



Data Technology Industry  
Innovation Institute  
DT大数据产业创新研究院(DTi)

赢在大数据系列丛书

# 赢

在

# 大数据



陈新河 © 主编

金融/电信/媒体/医疗/旅游/数据市场行业  
大数据应用典型案例

 中国工信出版集团



电子工业出版社  
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY  
<http://www.phei.com.cn>



Data Technology Innovation Institute  
DT大数据产业创新研究院(DTii)

赢在大数据系列丛书

# 赢

在



# 大数据



陈新河 © 主编

金融/电信/媒体/医疗/旅游/数据市场行业  
大数据应用典型案例

电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

## 内容简介

大数据已经由一个概念变成一种工具和行动,与我们的生活时时相伴,推动着产业转型升级。它正在逐渐成为一种思想和文化,影响我们的日常行为和做事方式,由经验式拍脑袋决策转变为让数据说话。从产业层面来看,中国的大数据产业从每年的千亿元级数据营销领域快速向万亿元级金融、保险、电信、制造、能源、工业、农业和政府等行业拓展。相对过去10年的数据营销领域而言,行业大数据应用才刚刚起步,缺乏可借鉴、可参考的案例已成为行业大数据应用的瓶颈。“赢在大数据系列丛书”通过收集整理,并从216个案例中遴选出覆盖18个行业的106个行业大数据最佳实践案例,再加上内容全面、系统、深入的《赢在大数据:中国大数据发展蓝皮书》一书,希望为业界提供应用大数据、实践大数据的参考和借鉴。

由DT大数据产业创新研究院(DTiii)集合上百家企业、聚集上百名一线业内专家学者编辑出版的第一批“赢在大数据系列丛书”共计5册。

- (1)《大数据领导干部读本(第二版)》;
- (2)《赢在大数据:中国大数据发展蓝皮书》;
- (3)《赢在大数据:政府/工业/农业/安全/教育/人才行业大数据应用典型案例》;
- (4)《赢在大数据:营销/房地产/汽车/交通/体育/环境行业大数据应用典型案例》;
- (5)《赢在大数据:金融/电信/媒体/医疗/旅游/数据市场行业大数据应用典型案例》。

本书精选金融、电信、媒体、医疗、旅游和数据市场6个行业的41个大数据应用最佳实践案例,通过翔实的案例剖析,系统分析了大数据在这6个行业中应用的技术、理论、实践和方法。无论是大数据从业者,还是政府产业管理者,抑或投资者,均可从本书中获得可借鉴的信息。

“赢在大数据系列丛书”将持续出版,以加速推进中国大数据生态发展。

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

赢在大数据.金融/电信/媒体/医疗/旅游/数据市场行业大数据应用典型案例/陈新河主编.—北京:电子工业出版社,2017.5

ISBN 978-7-121-31176-5

I. ①赢… II. ①陈… III. ①企业管理—数据管理—案例—中国 IV. ①F279.23

中国版本图书馆CIP数据核字(2017)第063971号

策划编辑:刘淑丽

责任编辑:刘淑丽

印刷:北京盛通印刷股份有限公司

装订:北京盛通印刷股份有限公司

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:22.25 字数:514千字

版次:2017年5月第1版

印次:2017年5月第1次印刷

定价:68.00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)88254888,88258888。

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

本书咨询联系方式:(010)88254199, sjb@phei.com.cn。

## 反侵权盗版声明

电子工业出版社依法对本作品享有专有出版权。任何未经权利人书面许可，复制、销售或通过信息网络传播本作品的行为；歪曲、篡改、剽窃本作品的行为，均违反《中华人民共和国著作权法》，其行为人应承担相应的民事责任和行政责任，构成犯罪的，将被依法追究刑事责任。

为了维护市场秩序，保护权利人的合法权益，我社将依法查处和打击侵权盗版的单位和个人。欢迎社会各界人士积极举报侵权盗版行为，本社将奖励举报有功人员，并保证举报人的信息不被泄露。

举报电话：(010) 88254396；(010) 88258888

传 真：(010) 88254397

E-mail: [dbqq@phei.com.cn](mailto:dbqq@phei.com.cn)

通信地址：北京市万寿路 173 信箱

电子工业出版社总编办公室

邮 编：100036

# 丛书序

## 数据驱动中国，创新成就未来

2017年1月6日，碧桂园集团信息管理中心的郑桂清通过微信询问我能否联系到上海脉策数据科技有限公司的CEO高路拓，接洽房地产大数据项目合作事宜。加上可视化项目、房地产解决方案项目，这已是本月他第三次来寻找大数据厂商信息了。2016年10月18日，A文数据公司的Gall在微信公众号“软件定义世界（SDX）”后台说，他们是一支从美国回来的大数据创业团队，在美国XN公司具有丰富的销售预测实战经验，在数据科学领域具有很深的积淀，已经开始做一些电商货品销售预测的项目，但对创业方向有些困惑，想寻找一些新的突破口，了解哪些行业目前进入比较好，以及估值10亿元以上的大数据企业都在做什么。你可能以为碧桂园的大数据应用之路才刚刚开始，其实它已经上马3年有余了；你或许认为A文数据公司刚成立，其实它已经拿到了数千万元的A轮融资了，团队也已有30多人。它们有的需求应该也是成百上千家企业的需要，这也是我们编纂“赢在大数据系列丛书”的初衷：行业用户需要学习借鉴行业最佳实践案例，来指导本企业；创业企业需要深入洞察行业、了解同行，进行深入交流合作。

类似这种用户寻找供应商、创业企业寻求市场或合作的事情近3年来对我来说天天都会发生。从某种程度上说，我已成了“大数据驿站”（Big Data Hub），缘由还要从3年前的大数据社区开始。2014年2月，中关村大数据产业联盟发起“大数据100分”活动。“大数据100分”聚焦为大数据社区活动，每晚21:00—22:40，由一名大数据创新创业企业的CEO、CTO或行业应用企业的CIO、CDO，在500人微信大数据专业社群分享PPT及文字内容，分享内容经嘉宾同意后，在微信公众号“软件定义世界（SDX）”公开发布。该500人微信群每人必须分享一次，坚持众智、众享的理念，谁愿意分享便会被邀入，不分享便被请出，经过数轮迭代，几乎聚集了国内所有主流大数据创业企业的CEO和CTO，以及部分行业用户的CIO。自从活动开始以来，我组织近300场“大数据100分”活动，微信群也由最初的一个500人大数据专业社群，扩展至金融、工业4.0、汽车、房地产、消费、旅游、农业、政府、个人信息保护等十多个大数据专业社群，个人微信号也不得不由一个扩展至三个，微信好友有12000多人，再加上微信公众号“软件定义世界（SDX）”近10万名专业订阅用户的需求与合作反馈，你能想象我的微信有多忙，我刚买3个月的128GB

内存的手机容量已经告警。“大数据 100 分”活动影响了数百万人，在某种程度上完成了中国大数据的启蒙，影响之广远超乎我的想象。2015 年 9 月，我在飞往上海的航班上，正在整理下午要用的 PPT，邻座的陌生人突然跟我打招呼：“您是‘大数据 100 分’的主持人吧！‘大数据 100 分’的活动内容我经常看，很受启发，希望能有机会进一步合作。”原来，他是西安陕西微软创新中心的王真，看到我在做 PPT 上的部分内容，希望在旅游大数据方面帮忙寻求合作方。

2014 年下半年开始，有几个出版社询问能否把“大数据 100 分”结集出版，我精选了 63 个案例，共计 890 页、29.89 万字，想来想去，还是决定不出版了，原因有很多。一是这些内容都已经在网络上公开，再出版意义不大。二是有些解决方案还难以落地，认知类内容多，实战案例内容少；洞察性案例多，解决业务实际需求少；探索性案例多，解决方案少；单一项目多，产品方案少。三是市面上已经有太多大数据的书籍了，选题不慎将导致销量有限。四是网络、报纸、新闻、媒体上的大数据信息铺天盖地，大数据的基本理念已经被灌输得差不多了。五是各种大数据的研讨会、沙龙、论坛、峰会多如牛毛，一些案例被重复讲来讲去。六是一般大数据类书籍也就 3 000~5 000 册的销量，还不如“软件定义世界（SDX）”一篇稿子的阅读量，再出版一本大数据的书有点画蛇添足。但之后发生的事情，让我对大数据类书籍出版有了新的看法。

2014 年 9 月，我有幸参与《促进大数据发展行动纲要》的编制，同期，参与的“大数据与国家治理”课题成果初具雏形。2015 年 4 月，中国行政体制改革研究会常务副秘书长、课题组组长王露建议一鼓作气，把课题组已经有的研究成果进行综合，再进一步深入调研、研究和分析，编写一本《大数据领导干部读本》。一方面，它可以配合国家大数据战略，为各级政府领导干部提供一本深入、系统的学习材料；另一方面，它也是课题组申请到的国家社科基金特别委托项目“大数据治国战略研究”课题的重要研究成果。同时，还可以通过调研更深入地了解国内大数据应用情况，为课题组下一步建言献策做好准备。我说出了我的担忧，当前市面上有 1 000 多本大数据书籍，除了迈尔-舍恩伯格（英）的《大数据时代》销量百万册、涂子沛的《大数据：正在到来的数据革命》销量在数十万册之外，一般的销量也就几千册，10 000 册以上就算销得不错的了；再说，认知类书籍已经过了黄金时间窗口，迈尔-舍恩伯格和涂子沛两本畅销书占据了最佳时间窗口。经过课题组的再三讨论，最后得出结果：出一本系统分析大数据内涵和外延、国内最新大数据在政府治理方面的实践案例的书应该能销上万册。于是，课题组组织 200 多人，经过 3 个多月的奋战，《大数据领导干部读本》终于面世了。这本书发行出奇的好，1 年销售近 8 万册，在当当、京东上的整体好评率超过 99.5%，很多读者对案例部分评价很高：“里面有彩图，质量不错，内容丰富，值得读”、“内容很丰富，高大上！彩色印刷，喜欢！案例也不错”、“一下买了一百本，打算送给客户的”、“经典的书籍，值得一看再看，推荐购买”、“喜欢，挺好的，个人觉得是领导干部读本里面写得最好的”。

因为主持“大数据 100 分”的缘故吧，我受邀到各类组织分享、交流和培训大数据比较多，一方面通过交流推广一下大数据的思想和理念，另一方面深入了解各行各业对大数据的一线需求，同时也为这方面的研究积累素材。3 年来，各类分享、交流和培训累计超过 200 场，既有国土资源部、苏州工业园、宁波市经信委、深圳市龙岗区、兰州市大数据社会服务管理局等政府组织，又有华为南京研究所、上海移动、北京联通、中国电信、华夏基金、南方基金等 IT 企业、运营商、金融企业，还有中国石油、国家电网、北控集团、花样年集团等传统企业，交流后大家都会问到这些问题：本企业也有大数据应用方面的需求，但苦于不知道找谁来实施，也不知道同行业其他企业应用大数据的情况，如果能有本行业的最佳实践案例做指导，或者其他行业的案例拿来借鉴就好了。这些反馈说明，一方面，大数据认知教育活动还不够，如果把“大数据 100 分”精华版结集成书，全面、系统、深入介绍大数据在各行业的应用，辅以内案例，推动应用的效果会更好些；另一方面，行业用户已经从认知教育阶段，开始进入尝试应用阶段，需要行业最佳实践案例来进行引导。

“大数据 100 分”虽然影响了数百万人，但还是一个小圈子，更具体地说是 IT 圈子，也就是乙方圈子，对项目的应用方——甲方影响有限，否则我的分享、交流和培训也不会多在甲方群体中了。信息化的发展历程也证明，乙方的觉醒和认知更快些，甲方往往是被推着走；但甲方被动地说服使用的进程，远不如主动要求“我要使用”快。对于乙方而言，目前大数据书籍介绍国外案例多，国内案例少；案例浅尝辄止，不深入、不具体、不详细，距离实际操作相距甚远，缺少借鉴的可能性；另外，也无法知道供应商是谁。除了 4 个“V”、只要相关不要因果，就是啤酒与尿布的故事，具体如何进行数据分析、建模、系统建设等信息寥寥。

在组织“大数据 100 分”活动及走访上百家大数据创新创业企业的过程中，我了解到很多大数据应用的创新案例，非常精彩，商业价值非常明显，用户非常认可，非常值得在全行业推广。尤其是 2015 年下半年以来，很多案例逐渐从项目行化成标准的产品，逐渐在行业用户中规模应用起来。上一波 IT 驱动流程的信息化 1.0 过程中，“金字工程”、“甩图板工程”的行业最佳实践案例起到了很大的“引路人”、“导航灯”作用；这一波 DT 驱动业务的信息化 2.0 也同样需要行业最佳实践案例，人类最善于模仿，历史很多是重复过去。

2016 年，在我近 20 年的软件产业研究基础上，以及近几年大数据研究、专家组织、业内资源积累的基础上，筹建起大数据智库——DT 大数据产业创新研究院（DTiii），专注于大数据生态体系研究和建设。DT 大数据产业创新研究院（DTiii）于 2016 年 11 月，在第三届世界互联网大会·乌镇峰会上发布《2016 中国大数据应用最佳实践案例 TOP50》，案例发布后引起业界强烈反响：行业用户想借鉴最佳实践案例的成功经验，大数据创新创业企业想学习同行的案例经验；风险资本想找好的苗子并做横向比较；政府想了解本地大数据企业在业界的位置；业界想通过案例进行深入学习；出版界想拿到最好的书稿。乌镇峰会后，我们又进一步广泛征集，行业从 15 个扩展至 18 个，最佳实践案例从 50 个扩展

至 106 个，这 106 个案例是从 219 个上报案例中经过初选、意见反馈、修改、专家评审最终确定的。

希望“赢在大数据系列丛书”能够加速中国大数据万亿元生态的快速形成，更希望成为数据驱动中国创新发展道路上的一盏明灯。

DT 大数据产业创新研究院（DTiii）院长、中关村大数据产业联盟副秘书长

陈新河

2017 年 1 月 31 日



# 前言

## DT驱动中国新世纪

DT时代已经扑面而来！2017年春晚，CCTV主持人董卿称“利用大数据，统计目前总共有1.04亿位观众观看了2017年春节联欢晚会，而移动端用户所占比例超过70%”。估计不少电视机前的观众听后，肃然起敬，内心不懂也不好意思多问边上的人，就把大数据当作20世纪80年代媒体上到处出现的电脑一样的高科技家族！很明显，大数据在此得到了运用，不少模拟电视用户、单向广播式接收的数字电视用户数量还只能用小数据的方式进行统计和估算。

暂且不论1.04亿位观众数量的准确性，毫无疑问，大数据已经与我们如影随形、时时相伴，与我们每个人的吃喝玩乐娱住行分不开了。当你打开微信，微信页面底部的广告条内容是基于大数据营销推荐引擎的；当你打开电脑浏览器浏览网页时，嵌入页面的广告内容是基于大数据推荐的，它会根据你过去一段时间的浏览记录在浏览器中的Cookie信息来向你推荐零食、服装或智能手机；出门叫个专车，看似简单的一个匹配，其实后面也蕴含着复杂的模型和算法，大数据在位置、距离、评分、拼车等多个参数约束下不停地运算。3年之后的2019年，你的车辆保险费很可能与你的驾驶行为数据相关，如果你的日常驾驶行为数据显示你是“三急”（急转弯、急刹车、急加速）用户，那么你的保险费用可能比别人要高30%，甚至更多；2020年的某天你在王府井步行街漫步时，你看到的路边广告可能与别人不同，广告系统能够自动捕捉你随身携带手机的硬件标识信息，然后根据该硬件标识信息，结合后台的大数据画像系统，从而显示你最关心的内容。

从产业层面来看，善用数据者，将占领竞争的制高点。在“赢在大数据系列丛书”房地产大数据应用案例篇，一家房地产企业通过利用大数据和不用进行对比，效果相差明显。其中，来电效果是2.5：1，到访效果是3：1，成交效果是3：0。效果对比后，客户立即决定2017年加大DT投入预算。从宏观层面来看，大数据已成为国家基础性战略资源，正日益对全球生产、流通、分配、消费活动，以及经济运行机制、社会活动方式和国家治理能力等产生重要影响。党的十八届五中全会明确提出要实施“国家大数据战略”，标志着大数据战略正式上升为中国的国家战略，政府和企业将利用大数据共同推进产业全面改革和升级。

大数据是什么？大数据是物理世界在网络世界的客观映射，是物理世界的DNA。如

果说过去几千年是人类生活的模拟世界，那么 DT 时代的到来，人类开始进入数据世界。大数据给予人类重新认识自我、认知世界、改造世界的新理念、新思想和新方法。小而言之，生活离不开大数据；大而言之，大数据重构产业，进而重塑世界新格局。

如果说过去百年中国在追赶欧美日工业强国的道路上，受困于既有路径难以超越，以及人均矿产、石油等不可再生资源的稀缺，那么毫无疑问，DT 时代，中国有望在东方崛起，成就中国新世纪。数据成为一种新的战略资源，永不磨灭；13 亿人无时无刻不在产生数据，从人口多的劣势转变为数据优势；中国人的数学智慧在数据科学领域的的数据能力有望超越在工业时代中国因后发劣势所导致的材料科学、制造工艺等工业能力的不足。

当前，中国大数据产业仅数据营销领域产值每年已超过 1 000 亿元，但整体而言仍处初级发展阶段，大数据应用的行业不断拓宽，深度日益加强，应用的重心从数据营销向行业领域扩展，万亿元行业大数据应用市场已经拉开帷幕。相对于已发展 10 年左右的数据营销领域，行业大数据应用才刚刚起步，即使在国外也没有太多成熟的案例可借鉴，基于此，DT 大数据产业创新研究院 (DTiii) 组织国内大数据创新创业企业、一线专家编写了这套“赢在大数据系列丛书”，丛书的编写汇集了近百家企业在 18 个行业领域的 106 个大数据最佳实践案例，在一定程度上反映了大数据在各行业的应用现状和发展趋势。撰写丛书的目的在于，梳理中国大数据应用现状及发展趋势，为行业用户、大数据创新创业企业、政府和从业人员提供参考和借鉴。

DT 大数据产业创新研究院 (DTiii)，聚集了上百名产学研用各界知名大数据专家，专注于大数据生态体系研究，发布《DTiii 版中国大数据产业地图》、《中国大数据发展蓝皮书》、《中国大数据创新创业 TOP100》、《中国大数据应用最佳实践案例 TOP50》、《中国大数据领军人物》等，全面深入洞察中国大数据技术、产品、市场、产业、人才、资本、政策、应用案例、创新创业等，深入研究国内外大数据最新发展动态，分享大数据最新发展模式，全面展示大数据应用最佳实践案例、创新创业企业风采，引领中国大数据产业发展方向，推进中国大数据生态建设。

“赢在大数据系列丛书”力求全面总结大数据行业发展情况，反映新特点、分析新问题、提出新趋势。第一批“赢在大数据系列丛书”共计五册，分别为《大数据领导干部读本》(第二版)、《赢在大数据：中国大数据发展蓝皮书》、《赢在大数据：金融/电信/媒体/医疗/旅游/数据市场行业大数据应用典型案例》、《赢在大数据：营销/房地产/汽车/交通/体育/环境行业大数据应用典型案例》、《赢在大数据：政府/工业/农业/安全/教育/人才行业大数据应用典型案例》。五本书不仅从宏观的行业环境角度为读者提供了全面的视角，也从具体的实践案例角度为读者提供了可以借鉴的方法和手段，以便读者能真正把握大数据的脉搏。

《大数据领导干部读本》(第二版)在第一版的基础上更新了一些案例，仍是强调对大数据内涵和外延的阐释，帮助读者深刻理解大数据的概念，以及大数据对经济社会方方面面的影响，着重强化了政府治理层面的大数据应用案例。考虑读者的多样性，也涉及了一

些行业案例，在政府数据开放、信息安全方面也做了重要安排。从京东、当当网对该书评价来看，该书的特点是通俗易懂、图文并茂、案例丰富。第二版再次继承了这些特点。

尽管该书名曰“领导干部读本”，其实读者对象不仅局限于政府人士，也可作为各行各业大数据入门培训系统教材。从销售跟踪的数据来看，不少组织把该书作为大数据培训教材。因该书出版在前，为保持该书书名的连贯性，仍旧保持《大数据领导干部读本》原名，没有添加“赢在大数据系列丛书”统一前缀“赢在大数据”。

《赢在大数据：中国大数据发展蓝皮书》是行业篇，主要从产业宏观发展和行业应用全貌进行分析和研究，力求产业描述全面、系统，行业覆盖广泛，行业各业务环节大数据应用深入。行业篇既有涵盖1311家大数据企业的详细名单和产业地图，也有产业发展现状和趋势预测分析，同时也对投融资、人才、数据交易，以及存在的问题进行了深入的分析。在每个行业领域，从行业的痛点出发，总结了大数据在各行业应用的概貌，分析了利用大数据解决行业痛点的手段和方法，也介绍了各行业已经成熟的大数据技术框架和模型、应用现状、存在的问题及发展趋势，从大数据产业整体发展情况进行详细分析，并描述产业未来的发展蓝图。对数据营销、金融、电信、房地产、政府、制造、农业、医疗健康、零售、影视、旅游、教育、人才、体育等14个行业的整体发展进行了深入、系统、完整的分析。

《赢在大数据：金融/电信/媒体/医疗/旅游/数据市场行业大数据应用典型案例》、《赢在大数据：营销/房地产/汽车/交通/体育/环境行业大数据应用典型案例》、《赢在大数据：政府/工业/农业/安全/教育/人才行业大数据应用典型案例》这三本书是翔实的案例分析，详细介绍了各行业如何结合行业及大数据的特点来解决行业痛点，如何形成具体的行业数据解决方案及这些方案所带来的优势等。大数据应用最早是从数据营销开始的，通过各种数据的收集形成客户画像，为数据营销和运营提供了最佳手段。目前，大数据已经在金融、电信、房地产、政府等行业得到了广泛的应用。例如，政府利用大数据来反恐，房地产企业利用大数据来拿地，工业方面利用大数据来进行设备的诊断，农业方面利用大数据来选种、培育作物等，城市管理方面利用大数据来进行城市规划、人口统计，媒体娱乐方面利用大数据来预测票房、推荐歌曲，教育方面利用大数据进行个性化的教学，等等。随着大数据在各行业中的运用，“大数据”这个词可能会消失，因为大数据可能逐渐成为各行业的“血液”，尽管你没有感觉到它的存在，实际上它却时时刻刻伴你左右。每个案例都按照统一的框架进行规范，内容基本依据行业痛点、大数据应用方式方法（技术原理、产品和实现方法）、大数据应用成效、重大意义未来前瞻、企业介绍等部分展开。

因出版时间限定，部分企业错过了这次案例提交的最后时间，没关系，“赢在大数据系列丛书”项目将持续进行，案例收集到一定数量便会结集出版，案例提交邮箱是Dtiiimail@126.com。

同时，DT大数据产业创新研究院（DTiii）还在准备一个在线交互的大数据产业地图，对不断增加的大数据创新创业企业进行更新。大数据产业地图更新每年进行一次，并提供PPT版下载，欢迎访问：[www.DTiii.org](http://www.DTiii.org)。

尽管中国大数据呈现蓬勃生机，但大数据的发展和应用还存在一些问题。尽管大数据在各行业的深入应用逐渐形成了相应的产业链，但也面临着“数据意识不到位”、“信息化程度参差不齐”、“数据就绪化差”、“数据难以整合”、“缺乏行业标准和规范”、“数据安全意识薄弱”、“数据管理制度缺乏”、“政府数据开放程度低”等诸多瓶颈。目前，国家大数据发展战略已经形成，随着对大数据价值认识的不断提高、企业转型压力的加大、创新性大数据技术产品的不断推出、政府数据的逐步开放、制度的不断完善，大数据将迎来持续高速发展时期，万亿元大数据产业指日可待。

“赢在大数据系列丛书”采用“众包”、“众智”、“众筹”、“众享”进行编辑出版，从2015年9月《大数据领导干部读本》出版发行时就开始筹划，直接参与的企业有190多家，参加编写的人员多达350人，因企业、人员太多，在此不再一一列出，特别是提供案例，但因各种原因最后没有成为本书作者的朋友们，真诚感谢大家的努力和支持。

由于我们的资料不全、深入实际调研不够、研究能力有限，“赢在大数据系列丛书”的内容与观点难免存在不全面、不准确等问题，敬请业内专家与广大企业提出批评和建议，共同推进中国大数据产业的发展。

DT 大数据产业创新研究院 (DTiii)

2017年1月31日

## 致谢人员名单

感谢以下人员对本书提供的帮助（按姓氏笔画排列）：丁冬、丁军、于莹、于笑博、马可、马斋、马超、马景辉、王二虎、王乐、王伟、王波、王建民、王炼、王晨、王清波、王琳琳、王筱东、王鹏云、文建平、左磊、石鹏、石磊、龙明盛、龙凯、付永晖、付海鹏、冯一村、宁笑娟、成立立、朱铎先、乔云云、乔玉川、伍邵旭、任广杰、任永杰、华保健、刘东日、刘宇航、刘奇、刘岩、刘星、刘虹、刘彬、刘翔、齐红威、闫旭东、米维聪、汤灏、许冬琦、孙岚、孙盛峰、孙静博、苏江文、苏萌、李少娜、李丹丹、李伟、李纪洲、李青龙、李雪鹏、李献坤、李楠、李颖超、李馥岑、杨帆、杨伟伟、杨佼杨、杨柏、杨紫胜、杨斌、连冰玉、吴子明、吴轶伦、吴爱国、何香萍、余莉、汪德嘉、沈贝伦、沈立勤、宋仲伟、宋志远、宋神桥、张弓、张文、张伟佳、张宇、张军、张昆仑、张春猛、张继生、张琳琳、张敬亮、张鹏英、张韶峰、张贻、陆薇、陈传林、陈庆华、陈军、陈思恩、陈娟、陈露妃、武新、茅明睿、林永汉、林松涛、林莉、林强、罗海英、金雅昭、周友鸿、周莉、郑林钢、郑颖尔、孟磊、赵龙、赵伟、赵华、赵博、赵婷、胡许冰、胡芳槐、胡嘉琪、饶翔、姜春玲、姚娟、贺群、骆飞、莫倩、贾信明、顾竹、钱家俊、倪时龙、徐连明、徐超、高岩、高瑞鑫、郭红岩、郭洁、郭尉、席强辉、唐日新、唐岳岚、陶涛、黄玉麟、黄东旭、黄代恒、黄岩、黄骞、盛尊阔、崔晓波、崔晓霞、崔晶晶、梁玫娟、梁笃国、彭作文、葛利鹏、韩正清、韩昱、韩涵、韩斌、谢鹏、蓝云鹏、赖星星、雷涛、詹俊、鲍忠铁、蔡劲松、廖常如、廖雅哲、熊薇、樊庆伟、魏晨曦。

# 目录

<b>第 1 章 金融大数据</b> .....	<b>1</b>
1.1 TalkingData : SmartDP 推动金融行业转型升级.....	1
1.2 数聚股份 : 大数据助力银行精准营销.....	12
1.3 百融金服 : 大数据推动信贷业务创新.....	16
1.4 百分点 : 智能营销 + 风控, 让银行的数据会“说话”.....	27
1.5 安讯科技 : 品值“让手机号等于信用”.....	33
1.6 银联智惠 : 大数据提升金融企业风控能力.....	40
1.7 蚂蚁金服 : 最高法院联手芝麻信用利用大数据让“老赖”无处可躲.....	48
1.8 南大通用 : 大数据助力反洗钱.....	53
1.9 数据堂 : 大数据让征信焕发新活力.....	63
1.10 百融金服 : 传统银行与互联网金融如何实现共赢.....	68
1.11 海翼知 : 全国企业知识图谱助力企业征信.....	74
1.12 天云大数据 : 大数据有效对银行信用卡申卡欺诈预警.....	84
1.13 深度数据 : 大数据助力企业风控.....	93
1.14 集奥聚合 : 大数据助力金融企业经营用户.....	101
1.15 优特捷 (日志易) : 大数据助力金融日志分析.....	108
1.16 联智科技 : 大数据助力保险转型.....	118
1.17 中金数据 : 广发银行呼叫中心语音大数据分析云应用介绍.....	135
1.18 普林科技 : 基于大数据技术的号码风控平台.....	140
1.19 博彦科技 : 大数据重构金融业历史数据管理和服务的能能力.....	146
1.20 智慧星光 : 银行客户情报系统.....	156
1.21 普林科技 : 金融风控解决方案.....	163
1.22 博彦科技 : 大数据重构银行影像文档存储管理.....	167
1.23 中科天玑 : 新型混合数仓及一站式分析引擎.....	172
<b>第 2 章 电信大数据</b> .....	<b>182</b>
2.1 智慧足迹 : 联通智慧足迹大数据——政府城市规划管理的抓手.....	182

2.2 南大通用：大数据深度分析助力电信运营 .....	189
2.3 灯塔大数据：运营商数据助力智慧城市建设 .....	197
2.4 试金石：中国移动大数据风险控制应用实践 .....	204
2.5 普林科技：大数据助力运营商运营 降低客户流失率 .....	209
2.6 上海奕行信息：大数据让客流管理更智能 .....	216
<b>第3章 媒体大数据 .....</b>	<b>224</b>
3.1 星红桉：大数据重构传媒新生态 .....	224
3.2 艾漫数据：影视因数据重生 .....	233
3.3 海量集团：大数据驱动下的史上第一次昆曲众筹 .....	246
3.4 拓尔思：基于大数据的智能传播平台 .....	257
3.5 天脉聚源：大数据为中关村管委会注入互动科技魅力 .....	266
3.6 中科点击：军犬舆情，大数据舆情全流程服务 .....	271
3.7 中视德尔：电视大数据营销开启新媒体精准营销制胜之道 .....	279
<b>第4章 医疗大数据 .....</b>	<b>289</b>
4.1 慧辰资讯：大数据 360 深度挖掘，建立医疗领域商业决策模型 .....	289
<b>第5章 旅游大数据 .....</b>	<b>297</b>
5.1 TalkingData：大数据在航旅行业的应用 .....	297
5.2 融信数联：大数据让旅游更具智慧 .....	313
<b>第6章 数据市场大数据 .....</b>	<b>320</b>
6.1 广州数据交易服务中心：数据资产交易与流通实践在中国 .....	320
6.2 聚合数据：大数据催生 API 数据经济 .....	329
<b>附录 A 行业大数据应用最佳实践案例一览表 .....</b>	<b>336</b>

# 第 1 章

## 金融大数据

### 1.1 TalkingData: SmartDP推动金融行业转型升级

在全民创业、万众创新的时代，“大数据”“互联网+金融”成为未来最具潜力的行业，国家对“大数据”“互联网金融”以及背后驱动的巨大产业寄予了高度的重视与支持。从2012年至今，金融行业涌现了新的细分行业，如P2P网贷、互联网证券、直销银行，以及新兴的商业模式和创新的产品，传统金融行业面临巨大的转型压力。面对日益激烈的同质化竞争，对于传统银行、券商、保险行业的挑战的本质是对于客户的了解的挑战，也是数据的挑战。传统金融企业或者深化对于大数据的应用，或者被边缘化。企业需要更加主动地运营、获取客户数据，从而了解客户，获取优质的客户资源。

金融企业开展主动获客营销活动时，通过外部大数据的应用，收集和掌握更为广泛的客户数据，还可以引入外部的客户画像标签，以及客户在电子商务、社交媒体、多终端媒介等一切渗透到客户生活中的各类非结构化信息。基于这些信息，可以为客户构建崭新的360度画像，分析粒度将从原有客户群体分析精细化到每个客户的个体分析，及时获知客户行为，洞察客户情感与情绪，预测客户期待，并及时组织好匹配的产品与个性化服务，快速响应客户的需求。

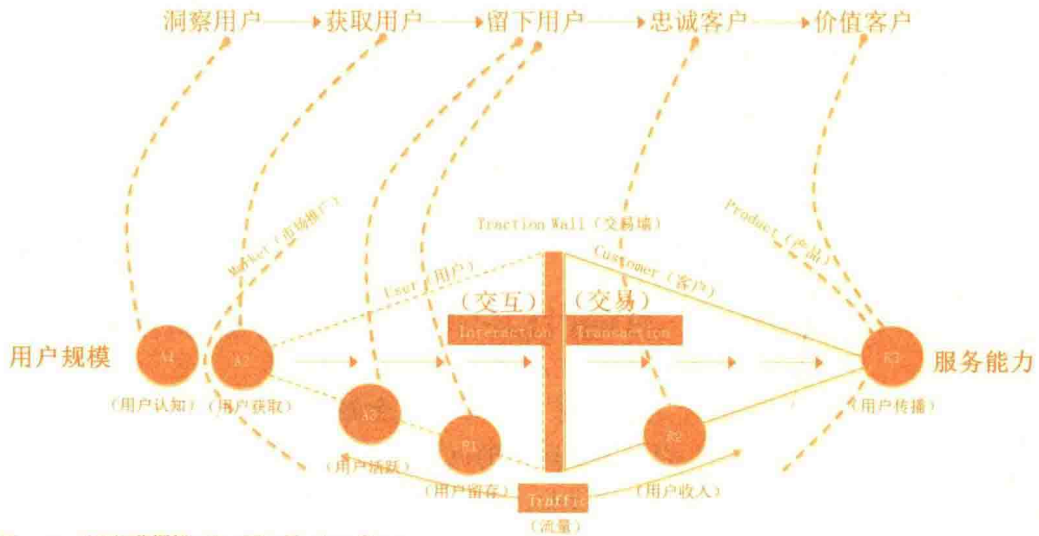
TalkingData作为国内领先的独立第三方移动数据服务平台，自2011年成立以来，始终专注于数据服务，用专业的数据服务产品线为客户提供全方位的大数据解决方案。从数据收集到数据挖掘，从数学模型到深度学习，从分析工具到数据平台，在数据应用领域积累了丰富的经验及海量数据，TalkingData逐步打造了以SmartDP平台为核心的数据生态体。随着国家对大数据产业的支持，2016年度T11峰会也受到了包括银行、证券、保险、互联网金融、地产、零售等行业的关注，中央电视台也对此次峰会进行了报道。

#### 1.1.1 金融行业的困惑

客户与传统金融企业进行数据交互，多数交易数据、日志、订单及客户交易行为（买卖、刷卡、查询、投诉等）通过企业内部的生产交易系统进行记录或“事后”进行收集结



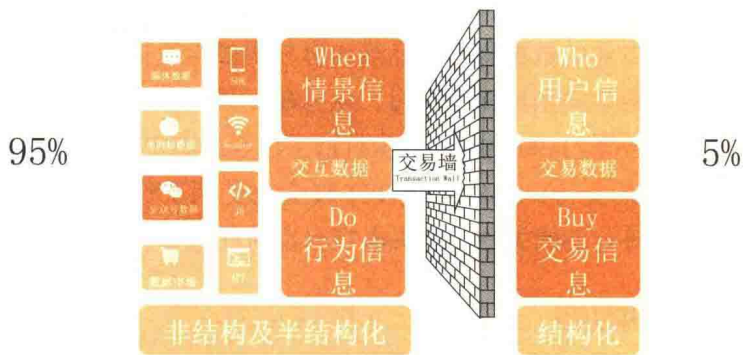
构化处理并存储在数据仓库，如图 1-1 所示，客户与企业的交易都发生在交易墙内，留下的忠诚客户通过购买产品和服务为企业创造了直接收益，价值客户后期通过向外传播为企业提供品牌和广告效益。



注：• TPU运营框架：Traffic/Product/User  
 • 核心指标：A1: Awareness; A2: Acquisition; A3: Activation; R1: Retain; R2: Revenue; R3: Refer.

图 1-1 客户与传统金融

在客户与企业进行直接交易之前，我们可以称为交易墙外数据。洞察到获取客户过程中的数据特点是以用户与企业的各种交互数据为主，如图 1-2 所示。数据本身代表了客户的行为，如位置、点击、浏览、企业 APP 内的操作行为、企业线下实体内的行为（购物中心内的到店足迹）等。此类数据开始出现大量非结构化，流式数据等多种形态。这部分非结构化数据维度丰富（人口属性、地理位置、消费偏好等），数据占到客户整体交互数据的 95%。传统企业如果需要收集并处理与客户交互的非结构化数据，需要花费极高的人力和资金成本。



丰富形态的数据无法装载到数据库

图 1-2 数据交互