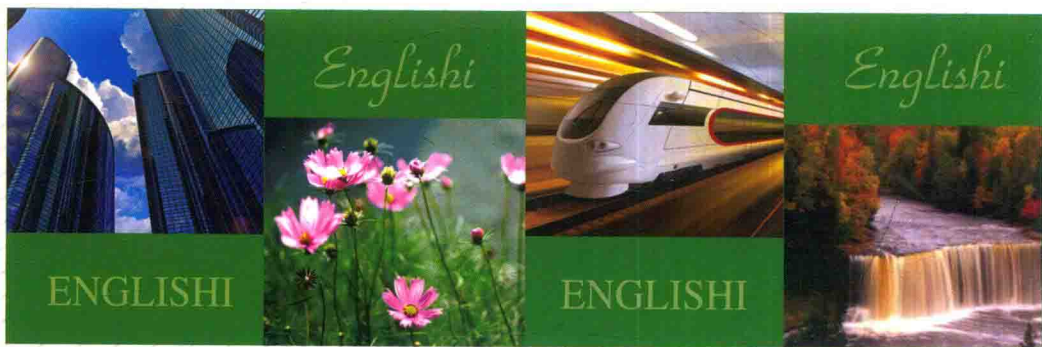




大数据时代的 高等外语教育创新与实践

● 王 鹤◎著



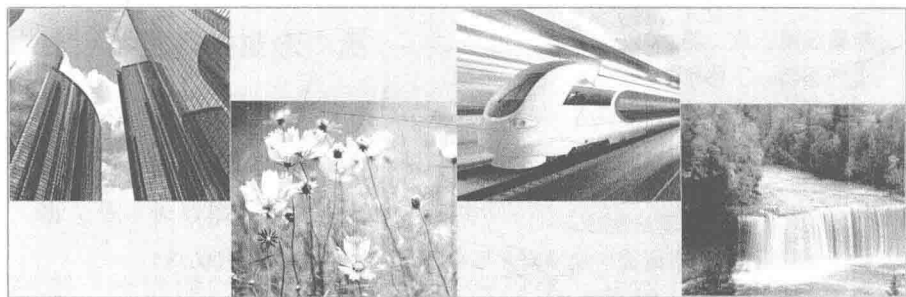
THE INNOVATION AND PRACTICE
OF FOREIGN LANGUAGE HIGHER EDUCATION
IN BIG DATA ERA



经济管理出版社
ECONOMY & MANAGEMENT PUBLISHING HOUSE

大数据时代的 高等外语教育创新与实践

● 王 鹤◎著



THE INNOVATION AND PRACTICE
OF FOREIGN LANGUAGE HIGHER EDUCATION
IN BIG DATA ERA

图书在版编目 (CIP) 数据

大数据时代的高等外语教育创新与实践/王鹤著. —北京: 经济管理出版社, 2016. 9
ISBN 978 - 7 - 5096 - 4518 - 5

I. ①大… II. ①王… III. ①高等教育—研究 IV. ①G64

中国版本图书馆 CIP 数据核字 (2016) 第 168941 号

组稿编辑: 张 艳
责任编辑: 王格格
责任印制: 黄章平
责任校对: 雨 千

出版发行: 经济管理出版社
(北京市海淀区北蜂窝 8 号中雅大厦 A 座 11 层 100038)

网 址: www.E-mp.com.cn
电 话: (010) 51915602
印 刷: 北京九州迅驰传媒文化有限公司
经 销: 新华书店
开 本: 720mm × 1000mm/16
印 张: 15.5
字 数: 238 千字
版 次: 2016 年 9 月第 1 版 2016 年 9 月第 1 次印刷
书 号: ISBN 978 - 7 - 5096 - 4518 - 5
定 价: 49.00 元

· 版权所有 翻印必究 ·

凡购本社图书, 如有印装错误, 由本社读者服务部负责调换。

联系地址: 北京阜外月坛北小街 2 号

电话: (010) 68022974 邮编: 100836

目 录

第一章 导论	1
第一节 大数据的社会价值	1
第二节 大数据的特性分析	4
第二章 大数据的概念内涵	7
第一节 大数据的基本概念	7
第二节 大数据的内涵探究	14
第三节 大数据的主体分类	19
第三章 大数据时代的高等外语教育概述	23
第一节 大数据促动教育变革	23
第二节 大数据与高等教育的关系及作用机制	28
第三节 大数据对高等教育发展的推动	32
第四章 高等外语学习理论及观念	43
第一节 行为主义学习理论	43
第二节 认知主义学习理论	47
第三节 建构主义学习理论	56
第四节 人本主义学习理论	67

第五章 大数据时代的外语学习理论——关联主义学习理论	75
第一节 关联主义学习理论的背景及相关研究	75
第二节 关联主义学习理论的概述	85
第三节 关联主义理论的主要观点	95
第六章 大数据时代多元外语教育模式及构建方法	117
第一节 Web 环境与外语教育	117
第二节 大数据时代的“智慧教育”	133
第三节 外语教育的网络交互模态探讨	139
第四节 基于“普适计算”的外语教育模式	147
第五节 云计算基础上的外语教育模式	153
第七章 大数据时代外语教育的实现途径	167
第一节 信息技术与外语深度融合的学科教学体系建构	167
第二节 大数据学习数据分析与外语教学	175
第三节 “慕课”与外语教育	183
第四节 “翻转课堂”与外语教育	189
第五节 “微课程”与外语教育	198
第八章 大数据在高等外语教育中的问题及策略	203
第一节 大数据在高等外语教育中的问题	203
第二节 提高大数据推动高等教育发展的对策建议	215
参考文献	229

第一章 导论

第一节 大数据的社会价值

大数据 (Big Data) 一词最早出现在 apache org 的开源项目 Nutch 中, 当时科学家用大数据这个术语描述在更新网络搜索索引的同时进行批量处理或分析的大量数据集。1980 年, 当代著名思想家阿尔文·托夫勒在其《第三次浪潮》中将大数据描绘为信息社会的重要篇章。人们对于数据的海量挖掘和大量运用, 不仅标志着产业生产率的增长和消费者的大量盈余, 而且也明确地提示着大数据时代已经到来。因此, 大数据正在影响着社会和科技的每个角落层面, 成为不可忽视的重大问题。

一、大数据重组战略资源

2009 年, 大数据开始在互联网领域及信息技术行业被广泛应用。美国互联网数据中心指出, 互联网中的数据每年增长 50%, 两年就会翻一番, 而目前世界上 90% 以上的数据被创造出来是近两年的事情。美国麦肯锡全球研究院 2011 年 6 月对大数据做出一份报告, 即《大数据: 下一个创新、竞争和生产力的前沿》, 指出数据是新时期的基础生活资料与市场要素, 重要程度不亚于物质资产和人力资本, 大数据将成为企业提高生产力和竞争力的主要方式与关键要素。数据成为资产、产业垂直整合、泛互联网化是大数据时代的

三大发展趋势。

奥巴马政府更是将大数据战略提升为国家战略。奥巴马政府为了进一步推动大数据相关产业发展,在2012年宣布投资2亿美元,将大数据作为未来的新资源,像石油一样重要,一个国家拥有的数据规模及运用的能力将成为综合国力的重要组成部分,对数据的占有权和控制权将成为陆权、海权、空权之外的国家核心权力。数据成为和土地、资本、人力同样重要的关键生产要素。

随着信息技术的飞速发展,人类社会进入数字信息时代。获取和掌握信息的能力已经成为衡量一个国家实力强弱的标志。海量数据又随时间持续产生、不断流动,进而扩散形成大数据,进而重组各类战略资源。只有经过分析大数据才能变成满足不同需求有价值的信息,才能使政府适应大数据时代重组战略资源格局的要求。

二、大数据推动社会变革

不断快速发展的大数据已成为智慧城市的核心资源,这成为大数据推动社会变革的重要表现。从我国整体投资规模来看,到2012年底全国开建智慧城市的数量超过180个,通信网络和数据平台等基础设施建设投资规模接近5000亿元,“十二五”期间智慧城市建设将成为大规模投资的重点,仅设备等基础投资将达1万亿元。建设智慧城市各方面各层级的决策都离不开大数据支持。

大数据引领城市变革体现在很多方面。城市规划需要挖掘自然地理与社会人文的大量信息,包括经济、人口、公共服务、气象等一切能够支持决策提高管理服务水平的海量数据,既要重视科学性,又要把握前瞻性。交通管理也需要全面采集各种交通方式的数据,实时挖掘及时整理,改善交通治理拥堵,集中处理及时响应,掌控突发状况,维护城市交通全天候运行。

大数据还影响到互联网监控和安全。一是舆情掌控需要及时发现搜索网络电台及智能分析网络关键词语义,加强舆情分析能力,及时全面了解公共议题,善于处理网络突发事件,对于网络犯罪依法处置。二是安全防灾离不

开采集相关大数据，挖掘大数据，第一时间预测自然灾害和社会恐怖事件，具备预警防范和应急抢险能力。随着社会发展，政府应该有勇气将手中的数据逐步开放，供给更多有能力的机构、组织或个人来分析并加以利用，以加速造福人类。

三、大数据引发思维创新

近年来随着 Map Reduce 和 Google File System (GFS) 等程序的问世，不仅用来描述数据的量非常巨大，还突出强调了处理数据的速度。所以大数据成为数据分析领域的前沿技术。数据成为当今每个行业和商业领域的重要因素。这也一定意义上标志着大数据正在引发一场思维革命，大数据正在改变人们考察世界的方式方法，以前所未有的速度引起社会、经济、学术、科研、国防、军事等领域的深刻变革，更引起思维领域随之不断创新发展。

思维创新体现在大数据的实用功能。大数据与人类息息相关，越来越多的问题可以通过大数据解决。不仅在数据科学与技术层次，而且在商业模式、产业格局、生态价值与教育层面，大数据都能带来新理念和新思维，包括政府宏观部门以及不同的产业界与学术界，甚至个人消费者。大数据与互联网一样，不仅是信息技术领域的革命，更加速了企业创新，在全球范围引领社会变革并启动透明政府的发展。

思维创新也体现在大数据产生的崭新理念上。数据技术变革能创造无限商机，促进行政管理和商业管理革新，推动政府信息公开、透明和社会公正。大数据时代随着信息技术的发展已经来临，创造、生产和分享的理念迅速传播。世界主要发达国家、国际组织和知名企业相继开展了大数据战略研究以应对挑战。

第二节 大数据的特性分析

一、大数据的主体性分析

大数据主体性主要是指在大数据时代数据作为一种独立存在的实体,其资产价值越来越突出,日益引起人们的重视。从具体的个人到形形色色的企业,从各国政府到各种组织都可以合法地去收集数据。例如,《经济学人》杂志数据编辑肯尼思·丘基尔认为,无论个人还是企业以及政府等都可以是数据的拥有者。个人使用全球定位系统的定位数据时,个人成为移动运营商收集大数据的来源,移动通信公司通过向用户提供更好的服务投资并收集大数据,政府可拥有如人口普查数据、城市规划、天气信息等特定的数据。

荷兰阿姆斯特丹自由大学计算机系高级研究员黄智生指出,今后个人隐私与数据归属权可能关系越来越少,欧洲民众要求政府信息公开的诉求极其强烈,民众有权向政府申请信息公开。除了涉及国家安全和个人隐私的公共信息外,欧盟和欧洲各国大部分政府信息已经公开。

二、大数据的社会性分析

大数据的社会性不仅指企业重视的新型资产,而且还指其有助于在社会公共领域解决大量问题。大数据无论在失业、教育还是医疗保健方面,将带来无穷无尽的社会效益和经济效益。牛津大学网络学院互联网治理与监管专业教授维克托·迈尔·舍恩伯格认为大数据可以促进信息社会发展,增进社会交流,提供人性化服务,支持做出更好的选择,随着未来社会更加信息化,我们将拥有更方便的科技设施和更先进的工具。

纽约市打击金融犯罪行动组主管迈克·弗劳尔斯认为,纽约市充分运用

大数据有助于市政执法、管理经济、防灾规划和灾后恢复，有助于优化行政资源、预防犯罪、使公共支出产生最大效益。工业和信息化部赛迪智库软件与信息服务业研究所所长安晖总结，大数据不仅能够帮助企业发展，获得经济回报，还可以促进社会管理水平提高，进而增进安全保障能力。

三、大数据的规范性分析

大数据的规范性是指数据流通需要经法律界定数据的权益，政府界定数据的类型，判断是否涉及国家安全等，做到有法可依。在确定数据权益的前提下，数据运用需要价值交换，数据可以有偿使用。

荷兰阿姆斯特丹自由大学计算机系高级研究员黄智生认为，对非法泄露他人隐私的行为，法律会进行严厉打击。牛津大学网络学院互联网治理与监管专业教授维克托·迈尔·舍恩伯格认为，需要设立数据使用者承担其行为责任的隐私保护模式。《经济学人》数据编辑肯尼思·丘基尔认为，关键在于大数据技术规范应该从强调监管大数据的收集转向重点监管大数据的实际使用。

复旦大学数据科学研究中心主任朱扬勇教授认为，数据不是自然就有的，当产生的数据涉及商业秘密和个人隐私时需要提供法律界定。中国人民大学信息院院长杜小勇认为，数据作为资产需要数据治理，应从法律上明确规定将窃取他人数据与攻击他人信息系统视为犯罪，开展国家间对话形成公约。

四、大数据的技术性分析

数据总量在最近增长态势惊人，全球在过去3年间产生的所有数据量已经超过以往总和。物联网的加速膨胀和移动互联网的迅速发展都使得新数据源大大增加，大量传感器不断产生新数据，获取、存储和处理数据的成本下降推动了数据持续大量膨胀产生。

工业和信息化部赛迪智库软件与信息服务业研究所所长安晖指出，信息技术以前只关注数据的体量与规模，而现在数据爆发，数据处理变得困难，

应用需求加大，大数据不仅数据的规模超过以往，更要求数据处理和应用的能力和速度提高，统一整合数据对象、技术与应用。

其他类似观点也是认为大数据不仅可以是如政府部门或企业掌握的数据库这种有限数据集合，也可以是如微博、微信、社交网络上虚拟的无限数据集合。大数据技术是采集数据进行存储管理，并且分析挖掘通过可视化等技术呈现的一系列技术集成。大数据应用通过整合大数据技术，处理获得的大数据集合，追求信息产生的价值。

五、大数据的经济性分析

大数据对经济的影响远远超出预计，不仅可以创造高额利润，还能为社会提供大量就业途径，出现一些新职业。一些欧洲公司通过网上数据，开始提升环境保护和医疗保健等领域的发展，并且用于科学实验。

欧盟委员会负责数字议程的委员尼丽·克洛斯认为，大数据是经济发展的动力和燃料，可以创造机遇改造社会，是使服务更加个性化和透明高效的重要工具。英国经济与商业研究中心指出，大数据能够改变创新途径，提升生产效益，为英国带来的经济收入高达 2160 亿英镑。

中国人民大学信息院院长杜小勇教授认为，需要耗费大量资源来获取数据与记录数据，因此数据具有资产的属性。SAS 公司大中华区总裁吴辅世指出，大数据改变了企业对数据的认知，开始重视数据、利用数据，希望通过收集和分析大数据来获取竞争优势和商业价值，数据将成为其核心资产。

第二章 大数据的概念内涵

第一节 大数据的基本概念

近年来，大型传感器与各种数码设备飞速递增，企业数据库和社交媒体网站日益扩张，90%的数据来自过去两年。这导致了从市场营销服务人员到政府的政策制定者都逐渐了解大规模数据集和大数据等概念。但是，关于大数据的界定、内涵和特征等仍然意见纷呈。下面笔者从历史的角度对大数据的概念做一番梳理。

一、大数据的发展历史

有史以来，处理各种不断增长的数据都是人类社会的难题。大数据的现代发展历史最早可追溯到美国统计学家赫尔曼·霍尔瑞斯，他为了统计1890年的人口普查数据，发明了一台电动机器来对卡片进行识别，该机器用1年时间就完成了预计8年的工作，促进了全球进行数据处理的新起点。1943年，第二次世界大战期间，英国为了快速解开纳粹设置的密码，组织工程师发明机器进行大规模数据处理，并采用了第一台可编程的电子计算机实施计算工作。1961年，美国国家安全局（NSA）首先应用计算机收集信号自动处理情报，通过数字化处理模拟磁盘信息。20世纪60年代，英国计算机科学家蒂姆·伯纳斯·李设计超文本系统，命名为万维网，使用互联网在世界范

国内实现信息共享。1965年,英特尔创始人戈登·摩尔(Gordon Moore)通过研究计算机硬件得出摩尔定律,认为同等面积的芯片每过一到两年就可容纳两倍数量的晶体管,能够使微处理器的性能提高两倍,或使之价格下降一半。近50年来,信息产品功能日趋强大,各种设备体积变小,存储器成本持续缩小了1亿多倍,能以很低的成本保存海量的数据。1988年,美国科学家马克·韦泽(Mark Weiser)指出,各种各样微型计算设备能随时随地获取并处理数据,被称为普适计算。今天,智能手机、各种传感器、RFID(射频识别)标签、可穿戴式设备等实现了无处不在的数据自动采集,为大数据时代的到来提供了物理基础。美国研究员大卫·埃尔斯沃斯和迈克尔·考克斯在1997年使用“大数据”来描述超级计算机产生超出主存储器的海量信息、数据集甚至突破远程磁盘的承载能力。

大数据时代的技术基础集中表现在数据挖掘,通过特定的算法对大量的数据进行自动分析,从而揭示数据当中隐藏的规律和趋势,即在大量的数据当中发现新知识,为决策者提供参考。现在的信息技术已经可以把一件产品的流向、每位消费者的情况都记录下来,再通过数据挖掘,为客户量身定制,把消费和服务推向一个高度个性化的时代。基于网络数据的挖掘,不需要制定问卷,也不需要逐一调查,成本低廉。更重要的是,这种分析是实时的,没有滞后性,数据挖掘将成为越来越重要的分析预测工具,抽样技术将下降为辅助工具。数据挖掘的优越性,也集中反映了大数据“量大、多源、实时”的三个特点。大数据的前沿和热点是机器学习,和数据挖掘相比,其算法并不是固定的,而是带有自调适参数的,也就是说,它能够随着计算、挖掘次数的增多,不断自动调整自己算法的参数,使挖掘和预测的结果更为准确,即通过给机器“喂取”大量的数据,让机器可以像人一样通过学习逐步自我改善、提高,这也是该技术被命名为“机器学习”的原因。除了数据挖掘和机器学习,数据的分析、使用技术已经非常成熟,并且形成了一个谱系,例如数据仓库、多维联机分析处理(Multidimension OLAP)、数据可视化、内存分析(In-memory Analytics)都是其体系的重要组成部分。

从2004年起,以脸谱网(Facebook)、推特(Twitter)为代表的社交媒体相继问世,互联网开始成为人们实时互动、交流协同的载体,全世界的网

民都开始成为数据的生产者，引发了人类历史上迄今为止最庞大的数据爆炸。在社交媒体上产生的数据，大多是非结构化数据，处理更加困难。2012年，乔治敦大学的教授李塔鲁（Kalev Leetaru）考察了推特上产生的数据量，他做出估算：过去50年，《纽约时报》总共产生了30亿个单词的信息量，现在仅仅一天，推特上就产生了80亿个单词的信息量。也就是说，如今一天产生的数据总量相当于《纽约时报》100多年产生的数据总量。

回顾半个多世纪人类信息社会的历史，正是因为1966年提出的摩尔定律，晶体管越做越小、成本越来越低，才形成了大数据现象的物理基础。1989年兴起的数据挖掘技术，是让大数据产生“大价值”的关键；2004年出现的社交媒体，则把全世界每个人都变成了潜在的数据生成器，这是“大容量”形成的主要原因。

2008年末，“计算社区联盟”（Computing Community Consortium）提出了独特的详细报告：《大数据计算：在商务、科学和社会领域创建革命性突破》，使人们不仅考虑机器的数据处理，而且在更广泛的领域发现大数据的社会意义，找到了更多的新用途和富有创见的新见解。社会领域的计算，也被很多学者称为“社会计算”（Social Computing），社会领域的计算、对类似知识和关系的捕捉，不仅能够有效推动社会治理，还能产生商业价值。

总的来看，对处理大规模信息的现实需求推动了大数据相关技术的迅速发展，起初国家安全是大数据技术的主要推动力，伴随着超级计算机的发明，大数据的存储和处理技术以及大数据分析算法的研发，最终导致大数据在教育、金融、医疗等许多方面开始实施，广泛应用。美国2012年3月的政府报告里明确要求每个下属的联邦机构都要制订一个“大数据”发展战略，奥巴马政府首先宣布投资2亿美元，立刻启动大数据研究与发展项目。

二、大数据的概念界定

大数据这个术语最早应用于apache org的开源项目Nutch，用来表达批量处理或分析网络搜索索引产生的大量数据集。谷歌公开发布Map Reduce和Google File System（GFS）之后，大数据不仅包含数据的体量，而且强调数据

的处理速度。在数据分析领域，大数据是前沿技术，大数据以及数据仓库、数据分析、数据安全、数据挖掘是 IT 行业时下最火爆的词汇，大数据的商业价值已经成为信息行业争相追逐的焦点。大数据包括各种互联网信息，更包括各种交通工具、生产设备、工业器材上的传感器，随时随地进行测量，不间断传递着海量的信息数据。

利用新处理模式，大数据具有更强的决策力和洞察力，能够优化流程，实现高增长率，处理海量的多样化信息资产。归根结底，大数据技术可以快速处理不同种类的数据，从中获得有价值的信息，处理速度快，且只有快速才能起到实际用途。随着网络、传感器和服务器等硬件设施全面发展，大数据技术促使众多企业融合自身需求，创造出难以想象的经济效益，实现巨大的社会价值。各行各业利用大数据产生了极大增值和效益，表现出前所未有的社会能力，而绝不仅仅是数据本身。所以，大数据可以定义为在合理时间内采集大规模资料、处理成为帮助使用者更有效决策的社会过程。

（一）大数据的分类

大数据一般分为以下四类：互联网数据、科研数据、感知数据和企业数据。

互联网大数据尤其社交媒体是近年大数据的主要来源，大数据技术主要源于快速发展的国际互联网企业。比如以搜索著称的百度与谷歌的数据规模都已经达到上千 PB 的规模级别，而应用广泛、影响巨大的脸谱、亚马逊、雅虎、阿里巴巴的数据也已达上百 PB。

科研数据存在于拥有极高计算速度且性能优越的机器的研究机构，包括生物工程研究以及粒子对撞机或天文望远镜，例如位于欧洲的国际核子研究中心装备的大型强子对撞机，在其满负荷的工作状态下每秒就可以产生 PB 级的数据。

移动互联网时代普及 LBS，基于位置的服务和移动平台的感知功能，感知数据逐渐与互联网数据重叠，但感知数据的体量同样惊人，并且总量或许可能不亚于社交媒体。

企业数据种类繁多，企业同样可以通过物联网收集大量的感知数据，增

长极其迅猛，企业外部数据则日益吸纳社交媒体数据，内部数据不仅有结构化数据，更多的是越来越多的非结构化数据，由早期电子邮件和文档文本等扩展到社交媒体与感知数据，包括多种多样的音频、视频、图片、模拟信号等。

（二）大数据技术

大数据技术包括大数据科学、大数据工程和大数据应用。大数据科学指在大数据网络的快速发展和运营过程中寻找规律，验证大数据与社会活动之间的复杂关系；大数据工程指通过规划建设大数据并进行运营管理的整个系统。

大数据需要有效地处理大量数据，包括大规模并行处理（MPP）数据库、分布式文件系统、数据挖掘电网、云计算平台、分布式数据库、互联网和可扩展的存储系统。当前用于分析大数据的工具主要有开源与商用两个生态圈，开源大数据生态圈主要包括 Hadoop HDFS、Hadoop Map Reduce、HBase 等，商用大数据生态圈包括一体机数据库、数据仓库及数据集市。大量非结构化数据通过关系型数据库处理分析需要大量时间和资金，因为大型数据集分析需要大量电脑持续高效分配工作。大数据分析常和云计算联系在一起，大数据分析相比传统的数据仓库数据量大、查询分析复杂。

大数据处理和存储技术源于军事需求，“二战”期间英国研发了能处理大规模数据的机器，“二战”后美国致力于数字化处理搜集得到的大量情报信息。计算机和互联网技术导致大数据处理问题出现，“9·11”事件后美国政府在大数据挖掘领域组建了大数据库用于识别可疑人，通过筛选通信、教育、犯罪、医疗、金融和旅行等记录，组建了基于网络的信息共享系统。

大规模数据分析技术源于社交网络，大数据应用使人们的思维不局限于数据处理机器，重要的是新用途和新见解，对大规模信息的处理需求从根本上推动了大数据相关技术的发展，超级计算机的发明、大数据的存储和处理技术以及大数据分析算法的研发最终导致了教育、金融、医疗等多方面大数据的广泛应用。

三、大数据的基本特点

第一，体量巨大，种类繁多。互联网搜索的发展、电子商务交易平台的覆盖和微博等社交网站的兴起，产生了无穷无尽的各种数据内容。国际数据统计机构 IDC 估计，2011 年和 2012 年的全球信息总量分别达到 1.8ZB、2.8ZB，到 2020 年将是 40ZB；思科公司预测全世界 2016 年产生的数据总量将达到 1.3ZB；谷歌前 CEO 施密特指出，从人类文明开始到 2003 年的近万年时间里人类大约产生 5EB 数据，而 2010 年人类每两天就能产生 5EB 数据。传感、存储和网络等计算机科学领域在不断前行，人们在不同领域采集到的数据量达到了前所未有的程度，收集大量数据的原因在于网络数据可以实现同步实时收集，包括电子商务、传感器、智能手机等，还有医疗领域的临床数据和科学研究，例如基因组研究将 GB 级乃至 TB 级的数据输送到数据库。数据总量的增长由于占到 85% 以上的非结构化数据的增长，增速比结构化数据快大概几十倍。对于存储和网络企业的投资者来说这类预测能提升信心，美国咨询公司麦肯锡从个体数据集的大体量定义大数据是指传统数据库软件工具难以采集、存储、分析管理的巨大的数据集。大数据的数据类型日益繁多，例如视频、文字、图片、符号等各种信息，发掘这些形态各不相同的数据流之间的相关性是大数据的最大优点。比如供水系统数据与交通状况比较可以发现清晨洗浴和早高峰的时间密切相关，电网运行数据和堵车时间、地点有相关性，交通事故率关联睡眠质量。

第二，开放公开，容易获得。采集大数据不是为了存储而是为了进行分析。大数据不仅存在于特定的政府机构和企业组织，而且是社会生活生产过程中自动产生存储的。电信公司积累客户的电话沟通记录，电子商务网站整合消费者的各种信息，企业通过挖掘海量数据可以增强自身能力，改善运营服务，提供决策支持，实现商业智能进而为企业带来高额经济效益回报，发现企业发展的特殊规律。今天，在一定规则开放性下，依靠应用程序接口技术和爬虫采集技术，越来越多的商业组织和政府机构开始向社会各界和研究机构提供自身采集储存的各种海量数据源，尤其是美国政府走在前列，主动