

科学数据共享意愿影响因素实证与仿真研究 *

■ 支凤稳^{1,2,3} 张萌¹¹ 河北大学管理学院 保定 071002 ² 中国科学技术信息研究所 北京 100038³ 河北省数字治理与协同治理研究基地 保定 071002

摘要: [目的/意义] 旨在探明科学数据共享意愿的影响因素、作用机理及动态演化规律, 以期为后续相关研究与实践提供参考。[方法/过程] 基于刺激机体反应理论, 构建科学数据共享意愿的理论模型。利用问卷调查获取 435 份有效样本, 并通过结构方程模型进行实证检验, 确定变量间路径系数。在此基础上, 构建系统动力学模型, 对变量间关系的动态演化规律进行仿真。[结果/结论] “信任→共享态度→共享意愿”路径是驱动科学数据共享的基本逻辑, 互惠、社会互动、数据质量对信任有显著正向影响, 其效应排序为“互惠>数据质量>社会互动”。同时, 感知行为控制对共享意愿有显著正向影响, 而主观规范对共享意愿的影响不显著。不同情境下的信任速率、初始感知行为控制、初始共享意愿均对科学数据共享意愿产生影响。最后, 针对主要研究结论提出促进科学数据共享的建议。

关键词: 科学数据共享 共享意愿 刺激机体反应理论 影响因素 仿真

分类号: G253

DOI: 10.13266/j.issn.0252-3116.2023.13.012

随着科学研究的深入发展, 大量的科学数据不断产出, 科研人员拥有的丰富数据资源是开展科学研究的基石, 也是展现科研人员研究价值的重要产物。科学数据共享是将科学数据从个人私有物转变为科技公共财富的重要手段^[1], 已成为数据密集型范式下驱动研究的重要保障。2020 年 12 月, 中国科学文献情报中心主办了“FAIR 原则与科学数据管理与应用服务”的研讨会, 提出了对于科学数据存储、管理、共享的明确要求。2021 年 6 月, 科学欧洲 (Science Europe, SE) 发布的《可持续科学数据实用指南》指出: 共享可重复利用的科学数据以支持科学进步正日益成为常态, 越来越多的组织机构已期望并鼓励科学数据的开放共享行为^[2]。科学数据共享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, 最大程度地提高研究透明度, 有助于降低科研成本, 使公共资助研究的结果为公众所用, 还可以实现各领域数据交叉利用, 提升研究和创新水平, 更好地发挥科学数据的价值。

科学数据提供者 (下文简称“提供者”) 是科学数据共享的基础条件^[3], 科研人员作为科学数据的主要生产者和提供者, 其共享意愿直接影响着共享行为

的发生和数据价值的发挥。有研究表明, 愿意使用共享数据的科研人员较多, 而愿意提供科学数据的人员较少^[4]。因此, 有必要进一步关注个体科研人员的数据共享活动, 理解其在科学研究中的数据共享决策。科学数据在各领域中的表现形式、收集目的以及研究背景存在较大差异, 部分科研人员担心共享的数据不能被合理解释, 也有人担心共享的数据质量存在问题, 故科学数据共享尚未各领域普遍实现。尽管科研人员数据共享意愿的一些影响因素已被发现, 但还有一些问题值得深入探讨, 比如是否可以从新的理论视角解释数据共享意愿的影响因素? 这些因素间的相互作用呈现怎样的变化趋势? 如何进一步促进共享意愿? 为此, 笔者从提供者视角出发, 结合刺激机体反应理论构建科学数据共享意愿影响因素的理论模型, 并进行实证与仿真分析, 进而揭示影响因素及其动态变化规律, 讨论科学数据共享的对策, 为后续研究提供理论参考和实证支持。

1 相关研究述评

1.1 科学数据共享意愿影响因素研究

意愿是人们进行某项活动之前的心理趋向, 能够

* 本文系河北省高等学校人文社会科学研究项目“元宇宙时代科学数据共享模式及其应用研究” (项目编号: BJS2022027) 研究成果之一。

作者简介: 支凤稳, 副教授, 博士; 张萌, 硕士研究生, 通信作者, E-mail: zhangme96@126.com。

收稿日期: 2022-12-30 修回日期: 2023-04-11 本文起止页码: 111-121 本文责任编辑: 徐健

对行为进行预测,科学数据共享意愿是指科研人员将科学数据进行共享的主观意愿的强烈程度^[5]。共享态度通过共享意愿间接影响共享行为^[4],因此,共享意愿在科学数据共享过程中起到重要作用,诸多学者结合多个维度对其影响因素展开研究。

个体维度中,王丹丹^[6]等发现年龄、性别、学科背景影响科学数据共享意愿,且女性的共享意愿比男性强烈,气候、生态、生物等领域的共享意愿比社会科学和医学领域更强烈。基于心理账户理论,沈婷婷^[7]发现熟人之间与陌生人之间的科学数据共享意愿存在差异。W. Z. Mltgen^[8]发现态度、共享经验影响社会学与政治学领域作者的数据共享意愿。此外,也有学者考虑到了其他维度。张晋朝^[9]从个体与制度维度研究发现,主观规范和信念对科学数据共享意愿有正向影响,而感知行为控制的影响不显著。从资源维度来看,数据来源可信度通过感知有用性和感知易用性正向影响科学数据共享意愿^[4]。刘嫣^[10]从个人、政策、环境、资源等多个维度对高校科研人员的共享意愿展开研究,发现其影响因素包括自我价值实现、信任、主观规范、机构政策、学科领域要求、平台系统安全性及易操作性。结合个人、制度、成本维度的研究发现,利益回报、数据素养、团体氛围、期刊出版压力、资助机构压力等会推动科研人员共享数据^[11-12]。

1.2 刺激机体反应理论应用研究

刺激机体反应理论(Stimulus Organism Response, SOR)是环境心理学家 A. Mehrabian 提出的广泛应用于行为预测的理论模型^[13],该模型解释了外部环境刺激对机体认知和心理产生作用的过程,并进一步对机体行为意愿和实际行为进行预测。

近年来,SOR 理论开始应用于信息行为领域的相关研究中,尤其在信息共享、知识共享等方面对个体行为进行了很好的解释。张海^[14]借鉴 SOR 理论,从系统刺激和环境刺激出发,构建科学数据开放意愿的理论模型;彭丽徽^[15]将动机、机会、能力作为刺激因素,将知识创新行为作为反应因素,研究虚拟社区知识创新行为。科学数据作为科学研究的重要产出,它在共享的过程中同样受到外部刺激的影响,而提供者在做出相应反应前也会经历认知和情感状态的变化。因此,SOR 理论为科学数据共享意愿研究提供了坚实的理论基础。

通过文献梳理发现,现有研究主要基于计划行为、技术接受、双路径等理论构建科学数据共享意愿

和行为的理论模型,虽然理论模型在不断发展,但理论视角较为单一。大多学者基于计划行为理论构建相对静止的模型,探究科学数据共享意愿和行为的影响因素^[16-17]。这些研究虽然充分考虑科研人员自身因素,较为深入地探寻了不同身份特征下心理状态对共享意愿和行为产生的影响,但研究的影响因素比较主观,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,忽略了个体行为过程中所产生的信任、依赖情感等情绪感受因素。SOR 理论可以体现外界刺激对机体认知和心理产生作用的过程,更好地解释个体受外部因素影响而产生的情感变化,进而影响个体行为的动态过程。此外,系统动力学仿真有助于分析影响因素间的动态演化规律,已被应用于预测科学数据共享安全系统的研究中,鉴于科学数据共享的复杂性和持续性,有必要从系统动力学的角度探究科学数据共享意愿的动态演化规律。然而,仅采用系统动力学仿真,难以保证存量流量图和仿真数据的科学性。因此,笔者基于 SOR 理论,采用结构方程技术与系统动力学仿真相结合的方法,构建科学数据共享意愿的系统动力学模型,以揭示各影响因素之间的关系及其对共享意愿演化的影响,有望丰富现有研究成果,完善国内相关研究的理论和方法,为促进科学数据共享提供参考。

2 研究假设与理论模型

2.1 刺激与机体

互惠作为社会资本创造和再生的基本驱动力^[18],是指在付出后希望得到相应回报。互惠的原则是对等,在科学数据共享中主要表现为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。诸多研究证实了信任对互惠的正向影响,如李静等^[19]从社会网络关系视角发现人际信任通过互惠的中介作用影响中青年的心理健康。科研人员作为独立的个体,同样是社会关系网络中的一员,在外部刺激到机体反应的过程,互惠作为外部刺激因素不仅可以培养科研人员之间的信任关系,还可以增强提供者对数据共享平台和共享工具可靠性的信任。因此,笔者提出以下假设:

H1:互惠显著正向影响提供者的信任。

社会互动是人与人之间进行信息交换和传播的通道,提供者通过与使用者、管理者等频繁互动,建立其在专业领域内的信任。个体对某一活动的认知和判别取决于个体所接受到的刺激,而这个刺激可以被

视为外部刺激（如社会互动）。根据社会资本理论，社会互动会影响人类的行为，而在科学数据共享中，社会互动的密切程度对数据共享具有重要影响。随着科研人员开展合作研究的需求日益增加^[20]，数据共享更强调科研人员之间的行为交互^[21]，以降低科研成本，加速科学成果产出，因此，笔者提出以下假设：

H2：社会互动显著正向影响提供者的信任。

数据质量是指在特定条件下使用数据时，数据特性满足明确和隐含要求的程度。科学数据的质量是科研人员进行数据共享关注的焦点^[22]，其属性多种多样，包括真实性、完整性、权威性、可用性和有效性。刘桂锋等^[23]发现，数据质量是最重要、最基础的影响因素。当数据质量较高时，可能会增强提供者对共享数据可靠性的信任，反之，可能会降低对数据共享的信赖程度。因此，笔者提出以下假设：

H3：数据质量显著正向影响提供者的信任。

2.2 刺激与反应

主观规范是指科研人员在选择是否进行科学数据共享活动时感知到的社会压力，它能反映出专家、同事或者合作者对共享决策的影响。计划行为理论认为主观规范是行为意愿的直接影响因素之一，有学者发现，它对我国科研人员的数据共享意愿也有积极影响^[5,24]。当科研人员受到专家、同行或者合作者行为或建议的影响时，会产生遵从动机而更可能进行科学数据共享。因此，笔者提出以下假设：

H4：主观规范显著正向影响提供者的共享意愿。

感知行为控制是科研人员对科学数据共享活动难易程度的感知，它反映科研人员是否具有科学数据共享所需的资源及能否把握共享机遇^[25]。依据计划行为理论，感知行为控制是影响行为意愿的最直接因素。当科研人员接受过科学数据共享的教育或培训，掌握了相关技能，并且具备数据共享的知识与经验时，更可能进行数据共享。相反，若科研人员不具备数据共享的基础条件，感知到数据共享比较困难，参与数据共享的可能性也会降低。因此，笔者提出以下假设：

H5：感知行为控制显著正向影响提供者的共享意愿。

2.3 机体与反应

在社会交换理论中，信任是人们进行社会交换的基础条件，影响着人们参与某项活动的意愿和行为。在知识获取和分享的研究中，K. Andrews 发现信任

比知识获取及分享的方式、操作程序更为重要，如果参与者对这一活动的信任程度较低，那么知识共享行为就不会发生^[26]。在调查服务业员工的知识共享态度和意愿的基础上，郑万松等发现信任对知识共享态度有显著正向影响^[27]。科学数据共享是一种社会交换行为，当提供者对使用者以及共享平台产生信任时，对数据共享的态度也会更积极。因此，笔者提出以下假设：

H6：信任显著正向影响提供者的共享态度。

2.4 共享态度与共享意愿

共享态度是科研人员在主观上对科学数据共享行为持有的积极或消极的心理倾向。共享意愿是提供者愿意共享数据的程度。在传统技术接受模型中，共享态度对共享意愿具有积极的影响^[28]。这一结论也在后续研究中得以证实，如何琳等^[5]发现科研人员科学数据的共享态度对共享意愿有显著正向影响。当科研人员对科学数据共享持积极态度时，理论上更可能参与科学数据共享；反之，其共享数据的可能性就会降低。因此，笔者提出以下假设：

H7：共享态度显著正向影响提供者的共享意愿。

SOR 理论可以体现科研人员面对外部刺激时，个体认知和心理状态的变化。基于 SOR 理论与前文的研究假设，构建本研究的理论模型，如图 1 所示：

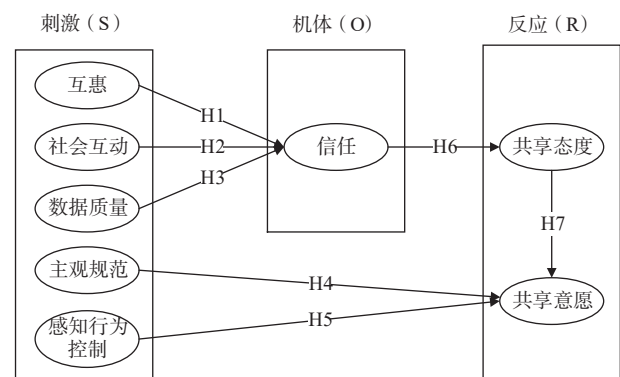


图 1 理论模型

3 实证检验

3.1 数据获取

笔者主要采用问卷调查方法获得样本数据，问卷第一部分是调查对象的基本信息，第二部分是量表，采用 5 级 Likert 进行测量，包含 8 个潜变量及 33 个题项。参考 C. M. Chiu^[29] 的研究，互惠变量设置以下 4 个题项：科研效率（RE1）、学术影响力（RE2）、经济效益（RE3）、长期合作关系（RE4）；

社会互动变量设置以下 3 个题项：花费时间交流（SI1）、建立密切联系（SI2）、主动进行互动（SI3）。参考 B. H. Wixom^[30] 和 V. McKinney^[31] 的研究，数据质量变量设置以下 5 个题项：真实性（DQ1）、完整性（DQ2）、权威性（DQ3）、可用性（DQ4）、有效性（DQ5）。参考余玲^[32] 的研究，主观规范变量设置以下 4 个题项：专家行为或建议（SN1）、同事行为或建议（SN2）、合作者行为或建议（SN3）、学术界宣传和倡导（SN4）。参考何琳^[5] 的研究，感知行为控量变量设置以下 4 个题项：教育或培训（PBC1）、数据共享的知识（PBC2）、数据共享的经历（PBC3）、数据共享的相关技能（PBC4）。参考 H. H. Chang^[33] 和张亚明^[34] 的研究，信任变量设置以下 5 个题项：数据价值（TR1）、共享途径或平台可靠性（TR2）、工具安全性（TR3）、制度保障有效性（TR4）、合理利用数据（TR5）。参考王春晓^[35] 的研究，共享态度变量设置以下 4 个题项：必要性（SA1）、明智性（SA2）、成就感（SA3）、科研信心（SA4）。参考郭宇^[36] 的研究，共享意愿变量设置以下 4 个题项：共享方式（SW1）、无偿共享数据（SW2）、数据重用（SW3）、宣传数据共享的价值（SW4）。

经过预调研和多轮讨论，形成最终问卷。选取高校及科研院所的科研人员为调查对象，通过多种方式收集数据，问卷收集时间为 2022 年 1 月，最终回收问卷 532 份，经过筛选去掉问卷填写不完整、题项重复率过高、填写时间较短、对科学数据了解程度较低的答卷，得到有效问卷 435 份，有效答卷回收率为 81.8%。

3.2 描述性统计分析

本次调查的 435 位科研人员中女性偏多，占比 59.9%，年龄集中在 40 岁以下。学历基本在本科及以上，且大部分有科研经历。43% 的科研人员具有职称，73.1% 的人员从事科研工作的年限在 5 年以下。这可能是由于一些科研人员从事科研工作时间较短，还未晋升职称。从工作单位类型来看，77.7% 的科研人员在高校或科研院所，在企业、政府部门或者其他类型的工作单位较少，说明高校及科研院所是大部分从事科研工作的人员的首选，基本与实际情况吻合。从学科背景来看，管理学领域占比 63.9%，工学占比 18.6%，理学占比 7.4%，经济学占比 3.0%，其他占比 7.1%。从对科学数据共享的了解程度来看，被调查者均对其有所了解，只是了解

程度有所差别。

此外，量表各题项极小值为 1，极大值为 5，样本均值在 3.25—4.2 之间，标准差在 0.627—1.038 之间，表明调查对象在问卷填写过程中存在主观差异，符合现实情况。统计结果表明，样本具有较好的离散程度，可以满足研究需要。

3.3 信度与效度检验

信度即可靠性，指采用同样方法对同一对象重复测量时，所得结果的稳定性、一致性和可靠性，即测量工具能否稳定地测量所测的事物或变量。学者们常用 Cronbach's α 对量表进行信度检验， α 系数范围在 0—1 之间。一般情况下，当 $0.7 \leq \alpha < 0.8$ 时，量表信度较高；当 $0.8 \leq \alpha < 0.9$ 时，量表信度很高；当 $\alpha \geq 0.9$ 时，量表信度非常高。SPSS 分析结果如图 1 所示（应该是表 1），8 个潜变量 Cronbach's α 均大于 0.8，说明量表各维度内部具有较高的稳定性、一致性和可靠性。

效度分为内容效度和结构效度，本研究采用国内外相关研究中的成熟量表，并进行多轮专家咨询，能够保证量表的内容效度。结构效度是量表中测量题项和被研究的理论概念一致性的程度，通常分为聚合效度和区分效度，一般使用因子分析的方法进行检验。在探索性因子分析时，采用主成分方法提取公因子方差，抽取 8 个因子，并且通过最大方差法对因子成分矩阵进行旋转。结果显示，公因子方差提取值均大于 0.5，表明提取的公因子对潜变量的代表性和解释率较好。在此基础上分析旋转后的因子成分矩阵，依据因子聚类情况以及测量题项的内容，剔除 RE1、TR1、SW1 三个因子，剩余因子的聚类关系基本明确。

通过 AMOS 对量表进行验证性因子分析，结果见表 1。聚合效度用于判断某一潜变量下的多个观测指标能否代表这一潜变量的特质。标准化因子载荷、平均方差抽取值（AVE）以及组合信度（CR）是衡量聚合效度的基本指标，其中，标准化因子载荷和 AVE 要大于 0.5，CR 确保在 0.7 以上。由表 1 可知，25 项指标全部符合标准，因此，量表均具有较好的聚合效度。

区分效度是为了验证不同潜变量之间的相关性，判断是否存在显著性差异，一般用变量之间的相关系数和 AVE 算数平方根的大小关系来判断。对比结果显示，AVE 的算数平方根均大于变量之间的相关系数，因此，量表具有较好的区分效度。

表 1 信度与效度检验结果

潜变量	观测指标	标准化因子载荷	Cronbach's α	AVE	CR
互惠	RE2	0.820	0.849	0.653	0.849
	RE3	0.791			
	RE4	0.812			
社会互动	SI1	0.861	0.839	0.655	0.850
	SI2	0.718			
	SI3	0.842			
数据质量	DQ1	0.813	0.917	0.702	0.921
	DQ2	0.808			
	DQ3	0.750			
	DQ4	0.919			
	DQ5	0.889			
主观规范	SN1	0.868	0.919	0.742	0.920
	SN2	0.887			
	SN3	0.895			
	SN4	0.792			
感知行为控制	PBC1	0.859	0.941	0.804	0.943
	PBC2	0.912			
	PBC3	0.874			
	PBC4	0.940			
信任	TR2	0.857	0.910	0.718	0.910
	TR3	0.865			
	TR4	0.841			
	TR5	0.825			
共享态度	SA1	0.801	0.901	0.698	0.902
	SA2	0.831			
	SA3	0.842			
	SA4	0.866			
共享意愿	SW2	0.716	0.841	0.660	0.853
	SW3	0.881			
	SW4	0.832			

3.4 结构方程模型检验

通常用 χ^2/df 、GFI、AGFI、NFI、IFI、CFI、RMSEA 等指标对模型拟合度进行检验, 拟合参数见表 2。比较观测值与标准值, 发现理论模型可以被接受。

表 2 模型拟合参数

模型	χ^2/df	GFI	AGFI	NFI	IFI	CFI	RMSEA
标准值	<3	>0.8	>0.8	>0.8	>0.8	>0.8	<0.08
观测值	2.660	0.855	0.826	0.908	0.940	0.940	0.062

对提出的 7 个假设进行验证分析, 根据 P 值的大小得出假设路径的相关性程度。如表 3 所示, * 表示 $P < 0.05$, 表示相关性显著; ** 表示 $P < 0.01$, 表示相关性比较显著; *** 表示 $P < 0.001$, 表示相关性非常显著。标准化路径系数 (β) 表示该假设是正向影响或者负向影响, 其数值越大, 表示影响越显著。结果显示, 除 H4 ($P=0.712>0.05$) 外的假设均成立, 且均为显著正向影响。

表 3 假设检验结果

假设	路径	β	P 值	结论
H1	互惠→信任	0.435	***	成立
H2	社会互动→信任	0.185	***	成立
H3	数据质量→信任	0.222	***	成立
H4	主观规范→共享意愿	-0.022	0.712	不成立
H5	感知行为控制→共享意愿	0.190	***	成立
H6	信任→共享态度	1.031	***	成立
H7	共享态度→共享意愿	0.688	***	成立

4 系统动力学仿真

4.1 存量流量图构建

系统动力学强调从整体动态发展的视角研究系统各部分要素的内在逻辑关系, 并根据要素间因果关系反馈信息, 发现系统内部要素相互作用产生的影响, 如阮冰颖^[37]通过系统动力学模型, 探究科学数据共享安全系统中不同因素间的动态规律。笔者

初步确定科学数据共享意愿的 8 个影响因素, 而实证结果表明, 主观规范对共享意愿的影响不显著, 故最终确定除主观规范外的 7 个影响因素。参考结构方程模型中的测量变量, 确定出系统动力学模型

中的 7 个状态变量、7 个速率变量、26 个常量。基于上文构建的理论模型和实证结果, 构建存量流量图, 为建立各要素间关系方程式奠定基础, 如图 2 所示:

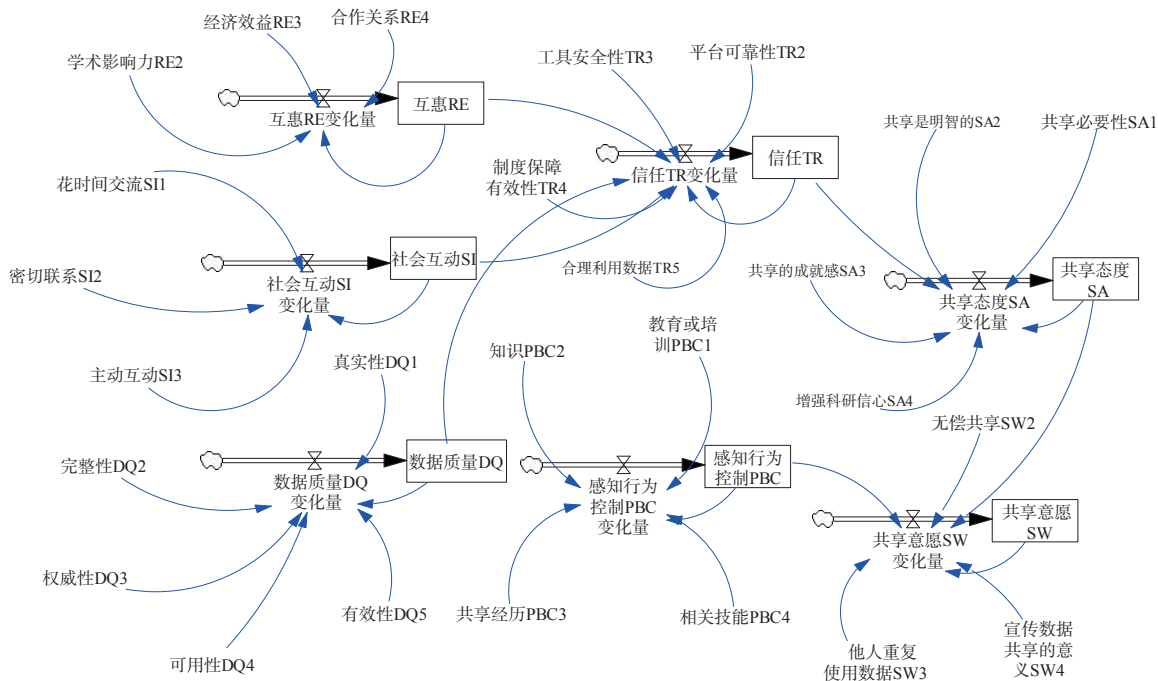


图2 科学数据共享意愿的存量流量示意

在科学数据共享意愿的系统动力学模型中,各变量的相关影响系数和权重系数,均参考结构方程模型实证过程中的因子载荷和路径系数。其中,常量取值为问卷中各题项的均值,状态变量依据标准化因子载荷计算权重,通过权重和常量值计算得出状态变量的初始值,具体如表4所示。

学者们通常在系统动力学模型仿真过程中引入**逻辑斯谛模型**^[38]，通过对系统动力学模型进行调试，更直观地展现仿真结果。据此，本研究设定 10 为科学数据共享实现的最高标准，52 周为仿真时长，0.5 为仿真步长。

4.2 自然状态下的科学数据共享意愿仿真

在进行仿真实验前,需要对模型的逻辑一致性以及系统稳定性进行测试,确保模型的可靠性和稳定性。在模型通过检验后,采用 Vensim 对数学方程式进行计算,输出仿真结果。在仿真周期内,科学数据共享意愿随着时间变化,呈现不断上升的趋势(见图 3)。到第 30 周之前,增长趋势较快,30 周之后逐渐放缓。在共享的前期阶段,互惠、社会互动、信任等增加,会对提供者的共享意愿产生较大的影响,上升趋势较快。经过前 30 周的积累,提供者掌握了丰富的数据共享知识,对科学数据使用者、共享平台、

表 4 系统动力学模型初始参数值

常量	题项均值	标准化因子载荷	权重	状态变量	初始值
RE2	4.076	0.820	0.338		
RE3	3.908	0.791	0.326	RE	3.997
RE4	4.016	0.812	0.335		
SI1	3.871	0.861	0.356		
SI2	3.522	0.718	0.297	SI	3.770
SI3	3.867	0.842	0.348		
DQ1	4.177	0.813	0.195		
DQ2	4.014	0.808	0.193		
DQ3	3.839	0.750	0.179	DQ	4.043
DQ4	4.087	0.919	0.220		
DQ5	4.074	0.889	0.213		
PBC1	3.248	0.859	0.240		
PBC2	3.423	0.912	0.254	PBC	3.327
PBC3	3.301	0.874	0.244		
PBC4	3.329	0.940	0.262		
TR2	3.986	0.857	0.253		
TR3	3.968	0.865	0.255		
TR4	3.982	0.841	0.248	TR	3.947
TR5	3.851	0.825	0.244		
SA1	4.078	0.801	0.240		
SA2	4.007	0.831	0.249	SA	3.995
SA3	3.908	0.842	0.252		
SA4	3.993	0.866	0.259		
SW2	3.526	0.716	0.295		
SW3	3.828	0.881	0.363	SW	3.757
SW4	3.871	0.832	0.343		

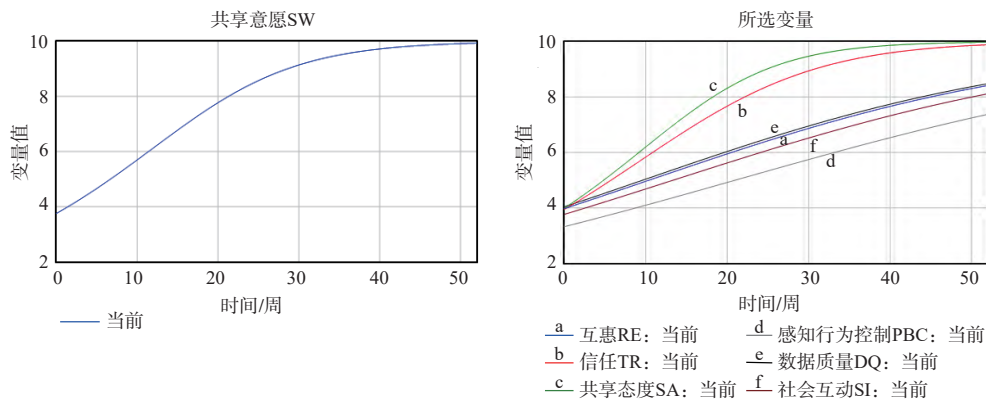


图3 科学数据共享意愿及各影响因素变化趋势

相关制度的信任达到峰值, 并且对数据共享所获得的经济效益和影响力也有了充分的了解和认识。此时, 提供者已具备较为成熟的数据共享能力, 后期的数据共享意愿增加趋势也会减缓。

由图3可知, 信任和共享态度的变化趋势都呈现出前期增长较快, 后期增长减缓的趋势。由于信任的变化率受到互惠、社会互动、数据质量因素的综合影响, 信任的增长趋势不能形成稳定的线性关系。随着时间的增加, 影响信任的各项因素达到顶峰, 后期的增长趋势会相应减缓。同样, 共享态度的变化率受到信任的直接影响, 互惠、社会互动、数据质量也会通过信任间接影响共享态度, 其增长趋势几乎呈线性增长, 且短期内较为稳定。

4.3 不同情境下的科学数据共享意愿仿真

灵敏度分析是通过调整变量的初始取值大小、变

量速率等, 分析不同情境下科学数据共享意愿的动态变化趋势。

(1) 改变信任速率。根据SOR理论, 机体接受外界刺激产生变化, 提供者的信任受到互惠、社会互动与数据质量的影响, 而不同提供者应对外界刺激的能力以及产生的变化存在差异。方案一将信任受到互惠、社会互动、数据质量的影响提升50%, 初始值降低30%; 方案二将信任受到互惠、社会互动、数据质量的影响降低50%, 初始值提升30%, 观察信任、共享态度、共享意愿的变化趋势。测试对外界反应敏感但初始信任较低的提供者, 以及对外界反应迟钝但初始信任较高的提供者, 并对比其在应对外界刺激方面存在的差异, 以及其信任的时间变化趋势, 从而进一步分析共享态度和共享意愿的差异。详情如图4所示:

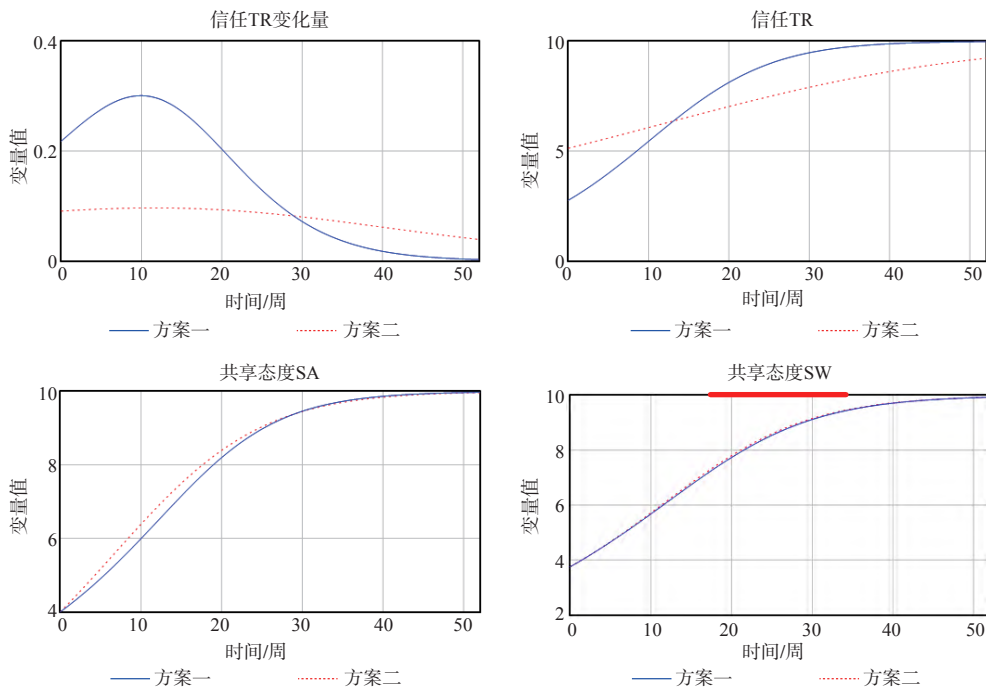


图4 不同信任速率对比

应该是共享意愿

由图4可知,在方案一中,应对外界刺激能力较强的提供者的信任增长速度在前10周非常快,后期速度逐渐减缓,且在前29周仍然高于应对外界刺激能力较低的提供者。在方案二中,虽然设置的信任初始值较高,但由于应对外界刺激的能力较弱,受到互惠、社会互动、数据质量的影响程度较小,到14周之后,信任程度持续低于应对外界刺激能力较强的提供者。对于数据共享平台、途径、工具的信任程度较高的提供者,其共享态度和共享意愿随着时间增加,一直高于信任程度较低的提供者。共享态度和共享意愿不仅受互惠、社会互动、数据质量的影响,还会受

信任以及感知行为控制的影响,最终呈现的变化趋势与信任不太一致。

(2) 改变感知行为控制初始值。感知行为控制对科学数据共享的影响未形成统一观点^[24,39],但本研究发现其对共享意愿具有显著的正向影响。考虑到提供者在教育、知识和相关技能方面的差距,以及科学数据共享经历的差异性,笔者设计了两个对比方案。方案一将感知行为控制的初始值设定为5,方案二将其设定为1,分别测试在科学数据共享教育或培训、知识、经历以及相关技能方面相对较多与较少的提供者。如图5所示:

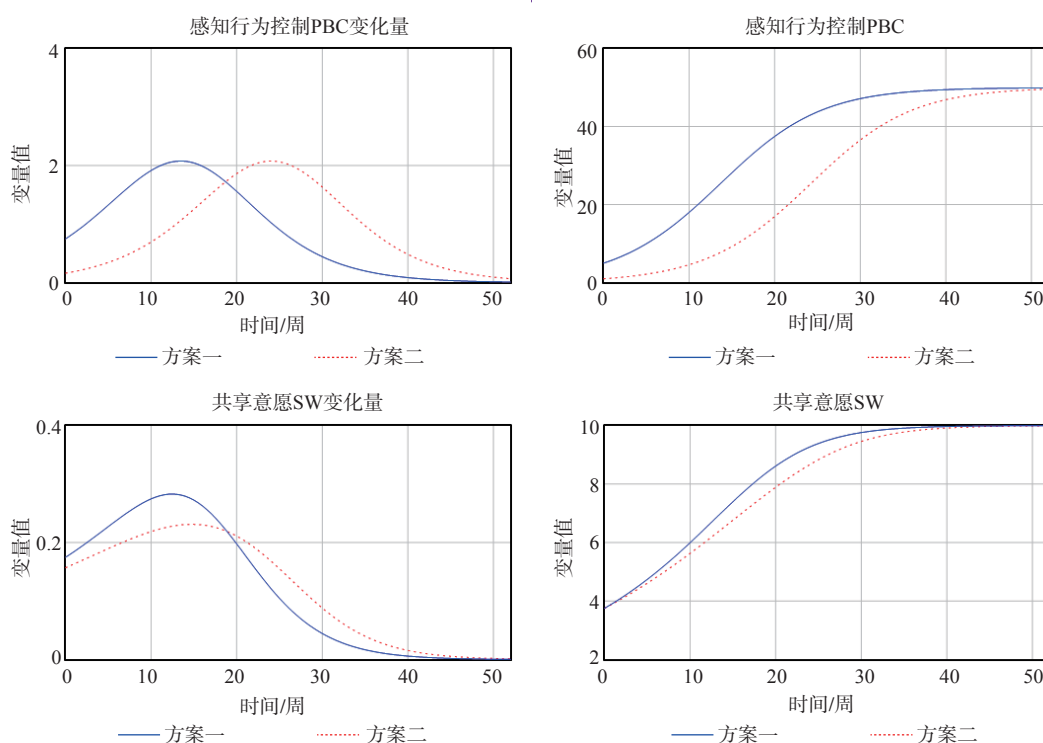


图5 不同感知行为控制对比

由图5可知,方案一中的提供者与方案二的提供者仅在感知行为控制初始值上存在差异,且方案一中提供者的感知行为控制初始值较高。最终仿真对比结果显示,方案一在第13周的增长速度达到高峰,而方案二在第23周到达了同样的顶峰。说明早期接受过较多共享数据教育或培训、具备丰富的共享数据知识和经历的提供者能够更快地适应数据共享活动,并且共享意愿较高。由共享意愿变化量的趋势可以看出,在第19周之前,方案一中提供者的共享意愿增加量一直高于方案二,而在19周之后,呈现出与前期相反的趋势。说明到仿真周期的第19周,提供者在科学数据共享教育或培训方面已经积累了丰富的知识和经验,其共享意愿也发展到了一个较为成熟的

阶段,反之,其数据共享意愿还具备较大的提升空间。因此,在19周之后,方案二的共享意愿变化量一直高于方案一。

(3) 改变共享意愿初始值。问卷中“在一定条件下,您愿意向他人提供自己的科学数据吗?”这一问题设置5个选项,意愿程度由强到弱。为了清晰呈现提供者在影响因素的作用下,共享意愿的时间变化趋势,设定3个方案。方案一将共享意愿的初始值设定为1,测试共享意愿弱的提供者;方案二将共享意愿的初始值设定为3,测试共享意愿一般的提供者;方案三将共享意愿的初始值设定为5,测试共享意愿强的提供者。对比如图6所示:

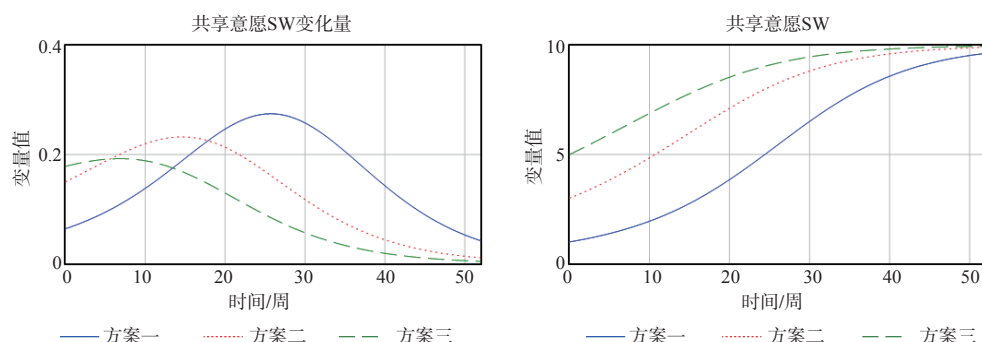


图6 不同初始条件下共享意愿变化趋势

由图6可知, 方案一、二、三中的共享意愿变化量依次降低, 分别在第22周、第15周、第10周达到峰值。仿真对比结果显示, 在前6周, 方案三的共享意愿变化量和共享意愿均高于方案一和方案二; 在6—16周, 方案二的共享意愿变化量高于方案一和方案三; 在16周后, 方案一的共享意愿变化量高于方案二和方案三。方案三的共享意愿始终高于方案一和方案二。方案一的提供者, 受到信任、感知行为控制的影响, 其在后期的共享意愿变化量较大, 具有较大的数据共享潜力。

5 结果分析与讨论

笔者对实证与仿真的结果进行讨论, 并尝试提出针对性的对策。

(1) “信任→共享态度→共享意愿”路径是驱动科学数据共享的基本逻辑。互惠、社会互动、数据质量对信任有显著正向影响, 效应排序为“互惠>数据质量>社会互动”。在知识共享方面的研究发现, 人们对奖励机制的感知会正向影响其信任^[40]。这种奖励包括经济效益、科研影响力、数据发表机会等, 科学数据共享与知识共享具有诸多相似之处。本研究发现, 互惠比数据质量、社会互动对信任的影响更强, 互惠是科研人员科学数据共享的基础条件, 而信任是在互惠的基础上进一步提升共享积极性的驱动力。Y. Kim^[40]通过设置开放性问题证实了科研人员会选择可信度高的同事、合作者共享自己的数据, 相关主体应重视互惠、社会互动、数据质量对信任的积极作用。也有研究对共享态度到共享意愿的影响关系进行验证^[24], 如果科研人员认为数据共享是有价值的, 则其对此持有积极的态度, 那么更愿意共享自己的数据。因此, 相关机构或部门应该积极宣传数据共享的重要性, 从一定程度上削弱科研人员对数据共享的消极评价, 进而提高其数据共享的参与度。还应明确各阶段的主体责任, 建立统一的数据标准, 确保数据的可用性和易用性, 对开放共享科学数据质量高、声誉好的科研人员,

进行物质或荣誉奖励, 反之, 对其给予适当惩罚^[41]。

(2) 主观规范对提供者的共享意愿影响不显著。这一结论虽然与计划行为理论的观点存在偏差, 但支持了Y. Kim^[42]关于社会科学家数据共享行为的研究结果。本研究在调查时发现, 虽然大部分提供者愿意共享自己的科学数据, 但实际参与共享的人员相对较少。这可能是提供者周围的专家、同事、合作者缺乏科学数据共享经验, 其行为或意见对提供者的影响较小, 因而导致主观规范对共享意愿的解释程度偏低, 甚至没有达到显著水平。

(3) 提供者应对外界刺激的能力越强, 其对数据共享途径、平台、工具、相关制度保障以及对使用者的信任水平越高, 进而对科学数据共享的意愿越强烈。问卷调查显示, 科研人员对科学数据共享途径或平台可靠性的信任均值最高, 而对数据使用者合理利用共享数据的信任均值相对较低, 表明与其他硬件设施、制度保障相比, 科研人员更担心使用者不能合理利用自己共享的数据。在各个领域中, 存在一些机会主义、不诚实、自私的人员, 而科研领域更需要实事求是的学术氛围, 故信任对共享态度的影响($\beta=1.031$)比其他路径的影响程度更强烈。此外, 通过测试对外界影响敏感但初始信任较低的提供者, 以及对外界影响迟钝但初始信任较高的提供者发现, 初始信任的高低对最终的共享意愿影响较小, 而对互惠、社会互动、数据质量等因素保持高敏感度, 有助于提高数据共享意愿。

(4) 感知行为控制显著正向影响提供者的共享意愿, 并且初期接受过较多共享数据教育或培训, 具备丰富的共享数据知识和经历的提供者, 其共享意愿始终高于初期没有受过此类教育或培训者。感知行为控制对意愿的影响会根据具体的研究环境而发生改变^[43]。如果科研人员接受过数据共享的教育或培训, 具备数据共享的知识、经历或者相关技能, 会更愿意与他人共享自己的数据。科研人员可以通过积极参加相关培训与讲座, 开展数据共享实践等方式, 识别出

自身在共享数据方面的薄弱环节,针对性地改善数据共享行为,提高数据共享能力。相关机构和部门要加强科学数据人才培养,并把人才渗透到科学数据集产生、共享、重用等各个环节,加快知识扩散的过程,满足飞速发展的科学数据管理和需求^[44]。

(5) 初期共享意愿弱的提供者,随着时间变化,受到刺激、机体维度的因素影响,其共享意愿显著提高,具有较大的数据共享潜力。初期共享意愿强的提供者,前期的共享意愿变化量较大,始终保持较强的共享意愿。不同意愿程度的提供者在影响因素的作用下,随时间而产生的意愿变化趋势符合事物发展的客观规律。初期共享意愿弱的提供者,随时间变化,其共享意愿的变化量近似于正态分布,其共享意愿呈生长曲线型,具有较大的数据共享潜力。根据“信任→共享态度→共享意愿”的影响路径,提高共享意愿的关键是增强对科学数据共享的信任,而对于共享意愿弱的提供者来说,他们的初始信任相对较低。互惠对信任的影响程度($\beta=0.435$)表明互惠是信任更为关键的影响因素,因此,可以通过互惠间接提高科学数据共享意愿。此外,科研资助机构和科学期刊可以制定相应的激励机制,明确提供者的贡献,并从经济效益和学术声誉等方面给予奖励,使科研人员受益于共享的数据,从而营造良好的数据共享氛围。

6 结语

本研究基于SOR理论,构建科学数据共享意愿的理论模型,从提供者视角实证科研人员共享意愿的影响因素,并引入系统动力学模型,设计3个不同情境下的测试方案进行仿真,探究科学数据共享意愿的动态变化规律,有望为后续研究提供一定的理论参考和实证支持。本研究尚存在一些局限,希望未来研究持续关注与完善,具体如下:①仅通过调查问卷的方式获取样本数据,未来研究可以通过多种途径进行数据收集,并考虑融合多种方法探索新的影响因素;②由于调研范围有限,管理学领域的样本居多,研究结论是否适用于其他学科领域,仍需进一步验证;③共享意愿并不能直接转化为共享行为,未来研究可以考察数据共享意愿、实际行为与预测结果之间的互惠循环,深入探索数据共享促进行为的策略。

参考文献:

- [1] 赵雨梅,刘雯.基于文献计量的我国科学数据共享研究态势分析[J].图书馆研究,2022,52(5):117-128.
- [2] Science Europe. Practical guide sustainable to research date[EB/OL]. [2023-01-14]. <http://www.science-europe.org/media/b3odxx3s/se-practical-guide-sustainable-research-data.pdf>.
- [3] 王勋荣,蒲筱哥.高校科学数据共享影响因素与策略研究[J].图书馆,2021(10):41-46,62.
- [4] 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.
- [5] 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.
- [6] 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.
- [7] 沈婷婷.主观感知对研究者科学数据共享意愿的影响[J].情报杂志,2021,40(12):158-162.
- [8] MLTGEN W Z, AKDENIZ E, KATSANIDOU A, et al. Factors influencing the data sharing behavior of researchers in sociology and political science[J]. Journal of documentation, 2018, 74(5): 1053-1073.
- [9] 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.
- [10] 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.
- [11] BEARTH A, SIEGRIST M. Psychological factors that determine people's willingness-to-share genetic data for research[J]. Clinical genetics, 2020, 97(3): 483-491.
- [12] 罗晓兰,李明.国内期刊论文科学数据共享政策与投稿意愿研究[J].中国科技期刊研究,2017(8):696-703.
- [13] MEHRABIAN A, RUSSELL J A. An approach to environmental psychology[M]. Cambridge: MIT Press, 1974: 24-30.
- [14] 张海,刘蕾.高校科研人员科研数据开放意愿的影响因素研究[J].新世纪图书馆,2020(11):57-64.
- [15] 彭丽徽,蒋欣.虚拟社区感知视角下用户知识创新行为影响因素研究[J].现代情报,2023,43(2):86-95.
- [16] 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.
- [17] HARPER L M, KIM Y. Attitudinal, normative, and resource factors affecting psychologists' intentions to adopt an open data badge: an empirical analysis[J]. International journal of information management, 2018, 41(8): 23-32.
- [18] 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.
- [19] 李静,高晓彩.社会网络视角下人际信任对心理健康的影响[J].甘肃社会科学,2020(4):68-73.
- [20] 盛小平,孙倩倩.国内科学数据开放共享主题的作者合作关系与合作研究内容分析[J].图书情报工作,2021,65(23):13-21.
- [21] 支凤稳,云仲伦,张闪闪.基于区块链的个人科学数据共享模式研究[J].现代情报,2021,41(12):69-78.
- [22] 支凤稳,赵梦凡,张萌,等.科学数据共享需求调查与关联挖掘[J].情报科学,2021,39(12):9-16.
- [23] 刘桂峰,濮静蓉,钱锦琳.科研数据共享影响因素分析及作用阐释[J].图书馆论坛,2018(11):10-17,26.
- [24] 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.
- [25] ROWHANI-FARIDA, ALLEN M, BARNETT A G. What incentives

- increase data sharing in health and medical research? a systematic review[J]. Research integrity and peer review, 2017, 2(12): 4-10.
- [26] ANDREWS K M, DELAHAY B L. Influences on knowledge processes in organization: the psychosocial filter[J]. Journal of management studies, 2000, 37(6): 797-810.
- [27] 郑万松, 孙晓琳, 王刊良. 基于社会资本和计划行为理论的知识共享影响因素研究 [J]. 西安交通大学学报 (社会科版), 2014, 34(1): 43-48.
- [28] 明均仁, 郭财强. 移动图书馆用户使用行为意愿的动态实证研究 [J]. 图书馆建设, 2018(10): 73-79.
- [29] CHIU C M, HSU M H, WANG E. Understanding knowledge sharing in virtual communities: an integration of social capital and social cognitive theories[J]. Decision support systems, 2007, 42(3): 1872-1888.
- [30] WIXOM B H, TODD P A. A theoretical integration of user satisfaction and technology acceptance[J]. Information systems research, 2005, 16(1): 85-102.
- [31] MCKINNEY V, YOON K, ZAHEDI F M. The measurement of web-customer satisfaction: an expectation and disconfirmation approach[J]. Information systems research, 2002, 13(3): 296-315.
- [32] 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.
- [33] CHANG H H, CHUANG S S. Social capital and individual motivations on knowledge sharing: participant involvement as a moderator[J]. Information & management, 2011, 48(1): 9-18.
- [34] 吴丹, 陈晶. 我国医学从业者科学数据共享行为调查研究 [J]. 图书情报工作, 2015, 59(18): 30-39.
- [35] 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.
- [36] 郭宇. 基于信息生态视角的新媒体环境下企业知识共享研究 [D]. 长春: 吉林大学, 2016.
- [37] 阮冰颖, 刘桂锋, 刘琼. 科研人员视角下科学数据安全影响因素仿真分析 [J]. 情报理论与实践, 2022, 45(2): 137-145.
- [38] 张宝生, 张庆普. 隐性知识流转网成员合作的共生关系、演化模型及仿真研究 [J]. 现代情报, 2020, 40(11): 34-43, 53.
- [39] 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.
- [40] KIM Y. Reputation, trust, and norms as mechanisms leading to academic reciprocity in data sharing: an empirical test of theory of collective action[J]. Proceedings of the Association for Information Science and Technology, 2018, 55(1): 244-253.
- [41] 邢文明, 杨泽芳. 基于声誉理论的开放科学数据质量提升激励机制研究 [J]. 图书情报工作, 2022, 66(21): 22-28.
- [42] 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.
- [43] AJZEN I. The theory of planned behavior[J]. Organizational behavior and human decision processes, 1991, 50(2): 179-211.
- [44] 杨宁, 张志强. 科学数据集知识扩散特征探析 —— 以基因表达数据集为例 [J]. 图书情报工作, 2022, 66(12): 82-91.
- 作者贡献说明:**
支凤稳: 提出论文选题和研究思路, 修改与审定论文;
张萌: 论文框架设计, 数据收集与处理分析, 论文撰写与修改。

Empirical and Simulation Research on the Influencing Factors of Scientific Data Sharing Willingness

Zhi Fengwen^{1,2,3} Zhang Meng¹

¹ School of Management, Hebei University, Baoding 071002

² China Institute of Science and Technology Information, Beijing 100038

³ Collaborative Digital Governance Research Base of Hebei, Baoding 071002

Abstract: [Purpose/Significance] The study aims to reveal the influencing factors, working mechanism and dynamic evolution law of scientific data sharing willingness to provide a reference for the following-up research and practice. **[Method/Process]** A theoretical model of scientific data sharing willingness was constructed based on the stimulus organism response theory. 435 valid questionnaire samples were empirically tested via structural equation model to get the path coefficients among variables, based on which a system dynamics model was conducted to emulate the dynamic evolution of the relationship among variables. **[Result/Conclusion]** The path of “trust→sharing attitude→sharing willingness” is the basic logic pushing scientific data sharing forward. Mutual benefits, social interaction and data quality have remarkable positive influence on trust, and their effects are ranked as “mutual benefits>data quality> social interaction”. Meanwhile, the perceived behavioral control has prominent positive influence on sharing willingness while subjective norms are insignificant. Under different circumstances, the trust rate, the initial perceived behavior control and the initial degree of sharing willingness all exert influences on scientific data sharing willingness. Finally, suggestions of promoting scientific data sharing were proposed based on main research conclusions.

Keywords: scientific data sharing sharing willingness stimulus organism response theory influencing factors emulation