

医学生数据素养评价体系构建及调查研究^{*}

王于心 贺培凤 于琦 贾晨晨

刘鸿齐

(山西医科大学 晋中 030600)

(山西医科大学第二医院 太原 030001)

〔摘要〕 介绍国内外数据素养研究现状, 构建医学生数据素养评价体系, 采用抽样方法对山西省医学研究生进行调查分析, 以期了解医学生数据素养水平与需求, 为医学高校开展数据素养课程改革提供建议。

〔关键词〕 医学生; 科研数据素养; 评价体系

〔中图分类号〕 R-056 〔文献标识码〕 A 〔DOI〕 10.3969/j.issn.1673-6036.2020.07.008

Building and Investigation of Data Literacy Evaluation System for Medical Students WANG Yuxin, HE Peifeng, YU Qi, JIA Chenchen, Shanxi Medical University, Jinzhong 030600, China; LIU Hongqi, Second Hospital of Shanxi Medical University, Taiyuan 030001, China

〔Abstract〕 The paper introduces the current status of data literacy study at home and abroad, builds data literacy evaluation system for medical students, investigates and analyzes medical graduate students in Shanxi Province by using sampling method, so as to understand the level and needs of medical students' data literacy, and provide suggestions for medical colleges and universities to carry out data literacy curriculum reform.

〔Keywords〕 medical students; scientific research data literacy; evaluation system

1 引言

在大数据技术发展前期, 国内外不少研究者对信息素养进行探索, 随着数据管理的深入, 逐渐发现仅用信息素养已不能完全概括大数据时代数据管理能力相关指标。数据素养以信息素养为基础, 但不限于信息素养, 包含信息素养核心要素且得到升

级, 更能够适应大数据时代需求^[1]。学界对数据素养主要的研究思路是“溯源——成长——精准——架构”, 对此全面的解释是: 数据素养溯源于信息素养, 在信息素养拓展中得到深化, 在信息素养不同诉求中精准概念, 最后在信息素养的教育中得到架构^[2]。本文所论述的数据素养主要是指科研及相关工作人员在从事科学研究时应具备的一种能力。

2 数据素养研究现状

2.1 概念尚未明确

2.1.1 国外 国外最先提出数据素养概念, 其中最为著名的是西班牙马德里卡洛斯三世大学的数据素养列表, 见表 1, 首次创新性地将数据素养模型上升到理论, 可以看出对于数据素养的要求集中在

〔修回日期〕 2019-12-19

〔作者简介〕 王于心, 硕士研究生; 通讯作者: 刘鸿齐, 博士。

〔基金项目〕 国家自然科学基金面上项目“基于多元分析的科研文献微观实体评价理论与实证研究——以生物医学为例”(项目编号: 71573162); 山西医科大学第二医院教学项目基金“医学生数据素养培养体系构建与应用研究”(项目编号: 201803-2)。

理解、获得、管理、处理、利用数据 5 大方面。Milo Schield 等^[3]指出数据素养是科研者能够根据自身科研需求,从不同数据来源中对数据进行获取、处理、总结,最终得到结论的一种能力。Tibor Koltay 等^[4]认为数据素养不仅仅是获取研究数据,而是能够对获取数据进行批判式的选择与利用。还有一些学者认为科学数据素养的核心能力包括正确理解、评价、管理、利用数据等^[5]。Clarence Maybee 和 Lisa Zilinski 通过分析目前现有的数据素养模型,结合相关理论,找出影响数据素养的关键因素,首次概括总结出数据素养应包含的 7 大类要素,提出数据素养通用模型,见表 2。

表 1 数据素养能力列表

一级指标	二级指标	具体涵义解释
理解数据	什么是数据	对数据的定义、内容进行了解,清楚数据的意义,以及有能力辨别清楚数据的各种变形形式
	社会中的数据	数据在社会中存在着不同的角色,需要对数据的产生、使用、生命周期、应用等进行了解。有能力意识到它们是如何产生的,并在使用中会产生什么意义
发现、获取途径	来源	有能力在多渠道多途径中找到数据来源,并在众多数据来源中辨析有用数据
	获取途径	获取数据的主要途径,面对需解决的主要或部分问题,能够采取新的途径获得有效数据
评估数据	解释评估	意识到数据的不同表现形式后,能够有效理解并将数据进行重现,能够有效对数据进行评估
	数据	
管理数据 利用数据	数据收集与管理	对元数据、数据库以及数据管理工具进行了解,认识到数据使用和保存的重要性
	数据处理	能够有效利用工具,如 R 语言、SPSS、EXCEL 等对数据进行分析处理
	数据重现	能够有效利用工具,如 Gap minder、Visually 等对数据的内容进行综合并有效重现
	数据伦理	数据伦理包括数据使用和共享道德,有意识地了解数据来源并进行标引,符合数据使用规定

表 2 数据素养通用模型要素

元素	元素涵义解读
意识	理解数据的存在意义,了解数据的价值
获取	了解如何定位、获取、利用数据,以及从数据库或其他形式中挖掘所需数据
从事	分析,解释和组织数据,并利用现有数据做出决策
管理	对数据进行开发和存储
交流	将数据利用工具进行可视化,包括图表和图示
道德	尊重数据来源,了解并遵从数据再利用所涉及的版权问题
保护	主要侧重于对需长期保存和利用的数据进行管理

2.1.2 国内 张静波等^[6]认为数据素养主要强调的是科研人员对数据采集、组织管理、分析处理以及协调共享的能力。孟祥保等^[7]认为数据素养包括 3 个层次:第 1 层是具有数据意识,能够意识到数据的重要性;第 2 层是具有处理、利用数据的技能;第 3 层是具有数据道德。杨晓琼等^[8]对数据素养概括为 4 个方面,分别是数据意识、评估、技能、道德。纵观国内外不同学者对数据素养概念的解释,可以看出学者们对数据素养中的意识、利用处理、共享、道德等方面较为关注。

2.2 缺乏统一评价体系

经过多年研究与探索,欧美国家在数据素养标准建立方面已很成熟。虽然亚洲国家通过借鉴欧美国家如 ACRL、SCONUL 等标准,结合本国实际情况制定相应的素养体系,但是仍未形成系统的具有权威性的数据素养标准。目前国外已有相关的数据素养标准或指标体系,但针对医学生的数据素养指标体系大多是基于现有的高等教育指标体系标准,如澳大利亚^[9]开展的高等教育数据素养测评中发现通用的高等教育指标体系并不适用于医学生。我国数据素养标准面临着与模式相互分离的问题,成熟的数据素养模型应能够协调标准模型之间关系,然而现有的大多数数据素养研究往往将标准与模式分离研究,二者并没有被有效联系并加以实质性运用^[10]。

2.3 缺乏数据素养教育

目前国内已有部分大学开展数据素养教育课程,但是开设院校较少且课程量不足,如哈尔滨工业大学和南京大学开设的大数据算法和用 Python 玩转

数据课程等，都偏向于编程或理论，未实现专业化^[11]。由于医学生专业特殊性，在课程安排上大部分是相关的专业课程，关于数据素养的课程不仅设置单一，且课程量仅为 18 ~ 36 个学时，难以满足现实需求^[12]。此外学生缺乏数据素养意识，在接触有关课程时会因其不是医学专业课或是意识不到重要性而忽视数据素养的培养。医学生应具备良好的数据素养能力，从而及时正确获取与分析数据信息，掌握学科最新的科研成果，为以后的科研奠定基础^[13]。

3 医学生数据素养评价体系设计

3.1 原始指标选取

调查对象为有科研数据需求的医学类博士研究生和硕士研究生。阅读相关文献并分析归纳现有成果构建指标，选取出现频次较高、专家认可度高、具有代表性的指标纳入评价体系中，最终构建一个包含 5 个 1 级指标、53 个 2 级指标的医学生数据素养能力评价指标体系。

3.2 评价体系分析

数据分析利用普通最小二乘法（Ordinary Least Squares, OLS）多元回归模型，设定若干控制变量，从 3 个层面递进展开，见图 1。第 1 层为客观现状与主观认识；第 2 层进一步分析 5 项 1 级客观现状指标以及相对应的自我认识指标，自我认识指标中采用李克特量表，分为 5 个等级：非常不符合、比较不符合、一般、比较符合、非常符合；第 3 层探讨 5 项 1 级客观指标与主观认识中各具体指标间的验证关系，确保该验证体系的客观性与真实性。

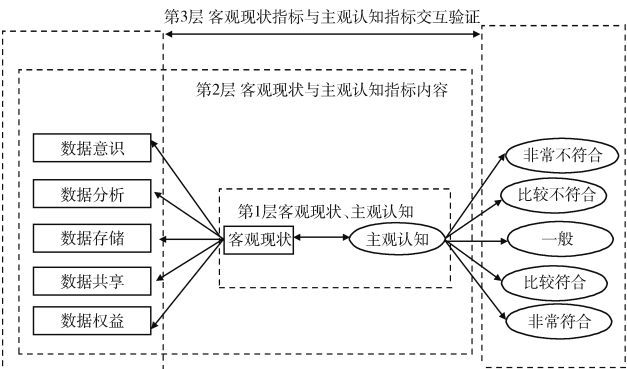


图 1 研究框架

4 研究对象与方法

4.1 研究对象

选取山西医科大学、山西中医药大学、长治医学院的在校研究生，采取方便抽样方法，共发放问卷 650 份，回收有效问卷 616 份，有效回收率 94.77%。

4.2 研究方法

问卷内容包括两个部分：第 1 部分是一般人口学资料，包括性别、学历等。第 2 部分是数据素养问卷，包括数据意识（6 题）、数据收集（4 题）、数据使用（4 题）、数据存储（6 题）、数据群益（5 题）5 个维度，满分 100 分，问卷 Cronbach's α 系数为 0.920，具有良好的信效度。

4.3 结果与分析

4.3.1 调查对象基本情况（表 3）

表 3 调查对象基本情况

项目	类别	人数	百分比（%）
专业	临床医学	212	34.4
	基础医学	40	6.50
	护理	33	5.36
	药学	26	4.22
	公共卫生	54	25.47
	中医	155	25.16
	影像	24	3.90
	其他	72	11.69
性别	男	116	18.83
	女	500	81.17
学历	硕士	533	86.53
	博士	83	13.47

4.3.2 单因素分析 调查结果显示不同学科之间数据素养总分在 55 ~ 68 之间，处于中等水平，但各学科之间水平参差不齐。其中综合得分最高的学科分别是基础医学（68.90 ± 17.81 分）和公共卫生

(67.63 ± 17.99 分)，最低的学科是影像（55.00 ± 19.92 分），见表 4，具有统计学意义（ $P < 0.05$ ）。

表 4 山西省医学研究生不同学科数据素养的单因素分析

因素（学科）	人数	综合（分）	数据意识（分）	数据收集（分）	数据使用（分）	数据存储（分）	数据权益（分）
临床医学	212	59.96 ± 18.61	16.28 ± 6.61	8.26 ± 4.55	11.75 ± 4.14	14.15 ± 6.30	9.51 ± 6.49
基础医学	40	68.90 ± 17.81 ^a	18.00 ± 6.01	8.30 ± 5.08	13.20 ± 3.97 ^a	16.70 ± 6.20 ^a	12.70 ± 6.00 ^a
护理	33	59.88 ± 13.84 ^b	15.27 ± 5.78	9.21 ± 4.53	12.97 ± 3.61	13.21 ± 5.24 ^b	9.21 ± 6.89 ^b
药学	26	62.92 ± 16.99	16.15 ± 5.71	7.23 ± 4.39	12.31 ± 3.74	15.85 ± 5.60	11.38 ± 6.47
公共卫生	54	67.63 ± 17.99 ^{ac}	16.44 ± 6.67	9.41 ± 5.39	13.04 ± 3.74 ^a	15.93 ± 5.63 ^{ac}	12.81 ± 5.89 ^{ac}
中医	155	65.88 ± 18.04 ^a	17.37 ± 6.71	9.24 ± 4.57 ^d	12.75 ± 3.78 ^a	15.12 ± 5.74	11.41 ± 6.62 ^a
影像	24	55.00 ± 19.92 ^{bdf}	14.00 ± 6.13 ^f	8.17 ± 4.79	11.17 ± 5.00 ^b	13.67 ± 6.23 ^b	8.00 ± 6.24 ^{bc}
其他	72	63.50 ± 15.25 ^g	16.11 ± 5.81	9.78 ± 4.97 ^{ad}	11.72 ± 3.99	14.22 ± 4.88 ^b	11.67 ± 5.77 ^{ag}
<i>F</i> 值	—	3.475	1.492	1.722	2.207	2.005	3.889
<i>P</i> 值	—	0.001	0.167	0.101	0.050	0.052	0.000

注：与临床医学相比，a 表示 $P < 0.05$ ；与基础医学相比，b 表示 $P < 0.05$ ；与护理相比，c 表示 $P < 0.05$ ；与药学相比，d 表示 $P < 0.05$ ；与公共卫生相比，e 表示 $P < 0.05$ ；与中医相比，f 表示 $P < 0.05$ ；与影像相比，g 表示 $P < 0.05$ 。

5 讨论与建议

5.1 体系构建有利于提高数据素养水平

通过评价体系结果可知医学研究生数据素养水平处于中等，且学科之间差异显著，尤其是在数据使用和数据权益方面。该体系构建能够较准确地了解该地区医学生数据素养能力，发现医学生在数据素养方面的优势和不足，从而促使医学生对数据素养多重内涵进行思考，激发主动提高数据素养的动力。

5.2 数据权益意识缺失

数据是一种财物，如果没有正确的使用与保护意识，轻则导致数据丢失，重则出现学术不端、财产受到侵害等问题^[14]。调查结果显示医学研究生数据权益意识不强，大部分专业的医学研究生不能清楚了解数据权益具体内容，更无法有效保护个人数据安全。如今每个高校针对研究生会开展防止学术不端系列讲座，本评价体系中针对数据相关权益设置了问题，能够了解医学生数据道德水平，同时高校可根据评价系统反馈结果适当进行课程改革，使

医学生在数据使用的开始阶段就养成正确的数据道德观，树立版权意识，培养数据管理、共享的大数据思维，提升数据素养能力。

5.3 构建信息化校园

医学类高校作为国家高水平医学科研人员的基地，应以处理分析相关研究课题为导向，围绕课题研究，培养医学生在实际科研过程中分析、处理、存储数据的能力以及数据共享意识与创新能力。通过评价体系反馈结果可知医学研究生在数据使用方面仍然存在较多问题，对此应加强学科数据化教学建设，搭建大数据交流与共享平台，减少实验数据重复收集成本，努力为学生交流、探索提供良好的环境。

6 结语

目前我国在数据素养评价体系建设方面仍处于摸索阶段，对数据素养标准讨论未形成统一定论，特别是在医学领域数据素养研究发展任重道远。数据素养应在不断学习中得以提升。本评价体系的构建在一定程度上可以帮助了解医学生数据素养情

况,通过测评结果对教学模式进行针对性改进,提升医学生数据素养能力水平。

参考文献

- 1 焦海霞. 由信息素养馆员向数据素养馆员转型: 动因、模式与路径 [J]. 图书馆学研究, 2018, 39 (23): 30-36.
- 2 黄如花, 李白杨. 数据素养教育: 大数据时代信息素养教育的拓展 [J]. 图书情报知识, 2016, 32 (1): 21-29.
- 3 Schield M. Information Literacy, Statistical Literacy and Data Literacy [J]. Iassist Quarterly, 2004 (2-3): 7-14.
- 4 Koltay T. Big Data, Big Literacies? [EB/OL]. [2019-02-15]. https://www.researchgate.net/publication/265334966_Big_data_big_literacies.
- 5 Calzada J, Ángel M. Incorporating Data Literacy into Information Literacy Programs: core competencies and contents [J]. International Journal of Libraries and Information Services, 2013, 63 (2): 123-134.
- 6 张静波. 大数据时代的数据素养教育 [J]. 科学,

2013, 65 (4): 29-32, 4.

- 7 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.
- 8 杨晓琼. 大数据时代高校数据素养教育的合作路径 [J]. 情报资料工作, 2015, 35 (3): 98-102.
- 9 马晓慧. 大数据背景下大学生数据素养教育探讨 [J]. 软件工程, 2019, 22 (2): 48-50, 41.
- 10 张晓娟. 信息素养: 标准、模式及其实现 [J]. 图书情报知识, 2009, 25 (1): 17-23, 29.
- 11 欧阳玲琳. 大数据背景下的医科院校数据素养教育 [J]. 中华医学图书情报杂志, 2017, 26 (6): 59-62.
- 12 吴曼翔. 医学院校图书馆信息素养教育模式研究 [J]. 湖北中医药大学学报, 2012, 14 (1): 78-79.
- 13 张意柳. 大学生信息素养教育的缺失与对策 [J]. 现代情报, 2006, 26 (5): 197-199.
- 14 马腾, 孙玲. 信息生态视域下高校大学生数据素养评价研究 [J]. 情报科学, 2019, 37 (8): 120-126.

(上接第 55 页)

量。因此构建护理质量控制数据库对于持续提升护理质量、提高各医疗机构护理管理水平、实现新型医疗的社会服务功能具有重要意义。采用信息化技术将所有护理数据进行收集、整合,实现数据信息资源共享,为相关研究开展提供数据支持。本文分析医院护理信息化建设现状及面临的问题,提出相关建议,以期推动我国医院护理信息化建设发展。

参考文献

- 1 汪牡丹, 成守珍, 李桂梅. 护理质量评价指标的研究进展 [J]. 中国护理管理, 2012, 12 (9): 40-43.
- 2 吴娅利, 蔡耘菲, 张恩全, 等. 护理信息化质量监控系统的临床应用及效果评价 [J]. 世界最新医学信息文

摘, 2018, 18 (88): 263-267.

- 3 薛晓英, 张利, 任成云. 平时移动护士工作站研发与应用 [J]. 海南医学院学报, 2012, 18 (10): 1500-1503.
- 4 秦林, 刘继终. 护理质量指标及数据库信息化建设的研究进展 [J]. 当代护士 (下旬刊), 2018, 25 (11): 23-25.
- 5 俞梦盈, 刘晓, 裴彩利, 等. 国内外护理安全信息系统应用研究进展 [J]. 护理研究, 2018, 32 (7): 997-1000.
- 6 韩斌如, 尉俊铮, 王燕秋. 信息化闭环管理助力护理安全 [J]. 中国数字医学, 2015, 10 (4): 87-90.
- 7 丁秀梅. 精准护理在外科急腹症患者护理中的应用效果观察 [J]. 社区医学杂志, 2017, 15 (7): 46-48.
- 8 赵瑜, 万长秀, 周琼. 大数据时代医院护理管理创新发展的思考 [J]. 全科护理, 2017, 15 (30): 3798-3800.