

治理框架视域下的数据治理研究

周振国

(中华人民共和国农业农村部 机关服务局, 北京 100125)

摘要: [目的 / 意义] 数据治理是提升数据价值的基础, 是数据时代的研究热点。对当前数据治理的研究成果进行系统梳理, 有助于更好地促进数据治理的理论研究与电子信息资源管理等方面实践的开展。[方法 / 过程] 运用文献分析法, 以数据治理框架为分析视角, 从治理目标、治理主体、治理客体、治理内容、治理过程、治理手段等 6 个维度系统分析了数据治理的研究进展。[结果 / 结论] 数据治理核心要素中, 对治理目标、治理主体、治理手段的认识较为一致, 对治理内容关键点的认识需要进一步统一, 对治理客体的研究仍需细化。数据治理过程中, 更好地区分数据管理和数据治理, 明确数据治理定位, 凸显数据治理重点与核心, 是后续研究需要注意的问题。

关键词: 数据治理; 信息资源; 治理框架; 述评

中图分类号: G252.8

文献标识码: A

文章编号: 1002-1248 (2020) 07-0057-06

引用本文: 周振国. 治理框架视域下的数据治理研究[J]. 农业图书情报学报, 2020, 32(7): 57-62.

Data Governance Research from the Perspective of Governance Framework

ZHOU Zhenguo

(Service Bureau of the Ministry of agriculture and rural affairs, Beijing 100125)

Abstract: [Purpose / Significance] As the basis of improving data value, data governance is a research hotspot in the data age. The systematic review of the current research into data governance will help to promote the theoretical research and the practice of electronic information resource management. [Method / Process] Using the methods of literature analysis, the paper systematically analyzes the research progress of data governance from six dimensions: the goal, the subject, the object, the content, process and means of data governance. [Results / Conclusions] Among the core elements of data governance, researchers have the consistent understanding of the goal, the subject and the means of data governance, while their understanding of the key points of the content of data governance needs to be further unified. The research into the object of data governance needs to be specified. In the future research into data governance, researchers should further distinguish between data management and data governance, clarify the posi

收稿日期: 2020-03-31

基金项目: 国家社科基金项目“基于信息治理视阈的档案服务质量优化控制研究”(16BTQ091)

作者简介: 周振国 (1989-), 男, 硕士研究生, 馆员, 研究方向为档案信息化

tioning of data governance, and place an emphasis on the focus and core content of data governance.

Keywords: data governance; information resource; governance framework; review

1 引言

数据治理是对数据资产管理行使权力和控制的活动集合。具体而言,是以实现数据标准化,提升数据价值为目标,以健全数据组织为保障,以数据过程管控为手段的一种数据管理体系。关于数据治理的研究,已经形成较为丰富的研究成果。张宁等采用文献调研方法,对数据治理的概念、体系、内容、应用实践等进行系统梳理,指出“现有研究集中于‘框架模型的设计’‘价值的探讨’和‘不同领域的应用’等,存在‘实证研究较少’‘数据治理框架模型设计欠缺优化’等问题,‘框架体系’‘政策标准’‘成熟度模型’‘数据质量’是未来研究应关注的重点领域”^[1]。刘桂锋等采用文献调研和网络调查方法,深入探析国外数据治理内涵、要素、模型与框架等,指出现有研究主要集中于理论探索、模型框架和实践应用3个方面^[2]。孙嘉睿采用文献调研法和文本分析法,系统梳理和分析了国内数据治理的相关文献,指出“当前研究主要集中在数据治理的体系架构、数据治理的有效保障以及数据治理在具体业务领域的实践探索等方面,数据治理要素梳理、数据治理标准构建和数据治理技术的应用是值得进一步研究的领域”^[3]。

尽管诸多学者从不同方面对数据治理研究做了大量有益工作,但总的来看,分析框架缺乏较高的系统性,形成对数据治理研究的全面认识有较大难度。且对比数据治理研究本身,目前对于数据治理研究现状相关整体性评价和分析较为缺乏,为数不多的有关研究成果,多着眼于企业数据、政府数据等中观业务层面。但与此同时,随着数据治理研究的推进,数据治理的相关成果不断涌现、研究内容不断丰富,也迫切需要从整体上梳理数据治理的研究成果,以更好地完善研究路径和内容,促进数据治理研究的开展。对此,笔者从数据治理的顶层框架设计着手,分析一个完整

的数据治理可能涉及的主要要素及其相互关系,并以此为基础系统分析数据治理的研究现状。

2 数据治理框架概述

数据治理框架,是指“为了实现数据治理的总体战略和目标,将数据治理领域所蕴含的基本概念(如原则、组织架构、过程和规则等),利用概念间关系组织起来的一种逻辑结构。”简单地说,数据治理框架是指数据治理要素组成及要素之间关系的一种结构。目前,较为认可的数据治理框架主要包括:DAMA数据治理框架^[4]、DGI数据治理框架、IBM数据治理框架、EDM模型^[5]、5层框架模型^[6]、决策域模型等。

从范围上讲,目前对数据治理框架的研究领域包括通用框架,政府、企业、高校等不同行业,以及特定领域的治理框架,适用领域各不相同;从整体看,着眼数据治理框架的终端应用,因数据治理应用广泛分布于科学数据、企业数据、政府数据等不同业务情境选择,各自治理要素存在较大差别,不同数据治理框架侧重点各不相同。政府数据治理框架一般着眼组织融合视角,结合政府数据治理过程中面临的主要问题与需求,谋求标准化与协同性。科学数据管理框架,追求数据治理的通用性、功能性和推广,如DAMA数据治理框架侧重于功能,DGI数据治理框架侧重于构建路径,5层框架模型侧重于顶层设计,CALib模型侧重于方法指导;企业级数据治理框架偏重实效性、质量性,如Dell EMC数据治理侧重于以大规模基础架构实现敏捷性、效率和管理设定。从治理的核心内容上看,不同框架关注核心分别包括数据质量、安全与隐私、规则角色等。

总的来看,不同的数据治理模型均为数据治理工作提供了有益的参考,但很显然,由于行业特点、内部需求、发展阶段等众多因素,很难出现一个能够广泛适配的数据治理框架。因此,对于数据治理研究而

言，关键在于从整体上去把握数据治理的核心要素，进而在具体环境中进行合理配置。从这个角度，系统梳理数据治理的核心要素，将有助于我们更好地指导数据治理的研究与实践。笔者在上述研究的基础上，探讨由点及面的治理逻辑和价值判断，试图初步提出基于治理目标、治理主体、治理客体、治理内容、治理过程、治理手段等涵盖数据治理前端、中端、后端全流程的6个维度，搭建一种新的数据治理框架。在此框架下，通过研究梳理各子维度的研究内容，从整体上把握数据治理的研究进展，探索建立数据治理机制的有效方式，进而更好地指导数据治理开展。

3 治理框架视域下的数据治理研究内容

3.1 治理目标

治理目标定位于治理活动的战略方向、活动指南和控制标准。DAMA^[7]指出，数据治理目标包括：落实数据政策、标准、计划、指南等；跟踪数据政策、标准和规则的遵从；解决数据质量问题；提升数据价值。DGI^[8]指出，数据治理的通用目标包括：更好地决策、减少运营摩擦、保护数据相关者的权益、建立标准化及可重复的流程、减少成本并增加效率、确保过程透明等。IBM^[7]指出，数据治理的目标为：数据风险管理和合规、价值创造。PANIAN认为，数据治理最常见的驱动力是增收、降本、合规^[9]。数据治理的目标是在管理数据资产的过程中，确保数据的相关决策始终是正确、及时、有效和有前瞻性的，确保数据管理活动始终处于规范、有序和可控的状态，确保数据资产得到有效管理，并最终实现数据资产价值的最大化。郑大庆等认为，治理的本质目标是价值和风险，大数据治理“追求实现价值与风险的均衡，建立鼓励‘实现价值’和‘管控风险’这两种期望行为发生的治理体系，追求可持续的价值”^[10]。国家标准《数据治理规范》（2018）明确数据治理的目标是保障数据及其应用过程中的运营合规、风险可控和价值实现。夏义堃指出，数据开放环境下政府数据治理目标为“建设开

放政府，提高公共参与和政府数据资产价值，拉动创新和经济发展”^[11]。

从以上研究路径和实践，较大部分研究分析认为，数据治理的目标定位于提升价值和管控风险，其中，提升价值主要是指能够带来收益或提升效率，这部分数据治理较多适用科学治理和企业治理范畴；管控风险主要是指降低合规风险，减少决策失误或经济损失，这部分数据治理较为适于应用于政府数据治理类别。

3.2 治理主体

治理主体是治理活动的决策者、组织者、协调者、实施者，或者说是数据掌握、管理、使用的关联部门或角色。PANIAN认为组织架构是数据治理最重要的因素之一。DGI指出，通常一个数据治理项目包括如下机构和人员：数据治理当局（主要职责为制定规则和重要决策）、数据治理办公室（主要职责为协调和决策支持）、与数据相关的业务人员（主要职责为落实数据政策和标准、报告数据质量）。鲍静等指出，数据治理的关联部门“可分为综合协调部门、数据采集生成部门、技术平台主管部门、数据管理部门、数据利用部门、技术支撑机构等”^[12]，并通过调研梳理了具体部门的设置及分工。严昕^[13]通过分析认为，数据治理角色包括决策者、数据管理人员、数据利益相关者。数据治理的典型机构和角色包括：数据治理委员会、数据治理制度委员会、数据治理办公室、数据治理相关角色。《银行业金融机构数据治理指引》（2018）指出，“银行业金融机构应当建立组织架构健全、职责边界清晰的数据治理架构，明确董事会、监事会、高级管理层和相关部门的职责分工”。纽约大学的数据治理组织架构分为决策、统筹、执行3层，具体角色包括：首席校务数据管理官、数据受托委员会、数据管理咨询小组、院校研究和数据整合办公室、数据受托人、数据主题域受托人、数据管理员、数据保管员、数据用户等^[14]。

结合以上分析，可认为数据治理的主体主要是指数据相关者。根据职责不同，数据治理主体一般包括决策层、协调层、执行层，其中决策层主要是指各类

数据治理委员会；协调层一般是指各类数据治理办公室；执行层主要指相关数据治理工作组或业务部门及其人员。

3.3 治理客体

治理客体是治理活动的实施对象，一般而言，治理客体主要指的是数据。包冬梅等指出，高校图书馆数据治理对象包括：业务数据、资源数据、用户数据、管理数据。鲍静等从来源角度将政府部门数据分为“政务数据资源（部门内部数据资源）、社会数据资源（社会组织生成数据资源）、互联网数据资源（网络虚拟组织数据资源）”三大类，其中政务数据资源又可分为：政府网站数据、业务系统数据、专门采集/检测/监测数据、工作存档归档数据等四小类。明欣等认为，“数据治理客体就是数据本身，即智慧城市建设中各种主体在生产、工作、学习、生活等各项业务流程中所产生的各项数据”^[15]。吴信东指出，数据治理的对象主要包括：结构化数据、半结构化数据、非结构化数据、流式数据和其他系统数据^[16]。

3.4 治理内容

治理内容是指数据治理的落脚点，也是数据治理的核心所在。DAMA指出，数据治理的主要内容包括：数据架构管理、数据质量管理、元数据管理、文档与内容管理、数据仓库与商业智能管理、主数据管理、数据安全管理等。DGI指出数据治理核心领域包括：政策、标准和策略；数据质量；隐私、合规和安全；架构和集成；数据仓库和BI；管理支持。《数据治理规范》指出，数据治理的内容包括数据管理体系和数据价值体系，其中数据管理体系包括数据标准、数据质量、数据安全、元数据管理和数据生命周期，数据价值体系包括数据流通、数据服务、数据洞察等^[17]。

数据治理包括数据获取治理、数据共享治理、数据重用治理、数据增值治理4个部分。郑大庆等指出，数据治理的核心内容是：数据质量管理、数据生命周期管理、数据安全与隐私^[18]。严昕等指出，数据治理内容包括：元数据管理、数据质量管理、数据跨系统

整合、数据安全治理、数据世系追踪反馈^[19]。张世明等指出，数据治理的主要内容包括：数据模型、数据服务、数据生命周期、数据框架、主数据、元数据、数据标准、数据安全、数据质量等^[20]。马广惠等通过实证分析将政府大数据平台数据治理具化为数据集成、数据处理、数据存储、数据共享、数据一致性5个方面，并指出业务驱动是跨部门跨系统大数据平台数据治理的关键^[21]。XIN认为，病理学数据治理侧重于数据准确性、完整性、元数据管理、数据可用、数据授权等方面^[22]。

虽然不同领域的治理内容各不相同，但从当前研究和质量性应用来看，一般出发点均共识于数据治理的主要内容需含有元数据管理、主数据管理、数据质量管理、数据安全治理4个基础方面。

3.5 治理过程

《数据治理规范》依据PDCA循环，指出数据治理过程包括：统筹和规划、构建和运行、监控和评价、改进和优化。鲍静等提出，政府部门数据治理的实践路径为：完善数据平台的技术支撑体系、优化政府部门数据外部流动机制、创新政府部门数据治理新形态。明欣等指出，“数据治理过程分为前端管理、过程管理、全程管理3个维度”。美国堪萨斯州数据治理过程可分为两步，“第一步，从数据源采集原始数据，根据元数据仓库定义的转换规则将从数据源采集到的数据转换为满足EDEN格式的数据，并存储到数据仓库；第二步，到数据仓库提取数据并生成向EDEN提交的文件，并将文件提交到EDEN数据系统^[23]。”吴善鹏等指出，政务大数据治理的主要流程为：数据源分析—数据治理规范—标准管理—治理策略管理—常态化治理实施—数据资源管理^[24]。

综上，数据治理遵循PDCA循环，大体包括规划、采集、分析、评估、改进、优化等。

3.6 治理手段

从现有研究范式和实践基础看，初步可以认为数据治理的手段一般包括数据治理的法规标准适用、数

据治理相关信息技术的嵌入、数据治理工具的匹配等等。

在法规标准方面。PANIAN 指出据治理政策标准主要包括数据定义和分类、主数据、访问控制、数据传递等业务规则。黄璜详细介绍了美国联邦政府在数据开放、信息公开、电子政务、信息安全、个人隐私保护和信息资源管理等六大数据治理领域的核心政策^[25]。何哲指出，“当前中国尚缺乏一部完备的《国家数据法》，导致目前对数据治理的各种原则、措施不够明确，对各个主体各种权利界限不清晰，对违法等现象的惩罚亦是空白^[26]。”李重照等详细分析了英国政府在个人数据保护、信息公开、政府数据开放、国家信息基础设施、信息资源管理与再利用、电子政务和网络信息安全等方面的数据治理政策^[27]。吴善鹏等指出，做好数据治理需要建立健全以下规范：元数据标准、数据元标准、数据分类编码标准、数据目录规范、数据质量标准、数据治理流程标准。

在数据治理信息技术方面。AI-RUITHE 指出，“技术是数据治理成功的关键因素^[28]”。戚学祥认为，“当政府数据治理面临海量、异质的数据与数据壁垒时，区块链不可逆的分布式账本系统、复杂数学算法、非对称加密技术等核心技术为解决数据的真实性、安全性、开放性等问题提供了可能性^[29]”。谭必勇等指出，“加拿大联邦政府从技术战略、技术安全管理和云技术应用管理等 3 个方面建立起了完整的技术治理模式^[30]”。吴信东等具体分析了数据规范、数据清洗、数据交换和数据集成 4 种技术在数据治理中的应用。吴善鹏等指出，“政务大数据治理中所用的主要技术包括数据检核引擎、ETL 工具、消息中间件、流程引擎、Hive 和 MPP DB 等，其中数据质量检核引擎的设计最具技术含量，也最为重要”。

在数据治理工具方面。鲍静等提出建立统一共享的数据共享平台，“解决不同类型信息采集渠道的重复建设，推进采集数据的有效交换和充分利用”。严昕等从数据层、资源层、工具层、应用层及展示层 5 部分搭建了城镇信息化数据治理平台系统。谭波指出，国家开发银行通过搭建混搭架构的大数据平台、分布式

的统一数据交换平台、外部数据平台、手工数据管理系统和数据归档系统，构建了完备的数据基础设施^[31]。张宇杰等搭建了政府大数据治理的成熟度评测模型，作为自我测评和改进的工具^[32]。吴信东等构建了包括数据源、数据接入模块、数据治理模块、数据服务模块等功能模块的“HAO 治理”模型，“该模型从大数据开始，为 HI（人类智能）、AI（人工智能）和 OI（组织智能）三者协同提供数据治理支持^[33]”。

4 结论与展望

通过对国内外总体研究动态的分析可以看出，针对数据治理的逻辑方法和实践中的各环节内容，国内外众多学者展开了大量研究，取得了丰硕的成果。以数据治理框架为分析视角，以治理目标、治理主体、治理客体、治理内容、治理过程、治理手段为分析维度，能够在现有理论和实践基础上，比较系统地较好搭建起了数据治理的研究和实践框架，为指导数据治理的开展提供一种新的较为清晰的思路。细分而言，数据治理目标主要是提升数据价值和防范数据风险；数据治理主体包括决策层、统筹层和执行层等数据相关者；数据治理客体主要包括各类数据资源，特别是有价值的的数据资源；数据治理核心内容包括元数据管理、主数据管理、数据质量、数据安全等；数据治理过程主要遵循 PDCA 循环；数据治理手段包括相关法规标准、信息技术和工具等。

这一框架下，以上数据治理核心要素中，针对治理目标、治理主体、治理手段等方面的认识，已经基本达成一致；针对治理内容的研究，仍相对比较多元，哪些内容是数据治理的关键点，仍有待进一步统一；针对治理客体的研究，总体上还比较抽象，有待进一步细化。同时，从数据治理的要素内容来看，其构成与数据管理存在较大趋同，在数据治理过程中，如何更好地区分数据管理和数据治理，更好地坚持数据治理的定位，凸显数据治理的重点与核心，仍是后续研究中需要注意的问题，是一个理论和实践方向。

参考文献:

- [1] 张宁,袁勤俭.数据治理研究述评[J].情报杂志,2017(5):129-134,163.
- [2] 刘桂锋,钱锦琳,卢章平.国内外数据治理研究进展:内涵、要素、模型与框架[J].图书情报工作,2017(11):137-144.
- [3] 孙嘉睿.国内数据治理研究进展:体系、保障与实践[J].图书馆学研究,2018(16):2-8.
- [4] NASCIO. Data governance partIII: frameworks-structure for organizing complexity[R/OL]. [2019-08-14]. <https://www.nascio.org/Portals/0/Publications/Documents/NASCIO-DataGovernancePTIII.pdf>.
- [5] ISO. Information technology-governance of it-governance of data part2: implications of ISO/IEC38505-1 for data management [S]. The standards policy and strategy committee, 2018(5).
- [6] SEINERRS. The non-invasive data governance framework-the details[EB/OL]. [2019-08-17]. <http://tdan.com/the-non-invasive-data-governance-framework-the-details/25031>.
- [7] NASCIO. Data governance partIII: frameworks-structure for organizing complexity[R/OL]. [2019-08-14]. <https://www.nascio.org/Portals/0/Publications/Documents/NASCIO-DataGovernancePTIII.pdf>.
- [8] The data governance institute. Goals and principles for data governance[EB/OL]. [2019-09-01]. http://www.datagovernance.com/adg_data_governance_goals/.
- [9] PANIANZ. Some practical experience in data governance[J]. World academy of science, 2010: 939-946.
- [10] 郑大庆,黄丽华,张成洪,等.大数据治理的概念及其参考架构[J].研究与发展管理,2017(4):65-72.
- [11] 夏义堃.试论数据开放环境下的政府数据治理:概念框架与主要问题[J].图书情报知识,2018(1):95-104.
- [12] 鲍静,张勇进.政府部门数据治理:一个亟需回应的基本问题[J].中国行政管理,2017(4):28-34.
- [13] 严昕,孙红蕾,郑建明.城镇信息化中的数据治理问题研究[J].情报科学,2017(9):30-35.
- [14] NYU. Administrative data management policy[EB/OL]. [2019-08-25].<https://www.nyu.edu/about/policies-guidelines-compliance/policies-and-guidelines/administrative-data-management-policy.html>.
- [15] 明欣,安小米,宋刚.智慧城市背景下的数据治理框架研究[J].电子政务,2018(8):27-37.
- [16] 吴信东,董丙冰,堵新政,等.数据治理技术[J].软件学报,2019(9):2830-2856.
- [17] 全国信息技术标准化技术委员会.信息技术服务治理第5部分:数据治理规范:GB/T34960.5-2018[S].北京:中国标准出版社,2018(6).
- [18] 郑大庆,范颖捷,潘蓉,等.大数据治理的概念与要素探析[J].科技管理研究,2017(15):200-205.
- [19] 严昕,孙红蕾,郑建明.城镇信息化中的数据治理问题研究[J].情报科学,2017(9):30-35.
- [20] 张世明,彭雪峰,黄河笑.开放大学数据治理框架研究[J].中国电化教育,2018(8):116-126.
- [21] 马广惠,安小米,宋懿.业务驱动政府大数据平台数据治理[J].情报资料工作,2018(1):21-27.
- [22] FUX, WOJAKA, NEAGUD, et al. Data governance in predictive toxicology: A review[J]. Journal of cheminformatics, 2011(1):1-16.
- [23] 李青,韩俊红.数据治理:提升教育数据质量的方法和途径[J].中国远程教育,2018(8):45-53.
- [24] 吴善鹏,李萍,张志飞.政务大数据环境下的数据治理框架设计[J].电子政务,2019(2):45-51.
- [25] 黄璜.美国联邦政府的数据治理:政策与结构[J].中国行政管理,2017(8):47-56.
- [26] 何哲.国家数字治理的宏观架构[J].电子政务,2019(1):32-38.
- [27] 李重照,黄璜.英国政府数据治理的政策与治理结构[J].电子政务,2019(1):20-31.
- [28] AL-RUITHEM, BENKHELIFAE, HAMEEDK. Data governance taxonomy: Cloud versus non-cloud[J]. Sustainability, 2018(10):95.
- [29] 戚学祥.区块链技术在政府数据治理中的应用:优势、挑战与对策[J].北京理工大学学报(社会科学版),2018(9):105-111.
- [30] 谭必勇,陈艳.加拿大联邦政府数据治理框架分析及其对我国的启示[J].电子政务,2019(1):11-19.
- [31] 谭波.做好金融大数据治理与服务—国家开发银行的大数据治理探索与实践[J].银行家,2018(7):65-67.
- [32] 张宇杰,安小米,张国庆.政府大数据治理的成熟度评测指标体系构建[J].情报资料工作,2018(1):28-32.
- [33] 吴信东,董丙冰,堵新政,等.数据治理技术[J].软件学报,2019(9):2830-2856.