

双色  
印刷

玩转移动互联网营销系列

案例多：通过近50个经典行业案例，全方位图解大数据营销与运营方法。  
剖析深：知晓16个行业领域引流技巧、运营手段、营销方法。

赚钱公司的  
大数据是  
这样玩的！



# 玩转大数据

商业分析 + 运营推广 + 营销技巧 + 实战案例

海天电商金融研究中心 编著

大数据营销、公司运营，一本就够！

- 价值挖掘  精准定位  客户体验  高效服务  风险管理  交通能源
- 网络通信  金融理财  娱乐传媒  企业管理  生产制造  特色餐饮



清华大学出版社

玩转移动互联网营销系列



# 玩转大数据

商业分析 + 运营推广 + 营销技巧 + 实战案例

海天电商金融研究中心 编著

清华大学出版社  
北京

## 内 容 简 介

一家公司，如何运用大数据进行客户定位、营销推广、品牌传播？

一个行业，如何运用大数据进行产业分析、前景预测、商机挖掘？

本书将全面揭秘大数据价值挖掘、平台构建、营销定位、生活服务、社会互动、风险管理在未来应用，特别是对销售行业、餐饮行业、网络通信、交通能源、医疗领域、娱乐传媒、生产制造、企业管理、金融行业、旅游行业、游戏行业和房地产行业等影响力大的行业或领域进行剖析，帮助读者知晓赚钱公司的大数据是怎样玩转的！

书中内容零基础、全图解，通过 12 个行业应用挖掘+17 章专题内容详解+400 多张图片图解展示，深度剖析了大数据的商业分析、运营推广、营销技巧和实战案例，让您一书在手，轻松玩转大数据！

本书对象为对大数据分析、营销感兴趣的所有读者，特别是企业经营者、管理者、互联网营销岗位的工作人员等。

本书封面贴有清华大学出版社防伪标签，无标签者不得销售。

版权所有，侵权必究。侵权举报电话：010-62782989 13701121933

### 图书在版编目(CIP)数据

玩转大数据：商业分析+运营推广+营销技巧+实战案例 / 海天电商金融研究中心编著. —北京：  
清华大学出版社，2017

(玩转移动互联网营销系列)

ISBN 978-7-302-44985-0

I. ①玩… II. ①海… III. ①企业管理—数据处理 IV. ①F272.7

中国版本图书馆 CIP 数据核字(2016)第 216216 号

**责任编辑：**杨作梅

**装帧设计：**杨玉兰

**责任校对：**张彦彬

**责任印制：**刘海龙

**出版发行：**清华大学出版社

**网 址：**<http://www.tup.com.cn>, <http://www.wqbook.com>

**地 址：**北京清华大学学研大厦 A 座 **邮 编：**100084

**社总机：**010-62770175 **邮 购：**010-62786544

**投稿与读者服务：**010-62776969, c-service@tup.tsinghua.edu.cn

**质 量 反 馈：**010-62772015, zhiliang@tup.tsinghua.edu.cn

**印 刷 者：**三河市君旺印务有限公司

**装 订 者：**三河市新茂装订有限公司

**经 销：**全国新华书店

**开 本：**170mm×240mm **印 张：**17.5 **字 数：**346 千字

**版 次：**2017 年 1 月第 1 版 **印 次：**2017 年 1 月第 1 次印刷

**印 数：**1~3000

**定 价：**49.80 元

# 序言

发生于 2015 年的各种轰轰烈烈互联网转型、跨界、创新、创业方兴未艾地渐渐远去了。2016 年，App、微信、大数据、Wi-Fi 的营销，已作为互联网及移动互联网最为典型也最为火爆的营销手段汹涌来袭！O2O、电商、微商、自媒体等，依旧在炮火纷飞中快速成长。

作者团队基于过去的经验、现在的经历、未来的趋势，特策划了本套“玩转移动互联网营销系列”丛书，力求从产业痛点、行业分析、运营推广、营销技巧、实战案例等角度，为大家做出一点分析，奉献自己的绵薄之力。本套图书具体书目如下：

- 《玩转 O2O：商业分析+运营推广+营销技巧+实战案例》
- 《玩转 App：商业分析+运营推广+营销技巧+实战案例》
- 《玩转微信：商业分析+运营推广+营销技巧+实战案例》
- 《玩转大数据：商业分析+运营推广+营销技巧+实战案例》
- 《玩转电商：商业分析+运营推广+营销技巧+实战案例》
- 《玩转微商：商业分析+运营推广+营销技巧+实战案例》
- 《玩转 Wi-Fi：商业分析+运营推广+营销技巧+实战案例》
- 《玩转自媒体：商业分析+运营推广+营销技巧+实战案例》

本套图书在编写时，时而会从行业产业的高度，进行商业模式的分析；时而会从公司经营的角度，讲解推广和营销的技巧；时而会从实战的角度，精选案例进行剖析；时而从最难点和痛点入手，专门解决客户最头痛的问题；最后，这些内容的核心，就是帮助用户解决某个问题或某类问题！

作为电商类的图书，有些经验技巧是互通的，比如 O2O 的商业模式，与 App 线上线下运营，可能极为相关，而大数据与电商的关系极为密切，如微商最大的痛点引流，却也是自媒体最重要的内容，因此，本套图书的知识内容是可以融会贯通、相互借鉴的。

借用某款口香糖的一句广告词：两粒一起吃，效果可能更好哦！同理，对电商、微商、自媒体等感兴趣的朋友，多参考两三本图书的内容，收益会更大。

本套图书在编写时，采用了 MBA 逻辑图解的办法，更加直观、形象地展示文字知识和逻辑关系，希望读者细心体会，学有所获。

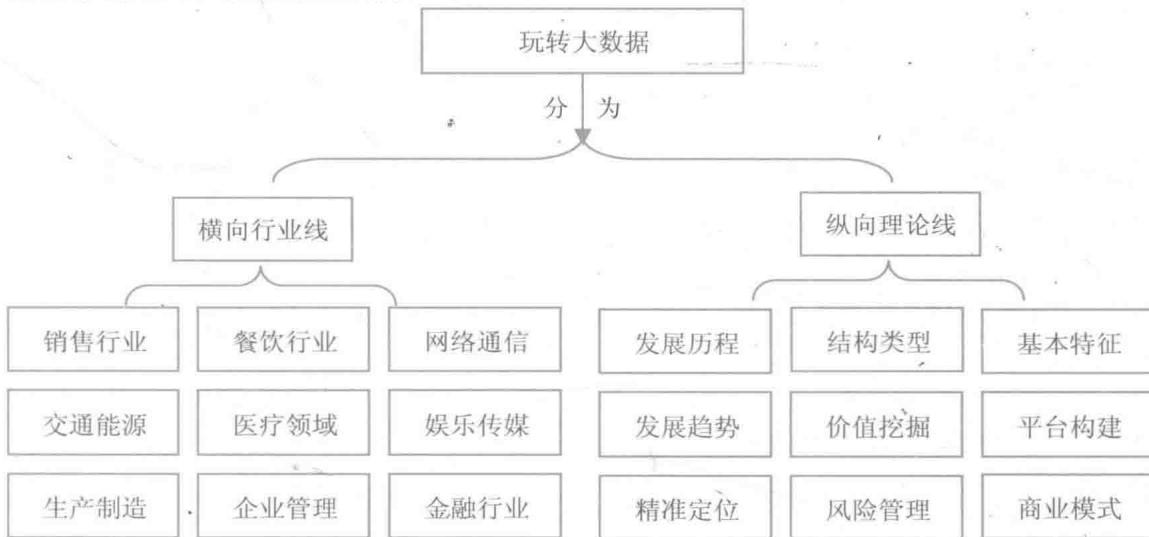
# 前言

## ■ 写作驱动

互联网和移动互联网的快速发展使海量数据得以产生，这些数据的分析、价值挖掘和应用给社会各行业带来了更多发展机会，引领它们走向“大数据+互联网”时代环境下经济发展的春天，开启一个数据化、智能化、信息化的新时代。

本书是以大数据为核心、以营销为根本出发点的专著，以图解的方式全面、深入地诠释大数据的主要特征、发展过程、价值挖掘、平台构建、营销定位、生活服务、社交互动、风险管理等行业应用，特别是结合了每一个与营销相关的行业内容，如大数据影响、发展趋势、应用策略、挑战、商业模式等，围绕相关内容全面解析了诸多行业的大数据应用。

本书紧扣大数据，采用集理论、案例和技巧于一体的结构框架，从横向行业线和纵向理论线全面剖析大数据(见下图)，让您轻松懂得怎样利用大数据创造价值，开拓新的市场空间，焕发企业活力。



## ■ 本书特色

(1) **技巧丰富，包含 100 多种营销方法。**本书巧妙地将 100 多种大数据分析、营销知识与运营方法嵌入行业案例中，生动形象地通过案例，将营销手段与运营方法表

述出来，让读者能够快速吸收、全面掌握大数据营销与运营的相关事宜，成为行家里手。

**(2) 实践性强，渗透近 20 个行业领域。**行业涉及销售、餐饮、网络、通信、交通、能源、医疗、娱乐、传媒、生产、制造、企业管理、金融等人们生活与工作的各个领域。列举大量案例，进行透彻的讲解和分析，使读者只需一本书就通晓整个行业产业的大数据营销与运营方法。

**(3) 易于理解，构建 350 多个逻辑图解。**对案例大数据进行专业的剖析，从大数据应用的方法、应用方法的好处等方面，通过形象的逻辑图解，将大数据营销手段与运营方法进行详细分析，从而推动读者进入和玩转大数据新时代！

## ■ 图解提示

本书是侧重大数据实际应用的实战专著，采取了全图解的方式进行分析。书中的 350 多个图解能够帮助读者快速掌握重点和了解核心知识，为降低读者的阅读成本作出了努力。但是，需要注意的是，读者需要在阅读过程中注意其逻辑关系，以便更好、更快地理解本书内容，从而感受阅读的知识性和趣味性。

## ■ 作者分工

本书由海天电商金融研究中心编著，参与编写的人员有周玉姣、刘胜璋、刘向东、刘松异、刘伟、卢博、周旭阳、袁淑敏、谭中阳、杨端阳、李四华、王力建、柏承能、刘桂花、柏松、谭贤、谭俊杰、徐茜、刘嫔、苏高、柏慧等人，在此一并表示感谢。由于作者知识水平有限，书中难免有错误和疏漏之处，恳请广大读者批评、指正，联系邮箱：feilongbook@163.com。

海天电商金融研究中心

# 目录

## 第1章 发展揭秘：全程破解大数据 ... 1

1.1	相关了解，全面分析 .....	2
1.1.1	大数据的发展历程 .....	2
1.1.2	大数据的4大特征 .....	4
1.1.3	大数据增长的结构类型 .....	5
1.1.4	大数据发展的三大趋势 .....	7
1.1.5	大数据视角下的世界 .....	9
1.2	深入了解，营销获益 .....	10
1.2.1	新型营销模式的形成 .....	11
1.2.2	新型业务模式的发掘 .....	12
1.2.3	存量客户的价值发掘 .....	12
1.2.4	新客户资源的高效获取 .....	13
1.3	核心建设，把握机遇 .....	14
1.3.1	4G时代下的大数据 产业链 .....	14
1.3.2	大数据营销机会的挖掘 .....	15
1.4	商业智能，价值转型 .....	16
1.4.1	大数据下的商业智能 概述 .....	16
1.4.2	商业智能的大数据基础 .....	17
1.4.3	商业智能的行业掘金 .....	17
1.5	综合利用，未来曙光 .....	19
1.5.1	必然走向的大数据 .....	19
1.5.2	大数据时代的业界生态 .....	19
1.5.3	大数据的未来应用 .....	20
1.5.4	大数据未来的发展要求 .....	21

## 第2章 价值获取：深度挖掘 大数据 ... 23

2.1	数据挖掘的相关知识 .....	24
2.1.1	数据挖掘的基本概念 .....	24

2.1.2 数据挖掘的商业解读 ..... 25

2.1.3 数据挖掘的具体计算 ..... 25

2.1.4 数据挖掘的一般过程 ..... 26

2.2 基础设施的建设与发展 ..... 27

2.2.1 云计算数据中心 ..... 28

2.2.2 存储服务器 ..... 28

2.2.3 全面虚拟化模式 ..... 29

2.2.4 虚拟化网络模式 ..... 30

2.3 互联网数据库的营销应用 ..... 31

2.3.1 免费Wi-Fi的客户数据  
搜集 ..... 31

2.3.2 用户头像的信息获取 ..... 32

2.3.3 Immersion的用户邮件  
挖掘 ..... 33

2.3.4 LinkedIn社交数据的商业  
分析 ..... 34

2.4 不同行业的大数据源 ..... 35

2.4.1 文本数据的用户情感分析 .... 35

2.4.2 电网数据的用户需求分析 .... 37

2.4.3 车载信息数据的风险评估  
分析 ..... 37

2.4.4 遥测数据的活动状况分析 .... 38

## 第3章 平台构建：大数据分布 计算 ... 41

3.1 分布式计算的相关概念 ..... 42

3.1.1 云计算系统的运行概述 .... 42

3.1.2 分布式文件系统的数据  
存储 ..... 43

3.1.3 分布式计算系统的劣势 .... 44

3.2 Hadoop分析技术 ..... 45

# 目录

3.2.1 Hadoop 的含义概述 .....	45
3.2.2 Hadoop 的 4 大特点 .....	47
3.2.3 Hadoop 的企业应用 .....	47
3.2.4 Hadoop 的拓宽应用 .....	48
3.3 平台搭建与营销效果 .....	50
3.3.1 大数据平台搭建 .....	50
3.3.2 英特尔的云生态圈构建 .....	51
3.3.3 公有云解决方案的应用 选择 .....	54
3.3.4 云创存储的智能门户 平台 .....	55
<b>第 4 章 精准定位：大数据策略 营销 .....</b>	<b>57</b>
4.1 做好细分，客户定位制胜关键 .....	58
4.1.1 客户属性细分 .....	58
4.1.2 精准定位的地位 .....	59
4.1.3 目标客户群定位 .....	60
4.1.4 企业客户细分 .....	61
4.1.5 二次细分与动态调整 .....	63
4.2 品牌传播，企业客户定位优选 .....	64
4.2.1 企业品牌的基本含义 .....	65
4.2.2 品牌定位的基本含义 .....	66
4.2.3 品牌的客户定位策略 .....	67
4.3 特征把握，行业客户定位技巧 .....	69
4.3.1 零售行业的个体特色 定位 .....	69
4.3.2 房地产行业的服务意识 定位 .....	71
4.3.3 汽车行业的品牌塑造 定位 .....	71

<b>第 5 章 生活服务：日趋便捷的移动 大数据 .....</b>	<b>75</b>
5.1 移动 LBS 的位置服务 .....	76
5.1.1 移动 LBS 的定义和特点 .....	76
5.1.2 移动 LBS 的生活服务 应用 .....	78
5.1.3 移动 LBS 的未来发展 .....	80
5.1.4 LBS 的移动大数据营销 .....	81
5.2 移动 O2O 的购物模式 .....	84
5.2.1 移动 O2O 模式的基本 概念 .....	84
5.2.2 移动 O2O 模式的发展 优势 .....	85
5.2.3 移动 O2O 模式的商业 用途 .....	87
5.2.4 O2O 模式的移动大数据 营销 .....	89
5.3 App 的各类生活应用 .....	90
5.3.1 App 的基本概念 .....	91
5.3.2 App 的营销优势 .....	93
5.3.3 App 的移动大数据营销 .....	93
5.4 二维码的扫码服务 .....	96
5.4.1 二维码的相关应用与价值 .....	96
5.4.2 二维码的营销优势 .....	97
5.4.3 二维码的移动大数据营销 .....	98
<b>第 6 章 社交互动：全天候的移动 大数据 .....</b>	<b>103</b>
6.1 微信的多样化互动 .....	104
6.1.1 微信互动的营销条件 .....	104
6.1.2 微信互动的营销含义 .....	106

6.1.3 微信互动的营销模式 .....	107	7.2.6 数据量偏重的误区 .....	135
6.1.4 【案例】南航的微信互动 服务体验 .....	111	7.2.7 他人经验轻忽误区 .....	135
6.2 移动微博的文本互动 .....	112	7.3 三大板块，管理优化 .....	135
6.2.1 微博互动的营销含义 .....	113	7.3.1 三大硬件设备管理 .....	135
6.2.2 微博互动的营销价值 .....	113	7.3.2 两类软件管理 .....	137
6.2.3 微博互动的营销原则 .....	114	7.3.3 两项认识调整 .....	138
6.2.4 微博互动的营销策略 .....	115		
6.2.5 【案例】京东的微博 引流营销 .....	118		
6.3 移动 QQ 的大范围沟通 .....	119	第 8 章 完整记录：销售行业的 大数据攻略 .....	139
6.3.1 移动 QQ 的营销平台 .....	119	8.1 大数据时代下的销售行业 .....	140
6.3.2 QQ 互动的营销优势 .....	120	8.1.1 大数据下的智能零售 形成 .....	140
6.3.3 QQ 营销的数据应用 .....	121	8.1.2 大数据下的零售业挑战 产生 .....	141
6.3.4 QQ 营销的互动技巧 .....	122	8.1.3 大数据下的零售业商业 价值 .....	141
6.3.5 【案例】西瓜的 QQ 空间 创意营销 .....	125	8.2 锁定客户的大数据实体零售 .....	143
<b>第 7 章 风险管理：大数据安全 应用 .....</b>	<b>127</b>	8.2.1 实体零售的信息化趋势 ....	143
7.1 五大风险，日益凸显 .....	128	8.2.2 【案例】精准定位的 “上品折扣” .....	144
7.1.1 企业数据管理风险 .....	128	8.2.3 【案例】服务转型的 富士通 .....	145
7.1.2 用户隐私泄露风险 .....	129	8.2.4 【案例】构建大数据 战略的朝阳大悦城 .....	146
7.1.3 企业成本控制风险 .....	130	8.3 大数据领域的电商零售方针 .....	147
7.1.4 网络数据安全风险 .....	131	8.3.1 金麦奖的实体零售方案 探索 .....	147
7.1.5 数据人才缺乏问题 .....	131	8.3.2 【案例】阿里巴巴的大数据 营销变革 .....	149
7.2 七大误区，问题丛生 .....	132	8.3.3 百度视频的大数据建模 ....	150
7.2.1 项目噱头应用误区 .....	132	8.4 大数据的广告营销引导 .....	152
7.2.2 成果过分夸大误区 .....	133	8.4.1 广告投放的一般法则 .....	152
7.2.3 项目盲目跟风误区 .....	134		
7.2.4 软件万能认识误区 .....	134		
7.2.5 项目应用僵化误区 .....	135		

8.4.2 【案例】投放精准的“泰一指尚” .....	153
8.4.3 【案例】亚马逊的 RTB 广告模式.....	154

## 第9章 市场定位：特色餐饮的大数据策略..... 155

9.1 大数据与餐饮行业的相关知识 .....	156
9.1.1 餐饮业市场的大数据需求 .....	156
9.1.2 餐饮业发展的大数据作用 .....	158
9.1.3 餐饮业经营的大数据应用.....	159
9.1.4 餐饮业管理的大数据挑战 .....	160
9.2 餐饮行业的大数据特色营销案例 .....	162
9.2.1 【案例】活力舒化：大数据和微博双助力 .....	162
9.2.2 【案例】美团美食：LBS 与大数据双联合 .....	163
9.2.3 【案例】海底捞订餐：大数据与 App 双选择 .....	165
9.2.4 【案例】食谱：大数据与创意双营销 .....	166

## 第10章 信息累积：网络通信的大数据变革 .....

10.1 大数据与信息行业的相关知识....	168
10.1.1 信息行业转变的大数据环境 .....	168

10.1.2 信息行业发展的大数据前景 .....	170
10.1.3 信息行业营销的大数据方案 .....	170
10.1.4 移动互联网的大数据分析 .....	171
10.2 互联网企业的大数据营销..... 172	
10.2.1 【案例】PPTV 聚力：大数据智能推送 .....	172
10.2.2 【案例】大众点评：大数据智能展现 .....	173
10.2.3 【案例】世纪佳缘：大数据智能判断 .....	174
10.3 通信行业的大数据应用手段..... 175	
10.3.1 【案例】中国移动：大数据信息化战略 .....	175
10.3.2 【案例】中国联通：大数据标准化进程 .....	177
10.3.3 【案例】湖南电信：大数据综合化推进 .....	177

## 第11章 智能监控：交通能源的大数据效益 .....

11.1 大数据与交通行业的相关知识.... 180	
11.1.1 交通行业城市发展难题 .....	180
11.1.2 交通行业的大数据应用.... 181	
11.1.3 交通行业的大数据优势.... 183	
11.1.4 交通行业的大数据挑战.... 184	
11.2 交通行业的大数据营销利器..... 185	
11.2.1 【案例】数据交流，行车安全 .....	185

11.2.2 【案例】信息服务， 丰田畅通 .....	186	12.3.2 【案例】“好大夫在线” 的大数据定位 .....	203
11.2.3 【案例】数据救援， 安联智能.....	187	12.3.3 【案例】康诺云的大数据 医疗服务 .....	204
<b>11.3 能源行业的大数据开发与应用....</b>	<b>188</b>	<b>第 13 章 迅速反应：娱乐传媒的大数据 冲击 .....</b>	<b>205</b>
11.3.1 电力行业的大数据应用 ....	189	13.1 大数据时代下的娱乐传媒.....	206
11.3.2 【案例】谷歌的漂浮数据 中心.....	190	13.1.1 娱乐传媒的大数据 意义 .....	206
11.3.3 【案例】UPS 的物流数据 中心 .....	191	13.1.2 娱乐传媒的大数据 挑战 .....	208
<b>第 12 章 高效服务：医疗领域的 大数据价值 .....</b>	<b>193</b>	13.1.3 娱乐传媒的大数据 策略 .....	209
12.1 大数据时代下的医疗营销 .....	194	13.2 娱乐传媒的大数据营销应用.....	211
12.1.1 医疗领域的大数据 价值.....	194	13.2.1 娱乐传媒的大数据 趋势 .....	211
12.1.2 医疗领域的大数据 应用 .....	195	13.2.2 娱乐传媒的大数据商业 模式 .....	212
12.1.3 医疗领域的大数据 前景 .....	196	13.3 娱乐传媒的大数据营销案例.....	213
12.1.4 医疗领域的大数据 挑战 .....	197	13.3.1 【案例】新影数讯的 大数据分析 .....	213
12.2 医疗领域的营销大数据 .....	198	13.3.2 【案例】《小时代》的 大数据分析 .....	214
12.2.1 医疗领域的大数据 增长 .....	198	13.3.3 【案例】《纸牌屋》的 大数据变革 .....	215
12.2.2 医疗领域的大数据 关系 .....	199	<b>第 14 章 供需调控：生产制造的 大数据支撑 .....</b>	<b>217</b>
12.2.3 医疗领域的大数据 方案 .....	200	14.1 大数据与生产制造业的相关 知识 .....	218
12.3 医疗领域的大数据应用案例 .....	201	14.1.1 生产制造业的大数据 挖掘 .....	218
12.3.1 【案例】“南湘雅”的 临床大数据系统.....	201		

# 目录

14.1.2 生产制造业的大数据 冲击 ..... 220	15.2.2 【案例】汉庭酒店的大数据预算管理 ..... 236
14.1.3 生产制造业的大数据 应用 ..... 221	15.2.3 【案例】机场的大数据预测管理 ..... 238
14.1.4 制造业的大数据商业 智能 ..... 222	15.2.4 【案例】国药集团的大数据平台 ..... 240
14.2 生产制造业的大数据价值体现 ..... 224	<b>第 16 章 线上线下：金融行业的     大数据竞争 ..... 241</b>
14.2.1 【案例】可口可乐的 大数据昵称捕捉 ..... 225	16.1 大数据时代下的金融行业 ..... 242
14.2.2 【案例】长虹的大数据 家电战略 ..... 226	16.1.1 金融行业的大数据变革 理念 ..... 242
14.2.3 【案例】欧派电动车的 大数据服务 ..... 227	16.1.2 金融行业的大数据应用 途径 ..... 244
14.2.4 【案例】长安汽车的 大数据制造应用 ..... 228	16.1.3 金融行业的大数据趋势 ..... 245
<b>第 15 章 对内经营：企业管理的     大数据战略 ..... 229</b>	16.1.4 金融行业的大数据挑战 ..... 246
15.1 大数据与企业管理的相关知识 ..... 230	16.2 银行业的大数据营销应用案例 ..... 246
15.1.1 企业管理的大数据内部 重塑 ..... 230	16.2.1 【案例】工商银行的大数据方案 ..... 247
15.1.2 企业管理的大数据 方法 ..... 231	16.2.2 【案例】招商银行的大数据战略 ..... 248
15.1.3 企业管理的大数据 智能 ..... 232	16.2.3 【案例】花旗银行的大数据服务 ..... 249
15.1.4 企业管理的大数据应用 关键 ..... 232	16.3 金融行业其他领域的 大数据应用 ..... 250
15.1.5 企业管理的大数据 要点 ..... 233	16.3.1 【案例】纽交所的大数据 系统 ..... 250
15.2 企业管理的大数据应用 ..... 235	16.3.2 【案例】基金业的大数据 预判 ..... 251
15.2.1 【案例】智慧商贸进销存的 企业管理 ..... 235	16.3.3 【案例】保险业的大数据 风险控制 ..... 251

**第17章 广泛发展：其他行业的  
大数据应用 ..... 253**

17.1 大数据时代下的旅游行业 ..... 254

17.1.1 旅游行业的大数据发展  
趋势 ..... 25417.1.2 【案例】黄山游客大数据  
引流 ..... 256

17.2 大数据时代下的游戏行业 ..... 257

17.2.1 游戏行业的大数据关联  
指导 ..... 25817.2.2 【案例】EA 游戏体验的  
大数据改进 ..... 258

17.3 大数据时代下的房地产行业 ..... 262

17.3.1 房地产行业的大数据  
营销 ..... 26217.3.2 【案例】万科地产的  
大数据战略 ..... 264



# 第1章

## 发展揭秘：全程破解大数据



学前  
提示

大数据这一 IT 新技术的兴起是互联网发展的结果，也是各行业日趋互联网化的重要表现。

全新 IT 时代下的大数据有其独特的魅力，吸引着人们去了解和探究。本章将带领读者一步步深入了解有关大数据的发展与影响的相关知识。





## 1.1 相关了解，全面分析

在互联网迅速发展的时代背景下，社会信息类数据量剧增，由此带来的数据累积效应明显。更重要的是，它带来了数据资产方面企业竞争的加剧，各行业的互联网化和大数据趋势日益明显。顾名思义，大数据即大量、海量的数据。本节将从 5 个方面对大数据的相关知识进行了解，具体如图 1-1 所示。

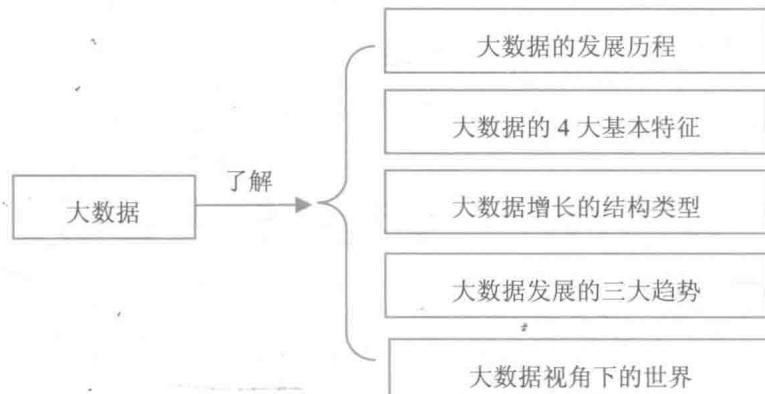


图 1-1 大数据的内容

### 1.1.1 大数据的发展历程

大数据的产生是互联网发展和信息数据化的结果。在新的时代背景下，数据不断发展和累积，如图 1-2 所示。

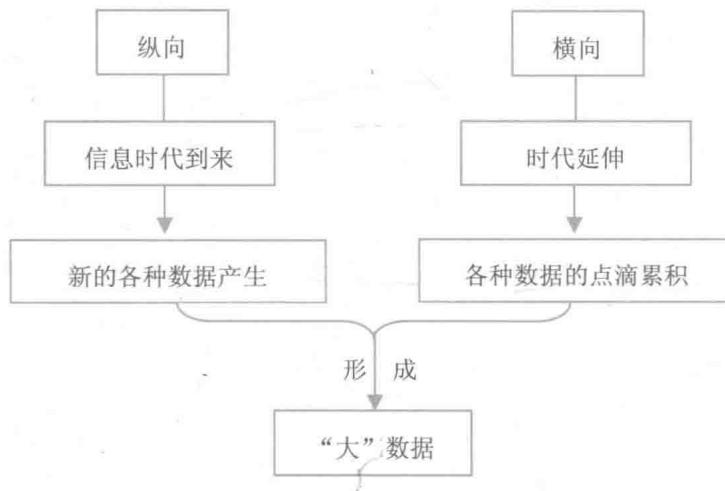


图 1-2 数据的累积

在这一发展环境下，多样性、海量的数据在人们生活中随时产生并存在，而这些数据化的信息若要加以综合利用，必须对其进行处理，大数据技术应运而生。其实关于数据的处理，早在 1890 年就已经出现，发展至今，已经形成了系统化和专门化的 IT 新技术，其发展历程如表 1-1 所示。

表 1-1 大数据技术的发展历程

年份	人物/机构	事件
1890	(美)赫尔曼·霍尔瑞斯	发明了一台用于读取数据的电动器，由此引发了全球范围内的数据处理新纪元
1961	美国国家安全局(NSA)	采用计算机自动收集、处理超量的信号情报，并对积压的模拟磁盘信息进行数字化处理
1997	(美)迈克尔·考克斯和大卫·埃尔斯	提出了“大数据问题”，认为超级计算机生成大量不能被处理和可视化的信息，超出各类存储器的承载能力。这是人类史上第一次使用“大数据”这个词
2009	印度身份识别管理局	扫描 12 亿人的指纹、照片及虹膜，分配 12 位的数字 ID 号码，并将这一数据汇集到生物识别数据库中
2009	data.gov 网站	该网站拥有超过 4.45 万的数据量集，利用网站和智能手机应用程序，实现对航班、产品召回、特定区域内失业率等信息的跟踪
2011	IBM	在智力竞赛节目中，其沃森计算机系统打败了人类挑战者，被称为“大数据计算的胜利”

随着 IT 产业的迅速发展，在新兴的 IT 供应商主导下，已有的计算机规范被重新定义，由此引起了以云计算、物联网为代表的新技术变革，大数据也是如此。

数据量的暴增是大数据产生的前提，而全球智能手机和移动设备的激增则是数据量爆炸的重要原因，如图 1-3 所示。



图 1-3 数据量爆炸的原因

由图 1-3 可以看出，数据处于迅速增长的趋势下。在这一社会基础上，在以“一切都被记录，一切都被数字化”为核心理念的数据化的发展趋势下，“大数据”应运而生，如图 1-4 所示。

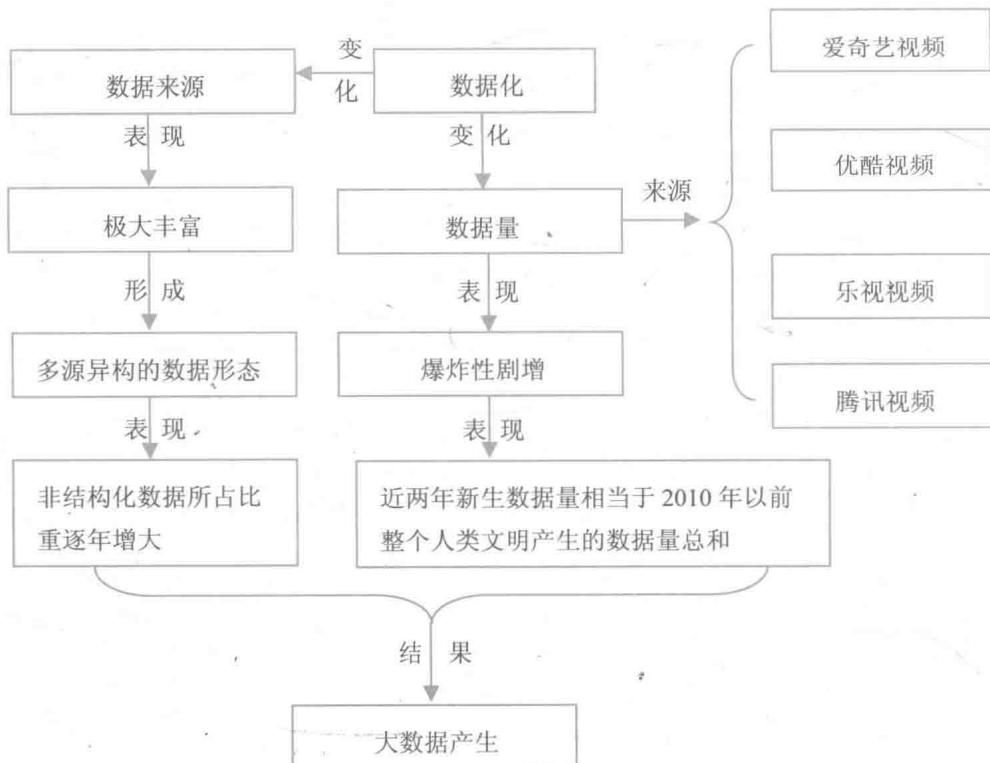


图 1-4 大数据的产生

综上所述，所谓“大数据”，即在传统数据库软件工具条件下，在一定时间内，对其内容无法进行挖掘、管理和分析处理的数据集合。

### 1.1.2 大数据的 4 大特征

大数据，总的来说，其特征可以用一个字来概括——大。从这一总特征入手，根据“大”这一修饰语在数据方面的大量、发展快的含义而言，大数据的基本特征主要表现在 4 个方面，具体内容如图 1-5 所示。



图 1-5 大数据的基本特征