

# 大数据产业革命：重构DT时代的企业数据解决方案

作者：（美）托马斯（Thomas, R.）

## 版权信息

COPYRIGHT

书名：大数据产业革命：重构DT时代的企业数据解决方案

作者：【美】罗伯特·托马斯；帕特里克·马博兰

出版社：中国人民大学出版社

出版时间：2015年5月

ISBN：9787300211909

本书由中国人民大学出版社授权得到APP电子版制作与发行

版权所有·侵权必究

# 本书赞誉

击破存在于数据管理员与企业管理者之间的壁垒，这是大数据时代面临的巨大挑战。《大数据产业革命》就是一枚书香炸弹，它告诉创业家、成熟企业的管理层，以及对大数据充满好奇的人们，只有打破壁垒，迈出第一步，才能领略大数据带来的巨大价值。

江青

中国统计信息服务中心大数据研究实验室主任

作者在本书中用讲述故事的笔触勾勒出未来的世界，在这个世界里，我们所熟悉的不少生产关系将被改写。本书适合引领多种类型的读者走近大数据。

韩亦舜

清华大学数据科学研究院副院长

通过《大数据产业革命》一书，你将看到各行各业是如何运用数据将不可能变为可能的。企业要用数据增强自身的实力，并坚信能够发现那些从未被开拓过的领域，将商业与工业推升到一个全新的高度。

李易

中国知名互联网专家、中国移动互联网产业联盟秘书长

《大数据产业革命》的作者罗伯特·托马斯是IBM大数据业务的掌门人，他对于数据有着精准的认识和深刻的洞悉。在书中，通过对九大行业中大数据实战运作的深度解析，他提出了30个数据要素和54项数据模型，可以助力企业从IT时代到DT时代成功进行升级转型。

薛载斌

赛迪研究院世界工业研究所所长

在大数据时代，快速准确的商业决策将使企业在激烈的竞争中脱颖而出。本书通过详实的案例，向我们展现了如何借他山之石，迎接属于你的大数据时代。

崔昊

WatchStor.com总编

本书将完全颠覆以往处理合作伙伴和顾客关系的游戏规则，数据已无比清晰地为企业勾勒出未来的商业发展方向，关键在于你是否能够迎头赶上。

刘宇

IBM大中华区系统与科技部主机及存储系统资深市场与策略经理

本书向我们展现了企业大数据的实战运作。这项技术的崛起正在快速改变我们的商业环境。如何在数据中迅速挖掘出极具价值的市场信息回馈，从而不断提升产品与服务的竞争优势，是每一位职业经理人都应当考虑的问题。

张瀚文

IBM大中华区系统与科技部存储系统市场经理，大华东华中区市场经理

本书不仅有跨各行业的生动翔实的故事，也有作者提炼的54项大数据模型。作者并未止步于故事和模

型，而是引领大数据实施和决策者进一步利用模型和各种精炼的体系架构为实施企业大数据创新提供方法论和指导，是企业决策者和相关管理人员认识、掌握进而实施大数据的指导书籍。

陈滢

慧科教育高级副总裁、慧科教育研究院院长

# 推荐序一

韩亦舜

清华大学数据科学研究院副院长

当“大数据”一词出现在李克强总理的政府工作报告中时，这个在国内已经被炒得沸沸扬扬的概念又进一步升温。有多少人意识到了大数据将会引领一场革命？

提到革命，我先想到的是一句耳熟能详的话：“革命就是解放生产力。”当旧有的生产关系或商业模式已经开始束缚生产力的发展时，革命就将成为必然。技术的进步导致了大数据时代的到来，使我们得以从更宽的视野、更多的维度去更仔细地观察事物的数据属性，从而发现里面所蕴涵着的新的生产力发展契机。

提到革命，我还想到了另一句颇为刺耳的话：“革命是要死人的。”当然，这里说的不是生物意义上的死人，而是指那些抱残守缺、不思进取的企事业单位必将受到大数据浪潮的冲击。淘宝使渠道扁平化，撼动了传统的批发零售系统；专车系统对传统出租行业由于垄断而获取的不合理利润进行了颠覆；过去凭经验拍脑袋的粗犷式经营必然被精细化经营所取代。所以，把这个“死人”解释为死一批商业模式和运行机制，也丝毫不为过。

《大数据产业革命》一书从九个行业的应用场景出发，介绍了不同行业的大数据实践。从农业到服务业，九个行业各有特点，但这九个应用场景的汇集焦点就是大数据的运用，给这九个看似传统的产业带来新希望的也正是大数据的运用。这对不少从事传统产业又对大数据应用感兴趣的读者，有很大的启示作用。从这些栩栩如生故事中，我们可以看到大数据思维与传统思想理念和传统运营模式之间的冲突，这和中国的情形完全一样。在大数据时代，我们确实面临着很多技术上的挑战，但最大的挑战还是来自观念的冲突。大数据时代的到来势不可挡，冲突的结果必然是传统的思考方法和传统运营模式被荡涤。

在介绍了应用场景之后，作者又对支撑各应用场景背后的数据模型和相关的要素进行了梳理。这30个数据要素和54个数据模型是不是涵盖了全部，我们姑且不论，由于受篇幅限制，书中对每个要素与模型的介绍基本上也是点到为止，但这种梳理对初学者来说，无疑是有所裨益的。作者并不想在探讨数据模型与数据要素上“恋战”，而是笔锋一转，着力用约三分之一的篇幅“动员”读者成为大数据产业革命的领导者。要想成为大数据浪潮中的弄潮儿，仅仅有大数据意识和对大数据时代商机的敏锐洞察力是不够的，作者指出了制定大数据战略有哪些关键步骤，如何认识数据资产，如何实施业务流程再造，如何搭建大数据的体系架构等问题，表现出两位作者的产业背景给其带来的厚实功底。

作者用讲述故事的笔触勾勒出了未来的世界，在这个世界里，我们现在所熟悉的不少生产关系将被改写。本书适合引领多个类型的读者走近大数据。

16世纪末到17世纪初，荷兰人发明了望远镜与显微镜，一改人类只能用肉眼观察世界的状况，引发了生物科学和天文学的发展，使人类对世界的认知能力得到了极大的延伸，巫术和地心说也因此被颠覆。而大数据时代的来临，就像望远镜与显微镜的出现，让我们对这个自然世界和我们人类自身的认知更多、更细、更精准。大数据正在改变着我们的工作和生活。

## 推荐序二

李易

中国知名互联网专家、中国移动互联网产业联盟秘书长

坦白地讲，在拿到《大数据产业革命》的样书之前，我对这本书颇不以为然。毕竟罗伯特·托马斯和帕特里克·马博兰这两个名字在中国都不是那么有名，甚至有些陌生。然而事实证明，我错了。在从北京前往上海的高铁上，我几乎是一口气读完了这本书的样书。这似乎是不可能想象的事情，但事实的确如此。

从用户体验的角度来看，这本书的绝妙之处就在于，罗伯特和帕特里克不是在给你讲大数据，而是在给你讲故事。当然，这两位作者完全有能力给你讲上几百个故事，但他们只是从中精挑细选了能够代表行业如何运用大数据的最具吸引力、最具典型性的故事讲给你听，其中涉及农业、医疗、保险、零售、客服、智能机器、民众与政府、天气与能源等九大行业。

更有趣的是，书中的这些故事在阅读时完全无需按主次顺序来说，我个人甚至建议你不妨先挑选一个你认为最有趣的故事开始阅读，然后再阅读剩余的部分。当然，对于我而言，作为一个保时捷的忠实用户，我就特别喜欢“从未停止创新的保时捷”这个故事。

从书中，我们不难看出，大数据的重要性不言而喻。

2015年3月5日，中国国务院总理李克强在年度政府工作报告中正式提出：制定“互联网+”行动计划，推动移动互联网、云计算、大数据、物联网等与现代制造业结合，促进电子商务、工业互联网和互联网金融健康发展，引导互联网企业拓展国际市场。在我看来，李克强总理的讲话透露出简单明了的逻辑：未来的中国属于“互联网+”，而未来的“互联网+”的计划实施，归根结底还依赖于大数据。没有大数据，就没有未来。因此，对于任何没有大数据背景知识的政府公务人员、实体经济管理决策层及管理层来说，《大数据产业革命》真的是一本不可多得的、为企业、政府提供切实可行的大数据解决方案的书。

没错，“大数据产业革命”的到来，意味着人类将使用数据来完成那些令人难以置信的壮举。而引领“大数据产业革命”浪潮的人物必须具备三项特质：

- 1.否定对任何可行性的质疑，为可能创造出新的定义；
- 2.具备内在化的模型识别的知识与洞察力，可以将某个行业或维度的模型运用到其他看似毫无关联的行业上去；
- 3.致力于改善与数据相关的每一个方面，哪怕只改善百分之一。

总之，我们运用这些看似各不相同的特性，将会创造出或许连我们都未曾想到的无限可能。企业要用数据增强自身的实力，并坚信能够发现那些从未被开拓过的领域，将商业与工业推升到一个全新的高度。

通过《大数据产业革命》一书，你将看到各行各业是如何运用数据将不可能变为可能的。通过《大数据产业革命》一书，你将发现，未来，一切皆有可能。

# 中文版序

1995年，我第一次来到中国。那时的北京首都机场很小，市内只有一家比较大的接待海外来访者的酒店——中国大饭店。当时中国的国内生产总值仅有7280亿美元，而今天这个数字已经变成9.2万亿美元了。可以说，中国经历了翻天覆地的变化。

在过去20年中，中国一直在不断实践各项最前沿和最尖端的科技。从半导体技术到生物医药，中国已经站在了推动科技创新的前沿领域。值此“十二五”计划的收官之年，数据这一广为人知又极具创新力的能量源渗透进了我们的生活，成为一种至关重要的数据资产。我十分期待中国能把握住这股足以引发巨变的力量，在各行各业中再创辉煌。

我在本书中所介绍的许多案例，从第一产业的农业到第三产业的零售业，都可以复制于中国市场。唯一的决定因素在于中国的商业领导者何时作出决策。我非常热切地盼望着能够很快看到这一天，并亲身经历这场发生在中国的大数据产业革命。

罗伯特·托马斯

# 前言

## 故事讲述者

故事是一种强有力的交流工具，本书汇集了我们在过去十年间所看到的各种故事、模型与方法论。我们一直专注于这个新兴的领域：大数据。本书中的故事都是根据真实事件所写，但有时会隐去真实的名称、事件或相关状况。

我讲述这些故事的目的是，通过我们所见证的事件，让我们感受到已经来临的大数据的挑战。我们的初衷是这样的：通过故事这样一种极佳的学习渠道，去借鉴那些领军人物是如何应对世界的变化，是如何在他们各自的行业中颠覆原有传统模式的。我们希望这能为读者带来灵感、勇气与知识，让他们能够拥抱大数据，在企业或组织内部点燃变革的燎原之火。

## 本书的目的

本书的受众范围涵盖了创业家、成熟企业中的管理层，以及那些对大数据会如何影响自己充满了好奇的人们。这本书旨在提供“弹药”击破存在于数据管理员与企业管理者之间的壁垒。数据是一种全新的知识产权，得之占尽先机，失之则追悔莫及。

对于那些故步自封的人而言，数据分析与决策制定之间毫无干系，阅读本书可能会让他们嗤之以鼻。无法利用数据资产的企业最终都将消亡。改善数据与行为之间的连接、机器与人类之间的关联，这需要在人力和财力方面投入巨大的资源。此外，还要求企业组织在这种转型的过程中具备足够的耐心。对于某些企业或组织，他们或许需要通过外部的协助或合作伙伴来充分评估大数据的潜在价值。

可惜的是，许多既得利益者会对大数据革命产生抵触，他们惧怕它对自身职业所产生的影响。这种消极的抵抗很可能将会是徒劳无益的，而那些积极投身于历史性变革的探索者将获得可贵可喜的机遇。正如弗雷（C.B.Frey）和奥斯本（M.A.Osborne）在《就业的未来》（The Future of Employment）一书中所提出的，目前半数职位都将面临着被自动化取代的风险。究其原因，是因为技术的迅猛发展，以及大数据和机器学习的应用和实践正在不断地发生。目前已经有许多日常事务被计算机化和自动化，而最新的科学进步又显示出可能会有更多非常规的知性工作，例如会计、法律工作、技术撰写以及其他许多白领职业也会被自动化所吞噬。

在本书中，我们将反复阐述，为了使企业管理战略在知识型社会中保持可持续发展，也为了不被用数据武装起来的竞争对手所超越，你必须先声夺人领先一步把握住大数据。转型是复杂的，它需要调整技术、流程与人们的行为。毫无疑问，从这种破坏性的变化中得以幸存实属不易。虽然没有适合所有类型与规模的万能良方，但我们提供了一条战略路线图，可以针对不同类型和规模的企业组织量体裁衣。

我们列举了许多实际案例，以帮助大家了解成功的企业或组织是如何改变自己并超越他们的竞争对手的。我们希望据此来阐述大数据时代的成功之道。

## 概要

### 第一部分 山雨欲来风满楼：九大行业的数据转型

第一部分所分享的故事主要用于诠释全书中的模型概念。根据已有的资料，我们可以找出数百个故事，不过我们只摘选出其中这9个最有吸引力的故事，将其作为我们所见到的模型的最佳诠释。阅读时，这些故事的顺序可以随意，我们建议你先挑选一个听起来最有趣的故事开始，然后再阅读剩余的部分。

在第1章“数据改变农业”中，我们首先快速浏览了农业历史和农业技术的发展。这一章深入探讨了这些年农业发展的局限性，以及我们应该如何克服，同时用大量篇幅谈到了数据将如何开始并持续地改变这个传统的行业。正如你将看到的，该行业中的许多参与者至今尚未感知到数据的影响，他们很快就会以迅雷不及掩耳之势被淘汰。

在第2章“被颠覆的医疗”中，我们强调了今天在医学领域中，许多决策都是基于主观判断而非事实，这将导致次优的结果。同时，我们也展现了医学领域的先驱者是如何运用数据逐步打破了传统的行业惯例的。其中还透露出了一个信息，医生的角色与技能将会在数据时代被彻底改变。

在第3章“保险精算师为什么会成为数据科学家”中，我们描绘了一场正在进行的针对如何更好地收集、访问和使用数据的彻底变革。承销与精算服务这些以往可以在很大程度上预测未来的职能或将让位，因为你已经可以实时监控正在发生的一切，并以此定价。全新商业模式的出现将彻底取代传统保险行业赖以生存取胜的技能和工具。

在第4章“个性化的零售与时尚行业”中，显示了数据如何将传统的基于客户群划分的零售行业，转变为一种更具个性化的行业。成千上万名各不相同的个人客户取代了以往的成千上万名客户。在零售行业中有一句座右铭：“只有半数的市场营销是有效的，但我无法辨别那有效的一半。”改变零售行业，不仅仅是利用数据来更加精准地找到客户，更重要的是通过数据转变零售商的态度，以个性化的方式看待每一位客户。

在第5章“彻底改写传统客户服务”中，我们将谈到如何运用数据来增加企业与客户之间的黏合度。通过改善客户数据的收集、关联和利用，企业将能够实时处理个人数据，从而达到客户满意度的最大化。除了单方面应答问题、解决麻烦，数据中包含的个人位置与喜好等信息还能够帮助企业改善服务，交付方案来实现个性化的用户体验，或者预测各项风险。

在第6章“智能机器能在不远的未来完胜人类吗”中，我们先从拜访丹麦与探讨风力涡轮机业务开始。这一章将继续讨论原先分散的机器是如何通过关联和数据走进我们的生活的。生活与机器的结合看似非常科幻，不过现实中该领域已经有了长足的进步，大量机器可辨识数据的生成有助于实现机器的智能化。在物联网中涉及许多能够共享信息的机器。通用电气宣布的工业网络、风力涡轮机的智能化、无人机广泛的应用前景以及特斯拉的车辆管理系统等都是网络中的数据点，他们将改变我们的世界。

在第7章“政府与社会之间游戏规则的变革”中，我们阐述了在加强民众与政府联系方面，数据有着相当大的潜力。相比选举、公民投票、民意调查所需花费的大量资金，社交媒体提供了另一种监督舆论、评估观念以及测试公共政策并进行微调的机制。与此同时，隐私风险是目前许多国家关注的重要问题，并且在一定程度上阻碍了数据开放访问的各项计划。找到一种可靠的办法，例如通过匿名化技术，在应对隐私风险的同时又不会降低数据的质量，这将是一项挑战。本章还将探讨以公私合作的方式运用大数据，从而带来社会经济福祉的潜力。

在第8章“大数据时代企业的可持续发展”中我们可以看到，人们通过互联网和社交媒体，对全球供应链以及我们每天所消费的服务和产品有了日益深入的理解和认识。在面对确保可持续发展的社会责任时，许多企业正在努力成为行业中的榜样，试图抢占先机。除了要具备部署企业可持续发展战略的信心外，还需要对未来场景中可能出现的风险进行评估。搅局者来自那些能够着眼于长期发展的企业。

在第9章“天气对能源的决定性影响”中，我们看到，尽管在数据、模型和技术领域取得了许多进步，天气预测一直以来都是并且今后也仍将是人类所面临的一项挑战。天气预报展示了人类如何通过科学技术预测未来并采取对策。气候与能源之间的紧密关系，为大数据应用于电力供电系统提供了平台，这其中与可再生能源的可变参数密切相关。虽然有着诸多变数，但是通过数据与技术创新更好的结合，会有助于平衡供需，确保能源的持续供应。

## 第二部分 为即将到来的大数据产业变革做好准备

本书的第二部分将第一部分中的9个故事提炼为一系列独立的模型。我们将探讨模型识别的概念在不同环境中的应用，以及对大数据时代所产生的影响。我们在这部分的结尾处详细探讨了我们所看到的54项大数据模型。

在第10章“模型识别的意义”中，我们强调了识别模型是构建系统进行预测、分类或者客户细分的第一步。辨识大数据模型，并对这些模型的轻重缓急进行排列的能力在数据时代至关重要。收集相关的数

据、清理数据、提取相应的特征，分类数据并且评价可信度都是这一流程中的各个组成部分。从理解贝叶斯定律到区分东京鱼市的水产种类，模型识别算法涵盖了从人们的直觉到复杂的机器算法的各个不同方面。

在第11章“迎来崭新的数据时代”中，大数据模型在不同行业中的出现有力地证实了这场革命的巨大潜力。在大数据时代，有三种突出的方式可以构建商业模式。在某些方面，数据造就了笑傲群雄的竞争优势。而在另一些方面，它可以提高现有产品或服务的质量。最后，通过找到合适的企业，提供特定的数据集访问权限，数据本身便可以成为一种产品。

在第12章“大数据的关键模型”中，我们强调了找到大数据模型会有助于那些在数据时代已经做好准备的企业。我们根据本书第一部分的故事，介绍了30项数据要素。这些要素能够更进一步地分解为54项大数据模型，旨在描绘数据时代前沿企业的最佳实践。

### 第三部分 成为大数据产业革命的领导者

本书的第三部分着重介绍了如何在你的企业内部掀起一场大数据革命。你可以从第一部分的各个故事中汲取他人的经验教训，开发出一套更优化的策略，结合自身企业的特定情况采取行动。我们推荐本书第二部分提取出的模型，至于是否适合你的企业，还需要深思熟虑的评估。商业领袖所面对的巨大挑战在于要在尽可能短的时间内启动变革的关键步骤。

第13章“数据带来的商机”着眼于机会往往更加能够激励行为的转变。通过观察石油和数据这两件商品的发掘、提炼与附加价值的过程，你会发现其中有着高度的相似性，这有助于诠释未来的价值创造过程。而经验主义角度看，贝恩资本在2013年发布的一份报告中，首次提到了专注数据的策略所带来的各项收益。早期接受大数据的使用者在他们的行业内拥有前25%的顶尖财务业绩表现，是同行业企业的两倍以上。

第14章“从未停止创新的保时捷”阐述了追求完美是成功的动力。罗伯特从意大利的一段经历开始，谈到保时捷的创始人费迪南德·保时捷的故事。这则故事告诉我们创新、适应性与毅力是成功打造这个誉满全球的跑车事业的关键所在。保时捷的定制设计确保了他们的汽车不仅能够满足不同客户的需求，同时还能针对性能、品质与价格进行优化。定制设计这无疑是数据时代企业领导人达成目标的最佳方式。

第15章“彪马与阿迪的恩怨大战”说明了努力是所有成功企业的特质之一。而彪马与阿迪达斯之间激烈的商业大战源于手足相争，并非一般意义上的商业竞争。1993年，约亨·蔡茨在接管彪马之后的13年间，使彪马的股价飙升了4000%。作为一位伟大的商业领袖，蔡茨运用大数据分析创造出一种环境盈亏核算系统，至今仍被积极提倡并被其他企业所广泛采用。

第16章“制定基于大数据的企业决策”阐述了该方法分为7个具体的步骤，为企业组织推荐了一种部署大数据模型的方式。这些方法一般是按顺序使用，有些企业组织或许已经完成（或者开始）了其中的几步。该方法论只是一个路线图，你也可以只选择自己需要的部分。

第17章“大数据体系架构”强调企业按照第16章的顺序，应全力以赴专注在落实执行上。第17章便是迈向执行的第一步，只有迈出第一步，才能领略大数据参考体系架构的全貌。其中的重点是参考体系架构的业务视图和逻辑视图。

第18章“参考体系架构中的业务视图”深入介绍了参考体系架构中业务视图的各项组成。我们介绍了一家虚构的零售商特伦克公司，用以阐述通向数据时代的征途。业务视图中包含分析网络、数据虚拟化、数据引擎、运维管理、数据管理以及用户界面、应用程序和业务流程等组件。

第19章“体系架构中的逻辑视图”说明一旦企业认识到其体系架构的业务视图，在实施过程中便需要相应的逻辑视图。逻辑视图在数据时代更细化地描述了如何具体执行，需要关注五项基本组成部分：数据摄取、数据分析、信息洞悉、操作型数据以及数据管理。

第20章“面向未来的体系架构”作为本书的结尾部分，详细介绍了特伦克这家在第18章中虚构的零售商的数据之旅，我们追踪了他们是如何应用方法论，部署参考体系架构的每一个具体步骤的。

我们之所以选择这种讲述故事的方式是因为这是一种最有效的沟通方式。书中的每个部分和章节都交叉贯穿着各种实例与配图，其目的在于增加趣味性，以便在读者的脑海中留下持久的印象。当你在大数据时代翱翔的时候，我们希望这些故事提供的不仅仅是经验与指南，同时也为你带来灵感。变革是困难的，而灵感往往会成为突破困难的力量。

这本书与大数据时代的其他事物一样，不应该是静止不变的。书中介绍的工具与环境也在发生着变化，因而在方式及技巧上同样需要经常调整。你可以通过浏览这本书的网站[www.bigdatarevolutionbook.com](http://www.bigdatarevolutionbook.com)来了解相关的变化、与本书相关的内容以及访谈。

欢迎访问：电子书学习和下载网站 (<https://www.shgis.cn>)

文档名称：《大数据产业革命：重构DT时代的企业数据解决方案》（美）托马斯（Thomas, R.）

请登录 <https://shgis.cn/post/923.html> 下载完整文档。

手机端请扫码查看：

