

数据密集型科研范式下馆员数据素养能力 培养路径分析*

董薇

(东北大学图书馆, 沈阳 110819)

摘要: 在数据密集型科研范式下, 馆员的数据素养正成为其主要核心能力, 本文从数据素养的概念及内涵入手, 在对国内外馆员数据素养现状进行分析的基础上, 提出针对馆员数据素养培养的框架体系, 并提出数据素养能力等级的定性划分, 即数据馆员、数据专家、数据科学家, 最后提出在政策引导下, 以提升方法技巧为目标, 通过嵌入科研项目培养数据能力的路径, 完成对数据密集型科研范式下馆员数据素养能力培养路径的分析与构建。

关键词: 数据素养; 数据密集型科研; 嵌入式科研; 馆员

中图分类号: G250

DOI: 10.3772/j.issn.1673-2286.2017.11.011

Negroponte在《数字化生存》中对信息技术为人类发展提供强有力的推动及其对人类社会强大塑造力倍加感慨, 认为信息DNA已经嵌入每个人的生活^[1]。在大数据时代, 获取数据、分析数据是基础能力。数据具有帮助人们发现问题, 并辅以完善或改进方案的作用; 有助于确认或怀疑某一实践是否正确, 还可发现问题根源并促其改进, 从而实现对资源分配的指导作用^[2]。随着数据密集型科研范式的兴起, e-Science与e-Social Science发展的深入, 数据管理成为图书馆业务创新的新领域。《不断生长的知识: 英国图书馆2015—2023战略》指出目前正处于数据创造、分析和开发的大数据时代^[3]; 美国新媒体联盟认为数字数据管理的发展, 将促进更准确的主题搜索发展, 从而帮助用户更有效地管理和展示相关信息资源^[4]。

数据密集型科研范式是以数据为基础, 利用大数据对科学研究进行模拟仿真并进行分析总结的研究范式。由于该研究范式以数据为基础, 故对数据素养能力提出更高要求。虽然数据素养并非数据密集型科研

范式下唯一的能力要求, 但作为最核心的能力, 对数据素养能力的研究和讨论理应成为图书馆学界研究的课题。对海量数据的处理能力成为科研活动最基础的能力, 馆员为服务科研活动而具备的数据素养能力成为亟须提高的能力。但现实情况却不尽如人意, 一方面, 科研活动尚无法从馆员处得到相应的数据支撑; 另一方面, 馆员也不能通过嵌入科研项目, 实现对科研活动的有效支持。面对需求与现实的差异, 本文尝试从数据素养能力培养路径的角度进行分析与构建, 以期科研活动的开展提供支撑。

1 数据素养能力现状分析

1.1 基于数据素养培训课程分析

王晰巍等选取美国与中国各4所高校(共8所)的素养教育培养模式作为样本, 分析两国数据素养能力的培养模式^[5](见表1)。

* 本研究得到辽宁省教育科学“十三五”规划立项课题“基于全程模式评价的慕课质量评价指标体系构建研究与实践”(编号: JG16DB123)、辽宁省高等学校图书馆情报工作委员会年度规划立项“基于VR/AR的泛在化图书馆元素养教育研究”(编号: LTB201708)和东北大学图书馆年度立项课题“泛在化智慧图书馆视域下教育研究”资助。

表 1 数据素养培养模式通览

国 家	学校名称	数据素养培养内容	培养模式
美国	加州大学洛杉矶分校	数据意识	通识教育
	西佛罗里达大学	数据意识和数据分析能力	基于软件服务平台
	普渡大学	数据分析能力	嵌入学科专业教育
	华盛顿大学	数据获取能力和数据分析能力	嵌入学科专业教育
中国	北京大学	数据意识	通识教育
	上海交通大学	数据意识、数据获取和分析能力、数据伦理道德	基于软件服务平台+嵌入学科专业
	浙江大学	数据挖掘能力	基于软件服务平台+嵌入学科专业
	吉林大学	数据意识	通识教育

高校或图书馆提供与数据素养培养相关的课程，不仅与其社会教育职能的特色相关，也与其自身所具备的培养条件有很大的关联性。量化研究在总体科研中的占比相对较小，对数据密集型科研范式的需求力度相对较差，故只提供意识层面的培养即可满足基本需求；软件服务平台或嵌入学科专业的教育模式进行数据素养培养的主体多以理工科见长，科研人员对科研数据的需求量较大，并注重从数据挖掘、数据分析、数据使用技巧等方面去培养数据素养能力，基于对上述两种情况的分析，发现此类图书馆馆员或从事数据素养培养的教师所具有的数据素养能力不尽相同。综合类图书馆馆员更侧重数据意识培养，理工类高校图书馆馆员更侧重数据技巧培养，目前并无图书馆在数据素养能力培养方面做到全面完善。需求差异化导致作为图书馆馆员核心能力的数据素养能力，没有得到全面推进和发展。

1.2 基于培训项目分析

武汉大学图书馆邀请美国雪城大学信息研究学院教授秦健作“科学数据管理在高校图书馆的应用”讲座，以提高图书馆服务人员的数据服务意识与素养^[6]；西南交通大学（九里堤校区）举办“图书馆数据科学家技术培训班”，除介绍大数据理念在实践中的应用外，还介绍信息搜集方法、利用Lucidworks Fusion对数据进行整合搜索和挖掘、数据可视化软件及应用、Xpath数据抓取软件的使用等有关数据收集、分析、处理的技巧方法^[7]；康奈尔大学图书馆开设了很多关于数据素养的讲座，并参与高校科学数据管理项目，以此发挥图书馆馆员在数据素养教育和数据管理服务中的作用^[8]。从中美图书馆数据素养培训模式可见，美国的数据素养培

训以输出为主；我国以输入为主，多邀请专家为馆员提供数据素养方面的培训，或鼓励馆员参加数据素养方面的培训。这种输入性行为可以从一定程度上看出国内在馆员数据素养培训方面的缺失，至少目前还不具备完善的数据素养培训体系，馆员的数据素养提升多以个体学习为主，群体化培训规模尚未形成，不具备培养“数据科学家”的条件。

1.3 基于嵌入科研项目过程分析

张军指出图书馆数据服务中心等机构或数据馆员等岗位，应以协助科研人员完成数据管理计划的起草、制定及修订等工作为宗旨，但由于身份和角度的局限，使其对科研项目缺乏深层次了解与感知，无法对所提供的数据服务做到准确无误^[9]。面对上述质疑，图书馆员不得不正视这样一个问题，即现有数据服务能力并没有很好地做到为科研活动服务，馆员所掌握的理念、方法、技巧在嵌入科研项目时，并没有发挥其最佳水平。从科研人员角度来看，馆员所提供的服务并不是科研人员迫切关注的，馆员对学科知识的盲区会在一定程度上限制其数据能力的发挥。当然，作为图书馆的数据馆员离“数据科学家”的要求仍相去甚远，北京大学图书馆招募了具有专利代理人职业资格的学科博士作为其嵌入式服务的中坚力量。这样的配置在图书馆中并不普遍，年龄偏大、学历偏低、专业知识老化、脱离科研项目时间过久、数据收集及转化利用水平低下、缺乏数据版权意识等诸多因素都将导致图书馆馆员在为科研活动提供嵌入式数据服务过程中的专业数据能力受到用户质疑，使高校图书馆为科研活动服务仍停留在简单、基础的文献保障服务层面，无法向更高层次的，为科研、决策提供支持的服务转型。

2 数据素养概述

维基百科对数据素养的定义为以信息的形式阅读、创造和交流数据的能力并将此过程正式描述出来。对数据的读写和教学方法所固有的技巧,已使得数据采集变得常规化。数据分析和大数据处理已经成为新闻、商业、政府和社会的普遍认知^[10]。百度百科将数据素养定义为对媒介素养、信息素养等概念的一种延续和扩展,至少包括五个方面的维度:对数据的敏感性,数据收集能力,数据的分析、处理能力,利用数据进行决策的能力,对数据的批判性思维^[11]。Stephenson等认为数据素养是为有效而恰当地发现、评估与使用信息和数据所具备的一种意识和能力^[12]。数据素养作为循证思维的一部分,其利用数据进行推理和论证^[13],而数据价值、数据技术、数据安全同样是数据素养着重考虑的问题^[14]。此外,对数据素养的培养还应注重其在共享和协同方面的应用,以促进创新的发展^[15]。因此,数据素养核心内涵可概括为意识、技能与利用数据解决问题的能力^[16]。

在数据素养概念流变的过程中,信息素养、媒介素养、计算机素养、网络素养、数字素养共同组成素养连续统^[17]。基于此,本文认为数据素养主要由数据基本意识、数据获取能力、数据管理能力、数据分析能力、数据存储安全能力和数据伦理意识六个方面构成,作为数据密集型科研范式下的图书馆员核心能力培养,也应从这六方面着手进行培养路径的分析与构建。

3 数据素养核心能力的内涵

3.1 数据意识

数据意识是数据认知层面的概念,从事不同学科领域研究的人有不同的认知,对从事数据素养教育的人员而言,就是指对数据的敏感度、信任度问题。具体可细化为:具有大数据意识、将数据视作科研活动中的重要因素、了解科研数据是有生命周期的、在科研过程中对待数据严谨认真、能够保证数据的敏感度和可信度,以及能将数据的需求细致化、具体化表达出来。

3.2 数据获取能力

数据获取能力主要指从海量数据中获取与科研相

关的数据,并能合理使用方法、技巧和工具获取与研究相关的科研数据。具体指能全面准确地从数据源处收集各类数据;熟练使用数据收集方法及工具,对于检索策略的制定、检索式的编辑能够兼顾查全率和查准率;能够准确解读各类显性、隐性数据。

3.3 数据分析处理能力

数据分析处理能力是数据素养最核心、最本质的能力,选择有效的数据处理和分析方法得出研究结论,是该能力要求的关键所在。对多学科领域数据分析工具的掌握能简化数据利用过程;利用可视化工具对数据进行多角度分析和表达,甚至进行数据建模均是对数据素养能力较高层次的要求;对数据的多种方式合理解读、重现和表达,并对数据分析结果进行认定、质疑和改进也是数据素养核心能力培养的目标之一。

3.4 数据转换能力

数据转换能力是数据利用能力的另一种表达形式,因为绝大部分数据一般无法直接使用,需要进行相应地转换,所以数据馆员需熟练运用多种学科领域的标准。此外,还需熟练使用数据转换平台及工具,能够利用数据分析结果支撑科学研究顺利进行;同时对数据实行重复利用,实现其使用的最大化,最终为决策提供数据支持。

3.5 数据评价能力

数据的评价能力指对所获数据能正确对待并客观评价的一种能力。数据评价关乎其在科研活动中的使用率、复用率,具体包括:能够验收数据正确性,并剔除错误、无效数据;认识到数据的局限性,并能够合理质疑数据。

3.6 数据伦理

数据伦理是数据素养必不可少的部分,指在数据的获取和利用过程中的法制观念及评判能力。在现实中,多样化的数据来源、类型和参差不齐的数据质量是数据伦理及个人隐私不容忽视问题^[18]。而规范数据著录的格式是数据伦理的另一个体现方面,为此,国内外

积极推进对元数据的研究和推广^[19]。

4 数据素养培养路径构建

4.1 嵌入科研生命周期的数据素养培养框架

宫学庆等认为数据密集型科研环境下科研人员面临的挑战主要来自科研生命周期管理和科学数据生命周期管理过程中的挑战,那么将数据素养的培养嵌入科研生命周期就可以同时解决两个挑战^[20](见图1)。

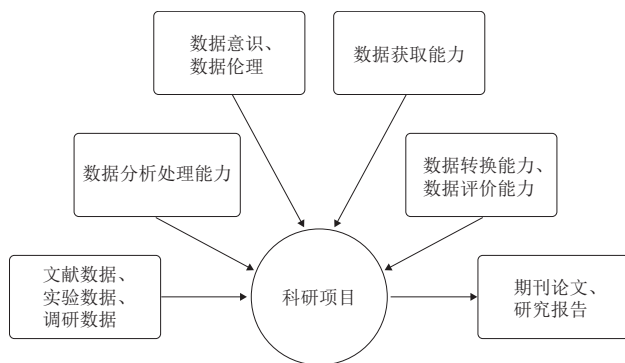


图1 嵌入科研生命周期的数据素养培养框架

数据素养能力培养以科研项目为生命周期,以文献数据、实验数据、调研数据作为科研输入端,以期刊论文、研究报告作为输出端,数据意识、数据伦理、数据获取能力、数据分析处理能力、数据转换能力、数据评价能力要求贯穿整个科研项目始终。

4.1.1 数据意识和数据伦理是无形资产

数据意识和数据伦理在数据素养培养过程中是抽象的,很难用量化的具体行为参数进行描述,故称为无形资产。在整个科研生命周期中,从科研项目的设计、实施到成果的产出、结题,数据意识和数据伦理贯穿始终,任何一个环节缺少了二者都将造成严重的后果,轻则使项目无法按照预期完成,重则将受到法律惩罚。

4.1.2 数据获取能力是前提

科研项目以阅读文献作为开始,已有研究成果中积累的大量经验和数据都是开展科研活动的前提。此外,掌握数据获取途径,熟练使用从数据库、年鉴、各类实验中获取数据源以及批量数据处理软件获取的海量数

据,均能为保证科研顺利进行奠定基础。

4.1.3 数据分析处理能力是核心要义

科研过程中产生的大量数据,既有有效数据,也有冗余数据,科研人员需要对其进行甄别、清洗后再加以使用。这就需要培养很强的数据分析处理能力,如对SPSS、Citespace、MATLAB等数据分析处理软件的熟练使用,从而发现隐藏在数据中的实质性问题,提出解决方案,推进科研项目,并增加研究深度。

4.1.4 数据转换及评价能力是续航保障

当科研项目进入尾声甚至结题之后,期间产生的数据仍然有用。科研过程是连续的,数据也存在继承关系,一个科研项目中已经分析处理好的数据会在后续项目中成为更便捷、准确的数据基础,将科研项目中有再利用价值的数据进行存档,为科研后时代作续航保障。为确保这种续航能力的稳定有效,需对馆员数据的转换和评价能力进行培养,以满足后科研项目时代对数据素养能力的需求。

4.2 数据素养能力等级标准划分

本文将数据素养能力划分为三个等级,分别是基础级的数据馆员、进阶级的数据专家、最高级的数据科学家。本文仅对三个等级的评价标准进行定性描述,有关定量的研究将在后续研究进行。

4.2.1 数据馆员

作为数据素养培养的基础目标,图书馆将数据馆员的培养作为最低标准。现实情况中,通常要求馆员会用至少一种以上的技术方法及软件就可以达到数据馆员的要求,因而数据馆员可以为科研活动提供基础的辅助服务,但无法承担嵌入科研项目的数据服务。但理论上,数据馆员应达到前述标准中6大类一级指标及其下设的22个具体二级指标,如表2所示。

4.2.2 数据专家

进阶级的数据素养培养目标即数据专家。数据专

表 2 数据素养培养模式通览

一级指标	二级指标
数据意识	大数据思维意识
	认为数据是科研过程中的重要因素
	认为科学数据具有生命周期
	对待科研过程中的数据态度严谨认真，能够细化、具化地表达数据需求
数据获取能力	全面准确地从数据源处收集各类数据
	熟练使用数据收集方法及工具
	准确解读各类显性、隐性数据
数据分析处理能力	选择有效的数据分析方法
	掌握学科领域数据分析工具
	数据结果多角度呈现及建模
	对数据分析结果的认定、质疑、改进
数据转换能力	熟练了解并运用本学科领域数据标准
	熟练使用数据转换平台及工具
	利用数据分析结果支撑研究结论
	复用数据并实现其使用最大化
	利用数据分析提供决策支持
数据评价能力	验收数据正确性，剔除错误、无效数据
	确保数据的敏感、可信
	认识到数据的局限性并合理质疑数据
数据伦理	对数据真实性负责
	规范数据著录格式
	遵守学术道德及法律准则

家相较于数据馆员，除学科背景外，还要求需辅修过3门以上与数学相关的课程，包括但不限于统计学、应用数理统计、高等数学、线性代数、模糊数学、离散数学等。除有相应的数学知识作为基础外，对6大类指标及下设的22个具体子目标的培养要求也比数据馆员更高。不仅要会用，还要能熟练使用，数据分析处理软件的掌握个数也从1种提升为至少3种，只有这样才能满足嵌入科研项目的需求，真正做好科研服务工作。

4.2.3 数据科学家

作为数据素养的最高级别人才，数据科学家是数据分析处理领域级别最高、技术最强、眼界格局最宽广、人数最少的科研群体。数据科学家至少要拥有1个以上的博士学位（最好拥有专业学科方面的学位），且从事数据研究及处理工作3年以上，具备数据素养要求的6大类及22个子类能力，对所有通用流行的数据分析

处理软件都悉数掌握，并能对各类新方法、新技巧、新工具在第一时间学习并掌握，对数据有极高的敏感度且有独立的数据应用理论及使用体系，同时，其理论在业界得到普遍认可，并作为权威标准使用。

4.3 数据素养培养路径设计

4.3.1 构建以政策、制度和理论为导向的数据馆员长效培训机制

馆员数据素养的培养及提升是一个长期的过程，需要长效的培训机制。培训机制的建立不仅受各图书馆办馆理念影响，还受图书馆学会、情报研究机构及高校图书情报专业的影响。图书馆学会从宏观角度影响图书馆对数据素养的发展导向，通过制定相应的数据素养培训标准，可以为图书馆发展战略起到引导作用；情报研究机构从微观角度为图书馆馆员进行数据素养培养，对图书馆工作提供“术”方面的引导；高校图书情报专业从先进理论、研究范式方面对图书馆馆员数据素养培养的具体业务实践提供理论指导。基于宏观和微观、理论和实践建立的学科馆员培训机制，在纵向上有继承性，在横向上有发展性，学科馆员数据素养能力培训的政策、制度及理论导向要以培养能够胜任实际工作的数据馆员为目标，而长效培训机制的建立正是从顶层确定培养路径的战略，为培训细则的制定提供基础性依据。

4.3.2 打造方法、技术过硬的数据馆员队伍

数据馆员数据素养最终要以服务科研为目的，在这个过程中，具备数据分析处理的方法、技巧是数据馆员的立足之本。对澳大利亚、新西兰、爱尔兰和英国四国图书馆馆员的调查显示，阻碍图书馆提供科研数据服务的首要原因是“需要额外的知识和技能”^[21]，数据馆员技能的匮乏成为图书馆提供高级服务的瓶颈；英国联合信息系统委员会指出图书馆在数据管理服务中的重要角色之一是“培养和提供数据图书馆馆员”^[22]，数据馆员队伍的建设不仅是近十年图书馆学界发展的要求，也是未来业界发展的长久趋势。由此可见，作为数据管理服务的提供者，数据馆员应着重提升自身在数据意识、数据平台使用、数据处理软件应用、数据分析方法及数据伦理精进等方面的技能，才能满足在数据密

集型科研范式下对数据馆员队伍业务能力的需求。

4.3.3 创新培养模式适应嵌入科研的现实需求

数据馆员的数据素养模式创新,应该以满足现实的科研需求为目标。传统培养模式以课堂培训、业务学习、案例实习为主,均是以模拟的形式进行培养,而围绕科研活动的生命周期,跟随科研项目提供可用数据服务的全实境模式下的培养方式并不多见。针对科研的不同阶段,开展形式多样的数据能力培训,图书馆在数据组织管理方面得天独厚的优势,通过学科馆员与所服务的研究项目间已形成的良性互动平台,学科馆员可以实现角色的升级和转换,即通过嵌入项目的实践转变为数据馆员,与此同时,以促进数据资源的交互为任务的数据馆员,在嵌入科研服务的过程中,积极掌握相关学科知识,结合自身数据素养能力的优势,更有利于数据专家,乃至数据科学家的培养。无论是从学科馆员向数据馆员转变,还是数据馆员通过嵌入项目向高级形式发展,都是以新的方法、思维和角度去适应数据密集型科研的现实需求。此外,哈佛大学图书馆还基于自身数据资源管理经验,实施了通过数据管理平台开展数据素养教育的实践,通过培训中介达到对外宣传数据管理知识的目的,为科研用户获取共享数据提供便利^[23]。

5 结语

数据密集型科研范式下的图书馆馆员需要具备数据素养能力,而这种能力的培养需要通过课程培训、嵌入科研项目等多种方式方法才能得以提升。数据意识、数据获取能力、数据分析处理能力、数据转换能力、数据评价能力及数据伦理是数据素养的整体要求,作为图书馆数据馆员的培养目标需要逐一掌握。在政策引导下,通过嵌入科研项目的历练,根据数据素养培养框架有条不紊地将数据馆员培养成数据专家乃至数据科学家,是当今图书馆在大数据时代背景下正在探索的艰难路径。

参考文献

- [1] NEGROPONTE N.数字化生存[M].北京:电子工业出版社,2017.
- [2] DATA LITERACY:Bridging the Gap[EB/OL].[2017-09-08].http://

- ideas.aetn.org/edweb/tess/data-literacy/TESS%20Data%20Literacy%20Presentation.pdf.
- [3] British Library.Livingknowledge:the British library 2015—2023 [EB/OL].[2017-10-16].http://www.bl.uk/aboutus/foi/pubsch/pubscheme3/living-knowledge-2015-2023.pdf.
- [4] JOHNSON L,ADAMS B S,ESTRADA V,et al.NMC horizon report:2015 library edition[EB/OL].[2017-09-22].http://cdn.nmc.org/media/2015-nmc-horizon-report-library-EN.pdf.
- [5] 王晰巍,张长亮,蔡佳铭,等.大数据环境下中美高校信息素养培养模式比较研究[J].图书情报工作,2016,60(11):29-35.
- [6] 魏攀.美国雪城大学信息研究学院教授秦健博士来我馆做讲座 [EB/OL].[2017-08-13].http://gzwl.lib.whu.edu.cn/pe/Article/Print.asp?ArticleID=765.
- [7] 伍芳芳.大数据时代图书馆数据科学家培养路径探析[J].科技风,2017(14):278.
- [8] Data Management Planning[EB/OL].[2017-08-14].https://data.research.cornell.edu/content/data-management-planning.
- [9] 张军.面向科研第四范式的科研人员数据素养培养研究[J].图书与情报,2016(2):133-136.
- [10] Data literacy[EB/OL].[2017-08-02].https://en.wikipedia.org/wiki/Data_literacy.
- [11] 百度百科.数据素养[EB/OL].[2017-08-03].https://baike.baidu.com/item/%E6%95%B0%E6%8D%AE%E7%B4%A0%E5%85%BB/922411?fr=aladdin.
- [12] STEPHENSON E,CARAVELLO P.Incorporating data literacy into undergraduate information literacy programs in the social sciences: a pilot project[J].Reference Services Review,2007,35(4):525-540.
- [13] VAHEY P,RAFANAN K,PATTON C et al.A cross-disciplinary approach to teaching data literacy and proportionality[J].Educational Studies in Mathematics,2012,81(2):179-205.
- [14] 缪其浩.大数据时代:趋势和对策[J].科学,2013,65(4):25-28.
- [15] 张静波.大数据时代的数据素养教育[J].科学:上海,2013,65(4):29-32.
- [16] 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.
- [17] 王佑镁,杨晓兰,胡玮,等.从数字素养到数字能力:概念流变、构成要素与整合模型[J].远程教育杂志,2013,31(3):24-29.
- [18] 宋文.“互联网+”时代图书馆员数据素养的构成与培养[J].河南图书馆学报,2016,36(10):91-92.
- [19] 黄如花,邱春艳.国内外科学数据元数据研究进展[J].图书与情报,2014(6):102-108.
- [20] 宫学庆,金澈清,王晓玲,等.数据密集型科学与工程:需求和挑战[J].计算机学报,2012,35(8):1563-1578.

- [21] CORRALL S, KENNAN M A, AFZAL W. Bibliometrics and research data management services: emerging trends in library support for research[J]. Library Trends, 2013, 61(3): 636-674.
- [22] SWAN A, BROWN S. The Skills, Role and Career Structure of Data Scientists and Curators: An Assessment of Current Practice and Future Needs[EB/OL]. [2017-10-23]. <http://www.jisc.ac.uk/media/documents/programmes/digitalrepositories/dataskillscareersfinalreport.pdf>.
- [23] Dataverse Project. Open source research data repository software [EB/OL]. [2017-04-22]. <http://dataverse.org/>.

作者简介

董薇，女，1982年生，馆员，研究方向：信息素养教育，E-mail: dongwei@mail.neu.edu.cn。

Analysis of Library Librarian Core Data Literacy Ability in Data Intensive Research Paradigm

DONG Wei

(Library of Northeastern University, Shenyang 110819, China)

Abstract: Data intensive scientific research paradigm, librarians' literacy has become the main aspect of the whole data of core competence. The article from concept and connotation of quality data, based on the current situation of domestic outside the staff on the analysis of data quality, puts forward the framework of data quality of librarian training and the data classification literacy level, namely data librarian, data experts, data scientists, and finally proposes the path through policy guidance to enhance the skills for the training goal, ability of data by embedding the research project, the completion of data intensive primary path analysis to cultivate librarians research paradigm and the construction of the core competence.

Keywords: Data Literacy; Data Intensive Research; Embedded Research; Librarian

(收稿日期: 2017-10-26)

■ 书 讯 ■

《中国高被引分析报告2016》

《中国高被引分析报告2016》按理、工、农、医、人文、社科等领域划分为50个学科，综合分析了各个学科的高影响力论文、研究热点与前沿、高影响力期刊、高影响力作者和高影响力科研机构，并以关联图谱的方式展现了多种学术关系，有助于科研人员及时发现并跟踪研究热点，有利于期刊编辑部监测本刊学术影响力，有利于科研管理机构评估科研能力，是高等院校、科研院所及期刊编辑部等相关单位和人员的参考工具书。

该书以“中国知识链接数据库”为依托，数据覆盖我国6 000余种期刊的论文及引文。书中分学科揭示了高影响力的学者、研究机构（大学、研究所、医院等）、地区（省/自治区/直辖市）、学术期刊、图书、外文期刊和会议录，并采用共词分析、共被引分析和合著分析等方法绘制出各学科的前沿主题分布以及作者、机构和期刊间关联的知识图谱。

《中国高被引分析报告2016》由中国科学技术信息研究所编制，曾建勋主编，科学技术文献出版社出版。《中国高被引分析报告2015》也已同步发行，欢迎业界同仁鉴阅订购。