



BIG DATA

大数据 分行业大解析

大数据创新创业实战案例

彭作文 刘宇航 © 著

倪光南院士等30位大数据专家强力推荐

深度解析大数据政用 | 商用 | 民用3大领域
覆盖 | 30个行业 | 50个应用案例

中国铁道出版社
CHINA RAILWAY PUBLISHING HOUSE



大数据 分行业大解析



彭作文 刘宇航◎著

中国铁道出版社
CHINA RAILWAY PUBLISHING HOUSE

内 容 简 介

本书从大数据应用的方向出发,从不同行业,不同维度解析大数据在各行
业中的落地应用。全书分为通路篇、现状篇、政用篇、商用篇、民用篇5大部分,
共8章内容,深入解读近50个应用案例,是对近年来大数据应用的一次重要盘
点;并且结合作者自身实践,总结提炼出大数据应用的通路,致力于为大数据
创业创新提供系统性、可操作性的参考。

重应用,干货多是本书的一大特色。介绍大数据概念的书籍已经有很多,
而大数据真正落地应用,才是行业的终极目标。本书正是从大数据自身实践与
行业应用出发,庖丁解牛式地解读大数据在各个行业、各个领域中的实践与价值,
对于大数据从业者或者有意进入大数据行业创业的人员来说,具有很强的借鉴
作用。

图书在版编目(CIP)数据

大数据分行业大解析 / 彭作文, 刘宇航著. —北京:

中国铁道出版社, 2016. 12

ISBN 978-7-113-22249-9

I. ①大… II. ①彭… ②刘… III. ①数据处理—产
业发展—研究—中国 IV. ①F279.244.4

中国版本图书馆CIP数据核字(2016)第272883号

书 名: 大数据分行业大解析

作 者: 彭作文 刘宇航 著

策划编辑: 苏 茜

责任编辑: 张 丹

责任印制: 赵星辰

读者热线电话: 010-63560056

封面设计: **MX** DESIGN
STUDIO

出版发行: 中国铁道出版社(北京市西城区右安门西街8号 邮政编码: 100054)

印 刷: 北京铭成印刷有限公司

版 次: 2016年12月第1版 2016年12月第1次印刷

开 本: 700mm×1000mm 1/16 印张: 15.5 字数: 174千

书 号: ISBN 978-7-113-22249-9

定 价: 49.00元

版权所有 侵权必究

凡购买铁道版图书,如有印制质量问题,请与本社读者服务部联系调换。电话:(010)51873174

打击盗版举报电话:(010)51873659

彭作文

中科点击创始人兼总裁，行业大数据应用专家，致力于大数据在各行业的管理及实践应用。军犬系列、慧数系列两大大数据品牌的缔造者。网络舆情商用定义者，曾负责北京奥运会、上海世博会舆情工作。中国传媒大学网络舆情专业客座教授，多次参与大数据行业峰会及主题演讲。

刘宇航

策划人，作家，自媒体人，中科点击市场品牌总监，品途商业评论、艾瑞、界面等专栏作家，发表了大量深度解读大数据行业的作品，对大数据行业应用有独到见解。多年市场与品牌策划工作，长于策略，精通创意，服务过数十个知名品牌整合营销。

拥有大数据是时代特性，解读大数据是时代任务，应用大数据是时代机遇。大数据作为一个时代、一项技术、一个挑战、一种文化，正在走进并深刻影响人们的生活。处在创新的时代，大数据必将引领未来生活新变化、孕育社会发展新思路、开辟国家治理新途径、重塑国际战略新格局。

近年来，大数据已成为人们谈论的热点话题，被誉为“21世纪的石油”，并已经在生产生活、社会管理和科研等诸多领域发挥作用，受到许多国家的普遍重视。正如2012年联合国发布《大数据时代与人类发展：挑战与机遇》白皮书中所指出的：“大数据时代是一次历史机遇，人们可以使用极为丰富的数据资源对社会经济进行前所未有的实时分析，帮助政府更好地完善社会和经济运行”。业内人士也普遍认为，大数据将对人类的社会治理、经济行为、生活方式和思维模式产生重大影响。IBM大数据专家谢国忠说：“在当今新一轮技术浪潮中，大数据始终是核心要素，像血液一样在云计算、物联网、移动和社交中流动”。他认为：数据是当今世界新的自然资源，并且取之不尽，用之不竭。未来将出现新的技术热点，如今形成舆论热点的物联网、人工智能、神经技术等，也必将以大数据技术和大数据思维为支撑。

目前，尽管国内外对大数据的定义、内涵和外延尚未形成统一认识，但普遍对大数据充满期待已不可否认。中科点击创始人、行业大数据应用专家彭作文先生，抢占先机，勇立潮头，多年来一直致力于大数据研究，从服务经济社会发展的实际出发，编撰形成了独具特色的《大数据分行业大解析》一书，以独到的眼界、多维度的思维方式，对大数据最主要的功能特征进行了论证，全面阐释了大数据具有的预测、比较和实时监控等社会管理方面的特征，已经得到许多业内人士的认可，形成了理论与实践的完美结合。

“得人者兴，失人者崩”。互联网领域是人才聚集地、资源富集区，“懂数据、会分析”的复合型人才缺乏是当下国内外面临的共同困难。彭作文先生“不唯学历、不唯论文、不唯资历，突出专业性、创新性、实用性”，广招天下“弄潮儿”，面对网络大世界、数据大时代、舆情大系统，推出创新巨著，传递新思想、新知识，新经验，堪称互联网领域的“怪才、奇才”。

“饥读之以当肉，寒读之以当裘，孤寂而读之以当友朋，幽忧而读之以当金石琴瑟也”。读其所著，体悟于心灵，修化于品行，践行于世事，服务于社会，可谓经世致用也！

甘肃省社会科学院研究员 熊俊贵

数据为伍的创业者

熟悉我的人都知道我是一个空中飞人，大学毕业后在报社工作，自然是飞来飞去，从报社下海后又选择了创业，早期创业非常辛苦，既当爹又当妈，还得当孙子，比在报社时飞得更多。

航旅纵横 APP 的大数据分析：我的 2015 年飞行里程达 207 291 公里，340 小时 45 分钟，飞行 156 次，飞行里程超过 99.5% 的用户，点评“登峰造极，您就是为飞行而生的”。其实我很感谢我的每次飞行，飞得越高，我越兴奋，越有灵感。飞行中，我要么速读一本书，要么完成一件重要的事。《大数据分行业大解析》这本书，对我来讲非常重要，我将这本书的自序也安排在 2016 年中秋回家的飞行途中完成，来回顾我的创业之旅，也回顾我与数据的一生情缘。

上大学时我就极其不安分，既要上学，还得打工养活自己和弟弟妹妹，大学四年基本都是早上六点起床，晚上宿舍关门后才回到寝室，周末则在校外打拼，干的事情很多：带群众演员演《还珠格格》，创办电脑学校自封“校长”，办份北京高校内部刊物——《塔内快

讯》，刊载一些培训、考研的广告，然后安排兄弟们一个宿舍一个宿舍的派送。学的是信息专业，大部分同学能做个网页，最厉害的是睡在我上铺的兄弟，居然会编程，在我看来他就是神。都快大四了，我尚未写出一个能满屏幕按指定轨迹飞的飞机（这可是我们学编程的第一个实例），平生第一次开始怀疑自己，一片茫然。刚从国外回国的宋铁英老师上我们的课，不是要求我们编程，而是要求我们对一个项目做 PPT，并且还要登台讲解，作为一个理科中的文科生，终于轮到我大显身手了，于是那门课成了我的最爱，宋老师也成为我的恩师，宋老师也改变了我对专业的看法：信息化专业不只是编程、算法，还可以有项目管理、产品策划，她特别强调“数据”与“信息”的价值，我清楚地记得，她说未来能把“数据”与“信息”的价值激活与运用的人，将有更大的前途。那天我回到寝室彻夜难眠，反复回味宋老师的话，天亮之前终于有了决定：编程是干不了了，但还是要信息化行业“混”（宋老师说这是未来最有前景和“钱”景的行业），那在信息化行业做什么呢？只有两条出路，要么是管编程的，要么是搞数据的，就这么定了。现在想起来，1999年的时候，电脑用的都还是 386，网速比蜗牛还慢，宋老师就对数据和信息有如此洞见，让我特别敬佩。后来宋老师当了我们学院的院长，只带博士和硕士，前几天教师节，十几位弟子去拜访恩师，我是唯一的一位本科生。

自那之后，我再也不去带人演群众演员了，不用让兄弟们在抗战片中趴在雪地上冷冷地装死，正式回归到信息专业中来，满脑子都是数据与信息。大四以后，从隔壁的清华、北大找了四五个人，用我搞电脑培训赚来的钱开始开发搜索引擎，模仿 Yahoo 的搜索，导航目录、爬虫、索引都做，记得兄弟们第一次把北大的 BBS 数据复制到自己的硬盘后，四五个兄弟在五道口的一家小饭馆，啤酒

配煮花生喝了个通宵，隆重地庆祝了一番，醒来后摇摇晃晃地回到校园。这一天，我把我在论坛上的网名（马甲）取名为“中国数据哥”。后来这套程序居然用上了，联想当时投资了一个门户网站叫FM365，经软件贩子若干倒手后，卖给了FM365，那时逢人便吹：“看没，那是我带人开发出来的。”这应该是国内较早的搜索引擎，尽管不成熟，但也算是我的处女作，最主要的是软件贩子（现在已经改名叫集成商了）给了我一张支票：20万元，我用来付了两套房子的首付……

毕业后我没有选择去软件公司，而是选择去报社搞数据库、信息化，算是换另外一个角度玩“数据”和“信息”，2003年曾放弃猎头用高薪挖我去百度搞搜索引擎，就这样与全球最大的中文搜索公司擦肩而过。

数据带给我财富的同时，也给了我极大的满足感！

2006年参与了北京奥运会的涉密项目，从2006年4月一直做到2008年10月，让我全面感受了数据的魔力，也系统性地完成了产品的架构，数据的采集、分析、串并、BI、可视化等，当各种数据最终呈现在大屏上是如此唯美，当各种数据最终被线下应用于各种实战时，简直令人叹为观止，我再次把自己这辈子的的工作锁定为数据为伍。

完成这个项目后，心生创业的念头，家里人全部反对，我以休假为名，用一月的工资4700元创业，如果创业成功就继续创业，如果一月把4700元花光了，就继续老实地回去上班。幸运的是，在一月内接到elong网廖总的项目，又是做数据，那时elong和ctrip竞争处于白热化阶段，我辅助elong采集ctrip、芒果网的酒店与机票价格信息，还采集各大航空网站的机票价格信息，为elong

做商业决策使用，据说 **elong** 根据这个分析结果，适当调整酒店价格，以比 **ctrip** 优惠 5 元的策略赢得了竞争。那时 **qunar** 还刚刚起步，同期还有陈华和吴世春创业搞了一个同类型的叫 **kuxun** 的机票比价平台。这一次收获了 26 万的收入，也是中科点击的第一份合同，当时完美时空的 **Tony** 拟投资 300 万元给中科点击，我觉得公司刚刚起步，尚不值钱，难以匹配，婉拒了 **Tony** 的投资，到今天这份只差我签字的投资协议还被我珍藏着，那代表 **Tony** 对我的认可，对数据产业的认可，这是一份信任和强心剂。又是数据的运用让我顺利启航创业。

2008 年，我们采集挖掘新闻、论坛、搜索、博客等信息，率先开发了“网络舆情”系统，创建了“军犬”这个品牌，在国内开拓出“网络舆情”行业，引得众多软件厂商模仿学习，“军犬”舆情算是数据在情报领域、媒体监测领域运用的典范。

“军犬”这个名字也是我在一次飞行中想到的：搞数据、搞情报，必须稳、准、狠、快，对数据必须嗅觉灵敏，才能对数据本身和价值预判感知，嗅觉灵敏最有代表的就是“犬”了；除此之外，搞数据、搞情报还得忠诚，能“精忠报国”，所以取名“军犬”。

2015 年，国务院印发《促进大数据发展行动纲要》，大数据一夜之间遍地开花，我们才醒过神来，一直以为是在搞信息化，原来是在搞“大数据”。又一次飞行途中，我和作涛商量，大数据一定得让数据“智慧”起来，于是就有了中科点击的另一个大数据品牌“慧数”，定位于做行业大数据应用，“慧数”——晓数据知行业。2016 年 5 月参加贵阳数博会，贵州省相关领导和参会嘉宾对于中科点击定位于做分行业大数据高度认可，中科点击正式全面拉开大数据在各行业、分领域、垂直化地应用。到本书截稿之日，中科点击以每投入 10 人、100 天、100 万元孵化一个分行业的速度打造了

10 个大数据应用平台，其中两个平台两月之内各获得 1 000 万元的天使投资。

成长的路上，一直与数据为伍，是数据成就了我，成就了中科点击，我心怀感恩于这个时代，这个属于大数据的时代。这一年来，我一直关注与研究大数据的应用，致力于大数据真正深入落地——政用、商用和民用，来促进社会治理、企业发展、服务百姓生活。于是借用本书，将大数据的分行业应用进行剖析与分解，权当抛砖引玉，与创业朋友探讨交流，基于数据的应用，碰撞出更多的火花。

2016 年 9 月 15 日 18 点 20 分，飞机将降落在深圳宝安机场，今天是中秋节，透过飞机舷窗，远远看到了明月，这样的夜空，应该用哪些数据来刻画呢？天宫二号也将于中秋之夜奔赴太空，中国迈向空间站时代，这些数据，这些历史又应该怎么被记录，怎么为后人应用？我又沉浸在自己的数据世界中。

一切皆数据，我是数据哥，数据为我魂！

彭作文

2016.9.15 日于 CA1357 航班

让大数据为我们服务

数据古已有之，只是没有任何一个时代如此“全球热恋”。

世上本没有“Big data”这个词，炒的人多了，便热起来。

从结绳记事到甲骨文，从竹简到帛书，从手抄到活字印刷，从现代印刷到计算机，从PC时代到移动互联网时代……数据从来都没有停止过，也从来都没有离开过我们的生活。它存在于一朝一夕，一举一动中，存在于世界的每一个角落。

正如有人的地方就有江湖一样，有人的地方就有数据。

当我们走进学校去读书求学，就产生了教育大数据；当我们走进餐馆去吃饭，就产生了餐饮大数据；当我们网购的时候，就产生了电商大数据；当我们去种地的时候，就产生了农业大数据；当我们出行的时候，就产生了交通大数据；当我们网上找对象的时候，就产生了婚恋大数据；当我们买房的时候，就产生了房屋大数据；当刮风下雨的时候，就产生了气象大数据……处处皆数据，数据无处不在。

我们吃过的每一顿饭，喝过的每一杯水，走过的每一个地方，说过的每一句话，见过的每一个人，读过的每一本书，打过的每一个电话，发过的每一条信息，爱过的每一个人，甚至每一场风花雪月的故事，都是大数据的组成部分，都构成了我们每一个人的“个人数据”。毫不夸张地说，每一个人的一生就是数据的一生。

也许有一天当我们死了，墓志铭只有一个简洁的二维码，扫一扫便是我们的一生。我猜想，当我的后人扫我的墓志铭，就能知晓我的一生：先父刘宇航，生于1986，一生放纵不羁爱自由，喝酒骑马走天下，生性狂放，好为文章……如果能借助正在兴起的VR技术，我的一生就像是一部纪录片一样呈现在我的后人面前，这不是一件很酷的事情？换句话说，一部《史记》算不算是一个牛气哄哄的大数据产品？

别迷恋概念，一切皆平常

从互联网+到+互联网，不管炒得多么火热的概念，都没有改变人类只有吃饭才能活下去的事实。在没有大数据这个概念的时候，人类其实早已经在应用大数据，只是“只缘身在此山中”罢了。每个人都在说大数据，这似乎是一个很时髦但又很有距离感的词。但是，你知道吗，你的每一个行为，都将运用到大数据，也都将组成大数据的一部分。

大数据并不神秘，就像吃喝拉撒一样，再平常不过

让我们暂且忘记那些虚无缥缈的概念，真正考虑大数据如何具体产生商业价值，如何被我们用在生活中，才是真正的王道。

一切新的东西终将会变成平常之物，一切平常之物也终将会变成历史，掉进尘埃里，被时间淡忘。我相信大数据也逃不脱这样的命运，只不过我们处在一个什么样的时代，就会做什么样的事，在

对的时间做对的事，一切都是再平常不过的事了。当我们想清楚了这些，我们会突然发现，大数据不再是高高在上的神。

就像我们的祖先刚开始时以捕猎为生，吃生肉，但是有一天突然发现食物还可以烤熟了吃，这是不是很新鲜的事情？但是现在我们觉得吃熟食是再平常不过的一件事情而已。

大数据不是我们吹牛时的谈资，也不是我们自我标榜的符号。做大数据应用也不是刻意去搜寻一大摞数据来堆砌。一切重在应用，如果我们认识不到“大数据”是为总结规律或证明结论服务的道理，而被不知所云的“大数据”概念唬住，那就会导致我们走入歧途，结果必定是搞出满篇花里胡哨的数据，却言之无物，令人看了昏昏欲睡。

因此，我们不会谈那些人云亦云，高高在上的概念，也不会迷恋于无法落地的理论。本着让数据干点实事的初衷，我们结合自身的实践与观察，来写这本书。让数据为我们所用，让数据为产生价值，走进普罗大众的心里，而不再是虚无缥缈，神乎其神的概念，也不再是忽悠人的说辞和道具，我相信这一天会很快就会出现。在这本书里，我们所讲到的都是我们亲身实践或者被市场验证过的事实。事实上，我们不能让每一个人成为大数据的专家，也许只会给你一点在大数据应用过程中的启发而已。

我认为真正将大数据落地做应用的人，更加值得尊重。就像中国的每一个古老的行业，真正让普罗大众受益的都是那些潜下心来手艺人。

让大数据走进生活，为我们所用

不应该对大数据膜拜，应该去征服它，利用它，为我们所用。当有一天我们觉得大数据并不神秘，并不高深的时候，就像是现在

我们对互联网的感觉一样，才算是大数据真正深入人心，真正为人人所用。

本书就是从我们所处的商业、生活等各个方面来解读大数据带给我们的改变，用事实说话，也当抛砖引玉，引出人们对大数据应用的更多思考。

刘宇航

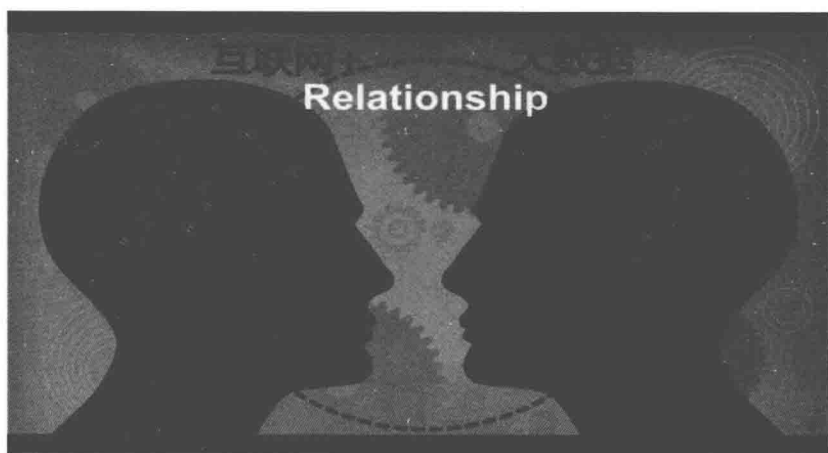
2016年9月

何为大数据

“忽如一夜春风来，千树万树梨花开”，大数据之花已然开遍了神州大地，开遍了全世界，成为一股不可阻挡的洪荒之力。

大数据是当今公众津津乐道的一个热词，人们纷纷在探讨大数据所带来的科技进步及所蕴含的巨大价值，甚至有人把大数据形容为未来世界的石油，未来的黄金。更有人宣称掌握了大数据的人，就可以像上帝一样来俯瞰整个世界。那么，到底什么是大数据？它将带给相关产业以及普通公众什么样的影响？

因为工作的关系，我经常参加一些关于大数据的论坛与活动，有一次在成都演讲后，有一位创业者跑来问我什么是大数据，尽管我几乎阅读了所有国内有关大数据的所有书籍，以及互联网、自媒体上有关大数据的论著和文章，但我差点被问住了，现场颇有些尴尬，我也只能用“海量多维集合”+5V来草草作答，我能感受到这位创业者一脸不知所云的状态，我顿时感觉到，大数据要普及、创业者要运用大数据创新创业，还需要将大数据的定义、原理、逻辑、应用深刻地理解后，深入浅出地转换为寻常百姓能听得懂，看得明白的内容。



不要说是普通的门外汉，即使对从事大数据这个行当的人来说，要用一句话很精练又恰到好处地来说明大数据是什么，也是勉为其难。也许有人说，大数据本就是一门高深的学问，一门尖端的科技，注定只有少数人能懂，不需要所有人都明白，这也许就目前来说有一定的道理。但是新鲜事物终会变成平常之物，正所谓太阳底下没有新鲜事，也许再过几年，大数据就如同现在的互联网、移动互联网一样，变得再平常不过。所以有必要，给大数据下一个人人都能听得懂的定义，让数据走进生活，让数据走进大众视野。

就目前对大数据的定义而言，更多的是从数据本身以及技术层面的解读。行业术语称为巨量数据集合。麦肯锡全球研究所给出的定义是：一种规模大到在获取、存储、管理、分析方面大大超出了传统数据库软件工具能力范围的数据集合，具有海量的数据规模、快速的数据流转、多样的数据类型和价值密度低四大特征。

不管是行业术语还是麦肯锡全球研究所给出的定义，都是从大数据本身出发，从数据的特征与处理解读，对于一般人来说是比较难以理解的，也不便于流传和记忆。大数据成为一种趋势，必将成