

2013被称为大数据元年。关于大数据的图书市面上就有20多种，其中最突出的是维克托·迈尔-舍恩伯格的《大数据时代——生活、工作与思维的大变革》（浙江人民出版社出版）。

舍恩伯格是最早洞见大数据时代发展趋势的数据科学家之一，早在2010年就在《经济人》上发布了对大数据应用的前瞻性研究，他的咨询客户包括微软、惠普和IBM等全球顶级企业，被誉为“大数据商业应用第一人”。

下面，我们循着舍恩伯格《大数据时代》的脉络，了解一下大数据究竟意味着什么。

世界的本质是数据

2009年，甲型H1N1流感爆发的前几周，谷歌的工程师在《自然》杂志上预测大型流感传播即将到来。不需分发口腔试纸或调查医生，他们建立了一个系统，在每天收到的数十亿条搜索指令中关注特定检索词条（如“哪些是治疗咳嗽和发热的药物”等）的频繁使用与流感传播之间的联系，及时判断流感从哪里传出。而疾控中心要到流感爆发一两周后才能确定。

谷歌的判断就建立在大数据基础上：即以一种特定方式，对海量数据进行分析，获得有巨大价值的产品和服务或深刻的洞见。

传统经济统计原则是以抽样表征总体，人类对世界的认识，很像打着手电筒，清晰地看到脚下的某一块石头；而大数据时代的信息统计包含的样本量是如此之大，如同一盏灯笼，可能细部不够准确，但能照出整个环境和道路通向哪里。那些不很准确的最原始、看似平凡无用的信息，经过恰当的分析，让我们越来越接近正确——

中英人寿通过分析爱好、常浏览的网页、常看的节目、收入估计等好几百种生活方式数据，找出更有可能患高血压、糖尿病和抑郁症的人。申请者不必提供血液和尿样，这个纯数据分析法只需5美元，可使保险公司在每人身上节省125美元。

网站内容设置依赖于数据而不是编辑的新闻敏感度，数据比有经验的记者更能揭示哪些是符合大众口味的新闻；

网上教育公司深入研究它收集的所有数据，比如学生重放讲座视频的哪个片段，从而找出不明晰或很吸引人的地方反馈给设计课程的团队……

这很像一场寻宝游戏，通过数据科学家的巧手，这些数据的潜在价值被挖掘出来，远超其最基本的用途。数据发出了自己的声音，带给我们惊喜。有了大数据的帮助，我们会意识到本质上世界是由信息构成的。

挖掘处理数据是“大数据”的真谛

一个男人冲进塔吉特商店，气愤地对经理说：“我女儿还是高中生，你却给她邮寄婴儿服和婴儿床的优惠券，你们是在鼓励她怀孕吗？”几天后，当经理打电话致歉时，男人语气变平和了：“我女儿的预产期是8月份，是我完全没有意识到这个事件的发生，该说抱歉的是我。”原来，塔吉特的分析团队发现，怀孕3个月的女性会买无香乳液，之后会买镁、钙、锌等营养品，有20多种关联物能使零售商较准确地预测预产期，寄出相应的优惠券招徕顾客。

在大数据时代我们可以预测未来。古代能够预言天气的人，往往被视为通神，而今需要的只是海量信息的占有和分析。

大数据不仅仅在于数据之大，挖掘处理才是大数据关键所在。一要有收集及开发数据的特定工具，二要有集黑客和定量分析员专长于一身的数据科学家。

随着相关技术的成熟，公共部门和私人企业过去积累的大量“垃圾”数据有可能重焕光彩。比如用微观居民和企业用电量数据指导智能电网建设、用交通事故和犯罪数据指导警力布局、用消费和税收数据指导收入分配、用客流量数据指导铁路和民航调配、用互联网关键词传播数据进行流行病预防等等。

沃尔玛可谓是大数据的玩家。在对每个顾客购物篮中的物品、具体购买时

群体的两面性

——我读《乌黑之众》

赵柒斤

信息时代，“人多力量大”不仅在网络上充分显示，还疯狂地向现实生活蔓延。许多“争吵事件”，最后占上风的往往不是“理”，而是粉丝人数；许多“闹剧”（比如抢购板蓝根、食盐、微博约架等）粉墨登场，不是凭事实真伪，而是大家都这么做……

法国著名社会心理学家、群体心理学创始人、有“群体社会的马雅雅利”之称的古斯塔夫·勒庞（1841—1931年）在20世纪创作的《乌黑之众》中认为，构成群体的人必须具备两个特点：首先是个性消失，其次大家的感情与思想都在关注同一件事。这与当下网络上的“群”和生活中的“圈”非常相似。一些偶然出现在同一个场合或大街上往来穿梭的人，称不上心理学意义上的群体。

在群体心理中，由于原本突出的才智被削弱了，表现出差异的异质化被同质化吞没了，最终，群体只有很普通的品质和智慧。故群体只能从事最低级的劳动，无法完成对智力要求较高的工作。但假如把不计名利、绝对服从、勇于献身等真实或虚假的理想称为美德，那群体必定是最具备这种美德的人。

因此，勒庞提醒大家注意6点：在涉及智慧方面，绝不能依赖群体；群体比个人确实更有力量，但表现极不稳定的，所以要注意发挥群体中每个人的智力和能力；群体行动是受感情激起并主导，其强弱程度，直接决定着群体的行为能力；群体表现得比个体好或差，完全取决于周边的环境；群体能做出什么，取决于影响群体的暗示具有何种性质；群体更多的表现是一个英雄主义的群体，但也是引发冲突、构成骚乱的“祸胎”。

为什么许多智商高、情商好的人在群体中也沦落为野蛮的玩呢？勒庞认为：由于群体是无名氏，而无名氏不需要为所做的任何事情承担责任。独处时，我们知道是非善恶，也知道趋利避害。一旦进入群体特别是受到某种暗示后，理性思考就无影无踪了，甚至喜欢干一些事后非常懊悔之事。



大数据，即将成为石油、矿产类的新能源、新生产要素和巨大的经济资产，开启一次重大的时代转型，预示着新一波生产率增长和消费者赢余浪潮的到来。它是管理国家的手段、商务经营的法宝、未来求职的热门，可能还是发达国家新一轮全球化竞争中的利器。

用“大数据”的眼光看世界看未来

张稚丹

间甚至购买日的天气分析后，研究人员发现，跟尿布一起搭配购买最多的竟是啤酒，而季节性飓风到来之前，POP-Tarts蛋挞被大量买走。于是沃尔玛把啤酒和尿布捆绑销售；在飓风警报发出后，将蛋挞和手电筒等放在一起。过去，总部人员需要先有想法，再收集数据来验证；如今，他们可以预测到当A出现，B即将出现，这已经足够珍贵。他们不再探求难以捉摸的因果关系，转而关注事物的相关关系。

搜集数据是否涉及隐私

一天早晨，警察冲进霍华德·马科斯的家里，后者正打算用剪刀刺杀他的妻子，因为他发现妻子给他戴了绿帽子。警察开始控制霍华德，霍华德大喊冤枉：“我什么都没做啊！”这是电影《少数派报告》中的场景。不受限制的大数据分析可能也会导致这样的场景：罪责的判定是基于对个人未来行为的预测。

这场即将到来的数据革命，将给企业、国家的发展模式带来前所未有的革新与挑战，必须从战略高度来认识。美国总统科学技术顾问委员会主席霍尔德伦说，像美国历史上对超级计算和互联网的投资一样，大数据计划将对美国的创新、科研、教育和国防产生深远的影响。美国的每一项立法和计划，都有一个数据库和信息管理系统与之对应。2012年3月，美国宣布将投入2亿美元启动“大数据发展研究计划”，以推动大数据的提取、存储、分析、共享和可视化。通用电气也将投资15亿美元在旧金山建立一个全球软件和分析中心，拟雇佣科学家400名。如同工业革命要开放物质交易、流通一样，开放、流通的数据是时代趋势的需要。

但滥用大数据也会带来危险，当零星存在的数据被汇总，危机就出现了——不仅仅是隐私的泄露，还包括了被预知的可能性——这些能预测我们可能生病、拖欠还款和犯罪的算法会让我们无法购买医疗保险、无法贷款，甚至在实施犯罪前被预先逮捕。过于依赖数据，我们同样会受到局限：因为数据量过于庞大，做出决策的将是机器而不是人类。

《大数据时代》一书的译者、电子科技大学教授周涛提醒说，大数据也不是能解决所有问题的万能法宝。哈佛大学访问学者冯煦明指出：大数据之于传统经济统计，是补充，而非替代。基于抽样、调查、汇总等程序获得的数据仍将在经济分析和政策制定中发挥重要的作用。横向来看，传统统计方法在经济增长、税收、贸易、收入分配等领域的统计上具有主导优势，而大数据在物价、通货膨胀、失业率、消费等方面的统计上更具有优势。

总之，这本书以丰富的事例生动展现了大数据的光芒如何照亮整个世界，而严谨平实的叙述框架又令人了解到大数据各个技术性方面。对大数据的了解令我们对未来增添了憧憬和信心，难怪宽带资本董事长田溯宁称它是“我看到的最好的大数据著作”。

用生态思想阐释社会经济

南开大学出版社近期出版了世界生态社会科学协会会长金建方的《社会生态通论》(第二版)。该书首次提出了“生态体”概念——即在一定空间范围内，所有生物物质和非生物物质，通过能量流动和物质循环过程形成彼此关联、相互作用的统一整体，“生态体”有三个层次：人体生态体、社会生态体和自然生态体，金建方研究了生态体法则、运行规律、运行机制，并由此步步深入，向生态社会、生态经济、生态货币延伸，展示了未来生态社会文明的魅力所在。（谷亚光）



老北京景象

对于一座城而言，文化就是它的灵魂。爱上一座城，必须真正身处其中，与它同呼吸、共命运。怀着这样的初衷，2005年8月8日，获得古根海姆奖、纽约市公共图书馆奖、怀享奖、洛克菲勒·白拉及尔奖等多个写作奖项的美国学者迈克尔·麦尔，在北京大栅栏的杨梅竹斜街安家落户。

此后两年，迈克尔亲身体会到四合院拥挤嘈杂、没有隐私、冬冷夏热、卫生堪忧的各种不便。同时，他吃着大娘包的饺子，喝着燕京啤酒，和大伙儿一起侃历史聊未来，倾听居民的各种烦恼，帮助他们解决拆迁的难题。为了和北京做比较，迈克尔还实地考察了河内、老挝、丽江的发展方式，想方设法去天津和冯骥才讨论文化遗产保护，坐通宵火车去平遥古城观察当地

开春以来，新书旧书，胡乱翻看。《蒋介石与现代中国》，中信出版社。作者陶涵（Jay Taylor），在哈佛兼职的国务院职业外交官，曾派驻台北多年，与台湾政界高层往来颇密。此书内容很不均衡，至少一半笔墨落在蒋统治下的中美、台美关系上，这也许是本书的学术价值所在。

冷战后美国国务院的中国问题专家，政治立场与二战时期的谢伟思、范宣德、戴维斯“三约翰”相比，左右倾向差异很大。作者的政治立场鲜明，使本书的学术色彩明显逊色于哈佛费正清中心其他纯粹学者的著作。从某种程度上讲，中国大陆近些年对蒋介石生平事迹的研究，如杨天石先生的著作，也比本书内容更有深度和广度。

陶涵声称，虽受台湾有关机构的资助和委托写作此书，但他事先表明将不为蒋隐恶扬善。可惜阅读此书的感受与此相悖。历史上蒋介石的任何错误和罪行，都被作者以充满同情和理解的态度对待，并给以圆场。

对大陆读者来说，本书最主要的吸引力，是想知道蒋从大陆撤到台湾后都干了些什么、是怎么干的。该书对蒋来台后所推行政策，概括评判多，具体描述分析少。或许陶涵另一本被坊间热推的《蒋经国传》，这方面欠缺稍少？

需要特别提到的是，此书译笔流畅。译者是台湾的大学教授。最近看了几本佶屈聱牙的译著，着实伤着了。

《美国与美国人》，费孝通著，三联书店1985年版。凡购于上世纪80年代的书，都盖有一方本人藏书印。本书无印，夺自他人？何时拥有？全无印象。辑录费孝通分别写于1945年、1947年、1980年的三本小册子。翻了一遍，感觉还是早年写的那两本好看点儿。联想到林达夫妇的“近距离看美国”系列，给人看美国的西洋景，景深是越来越大了。当年费孝通给国人展示的美国社会景象，尽管有其社会学视角，仍有隔山买牛之感。毕竟不像今天，中国与美国挨得这么近。

《美国与中国》第4版，费正清著，商务印书馆1989年版。再读，益发觉得费正清这本奠定其中国学泰斗地位的小书，不论学术性、概括性、水准，都比本尼迪克特的《菊与刀》要胜出不可道里许。二战和朝鲜战争时，两书都曾分别被美国军政高层指定为了解对手的必读功课，可怎么看也觉不出《菊与刀》经典性何在。美国人真的那么重视这本书吗？

去年底与郭毅兄结伴去上海，往返高铁上读完他带的基辛格《论中国》。读费正清这本书，就看出基辛格谈中国历史文化，明摆着是炒他人冷饭。而对当下中国的分析，尤其算是美国政界一家之言，而且是有被边缘化趋势的一家之言。

让知识的阳光照亮每一个人

近日，中国残联和国家图书馆联合主办的图书馆文化助残系列活动在国家图书馆启动。国图副馆长陈力代表主办方宣读倡议书，呼吁全国各级各类图书馆行动起来，深入开展面向残疾人的无障碍、针对性、多样化信息服务，构建覆盖全国的残疾人知识援助协作网，缓解并逐步消除残疾人的知识贫困，使残疾人能够平等、充分地参与社会生活。

国家图书馆与中国残联共同创立了“文津·阳光讲坛”，北京师范大学文学院教授郭英德作了题为《中国四大小说名著的文化价值》的首场讲座。

国家图书馆长期致力于残疾人知识援助工作，除为残疾人营造无障碍图书馆环境、划定盲人阅读区域、设数字加油站帮残疾人将书拷入U盘带回家阅览外，2008年还与残联合作，共同建设中国盲人数字图书馆，根据盲人的偏好和需求，每年选择1000本图书进行数字加工，盲人在家也可以登录数字图书馆，通过阳光阅读软件“听书”。让社会弱势群体能够普遍均等地享有公共文化服务。（谢万幸）

神州

生态环境……得到了真实的第一手体验。同样以书写当代中国而闻名的美国作家彼得·海斯勒这样评价他：“很少有作者能够真正活在一部作品里，融入当地的生活，并让这种探究走向深处。在当今的英语写作圈，没有人比他更懂这个世界。”

今年4月，迈克尔的《再会，老北京》由上海译文出版社出版，其副标题“一座转型的城，一段正在消逝的老街生活”，正是该书的主旨。“无形巨手”在胡同里穿梭来去，悄悄将一个“拆”画上四合院古老斑驳的墙。一条条胡同逐渐被大型超市、高层公寓楼和宽阔的街道所代替。不久前你还去吃过老字号饭馆、逛过的热闹街市、造访过的温馨社区，可能在短短几周内被夷为平地……新建的是房子，但不一定是家园。它丢失了个性和灵魂。这不能不令人想起陈占祥先生在1955年说过的话：“一个城市古老的东西不是凭空而来的，是生长起来的，要拆了再得到，谈何容易！”任何大城市的形成，从根本上来说，都是相应的都市社会创造有非凡意义的都市环境的过程。在这个过程中，通过长达几百上千年的人类居住形成都市风光，这风光深深影响着这个社会。毁掉老建筑带来的损失，不仅仅是肉眼所看见的拆毁古老砖石那么简单。

迈克尔在书中明确说明：“我不想做摇旗呐喊，喊着‘不要拆’的人。”并不是“拆，还是不拆”，而是如何拆，怎样拆的问题。中国建设和城市建设的危机，实际上是文化灵魂的失落。文化保护是一个长期的缓慢的过程，而摧枯拉朽的经济建设能够立刻看到效益。二者该如何兼顾，形成可持续发展？这不仅仅是城市规划的课题。

古老的印第安人告诫后人：“慢一点，让灵魂跟上来。”一个忙于前进、前进、前进的民族，在速度奔跑的同时，要记得带上灵魂。

慢一点，让灵魂跟上来

林 颐