

学习分析 教育信息化的新浪潮*

吴永和^{1,2} 陈丹³ 马晓玲⁴ 曹盼³ 冯翔^{1,2} 祝智庭^{1,2}

- (1.华东师范大学 教育信息化系统工程研究中心;
- 2.上海数字化教育装备工程技术研究中心;
- 3.华东师范大学 教育信息技术学系;
- 4.华东师范大学 信息学系,上海 200062)

[摘要] 教育信息化经历了学习管理系统(LMS)以及 Web2.0 应用的变革。新技术的深入应用带来了教育“大数据”的高速增长,挖掘这些教育数据潜在价值的迫切需求,使得学习分析应运而生。通过文献分析法,对国内外学习分析文献进行了分析和综述,首先对学习分析进行了概念界定和历史溯源,比较了与学习分析相关概念的区别和联系,之后针对学习分析作为教育信息化新热点,对其研究、发展、技术策略等方面进行了较系统地阐释,最后总结了学习分析目前面临的挑战和愿景,以期可以对学习分析进行全方位的阐述和梳理,并促进该领域的深入研究。

[关键词] 学习分析;教育数据;数据挖掘;智慧学习

[中图分类号] G434 **[文献标识码]** A **[文章编号]** 1672—0008(2013)04—0011—09

一、前言

随着信息技术的不断发展,教育中技术的应用正在经历一次又一次的进步和发展。从教育信息化大规模应用层面来说,教育信息化经历了学习管理系统(LMS)以及 Web2.0 应用层面的变革,这些新技术的深入应用也带来了教育数据爆炸的担忧。学习管理系统中存储着大量学生学习数据,学生各种在线学习保留着大量学习轨迹信息,社交网络中的学习行为更是呈现出急剧增长的数据流趋势。这些“数据集过大,难以被典型数据库软件工具捕捉、储存、管理和分析”的教育类“大数据”,目前已成为困扰教育机构的一大难题,研究者们逐渐开始关注这些教育数据潜在的利用价值。学习分析作为分析技术在教育领域中的应用和发展,正逐渐受到研究者的关注和重视。

2012年3月,在教育部正式颁布的《教育信息化十年发展规划(2011—2020年)》^[1]中提出,实现教育信息化的方法之一是要推动信息技术与教育教学的“深度融合”,这一“深度融合”的概念在《教育信息化规划》中先后出现达10次以上。学习分析就是促进这种“深度融合”的有效抓手。与此同时,智慧教育呼声渐涨,各国都开始对其予以高度重视。智慧教育需以智慧学习环境为技术支撑、以智慧学习为根本基石,而智慧学习分析为智慧学习环境的构建和智慧学习的发展

提供基础,因此,智慧教育也呼唤着智慧学习分析的发展。^[2]

国外的研究相比于国内,对于学习分析的应用有着更广阔的预见性。2012年4月,Bienkowski等学者联合向美国教育部教育技术办公室提交了关于学习分析的研究报告,报告中对学习分析相关教育应用,作了详尽的论述和预见,提出学习分析在美国教育领域中的应用刻不容缓,并期望美国在国家层面制定出更有效的政策和手段来在教育中应用学习分析技术。^[3]

由此可见,不管是国内还是国外,学习分析都已经成为引领教育信息化发展和变革的又一新浪潮。在这样的背景之下,本文尝试运用文献分析法对学习分析进行综述和梳理。国外文献分析主要有四个文献来源:一是利用笔者在美国访学的密苏里大学图书馆检索系统(<http://mulibraries.missouri.edu/>)以“Learning Analytics(学习分析)”、“EDM(教育数据挖掘)”、“Decision Making(数据决策)”等相近关键词对论文、图书、电子书、DVD、以及论文题目、图书题目和杂志题目,进行文献搜索,其中 Learning Analytics 检索论文,在42个数据库中共检索到909条,其中来源学术杂志401条、杂志82条、毕业论文50篇、图书48条、传统出版34条。二是通过 Google 学术搜索以“学习分析(Learning Analytics)”、“教育数据挖掘(EDM)”、“数据决策(decision making)”等相近关键词进行文

* 基金项目:本文为2011年度国家社科基金重点项目“电子课本出版与生态发展研究”(编号:11AXW001)、上海市哲学社会科学规划一般课题“电子课本生态发展的研究”(编号:2012BTQ001)、2010年度教育部人文社会科学研究规划基金项目“社会计算环境下 e-Learning 教育应用创新研究”(编号:10YJA880148)和上海市科委重大任务科研专项“跨平台智能数字化教育服务平台”(编号:11dz1504400)的成果之一。

献搜索,在搜索到的上百篇文献中,根据引用频率进行筛选。三是对可获取到的学习分析领域著名国际会议 LAK11(学习分析与知识国际会议)中的 27 篇、LAK12 中的 50 篇和 LAK13 的 46 篇等会议论文,根据作者文章引用频率和领域知名度进行了筛选。四是根据连续三年的 NMC《地平线报告》中关于“学习分析”部分的推荐、扩展阅读材料和两届 LAK 在线课程推荐资源列表,进行充分的选读和筛选。根据以上四个途径选取的文献资源,除去相同文献之外,最终获得相关文献 70 多篇。其中国内相关文献主要来源是“中国知网”等学术文献库中相关文献。

二、分析与学习分析

(一)什么是分析与学习分析

数据驱动的社会充斥着大量以不同目的命名的“分析”,比如,以特定兴趣主题命名的健康分析、安全分析、地理空间分析等术语,以活动目的命名的描述分析、预测分析、规定分析,或是对特定物体的分析(Twitter 分析、Facebook 分析、Google 分析)等。^[4]Watson 认为,“分析”术语的形成主要关注四点隐喻:在哪使用、谁来执行、需要的技巧和涉及的技术。^[5]从这种观点出发,在教育领域中使用,由教育者、学习者、管理者等利益相关者执行,运用分析技术对与学习相关的数据进行收集、处理和过程就可以称为“学习分析”。

目前,学术界对于学习分析的定义并未有统一的描述,比较常用的几个定义分别来自 LAK11、EDUCAUSE“下一代学习挑战”、NMC《地平线报告》、学习分析领域专家 George Siemens 等。

从时间顺序来看,最早的学习分析定义源于 EDUCAUSE 的“下一代学习挑战”,其中将学习分析定义为“使用数据和模型预测学生收获和行为,具备处理这些信息的能力。”^[6]这一定义被认为是对现有教育系统的升级,而不是修改和重构。因此,2010 年 8 月 25 日,学习分析领域专家 Siemens 在博客上贴出了他的学习分析定义,即“学习分析使用智能数据、学习者产生的数据和分析模型来发掘信息和社会联系,以对学习进行预测和建议”,^[7]这一版本的定义成为了之后学者广泛讨论的基础,引发了学术界对学习分析定义的探讨和深究。在 2011 年 2 月的 LAK11 会议上,LAK 对学习分析做了较权威的定义,即“学习分析是以理解和优化学习及其发生的环境为目的,对学习及其所处情境的数据进行的测量、收集、分析和报告。”^[8]美国学习分析研究协会 SoLAR(Society for Learning Analytics Research)也使用这一定义。在《地平线报告 2011》中,NMC 也阐释了学习分析的定义,NMC 认为,学习分析是以评估学业成就、预测未来表现、发现潜在问题为目的的,对学生产生和收集的大量数据进行阐释的过程。目前,不管是从定义发布机构的权威性,还是从定义使用频率上看,大多数研究使用的是 LAK11 和 NMC 的定义。

这些学习分析的定义,给我们较为清晰地描绘了学习分析的轮廓。我们认为,学习分析的执行准则可以概括为“先理解,后行动”,这就意味着学习分析的过程应该是先对数据进行“理解”,之后采取一定的“行动”。Siemens 提出的学习分析

过程,也体现了这个准则,他认为,学习分析经历了先分析,后预测,最后进行自适应、个性化和干预的整个过程。^[9]学习分析中的分析可以体现为一个周期,即先通过学习轨迹、话语分析等进行课程级的处理,之后采用预测建模、模型挖掘等挖掘教育数据,通过语义来定义课程资源的开发,最后进行自适应内容和学习。^[10]

(二)学习分析溯源

早期学习分析,被认为前身是社会网络分析(Haythornthwaite,1998;Wasserman&Faust,1994)、潜在语义分析(Landauer, Foltz, & Laham,1998)以及性格分析(Deakin Crick, Broadfoot, & Claxton,2002)。其主要原因有两点:一是这些领域的研究者之后转向研究学习分析,另一个原因是他们的成果被学习分析研究者引用。^[11]

学习分析是在多种数据挖掘和分析技术基础上发展而来的新兴应用。与学习分析相近的领域主要有商业智能、网络分析、学术分析、教育数据挖掘、行为分析等。^[12]学习分析可说是这些领域进一步发展和融合的结果。在总结大量文献基础上,我们将学习分析技术相关领域发展时间线情况总结如下(图 1 所示)。其中领域时间界定一般为该术语明确出现或基本的起源时间,也有部分其广泛的应用时间。

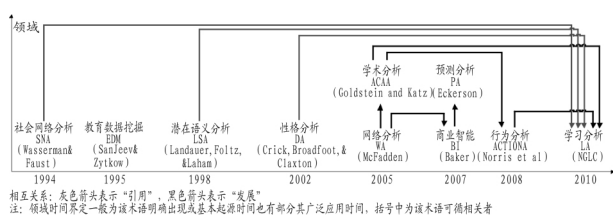


图 1 学习分析溯源时间线

数据挖掘应用于传统教育领域起源于 1995 年,之后随着学习管理系统的兴起和发展,教育数据挖掘在 21 世纪早期逐渐转向在线教育领域。^[13]近年来,对数据的分析在商业领域中再次兴起并逐渐发展起来,通过对网站访客和用户的相关数据进行收集、分析和报告的网络分析技术^[14]以及强调通过商业数据分析获取商业决策的商业智能^[15]成为了商业领域中数据分析的关键技术。这些技术的使用逐渐迁移到学习领域,于是产生了学术分析、预测分析、行为分析等分析技术。而学习分析通常被认为是从学术分析和行为分析分裂出来的技术,^[16]其概念的界定最早出现在 2010 年 EDUCAUSE 的“下一代学习挑战”中。可见,学习分析在经历了数据挖掘和分析技术的不断发展和演变后应运而生。

(三)学习分析相关概念的界定

学习分析被认为是数据分析技术在教育领域中的应用。从溯源学习分析可以看出,虽然学习分析的前身涉及到多种不同领域中的分析技术,但是,目前与之相近的在教育领域中的分析技术主要有教育数据挖掘、学术分析和行为分析三类。

1. 学习分析与教育数据挖掘

Bienkowski 等人对教育数据挖掘和学习分析做了详尽的区别。^[17]教育数据挖掘主要是寻找数据新模式、开发新算法和

新模型,而学习分析则是将已知预测模型应用到教学系统中。教育数据挖掘重点关注将学习分解成可被软件分析和影响的小块,使之可产生自动反应,适应于学生学习。而学习分析则是强调理解整个系统,产出的是如,适应性教学内容、危机学生干预和反馈等适应并支持决策。以下是学习分析与教育数据挖掘在历史来源、学科来源、目的等领域的区别比较(见表1)。

表1 学习分析与教育数据挖掘比较

项目	学习分析	教育数据挖掘
历史来源	学术分析、行为分析、预测分析	计算和心理学方法和研究
学科来源	信息科学、社会学、心理学、统计学、机器学习、数据挖掘	统计学、机器学习、数据挖掘
数据来源	教育管理和教学、教与学	教与学
目的	创建直接影响教育实践的应用	检测学习理论 报告教育实践
技术方法	预测、集群、关系挖掘、精化人类判断、用模型发现、社会网络分析、社会/“关注点”元数据	预测、集群、关系挖掘、精化人类判断、用模型发现

2. 学习分析与学术/行为分析

Long 和 Siemens 对学习分析和学术分析的区别进行了界定。学术分析侧重商业智能在教育领域中的应用,强调的是在机构、区域、国家/国际层面上的分析。而学习分析比学术分析更具体,主要关注点在学习过程,强调的是在课程、部门层面上的分析。^[18]因此,我们可以认为,学习分析更多的是面向个体和底层学习层级的,学术分析则主要关注的是顶层机构层面的分析和决策。高级别机构层面的政策决定,一般将学习分析和学术分析两者融合进行分析和使用。而行为分析的关注点也是机构层面,它可以认为是更综合的学术分析。^[19]因此,有些学者也将学术分析和行为分析视为同一范畴内的分析技术。

3. 学习分析相关概念粒度图示

Ferguson(2012)认为,当今学习面临着技术、教育、政治/经济三大挑战。而教育数据挖掘主要关注地是“如何从学习大数据集中提取有用信息”这一技术挑战,学习分析则主要关注“如何优化在线学习机会”这一教育挑战,而学术/行为分析则关注“如何在国家或国际层面上最大程度地改进学习机会和教育结果”这一政治/经济挑战。^[20]因此,我们发现,与学习分析相关的这几大概念,在分析面向对象粒度方面存在着



图2 学习分析相关概念粒度图示

较明显的区别。教育数据挖掘将学习分解成组块进行分析,学习分析则主要面向个体的学习,涉及部分机构层面内容;学术分析和行为分析都主要关注机构层面。因此,根据分析面向对象粒度的区别,我们可以将这几大学习分析相关概念总结如下(图2所示)。

三、学习分析成为教育信息化新浪潮

(一) 学习分析研究愈演愈热

美国新媒体联盟(NMC)连续三年发布的地平线报告(horizonreport)都将学习分析视为未来几年内影响学习科学技术发展的主要技术之一。在2010高教版《地平线报告》中,NMC第一次将学习分析写入,预测可视化数据分析(visual data analysis)将在未来4至5年内成为主流。^[21]而学习分析技术(learninganalytics)在2011年基教版和高教版的《地平线报告》中都被视为在未来4至5年内广泛采用。^{[22][23]}2012年的高教版报告,已将学习分析的采用预测提前到2至3年内。^[24]由此可见,对学习分析不断增长的关注度,以及各种先进分析技术的出现和整合,学习分析的预测采用期已从远期转为中期,足见其已成为技术促进学习领域发展较快的领域之一。2013年最新版本的《地平线报告》则是延续了2012年的预测,继续将学习分析视为2至3年内被采用的新技术。学习分析未能进入“1年内采用”分类的原因在于:一方面是由于学习分析发展迅速,另一方面也因为目前仍然无法被教育者完美掌控。^[25]

学习分析技术逐渐成为主流关注点,为了促进学术交流和学者之间的合作,2011年2月,学习分析与知识国际会议应运而生。学习分析与知识国际会议(International Conference on Learning Analytics & Knowledge, 简称为LAK)目前已召开三届,其中第三届于今年4月召开,主题为“学习分析辩证法”。LAK历年召开时间、地点和会议主题见表2。

表2 LAK 历年召开时间、地点和会议主题

LAK	会议召开时间	会议召开地点	会议主题
LAK11 ^[26]	2011年2月27日—3月1日	加拿大阿尔伯达省班芙市	学习分析技术在技术、社会和教学等维度的整合
LAK12 ^[27]	2012年4月29日—5月2日	加拿大温哥华市	学习分析技术、知识建模和表征、知识工作和分析等领域的现状及发展策略探索
LAK13 ^[28]	2013年4月8日—12日	比利时勒芬市	学习分析辩证法

今年召开的LAK13被认为是学习分析领域中至关重要的一次会议,LAK13强调了对学习和分析这两大领域的交叉空间的更多关注,并提出学习分析的研究趋势是将两者联系地更加紧密。与此同时,2013年也被认为是学习分析的关键年,从这一节点开始,学习分析作为一个新兴研究领域的身份逐渐稳固。^[29]

由于学习分析的迅速发展,使得一年一次的LAK会议难以解决学习分析领域中的诸多问题,因此,学习分析研究协会SoLAR(Society for Learning Analytics Research)启动了区域会议项目(SoLAR Flares)。SoLAR的第一次区域会议以“使用

http://dej.zjtvu.edu.cn

分析来创造实际问题的实际解决方案”为主题,已于2012年10月在普渡大学召开。11月又分别在英国的开放大学和澳大利亚的悉尼技术大学举行了两次区域会议,各国专家频繁开展研讨促进着学习分析不断发展。^[30]LAK和SoLAR所发布的大量学习分析相关报告和论文,目前很多都成为了领域中的引用较多、学术影响力较大的文献。

除LAK和SoLAR外,美国高校教育信息化协会EDUCAUSE、英国开放大学知识媒体研究所KMI等研究机构,也为学习分析的发展起到了推动作用。EDUCAUSE的学习创新计划(ELI)将学习分析作为其研究关注点,不仅周期性地发布研究报告和综述,还于2012年4月11日至12日针对学习分析开展了在线专题会议,^[31]会议召集了学习分析领域的专家针对学习分析中备受关注的意会及未来走向、隐私和伦理问题等主题展开了研讨和商议,这些研讨成果对于学习分析的进一步发展,起到了至关重要的推进作用。

但目前,国内学术界对于学习分析的研究仍不多,使用“学习分析”作为关键词在中国知网上进行检索,仅得到文献10多篇,分别来自中国学术期刊网络出版总库(9篇)和优秀硕士学位论文全文数据库(1篇)等。按照与学习分析相近的关键词“教育数据挖掘”、“学术分析”、“行为分析”来看国内学习分析领域的研究趋势(如图3所示),可以看出,教育数据挖掘和学术分析研究比较少,主要集中在2010年度;行为分析是研究较多的领域,而国内对于学习分析的研究从2006年才逐渐开始兴起,在2011年之后有明显增多的趋势。总体来看,我国学术界对于学习分析的研究还处于起步阶段,相关研究有待进一步挖掘和开展。

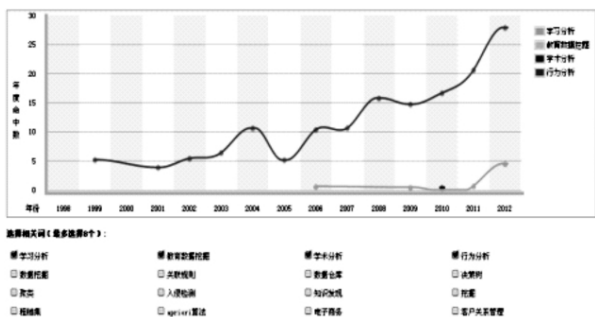


图3 学术分析相关论文年度发表情况

对国内外学习分析的文献主题分析可以看出,目前,国际上学习分析的研究方向主要集中于学习分析服务(包括开放数据集服务)框架、分析方法、工具与可视化工具领域;而国内目前主要集中于综述和应用分析。学习分析的研究可能聚焦在如下领域:(1)偏向于计算机科学与人工智能领域的数据挖掘、机器学习、知识工程;(2)侧重于教育技术领域的学习分析统计模型的设计;(3)侧重于人机交互与用户体验的学习分析结果可视化。

(二)学习分析的利益相关者

1.学习分析利益相关者组谱

利益相关者(stakeholder)是一个实体(人、组织等),与学习分析存在着一定的利益关系。利益相关者可以认为是学习

分析生态系统的受益者,也可以认为是学习分析的实施主体、面向用户等。

Brown认为,学习分析产生的报告对教学者、学生、管理者三方有着直接利益关系。^[32]教学者获得学生学习活动和进展的报告,学生获得自己学习进展的反馈,而管理者则了解到课程和学位完成情况的数据。这一观点与Long和Siemens是一致的,他们也将教师、学生和管理者/决策者视为三大利益相关者。有些学者也将研究者视为另一类利益相关者。^[33]我们认为,Brown提出的“教学者”不仅包括参与实际教学的教师,同时也包括对教学进行教育研究的教育研究者。因此,我们可以将学习分析中涉及的利益相关者分为三大类。(1)教学者:教师、教育研究者、教学设计者等;(2)学习者:学生;(3)管理者:政府、教育机构、学校管理部门等。

这三大类利益相关者与学习分析存在着直接的利益关系。在学习分析整个生态系统中,除了这三大利益相关者外,还涉及到与学习分析存在教育、金融、经济或道德等方面的间接利益关系,或者在系统中执行一个或多个功能的主体或机构。这些间接利益相关者包括:(1)家长:从家长角度对学生学习情况和相关信息进行了解和监管,适当进行学生干预;(2)企业:提供学习分析系统的软件和硬件的开发和管理的实体;(3)投资者(funder):这一利益相关者在海外较多出现,即对于一个项目和系统开展的投资实体。

这些利益相关者如图4所示。从图4中可以发现,学习分析三大利益相关者中学习者、教学者单个层面占的比重比“管理者”要高,主要原因是学习分析着重关注的是对个体层面学习的分析,机构层面的关注所占比重不如学术分析,因此在图示中,将学习者和教学者作为三大利益相关者中的主体部分。

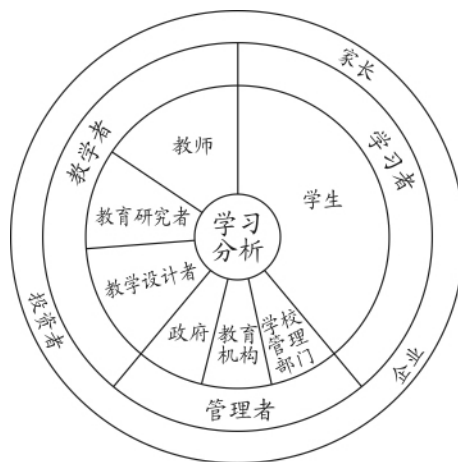


图4 学习分析利益相关者图示

2.学习分析服务于利益相关者

对学习分析利益相关者的梳理可以从整体上揭示学习分析系统应该和可以为利益相关者带来什么好处,即他们如何从学习分析中受益。根据学术层级来看,学习分析的三大利益相关者学习者、教学者、管理者三者之间又存在着金字塔式层级关系。^[34]学习者在金字塔最底端,学习分析的直接对

象为学习者,学习者层级的数据分析影响到上层的利益相关者。第二层为教学者,从下层学习者分析中获取信息以进行干预和计划。最上层为管理者,管理者从学习者和教学者中获得数据,并做高层决策。

(1)学习者。学习者作为学习分析获取数据的主要来源实体,同时也是学习分析的主要面向对象。学习者数据的来源不仅包括学习者的外显行为产生的数据,如,完成任务、参加考试,还包含了从学生内隐行为中收集到的数据,如在线社会交互、课外活动、讨论发帖等。^[35]

学习分析给学生提供实时行为和活动内容反馈,推荐社交网络信息等分析报告。学生可定制个性化和自适应学习过程和内容并获得及时推送,及时获得的反馈和建议不仅来源于教学者,还来源于与匿名同伴的比较。^[36]

(2)教学者。对于教师来说,追踪学生学业成就、理解怎样的教学实践更有效是学习分析的主要目标。教师不仅可以了解个体和整体学生学习进展情况,还可以获得包括学生的情绪状态、主题讨论参与程度和感兴趣程度等一系列活动状态在内的易理解的、可视化的报告。教师在对学生情况充分了解的基础上采取措施,可以对危机学生进行识别、警告和干预,也可以改进教学方案以便更好地开展教学。

对于教育研究者来说,学习分析将教育研究与真实学习数据贯一起来,从一定程度上解决了传统教育研究中仅凭经验进行质性研究的不足,给现有的一系列教学信仰和观点增添了科学性和实践性。^[37]从某种程度上说,教师也承担着教学设计者的任务,通过对学生的分析,决定如何为学生提供个性化资源和服务,如何选择合适而有效的学习模式等问题,都可以通过学习分析来解决。^[38]

(3)管理者。学习分析从一定程度上为管理者提供决策依据。从学校管理部门角度来说,通过对学习相关活动各种级别的追踪和分析,将会影响他们对于工具和学习空间资源的分配原则。他们获得的报告应该是高阶总结性的,课程完成度如何、资源如何分配、学校政策是否需要改变等问题都可以从中得出。^[39]政府和教育机构也在某种程度上从学习分析中受益,但是机构层面的分析一般结合学术分析和行动分析一起使用。

(三)学习分析的技术策略

1.学习分析技术手段

学习分析是通过技术、算法、教学理论等相互协作而完成的对数据的分析。^[40]其中,教学理论催生的学习系统体现了算法的有效性,而算法则需要深度技术知识的支撑。^[41]技术是学习分析的实现手段,它完成学习分析的整个数据分析和呈现过程。由上文表1我们可以发现,学习分析采用了一些技术手段,如,预测、集群、关系挖掘、社会网络分析等。学习分析采用的一些技术手段源于教育数据挖掘,并利用EDM的信息检索技术优势结合其他相关技术手段,如,社会网络分析、自然语言处理等,完成学习分析的整个过程。目前,常用的一些学习分析技术手段包括社会网络分析、话语分析、内容分析等。^{[42][43]}一些新的学习分析技术手段也在不断涌现,如,韩国檀国大学的教师研究出了一个多维分析法,它主要

是整合了常用学习分析技术手段,并结合在线交互可视化,从多个维度对学习数据进行分析。^[44]

与此同时,可视化数据分析技术也成为了学习分析的重要技术手段。作为2010年《地平线报告》推荐的主流趋势,可视化数据分析以其将高级计算方法和复杂图形引擎相融合,以呈现复杂图像的优越性,协助学习分析系统实现报告的可视化。可视化数据分析技术已发展的比较成熟,目前,完全可以实现变量的实时交互操作。作为学习分析报告呈现的必然趋势,可视化数据分析正向更好地呈现学习过程、社交进程等各种学习情况的方向发展。

Bienkowski等人将学习分析的应用领域分为用户建模(知识、行为、体验)、用户建档、领域建模(关键概念、知识组)、趋势分析等几类。^[45]这些不同领域学习分析的实现需要不同的数据来源和类型,同时也依靠不同的技术手段。例如,用户建档强调构建个性化和自适应的学习环境,就需要以用户分类技术和集群技术为依托进行用户基本信息、学习习惯、学习偏好等的分析。而在个性化和自适应过程中,对学习者的预测之后行为的趋势分析技术和序列分析技术也起到了至关重要的作用。

2.企业技术方案

学习分析技术层面的研究目前仍处于起步阶段,学习分析起源于商业分析手段,因此,一些企业对于大数据的技术解决方案,可以对学习分析层面的方案提供一定的参考。作为智慧教育推广的领先者,IBM公司的大数据分析平台强调对大量数据的快速、安全和可扩展的处理和分析,它的技术架构主要是四大组件的整合:智慧云平台分析组件、信息空间数据流计算组件、信息服务器组件和数据仓库组件。^[46]Oracle公司的技术方案定义了数据分析周期的四个阶段,即获取、组织、分析和决定,每一阶段由不同模块平稳协作完成。^[47]Intel提出的完整大数据分析解决方案,旨在对各种结构的数据进行分析,其主要架构核心是分析框架,配合数据仓库基础设施、数据流语言等各种组件整合使用,从一定程度上解决了分析中的技术难题。^[48]Google也提供了数据分析工具Google Analytics以帮助解决大数据分析中存在的相关问题。

这些大型企业对大数据分析的解决方案,对学习分析技术方案的提出有较大的启示和帮助。不管是商业领域中的分析,还是学习领域中的分析,技术层面上基本是一致的。这些解决方案也使得学习分析系统技术特征得到了体现。

3.学习分析系统

通过上文对学习分析利益相关者的分析,以及对学习分析相关技术手段的概述,对学习分析系统应该有怎样的架构,应该包括哪些内容等,都可以有了一些大体上的认识。我们认为,学习分析从技术策略上来说,应该不仅限于专用和通用分析工具的使用,而是应该扩展到整合各种分析模块的系统层面上去。

Wolfgang等学者提出了一个通用学习分析服务设计指导框架,为建设一个学习分析系统提供了思维蓝图。SoLAR的学者们设计了一种整合式模块化学习分析系统(如图5所示)。系统的核心分析工具(引擎)包括:学习分析引擎,自适

应内容引擎,干预引擎(推荐、自动化支持),仪表盘、报告和可视化工具。^[49]该学习分析系统中,分析引擎是核心,它是一个整合多种分析模块的分析框架,应用社交网络分析、话语分析等多种技术辨识和处理数据。自适应内容引擎重点关注学习适应和个性化,在出版商等开发者的帮助下,对学习内容进行适应性和个性化定制和推送。干预引擎主要是使用分析引擎产生的预测模型,进行学习进程追踪和自动化以及人工干预。整个系统的输出依靠一个仪表盘,根据用户角色的不同呈现不同的内容,内容主要以可视化、图形化的形式呈现。

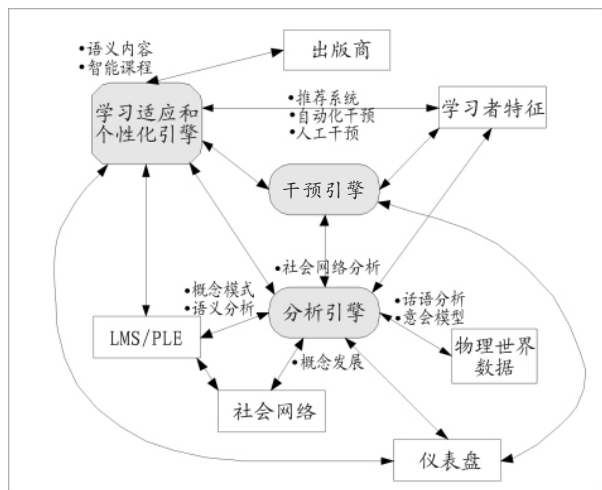


图5 SoLAR提出的整合式学习分析系统

我们认为,学习分析系统应该具有以下几种技术特征:

(1)功能模块化。单一的学习分析工具已无法满足智慧学习环境建构中对学习分析的多样化要求。学习分析系统要满足利益相关者的各种需求,就需要强调对于多种工具、引擎等的模块化聚合。不同功能和特征的模块集成为一个学习分析系统,模块内高内聚,不同模块间低耦合,针对不同结构的数据采用不同模块进行处理和分析。系统内模块可以有工具集、插件、微件等多种形式。

(2)报告可视化。学习分析结果面向主要对象一般是学习者和教学者,这就强调学习分析结果报告应是易于理解的,因此,可视化呈现是学习分析报告的必然结果,这就催生了“仪表盘(Dashboard)”这一概念。仪表盘显示的是可视化的学习分析结果,利益相关者根据需求可见不同显示界面的仪表盘,如,学习者界面、教学者界面、管理者界面等,不同的界面所显示的报告和内容是不同的,却是相关联、相联通的。报告可视化技术目前已经发展较成熟,一些优秀的学习分析可视化工具,如,梯度学习分析系统GLASS就可以提供优质的学习分析可视化报告。^[50]

(3)数据多源化。学习分析系统的数据来源已不再局限于各类数据化学习环境,如,Sakai、Moodle等,而是超越LMS尝试对多种来源的教育数据进行处理。这些数据来源于正式学习环境下的LMS数据、机构管理系统数据(学生档案、学生基本信息等),也可以是非正式学习环境下的各种学习行为痕迹采集,如,学生社交网络关联等,同时也期待可以对物

理学习环境中的数据进行获取和处理。

因此,根据学习分析系统应具备的技术特征以及学习分析过程,结合利益相关者的相关需求,参考SoLAR学习分析系统图示,我们从学习分析系统输入、输出层级角度将学习分析系统归纳为如图6所示系统信息建模,整个系统分为数据层、分析层、报告层和干预与适应层四层。数据层主要获取学习者相关学习数据,送入系统分析层由分析引擎进行数据分析,之后在仪表盘上按学习者、教学者、管理者三大利益相关者需求产生可视化报告,并在此基础上对学习进行干预,并完成自适应过程。

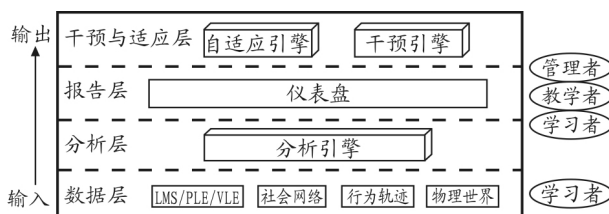


图6 学习分析系统信息建模图示

报告层中仪表盘尽管已经根据不同利益相关者进行了不同界面呈现方式的分类,但是针对不同利益相关者,特别是教学者和管理者,也可按照不同需求呈现不同显示模式,如,学习者个人、班级、区域、决策等多种呈现形式。Dyckhoff等人就针对教学者设计了一套学习分析报告呈现指标,其中包括学生活动行为、活跃领域、资源排行榜、论坛使用、采纳率等。^[51]这些报告呈现指标也反应了在分析层中分析引擎需要根据哪些分析指标进行学习数据的分析。因此,学习分析指标的设计也是急需解决的一个问题。

四、学习分析的挑战与愿景

1.学习分析面临的挑战

随着大数据的爆炸式增长,数据已经从结构化逐渐转变为半结构化、非结构化,并且有预测指出,在未来十年,非结构化数据量将占据全部数据量的90%。^[52]教育领域中的非结构化数据也占有极大的一部分比例。在LMS中,数据是以结构化的形式储存在关系数据库中的,但是,大量非结构化的数据在没有人工帮助下难以被计算机所处理和理解。如何从“大数据”中提取有价值的信息是目前学习分析面临的巨大技术挑战。尽管目前已有不少关于数据集框架的解决方法,如,dataTEL、LinkedEducation、PSLC dataShop在开放数据集服务方面提供了很好的借鉴,Verbert等学者设计了一套教育数据集框架^[53]也为解决数据集问题提供了参考。但是,数据集的研究仍有很多问题需要解决,如,数据标准化、多源、收集方式等。对于学习者在LMS之外的学习行为轨迹的捕捉,以及学习者物理世界数据的获取,都是目前学习分析面临的挑战。与此同时,不同数据系统中数据的互操作性缺失、政府数据系统中一致性的缺失等问题也困扰了很多研究者。^[54]

除了这些技术挑战外,学习分析中一直备受关注的安全性和隐私性问题也仍是一大挑战。Intel对IT领域中分析标

准的调研,为学习分析带来了一定的启示。对于使用者来说,数据安全是首要关注问题,其次是隐私性问题、数据互操作性、系统互操作性等。^[55]由此可见,安全性和隐私性问题是分析技术发展的重中之重。若这一挑战无法克服,将从很大程度上制约学习分析技术的广泛应用和发展。

与此同时,学习分析中特有的学习分析模型的设计问题也是研究者着重关注的重点。区别于一般分析技术,学习分析技术强调使用教学、学习理论进行分析模型设计,在模型技术上进行学习数据的分析,在这一方面目前有很大的潜力和发展空间。

另外,在技术层面还需要解决一些关键技术,目前,学习分析系统还缺乏一个清晰、可操作、可细化的信息模型与体系架构,学习分析的本体、元数据、核心数据集及相关标准还需要进一步研究。同时,如何解决保证学习分析源数据能准确全面体现原来的教育信息化 ITLET 系统、学习者、学习资源等全部特性,即“源息性”问题。

2. 学习分析愿景

从 B(Byte)、KB、MB、GB、TB,到 PB、EB、ZB、YB……数字宇宙的边界正在不断扩大,“大数据”时代的数据大爆炸,也是“大数据”时代的掘金之旅。^[56]目前,我们正处于数据大爆炸的起始阶段,如何迈好数据的掘金之旅?

(1)作为“三通两平台”的构建基石。2012年9月,教育部杜占元副部长在全国教育信息化工作电话会议上提出将“三通两平台”的建设作为今明两年教育信息化的重点工作。^[57]对学校、区域、国家各个层面的教育数据进行统筹挖掘,对学习管理系统、学生信息系统等各类系统学习数据进行整合分析。从学习分析中攫取高度科学和实践性的教育信息,可以更好地促进“三通两平台”的构建和实现。与此同时,“三通两平台”的构建也为学习分析提供了强有力的数据来源,平台构建时考虑到后期学习分析数据的可操作性和整合性,可以进一步促进学习科学的科学性和有效性。

(2)作为电子书包的优势支撑。电子书包系统包括电子课本、学习终端、虚拟学具和学习服务等部分。^[58]其中虚拟学具和学习服务支持,被认为是电子书包用于信息化教育的优势所在。^[59]而学习服务中所提供的学习进度、评价服务、智能学伴/代理功能等,都需要学习分析技术的支持。学习分析为学生、教师、家长等提供各具特色的报告、干预、推荐、建议等,特别是根据学习分析的结果对学生提供个性化和自适应的关联推送,对教师提供实时学生状态报告和干预建议等服务,都将从一定程度上解决电子书包目前仍存在的问题和挑战。作为关键技术的学习分析也从整体上为电子书包的建构,提供了科学性的支撑,更好地服务于国家电子书包的标准研制、产业发展和教育创新应用。^[60]

(3)作为智慧教育云的核心组分。智慧教育云服务强调在统一的智能开放架构云计算平台上,按需向用户提供服务,实现统计教育信息与数据、形成科学决策、实施教育评价等系列活。智慧教育云要求可以进行服务情景识别以及智能信息提取、处理、检索、推送的功能,这些过程即是学习分析技术的研究范畴。学习分析是教育云平台的核心组分,

在一定程度上教育云平台的构建也为运用学习分析技术进行数据分析提供了便利。同时,教育云服务高度嵌入的智慧学习环境也强调对学习情境、学习者特征等的全面感知,^[62]学习分析也逐渐成为了智慧教育的科学力量。^[63]这些未来相对较快实现的教育信息化新进展,也期待着学习分析技术的进一步发展和渗透。

学习分析对智慧教育的进步和发展起到了至关重要的作用,学习分析的潜力和价值也已经逐渐被人们所熟知。尽管学习分析面临着诸多挑战,但是研究者们已在通过不断地努力来直面挑战。移动设备、云计算等技术的快速发展,使得学习分析中的一些鸿沟正在逐步得到解决。目前,学习分析系统设计与开发、学习分析相关标准的研制、学习分析普适应用等研究都在如火如荼地开展,相信学习分析会进一步促进教育信息化的发展。

[参考文献]

- [1]中华人民共和国教育部. 教育信息化十年发展规划(2011-2020年)[EB/OL].[2012-03-13].http://www.edu.cn/zong_he_870/20120330/t20120330_760603.shtml.
- [2][60]祝智庭,贺斌.智慧教育:教育信息化的新境界[J].电化教育研究,2012,(12):5-13.
- [3][17][45][54] Bienkowski, M., Feng, M. & Means, B.(2012).Enhancing Teaching and Learning Through Educational Data Mining and Learning Analytics: An Issue Brief.
- [4]Barneveld, A., Arnold, K. & Campbell, J.(2012). Analytics in Higher Education: Establishing a Common Language. 2012-1.
- [5]Watson, "Business Analytics Insight."
- [6][32]Brown, M. Learning analytics: the coming third wave[DB/OL].<http://net.educause.edu/ir/library/pdf/ELIB1101.pdf>, 2011-4-1.
- [7]Siemens, G. What are Learning Analytics?[EB/OL].[2010-08-25].<http://www.elearnspace.org/blog/2010/08/25/what-are-learning-analytics/>.
- [8]Siemens, G. Learning and Knowledge Analytics- Knewton-the future of education?[EB/OL].[2011-04-14]. <http://www.learninganalytics.net/?p=126>.
- [9]Siemens, G. Educational Transformation: Openness And Learning Analytics [DB/OL]. [2010-10-15]. Presented to: Universidad del Sagrado Corazón, Retrieved on http://maestrias25.sagrado.edu/presentaciones_siemens/Siemens-Educational Transformation Openness AndLearningAnalytics.pdf.
- [10][18][33]Long, P., & Siemens, G. (2011). Penetrating the fog: analytics in learning and education. *EDUCAUSE Review*,46(5), 31- 40.
- [11][16][20]Ferguson, R. The State Of Learning Analytics in 2012: A Review and Future Challenges. [EB/OL].[2012-03-01]. <http://kmi.open.ac.uk/publications/pdf/kmi-12-01.pdf>.
- [12]Elias,T.LearningAnalytics: Definitions, Processes and Potential[DB/OL]. [2011-01-16]. <http://learninganalytics.net/LearningAnalytics-DefinitionsProcessesPotential.pdf>.
- [13]Romero, C., & Ventura, S.(2007). Educational data mining: A survey from 1995 to 2005. *Expert Systems with Applications*,33 (1), 135-146.
- [14]McFadden, C. (2005). Optimizing the Online Business Channel with



Web Analytics [blog post]. Retrieved October 5, 2010 from http://www.webanalyticsassociation.org/members/blog_view.asp?id=533997&post=89328&hhSearchTerms=definition+and+of+and+web+and+analytics.

[15] Baker, B. (2007). A conceptual framework for making knowledge actionable through capital formation. D.Mgt. dissertation, University of Maryland University College, United States—Maryland. Retrieved October 19, 2010, from ABI/INFORM Global. (Publication No. AAT 3254328).

[19] Donald Norris, Linda Baer, and Michael Offerman, "Why Action Analytics for Higher Education?" National Symposium on Action Analytics, St. Paul, MN, September, 2009, <http://www.edu1world.org/PublicForumActionAnalytics/23338>.

[21] The New Media Consortium. The Horizon Report 2010 Higher Education Edition [EB/OL]. [2013-01-21]. <http://www.nmc.org/pdf/2010-Horizon-Report.pdf>.

[22] The New Media Consortium. The Horizon Report 2011 K-12 Edition [EB/OL]. [2013-01-21]. <http://www.nmc.org/pdf/2011-Horizon-Report-K12.pdf>.

[23] The New Media Consortium. The Horizon Report 2011 Higher Education Edition [EB/OL]. [2013-01-21]. <http://net.educause.edu/ir/library/pdf/hr2011.pdf>.

[24][35] The New Media Consortium. The Horizon Report 2012 Higher Education Edition [EB/OL]. [2013-01-21]. <http://net.educause.edu/ir/library/pdf/hr2012.pdf>.

[25] The New Media Consortium. The Horizon Report 2013 Higher Education Edition [EB/OL]. [2013-4-28]. <http://net.educause.edu/ir/library/pdf/HR2013.pdf>.

[26] Siemens, G. 1st International Conference on Learning Analytics and Knowledge 2011 [EB/OL]. [2010-07-22]. <https://tekri.athabasca.ca/analytics/>.

[27] LAK2012 Home Page [EB/OL]. [2012-12-22]. <http://lak12.sites.olt.ubc.ca/>

[28] LAK2013: Site for the Learning Analytics conference series [EB/OL]. [2013-01-23]. <http://lakconference2013.wordpress.com/>.

[29] Surthers, D. & Verbert, K. (2013). Learning Analytics as a "Middle Space". Paper presented at LAK13: 3rd International Conference on Learning Analytics and Knowledge.

[30] SoLAR Flare [EB/OL]. [2013-3-3]. <http://www.solaresearch.org/flare/>.

[31] Diaz, V. & Brown, M. (2012). Learning Analytics: A Report on the ELI Focus Session [DB/OL]. [2012-05-12] <http://net.educause.edu/ir/library/PDF/ELI3027.pdf>.

[34][40] Greller, W., & Drachslar, H. (2012). Translating Learning into Numbers: A Generic Framework for Learning Analytics. Educational Technology & Society, 15(3), 42-57.

[36][37][39][49] Siemens, G., Gasevic, D., Haythornthwaite, C., Dawson, S., et al. (2011). [2011-7-28]. Open Learning Analytics: an integrated & modularized platform.

[38] Retalis, S., Papasalouros, A., Psaromiligkos, Y., Siscos, S., Kargidis, T. Towards Networked Learning Analytics—A Concept and a Tool [DB/OL]. [2011-10-26]. Proceedings of the 5th International Conference on Networked Learning, Lancaster, UK. Retrieved from <http://nlc.ell.aau.dk/past/nlc2006/abstracts/pdfs/P41%20Retalis.pdf>.

[41] Siemens, G., & Gasevic, D. (2012). Guest Editorial – Learning and Knowledge Analytics. Educational Technology & Society, 15(3), 1-2.

[42] 顾小清, 张进良, 蔡慧英. 学习分析: 正在浮现中的数据技术 [J]. 远程教育杂志, 2012, (1): 18-25.

[43] 李青, 王涛. 学习分析技术研究与应用现状述评 [J]. 中国电化教育, 2012, (8): 129-133.

[44] Kim, M., & Lee, E. (2012). A Multidimensional Analysis Tool for Visualizing Online Interactions. Educational Technology & Society, 15 (3), 89-102.

[46] Weed, B. IBM Offers Big Data Solutions for a Smarter Planet [EB/OL]. <http://www.intel.cn/content/dam/www/public/us/en/documents/white-papers/ibm-vendor-spotlight.pdf>, 2012-6.

[47] Billings, K. Oracle Takes You through the Four Phases of Achieving Big Data Insight [EB/OL]. <http://www.intel.cn/content/dam/www/public/us/en/documents/white-papers/oracle-vendor-spotlight.pdf>, 2012-6.

[48] Intel. Intel? Distribution for Apache Hadoop* 软件 [EB/OL]. [2013-01-29]. <http://www.intel.cn/content/dam/www/public/cn/zh/pdfs/hadoop-datasheet-080712-cs.pdf>.

[50] Gantz, J. & Reinsel, D. Extracting Value from Chaos [EB/OL]. [2011-06-13]. emc.com/collateral/analyst-reports/idc-extracting-value-from-chaos-ar.pdf.

[51] Leony, D., Pardo, A., Valentín, L., Castro, D. & Kloos, C. GLASS: A Learning Analytics Visualization Tool (2012). [EB/OL]. [2012-05-30]. Paper presented at LAK12: 2nd International Conference on Learning Analytics and Knowledge (30 April-2 May). Retrieved from http://dl.acm.org/ft_gateway.cfm?id=2330642&ftid=1262788&dwn=1&CFID=175337536&CFTOKEN=31090269.

[52] Dyckhoff, A. L., Zielke, D., Bültmann, M., Chatti, M. A., & Schroeder, U. (2012). Design and Implementation of a Learning Analytics Toolkit for Teachers. Educational Technology & Society, 15 (3), 58-76.

[53] Verbert, K., Drachslar, H., et al. (2011). Dataset-driven Research for Improving Recommender Systems for Learning. Paper presented at LAK11: 1st International Conference on Learning Analytics and Knowledge.

[55] 英特尔信息技术中心. 同行研究报告大数据分析 [EB/OL]. [2013-01-29]. <http://www.intel.cn/content/dam/www/public/cn/zh/pdfs/hadoop-datasheet-080712-cs.pdf>.

[56] 技术商业: 大数据, 下一个黄金增长点 [EB/OL]. [2013-02-22]. <http://www.intel.cn/content/www/cn/zh/it-management/intel-it-media-article-50n.html>.

[57] 杜占元. 今明两年落实“三通两平台” [EB/OL]. [2012-10-23]. http://www.edu.cn/li_lun_yj_1652/20121023/t20121023_859668.shtml.

[58] 吴永和, 祝智庭, 何超. 电子课本与电子书包技术标准体系框架的研究 [J]. 华东师范大学学报(自然科学版), 2012, (2): 70-80.

[59] 祝智庭, 郁晓华. 电子书包系统及其功能建模 [J]. 电化教育研究, 2011, (4): 24-27, 34.

[60] 吴永和, 马晓玲, 杨飞. 电子课本出版与生态发展的阐释与研究——基于标准研究、产业发展与教育创新的视角 [J]. 远程教育杂志, 2013, 31(214): 17-28.

[61][62]张进宝,黄荣怀,张连刚. 智慧教育云服务:教育信息化服务新模式[J]. 开放教育研究,2012,(3):20-26.

[63]祝智庭,沈德梅. 学习分析学:智慧教育的科学力量[J]. 电化教育研究,2013,(5):5-12,19.

[作者简介]

吴永和,博士,教育部教育信息化技术标准委员会委员及全国信息技术标准化技术委员会教育技术分技术委员会委员,全国电子课本与电子书包专题组秘书处负责人及总体专题项目组召集人,华东

师范大学教育信息化系统工程研究中心副主任,高级工程师,硕导,主要研究方向为教育信息化系统及标准、学习技术系统架构、教育新媒体、学习分析等;陈丹,曹盼,华东师范大学教育信息技术学系在读硕士生;马晓玲,华东师范大学信息学系讲师;冯翔,华东师范大学上海数字化教育装备工程技术研究中心助理研究员;祝智庭,教授、博导,教育部教育信息化技术标准委员会及全国信息技术标准化技术委员会教育技术分技术委员会主任,华东师范大学开发教育学院院长,教育信息化系统工程研究中心主任,主要研究方向为教育信息化、技术哲学、教育技术标准等。

Learning Analytics: The New Coming Wave of Education Information

Wu Yonghe^{1,2}, Chen Dan³, Ma Xiaoling⁴, Cao Pan³, Feng Xiang^{1,2} & Zhu Zhiting^{1,2}

(1. Shanghai Engineering Research Center of Digital Education Equipment, East China Normal University;

2. Shanghai Engineering Research Center of Digital Educational Equipment, East China Normal University;

3. Department of Educational Information and Technology, East China Normal University;

4. Department of Information, East China Normal University, Shanghai 200062)

[Abstract] Education information has witnessed the revolutions of LMS and web 2.0 applications. The deeper use of new technologies leads to a high-speed growth of the big data in education. The active demand of mining potential value of those makes the appearance of learning analytics. With the analysis of learning analytics at home and abroad by means of literature review, the different concepts of learning analytics and its origin are defined, and the related concepts of learning analytics are compared as well in the paper. In the latter part of this paper, recent researches on learning analytics, its development and technical strategies are discussed. Finally, the challenges and visions that learning analytics are facing are summarized in order to make an all-around discussion and classification towards learning analytics.

[Keywords] Learning analytics; Educational data; Data mining; Smart learning

收稿日期:2013年3月19日

责任编辑:陶侃

http://dej.zjtvu.edu.cn

国际开放、远程、灵活学习的新战略 第25届国际开放与远程教育理事会(ICDE) 世界大会10月在天津召开

【本刊讯】第25届国际开放与远程教育理事会(ICDE)世界大会将于2013年10月16日至18日在中国天津召开。

据悉,这次年会的主题是:国际开放、远程、灵活学习的新战略。为了全面和准确地体现年会的主题思想,本届年会将从文化、制度、技术、资源、质量、服务和残疾人教育(辅助学习)七个方面,集中展现全球开放远程教育研究的最新成果和最新进展。特别是把残疾人教育(辅助学习)作为年会的分主题,将为各国开放远程教育工作者和各界朋

友提供一个关注和研究社会弱势群体教育问题的新的交流平台。

两年一度的ICDE远程和开放教育世界大会,是世界范围内远程开放教育领域的顶级盛会,为前沿发展、网络建设和专业发展提供交流论坛。本届大会由国际开放与远程教育理事会主办,天津广播电视大学承办。除中国的参会者外,还将有众多来自世界各个国家和地区的代表与会。

(风余)