

引文格式：

樊建平, 孙婧, 李红辉, 等. 政务大数据管理技术研究进展 [J]. 集成技术, 2023, 12(3): 1-18.

Fan JP, Sun J, Li HH, et al. Research progress of government big data management technology [J]. Journal of Integration Technology, 2023, 12(3): 1-18.

政务大数据管理技术研究进展

樊建平¹ 孙婧^{1*} 李红辉² 张用川³ 朱利⁴ 仇阿根⁵ 马英⁶
杨晓睿² 段宇航²

¹(中国科学院深圳先进技术研究院 深圳 518055)

²(北京交通大学 北京 100044)

³(重庆交通大学 重庆 400074)

⁴(西安交通大学 西安 710049)

⁵(中国测绘科学研究院 北京 100039)

⁶(国家信息中心 北京 100038)

摘要 近年来, 互联网+政务蓬勃发展, 促进了互联网与政府公共服务体系的深度融合, 数字政府建设已初见成效。如何实现互联网+政务大数据的规范、高效、安全管理、共享和应用, 推动解决“三融五跨”问题, 是高质量建设数字政府、推进国家治理体系和治理能力现代化的关键。该文首先综述了国内外政务数据管理发展历程; 然后, 围绕互联网+政务大数据管理全生命周期中采集技术、存储技术、融合技术、计算技术、应用技术及安全技术进行阐述, 总结了主流数据管理技术演进历程; 最后, 探讨了互联网+政务数据管理中若干亟待解决的问题, 明确了互联网+政务大数据管理技术的未来发展趋势。

关键词 互联网+; 政务大数据; 数据管理

中图分类号 D 63; TP 311 **文献标志码** A **doi**: 10.12146/j.issn.2095-3135.20221205001

收稿日期: 2022-12-05 修回日期: 2023-01-13

基金项目: 国家重点研发计划项目(2019YFB2102500)

作者简介: 樊建平, 研究员, 研究方向为高性能计算机、云计算、并行与分布式计算; 孙婧(通讯作者), 高级工程师, 研究方向为模式识别、城市计算, E-mail: jing.sun1@siat.ac.cn; 李红辉, 研究员, 研究方向为大数据技术与工程、人工智能、网络与信息安全; 张用川, 讲师, 研究方向为智慧城市、地理空间大数据挖掘、社会感知计算; 朱利, 教授, 研究方向为跨媒体计算、机器学习和群体智能; 仇阿根, 研究员, 研究方向为政务地理信息服务、信息资源整合、时空地理信息大数据、应急地理信息平台等; 马英, 高级工程师, 研究方向为政务数据治理; 杨晓睿, 博士研究生, 研究方向为大数据技术与工程、机器学习; 段宇航, 博士研究生, 研究方向为大数据技术与工程、机器学习。

Research Progress of Government Big Data Management Technology

FAN Jianping¹ SUN Jing^{1*} LI Honghui² ZHANG Yongchuan³ ZHU Li⁴ CHOU Agen⁵
MA Ying⁶ YANG Xiaorui² DUAN Yuhang²

¹(Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China)

²(Beijing Jiaotong University, Beijing 100044, China)

³(Chongqing Jiaotong University, Chongqing 400074, China)

⁴(Xi'an Jiaotong University, Xi'an 710049, China)

⁵(Chinese Academy of Surveying and Mapping, Beijing 100039, China)

⁶(State Information Center, Beijing 100038, China)

*Corresponding Author: jing.sun1@siat.ac.cn

Abstract In recent years, the vigorous development of “Internet + Government Services” has promoted the deep integration of the Internet and the government public service system, and the construction of digital government has achieved initial results. How to achieve standardized, efficient, safe management, sharing and application of Internet + big data in government affairs, and promote the solution of the “Three Integrations and Five Crossings Issues”, is the key to building a higher quality digital government and promoting the modernization of national governance system and capacity. This paper first summarizes the development of government data management at home and abroad. Secondly, it expounds the collection technology, storage technology, fusion technology, computing technology, application technology and security technology in the full life cycle of Internet + government big data management, and summarizes the evolution of mainstream data management technology. Finally, this paper discusses some urgent problems in Internet + government data management, and the development trend of Internet + government data management technologies.

Keywords Internet +; big data of government affairs; data management

Funding This work is supported by National Key Research and Development Program of China (2019YFB2102500)

1 引 言

随着互联网、云计算、大数据和人工智能等网络信息技术的发展,利用新兴技术实现高效、便捷、安全的政务管理和智能服务,进而提高政务服务质量,提升政府治理体系和治理能力现代化水平,越来越受到学界和各地政府部门的重视^[1]。发达国家大力推进数字政府建设工作,美国推出了“美国数字政府服务(United States Digital Service)”和“Data.gov 计划”^[2],英国推

出了“政府转型战略”和“英国数字战略”^[3],新加坡推出了“智慧国建设计划”^[4]。我国也利用新型网络信息技术赋能、协同、重构政务活动,正经历从“电子政务”向“数字政府”阶段的跨越。2022年6月,国务院发布了《国务院关于加强数字政府建设的指导意见》(国发〔2022〕14号),为全面开创数字政府建设新局面作出了部署。

众多学者从社会形态演进^[5-6]、技术标准^[7-8]、法律法规^[9-10]、规划建设^[11]等角度出发,对数字

政府建设进行了探讨。其中, 支撑数字政府运行的政务数据和政务数据管理是核心议题。良好的政务数据管理不仅可为政务服务提供数据支持, 还可实现政府管理模式的转变, 优化政府内部组织架构, 简化各类政务服务流程, 推进数字化政府建设, 加快数字政府和数字经济发展。因此, 政务数据管理计划^[12]、管理方案^[13-16]、数据管理安全^[17-21]等成为研究热点。2016年, 国家印发《国务院关于加快推进“互联网+政务服务”工作的指导意见》(国发〔2016〕55号), 以“互联网+服务”推动我国政务信息化建设成为主旋律, 互联网+政务大数据应运而生。随着互联网+政务服务体系建设的不断成熟, 以数据汇聚融合、共享开放和开发利用为核心内容的政务大数据管理越来越凸显出重要性。2022年10月, 国务院办公厅印发了《国务院办公厅关于印发全国一体化政务大数据体系建设指南的通知》(国办函〔2022〕102号), 对政务大数据体系建设提出了新要求。鉴于当前研究缺乏从生命周期管理技术(政务大数据采集、预处理、存储、计算及应用等)角度的研究与总结, 为帮助学者尽快了解政务大数据管理方向的前沿研究, 促进政务大数据

管理技术的进一步发展创新, 本文对政务大数据管理过程中的相关技术进行了总结阐述。

本文首先介绍国内外电子政务发展历程中的数据管理演变过程; 然后, 介绍了政务数据生命周期管理的概念及相关模型, 针对政务数据的生命周期过程, 包括数据采集、存储、融合、计算、应用、安全等相关技术, 进行了现状总结; 最后, 探讨了政务大数据管理的研究挑战与展望。

2 互联网+政务大数据管理发展历程

政务大数据指国家组织机构通过现代信息技术, 向社会提供管理与服务过程中产生的数据^[22], 包括社会管理、公共服务、专项调查等活动产生的数据, 其数据类型丰富、数据量大、时间跨度久远、地域范围广阔^[23]。根据数据类型、数据属性、数据敏感程度, 对政务数据进行分类, 如表1所示^[24]。

结合国外电子政务的发展情况, 国外政务数据管理的发展可分为6个阶段: 人工管理阶段、穿孔卡片记录管理阶段、程序记录管理器阶段、

表1 政府数据分类

Table 1 Classification of government data

分类方式	类别名称	举例
数据类型	政府有权采集的数据	资源类、税收类、财政类等
	政府可获取的数据	生产建设数据、农业数据、工业数据等
	政府发起产生的数据	城市基建、医院、教育师资等数据
	政府监管职责所拥有的数据	人口普查、金融监管、食品药品管理等
数据属性	政府提供服务所产生的消费和档案数据	社保、水电、公安等
	自然信息类	地理、资源、气象、环境、水利等
	城市建设类	交通设施、旅游景点、住宅建设等
	城市管理统计监察类	如工商、税收、人口、机构、企业、商品等
敏感程度	服务与民生消费类	水、电、燃气、通信、医疗、出行等
	非敏感数据	公开数据
	内部数据	涉及用户隐私数据
	涉密数据	涉及国家秘密数据

在线网络数据库管理阶段、关系数据库和客户端-服务器计算方式阶段、多媒体数据库结合大数据技术阶段^[25-26]，如表2所示。

我国政务数据管理发展大致可分为5个阶段，目前正经历以服务数字政府与数字经济为主要任务的互联网+政务大数据管理阶段^[27-28]。我国政务数据管理方式发展过程如表3所示，包括：(1)人工处理数据阶段。20世纪80年代之前，我国对政务数据的管理主要依靠人工的方式，此时数据处于分散无序的状态^[29]。(2)数字化存储管理阶段。始于20世纪80年代中期，办公自动化的概念被引入中国^[30]，使我国政务数据的管理进入数字化存储管理阶段。在这一阶段，政府部门认识到政务数据的价值，开始重视数据质量，根据管理要求有目的地收集和使用数据，并建立数据管理团队负责数据管理。然而，政务数据依然保存在分散独立的管理信息系统中，虽然数据更新速度加快，但数据在机构内还不能共享。(3)数据库管理阶段。20世纪90年代中期至21世纪初，政府上网工程得到了各级政府部门的

积极响应，政府内部建立了专门的数据库管理机构，负责政务数据管理^[30]，为政务数据构建政务数据目录，并着手建立标准，力求打通数据的界限和壁垒。(4)网络化管理阶段。21世纪初，随着互联网和云计算技术的发展，我国政务数据管理进入新的网络化管理阶段，政府对数据质量的重视程度进一步加深，提升数据质量被视为政府机构的首要任务^[31]；数据共享被提上日程，如共享目录、数据标准体系等，各政府机构可以利用共享服务实现数据的跨部门使用。(5)互联网+政务数据管理阶段。2016年前后，国家出台加快推进互联网+政务服务指导意见，更加注重通过信息化手段实现政府部门之间网络的串联串通、数据的整合共享、业务流程的协同高效，数据的质量管理，以及数据资源的高效利用^[32]。该阶段的目标是全面加强政府数据资产管理，统筹规划、整合开放各类数据资源，运用大数据、人工智能等技术，促进数据要素有组织成规模地流动应用，使各类主体之间更容易获取信息、交互信息，充分发挥数据价值，为政府规划决策、企

表2 国外政务数据管理发展过程

Table 2 Development of foreign government data management

阶段名称	时间	特点
人工管理阶段	1900年之前	依靠人工统计
穿孔卡片记录管理阶段	1900—1955年	穿孔卡片方式代替手动记录
程序记录管理器阶段	约1955—1970年	磁盘存储
在线网络数据库管理阶段	约1960—1980年	使用专门的在线网络数据库管理系统
关系数据库和客户端-服务器计算方式	约20世纪末—21世纪初	面向集合访问、分布式批处理计算技术
多媒体数据库结合大数据技术阶段	21世纪以来	政府内部专门的数据库管理系统、大数据技术

表3 中国政务数据管理发展过程

Table 3 Development of domestic government data management

阶段名称	时间	特点
人工处理数据阶段	20世纪80年代之前	纸质化存储、人工办理业务
数字化存储管理阶段	始于20世纪80年代中期	磁盘存储、办公自动化改造
数据库管理阶段	20世纪90年代中期—21世纪初	使用专门的在线网络数据库管理系统
网络化管理阶段	21世纪初	面向集合访问、分布式批处理计算技术
互联网+政务大数据管理	2016年前后至今	数据共享、数据治理、支持政务协同办理

业创新发展、居民便捷办事提供强大助力^[33]。

纵观国内外电子政务的发展历程, 政务数据管理技术也在不断演变。最初, 政府部门缺乏政务数据管理的专门技术和方法, 数据获取方式主要为人工采集, 采集的数据没有统一标准, 文件格式多种多样, 政务数据存储在各自为政的管理信息系统中, 几乎没有数据分析; 随着数据库技术的发展, 开始出现统一的文件格式, 并开始注重开发、存储和使用元数据文件, 开发者开发出组织层面的共享目录、数据转换标准、语义模型等; 现阶段, 已有相关政策全面强调政务数据对管理和决策的驱动价值, 数据已成为管理和决策的赋能者。

3 互联网+政务大数据管理技术

政务数据生命周期管理, 是一种从数据输入到数据销毁整个生命周期管理数据的方法。段尧清等^[34]提出了政府开放数据全生命周期的概念, 构建了政府开放数据全生命周期概念模型。樊建平等^[35-36]提出了互联网+政务大数据透明管理概念, 构建了全生命周期的互联网+政务大数据透明管理技术模型。

本节将重点阐述政务大数据的管理技术和方法, 将诸多技术按照数据生命周期组织为互联网+政务大数据管理框架, 该框架由 6 部分组成: 政务数据采集管理、政务数据存储管理、政务数据融合管理、政务数据计算管理、政务数据应用管

理和政务数据安全, 如图 1 所示。



图 1 政务大数据管理技术框架

Fig. 1 Technical framework for government big data management

3.1 政务数据采集管理

政务数据采集是解决政务大数据管理的首要环节^[37], 根据政务数据采集的自动化程度, 可将其划分为人工采集、半自动采集和自动采集 3 类技术, 相关案例对比分析如表 4 所示。

在人工采集方面, 王薇等^[38]探究了人工统计方式在大数据领域中的应用, 但人工采集费时费力且易出现差错, 导致记录结果可靠性较差。为尽可能减少人为环节的误差, 半自动采集技术逐步得到了发展。武国雄^[39]提出利用 VBA 语言编程实现道路数据的半自动采集, 极大提高了数据

表 4 政务数据采集技术案例对比分析

Table 4 Comparative analysis of government data acquisition technology cases

案例分类	研究问题	所用技术	效果
人工采集	统计和大数据交叉的问题 ^[38]	人工截图整理, 录入工作表	人工采集费时费力, 且容易混入人工统计的误差, 可靠性不强
半自动采集	人工采集较费力, 结果误差较大 ^[39]	VBA 语言编程、自动分段自动排序+人工判读	极大提高效率, 稳定产品质量
自动采集	公交数据采集系统的研究 ^[40]	现有自动采集设备和数据仓库技术	实现公交运行数据、乘车人信息和 IC 卡数据高效采集和处理

采集效率，保证了测量精度。随着自动化技术的发展，数据自动采集技术逐渐发展成熟，并趋于完善。美国政府为提升政务数据管理能力和推动政府管理改革，提出了美国政府大数据计划^[40]，开发了“一站式数据下载”网站，将各联邦政府、州政府、部门机构和民间组织数据进行汇总和整合。

随着对政务数据采集实时性和精确性的要求不断提高，政务数据采集技术正由人工采集向半自动采集、半自动采集向自动采集的方向发展^[41-43]，逐步实现速度快、精度高、收集范围广的目标。

3.2 政务数据存储管理

稳定安全的数据存储是政务服务健康发展的基础，随着政务大数据管理的快速发展，数据呈现出来源广、规模大等特点，为充分发挥政务数据信息的应用价值，有效稳定的存储技术已成为领域内的研究热点^[44]。政务大数据存储技术可分为键-值存储、面向列存储、文件存储、图存储和云存储 5 种类型，各类政务数据存储技术案例对比分析如表 5 所示。

在键-值存储类型中，宋云奎等^[45]将网络边缘部分产生的电网相关数据传送至数据中心，运用 Redis 数据库技术，提出键-值存储的数据存储结构，利用 Redis 缓存机制优势，极大地降低了内存占有率，加快了数据存储速度。

在面向列存储类型中，符宁^[47]提出了基于分布式文件、关系数据、内存数据库、列式数据库的集合式政务数据存储方案，形成了全量数据存储和专题库存储，可辅助科技管理部门布局科技战略，提升个性化政务服务的能力。

在文件存储类型中，刘云汉等^[48]为存储多时态、数量大、存储规模不一的卫星影像数据文件，利用 MongoDB 分片复制集群的优势，结合自动分片技术，实现了海量卫星文件数据的安全高效存储。

图存储作为现代大数据存储的新兴研究方向之一，已有大量学者开展了相关领域的探索研究。张维冲等^[49]为对大规模、碎片化公文数据进行聚合，利用 Neo4j 图数据库在多粒度知识间分析和推理的优势，结合 Cypher 语言，实现了公文数据的查询、推理等操作。

为应对硬件存储空间有限和数据量无限的现实冲突，有关云存储的研究正逐步得到发展。秦晓东^[50]为构建较为完善的数据资源体系，将政务数据资源按照使用场景，以数仓分层形式进行整合，并将其存储在统一数据“资源云”中，以满足不同类型数据存储的需要，并为上层应用场景提供支持。刘应新等^[51]为响应政务云平台建设的政策方针，提出了 FC SAN 阵列存储和分布式块存储方案，为公众需要和计算区块提供了云存

表 5 政务数据存储技术案例对比分析

Table 5 Comparative analysis of government data storage technology cases

案例分类	研究问题	所用技术	效果
键-值存储	大量数据在网络边缘部分 ^[45]	Redis 技术	极大降低内存占有率，加快数据存储的时间
面向列存储	基于不同数据存储方案研究性能和体验 ^[46-47]	全量数据存储、专题库	辅助科技管理部门布局科技战略，提升个性化政务服务能力
文件存储	数据文件大小不一，单台服务器难以支撑海量存储 ^[48]	MongoDB 自动分片技术	实现卫星文件数据的安全高效存储
图存储	知识规模大、碎片化 ^[49]	Neo4j 知识图谱	支持多粒度知识间分析和推理
	构建资源体系 ^[50]	数据仓库技术	满足不同类型数据存储的需求支撑上层应用场景
云存储	建设政务云平台 ^[51]	IaaS 层云存储技术	降低建设成本，全面推进业务协同和信息共享，有效解决信息孤岛问题
	解决未来数据存储资源不够等问题 ^[52]	DNA 存储技术	使得“无限量”的政务数据存储成为可能

储资源, 推动了云计算、政务大数据在政府管理和公共服务等方面的应用。为解决数据存储资源不足的问题, 张大璐等^[52]提出一种基于 DNA 的数据存储技术, 以满足“无限量”的政务数据存储, 使数据无上限存储成为可能。

政务数据存储在政务大数据管理中作为中间层的支撑角色, 正逐步从本地直连向网络化、分布式方向发展^[53-54]。随着数据规模的不断扩大, 云存储技术正处于快速发展的阶段^[55]。此外, 政务数据存储的无限扩展性^[56-58]、安全可靠也随着存储技术的发展得到不断提升。

3.3 政务数据融合管理

政务数据来源广且收集标准不统一, 导致获得的政务数据存在信息源多样化、结构差异大、对齐难等问题^[59]。因此, 政务数据融合技术在政务大数据管理中至关重要。根据融合级别的不同, 可将数据融合技术分为数据级融合、特征级融合和决策级融合, 技术对比案例分析如表 6 所示。

数据级融合指直接把采集的多维原始数据进行拼接、关联和融合的过程。吴文莉等^[54]为打破多维传感器数据的融合局限性, 解决单维数据易受干扰的问题, 提出了基于相关性方差贡献率的数据级融合方法, 融合结果表明, 该方法具有良

好的抗噪性。杜刚^[60]为避免现场监测数据不准确, 提出基于 JDL 模型、Dasarathy 模型和 Boyd 控制环的监测数据融合模型, 有效缩短了应急响应时间。Khaleghi 等^[61]提出一种以数据为中心的数据融合方法, 并给出了该数据融合算法在机器人、传感器网络和图像处理等领域的应用案例。

特征级融合属于特征工程中间层融合的范畴, 在融合过程中, 首先对原始信息进行特征抽取, 然后对信息进行综合分析处理。沈王恒^[62]针对政务数据平台治理能力提升的需求, 利用 NLP、ETL 等工具, 实现了政务数据的深度特征融合, 为政务多源数据融合提供了系统化的实践经验和参考价值。王红霞等^[63]针对电力物联网数据融合需求, 提出一种使用电力结构化时序参量和非结构化图像参量的普适性融合框架, 实验结果表明, 该融合感知模型具有较好的容错性。徐济宣等^[64]提出一种基于深度神经网络的多源故障特征融合方法, 实验结果表明, 与单一特征相比, 融合特征的诊断精度提高了 10%。

决策级融合包括预处理、特征抽取、识别或判决步骤, 以建立对所观测目标的初步结论, 然后通过关联处理进行决策层融合判决, 最终获得联合推断结果。王长海等^[65]针对现有嵌入式平台行人检测准确率低和实时性差等问题, 提出一种

表 6 政务数据融合技术案例对比分析

Table 6 Comparative analysis of government data fusion technology cases

案例分类	研究问题	所用技术	效果
数据级融合	传感器易受干扰导致的数据质量问题 ^[54]	基于相关性方差贡献率的数据集融合技术	具有良好抗噪性, 可完整表现出整体动力特性
	现场监测数据不准确 ^[60]	基于 JDL 模型、Dasarathy 模型和 Boyd 控制环的监测数据融合	有效避免因失误导致应急响应时间增加
	数据不完善、虚假数据、数据冲突等 ^[61]	新的以数据为中心的数据融合技术	可扩展到多个应用领域
特征级融合	政务数据平台能力提升需求 ^[62]	NLP、ETL 等技术	完善政务多源数据融合体系
	现有异构多源融合技术薄弱、无法满足融合需求 ^[63]	神经网络	具有较好的容错性
	提高故障诊断精度 ^[64]	利用 AE、CNN、DNN 等构造融合网络	精度较单一特征高
决策级融合	准确率低, 实时性差 ^[65]	DBSCAN 算法、概率数据关联算法 决策级融合	提高检测精度, 用时大大降低
	情报之间不能有效融合 ^[66]	天网情报决策级融合分析技术	实现了多种情报数据有效融合, 为军事决策提供支持

决策级融合点云和图像的道路行人检测方法,采用 DBSCAN、概率数据关联等算法,将检测结果进行决策级融合,提高检测精度的同时,缩短了检测用时。杨晓冬等^[66]针对侦察情报和开源情报融合协同问题,提出一个天网情报决策级融合分析技术体系,建立不同关联关系库,构建天网情报时空图谱,为国防军事决策提供技术支持。

政务数据融合在政务大数据平台中尤为重要,其在多信息源、多平台和多用户系统内起着重要的处理和协调作用,保证了数据处理系统各单元与汇集中心间的连通性与实时性。

3.4 政务数据计算管理

政务数据计算即运用大数据手段对政务数据进行提取、加工和优化的信息化方法。数据计算技术在政务数据管理中发挥着重要作用,它既是数据存储的延伸,又是数据应用的基础。数据只有被整合和计算、挖掘潜在信息,才能充分发挥数据的价值,满足政务管理的需求。随着政务大数据规模的爆炸式增长,大数据计算方法可以总结为分布式计算、增量计算和绿色计算,相关技

术案例对比如表 7 所示。

针对分布式计算,2004年,Google公司提出一种适用于半结构化及非结构化数据的大规模数据批处理模型(MapReduce)。受 MapReduce 编程模型等一系列模型的启发,夏靖波等^[75]对论文中提及的思想进行了编程实现,Hadoop 自此诞生。2018年,基于改良 Hadoop MapReduce 的 Spark 被提出^[76],其具有速度快、易用性强等特点。Chen^[67]针对海量数据信息难以转化为有效数据资源的问题,引入了基于 Spark 技术^[77]的低时延集群分布式计算系统,实现了数据资源的有效提取和转化,但存在一定的局限性,如在某些特定行业,该技术不能满足精度的要求。

对于动态增长的数据集或离散数据,增量计算较为合适。Zhou 等^[70]针对推荐系统数据量大且数据具有动态性的难题,提出了基于奇异值分解的增量算法,该算法提高了海量动态数据的计算效率。Wang 等^[71]针对历史数据计算中,存在大数据计算冗余的问题,提出基于 Jacobi 旋转实现增量式高阶奇异值分解算法,为高效处理实时

表 7 政务数据计算技术案例对比分析

Table 7 Comparative analysis of government data computing technology cases

案例分类	研究问题	所用技术	效果
分布式计算	海量数据信息难以转化为有效数据资源 ^[67]	Spark 技术	实现有效提取和转化数据资源,确定并利用有价值的信息,促进社会经济的发展
	政务数据对时效性要求高 ^[68]	Storm、Spark streaming、Flink 等实时计算技术	及时计算和处理所采集的数据,快速解决政务需求,但存在特定场景精度无法满足的局限性
	政务数据海量、多源、难以治理 ^[69]	Apache Beam 技术和 ETL 工具中的 Kettle 组件	实现对政务大数据的分布式治理,挖掘政务大数据的潜在价值,提高政务数据资源的利用率,但存在无法处理非结构化数据的局限性
增量计算	推荐系统数据量巨大且数据具有动态性 ^[70]	基于奇异值分解的增量算法	提高了海量动态数据的计算效率
	大数据计算对历史数据进行冗余计算 ^[70]	基于 Jacobi 旋转实现增量式高阶奇异值分解算法	对张量表示的大数据进行降维、提取和降噪,为高效处理实时大数据提供了帮助
	大数据质量较低 ^[71]	增量张量流的计算方法	集成处理大数据,为人们提供有价值的服务
绿色计算	物联网制造流产生的海量待分析的复杂数据 ^[72]	基于增量计算的制造流分析的实时异常检测方法	提高了物联网应用的可扩展性
	计算资源需求快速增长,总体能耗不断增加 ^[73]	基于 Hadoop 集群计算机节点性能和能耗计算方法	在性能不变的前提下实现节能计算,从而降低了能源成本
	能源成本不断上升以及能源生产对环境产生危害 ^[74]	一套面向绿色计算的低功耗架构策略	有利于延长相应设备的使用寿命,降低运营成本和功耗,从而促进可持续发展

大数据提供了解决方法。Parto 等^[72]针对物联网制造流产生的海量待分析复杂数据, 提出基于增量计算制造流分析的实时异常检测方法, 提高了物联网应用的可扩展性。

绿色计算指在提高计算资源使用效率的同时, 实现能量消耗最小化。为应对计算资源需求快速增长、总体能耗不断增加的挑战, Ibrahim 等^[73]提出基于 Hadoop 集群计算机节点性能和能耗的计算方法, 降低了能源成本。Nayak 等^[74]针对能源成本不断上升及能源生产对环境产生危害等问题, 提出一套面向绿色计算的低功耗架构策略, 有助于延长相应设备的使用寿命, 降低运营成本 and 功耗, 促进可持续发展。

数据计算作为政务数据管理整个生命周期中最关键、最核心的环节, 是满足政务需求的基石。在计算中发现信息、挖掘知识, 满足上层应用需要, 是数据计算存在的核心价值。通过大数据计算技术, 可为政务数据管理构建自动化、智能化、定制化的服务, 激发政务大数据的潜在价值, 提高政务管理的水平和效率。

3.5 政务数据应用管理

政务大数据应用管理技术主要包括数据共享

和数据可视化两个方面。数据应用技术在数据管理框架中属于顶层应用的角色, 是数据价值创造的关键, 政务数据应用技术案例对比分析如表 8 所示。

政务数据开放共享的过程中, 政府各个部门强调数据的归属权和使用权, 但缺乏可信的数据资产权利确认方案, 导致实体之间的数据流通率较低, 使各实体逐渐成为“数据孤岛”^[86]。联邦学习作为一种新的机器学习范式, 在保证数据安全的前提下, 可解决数据孤岛问题^[87]。Chen 等^[79]针对大数据交换中存在的版权和隐私问题, 提出一种用于大数据交换的分散式解决方案, 有利于在数据交换市场中保护数据的版权和隐私。褚春燕^[80]针对政府数据共享理念薄弱、共享制度不完善、不同部门存在技术差异等问题, 提出了基于区块链技术的政府数据共享模型, 实现了政务数据信息跨地区利用, 有效改善了当前政府数据共享的现状。

数据可视化对于从海量数据中发现规律、增强数据表现、提升交互效率具有重要作用。目前, 数据可视化的概念及相关研究领域不断扩展, 在政务大数据管理领域也有诸多体现。例如, 在新冠肺炎疫情防控期间, 为避免“多头重复部署”等问题,

表 8 政务数据应用技术案例对比分析

Table 8 Comparative analysis of government data application technology cases

案例分类	研究问题	所用技术	效果
数据共享	迁移学习中存在的隐私泄露问题 ^[78]	联邦学习	在不泄露隐私的情况下, 从构建操作数据中学习可转移的知识, 解决了建筑能源管理中的数据孤岛问题
	大数据交换中存在的版权和隐私问题 ^[79]	区块链技术	有利于在数据交换市场中保护版权和隐私
	政府数据共享理念薄弱、共享制度不完善、不同部门存在技术差异 ^[80]	区块链技术	实现了政务数据信息跨地区利用, 推动了政府数据全方面共享, 有效改善了当前政府数据共享的现状
	医疗数据共享中存在的隐私安全问题 ^[81]	区块链技术	提高了医疗信息的安全性, 促进了医疗数据的开放共享
数据可视化	印尼政府政务数据质量较低, 不利于开放共享 ^[82]	数据清理、数据转换	应用于综合公立医院, 大大提升了医院数据集的数据质量, 促进了印尼政府政务数据的开放共享
	疫情防控期间“多头重复部署”问题 ^[83]	后台大数据在线自动比对核验	实时提供疫情形势综合研判可视化分析, 实现了统一管理, 信息共享
	电子政务公开难以得到有效提升和发展 ^[84]	条形图、词云图、可视化地图等数据可视化技术	实现政务信息领域数据的可视化, 符合当前电子政务平台政务公开发展方向, 也为群众提供了诸多便利, 但在可视化上仍存在信息增值力度弱、不深入等不足
	处理政府政务信息时难以精炼信息, 抓住内容要义, 合理把握表达数据的方向 ^[85]	信息动画技术	有利于进行信息分类, 建立信息秩序, 从而正确地传递信息给群众, 促进了政府政务公开和政策解读

福建省依托闽政通 App 平台^[83], 采用后台大数据在线自动比对核验, 实时提供疫情形势综合研判可视化分析, 实现了数据统一管理、信息共享。

数据应用技术作为数据管理中应用层的角色, 在政务大数据管理中发挥着重要的作用。运用大数据应用技术辅助科学决策和社会治理, 可推进政务数据管理相关工作的开展, 使政府决策更加科学、公共服务更加高效。政务大数据蕴含着巨大的经济与社会价值, 其开放共享和可视化的特点, 对于促进政府自身转型、社会需求获取模式转型、产业经济转型以及建设智慧城市具有重要意义。

3.6 政务数据安全

政务数据是政务大数据管理的重点, 在数据管理中发挥着不可或缺的保障作用。现有的数据安全可划分为安全通信、访问控制和安全

管理 3 个方面, 相关技术案例比较分析如表 9 所示。在安全通信方面, 数据加密和数据脱敏等技术已得到广泛应用。针对云平台的大数据安全, 常用的 AES、DES、3DES 算法等均有应用。韩

义森^[88]针对扬州市大数据中心海量政务数据密文信息熵较低的问题, 设计了基于区块链技术的政务信息大数据自动加密算法。该算法通过删除冗余数据, 显著提升了密文信息熵, 保障了智慧城市建设的

安全性。数据脱敏指将敏感信息根据一定规则进行处理, 以提高敏感数据的安全性^[98]。郭赞宇等^[89]针对政务大数据平台无法完全解决数据内部泄露、恶意攻击的问题, 引入数据脱敏机制, 保障了政务内部数据的安全, 但该平台缺少敏感数据地图和敏感数据搜索等功能, 政务数据安全管理有待完善。陈平刚等^[90]针对政府难以从海量网络信息中及时发现隐藏热点舆情的难题, 利用政务大数据智能脱敏技术, 建立了由大数据支撑的政府舆情预警系统, 实现了潜在热点事件的提前预警, 提高了政府决策的前瞻性。王庆德等^[91]针对政务信息资源共享可能遭遇网络攻击、政务数据云平台存在安全隐患等问题, 提出了基于云的电子政务信息资源共享安全保障体系。

在数据访问控制方面, 常用的数据安全技术有安全认证技术和防火墙技术。安全认证技术通

表 9 政务数据安全案例对比分析

Table 9 Comparative analysis of government data security technology cases

案例分类	研究问题	所用技术	效果
	扬州市大数据中心海量政务数据密文信息熵较低 ^[88]	区块链技术	显著提升了密文信息熵, 保障了智慧城市建设的
	政务大数据平台无法完全解决数据内部泄露、恶意攻击 ^[89]	数据脱敏机制	既保证了数据的可用性, 又完成了敏感内容与数据业务的解耦, 实现政务内部数据的安全保障
安全通信	政府难以从海量网络信息中及时发现隐藏热点舆情 ^[90]	大数据智能脱敏	实现潜在热点事件的提前预警, 提高政府决策的前瞻性
	政务信息资源共享可能遭遇网络攻击、政务数据云平台存在安全隐患 ^[91]	数据脱敏、数据加密、数据溯源	提高政务信息的安全性, 保障各政府部门政府数据资源互联互通, 实现政务数据安全地开放共享
	Hadoop 平台存储敏感数据可能存在泄露 ^[92]	Hadoop 上使用 ARIA 算法的 HDFS 数据加密	有效防止了存储在 Hadoop 上敏感数据被窃取或泄露
	电力监控系统面临病毒和敌对势力入侵 ^[93]	安全接入区域认证和访问控制	及时防范安全攻击, 进一步提高了电力监控系统的整体安全防护能力, 保障了重要电力数据的安全
访问控制	病人医保信息泄露、医疗数据传输泄露 ^[94]	风险访问控制	有效防止了医疗信息的泄露
	政务云数据存在安全隐患 ^[95]	数据库防火墙	实现数据库的访问行为控制、危险操作阻断, 为保障电子政务数据库中的隐私数据不被泄露提供了安全驱动力
安全管理	物联网领域存在用户安全意识不足、设备缺乏可靠的安全协议等问题 ^[96]	区块链技术	增强了系统的健壮性, 提高了物联网数据的安全性
	云计算容易受到外界恶意攻击 ^[97]	自适应同态区块链技术	提高了云环境的安全性, 但存在加密耗时较长的问题

过数字摘要、数字签名与认证、生物识别等技术,对数据信息源的正确性和信息的完整性进行检验与认证^[99]。防火墙技术则是通过数据过滤技术、代理技术、数据状态检测技术等,完成对系统访问日志与流量的检测分析,甄别非法访问用户,阻止外部网络病毒对内部网络的攻击,进而保障数据访问控制^[100]。张连营等^[95]针对政务云数据存在安全隐患问题,提出了电子商务云数据安全技术方案,其中包含基于主动防御机制的数据库防火墙技术,实现数据库的访问行为控制和危险操作阻断,为保障电子政务数据库中的隐私数据不被泄露提供了安全驱动力。

在数据安全方面,区块链结合密码学技术,采用链式结构存储数据的分布式数据管理技术逐渐得到了研究者的关注。Tchagna Kouanou等^[96]针对物联网领域存在用户安全意识不足、设备缺乏可靠的安全协议等问题,以智能家居为案例,利用区块链技术实现数据的安全保护,大大增强了系统的健壮性,提高了物联网数据的安全性。Devmane等^[97]针对云计算易受到外界恶意攻击的问题,提出利用区块链技术维护云环境下的数据安全,提高了云环境的安全性,但存在加密耗时较长的问题。

政务数据安全技术贯穿数据管理的每个环节,是政务信息安全的屏障。针对政务数据高安全性和高可靠性的要求,在数据管理和应用过程中需要时刻注重数据安全的内容,使用先进的安全保证技术保护公民信息和政务安全,从而提高政务服务平台整体的安全性和可靠性,充分发挥政务数据综合价值,改善民生服务。

4 研究挑战与展望

随着大数据和人工智能社会的到来,互联网+政务大数据管理研究取得了显著进展,但限于技术发展和行政管理体系现状,仍存在部分问

题:(1)数据融合与标准化问题。需要进一步研究跨部门、跨系统的数据融合技术,实现政务数据的统一标准化和整合,便于后续分析和应用。(2)数据安全性与隐私保护问题。需要更广泛地研究先进的加密、脱敏和匿名化技术,确保政务大数据的安全和隐私,防止数据泄露和滥用。(3)数据治理与质量评估问题,需要研究数据治理方法和评估指标,以保证政务大数据的质量、完整性和一致性,为后续数据分析和决策提供有力支持。

此外,随着信息技术的迅速发展,一些新技术和理念不断被引入政务数据管理中。(1)数据管理模式向区块链数据管理模式转变。区块链技术的出现为数据管理提供了一种新的方式,其在数据安全、数据隐私和身份认证等方面都得到了广泛的应用。(2)数据模型与业务模型向解耦转变。随着政务大数据建设的快速发展,越来越多的信息资源被整合利用,相关应用项目的系统结构也愈加庞大和复杂,业务功能和数据也越来越丰富,这就需要将数据模型和业务模型解耦,将整个项目分解到不同的单元,由不同的业务部门分别完成相应单元的任务,最终达到项目的整体目标。(3)分散管理向一体化统一管理转变。传统的数据资产管理建设一般由多个分散的管理活动和解决方案组成,易造成数据资产管理各个环节之间(包括开发与管理、管理与运营)的脱节,使数据从生产端到消费端的开发效率降低。如何增强多方角色协同,使数据资产管理成为一个有机整体,是未来数据资产管理的重要方向。

参考文献

- [1] 孙轩,王焕骁. 政务大数据安全防护能力建设: 基于技术和管理视角的探讨 [J]. 计算机科学, 2022, 49(4): 67-73.
Sun X, Wang HX. Capability building for

- government big data safety protection: discussions from technological and management perspectives [J]. *Computer Science*, 2022, 49(4): 67-73.
- [2] 姚水琼, 齐胤植. 美国数字政府建设的实践研究与经验借鉴 [J]. *治理研究*, 2019, 35(6): 60-65.
Yao SQ, Qi YZ. The practical research and experience of digital government construction in the United States [J]. *Governance Research*, 2019, 35(6): 60-65.
- [3] 张晓, 鲍静. 数字政府即平台: 英国政府数字化转型战略研究及其启示 [J]. *中国行政管理*, 2018, (3): 27-32.
Zhang X, Bao J. Government as a platform: the inspiration and reference of UK government digital transformation strategy [J]. *Chinese Public Administration*, 2018, (3): 27-32.
- [4] 沈霄, 王国华. 基于整体性政府视角的新加坡“智慧国”建设研究 [J]. *情报杂志*, 2018, 37(11): 69-75.
Shen X, Wang GH. “Smart nation” in Singapore: perspective of the whole of government [J]. *Journal of Intelligence*, 2018, 37(11): 69-75.
- [5] 戴长征, 鲍静. 数字政府治理——基于社会形态演变进程的考察 [J]. *中国行政管理*, 2017, (9): 21-27.
Dai CZ, Bao J. Digital government—a perspective from social form evolvement [J]. *Chinese Public Administration*, 2017, (9): 21-27.
- [6] 黄璜. 中国“数字政府”的政策演变——兼论“数字政府”与“电子政务”的关系 [J]. *行政论坛*, 2020, 27(3): 47-55.
Huang H. Policy evolution of “digital government” in China: concurrently discussing the relationship between “digital government” and “e-government” [J]. *Administrative Forum*, 2020, 27(3): 47-55.
- [7] 冯登国. 国外电子政务发展现状 [J]. *信息安全与通信保密*, 2002, (5): 43-48.
Feng DG. Current situation of e-government development abroad [J]. *Information Security and Communication Privacy*, 2002, (5): 43-48.
- [8] 王益民. 全球电子政务发展现状、特点趋势及对中国的启示——《2016 年联合国电子政务调查报告》解读 [J]. *电子政务*, 2016, (9): 62-69.
Wang YM. The development status, characteristics and trends of global e-government and its enlightenment to China—interpretation of the 2016 United Nations e-government survey report [J]. *E-Government*, 2016, (9): 62-69.
- [9] 张建军. 立足中国国情 建立标准体系 解读《电子政务标准化指南——信息安全》 [J]. *信息网络安全*, 2003, (6): 18-20.
Zhang JJ. Establishing standard system based on China’s national conditions interpretation of e-government standardization guide—information security [J]. *Netinfo Security*, 2003, (6): 18-20.
- [10] 李广乾. 电子政务相关法律法规的评价分析 [J]. *信息化建设*, 2005, (4): 19-21.
Li GQ. Evaluation and analysis of laws and regulations to e-government [J]. *Information Construction*, 2005, (4): 19-21.
- [11] 江文路. 数字文明新时代背景下的数字政府建设: 规划图景、改革路径与实践经验 [J]. *武汉科技大学学报(社会科学版)*, 2023, 25(1): 49-56.
Jiang WL. Planning prospect, reforming path and practical experience of digital government construction in the era of big data [J]. *Journal of Wuhan University of Science & Technology (Social Science Edition)*, 2023, 25(1): 49-56.
- [12] 刘骁. 成都市政务大数据管理存在的问题及对策建议 [D]. 成都: 电子科技大学, 2018.
Liu X. Problems and countermeasures of big data management in Chengdu government affairs [D]. Chengdu: University of Electronic Science and Technology of China, 2018.
- [13] Debnath AK, Chin HC, Haque MM, et al. A methodological framework for benchmarking smart transport cities [J]. *Cities*, 2014, 37(2): 47-56.
- [14] Kogan N, Lee KJ. Exploratory research on the success factors and challenges of Smart City projects [J]. *Asia Pacific Journal of Information Systems*, 2014, 24(2): 141-189.
- [15] Batty M. Smart cities, big data [J]. *Environment and Planning B: Planning and Design*, 2012, 39(2): 191-193.
- [16] Walravens N. Mobile city applications for Brussels citizens: Smart City trends, challenges and a reality

- check [J]. *Telematics and Informatics*, 2015, 32(2): 282-299.
- [17] Allard S. DataONE: facilitating eScience through collaboration [J]. *Journal of eScience Librarianship*, 2012, 1(1): 4-17.
- [18] Mazumdar S, Seybold D, Kritikos K, et al. A survey on data storage and placement methodologies for Cloud-Big Data ecosystem [J]. *Journal of Big Data*, 2019, 6(1): 15.
- [19] Lourenco RP. An analysis of open government portals: a perspective of transparency for accountability [J]. *Government Information Quarterly*, 2015, 32(3): 323-332.
- [20] 朱海涛. 政务大数据开放及共享安全问题研究 [J]. *网络安全技术与应用*, 2022, (6): 56-59.
- Zhu HT. Research on government big data openness and sharing security [J]. *Network Security Technology & Application*, 2022, (6): 56-59.
- [21] 郭赞宇, 张旗. 电子政务数据脱敏关键技术应用研究 [J]. *电子产品可靠性与环境试验*, 2022, 40(3): 66-70.
- Guo ZY, Zhang Q. Research on application of key technology of e-government data desensitization [J]. *Electronic Product Reliability and Environmental Testing*, 2022, 40(3): 66-70.
- [22] 许盛伟, 牟健. 基于机器学习的政务大数据定级技术研究 [J]. *保密科学技术*, 2021, (5): 39-44.
- Xu SW, Mou J. Research on classification technology of government big data based on machine learning [J]. *Security Science and Technology*, 2021, (5): 39-44.
- [23] 徐海明, 邓海春. 政务大数据中心建设关键技术及应用研究 [J]. *智能建筑与智慧城市*, 2022, (7): 94-96.
- Xu HM, Deng HC. Research on key technology and application of government big data center construction [J]. *Smart Buildings and Smart Cities*, 2022, (7): 94-96.
- [24] 曹军威, 袁仲达, 明阳阳, 等. 能源互联网大数据分析技术综述 [J]. *南方电网技术*, 2015, 9(11): 1-12.
- Cao JW, Yuan ZD, Ming YY, et al. Survey of big data analysis technology for energy internet [J]. *Southern Power System Technology*, 2015, 9(11): 1-12.
- [25] Gacitua R, Astudillo H, Hitpass B, et al. Recent models for collaborative e-government processes: a survey [J]. *IEEE Access*, 2021, 9: 19602-19618.
- [26] Matus A, Guerra E, Fuertes W, et al. On the development of an electronic invoicing solution to integrate SMEs with a tax-collection e-government-platform [C] // *Proceedings of the 2017 Fourth International Conference on eDemocracy & eGovernment*, 2017: 94-101.
- [27] 王中慧. 大数据背景下 S 省电子政务促进政府治理能力提升对策研究 [D]. 济南: 山东财经大学, 2022.
- Wang ZH. Research on countermeasures of S provinces's e-government to promote the improvement of governance capability under the background of big data [D]. Jinan: Shandong University of Finance and Economics, 2022.
- [28] 庞孟春. 遂宁市电子政务信息安全风险防范问题研究 [D]. 成都: 电子科技大学, 2022.
- Pang MC. Research on e-government information security risk prevention in Suining [D]. Chengdu: University of Electronic Science and Technology of China, 2022.
- [29] 金江军. 智慧政府: 电子政务发展的新阶段 [J]. *信息化建设*, 2011, (11): 16-17.
- Jin JJ. Smart government: a new stage in e-government development [J]. *Information Construction*, 2011, (11): 16-17.
- [30] 于施洋, 王建冬, 童楠楠. 国内外政务大数据应用发展述评: 方向与问题 [J]. *电子政务*, 2016, (1): 2-10.
- Yu SY, Wang JD, Tong NN. A review of the development of big data application in government affairs at home and abroad: directions and problems [J]. *E-Government*, 2016(1): 2-10.
- [31] Epstein B. Two decades of e-government diffusion among local governments in the United States [J]. *Government Information Quarterly*, 2022, 39(2): 101665.
- [32] Thomas MA, Cipolla J, Lambert B, et al. Data management maturity assessment of public sector

- agencies [J]. *Government Information Quarterly*, 2019, 36(4): 101401.
- [33] Haini SI, Rahim NZA, Zainuddin NMM, et al. Adoption of open government data in local government context: conceptual model development [C] // *Proceedings of the 2019 5th International Conference on Computer and Technology Applications*, 2019: 193-198.
- [34] 段尧清, 姜慧, 汤弘昊. 政府开放数据全生命周期: 概念、模型与结构——系统论视角 [J]. *情报理论与实践*, 2019, 42(5): 35-40+50.
- Duan YQ, Jiang H, Tang HH. Government open data full lifecycle: concept, model and structure: from system theory perspective [J]. *Information Studies: Theory & Application*, 2019, 42(5): 35-40+50.
- [35] 樊建平, 孙婧, 朱利, 等. 推进智慧城市构建进程促进城市“智理”水平提升——国家重点研发计划“互联网+政务大数据透明管理与智能服务平台”项目取得重要进展 [J]. *科技成果管理与研究*, 2022, 17(10): 79-81.
- Fan JP, Sun J, Zhu L, et al. Advancing the construction process of smart cities and improving the level of urban intelligence—important progress has been made in the national key research and development program “internet+government big data transparent management and intelligent service platform” [J]. *Management and Research of Scientific and Technological Achievements*, 2022, 17(10): 79-81.
- [36] 孙婧, 樊建平, 徐亦飞, 等. “互联网+”政务大数据智能服务平台设计与应用 [J]. *集成技术*, 2023, 12(1): 4-16.
- Sun J, Fan JP, Xu YF, et al. Design and application of “Internet +” government big data intelligent service platform [J]. *Journal of Integration Technology*, 2023, 12(1): 4-16.
- [37] 韩寒. 大数据技术在政务服务中的应用研究 [J]. *信息与电脑(理论版)*, 2020, 32(18): 43-44.
- Han H. Research on the application of big data technology in government service [J]. *China Computer & Communication (Theory)*, 2020, 32(18): 43-44.
- [38] 王薇, 胡晓丹. “大数据”在 CPI 调查中的应用初探——人工采集网络购买商品价格的研究及建议 [J]. *中国统计*, 2015, (12): 47-48.
- Wang W, Hu XD. The application of “big data” in CPI survey—the research and suggestion of manual collection of online commodity prices [J]. *China Statistics*, 2015, (12): 47-48.
- [39] 武国雄. 道路纵横断面测量数据的半自动采集方法 [J]. *城市勘测*, 2009, (3): 79-82.
- Wu GX. Study on the method semi-automatic collecting measurement data of vertical and horizontal section of road [J]. *Urban Geotechnical Investigation & Surveying*, 2009, (3): 79-82.
- [40] 杨光. 美国政府大数据计划出炉 [J]. *计算机与网络*, 2014, 40(6): 7.
- Yang G. The US government’s big data plan is out [J]. *Computer & Network*, 2014, 40(6): 7.
- [41] 卫祥, 罗发政, 杨浩, 等. 基于 WebService 的企业采购业务架构的设计与实现 [J]. *自动化技术与应用*, 2020, 39(2): 33-40+44.
- Wei X, Luo FZ, Yang H, et al. Design and implementation of enterprise purchasing business architecture based on WebService [J]. *Techniques of Automation and Applications*, 2020, 39(2): 33-40+44.
- [42] 董学润. 大数据分析 & 处理综述 [J]. *中国新通信*, 2021, 23(1): 67-68.
- Dong XR. A survey of big data analytics and processing [J]. *China New Telecommunications*, 2021, 23(1): 67-68.
- [43] 李汉鹏, 王阳, 郑君旋, 等. 多源异构检测大数据采集与存储方法研究 [J]. *电子质量*, 2021, (9): 53-55.
- Li HP, Wang Y, Zheng JX, et al. Research on data acquisition and storage methods based on multi-source heterogeneous testing big data [J]. *Mass of Electrons*, 2021, (9): 53-55.
- [44] 王晴. 数据中心存储技术研究综述 [J]. *信息与电脑(理论版)*, 2019, (4): 190-191.
- Wang Q. Review of data center storage technology [J]. *China Computer & Communication (Theory)*, 2019, (4): 190-191.
- [45] 宋云奎, 吴文鹏, 赵磊, 等. 基于 Redis 的分布式

- 数据存储方法 [J]. 计算机产品与流通, 2020, (8): 106.
- Song YK, Wu WP, Zhao L, et al. Distributed data storage method based on Redis [J]. Computer Products and Circulation, 2020, (8): 106.
- [46] 宋锴业, 徐雅倩, 陈天祥. 政务数据资产化的创新发展、内在机制与路径优化——以政务数据资产管理的潍坊模式为例 [J]. 电子政务, 2022, (1): 14-26.
- Song KY, Xu YQ, Chen TX. The innovation development, internal mechanism and path optimization of government data capitalization—taking the Weifang model of government data asset management as an example [J]. E-Government, 2011, (1): 14-26.
- [47] 符宁. 科技政务大数据管理与挖掘平台设计 [J]. 软件导刊, 2021, 20(5): 86-91.
- Fu N. Design of science and technology government big data management and mining platform [J]. Software Guide, 2021, 20(5): 86-91.
- [48] 刘云汉, 郭继光. 基于 MongoDB 分片集群的海量数据存储设计 [J]. 电子技术与软件工程, 2022, (7): 254-258.
- Liu YH, Guo JG. Massive data storage design based on MongoDB shard cluster [J]. Electronic Technology & Software Engineering, 2022, (7): 254-258.
- [49] 张维冲, 王芳, 黄毅. 基于图数据库的贵州省大数据政策知识建模研究 [J]. 数字图书馆论坛, 2020, (4): 30-38.
- Zhang WC, Wang F, Huang Y. Knowledge modeling of big data policy in GuiZhou province based on graph database [J]. Digital Library Forum, 2020, (4): 30-38.
- [50] 秦晓东. 基于大数据的省市政务数据共享开放一体化平台设计 [J]. 软件, 2021, 42(5): 107-109.
- Qin XD. Design of open integrated platform for provincial and municipal government data sharing based on big data [J]. Software, 2021, 42(5): 107-109.
- [51] 刘应新, 张磊, 姚鑫, 等. 政务云平台建设探讨 [J]. 广播电视网络, 2021, 28(6): 108-110.
- Liu YX, Zhang L, Yao X, et al. Discussion on the construction of government cloud platform [J]. Radio & Television Network, 2021, 28(6): 108-110.
- [52] 张大璐, 葛奇, 冯一博, 等. DNA 数据存储的科研概况国际对比与分析 [J]. 中国生物工程杂志, 2022, 42(6): 116-129.
- Zhang DL, Ge Q, Feng YB, et al. Comparison and analysis on scientific research programs on DNA data storage [J]. China Biotechnology, 2022, 42(6): 116-129.
- [53] 万婵, 魏理豪, 杨秋勇, 等. 电网行业元数据集成数据存储策略研究 [J]. 微型电脑应用, 2021, 37(1): 26-28+32.
- Wan C, Wei LH, Yang QY, et al. Research on data storage strategy for metadata integration in power grid industry [J]. Microcomputer Applications, 2021, 37(1): 26-28+32.
- [54] 吴文莉, 刘国华. 关系数据模型中函数查询的结构特征 [J]. 智能计算机与应用, 2020, 10(1): 22-27.
- Wu WL, Liu GH. Structural features of functional queries in relational data model [J]. Intelligent Computer and Applications, 2020, 10(1): 22-27.
- [55] 吴明礼, 张宏安. 数据存储技术综述 [J]. 北方工业大学学报, 2015, 27(1): 30-35+55.
- Wu ML, Zhang HA. Review of massive data storage technology [J]. Journal of North China University of Technology, 2015, 27(1): 30-35+55.
- [56] 祝君, 刘柏众, 余晟隼, 等. 面向 OceanBase 的存储过程设计与实现 [J]. 华东师范大学学报(自然科学版), 2016, (5): 144-152.
- Zhu J, Liu BZ, Yu SJ, et al. Designs and implementations of stored procedure in OceanBase [J]. Journal of East China Normal University (Natural Science), 2016, (5): 144-152.
- [57] 黄振业. 针对 Paxos 算法的全局唯一自增 ID 的生成方法 [J]. 科技资讯, 2021, 19(9): 1-5+12.
- Huang ZY. Generation method of global unique auto increment ID generation for Paxos algorithm [J]. Science & Technology Information, 2021, 19(9): 1-5+12.
- [58] Györödi CA, Dumșe-Burescu DV, Zmaranda DR, et al. Performance analysis of NoSQL and relational databases with CouchDB and MySQL

- for application's data storage [J]. *Applied Sciences*, 2020, 10(23): 8524.
- [59] 李志新. 国内政府数据开放研究综述: 2013-2016 [J]. *情报杂志*, 2017, 36(7): 156-161+187.
Li ZX. Literature review of domestic open government data: 2013-2016 [J]. *Journal of Intelligence*, 2017, 36(7): 156-161+187.
- [60] 杜刚. 基于多系统数据级融合的煤矿监测监控逻辑报警分析 [J]. *山西煤炭*, 2017, 37(5): 13-21.
Du G. Logic level alarm analysis of monitoring and measurement system in coal mines based on multisystem data fusion [J]. *Shanxi Coal*, 2017, 37(5): 13-21.
- [61] Khaleghi B, Khamis A, Karray FO, et al. Multisensor data fusion: a review of the state-of-the-art [J]. *Information fusion*, 2013, 14(1): 28-44.
- [62] 沈王恒. 浦东新区政务数据融合服务平台的探索 [J]. *信息技术与标准化*, 2021, (6): 13-18.
Shen WH. Exploration for Pudong New District government data integration service platform [J]. *Information Technology & Standardization*, 2021, (6): 13-18.
- [63] 王红霞, 王波, 董旭柱, 等. 面向多源电力感知终端的异构多参量特征级融合: 融合模式、融合框架与场景验证 [J]. *电工技术学报*, 2021, 36(7): 1314-1323.
Wang HX, Wang B, Dong XZ, et al. Heterogeneous multi-parameter feature-level fusion for multi-source power sensing terminals: fusion mode, fusion framework and application scenarios [J]. *Transactions of China Electrotechnical Society*, 2021, 36(7): 1314-1323.
- [64] 徐济宣, 马辉, 冯小凯. 轴承故障的多源异构数据特征级融合诊断方法 [J]. *机械设计与制造*, 2021, (9): 150-154+159.
Xu JX, Ma H, Feng XK. Bearing fault diagnosis method based on feature level fusion of multi-source heterogeneous spatial data [J]. *Machinery Design & Manufacture*, 2021, (9): 150-154+159.
- [65] 王长海, 陈倩, 唐欣, 等. 基于点云与图像决策级融合的道路行人检测 [J]. *广西大学学报(自然科学版)*, 2021, 46(6): 1592-1601.
Wang CH, Chen Q, Tang X, et al. Road pedestrian detection based on point cloud and image decision-level fusion [J]. *Journal of Guangxi University (Natural Science Edition)*, 2021, 46(6): 1592-1601.
- [66] 杨晓冬, 梁斯东, 张宝玉. 天网情报决策级融合分析技术体系研究 [J]. *电子技术与软件工程*, 2022, (1): 142-154.
Yang XD, Liang SD, Zhang BY. Research on the technical system of Skynet Intelligence decision-level fusion analysis [J]. *Electronic Technology & Software Engineering*, 2022, (1): 142-154.
- [67] Chen SD. Research on big data computing model based on spark and big data application [C] // *Proceedings of the 2021 Guangdong-Hong Kong-Macao Greater Bay Area Artificial Intelligence and Big Data Forum*, 2021: 012017.
- [68] 于海东. 大数据计算技术在电子政务领域的应用 [J]. *电子技术与软件工程*, 2019, (24): 127-128.
Yu HD. The application of big data computing technology in the field of e-government [J]. *Electronic Technology & Software Engineering*, 2019, (24): 127-128.
- [69] 任智. 面向政务多源大数据的分布式治理组件的设计与实现 [D]. 南京: 东南大学, 2020.
Ren Z. Design and implementation of distributed governance component oriented to government multi-source and big data [D]. Nanjing: Southeast University, 2020.
- [70] Zhou X, He J, Huang GY, et al. SVD-based incremental approaches for recommender systems [J]. *Journal of Computer and System Sciences*, 2015, 81(4): 717-733.
- [71] Wang XK, Wang W, Yang LT, et al. A distributed HOSVD method with its incremental computation for big data in cyber-physical-social systems [J]. *IEEE Transactions on Computational Social Systems*, 2018, 5(2): 481-492.
- [72] Parto M, Saldana C, Kurfess T. Real-time outlier detection and Bayesian classification using incremental computations for efficient and scalable stream analytics for IoT for manufacturing [J]. *Procedia Manufacturing*, 2020, 48: 968-979.
- [73] Ibrahim S, Phan TD, Carpen-Amarie A, et al. Governing energy consumption in Hadoop through

- CPU frequency scaling: an analysis [J]. *Future Generation Computer Systems*, 2016, 54: 219-232.
- [74] Nayak P, Bhat T. Low power architecture strategies for green computing [J]. *ECS Transactions*, 2022, 107(1): 16679.
- [75] 夏靖波, 韦泽鲲, 付凯, 等. 云计算中 Hadoop 技术研究与综述 [J]. *计算机科学*, 2016, 43(11): 6-11+48.
Xia JB, Wei ZK, Fu K, et al. Review of research and application on Hadoop in cloud computing [J]. *Computer Science*, 2016, 43(11): 6-11+48.
- [76] 吴信东, 嵇圣础. MapReduce 与 Spark 用于大数据分析之比较 [J]. *软件学报*, 2018, 29(6): 1770-1791.
Wu XD, Ji SW. Comparative study on MapReduce and Spark for big data analytics [J]. *Journal of Software*, 2018, 29(6): 1770-1791.
- [77] Gong SL, Qin X. Research on real-time learning prediction method based on spark [C] // *Proceedings of the 2019 IEEE 10th International Conference on Software Engineering and Service Science*, 2019: 354-357.
- [78] Li JY, Zhang CB, Zhao Y, et al. Federated learning-based short-term building energy consumption prediction method for solving the data silos problem [J]. *Building Simulation*, 2022, 15(6): 1145-1159.
- [79] Chen JC, Xue YZ. Bootstrapping a blockchain based ecosystem for big data exchange [C] // *Proceedings of the 2017 IEEE International Congress on Big Data*, 2017: 460-463.
- [80] 褚春燕. 区块链技术在政府数据共享中的应用研究 [J]. *电子技术与软件工程*, 2022, (11): 250-253.
Chu CY. Research on the application of blockchain technology in government data sharing [J]. *Electronic Technology & Software Engineering*, 2022, (11): 250-253.
- [81] Theodouli A, Arakliotis S, Moschou K, et al. On the design of a blockchain-based system to facilitate healthcare data sharing [C] // *Proceedings of the 2018 17th IEEE International Conference on Trust, Security and Privacy in Computing and Communications/12th IEEE International Conference on Big Data Science and Engineering*, 2018: 1374-1379.
- [82] Sunindyo WD, Amrita IDPDK. The development of data publishing tool for Indonesian open government data [C] // *Proceedings of the 7th International Conference on Electrical Engineering and Informatics*, 2019: 30-34.
- [83] 陈绍林. 推动大数据在数字政府建设中的应用——以数字福建建设为例 [J]. *海峡科学*, 2021, (10): 77-81.
Chen SL. Promoting the application of big data in digital government construction—taking digital Fujian construction as an example [J]. *Staits Science*, 2021, (10): 77-81.
- [84] 张增华. “津云”大数据政务公开模式研究 [J]. *数字技术与应用*, 2021, 39(9): 215-218.
Zhang ZH. Research on “Jinyun” big data government affairs open mode [J]. *Digital Technology and Application*, 2021, 39(9): 215-218.
- [85] 徐雅杰. 大数据时代下浙江省政府政务公开信息的可视化研究 [J]. *创新创业理论研究与实践*, 2020, 3(5): 10-11.
Xu YJ. Visualization research on public information of government affairs in Zhejiang province in the era of big data [J]. *The Theory and Practice of Innovation and Entrepreneurship*, 2020, 3(5): 10-11.
- [86] Moyano JP, Schmedders K. What managers need to know about data exchanges: the era of big-data silos is fading. Shared data is the future [J]. *MIT Sloan Management Review*, 2020, 61(4): 39-44.
- [87] Zhang C, Xie Y, Bai H, et al. A survey on federated learning [J]. *Knowledge-Based Systems*, 2021, 216: 106775.
- [88] 韩义森. 政务大数据在智慧城市建设中的应用探索——以“云上扬州”建设为例 [J]. *新型工业化*, 2021, 11(10): 38-44.
Han YS. Exploration on the application of government big data in smart city construction—taking the construction of “Yangzhou on Cloud” as an example [J]. *The Journal of New Industrialization*, 2021, 11(10): 38-44.
- [89] 郭赞宇, 张旗. 电子政务数据脱敏关键技术应用研究 [J]. *电子产品可靠性与环境试验*, 2022,

- 40(3): 66-70.
- Guo ZY, Zhang Q. Research on application of key technology of data desensitization in e-government [J]. *Electronic Product Reliability and Environmental Testing*, 2022, 40(3): 66-70.
- [90] 陈平刚, 蔡利华, 王玲丽. 政务大数据支撑的政府舆情预警研究 [J]. *现代商贸工业*, 2019, 40(34): 141-142.
- Chen PG, Cai LH, Wang LL. Research on government public opinion early warning supported by government affairs big data [J]. *Modern Business Trade Industry*, 2019, 40(34): 141-142.
- [91] 王庆德, 韩晓露, 张贺. 基于云的电子政务信息资源共享与安全保障体系及应用前景研究 [C] // 2022 年西湖论剑·网络安全大会——数字城市安全治理论坛论文集, 2022: 9-13.
- Wang QD, Han XL, Zhang H. Research on e-government information resource sharing and security assurance system based on cloud and its application prospect [C] // *Proceedings of the 2022 West Lake Debate and Cyber Security Conference—Digital City Security Governance Forum*, 2022: 9-13.
- [92] Song Y, Shin YS, Jang MY. Design and implementation of HDFS data encryption scheme using ARIA algorithms on Hadoop [C] // *Proceedings of the 2017 IEEE International Conference on Big Data and Smart Computing*, 2017: 87-90.
- [93] 张富川, 吴金宇, 陈刚, 等. 电力监控系统安全接入区接入认证与访问控制机制分析 [J]. *电子技术与软件工程*, 2022, (7): 9-12.
- Zhang FC, Wu JY, Chen G, et al. Analysis of access authentication and access control mechanism in security access area of power monitoring system [J]. *Electronic Technology & Software Engineering*, 2022, (7): 9-12.
- [94] 刘宇航. 基于风险访问控制的医院信息安全研究 [J]. *中国新通信*, 2021, 23(15): 119-120.
- Liu YH. Research on hospital information security based on risk access control [J]. *China New Telecommunications*, 2021, 23(15): 119-120.
- [95] 张连营, 周爻, 陈磊. 电子政务云数据安全技术方案设计 [J]. *电信科学*, 2020, 36(Z1): 81-88.
- Zhang LY, Zhou Y, Chen L. Design of cloud data security technology for e-government [J]. *Telecommunication Science*, 2020, 36(Z1): 81-88.
- [96] Tchagna Kouanou A, Tchito Tchapa C, Sone Ekonde M, et al. Securing data in an Internet of things network using blockchain technology: smart home case [J]. *SN Computer Science*, 2022, 3(2): 167.
- [97] Devmane V, Lande BK, Joglekar J, et al. Preserving data security in cloud environment using an adaptive homomorphic blockchain technique [J]. *Arabian Journal for Science and Engineering*, 2022, 47: 10381-10394.
- [98] 陈天莹, 陈剑锋. 大数据环境下的智能数据脱敏系统 [J]. *通信技术*, 2016, 49(7): 915-922.
- Chen TY, Chen JF. Intelligent data masking system for big data productive environment [J]. *Communications Technology*, 2016, 49(7): 915-922.
- [99] 罗军舟, 杨明, 凌振, 等. 网络空间安全体系与关键技术 [J]. *中国科学: 信息科学*, 2016, 46(8): 939-968.
- Luo JZ, Yang M, Ling Z, et al. Architecture and key technologies of cyberspace security [J]. *SCIENTIA SINICA Informationis*, 2016, 46(8): 939-968.
- [100] 马利, 梁红杰. 计算机网络安全中的防火墙技术应用研究 [J]. *电脑知识与技术*, 2014, 10(16): 3743-3745.
- Ma L, Liang HJ. Application of firewall technology in computer network security [J]. *Computer Knowledge and Technology*, 2014, 10(16): 3743-3745.