

● 包秦雯^{1,2}, 顾立平^{1,2}, 张潇月^{1,2}

(1. 中国科学院文献情报中心, 北京 100190; 2. 中国科学院大学图书情报与档案管理系, 北京 100190)

开放科研数据的行为影响因素研究^{*}

——以地球科学领域为例

摘要: [目的/意义] 开放科研数据面临着科研人员参与度和认同度不高的挑战。文章旨在明确科研人员开放科研数据的行为影响因素, 为各利益相关者更好地激励科研人员参与实践提供参考。[方法/过程] 研究基于计划行为理论建立模型, 探究行为信念、行为态度、主观规范和感知行为控制对地球科学领域科研人员开放科研数据行为的影响; 通过问卷调查收集数据, 采用偏最小二乘法结构方程模型 (PLS-SEM) 工具 SmartPLS3.0 分析数据验证模型。[结果/结论] 地球科学领域科研人员对开放科研数据态度直接显著影响行为; 行为信念 (包含感知职业利益与感知职业风险)、感知行为控制 (包含感知数据基础设施和数据管理培训实用性) 显著影响科研人员的态度, 且通过态度间接影响行为; 主观规范对科研人员的态度影响不明显, 但对行为存在显著影响。

关键词: 开放科研数据; 科研人员; 信息行为; PLS-SEM; 影响因素

Factors Influencing Open Research Data Behaviors: Earth Science Researchers' Perspective

Abstract: [Purpose/significance] Open research data faces the challenge of researchers' skepticism and low participation. To find out the motivational factors that influence researchers' open research data behaviors can help stakeholders better encourage researchers' participation. [Method/process] Based on the theory of planned behavior, a research model was established to explore how behavioral beliefs, attitudes, subjective norms, perceived behavioral control influence earth science researchers' open research data behaviors. A questionnaire survey was conducted to collect data, partial least squares structural equation modeling (PLS-SEM) tool SmartPLS3.0 was used to analyze the data and verify the research model. [Result/conclusion] The results show that earth science researchers' attitudes toward open research data significantly influence their actual behaviors. Behavioral beliefs (including the perceived career benefit and risk) and perceived behavior control (including the perceived available of data infrastructure and the perceived available of data management training) have significant influence on their attitudes, through which indirectly influence the actual behaviors. Subjective norms have significant influence on behaviors, but no influence on attitudes.

Keywords: open research data; scientific research personnel; information behavior; PLS-SEM; influencing factor

信息技术的发展和科学数据的海量增长, 催生了学术交流模式的变革和新的科学研究范式。科技和社会的创新越来越依赖于大量高质量和可信赖的科研数据。为了最大化利用科研数据的价值, 科研成果的开放获取成为全球趋势。2018年4月, 国务院印发《科学数据管理办法》, 要求对科研项目产生的数据进行存储积累和开放共享^[1]。各国科研教育机构、资助机构、出版商、图书馆及科研人员等多方利益相关者持续致力于推动开放科研数据 (Open Research Data, ORD), 呼吁构建一个开放、高效和跨学科的科研数据开放共享环境。

开放科研数据又称开放科学数据 (Open Scientific Da-

ta)。英国主要资助和研究机构在其联合发布的《开放科研数据协定》中定义 ORD: 科研数据可以被任何人自由地访问、使用、修改和分享, 并在需要时给予致谢认可^[2]。近期, FAIR 数据原则得到欧盟在内的世界主要科技先进国家的政策支持^[3], 本文据此综合修订后对 ORD 定义为: 除非有法律和道德的限制, 公开可访问的科研数据可被所有人自由获取、使用、修改和传播。

科研人员在推动 ORD 的各方利益相关者中处于关键核心地位。他们意识到 ORD 的重要性和诸多益处, 但仍存有许多担忧和疑虑, 真正进行的实践有限^[4]。科研人员所代表的利益相关者的声音和需求被长期忽视。由于缺乏沟通与理解, 他们对于大规模收割科研数据并且免费公之于众的系统项目存有疑虑。许多情况下, 急于发表科研成果也使得他们容易过早放弃从科研数据中获取更多价值的

^{*} 本文为中国科学院项目“大数据知识资源集成与服务的信息政策与法律法规议题”的研究成果, 项目编号: 院 1757。

机会^[5]。众多利益相关者权益博弈不平衡,国外资助机构机械式的政治干预混杂着不清楚的出版商和工程项目利益,就容易令人望而却步。近年来,越来越多的科研资助机构、教育机构和图书馆等都将激励科研人员作为推动 ORD 的优先事项^[6]。

本文基于计划行为理论,建立地球科学领域科研人员开放科研数据的行为影响因素模型,并展开问卷调查;数据分析采用偏最小二乘法 (Partial Least Squares, PLS) 结构方程模型 (Structural Equation Model, SEM),探究行为信念、主观规范和感知行为控制如何影响科研人员对 ORD 的态度,并进一步影响其实际行为。以期为进一步激励科研人员更好参与 ORD 提供具有实证研究结果支撑的参考信息。

1 文献综述

科研人员的态度与行为是科研数据开放共享领域的热点,研究方法以问卷调查和访谈为主,范围涵盖不同学科和年龄阶段的科研人员。随着 ORD 的推进,科研人员的认知与实践在逐渐改变。Fecher 等的调查表明不同学科的科研人员在对 ORD 认识和知识技能上有不同表现^[7]。Tenopir 等对比了数据共享不同发展阶段科研人员认知的变化,研究发现尽管科研人员对数据共享的接受度、意愿和实际参与行为都有所增加,但共享数据的风险和障碍依然存在^[8]。Kim 等的调研发现社会科学家的数据共享行为主要受个人动机 (感知职业利益和风险、感知耗费的精力、对数据共享的态度) 的驱动和主观规范的影响^[9]。

目前,国内对科研人员 ORD 行为的研究尚未形成成熟的体系。一些学者探究了影响科研人员数据共享意愿的因素。张晋朝的研究指出共享态度是影响共享意愿的最重要维度^[10]。何琳等发现感知行为控制、感知风险、感知有用性间接影响数据共享意愿^[11]。相关研究样本的学科分布较为分散,专门就特定学科展开的研究较少,孙晓燕和吴丹等分别探究了社会科学领域科研人员和医学从业者的数据共享行为^[12-13]。

本研究选取我国地球科学领域的科研人员作为调研对象。该学科历来重视科研数据的开放共享,且在国际合作、基础设施构建^[14]、元数据标准建设和数据关联出版^[15]等方面已有较为成熟的实践。国内有学者从地学数据共享平台的用户搜索和访问行为角度进行了行为预测和数据推荐的研究^[16],但少有研究涉及该领域科研人员对开放科研数据的态度与行为。ORD 的实现必须考虑到不同学科之间数据共享文化和科研活动的巨大差异^[17]。因此,探究地球科学领域科研人员开放科研数据的行为影响因素对进一步激励他们参与 ORD 有重要意义。

2 研究模型建构与假设

本研究基于计划行为理论 (Theory of Planned Behavior, TPB) 构建研究模型。TPB 认为,个体的行为受行为意向 (Behavioral Intention) 直接决定,行为意向受其行为态度 (Behavioral Attitude)、主观规范 (Subjective Norms, SN) 和感知行为控制 (Perceived Behavioral Control, PBC) 的影响^[18]。TPB 被广泛应用在多领域,且被证实能显著提高对行为的解释力和预测力^[19]。且国内外有许多学者将其应用在知识和数据共享行为研究中。

本文基于 TPB 和相关文献构建研究模型与假设 (见图 1),该模型旨在探究行为信念、行为态度、主观规范和感知行为控制如何影响地球科学领域科研人员 ORD 行为。行为信念 (Behavioral Beliefs) 是个体对行为发生可能性的预期,影响科研人员的态度,包含感知职业效益 (Perceived Career Benefit, PCB)、感知职业风险 (Perceived Career Risk, PCR) 和感知耗费的精力 (Perceived Effort, PE)。PBC 是不受个人主观意志控制的外部资源与机会,影响科研人员的态度与行为,包含感知数据基础设施实用性 (Perceived Available of Data Infrastructure, PA-DI) 和感知数据管理培训实用性 (Perceived Available of Data Management Training, PDMT)。近年来,其他学者对数据和知识共享行为的研究证实,态度可以直接影响人的实际行为^[12 20-21]。因此,本文假设科研人员对 ORD 的态度直接影响行为。

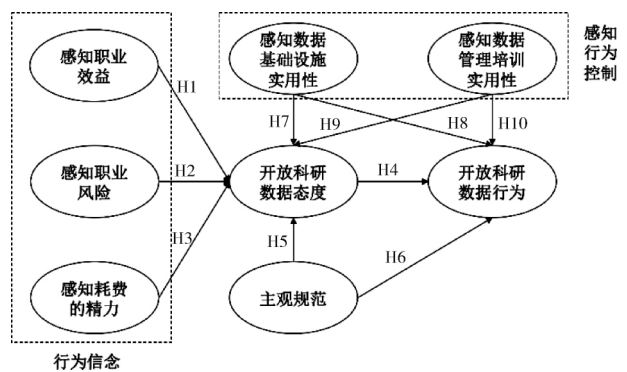


图 1 地球科学领域科研人员开放科研数据的行为影响因素研究模型与假设

2.1 感知职业效益

感知职业效益指科研人员认为 ORD 能为其职业发展带来的益处。例如,证明研究的可重复性、增加学术影响力、获得机构的认可与奖励以及更多的学术合作机会等。一些研究证实开放共享科研数据的论文比没有开放数据的论文被引频率更高^[22-23]。对许多科研人员来说,获得学术声誉和影响力是其参与 ORD 的重要动机。因此,

提出假设 1:

H1: 感知职业效益正向影响科研人员对开放科研数据的态度。

2.2 感知职业风险

感知职业风险指科研人员认为 ORD 可能为其职业生涯发展带来的风险。例如, 由于数据的复杂性和数据质量不佳导致的数据误读, 以及不遵守学科规范的数据不当使用等^[8]。这些潜在的风险可能会影响科研人员的学术声誉, 甚至影响他们的科研竞争优势, 科研人员倾向于不愿意冒险。因此, 提出假设 2:

H2: 感知职业风险负向影响科研人员对开放科研数据的态度。

2.3 感知耗费的精力

感知耗费的精力指科研人员感知到为参与 ORD 耗费的精力。例如, 科研人员在管理和存储可被重用的数据时, 需要花费时间和精力撰写细致的说明文档。Fecher 的调研发现数据共享最大的障碍之一是存储数据中耗费的大量精力^[7]。研究证实科研人员数据共享中耗费的精力会直接影响他们共享数据的态度^[9]。因此, 提出假设 3:

H3: 感知耗费的精力负向影响科研人员对开放科研数据的态度。

2.4 开放科研数据态度

行为态度是指个体对执行某一行为喜欢或不喜欢的程度^[19]。当科研人员对开放科研数据态度 (Attitude Toward Open Research Data, AT) 越积极, 认为 ORD 是有价值且愉悦的, 他们参与 ORD 的可能性就越大。如前所述, 本文舍去传统 TPB 模型中行为意向这一变量, 探究态度与行为的直接关系。因此, 提出假设 4:

H4: 科研人员对开放科研数据的态度正向影响开放科研数据行为。

2.5 主观规范

主观规范指科研人员所在的学术共同体中“重要他人”(Significant Others) 带来的规范性压力。例如, 科研合作者、专家、领导和重要学术团体的 ORD 相关经验和期望。国内学者的研究发现主观规范对科学数据共享意愿存在直接显著影响^[10-11]。Bock 等对知识共享行为意愿的研究也表明规范性信念会正向影响个体的态度与行为意愿^[24]。因此, 提出假设 5~6:

H5: 主观规范正向影响科研人员对开放科研数据的态度。

H6: 主观规范正向影响科研人员开放科研数据行为。

2.6 感知行为控制

感知行为控制 (包含感知数据基础设施实用性与感知数据管理培训实用性) 是影响科研人员的外部资源、

机会和条件。T. Darch 等认为基础设施和与数据相关的教育与培训有利于培养科研人员的共享观念^[25]。Kim 等的研究发现数据知识库的实用性对公共卫生领域科研人员数据共享行为有强烈的影响^[26]。数据密集型科研范式的转变向科研人员处理和分析海量数据的技能提出了更高的要求, 他们希望得到如数据管理计划的制定, 数据分析、处理和保存等的培训课程^[27]。因此, 提出假设 7~10:

H7: 感知数据基础设施实用性正向影响科研人员对开放科研数据的态度。

H8: 感知数据基础设施实用性正向影响科研人员开放科研数据行为。

H9: 感知数据管理培训实用性正向影响科研人员对开放科研数据的态度。

H10: 感知数据管理培训实用性正向影响科研人员开放科研数据行为。

2.7 开放科研数据行为

如前文对 ORD 的定义, 结合欧盟在更广泛传播和利用背景下对科学出版物和科研数据开放获取的描述^[28], 本文所指的开放科研数据行为 (Open Research Data Behavior, ORDB) 是: 将科研数据存储于数据知识库中 (公共数据存储平台或机构数据知识库), 或以出版物 (同行评议论文的附件材料或数据论文) 的形式开放共享, 并可被任何人免费获取和使用 (见图 2)。通过邮件、链接等形式共享给科研合作者等的个人数据共享行为在本研究中不属于 ORD 的形式。

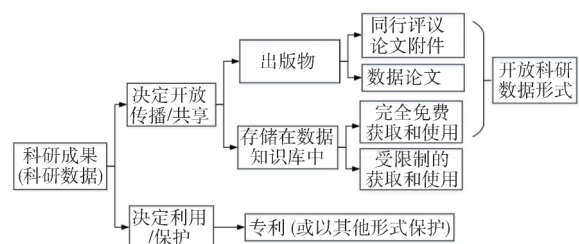


图2 科研人员开放科研数据形式

3 研究方法

3.1 问卷设计

调查问卷的整体设计遵循 TPB 研究的测量建议^[29], 各题项参考数据与知识共享行为研究成熟量表 (见表 1)。问卷中共有 28 个测量指标来解释潜变量, 所有观测变量采用 Likert 五分量表测量。预调研对象包括 40 名研究生与 6 名科研人员, 并根据反馈对问卷进行修正。对科研人员开放科研数据行为测量的题项结合前文对 ORD 行为的定义, 通过在问卷导语中的清晰说明使被受访者明确什么

是 ORD 行为。

表 1 潜变量与观测变量的测量项来源

| 潜变量 | 测量项来源 |
|-------------|--|
| 感知职业效益 | Bock 等 ^[24] |
| 感知职业风险 | Featherman 等 ^[30] , Kim 等 ^[21] |
| 感知耗费的精力 | Davis 等 ^[31] , Thompson 等 ^[32] , Kim 等 ^[21] |
| 感知数据设施实用性 | Kim 等 ^[21] , Taylor 等 ^[33] |
| 感知数据管理培训实用性 | Kim 等 ^[21] , Taylor 等 ^[33] |
| 主观规范 | 余玲 ^[34] , Bock 等 ^[24] |
| 开放科研数据态度 | Ajzen 等 ^[29] , Bock 等 ^[24] |
| 开放科研数据行为 | Kim 等 ^[21] |

3.2 数据收集

问卷设计完成后,采用抽样调查的方式,通过邮件和社交媒体向地球科学领域的科研人员发放。调查共持续 3 周(2018 年 6 月 19 日—7 月 10 日),回收有效问卷 314 份。在 SEM 分析中, Fornell 等^[35] 建议样本量至少为测量指标数量的 10 倍,本研究回收样本量满足 SEM 分析所需样本量。被调查者人口分布特征如表 2 所示。

表 2 被调查者人口特征分布

| | 类别 | 人数 | 占比 (%) |
|--------|----------|-------|--------|
| 性别 | 男 | 230 | 73.25 |
| | 女 | 84 | 26.75 |
| 年龄 (岁) | 20~25 | 114 | 36.31 |
| | 26~35 | 78 | 24.84 |
| | 36~45 | 61 | 19.43 |
| | 46~55 | 50 | 15.92 |
| | 56~65 | 10 | 3.18 |
| | 缺失 | 1 | 0.32 |
| 职位/身份 | 硕士研究生 | 134 | 42.68 |
| | 博士研究生 | 32 | 10.19 |
| | 博士后 | 4 | 1.27 |
| | 助教/实习研究员 | 3 | 0.96 |
| | 讲师/助理研究员 | 5 | 1.59 |
| | 副教授/副研究员 | 60 | 19.11 |
| 教授/研究员 | 76 | 24.20 | |
| 各部分总和 | | 314 | 100 |

4 数据分析与结果

本研究采用主成分形式结构方程模型 (Component-based SEM) 的偏最小二乘方法分析模型中各构念 (Construct) 之间的关系。PLS-SEM 对样本量要求相对较低,不要求样本数据服从正态分布,且可同时处理反映性指标和形成性指标,更适合探索性研究^[36],因此本研究采用了 PLS-SEM 的方法。数据分析的工具为 SmartPLS 3.2.7^[37],首先用测量模型评估测量指标反映同一个构念的程度;接着对结构模型各构念之间的假设关系进行显著性检验。

4.1 测量模型

测量模型反映了观测变量与潜变量的关联性,主要分

为反映性测量模型和形成性测量模型^[38]。开放科研数据行为这一潜变量由观测变量定义,测量指标为形成性指标 (Formative Indicator);其他潜变量的意义通过观测变量反映,测量指标为反映性指标 (Reflective Indicator)。

对测量模型的评估包含信度和效度检验。表 3 为各观测变量的因子负荷 (Factor Loading) 以及潜变量的组合信度 (Composite Reliability, CR) 和平均变异萃取量 (Average Variance Extracted, AVE)。内部一致性由 CR 值来评估,本研究中各潜变量的 CR 值在 0.798~0.905 之间,大于建议值 0.7,说明测量模型有较好内部一致性^[39]。效度检验包括对内容效度、收敛效度和区别效度的评估。本研究设计参考成熟量表,并进行了预调研,故具有较好内容效度。收敛效度由因子负荷、CR 值和 AVE 值评估,在探索性研究中,因子负荷最低接受值为 0.4,本研究各观测变量的因子负荷均大于 0.4,且潜变量 AVE 值均大于建议值 0.5,说明测量模型具有良好收敛效度^[40-41]。区别效度由各潜变量 AVE 值的平方根和相关系数检验,模型各潜变量的 AVE 值的平方根均大于相关系数 (见表 4),说明测量模型有很好的区别效度^[35]。

表 3 测量模型的信度与效度检验

| 潜变量 | 观测变量 | 因子负荷 | 组合信度 | 平均变异萃取量 |
|------|-------|-------|-------|---------|
| PADI | PADI1 | 0.602 | 0.804 | 0.583 |
| | PADI2 | 0.799 | | |
| | PADI3 | 0.866 | | |
| AT | AT1 | 0.827 | 0.905 | 0.761 |
| | AT2 | 0.916 | | |
| | AT3 | 0.871 | | |
| PCB | PCB1 | 0.808 | 0.878 | 0.645 |
| | PCB2 | 0.864 | | |
| | PCB3 | 0.722 | | |
| | PCB4 | 0.811 | | |
| PCR | PCR1 | 0.813 | 0.889 | 0.667 |
| | PCR2 | 0.843 | | |
| | PCR3 | 0.806 | | |
| | PCR4 | 0.806 | | |
| PDMT | PDMT1 | 0.437 | 0.798 | 0.588 |
| | PDMT2 | 0.913 | | |
| | PDMT3 | 0.859 | | |
| PE | PE1 | 0.801 | 0.878 | 0.645 |
| | PE2 | 0.760 | | |
| | PE3 | 0.762 | | |
| | PE4 | 0.883 | | |
| SN | SN1 | 0.811 | 0.865 | 0.682 |
| | SN2 | 0.833 | | |
| | SN3 | 0.833 | | |

4.2 结构模型

PLS-SEM 分析结果如图 3 所示,包括路径系数、 P 值、 R^2 和预测相关性 Q^2 。行为信念和感知行为控制对科

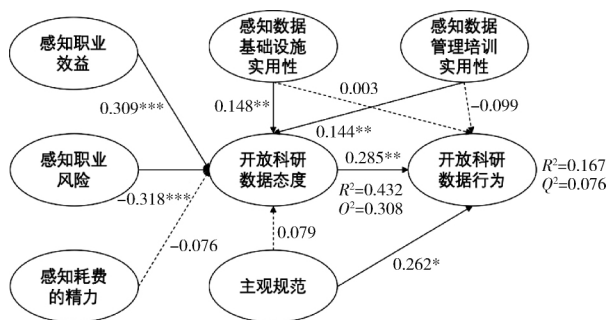
表4 AVE 值平方根及潜变量相关系数

| | PADI | AT | PCB | PCR | PDMT | PE | SN |
|------|--------|--------|--------|--------|-------|--------|-------|
| PADI | 0.764 | | | | | | |
| AT | 0.365 | 0.872 | | | | | |
| PCB | 0.359 | 0.511 | 0.803 | | | | |
| PCR | -0.049 | -0.420 | -0.166 | 0.817 | | | |
| PDMT | 0.396 | 0.335 | 0.425 | 0.069 | 0.767 | | |
| PE | -0.006 | -0.246 | -0.069 | 0.505 | 0.108 | 0.803 | |
| SN | 0.414 | 0.375 | 0.374 | -0.188 | 0.395 | -0.034 | 0.826 |

注：对角线为 AVE 值的平方根；非对角线值为潜变量之间的相关系数。

研人员态度的解释能力为 43.2%；Blindfolding 求得 Q^2 大于 0，表示模型具有预测相关性； Q^2 越大代表预测相关性越强^[39]，该模型具有良好的预测能力。

结果分析可得，10 个研究假设中有 6 个假设得到验证。科研人员开放科研数据态度对行为 ($p < 0.001$) 存在直接显著影响。感知职业效益对开放科研数据的态度 ($p < 0.001$) 存在正向显著影响；感知职业风险对态度 ($p < 0.001$) 存在负向显著影响；感知耗费的精力未对态度产生显著影响。主观规范对 ORD 行为 ($p < 0.05$) 存在积极显著影响，但对态度不产生显著影响。感知基础设施实用性和感知数据管理培训实用性对态度 ($p < 0.01$) 产生正向显著影响，但对行为不产生显著影响。



*** 表示 $p < 0.001$ ；** 表示 $p < 0.01$ ；

* 表示 $p < 0.05$ ， $Q^2 = 1 - SSE/SSO$

图3 结构模型假设检验结果

潜在自变量除了会对潜在因变量产生直接影响外，还可能通过其他变量对潜在因变量产生影响。使用自助法 (Bootstrapping) 检验开放科研数据态度的中介效应是否存在。分析结果显示 (见表 5)，PADI 和 PDMT 对行为存在

表5 开放科研数据态度的中介作用

| | Original Sample (O) | P 值 | 中介作用判定 |
|--------------|---------------------|-------|---------|
| PADI→AT→ORDB | 0.042 | 0.027 | 完全中介 |
| PDMT→AT→ORDB | 0.041 | 0.025 | 完全中介 |
| PCB→AT→ORDB | 0.088 | 0.004 | 部分中介 |
| PCR→AT→ORDB | -0.091 | 0.003 | 部分中介 |
| PE→AT→ORDB | -0.022 | 0.194 | 不存在中介效果 |
| SN→AT→ORDB | 0.023 | 0.127 | 不存在中介效果 |

间接影响，并完全通过开放科研数据态度这一中介变量实现；PCB 和 PCR 部分通过态度影响行为。

5 讨论

分析结果显示，科研人员对 ORD 的态度直接显著影响他们的实际行为，即当科研人员对 ORD 的态度越积极，他们参与的可能性就越大。这支持了其他学者指出的数据共享行为受数据共享态度直接影响的结论^[12 20-21]。而科研人员对 ORD 的态度受到感知职业效益、感知职业风险、感知数据基础设施实用性和感知数据管理培训实用性的影响。

感知职业效益和感知职业风险对科研人员 ORD 态度的影响高度显著，且间接影响实际行为。英国惠康基金 (Wellcome Trust) 对资助者的调研显示，增加论文被引、获得学术声誉和影响力是科研人员参与 ORD 的最主要个人动机，而最主要的障碍则来自于对数据滥用误读和科研成果被抢先出版的担忧^[42]。科研人员能获得的利于职业发展的好处越多 (如在职称评定和项目申请评估时承认科研人员为 ORD 做的努力与贡献)，担忧的潜在风险越少，他们对 ORD 的态度就越积极，进而参与的可能性就越大。实际上，一些潜在风险多存在于科学家的个人数据共享行为中，如通过邮件、硬盘等方式分享数据使科研成果被窃取。将科研数据存储于可信赖的数据知识库或以独立出版物、期刊的附加材料形式开放共享不会危及出版。因此，科研人员感知的职业风险可能不是实际的风险，需要加强他们对 ORD 内涵的认识，消除因不了解而产生的风险疑虑。

科研人员感知耗费的精力未对态度与行为产生显著影响。这可能是由于我国地球科学领域部分科研人员对 ORD 的认识不够深入。理论上，遵循 FAIR 原则的开放科研数据需要科研人员耗费时间和精力在数据准备工作中。例如，在项目初始阶段就制定详细的数据管理计划，为产出的数据集撰写细致的说明文档等。

感知数据基础设施实用性对科研人员态度有显著的积极影响，且通过态度间接影响行为。这与 Kim 等对美国 STEM 领域科研人员调查结果相似^[21]；而与孙晓燕对我国社会科学领域科研人员的调研结果不同^[12]。这说明外部资源的影响在不同学科表现不同。数据基础设施建设一直是我国地球科学领域科学数据管理和开放共享的重点。2019 年 1 月，地球大数据共享服务平台向全球发布，作为“地球大数据科学工程”的重大成果，该平台旨在通过建立数据、计算与服务为一体的数据共享系统，推动形成地球科学数据共享新模式^[43]。进一步扩大和完善现有的数据基础设施，能够积极促进科研人员参与开放科研

数据。

感知数据管理培训实用性同样对科研人员态度有直接显著的正向影响,且通过态度间接影响行为。荷兰莱顿大学科技评价中心(Leiden University's Centre for Science and Technology Studies, CWTS)和 Elsevier 对 1200 名科研人员如何共享数据的调查发现,科研人员缺少数据共享实践的原因之一是缺乏数据共享的相关培训^[17]。Williams 对拥有开放数据经验的农业科学家进行了深度访谈,受访者表示希望图书馆可以提供数据教育和培训以帮助培养具备数据管理能力的研究人员^[44]。开放科研数据并不意味着数据的简单公布,需要科研人员对数据进行恰当的管理使其易于被他人获取、理解、评估和重用。一些因数据质量不佳而产生的重用数据的风险也是由于缺乏良好的数据管理。因此,加强开展 ORD 相关教育与培训,如帮助科研人员制定数据管理计划,提供数据管理培训课程、科研数据重用的清单模板和可靠的数据知识库列表等,能有效发挥其与数据基础设施对科研人员的协同促进作用。

主观规范的压力对科研人员的态度没有明显影响,但对 ORD 行为存在显著影响。即科研合作者、同行、专家和学术社群等对 ORD 的倡导和实践经验分享(例如,通过线上线下的研讨会和论坛的形式宣传 ORD,分享数据共享最佳实践案例)会积极促进地球科学领域科研人员参与其中。

6 结束语

本文为探究科研人员开放科研数据的行为提供了可行的研究理论框架。通过实证研究得出,在激励地球科学领域的科研人员进一步参与开放科研数据时,积极转变科研评估方式,认可并奖励科研人员为 ORD 所做的努力与贡献,同时减少他们的风险疑虑是最为关键的因素;数据基础设施和数据管理培训是重要的外部促进因素,为科研人员提供了资源和机会;此外,学术共同体内的宣传推广和经验交流也不可或缺。这可为利益相关者在制定数据管理政策和提供数据管理服务等方面提供参考。未来研究可进一步对比该领域不同年龄科研人员的态度与行为;还可以考虑定性与定量相结合的方式深入研究影响科研人员开放科研数据行为的因素。□

参考文献

- [1] 国务院办公厅. 科学数据管理办法 [EB/OL]. [2018-07-18]. http://www.gov.cn/zhengce/content/2018-04/02/content_5279272.htm.
- [2] Higher Education Funding Council for England, Research Councils UK, Universities UK, et al. Concordat on open research data [EB/OL]. [2018-01-20]. <http://www.rcuk.ac.uk/documents/documents/concordatonopenresearchdata-pdf/>.
- [3] FORCE11. Guiding principles for findable, accessible, interoperable and re-usable data publishing version [EB/OL]. [2018-08-27]. <https://www.force11.org/fairprinciples>.
- [4] Digital Science, Figshare. The state of open data 2017 [EB/OL]. [2018-01-20]. https://figshare.com/articles/The_State_of_Open_Data_2017_Report_-_Infographic/5519155.
- [5] SENTURIASD. Guest editorial how to avoid the reviewer's axe: one editor's view [J]. *Journal of Microelectromechanical System*, 2003, 12 (3): 229-232.
- [6] European Commission. Open science policy platform recommendations [EB/OL]. [2018-10-27]. http://ec.europa.eu/research/openscience/pdf/integrated_advice_opspp_recommendations.pdf-view=fit&pagemode=none.
- [7] FECHER B, FRIESIKE S, HEBING M. What drives academic data sharing? [J]. *PLoS One*, 2015, 10 (2).
- [8] TENOPIR C, DALTON E D, ALLARD S, et al. Changes in data sharing and data reuse practices and perceptions among scientists world wide [J]. *PLoS One*, 2015, 10 (8).
- [9] 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.
- [10] 张晋朝. 高校科研人员科学数据获取意愿研究 [J]. *情报杂志*, 2013 (6): 70-75, 106.
- [11] 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.
- [12] 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]. *情报学报*, 2016 (10): 1062-1071.
- [13] 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 (18): 30-39.
- [14] 诸云强, 孙九林, 廖顺宝, 等. 地球系统科学数据共享研究与实践 [J]. *地球信息科学学报*, 2010, 12 (1): 8.
- [15] 刘闯, 郭华东, UHLIR P F, 等. 发展中国家数据出版基础设施与共享政策研究 [J]. *全球变化数据学报*, 2017, 1 (1): 8.
- [16] 王末, 王卷乐, 赫运涛. 地学数据共享网用户 Web 行为预测及数据推荐方法 [J]. *地球信息科学学报*, 2017 (5): 595-604.
- [17] CWTS, Elsevier. Open data-the researcher perspective [EB/OL]. [2018-09-28]. https://www.elsevier.com/_data/assets/pdf_file/0004/281920/Open-data-report.pdf.
- [18] AJZEN I. From intentions to action: a theory of planned behavior [M]. Springer, Berlin, Heidelberg, 1985.
- [19] 段文婷, 江光荣. 计划行为理论述评 [J]. *心理科学进展*, 2008: 6.
- [20] JEON S-H, KIM Y-G, JOONKOH. Individual, social, and organizational contexts for active knowledge sharing in communi-

- ties of practice [J]. *Expert Systems with Applications*, 2011, 38 (10): 9.
- [21] KIM Y, ZHANG P. Understanding data sharing behaviors of STEM researchers: the roles of attitudes, norms, and data repositories [J]. *Library & Information Science Research*, 2015, 37 (3): 189-200.
- [22] HENNEKEN E A, ACCOMAZZI A. Linking to data-effect on citation rates in Astronomy [EB/OL]. [2018-01-20]. <https://arxiv.org/abs/1111.3618>.
- [23] DORCHB F, DRACHEN T M, ELLEGAARD O. The data sharing advantage in astrophysics. [EB/OL]. [2018-01-20]. <https://arxiv.org/pdf/1511.02512.pdf>.
- [24] BOCK G-W, ZMUD R W, KIM Y-G, et al. Behavioral intention formation in knowledge sharing: Examining the roles of extrinsic motivators, social-psychological forces, and organizational climate [J]. *MIS Quarterly*, 2005, 29 (1): 87-111.
- [25] DARCH P T, KNOX E J M. Ethical perspectives on data and software sharing in the sciences: a research agenda [J]. *Library & Information Science Research*, 2017, 39 (4): 8.
- [26] KIM Y, KIM S. Institutional, motivational, and resource factors influencing health scientists' data-sharing behaviors [J]. *Journal of Scholarly Publishing*, 2015, 46 (4): 366-389.
- [27] HATE K, MEHERALLY S, MORE N S, et al. Sweat, skepticism, and uncharted territory: a qualitative study of opinions on data sharing among public health researchers and research participants in Mumbai, India [J]. *Journal of Empirical Research on Human Research Ethics*, 2015, 10 (3): 239-250.
- [28] European Commission. Guidelines to the rules on open access to scientific publications and open access to research data in Horizon 2020 Version3.2 [EB/OL]. [2017-09-28]. http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf.
- [29] AJZEN I. TPB questionnaires construction-constructing a theory of planned behavior questionnaire [EB/OL]. [2018-06-15]. https://people.umass.edu/ajzen/pdf/tpb_measurement.pdf.
- [30] FEATHERMAN M S, PAVLOU P A. Predicting e-services adoption: a perceived risk facets perspective [J]. *International Journal of Human-Computer Studies*, 2003, 59 (4): 451-474.
- [31] DAVIS F D. Perceived usefulness, perceived ease of use, and user acceptance of information technology [J]. *MIS Quarterly*, 1989, 13 (3): 319-340.
- [32] THOMPSON R L, HIGGINS C A, HOWELL J M. Personal computing: toward a conceptual model of utilization [J]. *MIS Quarterly*, 1991, 15 (1): 125-143.
- [33] TAYLOR S, TODD P A. Understanding information technology usage: a test of competing models [J]. *Information Systems Research*, 1995, 6 (2): 144-176.
- [34] 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.
- [35] FORNELL C, LARCKER D F. Evaluating structural equation models with unobservable variables and measurement error [J]. *Journal of Marketing Research*, 1981, 18 (1): 12.
- [36] HAIR J F, SARSTEDT M, RINGLE C M, et al. An assessment of the use of partial least squares structural equation modeling in marketing research [J]. *Journal of the Academy of Marketing Science*, 2012, 40 (3): 20.
- [37] RINGLE C M, WENDE S, BECKER J-M. "SmartPLS 3." boenningstedt: SmartPLS GmbH [EB/OL]. [2018-11-30]. <http://www.smartpls.com/>.
- [38] JARVIS C B, MACKENZIE S B, PODSAKOF P M. A critical review of construct indicators and measurement model misspecification in marketing and consumer research [J]. *Journal of Consumer Research*, 2003, 30: 20.
- [39] CHIN W W. The partial least squares approach to structural equation modeling [J]. *Modern Methods for Business Research*, 1988: 42.
- [40] HULLAND J. Use of partial least squares (PLS) in strategic management research: a review of four recent studies [J]. *Strategic Management Journal*, 1999, 20 (2): 10.
- [41] HAIR J F, BLACK W, BABIN B, et al. *Multivariate data analysis* (6th ed) [M]. Pearson, 2005.
- [42] VAN V, KNIGHT G, VLAD A, et al. Survey of wellcome researchers and their attitudes to open research [EB/OL]. [2018-01-22]. https://figshare.com/articles/Survey_of_Wellcome_researchers_and_their_attitudes_to_open_research/4055448/1.
- [43] 新华社. 中科院发布地球大数据共享服务平台 [EB/OL]. [2019-01-15]. http://www.gov.cn/xinwen/2019-01/15/content_5358075.htm?_zbs_baidu_bk.
- [44] WILLIAMS S C. Data sharing interviews with crop sciences faculty: why they share data and how the library can help [J]. *Issues in Science and Technology Librarianship*, 2013: 15.
- 作者简介: 包秦雯 (ORCID: 0000-0003-0887-7774), 女, 1993年生, 硕士生。研究方向: 用户研究与信息服务。顾立平 (ORCID: 0000-0002-2284-3856, 通讯作者), 男, 1978年生, 博士, 副研究员。研究方向: 科技信息政策, 用户研究与信息服务。张潇月 (ORCID: 0000-0001-8605-6491), 女, 1994年生, 硕士生。研究方向: 用户研究与信息服务。
- 作者贡献声明: 包秦雯, 设计研究方案, 数据收集与分析, 起草论文。顾立平, 提供研究思路, 辅助问卷量表设计, 修改论文。张潇月, 论文修改与校对, 辅助问卷调查。
- 录用日期: 2018-11-13