

全数据时代的 炼金师

顾佳峰◎著

ALCHEMISTS
OF THE WHOLE DATA ERA

中国工信出版集团



人民邮电出版社
POSTS & TELECOM PRESS

全数据时代的 炼金师

—— 顾佳峰◎著 ——

ALCHEMISTS
OF THE WHOLE DATA ERA

人民邮电出版社

北 京

图书在版编目 (C I P) 数据

全数据时代的炼金师 / 顾佳峰著. — 北京 : 人民
邮电出版社, 2017.3
ISBN 978-7-115-45184-2

I. ①全… II. ①顾… III. ①数据处理 IV.
①TP274

中国版本图书馆CIP数据核字(2017)第050332号

-
- ◆ 著 顾佳峰
责任编辑 吴娜达
执行编辑 唐名威
责任印制 彭志环
- ◆ 人民邮电出版社出版发行 北京市丰台区成寿寺路 11 号
邮编 100164 电子邮件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
三河市海波印务有限公司印刷
- ◆ 开本: 720×960 1/16 彩插: 6
印张: 12.25 2017 年 3 月第 1 版
字数: 198 千字 2017 年 3 月河北第 1 次印刷
-

定价: 59.00 元

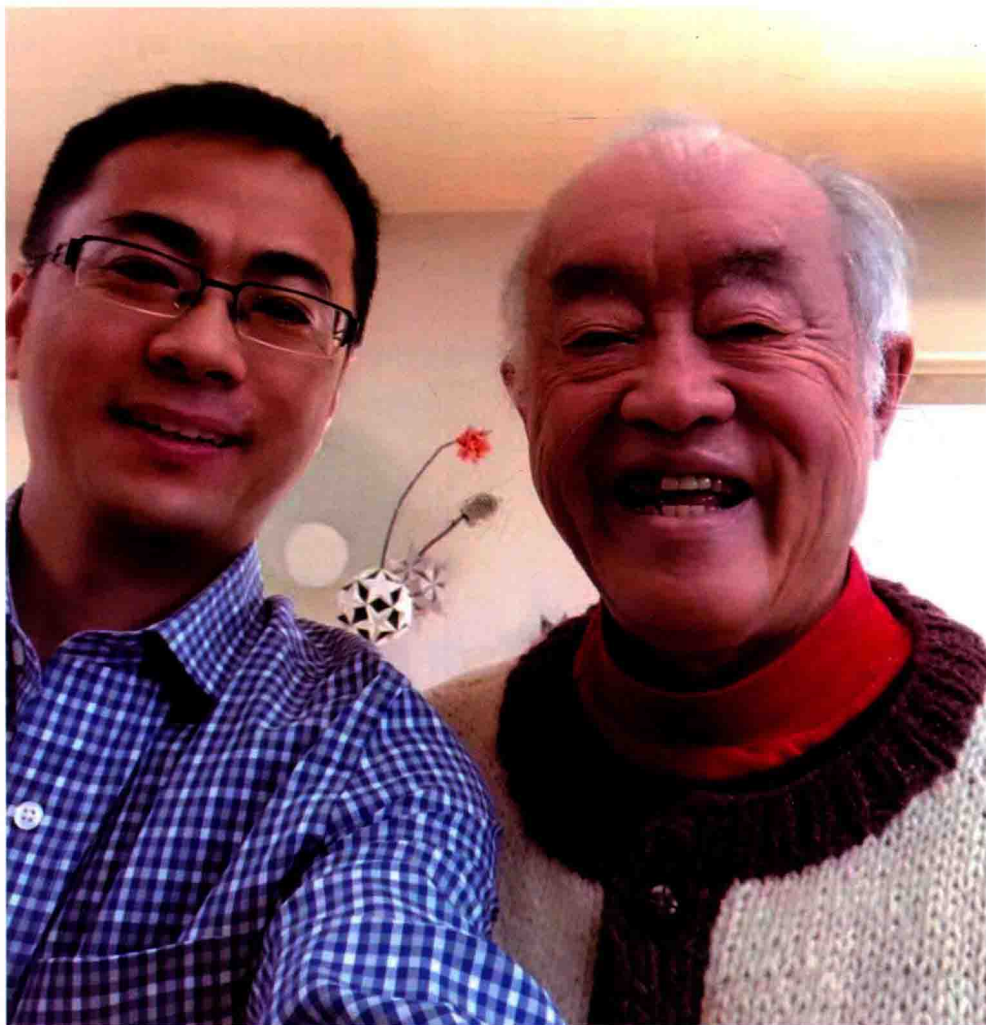
读者服务热线: (010)81055488 印装质量热线: (010)81055316
反盗版热线: (010)81055315

佳峰同学纪念

替中国科技争光

何毓琦

2017.1.21



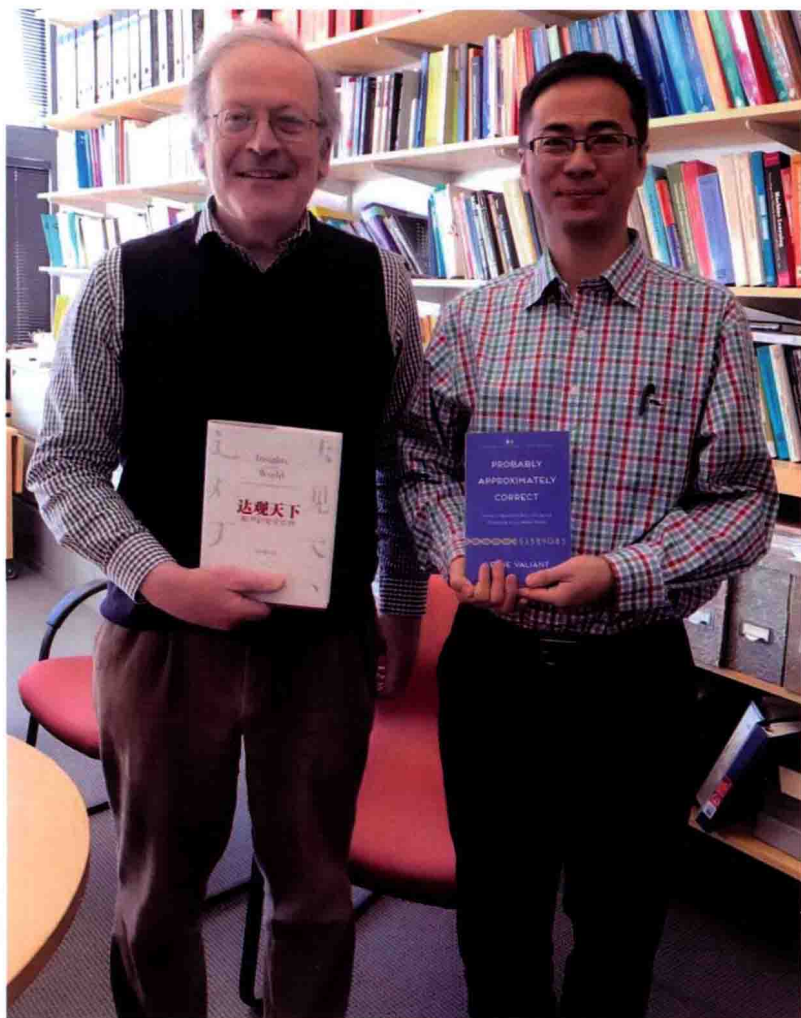
何毓琦教授与本书作者顾佳峰教授

顾佳峰教授拜访了著名华裔科学家，美国科学院院士，中国科学院、中国工程院两院外籍院士，IEEE Fellow何毓琦教授。顾教授向这位工程界国际权威、国际泰斗级人物汇报了本书的思想和观点，何老频频点头，并为本书写下了“替中国科技争光”的鼓励。



艾伯特-拉斯洛·巴拉巴西教授与本书作者顾佳峰教授

顾佳峰教授与全球复杂网络研究权威，无尺度网络的创立者、美国物理学会院士、匈牙利科学院院士、欧洲科学院院士艾伯特-拉斯洛·巴拉巴西（Albert-László Barabási）教授就本书内容进行了深入的探讨和交流，本书受到了他的高度肯定及推荐。



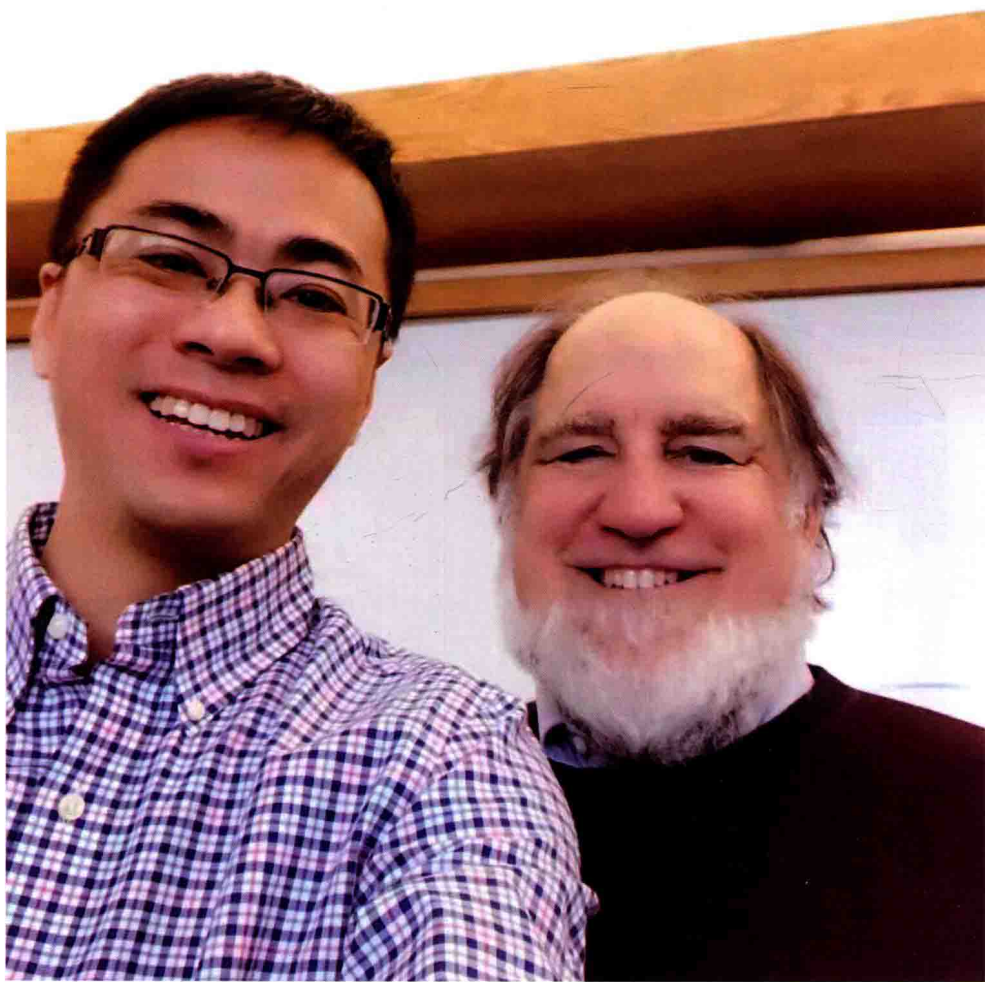
莱斯利·瓦伦特教授与本书作者顾佳峰教授

顾佳峰教授邀请了图灵奖得主、美国科学院院士、哈佛大学莱斯利·瓦伦特（Leslie Valiant）教授为本书撰写推荐语。莱斯利·瓦伦特教授就书稿中提到的大数据算法问题耐心地向顾教授进行了解答。



丹妮拉·鲁斯教授与本书作者顾佳峰教授

顾佳峰教授拜访美国科学家天才奖得主、美国科学院院士，MIT人工智能实验室（CSAIL）主任丹妮拉·鲁斯（Daniela Rus）教授，就本书提到的人工智能与机器人未来发展议题进行了畅谈，收获良多。



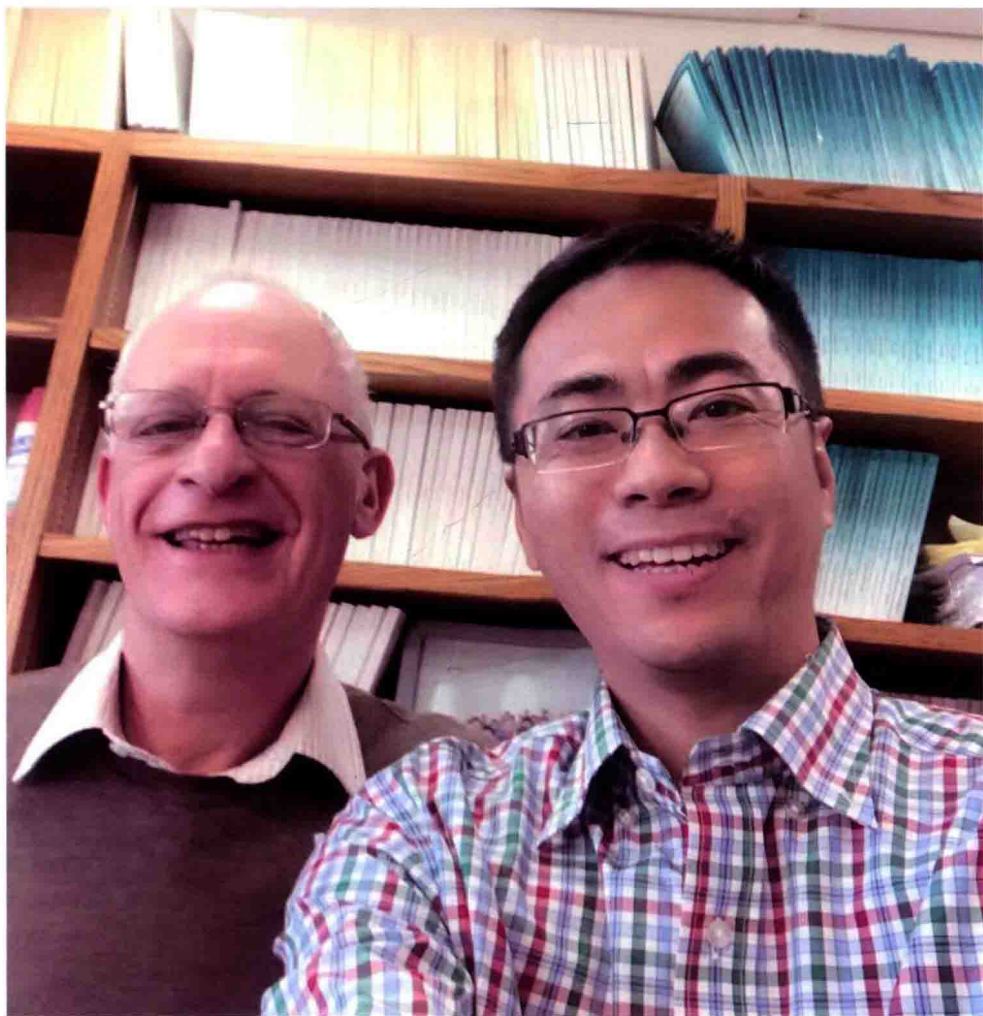
罗纳德·李维斯特教授与本书作者顾佳峰教授

顾佳峰教授拜访了计算机科学领域教父级人物MIT的罗纳德·李维斯特（Ronald L. Rivest）教授，他是美国科学院、美国工程院、美国艺术与科学院三院院士，图灵奖得主，密码学领域的三杰之首。罗纳德·李维斯特教授倾听了顾佳峰教授对本书的介绍，对书中的主要观点表示了赞同，建议可以出版英文版。



希尔维奥·米卡利教授与本书作者顾佳峰教授

顾佳峰教授拜访了美国科学院院士、美国工程院院士、美国工程院与美国艺术与科学院三院院士，图灵奖得主MIT的希尔维奥·米卡利（Silvio Micali）教授，就本书主要内容和观点和他进行了深入的探讨，获得了他的肯定。讨论很热烈，颇有相见恨晚的感觉。



奥利佛·哈特教授与本书作者顾佳峰教授

顾佳峰教授拜访2016年诺贝尔经济学奖得主奥利佛·哈特（Oliver Hart）教授，顾教授与这位经济学大师交流，探讨了数据科学与经济学研究之间的关系，大师的见解令人醍醐灌顶。



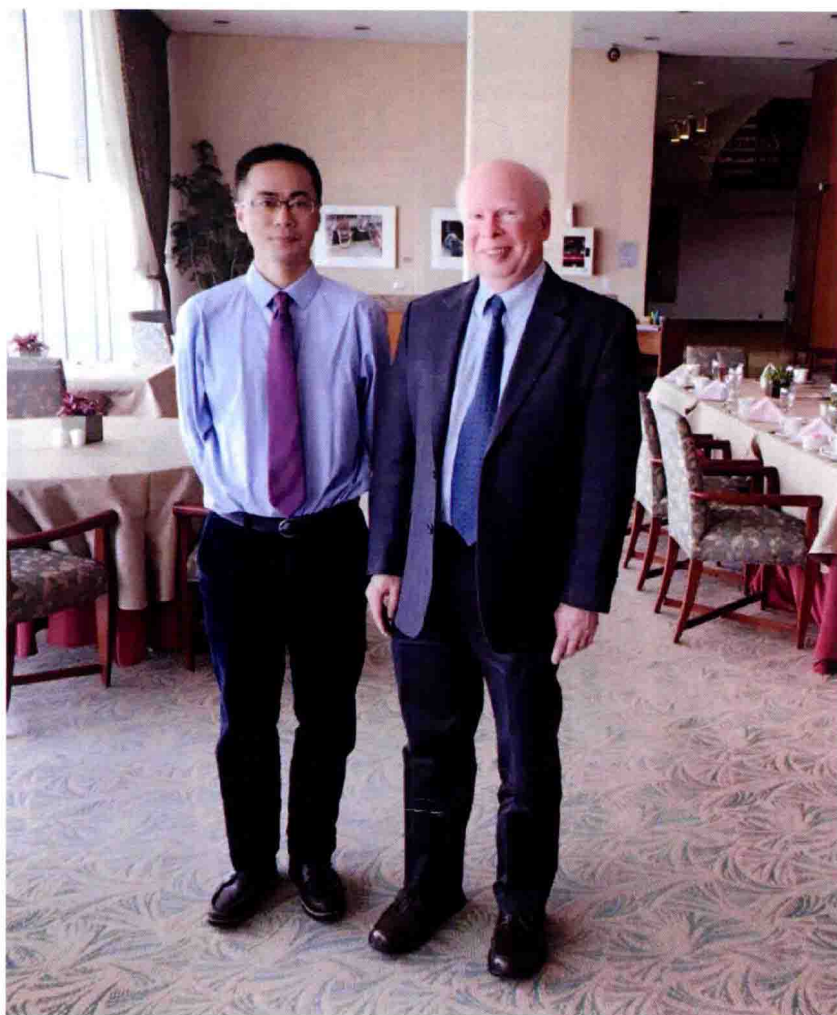
本格特·霍斯特罗姆教授与本书作者顾佳峰教授

顾佳峰教授拜访了2016年度诺贝尔经济学奖得主MIT的本格特·霍斯特罗姆（Bengt Holmström）教授。顾教授与这位经济学大师交流，探讨了数据科学与经济研究之间的关系。两个多小时的交流转瞬即逝，他的见解令人茅塞顿开，收获颇丰。



吴皓教授与本书作者顾佳峰教授

顾佳峰教授就本书中关于精准医学和医学大数据的内容和美国科学院院士、哈佛大学吴皓教授进行了深入的讨论，并听取了吴教授的意见。



罗伯特·特里斯特教授与本书作者顾佳峰教授

应波士顿联储银行副主席罗伯特·特里斯特（Robert Triest）教授的邀请，顾佳峰教授赴波士顿联储银行大楼最高层的联储内部餐厅与他单独共进午餐，席间与他深入探讨了金融大数据的发展和应用情况。



罗伯特·兰格教授与本书作者顾佳峰教授

顾佳峰教授拜访了集美国科学院、美国工程院、美国医学院三院院士于一身的罗伯特·兰格（Robert S. Langer）教授，他是麻省理工学院级教授。在交流过程中，与他就生物技术与医疗大数据的最新发展、精准医疗以及创业工程等进行了深入探讨。

内容提要

本书用生动活泼的语言，结合古今中外实例来分析数据科学发展历程及其趋势，探讨大数据和调查数据在现实社会中的妙用，便于人们全面认识各种数据的特征及数据融合的力量，提高人们数据利用能力。本书还深入论述了社会生活中人们如何通过“数据修炼”来训练全数据思维，掌握从浩繁数据中获取有价值信息的能力，成为数据时代的“炼金师”，使得数据为我所用。这是一部适合数据时代所有人阅读的书籍，相信广大读者都能从中获得启发和助益。

在当今社会的诸多领域，大数据已成为众人关注的焦点。这种关注热潮使得大数据相关的国家战略、学术研究和产业应用发展迅猛，然而，在大数据的理念和应用落地之间依然存在着巨大的鸿沟。虽然，当前各国已经将发展大数据纳入国家战略，工业界也对大数据的发展做出了重大的贡献，包括 Hadoop 平台的发布等，然而，现实中对大数据的应用需求还远远未被满足。大数据思潮已经兴起，但在这场思潮中，掌握大部分应用需求的参与者却仅仅停留于对照书本做事，从整体来看对大数据的应用还不成体系。理念与应用的脱节，导致了大数据的真正价值无法落地，在更多人看来，大数据是一个高高在上的玄妙概念，令人憧憬并且困惑。

顾佳峰教授的《全数据时代的炼金师》这本书，为大数据理念与应用落地之间的鸿沟架起一座有效的桥梁。作为社会调查专家，顾教授从调查数据角度入手，阐释了大数据和调查数据各自的起源、优势与缺陷，并用众多生动的案例诠释了大数据与调查数据的结合方式与价值。在我看来，顾教授此书不但阐明了大数据与调查数据结合的“炼金”方法，更是从一个重要的方向对大数据理念与应用间的鸿沟进行了填充和诠释，让读者对大数据有了一个更清晰全面的认识。

作为一个大数据的研究人员，在对大数据的认识方面，我的许多思路与顾教授不谋而合。大数据真正重要的意义不是对大量数据的存储与处理，而是通过大数据思维，重构商业和管理模式，重新定义问题。顾教授在这本书中，用很多案例和数据将这些观点娓娓道来。例如，在数据炼金的三大基本策略这一部分，作者提到了在美国加利福尼亚州的杰苏斯塔大火中，个人和组织自发建立地图站