

BIG DATA
HISTORY, REALITY
AND FUTURE

数据之巅

大数据革命，历史、现实与未来

涂子沛◎著

数据表示的是过去，
但表达的是未来。
所以看数据要有宏大的历史观。

政府施政，企业掘金。

中信出版集团

中信出版集团

数据之巅

大数据革命，历史、现实与未来

涂子沛◎著

BIG DATA
HISTORY, REALITY
AND FUTURE

图书在版编目 (CIP) 数据

数据之巅: 大数据革命, 历史、现实与未来 / 涂子沛著. -- 2 版. -- 北京: 中信出版社, 2019.1
ISBN 978-7-5086-9710-9

I. ①数… II. ①涂… III. ①信息经济学-研究
IV. ①F062.5

中国版本图书馆 CIP 数据核字 (2018) 第 245890 号

数据之巅——大数据革命, 历史、现实与未来

著 者: 涂子沛

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: 北京盛通印刷股份有限公司

开 本: 787mm×1092mm 1/16

印 张: 28

字 数: 328 千字

版 次: 2019 年 1 月第 2 版

印 次: 2019 年 1 月第 1 次印刷

广告经营许可证: 京朝工商广字第 8087 号

书 号: ISBN 978-7-5086-9710-9

定 价: 88.00 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

数据文化是尊重事实、强调精确、推崇理性和逻辑的文化。数据文化的匮乏，是中国之所以落后的一个重要原因；建设这种文化，中华文明的面貌将焕然一新。

给奕奕、捷克和炜婷，时时温暖的爱和快乐
To Yiyi, Jack and Weiting for the warm love and happiness

导 读

《数据之巅》全书共分为两个部分：小数据之历史和大数据的崛起。

全书以美国为主体，横跨东西方两个文明来阐述数据文化的形成、发展和未来。作者把美国 200 多年的历史划分为 7 个时代，第一个百年的 3 个时代——初数时代、内战时代、镀金时代——按时间点来划分，有起点也有终点；进入 20 世纪之后的 4 个时代分别为：量化时代、抽样时代、开放时代、大数据时代。一定程度上，这 4 个时代至今仍在延续并产生影响，所以这本书只按标志性的事件标记起点，而没有设定终点。

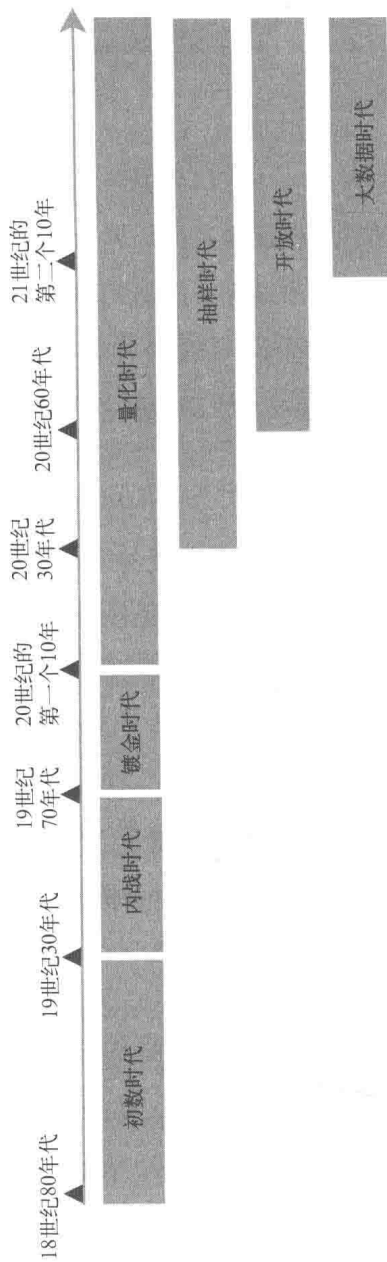
第一部分以美国的历史为主线，展现了在历史的长河中数据文化如何形成、数据技术如何兴起，数据治国的理念如何深入人心的宏伟画卷。这一部分按照时间序列共分为 5 章，每一章集中阐述了该历史阶段内发生的代表性事件，并对中国的相关问题进行了分析和思考。

第二部分共分为 3 章，分别是开放时代、大数据时代和智慧城市。作者详细地考证了大数据浪潮的来龙去脉，并结合第三次工业革命分析了大数据对商业运营、社会治理的深远影响。作者认为，有数据，还要有计算，更大的数据爆炸正在到来，计算型社会即将兴起，而人类使用数据的巅峰形式，是通过数据赋予机器“智能”。大数据浪潮最终将引领人类社会迈进一个新的形态：智能型社会，当下智慧城市的建设正是这一转型的要求和表现。作者

最后立足现实，对当前中美两国智慧城市建设中的若干问题和挑战进行了对比和探讨。

结语部分，作者总结了大数据的重要性，展望了数据文化的未来，认为世界范围内的一切竞争，归根结底都是国家之间国民素质和文化的竞争。没有一个健康、理性、与时俱进的文化，一个国家就难以变得强大。作者提出，中国社会要将大数据这个科技符号转变为文化符号，才能改变国民面貌、提高民族竞争力。

全书的篇章时序结构



专家热评

数据本无大小，运用数据的立场却分出大小，是谓大数据。数据表示的是过去，但表达的是未来，所以观察数据需要有历史观。涂子沛先生的这本新作《数据之巅》以大历史、大数据、大社会、大进步的格局，通过数据将遥远拉近，将悠久缩短，将巨大归聚，将过程简化。无论你从事何业或研究何题，这部巅峰之作都值得你认真阅读。

——朱小黄 中信银行前行长

涂子沛先生《数据之巅》一书的历史高度和现实意义，超出了他的《大数据》。如果说他上部书开启了中国人对大数据的认识，此书则从历史的角度来说明，对数据的理解和掌握是人类文明进步的标志。中国人在“初数时代”是被动的、落后的，黄仁宇在他的名著《万历十五年》中也将中国明朝未能走向资本主义归结为中国人不能在“数目字上管理国家”。在新世纪的大数据时代，我们能否成为参与者甚至领导者？子沛这本书中有诸多可贵的思考和启示，我向读者推荐阅读。

——田溯宁 中国宽带资本基金董事长

大数据的江湖有两派：国家治理派和产业升级派。涂子沛先生无疑是国家治理派的宗师，这本书在美国 200 多年的政界商海、刀光剑影中勾勒出雄浑的数据思维、文化、价值观和方法论。临到尾处，我的期待愈来愈强。他的第一本书《大数据》在通篇的历史视角、西方视角后，煞费苦心附上一篇“挑战中国”的尾声，把人拉回当下，令人掩卷忧思；这次果不其然，子沛以整整一章的浓墨重彩，挑战了“智慧

城市”这一兼具进行时和未来时的话题，而当涉及中国元素时，子沛的行文中出现了暖色调，那种欣赏和希冀让人宽慰之极、意犹未尽！期待《数据之巅》之后，山不穷，峰更高。

——吴甘沙 驭势科技联合创始人、CEO

涂子沛先生已经成为中国大数据和数据分析领域的前沿思想者和专家，如果你想了解大数据的历史、现实和未来，这本书将是你的常备书目。

——托马斯·H.达文波特 美国巴布森学院教授

麻省理工学院数字商务中心研究员

涂子沛先生不仅在写书，还在用书传递他的炽热激情：大数据时代是上天恩赐给中国人的，国人切切不可错过这个经济发展和社会进步的时代机遇。如其书中所言，在这个新的时代，我们不仅要“用数据说话”，还要“用数据创新”。

——刘鹰 清华大学社会科学学院教授

在全球化、信息化和市场化的今天，国际竞争本质上是软实力的竞争，即制度和文化的竞争。涂子沛先生倡导的客观、精确、理性和逻辑的“数据文化”理念，不仅是挖掘中国传统优秀文化、吸纳西方文明、重塑中华文化的新范式，也是推进国家治理体系和治理能力现代化的利器。

——周超 中山大学公共政策研究中心主任、教授

《数据之巅》以数据为轴线洞察人类社会的发展规律，描绘未来的趋势

和走向，力透纸背的还有作者严谨的治学风格及其赤子情怀。涂子沛老师的这本新书就是一座桥梁，实现了中西方之间、历史与现实之间、政治与商业之间的贯通融合，深入研读，可以帮助我们进入通达无碍的智慧境界！

——付伟 中国银联支付学院院长

身处国内喧嚣浮华的互联网金融盛景，此书恰似拨云见日，涂子沛先生跨洋而来的目光深邃地看清了中华文明与西方文明在虚与实上的认知鸿沟。民主的细节在于数据，金融的血液中流淌的也是数据，《数据之巅》，众心向往！

——郭宇航 点融网创始人、CEO

作为《大数据》最早的一批读者，我曾于2012年10月专程前往美国拜会涂子沛先生。涂先生的格局和视野、勤奋和雅静以及时代的责任感令我印象深刻。这本新书立意高远、数往知来，把数据从科技符号演进为文化符号，在今天这样一个尤其需要高瞻远瞩的大变革时代，相信《数据之巅》会再次成为热点话题，推动中国数据文化的普及。

——陈登坤 有米科技股份有限公司总裁

涂子沛先生的新作《数据之巅》揭示了在人类近现代文明进程中，在民主和共和体制的创立演化中那无形的数据之手。国内鲜有类似的作品。把数据与人类社会的进步、民主体制的演化联系起来，这是真正的人文关怀。当下的中国，从不缺“民主自由”的口号，但口号建不起先进的人类文明，数据才能。

——赵嘉敏 译言网联合创始人、CEO

通过《数据之巅》，涂子沛先生再次深入美国历史中的细节，以真正治学之人的责任与良心，发掘“数据”所凝聚的可以再生的思想文化资源。数据文化是一个全新的视角，非常值得教育工作者探索和思考。在智能型社会、人机共生的时代敲响鼓点、加速到来的时刻，对这本书的阅读，也是一场自我启蒙之旅，它将为我们推开那扇未来世界之门。

——潘江雪 上海真爱梦想基金会理事长

数据自古存在。这本书截取历史长卷中的数据剖面，或古今中外，或政经产学，案例生动，立意高远，令人心生钦敬。

——赵国栋 中关村大数据产业联盟秘书长

推荐序一 进入一个重要的现代文化园地

涂子沛先生所著的这部大作，是大数据时代的应时之作。

他将数据用在管理和研究方面的发展史，以数据在美国的发展过程作为主要内容。他将美国开国时期的人口普查作为起点，讨论民主制度如何经过数据的调查才能发展成为“一人一票”的制度。接着，19世纪中叶，美国向西开发，美国的工程兵团进行丈量和调查，使美国的地理状况和疆域都有明白的依据。20世纪“打孔”计算方法的发明，开创了后世计算器管理大量数据的技术。到今天，我们日常生活中，因为计算机和网络的普及，数据无处不在。以治病为例，要判断疾病的性质或药物的效果，必须将大量的数据作为诊断和治疗的依据。凡此，都是数据的应用。

美国的社会愈来愈复杂，在资本主义的国家，证券交易乃是一桩大事。单单用统计来管理证券，已经不太够用。最近30来年，许多大证券商为了预测经济的起落和某一种产业的兴衰，大量地使用不同产业之间的关系，也顾及国际贸易的情形。这些私营的企业，几乎都能相当精准地判断市场的情形。于是，美联储实际上就是美国的中央银行，也必须更细密地运用许多数据，以掌握经济的全貌，然后再决定

对市场供应的货币是从宽还是从紧。这才是大数据的第一次应用。

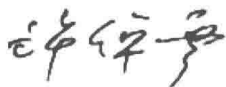
中国古代兵书《孙子兵法》就说过：“算则胜，不算则不胜；多算胜，少算则不胜。”此处的“算”字，就是指如何利用数字估计各种因素。一个能干的将领打一次战役，要考虑到天时、地利、人和，这都是可以用数字表现的。但是，一个治国的领袖，在上述因素以外，还得考虑许多其他的条件。1942年，日本偷袭珍珠港，以为可以一棒打死美国。他们没算到，美国工业的实力有充分的再生力量，三个月之内，美国立刻就能恢复足够的海空实力。这就是日本军人，他们只知道计算战役，不会计算战争。到今天，安倍野心勃勃，处处挑衅，他志在日本复兴。他的计算，大概又是计算自己现在的兵力和科技能力；他没有算到，自己的原料供应不够，能源不能自主，人口结构老化：这就是“算”得不够。

大数据之“大”，就在于将各种分散的数据彼此联系，由点而线，由线而面，由面而层次，让人瞻见更完整的覆盖面，也更清楚地理解事物的本质和未来的取向。人脑的结构，足够发挥联想力和推论。在每天的日常生活中，我们时时刻刻在不知不觉中做大数据的工作，将许多因素综合在一起，作为行动和决定的依据。只是，一个国家或一个社会的发展，不能全靠眼睛看得见的一些讯息。有许多事物，必须依靠全面和长期的发展情况，才能真实地反映当时一切决定的背景和条件。

今天，信息科学的发展已经能够产生、存储并实时地分析处理大量的信息，整合多个源头的的数据，形成全面的多项关系，指出线性的发展方向，引导我们拥有广阔的视野了。计算机今天处理数据的能力和速度，已经超过最聪明的个人。可是，使用计算机的还是“人”，如

何驾驭这些数字，还是“人”在设计。只是，我们必须要有此认识：今天的世界，已经千丝万缕，将各地、各种行业、各种条件，纠缠成一个复杂的全球网络。管理大企业和管理国家，必须要有足够的信息，了解多种多样的情况，以全面地理解各种问题及其彼此的关联。即使是对个人而言，因为越来越多的行为已经转变为电子化的记录，其生活也变得和大数据息息相关。大数据这个课题十分重要，我盼望有更多的学者在这方面写出更多的作品。

涂子沛先生的大作，是讨论大数据较早的中文作品；在此以前，还罕见讨论数据作为管理方式的书籍。这本书主要以美国社会中数据的使用为例。我知道他一心想用这些例子提醒中国的读者：在信息科学高度发展的今天，我们不能再忽略数据的使用。涂子沛先生开启了一扇大门，我相信，后面他会开启更多扇门，让大家进入这一个重要的现代文化园地。



著名历史学家、美国匹兹堡大学历史系荣誉讲座教授

2014年3月31日

推荐序二 一部精彩纷呈的时代杰作

好看的作品，出色的作家

认识涂子沛先生，源于他的第一本著作《大数据》。2012年，我在机场书店无意间看到这本书，一读起来，就不忍释卷。我当时很惊讶，没想到对信息技术产业内的大数据技术，竟然可以这样写。这本书如此自然流畅地与美国的社会发展、民主进程融合在一起，有观点、有故事，读来引人入胜，掩卷引人深思。

之后，我主动联络了涂子沛先生，邀请他来神州数码参观考察。涂先生欣然接受，他不仅给神州数码的员工做了非常好的演讲，而且，还与我们围绕中国智慧城市建设的课题，进行了深度碰撞，对我启发很大。

正是那个时候，他告诉我，他已经在构思下一本书，还是以大数据为主题，但会和中国有更多的结合。我非常兴奋，马上向他表示，神州数码非常愿意向他敞开大门，我们在大数据和智慧城市建设方面的所有思考、探索和实践，毫无保留地向他公开，欢迎他来了解、见证我们

的发展。在这之后一年多的时间里，涂先生果然多次来访神州数码，亲自走访了佛山、苏州、张家港、武汉等地，认真访谈了我们的业务负责人、技术带头人，也多次登门拜访各地的政府用户、企业和市民。他表现出的敬业精神和专业能力让我非常感动，也令我相信他一定能够再次完成一本具有社会影响力的作品。

亲眼看到《数据之巅》这部书稿时，我承认，我再次被震惊了。这本书再一次超出了我的预期，除了承袭《大数据》一书中科学历史观的叙事方式，这一次，涂先生还跳到了哲学思考的层面，以统计学的社会应用为切入点，解构数据文化在美国政治、经济乃至军事发展领域起到的关键作用，一环扣一环，构思精巧，故事生动，逻辑清晰，读起来实在“解渴”。而且，正如他曾经和我说的，“要和中国有更多的结合”，在每一章的最后，他都讲述了中国历史上相对应的数据事件。最后一章的视线更是完全转向中国，用独具中国特色的智慧城市建设案例，理性昭示着中华民族自己的未来。其间，亦庄亦谐的“子沛曰”，也体现了涂先生的幽默和智慧。

说涂先生是中国当代文坛最出色的科学作家之一，这毫不为过，在信息技术领域，他也是前沿的思想者。

捅破东西方哲学的窗户纸

众所周知，理性化、体系化、强调批判精神和实证精神，是西方哲学的特征；感性、体验、直觉，则是东方人的思维方式。中国的传统文化，喜欢用道、术、器对事物的本质进行模糊的归纳总结，而西方则在数据文化的基础上，形成了严谨、理性、体系化的实证科学，如统计