**《大数据时代》中的案例让我看到了数据分析的广阔应用领域**

**早该帮 编制**

《大数据时代》这本书主要描述的是大数据时代到临人们生活、工作与思维各方面所遇到的重大变革。

文中清晰的阐述了大数据的基本概念和特点，并列出明确的观点。不管对于产业实践者，还是对于政府和公众机构，都非常具有价值。作者将本书分为3个部分。第一部分提出了大数据时代处理数据理念上的三大转变：抽样等于全体；要效率不要绝对精确；要相关不要因果；第二部分作者从万事万物数据化和数据交叉复用的巨大价值两个方面，讲述驱动大数据战车在材质和智力方面向前滚动的最根本动力；最后一部分，作者描绘了大数据帝国前夜的脆弱和不安，包括产业生态环境、数据安全隐私、信息公正公开等问题。

本书观点掷地有声，作者观念高屋建瓴，从很多实例和经验中萃取普适性观念。例子详实丰富，囊括了进百个学术和商业实例。

引言提出了大数据将给生活、工作于思维带来重大的变革。一个例子是20xx年H1N1流行病毒背景下谷歌通过检测检索词条，处理了4.5亿个不同的数据模型，通过预测并与20xx年、20xx年美国疾控中心记录的实际流感病例进行对比后，确定了45条检索词条组合，并将其用于一个特定的数学模型后，预测的结果与官方数据的相关系数高达97%。按照传统的信息返回流程，通告新流感病毒病例将有一到两周的延迟。对于飞速传播的疾病，信息滞后两周是致命的。而谷歌运用大数据技术，以前所未有的方式，通过海量数据分析得出流感所传播的范围，为世界预测流感提供了一种更快捷的预测工具。此外，我联想到原淘宝董事长马云通过大量数据分析得出20xx年经济疲弱，为其商家提前做好迎接经济危机提供了时间缓冲。(补充并清晰描述详细)关于大数据在商业领域的应用，Farecast公司是一个成功的典型范例。该公司由奥伦·埃齐奥尼创办，利用机票的销售数据来预测未来的机票价格，旨在帮助用户在购买机票方面做出预测，并对机票价格走势预测的可信度标示出来供消费者查考。Farecast系统利用近十万亿条价格记录预测的准确度达75%，使得使用Farecast票价预测工具购买机票的旅客，平均每张机票节约50美元。而处理如此多的数据离开了大数据技术将无法进行。

也正是由于我们进入了一个前所未有的信息化时代，人们拥有了如此多的数据，才提供给我们利用大数据的分析处理手段，创造新的价值。也许有人以为我们大数据时代的还未来临。其实大数据技术早已渗透到我们中间，它被应用在垃圾邮件的过滤，新浪微博技术平台，谷歌翻译以及输入文字的自动纠错等。

文中提出的一个观点是，预测是大数据的核心。其实从过去的时代人们就利用掌握的数据进行各种分析，从而对经济等各方面进行预测、矫正。只是进入了大数据时代人们掌握的数据爆炸性的速度在增长，从而数据的存储和分析数据分方法成了释放大数据能量的关键。

关于不是随机样本而是整体数据中。作者指出了随机取样是小数据时代用最少的数据获取最大价值的做法。作者用大数据与乔布斯的癌症治疗例子说明了使用全部数据而非样本的意义。乔布斯成为世界上第一个对自身所有DNA和肿瘤DNA进行排序的人。乔布斯曾开玩笑说我要么是第一个通过这种方式战胜癌症的人，要么就是最后一个因为这种方式死于癌症的人。虽然最后难免死于癌症但这种获得所有数据而不是仅样本的方法将他的生命延长了几年。同样，从事跨境汇款业务的Xoom公司侦破一起犯罪集团的诈骗也是由于使用了整体数据。初此之外，他还列举了日本相扑等来证明使用全体数据的重要性。

作者同时也指出随着数据使用的越来越多，其得出的结果并一定能越来越精确，毕竟数据不能保证百分之百的正确，特别是大数据时代各种结构化与非结构化类型的数据聚集在一起难免导致结果的不太精确。大数据时代要求我们重新审视精确性的优劣。作者特别举了谷歌翻译成功的例子。谷歌翻译之所以优于IBM的Candide系统并不是因为它拥有更好的算法机制。和微软的班科和布里尔一样，谷歌翻译增加了各种各样的数据，并且接受了有错误的数据。(其语库来自于未经过滤的网页内容，会包含一些不完整的句子、拼写错误、语法错误以及其他各种错误)

在不是因果关系，而是相关关系的篇章中。作者指出在大数据时代往往知道是什么要比知道为什么来的更实在。作者列举了林登的XX推荐系统的成功，证实了大数据在分析相关性方面的优势以及在销售中获得的成功。XX也是充分利用并挖掘各类数据信息的先锋和代表，从以前广为人事的啤酒和尿布的案例，以及作者举的有关蛋挞和飓风天气的案例，都说明了掌握了相关关系对于其策略的帮助。建立在相关关系分析法基础上的预测是大数据的核心。XX保险公司利用几百种生活方式的数据，如爱好、长浏览网页等间接的预测出哪些人更可能患高血压、糖尿病和抑郁症。XX国家快递公司通过使用预测性分析检测其全美6万辆车队。进行防御性的修理，节约巨大得的成本。这些都充分显示了大数据在预测方面的优势。

本书第二部分讲的是大数据时代的商业变革。

作者用莫里绘制导航图的例子告诉我们，远在信息数字化之前，对数据的运用就已经开始了。莫里利用大量的人力去分析多年保存的航海记录，他从这些大量的数据中获取到新的利用价值。绘制的图表帮助商人节约一大笔钱，使年轻的海员们间接获取了成千上万名经验丰富的航海家的指导。日本先进工业技术研究所越水重臣教授通过安装压力传感器将人屁股特征数据化，进而形成对乘客身份的特征识别。这项技术为汽车防盗系统提供了方案。公司，致力于为顾客预测商品的价格，通过收集处理海量的价格信息，预测准确率高达77%，帮助顾客在购买一个产品时节约了大约100美元。MasterCardd.Advisor部门通过分析来自210个国家的15亿信用卡用户的650亿条交易记录，分析得出商业发展和客户消费趋势，如通过分析发现如果一个人下午四点左右给汽车加油的话，他很可能在接下来的一个小时内去购物或者去餐馆吃饭，且在这一小时里大约花费35到40美元。商家正可以利用这个分析结果，在加油的小票背面附加上附近商店的优惠券。

这些例子都证明了大数据蕴藏着巨大的商业价值。根据提供价值的不同来源，大数据价值链包括三大构成部分。包括第一种是基于数据本身的公司。这些公司拥有大量数据或者至少可以收集到大量数据，却不一定有从数据中提取价值或者用数据催生创新思想的技能。第二种是基于技能的公司。它们通常是咨询公司、技术供应商或者分析公司。它们掌握了专业技能但并不一定拥有数据或者提出数据创性用途的才能。比如说，XX和Pop-Tarts这两个零售商就是借助天睿公司的分析来获得营销点子，天睿就是一家大数据分析公司。第三种是基于思维的公司。皮特.华登，Jetpac的联合创始人，就是通过想法获得价值的一个例子，他通过用户分享到网上的旅行照片来为人们推荐下一次旅行目的地。对于某些公司来说，数据和技能并不是成功的关键。挖掘数据的新价值的创新思维才是这些公司脱颖而出的优势所在。

大数据成为许多公司竞争力的来源，未来可能整个行业的结构会发生改变，大公司和小公司最有可能成为赢家。如今的核心竞争力在于快速而廉价地进行大量的数据存储和处理。当然公司要根据自己的情况进行调整。大数据向小数据时代的赢家以及那些线下大公司(如XX、联邦快递、宝洁公司、雀巢公司、波音公司)提出了挑战。同时，大数据也为小公司带来了机遇。大数据也将会影响国家竞争力。当制造业已经大幅转向发展中国家，而大家都争相发展创新行业的时候，工业化国家因为掌握了数据以及大数据技术，所以仍然在全球竞争中占据优势，但这个优势很难持续。随着技术的发展，西方世界在大数据技术的优势将会慢慢消失。对于大公司而言，好消息是大数据技术可以加剧优胜劣汰。一旦公司掌握了大数据，它不但可能超过对手还可能遥遥领先。

文章第三部分讲了大数据带来无数好处的同时带来的不良影响以及如何面对这些影响。包括如数据的收益的处理问题以及数据中用户资料的隐私和决策过程带来的影响。作者在保护个人隐私方面提出了几种想法。一种是使用数据时征询数据所有个人的知晓和授权。第二个技术途径就是匿名化。作者同时也指出了这两种方式的难度。一方面收集到的数据可能会被后续的多次利用。另一方面，匿名化会在数据收集越来越多和数据的相互结合关联使用时变得无效。作者列列举电影《少数派报告》的情节说明越来越依赖数据时，大数据可能将我们禁锢在可能性之中。当然通过分析犯罪的常发地与常发时间，合理安排警力会对治安防范提供不小的帮助。作者还指出不能尽信数据的分析结果，因为不能保证获取分析结果来源的数据准确性。大数据在给我们生活提供便利的同时，也让隐私保护的法律手段失去了作用。我们必须杜绝对数据的过分依赖。

在高速迈进大数据时代的同时，人类信息管理准则需要重新定位，这将带动社会核心价值观的转变。大数据时代，对原有规范的修修补补已经不足以抑制大数据带来的风险。保护个人隐私就需要对个人数据处理器对其政策和行为承担更多责任。同时必须重新定义公正的概念，以确保人类行为的自由。作者提出了解决这些问题的方向。如个人隐私保护方面，可以让使用者承担更多的社会责任。将责任从民众转移到数据使用者有很多意义，也有充分的理由。因为他们更清楚将如何使用数据且是数据应用最大的受益者。关于公正方面简单的讲就是个人可以并应为他们的行为而非倾向负责。就像公司有内部会计和外部审计人员一样，大数据时代，公司将设置专门的人员--内部和外部算法师对大数据活动进行监督。还有可能出现第三方的机构对大数据行为进行监督和衡量。作者甚至考虑到对大数据存在的垄断情况进行分析并在反垄断反面给了建议。最后结语中作者提出大数据提供给人们的只是参考答案，提醒我们在利用这个工具时要铭记人类的作用是无法完全替代的。

大数据时代是信息化社会发展必然趋势，我们只有紧紧跟随时代发展的潮流，在技术上、制度上、价值观念上做出迅速调整并牢牢跟进，才能在接下来新一轮的国际竞争中摆脱受制于人的弱势境地，才能把握发展的方向，冲破与西方国家的差距。对于一个国家如此，对于一个企业亦是如此。在如此快速的到来的大数据时代，我们还有很多知识需要学习，许多思维需要转变，许多技术需要研究。公司的规划中，也需充分考虑到大数据对于公司的未来发展所带来的机遇和挑战。对于掌握大量数据的公司，需要考虑有多少数字化的数据，又有哪些可以通过大数据的分析处理而带来有价值的用途？比如国内目前的社交网站，购物网站等都掌握了用户的大量的数据信息。在大数据时代制胜的良药也许是创新的点子，也许可以利用外部的数据，通过多维化、多层面的分析给其他企业或个人带来价值。

关键词：大、数据、时代、中、的、案例、让、我、看到、了、数据、分析、数据分析、的、广阔、应用、领域、应用领域

参考文献：[1]早该帮https://bang.zaogai.com/item/BPS-ITEM-33025.html