

Kogan Page

# 数据战略

[英] 伯纳德·马尔 著  
(Bernard Marr)  
鲍 栋 译

如何从大数据、数据分析和  
万物互联中获利

# DATA STRATEGY

How to Profit from a World of Big Data,  
Analytics and the Internet of Things

让数据改变商业，帮助企业做数字化转型的战略地图！

Google、Visa、Uber、Amazon、IBM等公司

11位高管及CEO | 联袂推荐

 机械工业出版社  
CHINA MACHINE PRESS

Kogan Page

# 数据战略

[英] 伯纳德·马尔 著  
(Bernard Marr)

鲍 栋 译

如何从大数据、数据分析和  
万物互联中获利

## DATA STRATEGY

How to Profit from a World of Big Data,  
Analytics and the Internet of Things



数据正在彻底改变所有人做生意的方式，视数据为战略资产的公司才会生存并茁壮成长。本书是创建一个强大的数据战略的必读指南。

Copyright © Bernard Marr, 2017

This translation of Data Strategy is published by arrangement with Kogan Page.

This edition first published by Kogan Page, Great Britain and the United States in 2017. This translation is published by arranged with Kogan Page and Andrew Nurnberg Associates International Limited.

This title is published in China by China Machine Press with license from Kogan Page. This edition is authorized for sale in China only, excluding Hong Kong SAR, Macao SAR and Taiwan. Unauthorized export of this edition is a violation of the Copyright Act. Violation of this Law is subject to Civil and Criminal Penalties.

本书由 Kogan Page 授权机械工业出版社在中国境内（不包括香港、澳门特别行政区及台湾地区）出版与发行。未经许可之出口，视为违反著作权法，将受法律之制裁。

北京市版权局著作权合同登记 图字：01-2017-6676 号。

## 图书在版编目（CIP）数据

数据战略：如何从大数据、数据分析和万物互联中获利 /（英）伯纳德·马尔（Bernard Marr）著；鲍栋译．—北京：机械工业出版社，2018.10

书名原文：Data Strategy: How to Profit from a World of Big Data, Analytics and the Internet of Things

ISBN 978-7-111-61096-0

I. ①数… II. ①伯… ②鲍… III. ①数据处理—研究 IV. ①TP274

中国版本图书馆 CIP 数据核字（2018）第 230926 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：刘洁 责任编辑：刘洁

责任校对：李伟 责任印制：郜敏

北京圣夫亚美印刷有限公司印刷

2019 年 1 月第 1 版第 1 次印刷

145mm × 210mm · 7 印张 · 3 插页 · 174 千字

标准书号：ISBN 978-7-111-61096-0

定价：59.00 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

金书网：www.golden-book.com

封面无防伪标均为盗版

教育服务网：www.cmpedu.com

## 赞 誉

“计算机算法将控制我们生活的方方面面。本书将成为我们通过充分利用数据引导认识市场、参与竞争并在竞争中取胜的指南。”

——亨利克·范·希尔 (Henrik von Scheel)，Google 顾问委员会委员

“数据战略不仅适用于数据专业人士。伯纳德·马尔告诉我们，数据战略需要得到公司营销、顾客关系、产品和人才战略等部门的同等重视。按照我的经验，数据战略是所有部门取得成功的关键。我相信，伯纳德·马尔的这本书将为我们制定数据战略提供一个宝贵的起点。我认为，本书融合了我掌握的很多经验和教训。假如你还没有自己的数据战略，本书绝对值得一读。”

——大卫·珀迪 (David Purdy)，Uber 数据科学家

“如果你是一家公司的数据战略负责人，但依旧苦于无法明确这项任务的规模以及建立数据驱动文化会带来怎样的回报，那么，这本书将有助于你全面把握数据在当下商业世界中的方方面面，并为你提供一种应对这项任务的指导性架构。”

——杰克·威廉姆斯 (Jake Williams)，Amazon 零售战略负责人

“本书通过形形色色的案例——小到一家连锁店，大到全球家居知名品牌，让一个原本复杂深奥的话题变得通俗易懂。在最新著作《数据战略：如何从大数据、数据分析和万物互联中获利》中，伯纳德·马尔以深入浅出、娓娓道来的对话风格，为许多试图把握数据驱动型经济的人解开了它背后的诸多基本要素。因此，对那些想深入了解人类如何在数据和设备的辅佐下不断发展以及如何让自己的组织把握这种脉络的读者

来说，本书必将引起他们极大的兴趣。”

——加利斯·米切尔-琼斯 (Gareth Mitchell-Jones)，IBM 认知系统负责人

“在当下竞争激烈的商业环境中，本书为我们绘制了一张最大化数据价值的清晰、简明而且令人振奋的路线图。”

——拉尔夫·布洛尔 (Ralph Blore)，Visa 中央分析业务负责人

“当我们步入第四次工业革命时，伯纳德·马尔告诉我们，只有那些将数据视为战略性资产的企业才有可能生存和发展。他让我们认识到，数据是如何改进决策、改善运营并通过产品或服务实现货币化的。他的最新著作既是让我们了解大数据各个方面的参考手册，也是帮助我们利用大数据为组织创造价值的指南。不管我们如何看待不断增长的数据和自动化运用，马尔的专业化解读都将为我们用数据塑造未来助一臂之力。”

——安迪·鲁宾 (Andy Rubin)，Pentland Brands 董事长

“伯纳德·马尔化繁为简的能力令人不可思议，他总能让普通人读懂最复杂的话题。伯纳德在数据战略方面的见解更是让我们受益匪浅，毕竟，他真正理解如何创建一个全面的数据战略，更重要的是，他还知道该如何让我们接受这个过程。因此，我向那些想要拥有大数据领域实用指南的读者强烈推荐本书。”

——斯图尔特·B. 弗兰克尔 (Stuart B Frankel)，Narrative Science 首席执行官

“今天，利用数据驱动竞争优势已从一种选项变成一种需求，而伯纳德·马尔为组织领导者提供了一张全面的路线图，引导他们专注于绘制和评价自己的旅程。他在书中谈到了从数字企业到小企业的诸多案例，这些案例引人入胜，娓娓道来，无不凸显出数据战略之于企业的重要性。这不仅有助于读者提升企业价值，而且会帮助他们在技术、支持和数据

选项纷繁复杂的世界中乘风破浪，不断前行。”

——安德鲁·萨爾斯基 (Andrew Salesky)，  
嘉信理财 (Charles Schwab) 高级副总裁兼全球数据官

“伯纳德·马尔全面阐述了塑造数据战略的各项关键要素，并以生动鲜活、极富说服力的案例强调了数据驱动型企业在当今数字经济中的重要性。”

——布伦特·德克斯 (Brent Dykes)，大数据创业公司 Domo 数据战略总监

“在本书中，伯纳德·马尔谈及了影响所有企业的一种巨大的变革力量。拥有一套强有力的数据战略和一张数据科学路线图，已成为深入组织基因的核心构成要素。而本书则提供了路线图及大数据领域领导者对如何实现这个目标的诸多见解。”

——威廉·莫肯 (William Merchan)，DataScience 首席战略官

“在一个需要我们衡量一切、分析一切的时代里，伯纳德·马尔为我们完美诠释了成功企业必须回答的问题。”

——丹·莫里斯 (Dan Morris)，维亚康姆 (Viacom) 数据平台高级总监

谨以此书献给我生命中的四位至亲：我的妻子克莱尔，我亲爱的三个孩子索非亚、詹姆斯和奥利弗。

## 致 谢

能投身于这个如此富于创造力和快速发展的领域，我感到非常幸运。更让我感到荣幸的是，自己有机会与各个行业板块的企业和政府机构开展合作，协助它们以更有效的新方法去使用数据创造价值。这份工作让我从未停止过学习，否则，也不可能有这样一本书的问世。

在此，我想感谢很多帮助我走到今天的人。所有与我合作过的公司里的优秀人士，每当我为他们提供帮助的时候，他们都会给予我信任并回馈给我如此多的新的知识和经验。此外，我还要感谢所有与我分享思想的人，不管这种分享来自我们面对面的沟通，还是来自博客、书籍或者任何其他形式。感谢你们的慷慨，让我每天都能汲取新的养分！我还有幸认识了这个领域的很多关键的思想家和思想领袖，我真希望你们知道我多么感激你们给予的观点和意见！

最后，我还要感谢编辑和出版团队的帮助和支持。任何一本书从想法变成出版物，都是一个充满挑战的过程，真诚感谢你们的投入和帮助，感谢查理、大卫、艾美和露西。



## 作者简介

伯纳德·马尔 (Bernard Marr)，国际知名的商业畅销书作家，多家公司和多个政府机构的主题发言人兼战略顾问。他是商业数据领域的全球权威人士，被 LinkedIn (领英) 公认为全球前五大最具商业影响力人士之一。

伯纳德经常为世界经济论坛 (World Economic Forum) 撰稿，也曾为《福布斯》杂志和 LinkedIn Pulse 定期撰写专栏文章，他的专家评论经常出现于 BBC 新闻、天空新闻和 BBC 世界等电视媒体和广播，以及《泰晤士报》《金融时报》《CFO 期刊》《华尔街日报》等知名刊物中。

伯纳德·马尔撰写了大量开创性的书籍和数百篇引发轰动的报告及文章，其中包括国际畅销书 *Big Data in Practice: How 45 successful companies used big data analytics to deliver extraordinary results* (《大数据在实践中：45 家成功的公司如何使用大数据分析来提供非凡的结果》)、*Big Data: Using SMART big data, analytics and metrics to make better decisions and improve performance* (《智能大数据 SMART 准则：数据分析方法、案例和行动纲领》)、*Key Business Analytics: The 60+ business analysis tools every manager needs to know* (《关键业务分析：所有管理者都需要了解的 60 种业务分析工具》)、*The Intelligent Company* (《智能公司》) 以及 Dummies (傻瓜学) 系列丛书中的 *Big Data for Small Business* (《大数据专家：小企业也能用好大数据》)。

伯纳德·马尔曾与许多世界知名机构合作并提供咨询服务，其客户包括埃森哲咨询、阿斯利康制药、英格兰银行、巴克莱银行、BP (英

国石油)、思科、DHL(敦豪快递)、Fujitsu(富士通)、Gartner(高德纳咨询)、HSBC(汇丰银行)、IBM、Mars(玛氏)、英国国防部、微软、北大西洋公约组织、Oracle(甲骨文)、英国内政部、NHS(英国国家医疗服务署)、法国 Orange 电信、Tetley(泰特利茶业)、T-Mobile、Toyota(丰田汽车)、英国皇家空军、SAP、Shell(壳牌石油)、联合国以及沃尔玛等。

如果您仍需就咨询服务、演说邀请或培训需求与伯纳德·马尔交流,可通过如下网址联系: [www.ap-institute.com](http://www.ap-institute.com) 或电子邮件 [bernard.marr@ap-institute.com](mailto:bernard.marr@ap-institute.com)。

此外,您也可在 Twitter 上关注 @bernardmarr,他经常在上面分享他的观点,或是在他定期发表博客的 LinkedIn 上取得联系。

# 目 录

赞誉  
致谢  
作者简介

## 第 1 章 为何说当下业务无不是数据业务 // 1

- 1.1 大数据和物联网的惊人增长 // 1
  - 1.2 数据驱动的勇气新世界 // 2
  - 1.3 我们是否正在逼近人工智能 // 7
  - 1.4 数据正在如何彻底改变我们的商业世界 // 10
    - 1.4.1 数据在商业中的基本作用 // 10
    - 1.4.2 智能工厂与工业 4.0 // 14
    - 1.4.3 自动化及其对就业的现实威胁 // 16
    - 1.4.4 区块链技术：是否是数据和企业的未来 // 18
  - 1.5 所有业务都必须成为数据业务 // 19
    - 1.5.1 一切以数据战略为起点 // 20
    - 1.5.2 你的公司是否需要首席数据官 // 21
- 注解 // 23

## 第 2 章 战略性数据需求的确定 // 25

- 2.1 以数据提高企业的决策质量 // 26
  - 2.1.1 利用数据更好地了解客户和市场 // 27
  - 2.1.2 在一个意想不到的场景，让数据为你而动 // 29
- 2.2 利用数据改善运营 // 30
  - 2.2.1 通过数据获得内部效率 // 31
  - 2.2.2 亚马逊：如何以数据优化业务流程并增加销售额 // 33
  - 2.2.3 优步：如何以数据优化运输 // 34
  - 2.2.4 罗尔斯·罗伊斯：如何以数据驱动制造业成功 // 35
- 2.3 商业模式的转型：将数据作为企业资产 // 37
  - 2.3.1 如何以数据提升企业价值 // 37
  - 2.3.2 将数据转化为新的收入源泉 // 38

2.4 只有正确的数据才是有意义的,并非所有数据都是有价值的 // 40

2.5 为数据提供强有力的商业案例 // 42

注解 // 43

### 第3章 使用数据改善商业决策 // 44

3.1 明确你的关键性业务问题 // 44

3.1.1 好问题带来更好的答案 // 46

3.1.2 针对顾客、市场和竞争者的问题 // 47

3.1.3 有关财务的问题 // 52

3.1.4 有关内部运营的问题 // 55

3.1.5 有关人员的问题 // 57

3.2 数据的可视化及沟通洞见 // 60

3.2.1 是否每个人都应有权访问数据 // 61

3.2.2 告别电子表格,迎接数据可视化时代 // 63

3.2.3 以视觉与文字的融合发挥最大效果 // 64

3.2.4 虚拟现实和数据可视化的未来 // 65

注解 // 66

### 第4章 使用数据改善企业运营 // 67

4.1 利用数据优化运营流程 // 68

4.1.1 数据如何改善制造过程 // 68

4.1.2 如何以数据强化仓储和配送 // 70

4.1.3 如何以数据增强业务流程 // 71

4.1.4 如何以数据强化销售及营销流程 // 75

4.2 以数据改善顾客供应 // 78

4.2.1 为客户提供更优服务 // 78

4.2.2 提供更好的产品 // 81

注解 // 84

### 第5章 数据的货币化 // 85

5.1 增加企业价值 // 86

5.2 数据本身成为企业核心资产 // 87

5.3 由企业数据处理能力创造的价值 // 90

5.4 向顾客或利益相关者出售数据 // 91

5.5 理解用户生成数据的价值 // 96

## 第 6 章 数据的取得与收集 // 98

6.1 了解不同类型的数据 // 99

6.1.1 对“大数据”的定义 // 100

6.1.2 结构化数据的定义 // 101

6.1.3 非结构化数据和半结构化数据的定义 // 103

6.1.4 内部数据的定义 // 105

6.1.5 外部数据的定义 // 106

6.2 更多的新型数据 // 108

6.2.1 活动数据 // 108

6.2.2 对话数据 // 109

6.2.3 照片和视频数据 // 110

6.2.4 传感器数据 // 111

6.3 内部数据的收集 // 111

6.4 外部数据的访问 // 113

6.5 如果你需要的数据尚不存在 // 116

注解 // 117

## 第 7 章 将数据转化为洞见 // 118

7.1 分析技术的进化方式 // 119

7.2 了解不同类型的分析技术 // 120

7.2.1 文本分析 // 121

7.2.2 情感分析 // 122

7.2.3 图像分析 // 122

7.2.4 视频分析 // 123

7.2.5 语音分析 // 124

7.2.6 数据挖掘 // 124

7.2.7 业务实验 // 125

7.2.8 视觉分析 // 126

7.2.9 相关性分析 // 126

7.2.10 回归分析 // 127

- 7.2.11 情景分析 // 128
- 7.2.12 预测 / 时间序列分析 // 129
- 7.2.13 蒙特卡罗模拟法 // 129
- 7.2.14 线性规划 // 130
- 7.2.15 同期群分析 // 130
- 7.2.16 因子分析 // 131
- 7.2.17 神经网络分析 // 132
- 7.2.18 元分析 / 文献分析 // 133
- 7.3 高级分析：机器学习、深度学习和认知计算 // 133
- 7.4 以不同分析技术的结合追求成功最大化 // 137
- 第 8 章 技术和数据基础架构的创建 // 139**
  - 8.1 “大数据即服务”：能成为企业的一站式解决方案吗 // 140
  - 8.2 收集数据 // 143
  - 8.3 存储数据 // 146
    - 8.3.1 了解云基础 / 分布式存储系统 // 146
    - 8.3.2 Hadoop 概述 // 148
    - 8.3.3 Spark: Hadoop 的替代品 // 149
    - 8.3.4 数据湖和数据仓库的简单介绍 // 150
  - 8.4 数据的分析和处理 // 151
  - 8.5 提供数据访问服务 // 155
    - 8.5.1 倡导数据管家的概念 // 156
    - 8.5.2 数据的沟通 // 157
  - 注解 // 159
- 第 9 章 打造组织的数据能力 // 160**
  - 9.1 大数据技能短缺及其对企业的影响 // 160
  - 9.2 建立内部技能和竞争力 // 163
    - 9.2.1 五种基本的数据科学技能 // 163
    - 9.2.2 招募新人才 // 166
    - 9.2.3 为现有人员提供培训并提高其工作技能 // 168

9.3 将数据分析业务外包 // 170

9.3.1 与数据服务供应商合作 // 170

9.3.2 Kaggle：众包数据科学家 // 172

注解 // 175

**第 10 章 不要让数据成为负债：数据治理 // 177**

10.1 数据所有权和隐私方面的考虑 // 178

10.1.1 拥有还是外购 // 178

10.1.2 确保拥有合理权限 // 179

10.1.3 将数据最少化作为好的实践 // 180

10.1.4 理解隐私问题 // 182

10.2 数据的安全问题 // 187

10.2.1 数据泄露的重大影响 // 187

10.2.2 物联网的威胁 // 190

10.3 践行良好的数据治理 // 192

注解 // 194

**第 11 章 数据战略的执行和完善 // 196**

11.1 把数据战略付诸实践 // 196

11.1.1 态度是关键 // 196

11.1.2 数据战略为什么会失败 // 198

11.2 创建数据文化 // 201

11.3 重新审视数据战略 // 203

11.3.1 调整企业需求 // 204

11.3.2 持续演进的技术图景 // 204

11.3.3 遥望未来 // 207

注解 // 210

为配合读者阅读本书，作者还提供了一份在线参考资料——《超越大数据》(*Beyond the Big Data Buzz*)，要查询该电子书请访问网址：

[www.koganpage.com/beyond-the-big-data-buzz](http://www.koganpage.com/beyond-the-big-data-buzz)。

# 第 1 章 为何说当下业务无不是数据业务

数据正在以前所未有的速度改变着这个世界以及我们的生活和工作方式。不同的角度看法不同，我们或许正踏上一条无比激动的精彩旅程，抑或正在进入一个令人毛骨悚然的“大哥”时代——我们的一举一动在这个“大哥”眼中暴露无遗，甚至可以被它未卜先知，随你怎么想象，但两者确实皆有可能。然而，企业领导者和管理者几乎没有时间去怀疑数据的权威性。数据已开始彻底改变公司的运营模式，而且注定会在未来几年对组织来说越来越重要。只有那些将数据视为战略资产的公司，才有可能得以生存和兴旺。大数据和物联网的巨大增长，再加上数据分析方法的快速发展，必将不断提升数据在企业各个层面的重要性。

## 1.1 大数据和物联网的惊人增长

今天，我们只需两天时间即可创造出人类在 2003 年之前积累起来的全部数据。难以想象吧，只需要两天的时间！而且我们创造数据量的速度还在持续提高。到 2020 年，我们可以使用的数字信息量将从今天的大约 5ZB<sup>⊖</sup> 增长到 50ZB。我们的每一个动作都会留下相应的数字轨迹——无论是在线浏览、使用信用卡在实体店购物、发送电子邮件、拍照还是阅读一篇在线文章，甚至是在大街上行走，如果正好你带着手机或是你附近安装有闭路电视摄像头。

“大数据”一词是指这些数据的总体集合以及我们在商业等各个领

---

⊖ zettabyte, 泽字节, 1ZB=1 万亿 GB——译者注



域利用这些数据的能力。数据本身并不是什么新事物。实际上，早在计算机和数据库出现之前，我们就已开始使用数据来跟踪行动和简化流程——想想纸质的交易记录和档案文件。计算机，尤其是电子表格和数据库为我们提供了一种简单易行的存储和组织大规模数据的方式。忽然之间，我们只需单击一下鼠标，便可取得信息。

但直到最近，数据还仅限于电子表格或数据库，而这些数据必须是高度有序、整齐排列的。任何不易组织成行和列的数据都会因为实在难以处理而被忽略。而现在，存储和分析技术的发展都意味着，我们可以捕获、存储和处理非常多不同的数据类型。因此，今天的数据可以包括从电子表格到图片、视频、录音、书面文字和传感器数据等各种类型。

毋庸置疑，仅仅是我们所创造的数据量就已经无比巨大了。但说实话，我觉得自己从来就没有对“大数据”这个词满意过。这个词对数据的认识只停留在非常肤浅的层次上，因为关注的只是数据量，而不是数据所带来的无限机缘。我真希望能用一个更贴切的词汇来描述在我们技术、文化和世界中的这一巨大转型。也正因为如此，在这本书中，我才不遗余力地从各个角度来谈论“数据”这个概念，无论大小——因为你拥有多少数据都无关紧要，重要的是你能否将这些数据成功地为己所用。

## 1.2 数据驱动的勇气新世界

在大数据面前，你毫无隐私可言，因为它对你是无所不知的。谷歌知道你在网上搜索什么，脸书（Facebook）清楚你的朋友是谁，互联网服务供应商（ISP）会跟踪你在互联网上访问过的每一个网站，即使你采用隐私浏览模式也不能逃脱它们的视野，但“大数据”的能力还远不止于此。谷歌不仅知道你的年龄和性别（尽管你从未告诉过他们），而且你完全可以确信，它们拥有你最全面的个人画像，对你的兴趣一清二楚，因而可以确定该向你推送什么广告。脸书很清楚你的朋友是谁以及