



Measuring the Digital World:

Using Digital Analytics to
Drive Better Digital Experiences

衡量数字世界

使用数字分析达成更好的数字体验

◎ [美] Gary Angel 著

◎ 赵艳斌 张威 卢庆龄 李左续 译



中国工信出版集团



电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
<http://www.phei.com.cn>

衡量数字世界

使用数字分析达成更好的数字体验

[美] Gary Angel 著

赵艳斌 张 威 译
卢庆龄 李左续

电子工业出版社

Publishing House of Electronics Industry

北京 · BEIJING

Authorized translation from the English language edition, entitled *Measuring the Digital World: Using Digital Analytics to Drive Better Digital Experiences*, First Edition, 9780134195087 by Gary Angel, published by Pearson Education, Inc., publishing as FT Press, Copyright ©2016 Pearson Education.

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage retrieval system, without permission from Pearson Education, Inc.

CHINESE SIMPLIFIED language edition published by PEARSON EDUCATION ASIA LTD., and PUBLISHING HOUSE OF ELECTRONICS INDUSTRY Copyright ©2016.

本书中文简体字版专有出版权由 Pearson Education (培生教育出版集团) 授予电子工业出版社, 未经出版者预先书面许可, 不得以任何方式复制或抄袭本书的任何部分。

本书贴有 Pearson Education (培生教育出版集团) 激光防伪标签, 无标签者不得销售。
版权贸易合同登记号 图字: 01-2016-9470

未经许可, 不得以任何方式复制或抄袭本书之部分或全部内容。
版权所有, 侵权必究。

图书在版编目(CIP)数据

衡量数字世界: 使用数字分析达成更好的数字体验/(美)加里·安杰尔(Gary Angel)著; 赵艳斌等译. —北京: 电子工业出版社, 2017.9

书名原文: *Measuring the Digital World: Using Digital Analytics to Drive Better Digital Experiences*
ISBN 978-7-121-32535-9

I. ①衡… II. ①加… ②赵… III. ①数字技术—应用—企业管理—研究 IV. ①F270.7

中国版本图书馆CIP数据核字(2017)第202745号

策划编辑: 张迪 (zhangdi@phei.com.cn)

责任编辑: 张迪

印刷: 北京画中画印刷有限公司

装订: 北京画中画印刷有限公司

出版发行: 电子工业出版社

北京市海淀区万寿路173信箱 邮编 100036

开本: 720×1000 1/16 印张: 11.75 字数: 207千字

版次: 2017年9月第1版

印次: 2017年9月第1次印刷

定价: 49.00元

凡所购买电子工业出版社图书有缺损问题, 请向购买书店调换。若书店售缺, 请与本社发行部联系, 联系及邮购电话: (010) 88254888, 88258888。

质量投诉请发邮件至 zlts@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

本书咨询联系方式: (010) 88254469; zhangdi@phei.com.cn。

译者序

环顾我们生活的四周，从驱车上班到工作刷卡、从超市购物到款台结账、从读书学习到消费娱乐……我们都不可避免要触碰数字、使用数字，我们生活在数字的海洋里，我们生活在数字世界里。

当前，云计算、大数据、互联网+方兴正艾，所有这些新科技都离不开数字的支撑，当然更离不开数字分析的托举。当我有幸读到 Gary Angel 撰写的这本英文原著时，我才更真切认识到数字分析的重要性，它能够为我们生活在数字世界的我们带来很好的数字体验。这也催生我翻译这本原著的念头。

为了翻译好这本书，我查阅了许多相关资料，也向不少专家学者请教咨询，在这里向他们表示感谢。这本书是我的第二本译作，难免有些地方还够妥当，恳请各位专家和读者朋友们提出批评建议。

为了完成这本书的翻译工作，张威教授、卢庆龄教授不顾白天工作的劳累，常常深夜与我探讨里面的章节，其他译者李左续、赵洪彦、赵燕婷和魏扬澜也对这本书的翻译付出了很多心血，这里对他们表示诚挚的感谢。

衷心希望这本书能成为一把启迪钥匙，让你对数字世界测量有更进一步的理解。再次感谢所有帮助我完成编译这本书的老师和朋友们。

译者

2017年7月31日

鸣谢

在近 20 年的漫长时间里，无数人帮助和塑造我对数字测量的理解。有太多的人和太多的事情值得我回忆。确实有太多数量的人和太多的记忆。所以，我不能只提那些帮助我创建 Semphonic 数字分析实践公司的人们，Semphonic 几乎可以帮助所有数字分析想法的改进和成熟。对我来说，Semphonic 团队的每个人都具有特殊意义。他们帮助我建立一个公司和 Semphonic 实践，这些是我成年累月生活集中努力的证明，而且毫无例外地讲，他们一直都是我的好朋友。

当然，这确实是一个美妙出众的团队，在这里面，还要感谢我原来的伙伴和共同创始人乔尔·哈德利的努力。不管你是否对一场旅行感兴趣，但是陪伴你旅行的人会给旅途带来愉快。对我来说，很难想象可以有一个比乔尔·哈德利更好的合作伙伴，帮我建立公司并创建一套规章制度。

在我开启数字世界测量旅程时，必须提到陪伴我十多年的 3 名早期员工，他们是保罗·雷谷德、菲尔·克姆勒和杰西·格罗斯。他们不仅帮助我构建举世瞩目的实践项目，而且也启发了我在书中提到的思维和方法。保罗帮我塑造了许多分析性思考，尤其是分类问题、选择问题和偏差问题。最终，我很难说出我的想法起于何处，以及他的想法止于何处。对于菲尔，我感谢他提供的整体画面，他使分析远远超出数据处理的范畴，并使它专注于人们所提出的问题，以及如何解决这些问题。杰西从大学毕业就开始与我们一起工作，尽管他对数字分析的了解情况还不如我对梵文知道得多，但是他至少是一个完整的合作伙伴，我们一起探索客户声音和它们在数字测量中的作用。

虽然这些观点是我写出来的，但是有些想法通常都是他们提出来的。

当我开始涉足数字分析时，它还是一个小社区。像大多数小社区一样，它的成员也都是异类，他们都是最优秀的怪人。由于缺少凭证和等级，所以围绕数字分析研究的往往是那些愿意探索一门学科的人们，它既不提供可靠的基

础，也没有提示任何明确走向成功的方法。他们通常是最优秀的人。在这样一个社区成长，一直是一件快乐和荣幸的事情。在社区里面，我们争吵、打架和学习，并且我可以和一群真正热爱他们所做事情的人们分享观点，并努力把事情做得更好。

谢谢大家。

作者简介

目前，加里·安杰尔领导安永集团（EY）数字分析实践。EY 在 2013 年收购了加里先前的公司 Semphonic。作为董事长兼联合创始人，加里带领 Semphonic 公司从由 2 个人组成的咨询公司，发展成美国的一家领先的数字分析实践公司。鉴于加里写了一篇有影响力的博客（<http://semphonic.blogs.com/semangel>），所以他被数字分析协会评为行业最具影响力的贡献者，他也发表了大量的关于先进数字分析实践的白皮书，并经常活跃在行业活动中，他也是一名著名的演说家。在过去 20 年的时间里，他帮助创建并促进了最高水平的数字测量发展。

前 言

我们究竟该如何理解我们周围的世界呢？因为它不是很明朗的。

我们的感官将永无休止的数据送到我们的大脑。原始状态的数据是极其复杂的。我们吸收事物的形状、表面、大小、颜色、运动、气味和质感，而且这一过程似乎毫不费力。然而，即使最强大的超级计算机也无法实现实时处理。但是，对于具备适应性的人类婴儿大脑，在一些精心预布线的帮助下，毫无疑问，他们可以学习如何理解这些数据并解析它们，并且对于潮汐式的信息激增可以明智地迅速反应。

几个世纪以来，哲学家们认为无法核实我们自身之外的世界存在。我们只能从内部、从无尽的常量处理过程认识世界，我们把这些称为意识和明显的的数据流，我们认为这些数据是由我们的眼睛、耳朵、鼻子和手指产生的。物理世界中有些事情的存在，大大超出了我们的质疑能力。但是我们只能通过解释本质上的抽象数据了解它。

我们通过实施一些模式理解世界，在大自然无穷无尽的实验室中，已经有一些基础模式，留下深刻印记并深深植入我们的思想。大小、形状、颜色和运动只是人类对物质世界理解的几个核心构件。一旦意识数据漏斗进入大脑，我们便尽快进行检测，按照关键维度对它们分类，然后测量和比较所有数据。

这就是物理世界。进化可以确保我们知道如何理解它，即使我们自己都不知道我们是如何理解的。但是，我们不再把我们所有的时间都花费在物理世界。因为我们也生活在数字世界里，它有不同的规则、不同类型的数据、不同的参考帧，以及不同类型的测量。

在过去 17 年的时间里，我的工作一直是数字世界测量。为了从它呈现给我们的巨大数据流中收集数据，进行基本的分类工作十分必要。通过开发框架设备、维度和测量指标，我们可以像解析周围的物质世界那样，轻松地理解这个数字世界。这是很重要的工作，因为只有当我们理解它是如何工作的——看它如何为人们工作而不是从编程的角度看它如何工作（虽然我开始作为一名程序员），这样我

们才可以塑造和改善这个世界。

现在，我们花费大量时间、精力和金钱，努力改善数字世界。我们做这项工作的好坏可以决定公共政策（healthcare.gov）、我们钟爱的生命健康（eHarmony），以及我们知识世界状态（nytimes.com）的成功和方向。

这项工作并不容易，而且它还没有结束。我们仍然像刚出生的婴儿那样，需要学习解析数字传感器的数据。

数字世界是迷人的。而且不像物质世界，我们没有测量和理解数字领域的预布线。因为没有较大的预构建优势，所以我们解释错的概率要比正确的概率大得多。但是当我们正确解释后，我们至少知道我们做了什么。

接下来的章节将告诉你，我们目前已经掌握的关于如何测量和理解数字世界的知识。

目 录

第一章	数字意义	1
	数字挑战：我们的指标和测量缺乏意义	1
	杂货店里的无形顾客	1
	得到数字	4
第二章	两层分类	9
	创建数字测量基础	9
	两层分类概述	16
	噪声信号：两层分类的另一个视角	19
	理解两层分类	21
	两层分类示例	23
第三章	用例和访问意图	27
	洛杉矶塔可食品流动卡车：设计一个数字分类	28
	构建分类规则	32
	开发初始用例	36
	创建行为特征	39
	分类	44
	剩余部分分析	49
	定义成功	51
第四章	客户标识和分类	60
	加深对行为的理解	60
	功能主义	70
	构建分类	79
第五章	网站结构	82
	假设 2：假设 1 常常是错误的	82
	没有新问题：只是旧错误	86

	选择限制	87
	行为转化	90
	对照实验	91
	混合你自己的实验	96
第六章	态度和行为：配制具备更强大测量功能的“鸡尾酒”	99
	理解网站行为	100
	“粉饰”分类	102
	验证分类	103
	建立一个分类	105
第七章	客户声音、数字营销和成功测量	112
	数字网站成为零售点	113
	在线或通过呼叫中心采购	113
	在线播放和消费显示	114
	在线引领代际和销售转化	114
	在线或离线商业模式下的产品支持	115
	制药公司向潜在病人推销药物	115
	采样和偏差	116
	事前询问，过程中询问，事后询问	119
	测量上游旅程	119
	目标精度	121
	站外调查	123
	站点测量（过程中）	125
	稳定性	128
	何时询问：目标调查和样品的作用	128
	基准测试和跨站点比较	130
	测量下游价值（事后询问）	132
	再调查方法	133
	取样挑战	136
第八章	大数据和数字世界测量	138
	数字数据分析的真正不同之处	141

	大数据用例	146
	为什么每个人都理解错了	148
	用你所拥有的去获取	149
第九章	全方位渠道分析	151
	黄金路径的神话	152
	启示	155
	全方位渠道分析：旅程阶段	155
	为跟踪旅程搭建测量框架	157
	最近/频率/货币（RFM）	160
第十章	认知没有尽头	163
	评估框架	164
	一些谦逊时间：这里并不涵盖所有的事情	168
	期待：仍然有工作要做	169
	测量和它所蕴含的含义	170

第一章

数字意义

我们有专用的数字世界测量工具。毫无疑问，我们假设这些测量工具给我们的数据是正确的，但是事实却不是。甚至在人们还没有对数字测量有基本了解和认识之前，大多数数字分析工具就有了一套标准的网络指标，但是这些主要的分析工具，却未能充分利用这些广泛使用的数据。

数字挑战：我们的指标和测量缺乏意义

然而，最常见的数字指标几乎是无用的，它们以错误的方式测量错误的事情。在最基本的层面上，他们未能将数字世界发生的事情和我们对人们行为的理解连接起来。在这一章，你将意识到数字测量和分析面临的基本挑战，而且你将会明白，为什么共同的标准和报告不能轻易回答，这些是你可能会询问的有关数字世界的基本问题。

杂货店里的无形顾客

想象一个非常正常的杂货店，店内备有麦片、牛奶、啤酒、葡萄酒、鸡蛋、冰淇淋、罐头食品、蔬菜和水果，而且当然，通常在收银台旁有各式各样的美食。现在想象一下，你看不见顾客和他们的购物车。当顾客到达时，你会看到门自动打开；当他们离开时，你能听到收银机声响。虽然你知道他们买了什么，但其他

一切活动仍然隐秘。因此，你很难知道商店运作的情况，而且你不知道可以做什么使它运转地更好。消失的物品或品牌是消费者想要的吗？商店的储物和上货方式能方便顾客的日常生活吗？这有助于最大化顾客的购买行为吗？你给每个类型的商品分配适量的货架空间了吗？为了让一个顾客消费更多或者更忠诚，你可以做些什么事情？

多年来，这些类型的问题一直是营销专家研究、思考和从事的课题，它们在数字世界存在之前就进行了。有趣的是，他们发现可以回答这些问题，即使出于实用目的，客户看不见。同样有趣的是，他们发现，如果你不知道客户是谁，某些类型的问题更难回答，对于很多问题，数据可能帮助建议可能的答案，但是却很少提供明确的指导。

例如，假设在你的商店里，你发现高频购买的物品是牛奶、啤酒、鸡蛋和薯条。那么你可能会将所有这些东西一起放在商店前面的某个地方，它应该方便客户，使他们能够快速有效地找到所需要的东西。把顾客常买的物品放在最前面，你的超市是这种布局吗？

然而，真实的情况却几乎完全相反。这并不是因为你看不见！有两个原因解释不同超市的这种情况。消费者常常都过于玩世不恭，所以你可能确定出了第一个原因。杂货店并不是按照你的便利进行设置。它们经常把人们频繁购买的商品放在商店的后面，甚至有意识地试图将它们放得更远。如果在杂货店你从来没有过即兴购买，那这可能就是一件怪事。但是，如果像我一样，你沿着奶制品通道行走，你可能会向购物车里添加一些即时可烤的饼干，或者你在啤酒旁边扔入一袋薯片，所以这并不难理解为什么超市如此设置商品布局。通过尝试不同的店面布局和测量人们买了多少人商品（他们购物车的平均情况），店面设计师可以最大化总销售额。请注意，大多数杂货店指望你做出去哪儿购物的决定基于一些其他因素，而不是在商店内你需要多长时间找到你需要的物品。他们知道价格、选择和位置要比便利更重要。如果一家杂货店的隔壁新开了一家杂货店，它也有相同的选择和价格，那么商店可能会通过便利的布局进行竞争。但是，大多数商店将他们的布局当成是最大化他们利润的机会，而不是缩短你的时间。

杂货店没有按照便利性设计布局的第二个原因，要比追求利润最大化有趣和重要得多。杂货店有多个客户，你猜怎么着？他们都是不同的。当杂货店跟单员

开始研究人们购买什么商品时（他们仍不知道他们是谁，只知道同一票上购买了什么东西），他们发现截然不同的购物模式。在一个杂货店里，啤酒和牛奶可能是最常购买的两类物品，但是他们通常并不是一起购买。然而另一方面，薯片却常常和啤酒在一块购买。而且牛奶买家往往向购物车内添加麦片和鸡蛋。所以，一个杂货店，将常购买的物品聚集在一块可能效果并不好，或者它并不是对每个人都特别方便。

更重要的是，即使现在一个特定的布局设置运转起来很好，但是也许明天就不能适应人们的需求。当业务员只能看到每个购物者的收据，他们没有办法判断出人们的购物习惯和多样的购物模式。在他们的理解上存在着巨大的漏洞。为了绕过这个漏洞，杂货店创造了顾客忠诚度计划以换取折扣，这样他们可以了解你每一次购物之旅的实际情况。他们发现大多数人每次在杂货店的购物方式都不一样。当我们购买我们所需要的一切商品时，我们大多数人都有固定的购物探寻模式，沿着每条通道上升和下降。当我们须要遍历商店每一块区域时（或者，以我个人的购物经历，在我记忆中需要遍历两到三次），店面布局可能并不是一件大事。但是，有时我们去杂货店是因为我们刚刚喝完啤酒或牛奶，但愿不会如此。这时，我们可能停下来吃午餐或者购买特定配方成分（如面粉、袋装巧克力、香草精和鸡蛋）。这些商品属于非常不同的类型，所以如果杂货店可以让每次购物经历都完美（或者对杂货店来说完美），那么这将是一件伟大的壮举。商店也希望能够做到这样，但是当你已经走进门口，很难将这些商品摆在货架上。

我们不要忘记登记处放置的那些巧克力棒和女性杂志。很少有人为了买一块“士力架”巧克力和都市性（美国俚语，对一些杂志的称呼）去杂货店，但是我们中的许多人可能会被某些其他人的购买而驱动购买。对于那些在杂货店徘徊而无事可做的人，他们会被驱使购买吗？这就是这些“糖果”的去向（满足眼睛和胃的需求）。

当我们着手思考数字世界时，我们可以从这些杂货店跟单员那儿学到很多东西。最直的道路并不总是最好的。你的目标并不总和客户目标一致。并不是每件产品都一样，有些产品比其他产品对位置更敏感。商店没有理想的布局，因为它不只有一种类型的客户，而且客户并不总是做同样的事情。最后但也是最重要的一点，人们实际上做了什么，告诉我们很多关于他们是谁和他们为什么做这些行为的信息。

得到数字

我们对数字世界视而不见。我们很无助，因为没有办法知道我们的网站是挤满了访客，还是和商场关门时一样空空如也，我们不知道收银机是在忙碌还是保持固有沉默，我们不知道客户是年轻还是年老，我们不知道网站内容是被全神贯注地阅读还是几乎杂乱无章地跳读。我们需要“眼睛和耳朵”来帮助我们了解数字世界。某些工具备有这种功能，可以跟踪和可见的那个世界原本看不见的模式。这些数字分析工具既强大又丰富。它们可能包括成百上千份报告和似乎暴露数字行为方方面面的选项。我们太容易忘记如何依赖这些确切性工具，我们假设他们向我们展示的内容和展示方式就是数字世界存在的一切。

我们对物质世界的自然感官给我们带来很多好处。我们对所看到的一切保持一种坚定不移的信念。即使我们有身体感官，但它太容易忘记它所提供的进入世界的窗口是狭窄的。

还记得 2015 年早春，一件裙子图片像病毒一样传播散开吗？许多人认为这条裙子是黑色和蓝色，其他人认为它是白色和金色。如果你看的时间足够长，那么你可能会看到它呈现的各种颜色。如果你没有听说过这条裙子；而且你也不相信任何人都可以看到不同的呈现颜色，如果这件事情也深深触动了你，那么你可以浏览一下相关网页并仔细看一下，它一定会让你大吃一惊。

视觉错觉仅是眼睛误导我们的一个方面。尽管我们可以看到颜色（无论我们对此多么不同意），但我们却看不到热量。

为什么我们要看到热量呢？

好吧，我们为什么不能呢？

红外摄像机可以看到热量。热量只是另一种波长，在很多情况下，看到热量要比看到光有用得多（例如，在晚上打猎时）。对于这个问题，如果我们可以看到无线电波会怎么样？听力和视力似乎是完全不同的，但是它们都是我们体内不同器官使用的一组波。我们的无线电波眼睛会理解麦当娜的歌吗？可能并不会理解太多。

一个简单的事实是：我们的真实感受到我们体验它的工具限制。

这一切与数字分析有什么联系吗？我们的数字分析工具是进入数字世界的窗口。我们只认为它们可以跟踪或思考的事情是重要的。我们从服务器看到用户请求的页面，但是我们看不到需要多长时间完成页面加载。我们可以看到用户单击的链接，但是我们通常看不到用户滚动到页面的哪一部分。我们（有时）看到用户来自哪一个网站，但是我们（通常）看不到用户前往哪一个网站。这些选择深刻地影响着我们如何看待数字世界的方式，以及我们倾向于重视什么。

如果我们的工具不是很好该怎么办？如果对于真正的数字世界或者理解里面什么事情重要，而它们选择捕获的事件或者它们选择展现给我们的方式，只给了我们一个模糊的印象，那该怎么办呢？

从一开始，我就从事数字分析。甚至在一些小的方面，我亲眼目睹和帮助促成这些数字工具的进化。因为我经历了它们的发展历史，所以我知道跟踪数字世界里的事情和如何进行跟踪的决定通常是临时和肤浅的。

所建立的第一个数字分析工具是为了阅读博客。建立这些日志并不是为了理解和测量数字世界。创建它们是用来记录一个网络服务器正在做什么，这样 IT 专业人员可以跟踪许多操作问题（虽然很少使用这一功能）。这些日志记录了集中的 IT 信息，关于哪一文件内容被请求、处理它的精确时间、哪一个 IP 地址发出的请求、发送了多少内容，以及请求是否成功。

因为这些属于日志的记录字段，所以当我们第一次建立数字分析工具时，这些也都是我们使用的字段。因为我们是一个聪明的团队，所以我们从这些梗概组字段中插入了很多内容。我们通过设备请求，找到一种分组这些记录的方法（我们立即人格化访客）。通过查看请求之间的时间间隔，我们创建一个任意时间限制，可以将这些请求分组为不同的批次，并且我们标记这些访问请求的批次。然后我们可以看到在这些批次中，访问者最先看到的网页，我们称为入口页面。我们也可以看到最后处理的批次页面，我们将该页面称为退出页面。

重要的是，要意识到这些决定是多么的随意。当访问者第一次到达一个网站时，有时网站记录下访问者先前访问的网站——这个叫作推荐站点。通过保存每批记录（访问）的推荐站点，你可以了解哪些网站对你的页面生成流量。但这里有一个特点：通过定义 30 分钟作为任意时间限制对这些记录进行分组，有时我们创建的情况会有多个推荐站点；在其他情况下，访客有一个推荐网站是他所浏览