

Big Data

大数据

大价值、大机遇、大变革

李志刚 主编

朱志军 余丛国 闫蕾 等编著

转型时代丛书

中国电信北京研究院专家奉献

大数据

大价值、大机遇、大变革

主编：李志刚

编著：朱志军 余丛国 闫蕾 赵鹏 王若倪 杨明川

插图：罗勇



1044870



T1044870

电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

不管你愿意不愿意，我们每天都在制造和分享数据；
不管你关心不关心，我们已经生活在大数据的包围之中。

大数据，这个在 2012 年突然火爆起来的词汇，我们究竟该如何解读？

本书介绍了大数据产生的背景、特征和发展趋势，从实证的角度探讨了它对社会和商业智能的影响，并认为大数据正影响着商业模式的转变，将带来新的商业机会。能否对大数据进行处理、分析与整合将成为提升企业核心竞争力的关键！

大数据，既是一场大机遇，也将引发一场大变革！

未经许可，不得以任何方式复制或抄袭本书之部分或全部内容。
版权所有，侵权必究。

图书在版编目 (CIP) 数据

大数据：大价值、大机遇、大变革 / 李志刚主编；朱志军等编著. —北京：电子工业出版社，2012.10
(转型时代丛书)

ISBN 978-7-121-18560-1

I. ①大… II. ①李… ②朱… III. ①企业管理—数据管理 IV. ①F270.7

中国版本图书馆 CIP 数据核字(2012)第 221765 号

策划编辑：刘 皎

责任编辑：刘 皎

印 刷： 中国电影出版社印刷厂

装 订：

出版发行：电子工业出版社

北京市海淀区万寿路 173 信箱 邮编：100036

开 本：787×1092 1/16 印张：15.25 字数：350 千字

印 次：2013 年 2 月第 2 次印刷

印 数：4001~6000 册 定价：78.00 元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及邮购电话：(010) 88254888。

质量投诉请发邮件至 zlls@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线：(010) 88258888。

转型时代丛书

○ 指导委员会

主任委员：吴基传

副主任委员：杨杰

委 员：陈俊亮 李未 韦乐平

邬贺铨 张继平（按拼音顺序排序）

○ 编委会

主 任：李志刚

副主任：侯春雨 赵慧玲

委 员：毕奇 朱健 野永东 谢朝阳

陈自清 杨峰义 王晓平 张成良



转型时代丛书

总序

“变化，无论是突如其来的，还是循序渐进的，有时都会淘汰你认为理所当然的一切”。

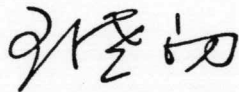
——《转型》

21世纪以来，信息化更加快速而深刻地改变着这个世界，大到全球经济社会的发展格局，小到每个人的日常工作生活。许多国家把数字化、信息化、智能化作为国家战略的关键主题，把信息基础设施建设作为后金融危机时代振兴经济的重要手段。同样，我国“十二五”规划也把全面提高信息化水平，特别是加快建设下一代国家信息基础设施、推动信息化和工业化深度融合、推进经济社会各领域信息化作为重要工作列入其中。

信息通信产业中新技术、新业务的不断快速发展不仅催生着新的经济增长点，造就了谷歌、Twitter、腾讯等一个又一个明星企业，引领整个行业及社会经济的发展方向，更重要的是它对人们生产、生活产生了深刻而久远的影响。我们的生产资料不仅仅是机器，还有电脑、手机、互联网；我们通过单击“百度”打开未知世界，通过“淘宝”购买商品，利用手机登录“Facebook”去了解彼此、评论时政，所有这一切都表明信息通信产业正在更广、更深地影响着我们每一个人，互联网 / 移动互联网已成为像水、电一样的生产、生活“必需品”。

环顾全球，整个信息通信产业正在朝着宽带化、移动化、智能化发展，特别是3G的普及和LTE的逐步成熟，使得移动互联网一跃成为整个行业中最前沿、最具革命性的领域。智能管道、物联网、下一代互联网、云计算等一个个新的理念、新的信息服务模式正在席卷全球，成为新热点。而这一切的变化，都将为从事信息服务的企业，包括电信运营商，带来前所未有的机遇和挑战。适者生存法则同样适用于多变的企业生态系统。无论是百年老店，还是创业新秀，只有顺应信息化时代发展潮流，重新审视并及时调整企业的商业模式，抓住信息化带来的重大机遇，才能在变化中顺势前进。

鉴于此，这套“转型时代丛书”既有对智能宽带网络、移动互联网、云计算等新技术、新网络的研究和实践总结，也有对商业模式、营销变革等现代管理中关键问题的长期探索。相信此系列书籍能帮助你了解趋势，廓清谜团，抓住机会，与信息化时代共同成长。



2012年1月

9/18/16



一切始于紧迫感

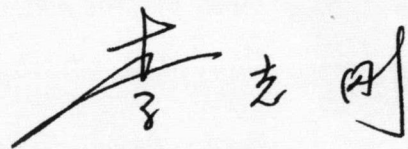
自2012年3月华盛顿宣布“大数据研究和发展倡议”，英国、日本、德国、加拿大等国纷纷效仿，推出与大数据应用相关的战略研究，自此一场关于“大数据”的战略争夺拉开战幕。

能源在消耗中面临枯竭，从而引发各国对能源的争夺和对危机的恐惧；数据不但不会因为占有而枯竭，还会随着使用和传播不断的丰富与生长。可以想象，数据将成为未来世界的DNA。但拥有的数据规模、质量和应用带来的马太效应将形成明显的数据资源区隔和竞争优势，因此大数据成为各国争夺的下一个战略前沿。这是一场没有硝烟的战争，随着“智慧经济”的崛起，采集数据、掌握数据、运用数据必将成为国家、企业的核心竞争力。

大数据正在为我们呈现一个全新的信息社会、智慧社会，世界经济在迎来新引擎的同时，格局也酝酿着新变化。面对历史的又一次机遇，我们将扮演何种角色、如何应对，这是值得我们每个人认真思考和探究的问题。

本书从社会、经济、商业和技术等多角度展现了大数据将带来的改变和促进，同时对其发展趋势、实施中的问题和风险也做了思考。

行胜于言，我们用最短的时间完成此书的编撰就是希望每位阅读此书的读者都能从中得到触动和借鉴，从而关注大数据及大数据下的发展。

A handwritten signature in black ink, consisting of stylized characters that appear to be '李志' (Li Zhi).

2012年9月

引言

“这是最好的时代，也是最坏的时代；这是智慧的时代，也是愚蠢的时代”。每天，有无数的企业在建立，也有无数的企业在消失。互联网让信息和数据快速积累和流动，世界变得更透明，更平坦。

谁能想象五年之后，一个不懂数据的公司如何生存和发展？数据革命正在瓦解已经建立的产业和商业模式，你、我都准备好了吗？

用数据打造冠军球队

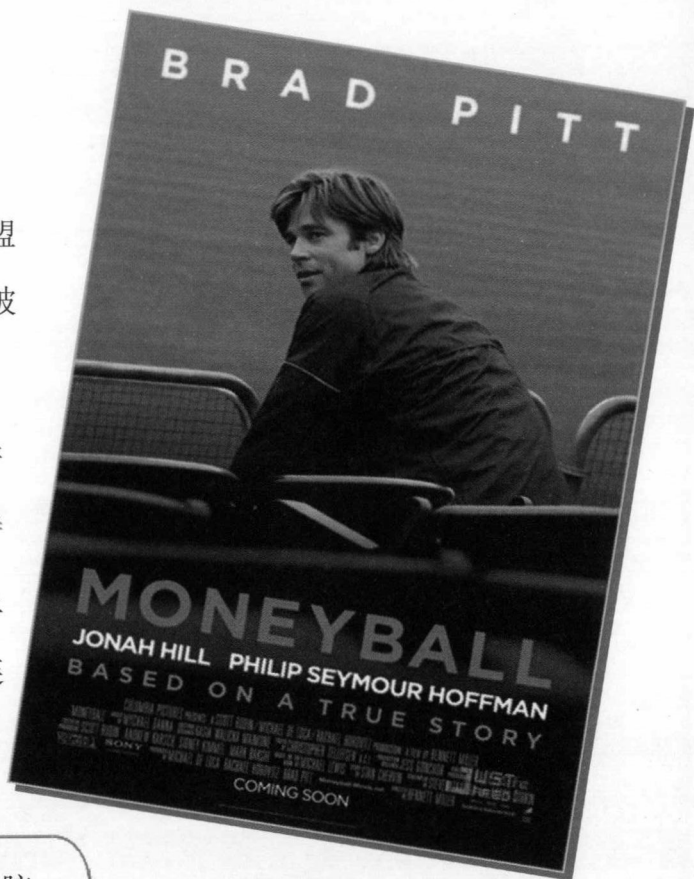
剧情背景

奥斯卡获奖影片《点球成金》讲述了一个美国职业棒球大联盟中的真实事件。奥克兰勇士队在刚有起色时，球队主力纷纷被重金挖走，资方投入严重不足，新赛季的前途非常渺茫。

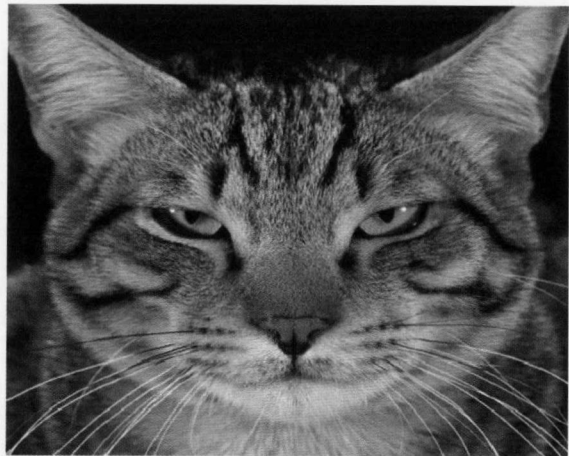
此时球队经理另辟蹊径，打破原有球队的经营理念，查询所有球员的历史数据，利用数学建模，定量分析不同球员的特点，合理搭配，重新组队，颠覆棒球界传统的靠重金挖明星球员组队的理念。在新的赛季中，奥克兰勇士队创造20场连胜的新历史。

颠覆

打破传统依靠明星球员的思维，运用电脑程序剖析每个球员的特点和比赛数据，利用数据分析寻找“性价比”最高的球员，在有限的资源下“运用数据”获得成功。



谁说机器不识猫



2012年6月26日，Google X部门通过1.6万片处理器构建了一个庞大的系统，用于模拟人类的大脑神经网络。

借助“谷歌大脑”，无需接受人类的任何培训和指令，就可以利用内在算法从海量数据中自动提取信息，学会如何识别猫咪。

通过“不需要借助大批研究人员找出事物之间的差异，只要为算法提供海量数据，让数据自己说话，让软件自动学习数据。”这一理念，“谷歌大脑”在大数据的海洋中不仅在发现知识，更在不断地生成知识和形成智能。

创新

谷歌的科学家将1.6万片电脑处理器连接起来，创造了全球最大的神经网络之一，不需要人为的教学，机器可以在互联网中“自学成才”。

目录

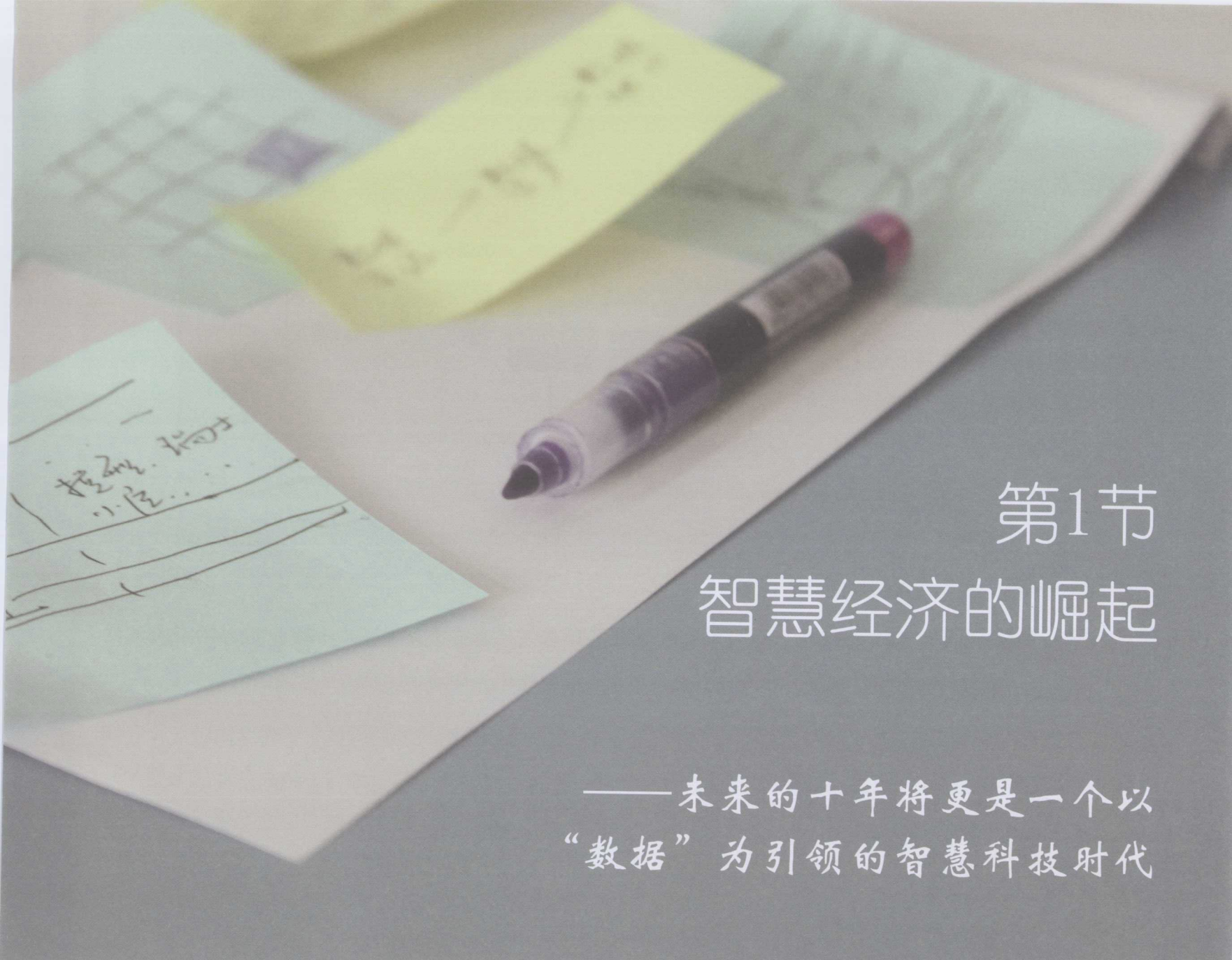
#1	第1章 大数据汹涌来袭	1	#5	第5章 带来数据处理新变革	149
	第1节 智慧经济的崛起	3		第1节 大数据的关键技术	152
	第2节 大数据时代已经来临	7		第2节 大数据与云计算	178
#2	第2章 大数据的前世今生	25	第3节 大数据技术的发展趋势	184	
	第1节 认识大数据	26	#6	第6章 未来之路	191
第2节 大数据能带来什么	37	第1节 未来：从信息化的发展趋势中探寻		193	
#3	第3章 引领社会进入“大数据时代”	49		第2节 创新：大数据带来的新应用和新领域	202
	第1节 大数据对国家、社会的作用	50	第3节 挑战：企业应用大数据的问题与风险	213	
	第2节 大数据将推动经济发展	57	参考文献	222	
#4	第3节 大数据将推动科技发展进程	86	补记	224	
	第4章 开启商业智能新阶段	97	后记	228	
	第1节 数据分析的发展——从数据到知识	98	作者简介	229	
	第2节 大数据如何让商业更智能	117			
	第3节 大数据应用案例	133			



第 一 章

Chapter One

汹涌来袭

A photograph of a desk with a purple pen, several sticky notes (yellow and light green), and a calendar. The pen is lying horizontally across the center. The sticky notes are scattered around, some with faint handwriting. The background is a dark, textured surface.

第1节 智慧经济的崛起

——未来的十年将更是一个以
“数据”为引领的智慧科技时代

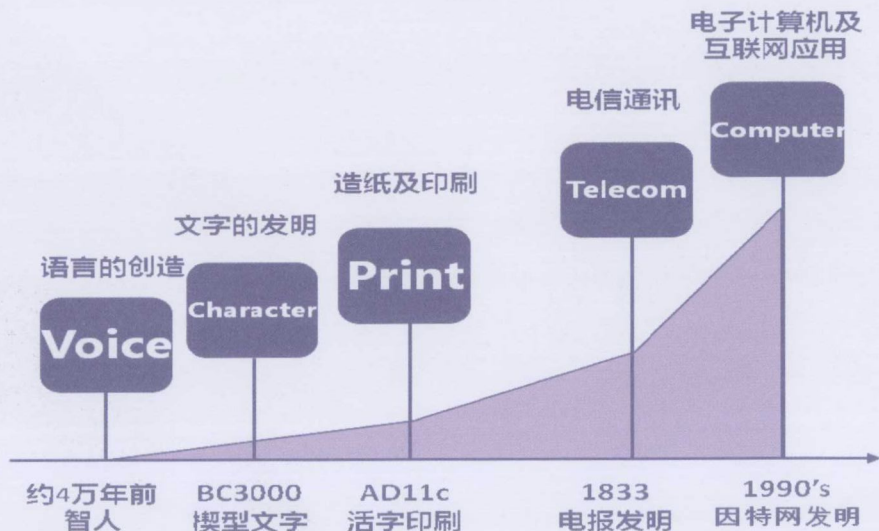
20世纪下半叶以来，信息技术日益显示出强大的生命力，在过去的50多年里，给人们的生活带来了巨大变革。信息经济逐渐成为国民经济的主导力量，成为推动整个社会前进发展的主要动力。据经济合作与发展组织（OECD）统计，2008年全球信息产业收入约3.76万亿美元，25%的GDP增长和40%的劳产率提升归功于信息产业的发展。

人类社会经过五次信息技术革命，信息的传播能力快速增强。计算机的发明和互联网的应用，正以空前的影响力和渗透力，不可阻挡地改变着社会的经济结构、生产方式和生活方式。

随着美国“智慧地球”概念的提出，通过更多的感知芯片装入，能够透彻地感应和度量世界的本质和变化；并通过互联网技术促进全面的互联互通，实现更深入的智能化，帮助企业获得更智能的洞察。

人类社会进入“后信息经济时代”，
智慧经济正在崛起。

人类社会经历5次信息技术革命，传播能力快速增长



作为后信息经济时代的新阶段，“智慧经济”正在崛起。智慧不仅反映了创新经济、创意产业成为整个社会经济的龙头产业，而且在微观层面上也反映了企业的所有商业决策过程更加智能。

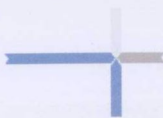
智慧经济是一种创新形态的萌生经济，不但表现在经济上各种要素（资金、劳动者、技术）、各个领域（生产、流通、交换、消费）、各个主体（国家、企业商、家庭或个人）间的最佳结合，而且还能实现经济与环境、能源、社会的最大限度的可持续发展。智慧经济的生产方式是以智力资本为中心的，所以对于信息乃至知识的掌握成为整个智慧经济的核心和基础。



数据的分析和应用是 “智慧经济”的基石

随着计算机的发明和应用，信息的承载从语言到文字再到现在的代码或数据，不仅提升了信息描述的准确性，更扩大了信息传递的广泛性。信息的数字形式是信息交流领域的“上帝”，无论什么样的信息，都要转化为数字形式，即0和1，它是信息世界的蒸馏器，信息的差别不见了，数字成了信息的结晶体，信息被吸收进去，然后再被抛出来。

在信息技术中，相对于信息（或数据）采集、存储而言，信息（或数据）的加工、处理和传递显得更为重要。信息的加工与处理（算法）能产生新信息，而传递则大大扩展了新信息的影响和接触范围，让信息在更广阔的空间产生价值。



计算机技术和网络技术的发展让信息技术有了根本性的革新，不仅有效地改变了以往高成本、低效率的信息处理局面，更加快了信息的流动和传播。

