

不管我们做什么，数字魔法师都可以把我们变为1、2、3、4、5、6、7……

——《长尾理论》作者克里斯·安德森

中信出版社·CHINACITICPRESS

当我们变成一堆数字

THE NUMERATI STEPHEN BAKER

图书在版编目 (CIP) 数据

当我们变成一堆数字 / (美) 贝克著; 张新华译. —北京: 中信出版社, 2009.7

书名原文: The Numerati

ISBN 978-7-5086-1586-8

I. 当… II. ①贝… ②张… III. 信息经济学 IV. F062.5

中国版本图书馆 CIP 数据核字 (2009) 第 100819 号

The Numerati by Stephen Baker

Copyright © 2008 by Stephen Baker

Published by arrangement with Stephen Baker c/o Levine Greenberg Literary Agency, Inc.

Simplified Chinese translation copyright © 2009 by China CITIC Press

ALL RIGHTS RESERVED.

当我们变成一堆数字

DANG WOMEN BIANCHENG YIDUI SHUZI

著 者: [美] 斯蒂芬·贝克

译 者: 张新华

策划推广: 中信出版社 (China CITIC Press)

出版发行: 中信出版集团股份有限公司 (北京市朝阳区和平街十三区 35 号煤炭大厦 邮编 100013)

(CITIC Publishing Group)

承印者: 中国电影出版社印刷厂

开 本: 787mm × 1092mm 1/16

印 张: 15.5 字 数: 208 千字

版 次: 2009 年 7 月第 1 版

印 次: 2009 年 7 月第 1 次印刷

京权图字: 01-2008-3562

书 号: ISBN 978-7-5086-1586-8/F · 1661

定 价: 32.00 元

版权所有 · 侵权必究

凡购本社图书, 如有缺页、倒页、脱页, 由发行公司负责退换。

服务热线: 010-84264000

<http://www.publish.citic.com>

服务传真: 010-84264033

E-mail: sales@citicpub.com

author@citicpub.com

The Numerati

引言

想象这样的场景。

有一家咖啡馆，嘈杂、喧哗。你在那儿，我也正好在那里写作。你旁边的座位上，一位年轻女士正在笔记本电脑上忙活着。你回过头来，看着她的电脑屏幕。她在网上“冲浪”，而你在一旁观战。

数小时过去了。

她在网上看了三篇关于中国的文章，搜索了周末晚上想看的电影，欣赏了《功夫熊猫》预告片的精彩片段，点击了一条声称能与她的高中老同学建立联系的广告。坐在那儿，你记下了她的一举一动。时间在一分一秒地流逝，你对她的了解也在不断增加。那么现在，就请想想让你坐观 1.5 亿人同时在网上浏览的情形。

这听起来似乎匪夷所思，可这差不多正是戴夫·摩根所做的事情。

你是燕子，我是驼背鲸

“什么是言情片影迷？”一个夏日里天色阴郁的下午，我和摩根坐在他位于纽约的办公室里，他问道。普通网民也许很难想象，我们对网页的浏览会令广告商兴奋得两眼放光。当我们从一个网页切换到另一个，他们悄悄地跟随我们移动的轨迹，就好像追踪一群燕子或驼背鲸。

最近，摩根就在窥探那些经常点击汽车租赁广告的人群，而他们大多数人都访问过在线死亡名单。摩根说，下雨时，观察雨点怎样噼里啪啦地敲打窗户，这样做自有道理，就像“有人去世了，你因此赶去参加葬礼，也就要租一辆车”。而让摩根颇费心思、琢磨得最多的，正是言情片影迷——这一人数高居第二位的广告受众。由于某些原因，对于许多似乎被阿拉姆汽车租赁公司（Alamo Rent A Car）的网络横幅广告所吸引的人，摩根还不敢说知之甚深。

43岁的摩根是一位天性快乐的男人。他梳着一边倒的发型，好像在年幼时，被妈妈用蘸了水的梳子特意打理过并保持至今。他在美国宾夕法尼亚州西部的克利尔菲尔德长大，那是距沙蝇镇^①仅有短途车程的一个小镇。每年2月的第二天（在冬至和春分之间），小镇周围便会聚集大量从冬眠中苏醒却依然懵懵懂懂的土拨鼠。人们赶到这个小镇，兴致勃勃地探究这种小动物对自己影子的回应^②。古老的凯尔特人认为，小小土拨鼠会告诉他们春天将扑面而来，还是会放慢脚步——直到3月底才姗姗而至。

① 沙蝇镇（Punxsutawney），美国宾夕法尼亚州西部靠近匹兹堡的一个小镇，1723年最早在这里定居的印第安德拉威人，相信土拨鼠是他们勇敢的祖先。沙蝇镇是印第安人起的地名。——译者注

② 2月2日是当地的“土拨鼠节”（Groundhog Day），土拨鼠会在这一天走出自己的洞穴。如果天空多云，它们看不到自己的影子，就知道冬天很快就要过去了；如果它们看到了自己的影子，就会转身回到洞里，因为冬天还得持续6个多星期。——译者注

家乡的民间预言一直影响着摩根在人生和事业上的迁徙。

他来到纽约创立的 Tacoda 网络广告公司，就是雇用擅长统计学的专家在互联网世界追踪网民，看他们如何浪迹“网”涯，下一步又将奔向哪里。作为互联网广告的先驱，摩根在互联网迅速发展的繁荣期就成立了 24/7Real Media 公司。而在互联网泡沫破裂时，他很快就创建了 Tacoda 网络广告公司，并将公司业务天衣无缝地转接到了他看好的下一桩大买卖：帮助广告客户精确定位那些最具商业价值的网上冲浪人群，并提供他们的相关信息。

整个 Tacoda 网络广告公司的生意完全依靠数据来运作。公司要面对数千个出版物的网站——从《纽约时报》到《商业周刊》。这些刊物的网站允许 Tacoda 网络广告公司获得一点被称为 cookie^① 的计算机代码而“进入”我们的电脑，这使得 Tacoda 网络广告公司能一路跟踪我们从一个网页转换到另一个。公司密切注视着我们在网上的一举一动，对顺便得知我们的名字或其他私人信息并不在意。这种行为或许会因为涉及个人隐私而引起人们的强烈反感，可 Tacoda 网络广告公司依然能获得关于网民的大量信息。比如说，公司知道你访问了《波士顿环球报》，阅读了“丰田普锐斯”栏目的文章，然后浏览了美国在线的汽车界面。

你进入汽车市场了——这可真是个好机会！

因此，Tacoda 网络广告公司就当你浏览一则购车广告时“抓”住你。你手握鼠标那么轻轻一点，Tacoda 网络广告公司既从广告客户那里得到报酬，也在这个过程中进一步地了解了你。要知道，这家公司在网上每天都要捕获有关网民浏览行为的 200 亿条线索。

摩根的团队有时会发现网上冲浪者对某些事显示出共同的兴趣，而要

^① cookie，或称 cookies，指某些网站为了辨别用户身份而储存在用户本地终端上的数据（通常经过加密）。cookie 让网页具有记忆功能，它是美国网景公司前雇员 Lou Montulli 在 1993 年 3 月发明的。——译者注

弄明白这些行为的动机，则相当困难。一旦搞清楚了这一点，广告商就可以加入人们在网上的旅行，将网民们正好需要的广告铺洒在他们所“走”过的“路”上。有些人爱好言情片，有些人对阿拉姆汽车租赁公司的广告感兴趣，而这两个人群之间究竟有什么奇特的联系，这需要进行研究。为了抓住这些特点，摩根和他的同事就得深入地收集和分析数据。

在汽车电影院，某类言情片的上映会令汽车租赁者数量激增吗？或许这些言情片颇具异国风情？这个群体的成员有共同青睐的其他网站吗？其答案依赖于我们的计算机发出的由“1”和“0”序列组成的数据。也许统计数字显示，影迷和租车者之间显而易见的链接只是统计学上的巧合；或许摩根的团队发掘出一个更为广泛的趋向——冒险故事与旅行、浪漫的渴望与漫游癖之间有着密不可分的关系。这可能会为提高各类广告的洞察力指引方向。不管在哪种情况下，摩根都要启动数百次的测试。每个测试都能为他了解我们收集到更多相关的信息，从而提高广告定点投放的精确度。他还对曾经投放的广告内容进行跟踪分析，并进行合理的替换。而我们这些互联网用户，则是他用做实验的土拨鼠，很难改变被利用的命运。

从数据到符号

在数据的制造方面，我们可是多产的作家！我们经常使用手机、电脑和信用卡——这些日常生活的行为，使我们的数据库每天都在扩充、膨胀。就拿我来说吧，这个春天的早晨，当我坐在新泽西一家咖啡馆里写作时，维里松无线通信公司（Verizon）能轻松锁定我所在的小小咖啡馆。维萨信用卡能证实我在昨晚8点19分购买了葡萄牙红酒，或许此刻酒劲刚过，正精神焕发。以这种状态赶去观看一场院校间的篮球比赛，真是再好

不过！而具有数字视频录像功能的 TiVo^①清楚地知道我在上半场结束时就退场了。靠近任何一家银行和便民小商店，安全探视镜头都能数着分秒地抓住我的身影，更别说打开电脑上网浏览了。在那里，我们所有的举动都被记录在案，为世界上无数的互联网出版人和广告商所用。我们在前面提到的戴夫·摩根，就是这个新奇的庞大群体中的一员。

20 世纪末，某国政府曾征募了好几万名间谍。可今天，还用得着什么间谍来侦察我们吗？每一分钟，我们都在自动地更新自己发出的电子数据。

这一切皆始于计算机芯片。

那些布满数百万微型晶体管的小小芯片在 20 世纪 80 年代以前，还是一种新鲜玩意儿。但是时光荏苒，随着科学技术的发展，它们的使用成本在下降，应用价值却在不断提高。现在，几乎制造业的所有产品都受益于那些小巧时髦的芯片。它们令我们的移动电话、汽车的控制设备、数码相机和计算机功能强大。在每个外出旅行的假期，当我们打开行囊，都会发现我们的生活中又增添了更多离不开芯片的物件。这些芯片接收指令，完成它们该做的工作。它们还是一群苛刻、挑剔的记录者，细致地记下我们生活中琐碎的方方面面。单独地阅读它收录的每条信息几乎毫无意义，但将它们放到一起而组成的行为模式，就会清晰地描述我们的品位，透露我们的种种表象，揭示我们的工作规律，并勾画出我们到购物中心和超级市场的消费图。

这些如潮水般涌来的数据在全球川流不息。要是你用手机给朋友发送了一张笑脸，在你将手机放回口袋之前，你以一个微不足道的小动作发送的信息会与千百亿个其他信息一起，高速冲向光导纤维电缆，飞向人造卫星，然后再返回地面，进入新加坡某公司的服务器。生活在信息“满天飞”

① TiVo 被称为硬盘数字录像机，1997 年产于美国。它既是一种数字录像设备，又能帮助人们非常方便地筛选和录下电视上播放过的节目。这种能个性化录制电视节目的硬盘录像机短短几年内就在美国风靡起来。——译者注

的世界，就连我们呼吸的非同寻常的空气里，也充满了信息的微粒。

假如有人能汇聚和组织这些漫无边际的电子语言，我们生活的碎片将被迅速地集中，组合成一幅不断变幻、更新的人类行为的拼图。这个光明前景足以令商人们兴奋得手舞足蹈。一旦把握了能收集和分析我们数据资料的方法，他们就可以解析我们的欲望、担忧和需求，然后准确地向我们兜售我们正好渴望的东西。

但是，这一切说起来容易做起来难。要通过筛选人们在网上收发电子邮件、下载色情作品等行为的数据来推销商品，就会搅动互联网数据的海洋。根据 comScore 这家网路研究公司 2008 年的分析报告，在一个月的时间里，仅雅虎网站就收集了关于消费者的 1 100 亿个数据资料。这就是说，在访问雅虎网站广告的人群中，平均每人留下了 2 520 条消费线索。或许你以为，要将这些零敲碎打的数据按照购物者、旅游者和工薪族的类别组合起来好似举手之劳。然而实际上，进行如此清晰的分类其实是个相当艰苦的过程。

在我拜访雅虎公司研究部负责人普拉布黑克·拉格万时，他告诉我，他们发现的大多数数据都是一些数字垃圾。他把它们称为“噪音”，并说这些数据噪音很容易就在整个雅虎的计算机系统泛滥成灾。在进行数据搜索时，如果拉格万手下的一位科学家给计算机下达了不够严密的指令，公司的服务器就会一连好几天发疯似的呼呼作响。但是如果对指令作出及时纠正，就能将数据搜索的速度提高 3 万倍，将一个需要 24 小时才能完成的搜索过程缩减到 3 秒钟。拉格万认为，正确运用聪明才智的人，几乎能够从浩瀚无边的数据的海洋中收集自己想要的所有东西。虽然不太容易，但他们能够做到。

唯一能搞懂我们制造的数据并加以利用的人，是高明的数学家、计算机科学家和工程师。他们懂得如何把人们生活的细节转变成象征性的符号。

为什么一定要进行这样的转变呢？

假设你打算对自己日常的饮食进行一年的跟踪记录，而你的年龄差不多与我相仿，进入人生的第4个10年。你走进文具店，买了一大摞记事卡片。然后，你将每一餐进食的食物都写在一张空白卡片上：肉馅糕、菠菜、木薯布丁……

加油啊，别停下！

一段时间后，你就积累了一大堆还在继续增加的卡片。问题也就来了：你无法统计和分析它们！它们只是些写在白纸上的文字。当然，它们也是一些符号，每个都代表了一种事物或一个健康的观念，但是你几乎不可能对它们进行加减，或放到一张图表里说明存在的某种倾向。专家们把这些堆积在一起的文字称为“非结构化数据”（unstructured data），对计算机而言，它们只是“一团乱麻”。

要处理这些数据，最好是给它们加标签：为肉类贴上M，将绿色蔬菜标为G，把所有的碳水化合物食品放进C……以此类推。一旦将这些文字简化成了符号，你就可以把它们导入电子数据表，计算出在一个特定的星期里，你吃了几次肉、几次碳水化合物。然后，你还可以做一个图表，将你的食谱与体重的变化，或脸上粉刺数量的增减联系起来。

这个过程的关键是要找到相似之处和共同的模式，而发现它们，恰恰是我们人类的本能。这就是数万年前，我们的祖先之所以能够找到可吃的植物、能够相互交谈的原因。那时候，当许多人将注意力放在迎接生存的挑战时，另一些人正在进行更具象征性的思考。

我可以想象这样一幅画面：人类的祖先围坐在火边。有的人在为一块最大的肉而格斗，有的人在一旁忙着调情、交配。而在远处，一些有点头脑的人则在摆弄着一些石头，寻思着：“要是每个小鹅卵石（pebble）都代表一头猛犸象，那么，这个大块的石头又代表什么呢……”

后来，托拜厄斯·丹特齐克在《数字：科学的语言》（*Number: The Language of Science*）中指出，“罗马人用他们的词汇‘calcula’来表示

‘pebble’，给人类祖先的思维过程赋予了一个名称。而小鹅卵石的出现仅仅只是个开始，计算的本质从物质的计量，发展到了不断拔高的抽象推理领域。”

Numerati——数字魔法师大驾光临

科学发展的脚步跨过了一个又一个世纪。如今，我们的专家们正舒舒服服地运算着那些十亿、万亿等大得令人发笑的数字——这些连接着许多个“0”的数字对其他人来说，既难以想象，又毫无关系。

科学的继承者将我们现实中的每一天转变成符号。人类制造的数据继续暴增，计算机世界便日益强大，令这个领域的艺术大师获得了无尽的能量。他们中间的两个人在20世纪90年代后期创建了Google公司并大获成功。在我们迈步走进的21世纪里，Google成为一家巨无霸公司，它的宏伟事业几乎完全建立在数学的基础之上；它的目的，就是要帮助人们在广阔的互联网世界追逐并抓住自己需要的数据。Google公司将简单的搜索引擎打造成媒体业巨人的突破性进展，以及发现我们在网页上搜索录入的文字所体现的需求对广告商的巨大价值实难估量。于是他们寻思，怎样才能将川流不息的数据变成滚滚财源，而许多其他的人也在盘算着同样的事情。

我们制造的数据如呼啸的子弹，嗖嗖地掠过互联网世界的上空，落在生物、医学、广告、体育和政治的领地。它们将这些数据汇总——我们则被量化。

当这个壮美的情景在半个世纪前初现端倪的时候，第一批计算机的体积有垃圾箱那么大，它与我们普通人保持着可观的距离，远远地在拥有空调的屋子里隆隆地运转着。在它发展的初级阶段，人类这种生物对于它们来说过于复杂；与我们对弈时，它们是我们手下的败将。然而在某个数字

的领域，它们给我们带来了希望。

那就是，一项早期的测试涉及了消费信贷问题。

1956年，两位美国斯坦福大学的毕业生——数学家比尔·费尔和他的工程师朋友厄尔·艾萨克突发奇想，要用计算机来代替信贷员的工作。

笨重的计算机对消费信贷一无所知，不清楚申请者依靠什么来生活，当然也不懂得人们涨了工资或提出了离婚申请是怎么回事。相反，那些众多的信贷员则掌握着许多申请者的数据资料。他们通常了解信贷申请者的家庭情况，知道有多少信贷申请者在高中时代艰难拼搏，还知道其中一位男士的婚约是怎样可能由于酗酒而搞砸了（如果他什么事都以他那不争气的叔叔为榜样）。假如信贷员对所属市镇的那些家庭有自己倾向性的观点，他们掌握的资料就足以撰写出社会学的论著。但是，他们缺乏一个科学的系统来作出全面的分析。而银行家在很大程度上皆依赖于内部的信贷员。

相比之下，用计算机进行管理的方法瞄准的只是数字中的一小部分，而不用将其他大部分涉及银行结余、债务和支付历史的数字囊括其中。最重要的是，费尔和艾萨克创建了一家公司来分析那些数字的模式。他们找到了一种方法来判定每位消费者履行信贷合同的概率，对每个人都得出了一个数字。风险评估的得分证明，与那些拥有内部信任的信贷员相比，这种评分方法能更好地预测信贷风险，发现大部分获得高信用评分的借贷人在消费信贷方面都表现良好，从而使更多的人有资格获得信贷服务。毕竟，除了区别数字，计算机没有对信贷者的基本条件产生任何歧视，在银行信贷业务中提供了一个公平的机会。因此，像许多分析系统一样，它更加公正。它将信贷发放的狭窄、沮丧状况，出人意料地拉回到宽宏大量的皆大欢喜。许多被信贷员怀疑的人，成为了信誉良好的信贷者。于是，消费市场迎来了信用卡发放的扩张。

尽管如此，计算机依然清楚自己的位置。它在数字的世界里如鱼得水，它就待在那里，一些在语言、音乐和图像方面有特殊才能的人几乎还没有

注意到这个前途无量的小家伙。

然而，在接下来的几十年里，计算机成为现代社会的宠儿。它的功能日益强大，在每一毫秒里贪婪地攫取更多的“1”和“0”。它变得更便宜、更小巧，将生活在世界各地的人更紧密地连接在一起。从人类的视点（包括历史的角度）来看，计算机吸纳了全部科学技术的精华，它制造了令人震惊的工作效率。它将老式打字机扫地出门，带着帝王的威严向前挺进；它让电唱机闭嘴，令胶片摄影机停工，还插手了强大的电话业务。终于，在 20 世纪 90 年代，甚至我们这些长期以来将计算机看做异类的人，也开始将计算机从书呆子世界的地下室迎进了我们温馨的家庭和明亮的办公室。我们学会了使用这些机器，与整个世界分享我们的文字、电影和照片。

事实上，我们别无选择。

老式的行为方式在缓慢地渐行渐远，但是告别过去有一个条件：我们不得不将通过计算机呈递的一切——那些我们生活的特殊素材，都翻译成由无数个“1”和“0”组成的计算机语言。我们将丰富的数据资料 and 在世界范围内进行交流的关键，都交给了那些符号语言的魔法师。如今，正是这些数学家和计算机科学家，统领着我们生活的数字化信息时代。

我把他们称为“数字科学家”（Numerati）。

工厂生产差异，消费引导广告

夏季，纽约。一个酷热难耐的下午，戴夫·摩根坐在他斯巴达式的简朴办公室里。从那儿，可将第七大道的街景尽收眼底。办公室安装了遮阳篷以阻挡强烈的日晒，他可不想再打开荧光灯平添暑热。背着光坐在阴影里，他告诉我市场在过去的二三十年里发生了怎样的变化。

“多少年来，”他说，“商人们一直关注的是那些大的消费群体。”在他们看来，与通用汽车公司的生产线相比，我们消费者也并没有太大的区别。

比如，公司为富人和赶超时髦者制造凯迪拉克和别克，为中产阶级生产雪佛兰，为年轻气盛的飙车者打造庞蒂克，为农场主选择大卡车。除此之外，他们不需要知道得太多。因为在 20 世纪中叶，不管是生产牛仔裤还是花生酱，美国的工厂一直采取大批量制造的方式，因为小批量生产会付出更多的精力和金钱。当然，某些尖顶城堡组成的社区会吸引一些驾驶着外国车、穿着吊带式皮短裤、戴着贝雷帽招摇过市的古怪之人。但是对我们大多数人来说，我们吃的东西、穿的衣服和开的汽车，都是工厂大批量制造的产物；我们对这些产品的了解也来自于大众媒体。美国创造了这一生产和消费的模式，并在第二次世界大战之后的几十年里，将它传遍了欧洲以及亚洲和拉丁美洲的大部分国家。这是一条让机器制造的产品走进千百万消费者生活的有效途径。

“广告营销在这种产业的综合生产模式中是很简单的。” 摩根说。你根据消费者的收入、性别、居住的社区、读什么杂志、看什么电视节目，将消费者划分成五六个人口统计学意义上的人群，把你的广告发送到相关的杂志和电视节目就可以了。实际上，在产品差别不大的年代，品牌就显得至关重要。后来，这一切都发生了变化。“在 50 年之内，我们就从指令经济，走向了以消费者为导向的经济。” 摩根在黑暗的办公室里对我说道。

为什么会产生这样的变化呢？

首先，计算机得以进入工厂，它为制造业的老板带来了新的机动性。他们可以轻易地调整谷物或碳酸水的生产，创造多坚果的麦片或多柠檬的混合饮料。发出一个简单的指令，让织布机的编织模式从条纹变为格子花呢，这并不比我在计算机上撰写这一章节时将字体从 Times 变为 Papyrus 费劲多少。这种改变意味着我们的产业能够创造出成千上万个新的差异。与此同时，来自世界各地的全球性倾销的产品，实际上已经堆到了我们的家门口。

今天，我们面对的选择几乎是无穷尽的。

要想在这个拥挤不堪的市场竞争中胜出，它对我们的要求远远超过工业生产所强调的效率。而当今获胜的诀窍，就是要把购买人群所期望的风格、质地、颜色和恰当的价格等产品信息，精确地传达给每个消费者。“是消费者在主持产品的展示，”摩根说，“而不是那些进行制造和生产的企业。”

这就是说，商人们必须以区别对待的方式来面对消费者。你可以派遣大批精通心理学的人员去登门拜访他们，但在实际中，这样做“此路不通”。研究消费者的明智之道，是追踪和分析他们一路走来，却一刻也没有停止过抛洒在身后的数据。而摩根的研究大大超越了这个基本的方法。他告诉我，他的研究团队正在将实验发展到监控人们在线观看广告时大脑识别所产生的生物电波。这个测试关注 p300 脑电波^①的情况（美国海军也曾经做过类似的实验，以观察飞行员怎样在空中从敌军阵营里辨别朋友）。假如当受试者在看到一幅广告信息时，在不到 1 秒钟的时间内被测到 p300 脑电波，Tacoda 网络广告公司的研究团队就会认为，这个观看者不仅看到了这个广告，而且产生了共鸣。

接下来要做什么呢？要找出什么类型的人对什么样的广告感兴趣。就像活跃在广泛的工业领域的其他数字奇才那样，戴夫·摩根正在仔细地审视人类，搜寻未知的相关性。“我们要做什么？”他问道，“或许就是找出我们下一步的行动计划。”

数据世界：巨大无比的人类行为实验室

当我向人们介绍这本书时，他们通常的反应是：“我们就要变成数

① p300 脑电波于 1965 年被科学家发现，它是研究人的认知活动和判断人的认知能力的有效指标。在测试时，刺激物与被试者的利害关系及被试者的情绪都在 p300 电位上有所反映，而且是被试者无法自控的。——译者注

字了！”

“是的。”我回答道，“其实长期以来，我们就已经成为数字了！”想想在墨西哥装配厂数不清的流水作业线上，那些仿佛被电缆穿在一起的工人；或者在凡尔登战役^①的枪林弹雨中冲锋陷阵的千千万万个士兵；甚或那些在“感恩而死”音乐会（Grateful Dead concert）现场挤过旋转门的欣喜若狂的观众。

在管理者、指挥者和组织者的眼里，我们所有出现在这些场景中的人，也都不过是一些无名、无脸部特征的生物，就像一些可以相互置换的零件。早在工业时代，人们就被看做一目了然的数字。这就是人类社会昨天的故事。

在今天，数字科学家对我们可是有着雄心勃勃的计划。他们绝不会将我们看做单纯的数字，而是当成一座由数字和方程式构建的庞大而复杂的迷宫。数十年来，科学家利用这些数学模型，来模拟从大卡车组成的车队到原子弹的一切事物。他们依据数据之间的广泛联系，以及每个数据所代表的事实和概率来构筑模型。每个模型都须以数字来反映物质的真实：它的大小、重量，它的材质和可塑的特性，以及在气压和温度发生变化时作何反应。复杂的模型有数千或数百万个变化。在数学魔法的世界里，它们相互作用、相互配合，恰如它们在真实世界中的千变万化。

构建模型是一项艰辛的工作，有时会遭遇失败。而戏剧性变化的市场给2008年带来的动荡，就是由于错误的模型掩盖了房贷市场的复杂和风险。

尽管脚下的路磕磕绊绊，今天的数字科学家依然在奋力前行，他们已经将我们的数据融进了预测的模型，一切准备就绪。未来的10年，我们的模型将开始孵化、繁殖，不知不觉地遍及我们生活的每一个角落。我们会

^① 第一次世界大战中，德军和法军在法国凡尔登进行的战役。此役是典型的阵地战、消耗战。双方参战兵力众多、伤亡惨重。法军损失54.3万人，德军损失43.3万人，故此役有“绞肉机”、“屠宰场”之称。——译者注

被制作成工薪族、病患者、士兵、恋人、购物者和选民。尽管最终的目的，是要书写解读人类的译本——它就像我们人类本身一样复杂，一样五彩缤纷。但是在这个过程的早期阶段，许多模型仍然比较原始，让我们看起来就像是一些火柴棍做成的小人。

将所有的努力汇集到一起，我们将亲眼目睹并亲身体验人类被数字化塑造的过程。它将成为 21 世纪一项伟大的事业，不断成长、扩展到物质世界的广阔天地，就像数学家掌控新的数据流——从大气传感器的布局，到数百万安全探头的使用。一个并行不悖的世界正在形成，实验室用于发明创造，并发现由数字、向量和运算法则组成的世界。

而你和我，就站在这个世界的中央。

当数字科学家把我们转变成令人眼花缭乱的数字的组合体，他们要了解我们什么呢？

首先，他们得知道我们是谁。

例如，你是一位纽约北郊 SUV^① 的潜在购买者；或一位经常出入教堂、反对堕胎，居住在美国新墨西哥州中部大城市阿尔伯克基的民主党人；或是个准备在印度南部城市海得拉巴定居的 Java 程序设计师；或是一位酷爱爵士乐，渴望着依偎在斯德哥尔摩的壁炉边，惬意地品着意大利基安蒂红葡萄酒的射手座人士；或者是一个腰缚炸弹，迫不及待地要登上一辆公共汽车实施自杀性爆炸的恐怖分子——愿上帝帮助我们发现你！

林子大了，什么样的“鸟”都有，而无论你是谁，公司和政府都想辨别你、定位你。想想看，Google 通过帮助人们找到他们想要的网页，已匪夷所思地成长为一家价值数十亿美元的公司。在每个你可以想到的产业找对人，这样的事将会价值几何？信息就是财富！而我们在互联网抛洒的个人数据资料，描绘了数不清的路径，直接通向我们的家门口。即使你将自

① SUV 是“Sports Utility Vehicle”的简称，即运动型多功能车，俗称休闲越野车。这种四轮驱动厢式车起源于 20 世纪 80 年代的美国。——译者注