科研机构科学数据治理路径初探

彭洁
（中国科学技术信息研究所，北京 100038）

摘要：科研机构作为科学数据生产和管理的重要部门，有效实施科研机构科学数据治理是实现其科学数据高效利用的重要手段。本文以数据治理框架为基础，构建科研机构科学数据治理模型，并以某科技信息研究机构数据治理为例进行分析，为落实《科学数据管理办法》提供参考。

关键词：科研机构；科学数据；数据治理；数据服务
中图分类号：G350　DOI：10.3772/j.issn.1673-2286.2021.12.003
引文格式：彭洁. 科研机构科学数据治理路径初探[J]. 数字图书馆论坛，2021（12）：15-21.

科学数据是科技创新的重要战略资源。科学数据主要包括在自然科学、工程技术科学等领域，通过基础研究、应用研究、实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]。它既是支撑科学研究的重要基础，也是科学研究的重要产物和成果；既有一般科学数据和大数据的特征，也有其自身独有的特征。

科研机构是科学数据生产和利用的重要主体，是科学研究的重要执行基地，在研究过程中产生大量有价值的数据。对科研机构科学数据高效管理是落实《科学数据管理办法》的基本要求，也是提升科学数据开放共享水平、发挥科学数据价值的重要内容。数据治理（data governance）是各类组织对其数据使用的一整套管理行为，本研究主要基于数据治理理念，针对当前科学数据管理和利用过程中的问题，提出科研机构开展科学数据治理的基本框架和实施路径。

1 相关研究

随着信息技术的发展，数据量日益增加，数据治理成为信息研究的重要领域。Watson等[2]最早提出“数据仓库治理”的概念，开始关注数据治理这一研究领域。许多学者对数据治理的内涵进行了深入探讨：Wende[3]认为数据治理是鼓励理想的使用数据行为的决策权和责任框架；Weber等[4]认为数据治理作为实施数据责任的通用方法，适合所有数据和组织的需求；Thomas[5]认为数据治理是对于信息相关流程的决策权利和责任制度，按照商定的模型执行，描述谁可以采取什么行动。可以说，有效的数据治理已被认为是获取数据使用价值的关键[6]。本研究认为，数据治理是对于数据相关流程的决策权利和责任制度，按照商定的模型执行，描述谁可以采取什么行动。数据治理的核心是数据资产管理的决策权分配和职责分工。

数据治理需要巩固和结合整个公司、机构的数据，其重要性得到了从业人员的认可，并且认为数据治理一直也必将是企业信息管理的新趋势。在此基础上，在许多重要会议上数据治理都成为重要主题，如TDWI（The Data Warehousing Institute）World Conference、DAMA（Data Management Association）International Symposium、DG（Data Governance）Annual Conference和MDM（Master Data Management）Summit。《DAMA数据管理知识体系指南》认为数据治理是对数据资产管理行使权力和控制的活动集合[7]。2004年IBM制定了包括四大领域11个要素的数据治理框架和方法来指导数据治理工作的开展。近年来，国内外研究逐渐增多，国外有Wende[3]等进行了数据治理模型方面的研究；国内来看有李振[8]、徐建忠[9]、孟晓峰[10]等分别对大数据

治理的模式、方法及隐私保护进行了相关研究。

总体来看，数据治理是一个新兴的热点研究领域，是数据管理的补充，是IT治理的高级阶段和内容深化。数据治理是鼓励数据创造、使用、存储、归档和删除而涉及的组织结构、规则、决策权及责任。当前已经有许多学术组织、企业和学者提出了一些数据治理模型、重点内容、工作原则、角色职责等。将数据治理理念用于科研机构科学数据管理领域的研究较少，有针对科研机构数据政策[11-12]、管理现状分析[13]等尚未成熟的治理体系和方法路径。

2 科研机构数据治理模型及实施路径

科研院所是由国家或部门根据经济社会发展需要建立并资助，由占有一定空间与设备条件的各级各类科研人员所组成的科研工作单位，是国家科技创新体系的基本力量[14]。中国科研机构众多，在物理、地理、化学、数学、机械、电子等理工科及人文和社会科学、医学、农学、艺术学等各个方面都有涉及。截至目前，中国有各大科研机构近1 300所，涵盖中国科研的各个方面。在科研机构的数据库建设方面，几乎每个科研机构都有自己的专业数据库，只是在专业程度、规模、建设层次和水平等方面有着较大的差异。

长期以来，科研院所集聚了大量国家投入形成的科技资源。其中，科学数据库是最重要的科技资源之一，这些分布在科研院所内部的数据库是一笔巨大的财富，它们既是科学研究的基础，也是科学研究的成果，在大数据时代，更是一种重要的国家战略资源，同时具有很高的经济价值，应该得到充分的共享和使用。但是，分布在科研院所内部的大量科学数据库的共享一直是难点，迄今未得到有效解决，严重制约了中国科研和科研经费使用效率的提高。科研院所内部数据库的共享和利用是一个复杂的系统问题，既涉及个人和单位的利益，也涉及知识产权、机制和技术等诸多因素，需要对这些因素进行研究，并以政策的形式对诸多因素进行协调，开展数据治理。

2.1 治理目标

在数据治理中，非常关键的是进行数据治理总体目标与蓝图规划。对于数据治理的核心目标，一是增加决策制定的一致性和保密性，二是减少调节的风险，三是提升数据安全性，四是将数据产出最大化。其包含组织结构和过程来确保组织的数据资产维持并拓展组织的战略和目标。Rifaie等[15]指出数据治理有4个目标：数据价值和对齐（data value and alignment）、任责/问责（accountability）、绩效测评（performance measurement）和风险管理（risk management）。

科研机构开展数据治理的目的是实现科研机构内部已有和新增的多源异构科学数据的有效共享与高效利用，并能够根据机构的业务定位，面向不同场景，提供基于数据融合的各类服务。

（1）在数据使用意愿方面，需要制定形成机构内部的数据管理办法，保证机构内部各业务部门、研究团队、科研人员能够主动地汇聚科学数据，建立汇交数据的各种奖励措施，保证能够按照机构战略实施数据共享，促进数据利用，形成一种愿共享、想被用、促增值的良好文化氛围。

（2）在数据质量方面，需要形成对各类科学数据元数据质量、数据质量、数据汇交质量等管理的标准，并根据标准进行比对检查，保证提交数据的准确性、一致性和完整性。同时建立基于科研机构业务的数据模型，形成数据的关联。建立机构内部科学数据获取、加工、组织、关联、汇聚、共享服务等全生命周期的管理规范、办法等，保证数据的有效管理。

（3）在数据使用效益方面，需要根据信息技术的发展，建立适用于机构内部各类数据汇聚和保存的数据仓储和数据服务门户，利用先进的IT技术和数据技术，保证服务响应周期内面向机构内用户共享和利用，保持用户服务的可拓展性。

2.2 治理模型

Moseley[16]指出数据治理的5个核心内容包括政策（policies，有关组织层面数据治理操作程序、决策制定和数据管理活动的方向）、过程（processes，详细的过程定义了步骤和产出）、业务规则（business rules，用于实施数据治理的组织政策和决策制定过程）、人员（people，具有专门职责的角色，直接度量和对数据治理产生贡献）、技术（technologies，数据质量管理工具、中间件、工作流和主数据管理软件等）。

对于数据治理的体系框架，由于机构背景、治理目的、研究视角等不同，不同的机构和研究者提出了不同的数据治理框架，例如：Thomas[17]提出的重点在于关

注数据仓库和商业智能的治理框架；Weber等[4]提出的面向企业的角色、决策任务、责任三位一体治理框架；Khatri等[18]提出强调数据原则和数据生命周期的治理框架。此外，还有IBM数据治理模型（支撑域、核心域、促成因素和成果4个层级11个要素）[19]、《数据治理白皮书》的治理框架（原则、范围、实施和评估）[20]、ISACA的体系框架（提出数据治理的5项基本原则）、Gartner数据管理参考架构（规范、规划、构建和运行）等。

综上，本研究认为，对于科研机构来讲，数据治理可采用《数据治理白皮书》框架，包括3个维度：原则、范围、实施与评估（见图1）。原则维度给出了科研机构数据治理工作所遵循的、首要的和基本的指导性法则，即治理战略与机构战略一致、治理行动合规、服务绩效提升。范围维度（决策域）描述了科研机构数据治理的关键域，即科学数据治理决策层应该在哪些关键领域作出决策。治理目标是在现有数据资源基础上，实现科研机构数据的高效利用与共享，在这个过程中，需要对实现这些目标的各个决策域开展工作，并进行评估、监督和指导，以保证这些流程和行动合规。在数据治理实施过程中，需要明确参与数据治理各个角色的权责范围，根据目前科学数据资源及利用现状，开展现状评估，梳理确定开展数据治理的关键重点领域，并将关键重点任务分配给各个角色负责实施，在此过程中定期开展治理状况的评估与监督，根据评估反馈不断完善治理实施路径，以达到科学数据高效利用状态。

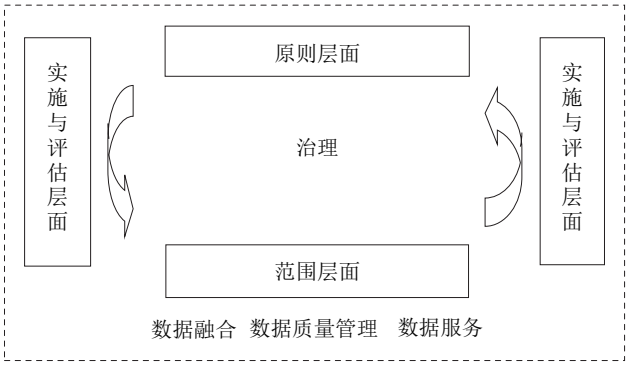

图1　科研机构科学数据治理模型

科研机构开展科学数据治理的基本方法和步骤包括以下4个方面。

（1）了解科研机构科学研究及其产生和汇聚的科学数据现状，分析在其利用过程中存在的问题，形成科学数据治理的目标（确立目标和指导原则）。

（2）根据科研机构特点，建立科学数据治理的组织体系，对涉及的各个角色进行职责分配，并确立奖惩机制（组织保障）。

（3）针对治理目标，开展治理活动（决策域）。重点形成科学数据治理的元数据整合汇聚框架和质量提升策略，构建融合科研机构数据的数据模型和数据应用服务门户架构，设计科学数据共享及应用服务的主要场景和服务内容。

（4）开展治理成熟度评估和科学数据资产审计，反馈治理效果，并不断优化针对各个决策域的治理活动（评估、监督和指导）。

3　科研机构数据治理案例

本研究以科技信息研究机构为例，按照本研究提出的科研机构科学数据治理模型，对其科学数据治理现状进行剖析，重点针对数据治理中的组织体系建设以及治理过程中的元数据集成、场景化的数据服务等关键问题进行案例分析。

3.1　科学数据治理现状及问题分析

某科技信息研究机构是一个面向科技管理决策提供数据和信息支撑的智库型研究机构，同时也是为创新主体提供科技知识服务的服务机构，建有科技文献数据库、专利数据库、科技人才数据库、科研机构数据库等近50个数据库，由下属各业务部门分别建设。该机构科学数据建设和利用现状如下。

（1）缺乏支撑数据高效利用的政策体系。该机构制定出台了《自建数据库资源使用暂行管理办法》，主要对数据库资源在机构内各部门间共享的审批流程、共享数据的使用、数据共享过程中的责任划分做出了规定。但随着时代的进步，科学数据共享资源标准的提高，该办法一方面无法完全满足科研工作者对于数据共享的需求；另一方面无法充分调动相关部门对于数据利用的积极性，在一定程度上制约了科研工作的高效开展。

（2）数据资源标准化程度不高。同一主题类型的数据库存在数据标准不统一、相互冲突，而数据库建设标准是实现数据共享利用的前提条件，同时也是科学数据平台建设的基础和支撑。只有将原始数据的来源、表现方式、存储形式等量化并制定统一标准，才能更好地进行数据的融合与共建共享。例如，同一研究对象，

由于采用的相关理论、研究手段、研究范围等没有统一的标准作为规范，在理论指导层面不尽相同，各个研究机构调研数据呈现方式也不同，同一研究出现多种数据结果等。这就导致数据共享时数据内涵和对数据的理解出现混淆，为数据利用和深入处理的后期工作带来应用困难和不确定性等问题。缺乏共享建设标准直接导致数据顶层设计难以进行，现存的数据治理与共享体系建设需求、现实科研和科研管理需求的匹配性有待提高。

（3）缺少统一的数据资源目录。有10余个数据库是以在线服务系统的形式体现，其余则是各种离线的异构数据库，存储于机构内部各科研人员机器或相关服务器上；已经建设的一些数据系统中，有时会出现数据无法浏览、无法下载、数据链接不存在等情况，大数据平台的稳定性和可通达性较低，甚至有些数据服务平台处于后台服务和更新停滞的状态，没有形成基于全局发展战略和业务目标的数据资源目录，给科研人员及数据使用者造成不利影响。

（4）数据资源利用的利益分配机制缺失。对于数据开放与共享方面，大多数部门都有使用其他部门数据的现实业务需求，认为目前数据资源共享利用效果不佳的主因是利益分配及共享机制问题，各部门希望并接受以利益分成的形式进行数据资源共享。

为了解该机构科学数据治理的现状和理想目标间的差距，以便给科学数据治理领导决策层提供决策依据，本研究开展了科学数据治理成熟度测评。本次测评选择6个内设部门中参与数据采集加工、数据管理和服务的项目团队、数据典型用户等开展内部访谈和调研，借鉴卡内基梅隆大学的能力成熟度集成模型（Capability Maturity Model Integration，CMMI）的过程改进方法，从领导、范围、测量和管理4个维度开展科学数据治理的成熟度评测，分别将成熟度的6个阶段对应1~6的数值进行打分，然后通过综合平均计算，得到各维度的平均评测值（见表1）。

表1 各部门数据治理成熟度测评情况

数据库所属部门	领 导	范 围	测 量	管 理	合 计
部门A	2.50	1.50	2.20	3.20	9.40
部门B	1.60	1.40	2.30	3.10	8.40
部门C	1.60	1.60	2.40	3.40	9.00
部门D	2.40	1.80	2.50	4.50	11.20
部门E	1.80	1.40	2.20	3.20	8.60
部门F	2.10	1.50	2.20	3.50	9.30

针对数据治理的领导、范围、测量和管理4个维度对机构内部业务部门进行测度后，可以发现两方面问题。一是各部门数据治理成熟度不一。有些部门形成特色数据库产品和系统平台，具有较好的数据治理水平。大多数部门业务与数据库建设、系统平台搭建关联性不大，相关业务系统应用单一，缺少成熟稳定的治理应用，表现欠佳。二是从单个数据库来看，有的数据库在资金支持、人员配备、数据体系建设及服务等方面都极具战略性，数据治理成熟度较高。部分数据库面临共享程度低、使用率低、数据更新缓慢或停滞等各种状态，数据治理成熟度较低。

3.2 科学数据治理的组织体系建设

数据治理的重要内容是确保所有权、客观性、有效的决策和行动，因此需要定义数据治理的角色与职责[21]。主要的治理角色包括数据治理委员会、数据治理指导组、数据治理工作组及成员等。按照治理需求建立科学数据治理组织体系，主要是保证各个涉及科学数据的角色能够各司其职，相互协调，共同完成治理任务。其中数据治理委员会是研究机构内开展数据治理工作的最高决策组织，主要成员包括机构领导、各业务部门领导和外部专家顾问。数据治理指导组受数据治理委员会的领导，按照治理实施要求，可以划分为不同的工作团队，分别负责数据质量管理、数据建模、数据隐私和安全保密、数据服务等工作内容和方案制定，并提交数据治理委员会审议，成员由该机构熟悉上述业务内容的工作人员组成。数据治理工作组的每项具体工作一般由数据治理指导组成员牵头推进，各业务部门有至少一位熟悉业务的员工参与其中。各个角色的主要职责如下。

（1）数据治理委员会。其主要职责是从战略角度来统筹和规划本机构数据治理内容、方法和要求，明确数据治理各角色在数据使用和管理过程中的流程及职责。根据相应的标准和原则开展对各角色工作内容的监督评估，保证数据治理顺利实施。

（2）数据治理指导组。其主要职责是指导和协调本机构的数据治理工作，制定本机构的数据管理办法、数据治理考核标准、流程及奖惩措施等，进行数据治理过程的监测与评估。

（3）数据治理工作组。数据治理工作组包括数据业务分析员、数据质量分析员、数据管理员、系统开发

人员等。不同角色在数据治理过程中承担的职责不同，但相互之间需要协同工作。工作组的主要职责包括：对数据质量进行评估、提升；在数据管理流程中根据相关要求进行元数据管理、数据模型构建、数据产品设计等数据处理工作；根据数据治理指导组意见修正工作内容；提供数据服务，作为完善数据治理内容的重要参考。

3.3 科学数据治理实施

科学数据治理实施包括4部分内容，分别是制定数据内容标准、完善科学数据技术支撑、构建科研机构科学数据开放服务平台、构建科研机构数据应用体系。

（1）制定元数据标准、数据标准术语、交换格式标准、数据转换标准和数据参考模型等数据内容标准，制定数据整理标准和分析标准等科研数据处理标准，制定科学数据平台架构标准，建设数据库开放应用接口标准、数据可视化工具标准等。

（2）完善科学数据技术支撑。包括科学数据的来源与筛选技术，科研数据资源编目技术和数据格式规范技术等。

（3）构建科研机构科学数据开放服务平台，明确科研机构层面大数据平台和服务平台的建设任务，厘清服务平台与科研机构已有各个业务数据平台和系统之间的衔接配合关系，制定各部门数据范围边界，明晰各部门数据管理的义务和权利。推动与外部各级科研机构信息系统和公共数据互联共享、制度对接和协作协同。

（4）构建科研机构数据应用体系。围绕科学数据服务水平提高，建立科研机构数据产生、交换、运算，以及数据融合、平台建设、数据治理等要素的监测体系；建立科研机构在人员配备、数据采集、后期采编、数据传送等方面的运行机制。

数据治理主要业务流程如下。

（1）获取和创建流程。数据库及相关子库是系统的原始资源，首先对这些数据库采用不同的技术手段导入中间数据库，导入过程中控制需要监测和审计导入的字段。对中间数据库进行进一步深入融合和分析。首先进行该机构人员、项目、机构信息的比对、去重、标识，形成主数据库，在此过程中可能需要借助文本挖掘技术。通过主数据库匹配关联策略，生成关联关系数据库；对中间数据库的成果信息进行提取比对去重，可以生成成果数据库。

（2）其他应用调用流程。其他应用调用数据流程包括应用系统提交资源访问请求，进行资源需求及用户标识（根据用户信息判定是否给其相应资源）；根据请求情况，分别向主数据库、关联关系数据库、成果数据库提交请求；各数据库返回相应结果，按照约定的格式提供给其他应用。

（3）科学数据目录体系。科学数据目录体系主要包括科研机构内所有的数据目录及可以用来汇聚的数据资源目录，从而帮助未来使用该机构科学数据服务平台的用户发现和定位数据资源。

3.4 基于数据治理的数据服务

3.4.1 元数据集成体系

根据该科研机构数据特点，元数据集成设计分为数据抽象层和业务逻辑层，数据抽象层为核心数据模型，包括人员信息、机构信息、项目信息等基本实体，也包括出版物、专利、产品等产出实体，还有项目资金、语言、货币等二级实体；业务逻辑层为核心数据模型的衍生层，选取核心数据模型的一部分来完成一项业务数据存储，根据业务方向的不同，划分为人才数据库、机构数据库、项目数据库等不同类型的业务域，梳理形成各类实体相互关联的E-R图[22]（见图2）。

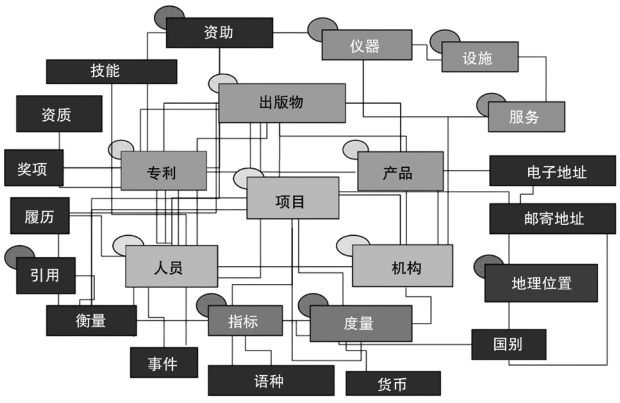

图2 数据元素抽象图

3.4.2 数据服务场景示例

在该机构各个业务部门充分数据共享的前提下，结合关于科技多要素融合后的模型查询和分析需求，应用知识图谱的技术，将融合后的模型转换为具有图

结构的RDF数据，使得科技各要素之间的关系显性化，实现基于此类数据的查询和基于关系的画像分析场景，进一步挖掘融合后的各要素数据存在的价值。

以科技人员数据为例，将人物基本信息，科研成果信息，与其他人、机构、项目、科研成果等的关系图谱信息，以及分析统计信息和人物标签、合作关系展现等汇聚成科研人员画像。在科研人员基本信息和关系图谱构建的过程中，按照科学数据治理中的隐私相关内容，依照脱敏的规范要求，对可能涉及的隐私进行脱敏处理。

以研究人员为中心点，对其和其他实体之间的关系进行图谱展示，其中包含的元素有机构、人员、项目、科研成果等，可以看到人员和人员、人员和机构、人员和项目、人员和专利、人员和期刊等之间的关联信息，还可以将本身有联系的信息汇聚在一起显示，可以更加清晰地看到他们之间的联系（见图3）。

图3 人物关系关联图

通过这种元数据集成和可视化展示，可以更加精准帮助科技信息研究机构开展科研人员成长规律分析、科研团队特征分析、科研人员评价等。

基于该信息研究的主要服务内容和对象，通过对数据资源管理现状的问题剖析，开展数据治理，形成针对具体场景的各类数据服务，有效提升了数据服务效能。从科研机构数据治理实施来看，还应循环进行科学数据治理成熟度测评，针对评测结果不断优化治理实施，即不断进行元数据管理与集成、数据建模、服务产品开发，从而使得数据治理保持可持续发展。

4 结语

科研机构是我国科技创新活动的基本单元，围绕数据现状、治理组织和权责体系等主题，进行清晰的数据治理规划，开展数据治理，符合当前科学数据高效管理和利用的需求。通过数据治理，对科研机构科学数据采集、加工、处理、交流、传递、共享、利用等过程进行重塑，进而实现数据整体价值增值。同时，科研机构的数据治理也应结合本机构业务特点、数据特征等灵活进行治理范围和内容的优化，保证治理效果。

参考文献

[1] 国务院办公厅关于印发科学数据管理办法的通知[EB/OL]. [2021-11-05]. http://www.gov.cn/zhengce/content/2018-04/02/content_5279272.htm.

[2] WATSON H J, FULLER C, ARIYACHANDRA T. Data warehouse governance: Best practices at Blue Cross and Blue Shield of North Carolina [J]. Decision Support Systems, 2004, 38(3): 435-450.

[3] WENDE K. A Model for Data Governance Organising

Accountabilities for Data Quality Management [C] // Australasian Conference on Information Systems. OAI, 2007.

[4] WEBER K, OTTO B, HUBERT S. One size does not fit all - a contingency approach to data governance [J]. Journal of Data & Information Quality, 2009, 1 (1): 1-27.

[5] THOMAS G. The DGI Data Governance Framework [EB/OL]. [2021-11-26]. https://www.docin.com/p-1718904829.html.

[6] KAREL R. Data governance: what works and what doesn't [EB/OL]. [2021-11-26]. https://www.mendeley.com/research/data-governance-works-doesnt/.

[7] 李善斌. 银行业数据治理分析 [J]. 金融科技时代, 2017 (4): 36-38.

[8] 李振, 鲍宗豪. "云治理": 大数据时代社会治理的新模式 [J]. 天津社会科学, 2015 (3): 62-67.

[9] 徐建忠, 张亮, 李娇娇. 数据智能分类技术在数据治理中的应用研究 [J]. 信息安全与通信保密, 2016 (6): 88-90.

[10] 孟小峰, 张啸剑. 大数据隐私管理 [J]. 计算机研究与发展, 2015, 52 (2): 265-281.

[11] 郭春霞. 科研机构数据管理与共享政策研究 [J]. 情报杂志, 2015, 34 (8): 147-151.

[12] 薛秋红, 徐慧芳. 西方国家科研机构科学数据管理政策要素研究 [J]. 情报理论与实践, 2021, 44 (7): 124, 191-196.

[13] 邢文明, 杨玲. 我国科研机构科研数据管理现状调研 [J]. 数字图书馆论坛, 2018 (12): 27-33.

[14] 张雅群. 公共科研机构技术商业化能力影响因素研究 [D]. 合肥: 中国科技大学, 2015.

[15] RIFAIE M, ALHAJJ R, RIDLEY M. Data governance strategy: a key issue in building enterprise data warehouse [C] // iiWAS'2009. DBLP, 2009: 587-591.

[16] MOSELEY M. Improving Data Quality through Agile Data Governance [EB/OL]. [2021-11-26]. http://tdwi.org/~/media/5A2DC52A7C154F51ABFE165F9821774C.pdf.

[17] THOMAS G. Alpha Males and Data Disaster [M]. Boston: Brass Cannon Press, 2006.

[18] KHATRI V, BROWN C V. Designing data governance [J]. Communications of the Acm, 2010, 53 (1): 148-152.

[19] 谈谈典型的数据治理体系框架 [EB/OL]. [2021-12-24]. http://www.360doc.com/content/21/1224/14/78237952_1010147429.shtml.

[20] 张明英, 潘蓉. 《数据治理白皮书》国际标准研究报告要点解读 [J]. 信息技术与标准化, 2015 (6): 54-57.

[21] STOCKDALE S. Deconstructing data governance [EB/OL]. [2021-11-26]. http:digitalrepository.unm.edu/cgi/viewcontent.cgi?article=1010&context=hslic-posters-presentations.

[22] CERIF in Brief [EB/OL]. [2021-11-26]. https://www.eurocris.org/eurocris_archive/cerifsupport.org/cerif-in-brief/index.html.

作者简介

彭洁, 女, 1965年生, 博士, 研究员, 研究方向: 信息资源管理, E-mail: pengj@isitc.ac.cn。

An Approach to Governance of Scientific Data in Scientific Research Institutions

PENG Jie
(Institute of Scientific and Technical Information of China, Beijing 100038, P. R. China)

Abstract: Scientific research institutions is the important branch of scientific data production and management, effective scientific data governance is the important means that makes its scientific data to be used efficiently. Based on the data governance framework, this study constructed a scientific data governance model for scientific research institutions, and carried out a case study of data governance in a scientific information research institution, it will provide an important reference for the implementation of *scientific Data Management Measures*.

Keywords: Scientific Research Institutions; Scientific Data; Data Governance; Data Services

（收稿日期：2021-11-11）