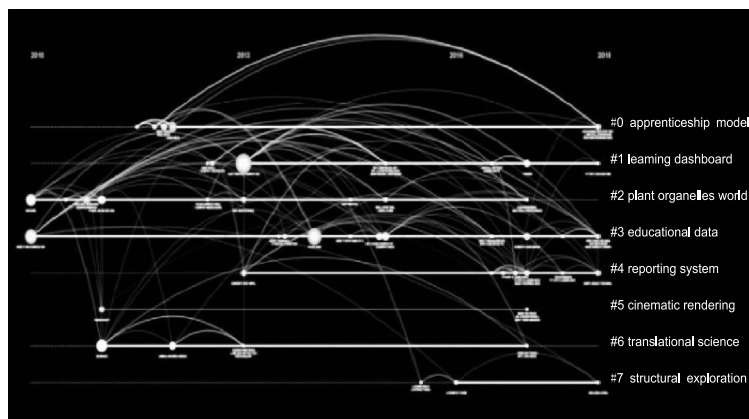


图 2 时间线图

由图中可以看出教育领域的可视化研究主题主要是有关 E-learning 的研究,比如学习仪表盘(learning dashboard)、报告系统(reporting system)等。这种结果是意料之中的,数据可视化的前提是要有大量的数据,而借助于电子设备和数据库可以方便地收集用户的数据,并进行在线分析和可视化。研究热点之一的学习仪表盘可以说是进行学习分析的有效的一种现代电子设备,近些年来关于在线学习系统包括学习仪表盘的发展速度也很快。将上面的聚类结果以时间线模式(Timeline View)呈现,可以从时间维度探究教育可视化领域研究主题的演变过程。如上图所示。通过解析时间线图可以发现,早于 2010 年之前就已经有学者开始研究植物细胞器和教育数据相关的主题了,这表明可视化早期较多地应用于生物学或医学领域,例如用于植物细胞结构的可视化,同时由于电子化设备的应用,已经形成了教育数据的概念。2012 年维克托·迈尔-舍恩伯格的《大数据时代》一书的问世,引起了业界人士对大数据的重视,维克托·迈尔-舍恩伯格在书中前瞻性地指出,大数据带来的信息风暴正在变革我们的生活、工作和思维,大数据开启了一次重大的时

代转型,事实也确实如此。在随后的2013年,有关数据可视化的典型应用——学习仪表盘的研究蓬勃发展,这也正是本文的研究重点。大数据的应用发展,可以给人类的生活带来前所未有的量化维度,对于教育领域的海量数据,利用数据挖掘和数据可视化技术,可以给学生的学习、教师的教学以及教育管理者的工作带来极大的便利。将聚类结果以内容类别呈现,可以看出教育数据涉及众多层面,主要分为三类:对学生学习相关的教育知识数据可视化、对师生行为相关的教育数据可视化、以及对知识结构与主体行为融合的教育数据可视化。对学生学习相关的教育知识数据可视化,可以向学生展示复杂知识的内在特征与发展态势,还可以结合深度学习等手段发掘其背后的潜在价值,即可以通过将分散的知识进行系统化的处理,帮助他们用结构化的逻辑从整体到部分分层思考。所以,对教育数据可视化分析可以培养学生的元认知策略和思维能力。对师生行为相关的教育数据可视化,比如用图表展现师生课堂行为分布、学生课堂行为表现以及学生成绩变化,可以帮助教师快速清晰地了解学生知识掌握情况、学习变化趋势等,同时还可以让教师反思其教学设计,并据此对教学工作做出改进和优化。所以,对教育数据可视化分析可以帮助教师进行深度反思和教学问题改进。对知识结构与主体行为融合的教育数据可视化,可以帮助学校教学管理者查看学生学习状况和教师教学水平,促进教师对教学工作的优化和改进,从而提升学生学业表现与就业技能,同时促进教师的职业发展,并提升院校教学管理效能。而对全国教育数据进行可视化,可以帮助教育决策者快速聚焦于某一领域的突出问题,经过及时讨论和协商,制定相关教育决策,提高执行效率。

四、教育数据可视化应用载体——学习仪表盘

学习仪表盘也称数字化学习仪表盘,是一类智能化、集成化的典型在线学习系统,其集成了可视化技术,实现在线数据的教育价值。在线学习容易实现师生之间一对一交流,充分尊重学生个性,激发学习动机,学生可以不受时间、地点和空间的限制随时与老师进行互动。在线学习同样也存在很多缺陷,比如缺乏人性化的沟通。网络人为地拉大了人与人之间的距离,同时阻碍了直接情感交流,在线学习的师生、生生之间缺乏及时有效的情感交流。普通的在线学习平台中存在缺少“人性交互”的问题,学习仪表盘利用数据可视化技术可以打开学习者学习过程中的“黑匣子”,有利于教与学,同时能促进自我意识、自我反思、自我行为监控和学情追踪,培养和发展学生的高阶思维能力,进而帮助教师优化和设计学习过程^[20]。

学习者经过意识、自我反思到意义构建,然后影响自身学习,形成一种意识后又进入下一轮循环,学习仪表盘构建的正是这样一种反馈型学习机制。大数据的价值被发现后,教育大数据的价值也被逐渐挖掘。张悦在其文章中指出现代课堂的重要标志就是用数据来读懂学生^[21],而学习仪表盘就是利用在线课堂生成的教育大数据来读懂每个学生,实现个性化的教学。学习仪表盘的数据可视化过程分为数据收集、数据分析处理以及数据的可视化呈现。一般来说,学习仪表盘针对的对象分为学生、教师以及学生和教师。

第一类服务对象主要为学生,比如Khan Academy。这类学习仪表盘的数据来源主要有登录时间、频次、练习和测试、学习进度。Khan Academy于2013年9月推出数学课程的学习仪表盘,记录每个学生看视频的时间,花费时间比较多的单元章节和做错的题目等数据,并对这些数据进行分析,作为补充教学依据。该学习仪表盘把教学内容精细地切割为上百个知识点,用小格组成“任务

进度”图,每个小格代表一个知识点,并链接到相应的学习任务,用颜色深浅表示学习者对该知识点的掌握程度,学习者可以自行选择符合个人学习风格的学习路径和知识点^[22]。

Duolingo 学习仪表盘用树图形象化地表现学习者学习进度,以及学习者当前已有的知识构架^[23]。用树图呈现出当前的学科体系及学习任务,展示学习者的个人知识架构,提高学习者的元学习能力。此外,Duolingo 还融入了各种游戏体验,借此激发学习者的学习动力,让学习者能够坚持学习,实现教育一直追求的快乐学习。

此类学习分析仪表盘的功能大多为学生定制,登录时间和登录频次一般用条形图或折线图,直观地展示学习者每天的学习时间,一段时间内的学习时间波动趋势。学习进度用于显示个人的学习效率,练习和测试用于帮助学习者测试学习效果等。

第二类服务对象主要为教师,学习仪表盘有 Moodle、Teacher ADVisor 等等。其中 Moodle 是一个开源课程管理系统(CMS),又称为学习管理系统(LMS)。教师页面在系统顶部显示有关课程的一些数据:课程的缩写和全名,学生人数等的名称,以及他们的电话号码和电子邮件。左上方最明显的区域以图表显示出勤率和提交的作业情况。在右上方,一个大区域用于详细展示每个学生当前状态和未来预测情况。出勤率、作业成绩、测验、期中考试和当前最终成绩的数据均用图表和数字可视化呈现。教师可以在警报图标的帮助下发现需要注意的学生,比如作业成绩较低或者考勤稀少的情况^[24]。

Class View 学习仪表盘主要由六个主要信息区组成:学生徽章矩阵、随时间变化的活动/徽章、活动清单、活动详细信息字段、徽章概述和过滤区域。学生徽章矩阵横轴上显示学生姓名,纵轴显示授予学生的徽章次数,圆圈越大表示徽章次数越多。随时间变化的活动视图包含五个图表,显示学生的总活动量、博客文章、博客评论和推文,此区域内的条形图还可以互动,单击可更新活动列表。活动详细信息字段显示连接到活动的内容,徽章概述则是获奖徽章的另一种可视化,便于学生或小组比较。过滤区域中有多个过滤选项,可以从列表中选择学生并分配 Set 1(蓝色)和 Set 2(红色),其中时间滑块允许用户修改仪表板上显示数据的时间范围^[25]。

Course Vis 是一种从课程管理系统获取跟踪数据并将数据转换为便于处理的形式工具。CourseVis 系统生成便于课程教师探索的图形图表,以检查远程学生的社交、认知和行为等方面,通过处理和图像化表示 CMS 收集的数据来帮助教师了解学习者的需求并有效地管理远程课程。CourseVis 收集在线远程学习课程的学生行为记录,包括所访问页面的历史记录、学生每次测验收到的信息以及发布到论坛的消息等。教师可以通过查看由学生的姓名和课程日期组成的简单矩阵了解全球学生对课程的访问情况,矩阵明确显示出了模式和趋势,以帮助教师了解远程学生的具体学习行为。CourseVis 在远程学习中可能存在的一种作用是帮助教师识别需要注意的个体,发现距离群体的模式和趋势,并反思他们自身的教学实践^[26]。

这类教师用仪表盘的数据来源有文档和工具使用情况、社会交互、产生式资源、学生的活动情况、肢体动作、登陆次数、在线时间、练习和测试。教师通过登录时间和频次查看个别学生的学习时间情况,还可以矩阵散点图了解整个班级的学习情况。通过学生的活动和社会交互情况了解学生之间的人际交流,通过练习和测试的结果确定学生的学习效果,肢体动作则可用于分析学习者的心理活动,学习实际掌握情况等。这些数据分别以表格、条形图、社会关系网图、表矩阵、指针图等来可视化显示。

第三类为教师和学生共同服务的学习仪表盘,通过不同的登录入口或设置权限来区分教师和学生,比如快乐学、SAM等。快乐学旨在帮助老师提高教学效率,同时帮助学生诊断学习难点,快速提升学习成绩并增强学习者的自信心。快乐学包括四个模块:知识模型、优质内容、智能引擎和知识足迹。智能引擎是整个平台的核心,它从学习要点、学习方法和学习能力等多个维度挖掘与学习需求相匹配的学习资源^[27],给学生推荐个性化的学习资源,鼓励学生自主学习,逐个排除学习难点^[28]。当学生学习完某个章节后,知识足迹模块将从题库里挑选出同类型的优质题目推送给学生,以检验学生对知识点的掌握情况。待学生练习完后,再将他们的练习结果与学科知识模型进行比对,便可映射出学生的知识足迹,从而找出学生自身知识结构中存在的薄弱环节,及时督促用户进行自我诊断练习。系统还可以生成个性化的学习方案,学生通过进行针对性的练习能将单位时间内的学习效率最大化^[29]。快乐学为不同的需求提供了不同的入口,比如学生入口的页面主要显示学生在练习过程中的错题类型与数量,系统可据此生成针对个人盲点的错题本,帮助学生进一步强化和提高。教师页面可支持师生间一对一地进行交互,帮助实现个性化教学。家长进入系统后可看到学生的总体学习指数,页面根据学生具体练习情况显示关键知识点的掌握情况,基于此家长能够有效地指导学生的学习。

SAM(Student Activity Monitor)学生活动监测系统主要用于支持学习者的自我监控和教师的教学意识,该系统收集四个方面的数据:课程花费的总时间,平均浏览文档的时间,下载使用的文件数量和学生工作的平均时间^[30]。用条形图、统计图、折线图 etc 来可视化相关数据。

这类仪表盘的数据来源就包括学生用仪表盘和教师用仪表盘的所有数据,数据收集的重点是文档与工具的使用情况和产生式资源,比如某种功能的使用频率,文档的下载情况和浏览时间等。这类数据的可视化方式也是一些基本的图表:柱状图、饼状图、标签云等,此外还有时间线图。

本文收集了23个学习仪表盘,它们的服务对象及可视化方法多种多样,其中有10(43.49%)个主要服务教师,8(34.78%)个主要服务于学生,还有5(21.74%)个同时为教师和学生服务。收集的数据来源一般都有登录次数、登录时间、浏览时间、文档下载和使用情况等,不同的仪表盘还有一些各自特别的数据,有的系统还会收集社会交互情况、产生式资源、肢体动作、面部表情等,用于分析学习者的心理变化情况等。不同的在线系统收集的信息维度和数据分析的深度也不同,一般来说都会收集登录次数、在线时间等一些基础元信息,相对应的就用一些条形图、折线图基本图表来可视化登录频率、在线时间等,其中一些仪表盘会基于收集的社会交互、眼动追踪信息等来发掘学习者的兴趣爱好、个人偏好、性格特点等深层次的信息。通常用一些图形和图表来可视化学习者信息,一般有条形图、折线图、饼状图等,特别的还有信号灯、雷达图、2D3D立体图、气泡散点图、热网格显示、甘特图以及图表的组合使用等,还有用各种植物图隐喻可视化,比如用一盆植株的生长来意象化学习进度。部分学习仪表盘嵌入了游戏元素,通过获取礼物之类的方法来激励学习者学习,以保持学习的动力,如可汗学院和Duolingo多国语言平台在学习过程中嵌入了很多游戏人物和游戏术语,使学习者在学习时也能感受到玩游戏的兴趣。不同的图形图表起到的可视化效果也不尽相同,具体要依据数据的类型、学习者或教师的个人风格、要达到的目的或者要凸显的重点而言。关于这一点,相关教育行业可制定基于在线学习系统的数据标准、可视化的标准规范等,类似于工业标准,这样教育数据可视化就有了参照模板,而不再是一种数据对应着合适或不合适的各式各样的图表。其中,可视化的标准规范可以根据数据类型和维度或者用户的需求来制定。

表 2 不同图表的特点^[31]

图表类型	适用数据及呈现特点
折线图	观察数据的变化趋势、走向
曲线图	直观显示数据变化情况
直方图	显示各组频数分布情况,显示各组之间频数的差别
柱状图	比较各组数据
散点图	适用连续规模的数据,观察变量之间的相关性及群聚效应
饼状图	表示部分占总体的比例
3D图	展示高维数据,视觉表现力强
社会关系图	查看用户的社交网络情况
矩阵表	用于多因素综合思考,探索问题
雷达图	体现各项指标的比率情况以及变动情形和好坏趋向
热图	用特殊高亮的形式将用户在网站上所点击的位置通过热图进行显示。聚集的点击越多,显示的颜色越热。可直观反映用户的探索需求或进入下一步的动机。通过各元素的点击热度可快速了解页面上吸引用户互动的内容
标签云	表示内容的权重

学习仪表盘能追踪和记录用户各个维度的数据信息,并以可视化的方式呈现出来,具有良好的视觉效果。基于对数据的分析和可视化,能支持多对象学习或工作。

于教师而言,学习仪表盘为教师提供了一个全面监测学习者行为的平台。教师通过这个平台可以了解到学生各个维度的数据,如学生的登录次数、在线时间、文档的下载与使用情况、论坛讨论情况、练习测试效果、个人进展、学习风格、情感体验等。何克抗教授的“大数据面面观”一文中提到:大数据背景下可以建立起前所未有的学习预警系统^[32]。通过这些信息,教师可以很容易地识别出处于危险状态的学生,这些信息就起到了预警的作用,教师及时给予有效干预,对学生进行个人指导和情感关怀,就能防止放弃学业,降低辍学率。学习仪表盘还可帮助教师进行个性化地教学,实现因材施教。以往教师都是基于自身经验来判断学生,对于刚上岗的新教师来说则是无处了解,而现在则可以借助学习仪表盘转换为对数据的分析。梁文鑫等人^[33]就提出过类似的观点:大数据促使教师的教学从经验走向对海量数据的分析,从群体教育转向对学习者的个性化教育。

于学习者而言,学习仪表盘为学习者提供了一个自我监控的学习平台,学习者通过这个平台记录的相关数据能更好地感知自己的学习行为,有利于学习者本身进行自我监督学习,通过可视化的数据呈现,学习者还可以轻松得知自己的当前学习状态,和班上其他人的差距,自己的学习风格,找到合适的学习资源和学习方法,针对性地练习自己的薄弱环节,培养自己的元学习能力,进而养成良好的学习习惯。平台根据需要可推送相关的学习资源,适合的学习路径来促进学习的提高。用户还可以方便地与其他人进行在线沟通,提高社交能力,发现学习的快乐,有目标有计划地自我调节学习,自我认知,增强约束力,消除孤立感。

对于家长来说,学习仪表盘提供了一个监管孩子学习的平台,家长通过这个平台可全面了解孩子在学校的学习情况,不必向以前一样需要频繁去学校问老师,了解自己孩子的学习方式学习风格以及学习习惯等,这样就可以在合适的时机实施有效干预,购买合适的学习资料,还可以了解到班上其他学生的学习情况,便于与自己的孩子比较分析,更好地帮助孩子学习。

对于研究者而言,以往的研究与实践往往是分开的,通过学习仪表盘的使用,研究者可以结合理论研究和教学实践,深入了解学生的实际学习情况,这样便于发现教学中存在的问题以便及时改善系统,修正理论上的错误,同时产出相关科研成果,有利于更进一步的深入研究。

对于教育管理者,他们一般关注的是宏观层面的数据,比如整个年级、整个学校以及整个地区的教学质量,通过区域间的教学质量比较,有助于合理分配教学资源,平衡教学优质资源分布,实现教育公平,整体提升教学质量,基于现有的情况设置课程安排,提供决策参考。决策过程过于依赖经验、缺乏数据的支撑,一直是教育界中颇受诟病的不足之处。在决策领域,大数据时代电子数据以海量、多样、高频的方式爆炸性增长,这就为决策者追踪事物的发展过程并调整现行政策提供了可能性,而不再是之前那样仅仅以经验来判断。基于数据的教育决策还可以帮助学校、地区精确地制定发展计划,提供改进的途径^[34],并通过数据的应用构建一种良性循环的校园文化^[35]。

对于开发方,学习仪表盘收集的反馈情况可有助于他们了解用户的需求,发现系统的漏洞以及不足之处,以便不断改善系统性能优化系统,扩大应用市场,开发出符合实际需求的系统。

综上所述,借助于仪表盘载体的教育数据可视化应用对各方都有益处,但同时也不可避免地存在一些弊端。其中最突出的问题就是数据隐私保护,经常有新闻报道某某学校学生数据泄露,家长频繁接到推销电话,不堪其扰。其次,学生性格迥异,统一的可视化方法并不是所有人都受用,在设计可视化的同时应注意到个性化问题。数据有不同的来源、类型和维度,要想做到对学生情况的全面了解,在线学习平台应该不仅仅是收集平台内部数据,还应融合外部数据,如面部表情、脑电之类的生物数据。针对不同类型不同维度的数据,应该用多种可视化方法去解读,给用户可选的权利,提供更人性化的服务。此外,随着数据量的不断增长,未来对平台及可视化技术的要求会越来越高,科技要做到与时俱进。

表3 多对象支持

受益对象	相关具体表现
学生	培养元学习能力、促进自我监督、自我反思、自我调节学习、提高效率、参与度、成就感
教师	因材施教、制定教学方法、有效干预危险学生、方便和外界交流
家长	同步监测、管理学习者、提供有效实时帮助与指导
教育管理者	有效决策、分配教育资源、不同学科和地区的教育水平对比
科研工作者	研究与实践结合、改进系统设计、深入研究教与学
开发方	了解用户的需求,不断改进、提升系统质量,扩大市场

五、教育数据可视化的未来发展

教育大数据的可视化涉及到方方面面,尤其是作为典型应用的在线学习系统,涉及到数据的来源收集、在线平台的构建、可视化技术、学习者隐私等一系列问题。

(一)数据隐私保护

教育大数据对教育发展有着潜在价值的同时,也存在潜在的隐私风险。教育大数据的应用需要注意风险预防,培养学生的数据隐私保护意识。政府方面可以订制相关法律法规来约束教育大数据的应用,形成学生数据隐私的立体保护体系,力求在大数据运用和隐私保护这两点间寻求一个平衡点。美国政府针对此类学生数据隐私侵犯问题,通过了立法和政策引导以及行业协会的自律性政策,基本构建起大数据时代的数据隐私保护法规体系,运作方面也形成了数据隐私治理体

系^[36]。除了数据隐私风险外,大数据应用存在的风险在人工智能技术自身、利益相关主体的问责等问题上面临数据与算法的缺陷,影响教育大数据潜在价值的挖掘,这些问题都需要教育者正视并注重^[37]。Johnson 于 2014 年阐述了大数据技术在高等教育领域存在的伦理道德问题和社会挑战,并提出应用大数据技术时应保持审慎的态度,这样才能保证数据挖掘的科学性和规范性。

(二)关注个人学习风格

世界上并没有放之四海皆准的道理,数据可视化也同样如此。并非所有的数据可视化都面向同样的用户,因此把握好服务对象非常关键。不同学习者的学习风格是各不相同的,对于学习积极性高的学生不需要系统给予太多提示和干预甚至是入侵式地提醒,这类学习者一般都会主动完成自己的学习任务,积极探索自己的需求,有计划地追求高目标高质量的学习。如果系统提供太多干预可能会引起他们的反感或不适,特别是当系统给出的提示不符合学习者个人的学习习惯时,更会引起学习者的不满。而对于学习主动性低的学生,系统的提示才有可能起到积极的正面效果。东北师范大学的姜强教授在其“基于大数据的学习分析仪表盘研究”一文的研究中发现,学习仪表盘对成绩弱一些的学生作用效果更好。这一研究结果证实了在线学习系统对于动机性不强的学生作用效果会更好一些。对于这一点,系统应该有自适应的学习功能。如果系统检测到学习者经常登录网站学习,做一些练习、测验,并积极主动地下载浏览学习资源时,系统应该降低干预度,反之则应该提高系统的干预度。其次,不同的学习者有着不同的视觉偏好,同一种数据,有的人可能喜欢柱状图表示,有的人可能对直方图比较受用,系统对于同一种数据应尽可能用多种合适的可视化方法展示,供用户自行选择。只有充分考虑到学生的个人性格特征,系统才能做到智能化地因材施教,最大化地贴近用户的需求。关于学生个人数据对外是否公开的问题,这个要看学生自己的意愿,优秀的学生或许更愿意公开自己的学习情况,但学习较差的学生却不一定想让别人知道的情况,和同伴进行比较也是能够促进学习者学习的,但比较却不一定暴露,关于这一点系统应该要给与个人选择的权利,个人可以有选择地公开个人数据。其次,用户界面的颜色和字体可能会影响到呈现的视觉效果,因此仪表盘在可视化时可以对视觉表征方面进一步细化。例如用户可以根据自己的偏好来调整视觉表示,这需要仪表盘具有用户控件的高度交互功能,允许可视化的个性化,用户可调节字体、颜色、饱和度、亮度、大小、标签、纹理等属性。

(三)数据可视化发展方向

大数据时代数据日益丰富,数据量更是以指数级增长,如此数据分析的速率必须能够跟上数据的增长速率才能适应未来的发展需求。新型的数据可视化技术必须满足互联网爆发的大数据需求,应该能够快速收集、筛选并分析和总结出决策者所需要的信息,还能根据新增的数据进行实时更新,这样才能满足用户的需求。未来的数据可视化应该朝着高度交互式的动态可视化发展,交互的意思是用户可以通过硬件设备如鼠标或键盘和计算机进行互动,动态就是用户与计算机进行交互后能实现页面的快速刷新,比如用户通过改变数值来预测未来某一领域的趋势情况,系统应该能快速把新的数据作为输入显示出变化后的实时的可视化数据图形或图表。这个对计算机的性能要求会比较高,不过随着科技的进步相信未来都能够实现。

(四)关注用户反馈情况

一般研究只关注如何对数据进行可视化,如何展现更好的可视化效果问题。这个研究层次仅仅停留在怎样去做的点上,很少有追踪用户使用系统后的反响情况如何,学生是否如我们预期的那

样感受到了可视化的效果,教师又是否有感觉到可视化技术给教学带来的益处,以及不同数据可视化方法带来的效果之间的差别问题,也就是探索可视化的一般规律或规范问题,探讨不同类型的数据对应的可视化方法以及不同性格的人对图形视觉的偏好问题,这些都是未来的在线平台进行可视化时需要注意的问题。如能发现可视化的一般规律以及不同性格的人的可视化效果偏好问题,这将有助于未来系统的数据可视化工作做得地更好,更贴近于人性化的服务,如此用户的体验感会更好。系统可以内置问卷调查模块,用于收集用户对系统的看法和建议,基于问卷数据分析用户的偏好,有利于系统的进一步完善。

(五)关注学生线下数据

目前的教育数据可视化的数据来源一般都是学生的在线学习数据,是基于电子设备存储的静态数据,这种数据方便收集,也方便利用计算机进行可视化分析。但是学习更多的发生在线下环境中,放眼中小学校,学生大部分时间是进行线下学习如坐在教室里听课,线下学习则包含了学生的面部表情、肢体动作以及学习环境等数据信息。这些数据也能反映或影响学生的学习情况,在仪表盘的背景下,如若能把这些线下数据和线上数据集成起来,线上线下结合分析学习者的学习情况,增强仪表盘反馈的效度,则可以更准确地捕获学习者的当前状况,从而更好地帮助学习者提升学习成绩^[38]。比如近期一些研究显示,人们对探索从各种可穿戴传感器和音频视频流收集的生物反馈的多模态数据越来越感兴趣^[39]。国内已有试点学校用摄像头来捕捉学生的面部表情,分析学生的上课认真程度,但由于各种因素,现今还未能推广开来。因此,未来的学习分析仪表盘需要具备可扩展性功能,收集从各种来源、各种数据格式的大量实时的学习者数据(如可穿戴传感器的数据),并对学习情况进行综合分析。

六、总 结

数据的采集、提取和理解是人类感知和认识世界的基本途径之一,数据可视化为人类洞察数据的内涵、理解数据蕴藏的规律提供了重要的手段。随着数据时代的来临,大数据的分析、挖掘与可视化已经成为信息技术发展的迫切需求^[40]。自大数据时代的兴起以来,用在线学习平台来辅助教育教学已经成为一种趋势。利用在线系统和可视化技术,可真正实现因材施教,而这是教育领域一直存在的一个痛点问题,在线学习平台及可视化技术的出现可在一定程度上缓解这个问题。虽然目前国内的知名在线学习系统数目不多,但随着 Khan Academy、雨课堂等众多在线学习平台的兴起,国内将会出现更多的在线学习平台,用以辅助教育教学,引领教育发展的潮流。本文大致梳理了数据可视化的发展历史和来源,重点介绍了教育数据可视化的研究热点即国内外二十多种学习仪表盘,从数据的选取、数据处理、可视化方法、受益对象多角度分析在线系统的数据可视化,最后提出了在线学习系统可视化过程中存在的问题以及发展趋势。大数据时代的到来使得教师的教学有迹可循、有据可依,是未来的教育基础,相信今后数据可视化技术在教育领域的应用将会逐渐深入并逐步扩大,帮助实现教育质量质的飞跃。

[参 考 文 献]

[1] 王兆鹏. 文学研究的可视化趋势[N]. 社会科学报, 2021-03-11(005).

[2] Ward, M., Grinstein, G., & KeimD. (2010). Interactive data visualization: Foundations, techniques,

- and applications[M]. Florida: CRC Press:36.
- [3] 李芒,蔡旻君,蒋科蔚,等.2013.可视化教学设计方法与应用[J].电化教育研究(03):6—22.
- [4] 陈为,张嵩,鲁爱东.2013.数据可视化的基本原理与方法[M].北京:科学出版社.
- [5] 刘海,李姣姣,张维,张昭理,易泽顺.面向在线教学平台的数据可视化方法及应用[J].中国远程教育,2018(01):37—44.
- [6] 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,2015,36(09):54—61+69.
- [7] Friendly M. A brief history of data visualization[M]//Handbook of data visualization. Springer, Berlin, Heidelberg, 2008:15—56.
- [8] Friendly M. A. -M. Guerry's "Moral Statistics of France": Challenges for Multivariable Spatial Analysis[J]. Statistical Science, 2007,22(3):368—399.
- [9] Quetelet A. Recherches sur le Penchant au Crime aux Différens Ages[M]. Hayez, 1833.
- [10] Palsky G. Des chiffres et des cartes: naissance et développement de la cartographie quantitative française au XIXe siècle[M]. Comité des travaux historiques et scientifiques-CTHS, 1996.
- [11] Peddle J B. The construction of graphical charts[M]. McGraw-Hill Book Company, Incorporated, 1919.
- [12] Brodetsky S. The Use of Graphs in Commerce and Industry[C]// Mechanical Engineering Congress, 2013.
- [13] von Huhn R. Charts and Graphs. An Introduction to Graphic Methods in the Control and Analysis of Statistics[J]. 1925.
- [14] Warne F J. Warne's Book of Charts, A Special Feature of Warne's Elementary Course in Chartography[J]. FJ Warne, Washington, DC, 1916,3.
- [15] Ayres L P. The war with Germany: a statistical summary[M]. US Government Printing Office, 1919.
- [16] Shewhart W A. Economic control of quality of manufactured product[M]. ASQ Quality Press, 1931.
- [17] Eells W C. The relative merits of circles and bars for representing component parts[J]. Journal of the American Statistical Association, 1926, 21(154):119—132.
- [18] Friendly M, Denis D J. Milestones in the history of thematic cartography, statistical graphics, and data visualization[J]. URL <http://www.datavis.ca/milestones>, 2001,32:13.
- [19] 维克托·迈尔-舍恩伯格,能思.库克耶.大数据时代生活、工作与思维的大变革[M].杭州:浙江人民出版社,2013.
- [20] 姜强,赵蔚,李勇帆,李松.基于大数据的学习分析仪表盘研究[J].中国电化教育,2017(01):112—120.
- [21] 张悦.让数据成为课堂的“学习仪表盘”[J].中小学信息技术教育,2018(05):32—35.
- [22] 张振虹,刘文,韩智.学习仪表盘:大数据时代的新型学习支持工具[J].现代远程教育研究,2014(03):100—107.
- [23] Munday P. The case for using DUOLINGO as part of the language classroom experience[J]. RIED: revista iberoamericana de educación a distancia, 2016,19(1):83—101.
- [24] Podgorelec V, Kuhar S. Taking advantage of education data: Advanced data analysis and reporting in virtual learning environments[J]. Elektronika ir Elektrotehnika, 2011,114(8):111—116.
- [25] Charleer S. Designing and Evaluating Student-facing Learning Dashboards: Lessons Learnt[J]. 2017.
- [26] Riccardo MAZZA V D. Course Vis: Externalising Student Information to Facilitate Instructors in Distance Learning[J]. Artificial Intelligence in Education: Shaping the Future of Learning Through Intelligent Technologies, 2003,97:279.
- [27] 李宝,张东红.教育大数据下个性化资源推送服务框架设计[J].中国远程教育,2017(09):62—69+80. DOI:10.13541/j.cnki.chinade.20170915.003.
- [28] 牟智佳,俞显.教育大数据背景下智能测评研究的现实审视与发展趋向[J].中国远程教育,2018(05):55—62. DOI:10.13541/j.cnki.chinade.20180514.006.
- [29] 蒋鑫,洪明.国际教育大数据研究的热点、前沿和趋势——基于 WOS 数据库的量化分析[J].中国远程教

- 育,2019(02):26—38. DOI:10.13541/j.cnki.chinade.20190127.009.
- [30] Duval E. Attention please!: learning analytics for visualization and recommendation[C]//Proceedings of the 1st international conference on learning analytics and knowledge. ACM, 2011:9—17.
- [31] 杨兵,卢国庆,曹树真,Tiong-Thye Goh. 在线学习系统数据可视化评价标准研究[J]. 中国远程教育,2017(12):54—61+80.
- [32] 何克抗. 大数据面面观[J]. 电化教育研究,2014,35(10):8—16+22.
- [33] 梁文鑫. 大数据时代——课堂教学将迎来真正的变革[J]. 北京教育学院学报,2013(01):14—16.
- [34] 冯仰存. 数据驱动的教师教学决策研究综述[J]. 中国远程教育,2020,41(04):65—75. DOI:10.13541/j.cnki.chinade.2020.04.009.
- [35] 顾小清,黄景碧等. 让数据说话:决策支持系统在教育中的应用[J]. 教育科学文摘,2010,16(16;5):79—80.
- [36] 王正青. 大数据时代美国学生数据隐私保护立法与治理体系[J]. 比较教育研究,2016,38(11):28—33.
- [37] 钱小龙,张奕潇,宋子昀,李强. 教育人工智能系统的伦理原则与困境突破[J]. 江南大学学报(人文社会科学),2021,20(06):96—104.
- [38] 吴南中,黄治虎,曾靓,谢青松,夏海鹰. 教育大数据生态圈构建:“3+3”模型的逻辑与实践[J]. 中国远程教育,2019(07):77—85. DOI:10.13541/j.cnki.chinade.20190226.005.
- [39] Sedrakyan G, Malmberg J, Verbert K, et al. Linking learning behavior analytics and learning science concepts: Designing a learning analytics dashboard for feedback to support learning regulation[J]. Computers in Human Behavior, 2018.
- [40] 陈为,沈则潜. 大数据丛书:数据可视化[M]. 北京:电子工业出版社,2013.

(责任编辑:闫卫平)

Visualization of Educational Data: Development History, Research Hotspots and Typical Applications

WANG Yi-fan^{1,2}

(1. School of Education, Central China Normal University, Hubei, Wuhan 430079;

2. Wuhan Academy of Educational Sciences, Hubei, Wuhan 430021)

Abstract: With the rise of the era of big data, there has been a wave of big data analysis in all fields, and the field of education is no exception. The rise of E-learning in recent years has made it possible to mass-store educational data, and at the same time it has formed educational big data. The visual analysis of educational big data can bring great convenience to people of different roles such as students, teachers and parents. In the field of education, there is huge potential value. For example, it can realize the teaching of students in accordance with their aptitude and achieve individuality. Based on the literature analysis method, this paper traces the historical origin of data visualization from ancient times to the present, analyzes the development trend and hot topic research of data visualization in the field of education, and focuses on the typical representative of data visualization — the application of learning instrument panel in education field, so as to provide a feasible reference for the development of educational big data visualization.

Key words: big data; visualization; learning analysis; online learning system