



对国外 iSchools 核心院校研究生项目数据科学类课程的调查与思考*

□赵蕊菡 王敏桦

摘要 在大数据时代,发展数据科学教育、培养数据科学人才已成为图书情报专业提升专业核心竞争力、优化人才培养质量的重要路径。对 iSchools 学院的数据科学类课程设置现状进行系统深入研究,将有力推动我国数据科学教育,有利于提升图书情报专业核心竞争力。文章以国外 31 所 iSchools 核心院校的研究生项目为对象,基于“数据生命周期—人才培养目标—课程体系”的对应关系对其数据科学类课程内容进行统计,总结 iSchools 核心院校数据科学类课程的特征,并建议我国图书情报专业设置数据科学类课程要体现专业特色、面向国家战略和社会需求探索研究生分类课程教学体系、拓展课程知识应用领域、丰富课程实践教学形式。

关键词 iSchools 数据科学 课程设置 数据生命周期

分类号 G250

DOI 10.16603/j.issn1002-1027.2023.02.009

1 引言

近年来,伴随着数据密集型产业的发展,数据科学取得了飞速发展,社会各行业都急需能够高效处理、分析并深入利用大数据的专业人才。据美国劳工统计局(U.S. Bureau of Labor Statistics)预测,到 2028 年,数据科学专业学生的就业机会将增长 16%,远高于所有职业的平均水平^[1]。世界各国充分认同数据科学人才的重要价值,并高度重视其学科发展和人才培养。2015 年和 2017 年,我国教育部先后新增“数据科学与大数据技术”“大数据管理与应用”两个与数据科学密切相关的本科专业,多所院校开始招收数据科学方向的硕士研究生。这些数据科学相关专业多数位于计算机学院、统计学院或经济管理学院,分布在数学、软件工程、信息与通信工程、管理科学与工程、计算机科学与技术、图书情报与档案管理学科中,人才培养范式向更多元宽泛的应用方向拓展^[2]。

作为较早开始数据科学研究的人文社会学科之一,图书情报专业应承担起培养数据科学专业人才的职责,将数据科学教育作为人才培养和科学研究

新的增长点^[3]。为此,2019 年,全球拔尖信息学院联盟(iSchools)组建 iSchools 数据科学教学指导委员会(The iSchools Data Science Curriculum Committee, iDSCC),以更好地发展和推进图书情报学科视角下的数据科学教育^[4]。目前,全球接近 1/4 的数据科学教育项目来自 iSchools 成员,但我国目前仅有武汉大学信息管理学院形成了数据科学本科、硕士、博士完备的多层次人才培养体系^[5]。

通过文献调研可知,学者已经认识到数据科学教育对图书情报学院持续发展和人才培养质量提升的重要意义。陶俊等认为嵌入数据分析能力有助于提升图书情报专业核心竞争力^[6];王晰巍等认为培养研究生的数据素养能力、学科融合能力及理论与实践相结合能力是未来图书情报学科研究生人才培养的趋势^[7];吴丹等对比 iSchools 院校数据科学人才培养的相关数据,认为应审时度势、考虑学科基础并提供线上线下混合的人才培养模式^[8]。而伴随着“新文科”的发展趋势,宋(Song Il-Yeol)等认为 iSchools 应抓住数据科学教育中的机遇,培养学生使用数据工具解决问题以及创造价值、信息和知识

* 2021 年郑州航空工业管理学院教育教学改革研究与实践项目(编号:zhjy21-21)、“面向图情专硕人才培养的‘航空特色’思政教学案例库建设研究”(编号:2021YJSJG07)的研究成果之一。

通讯作者:赵蕊菡,邮箱:xiaoye713@126.com。



的能力^[9];朝乐门等基于对数据科学发展阶段的梳理,认为数据科学的发展离不开专业人才培养的体系化^[10]。针对图书情报学科如何开展数据科学人才培养和课程设置,吴丹提出“以人为本”将是图书情报学科视角下开展数据科学教育的核心特征,具体落脚于学生数据意识和数据素养的塑造、以问题与情境驱动解决数据问题的能力和方案等^[11],其团队基于图书情报、经济与工商管理、计算机科学与技术、数学与统计等四个学科的数据科学相关硕士专业特征进行了对比分析^[12],构建了包括数据驱动的方法和技术、领域知识、数据的法律、道德和伦理、塑造发展个人特质等四个模块的数据科学研究生课程模式^[13]。此外,学者们借鉴国外图书情报院校在数据科学专业建设^[14]、教育项目^[15]、课程体系设置^[16]、人才培养模式^[17]、MOOC 课程体系^[18]等方面的经验,认为数据科学为图书情报学科体系设计和学科特色提供了新的发展方向,有助于更好地适应国家战略和社会应用的多方现实需求^[19],图书情报学科应将数据科学作为新的契合点和增长点^[20]。

但当前研究多是基于已有的数据科学课程体系开展分析,并未从数据生命周期的视角整体审视现有课程体系存在的不足,也未能体现人才培养目标与课程体系的对应关系。因此,本文将在已有研究的基础上,调查 iSchools 核心成员院校研究生项目中数据科学相关课程的设置情况,并基于数据生命周期理论分析其在培养目标和课程体系上的特点,为我国图书情报专业开设系统的数据科学相关课程、培养具有扎实理论基础和专业技能的数据科学人才提供参考借鉴。

2 调查设计

2.1 调查对象和数据收集方法

iSchools 按机构层级分为 6 个级别,其中,iCaucus 作为 iSchools 的核心成员,在联盟中发挥着领导核心和决策作用,其发展现状代表了全球图书情报领域的最高水平,截至 2021 年 10 月,iSchools 共有 39 所 iCaucus 院校(8 所位于中国)。因此,本研究通过网络调查法,对国外的 31 所 iCaucus 院校网站进行访问,以其开设的与图书情报领域高度相关的研究生项目作为主要调查对象。具体调查步骤为:(1)浏览调查院校的官方网站,获取研究生项目目录;(2)获取各研究生项目人才培养相关信息;(3)

获取各研究生项目课程相关信息,并根据培养要求区分核心课程(Core Courses)和选修课程(Elective Courses)。最终,本研究共收集了 31 所 iCaucus 院校面向研究生开设的 4117 门课程(核心课程 901 门,选修课程 3216 门),其中有 795 门课程(核心课程 193 门,选修课程 602 门)与数据科学相关,具体情况见表 1。

2.2 数据分析方法

教育的目的是培养人才,人才培养目标是教育教学活动的出发点和归宿,不同的培养目标对学生的知识、能力和素质结构提出了不同的要求。而课程体系是实现培养目标的重要载体和手段,课程体系的设置和实施情况将直接影响到人才培养目标的实现。因此,本文依据“数据生命周期—人才培养目标—课程体系”的对应关系,分析国外 iSchools 核心院校研究生项目所开设数据科学类课程的特征。

(1) 文本内容分析

iSchools 教育理念以“信息(数据)”“技术”“人”三要素之间的关系为核心^[21],并围绕该理念设计人才培养目标的知识 and 能力结构。因此,本研究依据“信息(数据)—技术—人”三要素对 iCaucus 院校的研究生人才培养相关内容(如项目介绍、项目目标、毕业生核心竞争力等)进行文本内容分析,并将解析出的知识目标和能力目标按照数据生命周期进行归类,以识别人才培养目标与课程体系的对应关系。

文本内容分析是社会科学领域常用的研究方法之一,通过将文本内容划分为特定类目,计算每类内容元素出现频率。该方法结合了定性和定量研究方法,能够客观、系统地描述文本内容特征。本研究利用 NVivo 质性分析软件(版本 12.0)作为内容分析工具,通过内容编码发现人才培养目标文本的内在逻辑关联,通过节点数目发现 iCaucus 院校数据科学人才培养的重点环节。分析过程由两位作者分别进行编码,保证分析结果的有效性和可靠性。

(2) 描述性统计

数据生命周期用来描述数据从产生到组织、发布以及再利用的数据增值过程^[22]。当前国内外已经建立多个具有代表性的数据生命周期模型,如 DataONE 数据生命周期模型^[23]、DCC 数据管理生命周期模型^[24]等。学者还基于不同的侧重点构建数据生命周期,如基于 ACRL 信息素养能力标准将数据生命周期划分为数据收集、数据组织、数据分析、



表 1 iCaucus 院校开设数据科学相关课程基本情况

iSchools 院校		数据素养能力课程开设情况			开设层次	
大学	学院	开设数量 (门)	总课程数量 (门)	数据素养能力课程占比 (%)	必修课占比 (%)	选修课占比 (%)
匹兹堡大学	计算机与信息学院	71	341	20.82	9.86	90.14
印第安纳大学伯明顿分校	信息学、计算机与工程学院	61	350	17.43	59.02	40.98
佐治亚理工学院	计算机学院	51	309	16.50	25.49	74.51
卡耐基梅隆大学	海因茨学院	49	309	15.86	26.53	73.47
亚利桑那大学	信息学院	45	108	41.67	6.67	93.33
密西根大学	信息学院	44	101	43.56	77.27	22.73
加州大学欧文分校	信息与计算机科学学院	44	303	14.52	2.27	97.73
北卡罗莱纳大学	信息与图书馆科学学院	36	152	23.68	8.33	91.67
康奈尔大学	计算机与信息科学学院	36	160	22.50	36.11	63.89
德雷塞尔大学	计算机与信息学学院	31	96	32.29	32.26	67.74
马里兰大学	信息研究院	30	164	18.29	10.00	90.00
巴尔的摩马里兰大学	信息系统系	28	117	23.93	17.86	82.14
华盛顿大学	信息学院	26	133	19.55	3.85	96.15
北德克萨斯大学	信息学院	26	153	16.99	15.38	84.62
蒙纳士大学	信息技术学院	25	45	55.56	20.00	80.00
雪城大学	信息研究院	25	136	18.38	28.00	72.00
伊利诺伊大学香槟分校	信息科学学院	22	103	21.36	4.55	95.45
不列颠哥伦比亚大学	信息学院	18	130	13.85	5.56	94.44
加州大学伯克利分校	信息学院	17	98	17.35	35.29	64.71
谢菲尔德大学	信息学院	13	48	27.08	61.54	38.46
佛罗里达州立大学	传播与信息学院	13	91	14.29	46.15	53.85
加州大学洛杉矶分校	教育与信息研究院	13	118	11.02	0.00	100.00
罗切斯特理工学院	计算与信息科学学院	12	61	19.67	41.67	58.33
罗格斯新泽西州立大学	传播与信息学院	11	81	13.58	18.18	81.82
宾夕法尼亚州立大学	信息科学与技术学院	10	56	17.86	0.00	100.00
圣何塞州立大学	信息学院	9	61	14.75	22.22	77.78
田纳西大学	信息科学学院	9	72	12.50	11.11	88.89
肯塔基大学	传播与信息学院	7	51	13.73	0.00	100.00
柏林洪堡大学	柏林图书馆与信息科学学院	6	25	24.00	16.67	83.33
德克萨斯大学奥斯汀分校	信息学院	4	56	7.14	50.00	50.00
密苏里大学	信息科学与学习技术学院	3	89	3.37	0.00	100.00
总计		795	4117	21.37	24.28	75.72

数据可视化以及数据再利用等阶段^[25];基于图书馆数据管理流程划分为数据收集、分析、获取和复用四个阶段^[26];基于数据科学教育内容划分为业务理解、数据理解、数据准备、模型规划、模型构建、评估、部署以及审查监控等八个阶段^[27];基于具体学科(如生命科学)划分为数据管理计划、数据采集、数据组织、数据保存、数据共享利用五个阶段^[28];基于数据生态链划分为数据产生或获取、数据描述与组织、数据处理与分析、数据保存或归档、数据出版或共享

等五个阶段^[29]。此外,基于项目运作迭代规律,数据科学生命周期还可以划分为了解业务、数据采集和理解、建模、部署、客户验收等阶段^[30]。

通过对代表性数据生命周期模型的梳理发现,尽管不同数据生命周期模型的使用领域和划分细节存在差异,但都紧紧围绕着数据从产生、保存到利用的整个过程。结合数据科学人才培养的知识目标和能力目标,数据生命周期不仅应包括数据管理过程(如数据采集、数据组织、数据分析、数据存储和数据

共享等),还要能够全面综合体现利用数据开展的知识抽取过程^[31],包括数据意识、数据表达、数据决策、数据重用及数据科学的道德与法规知识等。因此,结合数据科学教育知识更新快、技术要求高、专业交叉融合的特点,本研究将知识抽取过程与数据管理过程相结合,以数据生命周期中的数据意识与计划、数据采集与处理、数据分析与揭示、数据保存与安全、数据利用与共享等五个阶段为切入点,调研 iCaucus 院校数据科学类课程设置的合理性和完整性,并探讨在每个数据生命周期阶段对学生的知识和能力要求。数据生命周期主要阶段及其内涵见图 1。



图 1 数据生命周期主要阶段及其内涵

3 调查结果

3.1 iCaucus 院校人才培养目标分析

培养目标是对本专业人才的基本素质、服务面向、人才定位的总体描述。本研究围绕 iSchools 教育理念“信息(数据)—技术—人”三要素,对 iCaucus 院校的研究生人才培养目标文本内容进行编码,并与数据生命周期阶段进行对应,分析结果见表 2。

在“信息(数据)”要素中,iCaucus 院校从数据生产的视角对学生的能力目标提出具体要求,学生要具备数据意识、数据收集、数据组织、数据分析、数据存储、数据决策、数据共享和数据重用等复合能力,培养目标对应于数据生命周期全阶段。如加州大学伯克利分校的数据科学硕士学位要求学生具有处理多学科大规模数据集的能力,能够从整个数据生命周期视角开展数据科学的全过程研究(包括识别问题、数据检索、清理、建模和交流结果)。在此要素的目标导向下,iCaucus 院校培养的人才将在非营利的博物馆和档案馆等文化机构和网络搜索门户等信息密集型环境中从事数据管理、数据科学和分析、数字档案、数字人文科学等工作,主要职位包括编目/元

数据馆员、高技能数据科学和数据分析专业人员、数据管理首席规划师等。

iCaucus 院校认为“技术”要素是传递信息和数据的手段,为数据科学提供工具性的支持,因此要求学生在知识目标上掌握大数据管理和数据安全知识,在能力目标上具备数据工具、数据挖掘和数据评估能力。如北卡罗莱纳大学要求信息科学硕士学位获得者掌握系统分析和设计、数据库管理、信息检索和挖掘、知识管理系统、用户界面设计、用户体验等多种新兴技术。技术要素相关的人才培养目标不涉及数据共享与利用阶段,在此要素的目标导向下,其毕业生将在出版社、创业企业、互联网服务提供商等组织中从事信息系统管理、分析挖掘和算法、可视化与交互相关工作,主要职位包括信息系统专家、计算机专家、图书馆技术服务员、交互设计师、用户体验设计师、信息架构师等。

“人”的要素是 iSchools 教育的最终归宿和落脚点,因此 iCaucus 院校要求学生能够将合适的信息以合适的方式在合适的时间传递给合适的人,包括数据科学、数据规则和数据伦理等知识和数据表达能力。如德雷塞尔大学认为其数据科学专业学生应从海量数据中发现隐藏趋势、制定和实施数据行动方案,并有效地管理和展示数据。人的要素相关的人才培养目标不涉及数据采集与处理阶段,在此要素的目标导向下,其毕业生将在各类型图书馆、政府机构、高等教育机构中从事参考咨询、用户教育、信息服务等以人为本的工作,主要职位包括图书馆/博物馆/档案馆馆员、研究员、信息服务人员等。

3.2 iCaucus 院校数据科学类课程分析

iCaucus 院校所开设的数据科学类课程在总课程中所占比例达到 21.37%,较之前有较大提升^[6],近半数院校的数据科学类课程占比都超过了 20%。课程对人才培养目标起到支撑作用,因此其设置和建设都要紧扣人才培养目标,明确课程内容与人才培养目标的直接关联。如前所述,本研究按照“数据生命周期—人才培养目标—课程体系”的对应关系对数据科学类课程的数量和主题进行统计,结果见表 3。其中部分课程涉及多个人才培养目标,在计数时进行分别统计,如巴尔的摩马里兰大学开设“数据分析:R 在医疗保健中的应用”课程,其内容涉及了数据分析能力和数据决策能力,在统计时分别在对应目标的课程数量上加 1 进行计数。



表 2 iCaucus 院校研究生人才培养目标文本分析

教育理念要素 (节点))	人才培养目标 (子节点)	数据生命周 期对应阶段	节点 数量	代表性描述内容
信息/数据	数据意识能力	数据意识与 计划阶段	26	毕业生能够为各种领域创造价值,包括医疗保健、金融和业务运营
	数据收集能力	数据采集与 处理阶段	13	检索跨广泛信息领域的多种媒体中的数据和信息
	数据组织能力		30	使数据和元数据可查找、可访问、可互操作
	数据分析能力	数据分析与 揭示阶段	31	运用数据分析深入研究数据科学的业务应用
	数据存储能力	数据保存与 安全阶段	17	处理数字材料和模拟材料的数字化副本的保存、组织和存储
	数据决策能力	数据利用与 共享阶段	24	将数据转化为应对现实挑战的解决方案
	数据共享能力		2	共享数据和代码的工作流
	数据重用能力		2	使数据在……可用/可重用方面发挥着关键作用
技术	大数据管理知识	数据意识与 计划阶段	43	在复杂系统中管理海量、“嘈杂”和快速增长的数据集
	数据评估能力	数据采集与 处理阶段	17	对选定的案例研究和数据进行科学的评估
	数据工具能力	数据分析与 揭示阶段	31	掌握强大的软件工程技能,能够有效地开发数据科学应用程序
	数据挖掘能力		8	学习如何使用数据挖掘
	数据安全知识	数据保存与 安全阶段	5	确保负责任地使用数据
人	数据科学知识	数据意识与 计划阶段	19	应用数据科学方法、开发工具和算法以及管理数据和信息来解决以数据为中心的问题、获得洞察力和交流
	数据表达能力	数据分析与 揭示阶段	18	学习如何有效地向关键利益相关者展示结果
	数据规则知识	数据保存与 安全阶段	2	数据管理的法律框架
	数据伦理知识	数据利用与 共享阶段	10	对数字数据的伦理处理,以避免可能影响结果的偏见

(1) 数据意识与计划阶段

该阶段相关课程主要介绍数据科学和大数据管理的基本知识,培养学生利用数据识别复杂问题的数据意识能力。

数据科学知识相关课程介绍数据科学基本知识,如伊利诺伊大学香槟分校的“IS407 数据科学导论”和“IS477 数据管理、保存和再现”课程,讲授数据科学的管理方法和常用技术,培养学生创建数据科学管理计划的能力;华盛顿大学更是为信息管理科学硕士开设四门数据科学导论课程,系统介绍数

据科学的理论技术、技术方法、应用与道德等关键知识。

大数据管理知识是数据科学的重要内容,代表性的如印第安纳大学伯明顿分校的“INFO-I535 大数据和复杂数据的管理、访问和使用”课程,重点介绍大数据的知识表示、数据管理和数据管理等内容,“ILS-Z645 大数据的社会和组织信息学”课程则从社会和组织信息学的角度介绍大数据创建、传播和使用过程中的组织、法律、政治和社会知识。

数据意识能力培养学生在具体应用情境(如健康、



表 3 数据生命周期—人才培养目标—数据科学类课程体系的关联矩阵

数据生命周期	人才培养目标	数据科学类课程体系	
		课程数量(门)	课程主题
数据意识与计划阶段	数据科学知识	48	数据科学的关键概念、技能和技术等基本知识,培养学生的数据意识
	大数据知识	57	大数据的基础知识和核心技术,能够通过大数据系统进行实践
	数据意识能力	60	利用数据识别实际应用中的复杂问题(如政策、环境、营销、经济等)
数据采集与处理阶段	数据收集能力	42	数据检索和整理;数据采集和预处理的技术
	数据组织能力	129	数据组织方法和工具;数据组织需求评估
	数据评估能力	0	考察和评价数据质量,评估的框架、指标和方法
数据分析与揭示阶段	数据工具能力	165	常用的数据处理和数据分析的工具、技术和应用场景
	数据挖掘能力	42	数据挖掘的概念和技术,包括数据预处理、关联的规则,数据挖掘趋势和应用情境
	数据分析能力	138	数据分析基础知识和技术,包括建立数据分析方案,应用分析方法和工具,进行实验数据分析,评估分析结果等
数据保存与安全阶段	数据表达能力	59	对数据分析结果进行解释和评估;使用数据可视化开展数据表达和交流
	数据规则知识	32	数据环境建设;数据政策和法律
	数据安全知识	31	数据防护、处理和存储的安全措施,数据安全要求(例如访问权限、存储介质等)
数据利用与共享阶段	数据存储能力	66	根据不同的应用环境需求,有效保存数据
	数据伦理知识	24	在数据收集、存储、处理和利用过程中产生的伦理挑战的应对办法和工具
	数据决策能力	52	从数据中提取信息的优先级;将数据转化为可操作的信息并执行决策或解决方案
	数据共享能力	0	共享方案设计、共享平台构建,共享政策、方式、流程,数据整合和开放等
	数据重用能力	0	数据可重用的解决方案,相关政策、法律、技术、成本控制,数据监护人员教育、数据重用参与机构管理等

商业、政府、社交媒体、地理空间等)中利用数据发现问题的能力。如田纳西大学的“INSC543 空间数据管理”课程,主要介绍空间数据相关的采集、管理和元数据创建等知识;圣何塞州立大学的信息学理学硕士可以选择健康学研究方向,并学习“INFM214 健康数据和分析”课程,掌握医疗数据的挖掘和管理等知识,用以解决实际医疗业务和临床智能问题。

(2) 数据采集与处理阶段

随着异构多源数据的爆炸式增长,该阶段关注培养学生对数据进行智能采集、组织加工和质量评估的相关能力。

数据收集能力相关课程培养学生执行数据解析、采集、清理和检索的能力,如德雷塞尔大学开设的“DSCI511 数据采集和预处理”课程,从项目生命周期的角度介绍数据采集、获取和预处理知识;伊利诺伊大学香槟分校除了传统信息检索课程外,还开设了“IS537 数据清理的理论与实践”课程,使学生掌握数据预处理和数据清理的相关知识和工具。

数据组织能力相关课程培养学生掌握数据库管

理、信息组织和元数据等工具以组织数据的能力,多数为 iSchools 传统课程,也有部分课程聚焦数据环境下的数据组织新进展,如不列颠哥伦比亚大学的“LIBR514 元数据”课程,主要关注在开放数据环境中的元数据管理模式;印第安纳大学伯明顿分校的“ILS-Z63 数据语义”课程,着重探索语义 Web 技术在数据组织中的应用。

(3) 数据分析与揭示阶段

本阶段相关课程培养学生对数据进行分析、处理等操作进而提炼数据价值、揭示数据中的情报和知识的能力,课程数量和内容都较多。

数据工具能力相关课程培养学生根据需要使用合适的工具或技术的能力,课程内容为 R、Python、机器学习、自然语言处理、云计算和人工智能等常用的数据工具。相关课程将讲授如何使用 R 和 Python 等编程语言完成数据识别、分析和提取,如加州大学伯克利分校开设的“InfoW18 数据科学的 Python 基础”课程;介绍在数据科学中常见的机器学习方法,如北德克萨斯大学的“INFO 5505 面



向数据科学家的应用机器学习”课程;讲授人工智能的理论、算法和应用等,如康奈尔大学的“INFO 5600 医疗保健人工智能”课程。此外,用以处理海量非结构化的网络数据的自然语言处理课程也较多,亚利桑那大学就从情感分析、统计建模等角度开设多门自然语言处理课程。

数据挖掘能力相关课程培养学生利用编程语言执行数据分类、聚类、关联和预测的能力,如匹兹堡大学的“INFSCI2160 数据挖掘”课程,重点介绍数据挖掘原理、关联规则、决策树和聚类分析等数据挖掘技术。该能力注重在真实情境中挖掘出信息和知识,如德雷塞尔大学“INFO634 数据挖掘”课程,重点介绍数据挖掘技术在商业、政府、医疗等信息密集型环境中的功能。

数据分析能力相关课程培养学生利用工具建立数据分析方案,对数据进行汇总、理解和分析的能力,这类课程既是 iSchools 院校传统优势课程,也是数据生命周期的核心环节。北德克萨斯大学开设了“人类语言技术中的数据分析 I&II”“数据分析与知识发现”和“信息专业人员的数据建模”等多门数据分析课程,系统讲授开展超大数据集分析和预测的相关知识;密西根大学特别针对健康信息学方向开设了“LHS610 健康探索性数据分析”课程,学生使用真实医疗数据集开展数据分析和推断。随着数据挖掘和数据分析的联系越来越紧密,也出现了讲解如何使用编程工具(如 SAS、R 等)进行数据分析的课程。

数据表达能力相关课程培养学生对复杂高维数据的分析结果进行解释、评估、呈现和交流的能力,着力培养数据可视化能力。如伊利诺伊大学香槟分校开设多门相关课程讲授数据创建、可视化聚合和解释的技术;亚利桑那大学利用来自 Netflix、IMDB、DBLP 等公开真实大型数据集开展数据可视化课堂实践;密西根大学开设“SIADS523 交流数据科学成果”课程,教授学生构建有效的交流策略以展示数据分析成果。由于数据分析与表达存在密切的能力递进关系,许多学校直接将这两种能力的内容合并在一门课程中,如肯塔基大学的“ICT662 数据分析和可视化”课程,内容涵盖数据分析和可视化的基本方法和关键实践技能。

(4)数据保存与安全阶段

本阶段要求学生掌握数据规则知识和数据安全

知识,同时具备数据存储能力。

数据规则知识主要指为了规范数据处理活动,保障数据安全,促进数据开发利用而制定的数据政策、数据主权、数据标准等知识。代表性的如伊利诺伊大学香槟分校“IS467 数据科学伦理与政策”课程,主要介绍组织、机构、政府和超国家层面的一系列数据相关政策;华盛顿大学的“LIS555 数据主权和本土知识系统”课程,主要讨论数据产权的建立机制和协议,以促进区域内数据的管理、保护和传播。

数据安全知识相关课程要求学生掌握数据防护、处理和存储的安全措施,并能够根据数据安全要求评估访问权限和存储介质。如佐治亚理工学院在传统的信息安全课程的基础上,又增设了“INTA 6450 数据分析和安全”课程,研究数据保密、数据完整性、数据权属、身份认证等数据安全问题。

数据存储能力培养学生根据不同的应用环境需求,有效进行数据存储、归档和维护的能力。伊利诺伊大学香槟分校开设了一系列相关课程,如“IS547 数据监护基础”课程讲授数据监护服务的基础知识;“IS543 数字保存”课程研究复杂数字对象的保存、标准规范、风险评估和管理保存活动。

(5)数据利用与共享阶段

本阶段介绍数据伦理知识,着力培养学生使用数据进行决策、共享和重用的能力。

数据伦理知识相关课程主要培养学生掌握与数据相关的法律和伦理问题的能力,往往使用真实数据开展研究。如北卡罗莱纳大学开设了“INLS584 信息伦理”“INLS384 信息与计算机伦理”和“INLS774 应用数据伦理”等三门与数据伦理相关的课程,从知识产权、隐私、访问/审查和道德行为准则等角度研究数据科学家在创建、收集、管理和使用数据时面临的道德问题;加州大学伯克利分校的“INFO188 数据背后:人与价值观”课程,使用包括Facebook的“情绪传染”实验、社交媒体算法、自动驾驶汽车等真实数据伦理案例,引导学生讨论解决数据伦理问题的框架、流程和工具。

数据决策能力相关课程培养学生在复杂情境下对数据的利用能力,包括从数据中提取优先级、将数据转化为可操作的信息、评价数据质量并据此制定决策或解决方案。如亚利桑那大学的“LIS634 医疗系统中的数据管理”课程,使用数据解决医疗保健问题,“INFO514 计算社会科学”课程,使用社交网络、在



线社区、电子商务和数字营销等数据,解决多领域的社会问题;匹兹堡大学的相关课程如“INFSCI2809 研究空间数据”和“LIS2188 开放政府数据”等,主要介绍特定领域的数据驱动决策过程。

4 iCaucus 院校数据科学类课程设置特征

4.1 课程体系覆盖数据生命周期阶段,但部分人才培养目标缺乏对应课程

依据本研究提出的数据生命周期模型,按照“数据生命周期—人才培养目标—课程体系”的对应关系构建数据科学类课程体系(见图2)。括号中数字为 iCaucus 院校已开设相关课程数量,后为代表性课程名称;如数字为 0,则说明暂未开设与该培养目标相关的课程。

从图2可见,当前开设的数据科学类课程已经充分覆盖了完整的数据生命周期,这说明 iCaucus 院校在课程设置时遵循数据科学运作规律。加州大学伯克利分校认为,数据科学家必须掌握数据科学生命周期的全部内容,并依据数据科学项目运作规律在捕获、维护、过程、分析、沟通等数据生命周期阶段获得收益回报^[32];匹兹堡大学、密苏里大学、华盛顿大学等 iCaucus 院校的人才培养相关信息里也都提到依据数据(信息)生命周期设置课程,以使学生在复杂的信息环境中获得洞察力。

但从“人才培养目标—课程体系”的对应关系来

看,各培养目标对应的课程数量严重不均,特别是与数据伦理、数据规则和数据安全等知识相关的课程数量较少。一方面,信息科学领域已有相关课程,如“信息伦理”“信息安全”等,但课程内容未向数据科学领域转型;另一方面,这些课程数量相较先前调查^[33]已有增长,但整体数量仍较少,并不能解决当前数据伦理困境。在 iSchools 教育理念引领下,人的要素是人才培养的落脚点,而近年来数据开发利用背后隐藏的安全与伦理风险日益引起关注,数据泄露、过度采集、非法交易等乱象严重阻碍产业的数智化进程,因此,图书情报专业应该更加重视数据科学课程体系中的价值观要素,通过培养学生的数据意识,树立数据伦理来权衡技术理性与价值理性^[11]。

iCaucus 院校人才培养目标还包括培养学生的数据评估、共享和重用能力,但并未开设相应的独立课程。数据评估能力可以使学生从数据的完整性、准确性、时效性和一致性等角度对数据的采集、存储和产出质量进行全面的考察和评价,从而提高数据的可信度和有效度,为数据决策提供有利的基础,但相关知识分散于数据采集处理、分析揭示和保存安全等阶段的课程内容中;数据共享和重用可以避免不必要的重复试验,加速新发现,涉及了数据共享管理(如共享的持续性、查找性、可访问性、互操作性和盈利的透明度)和再利用(如原始数据的完整性、信任和评估、二次使用、数据监督)等知识^[34],但零散

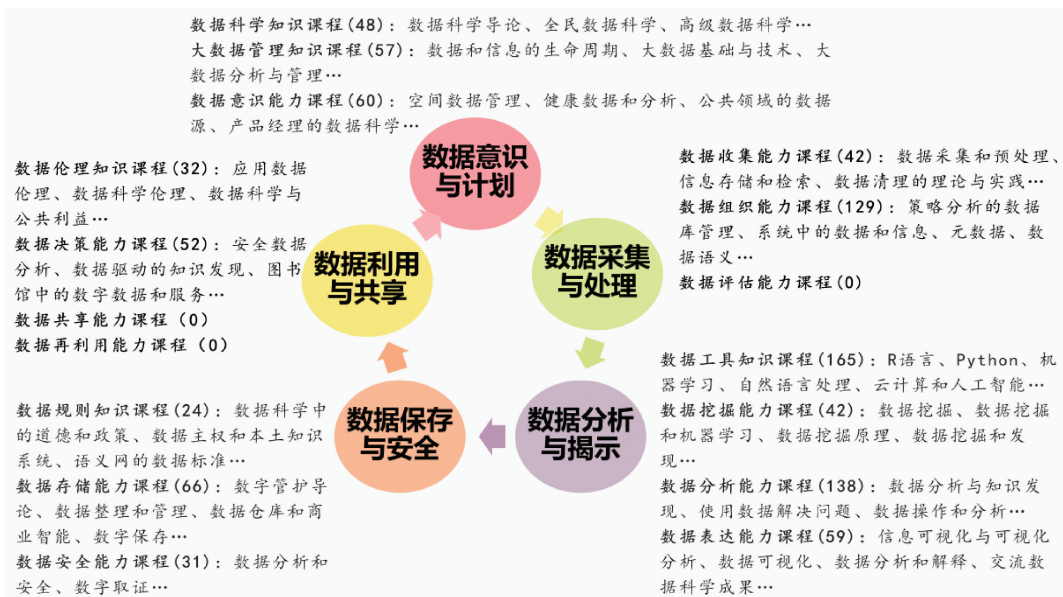


图2 基于数据生命周期的数据科学课程体系



分布于数据意识计划、数据利用共享等阶段的课程内容中。独立课程的缺失将难以支撑培养目标的实现,缺乏对相关知识的系统掌握也将限制其毕业生的能力。

4.2 侧重于大数据环境下的图书情报专业核心能力纵深发展

根据对大数据专业培养内容的分析,数据科学已经融入了信息组织、信息检索和信息分析的各个层次^[2]。调查显示,iCaucus 院校开设数量最多的数据科学类课程集中在数据组织处理和分析揭示阶段,对应培养目标中的数据工具、数据组织和数据分析能力。而这些能力正是图书情报与数据科学交叉融合时对本学科核心知识体系的深化和拓展。

作为数据密集驱动的研究,数据科学教育的核心任务就是通过对已有和待预测的数据进行分析,产生对大数据的理解、利用和管理,最终将数据转化为知识^[35]。因此,数据科学家的核心能力是数据分析和机器学习,特别是掌握基于数据的深度学习模型。iCaucus 院校要求学生掌握数据分析工具,如 Matlab、Stata、SPSS 等,同时掌握 Python、R 语言、SQL、Hadoop 和 Spark 等数据科学的主要语言和人工智能、深度学习等机器学习技能。从课程性质上来看,与数据挖掘、数据表达能力相关的课程最常被设置为必修课,体现了 iSchools 院校从“以人为本”出发,培养学生扎实的数据挖掘能力以解决“数据丰富但信息匮乏”的现状,培养学生构建高度交互的用户界面和可视化结果的能力,以便于用户与系统交互。

信息组织课程体现图书情报专业对文献、信息、数据进行描述、标引、揭示,使之有序化的过程,培养学生将无序的、分散的数据和信息整理成有序的信息资源的能力,是图书情报专业的核心能力之一。因此,数据结构与算法、数据语义、非结构化数据库、分类与标引、元数据等与数据组织能力相关的知识内容成为 iCaucus 院校的核心课程内容。学生使用真实数据集开展数据的分类、语义分析、受控词表创建、元数据标准创建和使用等学习和实践,体现图书情报专业知识在数据的分类、描述、约减、评估、交换共享等方面的重要作用。

4.3 强调情境驱动的数据科学教育,形成阶梯式课程组

随着大数据的广泛应用,数据科学的数据来源包括了科学数据、社交媒体数据、政务数据、健康医

疗数据、地理空间数据、天文数据等多个领域。iSchools 院校重视培养其学生在复杂真实数据情境下的数据意识、数据分析、数据争论、数据表达、数据规范和数据伦理等核心能力,以提高学生在大数据时代的适应性。如印第安纳大学伯明顿分校允许其数据科学硕士生选修生物信息学相关课程,学生将通过大规模生物数据集的建模、分析和管理,推动生物学、农业和生物技术等领域的进步;卡耐基梅隆大学基于 Covid-19 数据集,开设了“94-892 与 Covid 交流:利用数据和可视化建立信任并采取行动”课程,学生使用 Tableau 等数据工具开展数据分析、可视化、解释和交流,为疫情下的行为设计和政策制定提供参考。以情境驱动的数据科学教育既能提高图书情报学科的整体学科活力,又能增强学生的核心能力,拓展数据素养高级技能,开展学科化数据素养教育^[36]。

研究生专业基础知识参差不齐,为此,iSchools 院校设置阶梯式的课程或课程组,便于学生逐步掌握数据科学相关知识。如巴尔的摩马里兰大学开设有关数据库和数据挖掘系列课程,课程内容涉及了大数据管理知识、数据组织和数据挖掘能力等;亚利桑那大学的情报学硕士下设有“以人为本的计算”和“机器学习”两个方向,分别设置不同的核心课程组,如机器学习方向要求学生在学习“信息基础”“机器学习简介”和“数据分析和可视化”三门必修课之后,在“贝叶斯建模和推理”“数据挖掘和发现”“统计自然语言处理”“人工智能”“应用自然语言处理”和“神经网络”等课程中任选三门作为选修课。这些核心课程组能够帮助学生由浅入深逐步了解机器学习及其在数据科学中的应用。

5 对我国图书情报专业开展数据科学类课程建设的启示

数据科学是信息科学在大数据时代的拓展和延伸,图书情报学开展数据科学教育更是图书情报专业教育的拓展和提升。本文通过对 iSchools 核心成员院校的研究生项目开设数据科学相关课程现状的分析研究,可为我国图书情报专业设置数据科学类课程、培养具有扎实理论基础和专业技能的数据科学人才提供借鉴。

5.1 数据科学类课程设置要体现图书情报学科交叉特征

2019年,我国启动了新文科建设工程,新文科



起源于把新技术融入传统文科课程中,为学生提供综合性的跨学科学习。学科交叉融合是当前科学技术发展的重大特征,是社会经济发展的内在需求,也是创新型人才培养的有效路径。数据科学正是一门专业融合特质明显的交叉学科,它涉及统计学、数学、计算机、人工智能、机器学习、数据库、模式识别、可视化技术等多学科领域知识,重视培养学生的基础数理能力、数据科学思维能力和数据科学实践能力^[2]。

图书情报专业开展数据科学教育正是顺应当代学术研究交叉与融合的趋势,打破学科壁垒,在数据管理、数据科学专业教育、人才培养、指标评价等领域进行的积极探索^[37]。注重数据生命周期和以人为本是图书情报学科开展数据科学教育最重要的特征^[11],也是 iSchools 核心教育理念在数据时代体现“跨学科性”的重要方式^[38]。基于此,本文根据调查结果,以基于数据生命周期的课程体系为核心,面向图书情报专业设计数据科学类课程模型(见图3)。该模型面向国家大数据和数字经济重大发展战略,依据新文科理念和 iSchools 教育理念设定人才培养目标,基于数据生命周期规律优化和完善数据科学类课程体系,以满足行业和市场对数据科学人才的能力、素养等需求,同时坚持交叉融合,突出图书情报学科研究特色和所属院校办学特色,致力于培养德才兼备的数据科学高水平人才。

课程体系的完备,推动图书情报专业培养的学生在数据意识和计划、采集处理、分析揭示、保存安全和利用共享等各个数据生命周期阶段的知识学习和能力提升。另一方面,坚持交叉融合,既在数据组织处理、分析揭示等阶段加强图书情报专业优势,深化学生对数据进行组织、整合、分析、揭示的能力,同时坚持“以人为本”,从数据共享、数据表达、数据存储等能力维度入手,教导学生充分了解并理解数据科学的内涵,并利用技术工具破解数据治理困境,使学生能够利用数据能力优势增强竞争力。

5.2 面向国家战略和社会需求探索研究生分类课程教学体系

与本科生着重于知识传授和能力培养的课程设置不同,研究生课程教学注重“研究性”这一核心目标,是学生从接受式学习向研究性、自主性学习转变的关键期,对学生研究能力和思维能力形成具有不可或缺的作用。结合上述数据科学类课程模型,各图书情报人才培养单位需要结合数据科学研究生教育的特点,统筹构建适应学术学位和专业学位协调发展的课程教学体系。

学术学位研究生的培养以提高创新能力为目标,其课程教学体系要以学术创新为导向。在数据密集驱动的研究范式下,图书情报学科正在借用数据科学的理论和方法推动学科变革^[39]。2022年9月13日,国务院学位委员会、教育部印发《研究生教育学科专业目录(2022年)》,“图书情报与档案管理”一级学科更名为“信息资源管理”^[40]。此次更名将重新确立学科定位和发展目标,推动学科的改革与创新^[41]。因此,图书情报专业在学术学位研究生培养过程中,要面向国家大数据战略和高素质数据人才需求,大力开展数据科学教育创新研究,丰富数据科学教育的本土化研究。在教学过程中,要抓住数据时代的机遇,将开放、共享的数字生态理念贯穿于数据科学类课程教学过程中,实现教学与科研紧密结合,构建图书情报学科在数据科学研究中的交叉创新话语体系,体现学科在数智赋能环境下的积极应对。同时,通过“研”字当头的课程设计,把数据科学的科研成果融入研究性教学体系,提高课程的前沿性、高阶性和挑战度,不断提升学生的科研创新能力,以数据赋能人才发展,持续提升人才培养质量。

专业学位研究生的培养以提升职业能力为导向,其课程教学体系要以实践创新为导向。数字经

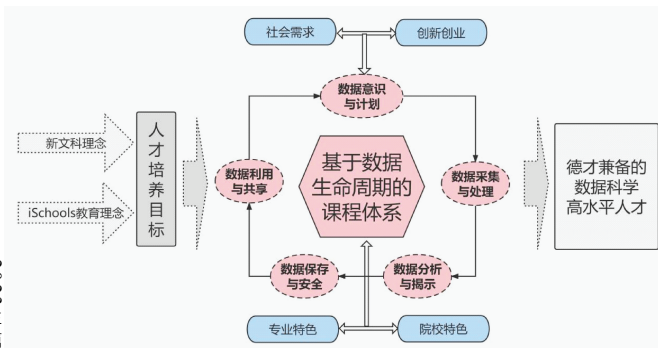


图3 图书情报专业数据科学类课程模型

在此课程模型的基础上,图书情报专业要积极应对数据密集型研究范式带来的机遇和挑战,利用数据科学的发展趋势拓展学科边界,从理论深化和实践拓展两个层面将图书情报人才培养工作全面推向数据时代。一方面,要从数据生命周期视角对当前数据科学类课程的内容体系进行逻辑梳理,推进



济环境下,云计算、物联网、人工智能都和数据科学密切相关,互联网企业、大型国企、政府机关、事业单位等用人单位对具有数据思维、掌握数据技能的高素质应用型人才需求非常旺盛。图书情报专业应面向国家和社会对高学历数据科学人才的需要,创新图书情报专业硕士培养方案。例如,自2021年起中国人民大学已开始招收数据管理方向的图书情报专业学位研究生^[42]。图书情报专业硕士培养单位要根据其生源特点,在课程设置时强调工具属性,把Python语言、分布式数据库、机器学习、自然语言处理、云计算和人工智能等数据工具类课程嵌入到图书情报专业课程体系中^[43],培养学生利用数据范式解决信息资源的深度加工、规范性分析、快速洞见、产品化研发和资产化管理等问题的能力^[44]。同时,模拟现实大数据环境,设定相关的数据管理实践任务,并从科研和企业一线引入真实数据集和数据项目,将项目式学习(Project-Based Learning, PBL)和案例分析法融入课堂,鼓励学生在真实情境和具体行业中解决复杂的数据管理与应用问题,以更好地培养学生的数据科学实践能力。

5.3 拓宽课程知识应用领域,丰富课程实践教学形式

情境驱动的数据科学研究极大地提升了 iSchools 学生在数据丰富的工作环境中开展数据收集、探索、操作、存储、分析和演示方面的能力。本次调研显示, iSchools 课程涉及医学、出版、传播等不同学科,利用数据识别非图书情报学科领域问题和解决问题的课程数量多、融合度深。“跨学科性”是 iSchools 院校的显著特征之一,当前图书情报学仍是 iSchools 院校最重要的学科,但数据科学、传播学等研究也是其知识产出与流动的主要领域^[45]。因此,一方面,要求图书情报专业在培养目标设置上就注重培养学生的数据意识,将信息牵引的学科研究内容转移到数据牵引的学科研究内容中;另一方面,强化行业需求,聘请行业领域内具有重要影响力和突出业绩的行业导师,充分发挥他们的指导作用,利用来自一线的行业数据或数据难题来指导学生的学习和研究。

我国图书情报专业对于硕士的实践实习有一定要求,如全国图书情报硕士专业学位研究生教育指导委员会拟定的《图书情报硕士专业学位设置方案》要求“突出图书情报实践导向……实践教学时间不

少于半年”^[46]。但2021年我国图书情报专业硕士的培养单位中只有2所具有10个以上的实践基地,有25所培养单位的实践基地数量少于或等于3个^[47],很难满足硕士的实践实习要求。而 iSchools 院校则设置包括课程实践、专题实践、学期实践、研究实践项目等多种形式的实践教学,特别是普遍设置了“顶点项目”(Capstone Project)。顶点项目研究能够支持学生的深度学习,将数据科学知识从课堂向实际工作过渡,实现数据科学相关知识整合、建构和迁移运用,是一种有效评估学生学习效果和评价工具。因此,我国图书情报专业在开展数据科学类课程教学时,也可以借鉴这一模式,构建“课堂—实验室—实习实训基地—社会”“四位一体”的研究生实践教学互动体系,培养学生知行合一,将数据科学知识转化为数据问题解决的能力。

6 结语

文章通过对 iSchools 核心院校研究生项目开设数据科学类课程情况的调研,从“数据生命周期—人才培养目标—课程体系”的对应关系总结了 iSchools 院校数据科学类课程的设置特征,并构建了基于我国国情的图书情报专业数据科学类课程模型。未来融合数据科学的图书情报学科发展无法避免挑战与困境,但也意味着机会与创新同在。因此,我国的图书情报专业在开设数据科学相关课程时也要适应趋势变化,从数据生命周期视角出发,强化培养学生的数据意识、采集、分析、揭示、存储、共享和重用等综合能力,为数据科学的发展贡献图书情报学科的力量,为国家的大数据战略培养更多优秀的数字人才。

本研究也存在一定的局限性,如在研究样本的选择上,数据科学当前并未形成统一的研究范式,因此 iSchools 院校的课程开设情况一直在变化,例如卡耐基梅隆大学的“94—892 与 Covid 交流”课程就是2021—2022 学年基于疫情数据所新增的课程。未来随着数据科学研究的深化, iSchools 院校可能还会出现更多适应新数据应用情境的课程。其次,本研究对课程内容的分析还仅停留在表层,未来将针对更具体的数据科学类课程(如数据组织、数据分析等)进行更深层次的内容分析,以为我国图书情报专业相似课程的设计和教学方式等提供更多借鉴。



参考文献

- Drexel University College of Computing & Informatics. Master of Science (MS) in data science program[EB/OL].[2022-04-04]. <https://drexel.edu/ci/academics/graduate-programs/data-science/ms-in-data-science/>.
- 杨杰,赵星.大数据专业培养内容的主题分析及对图情档学科的启示[J].图书情报工作,2022,66(2):109-116.
- 柯平.新图情档——新文科建设中的图书情报与档案管理一级学科发展[J].情报资料工作,2021,42(1):15-20.
- Zhang Y, Wu D, Hagen L, et al. Data science curriculum in the iField[J/OL].Journal of the Association for Information Science and Technology,2022;1-22[2022-08-29].<https://asistdl.onlinelibrary.wiley.com/doi/10.1002/asi.24701>.
- 武汉大学新闻网.吴丹教授 JASIST 发文探讨全球数据科学教育[EB/OL].[2022-10-04]. <https://news.whu.edu.cn/info/1015/67697.htm>.
- 陶俊,何晓东.面向图书情报的数据科学专业课程结构比较研究[J].图书馆学研究,2019(6):10-16.
- 王晰巍,李玥琪,刘宇桐,等.大数据及人工智能时代背景下国外图书情报专业研究生人才培养趋势研究[J].图书情报工作,2019,63(11):5-14.
- 吴丹,陈曦,徐小芸.国内外高校数据科学人才培养调查研究[J].文献与数据学报,2019,1(2):24-38.
- Song IY, Zhu Y. Big data and data science: opportunities and challenges of iSchools[J]. Journal of Data and Information Science,2017,2(3):1-18.
- 朝乐门,张晨,孙智中.数据科学进展:核心理论与典型实践[J].中国图书馆学报,2022,48(1):77-93.
- 吴丹,许浩.“以人为本”的数据科学教育:图书情报学科的新发展[J].图书情报知识,2021,38(6):12-20.
- 吴丹,孙雅琪,许浩.数据科学研究生教育的多学科比较研究[J].图书馆论坛,2021,41(11):108-117.
- Wu D, Xu H, Sun Y, et al. What should we teach? a human-centered data science graduate curriculum model design for iField schools[J/OL].Journal of the Association for Information Science and Technology,2022;1-18[2022-08-29].<https://asistdl.onlinelibrary.wiley.com/doi/full/10.1002/asi.24644>.
- 杨瑞仙,吴东昌.日本数据科学专业建设情况调查研究[J].情报理论与实践,2020,43(8):195-201,166.
- 闫慧,张钰浩,张鑫灿,等.iSchools 联盟数据科学教育项目现状调查[J].情报资料工作,2018(4):95-100.
- 苏日娜,杨沁.LIS 学科中数据科学课程体系设置研究——以 iSchools 高校课程调研为中心[J].图书馆论坛,2019,39(4):40-49.
- 曾粤亮.国外 iSchools 数据科学项目人才培养模式的特点与启示[J].图书情报知识,2018(4):109-118.
- 桂思思,任珂,王玉琦,等.基于 MOOC 平台的国内外数据科学课程建设对比分析及经验启示[J].图书情报知识,2018(3):119-128.
- 张胡,李树青,丁晓蔚,等.面向数据科学的情报学学科研究方法
- 及学科教育内容设计[J].图书与情报,2021(5):115-122.
- 程铄,刘桂锋,刘琼.当图书情报学科遇上数据科学:交叉与拓展[J].图书馆论坛,2022,42(11):94-100.
- 洪亮,刘宣治,侯雯君.基于 iSchool 教育理念的图书馆学专业课程模型[J].图书馆论坛,2019,39(8):47-56.
- 张宇,刘文云.基于数据生命周期的高校机构知识库科研数据服务研究[J].图书馆学研究,2021(3):71-80.
- DataONE. Data life cycle[EB/OL].[2022-04-04]. <https://old.dataone.org/data-life-cycle>.
- DCC curation lifecycle model [EB/OL]. [2022-04-04]. <https://www.dcc.ac.uk/sites/default/files/documents/publications/DCCLifecycle.pdf>.
- Carlson J, Fosmire E M, Miller C C, et al. Determining data information literacy needs: a study of students and research faculty [J]. Portal: Libraries and the Academy, 2011,11(2):629-657.
- University of Alberta Libraries. Data management [EB/OL]. [2022-10-04]. <https://www.library.ualberta.ca/research-support/data-management/>.
- Song IY, Zhu Y. Big data and data science: what should we teach? [J]. Expert Systems,2015,33(4):364-373.
- 夏义堃,管茜.基于生命周期的生命科学数据质量控制体系研究[J].图书与情报,2021(3):23-34.
- 聂云贝,刘桂锋,刘琼.数据生态链视角下科学数据生命周期运行过程分析[J].信息资源管理学报,2021,11(2):69-77.
- Microsoft Learn. Team data science process 生命周期[EB/OL]. [2022-10-04]. <https://learn.microsoft.com/zh-cn/azure/architecture/data-science-process/lifecycle>.
- 齐乾坤,王文龙.基于数据生命周期的高校研究生数据素养评价研究[J].情报科学,2021,39(9):125-130,145.
- Berkeley School of Information. What is data science? the data science career path [EB/OL]. [2022-10-04]. <https://ischoolonline.berkeley.edu/data-science/what-is-data-science-2/>.
- 司莉,姚瑞妃.图书情报专业研究生数据素养课程设置及特征分析——基于 iSchools 联盟院校的调查[J].图书与情报,2018(1):28-36,101.
- 罗娇,刘细文.知识产权视角下科学数据安全管理的策略选择[J].图书情报工作,2021,65(12):38-46.
- 秦小燕,初景利.科学数据素养内涵结构研究[J].图书情报工作,2019,63(18):30-39.
- 张璇,孟祥保.面向数字人文的高校数据素养教育案例研究[J].大学图书馆学报,2019,37(5):87-94.
- 刘敏.我国图书馆科学数据素养研究综述[J].图书馆学研究,2020(12):17-23.
- 肖雪,曹羽飞,曹建业,等.数据时代美国 iSchools 图书情报专业课程调查与思考[J].情报理论与实践,2022,45(2):10-17.
- 金波.紧抓新文科建设机遇 推动学科转型发展[J].图书与情报,2020(6):6-9.
- 教育部.国务院学位委员会 教育部关于印发《研究生教育学科专业目录(2022年)》《研究生教育学科专业目录管理办法》的通知[EB/OL].[2022-10-04]. <http://www.moe.gov.cn/srcsite/>



- A22/moe_833/202209/t20220914_660828.html.
- 41 初景利,黄水清.从“图书情报与档案管理”到“信息资源管理”——一级学科更名的解析与思考[J].图书情报工作,2022,66(14):3-9.
- 42 中国人民大学信息资源管理学院.信息资源管理学院 2022 年图书情报(数据管理方向)专业学位研究生招生简章[EB/OL].[2022-10-04].https://irm.ruc.edu.cn/displaynews.php?id=6867.
- 43 沈思,左明聪,王东波,等.基于课表知识抽取的情报学课程设置启示研究[J].情报学报,2020,39(12):1253-1263.
- 44 朝乐门.信息资源管理理论的继承与创新:大数据与数据科学视角[J].中国图书馆学报,2019,45(2):26-42.
- 45 杨思洛,张一鸣.iSchools 院校研究的跨学科特征:文献计量分析的视角[J].中国图书馆学报,2020,46(6):55-74.
- 46 全国图书情报硕士专业学位研究生教育指导委员会.图书情报硕士专业学位设置方案[EB/OL].[2022-10-04].http://mlis.whu.edu.cn/news.jsp?id=3.
- 47 段宇锋,顾思晨.图书情报硕士专业学位(MLIS)教育发展状况调查报告(2021)[J].图书情报知识,2022,39(3):95-102.
- 作者单位:郑州航空工业管理学院信息管理学院,河南郑州,450003
收稿日期:2022年4月4日
修回日期:2022年11月12日
- (责任编辑:关志英)

Investigation and Inflection on Data Science Curriculum of Graduate Degree Programs in iSchools at Abroad

Zhao Ruihan Wang Minhua

Abstract: The perfection of the data science education system and the cultivation and training of data talents are the keys to the development of library and information science in big data era. A systematic and in-depth study on the curriculum of data science in iSchools is conducted to promote data science education on postgraduate and enhance LIS core competitiveness. Based on the graduate degree programs in 31 iSchools at abroad, this paper uses the corresponding relationship between data lifecycle, talents cultivation goal and curriculum to investigate the setting and content of data science courses and summarizes the characteristics of the data science courses in iSchools. Therefore, the curriculum of data science in Chinese library and information schools should embody its special characteristic, carry out the division teaching strategy to meet the needs of national strategic and society, and improve the application of professional knowledge, and enrich the practical teaching methods.

Keywords: iSchools; Data Science; Curriculum; Data Lifecycle