

全球顶级杂志《自然》(Nature) 和《科学》(Science) 同步推荐

THE BLACK BOX SOCIETY

The Secret Algorithms
That Control Money and Information

黑箱社会

控制金钱和信息的数据法则

[美] 弗兰克·帕斯奎尔 (Frank Pasquale) 著

赵亚男 译



中信出版集团 · CHINACITICPRESS

THE BLACK BOX SOCIETY

The Secret Algorithms That Control Money and Information

黑箱社会

控制金钱和信息的数据法则

〔美〕弗兰克·帕斯夸尔（Frank Pasquale）著

赵亚男 译



图书在版编目 (CIP) 数据

黑箱社会: 控制金钱和信息的数据法则 / (美) 帕
斯奎尔著; 赵亚男译. —北京: 中信出版社, 2015.12
书名原文: The Black Box Society
ISBN 978-7-5086-5520-8

I. ① 黑… II. ① 帕… ② 赵… III. ① 经济信息-信
息安全 IV. ① F208

中国版本图书馆 CIP 数据核字 (2015) 第 227358 号

The Black Box Society by Frank Pasquale
Copyright © 2015 by the President and Fellows of Harvard College
Published by arrangement with Harvard University Press
through Bardon-Chinese Media Agency
Simplified Chinese translation edition copyright © 2015 by CITIC Press Corporation
ALL RIGHTS RESERVED
本书仅限中国大陆地区发行销售

黑箱社会: 控制金钱和信息的数据法则

著 者: [美] 弗兰克·帕斯奎尔

译 者: 赵亚男

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

(北京市朝阳区惠新东街甲 4 号富盛大厦 2 座 邮编 100029)

(CITIC Publishing Group)

承 印 者: 北京通州皇家印刷厂

开 本: 880mm×1230mm 1/32

印 张: 10 字 数: 187 千字

版 次: 2015 年 12 月第 1 版

印 次: 2015 年 12 月第 1 次印刷

京权图字: 01-2015-6404

广告经营许可证: 京朝工商广字第 8087 号

书 号: ISBN 978-7-5086-5520-8/F · 3493

定 价: 49.00 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: 010-84849555 服务传真: 010-84849000

投稿邮箱: author@citicpub.com

用数据深入阐释、揭示互联网时代的信息和金钱法则

Google、亚马逊、Facebook、苹果等硅谷和华尔街公司的秘密算法大白于天下

从这本书开始，互联网公司也许再也没有商业秘密可言

致雷 (Ray)

谨以此书纪念我的父母

醒时，我们共处一个世界，睡去，人们进入各自的世界。

——赫拉克利特



百万丛自然的篝火被点燃，熊熊燃烧。

却骤然熄灭她最迷人的，她最心爱的，她最清晰为自我的火花。

人啊，他的火痕，他心灵上的印记，消逝得多么迅速！

两者都在一种不可探测中，一切都在一个巨大的黑暗中淹没。

——杰拉尔德·曼利·霍普金斯

曾经在 IT 历史上显赫一时的美国太阳微系统公司 (Sun Microsystems) 似乎总有着超人的先知先觉能力，在公司成立之初的 1982 年，就接受了约翰·盖奇 (John Gage) 的建议，将公司的口号定为“网络就是计算机”。虽然这句话在目前看来是这么的理所当然，但在互联网还没有走入大众生活，甚至连“信息高速公路”的概念都还没有提出的时候说出这句话，其实是很令人匪夷所思的。而现在，该理念被认为与云计算思想不谋而合，但这已经是 20 年之后的事儿了。无独有偶，1999 年，面对正如火如荼发展的互联网技术，时任该公司 CEO 的斯科特·麦克尼利 (Scott McNealy) 在一次发布会上对台下众多媒体记者和分析师说道：“你的隐私只剩零了，想开点吧。”他认为互联网的分享将彻底“杀死”隐私。自此，关于互联网时代隐私的争议不绝于耳，而 2013 年爱德华·斯诺登披露的“棱镜门”事件更是将公

众对隐私问题的关注推到了一个新的高度。有不少人呼吁人们要学会适应这个越来越透明的社会，甚至有人提供一些技巧来拒绝某些互联网服务。其实，这样的认识存在着一个严重的问题，就是将用户个人信息的收集过程等同于对隐私的侵犯。

我也曾对此类话题充满了兴趣。我认为社会已经发展到这一步——人类目前并没有选择信息的自由，相反，通过放弃使用数字科技来实现隐私保护是不可取的。现在的智能手机平均配有 7 个传感器，只要人们还在使用互联网，相关数据就会被采集，而且会越来越普遍，几乎可以说是无所不在，无所不包。纵观人类历史，每当一个新技术来临，总有一部分人决定不参与其中，但残酷的现实告诉我们，最有可能享受新技术带来的利益的是那些总是追随技术的忠实客户。在这个新的时代，人们正在发明新的隐私保护方式。例如，在电子交易中，网上支付借助风控大脑的帮助会变得越来越安全，而且是用得越多的用户越安全，因为支付系统已经找到了一种更接地气的创意——通过场景来“认人”。

最早关注弗兰克·帕斯奎尔（Frank Pasquale）的《黑箱社会》这本书，是源于今年开年之处《自然》（*Nature*）和《科学》（*Science*）几乎同时对该书推荐。究竟是怎样的一本好书呢？居然得到了这两本被学术界称为最高大上刊物的顶级杂志的青睐，我充满了好奇！书名中的“black box”，让人联想到两个概念：飞机中的黑匣子和系统工程中的黑箱模型。看完这本书，读者会发现，这个词的确是一语双关的。一方面，物联网、大数据

和无处不在的传感器网络记录着我们的世界和世界中的人群，就像我们人人都拥有一个自己的黑匣子，随时可以查阅其中的信息而对未来采取更好的对策。人类行为的大规模数据集有可能从根本上改变我们对抗疾病、设计城市或进行科学研究的方式，有人甚至将这种大规模数据集的使用行为与显微镜的发明相提并论。另一方面，我们个人却不知道有多少信息被记录了，这些信息会传播到什么地方，也不知道哪些人会使用它，使用这些信息的目的何在，更无从知道这些信息的泄露会产生怎样的后果，依据这些信息所做的判断是否准确，是否存在偏见或破坏性？这对目前的人类社会来说，就是一个神秘的黑箱运作机制！我们正处于一种迫切希望采集更多的数据，又害怕数据被采集太多的纠结之中，本书主要针对这种纠结进行探讨，提供了不少新思维，阅读该书可能推进我们对于隐私争议的理解和认识上的更新。

《黑箱社会》这本书最可取之处是脱离了以往那些空谈隐私保护的话题，而是时不时给出了问题的症结所在和可能的解决方案。早期的隐私专家只是强调信息威力的力量失衡对社会造成的后果，那些所谓的经典补救措施也就是要求个人知情同意。问题是，只要告诉用户信息收集者会如何使用这些数据，只要他们同意，隐私就得到保护了吗？帕斯奎尔认为这远远不够。知情同意面对目前的社会已经退化为一个呆板的闹剧，因为我们大多数人既没有时间也没有足够的专业知识来解读这个潘多拉魔盒。虽然大数据能够从大量多维数据和组合中获得新的洞见，但在如此

背景下也只有富人能承担数据使用的费用，总体上只会加剧系统本身的缺点。因此，帕斯奎尔主张隐私保护制度的根本转变，放弃知情同意这种形式上的做法，去严格监管实际使用这些数据的企业和政府机构。难怪《科学》杂志在书评中称之为“史诗般的变化”。

对待信息的角力，帕斯奎尔在书中强调了三个关键词：信誉、搜索和金融。信誉是被他人认识的结果，搜索为了认识他人，金融则是帮助人们做出经济决策。各公司都在为挖掘潜在的客户详细信息费尽心思，但在面对监管时，却极尽手段隐藏它们的统计数据和操作规程，同时阻止用户对由此生成的数字档案进行任何操控，这其实就是在制造黑箱。试想，如果这些东西是隐藏的，那么错误的数据、无效假设和有缺陷的模型就不能得到纠正，这种黑箱最终可能会危及我们所有人。我们如今的生活已经受到了信息的影响，制定一个紧跟时代的信息管理框架已迫在眉睫，其中政府公共部门应发挥什么作用呢？也许，我们需要一个公共机构来利用全面的数据提供信用评分，用一个开放的、可解释的算法决策工具来代替黑箱系统，而不是靠众多私营企业自己搜集的数据和评定代码。真正将黑箱社会变成显式社会，确保最重要公司的关键决策是公平的，一视同仁的，并能接受批评。

赵 斌

复旦大学生态学教授、博士生导师

中国的互联网公司在竞相作用于信息发布和在线交易的未来。阿里巴巴抵制了腾讯公司的某些短信业务，而在快的打车与滴滴打车（腾讯投资）合并之前，腾讯公司曾禁止快的打车利用微信平台来发放叫车服务的优惠券。它还禁止美国出租车公司 Uber 出现在微信上面。

我本人也不是很认同 Uber 公司的经营理念。在读过张夏准的作品（尤其是《富国的糖衣》）之后，我也认为每个主权国家都应当促进本国科技在搜索、金融和信誉等领域中的发展。只有科技公司、数据经纪商和银行在全球呈现多样化趋势，各国才能够维持响应其国民需求的政治秩序。

任何一个领域中的主导性平台都需要得到更好的监管，并且要对用户透明，同时尽可能公平地对待其合作商户。微信拥有约 5.5 亿活跃用户，在日常生活中扮演着重要的角色。用户可以通

过微信菜单获取优惠券、医疗、娱乐、新闻、美食、交通等诸多信息。这是一个万能的平台，可以实现很多功能，例如查看航班、支付水电费等等。随着大型公司逐渐渗透到我们生活中的方方面面，它们在选择商业伙伴和利用编辑数据时应当做到公开透明。

“公平竞争”这个概念在商界中似乎不太现实。大家都在混乱无序地争夺市场份额、用户和收入。然而，企业越是用黑箱掩盖自己的所作所为——避开公民或监管机构的审查——它们就会有更多的机会来要挟我们。在 2008 年，百度为了达到营销目的，威胁其他企业购买百度广告，否则就降低其搜索排名，我听说这件事情的时候非常困惑。这种商业模式将导致的最终结果可能是，那些真正制造产品和服务的企业所获得的利润会逐渐地被各大互联网公司盘剥殆尽。

排名、评级、查询和融资公司不应该在经济活动中占据主导地位。它们是很重要，但不是最重要的。这些公司的特点是，规模越大，收益越多。当它们在 GDP 中占据的份额过大时，不平等现象就会加剧。信誉、搜索和金融领域中的公司巨头（分别在本书第二、第三、第四章有所论述）夸大了其自身功能的重要性。它们充其量也就是实体经济的仆役，可却妄图反客为主。

本书的一大主题是社会利益和责任的合理分配。美国从中国的经济腾飞和绿色能源部署方面学到了很多，而中国的经济学家和监管机构也可以从美国的成败中有所收获。长久以来，GDP 的增长只能依靠“投资”。当前中国国内太多的信托和理财产品都

属于不可持续的发展项目，其资产负债表极不透明。很多投资者对其投资项目并没有真正了解。随着股市的大起大落，政府面临着一个艰难的选择：要么放任自流，甘冒经济失控的风险；要么用现实资源来维系黑箱金融。第一个选择会使穷人受损；而第二个选择会使富人更富，有损公平。两者都无法实现经济的可持续发展。

真正能够促进社会和谐和繁荣的政治经济必须是实体经济，而非简单的量化经济。它必须能够为公民带来医保、教育、艺术、公园、新鲜空气和清洁用水，以及休闲项目。单纯用数值来衡量收入或生产力的做法只是以偏概全。事实上，一些外界观察家就担心，这些数值可能会受当局操控并被算作“生产力”，这是社会弊病的演变态势。

针对这一忧虑，社会科学和人文科学所描绘的世界（及其可能的未来）图景可以为我们指点迷津。这两门科学试图以事件为背景进行解释、梳理、分辩和争论，目的不是像模拟原子和分子原理那样来塑造人们的认知，也不是用数值去勾勒一幅虚假的繁荣景象。这里的人文科学不是自然科学，而是知识与洞察力相结合的有效产物。

太多的行为学家和人工智能研究人员只注重分析（并最终模仿或操纵）人们的表象行为，而很少或根本不去关注内部的思维过程。他们把大脑当作一只黑箱来处理，这是病态的社会科学和非人道社会的表现。我们必须行动起来。如果只是被动地去促进生产力的最大化，而不去追问前景或愿景，即使最大限度地提高

了 GDP，也会迷失在未来。

20 世纪 70 年代初期，查尔斯·泰勒等哲学家指出了人工智能和行为主义在目标与方式上存在的弊端，帮助人们改变了以往的错误认知。但随着“大数据”的风靡，这两种思维方式再度被吹捧为了解和控制社会的途径，而对于数据的源头（或思维模式），人们却不予关注。金融公司可能会做出诸如此类的辩解，“我们之所以对有过违约记录的用户收取更高的利息，是因为以往的数据显示，这些人常常会再次违约”，但在这一过程中，违约者曾经未能及时还款的原因却是未知的，可能他拒绝支付，也可能是因为家里出了紧急事故。

然而，如果是因为对以往违约者收取过高利息的做法或是歧视性的贷款行为而导致了再次违约呢？如果每一项决策都加大了未来违约的可能性呢？那么大数据所承诺的“社会科学”就会演变成对身背污点者进行压迫的工具，而受压迫者却丝毫没有还击之力。用以做出判断的算法程序变得越来越不客观，它们将主观偏见伪造成看似中立的分数，从而逐渐成为分配机遇和施加惩罚的仲裁者。由于技术系统将人看作数据点的集合，所以受此影响的人们将不再被当作个体来对待和理解。

一个人不应该仅仅因为生了病就遭到公司解雇，或无法获得贷款或保险（在这个数据驱动的时代，更有可能的原因是，一些秘密算法所加权的健康指标认为该人“可能生产力低下”或“不稳定”）。为了在保密性和透明性之间取得更好的平衡，本书为相

关法律的修订提供了许多建议。本书还是一部探讨社会理论的著述，揭露了金钱、权力和数据交易的基本模式。书中指出有些人在肆无忌惮地监控数百万人，而其他人连自己受到了谁的审查都不知道。监控者和被监控者的划分往往源自旧有的阶级划分，这也在意料之中。

虽然经济变化日新月异，但权力和财富的既有层级却比以前更加固化、稳定、坚若磐石。我从这一问题追溯至两种相互强化的趋势：数据的金融化和金融的数据化。当货币反映的仅仅是一个精心设计的算法游戏，以衡量人们的购买力（而不是现实的生产力）时，其价值系统便已坍塌，货币分配也失去了合理性。但在黑箱的掩盖下，权力可以被神秘地解读成为知识，运气成为技能，偶然性成为必然性，垄断成为竞争，强制成为同意。

比较无知学（即对知识的抑制或破坏）是本书的一个中心议题，这是一个很难给予系统性解释的话题。但我已经尝试论证了一些非常重要的症结点，指出了它们之间存在的广义关联，并说明了哪些法律需要修改，以便帮助我们真正了解主流公司的逐利目标。虽然黑箱有些骇人听闻，而且似乎坚不可摧，但并非不可战胜。人们很可能迎来一个可被了解的社会，但首先我们必须看清这个社会已经在多大程度上变成了一只黑箱。

推荐序 // XI
中文出版序 // XV

第一章

我们需要知情 // 001

保密优势 // 006
无处不在的大数据 // 008
保密性和复杂性 // 010
警惕软件 // 013
单向镜 // 014
企业保密和保密企业 // 016
回首 // 017
本书结构 // 021
可以自我破解的预言 // 024

第二章

数据失控时代的网络信誉 // 027

数据（和数据之外）// 032

失控的档案 // 044

工作中的大数据 // 049

种族偏见的幽灵 // 055

监控国家的诞生 // 061

“国家”与“市场”的界限正在消失 // 069

到处是威胁，到处是危险，到处是信息？ // 071

持久的不透明 // 073

没有出口 // 074

未来大曝光 // 079

第三章

隐藏在搜索中的逻辑学 // 083

搜索和透明度 // 088

搜索、信任和竞争 // 115

搜索和薪酬 // 120

搜索和控制 // 127

内容、传播渠道和搜索：新兴竞合 // 131

迈向数字新政 // 138

第四章

金融算法：皇帝的新代码 // 143

预警信号 // 148

机器梦 // 151

次贷的修饰部门 // 153

评级机构的失职 // 154

公司封杀异议 // 158