

算法时代
新经济的新引擎

How Algorithms
Solve All Our Problems...
and Create More

Luke Dormehl
[美] 卢克·多梅尔◎著
胡小锐 钟毅◎译

THE
FORMULA



中信出版集团 · CHINA CITIC PRESS



算法时代 新经济的新引擎

[美] 卢克·多梅尔 (Luke Dormehl) © 著
胡小锐 钟毅◎译



THE FORMULA

How Algorithms
Solve All Our Problems...
and Create More

图书在版编目 (CIP) 数据

算法时代 / (美) 多梅尔著 ; 胡小锐, 钟毅译. —

北京 : 中信出版社, 2016.3

书名原文 : The Formula

ISBN 978-7-5086-5753-0

I. ①算… II. ①多… ②胡… ③钟… III. ①电子计

算机—算法理论—影响—世界经济—研究 IV. ①F11

中国版本图书馆 CIP 数据核字 (2015) 第 297293 号

The Formula by Luke Dormehl

Copyright © Luke Dormehl 2014

First published by W H Allen, an imprint of Ebury Publishing, A Random House Group Company. Rights arranged by Peony Literary Agency.

Simplified Chinese translation copyright © 2016 by CITIC Press Corporation

All rights reserved

本书仅限中国大陆地区发行销售

算法时代

著 者 : [美] 卢克 · 多梅尔

译 者 : 胡小锐 钟 毅

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 : 787mm × 1092mm 1/16

印 张 : 15.5 字 数 : 200 千字

版 次 : 2016 年 3 月第 1 版

印 次 : 2016 年 3 月第 1 次印刷

京权图字 : 01-2013-8586

广告经营许可证 : 京朝工商广字第 8087 号

书 号 : ISBN 978-7-5086-5753-0/F · 3556

定 价 : 59.00 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



在当前的大数据狂热中，自然科学家的声音正在显著压倒人文学者的声音。在大数据面前，人的一切都被预测决定，而自由意志失去栖身之地。在技术面前，人难道真的只有招架之功，没有还手之力了吗？人在 12 级台风中，需要一个保存自我的家园。《算法时代》将带你去寻找这一家园，这就是你需要阅读这本书的理由。

多梅尔是一位“数字人文”类型的作家。他眼中的技术明显有别于工程师眼中的技术，他更多地是从人的角度去思考技术。《算法时代》最关注的问题是算法技术对人的影响。

《算法时代》的主题词是“公式”（The Formula）。

作者认为，公式是技术理性的象征，“公式决定人”的技术决定论“反映了围绕客观性建立起来的某种社会秩序”。与这种客观性相对的，则是自我、爱、正义、创造性这些“主观性”的东西，围绕它们建立的社会秩序也应是主体性的。以此推论，“公式”的反义词应是自由意志，认同、爱、正义和创造力可以被看作自由意志的4个分身。

多梅尔在《算法时代》中，分别用4章的篇幅——探讨了“客观性”的公式与“主观性”的认同、爱、正义和创造力是否合拍的问题。

第1章讨论公式与自我认同的关系。针对大数据根据过往记录将人们归入不同细分类别，作者援引德勒兹与瓜塔里的《千高原》一书，指出公式把人变成了“可分解动物”；而人本应具有多种自我，“多种主观性并存”，“以错综复杂、粒度极高而且常常十分隐晦的方式相互影响”。

第2章讨论公式与爱的关系。针对图灵测试，作者引用珀森的话说，“在‘浪漫’的眼中，理性主义爱情观是没有全心全意投入的爱情”。浪漫主义者往往认为理性主义者的感情非常肤浅、受到抑制，而且他们害怕产生激情，小心翼翼地把自己囚禁起来，认为“再先进的算法也无法复制真爱”。

第3章讨论公式与正义的关系。法律主持正义，算法可以主持正义吗？作者认为，算法会倾向于死板地遵循法律条文，不知变通，而人更倾向于体现“法律意图”。

第4章讨论公式与创造力的关系。对艺术创作，作者不否认像《纸牌屋》那样利用数据分析可以推出热门产品，但提出的疑问是：“艺术品是否有统一的衡量标准”。

决定论与自由意志（非决定论）之争，是人类争论的永恒主题。纵观全书，《算法时代》是在大数据背景下重新讨论或者说升级了这一话题。

这不是老调重弹，而是透着时代精神。

首先，算法问题是时代的问题。如果说在原子时代对人起决定作用的是生产方式，那么在比特时代对人起决定作用的则变成了思维方式。因此，在信息时代讨论决定论与非决定论，当然应把重心放在思维方式上。算法是思维方式的技术基础，有什么样的技术基础，就有什么样的思维方式。这是在讨论算法问题的现实价值。

其次，算法问题也是人生的问题。生产方式与思维方式有一个共同之处，那就是“通过可以利用的资源实现价值的整套方法”。利用资源是手段，实现价值是目的。《算法时代》讨论算法的方式不同于工程师，是把它当作一个人生问题来看待。

当作者从认同、爱、正义和创造力这4个非决定论的方面质疑“公式”时，把算法的内在结构置于自由意志的对立面，即技术理性。算法真的只能属于技术理性吗？这可不一定。

图灵当初提出算法问题，是从机器与人两个方面立论的，并非只谈机器和决定论。其实我们可以反向设问，如果实现价值需要的话，为什么不能调整算法，从而建立新的思维方式——比如互联网思维呢？

至少有三个视角，可以帮我们延展思考《算法时代》关于算法的内在结构的问题。

一是“人-机”算法与“人-人”算法的视角。谷歌算法是典型的人-机算法，属于迷信数学决定论的方法。也就是多梅尔所说的技术理性算法。10年前，我有一次同中国科学院搞软件的数学家讨论谷歌算法时，数学家们提出一个悖论：你说脸谱网是什么算法？脸谱网的算法从数学角度不值一提，但计算结果不比谷歌差。道理很简单，这是一个脑筋急转弯。谷歌的模式是

人问机器，算法只在机器一端。而全世界的计算机加在一起，也达不到一个人脑的潜在计算力，为什么不能让人问人？这样的算法显然就不是技术理性了，而是“自由意志 – 自由意志”之间的 P2P 计算。这是价值观反作用于算法的例子。

二是全局最优算法与情境最优算法的视角。傻瓜在什么情况下会战胜聪明人（臭皮匠在什么情况下会战胜诸葛亮）？这又是一个脑筋急转弯。佩奇实验实际上把技术理性的算法颠覆了。答案应是，当人算不如天算时，傻瓜会战胜聪明人，因为傻瓜具有解决方案的多样性。从数学规划角度讲，当全局有唯一最优解时，诸葛亮往往胜于臭皮匠；但如果最优解只在节点，即情境最优时，臭皮匠往往又会胜于诸葛亮。这是价值观反作用于算法的又一个例子。

三是“深刻”算法与“肤浅”算法的视角。对技术理性而言，算法肯定是深刻好，而肤浅不好；能算到本质，肯定比只算现象要强。尼古拉斯·卡尔发现，互联网可以让人们的思维变得“浅薄”，但这不一定是坏事。自胡塞尔提出现象学之后，注重当下、此在的价值，就成了一种新的思维方式。顾客来到小摊，首先想要的问题是，面前这个小贩是不是个骗子？深刻的教授未必能回答这个问题，拥有大量咨询模板的麦肯锡也没法回答这样的问题。最适合回答这个问题的人，恐怕是头一天上了一回当的保姆。那么如果有种算法（比如小世界网络算法）能在一秒内找到保姆，这样的算法可能并不比“深刻”的算法差。

读《算法时代》，倒不一定非要得出什么确切的结论。从中获得启发，在某种众口一词的一般见识中看出不同的道理，就很值得。

姜奇平

中国社科院信息化研究中心秘书长，《互联网周刊》主编



算法的影响力

从根本上看，算法比一系列的步进指令略显复杂，通常由计算机执行。不过，尽管算法描述起来非常简单，它们的内部运行机制以及对我们生活的影响却绝不简单。

每天，算法都会对展示在我们眼前的信息进行分类、筛选与取舍。我们看到的谷歌搜索结果、脸谱网（Facebook）上高亮显示的好友信息，以及亚马逊网站预测我们很可能购买的产品类型等，都是算法运行得出的结果。此外，电影、音乐等娱乐方式的发展，通过预测浪漫关系的结果来选择伴侣，甚至法律的执行

与警察办案等，也将越来越依赖于算法。算法在扫描我们的元数据之后，可能会评价我们有希望成为一名勤劳的员工，也可能指控我们犯有某种罪行，或者认为我们不适合驾驶汽车。通过这些过程，算法正在深刻地改变我们看待生活、看待宇宙乃至看待一切事物的方式（引自道格拉斯·亚当斯）。

在人们针对技术发表的各种评论中，我最喜欢的是：“船只的发明同时带来了海难。”人们通常认为这是文化理论家保罗·维利里奥的观点。当然，我们也可以把这句话反过来，说海难的始作俑者就是船的发明者。在海难悲剧中，算法难辞其咎，但算法同时也发挥了极其重要的作用，帮助我们在每天产生的多达 2.5 艾字节（是人脑信息储存量的 100 万倍）的数据海洋中航行，并得出切实可行的结论。

这里，我将采用剥茧抽丝的方式，从本书的主要内容谈起。首先，这本书不宜用作计算机科学的算法教程。适合用作算法教程的书有很多，其教学效果也远胜这本（而且，很多作者比我更有资格编写这样的教程）。

其次，本书讨论的也不是算法概念的历史。我曾经也想做这种尝试，但是考虑到工作量太大，而且最终的成书（毫无疑问，如果由合适的人选来完成它，必将深受欢迎）与我敬而远之的教材并没有多大区别，因此我放弃了这个想法。我并不是说介绍算法历史的书与算法教程必然没有区别，而是指对算法（一度是小众数学概念）历史感兴趣的人，可能只限于那些已经熟悉这个概念的数学家或者计算机科学家。

算法通过各种各样的方式（有的非常微妙，有的则比较明显），影响着我们生活的方方面面，包括我们喜爱的娱乐活动与我们的爱情观。因此，在本书中，我希望揭示算法的这些影响，还要告诉大家以下事情的共通之处，比如，如何成功地吸引性感迷人的约会对象，好莱坞失败之作的产生过程，

将自己的排泄物装袋处理，减少律师从我们身上赚钱的机会等等。也就是说，在本书里，我要告诉大家的是算法对我们眼下生活的处理方式。

我白天的工作是为《快公司》杂志写作“数字人文”领域的文章。上班时我常常思考，“算法”文化以及认为所有问题都可以选择适当的算法予以解决的观点（未必都不正确），可能会带来哪些后果。

比尔·唐瑟尔于2009年出版的《在线为王》一书中有一个典型的例子，可以帮助大家理解本书的意图。唐瑟尔[至少有一位作者称他是“世界上最卓越的网上（行为）专家”]在书的开头介绍了他某天驾车时收听到的访谈。接受采访的是一位英国心理学家，他说自己研发了一种数学公式，可以找出一年中最令他沮丧的是哪一个星期。经过反复研究，他发现1月的第三个星期最令他感到沮丧，因为他同时遭遇了三个问题：新年决心未能实现，放假期间信用卡债暴增，没完没了的阴沉天气。唐瑟尔称他对此有所怀疑，但这并不是因为他认为一个简单的公式不可能为沮丧这种异常复杂、涉及多种因素的问题提供正确答案，而是因为他相信自己可以设计出一个更好的公式。

也就是说，他怀疑的不是最终结果，而是这个算法的工作原理。

多年来，我经常听到类似的说法，即人们认为技术可以将所有问题化繁为简，直至最适合用公式计算的程度，然后给出问题的客观答案。受他们的启发，我萌生了创作本书的念头。也正因为如此，我用大写来表示“公式”（**The Formula**），而不是把它简单地看成本书所描述的所有技术的笼统表达。对于算法及其相关技术，我是从更广义的角度来理解的，认为它们可以体现某种形式的技术理性，可以反映围绕客观性建立起来的某种社会秩序的特点，因此，“公式”这个概念暗含着某种意识形态的成分。这与已故美国政治学者、传播理论家哈罗德·拉斯韦尔对“技艺”（**technique**）一词的使用

颇有相似之处。拉斯韦尔认为，所谓技艺，就是“通过可以利用的资源实现价值的整套方法”，而本书中的“公式”既表示一种应用及其应用范围，还表示潜藏在表面之下、有待我们利用合适的数据挖掘工具去梳理发掘的客观事实。

技术作家（除了几位著名作家以外）的世界观往往包含了强烈的乌托邦思想，认为所有的技术进步都是积极的。于是，在新的发明问世之后，这些技术作家往往会欢呼雀跃，认为这项发明标志着一个新“时代”的开始。因此，人们以一种蔑视的口吻，称之为“网络（技术）万能论”。尽管本书也可以随大溜，为书架上长篇累牍的关于算法与大数据的著作锦上添花，增添一些新内容，但我真正感兴趣的不是讨论互联网或者个人电脑时代的诞生过程，而是希望做进一步研究。

20 世纪 60 年代前半叶，法国社会学家（同时也是基督教无政府主义者）雅克·埃吕尔在他的著作中描述了所谓的“技术人”，认为他们“对统一标准付诸实施之后所带来的即时结果感到心驰神往……孜孜不倦地寻找‘最佳方式’以实现任何指定目标”。他们希望实现的目标有时会变得含糊不清，但这不是因为他们对伦理道德的关注难以测量，而是因为他们对目标实现方法的热情过度，以为人类心灵手巧，针对任何问题都可以凭空想象出“效果惊人”的解决办法。

埃吕尔的研究表明，持有这种观点的人早已有之，谷歌创始人及本书讨论的那些高科技公司的领导者也不是“求序意志”（will-to-order）的先驱。这个概念是由已故美国社会学家刘易斯·芒福德提出的，它的意思是，人们希望借助公式了解身边世界。20 世纪 30 年代，早在现代计算机问世之前芒福德就在著作中指出，自动化的另一个目的是“放大人体的机械或者感官能

力，将人类生命的各个阶段简化为可以测量的秩序与规律”。他认为，我们对人类生命进行简化处理以便了解其全貌；为了让人类智力等抽象概念便于量化分析，我们又对这些概念做进一步提炼，去除各种复杂因素，再赋予看似任意的某个数值，最后得到某个人的智商值。

但是，这些著作还没有触及上述观点在现代社会中的影响范围。目前，算法与“公式”的影响无处不在，几乎涉及工作与休闲领域的方方面面。因此，本书试图探讨这种局面的形成原因，研究算法在人类创造力、人际关系（更具体地说，是人们的浪漫爱情）、身份概念与法律问题等方面的影响力。

在这些方面，算法都可以为我们提供答案。

但问题在于，算法提供的答案真的是我们想要的吗？

推荐序 IX

引言 算法的影响力 XIII

第1章 算法程序与人类的自我认知 001

通过数据认识自我 007

量化的世界 009

购物也是创造 011

用户的喜好就是我们的商机 012

第三次浪潮 015

请不要挂机，马上为你转接 015

谷歌的乌比冈湖策略 018

量化人的职业潜能 023

通过推特和脸谱网窥视用户的生活 026

- 近乎专制的员工管理方式 031
- 提升员工的工作效率与幸福度 034
- 简约现实聚合器与过滤气泡 036
- 清道夫级用户和差别定价法 039
- 圆形监狱与人类的数字身份 042
- 算法赋予我们的身份类别 046

第2章 理性的算法与浪漫的爱情 049

- 用公式预测婚姻关系的持续时间 056
- 算法时代的婚介业务 059
- 互联网为每个人都准备了一个公式 063
- 算法时代的爱情观 066
- 你确定要删除这段感情吗？ 069
- 抓住技术创造的浪漫邂逅机会 071
- 形形色色的社交发现软件 074
- 触碰手机屏幕就能知晓一切 076
- 以前所未有的方式思考性生活 078
- 跟故去的亲人互道“我爱你” 083
- 人与机器人之间的浪漫爱情 086
- 爱情的算法化和编码化 093
- 再先进的算法也无法复制真爱 095

第3章 算法与人类执法者的较量 099

- 用算法预测并制止犯罪行为 104

犯罪统计学和道德统计学	108
现实版《少数派报告》	112
谁实施犯罪行为的可能性更大	114
未来的我们不需要律师	118
计算机算法属于颠覆性创新	120
透明的执法者	122
谷歌地图与社会偏见问题	124
技术给人类撰写的“剧本”	126
令人瞠目结舌的 500 张罚单	129
规则与标准的冲突不断升级	131
公正的算法与带有偏见的算法	133
算法犯下的种种错误	136
算法解决方案的透明度问题	139
法官、陪审团与数据挖掘技术	142

第 4 章 用科学的算法进行艺术创作 147

公式和规律随处可见	151
能预测电影票房收入的“秘密武器”	153
艺术家与科学家之间的鸿沟	158
“平行宇宙”实验与流行歌曲排行榜	161
依据吸引力的大小量化艺术作品的价值	163
从二维照片变成三维物体	168
隐藏在艺术作品中的“真相”	170
《纸牌屋》成功背后的最大功臣	175
如果每部电影都有 4 种结局	176

可以一直演下去的电视剧	180
数字看门人巨大的隐性权力	183
艺术品位是否有统一的衡量标准?	184
计算机也有可能成为艺术家	188

结 语 预测未来 193

算法在我们的生活中变得越发重要	195
算法能否处理一切事务?	196
每周工作 20 个小时	200
人类和计算机的角色分工	203
后机械时代的客观性	205
算法会造谣中伤他人吗?	210
所谓的客观性不过是一种假象	213
人类世界被植入了计算机逻辑	217
在算法时代必须养成提问的好习惯	221

致 谢 227



The Formula

第 1 章

算法程序与人类的自我认知