

科学数据共享研究的多视角分析与整合框架构建

支凤稳^{1, 2, 3} 张 萌² 郑彦宁¹

(1. 中国科学技术信息研究所, 北京 100038; 2. 河北大学管理学院, 河北保定 071002;
3. 河北省数字治理与协同治理研究基地, 河北保定 071002)

摘要: 科学数据是国家科技创新、经济社会发展和国家安全的基础性战略资源, 而共享是其价值充分发挥的重要途径。在系统文献调研的基础上, 首先筛选出152篇密切相关的实证文献, 从社会学、心理学、管理学3个视角回顾其研究重点, 进而分析比较各视角下相关研究的优势和局限。然后, 从理论基础、研究方法、研究对象、研究内容4个方面构建整合理论框架, 对科学数据共享的动因、方式与价值展开进一步讨论。最后, 归纳现有研究的不足, 指出未来可能的研究方向。

关键词: 科学数据共享; 社会学视角; 心理学视角; 管理学视角; 整合框架

DOI: 10.3772/j.issn.1674-1544.2023.02.001

CSTR: 15994.14.issn.1674-1544.2023.02.001

中图分类号: G350

文献标识码: A

Multi-Perspective Analysis and Integrated Framework Construction of Scientific Data Sharing Research

ZHI Fengwen^{1, 2, 3}, ZHANG Meng², ZHENG Yanning¹

(1. Institute of Scientific and Technical Information of China, Beijing 100038; 2. School of Management, Hebei University, Baoding 071002; 3. Collaborative Digital Governance Research Base of Hebei, Baoding 071002)

Abstract: Scientific data is a basic strategic resource for national scientific and technical innovation, economic and social development and national security, and sharing is an important way to give full play to its value. On the basis of systematic literature research, 152 closely related empirical literatures were selected in this study firstly, and their research focuses were reviewed from the perspectives of sociology, psychology and management, and then the advantages and limitations of relevant researches from each perspective were analyzed and compared. Then, it constructs an integrated theoretical framework from four aspects: theoretical basis, research method, research object and research content, and further discusses the motivation, way and value of scientific data sharing. Finally, the shortcomings of the existing research are summarized and the possible research directions in the future are pointed out.

Keywords: scientific data sharing, sociological perspective, psychological perspective, management perspective, integrated framework

作者简介: 支凤稳 (1987—), 女, 中国科学技术信息研究所博士后, 河北大学管理学院副教授, 研究方向为科学数据管理与共享、竞争情报; 张萌 (1996—), 女, 河北大学管理学院硕士生, 研究方向为科学数据管理与共享 (通信作者); 郑彦宁 (1965—), 男, 中国科学技术信息研究所研究馆员, 研究方向为竞争情报、情报学理论方法。

基金项目: 中国博士后基金面上资助项目“科学数据共享模式及驱动机制研究”(2018M641446); 河北省高等学校人文社会科学基金项目“元宇宙时代科学数据共享模式及应用研究”(BJS2022027)。

收稿日期: 2022年9月12日。

0 引言

科学数据主要是指在自然科学、工程技术科学等领域,通过基础研究、应用研究、试验开发等产生的数据,以及通过观测监测、考察调查、检验检测等方式取得并用于科学研究活动的原始数据及其衍生数据^[1]。科学数据是一种具有潜在价值的战略资源,各国通过科学数据的开放共享提升科研创新水平,从而保持国家竞争力。科学数据共享本质上是不同角色围绕科学数据资源,从满足数据用户需求出发,利用各种方式实现科学数据从数据提供者到数据用户对于其拥有和使用权利的让渡,达到科学数据资源满足科研活动需求的状态^[2]。随着科学研究的深入发展,大量的科学数据不断产出,这既是开展科学研究的基础,也是重要的科研产物。2021年6月,Science Europe^[3]指出共享可重复利用的科学数据以支持科学进步已成为常态,越来越多的组织机构期望并鼓励科学数据开放共享的行为。开放共享实现科学数据重用,可以促进科研人员利用现有研究数据提出新的观点和问题,最大限度地提高研究透明度^[4],并且有助于降低科研成本,使公共资助研究的结果为公众所用。不同科学领域在数据共享和重用文化上存在差异,社会科学数据不同于来自实验室或其他生命、物理科学研究的数据^[5],其共享实践尚未广泛开展。因此,通过数据的开放共享实现各领域数据交叉利用仍需各方的不断努力。

我国高度重视科学数据共享。从21世纪初开始全面加强科学数据开放共享服务工作,在较短时间内迅速发展,缩小了与欧美国家的差距^[6]。近年来,学者们也意识到了科学数据共享的重要意义,并基于不同理论对其模式、行为、驱动机制等进行了系列探索。如盛小平等^[7]从科研人员自身、技术、平台、法律、政策、数据、资金、组织等8个方面总结了科学数据共享的影响因素,并提出了当前研究对某些现象或问题的争议点;万莉等^[8]侧重于科学数据共享行为前因变量作用机制的差异评价,挖掘出差异产生的潜在原

因,并在此基础上提炼出个人、资源以及制度环境3个维度的潜在因素;王丹丹等^[9]针对当前不同研究结果之间的差异,采用元分析方法对科学数据共享行为影响因素进行系统检验。文献调查发现,相关研究成果在不断积累,尤其是针对科学数据共享影响因素的实证研究居多,且理论模型丰富,也不乏相关的研究综述。但大多是从内容角度对影响因素进行归类,分析产生差异的原因,却鲜有学者以理论视角对相关实证文献进行分类与整合,不同学科视角下的理论框架尚不清晰。鉴于此,本文从理论视角出发,对现有实证文献进行系统梳理,分析其在社会学、心理学、管理学视角下的优势和局限性,构建理论整合框架,并指出未来可能的研究方向,以期后续学者提供一定的理论参考。

1 不同视角下科学数据共享研究回顾

本文选择CNKI、坤舆发现、NSSD、Web of Science、Emerald、Elsevier、Wiley为主要检索数据库,中文以“科学数据”“科研数据”“研究数据”分别与“开放共享”组配为主题词,外文以“scientific data”“Scientific research data”“research data”“Scientists’ data”分别与“sharing”“open”组配为主题词,检索时间为2022年7月1日,最终得到中文实证研究文献63篇,外文实证研究文献89篇。逐一研读152篇实证文献,并从社会学、心理学、管理学视角对各变量及研究结论展开论述。由于理论具有一定的普适性,个别理论在应用中存在交叉现象,如社会交换理论可被称为行为主义社会心理学理论。根据其在科学数据共享研究中的具体应用,本文则将其归为社会学视角。

1.1 社会学视角

社会学视角下的研究,主要从外部环境和人际关系交互两个方面展开,并将科学数据共享视为社会化交互的过程,应用的理论主要有社会交换理论、制度理论、仿真实论。

1.1.1 基于社会交换理论的研究

社会交换理论源于经济学交换理论在社会

学中的应用和发展，布劳将其界定为“当别人做出报答性反应就发生，当别人不再做出报答性反应就停止行动”^[10]。基于社会交换理论的科学数据共享研究，将共享行为视为人与人之间资源交换的社会活动，通过实用价值衡量报酬和代价对行为产生的影响。报酬不仅包括金钱、信息等直观的回报，还包括社会地位、服务、情感价值等抽象的回报，而代价往往是社会活动中所产生的消极结果，在科学数据共享中可能要付出数据泄露、观点盗取等代价。实用价值强调个人的行为受自我价值利益的引导^[11]，而科研人员在进行科学数据共享活动前考虑到报酬大于代价，故实用价值驱使科研人员科学数据共享。

相关研究多关注于感知成本和收益，如李逸飞^[12]在考察绩效期望对高校科学数据共享意愿的影响时，将自我效能和预期回报作为具体的表现形式，且把预期回报作为个体特征变量的测量项^[13]。互惠作为社会资本创造和再生的基本驱动力^[14]，其与预期回报具有相似性，是指在付出后期望有相应回报的心态，不同之处在于互惠的规则是相互受惠，在科学数据共享中主要表现为数据提供者共享自身的科学数据所获得的回报与付出的价值相匹配。邹威威^[15]研究发现互惠对科学数据共享行为意愿的影响并不显著，而Kim^[16]的研究显示学术互惠对韩国生物技术研究人员的共享意愿有显著影响，并且学术声誉对其有显著的调节作用。不同国家学术奖励系统的发展和完善程度可能存在较大差异，而互惠作为影响科学数据共享意愿的争议因素，需要通过学术奖励系统呈现具体的形式。

1.1.2 基于制度理论的研究

Bell^[17]将制度定义为一组规则、守则或塑造行为的默契。制度理论的核心是为实现自身稳定而对组织施加压力，以期优化组织的结构和流程。制度理论将科研人员的科学数据共享行为根植于大的体制背景下，依靠结构化的社会期望和行为规范培养科研人员的共享行为，在奖励机制和惩罚机制的双重作用下塑造履行规则的社会行动者。制度理论下的科学数据共享受到体制力量

的影响，不同体制环境下的社区文化是学者们关注的一个情境因素^[11]。面对不同制度和规则，科学数据共享情境会随之变化，如科研人员受到资助机构和刊物出版商的调控压力，以及行业规范性压力的影响^[18]。

不同科学领域之间的科学数据共享较难实现。由于科学数据类型和大小的多样性，以及学科文化的差异性，科研人员对科学数据共享的看法和实践可能会随着研究环境的不同而有所差异，如天文学等大型科学领域将数据共享作为标准实践，而社会科学领域的科研人员却没有广泛参与共享活动^[4]。Kim等^[19]针对生物科学领域科研人员的数据共享行为，探讨了制度环境对科学数据共享的影响，包括资助机构和期刊出版商的监管压力，以及机构资源等因素。孙晓燕^[18]研究发现科研人员的共享行为显著受到行业规范压力的影响，但资助机构压力和刊物出版商压力对共享行为的影响并不明显。目前，相关资助机构和刊物出版商对数据共享提出了一系列强制要求，如《数据分析与知识发现》编辑部要求投稿论文必须提交支撑研究结论的科学数据，并且对所录用的论文支撑数据进行公共保存^[20]。因此，新的制度环境下科研人员受到的影响也会发生转变，有必要进一步完善制度理论下科学数据共享的研究。

1.1.3 基于仿真实验的研究

博弈论将科研人员的科学数据共享行为分为共享科学数据和不共享科学数据两种，并将这两种行为放入演化博弈的理论框架中，进行双方博弈，进而分析科研人员的共享行为和共享策略。庄倩等^[21]构建了科研人员之间进行科学数据共享的博弈模型，将信任作为影响科学数据共享的关键因素，发现在信任程度高的情况下双方共享数据可以达到收益最大化，反之共享一方要比不共享一方付出更大的代价，这样不利于信任互惠的形成，极大地降低了科学数据共享的作用。李思宇^[22]从成本和收益的角度，用进化博弈模型研究科学数据共享的充分条件，对数据共享提出了“成本侧”和“收益侧”的促进策略。

系统动力学理论与博弈理论均是对科学数据共享的过程进行仿真模拟,通过动态的方式展现科学数据共享受相关因素影响的程度,而系统动力学理论更加强调从系统的角度研究影响因素的作用机理。是沁等^[23]对科学数据共享保障机制展开研究,模型仿真发现,政策法规建设、组织服务、技术支持、资金支持等均对科学数据共享有正向影响。目前,科学数据共享处于发展阶段,建立高度的信任关系需要科研人员参与共享活动,合理地使用科学数据,形成信任互惠的良性互动,从而衍生出满足更高层次需求的新数据。

1.2 心理学视角

心理学视角下的相关研究重点关注科研人员的心理状态,以及个人动机对科学数据共享意愿与行为的影响,应用的理论主要有社会认知理论、计划行为理论和双路径理论。

1.2.1 基于社会认知理论的研究

心理学家最早提出的社会学习理论用于解释人类的间接学习行为。随后,为了解释具有复杂动机的人类活动,在此基础上引入了自我效能而发展为社会认知理论^[24]。这个理论阐释了人类行为、个体认知和外部环境之间的因果关系,并且认为个体和环境共同决定着人类的意愿与行为。在科学数据共享过程中,无论是科学数据提供者还是科学数据使用者,都会受到认知因素的影响。

在相关研究中,通常将组织氛围作为外部环境因素,积极情绪和自我效能作为个体认知因素^[25]。个体认知因素也可从感知职业收益、感知风险、感知努力等方面进行微观考察,如Kim等^[26]研究科学、技术、工程和数学领域科研人员的科学数据共享行为,发现个体认知因素对相关领域科研人员的数据共享态度具有显著影响。此外,个体的认知能力和调节能力可以概括为自我效能,具体表现为结果预期和效果预期。不同的利益相关者,在结果预期方面存在较大差别,如科学数据提供者更多地关注于奖励与荣誉,同时考虑数据泄露、观点盗取等问题,而科学数据使

用者更关注于及时获取数据、方便学术交流等。因此,针对不同的利益相关者,有必要分别探讨自我效能对共享意愿的影响。

1.2.2 基于计划行为理论的研究

计划行为理论源于Ajzen等^[27]对理性行为理论的发展,它是预测和解释人的行为改变研究中的一个较为成熟的理论,认为人的行为是经过深度思考作出的选择,核心要素是意愿对行为的影响。然而,意愿又受到感知行为控制、主观规范以及态度的影响,认为进行某项活动的难易程度属于感知行为控制,其与社会认知理论中的效果预期具有相似之处,属于认知驱动的主要因素。进行某项活动时个人感受到的外部压力属于主观范式,属于社会价值驱动的主要因素,而态度主要是指对进行某项活动正面或者负面评价的程度。

诸多学者在对科学数据共享的研究中引入了计划行为理论,探讨个人层面的影响因素。早期研究发现,社会科学领域科研人员的主观规范对共享行为和共享意愿有显著正向影响^[28-29],而Kim等^[30]在后期的研究中打破了这一假设。因此,主观规范这一因素对科学数据共享的影响具有争议性。感知行为控制对科学数据共享的影响也未形成统一观点^[31-32],其对意愿的影响会根据不同的研究环境发生改变^[33],主观规范和感知努力也会受到数据重用经验的影响^[34]。早期对科学数据共享的研究无论是制度环境还是主观规范方面,与当下研究环境存在较大差异。随着科学数据共享强制策略的提出,以及人们数据素养和共享意识的提高,科研人员的感知行为控制可能会发生改变。

1.2.3 基于双路径理论的研究

双路径模型也称精细加工可能性模型(Elaboration Likelihood Model, ELM),最早由Petty等^[35]于1986年提出。个体在接触某信息时,如果需要对其进行精密的认知思考和理性研究,从而形成对目标行为的认知判断,属于中枢路径的范畴;如果只通过对信息外在表象特征的简单分析和推理来判断目标行为,而不涉及深度

的认知思考,则属于边缘路径的范畴。人们进行科学数据共享活动时,对其进行精密的认知思考和理性研究,或者对外在表象特征进行简单分析和推断,最终可通过中枢路径和边缘路径计算发生共享的可能性。

目前,双路径模型在知识共享领域的实证研究中应用较多,而在科学数据共享领域应用较少,但这个模型在判断科研人员科学数据共享意愿和行为方面具有较好的适用性。支凤稳等^[36]从科学数据使用者的角度调查共享意愿和共享行为,将科学数据质量和来源可信度作为关键变量,研究结果显示来源可信度间接正向影响科学数据共享意愿和共享行为,而科学数据质量对共享意愿的影响并不显著。事实上,科研人员在使用科学数据的过程中需要对数据质量进行甄别,但面对不同类型的研究数据,甄别过程可能比较复杂,需要科研人员在具备良好的数据素养及专业知识能力的条件下,通过合理使用科学数据发现其质量的参差情况。根据双路径模型,将科学数据质量作为中枢路径,甄别数据的真实性、完整性、有效性、权威性、价值性等特点,这一过程更多地是在获取科学数据后的呈现。

1.3 管理学视角

管理学视角下的相关研究重点关注科研人员在数据共享过程中对相关技术的接受程度,以及不同参与主体的利益诉求和行为逻辑,并基于技术接受模型、利益相关者理论进行实证分析。

1.3.1 基于技术接受模型的研究

技术接受模型由Davis^[37]于1989年提出,认为人们接受信息系统的行为是由态度和感知有用性决定的,而态度受到两大重要因素(感知有用性和感知易用性)的影响,其在有关新事物接受的预测中得到了广泛应用。

诸多学者在对科学数据共享的研究中引入了技术接受模型,而感知有用性作为其中的一个关键因素,其对科学数据共享的影响作用存在争议^[38-39]。王春晓等^[40]构建了科学数据共享接受模型,发现感知有用性对共享意愿的影响不显著,仅对共享态度有正向影响,并且更多地被调查者

认为共享科学数据对提高自我学术地位具有很大帮助。与前期何琳等^[39]的研究相比,被调查者不仅感知到获取科学数据所带来的回报,而且感知到提供科学数据所带来的价值。由此可见,随着社会环境以及政策制度的发展,人们对于科学数据共享有用性的感知得到了明显的提高。

1.3.2 基于利益相关者理论的研究

利益相关者理论的核心思想是组织中的利益相关者直接影响组织目标的实现,应充分考虑和满足不同层次利益相关者的诉求^[41]。这个理论可以用来明确科学数据共享中不同利益主体的身份特征,如科学数据提供者、使用者、管理者等,进而根据主体间的差异性和互补性,解释各主体参与共享的驱动因素。

相关学者针对不同利益主体在科学数据共享中的不同表现进行了探讨。如盛小平等^[42]从不同利益相关者角度分析驱动因素,划分了科学数据生产者、利用者、管理者、资助者、组织者、传播者、发布者7类利益相关主体。也有学者对多元主体驱动机制展开研究,基于各主体定位和诉求总结了利益共赢拉动力、制度环境引导力、社会环境推动力、组织环境支持力、技术环境保障力5个方面的驱动因素^[43]。李宜展等^[44]发现科学数据提供者在数据共享过程中具有自愿参与权,同时遵守知情同意的伦理道德准则。此外,需要保障数据安全和数据隐私。由此可见,利益相关者理论视角的研究体现了利益主体的身份特征不同,其责任、诉求与动机也有所差异。

2 不同视角下研究的比较分析

不同视角下实证研究均对科学数据共享意愿或行为进行探索,但各视角下的理论依据与具体因素不尽相同,且各个理论既有优势也有自身的局限性。具体如表1所示。

诸多学者从社会学、心理学等多个视角,基于制度理论、计划行为理论等多个理论构建模型,用于解释和预测社会科学领域的科学数据共享行为^[18]。因此,相关研究中存在不同视角共同发挥作用的情况。然而,也有学者基于某一视角

表 1 研究视角比较分析

视角	理论基础	具体因素	优势	局限
社会学视角	社会交换理论 制度理论 仿真理论	绩效期望、精神奖励、互惠、资助 机构压力、刊物出版商压力、行业 规范压力、信任、技术支持等	从团队或组织的角度考虑 社会因素，关注人与人之 间构成的社会网络关系以 及相互作用	忽略了个人心理特征变化 的动态性以及技术因素
心理学视角	社会认知理论 计划行为理论 双路径理论	组织氛围、自我效能、结果预期、 效果预期、主观规范、感知行为控 制、数据质量、数据来源可信度等	关注科研人员心理因素的 动态变化在科学数据共享 中的主导作用	忽略了社会环境、利益、 技术平台等客观因素
管理学视角	技术接受模型 利益相关者理论	感知有用性、感知易用性、系统特 征、数据安全、数据隐私等	从信息系统和组织的角度 对科学数据共享进行客观 性分析	研究角度较为单一，因素 较少，研究结论容易受到 主观感知的影响

下的单个或多个理论构建理论模型，如基于理性行为理论揭示科研人员获取科学数据的信念、主观规范以及态度对共享意愿的影响作用^[29]。基于此，本文将以社会学、心理学、管理学 3 个研究视角进行比较分析，为未来根据不同需求研究视角融合提供借鉴。

2.1 基于社会学视角

社会学视角下的研究体现了科研人员的社会互动和利益关系，如绩效期望表现为科研人员认为自身参与科学数据共享可获得有利结果。无论是经济效益，还是信任关系的建立，都是科研人员考虑社会因素选择参与共享活动得到的益处。用社会学视角研究的优势在于科研人员或科研团队可以从社会整体角度，充分考虑人际关系和社会网络的影响，其局限性在于容易忽略科研人员个人心理的动态变化以及系统或平台方面的客观问题。

2.2 基于心理学视角

心理学视角下的研究关注了科研人员的情感和能力认知，主要从主观心理感受出发判断是否参与科学数据共享活动，如自我效能表现为科研人员对自身实际完成科学数据共享行为能力和信心的主观估计情况。用心理学视角研究的优势在于充分考虑科研人员自身情感因素，更加深入地探寻了不同身份特征下心理动态对共享意愿和行为产生的影响，其局限性在于研究的影响因素较为主观，且不同个体对自身情况的估计可能会出现高估或低估的情况，难以发挥社会环境、政策文化以及技术平台等方面的作用。

2.3 基于管理学视角

管理学视角下的研究体现了科研人员参与科学数据共享活动的难易程度，以及不同利益主体的责任与诉求，主要从系统或组织的角度分析利益主体的共享态度与意愿之间的关系。用管理学视角研究的优势在于分析了系统平台等方面的客观因素的作用，强调不同利益主体的意愿与行为之间的差异，但研究结论过于依赖个体同利益主体的主观感知。

综上所述，各视角下的理论虽有一定的适用性，但也有一些自身的局限性。科学数据共享是一个较为复杂的过程，仅基于某一视角下的理论进行研究，容易忽略实践中面临的一些问题。如以仿真理论为基础的研究主要通过构建两方博弈模型以及系统动力学模型，对科学数据共享进行动态展示。但由于建模过程中对模型各方面条件的限制，多采用假设条件达到技术要求，可能忽略了实际共享过程中的一些突发情况，不能从实际样本中反映科研人员的真实状况。此外，科学数据共享既存在个人行为，也存在团体行为，而团队成员间心理状态的变化因素也不能通过仿真实验很好地模拟出来。因此，特定场景下多理论的融合使用有助于取长补短，更好地解释与预测科学数据共享行为。

3 整合框架构建

基于前文研究视角的总结与大量相关文献的内容分析，本文从理论基础、研究方法、研究对象、研究内容 4 个方面构建科学数据共享研究的

整合框架。具体如图 1 所示。

在图 1 中，科学数据共享研究的理论基础包括社会学、心理学和管理学 3 个视角下的 8 种理论。研究方法可以简单地划分为数据收集方法和数据分析方法。前者主要有文献调研、问卷调查、访谈等，后者主要有描述性分析、相关分析、因子分析、多元回归分析、结构方程建模、文本分析等。从研究对象来看，数据提供者、使用者、管理者等不同主体的行为特征被很多学者所关注，这也许因为共享参与主体是多元与动态的，其动机是个体理性而非集体理性的。从研究内容来看，学者们遵循“起因—经过—结果”的基本逻辑思路，研究重点集中于科学数据共享的动因、实现方式与各主体的共享结果。

3.1 科学数据共享的动因

基于相关理论的内涵与特点，以及其在揭示科学数据共享意愿或行为影响因素中的具体应用，本文把科学数据共享的动因归纳为价值、认知和情境 3 个维度。

从哲学角度来讲，价值属于关系范畴，泛指客体对主体反映出来的积极意义和有用性。在科学数据共享中，感知价值是科研人员对这一活动的主观评价，具体可分为共享时所体验到的价值和共享结果所带来的价值。社会交换理论主要通过绩效期望、预期回报等来反映科研人员对活动的评价。具体来讲，科研人员的学术声誉、共享过程中形成的人际信任以及共享结果所带来的互惠效益，均可反映价值因素的作用。

自我效能是社会认知理论的核心要素，它反映了科研人员能否成功进行共享的主观判断，与计划行为理论中的感知行为控制具有相似性。科研人员无论是对自身能力、情绪的认知，还是对外部环境的认知，均对其共享行为有一定的影响。当然，诸多因素的划分也存在交叉现象，如自我效能在社会交换理论和社会认知理论中均有体现，而在社会认知理论中，具体表现为结果预期和效果预期，在相关研究中常用于考察认知因素的作用。

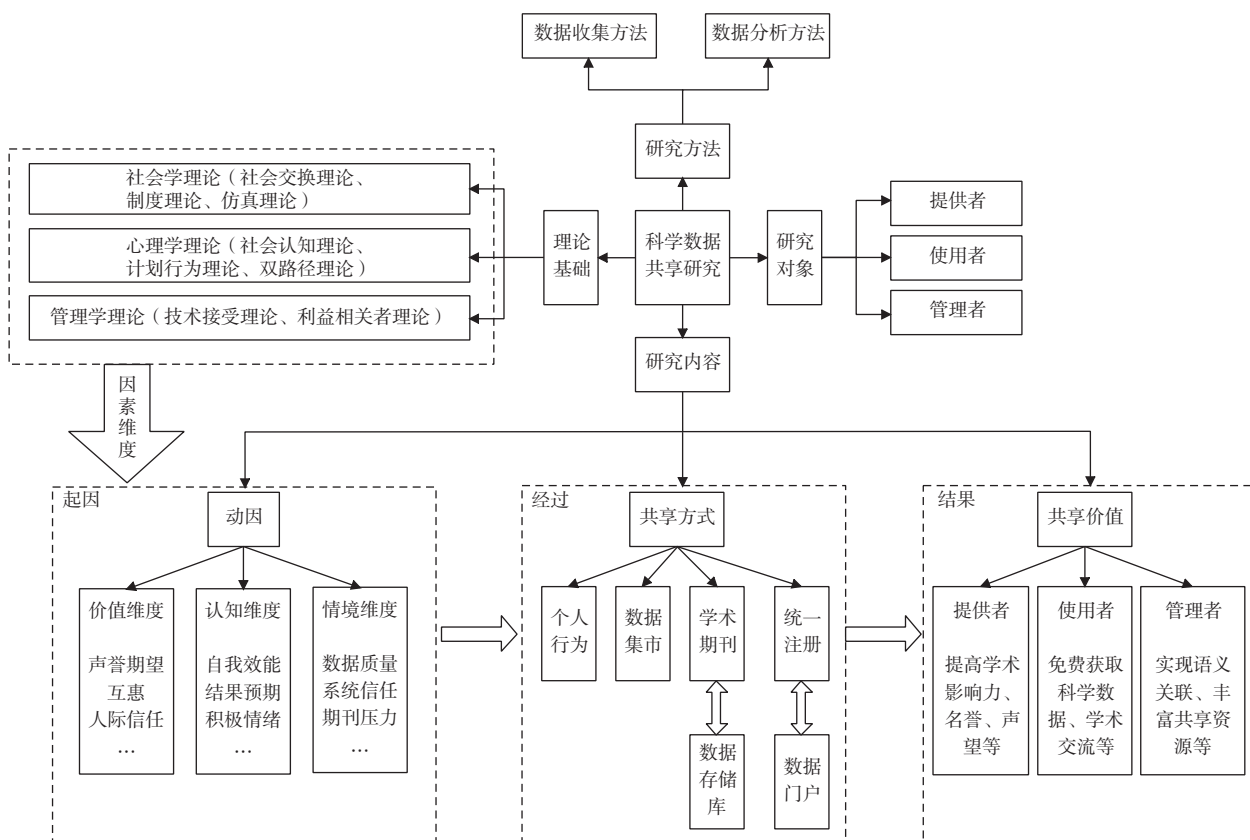


图 1 科学数据共享研究的整合框架

除了价值、认知维度的因素,情境因素对科学数据共享也有一定的影响。制度理论通过资助机构压力、刊物出版商压力、行业规范压力等来反映科研人员遇到的管理情境。在技术接受模型中,感知有用性和感知易用性是用于衡量科研人员感知技术平台的关键变量,对科学数据平台和系统的接受程度,体现了科研人员对数据来源的信任程度。此外,双路径模型中的数据质量和来源可信度也属于情境因素。

3.2 科学数据共享的方式

相关研究表明,科学数据可以在科研人员间直接进行共享,也可以通过第三方机构或平台进行共享。主要呈现出以下4种共享方式:一是基于信任关系的直接共享。科研团队成员间进行长期的科研合作,建立良好的信任关系,科学数据可作为团队科研成果在成员间进行共享。二是基于数据集市的间接共享。科学数据提供者和使用通过数据集市完成科学数据的发布和获取,这一过程需要数据中心和数据存储库的支持。由于这一方式主要依靠经济效益驱动,数据存储库的经费投入难以持续^[45],数据共享的广度和深度受到限制。三是基于学术期刊和数据存储库的共享。近年来,学术期刊对论文中研究数据的共享提出了明确要求,科研人员需要通过期刊发表将科学数据存放到数据存储库。通过相关政策和文化的驱动,科学数据提供者根据个人身份特征和数据应用范围,可将科学数据存放到机构存储库、学科存储库和通用存储库^[46]。虽然数据存储库会对提交的科学数据进行必要的质量控制,但机构存储库和学科存储库面向的用户群体较为有限,而通用存储库的专业性和覆盖范围却具有局限性。四是基于数据门户的共享。科研人员可以按照相关标准和规范将科学数据进行集中注册,由第三方数据门户负责连接注册科学数据资源,为使用者提供一站式检索服务。在科学数据共享政策的驱动下,广域的合作协议和合作计划推动了分布式数据服务系统的建立和完善,为全球数据用户提供了免费的数据获取服务。

3.3 科学数据共享的价值

科学数据共享本身无法直接满足主体需求,需要主体通过数据共享建构和挖掘数据资源^[47],因此不同利益相关者进行科学数据共享的结果不同,可根据科学研究效率、绩效以及科学数据共享实现的效果来衡量共享结果。

在科学数据共享中,提供者通过提供自己的科学数据寻求身份认同,识别相关群体,并在互动过程中进行身份建构。当其受到期刊政策、同行规范或经济效益的驱动时,将自己的科学数据共享到数据存储库或直接提供给科研合作者。在保证科学数据质量的前提下,这不仅可以获得相应的经济效益,还可以提高自身的学术影响力以及在科研领域的名誉和声望等。使用者通过获取数据满足科研需求,以参与者间的数据流动为基础建立信任关系。当其受到自我效能、结果预期的驱动时,通过科研合作者直接获取或者从数据集市、数据存储库中有偿获取所需数据。这不仅可以建立科研合作者间的信任关系,还可以节省科研的时间和精力。科学数据共享也离不开科学数据管理者,如图书情报机构、数据中心、政府等^[42]。科学数据管理者汇集了大量的科学数据,可通过专业人员检测增加科学研究的透明度,并且根据科学数据的研究主题和应用领域实现语义关联,丰富共享资源。

4 研究结论与展望

本文深入分析了152篇科学数据共享的实证文献,从社会学、心理学、管理学3个视角回顾研究重点,对比分析各视角的优势和局限性,并从理论基础、研究方法、研究对象、研究内容4个方面构建了理论整合框架,进而对科学数据共享的动因、方式与价值展开进一步讨论,研究成果希望能够为后续相关学者提供一定的理论参考。

本文研究发现,科学数据共享的研究成果比较丰富,形成了较为完善的研究框架,但仍存在一些问题,需要在未来研究中持续关注与探索。第一,理论基础依然不够丰富,已有研究主要应

用社会学、心理学、管理学视角下的相关理论构建模型, 并且管理学视角下的理论应用较少。面对海量科学数据资源重复投资生产与闲置的现象, 应该优化资源配置, 而目前资源配置理论应用极少。未来研究可考虑融合多种理论构建模型, 不仅要关注社会学、心理学、管理学视角下的诸多理论基础, 还要探寻经济学、伦理学等视角下的相关理论, 丰富理论框架和研究内容。第二, 众多影响因素被发现, 但同一作用路径的结果不尽一致。针对不同科研领域、情境的科学数据共享实践研究有待于进一步加强, 以提升研究结果的针对性和现实意义。第三, 相关研究多采用定量的方法, 样本数据多来源于作者采集, 科学数据复用的情况较少, 且变量测量标准和样本选择的代表性难免有一定的主观性。未来研究可尝试利用大规模客观数据集进行跨学科领域的交叉探索, 以增强研究结论的可靠性。

参考文献

- [1] 国务院办公厅. 科学数据管理办法[EB/OL].[2022-01-01].<https://baike.so.com/doc/27159694-28552208.html>.
- [2] 屈宝强, 彭洁, 刘蔚, 等. 科学数据共享及其发展趋势[J]. 情报学进展, 2020, 13(1): 381-420.
- [3] Science Europe. Practical guide sustainable to research data[EB/OL].[2021-06-14].<http://www.science-europe.org/media/b3odxx3s/sc-practical-guide-sustainable-research-data.pdf>.
- [4] CURTY R G, CROWSTON K, SPECHT A, et al. Attitudes and norms affecting scientists' data reuse[J]. Plos one, 2017, 12(12): 1-22.
- [5] YOON A, KIM Y. Social scientists' data reuse behaviors: exploring the roles of attitudinal beliefs, attitudes, norms, and data repositories[J]. Library & information science research, 2017, 39(3): 224-233.
- [6] 张贵兰, 王健, 潘尧, 等. 科学数据共享服务模式及其演化研究[J]. 情报理论与实践, 2022, 45(2): 70-77.
- [7] 盛小平, 袁圆. 国内外科学数据开放共享影响因素研究综述[J]. 情报理论与实践, 2021, 44(8): 173-179.
- [8] 万莉, 程慧平. 基于元分析结构方程模型的科学数据共享行为因素研究[J]. 情报理论与实践, 2021, 44(7): 125-130, 158.
- [9] 王丹丹, 刘清华, 王晓梅. 科学数据共享行为影响因素的元分析[J]. 图书馆学研究, 2021(22): 74-84.
- [10] 钟苑婷. 布劳社会交换理论与科尔曼理性行动理论比较分析[J]. 法制与社会, 2017(35): 229-230.
- [11] 赵文军. 虚拟社区成员知识共享行为的驱动机制研究综述[J]. 现代情报, 2015, 35(11): 164-170.
- [12] 李逸飞. 高校科研人员科学数据共享影响因素研究[D]. 湘潭: 湘潭大学, 2021.
- [13] 刘嫣. 高校科学数据共享主要影响因素研究[D]. 南京: 南京航空航天大学, 2019.
- [14] OKOLI C, OH W. Investigating recognition-based performance in an open content community: a social capital perspective[J]. Information & management, 2007, 44(3): 240-252.
- [15] 邹威威. 高校师生科学数据共享行为意愿的影响因素研究[D]. 贵阳: 贵州财经大学, 2020.
- [16] KIM S. An exploratory study of biotechnology scientists' research data sharing intention: the moderating effects of academic reputation[J]. Journal of the Korean society for information management, 2022, 39(1): 45-68.
- [17] BELL S. Institutionalism[J]. Government, politics, power and policy in Australia, 2002(7): 363-380.
- [18] 孙晓燕. 科学数据共享行为的理论模型构建及测度实证研究[J]. 情报科学, 2016, 35(10): 1062-1071.
- [19] KIM Y, BURNS C S. Norms of data sharing in biological sciences: the roles of metadata, data repository, and journal and funding requirements[J]. Journal of information science, 2016, 42(2): 230-245.
- [20] 《数据分析与知识发现》编辑部. 论文支撑数据公共保存与共享暂行办法[J]. 数据分析与知识发现, 2022, 6(2): 4.
- [21] 庄倩, 何琳. 科学数据共享中科研人员共享行为的演化博弈分析[J]. 情报杂志, 2015, 34(8): 152-157, 168.
- [22] 李思宇. 科学数据共享问题: 一个博弈论视角的讨论[J]. 自然辩证法通讯, 2020, 42(12): 76-79.
- [23] 是沁, 储节旺. 基于系统动力学的科学数据开放共享保障机制研究[J]. 情报杂志, 2018, 37(11): 143-149.
- [24] BANDURA A. Human agency in social cognitive theory[J]. American psychologist, 1989, 44(9): 1175.
- [25] 余玲. 科研人员科学数据共享意愿的影响因素研究[D]. 衡阳: 南华大学, 2016.
- [26] KIM Y, ZHANG P. Understanding data sharing behaviors of STEM researchers: the roles of attitudes, norms,

(下转第33页)

- [23] 徐明波, 荀渊. 高校科技成果转化机构定位、职能及其影响因素研究: 基于中美研究型大学科技成果转化机构的对比分析[J]. 高教探索, 2021(11): 34.
- [24] 郁亚男. 基于 Android 平台的人机交互的研究与实现[D]. 北京: 北京邮电大学, 2011.
- [25] 孟展, 赵希男, 周岩. 研究型大学“双创”生态系统营造探讨: 以荷兰埃因霍温理工大学为例[J]. 中国高校科技, 2021(10): 64.

(上接第9页)

- and data repositories[J]. Library & information science research, 2015, 37(3): 189–200.
- [27] AJZEN I, FISHBEIN M. Attitude-behavior relations: a theoretical analysis and review of empirical research[J]. Psychological bulletin, 1977, 84(5): 888–918.
- [28] KIM Y, ADLER M. Social scientists' data sharing behaviors: investigating the roles of individual motivations, institutional pressures, and data repositories[J]. International journal of information management, 2015, 35(4): 408–418.
- [29] 张晋朝. 我国高校科研人员科学数据共享意愿研究[J]. 情报理论与实践, 2013, 36(10): 25–30.
- [30] KIM Y, STANTON J M. Institutional and individual factors affecting scientists' data-sharing behaviors: a multilevel analysis[J]. Journal of the association for information science and technology, 2016, 67(4): 776–799.
- [31] 邓灵斌, 肖洪伟. 我国高校科研人员科学数据共享意愿之实证研究[J]. 新世纪图书馆, 2021(5): 17–22.
- [32] 马玲. 高校科研人员科学数据共享机制研究[J]. 情报科学, 2021, 39(9): 80–83.
- [33] AJZEN I. The theory of planned behavior[J]. Organizational behavior and human decision processes, 1991, 50(2): 179–211.
- [34] YOON A, KIM Y. The role of data-reuse experience in biological scientists' data sharing: an empirical analysis[J]. The electronic library, 2020, 38(1): 186–208.
- [35] PETTY R E, CACIOPPO J T. The elaboration likelihood model of persuasion[C]//LEONARD B. Communication and persuasion advances in experimental social psychology. Salt Lake City: Academic Press, 1986: 123–205.
- [36] 支凤稳, 张萌, 赵梦凡, 等. 双路径视角下科学数据共享行为的影响因素研究[J]. 信息资源管理学报, 2021, 11(6): 40–50.
- [37] DAVIS F D. Perceived usefulness, perceived ease of use, and user acceptance of information technology[J]. MIS quarterly, 1989, 13(3): 319–341.
- [38] KIM Y, NAH S. Internet researchers' data sharing behaviors an integration of data reuse experience, attitudinal beliefs, social norms, and resource factors[J]. Online information review, 2018, 42(1): 124–142.
- [39] 何琳, 常颖聪. 科研人员数据共享意愿研究[J]. 图书与情报, 2014(5): 125–131.
- [40] 王春晓, 陈姝彤, 徐坤. 研究生科学数据共享态度与共享意愿关系研究[J]. 情报科学, 2020, 38(12): 78–84.
- [41] FREEMAN R E, EVAN W M. Corporate governance: a stakeholder interpretation[J]. Journal of behavioral economics, 1990, 19(4): 337–359.
- [42] 盛小平, 吴红. 科学数据开放共享活动中不同利益相关者动力分析[J]. 图书情报工作, 2019, 63(17): 40–50.
- [43] 胡佳琪, 陆颖. 开放科学数据利益主体协同机制研究[J]. 图书情报工作, 2020, 64(21): 26–33.
- [44] 李宜展, 刘细文, 李泽霞, 等. 科学数据安全边界概念模型研究: 基于利益相关者视角[J]. 中国科学基金, 2022, 36(2): 339–347.
- [45] 袁雅琴, 胡晓彦, 佟继周, 等. 大数据开放背景下的我国空间科学数据出版实践[J]. 中国科技资源导刊, 2022, 54(1): 89–96.
- [46] 李成赞, 张丽丽, 侯艳飞, 等. 科学大数据开放共享: 模式与机制[J]. 情报理论与实践, 2017, 40(11): 45–51.
- [47] 赵丽梅. 科学数据共享的价值及其表征: 基于主体性的分析视角[J]. 自然辩证法研究, 2022, 38(5): 116–122.