

SE.
MAN
TIC

WEB3.0

互联网的语义革命

PULL THE POWER OF THE SEMANTIC WEB
TO TRANSFORM YOUR BUSINESS

〔美〕大卫·西格尔 (David Siegel) 著
管策译

WEB

计算机和网络将理解人类的语言
数据和数据之间将自由对话
在大数据时代
我们依靠语义结网
勾勒出下一次颠覆性商业浪潮的蓝图

 科学出版社

WEB3.0

互联网的语义革命

PULL THE POWER OF THE SEMANTIC WEB
TO TRANSFORM YOUR BUSINESS

[美] 大卫·西格尔 (David Siegel) 著
管策 译

科学出版社

Pull: The Power of the Semantic Web to Transform Your Business by David Siegel

Copyright© 2012 by David Siegel

Published by arrangement with Portfolio, Penguin Books LLC through Andrew Nurnberg Associates International Limited.

本书中文简体字版由Portfolio, Penguin Books LLC通过Andrew Nurnberg Associates International Ltd. 授权在中华人民共和国境内独家出版发行。未经出版者书面许可, 不得以任何方式抄袭、复制或节录本书中的任何部分。

北京市著作权合同登记 图字 01-2012-4219

图书在版编目(CIP)数据

Web 3.0: 互联网的语义革命/(美)西格尔(Siegel, D.)著; 管策译. —北京: 科学出版社, 2013. 8

ISBN 978-7-03-034909-5

I. ①W… II. ①西… ②管… III. ①互联网络 IV. ①TP393.4

中国版本图书馆CIP数据核字(2012)第130711号

责任编辑: 徐 烁 张晓雪 / 责任校对: 宣 慧

责任印制: 张 倩 / 封面设计: 柏拉图

科学出版社出版

北京东黄城根北街16号

邮政编码: 100717

<http://www.sciencep.com>

北京通州皇家印刷厂印刷

科学出版社发行 各地新华书店经销

*

2013年8月第一版 开本: B5 (720 × 1000)

2013年8月第一次印刷 印张: 14

字数: 200 000

定价: 45.00元

(如有印装质量问题, 我社负责调换)

前 言

想想你填过的每一个表格，想想每一张附有你名字的报表。你曾填过的表格，书面的也好，在线的也罢，如果全部打印出来，可能要堆满整个房间。想象一下，你可以看见全球所有文件柜和数据库中的文件：医疗单据、处方、测验、信用报告单、收据、入学登记表、政府报表、包装单、备忘录、邮件、酒店入住单、标志牌、票据、合同、认购书以及从电子高速公路上散落下来的小块数据。如果你有机会俯瞰个人数据的分布情况，最好不要吃惊于自己的信息竟然无所不在、满坑满谷。

在地图上有你信息存在的地方放一个小点，地图会被数以千计的小点所覆盖。退后几步看，你会看到有许多点分布在全国各大城市里，甚至在菲律宾、印度、新加坡、南非、比利时、英国、加拿大、冰岛和墨西哥也有你的信息存在，而这些地方你可能一个都没去过。到网上去看看，你会发现自己的信息存放在遥远的数据库中，存放在你从没访问过的网站中，存放在你从未听说的公司中，乃或黑客的住所中。

在带领你游览不断变化的信息世界前，我首先要向你展示的“幻灯片”就是这张有关你信息的互动分布图。如今，信息生态系统的联系很松散，自组织的数据拼图并不严密，而是正由搭档、捕食者和寄生虫所完善着。对于企业和个人生活而言，信息已然失控，是时候对信息采取行动了！

警告：商业加速前进！

本书将为你描述拉取时代(pull era)：消费者按照需求将一切事情——产品、服务、信息、知识和建议——拉到自己身边。使“拉取”行为得以实现，主要靠的就是人们所说的“语义网”。语义网是将信息打包，从而让信息变得更有用，重复利用性更高的新方法。在未来10~20年时间里，语义网会将“引领-推送”(lead-push)的商业模式改造成与消费者互动的“拉取-追随”(pull-follow)模式。

我将带你一道，深入查看目前正在搭建中的语义网基础结构。我们将看到，几年之后世界将会变得大不相同。本书中的大部分讨论都将是技术性的，请做好接受新术语和新概念的准备。书中还会出现少量新行话，不过我会尽可能避免它们的出现。注意 10^{15} (quadrillion) 和 10^{18} (quintillion) 这些词。你可能会说“这不可能”，问问自己，如果还身处1990年，在万维网永久改变商业环境之前，你会对未来作出怎样的预测？与其寻求已被证明的战略，还不如请你考虑一下未来的可能性。你现在面对的是改变这个世界游戏规则的技术，请不断追求管理思维的转变，以便勇敢地闯入我们的拉取世界。

在第一部分中，我会给出“拉取信息”的基本定义和原则。接着，我会展示三个走在潮流尖端的行业，学习他们由推变拉的方法。第二部分描述了未来5~10年内改变信息使用方法最重要的原则和机制，这一过程将由消费者和商家共同参与，消费者将以需求引路，而商家则必须对需求做出回应。第三部分探讨了当我们开始拉取信息时，各个行业将如何变化，也许有些国家和地区将会改进法律制度，让病人真正享受到医疗保障，并且消灭个人所得税。从这一全新的角度出发，未来的经济将会是绩效经济。最后，我将谈到如何让你的公司创建从推到拉的路线图。

采用这些原则，既有障碍也有风险，但我认为不这么做的话，风险将会更高！早起的鸟儿有虫吃，那些依赖传统业务流程而逐渐落后的企业，将会发现饕餮大餐早就被先行者风卷残云了。

我回答的问题肯定没有由此引发的多。在读完本书之后，请访问 ThePowerofPull.com 了解更多内容，看看大家都在探讨些什么，并加入到我们的对话中来。本书中还会提及许多网站，ThePowerofPull.com 上有每一章所提及的网站的链接。

黑体字是本书的结构标记，以便读者朋友们在重温部分章节时更加方便。

目 录

Preface 前言

Part 1 拉动的权力变革

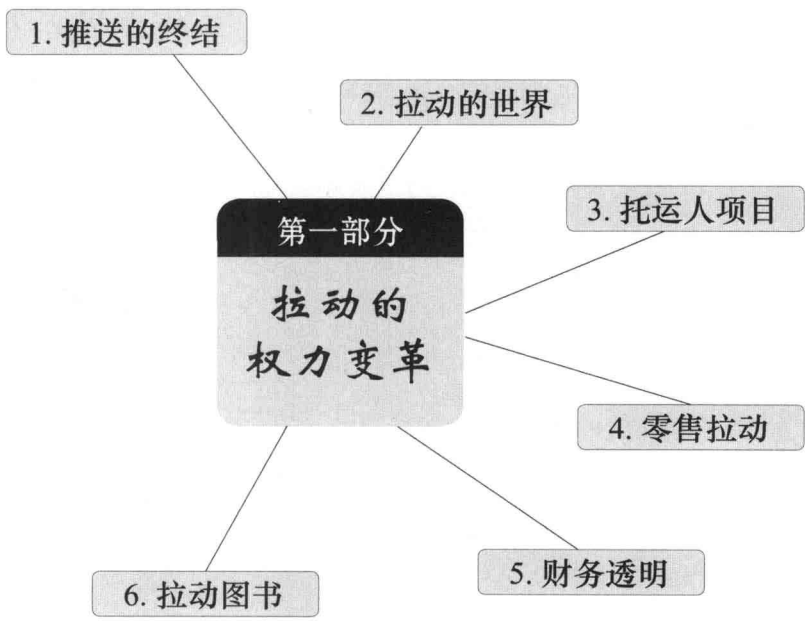
第一章 推送的终结	003
第二章 拉动的世界	006
第三章 托运人项目	017
第四章 零售拉动	023
第五章 财务透明	033
第六章 拉动图书	044

Part 2 基础

第七章 大数据迁移	061
第八章 被动商业	085
第九章 未来房屋	105
第十章 语义搜索	119
第十一章 拉营销	129
第十二章 知识危机	140

Part 3 领域

第十三章	法律拉动	159
第十四章	患者拉动	168
第十五章	公平税	182
第十六章	绩效经济	187
第十七章	一大步	201



第一章 推送的终结

拉的概念：渐进式提高，此路不通

身处变化无常的时代，学习者
承继未来。照本宣科者会发现，以
往自己具备能力生存的世界已然不
复存在。

——美国哲学家 埃里克·霍弗

固步自封带来的真正问题是：我们经营商业的方法不再奏效，但我们却基本上没有注意到这一点，要么对此感到困惑，要么加以否认。如果我们转头回顾过去十年的经商之路，到如今，再展望商业形势，同样的路数已经不能引领我们跨越下一个十年。我们现在处理问题的方法将无法满足不同客户持续增长的需求。

让我们先从几个角度来看这个问题。我们先听听下面这些管理者的心声。

首席技术官：“一年前，维护成本占到了我们IT预算的80%，我把这一比例降到了75%。如果要将大部分IT预算花在保持公司不崩溃上，我们怎么可能尝试新东西呢？我们的供应商已经将我们锁定在生命补给上了！”

分销副总裁：“我刚从欧洲的供应链会议上回来，我们被落得很远。由于使用信息的方法不同，他们店铺的生产力是我们的十倍。我估计，仅仅因为客户问起时商品无货或上架，我们就损失了12%的销售额。”

营销副总裁：“我们的市场份额正在不断萎缩。我们不再是5年前的那个权

问题，而且它还赖着不走了。”

执行副总裁：“我们的管理人员素质非常好，对任何事情的执行力都没问题。但我们没有领导者。我们失利是因为没有在精神上激励员工和客户，就是这样。”

风险投资人：“我有几家小公司位于web2.0和语义网的过渡区间中。我知道的是——我该在什么时候全力投资语义网？似乎语义网永远在地平线上，却从未到来。”

企业家：“互联网支离破碎。我们总是用不同的网站来干不同的事情。我们整天去这个网站，上那个网站，登录，填写身份信息和信用卡号，一遍又一遍搜索，试图找到我们希望找的东西。我想要建一个网站来解决一个特定的问题，但我的客户不需要另一个网站了。他们需要更少的网站和更多的功能！”

对你而言，最大的问题是什么？是什么妨碍了你的公司继续增长和赢利？我希望你能带着自己的问题来读这本书，看我是否提到了它。我们的目标并非实现渐进式提高，而是要找到新的学习曲线，追随客户进入截然不同的未来商业，重塑企业。因为我们要做的，正是由推送信息变为拉取信息。

第二章 拉动的世界

拉动概念：自动生成元数据

我们需要能表达不同数据之间关系的语言：一个将让所有数据被视为一个大数据库的“语义网”。语义网将是电子商务的革命。

——互联网之父 蒂姆·伯纳斯·李

拉取信息指的是在我们需要信息时，我们就能得到它。这一过程自1980年开始，将在2030年左右大致完成。这是人类使用信息的一种新方法，在我们接收和管理业务的方式上，拉取信息有着巨大的应用能量。

下面我来举几个例子，展示在我们开始拉取信息时，我们的生活和商业会显现出如此的天壤之别。接下来，我还会给出后续章节中将会出现的几个专业术语的定义。

百货商店

让我们从一道问答题开始。

假设你经营着一家庞大的百货店，店里有多台收银机，其中有4台专为购

物10件及以下的顾客而开设，以鼓励顾客进店小额购物；否则，如果等待结账的队伍排得太长，顾客就可能会畏惧进店消费而转身离开。于是，你找了个咨询师来考察自己的业务。咨询师说：“这么做，其实对你最优质的顾客很不公平。为购物件数少的顾客开设专用结账通道是件好事，但你并没能从他们身上赚到很多钱，其实街边的便利店就可以满足他们的需求。而且，你让那些买了满满两购物车东西的顾客情何以堪呢？难道这样的顾客不应该被优先排到队伍前面吗？这些顾客就不是你的上帝吗？我认为，应该开设3台收银台来为购物10件及以下的顾客结账，再开设1个收银台专为购物满两车及以上的顾客结账。”

现在的问题是：你是否会接受咨询师的建议呢？先想想，再对比我的答案。

我的答案是，如果考虑单个等待结账的顾客，这样的顾客可太多了。咨询师想的是让一个旧系统变得稍显人性化一些，更加以客户为中心，但新技术将完全淘汰收银机，既能帮你省钱，又能给顾客提供应享的服务。

那么，该如何让顾客购物结束时，不用自己去结账呢？这个问题的答案在本书第四章。现在，让我们来看一个权力已经发生转移的行业。

保龄球馆

19世纪末，保龄球这项运动在美国东海岸的多个州普及开来。那时，每一家保龄球馆都有自己独特的规则、保龄球、球道和球瓶。在打完一次保龄球后，摆球瓶的球童要捡起球瓶，摆放到和原来差不多的位置上。1895年，乔·萨姆(Joe Thum)和一群来自各州的保龄球手成立了美国保龄球协会，将保龄球运动标准化。美国宾士域(Brunswick)公司于1914年推出了直径为8.5英寸的橡胶保龄球，取代旧式木质保龄球。大部分保龄球馆采纳了新标准，球道也变得一样了，保龄球锦标赛变得流行起来。每年，美国有大约7千万人参与保龄球运动。

1952年，第一台AMC球瓶放置器生产下线并迅速取代人力，球瓶放置质量得到提升，成本有效降低。但从信息的角度来看，保龄球业仍旧处于黑暗时代。

我开始打保龄球是在20世纪60年代末。和朋友到保龄球馆前台登记后，我们领了鞋子和球。当被问及谁负责记分时，所有人都看向了我，于是我另外又领到了一张大纸和一支带橡皮的铅笔。我把所有人的名字写下来，轮到谁时就叫谁。在他们投完球后，我按照记分规则，用铅笔把他们的得分记在纸上。保龄球记分规则有些地方很容易让人搞混，我时不时就会犯错，不得不重新算分。最后，满满一大张纸上全都是我记下的分数。

针对这种复杂计分方式的变革发生在20世纪80年代中期，那时，保龄球馆采用了能向电脑报分的球瓶放置器，每个人投完球后都会自动显示分数。人们再也不必担心记分了，只需输入打球人的姓名就可以开始了，把球往球道里一滚，就有了分数。保龄球的复兴就得益于这一创新（当然也离不开激情的音乐和紫色灯光）。但人们仍然需要在打球前输入姓名。

保龄球业转变的最终阶段将是，你走进任何一家保龄球馆，从墙上取下一双鞋，找到一条闲置的球道，每打一轮，球馆就从你的账户上自动扣费，并将分数添加到你的在线个人数据库（第八章将会提及）中去。你所有的保龄球运动数据都保存在那里，不管你在哪儿打球，每一轮的成绩都会记录到你的在线数据库中。你可以在个人数据库中制定自己的练习进度，利用在线工具统筹自己的保龄球生涯。得益于全球元数据标准，你可以在同一天里和来自全球数百个城市的数千名保龄球手进行比赛。软件会比较你的分数，宣布谁是赢家。

这就是拉取信息。保龄球业已经经历了从顾客告诉系统分数是多少，到系统告诉顾客得分的关键性转变。当其他行业依旧陷于19世纪的信息模式时，保龄球业是少数几个已经度过这一重要转折点的行业之一。而要进一步为保龄球运动锦上添花，在本世纪中开发出促进行业发展的东西，就需要日积月累了。

高尔夫球场

现在，我们再来看看高尔夫球。

我们想要高尔夫球不那么令人沮丧，不是通过让高尔夫球车跑得更快，或加大球洞直径来实现，而是通过逆转记分机制。这需要数年时间来发展，但最终球杆、高尔夫球和球洞之间将能彼此沟通，并将信息传递到你的个人高尔夫球数据库，并在数据库中立即显示分数。你可以在和手机非常相似的电子记分卡上看到自己的分数。事实上，这个电子记分卡就是你的手机。你不用告诉显示屏你的分数是多少，显示屏会主动告知你。你不用数自己的杆数，你的手机会告诉你你挥了多少次杆，球离球洞还有多远。手机还会告诉你，你的某支球杆不在球杆袋里，并在地图上显示那根球杆的位置。当天气变化或有重要消息时，手机会提醒你。手机还能找到淹没在茂密草丛里的球。你可以查看每一次挥杆的轨迹图，可以将针对同一个球洞的击球与之前的比赛或其他选手进行比较。如果你愿意，其他远程观看的人也可以实时地获得你的数据。

从信息的角度来看，保龄球在曲线之前；高尔夫球在曲线之后。我们依旧在高尔夫球上推送信息，正如我们在生活的绝大多数领域中所做的事情一样。高尔夫球是少数几个还没应用提升信息流动技术的领域之一。我们仍然在沿用铅笔和纸质记分卡，高尔夫球运动要达到拐点至少还需要10年时间。许多行业也需要这么长时间才能到达拐点，每一个行业都有其模糊的拐点。

Facebook

有一天，我打电话到父亲家里，继母接了电话，我们聊了一会。然后父亲接过电话说道：“你祝你继母生日快乐了吗？”我的错！我完全忘记了她的生日，跟她聊天时，我啥都没说。她肯定心里有数了！

我当时想，这不公平！如果你是朋友或亲戚，而今天又是你的生日，我的Facebook主页会告诉我。如果你在我的Facebook朋友列表里，Facebook会告诉我今天是你的生日。为什么我必须去维护生日和周年列表呢？为什么不在Facebook上和我互加朋友，让在线关系反映线下关系呢？如果我将你从朋友列