

项目结题验收单

专家验收表（主持人所在单位组织 3-5 名专家对项目进行验收、自评。）

项目名称	大数据生态下的科研人员数据素养能力与评价研究				
主持人	王丹	职务/职称	研究馆员		
所在单位	东北农业大学图书馆（加盖单位公章）				
专家意见	<p>课题研究始终按预定计划有序进行，针对研究主题进行了可靠文献参考与实际研究，报告整体行文流畅，内容详实，思路清晰，较好的完成了项目研究工作。报告针对大数据环境下的科研人员数据素养能力提升与效果评价进行了可靠的设计，并进行相关内容实践，保证研究的合理性与可行性，为数据素养能力与评价的后续研究与发展提供数据支持，研究有较强的现实意义。经讨论，专家组一致认为，项目研究内容清晰完整，并具有较强现实意义，可以结题。</p> <p style="text-align: right;">（如需要可增加页数）</p>				
专家签字					
职务/职称					



项目编号：2020007

注：项目编号请查看立
项通知，也可缺省

CALIS 全国农学文献信息中心研究项目 结题报告

项目名称： 大数据生态下的科研人员数据素养能力与评价研究

项目关键词： 大数据 数据素养 能力评价

项目单位（盖章）： 东北农业大学图书馆

通信地址：(详细地 黑龙江省哈尔滨市香坊区长江路 600 号 东北
址含邮编) 农业大学图书馆 317 室 150030

项目主持人： 王 丹

联系电话： 15046091517 0451-55191598

电子邮件： Wangdan598@163.com

提交日期： 2021 年 4 月 25 日

(结题报告含有以下 5 部分内容，其他内容根据项目情况可增加，
字数不少于 4000 字)

报告题名：大数据生态下的科研人员数据素养能力与评价研究

关键词:大数据 数据素养 能力评价

1 研究背景、目的及意义

科学数据是知识发现与科技创新的重要基础。随着数据密集型科研和第四科研范式的兴起，科研工作者面临一系列数据管理与使用问题。为了进一步加强和规范科学数据管理，保障科学数据安全，提高科学数据开放共享水平，2018 年 3 月，国务院办公厅正式出台了《科学数据管理办法》，2019 年 2 月中国科学院随即发布了《中国科学院科学数据管理与开放共享办法（试行）》，我国在科学数据开发利用、开放共享和安全保护等方面迈出了坚实的步伐。随着国家对科学数据管理与共享的深入推进，对科研人员科学数据素养的要求越来越高，科学数据素养成为科研人员学术研究与交流的必备能力之一。

科学数据素养能力，是指在科学研究过程中，科研人员作为数据的生产者、使用者和管理者，收集、处理、管理、评价和利用数据进行科学研究所涉及的思维、知识和技能，以及在数据生命周期中普遍遵循的伦理道德与行为规范。其重点强调对科学数据的理解、利用和管理能力，目的是将数据转化为知识。

科学数据素养是新兴的跨学科研究领域，国外在科学数据管理与科学数据素养教育方面的研究起步较早。2007 年美国学者冈特（G. A. Gunter）和卡尔森（J. Carlson）等人就指出了数据素养的重要性，并规划了数据素养的发展策略。而国内关于数据素养的研究则是近几年才开始，缪其浩、何涛、张静波等人分别从认识角度、科研角度、数据周期角度定义了数据素养的含意。但由于数据素养研究缺乏统一的理论基础体系，大多数研究都局限于侧重探讨数据素养的培养内容与模式，关于我国科研人员科学数据素养的知识结构和能力体系框架尚未完全建立，因此，本研究将重新把握科研人员的科学数据素养能力特征，从而构建符合当科研人员实际的科学数据素养能力及评价体系，为科学数据素养相关研究提供理论基础。

在当今数据密集型的科研环境中，科学数据的作用越来越重，然而，在“重论文、轻数据”的学术环境中，当前高校研究者的数据素养水平有限，数据意识淡薄，使得高校科学数据管理服务困难重重。科学数据管理不仅是政策、技术的问题，更是数据素养的问题，科研用户必须将数据态度、数据意识、数据知识、数据技能、数据伦理等几个方面的要素直接渗透到科学研

究生命周期中，才能使数据发挥出创新的魅力。大数据生态下，数据素养教育的内容和特征需要深入探讨，从而制定出科学的数据素养能力的标准和评价体系。

2 研究内容及方法（思路、方法、具体内容）

2.1 数据素养能力的内涵

大数据时代的到来，推动着数据成为了社会发展的主要驱动力，掌握对数据的获取、管理和应用能力，即具备一定的数据素养也便成为大数据时代社会人的基础要求。传统的信息素养教育也难以适应社会发展的需求，将信息素养教育向数据素养教育方向发展也成为新时代 LIS 教育的学科建设发展态势。但目前，学界对数据素养的定义并未达成统一认识。如 Calzada 等认为数据素养“包含获取、理解 数据、批判性评估和管理数据的能力，并在使用数据的过程中遵守道德规范”。伊利诺伊香槟分校的 Hogenboom 等将其定义为“阅读、解释、分析、批判性思考以及将统计数据作为证据的能力”。国内胡卉提出“数据素养是正当地发现和获取数据、批判地选择和评估数据、规范地管理和处理数据、合理地利用和共享数据的意识和能力”。英国数据档案馆根据数据生命周期提出了各阶段的数据素养内容，包括创建数据、处理数据、分析数据、保存数据、共享数据和复用数据等 6 个方面；LIS 学科中对数据素养的要求主要包括数据知识、数据获取、数据解释与评估、数据管理、数据应用五个方面。综合各方观点，笔者认为数据素养强调的是数据生命全周期内的数据发现、获取、选择、评估、管理、保存、共享等一系列所需的意识和能力。

2.1.1 数据的发现与获取

数据的发现与获取能力是科研人员利用数据的基础能力。从行为方式角度来说，就是指点用户对于数据的感知并收集数据的能力，即对数据的敏感性。用户能够认识到数据的价值，包括对科学数据的使用意识、需求意识和法律意识，这一能力的形成需要长期的自我学习和有效的社会引导，数据使用者拥有这种能力后才能在海量数据中实现“低成本大产出”。大数据并不是讲究“大”，如果不具备对大数据的感知与获取能力，大数据就会成为影响科研工作与学习交流的障碍。学界经常会谈到大数据的核心价值，那么大数据的这种支撑与趋势价值的实现，对于科研人员发现与获取能力的要求也是比较高的，“数据驱动决策”，有效发现并甄别数据真伪在数据爆炸的今天值得重视，并不是每一个用户都具备真实有效的数据信息筛选能力，因此，做到对数据的敏感与发现是数据能力的基础，只有拥有数据敏感能力的人才能从这些数据中提取信息、发现知识和找准趋势。

2.1.2 数据的组织与存储

大数据时代，科学数据的特点更明显的表现为数据量大、变化快、分布式异构且多源、数据多维关联、计算分析一体化、跨区域协作、学科差异较大等，因此，按照数据属性与学科特点，构建适合数据组织与存储的科学数据管理平台成为数据利用前的最要操作。在我国，科学数据管理平台建设已取得一定的进展，形成了各自的特色，为科学数据的开放共享发挥了重要的作用。在科学数据的存储与获取方面，科研数据管理人员按照标准统一的元数据描述进行数据管理，提高了科学数据的可发现性，更便于用户对科学数据的发现和获取。目前科学数据的存储大致分为两个层面：一是采用分布式文件系统 HDFS 存储非结构化数据；二是采用加密算法和云存储技术相结合，对存储的数据进行加密，进而存储于云服务器上，保证了数据的安全存储。不同数据平台按照统一的标准规范对科学数据进行整合后，实现了数据资源物理分布、逻辑集成、统一访问和服务的功能，为用户提供一站式、集成的数据资源发现、检索与下载服务。

2.1.3 数据的分析与利用

数据分析是指用适当的统计分析方法对收集来的大量数据进行分析，提取有用信息和形成结论而对数据加以详细研究和概况总结的过程，是为了寻求问题的答案而实施的有计划、有步骤的行为。利用数据分析来总结经验、发现规律、预测趋势，发展决策才能更加科学、精确、合理。对于科研工作来说，数据分析与利用是整个数据周期中最重要的一环。无论是数据发现与获取，还是存储与管理都是为了最后的数据分析与利用。数据分析的目的是把隐藏在一大批看来杂乱无章的数据中的信息集中提炼出来，从而找出所研究对象的内在规律。在实际科研工作中，数据分析可帮助用户做出正确的判断。如何有组织有目的地收集数据、分析数据，使之成为信息并能有效利用的过程是科研人员数据素养的闭环所在，也是科研人员更行之有效进行科学研究的必备素质。

2.2 大数据生态下数据素养能力的教育与发展

2.2.1 数据意识的培养

数据意识是数据素养能力基本构成部分，其要求用户对于现实环境的需求和变化保持高度的敏感性，并尽可能激发出自身的探索需求，其不仅涵盖本学科数据内容的基础意识，也涉及到多元领域和跨学科层面上的数据拓展意识，同时涵盖贯穿其生命周期全过程的伦理道德上的内容。从收集数据开始，对数据进行有效存储及管理，重复和连续使用及分析数据，在此基础

上培养用户的数据素养能力。在高校层面，除专门的信息素养教育课程教育外，图书馆应成为推动落实数据素养的重要场所，通过革新机制与内容，帮助用户成为数字时代的创造者和消费者，发挥图书馆知识管理与知识服务的专业优势，创造条件确保图书馆发挥自身技术力量帮助用户培养数据素养。图书馆承担着社会教育和文化资源保存的职能，大数据时代要求图书馆在服务方式和内容上必须做出调整，不仅要成为社会的数据中心，也必将成为大数据时代数据素养的培养机构，针对不同用户群体的需求，开展个性化的培训及数据素养教育活动。

2.2.2 数据素养的提升

国外高校开展数据素养教育的主要模式包括通识数据素养教育和学科专业数据素养教育。通识数据素养教育主要面向学生，多为由大学图书馆主导，与其他部门合作开展。学科专业数据素养教育主要是针对具体学科专业，将该专业需掌握的数据素养技能嵌入到专业课程当中。无论是 IMLS、ALA、JISC 等机构资助国外高校开展数据素养项目，还是各高校具体开展的数据素养教育，都充分体现了合作的精神，所以，丰富数据素养教育内容与多部门协作是更容易达成的数据素养教育提升方式。同时，制定数据素养框架能够帮助区分数据素养与信息素养的差异，明确数据素养的内涵并且指导数据素养教育的开展与提升。面对深刻的社会信息环境变革，我国政府、教育部门亟需出台相关的政策和实施办法来提高用户数据素养教育水平，以帮助其适应新的数字化生存环境，从而推动国家社会经济发展。

表 1 用户数据素养发展阶段性要素

培养阶段	课程种类	技能形成
初级阶段	通识性教育课程；	数据素养教育基本知识完善；
	启发性课程；	数据素养的元技能基本形成；
中级阶段	数据素养基本知识课程；	数据素养意识形成；
	数据素养管理课程等；	学会发现分析使用数据； 掌握一定的数据可视化分析工具使用；
高级阶段	数据素养专业研究领域课程；	数据查找、存储、分析利用能力形成；
	数据处理及应用课程； 数据伦理与数据规范课程；	数据驱动决策思维形成； 数据质量评估思维形成；

2.3 数据素养能力评价

2.3.1 用户数据素养能力的特点

当前面向多元环境的形成和变化，用户的数据素养逐步呈现出适应性和进步性的特点，用户在数据演化全过程中表现出协同性、复杂性和动态性。用户一方面需要进行策略性的数据采集并在此基础上开展相应的数据使用行为，进而服务于自身综合发展的需求；另一方面其也在利用各种途径为第三方提供数据并在渐进累积中完成数据的整合，以实现数据生命周期内其效用和价值的转化，同时还不断开展数据的反思、重用和创新，进而在实践层面上完善储备技能和进行数据效能挖掘；最后，在数据密集型科研背景下，用户需要在数据伦理规范、挖掘利用、引用评价等能力方面予以提升，进而在科学数据开放共享基础上实现知识发现和科技创新。因此，在多维视阈下围绕核心发展目标，依托现代信息技术的发展和信息学科的深度整合，以跨学科思维不断拓宽人才培养核心内容的深度和广度，保持新时代 LIS 教育生态体系的动态平衡，才能实现面向社会需求和国家人才战略目标的良性发展。具体内容如图 1 所示。

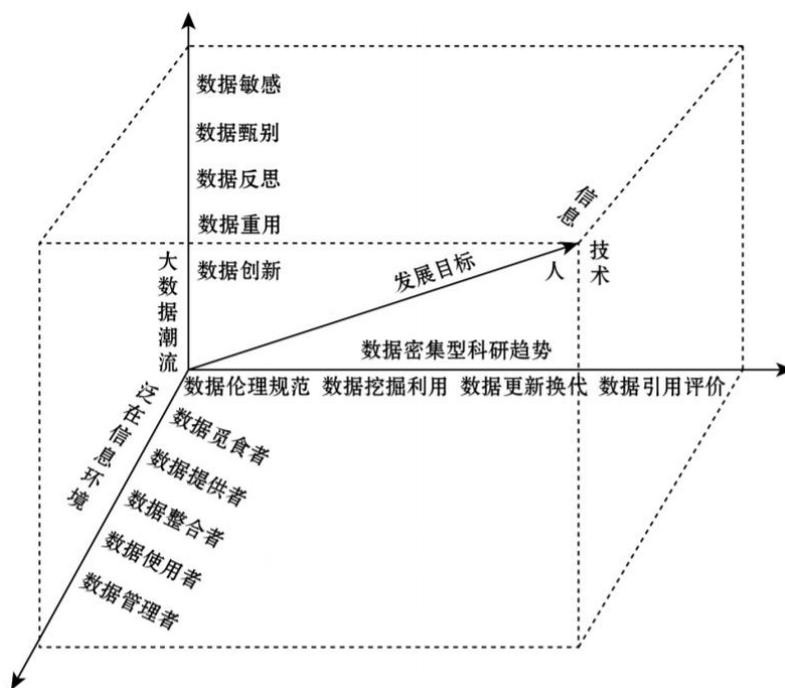


图 1 多元环境下用户数据素养新特征图

2.3.2 数据素养能力的评价指标体系

按照国家战略和科研工作实际需求，从大数据环境基础考虑用户数据素养的评价体系，本文设计并拟建了包含 4 个一级评价指标及 28 个二级评价的评价体系，具体指标如表 2 所示。

表 2 用户素养能力评价指标体系

一级指标	二级指标
数据意识	了解什么是数据、数据与其他概念之间关系 能因数据表象激发探索冲动并探寻其深层内容 谨慎应对个人数据获取、数据分析处理的全过程 能够清楚并以适当方式表达对所处学科的数据需求 能够利用已有的数据库及专业搜索引擎获取研究数据 关注学科相关的以及跨学科的数据内容 注重对数据外延内容和信息技术的学习 知晓数据伦理及道德的基本概念和要素
数据获取与整合	能够依据不同情境有针对性地制定数据检索计划 掌握一种或多种数据检索工具和技能 能够有效利用检索过程中的数据线索 能够对数据获取全过程进行效果评估和风险规避 注重将多源异构数据转化为同构数据 掌握一定的数据集成方法
数据管理与应用	掌握一种或多种数据存储方式 周期性对已有数据进行备份并能采取措施以保护数据安全 掌握数据分析的操作流程，并能完整规范操作 能够利用数据分析工具对数据进行深入分析并得到有效结论 了解数据可视化的概念等内容 对数据的蕴藏价值进行挖掘 能够将已有的数据进行实践应用转化
数据交流与创新	注重与其它个人与机构的数据共享 经常通过共享数据实现自身与他人的数据需求 注重对学科数据的聚集、优化和动态分配 注重多学科的数据联合、组合和评价 能够保持客观辩证态度对数据进行合理质疑 认识到多学科融合环境下数据创新的重要性 曾利用数据进行创新实践

数据意识要求用户对于现实环境的需求和变化保持高度的敏感性，并尽可能激发出自身的探索需求，涉及到多元领域和跨学科层面上的数据拓展意识，同时数据伦理道德上的内容。

数据获取与整合不仅仅完成基本的数据采集，还需要面向数据演化的全过程以专业性、精确性的数据效用进行策略性的数据获取，并将多源异构数据进行集中整合。

数据管理与应用是在采取多样且安全的存储方式的基础上，利用专业化工具并依照规范化流程开展数据分析，以帮助用户在实际应用中辅助数据决策和完成数据效能的价值转化。

数据交流与创新是跨学科的思维完成数据的联合和评价，进而实现数据的高度共享、深层交流以及合理质疑基础上的数据创新。

3 结论与建议

目前，我国对于数据素养能力相关内容研究还处于初级阶段，虽然部分高校或科研机构借鉴国外已开展情况进行了相应的改革与发展，但整体人员数据素养能力与评价内容并不完善，无标准性可言，数据素养的环境氛围还没有形成，有关部门的相关政策与指导意见支持还比较欠缺，所以，现阶段主要的提升与评价手段都在探索过程中。参照国外各高校及机构发展内容及行为，开展全方位的数据素养教育是当前提升用户数据素养能力切实可行的有效途径。结合目前高校科研人员实际情况，建议从以下几方面探索用户数据素养教育的研究发展。

3.1 制定相关教育管理政策

对于科研人员聚集的高校系统，应该积极的建立关于数据素养教育的实施细则与管理办法，通过专门机构的指导与规划，加快高校科研人员对于数据素养的重视与认可，并通过一定的评价体系对所有人员进行可持续的数据素养评价，保证环境影响下的政策的可持续发展，使更生更具操作性与科学性。

3.2 设立数据素养教育课程专题

以信息素养教育课程为基础，设立数据素养教育课程专题，与科研活动相结合，从科研周期中使用户快速认识数据素养提升的重要性，以科研活动为载体，引导用户自主走向数据素养评价，完成数据素养提升过程并实现能力评价，为科研创新与发展的细粒性挖掘做好准备。

3.3 加强图书馆数据人才的培养

高校图书馆特有的教育与服务职能要求图书馆人员必须具有数据特质。做为科学数据管理的重要场所，学校信息素养教育的主阵地，图书馆在学校科研人员数据素养提升与评价过程中起着重要的作用。图书馆员应持续参加“数据科学家”培训与考核，通过继续教育，丰富并提升馆员的数据素养，为其它科研人员的指导与培训做好准备。

3.4 促进数据共享与交流

无论是数据使用个体还是数据管理团体，加快数据共享的行为对于数据素养能力的提升都是高效的途径。数据的共享与交流更容易激发数据潜在价值的显现，提升科研工作进程的同时

找到科研增长点，为科研工作的深入开展指明方向，同时为科学决策发展提供有力支持。

综上所述，数据素养研究的发展，无论是对当前的科研工作还是对未来的人才战略都是十分重要的，提高重视才能迎接挑战。

4 项目成果（发表的文章、开发的软件、取得的实践效果等）

项目研究形成科研论文 2 篇：

王丹、王玉龙《大数据生态下高校科研人员数据素养评价——以东北农业大学为例》，投稿中；

王丹、张杨《农科“双一流”高校数据素养教育发展调研》，投稿中

5 参考文献

- [1]雷春蓉,陈梦.国内外高校图书馆数据素养教育比较研究[J].图书馆,2021(04):47-51.
- [2]张群,印熙媛.国内外数据素养教育研究现状及启示[J].数字图书馆论坛,2020(12):65-72.
- [3]梁宇,郑易平.高校博士生数据素养的影响因素及应对策略[J].情报理论与实践,2021,44(04):146-152+137.
- [4]贾璞,宋乃庆.大数据时代中学生数据素养：内涵、价值与构成维度[J].电化教育研究,2020,41(12):28-34+58.
- [5]王焱,袁先文,林晨晨,程结晶.高校图书馆数据素养教育的现状及其要素分析[J].数字图书馆论坛,2020(11):42-49.
- [6]黄如花.面向高质量发展的数据素养教育[J].图书馆建设,2020(06):26-29.
- [7]卢祖丹.我国高校图书馆数据素养服务供给实证评价研究[J].图书馆杂志,2020,39(10):63-70.
- [8]代沁泉,熊回香,陈琦,李昕然.中国内地 iSchool 院校学生数据素养评价研究及启示[J].图书馆学研究,2020(17):16-27.
- [9]卢祖丹,余达淮.我国高校图书馆数据素养教育内容分析[J].数字图书馆论坛,2020(08):36-41.
- [10]秦小燕,初景利.科学数据素养能力评价指标体系构建研究[J].图书与情报,2020(04):56-66.
- [11]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.
- [12]张娟,杨文建.国外数据素养教育研究及其启示[J].数字图书馆论坛,2020(04):60-67.
- [13]邢变变,杨晴晴.大数据时代档案专业人员数据素养提升探析[J].档案与建设,2020(04):38-42.
- [14]余维杰,周娅莉,吴锦池.我国研究生在科研活动中的数据素养现状研究——以双生命周期理论为视角[J].图书情报工作,2020,64(07):84-93.
- [15]焦皎.大数据时代高校图书馆数据素养教育研究[J].图书馆工作与研究,2020(03):81-84.
- [16]张路路,张群,贾洁.国内外数据素养评价研究进展[J].图书馆理论与实践,2020(02):56-61.

- [17]付超.大数据时代我国数据素养分类研究[J].图书馆理论与实践,2020(02):68-74.
- [18]余维杰,陈思琪,陈序.国内外数据素养研究的文献计量分析[J].图书馆理论与实践,2019(12):32-39.
- [19]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.
- [20]秦小燕,初景利.面向我国科研人员的科学数据素养能力评价研究[J].情报理论与实践,2020,43(02):21-27.
- [21]周志强,王小东.大学生数据素养提高的途径与对策研究[J].情报科学,2019,37(09):79-84.
- [22]陈媛媛,王苑颖.加拿大数据素养教育实践及启示——以达尔豪斯大学为例[J].情报理论与实践,2019,42(06):166-171.
- [23]郭倩,李建霞.基于多元主体的高校数据素养教育生态模式构建研究[J].图书馆理论与实践,2019(05):49-57.
- [24]张长亮,王晨晓,李竟彤.大数据时代中美高校数据素养教育比较研究[J].情报理论与实践,2019,42(08):131-137.
- [25]宋甲丽,程结晶.高校图书馆科学数据素养教育现状调查及建议[J].图书馆学研究,2018(19):2-9.
- [26]付超.大数据背景下公民数据素养提升策略探析[J].图书馆理论与实践,2018(08):7-11.
- [27]潘雪,陈雅.我国高校数据素养教育评价指标体系探析[J].新世纪图书馆,2018(07):31-34.
- [28]张新红.国内外高校图书馆科学数据素养教育比较分析[J].图书馆学研究,2018(11):9-13+39.
- [29]肖希明,倪萍.中美 LIS 教育中数据素养教育的调查与比较分析[J].图书与情报,2018(01):20-27.
- [30]刘雅馨,杨现民,李新,田雪松.大数据时代教师数据素养模型构建[J].电化教育研究,2018,39(02):109-116.
- [31]杨文建.大数据环境下的教师数据素养研究[J].图书馆理论与实践,2017(11):102-107.
- [32]邓李君.国内外数据素养教育发展现状研究[J].图书馆理论与实践,2017(07):30-33.
- [33]王晓文,沈思.国外科研人员数据素养教育述评及启示[J].情报资料工作,2017(03):102-106.
- [34]李立睿,邓仲华.“互联网+”视角下面向科学大数据的数据素养教育研究[J].图书馆,2016(11):92-96.
- [35]郝媛玲,沈婷婷,高珊.高校数据素养教育实践的思考和建议——基于哈佛大学案例和我国图书情报人员访谈的分析[J].图书情报工作,2015,59(12):44-51.