

· 高等教育 ·

大数据时代高校信息管理与决策机制研究*

舒忠梅, 屈琼斐

(中山大学 教育学院, 广东 广州 510275)

摘要: 高等教育机构正处于数据爆炸但知识贫乏的大数据时代, 多年积累的海量教育教学数据为高校的管理与决策带来了挑战。分析了高校海量数据现状及其存在的问题, 对如何有效利用数据以提升管理与决策支持机制进行了探索, 提出高校要进一步推进海量数据信息集成建设, 通过信息技术对教育发展进行变革, 从信息技术上采用商业智能等大数据处理技术、从行政管理上落实全面的数据管理机制, 构建完善的高校教育教学数据质量保障体系, 从根本上解决和满足大数据时代高校对信息的需求, 从而提升管理效益。

关键词: 高校信息化; 大数据; 数据管理; 商业智能

中图分类号: G647 **文献标识码:** A **文章编号:** 1009-055X(2013)06-0096-06

2010年7月颁布的《国家中长期教育改革和发展规划纲要(2010-2020年)》提出“信息技术对教育发展具有革命性的影响, 必须予以高度重视”。^[1] 信息技术的高速发展, 为高等学校的教育教学及其服务管理提供了新的手段和丰富的资源。^[2] 随着高等教育的发展和教育体制的改革, 高校积极进行管理信息化建设, 实施数字化校园或智慧校园等举措, 广泛采用了各类信息系统辅助日常行政管理与教学科研工作等。信息技术在为学校师生带来便利的同时, 也产生了大量数据。国际数据公司 IDC 研究表明, 数字大学仅 2007 年一年创建或采集 2.25×10^{21} 位数据, 并预计将以每年 60% 的增速持续至今, 大数据时代已经来临。^[3] 高等教育机构正面临着大数据的挑战, 本文从信息技术的有效利用与高等学校的管理与决策机制出发, 探讨如何有效利用当前高校信息系统中产生的海量数据, 提升高校的管理效益和科学决策水平。

一、大数据对高校管理决策的挑战

随时大数据时代的到来, 高校广泛采用的信息系统产生和存储了大量教育教学数据, 这些数据记

录着学校的有形资源及其历史演变。研究这些数据可以发现学校潜在的问题, 通过数据分析支持科学的决策管理, 提高管理效益。然而, 由于高校数据内容繁杂、数据量大、数据相关关系混乱, 管理与决策在大数据时代面临着新的挑战。

(一) 如何有效利用操作型业务管理数据

计算机和网络技术的快速发展, 为院校信息管理提供了新的手段和丰富的资源。随着教育机构的发展、信息化建设的推进和教育活动的增加, 计算机在我国教育行业的应用经历了一个快速发展的过程: 从最原始的日常事务人工管理到某些独立运行的管理系统、直到最终实现的办公自动化系统、教务管理系统、科研管理系统、校园图书馆信息系统、资源设备数据库管理系统、人事管理系统、财务管理系统和在线远程教育系统等等。这些信息系统在日常事务处理应用中积累了相当多的操作型业务数据, 这类数据在很大程度上是结构化的 (Structured) 数据, 即根据数据模式定义数据的大小、类型及其合理的取值范围。如何运用这些数据进行分析、提升和促进高校管理能力与科学决策水平, 在教育竞争中越来越凸显出重要的作用。

收稿日期: 2013-09-05

* 基金项目: 国家自然科学基金项目: ITS 中基于有向超图的个性化的学习过程及其支持资源的优化 (61202345)

作者简介: 舒忠梅 (1974-), 女, 湖北荆门人, 讲师, 博士, 研究方向为教育信息化、院校研究等。屈琼斐 (1973-), 女, 浙江台州人, 副研究员, 博士, 副院长, 研究方向为院校研究、教育社会学等。

（二）如何利用 IT 技术处理庞大的内容服务数据

由于信息交流的崭新方式（如 Web）和各种多媒体的相互作用，内容服务数据是一个极为含糊的概念，但对于有效管理和服务而言，内容服务数据是不可缺少的必要因素。它不仅包括校园网站上的各类资料，也包括备忘录、电子表格、出版物、网上教学资源、图书馆及博物馆的收藏等。其大多数的内容服务数据是非结构化的（Unstructured），即与传统业务应用中产生的建模数据不同，含有比较随意形式的信息。由于内容服务数据将应用于特定的网络模板或数据库调用中，因此越来越多的内容采用半结构化的（Semi-structured）数据形式。有些内容服务数据，如政策声明、描述学术课程网站、课程简介等，需要官方批准；同时也需要以正规方式进行描述以便于追踪、搜索和随时可用。而另外有些内容服务数据却是临时性的数据垃圾。为此，高校需要提供相应的工具，帮助合适的用户提供创建、发布、查找、保存或选择合适的内容服务数据。

（三）如何以院校研究为切入点管理研究数据

现代高校每天进行着大量的科学实验或计算，衍生了海量数据。相对于传统的数据捕获形式，数字数据更易于携带和搜索。研究数据除了和日常操作数据及内容所面临的共同问题之外，还另有关于研究数据的归属、保留及解释等问题。这些问题若能得以解决，研究效率就可以大大提高。否则，花费了大量精力收集的研究数据最终将因为退化或无人问津而丧失其价值。

这些挑战与机遇并存。一方面，高等教育管理正面临信息爆炸的危机，高校管理系统正在被数据所淹没；但另外一方面，海量数据也给管理者带来了科学决策的契机。我国高校管理系统中，高等教育研究一直尚未涉足对高校管理有关的数据信息的综合、定量、系统地研究。而在国外大学的管理系统中，一个完整的院校研究部门是必不可少的，其基本的功能就是对信息的管理和分析。对数据展开研究，是增强高校管理系统职能的一个重要环节。而且，随着国家逐渐放给大学越来越多的办学自主权的发展趋势，高校管理系统将因重视信息的管理和分析而衍生出一个新的功能，或者吸引高等教育研究人员深入关注信息管理与分析问题，从而更加

实际地支持高校管理提升，是已萌芽并可预见的发展趋势。

二、大学管理决策数据的利用现状

（一）综合性的信息分析与决策支持系统欠缺

高校各类信息系统的建立和完善逐渐提高了高校日常工作的办事效率，同时为高校办学质量评价与监控带来了显著效益，然而在实际运用中却存在以下问题。

1. 信息系统缺乏整体统一规划

部分高校的业务部门分别开发满足自己需要的独立系统，造成环境的异构、数据的冗余，使得信息处在零散的孤岛状态。部门之间横向信息流通不畅，各类信息保存在各个单位，而其他相关单位得到某公共信息较为困难，缺少一种快速、有效的方法对信息进行共享。学校每年都要完成相关办学信息的总结工作，从基层部门经层层填表整理，直到最终数据的收集统计，往往都是基于电子表格文件传输或网上系统填报，耗费大量的人力、物力和时间。

2. 信息系统欠缺分析和决策支持功能

目前大多数高校的业务信息系统基本上属于联机事务处理（OLTP）系统，无法直接提供各种综合程度的数据采集、分析及综合利用的功能。尽管关系数据库和企业软件应用包使得这些日常业务数据的操作更加快捷灵活，但仍无法提供广泛的分析和功能。例如，通过高校现有的教育教学信息系统，可以轻易地为学生设置合适的课程，然而要综合利用学生和课程的聚类数据，完成招生录取和课程规划却实属不易。由于这些信息系统不具备决策支持能力，系统缺乏进一步扩展应用的能力，无法直接从各级各类办学信息系统中采集数据并加以综合利用；各级教育系统产生的大量数据也不能进一步提炼升华为知识，及时提供给决策部门。一项教育决策的制定往往需要查询多个基于各种异构数据源的教育信息系统和相关的外部系统，进行大量数据分析后才能做出此决策。这种工作量大且容易出错的行为过程必然会影响决策质量。

（二）数据质量和有效利用率亟待加强

在教育竞争日益激烈的今天，数据资源的管理和应用是现代大学取得竞争优势的关键。然而，尽管时至今日高校的数据量巨大，而其中真正有价值的信息却不多。^[4]数据能否带来价值和转化为决策

效益不仅仅取决于量的多少,更重要的是取决于数据的可用性和对其合理的使用。现代大学所探索的问题除了包括“是什么”、“怎么样”外,更加侧重于“为什么”以及将来“怎么办”,正确合理的数据能给出科学的答案,因而对数据的依赖程度越来越高。这种依赖既包括数据的“量”,也包括数据的“质”。^[5]

然而,高校现有的大部分信息系统中存储的数据质量堪忧。数据质量是信息系统所表达的数据视图与客观世界中同一数据之间的距离,数据质量目标模型可通过一个多维的、层次结构的方式来描述,包括可访问性(accessibility)、可理解性(interpretability)、可信度(believability)以及可用性(usefulness)等4个方面。不少高校通过信息化建设积累了大量数据,但是这些数据存在着不少数据质量问题,无法确保其可访问性、可理解性、可信度以及可用性,导致不能被有效利用。通过调查高校数据管理现状,不难发现其影响因素主要来源于技术和管理两个方面。

1. 技术因素

高校信息系统设计与开发过程考虑不周,如数据元素定义不清、数据输入格式缺乏控制、数据存在冗余等;有些数据项在设计时为开发方便采用缺乏缜密的结构化处理;有些数据缺少更新时间、维护人员和操作记录等重要属性,导致查找或确认数据时缺乏可用的支撑数据。

2. 管理因素

不少高校部门管理人员尚不能从手工或使用电子表格等简单软件的工作方式转变为使用信息系统的工作方式,缺乏共享意识和整体观念,在对系统的数据进行录入或维护管理时,出现部分数据的缺失或者错误。另一方面,同一数据被学校多个部门维护、数据归属责任不清、各部门交叉更新、重复采集、数据定义不清、参照不完整等,造成数据质量不高、不能综合利用等诸多问题。

在数据利用上,我国高校往往存在下面问题:一是对数据“断章取义”和片面使用。随便举出几个例子或片面寻找几个数据以佐证某种事物优于另一种事物的研究方式不仅违背统计学研究方法中的系统性原则,也和一般科学研究的方法是相悖的。二是校内各部门之间的沟通与联系较少,数据统计指标定义不规范,统计范围不明确,常常“数出多头”,重复统计。三是信息垃圾。数据的过剩和不足是同时存在的一对矛盾现象。数据搜集

和日常工作积累了大量的冗余信息,我们经常只关注数据的不足,而对冗余信息处理不够及时。四是管理决策缺乏数据支持。部分高校领导在面临管理决策问题时找不到可用的综合数据进行相应的支持。

显然,简单的数据管理已不能满足大数据时代现代大学管理与决策的信息需求。随着计算机和网络技术的发展,为信息过剩提供了技术上的可能,现在我们能够收集和储存的数据之多、规模之大,是前人无法想象的。因此,面对海量的教育教学数据,若不能加以有效利用,就会降低对数据的使用效益并使管理和决策更加困难。

三、商业智能技术与行政管理机制协同解决方案

大数据时代的高校信息管理与决策问题不仅是一个技术问题,更是一个管理问题。在海量数据的有效管理与利用的迫切需求背景下,高校要进一步推进信息化建设,通过信息技术对教育发展进行变革,其关键在于从信息技术和行政管理上采取协同保障机制。

(一) 从行政管理上落实全面的数据管理机制

部分高校在数据管理体制与流通机制上缺乏切实有效的总体构思与全面设计,信息管理工作人员亟待加强相关的信息系统使用指导与信息工作规范培训。因而需要从高校的行政管理层面上建立全面的数据管理保障机制,包括:

成立学校数据管理委员会。建设一个强有力的校级数据管理委员会,可以如国外大学的设置,设计为由学校主管领导牵头、所有数据应用和管理的职能部门及学科领域的主管领导组成。该委员会指导数据工作小组形成学校数据质量管理规范,包括数据质量标准、数据管理流程、数据监测体系、数据质量评估等优化方案,建立可操作的数据质量责任管理制度、监督数据工作小组进展,为数据质量提供最好的实践、度量和方法。

组建学校数据工作小组。数据工作小组由数据管理委员会领导,由所有数据应用和管理的职能部门及学科领域的数据管理人员组成。该工作小组成员从各自的业务需求出发,承担着明确的数据管理责任,相互协同工作、建立完善的数据协调机制,负责确定监控数据质量的度量、指派数据质量的责任、调整数据管理实践或业务信息系统以提升和维

护数据质量。

建立完善的数据管理培训与激励机制。为确保学校数据质量，提升数据工作人员素质，定期举办有关信息系统使用、数据质量管理规范等方面的培训，建立健全学校数据工作奖惩与激励制度，减少人为因素导致的数据质量问题。

（二）在信息技术上采用商业智能技术

随着云计算和云存储的推广，可以收集和利用越来越多的数据。为了高效地处理和有效地利用各种形式的海量数据，信息科学技术的飞速发展和企业界不断提出的新的需求，以数据仓库（Data Warehousing, DW）、数据挖掘（Data Mining, DM）、联机分析处理（On-line Analysis Processing, OLAP）等为核心的商业智能（Business Intelligence, BI）技术应运而生。

商业智能这一概念由加特纳·格鲁派于1989年提出，可将商业智能理解为一种能将机构现有的数据转化为知识，帮助机构通过基于事实和电脑化系统，做出科学的业务决策的工具，已成功地应用于银行、电信、保险、制造业和零售业等行业。

1. 国内外高校管理中的商业智能应用

通过调查了解世界各地的高等教育部门采用商业智能技术的状况发现，高等教育已明显落后于其它行业。然而，正如 Dave Wells 在文献^[6]中所描述的，越来越多的高校正在关注商业智能应用这一主题，高等教育在技术方面到了接受商业智能的时候。

北美高等教育体系较早关注商业智能应用的重要性，美国高等教育数据仓库论坛（HEDW）的宗旨是“致力于分享知识和最好的大学知识管理实践经验的高等教育同行的网络组织，包括建设数据仓库、发展院校报告策略、提供决策支持”。^①根据调查了解，不少高校已经采用或正开始采用商业智能技术，对学校的各种教育教学数据进行整合与分析，为学校的科学决策与管理提供信息支持。如宾州州立大学、密执根大学均于2005年提出建设BI项目的倡议，佛罗里达州立大学、华盛顿大学分别于2007年、2008年开始利用BI项目分析学校整合数据进行决策支持，普渡大学于2008-2014期间建设学校新的协同集成系统，印第安纳大学2009年提出BI建设路线图，斯坦福大学商业智能

中心2009年提交的BI架构及方案获批准、2010年正式启动项目建设，加州大学（总部）建设StatFinder系统，伊利诺斯大学建立学校决策支持数据仓库，亚利桑那州立大学为支持科学决策建设了仪表盘（dashboard）等。其中，佛罗里达州立大学建设的BI系统被认为是较成熟系统，目前用Oracle系统建设了一个统一标准化平台，号称现已拥有1600个用户、189个仪表盘、55个主题区域，其BI建设逻辑是：让终端用户使用尽可能多的操作系统，以便让更多部门通过培训可以自行处理大量的报告和仪表盘。^②同时基于商业智能技术产出的数据，也形成一批对高等教育和高校自身进行深入分析的研究报告。如美国教育研究的主要组织院校研究学会，年会报告除了涵盖数据管理、数据仓库等关于计算机技术本身的报告之外，关于评估、资源、学生、合作和分析的报告，较多地来自于商业智能系统数据的研究结果。^[7]

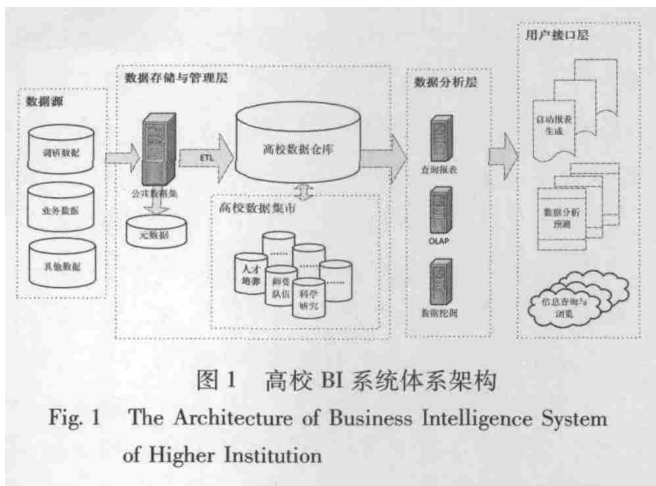
在欧洲、亚洲等其他地方的高等教育体系中，商业智能技术的应用才刚刚起步。我国少数高校也开始迈出了建设BI系统的步伐，如中山大学2010年开始提出学校BI系统建设倡议、现已初步完成BI系统的总体设计并同步进行数据仓库系统的开发，上海交通大学2011年初步推出BI项目的子系统——财务管理驾驶舱系统，复旦大学为学校师生在校生命周期实现管理信息化“全覆盖”，中国人民大学推出综合数据填报、数据存储、数据管理和数据展示等四大功能的数据平台，浙江大学为优化资源配置建设共享数据中心进行数据集成，上海财经大学建设校务决策支持系统，常熟理工学院建立高校决策支持系统等。

2. 高校管理与决策支持的商业智能架构

商业智能技术应用到高校信息化管理与决策支持中，其关键是建立综合层面上的、能反映高校整体教育教学质量和状态的信息集成系统（下文简称高校BI系统平台）。高校BI系统平台体系架构由数据源、数据存储与管理层、数据分析层和用户接口层四部分组成，如图1所示。

① 来源于美国高等教育数据仓库论坛：<http://www.hedw.org/>。

② 来源于美国院校研究协会网站：<https://www.airweb.org/>。



高校决策者常常希望从不同的角度来审视高校各类教育教学数据,比如从时间、区域、学科、教学或科研成果、课程建设、学生层次、交流合作、办学资源等维度全面了解学校的教育教学质量和状态。高校BI系统平台从学校现有的业务处理信息系统数据库中或其他相关数据源中提取出有用的数据并进行清理整合,以保证数据的正确性、一致性和完整性等,随后经过抽取(Extract)、转换(Transform)、装载(Load),即ETL过程处理使之成为统一、集中、稳定的数据,在元数据库中保存数据描述、数据转换及映射等过程。并根据高校常见的分析和统计主题,建设校级数据仓库包含人才培养、师资队伍、科学研究、办学资源、交流合作等主题的数据集市,从而得到学校办学数据的全局视图。

在此基础上,数据分析层主要利用SQL快速查询、Hadoop深层分析查询等工具、数据挖掘工具、OLAP工具等,从大量的教育教学数据中经过多维度分析,获得有利于高校运作、提高竞争力的信息。在高校BI系统平台中,商业智能技术为高校管理主要提供以下数据查询、分析与利用方式。

固定的报表:如定期反映学校招生人数、毕业生人数、师生比、教学科研成果等数据,以及国家和地方教育部、教育厅等上级部门需要收集学校的相关办学数据报表等。

即席查询:支持灵活地查询分析,满足报表之外的信息需求。

统计分析:如通过回归分析,在学生、课程、教师、教学和学习成果等因素中找出影响学生满意度的关键因素。

多维分析:从多个角度分析查询,如从学科专业、时间、校区、培养类型等多维度寻找人才培养

质量下降的原因,从招生录取情况、毕业就业情况、师生比、教学科研成果、国际交流等多角度分析影响学校声誉的短板,通过学科、专业、校区、学生年级等各类内部结构特性针对院系、专业、年级等进行全校整体办学信息的内部结构分析等。

预警功能:例如某年度人才培养质量或科研成果下降时可以自动发送提醒消息进行重点关注。

预测分析:根据历年办学信息,分析学校各项指标增长情况,并预测未来的趋势,从而为合理调配资源提供支持等。

数据挖掘建模分析:基于学生的入学状况、家庭背景、学习兴趣、学习投入、学习成果等学生学习行为状况,细分学生类型,寻找学生学习成功的行为特征以及潜在需求,推送个性化的选修课程建议等信息和服务。

优化分析:如根据学校现有学生、师资、资源、科研和人才培养等,对学校的资源情况作配置效益和投入产出情况分析等。

高校信息化管理与商业智能技术相结合,是当今信息时代高教研究不可或缺的重要组成部分。将商业智能技术应用到高校信息化管理中,对高校各类数据进行搜集、管理和分析,使高校的各级决策者获得知识和洞察力,帮助他们做出对学校发展更为有利的科学决策;同时为高校的师生提供集成的、共享的、个性化的信息服务,帮助提升学校的教育教学、科学研究水平,提高人才培养质量。

四、总结

数据资源的管理和应用是取得竞争优势的关键。如何从大量的办学数据中获取能够真正反映教育状况的有效信息,如何系统有序地管理资源的调度、管理信息资源、采取合适的招生策略、降低校园管理成本、提高教学效果,从而为领导决策提供科学支持,是教育部门普遍关心的问题。随着教育改革的逐步深入和高等教育的国际化,高校面临着质量评估、绩效考核、社会问责、大学排名等现实问题,如何有效地管理海量的教育教学数据,并从这些数据中获取有用的信息,提升教育教学质量、科研水平和管理效益,是高校信息化管理与深入发展的重要研究课题。

为此,在分析高校海量数据所带来的挑战同时,对现代大学管理注重利用数据以提升高校管理效益进行解读,指出高校要进一步推进海量数据信

息集成建设,通过信息技术对教育发展进行变革,其关键在于从信息技术上采用商业智能、从行政管理上落实全面的数据管理机制,构建完善的高校教育教学数据质量保障体系,努力促成用数据来说话、用数据来管理、用数据来决策,从根本上解决和满足大数据时代现代大学对信息和数据的需求。

参考文献:

- [1] 国家中长期教育改革和发展规划纲要工作小组办公室. 国家中长期教育改革和发展规划纲要(2010-2020) [N]. 中国教育报, 2010-07-30 (1-3).
- [2] 熊才平, 何向阳, 吴瑞华. 论信息技术对教育发展的革命性影响 [J]. 教育研究, 2012 (6): 22~29.
- [3] P D John, F Gantz, C Chute, et al. The Diverse and Exploding Digital Universe: An Updated Forecast of Worldwide Information

Growth through 2011 [EB/OL]. (2008-08-10) [2012-11-30] <http://www.emc.com/collateral/analyst-reports/diverse-exploding-digital-universe.pdf>.

- [4] Yanasky R. Institutional Data Management in Higher Education. EDUCAUSE Center for Applied Research [EB/OL]. (2009-12-30) [2013-01-30] <http://www.educause.edu/library/resources/institutional-data-management-higher-education>.
- [5] 常桐善. 构建院校智能体系: 院校研究发展的新趋势 [J]. 高等教育研究, 2009 (10): 49~54.
- [6] Dave Wells. Institutional Intelligence: Applying business intelligence principles to higher education. Campus Technology [EB/OL]. (2007-04-08) [2012-11-30] <http://campus-technology.com/articles/2007/04/institutional-intelligence.aspx>.
- [7] S ElAtia, D Ipperciel, A Hammad. Implications and Challenges to Using Data Mining in Educational Research in the Canadian Context [J]. Canadian Journal of Education, 2012 (2): 101-119.

Research on Higher Institutional Mechanism of Information Management and Decision Support in Big Data Era

SHU Zhong-mei, QU Qiong-fei

(Education School, Sun Yat-sen University, Guangzhou 510275, Guangdong, China)

Abstract: Higher Education Institution is faced with the big data era which is full of data but lack of knowledge. The digitization of information is dramatically changing the landscape of its management. The challenges and opportunities of big data are given rise to the modern universities, which pay much attention to effective and intelligent use of data to promote the efficiency of institutional management. However, the complexity associated with managing data is increasing; attention must be paid first and foremost to data quality. Therefore, universities should manage an infrastructure for data from the administrative system, and secure the advanced business intelligence from the technology as well, to get full value from the huge institutional data, to mine the data to promote better institutional outcomes, and to improve institutional decisions.

Key words: higher education information; big data; data management; business intelligence

(责任编辑: 邓泽辉)