

# 大数据时代信息类专业建设的思考\*

明小波, 孙艳玲, 刘杨, 苏谦

(成都信息工程学院 物流学院 成都, 610103)

**摘要** 主要以“2013年大数据时代信息类专业建设与人才培养研讨会”专家演讲和教师讲解为素材, 借助近期相关研究作补充, 探讨大数据时代下信息类专业建设的可行路径, 包括大数据时代信息类专业专业的地位、核心能力、综合能力、实践背景特色及专业身份认同等问题。

**关键词** 大数据, 信息学科, 专业建设

**中图分类号** C931.6

## 1 本文背景

2013年7月26—27日, 中国电子教育学会高等教育分会、电子工业出版社、北京致远协创软件有限公司共同举行“大数据时代专业建设面临的挑战——信息类专业建设与人才培养研讨会”, 来自教育部管理科学与工程类学科教学指导委员会等学术机构和高等院校的9位同行专家围绕大数据时代信息类专业建设的新趋势、新问题、新思路和新方法系统地阐述了自己的认识, 对于未来信息类专业建设具有很强的指导意义。各位专家站在不同视角阐述自己的观点, 有着独立的逻辑思考, 能够从不同角度启发大家深入探索某些局部性问题。可是, 通过完整记录专家们的演讲(全部录有视频), 并对这些主题内容进行细致分析, 发现各个主题之间有着明显的强相关性, 如果把它们的相关要点进行梳理, 则会对信息类专业专业的历程、机遇、地位、价值、能力需求、知识体系及实践特色等问题形成系统性认知, 更有利于信息类专业专业的教师与从业者站在全局的视角理解大数据时代下本类专业专业的发展路径。为此, 本文主要以参会专家演讲和部分教师讲解为素材, 借助近期相关研究作补充, 探讨大数据时代下信息类专业专业建设的可行路径, 为促进高校同行达成共识提供参考。

## 2 大数据时代凸显信息类专业地位

### 2.1 大数据时代为信息类专业专业发展创造机遇

如同20世纪60年代大型机、80年代个人机、90年代互联网和2000年Web 2.0一样, 每一次技术进步都对社会和个人带来挑战和机遇, 当今的云计算/物联网/大数据也对信息类专业专业提出挑战, 并且对信息类专业专业发展是一次极好的机会。过去的历程说明, 计算机网络发展经历了数据集中(大型机)、数据分散(个人机、互联网、Web 2.0)和数据再次集中(云计算/物联网/大数据)的各个阶段, 社会和个人在进入每一个阶段时都要开始新一轮的学习过程, 信息类专业专业建设同样面临新一轮的学习过程<sup>[1]</sup>。目前处于数据再次集中(云计算/物联网/大数据)的阶段, 信息类专业专业将适应大数据

\* 通信作者: 明小波, 男, 成都信息工程学院教授。研究方向为电子商务和企业信息化。E-mail: mingxb@foxmail.com。

分析的要求,凸显其在应用和学术上前所未有的重要地位,这是在前两个阶段从未出现过的历史机遇。近年国外学者有关大数据的研究明显增多,从2008年对大数据研究兴起到现在,在ISI Web of Science上的“big data”关键词下可检索130个相关条目和94位作者,而这一数据在2008—2011年其平均水平分别是16、16、13、26位<sup>[2]</sup>。因此,在大数据背景下,信息类专业专业的价值迅速提升,将有利于消除就业“黄牌”的困惑(这是“小数据”时期留给信息类专业专业的困惑)。

## 2.2 大数据产生的物质基础为信息应用提出新问题

物联网是大数据的主要来源之一,云计算是大数据价值提升的重要途径<sup>[3]</sup>,大协同是产生大数据的社会化活动<sup>[4]</sup>。如果说互联网重点解决了人与人之间信息交互问题,那么在物联网基础上将会完成人与物、物与物的信息交互,导致产生更大的数据量。除了数据产生的大量性(Volume)特点外,大数据还体现在数据类型的多样性(Variety)、数据传输的高效性(Velocity)和数据利用的价值性(Value)上,即“4V”<sup>[1,3]</sup>。由这些特性所带来的信息系统应用问题都有赖于信息类专业专业发挥更大的作用。

具体而言,大数据主要产生于三个方面的商业应用中。

(1) 交易数据:ERP、SCM、CRM和Web交易程序等经典的交易处理系统,这些系统通常用SQL数据库来存储高度结构化的数据。

(2) 互动数据:“互动”指人们或事物相互之间,与企业之间的互动。网站日志、用户点击流、社交互动 & Feeds,以及UGC用户产生的内容都是典型的互动数据。

(3) 观测数据:观测数据主要来自物联网,包括观测热量、运动、压力的传感器,以及来自移动设备中RFID和GPS芯片的数据,甚至ATM和飞机引擎都是物联网中产生观测数据的“物”。<sup>[5]</sup>

为应对大数据挑战,信息应用将面临如下主要问题:实施SQL专业分析;扩展对邮件、论坛、博客、社交网络、POS系统和机器生成的非结构化数据的解决方案;对大数据进行优化检索;创建基于ETL(数据提取、转换和加载)的数据整合解决方案;运用实时信息流处理“大数据运动”;增强数据挖掘能力,支持统计、文本、图像、情感分析等;推动“R”开源统计语言处理复杂的统计分析问题等。<sup>[6]</sup>

## 2.3 大数据时代为信息类专业专业提升社会价值

大数据时代下信息类专业专业地位的提升主要体现在以下三方面。

(1) 大数据分析的价值提升。麦肯锡公司2011年报告推测,如果把大数据用于美国的医疗保健,一年产生潜在价值3000亿美元,用于欧洲的公共管理可获得年度潜在价值2500亿欧元;服务提供商利用个人位置数据可获得潜在的消费者年度盈余6000亿美元;利用大数据分析,零售商可增加运营利润60%,制造业设备装配成本会减少50%<sup>[7]</sup>。

(2) 大数据人才的需求提升。大数据需要专业化的技术和管理人才,包括大数据解决方案的设计和实施,需要专业化分析复杂数据集的工具和技术。企业、组织需要大量既精通业务又能进行大数据分析的人才,以及具备理解和基于大数据研究做出决策的经理和分析师人才<sup>[8]</sup>。

(3) 重视大数据人才的培养。IBM中国开发中心信息管理总经理朱辉进一步证实了信息类专业专业的社会需求:“当前,越来越多的组织已意识到大数据时代下信息管理人才的重要性。IBM与北京交通大学发布了信息管理联合人才培养计划,共同开发的信息管理方向系列课程,引入行业实践,帮助学生掌握社会和产业真正需要的技术能力。”<sup>[9]</sup>

### 3 “分析与处理”是信息类专业核心能力

“大数据”亟待分析和利用,否则就会成为“信息垃圾山”而危害信息社会。有效利用“大数据”的前提是对数据进行快速的分析与处理,因此在大数据驱动下对信息类专业在数据分析能力方面的要求格外突出。如果说以前各个时期对信息类专业能力培养的定位还不够明确一致,那么现今大数据时代下信息类专业的核心能力则可以明确为“分析与处理”。即是,大数据时代促成了信息类专业的明确定位。

#### 3.1 “四维需求”与信息类专业的核心能力

“分析与处理”对大数据利用的意义主要体现在两方面:一是提升大数据链的整体价值;二是支撑物联网与互联网上的各种应用和服务。信息类专业的“分析与处理”能力主要体现在两个维度中,一个维度是“用”;另一个维度是“造”<sup>[1]</sup>。“用”和“造”的对象是信息系统,信息系统将信息管理与信息技术实现深度融合<sup>[10]</sup>,而深度融合的途径则正是“分析与处理”过程。

在“用”的维度里,重点开展系统分析与数据分析工作。不管是进行系统分析还是数据分析,首先都要对实际业务流程进行分析,只有在充分理解业务流程的前提下实施的信息系统才能给企业运营带来价值<sup>[11]</sup>。数据分析的内涵是“管理大数据,理解大数据”,云计算本身就是一种管理和理解大数据的业务模式,体现大协同下的数据处理机制<sup>[4]</sup>。在“造”的维度里,重点研究如何采用分布式处理方式高效地分析和处理非结构化数据和半结构化数据(目前占据信息资源的 80% 以上),以及如何采用先进的移动终端设施采集和处理包括情景感知在内的非结构化数据<sup>[3]</sup>。

在实际应用和研究中,信息分析与处理能力除了以“用”和“造”的维度划分外,还可以通过增加“受”和“予”两个维度做出进一步的细分<sup>[1]</sup>。这样,就构成信息系统的四维功能结构,如图 1。

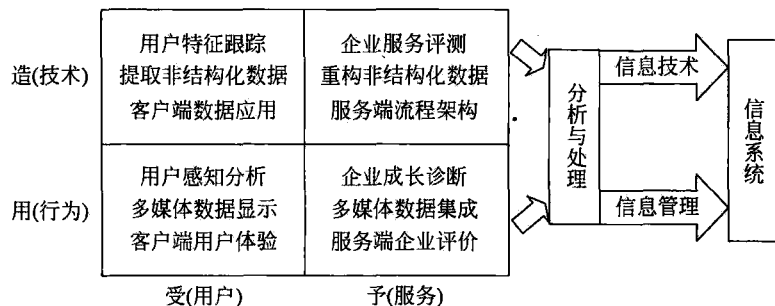


图 1 信息系统的四维功能结构

图 1 说明:

(1) “造”的维度代表从信息技术角度能够实现的应用功能;“用”的维度代表从使用行为角度能够实现的应用功能;“受”的维度代表面向客户端或用户的服务功能;“予”的维度代表面向服务端或企业的服务功能。

(2) 在“造-受”维度中,表明客户端在技术上主要满足用户特征跟踪、提取非结构化数据和数据应用的需要;在“用-受”维度中,表明信息系统能够实现用户感知分析、多媒体数据显示和客户端的用户体验等一些主要功能;在“造-予”维度中,表明服务端在技术上主要满足企业服务评测、重构非结构化数据和服务流程架构的需要;在“用-予”维度中,表明信息系统能够进行企业成长诊断、多媒体数据集

成和对服务企业的评价等工作。

(3) 一方面,上述“四维功能”是对“信息分析与处理”能力的具体要求;另一方面,通过“分析与处理”能力的作用,可以将信息技术和信息管理融合成为信息系统,使信息系统具备“四维功能”的能力。

### 3.2 大数据分析处理的关键技术和需求趋势

适用于大数据的关键技术包括遗传算法、神经网络、数据挖掘、回归分析、分类分析、聚类分析、关联规则学习、数据融合与集成、机器学习、自然语言处理、情感分析、空间分析、分布式文件系统、分布式缓存、分布式数据库、非关系型数据库系统、可视化技术。<sup>[8]</sup>它们说明在“用”和“造”两个维度中关于分析与处理范畴的更为详尽的能力与知识构成。

摩根斯坦利发布报告:2013年大数据十大最快增速领域是医疗(Healthcare)、娱乐(Entertainment)、互联网/媒体(Com/Media)、制造业(Manufacturing)、金融业(Financial)、商业服务(Business Services)、运输业(Transportation)、Web技术(Web Tech Web)、通信(Distribution)、工程(Engineering)。美国技术专家 Jim Kaskade 认为 2013 年大数据分析的基本组件包括实时流处理(real-time stream processing)、趋近实时分析(基于 NoSQL 和 NewSQL 的数据存储)(Ad-hoc near real-time analytics (see NoSQL and NewSQL data stores))、批量处理与分析(batch analytics),并且同时具备三方面的能力才是大数据真谛所在<sup>[12]</sup>。表明大数据时代对数据分析与处理及其技术需求的总趋势。

## 4 重构信息类专业能力知识体系

### 4.1 “使能创新”能力与核心能力的关系

在北美地区,信息类专业人才的就业状况已经走出低谷,并对专业人才有较高的综合素质要求,应当具备 7 种高端能力(即组织流程改进能力,利用 IT 创新机遇的能力,理解和处理信息需求的能力,设计和管理企业架构的能力,辨识和评价解决方案和资源配置方案的能力,数据和基础设施安全保障能力,认识、管理和控制 IT 风险能力)。这 7 种高端能力归结在一起,就是 IT 使能创新(IT-enabled innovation)的能力。华盛顿州立大学的 Joshi, KD 等人调查了 110 家企业,收集有关创新成果的相关数据。利用整理后的二次数据,他们发现使能创新的知识能力(IT-enabled knowledge capabilities and innovation)对企业技术创新的影响非常显著<sup>[13]</sup>。据此,信息类专业围绕“使能创新”,并结合实际教学的可行性,可以将该专业的综合素质简化为四个方面的能力构成,即构架能力、应用能力、分析能力和创造能力,而对四种能力的培养则要做到均衡发展<sup>[14]</sup>。

综合能力寓于核心能力之中,如图 2。分析与处理行为在架构能力上的表现结果是做出流程设计,在应用能力上的表现结果是解决管理问题,在分析能力上的表现结果是提出解决方案,在创造能力上的表现结果则是创新商业模式或运营模式。

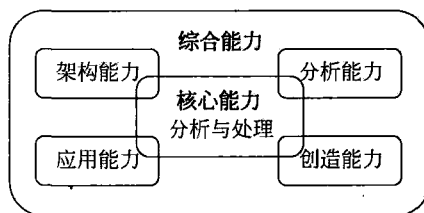


图 2 信息类专业核心能力与综合能力关系示意

## 4.2 “数据分析师”类型及其能力要求<sup>[15]</sup>

O'reilly Strata 于 2013 年 6 月发布《Analyzing the Analyzers》报告,阐述了数据科学部门所需要的不同角色及其技能。经问卷调查,目前“数据分析师(data scientists, or data analyzers)”主要分为 4 类:“商务数据师(data businesspeople)”、“数据创意师(data creatives)”、“数据处理师(data developer)”和“数据应用师(data researchers)”。

其中,“商务数据师”专注于组织管理和如何从数据项目中产生利润;“数据创意师”能够综合使用各种工具和技术解决问题,并像黑客马拉松那样创新产品原型,是典型的万事通型人才;“数据处理师”专注于解决数据管理的技术问题,即如何获取、存储及学习;“数据应用师”侧重利用数据了解复杂过程。大多数“数据分析师”既有丰富的数据分析专业经验,还至少在数据学科的某一方面进行深入钻研,如社会学、物理学、统计及业务沟通等,形成“T 形能力”模式。“T”字横画表示广博技能,竖画表示在某个区域的深度钻研。经典计算机科学(机器学习)和应用统计学知识是成为上述各类“数据分析师”的基点。

## 4.3 信息类学科专业能力知识体系新构架

据此,可以构建现代信息分析师的能力知识体系,如图 3。这个能力知识体系可以作为制定信息类学科专业培养方案的基本依据。

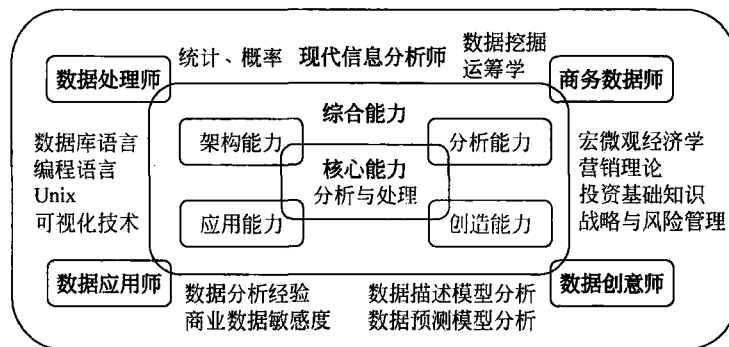


图 3 现代信息分析师能力知识体系

图 3 表明:

(1) 现代信息分析师按照他们工作的侧重点不同,主要分为 4 个职业方向,即数据处理、数据应用、商务数据分析和数据产品创新,并由此形成 4 个新的职业岗位,包括数据处理师、数据应用师、商务数据师和数据创意师。现代信息分析师是这 4 个新的职业数据师的统称。

(2) 现代信息分析师不论职业方向侧重如何,都应具备“分析与处理”核心能力与“架构能力、应用能力、分析能力和创造能力”相兼容的综合能力。在此基础上,再突出职业方向上的能力知识培养。

(3) 现代信息分析师必须具有统计、概率、数据挖掘和运筹学等通识性专业基础理论知识。

(4) 不同职业方向上的能力知识分别是:数据处理师突出对数据库语言、编程语言、UNIX 和可视化技术的学习;数据应用师突出在数据分析实践和商业数据价值分析方面进行培养;商务数据师突出对宏观经济学、营销理论、投资基础、战略与风险管理方面知识进行深入学习;数据创意师突出在数据描述模型分析和数据预测模型分析方面的能力培养,即掌握上面所提及的大数据关键理论方法等。

## 5 实践环节是各信息类学科专业的特色所在

基于上述对核心能力和综合分析的能力,各高校信息类学科专业的培养目标还应当根据自身所处地区、行业背景和教学师资条件的特点实行多元化定位,将通识教育与专业教育相结合、通用型人才与专业型人才培养相结合,设立柔性化课程体系,让学生做出个性化方向选择。无论采用哪种培养模式,“实践学习”或“实与习”应该成为常态<sup>[14][16-17]</sup>。由于实践环节与地区需求、行业背景结合最为紧密,各高校信息类学科专业的差异化或者说专业特色最能够在实践环节上得以体现。

可以根据各高校的内部条件不同实施多元化教学,如清华大学等引入网络化实践课程,推荐MOOC(massive open online course,即“魔克平台”——大规模开放在线课程),将教室转化为实验室,设立教学实验基地,将实践领域专家引进课堂<sup>[14]</sup>;北京师范大学引入iSchool运动教学体系<sup>[18]</sup>,分享iSchool学校群在信息、人员和技术上的资源等。在学生个性化选择培养方向的前提下,强调对学生实践作品进行考核,将学科专业的特色落实到实践环节上<sup>[14][19]</sup>,例如,广东药学院引导学生在医学数据中完成实践作品<sup>[16]</sup>;东华大学要求学生在教师引进的企业项目中完成实践作品<sup>[19]</sup>;北京师范大学让学生在iSchool运动中独立完成一个实践项目<sup>[18]</sup>;成都信息工程学院引导学生以项目为载体完成CDIO(Conceive-Design-Implement-Operate,构思-设计-实施-运行)工程实践过程<sup>[20]</sup>。

## 6 信息类学科专业的身份认同有赖于努力建设

长期以来,由于信息类学科专业缺少本身的核心理念、核心理论和基本有力的概念(absence of core theories and a fundamental and powerful concept),信息系统研究缺乏核心对象,缺乏实际相关性,研究呈现碎片化特征(fragmented nature of IS research),信息系统相关利益者得不到足够的身份认同(identity)<sup>[14]</sup>,造成信息类学科专业不易得到社会的认同和接受、在高校内其学科专业的地位不能凸显、专业教师的归属感不强、专业学生就业方向不能确定等窘境。(信息系统相关利益者包括其研发人员、管理人员和咨询人员,以及信息类学科专业的教师们。)

目前,数据分析与处理的关键技术与方法已经积淀有一定基础,面对大数据时代的到来,需要依托这些技术与方法进一步发展和提升信息类学科专业的理论体系,这对彰显信息类学科专业的社会价值和地位极有必要,也极为有利,能破解信息系统相关利益者身份认同的问题。当然,信息类学科专业的理论体系建设还需要信息系统的研发人员、管理人员、咨询人员以及广大教师付出艰苦的努力,要积极面对大数据下的难题,“确立大数据地位,夯实大数据基础,创新大数据应用”<sup>[21]</sup>,变数据仓库为“大数据池”,切实解读社会科学问题,突破基于大数据的社会研究范式,拓展全新的研究视野,创建核心理论体系<sup>[22]</sup>,涌现一批信息分析与处理咨询师、信息系统架构专家等,得到社会广泛认同,才能最终牢固地确立信息类学科专业的身份地位。

## 7 信息类学科专业建设的启示

信息类学科专业建设的启示如下。

(1) 学科专业的“合与分”<sup>[23]</sup>。信息类学科专业主要体现为三个方面的融合:IT与应用的融合(包括信息技术与生活、商务及业务应用的融合);企业内外数据的融合(从关注企业内部信息转向关注企业外部信息);跨学科的融合(传统经济管理领域都在考虑OnLine的问题,学科在支撑全领域应

用,导致学科细分调整)。另一方面,在实际应用趋于高度融合的同时,更需要注意在学科专业建设中保持“分野”的态势,求同存异,坚持各高校自身的专业特色、培养模式和成功经验,从适应企业和社会需求上形成各高校学科专业的差异性。

(2) 正确处理教学内容关系<sup>[23]</sup>。信息类学科专业的本科教学中,注意处理好基础内容、主流内容和把握发展趋势的关系。基础内容是指掌握经典的、成熟的知识,如ERP、CRM等。主流内容是指对当前主流的技术和方法要重点学习,如云计算、物联网、大数据、商务智能、数据挖掘等。把握发展趋势是指对未来5~10年学科专业方向有所预见,例如信息技术、业务应用等将有什么变化。清华大学在处理这三者学时的比重时,大致做到基础内容占总学时的30%,主流内容占40%,发展趋势内容占30%。对于研究生教学,则适当加重后两者的比重。

(3) 我们也留意到外国学者在注重课堂生动性和调动学生积极性上做出的一些研究,例如:Thomas Carter和Peggy Hauselt等提出根据学生的性格和期望不同,量身安排课程和实践活动,充分调动其积极性<sup>[24]</sup>。在师资培养方面,要注意“研与教”结合:学科专业教师应当做与自己教学和研究相关的事情,特别要重视开展应用研究或者理论研究,敏锐地感知实际中和学术上的动向,及时将这些感知投射到教学中,让教学既有深度,又有现实性,使课堂内容更加生动<sup>[23]</sup>。

## 参考文献

- [1] 陈国青. 大数据时代下管理挑战与专业建设思考. 主题演讲, 清华大学经济管理学院.
- [2] Snijders C, Matzat U, Reips U D. Big data: big gaps of knowledge in the field of Internet Science[J]. International Journal of Internet Science, 2012, 7(1): 1-5.
- [3] 吕廷杰. 大数据时代高教的机遇与挑战——信息化发展与教学体系创新. 主题演讲, 北京邮电大学经济管理学院.
- [4] 刘古全. 大数据背景下的校企合作如何共生共赢. 主题演讲, 致远协同研究院.
- [5] Connolly S. 7 Key Drivers for the Big Data Market[EB/OL]. <http://hortonworks.com/blog/7-key-drivers-for-the-big-data-market/>, May 14th, 2012
- [6] Zurek B. 7 Big Data Trends That Will Impact Your Business[EB/OL]. <http://smartdatacollective.com/bob-zurek/125591/7-big-data-trends-will-impact-your-business>, May 24, 2013.
- [7] 邬贺铨. 大数据时代的机遇与挑战[J]. 信息安全与通信保密, 2013, 3: 9-10.
- [8] 严霄凤, 张德馨. 大数据研究[J]. 计算机技术与发展, 2013, 23(4): 168-172.
- [9] 刘海永. “助力智慧成长, 培养大数据时代高端人才”[J]. 中国教育信息化, 2012, (19): 89.
- [10] 梁昌勇. 新兴信息技术发展背景下信管专业建设的思考. 主题演讲, 合肥工业大学.
- [11] 谢康. 企业信息技术与业务匹配: 动态能力视角. 主题演讲, 中山大学管理学院人力资源系统实验室.
- [12] Gartner公司. 2013年大数据真谛: 实时分析与批量处理. 2013. 8.
- [13] Joshi K D, Lei C. Changing the competitive landscape: Continuous innovation through IT-enabled knowledge capabilities[J]. Information Systems Research, Sep2010, Vol. 21 Issue 3
- [14] 黄丽华. 大数据时代信息管理与信息系统专业的人才培养. 主题演讲, 复旦大学经济管理学院信息管理系.
- [15] Harlan H, Sean M, Marck V. An introspective survey of data scientists and their work, analyzing the analyzers [J]. O'Reilly, June 2013.
- [16] 周怡. 大数据背景下教学方法出现的新动向——面对医学数据堆积. 主题演讲, 广东药学院.
- [17] 侯卫真. 大数据背景下的信息资源管理专业建设设想. 主题演讲, 中国人民大学信息资源管理学院.
- [18] 耿睿. 面向应用驱动的信息管理专业教学探讨. 主题演讲, 北京师范大学政府管理学院.
- [19] 王素芬. 东华大学信息管理与信息系统专业建设特色. 分组讨论会发言记录, 东华大学.
- [20] 孙艳玲. 成都信息工程学院信息管理与信息系统专业建设特色. 分组讨论会发言记录, 成都信息工程学院.
- [21] 刘琼. 大数据时代的美国经验与启示[J]. 人民论坛, 2013, 15: 30-31.

- [22] 徐磊. 大数据基础上的社会认知[J]. 中国电子科学研究院学报, 2013, (1): 23-26.
- [23] 陈国青. 大数据时代下信息类专业建设的几点启示. 大会总结, 清华大学经济管理学院.
- [24] Thomas C, Peggy H, Melanie M, Megan T. Building a big data research program at a small university[C]. CCSC: Rocky Mountain Conference, JCSC 28, 2 (December 2012): 95-102.

## Construction of the Information Systems Disciplines in the “Big Data” Era

MING Xiaobo, SUN Yanling, LIU Yang, SU Qian

(Chengdu University of Information Technology, Chengdu, 610103)

**Abstract** Based primarily on the expert speakers and teachers speaking on “2013 Seminar of Information Disciplines construction and personnel training in era of big data”, and supplemented with recent studies to explore the relationship among those lecture topics, try to find a viable development path of following information disciplines in “big data”, including the position of information disciplines, and its core competencies, comprehensive ability, characteristics of practice, professional identity and other issues.

**Key words** Big Data, Information Disciplines, Construction of Disciplines

### 作者简介

明小波(1959— ),男,成都信息工程学院教授。研究方向为电子商务和企业信息化。E-mail: mingxb@foxmail.com; mxb@cuit.edu.cn; mingxiaobo@tsinghua.org.cn。

孙艳玲(1971— ),女,成都信息工程学院教授。研究方向为信息资源管理。E-mail: sunyanling@cuit.edu.cn。

刘杨(1988— ),男,成都信息工程学院硕士研究生。研究方向为电子商务。E-mail: ly5628505@foxmail。

苏谦(1971— ),男,成都信息工程学院讲师。研究方向为信息管理与数据挖掘。E-mail: sla@cuit.edu.cn。