

BIG DATA IN

Towards the future

i n d i v i d u a l i s t i c e d u c a t i o n

EDUCATION

教育 大数据

迈向共建、共享、开放、
个性的未来教育

方海光◎著


中国科学院院士 张景中

北京师范大学教育学部教授 黄荣怀

倾情作序



机械工业出版社
CHINA MACHINE PRESS



本书在分析大数据技术发展及其对教育产业之潜在影响的基础上，应用丰富的案例，对教育大数据的生态体系进行了深入的梳理。教育大数据将促进共建、共享、开放、个性的未来教育模式的建立，对教育、学习和管理都将产生深远的影响。本书为有志于应用大数据升级教育产业的相关实践者提供了用于数据搜集、处理、分析和呈现的相关资源和工具，并对教育大数据的未来发展前景进行了大胆展望。

BIG DATA IN
EDUCATION

教育 大数据

迈向共建、共享、开放、个性的未来教育

BIG DATA IN

Towards the future

i n d i v i d u a l i s t i c e d u c a t i o n

EDUCATION

方海光◎著



机械工业出版社
CHINA MACHINE PRESS

教育 大数据

迈向共建、共享、开放、个性的未来教育

方海光◎著

BIG DATA IN EDUCATION

Towards the future

i n d i v i d u a l i s t i c e d u c a t i o n



机械工业出版社
CHINA MACHINE PRESS

本书在分析大数据技术发展及其对教育产业之潜在影响的基础上，应用丰富的案例，对教育大数据的生态体系进行了深入的梳理。教育大数据将促进开放、共享、个性的未来教育模式的建立，对教育、学习和管理都将产生深远的影响。本书为有志于应用大数据升级教育产业的相关实践者提供了用于数据搜集、处理、分析和呈现的相关资源和工具，并对教育大数据的未来发展前景进行了大胆展望。

图书在版编目（CIP）数据

教育大数据：迈向共建、共享、开放、个性的未来教育/方海光 著.
—北京：机械工业出版社，2016.9（2018.1重印）
ISBN 978 - 7 - 111 - 55003 - 7

I. ①教… II. ①方… III. ①教育产业 - 产
发展 - 研究 - 中国 IV. ①G522.8

中国版本图书馆 CIP 数据核字（2016）第 238150 号

机械工业出版社（北京市百万庄大街 22 号 邮政编码 100037）

策划编辑：坚喜斌 责任编辑：刘林澍 杨 冰
责任校对：郭明磊 版式设计：张文贵 责任印制：常天培

北京圣夫亚美印刷有限公司印刷

2016 年 11 月第 1 版·第 2 次印刷

145mm × 210mm · 7.25 印张 · 3 插页 · 165 千字

标准书号：ISBN 9787 - 111 - 55003 - 7

定价：45.00 元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换

电话服务

网络服务

服务咨询热线：(010) 88361066

机工官网：www.cmpbook.com

读者购书热线：(010) 68326294

机工官博：weibo.com/cmp1952

(010) 88379203

教育服务网：www.cmpedu.com

封面无防伪标均为盗版

金书网：www.golden-book.com

推荐序1 大数据时代，教育的未来

“大数据”现在很热。有人甚至说掌握大数据的人可以像上帝一样俯瞰整个世界。当然，若有这样的人，大数据就不叫大数据了。

大数据概念太新，在这个领域，我其实算是外行。听说所谓大数据（big data）之“大”，指的是数据量大到难以用当前的信息技术有效地处理。对难处理的东西进行处理，当然是富有挑战性的任务。而且大家相信，大数据里蕴含着宝贵的、有用的信息。于是，大数据的分析和挖掘，吸引了越来越多的人关注。各行各业更加意识到，谁对大数据的挖掘更为深刻，谁就将抢占未来先机。

这本书，说的是教育大数据。

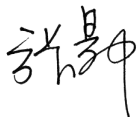
在信息时代，什么都可以用数据来表示。有些可以用小数据表示，有些则要用大数据表示。比如说，一个班里的考试成绩算是小数据，但把一个学生成长过程中多年的作业情况都放在一起，从这个数据集中分析他的成长过程，这就涉及到大数据了。研究大数据，就是设法在海量数据里抽取有用的信息，变大为小。数据量看上去很“大”，但本质的东西可能很“小”。怎么从大量的数据里提炼出本质？就是要靠数学，回归到数学工具。信息时代的本质是数学时代，大数据分析需要数学基础。

未来的教育和学习将进入大数据驱动的新时代。大数据时代，教育将从“用经验说话”转到“用数据说话、用数据决策、用数据管理、用数据创新”。教育从业者通过大数据来分析学习进程和结果，进一步改善教学的方式与方法，有助于实现“改善学习成果、促进

自主学习”，有利于大面积地提升教与学的效果，提高学生的素质。

在本书中，作者提出了对教育大数据的三层理解：全学习过程、教育大数据系统和教育大数据共建共享思想，并用通俗易懂、颇有趣味性的语言和丰富的实际案例来佐证自己的观点。书中讲了可汗学院的学习分析系统、学生利用 Knewton 实现适应性学习、基于大数据的个性化自适应学习过程等案例，并详细介绍了教育大数据处理方法、过程和工具。此外，作者从系统科学的角度对教育大数据发展的社会环境、教育大数据的技术体系框架、教育大数据的应用服务以及教育大数据的载体及应用场景进行了阐述。

作者方海光博士是位年轻有为的学者。他在移动学习、智慧教育、教育大数据等领域耕耘多年，有诸多研究成果和佳作面世。这本《教育大数据》是他的新近力作。读来观点新颖，案例丰富，语言幽默，可读性强，可供关心大数据的研究者参考，更值得广大一线教师、学校管理者和相关行业人士阅读浏览。



2016年10月12日

推荐序 2 教育大数据：开启智慧教育之门

大数据正日益对全球生产、流通、分配、消费活动以及经济运行机制、社会生活方式和国家治理能力产生重要影响。近年来，国家高度重视大数据的发展，政府也全面提升对大数据产业的支持。国务院于2015年8月印发《促进大数据发展行动纲要》，明确指出要全面推进我国大数据发展和应用，加快建设数据强国，并系统部署了在未来5至10年内大数据在各领域的具体应用与任务工作。工信部已完成《国家大数据产业“十三五”发展规划》编制，全国30多个省市出台了大数据相关政策文件，以支持大数据技术和产业创新发展，提升产业支撑能力，培育新业态、新模式。我国大数据产业已步入高速发展阶段。大数据技术正在驱动社会各行业创新并持续激发商业模式创新，催生新业态，已成为促进传统企业转型升级、业务创新增值，提升企业核心价值的重要驱动力。

在互联网技术的支撑下，大数据的潜在价值不断被分析、挖掘和预测。它在悄然改变世界，也在改变教育。随着《关于“十三五”期间全面深入推进教育信息化工作的指导意见》的发布，教育大数据备受瞩目。

教育发展追求的是学生的个性化发展，而不是标准化发展，世界各国的教育改革都倡导针对学生个体差异实施个性化教学，力争做到像医学一样根据学生的学习行为大数据来调整教学策略，在课堂教学中采用基于证据的教学，这一直是人们所追求的未来教育模式。而目

前教育信息化面临最严重问题是应用不深入，需要深度融合与创新。教育大数据是教育产业发展的重要战略方向，可以通过教育大数据实现用科学的数据来引导教育发展。

教育大数据可以真正实现高度集成的资源共享，为教育决策提供支持，改变传统的课堂教学和学习，实现多样化的教学工具与教学方式，实现无地域、无时间限制的在线学习，提升互动体验，真正实现学生个性化的学习、发展，最终实现“智慧教育”。

过去，在学校的教育教学活动中开展实证教学受到技术水平的限制，要收集和处理学生在教学活动中产生的大量数据比较困难。而今，随着全社会移动终端的普及、云计算服务的发展和大数据分析技术的突破，基于学生行为大数据分析的教学将逐步变成现实。教育大数据技术的应用，其战略意义不在于掌握庞大的数据信息，而在于对这些数据进行专业化处理，使其得以增值。在“智慧教育”时代，需要的正是这样的技术。教育大数据可以让教育者真正读懂学生，让“智慧教育”更具智慧。大数据使“经验式”的教学模式变为“数据服务式”的教育模式，将推动传统“以教师为中心”的教学方式向“以学生为中心”的教学方式的转变，推动教师从“演员型”向“导演型”转型，从宏观群体走向微观个体。教师可以应用数据关注个体学生的微观表现，通过个体学生相关数据的分析，有针对性地调整教育方案，从而实现个性化教育。

在本书中，作者创新性地提出了对教育大数据的三层理解：全学习过程、教育大数据系统和教育大数据共建共享思想，介绍了教育中运用大数据技术的实例和教育数据处理的相关流程和方法，介绍了常用的教育大数据支持工具，深入讨论了大数据给教

育产业带来的深刻影响，为读者展示了大数据改变教育教学的美好前景。

全书内容丰富，观点新颖，可供关心新兴技术对教育发展之影响的研究者参考，更值得广大一线教师阅读学习。

黄荣怀

北京师范大学教授

2016年10月10日

前 言

互联网和信息技术的迅速发展及普及，标志着数据技术时代的来临。人们生活在一个充满数据的世界里，信息数据量正呈指数增长，由此开启了大数据时代：通过挖掘和分析海量数据，可以获得其背后隐藏的巨大价值，从而促进应用服务的细分化和精准化。因此，如何管理和使用大数据，如何在垂直领域深度应用，已经成为国内外关注和研究的焦点。

教育大数据是大数据在教育领域的具体表现形式，可以理解为教育领域的大数据，也可以理解为大数据在教育领域的应用，为新时代的教育教学创新提供了新的思路和方法。教育大数据直接产生于各种教育活动中，与传统教育数据相比，教育大数据的采集具有更强的实时性、连贯性、全面性和自然性，其分析处理更加复杂和多样，应用更加多元、深入和个性化。通过对数据有选择性地采集、存储与分析，可以在提升教育质量、促进教育公平、实现个性化学习、优化教育资源共享、辅助教育科学决策等方面发挥有效作用。

实际上，随着人工智能、数据挖掘、机器学习、移动互联网等技术的发展，教育领域的数据处理与应用正逐步迈向成熟。国家已经相继出台了多份教育大数据相关的文件和行动指南，各省市也在逐步建设各级各类教育大数据中心，为数众多的中小学和幼儿园正逐步接触到数据驱动的教学设计、学习分析技术、数据可视化等多种创新应用模式。通过 ICT 技术驱动教育教学改革，特别是互联网和移动互联网与教育的深度融合，使得诸多新型教育模式更离不开教育大数据的支

持。本书作为教育大数据的先导类书籍，较为全面地对教育大数据相关内容进行了介绍和阐述。希望本书能够起到抛砖引玉的效果，对教育工作者有所帮助，也能为相关研究者在今后深入这一领域奠定一定的基础。

本书汇集了作者团队及数字化学习实验室过去多年的研究积累，特别是在 MOOC、教育资源标准、微课程、绩效技术、个性化学习、智慧学习环境、平板教学、移动互联网教育等研究领域的成果。本书内容新颖，深入浅出，力求理论与实际相结合，有一定深度、广度，又不失通俗性，既有对教育大数据的全面概括介绍，又有对典型应用场景的详尽案例剖析，很适合来自中学小学、幼儿园的教师及管理者参考学习，也适合各级各类教育管理部门和行政单位的老师们参考，还适合广大企业和研究机构的教育工作者们使用。同时，对于高等院校教育学本科生、研究生来说，本书可以作为教育大数据课程的入门教材。

作者深受中国科学院院士张景中先生相关研究以及相关思想的启发和影响，这些先进理念为本书的编写奠定了坚实的基础。同时，在本书策划以及编写过程中，承蒙北京师范大学教育学部黄荣怀教授的不断鼓励，并提出了许多宝贵的修改意见。此外，本书的编写得到了来自北京师范大学智慧学习研究院、首都师范大学教育技术系、中国科学院、国家开放大学等单位同仁的大力支持，也特别感谢作者团队成员吴淑苹、罗金萍、陈俊达、杜婧敏、仝赛赛、高洁和李维杨等的奉献和付出。最后，还要感谢作者的父母以及妻子儿女的默默付出。谨以此书作为向以上各位表示衷心感谢的小小纪念！

当然，本书难免有不当之处，敬请广大读者不吝赐教！

方海光

2016年10月12日于北京

目 录

推荐序 1

推荐序 2

前言

第 1 章 绪论	001
1. 智慧城市，如火如荼地在路上	001
2. 智慧教育，步履稳健地前进着	002
3. 智慧学习环境，那是一个关心我学习感受的教师	003
4. 互联网教育带来了什么	005
5. MOOC、可汗学院和教育大数据是怎么回事	007
6. 我们未来的学习会是什么样子	009
第 2 章 教育大数据	011
1. 轰轰烈烈的大数据要干什么	011
2. 教育大数据就像蒙娜丽莎的微笑一样神秘吗	016
3. 教育供给侧改革和教育大数据相关吗	027
4. 国内外专家和实践者是如何评论教育大数据的呢	028
第 3 章 全学习过程——教育大数据的第一层理解	031
1. 什么是学习行为	031
2. 自己在学习的时候，有人会偷偷跟踪我的学习吗	032
3. 学习像蚂蚁爬行一样，会留下学习轨迹吗	036
4. 学习不是走迷宫，你可以找到自己的学习路径	037
5. 电子学习档案袋，也许留下了一丝云彩	038
6. 知识地图是一个导航系统	039
7. 数据不整合，就是一团乱麻	040

第 4 章 教育大数据系统——教育大数据的第二层理解	··· 042
1. 云计算就像使用电一样，插上插座就可以使用互联网	··· 042
2. 数据中心是一个大机房吗	····· 043
3. 是时候把数据中心换一个更好的地方了	····· 044
4. 教育大数据中心的数据是一个可以提取价值的金矿	····· 046
5. 一个非教育领域的综合例子：京东大数据平台	····· 047
6. 一个教育领域的综合例子：美国“希维塔斯学习”	····· 049
第 5 章 教育大数据共建共享思想——教育大数据的第三层理解	····· 051
1. 大数据和云计算其实是一对孪生兄弟	····· 051
2. 钱学森倡导的系统科学也包含着大数据的思想	····· 052
3. 共建共享思想需要打开窗户迎接开放	····· 053
4. 共建共享是大数据时代教育变革的灵魂	····· 055
5. LRS 是一个能够共享自由学习记录的裁判	····· 056
6. xAPI 是一个让你能够自由学习的默默助手	····· 056
7. SCORM 不擅长自由学习，但是共建是它的强项	····· 058
8. CELTS 是成长在这片土地上共建共享的希望	····· 060
第 6 章 教育大数据应用场景	····· 062
1. 未来的教学，躲不开也避不掉大数据了吧	····· 062
2. 大数据为我们的学习带来了怎样的变化	····· 069
3. 管理者准备好使用大数据了吗	····· 073
第 7 章 教育大数据给教育带来哪些变革	····· 076
1. 大数据时代，教师怎么教学	····· 076
2. 大数据时代，学生怎么学习	····· 078
3. 大数据时代，学校怎么管理	····· 082

第 8 章 教育大数据系统环境	086
1. 区域教育信息化与教育均衡发展	086
2. 教育大数据的技术体系框架	089
3. 教育数据挖掘与学习分析	090
4. 教育大数据的应用服务：个性化学习环境	094
5. 教育大数据的重要载体：自适应学习系统	098
第 9 章 教育数据采集	101
1. 教育数据有哪些分类	101
2. 教育数据的结构模型是什么	102
3. 学校、班级、课本和课程是重要的数据平台	104
4. 如何进行数据采集	105
5. 系统日志采集法	108
6. 网络数据采集法	111
7. SQL、MySQL 与数据采集	114
第 10 章 教育数据处理	116
1. 教育数据处理是教育大数据的核心环节	116
2. 数据仓库是数据查询和分析的基础	117
3. 云存储是以数据存储和管理为核心的云计算系统	119
4. 分布式文件系统	121
5. 没有听说过的数据湖泊（Data Lakes）	123
6. 不应该忽视的网络	127
7. 非结构化数据处理	129
8. 数据清洗、数据交换以及数据算法	131
第 11 章 教育数据分析	140
1. 数据分析方法和工具：传统与新兴	140
2. 五大技术助力教育数据分析	142
3. 数据分析：可汗学院的大脑	143

4. 你所不知道的教育数据挖掘	145
5. 五类技术让你玩转数据挖掘	146
6. 聚类分析最常见的应用：用户分类	148
7. 文本挖掘与学习分析	151
8. 学习分析：实现差异化及个性化教学的基础	151
9. 学习分析：周期、层次和关键技术	153
10. 多用户视角下 MOOC 学习分析模式	154
11. 知识图谱是靠谱的	156
第 12 章 教育数据呈现	160
1. 什么是数据可视化	160
2. 不同的数据可视化方法	161
3. 数据可视化需要分几步完成	162
4. 优秀数据可视化技术的核心理念	163
5. 规避错误，建立成功的数据可视化策略	166
6. 数据可视化工具初级篇——零编程工具	167
7. 数据可视化工具开发篇——JavaScript 库	171
8. 数据可视化，推动教学过程的优化	179
第 13 章 教育大数据支持工具	183
1. 新手必备的护法软件，左 Excel 右 SPSS	183
2. 数据处理工具 Hadoop，大数据的代名词	185
3. 你贴心的辅助开发工具，大数据强有力的技术支持	193
第 14 章 教育大数据的发展：机遇与挑战并存	202
1. 任重道远：教育大数据发展的几重挑战	202
2. 见招拆招：教育大数据发展的应对策略	206
3. 路在何方：教育大数据发展趋势	210
参考文献	213