

大数据供应链：构建工业4.0时代智能物流新模式

作者：（美）娜达·R·桑德斯

版权信息

COPYRIGHT

书名：大数据供应链：构建工业4.0时代智能物流新模式

作者：【美】娜达·R·桑德斯

出版社：中国人民大学出版社

出版时间：2015年6月

ISBN：9787300214559

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

推荐序

魏祁蔚

广东省采购与供应链协会秘书长

很荣幸获邀为《大数据供应链》写序，书稿拿到手中来回看了两遍。娜达·桑德斯的确是一位了不起的学者，作为美国供应链排名前8%的行业专家，她所著的《运营管理》早就被行业中的一些翘楚奉为主桌，遗憾的是那部巨著至今没有完整的中文版。在国内缺乏专业供应链人才的现实下，一本最前沿的关于供应链解决方案的新书无疑是甘露。

最近有另外一本书火得不行，是马云推荐的《互联网+：从IT到DT》，风清扬大侠已经很清楚地告诉大家，大数据是巨大的宝藏。而国内对大数据的书籍虽然不少，但大部分处于普及型的书籍，真正给予实际应用的书籍少之又少，感谢遇见这本书，感谢出版社能快速地翻译出版。

在信息化时代，谷歌每天有数十亿的搜索量，这个数量是美国国会图书馆纸质材料的上千倍；Facebook每小时图片上传量达到数千万；沃尔玛每小时交易量达到100万；纽约证券交易所每天股票交易量近10亿。

在未来，得数据者得天下。谷歌、Facebook、雅虎、亚马逊、苹果，他们都会收集大量的数据。国内的阿里巴巴、腾讯、百度、小米也各显神通地展开了数据争夺战。除了广告点击量和你的搜索关键词之外，这些企业还会收集你的地理位置信息、你的IP、互联网提供商、你使用的浏览器、你的邮箱地址和手机号码，甚至他们还会收集你的面部信息数据、健康指数。为了获取数据，国际、国内巨头无所不用其极，先是Google宣布气球Wi-Fi计划，后有Facebook的无人机Wi-Fi计划，当然还有很多巨头虎视眈眈。未来最了解你的不是你的父母或者伴侣，而是手机和电脑。

信息时代的今天，我们拥有了越来越多的数据源与数据种类，文字、数字、图片、语音、影像……，然而拥有众多数据就是大数据吗？答案当然是：不是。大数据应用的关键还是在于大数据分析，大数据的出现给我们提供了一种洞悉世事的新视角：透明的因果！

例如家喻户晓的美国塔吉特百货（Target）识别孕妇的案例。大数据分析能够识别已怀孕4~6个月的妇女所购买的商品。当塔吉特百货向一位16岁的小女孩赠送婴儿优惠券时，女孩的父亲对此非常生气，因为这意味着他的女儿竟然怀孕了。这位消费者愤怒地投诉了塔吉特百货。然而，事实却证明他的女儿的确怀孕了，他也因此向塔吉特百货道歉。塔吉特百货大数据分析借此名扬天下。

供应链管理作为竞争的核心战略，对于未来用户需求个性化和过去生产制造规模化的矛盾提供了不可替代的解决方式，譬如柔性供应链、敏捷供应链。而要成为洞悉先机的领先者，必须在数据的世界里获取竞争优先权；要成为首屈一指的企业，必须采用“大数据+高科技”的双轮驱动同步管理供应链。从大数据分析中你将获得敏锐的洞察力和全新的价值观，赢得改变市场、组织和商业关系的先机。

通过检验大数据中两种变量之间可能的联系，对事物进行因果检验，能够让我们清楚进一步检验的方向，提供信息以便建立模型、分析因果关系、优化策略，以便运用等级分析法（如供应链优化）对数据作进一步分析，从而为整个社会及商业活动带来前所未有的供应链新图景。

大数据使得建立数据规则成为可能，这种规则能够识别不符合规范的数据，即异常值。这种智能在风险管理中将改变游戏规则。

大数据分析法可以让更多中小企业获得和大公司一样的资源与机会。未来的竞争将会对数据源展开激烈的争夺。且分析数据的能力将获得前所未有重视，因为从数据的角度去验证和预测未来会成为一种能力。当这种能力应用于供应链将给我们的商业和生活带来神奇的变化。这也意味着企业将以前所未有的方式洞察商机，并拥有更好地解决问题的能力。未来几乎人人都可以借助大数据解决问题，人们能更客观更实事求是地多维度验证商业与用户的方方面面。

当今世界的数据量与日俱增，在无处不在的海量数据中，分析后所得到的资讯将是一笔巨大的商业财富，它可以让你的企业拥有无可匹敌的优势。而对于学习如何处理这些数据，商业界才刚刚起步。

若要成为大数据的赢家，要时刻小心，不要被数据蒙蔽。谷歌曾经尝过极端数据独裁的滑铁卢。应用大数据而不能过分依赖数据，否则可能会导致供应链的灾难，如同希腊神话中伊卡洛斯（Icarus）所犯的错误——他搭着鸟羽和蜡制成的飞行翼翱翔天际，但却太相信自己的飞翔技能，飞得太高被太阳融化翅膀而折翼坠海。

前言

为什么会存在这些问题呢

与淘金热的兴起相似，相关业界和学术界对“大数据分析”这一话题的兴趣犹如潮涌——引起了太多的骚动，但却鲜有人知到底该去何处寻找实质性的回报。许久以来，应用新兴技术对于企业竞争力的保持至关重要，这已成为不争的事实，然而，技术本身在向构造型方式的转变过程中，需要思想领袖紧跟其步伐。

大部分企业尚未学会运用大数据分析来指导其供应链运作。有些企业淹没在数据的海洋中，不确定如何使用它们驱动供应链；有些企业仅仅在部分环节利用或实施了大数据分析，而缺乏系统的、协调的整体调控，其结果只能是部分获利，缺乏见解和竞争力，整个供应链到处存在着低效率和成本超限的现象；而其他企业，尤其是一些中小型企业，甚至对于从何处开始着手大数据分析都一头雾水。

为什么会存在这些问题呢？原因是企业对于如何能够以有意义的、有效控制成本的方式实施大数据分析法缺乏一幅明晰的路线图指导。

本书尝试着解决上述问题，以期为企业如何在整个供应链中实施大数据分析法，从而将信息转化为情报并获取竞争优势提供一套系统框架。这种应用大数据分析法端对端的视角为企业提供了在大数据方面亟须用来连接战略和战术的概念框架。此外，本书的路线图为企业领导者展示了实施大数据分析法的企业在组织结构上需要做出的改变。

第一部分

“大数据”驱动的供应链

第1章 游戏规则的改变者

当今世界的竞争方式与以往相比已截然不同，它就像一场海啸，在转变整个世界的同时也引发许多灾难。正如古登堡（Gutenberg）发明的印刷机改变了世界一样，大数据潮流也同样在促使商业和社会发生结构性变化。身处这样的时代，不转变就只能落后。

以美国图书零售巨头鲍德斯（Borders）的遭遇为例。1971年，这家公司在密歇根州的安阿伯市（Ann Arbor）成立了第一家书店，那时图书业的发展是另一番景象。然而，时至40年后的2011年，这家连锁实体书店却倒闭了。究其原因还是鲍德斯没有跟上互联网和数据时代的步伐。由于没有意识到游戏规则已经改变，鲍德斯只能将其网上图书销售业务外包给亚马逊（Amazon），因此，只要一浏览鲍德斯的网页就会转到亚马逊的网站上。从以往的游戏规则来看，这种做法似乎是个明智之举。然而，在新时代环境下，这一举措却大有问题。

鲍德斯摒弃自己的竞争优先权去追随亚马逊的这一做法没有考虑到数字世界的竞争本身就是获取竞争优先权，把控制权移交给另一公司的做法只会减少自己的客户群。另外，鲍德斯全然没有意识到世界的数字化进程，不能像亚马逊和巴诺书店（Barnes & Noble）那样欣然接受电子书籍。这导致走进鲍德斯就如同走进往昔的书店一般，鲍德斯的结局可想而知。

在鲍德斯所处的传统实体书店的竞争世界里，图书销售商可以追踪到已售出与未售出的书籍，客户忠诚计划可以将书籍购买与单个顾客联系起来，但是，仅此而已。然而，当顾客网上购书后，网上零售商能够追踪到顾客的任何购买行为，包括获取顾客的浏览信息、浏览方式、浏览某一网站时长以及对促销活动和图书页面排版的反应等。如今，基于多种多样的顾客特征，网上零售商能够建立起微观顾客群，为每位目标消费者量身定制促销计划。他们开发算法以预测每个顾客想要阅读的书籍，一旦顾客对某一推荐书目作出回应，这些算法就能够获取这一信息，进行自动处理并进一步得到优化。而诸如鲍德斯在内的传统图书零售商是无法获取这些信息的，自然也无法与网上零售商展开实时竞争。

那么亚马逊又采取了怎样的做法呢？这家公司推出了电子书籍阅读器，并说服了许多出版商在其阅读器上发行新书以垄断市场。亚马逊将图书数据化，也就是将图书转化成可列表分析的量化格式，这一创举使得该公司无所不能。亚马逊不仅可以向读者推荐书籍，还可以利用各类算法发现不同书籍主题之间存在的那些并不显而易见的关联。起初，亚马逊的目标是发展成为最大的网上书店，而其融入互联网时代，发展新科技，运用数据驱动的种种创举却早已超越了这一目标。眼下，亚马逊正朝着成为世界上首屈一指的零售商的目标迈进。亚马逊深知，要达成这一目标，意味着其必须采用大数据和高科技同步管理整条供应链。众所周知，亚马逊的首席执行官杰夫·贝索斯（Jeff Bezos）要求，只有在精确量化消费者体验后，才能推出新产品。数据与科技已经用来整合供应链上的每一环节，从顾客订单到满足需求、库存管理、劳务、仓库、运输以及交货。

亚马逊并不是唯一一个采用大数据来发展业务的公司。全球的企业巨头们都已享受到利用大数据所带来的成功。在这个日新月异的时代，沃尔玛、ZARA、UPS公司、特易购（Tesco）、哈拉斯（Harrah's）、前进保险公司（Progressive Insurance）、第一资本金融公司（Capital One）、谷歌、易趣（eBay）等企业勇于变革、引领潮流，从而得以制胜。这些企业从大数据分析中获取了敏锐的洞察力，建立了全新的价值观，改变了整个市场、组织和商业关系。

大数据概要

要全面了解大数据分析法对企业的影响，我们首先要明白什么是大数据。本小节主要阐述大数据的基本内容，包括大数据分析法的重要概念，以及这些概念如何在商业中得到创新应用。本书接下来会讨论大数据分析法在供应链管理中的应用，本节内容是为后续的讨论做准备。

大数据

“大数据”这一术语特指大量数据集，这些数据集的规模巨大，以至于无法存入电脑用来处理信息的常规存储器中。这种数据可以被获取、存储、传达、整合、分析。目前，对于大数据的规模并没有特别的定义，例如，多少太字节（terabytes）或多少千兆字节（gigabytes）可以称为大数据。这是由于大数据并不是一成不变的，科技在不断发展，人们称之为大数据的数据集规模也在不断增大。

不断增多的数据量需要不断更新的分析工具，鉴于此，诸如谷歌的MapReduce软件和与其并驾齐驱的开放源代码软件Hadoop等新型的信息处理技术才得以发展，使得企业能够处理规模不断扩大的数据。更为重要的是，如今数据集的处理技术与传统技术不同，并不需要数据排成整齐的行列就可以进行分析。

大数据的不同形式可能来自各类不同的资源。它可以结构化，也可以非结构化；可以是一个数字序列、一段音频、一段文字或者一段对话；可以是销售点（POS）数据、射频识别（RFID）数据、全球定位系统（GPS）数据；也可以是位于Twitter、Facebook、呼叫中心或者消费者博客中的数据。无论以何种形式存在，如今的高级分析工具都能够解读这些数据。

分析法

分析法运用数学和统计学分析这些海量数据，这种方法通常称为“大数据分析法”，它能够让我们以前所未有的方式深入探索周围的世界，并运用它来推断未来某一事件发生的概率。

生活中，人们对大数据分析法实际上早已习以为常。我们习惯利用电子邮件过滤器判断一份邮件是否为垃圾邮件，或者判断键入单词“teh”时的本意是否应为“the”。这些系统的关键在于其预测结果是建立在源源不断的数据分析的基础之上的。不仅如此，随着更多的数据输入，这些系统能够通过追踪优质信号与模式不断进行自我优化。想象一下，你可以通过标记相似邮件的方式告诉电子邮件过滤器哪种类型的邮件是垃圾邮件。

大数据让沃尔玛了解到消费者在刮飓风期间，更喜欢囤积果酱小饼；让易趣识别出哪种网页设计能够创造最高的销售量；让前进保险公司学会了如何通过区分风险类别使得保险费最大化。

除了上述这些大型企业外，即使是中小型企业也可以从大数据中受益。在线音乐设备零售商“音乐家的朋友”（The Musician's Friend）运用基本分析法，通过比较其网页的不同版本来确定消费者偏好。与最初的网站主页相比，更受欢迎的网页使得设备销售量增加了35%。这一小小的改变使得该企业的投资回报率得到了很大提高。

完美的二重奏

需要说明的是，没有分析法的大数据仅仅是一堆数据，数年来，我们已经积累了大量数据；而没有大数据分析法也只是一些数学和统计工具，例如，相关性和回归分析早已存在数十年。实际上，谷歌的研发负责人彼德·诺米格（Peter Norvig）说得好：“没有更高级的电脑程序，只有更多的数据。”

然而，为世界带来更大影响的却是两者的结合。大数据与分析法相结合能够让我们对信息有着深刻的认识，并把信息转化成商业智能（见图1-1）。不仅如此，大数据和分析法互相依存，不断地应用简单的分析工具能够使它们的性能得到持续的提高与优化。试想，随着你越来越多地找出垃圾邮件，电子邮件过滤器会自动学习这一过程，并能够越来越准确地识别出垃圾邮件。正因如此，“大数据分析法”这一术

语将贯穿全书，用来表示“运用分析法分析大量数据集”。

□

图1-1 将信息转化成情报

现代化的运算能力

企业如何从海量数据中获取信息呢？如今，企业能够以前所未有的廉价成本获取大规模运算能力，这使得从大量数据中获取有用的信息有了可能。海量数据与大规模、廉价的运算能力结合，能够实现更大规模的数据处理，这是小规模运算能力无法完成的。运算能力的不断提高使得计算机性能不断优化，这使得高级分析法能够应用于处理大量非结构化数据集。

正如摩尔定律所言，计算机处理能力会不断提高。这一定律是以英特尔（Intel）共同创始人之一高登·E·摩尔（Gordon E. Moore）的名字命名的。该定律指出，同等数量的资金可以买到2倍于两年前的运算能力。事实证明了这一规律的正确性。越来越快的计算机信息处理速度和越来越大的存储容量都是有目共睹的事实。同样，云计算也使得存储空间不断扩展。云计算是指通过互联网处理高度可伸缩计算资源的能力。由于资源可以共享，云计算通常比安装室内电脑更廉价。此外，驱动大量系统运行的算法性能也在不断提升。因此，大数据在数据集规模不断提升的同时还伴随着迅速提高的信息处理能力和算法性能。

作为开放源代码技术平台，Hadoop软件因采用了大数据解决问题而受到广泛关注。复杂和结构化数据混合的大数据通常不能与表格完美结合，Hadoop正是为处理这种数据而量身打造的。率先将Hadoop应用于网页检索以及用户行为监督以改善算法的是谷歌，其后，雅虎则进一步将其应用于公司层面。Hadoop采用的是位于众多服务器上的分布式应用程序，其数据库也分布于大量电脑中。分散在多个服务器上的数据大大提高了运算速度，同时减少了数据库中每个表格的行数，从而大大缩小了检索范围，提高了搜索效率。数据库通常被分隔到若干个被称为“数据库碎片”的硬盘分区中。一个数据库碎片可保存在单个硬盘上，数据库碎片集则保存于多个硬盘。数据库碎片可以使运算性能得到极大的优化，碎片内的数据库分区可以建立在现实世界分区的基础上，这不但有助于分析——例如，区别加拿大和美国的用户分析，而且能够进一步简化特定数据分区检索以及跨区比较评估。

解决问题的新方案

就在几年前，分析法和运算能力还只是一些数据狂人才关注的事情。然而今天，大数据已成为从医疗保健到产品制造等各行各业的领军企业不可或缺的信息资源。几乎所有组织都已具备了对数据进行采集、存储、集中、整合并作出深入分析的能力。随着运算能力、数码存储以及云计算成本的持续走低，这一趋势还将不断发展。科技进步会进一步打破技术壁垒，给企业，尤其是中小型企业创造公平竞争的环境。试想一下，现如今，个人用户以不到600美元的价格就可以购买一块足够存储世界上所有音乐的磁盘驱动器。存储1TB数据的成本已由20世纪70年代的100万美元降到了今天的50美元。

这对企业来说意味着什么？简单来说，分析技术的不断精细化和计算机性能的日益提升意味着企业能够以前所未有的方式来洞察商机。大数据不仅仅是数据那么简单，它代表着一种能够更好地解决问题的能力。如今，几乎人人都能够运用大数据解决问题。

有何不同

数年来，各种企业都在通过不断收集、分析数据来获取市场信息。因此，我们自然想知道，当今世界与之前相比到底有何不同？

不同之处在于规模。这里的规模指的是数据量与分析数据的运算能力，两者结合能使企业客观、实事求是地洞悉商业的方方面面。这不仅使企业面临更高层次的竞争，而且这种智能应用还可以使企业获得竞争优势。我们生活在数据爆炸的时代，最近《纽约时报》的一篇文章写道：“就像煤炭和铁矿在工业革命中所起的作用一样，数据是信息经济不可或缺的原料；而对于学习如何处理这些数据，商业世界才刚刚起步。”

大数据的3个特点是：大容量、高速率与多样性。这3个特点将大数据与以往的数据区别开来（见图1-2）。

□

图1-2 大数据的三大特点

大容量

如今的数据量已经大到了可以被称为海量数据，且无处不在。谷歌每天有30亿的搜索量，这个数量是美国国会图书馆纸质材料的上千倍；Facebook每小时图片上传量达到1000万；沃尔玛每小时交易量达到100万；纽约证券交易所每天股票交易量达到4.9亿。当今世界的的数据量与日俱增，据预计，全球射频识别标签销售量将从2011年的1200万个增长到2021年的2090亿个。

生产商与零售商从供应链获取数据，包括销售点数据、全球定位系统与射频识别数据、设备传感器发出的数据以及社交媒体源数据。几乎所有公司都在后台安装信息技术系统。我们生活在数据时代，据估计，若把太字节和艾字节（exabytes）转换成有意义的术语，公司与个人产生与储存的数据量能够填满6万座美国国会图书馆。这些数据都去了哪里呢？它们不断累积，而且每时每刻都在增长。这就是大数据。但是，很多公司却对他们所处理的数据一无所知，也并不知晓其价值所在。数据可以用来交易，并具有经济价值，是一种新型资产。

高速率

数据量的确很大，而且在飞速增长。

数据如同滔滔洪水倾泻到商业与日常生活的方方面面。企业不断获取与日俱增的可交易数据以及消费者、供应商、运营商的信息。日常生活中，许多像移动电话、智能能量计、汽车、工业机器的设备都安装了上百万种传感器，这些设备在所谓的“物联网时代”获取和交流数据。

企业与个人产生了大量的废弃数据，这种数据是日常活动的副产品。消费者每天发短信、与人交流、浏览网页、检索信息、购买商品都会产生数据痕迹，企业则能够获取、监视和分析这些数据。社交媒体、智能手机、个人电脑与台式电脑使世界上许多人可以不断增加可用的大数据。多媒体内容数量的不断增加对高速增长的大数据产生了重要影响。例如，高清视频每秒钟产生的字节是存储一页文字所用字节的2000倍。

企业在以持续增加的粒度与频率收集数据，他们可以获取每一笔交易，而且会在每笔交易中增加越来越多的个人信息。因此，企业在不同环境中不断地收集着消费者的行为信息，然后将这些数据整合或者拆分成无限种组合，从而系统地做出最优市场战略决策、为消费者量身定制购物体验计划、设计灵活的目标市场产品、优化交货系统。

多样性

大数据有着多种表现形式。通常人们认为数据只能整齐地排列在一栏或一行中。而事实上大数据的形式多样，而且这种多样性在不断增加。它可能是结构化的数据，并以我们熟知的形式呈现出来；也可能以传感器读数、手机全球卫星定位系统信号、售货点以及射频识别数据的形式呈现；大数据还可能非结构化的，例如，手机短信、语音信息、社交网络源和消费者在博客中的交流信息等。

大数据的来源多种多样——它可能来自射频识别传感器与零售商销售点、全球定位系统传输的地理定位信息，可能来自设备上的振动与温度传感器，也可能来自社交媒体源（例如，Twitter的“转发”功能与Facebook的“点赞”功能）、产品维修记录与客户投诉（见图1-3）。大多数智能手机安装了全球定位系统，具备追踪功能。轿车中的芯片、传感器与软件无处不在，这些设备可以在轿车保养时将其性能以数据的形式上传到生产轿车的电脑里。目前，一辆典型的中档汽车装有约40个微处理器与电子器件，这些设备占了整辆车成本的1/3。那么，汽车制造商除了通知你去保养爱车之外还需要告诉你什么呢？

□

图1-3 数据源无处不在

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

文档名称：《大数据供应链：构建工业4.0时代智能物流新模式》（美）娜达·R·桑德斯 著.ep

请登录 <https://shgis.cn/post/924.html> 下载完整文档。

手机端请扫码查看：

