

中国大数据治理研究述评

宋姗姗^{1,2}, 白文琳^{3*}

(1. 中国科学院 武汉文献情报中心, 武汉 430071;

2. 中国科学院大学 经济与管理学院图书情报与档案管理系, 北京 100049; 3. 天津师范大学 管理学院, 天津 300387)

摘要: [目的/意义]数据已成为关键生产要素和战略性资源, 大数据治理是充分释放数据价值、实现国家治理体系和治理能力现代化的重要途径。[方法/过程]本研究在文献计量与内容分析的基础上, 对国内大数据治理的相关研究进行了系统梳理和内容评述, 发现近年来有关大数据治理的研究热度不断上升, 并将相关研究主题概括为概念体系、实施评估与实践挑战3个方面。[结果/结论]研究发现大数据治理体系基本是从治理目标、治理主体、治理客体以及治理工具4个方面进行构建, 治理的实施评估重点围绕数据生命周期、数据质量、数据安全、数据治理成熟度、数据审计等内容, 实践挑战主要表现在大数据治理领域所面临的管理和技术问题。

关键词: 大数据; 大数据治理; 治理体系; 治理实施评估

中图分类号: G203 **文献标识码:** A **文章编号:** 1002-1248 (2022) 04-0004-14

引用本文: 宋姗姗, 白文琳. 中国大数据治理研究述评[J]. 农业图书情报学报, 2022, 34(4): 04-17.

1 引言

伴随着新一轮信息技术革命浪潮的出现和大数据时代的到来, 大数据已经成为关键性生产要素和国家基础性战略资源, 利用大数据技术手段提升国家治理能力现代化已经成为趋势。党的十八届三中全会明确将“推进国家治理体系和治理能力现代化”作为全面深化改革的总目标, 十八届五中全会提出要实施“国家大数据战略”, 同时《促进大数据发展行动纲要》的实施, 使中国大数据事业开启了快速发展模式。2020

年党的十九届四中全会首次将“数据”列为生产要素, 明确提出要“加强数字社会、数字政府建设, 提升公共服务、社会治理等数字化智能化水平”。

大数据价值的释放有赖于一套科学完善的大数据治理体系, 因而“大数据治理”在近几年已经成为学术界所关注的焦点问题, 国内学者开展了大量有关大数据治理的理论探索和实践创新研究。基于此, 本文对国内大数据治理的已有研究进行发展脉络梳理和文献述评, 进而试图厘清大数据治理的概念体系、实施评估及应用实践内容, 以期更好发挥大数据作为新型生产要素和社会治理手段的正向实践功能。

收稿日期: 2022-02-22

基金项目: 中国人民大学科研基金重大项目“政府大数据治理与综合利用的标准化协同管理体系构建及其应用研究”(21XNL019)

作者简介: 宋姗姗, 博士研究生, 中国科学院武汉文献情报中心, 研究方向为数据治理、智库研究

*通信作者: 白文琳, 博士, 讲师, 天津师范大学管理学院, 研究方向为政府数据治理、公共信息服务、电子文件管理、档案管理。E-mail: wenlin-bai@163.com

2 大数据治理研究总体发展态势

2.1 文献分布与走势

本文以中国知网 (CNKI) 为来源数据库, 以 $TI=$ ('大数据*'治理') or '大数据治理' 为检索式进行中文期刊文献检索 (检索时间: 2021 年 12 月 31 日), 论文发表时间不限, 检索结果为 1 704 篇。经过人工清洗与去重合并, 剔除新闻报道、图书推介及与研究主题明显不符的文献, 最终得到 1 635 篇有效论文作为本研究的原始数据来源。

从文献数量的时间分布图 1 可知, 在 2013 年以前, 学界对于大数据治理的相关研究还没有进行, 这一领域还属于研究空白。从 2013 年起至今, 年度论文、参与机构、研究人员的数量都开始逐渐上升, 大数据治理的相关研究呈现出一个迅猛发展的势态。其中有两个关键的节点时间是 2013 年和 2015 年。2013 年 11 月党的十八届三中全会通过《关于全面深化改革若干重大问题的决定》提出要“创新社会治理体制”; 2015 年印发《关于积极推进“互联网+”行动的指导意见》《促进大数据发展行动纲要》, 以及《关于运用大数据加强对市场主体服务和监管的若干意见》, 明确提出“运用大数据技术推动政府精准治理、市场资源整合和经济转型发展, 形成多方协作的社会治理新

模式”。基于以上政策的实施, 引领着学界特别是公共管理学科的学者更多地把研究聚焦在大数据治理的相关领域上来。从研究涉及的学科领域来看, 国内关于大数据治理的学科分布相对比较宽泛, 发表相关论文较多的学科有行政管理、计算机科学、信息经济学、新闻传媒以及社会统计学等; 发文较多的期刊包括《电子政务》《情报杂志》《中国行政管理》《情报理论与实践》《图书情报知识》等; 发文较多的代表性学者包括中国人民大学安小米、上海财经大学郑大庆、武汉大学夏义堃、清华大学孟天广以及北京大学黄璜等。

2.2 研究内容及热点

本文通过对国内有关“大数据治理”的文献关键词进行聚类分析可以看出, 近年来关于大数据治理的研究热度不断上升, 且受到较多领域的广泛关注, 涉及的核心关键词包括: 大数据、数据治理、社会治理、政府治理、大数据技术、数据资源、网络舆情、风险治理、数据安全等, 研究主题聚类结果如图 2 所示, 具体聚类主题详见表 1。

通过文献计量分析可以看出, 国内关于大数据治理的研究已有一定成果积累, 相关主题涉及大数据治理概念、政府治理能力、数据开放与共享活动、大数据技术运用、数据安全与隐私问题等, 据此, 本研究

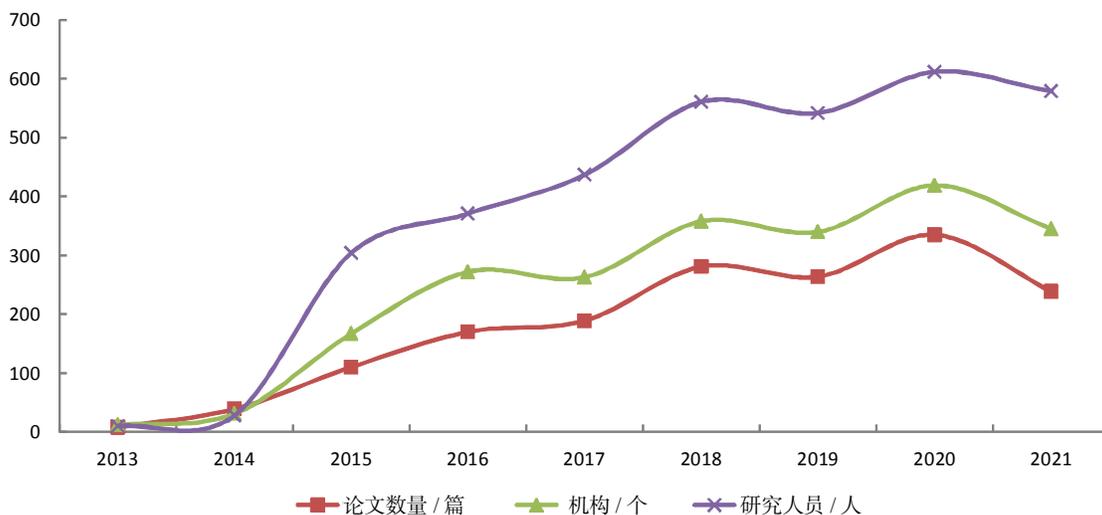


图 1 中国大数据治理领域研究发展趋势

Fig.1 Distribution of big data governance research in China

学者 Sunil Soares (桑尼尔·索雷斯) 在 *Big Data Governance: An Emerging Imperative* (《大数据治理》) 一书中提出, 他认为“大数据治理是广义信息治理计划的一部分, 它通过协调多个职能部门的目标和利益来制定有关大数据优化、隐私保护以及数据资产变现的策略^[2]”。基于此, 国内学者从不同角度对大数据治理概念进行定义, 核心内容上具有一定的共同点。相关研究可以归纳为以下 4 个方面。

(1) 大数据治理与数据治理的关系。大数据治理形成于大数据时代, 是在“数据治理”概念基础上发展而来。郑大庆^[3]认为数据治理与大数据治理的目标相同, 都是为了“实现价值”和“管控风险”, 不同在于数据治理侧重于效率提升, 大数据治理强调效益提升和风险管控。黄璜^[4]阐述了数据治理是为有效地使用数据资产而组织实施的政策、流程和标准的实践。夏义堃^[5]从数据治理部署的角度, 指出政府数据治理是综合运用法律制度、人员组织、技术方法、流程标准等手段, 对政府数据的可用性、完整性、安全性等进行全面有效管理。

(2) 大数据治理与数据资产变现。现有大数据治理研究都重点关注了数据资产变现问题。安小米^[6]从提出政府大数据治理是政府机构进行大数据决策权分配和职责分工的活动过程, 旨在促进政府数据资产价值最大化。朝乐门^[7]认为大数据治理是为了帮助组织更好地管理数据资产, 改善组织在整个业务领域的的数据管理实践。张绍华^[8]提出大数据治理是对大数据的管理和利用进行评估、指导和监督的体系框架, 其目的之一是实现数据的价值创造, 并提供创新的大数据服务。

(3) 大数据治理与信息技术支撑。信息技术是实现大数据治理的重要依托工具。向芳青^[9]认为大数据治理要借助信息技术获取、互联网平台汇聚, 先进的技术手段和方法进行数据分析, 提升科学高效的治理能力。梁芷铭^[10]提出大数据治理是运用技术工具进行大数据收集、整理、存储、分析及挖掘其价值的行为。此外, 也有学者认为大数据治理不仅需要信息技术工具的支持, 还需要关注大数据治理形成的规则和政策框架, 实现大数据治理工具从技术维度到管理维度的

转变^[11]。

(4) 大数据治理与国家治理变革。大数据治理被视为国家治理的关键组成部分。戴香智^[12]提出大数据为社会治理带来了多中心治理范式、“用数据说话”方法论、全数据关联视角、主动预见型治理流程、治理技术能力提升。石火学^[13]认为政府大数据治理正朝着治理主体协同化、治理内容预防化、治理过程透明化、治理技术迭代化、治理方式数据化和治理结果精准化等方向发展。张明斗^[14]提出大数据治理通过数据共享、法律约束、绩效评估、运营保障 4 个机制推动城市治理现代化。

3.2 大数据治理体系

大数据治理是一项复杂、多层次、多维度的系统化工程, 需要充分调动相关主体资源和工具手段, 形成一个普遍科学有效的大数据治理体系, 更好地支持大数据治理应用实践。马广惠^[15]提出大数据治理体系由人、物、技术、数据构成。刘越男^[16]提出了涵盖治理主体、治理客体、治理方法和流程的政府大数据治理理论框架。据此, 本研究从大数据治理目标、治理主体、治理客体、治理工具 4 个层面对大数据治理体系研究进行归纳, 具体如表 2 所示。

(1) 大数据治理目标。在大数据治理实施过程中, 首先需要明确大数据治理的目标, 从而让决策者对大数据治理的实施制定总体规划。从宏观层面来看, 实现大数据的风险管控、运营合规和价值创造, 确定数据的资产地位, 是大数据治理与数据治理的根本区别, 也是当前研究较为公认的大数据治理目标。从微观层面来看, 吴善鹏^[16]将政务大数据治理目标归纳为: 提升数据质量、促进数据融合、规范数据流程以及发掘数据价值。大多数学者在大数据治理目标上都重点关注了数据价值创造的话题, 数据本身不产生价值, 但是从庞杂的数据背后评估和预测相关利益者的需求, 并有针对性地产生一系列创新的大数据服务, 才能达到价值共创的结果。

(2) 大数据治理主体。治理主体是治理活动的决策者、组织者和协调者。在大数据治理主体的研究中,

表2 大数据治理体系及构成要素梳理

Table 2 Big data governance system and constituent elements

序号	代表学者	治理体系要素	治理目标	治理主体	治理客体/对象	治理工具/手段
1	郑大庆等 ^[3]	治理目标; 权利层次; 治理对象; 解决的实际问题	实现价值; 管控风险	大数据利益相关者(个人、数据收集者、数据处理者、数据应用者、管理监督者)	大数据治理规则; 制度; 流程	决策机制(决策权、组织结构); 监督机制; 激励与约束机制
2	梁芷铭 ^[10]	治理理念; 治理主体; 治理客体; 治理工具; 治理目标	大数据价值的挖掘; 大数据治理的应用	企业界; 与大数据相关部门; 机构和组织	大量数据; 数据结构; 数据容量; 数据形态; 数据性质; 数据传播; 数据价值; 数据安全	组织结构; 相关政策法规; 计算机等硬件设施
3	张绍华等 ^[8]	原则子框架; 范围子框架; 实施与评估子框架	风险可控; 安全合规; 绩效提升; 价值创造	大数据治理利益相关者; 大数据治理委员会; 大数据管理者; 数据专家	大数据质量; 大数据生命周期; 大数据架构; 大数据安全隐私与合规	组织要素(战略方针、组织架构、责任分工); 促成因素(环境与文化、技术与工具、流程与活动); 实施过程; 成熟度评估; 审计
4	安小米等 ^[6]	治理目标; 治理主体; 治理客体; 治理活动; 治理工具	运营合规; 风险控制; 价值创造	大数据治理委员会; 治理办公室; 管理者; 数据专家	数据; 信息; 人员	信息基础设施; 大数据技术; 检测工具
5	程广明 ^[22]	人与组织; 策略; 能力	数据优化; 隐私保护; 数据变现(出售数据、利用数据开发业务)	大数据治理利益相关者(数据产生者、管理者、监督者); 治理委员会; 管理委员会; 内部员工	元数据管理; 数据质量管理; 业务流程整合; 主数据管理; 信息生命周期管理	治理的使命和愿景; 数据治理规则和定义; 数据治理指标; 组织和个人的权利职责; 治理风险的控制措施
6	吴善鹏等 ^[16]	组织架构; 标准规范; 流程机制; 功能软件; 法律机制	提升数据质量; 促进数据融合; 规范数据流程; 发掘数据价值	数据治理方; 数据提供方; 数据使用方	数据源管理(部门与网络数据); 贴源层治理(建模、标准化、技术检核); 中心层治理(数据关联融合、业务数据检核); 数据资源中心(数据资产目录、基础库、主题库)	主要技术支撑(Hive、ETL、流程引擎等); 数据规范规范体系(元数据标准、数据分类标准、数据质量标准等); 大数据安全与隐私管理体系; 数据治理平台系统
7	杨琳等 ^[23]	治理目标; 治理保障; 治理域; 治理方法论	战略一致; 风险控制; 运营合规; 价值实现	根据具体情况建立大数据治理组织结构	数据架构; 元数据管理; 主数据管理; 数据生命周期管理; 数据质量管理; 数据安全与合规管理	方法论(评估、监督、指导); 促成因素(文化、技术工具、流程活动); 内外部环境(政策法规、市场需求等)
8	夏义堃 ^[5]	组织管理; 法律制度; 数据资产; 风险管理	数据资产的保值增值	政府部门、社会组织、企业、公众、黑客	数据内容; 大数据技术及平台; 业务流程; 信息基础设施; 人员和内部管理等	数据管理法律制度; 人员组织; 技术方法; 流程标准
9	刘驰等 ^[24]	数据治理原则(有效性、价值化、统一性、开放性、安全性)、治理范围; 实施与评估	业务创新; 价值获取; 风险控制	战略决策者; 业务管理者; 业务操作员	大数据生命周期; 大数据架构(数据存储、元数据、数据仓库、业务应用); 数据安全与隐私; 数据质量; 数据服务创新	治理目标与业务需求; 治理的软硬件环境; 实施流程体系和规范; 实施技术和工具

表 2 (续)

Table 2 Continue

序号	代表学者	治理体系要素	治理目标	治理主体	治理客体/对象	治理工具/手段
10	尧淦等 ^[25]	组织结构; 标准规范; 技术支持	数据的完整性、一致性、合规性; 保障数据安全和质量; 大数据价值实现	大数据治理委员会; 大数据治理办公室; 大数据管理者; 监察委员会	规范管理; 数据过程; 技术标准	基础设施和大数据技术; 人才/人员; 平台建设; 应用示范
11	马广惠等 ^[11]	治理主体; 治理客体; 治理工具; 治理过程(大数据汇聚、融合、应用)	业务数据应用; 场景化数据整合; 数据价值挖掘	大数据领导小组; 大数据局; 大数据治理公司	大数据资源; 大数据治理的过程和方法; 大数据治理的社会功能	数据资源整合数据库; 数据共享交换平台与机制; 数据处理标准与规范; 数据安全保障与监控体系
12	向芳青等 ^[9]	数据源; 数据汇聚; 数据分析; 治理能力提升	实现数据价值; 改变决策方式; 提升治理能力	政府; 社会; 公众	数据公开(部门数据、共享数据、社会数据); 数据汇集(数据汇集、关联、更新); 数据分析(数据统计、挖掘、预测)	大数据思维; 大数据整合与共享平台; 数据开放共享制度; “互联网+”技术手段

学界基本统一认为大数据治理主体主要涉及政府, 企业、社会组织和公民个人等, 同时也针对这些主体在大数据治理中的地位、作用及效果进行了探讨^[7]。在具体组织内部, 大数据治理团队是大数据治理的促成要素, 是组织机构中参与决策的群体, 包括大数据利益相关者、大数据治理委员会、大数据管理团队以及大数据专家等, 有效的组织架构和明确的职责分工是大数据治理目标实现的有力保障^[18]。

(3) 大数据治理客体。大数据治理客体即治理的对象和内容。在对大数据治理客体理解上, 学者们还尚未达成共识, 多数学者都尝试采用“归类”与“分层”的方法进行大数据治理客体的研究。从狭义上理解, 大数据治理范围既涵盖传统的业务数据, 也包括大量非结构化的网络数据^[19]。从广义上说, 大数据治理的核心客体是数据本身, 同时也包含了一系列与大数据治理相关的活动和要素, 涉及数据生命周期、数据质量、数据技术与平台、信息基础设施、数据隐私与安全、组织人员和业务管理等^[5]。

(4) 大数据治理工具。治理工具是实现大数据治理活动的有力保障, 具有多样性特征。在现有研究中, 可以将大数据治理工具归纳为环境因素和促成因素两个方面^[20], 环境因素包括国家政策、行业规范、市场

需求、企业制度等内外部环境; 促成因素包括管理、组织、技术、标准、流程、文化等。也有学者提出大数据治理的工具大部分也适用于大数据治理, 所以大数据治理既包括元数据、数据质量控制、数据标准制定等传统数据治理方法, 也需要制定特殊的管理规则, 综合运用物联网、云计算、关联数据、机器学习等大数据技术手段来管理大数据生命周期^[21]。

通过对国内大数据治理体系要素构成的文献梳理结果表明, 近年来有关大数据治理体系构建已经取得了较多的研究成果, 但还存在维度单一、层面隔离、缺少多样化治理场景的研究局限, 对于建立一个行之有效、灵活通用、覆盖多主体、多学科、多维度、多应用场景的大数据治理体系研究还比较缺乏, 所以大数据治理体系的构建是一项长期性、计划性和连续性的工作, 也是一个不断持续改进的过程。

4 大数据治理的实施及评估

大数据治理实施与评估的重点就在于建立大数据治理体系。杨琳等^[22]认为政府大数据治理活动包括组织规划、具体实施以及评估优化 3 个阶段, 其中组织规划阶段包括制定实施方针和构建治理组织, 具体实

施阶段包括元数据与主数据管理、数据质量与数据安全保障等，评估优化阶段是对数据治理结果进行评估分析，并根据反馈结果持续优化。综上，本研究将大数据治理实施与评估需要重点关注的内容归纳为图3所示。

4.1 大数据治理实施

治理实施描述了在大数据治理过程中需要重点关注的内容，从而为组织进行大数据治理提供指导性方案，治理重点包括大数据生命周期、大数据质量管理以及大数据安全与隐私保护3个方面。

(1) 大数据生命周期。从大数据治理涉及的数据生命周期来看，大数据治理是在大数据战略的指导下，从治理问题的发现与理解、数据采集与整合、数据建模与分析以及解决方案的提出等一系列过程。与传统的数据生命周期不同在于，大数据生命周期管理更关注在风险管控的条件下，最大程度发挥数据的潜在价值。吴善鹏^[16]指出政务大数据治理活动涉及数据源分析、治理活动规划、数据资源和标准管理、数据治理策略管理以及治理活动实施。从大数据治理的具体实践情况来看，数据质量、个人隐私保护和信息安全是研究中的重要议题，马广惠^[26]面向跨系统和跨部门的政府大数据平台，提出大数据治理活动需要重点关注5项具体内容：数据集成、数据一致性、数据处理、数据存储和数据共享。

(2) 大数据质量管理。当前大数据治理领域对于数据质量的研究主要集中在数据质量的评估和提升两个方面。数据质量评估是数据质量管理的基础，关键在于评估指标的选取，大数据治理评估指标的选取要基于数据质量特征、数据生命周期、数据使用问题、数据应用情景等^[27]。莫祖英^[28]认为数据质量要求应包括客观数据质量范畴（规范性、安全性、增值性）和主观数据质量范畴（可用性、可靠性、及时性、可获取性、可理解性）。朱建平^[29]从数据收集（客观性、适用性、准确性），数据处理（方法的健全性、可靠性、可比性），数据公布（及时性、完整性、可获得性）3个环节阐述了数据质量评价构想。数据质量的提升侧重于从管理和技术两个视角出发，续瑾成^[30]从管理角度来看，数据质量的提升重点在于建立相关组织架构、设计数据质量标准、加强对主数据的管理等策略实施，宗威^[31]从应用数据库技术、数据分析技术、数据质量监测技术的技术角度研究了数据质量的提升路径。

(3) 大数据安全与隐私保护。当前，大数据安全管理问题日渐凸显，大数据的4V特征、数据的跨组织和跨系统流通、数据应用的复杂性以及数据分析挖掘技术的迅速发展，进一步加速了数据被盗用和滥用的安全风险。数据安全防护主要通过区块链技术来实现数据的安全存储与安全计算，如分布式存储、数据异地备份、加密手段、身份统一认证、权限控制技术、入侵检测与防御、高危操作防护等。肖人毅^[32]总结了

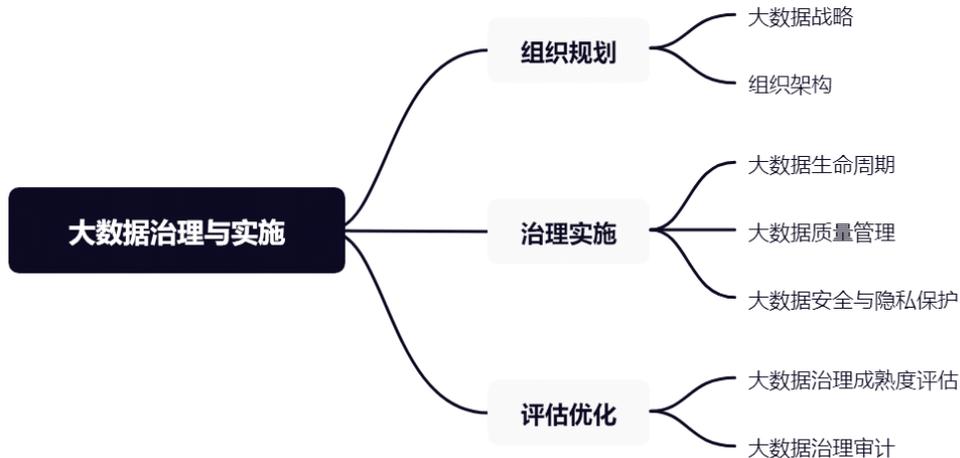


图3 大数据治理实施与评估的重点内容

Fig.3 Content of big data governance implementation and evaluation

基于不经意随机访问存储器 (ORAM)、对称加密、公钥体制以及文档的排名查询和模糊查询 4 个方面的数据保护技术。数据隐私保护的主要目的是解决在不泄露用户隐私的前提下、提高数据利用率及价值, 方滨兴^[33]提出数据隐私保护涉及匿名、加密存储、第三方审计、访问控制、数据脱敏技术等。此外, 大部分学者认为大数据的安全与隐私保护还需要非技术性层面的协同支持, 如专门的数据安全管理组织、全面的数据管理工具与机制、相关法律法规保障、标准统一的数据安全服务等^[34]。因此如何在合理监管的基础上保证数据获取与利用的可靠性, 实现数据价值挖掘, 是大数据治理研究领域面临的巨大挑战。

4.2 大数据治理评估

通过治理成熟度评估可以详细了解当前大数据治理的实施状态, 实现组织对大数据管理能力的量化评价, 是大数据治理体系规划的起点, 大数据治理审计有利于促进大数据治理的规范合规性, 为实现数据价值最大化提供依据。

(1) 大数据治理成熟度评估。当前比较通用的是 IBM 数据治理成熟度模型, 该模型包括成效 (数据风险管理、业务价值创造), 支持要素 (组织结构与文化、数据管理、政策), 核心准则 (数据质量管理、信息生命周期管理、信息安全与因素), 支撑准则 (数据架构、分类与元数据、信息审计记录与报告) 4 个相互关联的组及 11 个关键要素。中国于 2018 年发布了国家标准《数据管理能力成熟度评估模型》, 该模型包含了组织层面 (数据战略、数据生命周期), 技术层面 (数据应用、数据安全), 制度层面 (数据治理、数据架构), 流程层面 (数据质量、数据标准) 4 个评估层面及 8 个关键过程域。国内学者程广明^[22]提出了包括初始级、基本级、定义级、管理级和优化级 5 个评价等级、15 个具体评价指标的大数据治理成熟度评估模型。张宇杰^[35]提出大数据治理成熟度评估内容主要集中在战略规划、数据管理、组织保障、制度保障、技术架构、治理能力 6 个方面。

(2) 大数据治理审计。开展大数据治理审计是对

大数据治理活动进行综合检查和评价, 发现存在的潜在风险, 并给出详细、有价值的审计意见, 促进大数据治理的持续优化和规范发展。张绍华^[6]提出大数据治理的审计范围涵盖了大数据治理的整个生命周期过程, 以及在该过程中产生的中间产物和相关的治理环境, 具体可对应大数据治理目标进行检查评估, 分别为战略一致审计、风险可控审计、运营合规审计、绩效提升审计, 同时还包括在大数据治理过程中的组织机构审计、数据生命周期审计、数据架构审计、数据安全与隐私管理审计、数据质量审计。郑伟^[36]提出大数据的参与创新了基于数据式的审计模式, 并从逻辑流程、网络构建和应用架构 3 个方面阐述了大数据审计的完善路径。

5 大数据治理的实践挑战

5.1 大数据治理的应用实践

当前, 中国通过制定大数据发展战略, 部署相关行动计划, 积极将大数据应用于公共管理、政府治理和产业经济发展等领域, 取得了一定成效。国内学者也均从不同领域的实践出发, 强调了大数据治理的重要价值, 主要研究集中在中国政府积极运用大数据提升科学决策能力、社会管理水平及公共服务建设 3 个方面。

政府的科学决策是建立在充足的数据支撑、科学的决策程序和有效的决策手段基础之上。胡税根^[37]认为基于大数据的智慧公共决策具有全面感知、客观透明、实时连续、自主预置和多元共治等新的特征。单志广^[38]提出大数据的快速发展为社会管理带来了思维方式和模式的变革, 要积极把大数据技术与思维运用到政府治理能力与决策之中。马琳^[39]总结在大数据环境下, 社会管理从静态化向动态化、从碎片化向协同化、从单一政府管理向多元公众参与转变。赵亚珠^[40]认为大数据应用保障政府决策的科学化和民主化、实现公共利益表达的自由化和平等化、推动社会管理的透明化和精细化。此外, 基于大数据衍生出的智能

公共服务与个性化公共服务已经成为当前的发展趋势，受到大数据影响的社会服务领域包括工业企业、商业金融、科技教育、农业环境、医疗卫生、城市管理等重点领域，具体如表 3 所示。特别是在新冠疫情防控工作中，大数据在疫情监测、行程跟踪、资源调配、物流运输等方面得到了广泛应用。

5.2 大数据治理的应用挑战

尽管当前各行业和领域都十分重视并大力推进大数据技术在社会治理中的应用探索，但大数据作为近几年逐渐发展的新兴事物，中国对于大数据的实践应用整体来说还处于探索发展阶段。国外研究发现大数据治理面临的主要挑战在于技术障碍，社会、组织和经济障碍，法律和政策障碍，其根本还是管理问题^[51]。总体而言，大数据治理挑战主要包括管理和技术两个方面。

管理挑战主要表现为：①数据垄断问题。数据依赖平台存在，目前一些拥有大量数据的支配组织对其他经营者的数据访问行为进行限制，意图垄断数据资源，因而有序推进数据开放、加强数据法律体系建设、构建良好的数据生态的过程还很漫长。②数据权属问题。数据作为一种全新的无形虚拟资产，具有巨大的潜在应用价值，数据权属不清一直是影响数据资产化和数据交易的首要问题，因此“确认数据权属”和“保

护数据产权”对于数据资产化具有重要意义^[52]。③数据隐私问题。如何在数据安全隐私保护和提高大数据利用价值之间实现有益平衡，是目前大数据治理研究的关键问题之一，因此未来涉及国家安全、网络安全、数据安全和个人隐私保护的数据标准和数据立法还需要更多研究。④数据壁垒问题。传统信息管理多头交叉容易导致“信息孤岛”现象，“信息孤岛”现象使得大数据治理无法获得多样化的丰富数据，阻碍了数据效能的提升，因此优化与数据共享、交易、流通相关的制度和服务体系势在必行。⑤数据质量问题。由规模巨大、类型复杂、来源多样的数据构成的大数据海洋，由于缺乏一致的数据质量标准，使得数据的完备性、有效性、标准化、权威性不足，大数据治理一直处于一种无序状态，给数据的整合共享造成困难^[53]。

技术挑战主要表现在：①大数据技术滞后。当前中国虽然在大数据相关技术方面有所突破，但在海量数据存储与组织、数据挖掘与处理、以及融合云计算、人工智能、区块链等技术进行大数据治理方面与国外先进水平还存在明显差异，对一些前瞻新兴领域的数据分析还存在不足。②“暗数据”广泛存在。数据规模呈几何数高速增长，及大量异质、半结构化和非结构化数据的存在，显著增加了数据存储、整合和处理的困难，数据资源迅猛增长与数据处理能力提升之间的“剪刀差”不断扩大，广泛产生了未被有效利用的

表 3 大数据治理应用领域研究开展情况

Table 3 Application of big data governance research

研究领域	大数据治理的实践影响
工业	工业过程优化；故障监测与诊断；产品性能提升；产品生产率预测 ^[41]
商业	商业模式创新；资源与技术的工具化运用；推动大数据产业链形成；引发行业跨界融合 ^[42]
金融	金融数据聚合与价值挖掘；风险管理；消费智能；智能运营；品牌创新 ^[43]
企业	战略理念创新；关键业务创新；工作流程创新；收益模式创新；关系与价值网络重构 ^[45]
科技	科技数据打通融合；科技情报洞察与前瞻预测；科技项目查重查新；科技数据监测与价值评估 ^[44]
教育	过程数据化；决策科学化；模式个性化；管理可视化；评价综合化；服务人性化 ^[45]
医疗	智慧医疗（临床诊断、疾病预警、远程医疗、成本控制）；医药研究（临床科研、药品研发）；智能决策（准确决策、科学临床决策） ^[46]
农业	农业生产率提高；农产品交易创新；农业管理标准化；农业信息服务支撑 ^[47]
环境	数据的收集与共享；预警和环境政策制定；目标设定与综合管理；立体化的环境治理 ^[48]
城市管理	城市计算技术体系建设；数据驱动的智慧交通；城市人类活动统计 ^[49]
疫情防控	治理主体的协同性；行为决策的科学性；风险预测的前瞻性；政府回应的权威性 ^[50]

“暗数据”。③“通用”大数据技术体系缺乏。目前大数据技术应用平台已有大量研究,针对大数据治理特定领域和特定问题的解决也已有完善的技术体系作为保障,但尚缺乏“领域通用”的大数据技术体系,数据的融合、流动、监管和应用还面临一定挑战。

6 结 语

近年来,大数据作为核心战略资源和治理能力现代化的重要抓手,在重塑国家治理理念、推动数字经济发展、提高政府科学决策、加快智慧城市建设、营造良好社会生态等方面发挥了关键性作用。大数据治理作为一种制度与技术双驱动的治理模式,将从治理理念到治理实践对国家治理现代化产生变革性影响。本文利用文献计量和综述对中国有关大数据治理的研究进行分析,从文献分布与研究主题整体来看,大数据治理研究受到了国内学术界的广泛关注,其重要性日渐凸显,相关研究主题也展现了广泛多样性和多学科交融等特征。

但大数据治理作为一个开放性议题和动态性过程,目前国内有关研究还是多以概念性、描述性的定性分析为主,偏重于理论的论述和经验的总结,针对大数据治理面临的实际关键问题,还缺少相关的实践案例支撑和技术应用研究,很难对大数据治理的整体情况进行全面反映。此外,大数据治理体系的设计和评估还有待完善和优化,有关大数据治理价值释放、成熟度评估、数据政策标准、数据质量、数据安全隐私等主题需要被给予更多的关注,大数据治理这一问题的研究也应被学术界给予更深的思考。

参考文献:

- [1] 韩兆柱, 马文娟. 数字治理理论研究综述[J]. 甘肃行政学院学报, 2016(1): 23-35.
HAN Z Z, MA W J. Digital governance theory literature review[J]. Journal of Gansu administration institution, 2016(1): 23-35.
- [2] 桑尼尔·索雷斯. 匡斌, 译. 大数据治理[M]. 北京: 清华大学出版社, 2014.
SUNIL S. KUANG B(trans). Big data governance[M]. Beijing: Tsinghua university press, 2014.
- [3] 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, 29(4): 65-72.
ZHENG D Q, HUANG L H, ZHANG C H, et al. Concept and reference architecture of big data governance[J]. Research and development management, 2017, 29(4): 65-72.
- [4] 黄璜. 对“数据流动”的治理——论政府数据治理的理论嬗变与框架[J]. 南京社会科学, 2018(2): 53-62
HUANG H. Governance of "flows of data": Theory development and framework of government data governance[J]. Nanjing social science, 2018(2): 53-62.
- [5] 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.
XIA Y K. Analysis on government data governance in open data environment: Conceptual framework and main issue[J]. Documentation, information & knowledge, 2018(1): 95-104.
- [6] 安小米, 白献阳, 洪学海. 政府大数据治理体系构成要素研究——基于贵州省的案例分析[J]. 电子政务, 2019(2): 2-16.
AN X M, BAI X Y, HONG X H. Research on the constituent elements of government big data governance system based on the case analysis of Guizhou province[J]. E-government, 2019(2): 2-16.
- [7] 童楠楠, 朝乐门. 大数据时代下数据管理理念的变革: 从结果派到过程派[J]. 情报理论与实践, 2017, 40(2): 60-65.
TONG N N, CHAO L M. Change of data management concept in big data era: From result-oriented to process-oriented[J]. Information studies: Theory & application, 2017, 40(2): 60-65.
- [8] 张绍华, 潘蓉, 宗宇伟. 大数据治理与服务[M]. 上海: 上海科学技术出版社, 2016: 17.
ZHANG S H, PAN R, ZONG Y W. Big data governance and services[M]. Shanghai: Shanghai science and technology press, 2016: 17.
- [9] 向芳青, 张翊红. 政府实施大数据治理的应用框架构建[J]. 凯里学院学报, 2018, 36(2): 32-38.
XIANG F Q, ZHANG Y H. Construction of the application framework for government to implement big data governance[J]. Journal of Kaili university, 2018, 36(2): 32-38.
- [10] 梁芷铭. 大数据治理: 国家治理能力现代化的应有之义[J]. 吉首

- 大学学报(社会科学版), 2015, 36(2): 34-41.
- LIANG Z M. Mega data governance: An essential approach to the modernization of state governance[J]. Journal of Jishou university (social sciences), 2015, 36(2): 34-41.
- [11] 马广惠, 安小米. 政府大数据共享交换情境下的大数据治理路径研究[J]. 情报资料工作, 2019, 40(2): 62-70.
- MA G H, AN X M. Research on big data governance path in government big data sharing and exchange situation[J]. Information and documentation services, 2019, 40(2): 62-70.
- [12] 戴香智, 马俊达. 大数据时代下的社会治理创新: 概念、关系与路径[J]. 中国科技论坛, 2016(10): 39-44, 52.
- DAI X Z, MA J D. Social governance innovation of big data era: Concept, relationship and path[J]. Forum on science and technology in China, 2016(10): 39-44, 52.
- [13] 石火学, 潘晨. 大数据驱动政府治理变革[J]. 电子政务, 2018(12): 112-120.
- SHI H X, PAN C. Big data-driven government governance reform[J]. E-government, 2018(12): 112-120.
- [14] 张明斗, 刘奕. 基于大数据治理的城市治理现代化体系研究[J]. 电子政务, 2020(3): 91-99.
- ZHANG M D, LIU Y. Research on the modernization system of urban governance based on big data governance[J]. E-government, 2020(3): 91-99.
- [15] 刘越男, 闫慧, 杨建梁, 等. 大数据情境下政府治理研究进展与理论框架构建[J]. 图书与情报, 2017(1): 87-93.
- LIU Y N, YAN H, YANG J L, et al. Governance in the context of big data: Research advances and theory framework[J]. Library & information, 2017(1): 87-93.
- [16] 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.
- WU S P, LI P, ZHANG Z F. Design of data governance framework in government big data environment[J]. E-government, 2019(2): 45-51.
- [17] 曾小锋. 大数据时代政府治理面临的双重境遇与突破路径[J]. 领导科学, 2016(8): 18-20.
- ZENG X F. The dual situation and breakthrough path of government governance in big data era[J]. Leadership science, 2016(8): 18-20.
- [18] 安小米, 白献阳, 洪学海. 政府大数据治理体系构成要素研究——基于贵州省的案例研究[J]. 电子政务, 2019(2): 2-16.
- AN X M, BAI X Y, HONG X H. Research on the constituent elements of government big data governance system based on the case analysis of Guizhou province[J]. E-government, 2019(2): 2-16.
- [19] 翟云. 中国大数据治理模式创新及其发展路径研究[J]. 电子政务, 2018(8): 12-26.
- ZHAI Y. Research on the innovation model and development path of big data governance in China[J]. E-government, 2018(8): 12-26.
- [20] 黄敏杰. 基于大数据环境下的数据治理框架研究及应用[J]. 信息系统工程, 2018(11): 61.
- HUANG M J. Research and application of data governance framework based on big data environment[J]. China CIO news, 2018(11): 61.
- [21] 刘叶婷, 唐斯斯. 大数据对政府治理的影响及挑战[J]. 电子政务, 2014(6): 20-29.
- LIU Y T, TANG S S. Impacts and challenges of big data on government governance[J]. E-government, 2014(6): 20-29.
- [22] 程广明. 大数据治理模型与治理成熟度评估研究[J]. 科技与创新, 2016(9): 6-7.
- CHENG G M. Research on big data governance model and governance maturity evaluation[J]. Science and technology & innovation, 2016(9): 6-7.
- [23] 杨琳, 高洪美, 宋俊典, 等. 大数据环境下的数据治理框架研究及应用[J]. 计算机应用与软件, 2017, 34(4): 65-69.
- YANG L, GAO H M, SONG J D, et al. Research and application of data governance framework in big data environment[J]. Computer applications and software, 2017, 34(4): 65-69.
- [24] 刘驰, 胡柏青, 谢一, 等. 大数据治理与安全. 从理论到开源实践[M]. 北京: 机械工业出版社, 2017: 7.
- LIU C, HU B Q, XIE Y, et al. Big data governance and security, from theory to implementation[M]. Beijing: China machine press, 2017: 7.
- [25] 尧滢, 夏志杰. 政府大数据治理体系下的实践研究——基于上海、北京、深圳的比较分析[J]. 情报资料工作, 2020, 41(1): 94-101.
- YAO G, XIA Z J. Research on practice government big data governance system: Comparative analysis of Shanghai, Beijing and Shenzhen[J]. Information and documentation services, 2020, 41(1): 94-101.

- 101.
- [26] 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.
MA G H, AN X M, SONG Y. Business-driven government big data platform data governance[J]. Information and documentation services, 2018(1): 21-27.
- [27] 刘冰, 庞琳. 国内外大数据质量研究述评[J]. 情报学报, 2019, 38(2): 217-226.
LIU B, PANG L. Review of domestic and international research on big data quality[J]. Journal of the China society for scientific and technical information, 2019, 38(2): 217-226.
- [28] 莫祖英, 白清礼, 马费成. “政府公开信息质量”概念及内涵解析[J]. 情报杂志, 2015(10): 151-155, 116.
MO Z Y, BAI Q L, MA F C. Quality of government public information: the concept and its connotation[J]. Intelligence journal, 2015(10): 151-155, 116.
- [29] 朱建平, 陈飞. 统计数据质量评价体系探讨[J]. 商业经济与管理, 2010(12): 77-81.
ZHU J P, CHEN F. The research on evaluation system of statistical data quality[J]. Journal of business economics, 2010(12): 77-81.
- [30] 续瑾成. 浅谈数据治理在数据质量管理中的作用[J]. 中国管理信息化, 2016, 19(18): 192.
XU J C. The role of data governance in data quality management[J]. China management informatization, 2016, 19(18): 192.
- [31] 宗威, 吴锋. 大数据时代下数据质量的挑战[J]. 西安交通大学学报(社会科学版), 2013, 33(5): 38-43.
ZONG W, WU F. The challenge of data quality in the big data era[J]. Journal of Xi'an Jiaotong university (social sciences), 2013, 33(5): 38-43.
- [32] 肖人毅. 云计算中数据隐私保护研究进展[J]. 通信学报, 2014, 35(12): 168-177.
XIAO R Y. Survey of privacy preserving data queries in cloud computing[J]. Journal of communications, 2014, 35(12): 168-177.
- [33] 方滨兴, 贾焰, 李爱平, 等. 大数据隐私保护技术综述[J]. 大数据, 2016, 2(1): 1-18.
FANG B X, JIA Y, LI A P, et al. Privacy protection in big data: A survey[J]. Big data research, 2016, 2(1): 1-18.
- [34] 刘桂锋, 阮冰颖, 刘琼. 加强数据安全防护提升数据治理能力——《中华人民共和国数据安全法(草案)》解读[J]. 农业图书情报学报, 2021, 33(4): 4-13.
LIU G F, RUAN B Y, LIU Q. Enhance data security governance capability: Interpretation of Data Security Law of the People's Republic of China (Draft)[J]. Journal of library and information science in agriculture, 2021, 33(4): 4-13.
- [35] 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.
ZHANG Y J, AN X M, ZHANG G Q. Construction of maturity evaluation index system for government big data governance[J]. Information and documentation services, 2018(1): 28-32.
- [36] 郑伟, 张立民, 杨莉. 试析大数据环境下的数据式审计模式[J]. 审计研究, 2016(4): 20-27.
ZHENG W, ZHANG L M, YANG L. Research on data-oriented audit mode in a big data environment[J]. Auditing research, 2016(4): 20-27.
- [37] 胡税根, 单立栋, 徐靖芮. 基于大数据的智慧公共决策特征研究[J]. 浙江大学学报(人文社会科学版), 2015, 45(3): 5-15.
HU S G, SHAN L D, XU J R. Characteristics of big-data-based smart public decision making[J]. Journal of Zhejiang university (humanities and social sciences), 2015, 45(3): 5-15.
- [38] 单志广, 房毓菲, 王娜. 大数据治理: 形势, 对策与实践[M]. 北京: 科学出版社, 2016: 141.
SHAN Z G, FANG Y F, WANG N. Big data governance: Situation, strategy and practice[M]. Beijing: Science press, 2016: 141.
- [39] 马琳. 大数据时代下的社会管理创新[J]. 领导科学, 2013(32): 23-24.
MA L. Social management innovation in big data era[J]. Leadership science, 2013(32): 23-24.
- [40] 赵亚珠, 张晓. 大数据时代的社会公共管理模式创新[J]. 经济研究导刊, 2017(28): 182-183.
ZHAO Y Z, ZHANG X. Social public innovation management mode in big data era[J]. Economic research guide, 2017(28): 182-183.
- [41] 苏鑫, 吴迎亚, 裴华健, 等. 大数据技术在过程工业中的应用研究进展[J]. 化工进展, 2016, 35(6): 1652-1659.
SU X, WU Y Y, PEI H J, et al. Recent development of the application of big data technology in process industries[J]. Chemical industry

- try and engineering progress, 2016, 35(6): 1652-1659.
- [42] 张清辉, 陈昊. 大数据背景下商业模式研究回顾与展望[J]. 中国管理信息化, 2015, 18(7): 159-161.
- ZHANG Q H, CHEN H. Research review and prospect of business model in big data era[J]. China management informatization, 2015, 18(7): 159-161.
- [43] 柴洪峰. 金融大数据及银行卡产业大数据实践[J]. 上海金融, 2013(10): 27-29, 116.
- CHAI H F. Big data in finance industry and the big data practice in bank card industry[J]. Shanghai finance, 2013(10): 27-29, 116.
- [44] 彭作文, 苏鸿. 大数据在科技管理工作中的应用[J]. 科技风, 2018(19): 55.
- PENG Z W, SU H. Application of big data in science and technology management[J]. Technology wind, 2018(19): 55.
- [45] 王明平. 基于学生核心素养的智慧校园构建研究与实践[J]. 教育信息化论坛, 2017(1): 7-9.
- WANG M P. Research and practice on the construction of smart campus based on students' core literacy[J]. Education information forum, 2017(1): 7-9.
- [46] 张振, 周毅, 杜守洪, 等. 医疗大数据及其面临的机遇与挑战[J]. 医学信息学杂志, 2014, 35(6): 2-8.
- ZHANG Z, ZHOU Y, DU S H, et al. Medical big data and its opportunities and challenges[J]. Journal of medical informatics, 2014, 35(6): 2-8.
- [47] 孙坦, 黄永文, 鲜国建, 等. 新一代信息技术驱动下的农业信息化发展思考[J]. 农业图书情报学报, 2021, 33(3): 4-15.
- SUN T, HUANG Y W, XIAN G J, et al. Considerations for the development of agricultural informatization driven by a new generation of information technologies [J]. Journal of library and information science in agriculture, 2021, 33(3): 4-15.
- [48] 李娜, 田英杰, 石勇. 论大数据在环境治理领域的运用[J]. 环境保护, 2015, 43(19): 30-33.
- LI N, TIAN Y J, SHI Y. Research on the application of big data in environmental governance[J]. Environmental protection, 2015, 43(19): 30-33.
- [49] 王静远, 李超, 熊璋, 等. 以数据为中心的智慧城市研究综述[J]. 计算机研究与发展, 2014, 51(2): 239-259.
- WANG J Y, LI C, XIONG Z, et al. Survey of data-centric smart city[J]. Journal of computer research and development, 2014, 51(2): 239-259.
- [50] 熊文景. 重大疫情防控视野下的数据治理: 主要价值、现实困境与优化路径[J]. 山西档案, 2020, 4(3): 22-28.
- XIONG W J. Data governance from the perspective of major epidemic prevention and control: Main value, realistic dilemma and optimization path[J]. Shanxi archives, 2020, 4(3): 22-28.
- [51] SAYOGO D S, PARDO T A. Exploring the determinants of publication of scientific data in open data initiative[C]//International conference on theory and practice of electronic governance, ACM, 2011: 97-106.
- [52] 戴炳荣, 闭珊珊, 杨琳, 等. 数据资产标准研究进展与建议 [J]. 大数据, 2020, 6(3): 36-44.
- DAI B R, BI S S, YANG L, et al. Research status quo and suggestions on data assets standardization[J]. Big data research, 2020, 6(3): 36-44.
- [53] 梅宏. 数据治理之论[M]. 北京: 中国人民大学出版社, 2020: 44.
- MEI H. Data governance[M]. Beijing: Renmin university press, 2020: 44.

A Review of Big Data Governance Research in China

SONG Shanshan^{1,2}, BAI Wenlin^{3*}

(1. Wuhan library, Chinese Academy of Sciences, Hubei Wuhan 430071; 2. Department of Library, information and Archives Management, School of Economics and Management, University of Chinese Academy of Sciences, Beijing 100049;
3. Management School, Tianjin Normal University, Tianjin 300387)

Abstract: [Purpose/Significance] With the rapid development of information technology and the arrival of the era of big data, big data has become a key production factor and a strategic fundamentally resource for a country. Big data governance has become an important way to fully release data value and realize the modernization of national governance system and governance capacity. This study reviews the related research on big data governance, in order to better play its positive role as a new production factor and a means of social governance. [Method/Process] On the basis of bibliometrics and content analysis, this study systematically reviews the relevant research on big data governance in China, and finds that the research on big data governance has been rising in recent years. The relevant research topics have also displayed characteristics such as wide diversity and multidisciplinary integration. The core topics involved in the research include big data governance, social governance, government governance, big data technology, data resources, network public opinion, risk governance, and data security. The research on the big data governance system of the whole process is the most critical part. [Results/Conclusions] This study classified the relevant research topics into three aspects: theoretical system, implementation evaluation and practical challenges. The big data governance system is a complex, multi-level and multi-dimensional systematic project, which consists of governance goals, governance subjects, governance objects and governance tools. Big data governance implementation and evaluation includes three stages: organizational planning, specific implementation, evaluation and optimization. The organizational planning stage includes formulating implementation guidelines and constructing governance organizations; the specific implementation stage includes metadata management, master data management, data quality and data security; the evaluation optimization phase includes big data governance maturity evaluation and big data governance audit. The practice of big data governance focuses on three aspects: the Chinese government actively uses big data to improve scientific decision-making ability, social management level and public service construction. The practical challenges are mainly reflected in the management and technical challenges faced by data-centered fields. As an open topic and dynamic process, big data governance is still mostly based on conceptual and descriptive qualitative analysis, and focuses on theoretical discussion and experience summary. There is still a lack of relevant case support and technology application research, and the design and implementation evaluation of the big data governance system still needs to be improved and optimized.

Keywords: big data; big data governance; governance system; governance implementation evaluation