

中国大数据专业建设的跨学科模式研究

宁慧聪

(中国电子学会 北京 100036)

摘 要 随着以大数据、云计算、人工智能为代表的新一代信息技术的蓬勃发展,数字经济成为了带动中国经济增长的重要引擎,加快大数据专业建设、培养新一代信息技术人才意义重大。当前,国内外已经有多所高校和研究机构开展了大数据人才培养工作,但在如何开展大数据专业建设的问题上还没有成熟的模式。为此,首先梳理了国内外大数据专业建设现状,接着采用德尔菲法(专家调查法)和案例分析法对其进行分析,最后结合跨学科研究和人才培养机制,提出一种“点”“线”“面”“立体”相结合的中国大数据专业建设的跨学科模式,以期为中国大数据专业建设的跨学科发展研究提供有益的参考。

关键词 大数据,数据科学,跨学科,专业建设

中图分类号 TP391 **文献标识码** A

Study on Interdisciplinary Model of Construction of Big Data Discipline in China

NING Hui-cong

(Chinese Institute of Electronics, Beijing 100036, China)

Abstract With the vigorous development of new generation of information technology represented by big data, cloud computing and artificial intelligence, digital economy has become an important engine to drive China's economic growth. It is a great significance to speed up the construction of big data discipline and train new generation of information technology talents. At present, there are many universities and research institutes at home and abroad to carry out the training of big data talents, but there is no mature model on how to carry out the construction of big data discipline. Therefore, this paper summarized the existing achievements of the construction of big data discipline at home and abroad, and used the Delphi method (expert investigation method) and case analysis method to conduct analyse. Lastly, combined with interdisciplinary research and personnel training mechanism, the interdisciplinary model of “point” “line” “lane” and “three-dimensional” in the construction of big data discipline in China was proposed to provide useful reference for the interdisciplinary development research of the construction of big data discipline in our country.

Keywords Big data, Data science, Interdisciplinary, Discipline construction

1 引言

当前,新一代信息技术正在全球孕育兴起,在学术、科技、产业等科技工作者的共同努力下,大数据的融合创新正发生着重大变革。*Nature* 和 *Science* 杂志分别于 2008 年和 2011 年以专刊形式报道人类已迈进 PB 规模的大数据(Big Data)时代。世界各国非常重视大数据技术和产业的发展,陆续出台国家战略。例如:2012 年 3 月,美国总统科技政策办公室 OSTP(Office of Science and Technology Policy)宣布了每年投资两亿美元支持“大数据研究计划”(Big Data R&D Initiative)。同天,我国科技部发布的“‘十二五’国家科技计划信息技术领域 2013 年度备选项目征集指南”把“大数据研究”列在首位。我国在 2015 年由国务院印发《国务院关于印发促进大数据发展行动纲要的通知》,与此同时世界主要发达国家如英国、法国、德国、日本等相继出台大数据相关政策,这些都标志着大数据已经上升到国家战略层面。

为适应大数据快速发展的特点和对社会产生的重要影响

力,我国高度重视大数据专业人才的培养。根据 2018 年初发布的《教育部关于公布 2017 年度普通高等学校本科专业备案和审批结果的通知》,我国新增“数据科学与大数据技术”本科新专业 250 个,自 2016 年教育部批准 3 个第一批“数据科学与大数据技术”专业以来,到 2018 年我国已有 283 所高校获批数据科学与大数据专业。其中包括“数据科学与大数据技术”专业、“大数据管理与应用”专业和“大数据技术与应用”专科专业,分别授予其工学、理学和管理学学位^[1]。

开展大数据专业建设已经成为共识,如何开展大数据专业建设成为学术界和产业界共同关注的焦点。为专注本文研究内容——大数据专业的跨学科模式研究,本文不对多个专用名词的差异进行详细分析,仅将专用名词进行简化。对前文提到的“数据科学与大数据技术”专业、“大数据管理与应用”专业和“大数据技术与应用”专科专业等名词,如无特殊情况,后文统一简称为“大数据专业”;对“跨学科”^[2-5]和“交叉学科”^[2-4]等名词,如无特殊情况,后文统一简称为“跨学科”。

本文受中国科学技术协会新一代信息技术人才培养项目资助。
宁慧聪(1983—),男,博士,工程师,CCF 会员,主要研究方向为大数据、云计算、人工智能,E-mail:18211103893@163.com。

2 研究内容与研究方法

本文的讨论分为两大部分:第一部分是是对国内外近年来大数据专业建设的情况做研究综述,即通过梳理,分析大数据专业建设的跨学科挑战,见第 3 节和第 4 节;第二部分提出大数据专业建设的跨学科模式,即在第一部分的基础上对“大数据”的“跨学科建设”的系统化呈现,是本文的主要结论,见第 5 节。

2.1 研究内容

本文主要研究我国大数据专业建设的跨学科模式。先对我国大数据专业建设的现状、存在的挑战进行详细讨论,分析大数据专业建设模式中存在的不足,阐述大数据的跨学科特性。通过多种研究方法,分析出国内外大数据专业建设的挑战,提出一种我国建设大数据专业的跨学科模式。

2.2 研究方法

本文以理论和实践结合为基础,一方面采用德尔菲法(专家调查法)和案例分析法对我国大数据领域的院士和知名高校开课专家进行理论和实践探讨,同时在举办“全国高校《大数据导论》师资研修班”的基础上,对多种类型约 40 个高校的近 200 位一线教师进行德尔菲法(专家调查法)调查,以跨学科研究和人才培养机制的方法为指导,综合使用归纳推理、分类讨论等方法,进而探析大数据专业跨学科建设的模式。

3 国内外大数据专业建设的现状

3.1 国外大数据专业建设概览

2010 年,Drew Conway 提出了第一张揭示大数据学科地位的维恩图,如图 1 所示,这对数据科学的发展产生了非常重

要的影响;明确探讨了数据科学的学科定位问题,他认为数据科学处于统计学、机器学习和领域知识的交叉之处。

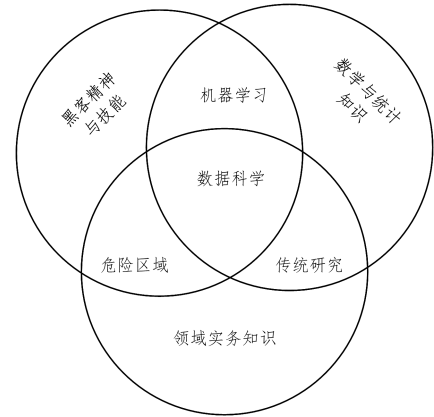


图 1 数据科学韦恩图

从图 1 可以看出,数据科学是多个学科的交叉,本文拟从跨学科角度对国外大数据专业建设现状进行分析,如表 1 所列。目前仅有少数学者对大数据专业建设进行研究,我国学者朝乐门等^[6-7]重点从特色课程角度对数据科学的内涵和发展过程等进行分析研究,提出了特色课程建设的建议 and 对策。大数据专业涉及多个学科,针对学者们多从学科层面对现有大数据专业建设进行研究,并且国内外学科目录划分有差异,本文将侧重从大的相关学科方面进行阐述,不细分每个学科的具体差异,为此将以计算机相关学科为主的专业简写为计算机,将以统计、数学学科为主的专业简写为数学 & 统计类,将以经管类学科为主的专业简写为经管类,并将与经管类相似的交通、生物等学科简称为行业类。

表 1 国外大数据专业建设的情况

地域	高校名称	专业名称	侧重方向
美国	卡内基梅隆大学	计算数据科学硕士	计算机
	斯坦福大学	数据科学	数学 & 统计类
	佐治亚州亚特兰大	分析科学硕士	经管类
	华盛顿大学	数据科学理学硕士	数学 & 统计类
	哥伦比亚大学	数据科学理学硕士	计算机
	纽约大学	数据科学理学硕士	计算机和行业类
	明尼苏达大学	数据科学理学硕士	计算机和数学 & 统计类
	西北大学	分析科学硕士	计算机
	北卡罗莱纳州立大学	分析科学硕士	计算机
	康奈尔大学	应用统计学专业硕士	数学 & 统计类
英国	德克萨斯大学奥斯汀分校	计算机科学理学硕士	计算机
	芝加哥大学	分析科学硕士	经管类
	德比大学	大数据分析	计算机
	伦敦大学	数据科学	计算机
	拉夫堡大学	网络安全与大数据	计算机
	剑桥大学	大数据	计算机和数学 & 统计类
德国	耶鲁大学	统计与数据学院	计算机和数学 & 统计类
	慕尼黑工业大学	数据科学中的数学	数学 & 统计类
	弗莱堡大学	弗莱堡数据分析与建模中心	数学 & 统计类
	海德堡大学	应用计算机科学	计算机
	柏林洪堡大学	商业分析与数据科学	经管类
印度	印度理工学院,孟买	数据科学	计算机和数学 & 统计类
	德里大学	统计系	数学 & 统计类
	安娜大学	计算机科学与工程系	计算机
意大利	天主教圣心大学	数据科学管理	计算机
	热那亚大学	计算机科学	计算机
	帕多瓦大学	软件可靠性和网络安全	计算机
	都灵大学	随机数据和数据科学	数学 & 统计类

(续表)

地域	高校名称	专业名称	侧重方向
澳大利亚	维多利亚墨尔本大学	商业分析	经管类
	昆士兰大学	数据科学	数学 & 统计类
	蒙纳士大学	数据科学	计算机
	拉筹伯大学	数据科学	计算机
加拿大	多伦多大学	数据科学	经管类
	麦基尔大学	数据科学	社会学等
	英属哥伦比亚大学	数据科学	多学科交叉
	滑铁卢大学	数据科学	计算机和数学 & 统计类
日本	东京大学	信息科学与技术系	计算机
	京都大学	数据科学	行业类
	早稻田大学	数据科学中心	多学科交叉
	神户大学	数学和数据科学中心	计算机和数学 & 统计类

注:数据来源于高校网址

3.2 国内大数据专业建设的概览与分析

我国大学的大数据专业建设呈现工学、理学和管理学发展较快而其他学科逐步发展的现状,逐渐形成了以单一学科为主而其他学科为辅的院系设置局面。其中,工学以计算机相关专业为主;理学以统计、数学等相关专业为主;管理学以经济、管理类相关专业为主。根据教育部大数据和数据科学专业申请情况和我国大数据专业建设现状,大数据专业需要工学、理学和管理学等多学科共同融合发展,但当前我国大数据专业建设呈现以单一学科的背景、理论、方法作为建设指导思路,似乎无论在制度上还是在学术资源的分配上都束缚了跨学科建设,跨学科成为现有学科的知识边缘。

综上,国内外多所高校已经开展了大数据专业建设工作,为大数据专业建设提供一手资料。通过对国内外大数据专业建设现状的分析,得出国内外大数据专业建设的异同点,如表 2 所列。

表 2 国内外大数据专业建设的分析

地域	国外	国内
是否跨学科	是	是
培养层次	硕士为主	本科大专为主
专业年限	2 年为主	3~4 年为主
课程设置	应用型为主,理论研究开设少	应用型为主,理论研究开设少
是否有成熟的跨学科课程体系	正在实践	正在实践

4 国内外大数据专业建设存在的挑战

4.1 大数据本身存在的挑战

一方面,大数据研究本身尚未成熟。国内外专家学者提出海量性、多样性、高速性和易变性等多种阐述,但大数据的内涵与外延尚未成熟,随着数据量的不断增大,种类不断增多,大数据在各个行业中的需求快速变化,为大数据科研、教学和行业标准的制定都带来多方面的挑战。在信息通信技术的快速发展下,与科学研究有关的事情都在发生巨大变化,从而兴起了数据密集型的科研范式。数据密集型科研范式还面临着一系列问题^[9]。而大数据属于数据密集型科研范式,因此对大数据的认识还存在挑战。

另一方面,大数据专业建设还需要满足“双一流”高校建设和为产业升级服务的功能。建设“双一流”高校为大数据专业建设提出新的挑战,根据文献[8]对一流学科建设的研究,我国高校大致分为以下 4 类,如图 2 所示。高起点的专业建

设为大数据专业建设带来了更大的挑战。

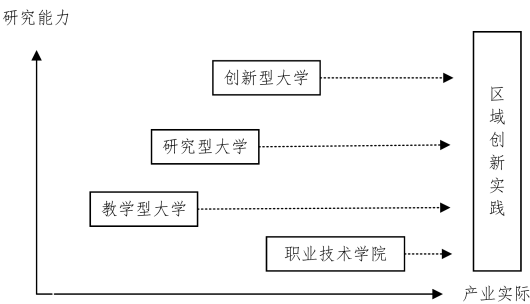


图 2 各类型大学与区域创新实践的关系

4.2 跨学科建设存在的挑战

跨学科理论研究有初步进展。国内外跨学科文献尽管有诸多差异,但从跨学科一词的基本用法来看,可以把目前的跨学科文献归结为两种类型:一种类型是把跨学科看作研究工具,相关的研究内容通常称为跨学科研究;另一种是把跨学科看作研究对象,相关的研究内容通常称为跨学科问题研究^[3]。根据文献[3]对跨学科表现形式的多样性研究,跨学科可以分为 4 类,如表 3 所列。

表 3 跨学科的分类

跨学科的分类	特征
工具借用	把从其他学科借来的工具用于解释和回答本学科的问题而产生出新的甚至重大的知识成果
多学科	出于解决实际问题的需要,虽然跳出了单个学科的范围,但是涉及到的几个学科之间是并列的关系,彼此之间几乎不发生联系,这种情况叫多学科
交叉学科	多学科团队开始彼此交流,师徒建立不同学科之间的联系,互动和整合变得活跃,涉及的学科之间发生了交叉,多学科就迈入了交叉学科阶段。这里的交叉学科是指一种研究方法,接近于文献中经常出现的跨学科,即狭义的跨学科,表示学科之间的互动关系,不同于表示新兴学科的交叉学科
超学科	有两种解释,一种是通过全面综合从而超越学科时间观的跨学科,它与跨学科的显著区别在于超学科不以学科为基础。第二种是从国家层面推动科研发展,类似于二战期间的官产学研合作科研项目

跨学科建设的机制尚在探索,存在政策支持、教育地位和传统学科体制等宏观支撑条件缺乏,院系学科融合难、人才培养的评价和激励机制不健全等中观层面的大学组织和管理机制仍未配套实行,以及人才培养的学分制、导师配备和毕业学位等微观方面的问题^[4]。

5 我国大数据专业建设的跨学科模式

大数据是科学、技术与工程的结合,大数据专业建设具有

跨学科的特性。我国大数据专业建设的难点主要在于两个方面:一个方面是如何界定学习大数据技术的范畴;另一个方面在于如何界定跨学科中学科的边界。克服大数据专业建设的难点,重点是要建立一种划分标准,该标准应该满足大数据专业建设的不同目标。为此,文章提出一种以现有 4 种类型(应用型、教学型、研究型和创新型)大学为基础,以“点”“线”“面”和“立体”相结合的我国大数据专业建设的跨学科模式,如图 3 所示。其中,横坐标 X 轴表示大学类型,对专业建设的要求逐渐递增;纵坐标 Y 轴表示跨学科程度,跨学科程度逐渐递增;Z 轴表示大数据专业知识,难度逐渐递增。该模式解决了我国大数据专业建设存在的涉及学科广、参与高校类型多,以及具体到某一高校时,该学科不同阶段(本科、硕士、博士)能力要求不同的建设需求。

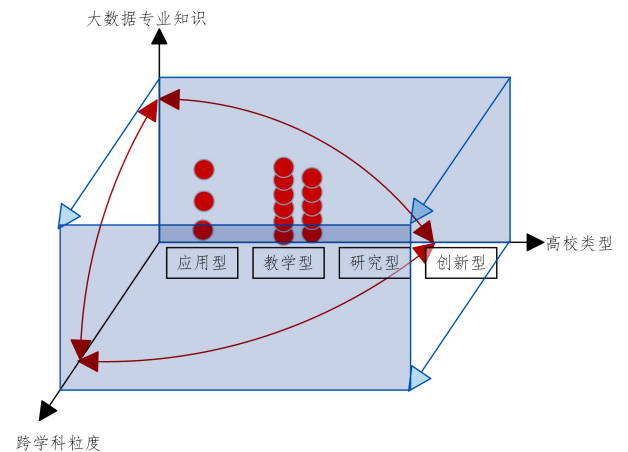


图 3 我国大数据专业建设的跨学科模式(电子版为彩色)

应用型大学的大数据专业建设,只采用大数据的实用知识作为本专业教学,不涉及跨专业,更多的是在自己的专业借用其他专业的成熟工具开展工作,以红色离散的圆点表示;教学型大学的大数据专业建设,应结合本专业的基本知识,根据 need 掌握的大数据获取、存储与管理、分析、可视化,以及安全和平台技术的某一个链条知识点,使学生能解决本链条大数据的具体应用,以红色圆点链接成线表示;研究型大学是图 3 中枣红色的 3 个曲线围城的三维曲面上的点,大学应该根据自己的特色、优势专业,结合学生情况,选取适合的大数据专业知识点进行跨学科的大数据专业教学、研讨和研究,培养学生开展研究工作的能力;创新型大学则应努力培养跨学科人才,培养学生在学习、科研和创新等多个方面的能力,图中以深蓝色的立体表示。4 种类型大学分别对应的大数据专业跨学科建设如表 4 所列。

表 4 不同类型大学与跨学科大数据建设的建议

大学的分类	跨学科特征	专业建设
应用型大学	工具借用	点
教学型大学	多学科	线
研究型大学	交叉学科	面
创新型大学	超学科	立体

6 未来工作展望

大数据专业的跨学科人才培养模式是下一步项目组研究

的重点内容。大数据跨学科专业建设的目标是培养能促进新一代信息技术与经济发展的人才。然而,大数据技术的发展与云计算、人工智能专业^[9]等新兴技术融合交织,同时与原有的计算机学科、统计学科、数学学科、经管类学科存在跨学科的关系,因此探究跨学科人才培养模式对于促进人才的培养具有重要意义。

结束语 大数据跨学科专业建设是一个系统、复杂的工程,需要事先评价、事中评价和事后评价相结合,建立切实可行的大数据跨学科专业建设模式,不仅需要与现有专家团队进行长期持续的实践与思考,还需要结合教育学、社会学等多方面的智力对课程建设等^[10]构建一种适合的新的探究范式。

致谢 特别感谢以下专家和老师。

- (1)中国科学技术协会《新一代信息技术人才培养》项目顾问:怀进鹏院士、梅宏院士等多位专家。
- (2)中国科学技术协会《新一代信息技术人才培养》项目的 3 本导论图书中《大数据导论》的主编张尧学院士、执行主编北京航空航天大学胡春明,以及编委团队的王宏志、王建民、唐杰、袁晓如、朱跃生、吴中海、吕金虎、王晨、陈恩红、刘闯、王德庆、马民虎等多位教授和老师,特别感谢本书顾问李德毅院士和百度高级副总裁王海峰老师。
- (3)中国科学技术协会《网络与大数据学科方向预测及技术路线图——大数据学科发展研究报告》项目,北京大数据研究院鄂维南院士、大数据分析技术创新中心主任傅毅明及其团队成员。
- (4)“全国高校大数据师资研修班”(两期)项目中近 200 位学员老师和讲课专家。
- (5)中国电子学会林润华副秘书长,以及为完成该项目的各位专家和老师的团队团队成员。

参 考 文 献

[1] 教育部. 教育部关于公布 2017 年度普通高等学校本科专业备案和审批结果的通知[OL]. http://www.moe.gov.cn/srcsite/A08/moe_1034/s4930/201803/t20180321_330874.html.

[2] 傅存良. 如何进行跨学科研究[M]. 北京:北京大学出版社, 2016.

[3] 刘小宝. 论“跨学科”的谱系[D]. 合肥:中国科学技术大学, 2013.

[4] 伍红军. 学科框架内的有限颠覆——交叉学科研究生培养机制研究[M]. 上海:上海交通大学出版社, 2015.

[5] 中国科学院国家自然科学基金委会. 未来 10 年中国学科发展战略——总论[M]. 北京:科学出版社, 2012:114-115

[6] 朝乐门,邢春晓,张勇. 数据科学研究的现状与趋势[J]. 计算机科学, 2018, 45(1):3.

[7] 朝乐门,邢春晓,王雨晴. 数据科学与大数据技术专业特色课程研究[J]. 计算机科学, 2018, 45(3):189-206.

[8] 李春林. 基于知识网络的创新型大学一流学科生产机理研究[D]. 哈尔滨:哈尔滨工业大学, 2013.

[9] 扬博雄,李社蕾. 新一代人工智能学科的专业建设与课程设置研究[J]. 计算机教育, 2018(10).

[10] 郭华. 课程·教材·教法[D]. 北京:人民教育出版社, 2018.