

面向青岛西海岸新区的大数据人才培养模式研究与实践
(青岛市社科规划项目, 项目编号: QDSKL1601116)

大数据必修课

BIG DATA REQUIRED COURSE

孔 环 梁 虹 陈宗智 © 著



中国海洋大学出版社
CHINA OCEAN UNIVERSITY PRESS

大数据必修课

孔 环 梁 虹 陈宗智 著

中国海洋大学出版社

• 青 岛 •

图书在版编目(CIP)数据

大数据必修课 / 孔环, 梁虹, 陈宗智著. —青岛:

中国海洋大学出版社, 2017. 7

ISBN 978-7-5670-1503-6

I. ①大… II. ①孔… ②梁… ③陈… III. ①数据处
理 IV. ①TP274

中国版本图书馆 CIP 数据核字(2017)第 168729 号

出版发行	中国海洋大学出版社	
社 址	青岛市香港东路 23 号	邮政编码 266071
出 版 人	杨立敏	
网 址	http://www.ouc-press.com	
电子信箱	pankeju@126.com	
订购电话	0532-82032573(传真)	
责任编辑	潘克菊	电 话 0532-85902533
印 制	蓬莱利华印刷有限公司	
版 次	2017 年 8 月第 1 版	
印 次	2017 年 8 月第 1 次印刷	
成品尺寸	170 mm × 230 mm	
印 张	8	
字 数	300 千	
印 数	1 ~ 1500 册	
定 价	28.00 元	

大数据的价值在于它能够借助云客户端、互联网植入传统产业,使传统产业转型升级,使企业、政府以及公众机构获取更多的信息资源。互联网造就了一个规模化生产、智能制造、分享和应用大数据的现代信息化时代,并逐步或者说已经形成了大数据产业。包括德国的“工业4.0”、我国的“中国制造2025”也渗透着对大数据的倾心。政府对大数据发展高度重视,政策导向明显,产业布局合理;企业对大数据应用积极参与,敢于资金投入,摸索技术创新,探究产业升级;科研机构对大数据产业兴趣浓厚,设立科研院所、申请专业培养、设置专业课程,还通过举办大数据产业峰会吸引大数据产业巨头、对接大数据高端理论,等等。大数据已经在很多地区落地开花。

大数据时代将对经济学、管理学、政治学、社会学、组织学等很多学科领域产生巨大甚至根本意义上的改变。大数据的开发、应用翻开了新时代的篇章。我们在投入、拥抱、适应、推动大数据产业的过程中,也给大数据会计、大数据统计带来新的课题。对大数据合理地进行确认和计量;正确地对大数据进行会计处理,公允地将其体现在会计报表上,使大数据的应用成为一项新型企业资产研究的课题。站在大数据逐步产业化的层面,立足于雨后春笋般诞生的众多大数据企业和大数据研究机构的角度,让大数据资产走进大数据企业会计报表之中具有现实意义和深远的影响。

追溯大数据的专业渊源,我们能在传统统计专业里发现其蛛丝马迹。从大数据人才培养的角度来看,一些高校在统计或者相关专业的基础上开设大数据课程,这是方法之一。一些有能力的高校,看到日益崛起的大数据产业的无限



前景,单独开设大数据专业或者数据工程专业也未尝不可。

本书从大数据时代的特征入手,对大数据及其应用领域所涉及的“工业4.0”“互联网+”以及《中国大数据行动纲要》作了简要的分析。对大数据会计确认、计量、记录、报告进行了畅想式的模拟,对大数据时代给统计专业带来的机遇和挑战进行了分析,并对大数据人才的培养提出了框架式的建议。

本书编写初期曾取名为“大数据会计与统计必修课”,后追溯会计复式记账的数字处理根源和统计专业理论与实践渊源,认为会计与统计的指向就是大数据,并且高校大数据专业已经初现端倪,最终将本书直接定名为“大数据必修课”。

本书适合于想了解大数据、大数据资产、大数据产业的人士,大数据企业管理人员和财务人员,大数据、会计、统计、数据工程等相关专业的研究人员阅读与使用。

在本书编写过程中,参考了一些书籍和有关资料上的内容,得到了相关机构和人员的帮助,在此一并表示感谢。鉴于水平有限,书中尚有许多不足之处,敬请广大读者批评指正。

编 者

2017年5月

Contents 目录

第一章	大数据时代	1
第一节	大数据好玩且有用	5
第二节	大数据的时代脉搏	9
第二章	大数据与“工业 4.0”	25
第一节	“工业 4.0”剖析	29
第二节	大数据奠定“工业 4.0”未来发展	38
第三节	德国“工业 4.0”战略与大数据推进启示	42
第三章	大数据与“互联网 +”	47
第一节	“互联网 +”让云计算和大数据走进终端用户	51
第二节	“互联网 +”借助大数据升级传统产业	58
第四章	大数据在行动	63
第一节	发展大数据的意义和目标	67
第二节	发展大数据的任务和机制	70



第五章	大数据资产会计畅想	83
第一节	大数据会计的背景培育与技术支持	87
第二节	大数据资产的会计处理	93
第六章	大数据专业梦想	101
第一节	大数据与统计专业渗透碰撞	105
第二节	大数据专业人才培养	105
第七章	大数据落地之道	111
第一节	政府搭台,企业唱戏	115
第二节	高校研究峰会渗透	117

大数据时代

“大数据”一词,当下炙手可热。在百度词条搜索里,不是以多少条计数,而是需要在电脑里翻多少页的问题。你若不了解,那只能证明你已经跟不上时代前进的步伐,被“OUT”了。

其实,多数没被“OUT”的人也只是道听途说,因为关于大数据的段子形象逼真,借助互联网、手机朋友圈已经飞到了千家万户甚至每一个人的手里。从这个意义上讲,大家没有被时代落下,但是真正能全面系统地了解大数据并能够投入其开发应用或者说从中获益的人却很少。当下,大数据已经渗透到第四次工业革命(“工业4.0”)中,利用在“中国制造2025”中;大数据会计也将呼之欲出,并且已经有人开始展望大数据的未来、勾勒中国统计的美好蓝图了。投资数亿元人民币的惠普大数据产业园在青岛奠基;中国的高等学府之一——清华大学设立了大数据研究院、工业大数据研究中心;北京大学设立了大数据技术研究院,开设了大数据专业课程;很多高校、企业、科研机构更是瞄准了大数据、数据工程、数据产业。可以说大数据时代已经来临;或者说,我们已经处在一个“大数据”时代。

我们有理由也有必要去了解、学习、认识大数据,作为处在大数据这个产业、行业及其相关领域的人更应该把大数据当作一项必修课去学习和掌握。作为高校,应该将大数据内容融入相关学科的教学和科研中,让更多的人跟上时代的步伐。作为从事大数据、会计、统计、数据工程等理论与教学的专业,更应该深入研究大数据并将其作为学科建设的基础,开好大数据必修课,让学生学好大数据必修课。

第一节 大数据好玩且有用

最近流传着这样一个段子，题目是“什么是大数据”。这个段子虽然没有给大数据下一个严格的定义，但是也形象逼真地描述了一些大数据的特征。既然要研究大数据，我们不妨也传播一下这个段子。

什么是大数据

某必胜客店的电话铃响了，客服人员拿起电话。

客服：必胜客。您好。请问，有什么需要我为您服务的吗？

顾客：你好，我想要一份……

客服：先生，烦请先把您的会员卡号告诉我。

顾客：16846146***

客服：陈先生，您好！您是住在泉州路1号12楼1205室，您家电话是2624***，您公司的电话是4666***，您的手机号是1391234****。请问，您想用哪一个电话付费？

顾客：你为什么知道我所有的电话号码？

客服：陈先生，因为我们联机到CRM系统了。

顾客：我想要一个海鲜比萨……

客服：陈先生，海鲜比萨不适合您。

顾客：为什么？

客服：根据您的医疗记录，您的血压和胆固醇都偏高。

顾客：您可以试试我们的低脂健康比萨。

顾客：你怎么知道我会喜欢吃这种比萨？

客服：您上星期一在国家图书馆借了一本《低脂健康食谱》。

顾客：好。那我要一个家庭特大号比萨，要付多少钱？

客服：99元，这个足够您一家六口吃了。但您母亲应该少吃，她上个月刚做了心脏搭桥手术，还处在恢复期。

顾客：我可以刷卡付费吗？

客服：陈先生，对不起。请您付现款，因为您的信用卡已经刷爆了，您现在还欠银行4807元，而且还不包括房贷利息。

顾客：那我先去附近的提款机提款吧。



客服:陈先生,根据您的记录,您已经超过今日提款限额。

顾客:算了,你们直接把比萨送到我家吧,家里有现金。你们多久能送到?

客服:大约 30 分钟。如果您不想等,可以自己骑车来。

顾客:为什么?

客服:根据我们的 CRM 全球定位系统的车辆行驶自动跟踪系统记录,您登记有一辆车号为 SB-748 的摩托车,而且目前您正在解放路东段华联商场右侧骑着这辆摩托车。

顾客当即晕倒。

读完这个段子,给我们的第一印象自然是大数据很好玩。但是,这也告诉我们大数据给我们的生活带来了很多便利,从客服人员的服务里我们也感受到了企业服务的“无微不至”。其实,从这个层面上讲,大数据为提高服务行业的服务水平、服务效率做出了贡献;从更高层面去理解,大数据实现了传统服务业的升级。而事实上,大数据的魅力不仅仅是服务行业,未来几年甚至当下,大数据已经不断植入各行各业,让更多的传统产业具备了转型升级的可能。也可能有人提出这样的疑问:我的住宅地址、电话号码、健康状况是我的个人隐私,怎么能随意让他人获取呢?但是,我们的这些所谓个人隐私,如果不被作为非法利用而是让我们生活更加便利、更具品质,也就未尝不可。这些所谓隐私,隐起来又有什么用处呢?是的,假设让一些不该知道的人知道,这些所谓隐私的暴露会让我们不胜其烦,但是我们国家已经在制定大数据行动纲要,其中一个重要的内容就是数据信息安全,况且这些数据信息仅仅是大数据的冰山一角。

一、什么是大数据

言归正传,下面我们介绍一下到底什么是大数据。国际数据公司界定了大数据的四大特征:海量的数据规模(vast)、快速的数据流转和动态的数据体系(velocity)、多样的数据类型(variety)和巨大的数据价值(value)。这也是我们在很多大数据报刊、图书、网络上看到的所谓 4V 大数据概念。

从经济学的角度看,大数据是经过系统整理,储存在现实或虚拟空间里,能够提供一定价值的信息资源;狭义到会计学的层面,这些信息资源是大数据企业或大数据研究机构通过过去交易或事项合法取得,能够拥有或控制,并可以

带来经济利益的资产。

本书认可关于大数据的 4V 特征,同时将大数据定义为信息资源,并认定其为一项资产。当然,认定大数据为资产是需要建立在大数据企业、政府部门、高校团体、研究机构经过数字、信息、数字化、信息化、数据化、大数据化到大数据资产的过程中进而发挥其价值的基础上的。

从海量的数据规模来看,根据报道,全球 IP 流量达到 1 EB 所需的时间,在 2001 年是 1 年,在 2013 年仅为 1 天,到 2016 年则仅为半天。全球新产生的数据年增 40%,全球信息总量每两年就可翻番。而根据 2012 年互联网络数据中心发布的《数字宇宙 2020》报告,2011 年全球数据总量已达到 1.87 ZB (1 ZB = 10 万亿亿字节),如果把这些数据刻成 DVD 盘,将这些盘一张接一张排起来的长度相当于从地球到月亮之间一个来回的距离,并且数据以每两年翻一番的速度飞快增长。由此看来,大数据真够“大”的。预计到 2020 年,全球数据总量将达到 35 ~ 40 ZB,10 年间将增长 20 倍以上。

事实上,所谓大数据并不仅仅指数据海量,而更多的是指这些数据都是非结构化的、残缺的、无法用传统的方法进行处理的数据。也正是因为应用了大数据技术,美国谷歌公司才能比政府的公共卫生部门早两周时间预告 2009 年甲型 H1N1 流感的暴发。也就是说,大数据需要量化并进行不断的开发、分析和应用。大数据需要量化而不是数字化。所谓量化是指从错综复杂的大量数据中不断地提取、整理,把现象转变成为可以分析应用的形式。笔者想给大家说的就是大数据好玩,但不是用来玩的;大数据有用,应该体现其价值所在。

大数据将带来前所未有的变革,这也是我们说大数据的到来使我们进入大数据时代的原因。就像电力技术的应用不仅仅像发电、输电那么简单而是引发了整个生产模式的变革一样,基于互联网技术而发展起来的“大数据”应用,将会对人们的生产过程和商品交换过程产生颠覆性影响,数据的挖掘和分析只是整个变革过程中的一个技术手段,而远非变革的全部。“大数据”的本质是基于互联网基础上的信息化应用,其真正的“魔力”在于信息化与工业化的融合,使工业制造的生产效率得到大幅度提升。那么,信息化与工业化的融合就恰恰是我们“中国制造 2025”的精髓,因此我们后面的章节还会和大家一起研究大数据在行业变革中的应用。

大数据并不能生产出新的物质产品,也不能创造出新的市场需求,但能够



让生产力大幅度提升。正如《大数据时代：生活、工作与思维的大变革》作者肯尼思·库克耶和维克托·迈尔-舍恩伯格指出的那样，数据的方式出现了三个变化：第一，人们处理的数据从样本数据变成了全部数据；第二，由于是全样本数据，人们不得不接受数据的混杂性，而放弃对精确性的追求；第三，人类通过对大数据的处理，放弃对因果关系的渴求，转而关注相互联系。这一切代表着人类不再总是抱有试图了解世界运转方式背后深层原因的态度，而仅仅需要弄清现象之间的联系以及利用这些信息来解决问题。

二、大数据在国家层面的认可

在 2016 中国大数据产业峰会上，面对 3 000 多位海内外嘉宾，国务院总理李克强发表了重要讲话，高度认可信息技术为全球发展带来的助力作用，同时将大数据誉为“钻石矿”，并从三个层面深度分析了大数据技术给社会经济带来的关键意义。

（一）推动信息联通

总理开篇谈到，当今世界，信息化潮流席卷全球，地球村从概念变成现实，大数据在其中起了关键作用。

随着互联网、移动互联网产业的快速发展，信息的快速流通与交互拉近了人与人之间的距离。国家之间的快速联通，使得“地球村”的概念快速落地。这其中，大数据的助力作用不容忽视。

“谁掌握了先机，谁就掌握了未来。”总理认为，中国曾经有过与世界科技革命失之交臂的教训。当下随着创新技术的快速发展，所有国家和地区只要通过努力，都可以站在同一条起跑线上，以前落后的地方同样可以抢占先机。

一些欠发达的地方，能够跃上这个高地，依靠大数据、云计算、物联网所代表的新一代创新技术，将发展新经济作为主要方向，从依赖自然资源到依赖人力资源，实现可持续发展。中国有 14 亿人口，我们有发展压力，但是把压力变成资源后，压力就是潜力。数十亿人都在生成数据、加工数据、处理数据，数据将成为巨大的新资源。

（二）实现产业变革

创新技术的快速迭代，带来了新兴产业的兴起。近年来，中国在互联网与

移动互联网领域的技术发展与商业模式的创新,实现了对欧美国家的赶超,与互联网相关的产业都在高速发展。

伴随着国家“互联网+”战略的逐步深入,传统产业转型升级的步伐也在不断向前迈进。总理指出,这其中大数据带来了深刻影响,同时带动了产业变革。以货运行业为例,有一家货运企业拥有会员货车 170 多万辆,通过大数据技术进行信息收集、数据交互,大大降低了空驶率。把大数据与传统行业的工匠精神结合起来,就能融合虚拟世界和现实世界,实现新旧动能转换,实现价值链、产业链和供应链的变革。

(三) 带动经济增长

总理指出,今天的中国,完全可以把握住历史发展的机遇,推进供给侧改革,不断提升劳动生产力,通过发展新经济推动产业转型升级。

数据可以比作钻石矿,应把互不相连的信息孤岛连接起来。总理认为,也正因为有共享,数据才能无限放大。共享经济正作为新的经济增长点,带动中国经济进入一个新的发展周期。“共享经济可以利用闲置资源和过剩产能,提高效率,缩小区域、城乡、人群之间的差距和数字鸿沟。”

目前 80% 的信息资源掌握在政府部门手中,政府就要发挥作用,打破信息孤岛,除涉及隐私之外,其他信息都应该向公众和社会开放,形成“人在干,数在转,云在算”的局面。

第二节 大数据的时代脉搏

大数据在中国的发展,经历了前期的基础认识、理论研究阶段,后期的大数据市场渗透阶段,目前进入产业初期应用阶段。从产业经济学的角度看,目前大数据产业仍然处在初期阶段。在这样一个阶段,作为一个产业的形成,首先需要深入地开展更高层次的大数据理论研究;其次需要培养大数据人才和大数据从业人员;再次要建立大数据产业机制,完善大数据产业政策;第四应构建大数据产业集群,扩大大数据产业规模。

十多年前,大数据已经进入中国的理论界的视野并有人开始研究,典型的大数据产品和服务也已经开始崭露头角。互联网对大数据的应用率先落地。



紧接着,也就是在几年前,大数据的概念迅速普及,引起了资本市场的密切关注,一些拥有数据信息资产的企业已经开始谋求利用大数据进行产业转型升级。伴随着中国近两年“互联网+”概念的推出和“大众创业、万众创新”发展策略的实施,大数据应用最为广泛的还是商品流通行业,它们利用大数据进行市场细分,逐步获得差异化竞争的先机;进而,大数据应用向交通、物流、医疗、体育、传统服务业甚至农业等领域渗透。

一、大数据发展的脉络

大数据的特征是数据信息量巨大、数据类型多样化、时效性强、真实性差。所以在大数据发展的过程中,面对庞大、多样的数据信息如何进行数据分析整理,面对转瞬即逝的时效性“财富”如何迅速识别是“财富”还是“误导”,这些成为大数据发展与应用所面临的挑战。

某报道称,目前全球教育、交通、消费、电力、能源、健康、金融等七个大数据重点应用领域,大数据应用价值在3万亿~5万亿美元。因此,世界各国都对大数据发展高度重视。美国制定了《大数据发展研究和发展计划》,欧盟出台了《数据价值链战略计划》,英国实施了《英国数据能力发展战略规划》,亚洲的日本和韩国则分别颁布了《创建最尖端IT国家宣言》和《大数据中心战略》。

中国在2015年“两会”期间提出“互联网+”的概念,紧接着制定了《“互联网+”行动计划》;同时,积极筹划、编制“软件与大数据产业十三五规划”和推进“云计算与大数据体系建设”;2015年7月国务院办公厅下发了运用大数据加强对市场主体服务和监管的若干意见。

附:国务院办公厅关于运用大数据加强对市场主体服务和监管的若干意见

国务院办公厅关于运用大数据 加强对市场主体服务和监管的若干意见

国办发〔2015〕51号

各省、自治区、直辖市人民政府,国务院各部委、各直属机构:

为充分运用大数据先进理念、技术和资源,加强对市场主体的服务和监管,推进简政放权和政府职能转变,提高政府治理能力,经国务院同意,现提出以下