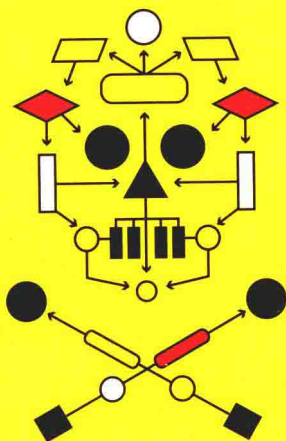


WEAPONS OF MATH DESTRUCTION

How Big Data Increases Inequality
and Threatens Democracy

CATHY O'NEIL

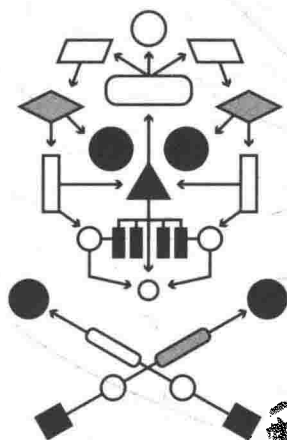


算法霸权

数学杀伤性武器的威胁

[美] 凯西·奥尼尔 著 / 马青玲 译

中信出版集团



算法霸权

数学杀伤性武器的威胁

[美] 凯西·奥尼尔 著 / 马青玲 译

中信出版集团 · 北京

图书在版编目 (CIP) 数据

算法霸权 / (美) 凯西·奥尼尔著; 马青玲译. --

北京: 中信出版社, 2018.9

书名原文: Weapons of Math Destruction: How Big

Data Increases Inequality and Threatens Democracy

ISBN 978-7-5086-9206-7

I. ①算… II. ①凯… ②马… III. ①数据处理

IV. ① TP274

中国版本图书馆 CIP 数据核字 (2018) 第 148250 号

Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy

Copyright © 2016 by Cathy O'Neil

Simplified Chinese translation copyright © 2018 by CITIC Press Corporation

ALL RIGHTS RESERVED

本书仅限中国大陆地区发行销售

算法霸权

著 者: [美] 凯西·奥尼尔

译 者: 马青玲

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)

承 印 者: 北京楠萍印刷有限公司

开 本: 880mm × 1230mm 1/32

印 张: 9 字 数: 220 千字

版 次: 2018 年 9 月第 1 版

印 次: 2018 年 9 月第 1 次印刷

京权图字: 01-2018-4671

广告经营许可证: 京朝工商广字第 8087 号

书 号: ISBN 978-7-5086-9206-7

定 价: 69.00 元

版权所有·侵权必究

如有印刷、装订问题, 本公司负责调换。

服务热线: 400-600-8099

投稿邮箱: author@citicpub.com

这本书献给数字难民

本书所获赞誉

见解深刻，发人深省！

——《纽约书评》

《算法霸权》对于正被广泛误用于我们生活的方方面面的数学是一种必要的批判。

——《波士顿环球报》

这本书具有启发性……奥尼尔证明，算法依赖使我们走向极端。

——美国《大西洋》杂志

《算法霸权》是一声响亮且直白的战斗号角。这本书承认，模型不会消逝：模型作为一种定位弱势群体的工具其发挥的作用非常令人震撼，但是如果用它来惩罚和剥夺公民权，那就是一个噩梦了。凯西·奥尼尔这本书之所以重要就是因为她相信数据科学。想要知道我们为什么一定要审查身边的各种模型，对这些模型提出更高

的要求，就来读读这本重要的书吧。

——《弟弟》(*Little Brother*) 作者、美国知名综合类博客

Boing Boing 合作编辑 科利·多克托罗

许多算法服务于对权力和偏见的平等操控。如果你不想被这些算法操控，就去读凯西·奥尼尔的《算法霸权》，解构如今在市场上横行霸道的大数据模型。

——《任何速度都是不安全的》(*Unsafe at Any Speed*) 作者

拉尔夫·纳德

下次再听到有人夸夸其谈大数据的神奇时，你就给他看《算法霸权》。他一定会受益终身的。

——美国文化新闻网站 fusion 资深编辑 费里克斯·萨尔蒙

帮人们找工作、找伴侣，预测模型潜移默化地塑造和控制着我们的命运。凯西·奥尼尔带我们踏上了一场愤怒与惊奇之旅，这本书的语言简单得就像是在对话，但它是重要的对话。我们对技术要小心谨慎。

——《当收入只够填饱肚子》(*Hand to Mouth: Living in Bootstrap*

America) 作者 琳达·提拉多

这本书不可或缺……尽管“数学杀伤性武器”这个话题在技术上

很复杂，但是这本书能够清晰地引导读者理解这些模型系统……奥尼尔的书很优秀，展示了大数据和一个如此仰赖算法的世界的伦理道德风险……以及大数据是如何为自己、为那些对大数据是如何重新塑造我们这个世界感兴趣的人带去利益的，阅读《算法霸权》是十分必要的。

——《加拿大国家邮报》

如果你曾经也怀疑对大数据深信不疑可能存在危害，但又缺乏数学知识弄明白问题究竟出在哪儿，那么，这本书就是为你而来。

——《沙龙》网络杂志

奥尼尔是写这本书的理想人选。她是一个数学学者，后来转型成为华尔街金融工程师，又变身为数据科学家，曾参与过“占领华尔街”运动，最近还创办了算法审计公司。她是强烈呼吁限制算法影响我们的生活方式的人士之一……虽然《算法霸权》这本书里的硬知识和统计数据俯拾皆是，但这本书易于理解，甚至饶有趣味。奥尼尔的写作直接、易读，我一下午就读完了。

——《科学美国人》杂志

凯西·奥尼尔，一位转型成为数据科学家的数学理论家，传递了一个简单但重要的信息：统计模型无所不在，它们对我们日常生活许多方面的影响正日益增加……《算法霸权》提醒我们，我们生

活中的一些领域已经被无形的算法掌控了。随着大数据巨头的继续扩张，来自奥尼尔的警示和提醒是受欢迎且必要的。

——《美国瞭望》(*American Prospect*)

奥尼尔自认数学迷，她的生动描述写出了大数据对我们生活的影响。

——《旧金山纪事报》

通过追踪算法在每个阶段是如何塑造人们的生活的，奥尼尔证实，我们的“机器人霸主”使用数据来区分群体，制造不公平以及干预民主选择。如果你需要和数据打交道，或者哪怕你只是大量在线数据的普通创造者之一，这都是一本必读书。

——美国知名科技博客媒体 Ars Technica

清醒，惊人，有价值……（奥尼尔）的写作清晰而准确，因为她将自己的论点瞄准了非专业读者，这使得这部分读者也能够快速阅读这本有关算法的书。生活会受到大数据影响的人都可以读读《算法霸权》这本书——换句话说：这本书是人们用得上的。

——《亚斯本时报》

《算法霸权》通过现实生活中的案例以及生动的讲述，揭示了政府和大企业利用隐秘的算法和复杂的数学模型削弱公平、增加私权

的事实。用明确打败隐秘，用明晓战胜困惑，这本书的出现能帮助我们及时改变航向。

——《人的舞台》(*The People's Platform*) 作者

阿斯特拉·泰勒

凯西·奥尼尔的陈述非常好，她凭着自己的数学特长以及对社会正义的热情找出了大数据的漏洞。她提出了一个令人信服的观点，即数学模型被当作武器压榨社会边缘群体，扩大社会不公正。她的分析极好，行文引人入胜，同时，她的很多发现也可能会引起人们的不安。

——数据与社会机构创始人、《复杂》(*It's Complicated*) 作者

达纳·博伊德

我虽然是一个数学家，却直到读了这本书才知道大数据武器是如何发挥作用的。尽管觉得恐惧，但这仍是一次相当愉快的阅读体验：凯西·奥尼尔眼中的算法世界既有黑色幽默，又充满愤怒情绪，就像《奇爱博士》或者《二十二条军规》。这本书能开阅读者眼界，引起读者忧思，非常重要。

——康奈尔大学教授，《X的奇幻之旅》作者

史蒂夫·斯托加茨

易读，生动……简洁而有说服力……《算法霸权》是我们这个时

代的“屠宰场”……所有数据科学家以及所有相信数学模型可以取代人类判断的企业决策者都应该阅读这本书。

——《数据和社会：点》(*Data and Society: Points*) 作者

马克·凡·霍勒贝克

前言

小时候，我常常盯着车窗外的车流，研究每辆车的车牌号。我会把每个车牌号分解成素数，如： $45 = 3 \times 3 \times 5$ 。这叫作因式分解，是我最喜欢的消遣活动。我这个小数学迷对素数特别感兴趣。

我对数学的爱好逐渐发展成热爱。14岁时我参加了一次数学夏令营，带回来一个心爱的魔方。数学使我摆脱了现实世界的混乱。经过数学家们的一步步证明推导，数学不断发展，其覆盖的知识领域不断扩大。我也加入了数学领域，在大学时期主修数学，后来取得数学的博士学位。我的论文方向是代数数论，这根源于是我从小就喜欢的因式分解。最后，我成为巴纳德学院的终身教授，该学院的数学系是与哥伦比亚大学联合创办的。

后来，我做了一个重大的决定：从大学离职，到顶尖对冲基金德劭集团（D.E.Shaw）担任金融工程师。我离开学术界进入金融领域，把抽象的数学理论应用到金融分析的实践中。我们所做的数据分析为一个又一个账户实现了总量达到数万亿美元的变现。

起初，在新的研究室研究全球经济让我感到既兴奋又震撼。但就在我在那儿工作了一年多的时候，2008 年秋，全球金融危机爆发了。

显然，金融危机使得我曾经的庇护所——数学不仅卷入了这个世界性的问题，还助推了其中许多问题的发生。房地产危机，大型金融机构倒闭，失业率上升，在幕后运用着神奇公式的数学家们成为这些灾难的帮凶。而且，由于数学的功能特别强大（这是我热爱数学的原因之一），一旦其与科技相结合，其所造成的混乱和不幸也会成倍增长，它使得一个有着巨大缺陷的系统加速运转，进一步扩大规模，这些都是我原来不曾意识到的。

要是我们当时头脑清醒的话，就会后退一步思考，数学是怎么被我们误用的？我们该如何防止未来发生同样的灾祸？但是，金融危机发生以后，新的数学技术变得比以往更热门，其应用甚至延伸到更多的领域，每时每刻都在搅动着海量数据，其中大多数数据都是由社交媒体或者电子商务网站从使用者那里搜刮而来的。而且，数学逐渐不再关注全球金融市场动态，而是开始关注我们人类本身。数学家和统计学家一直在研究我们的欲望、行动和消费能力，一直在预测我们的信用，并用结果来评估我们作为学生、职员、情人的表现以及是否有变成罪犯的潜力。

这也就是我们所说的大数据经济，其收益前景非常可观。一个电脑程序可以在 1~2 秒内快速扫描成千上万份简历或是贷款申请，然后将结果整理成清晰的列表，让最有潜力的申请者位居前

列。这不仅节约时间，而且公平客观。毕竟，电脑程序不像人类带有个人偏见，它只是一台处理数字的无情机器。到2010年左右，数学已深刻地介入人类事务，公众对数学这一工具的出现表示出了极大的热情。

然而，我看到的是危机。数学应用助推数字经济，但这些应用的建立是基于不可靠的人类所做的选择。有些选择无疑是出于好意，但也有许多模型把人类的偏见、误解和偏爱编入了软件系统，而这些系统正日益在更大程度上操控着我们的生活。这些数学模型像上帝一样隐晦不明，只有该领域的最高级别的牧师，即那些数学家和计算机科学家才明白模型是如何运作的。人们对模型得出的结论毫无争议，从不上诉，即使结论是错误的或是有害的。而且，模型得出的结论往往会惩罚社会中的穷人和其他受压迫的人，而富人却因此更加富有。

我为这些有害模型提出了一个名称：“数学杀伤性武器”（Weapons of Math Destruction，简写成WMD）。接下来，我将用一个例子向你们阐明这种模型的破坏性。

这个案例中的模型和很多其他的案例一样，其出发点是好的。2007年，华盛顿特区新上任的市长艾德里安·芬提下定决心对本市教学质量不佳的学校进行改革。当时，几乎每两个中学生中就有一个九年级学生是勉强毕业，只有8%的八年级学生在数学上的表现达标。为此，市长芬提设立了一个新的职位——华盛顿市教育总督，并聘用知名教育改革者李阳熙担任该职务。

当时流行的理论是：学生学得不够好是因为老师教得不好。所以，在 2009 年，教育总督李阳熙落实了一项旨在开除教学表现差的教师的计划。这符合当时全美教学质量差的地区所进行的改革的一种趋势，而且从系统工程学的角度看，这种想法非常有意义：评估教师。开除最差的教师，把最好的老师调到需求最紧迫的地方发挥他们的才干。用数据专家的话来说，就是“优化”学校的教师系统，尽可能保证给孩子们提供好的教育。除了那些“差”教师，谁会反对这项提议？教育总督李阳熙开发了一个叫作 IMPACT 的教师评估工具，至 2009~2010 学年末，华盛顿特区开除了评估结果垫底的 2% 的教师。第二学年末，又开除了 5%，也就是 206 名教师。

华盛顿特区一所公立中学的五年级教师萨拉·韦索基似乎没有任何理由为此担心。她在麦克法兰中学仅任教了两年就得到了校长和学生家长的一致好评。校长表扬她对学生们的教育极负责任，学生家长纷纷称她为“接触过的老师中最好的一个”。

但是在 2010~2011 学年末，韦索基的 IMPACT 评分很低。她的问题出自一个叫作增值模型的新评分系统，该系统用于评估数学教学和语言技能教学的效果。该算法给出的评分权重占她最终评分的一半，超过了学校领导和社区的评价。华盛顿特区别无选择，只好开除了她，以及另外 IMPACT 得分在最低限度之下的 205 名教师。

这看起来不完全像是一种政治迫害或者分数决定论。该学区

的这一评估办法确实是有其内在逻辑的。毕竟学校领导也有可能是糟糕教师的朋友。他们可能只是喜欢这些教师的个性或是表面上的尽心尽力。糟糕教师很可能从表面看来是个好教师。所以，像许多其他的学校系统一样，华盛顿特区愿意减少人为偏差，更加注重评估得分，因为这一分数是根据实实在在的数学和阅读成绩计算得出的。华盛顿特区官员承诺，分数可以清楚地说明问题。分数更能体现公平。

韦索基当然觉得这些数字极其不公平，她想知道这些分数是怎么得来的。她后来告诉我说：“我认为没有人能理解这些分数。”一个优秀的教师怎么会得到如此低的分数呢？增值模型评估的到底是什么？

她所知道的就是，评估模型很复杂。华盛顿特区聘用麦斯迈提卡政策研究机构（Mathematica Policy Research）研发评估体系。该机构遇到的难题是测量特区学生在学业上的进步，然后计算学生的进步或退步在多大程度上归因于他们的老师。这当然不容易。研究人员知道，许多变量，包括学生的社会经济背景、是否存在学习障碍等，都会影响学生的学习成绩。评估算法必须要考虑到这些个人差异，这就是评估模型往往十分复杂的一个原因。

试图将人类行为、表现以及潜力归纳为某个算法或模型确实不是一件容易的事情。要想理解麦斯迈提卡政策研究机构处理的是什么问题，你可以想象一个住在华盛顿特区东南部贫困区里的10岁小女孩。在一学期的学习之后，她要参加五年级的标准化测

试。然后她的生活将继续下去。她可能正面对着家庭纠纷或是家庭经济困难，也许她正在搬家或是在担心她品行不良的哥哥，也许她不满意自己的体重或是在学校总被欺负。无论她在生活中经历了什么，下一学年她都要参加六年级的标准化测试。

如果你比较一下这个女孩两次测试的结果，最可能的情况是分数持平，当然更好的是分数提高了。但是如果分数下降，你能很容易地计算出她和那些优秀学生在两次测试的分数差距上差了多少。

但是，老师该为这一差距负多大的责任呢？这很难计算，而且麦斯迈提卡政策研究机构的教学评估模型只有少许数据可供比较。与之相反，像谷歌这样的大数据公司，研究人员会不断测试、监测成千上万个变量。他们可以把任一广告的字体的蓝色改为红色，将不同的版本分别投放给 1000 万名用户，然后追踪哪个版本获得的点击率更高，随时根据用户的反馈微调算法和操作。虽然我对谷歌公司有许多意见（接下来我将会在本书中做具体探讨），但谷歌的这种测试方法可以说是对数据的一种有效利用。

而想要计算一个人在一个学年内对另一个人的影响则复杂得多。韦索基表示：“学习和教学中有太多的不确定因素，很难一一评估。”而且，试图借助对二三十个学生的考试成绩的分析评估一名教师的教学水平，从统计学上来说也是不可靠的，甚至是很可笑的。样本量太小了，一切皆会出错。如果要采用严格的统计学

标准分析教师的教学效果的话，我们必须随机挑选几千个甚至数百万个学生参加考试。统计学家需要大量的数据平衡例外和反常情况。（我们在后文将会看到，数学杀伤性武器惩罚的个体往往是多数人中的例外。）

同样重要的是，统计系统需要反馈通路，以保证系统出差错时运行者能觉察到。统计学家会不断用差错训练模型，使之更加智能。若亚马逊的推荐模型的相关性计算出错，给十几岁的女孩推荐了草坪修剪的工具书，则其网站的点击量必然会发生骤降。为此，亚马逊公司就需要不断调整模型，直到用户相关性推荐的算法运作正常为止。但是，如果没有错误反馈，大数据模型就会持续输出错误的结果，而没人试图对此加以改进。

我将在本书中探讨的许多数学杀伤性武器都属于后者，包括华盛顿学区的教师评估增值模型。许多数学杀伤性武器都是依靠自己的内置逻辑来定义其所处理的情况，然后再以其自己的定义证明其输出结果的合理性的。这种模型会不断地自我巩固、自我发展，极具破坏力——而且在我们的日常生活中很常见。

在麦斯迈提卡政策研究机构的评分系统给予韦索基和其他 205 名教师差评之后，华盛顿特区开除了这些教师。但是该评分系统如何知道其决策是否正确呢？无从知道。评分系统确定这些教师是不合格者，那么别人就会认为他们是不合格者。206 名“差”教师走了。仅仅是这一事实就表明了该评估增值模型的效果——该模型正在清理华盛顿特区的不合格教师。比起探索教学质量不佳