

大数据背景下数据素养教育研究^{*}

黄如花¹, 林焱²

(1. 武汉大学信息资源研究中心, 武汉 430072; 2. 武汉大学信息管理学院, 武汉 430072)

摘要: 讨论信息素养教育拓展到数据素养教育的原因, 以及相关教学课程的设计, 从而为我国相关研究提供参考。采用网络文献调查法对相关主题的研究进行调研, 认为数据素养教育应采取分阶段循序渐进的方式教学, 分为“初级阶段”“进阶阶段”以及“高级阶段”, 并对每个阶段的教学内容进行设计, 以期提升学员的数据素养综合能力。

关键词: 大数据; 数据素养; 信息素养; 教育

中图分类号: G254

DOI: 10.3772/j.issn.1673-2286.2016.5.004

在大数据时代背景下, 信息素养 (Information Literacy) 的概念和内容正在泛化, 目前包括媒介与信息素养、视觉素养和数据素养等^[1]。纵观大数据时代的情状, 国际权威机构正积极推动数据素养的提升, 美国大学与研究图书馆协会 (Association of College & Research Libraries, ACRL) 发布的《面向高等教育的信息素质框架》新增了数据素养 (Data Literacy) 的提法^[2]; 联合国教科文组织 (United Nations Educational, Scientific, UNESCO) 发布的《全球媒体与信息素养评估框架》(Global Media and Information Literacy Assessment Framework) 涉及到数字素养教育; 美国博物馆与图书馆服务机构 (Institute of Museum and Library Services, IMLS) 为满足电子科研者需求, 设立“数据信息素养” (Data Information Literacy, DIL) 项目^[3]; 图书馆信息资源委员会 (the Council on Library and Information Resources, CLIR) 发布的《数据问题》(the Problem of Data) 对数据素养教育中数据管理方面内容展开具体讨论^[4]。

相较于全球学科领域在大数据环境下对数据素养人员的需求大、人才地位日益重要的趋势, 而极为不对等的是, 现有的信息素质教育往往忽视对数据的检索、再利用、数据挖掘、管理和评价等能力的培养, 因而显得

数据素质教育需求更为迫切。基于上述情况, 本文拟研究讨论为何信息素养教育要拓展到数据素养上, 并提出相应课程设计建议, 以此为我国数据素养教育提供参考。

1 数据素养概念内容

对于数据素养的定义, 学术界尚未有一致性的界定。根据调研, 国外对数据素养的定义主要有: Hogenboom等提出“数据素养是一种阅读、翻译、分析和批判性思考关于统计数据以及利用统计数据作为证据资料的能力”^[5]; Stephenson和Caracello提出“数据素养是一种查找、评价以及高效地、符合伦理道德地使用信息 (包括数据资源) 的能力”^[6]; Mandinach与Gummer主张“数据素养是理解和使用数据的能力, 能够有效地利用信息指导决策制定”^[7]。

国内对数据素养的观点: 张静波提出“数据素养是一种对数据的理解、创造和使用的能力, 是对媒介素养、信息素养等概念的延伸和补充”^[8]。孟祥保等认为“科学数据素养”包括三个层次: 一是具有“数据”意识, 能够认识到数据是一种资源; 二是具备数据基本知识与技能, 熟悉数据生命周期与科学研究生命周期, 在

^{*} 本研究得到武汉大学自主科研 (人文社会科学) 基金项目“促进我国数据开放共享的对策研究”、中央高校基本科研业务费专项资金以及武汉大学和湖北省教改项目“‘慕课’背景下信息素质教育课程的变革研究” (编号: 2015009) 资助。

特定专业领域内具有查找、管理、利用与共享数据的能力,了解相关数据管理政策法规和数据伦理等知识;三是在大数据时代,能够利用数据资源发现问题、分析问题与解决问题^[9]。程萌萌等认为就概念本身来说,数据素养与媒体素养十分接近,主要体现在帮助使用者安全和道德地进行社交与合作^[10]。在一些文献中提到数据素养与计算机、通讯技术科学以及统计技能联系紧密。

综上所述,对数据素养内容的观点,比较认同Stephenson和Caracello的提法。研究认为这种提法不仅将数据素养的能力内容较为全面地展现出来,且该提法体现出数据素养与信息素养的本质区别:其中明确提到数据素养的施用对象“包括数据资源”,而信息素养强调的是“信息”。“数据”可看作一种具体形式的“信息”,是可直接处理的数字对象;而“信息”的内涵更广且更抽象。此番将“信息”锁定到“数据”的范围,明确体现出“数据素养”是作为“信息素养”的一个子集,较之“信息素养”面向数据时代普通大众应具备的能力素养,“数据素养”教育更趋向于面向科研人

员的专业技能培养,是“信息素养”教育进一步的提高和深化。

2 数据素养教育课程研究进展

对于数据素养教育的内容,Carlson等认为信息数据素养教育的基本内容包括数据库及数据格式介绍、数据发现与获取、数据管理与组织、数据转换与互操作、数据质量保证、元数据、数据管理与再利用、数据保存、数据分析、数据可视化和数据伦理(包括数据引用)等^[11]。

2.1 数据素养课程开设的现状

基于“信息素养教育”(Information Literacy Instruction, ILI)的宽泛化趋势,数据素养教育作为信息素养教育新兴的重要热点研究领域,社会对其需求度不断增大,目前国内外图书馆及高校都开设有相关的数据素养课程(见表1)。

表1 数据素养教育课程的开设情况

分 布	开设机构	开设时间	课程名称	课程内容概要	教学方式	语 言
国外	耶鲁大学	2013-07-08 (5天)	数据素养——数据时代下中小学课堂文化教育(Data Literacy ©Literacies for the Digital Age to Teach in the K-12 Classroom)	数据思维、数据挖掘、数据驱动决策、材料学习、数据软件学习	讲座	英语
	内华达州立大学拉斯维加分校	2013-05-29	信息素养教育中数据资源的使用指南(Piloting the Use of Data Sources in Information Literacy Instruction)	数据检索及工具、数据评价、数据分析、数据交流和引用	大学图书馆	英语
	德克萨斯大学阿灵顿分校	2014-10-20 (9周)	数据,分析和学习(Data, Analytics and Learning)	利用专有和开源工具、分析数据周期、网络学习(翻译社交网络分析数据)、Gephi可视化分析结果、利用点击流数据评估分类器、数据评估、使用LightSIDE处理文本挖掘领域的问题、自动协同学习过程分析	edX	英语
	约翰·霍普金斯大学	2014-10-27	数据科学(系列课程)(DataScience) ^[12]	数据科学家的工具箱、R语言编程、获取和清理数据、探索性数据分析、可再生研究、统计推断、回归模型、实用机器学习、开发数据产品、数据科学顶点课程	Coursera	英语
	多伦多大学	2013-04-01 (47周)	统计:让数据有意义(Statistics: Making Sense of Data) ^[13]	数据概念、数据收集、或然率、置信区间、数据检验、一元线性回归、统计调查过程	Coursera	英语

续表

分 布	开设机构	开设时间	课程名称	课程内容概要	教学方式	语 言
国外	德里印度理工学院	2014-04-20 (9周)	网络智能和大数据 (Web Intelligence and Big Data) ^[14]	数据搜索、数据意义再现、数据库及其进化、大数据技术和趋势、多核处理编程、数据分类与聚类、数据挖掘、信息提取、数据推理、贝叶斯推理用于医疗诊断、数据预测、数据分析	Coursera	英语
	Google	2014-03-18 (17天)	数据的分析和解读 (Making Sense of Data)	数据分析软件、Google fusion table	Fun	英语
	奥克兰大学	2014-10-06 (8周)	洞察数据: 数据分析导论 (Data to Insight: An Introduction to Data Analysis) ^[15]	数据探索和发现、数据的局限性、数据可视化、iNZight软件学习	Canvas	英语
	加州大学伯克利分校信息学院	任意时间或按在校开课时间	信息和数据科学硕士 (Master of Information and Data Science, MIDS) 课程	数据和分析的研究设计及应用程序、探索和分析数据、保存和检索数据、应用机器学习、数据可视化和交流、顶点课程、数据背后的人性价值观、实验和因果推论、大数据、大规模机器学习、回归应用和时间序列分析	在线或在校学习	英语
	印第安纳大学情报与计算机学院	按在校开课时间	硕士研究生课程: 数据语义学 (Data Semantics) 元数据 (Metadata)	数据语义学、元数据	课堂学习	英语
	爱荷华大学图书情报科学学院	按在校开课时间	图书馆学中基于证据的实践 (Evidence Based Practice in LIS) 元数据理论与应用 (Metadata Theories and Applications) 设计、可视化和三维环境的映射 (Design, Visualization, and Mapping 3-D Environments)	数据分析、元数据理论、数据可视化	课堂学习	英语
国内	南京大学	2016-04-12 (8周)	用Python玩转数据	Python基础、数据获取与表示、强大的数据结构和Python扩展库、Python数据统计和可视化、面向对象和图形用户界面	中国大学MOOC	中文
	哈尔滨工业大学	2015-03-02 (18周)	大数据算法	大数据算法、亚线性算法、外存算法、外存查找结构、外存图数据算法、MapReduce的并行算法设计、非MapReduce的并行算法设计、众包算法	中国大学MOOC	中文
	哈尔滨工业大学	2016-04-01 (14周)	概率论与数理统计	处理随机数据	中国大学MOOC	中文
	哈尔滨工业大学	2016-03-01 (20周)	数据库系统 (上、中、下)	SQL语言、数据建模	中国大学MOOC	中文
	浙江大学	2016-02-29 (16周)	概率论与数理统计	处理“随机”事件、发现和处理各种数据资料、回归分析	中国大学MOOC	中文
	国防科技大学	2016-02-29 (16周)	概率论与数理统计	一元线性回归	中国大学MOOC	中文
	中国海洋大学	2016-03-07 (23周)	数据结构	线性表、栈和队列、串、数组和广义表、树和二叉树、图、内部排序	中国大学MOOC	中文

由表1可见,国外数据素养教育涉及面广,涉及的课程内容包括课程的面向对象、数据素养的多种能力培养(例如数据收集、挖掘和分析等)、应用数学知识处理数据等。有明确的“数据素养”主题,将信息素养缩小到数据素养的范畴,例如内华达州立大学拉斯维加分校的“信息素养教育中指导对数据资源的利用”课程。与国外开设的数据素养教育课程相比,国内开设的数据素养相关课程呈现的数据能力主要是数据分析,集中在利用数学、计算机科学处理数据,偏重数学理论、编程设计等能力培养,而鲜有强调对数据为主体进行的管理、开发等综合数据素养培养。尽管这些课程开设的目的是希望通过数学方法对数据进行分析与利用,但从课程知识内容及学时分配情况,可以看出在数据素养能力培养方面多是一带而过,涉略尚浅。

2.2 数据素养教育存在的问题

从总体来看,国内外关于数据素养教育课程的开设情况目前存在的普遍问题有三个,分别是开设机构数量不足、综合能力培养不足和教学监督力不足。

2.2.1 开设机构数量不足

国内外针对数据素养教育所开设的教学机构屈指可数,但国外相对多于国内,反映出我国当前对数据素养教育重视程度不足,该领域办学能力与教学资源缺乏。另外从数据素养课程开设机构类型来看,国外在数据素养教育方面相继开设相关课程,以提高用户的数据处理能力,其中以高校图书馆为主力军;而相较于国外的课程开设情况,国内开设数据素养教育课程主要集中在理工科类高校,且以“数据素养”为专题的课程针对性不强。

2.2.2 综合能力培养不足

已开设的数据素养教育课程对数据素养综合能力的培养不系统,教学时间较短且教学内容涉及范围有限,缺乏系统化教育。

从教学内容来看,数据素养教育存在能力培养不均衡、教学深度和力度不足的问题。国外开设的数据素养教育课程较为系统,但投入量最大的方向是数据处理与分析能力培养,各大教育单位都不遗余力地推

行对数据统计分析工具、数据计算方法模型以及相关的编程方法的教学,但是仍待商榷的是其中数据处理与分析教学中讲授的统计分析工具专业性、规范性,例如德克萨斯大学阿灵顿分校讲授的LightSIDE工具。另外,数据素养作为一项综合数据管理能力,若对数据统计分析工具讲授比重太大,则在学时有限的情况下势必会削减其他能力的培养,而数据素养中对数据的敏感度、准确获取、全面评价和有效保存等特别是数据解释与表现都是必不可少的能力。国内开设的数据素养教育课程甚是稀缺,目前已开设的课程也仅仅是针对数据素养某一方面的特定技能进行培养,且教学深度和广度极为不足;更令人遗憾的是目前尚无明确的“信息素养”专题课程开设,更没有“数据素养”的专题课程。

2.2.3 教学监督力不足

一方面,教学方式多为在线学习,虽然这种方式可以方便学员自主安排时间,但是没有学位授予或过关考试,片段化的学习模式对学生的自觉性要求极高,这容易造成教育落实到学生身上程度低的隐患;另一方面,从教学时长来说,国外针对数据素养教育课程设计的时间除了几所设有信息管理学院或相关学位授予点的高校外,其他办学机构课程多为短期项目,这就存在学生走马观花式学习、不能在实际中灵活应用的问题。

3 数据素养教学内容的设计

加强数据素养教学内容将使信息用户更好地适应大数据环境,并拥有对数据综合应用的能力,从而在学习生活、工作竞争中利用数据取得便利和有利条件。

Carlson等^[16]认为数据素养教育应主要面向本科生、研究生和教员开设;而英国在2012年“数字内容管理课程II项目”(Digital Curation Curriculum Project, DigCCurr II)中提出数据素养教育应该扩大到博士生和从业者^[17];而一些学者认为数据素养还应该向行业从业人员开放。综上所述,研究认为数据素养的专业技能属性更为显著,更趋向于科研工作人员在从事科学研究时对科学数据处理的能力,因此课程拟面向科研人员的数据素养能力培养而设计。

根据美国信息科学技术学会(American Society for Information Science and Technology, ASIS&T)的

“数据访问和保存研究”(Research Data Access and Preservation, RDAP)峰会委员会对硕士研究生定义的12个关键数据能力,以及加州大学圣地亚哥分校对数据生命周期的描述(见图1),并基于数据素质能力重要性评估调查情况,研究认为数据素养教学课程中数据素养的能力培养应该包括数据处理与分析、数据引用与重用、数据管理与共享、数据可视化与表达、数据伦理与数据贡献说明、数据评价与数据证据、数据整合与组织、数据转换与互操作、元数据与描述、数据发现与检索、数据保存和实践操作。

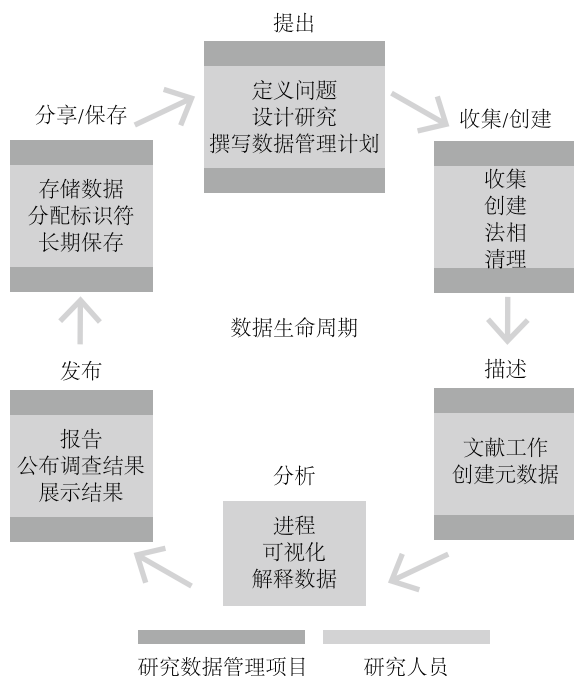


图1 UCSD数据生命周期^[18]

课程目的在于综合培养高素质数据人才,分为三个学习梯度,分别是初级阶段课程、进阶阶段课程以及高级阶段课程,各阶段针对不同的数据素养进行专项教学,例如初级阶段主要培养“数据伦理”与“数据检索”能力。

将课程分为阶段学习,是为使学员能够由浅入深地掌握数据素养的各项技能,根据初级、进阶到高级的划分,学员可以根据自身情况选择所需到达的技能水平。初级阶段属于基础性必修课程,主要用于树立数据道德伦理意识,同时培养数据获取的能力;进阶阶段培养较为精深的数据分析、评价及组织能力;高级阶段则重在培养学员实际操作数据的能力,包括利用数据进行交流、再利用的能力,以及强调数据保存能力。

3.1 初级阶段课程

3.1.1 数据伦理

数据伦理(Data Ethics)课程主要培养在数据利用中所应遵循的数据伦理道德规范。数据伦理即指基于数据资源的决定制定过程中的逻辑和道德伦理,以及通过设计注重交流数据。

根据RDAP委员会的“数据素养项目”对教师与学生列出的12项数据素养重要性的评价(见图2),“伦理与贡献”项师生评价的分值都居高,因而它作为整个数据素养的基础是合理且必须的。

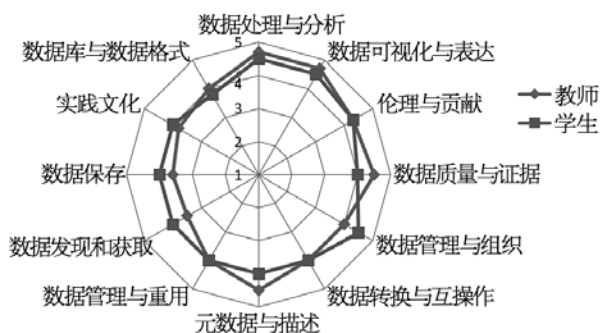


图2 师生对数据素养各部分能力重要性评价^[19]

(注: 分值意义中5为非常重要, 4为重要, 3为较为重要, 2为一般, 1为不重要)

在这部分课程中,主要对数据采集、数据决定制定和元数据描述等过程中涉及的伦理道德以及数据贡献说明进行教学。由于参与课程的学生具有不同学科背景,而数据素养源自计算机科学和信息学,因此课程还将简要介绍计算机科学在社会学、经济学、商业、法律、图书馆研究、认知科学和心理学等方面的运用,为学生进行更高层次的课程学习作铺垫。课程涵盖数据道德内涵、网络和网络安全、法律和政策、商业模式和创业精神。数据伦理课程内容设计见表2。

3.1.2 数据采集

数据素养的一项基本技能便是拥有对数据的敏感度和采集能力,因此将“数据采集”纳入初级阶段课程中。课程内容应包括了解关系数据库的概念,如何查询这些数据库,并熟悉其标准数据格式和类型,以及确定不同的研究问题所适合的格式和数据类型。课程主要通

表2 数据伦理课程内容设计

课程单元	教学内容的	学 时
第一单元	对数据素养内容进行整体介绍,对数据伦理内涵以及在商业、科研工作中的重要性进行教学。了解知识产权的发展,隐私和保密问题,在共享和管理数据过程中的道德准则。恰当地确认从外部来源获得的资料,避免提供数据时误导或表达模糊 ^[20]	1
第二单元	介绍网络基础技能与网络安全相关法律与政策;介绍数据在商业模式和创业中的重要性,同时强调数据伦理在各项工作中的原则。分析涉及数据伦理的案例,指明其中所应遵循或已违反的伦理道德,并介绍目前违反数据道德规范所对应的惩处规则;介绍数据贡献说明的规范方式	1
第三单元	提出一项数据伦理案例进行课堂分析,并让学生讨论制定维护数据伦理可行性方针。学生通过小组工作形式对案例进行研究、评估和分析,并演示所制定的解决方法,从而增强教员与学生、学生与学生之间的课程交流,加深对数据伦理道德的认识和理解	1

表3 数据采集课程内容设计

课程单元	教学内容的	学 时
第一单元	通过社会书签引入数据发现的重要性介绍,同时介绍提升数据敏感度的方法;此外介绍目前主要的检索工具、检索平台,包括常用网络搜索引擎(百度、谷歌、必应和雅虎等)及其各自的分类检索功能;介绍常用期刊论文数据库(CNKI、SSCI、Elsevier、SDOL和WorldCat等)	1
第二单元	讲解检索方法与技巧。主要介绍网络信息检索常用方法,包括布尔逻辑检索及其逻辑运算符、邻近检索、截词检索等;讲解检索式的构造;不同数据库的特殊检索命令;检索技巧包括调整检索策略控制检索范围,改变检索式构造或规范检索词等	1
第三单元	生活中数据检索实例。学生通过课堂研讨会的形式,首先了解搜索引擎工作的基本知识,然后探讨搜索引擎设计如何影响商业和文化。主题包括搜索广告和拍卖,搜索和隐私,搜索排名,反垃圾邮件,本地搜索,点对点的搜索,搜索博客和在线社区等丰富的主题内容	1
第四单元	实习课。设置检索题目安排学员上机完成检索任务,实现对数据发现和检索能力的初步实践	1

过介绍检索工具、检索平台(网络搜索引擎及数据库等)和检索方法,旨在培养对数据发现、传送及数据格式处理的能力,并对上述能力进行初步实践。数据采集课程内容设计见表3。

3.2 进阶阶段课程

3.2.1 数据分析与评价

数据素养较为高层次的能力便是对数据的分析与评价。由于数据分析与评价需要数学软件统计分析,因而课程将倾注介绍数据分析统计软件。数据分析技能应包括熟悉基本的数据处理和分析工具以及技术,了解这些工具对数据可能产生的作用。使用适当的工作流管理工具来自动化重复分析数据。通过介绍分析工具和数据分析方法,旨在培养对数据的批判性认识,从而能够对数据进一步处理与分析,掌握元数据与描述知识,评价数据质量。数据分析与评价课程内容设计见表4。

表4 数据分析与评价课程内容设计

课程单元	教学内容的	学 时
第一单元	数据分析及统计工具介绍(上)。介绍数据分析工具InspireData	1
第二单元	数据分析及统计工具介绍(下)。选择性介绍数据统计工具(ArcGIS、Stata、SAS、SPSS、Stata、R及Python等)	1
第三单元	数据质量评价。主要讲授如何定位和利用学科数据存储库,使得学生认识到数据质量可以从外部来源评估,且在标识适当的外部数据源时也可引入数据,并在必要的时候将其转换供本地使用。塑造学生识别、证明以及解决任意数据集明显的样式、半成品或讹误问题的能力;并利用元数据来帮助理解潜在的数据集问题。培养学生利用足量的文件数据使研究成果实现再生产,利用他人研究数据的能力;跟踪数据出处和清晰描绘数据集描绘和表示版本的能力	1
第四单元	实习课。实际操作数据分析软件以及完成数据质量评价任务	1

3.2.2 数据组织

课程培养目标是通过理解数据生命周期,开发数据管理计划,跟踪子集关系或原始数据集处理过的数据;

创建数据管理程序和文档的标准化操作模式; 综合培养对数据以及数据集的整合、组织与重组的能力。数据组织课程内容设计见表5。

表5 数据组织课程内容设计

课程单元	教学内容的设置	学 时
第一单元	介绍数据生命周期及数据管理计划工具DMPTools的使用和功能	1
第二单元	利用DMPTools制定数据管理计划(上)。设定研究项目, 演示数据管理计划工具DMPTools开发数据管理计划的过程	1
第三单元	利用DMPTools制定数据管理计划(下)	1

3.3 高级阶段课程

3.3.1 数据交流

作为数据素养高级阶段的课程, 数据交流不仅是一门课程, 也是一项偏向实际操作的能力, 它对数据素养中数据引用与重用、数据可视化与表达、数据转换与互操作和数据管理与共享等能力都有一定要求, 是数据素养综合能力的表现。因而这门课程将分别就上述能力进行分单元教学。数据交流课程内容设计见表6。

表6 数据交流课程内容设计

课程单元	教学内容的设置	学 时
第一单元	数据引用与重用。介绍元数据创建过程和方法(参见ICPSR数据引用指南) ^[21] ; 介绍元数据创建工具DataCite ^[22] 以及EZID ^[23]	1
第二单元	数据可视化与表达。通过实现数据的可视化展示, 解释数据意义, 实现使数据成为证据的转变。熟练使用基本的学科可视化工具, 避免在表格、图表等陈述数据时误导或含糊表达。选择适当可视化类型, 如地图、图表、动画或视频, 根据对原因或目的的理解用可视化的方式显示数据。介绍可视化软件board、Stratum和Tableau ^[24]	1
第三单元	数据转换与互操作。培养学生能够精通数据在格式间的迁移, 并理解由于数据格式变化所造成的风险、潜在的损失或讹误信息; 了解在标准数据格式促进后续使用的好处	1
第四单元	数据管理与共享。介绍数据管理计划共享资源PUBLIC DMPS ^[25] 以及资源数据共享平台SHAROE ^[26] 和CASFI ^[27] 。要求学员能够通过课程学习养成数据管理和共享的意识及能力	1

3.3.2 数据保存

数据保存对于数据管理、数据检索、数据重用等意义重大, 然而数据保存的成本、技术、保存完整度和保存周期一直是学界探讨的问题。因而数据素养教育课程介绍数据保存的技术与工具, 培养数据保存的能力是极为重要的。数据保存课程内容设计见表7。

表7 数据保存课程内容设计

课程单元	教学内容的设置	学 时
第一单元	数据保护的收益和成本; 理解数据保存的技术、资源和组织组件, 在数据准备中利用最佳实践在其活跃的生命周期期间最终保存。清楚地阐释数据对自己或他人长期潜在价值, 从而能够确定适当的保存时间。理解发展保存政策的需求, 能够识别这些政策的核心元素	1
第二单元	介绍数据保存工具。包括: 加州大学数字图书馆的EZID与Merritt; 加州大学洛杉矶分校(UCLA)图书馆网站的数据保存和共享(Data Storage and Sharing)功能; “数据保存系统”(Data Preservation System) ^[28]	1

4 结语

大数据时代数据的价值受到越来越多的重视, 对挖掘数据价值、提高处理数据和分析数据能力等的需求也随之增加, 这就要求更多的数据素养人才投入。然而目前数据素养教育在学科领域内才刚起步, 课程设计、资源配置等都还不完善, 研究基于此情况展开对数据素养教育课程的设计, 希望能够拓展大数据信息素养教育体系外延, 参与讨论数据素养培养的具体操作, 将数据素养教育课程设置为梯度模式, 分阶段渐进式地对数据人员进行授课, 体现数据素养所包含的各种能力间的关系和重要性关系, 为相关教育研究者提供借鉴。然而研究中难免有纰漏或偏狭之处, 在此欢迎业内学者不吝赐教, 共同为我国大数据事业谋划发展。

参考文献

- [1] 黄如花, 李白杨. MOOC背景下信息素养教育的变革[J]. 图书情报知识, 2015(4):14-25.
- [2] OAKLEAF M. A roadmap for assessing student learning using the new framework for information literacy for higher education [J]. The Journal

- of Academic Librarianship, 2014, 40(5): 510-514.
- [3] Data Information Literacy [EB/OL]. [2016-04-26]. <http://www.datainfolit.org/>.
- [4] JAHNKE L, ASHER A, SPENCER D C, et al. The problem of data[EB/OL].[2016-04-23].<http://www.clir.org/pubs/reports/pub154/pub154.pdf>.
- [5] HOGENBOOM K, PHILLIPS H C M, HENSLEY M. Show me the data! Partnering with instructors to teach data literacy[EB/OL].[2016-05-07]. http://www.ala.org/acrl/sites/ala.org.acrl/files/content/conferences/confsandpreconfs/national/2011/papers/show_me_the_data.pdf.
- [6] STEPHENSON E, CARACELLO P S. Incorporating data literacy into undergraduate information literacy programs in the social sciences: a pilot project [J]. Reference Service Review, 2007, 35(4): 525-540.
- [7] MANDINACH E B, GUMMER E S. A systematic view of implementing data literacy in educator preparation [J]. Educational Researcher, 2013, 42(1):30-37.
- [8] 张静波.大数据时代的数据素养教育[J].科学,2013, 65(4):29-32.
- [9] 孟祥保,李爱国. 国外高校图书馆科学数据素养教育研究[J]. 大学图书馆学报,2014,32(3):11-16.
- [10] 程萌萌,夏文菁,王嘉舟,等.《全球媒体和信息素养评估框架》(UNESCO)解读及其启示[J]. 远程教育杂志,2015(1):21-29.
- [11] CARLSON J R, FOSMIRE M, MILLER 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.
- [12] 数据科学 [EB/OL].[2016-01-07].<https://www.coursera.org/specializations/jhudatascience>.
- [13] 统计: 让数据有意义-多伦多大学[EB/OL].[2015-12-31].<https://www.coursera.org/course/introstats>.
- [14] COURSERA: 一流大学免费在线课程平台[EB/OL].[2015-12-31]. <https://www.coursera.org/course/bigdata>.
- [15] 黄如花,钟雨祺,熊婉盈. 国内外信息素养类MOOC的调查与分析[J]. 图书与情报,2014(6):1-7.
- [16] CARLSON J, FOSMIRE M, MILLER C, et al. Data literacy essay » ACRL scholarly communication toolkit [EB/OL].[2016-04-23].http://acrl.ala.org/scholcomm/?page_id=330.
- [17] DigCCurr[EB/OL].[2016-04-23].<http://ils.unc.edu/digccurr/>.
- [18] Data Management - The Library [EB/OL]. [2016-05-05].<http://libraries.ucsd.edu/services/data-curation/data-management/index.html>.
- [19] CARLSON J, JOHSTON L R. Data information literacy: librarians, data, and the education of a new generation of researchers [M]. Indiana: Purdue University Press, 2015:53.
- [20] CARLSON J, NELSON M S, BRACKE M S, et al. Data information literacy toolkit[M]. Indiana: Purdue University Libraries , 2015:1-83.
- [21] Data Citations[EB/OL].[2016-05-05].<http://www.icpsr.umich.edu/icpsrweb/ICPSR/curation/citations.jsp>.
- [22] Cite your data.[EB/OL].[2016-01-02]. <https://www.datacite.org/services/cite-your-data.html>.
- [23] EZID: EZID Home[EB/OL].[2016-01-03].<http://ezid.cdlib.org/>.
- [24] Compare Data Visualization Software [EB/OL]. [2016-01-03].<http://www.softwareadvice.com/bi/data-visualization-comparison/>.
- [25] Public DMPs: DMPTool[EB/OL].[2016-01-03].https://dmptool.org/public_dmps.
- [26] SINGH A, LIU L. Sharoos: a data sharing platform for outsourced enterprise storage environments[EB/OL].[2016-01-03]. <http://www.doc88.com/p-399308729189.html>.
- [27] CASFI Data Sharing Platform[EB/OL].[2016-05-07].http://www.academia.edu/377048/CASFI_Data_Sharing_Platform.
- [28] Data preservation systems: automated online offsite data backup services for small to mid-sized business[EB/OL].[2016-01-05].<http://www.datapreservation.com/>.

作者简介

黄如花,女,1968年生,博士,武汉大学信息管理学院副院长、教授,博士生导师,研究方向:信息检索, E-mail: rhhuang@whu.edu.cn。
林焱,女,1993年生,硕士研究生,武汉大学信息管理学院硕士研究生,研究方向:信息检索, E-mail: yolandelin@163.com。

Data Literacy Education Research under the Background of Big Data

HUANG RuHua¹, LIN Yan²

(1. the Center for the Studies of Information Resources, Wuhan University, Wuhan 430072, China;
2. the School of Information Management, Wuhan University, Wuhan 430072, China)

Abstract: We discuss the cause for which the education of information literacy should expand to data literacy, as well as the design of the curriculum, thus to provide references for related research in China. We use Internet survey method and literature survey method to conduct the investigation, and deem that the data literacy education should be taken in stages, teaching progressively, to be divided into "primary stage" "capstone stage" and "advanced stage". In order to enhance students' comprehensive data literacy, the design for teaching contents of each stage is included in our discussion.

Keywords: Big Data; Data Literacy; Information Literacy; Education

(收稿日期: 2016-04-08)