

政府大数据治理：政府管理的新形态

杨孟辉^{1,2}, 杜小勇^{2,3}

1. 中国人民大学信息资源管理学院, 北京 100872; 2. 数据工程与知识工程教育部重点实验, 北京 100872;
3. 中国人民大学信息学院, 北京 100872

摘要

政府大数据治理成为政府关心的一个重要议题。采用文献调研等方法, 对当前主要的国家政府大数据实践进行了分析, 认为政府大数据应用仍处于早期阶段。对政府大数据治理的概念进行了溯源, 认为政府大数据治理一方面继承了公共和行政管理领域的治理理论和行政管理手段, 另一方面也借鉴了公司治理中IT治理、数据治理和大数据治理的框架和方法。对政府大数据治理的概念、特征进行了阐释, 提出了政府大数据治理的新挑战、新目标和新机制, 最后指出大数据治理是政府管理的新形态。

关键词

政府治理; 数据治理; 大数据治理; 政府大数据治理

中图分类号: C931.2

文献标识码: A

doi: 10.11959/j.issn.2096-0271.2020010

Big data governance in governments: a new form of the government administration

YANG Menghui^{1,2}, DU Xiaoyong^{2,3}

1. School of Information Resource Management, Renmin University of China, Beijing 100872, China
2. Key Laboratory of Data Engineering and Knowledge Engineering, Renmin University of China, Beijing 100872, China
3. School of Information, Renmin University of China, Beijing 100872, China

Abstract

Exploring big data governance in governments becomes an important issue for government sectors. By using literature survey, this study made a comprehensive analysis of big data practice in governments. Results shown that the big data application projects in governments were still at the early stage. The concept of big data governance in governments was traced. The concept of big data governance in governments inherited the governance theories and administrative management means in the public management and the administrative management areas, and also borrowed the frameworks and approaches, from IT governance, data governance, and big data governance in corporate governance. The concept and characteristics of big data governance in governments were explained. New challenges, new goals and new mechanisms of the big data governance in governments were proposed. It was pointed out that big data governance in governments is a new form of government management.

Key words

government governance, data governance, big data governance, big data governance in governments

1 引言

大数据是从各种渠道收集的大量数字数据的总称,由于它数量巨大、未经处理、非结构化,无法通过目前主流的软件工具在合理时间进行采集、存储和处理。据IBM统计,全球近90%的数据是在过去的2年中生成的,每天增加2.5 EB(1 EB=1 024 PB)数据,其中,约90%的数据是非结构化的。来自网络和云平台的大数据为价值创造和商业智能提供了新的机会,可以为任何组织提供决策支持。尽管如此,大数据也意味着新的挑战,涉及复杂性、安全性和隐私风险以及对新技术和技能的要求。虽然商业企业在大数据应用和技术开发领域处于领先地位,但是政府部门也已开始利用来自网络、生物和传感器、视频等多种渠道的快速增长的动态数据,以获得洞察力和进行及时决策。面对大数据带来的机遇和挑战,各国政府都想在这个领域占得先机,虽然实施方案有所差异,但在决策理念上却几乎一致,即在宏观层面部署符合本国国情的大数据战略,在中观层面设计大数据治理的体制机制,在微观层面考虑大数据管理的应对策略、程序和行动。

2009年,美国政府上线了data.gov,这是全球第一个开放政府数据项目,它是一个数据仓库,目前它拥有近500万个数据集,涵盖交通、经济、医疗、教育和公共服务以及多种应用领域的的数据。2012年,美国政府宣布了“大数据研究与开发计划(big data research and development initiative)”,投资2亿美元,主要目标是发展最先进的核心大数据技术,加快科学和工程学的发现速度,加强国家安全,转变教学方式,扩大开发和使用大数据技术^[1]。2012年,美国国家科学基金会

(National Science Foundation, NSF)和美国国立卫生研究院(National Institutes of Health, NIH)联合启动了“发展大数据科学与工程核心技术(core techniques and technologies for advancing big data science & engineering)”项目,旨在发展核心科学技术手段,以提取、管理、分析、可视化来自大量、多样、分布式、异构数据集集中的有用信息^[2]。2014年,NIH已在亚马逊云服务(AWS)平台上累积了数百TB的人类基因数据,研究人员无需开发自身的超级计算功能即可访问和分析其中的大量基因数据^[3]。2010年,欧盟启动“欧洲数字议程(digital agenda for Europe)”战略,以解决如何通过快速和超快速的互操作互联网应用程序在单一数字市场为欧盟公民带来可持续的经济和社会利益的问题。2012年,欧盟委员会在“欧洲数字议程以及2012挑战(digital agenda for Europe and challenges for 2012)”计划中详细介绍了大数据战略,为确保欧盟成员国之间的网络安全和数据互联安全,该战略强调了欧盟成员国之间公共数据的安全性以及包括政府等在内的公共机构管理下的大数据的价值潜力,并明确该战略是欧盟数字化议程的一部分^[4]。欧盟20个成员国与IBM合作发起了一项名为穹顶(DOME)的大数据项目,该项目由荷兰射电天文学研究所(ASTRON)和IBM合作提供Exa超级计算机技术,旨在解决有关可观测宇宙的一系列科学问题,每天能够处理从平方公里射电阵(square kilometre array, SKA)射电望远镜产生的超过1 EB的数据^[5]。英国是欧洲较早实施大数据计划的国家之一,于2004年成立了英国地平线扫描中心(Horizon Scanning Center, HSC),以提高政府应对跨部门和多学科挑战的能力^[6]。英国政府于2009年创建了开放政府数据网站data.gov.uk,为公共部门的数据利用和数据共享提供一站式数据服

务。韩国于2011年发布的“大数据计划(big data initiative)”旨在通过大数据促进知识融合和行政大数据分析^[7]。2005年到2011年,日本教育、文化、体育、科学和技术部(MEXT)与大学以及其他研究机构合作,启动了“信息爆炸时代网络基础设施(cyber infrastructure for the information-explosion era, Info-plosion)”项目^[8];MEXT与日本国家科学基金会合作,加强研究并利用大数据技术预防、缓解和管理自然灾害;日本内政和通信部下属的信息通信委员会和信息通信技术(information communication and technology, ICT)战略委员会将“大数据应用”作为日本2020年的关键任务。

在世界各国纷纷借助大数据技术发展提升国家治理能力和战略能力的同时,我国对大数据的认识也不断加强。2006年我国发布了《国家电子政务总体框架》,明确要建设国家电子政务网络、政府信息资源目录体系与交换体系(government information resource catalog system and interchange system, GIRCSIS)和信息安全基础设施。2009年至今,已经建起全国性的政府信息资源目录体系与交换体系基础设施。2015年中华人民共和国国务院通过了《促进大数据发展行动纲要》,指出大数据在日常生活中发挥着越来越重要的作用。在中国共产党第十八届中央委员会第五次全体会议上,与大数据相关的概念被反复提及,“实施国家大数据战略,推进数据资源开放共享”的建议被正式提出。李国杰^[9]认为大数据可能是中国信息产业从跟踪走向引领的突破口。长期以来,我国政府数据仅被视为管理类项目的下属内容,在数据管理方面暂未建立政府机构体系下符合关联部门合作实情的治理体系。

不同国家发展大数据战略和实施大数据项目选择的领域、侧重的问题不同,为了

便于理解和比较,笔者将这些政府大数据项目按照其数据特征和应用领域分为4个象限,如图1所示。其中,横坐标表示应用领域,左边表示企业应用,右边表示政府应用;纵坐标表示数据特征,上面表示大数据,下面表示政府数据。

从图1可以发现3点。第一,正在实施和已经运行的大数据项目大多采取共享存储的方式与政府或者企业共享存储在数据库中的结构化数据,多数大数据项目没有使用实时、动态以及非结构化或半结构化数据,例如美国的“开放政府数据”项目data.gov,英国的“开放政府数据”项目data.gov.uk,我国的GIRCSIS项目。第二,政府部门启动的大数据计划均利用了大型和复杂的政府数据集,政府希望大数据能够增强政府为社会服务的能力,应对经济、卫生健康、自然灾害、恐怖主义等重大问题的挑战。例如,欧盟和IBM合作的穹顶项目,美国疾病控制与预防中心的用于细菌和疾病暴发的联网基因组学(networked phylogenomics for bacteria and outbreak ID, NPBOID)项目、美国国家科学基金会和美国国立卫生研究院的发展大数据科学与工程核心技术项目、美国亚马逊云服务中人类基因数据(Genmone on AWS)项目、英国地平线扫描中心项目、韩国反腐败和民权委员会的投诉信息分析中心(complaints information analysis center, CIAC)项目、韩国的国家DNA管理(national DNA management, NDM)项目、韩国的预防口蹄疫综合征(preventing foot and mouth disease syndrome, PFMDs)项目、日本智能交通系统(intelligent traffic system, ITS)项目以及“信息爆炸时代网络基础设施”项目,这些大数据计划大部分才刚刚开始,这也意味着政府大数据计划仍处于早期阶段^[10]。第三,大多数大数据应用于政府改善公

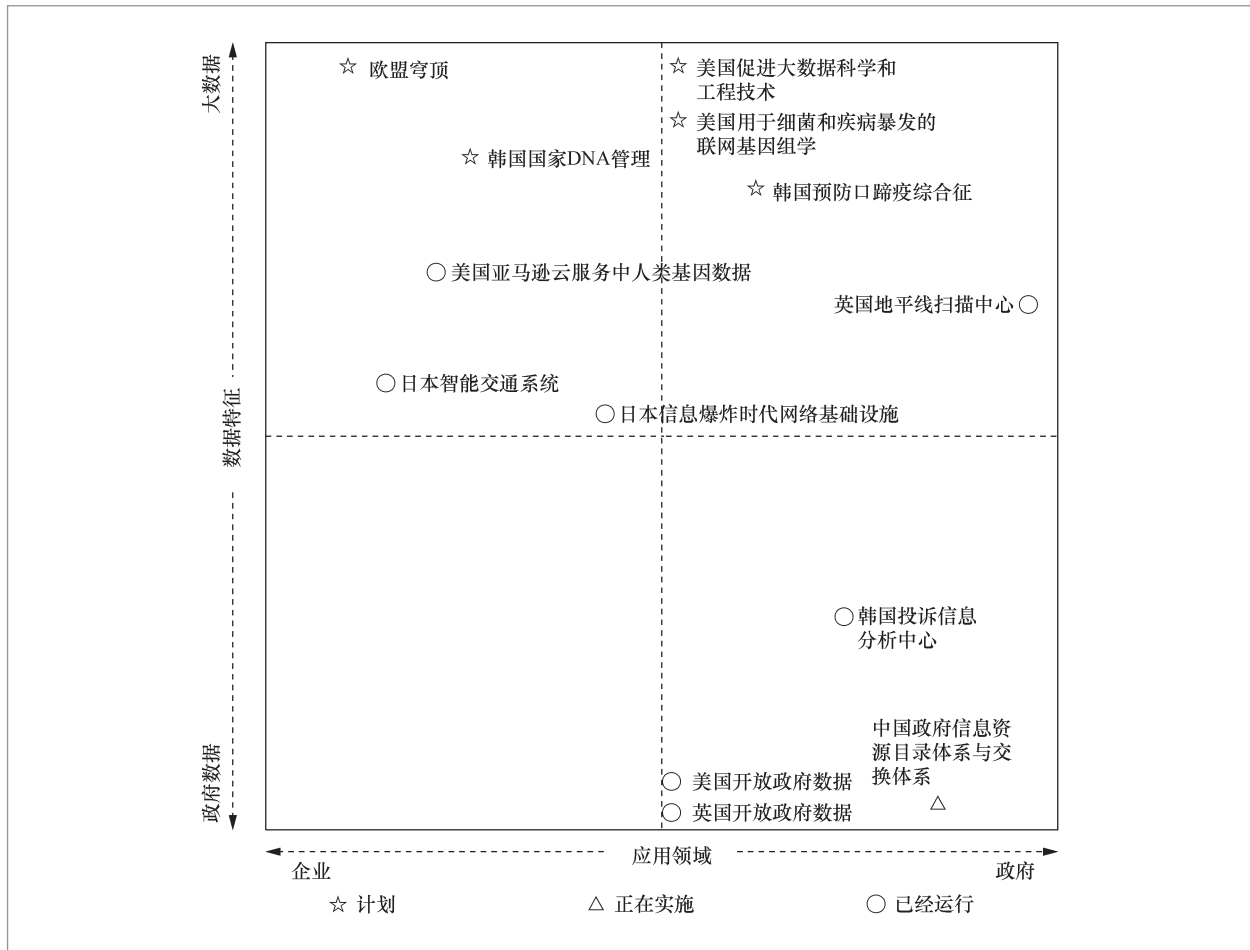


图1 政府数据以及政府大数据计划与实践

民参与公共事务、政府与商业企业合作领域，而政府内部的应用较少。

2 政府大数据治理概念溯源

“治理危机 (crisis of governance)”首次出现在1989年世界银行关于撒哈拉以南非洲的发展报告中，该报告将该地区的危机描述为“治理危机”^[1]。此后“治理”的概念被广泛用于各种体制问题的研究中，它首先被使用在政治学领域中，然后扩展到社会、经济领域。因此，“治理”是一

个概念性术语，具有极为广泛的含义。在政治学领域，“治理”意味着“政府治理”；在商业管理领域，“治理”指“公司治理”。这2个领域的研究对政府大数据治理的概念的形成起着主要的作用。

公司治理 (corporation governance) 被定义为“影响公司领导、管理或控制方式的一系列过程、惯例、政策、法律和制度”。公司治理下的实践领域之一是信息技术 (information technology) 治理，即IT治理。IT治理被定义为“组织领导、组织结构与组织过程，确保组织的IT支撑，并维护组织战略和组织目标”。国际

标准化组织(International Organization for Standardization, IOS)等行业组织已经建立了详细的IT治理标准和流程,许多公司和政府组织已将其付诸实践。这些IT治理实践大多数由以企业资源计划(enterprise resource planning, ERP)或客户关系管理(customer relationship management, CRM)等软件应用程序为中心的常见IT组织结构和使用信息系统结构的模式演变而成。从大型机到客户机、服务器体系结构,再到Web应用程序,IT部门一直专注于应用程序和应用程序的用户交互,而很少关注驱动应用程序的数据。通常,管理数据是根据数据如何服务于特定应用程序进行的,而不是根据它们如何服务于整个组织的战略利益进行的。因此,IT治理实践主要围绕组织的应用程序而不是其数据进行设计。大多数组织中均有明确定义的所有者、流程和策略对企业软件应用程序进行管理。但是,许多组织缺乏针对其最有价值的核心数据资产的等效治理结构。为了增加数据作为商业资产的价值,公司和政府组织需要建立用于使用、开发和管理数据的标准、策略和流程,以创建正确的组织结构和开发相应的技术基础设施。数据治理(data governance)作为一种根植于公司治理和IT治理的实践,被定义为“管理和确保企业数据的可用性、可访问性、质量、一致性、可审计性和安全性所需的流程、策略、标准、组织和技术”^[12]。只要有需要,数据治理就可以向合适的人提供合适的数据集,以帮助决策者做出正确的决策。国际标准化组织已经建立了数据治理标准和流程,数据治理研究所(Data Governance Institute)制定了数据治理框架,IBM提出了数据管理统一流程。由于数据的大量增长,大数据管理已成为IT和业务领域的挑战。一方面,大数据通常是结构化、非结构化和半结构化

的数据,而现有的数据治理不适用于半结构化和非结构化数据。另一方面,大多数大数据技术不提供数据治理功能。由于缺少对大数据治理的设计和考虑,因此从现有数据集向大数据过渡时,不仅数据完整性和数据质量受到威胁,现有的数据治理也面临着挑战。Al-badi等人^[13]提出了大数据治理框架,并对该大数据框架与ISO提出的数据治理框架进行了比较,比较结果表明,其提出的大数据治理框架满足ISO 8000标准的87%。

政府治理(government governance)突出了政府的主导性作用,它沿袭了新公共管理(new public management, NPM)的概念,政府治理由政府掌舵,政府与非政府组织的关系是委托与代理的关系,非政府组织发挥工具性作用^[14]。网络化治理(network governance)、整体性治理(holistic governance)和数字治理(digital governance)是目前主要的3种政府治理模式。网络化治理主要是指政府的工作不太依赖传统意义上的政府雇员,而是更多地依赖网络。整体性治理是指在组织和部门明确分工的前提下,将整合与协调结合,既满足了组织强制力下的团结,同时又相互协调各自的目标,解决组织间目标的冲突,从而形成一种持久的有机团结。数字治理也被称为数字时代治理(digital-era governance),是指在政府与自身或社会的互动中运用数字技术,简化行政程序和公共事务处理程序,提高民主化程度的治理模式^[15]。电子政务(E-government)是一般意义的数字治理。电子政务作为IT治理概念在政府部门的具体应用,是指政府机构采用信息技术转变政府与民众、企业以及其他政府组织之间的关系,由此带来的好处包括减少腐败、增加政府透明度、提高政府服务的便利性、降低政府运行的成本。我国的电

子政务治理模式继承了IT治理模式的思路,但是在治理结构的安排方面却依然保守化,未能与现有政府管理模式有效匹配。尽管信息化领导小组、部级联席会议制等发挥了一定作用,但是由于缺乏长期性的合理安排,存在多种形式的利益冲突,工作实施过程中仍然存在困难^[16]。

政府拥有民众、企业以及政府各级机构的大量数据,并利用这些数据进行决策和服务社会。民众通常认为政府应该在使用个人数据时保护个人隐私、保证数据质量,但是实际情况是由数据使用不当或者数据错误导致的公共数据管理问题时有发生。这些问题通常不是源于现有的政府业务规则或者技术本身,而是由缺乏数据治理机制和能力导致的^[17]。政府数据治理(government data governance)作为政府治理的重要组成部分,需要充分结合数据管理的法律、组织、技术、标准等多种因素,确保政府数据的使用合法合规,并保障数据的安全,以维持政府数据资产的有效价值。政府数据治理着眼于政府部门之间的内部数据流动关系,强调数据处理过程中的强关联性,主张数据治理从分散走向集中、从部分走向整体、从碎片走向整合,要求系统化配置数据资源,并建立一体化的治理内容框架^[18]。

据统计,有31%的政府数据是非结构化的数据,这些数据具有大数据特征。政府在管理和利用大数据方面的挑战除了数据量大、数据种类繁多、数据变化快之外,还有数据孤岛、数据安全性和数据合法性。每个政府部门通常都有自己的机密信息和公共信息仓库,通常不愿把他们认为专有的数据共享给他人。政府的每个系统都将其数据与其他系统隔离,使得政府机构和部门之间共享数据变得更加复杂。数据安全是政府大数据的主要属性,因此收集、存储和使用数据需要特别注意安全

问题。但是,当今大多数大数据技术(包括Casandra和Hadoop)缺乏足够的安全工具,这使数据安全成为政府的另一项挑战。同时,政府还必须在使用数据时满足相关的合法性和合规性要求。发挥政府大数据的潜力需要新的大数据治理框架。政府大数据治理(government big data governance)是政府行政管理及政府治理的重要组成部分,原因如下:其一,政府通过引入大数据及相关技术改变其治理方式;其二,政府作为信息化建设的主体,关注并协调其管辖范围内与大数据建设相关的主体,并基于数据治理的基本特征优化现有信息治理条件。就治理对象而言,既包括政府数据及技术、基础设施及应用,也包括资源价值积累过程中涉及的权责利关系;就治理手段而言,既包括元数据等传统数据管理方法,也包括投资收益管理等行政管理手段。若要提升政府大数据治理水平,需同时考虑治理对象和实际治理的动态变化与过程^[19]。

综上所述,IT治理、数据治理和大数据治理均是公司治理的一部分。数据治理是IT治理的一部分,大数据治理是数据治理的一部分。电子政务、政府数据治理和政府大数据治理一方面继承了公共和行政管理领域的治理理论和行政管理手段,是数字治理的一部分,另一方面也借鉴了公司治理中IT治理、数据治理和大数据治理的框架、方法和工具,相关治理概念之间的关系如图2所示。政府数据治理是电子政务的一部分,政府大数据治理是政府数据治理的一部分。

3 政府大数据治理的涵义和特征

政府和商业企业性质的主要差异见表1。企业的主要目标是通过提供商品和服

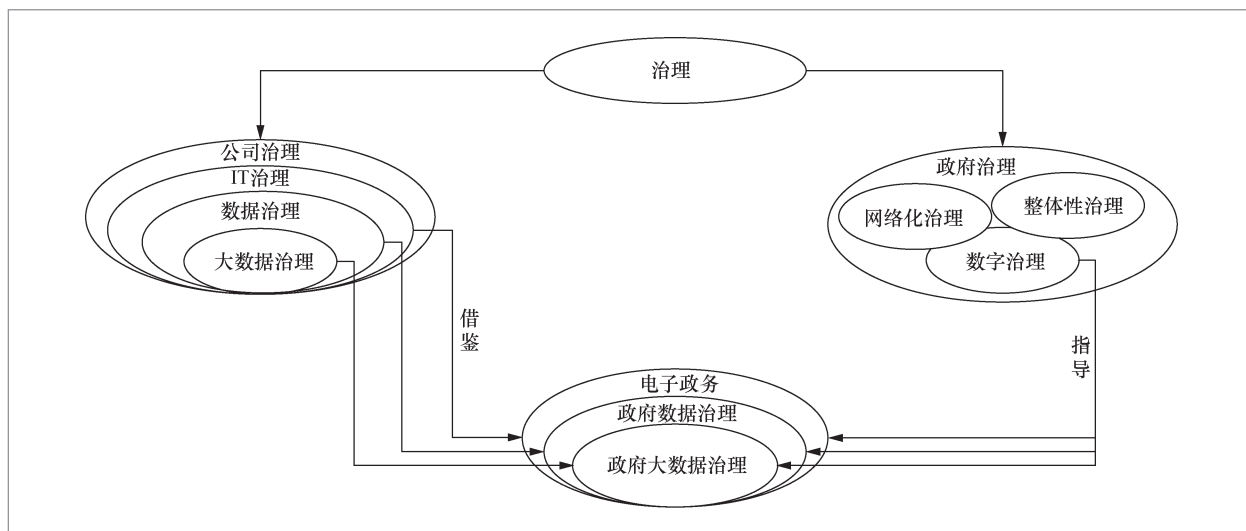


图2 政府大数据治理概念溯源

表1 政府和商业企业性质的主要差异

属性	政府	商业企业
目标	维持国泰民安,可持续发展	利益相关者获利
任务	保障公民基本权利,促进国民福利、经济的增长	发展竞争优势,提升客户满意度
决策类型	长期决策	短期决策
决策者	各种决策者	少数决策者
资金来源	税收	公司收入
活动性质	合作与监督	竞争和参与

务、开发/维持竞争优势来获取利润,通过提供价值来满足客户和其他利益相关者的需求。政府的主要目标是维持国家稳定、社会安宁,实现可持续发展,确保公民的基本权利,促进总体福利和经济的增长。大多数企业处在竞争激烈的市场环境中,决策由少数人参与,且决策时间短。政府的决策通常需要更长的时间,要大量参与者的协商和相互同意才能进行决策。因此,需要采取许多明确的步骤来降低风险,提高政府决策的效率和效力。商业企业的动力来自其参与市场竞争获得的收入和利润,因此企业通过提高生产效率,降低生产成本,开发具有竞争力的产品和服务,达到增加收入的目的。政府的收入主要来自

税收,其主要职责是协调各方利益,监督各方活动是否合规合法,以提高社会整体福利。因此,政府与企业之间的大数据属性以及应用也有所不同。

由于政府和企业性质存在差异,政府大数据和企业大数据的性质自然也存在差异,政府大数据和企业大数据性质的主要差异见表2。政府不仅要面对、处理由源头多、格式和价值密度不同导致的大数据集成的一般问题,而且还面临一些特殊的挑战。

- 数据收集困难。数据不仅来自多个渠道(如政府传统业务系统、社交网络、Web和众包),而且来自不同的源头(如政府机构、企业和民众)。

表2 政府大数据和企业大数据性质的主要差异

属性	政府	企业
容量	政府的数据仓库中积累了大量的数据	传统业务数据和机器生成的数据呈指数增长
种类	数据收集渠道多、来源广、格式多样，包括结构化、半结构化、非结构化数据	格式多样，包括结构化、半结构化、非结构化数据
速度	数据产生、传递、处理速度快	数据产生、传递、处理速度快
数据烟囱/数据孤岛	严重	少
数据集成、共享	困难	不难
数据安全性要求	高	不高
数据合法性要求	高	不高
数据合规性要求	高	不高

● 政府部门之间数据集成和共享困难。每个政府机构或部门通常都有自己的机密信息或公共信息仓库，各机构通常不愿共享他们认为专有的数据。每个系统都将其数据与其他系统进行隔离，即产生了数据孤岛，使得在政府机构和部门之间整合、共享数据变得更加复杂。

● 数据安全是政府大数据的主要属性，因此收集、存储和使用数据需要遵守政策法规。但是，当今大多数大数据技术（包括Cassandra和Hadoop）缺乏足够的安全工具，这使安全成为政府的另一项挑战。

● 高度监管的行业（如金融服务和卫生医疗）的合规性仍然是大数据政府项目收集和处理数据的另一个挑战，因为这些行业的数据收集和使用必须遵守国家的相关法律、法规和标准，如隐私保护、健康信息保护等。

企业大数据的计划和实施一般不受太多非技术因素的干扰，而政府大数据的应用无法突破政府中现有的政治和结构限制，需要政府制定战略、出台政策、制定法规、实施项目和评估影响等。政府大数据治理有2个视角。一个视角是对政府大数据的采集、存储、共享、管理和监督。首先，政府数据是大数据的一种，具有全面、细粒度、数量庞大、可被发现和重新利用等

特点。政府长期利用自身的行政数据和绩效数据实施基于证据的决策，如人口普查数据在数据多样性和变化速度上具有大数据的特征，因此常被看作大数据，并且作为政府政策分析的基础。其次，政府对自身的数据缺乏管理。Thompson等人^[17]指出政府机构中广泛存在数据管理不当现象，例如被评估为不适合拥有枪支的人拥有枪支、急诊科信息系统无需身份验证等限制就可以访问病人的隐私信息，他认为这些问题通常不是由现有规则或技术本身引起的，而是由缺乏完善的数据治理引起的。因此，政府需要制定政府数据治理标准，并评估其对政策和标准的遵守情况。另一个视角是针对特定政策问题的政府大数据利用。首先，政府缺少足以解决特定政策问题的数据，即负责机构通过常规手段收集的数据无法解决政府特定政策中提出的问题。其次，政府收集的数据主要用于管理和审计，而不是用于解释意料之外的某个问题的原因。因此，政府需要使用大数据解决特定政策问题，而现有的常规数据不能满足要求。世界银行组织发布的“政府大数据在行动（big data in action for government）”报告指出，政府利用大数据主要是为了提升公共服务效率和效能、提高政策决策水平、增加民众参与政府治

理的深度和广度。范灵俊等人^[20]提出的政府大数据治理指的也是政府大数据利用。

综上所述,本文研究的“政府大数据”是指政府机构通过多种渠道直接或间接收集的来自多个源头的多样性数据,包括结构化的、半结构化的和非结构化的数据。政府大数据具有大数据的特征,同时还具有其他大数据(如企业大数据)没有的特征,包括“数据孤岛”现象严重,整合和共享困难,数据的安全性、合法性、合规性要求高。“政府大数据治理”包含两方面含义。一是对政府大数据的采集、存储、共享、管理和监督。政府负责制定策略、出台政策、制定法规和标准,通过新技术平台和工具方法管理政府大数据。针对跨部门、跨领域的数据,通过制定自上而下的大数据管理制度和流程进行数据的管理、集成和共享,以整合政府各部门“数据烟囱”中积累的结构化或非结构化数据集。二是在各级政府部门以及各种政策领域应用政府大数据解决政府面临的挑战和问题。政府大数据应用程序的主要关注点在政府数据安全性、合法性、合规性、互操作性等方面;政府应用大数据的特定政策领域包括突发公共安全、突发卫生健康、犯罪、自然灾害、恐怖主义、国防等。

4 政府大数据治理的新挑战与新目标

政府大数据治理一方面是管理政府大数据,另一方面是利用政府大数据。政府数据的唯一性、高质量和高价值密度是政府决策和向公众提供公共产品和服务的重要支撑。除了政府部门,社会组织和商业企业也在积极探索如何利用政府数据。越来越多的企业提供的产品和服务依赖政府数据,如企业提供的天气服务、地图服务、金融服务、医疗健康服务、生活服务等。政

府数据开放获取和政府数据再利用一直是社会组织和企业关注的问题,但是政府部门暂未给出明确的答案。

从历史上看,绝大多数政府严格控制其数据/信息资源,发布的数据/信息通常采用印刷文件这种不方便在线传送且费用高的方式;对于数据/信息的再利用通常采用限制性许可或者通过交易基金等方式进行限制;提供的内容通常不是全部内容,只是部分内容。政府限制对政府数据的获取一方面源于政府保密文化,另一方面是政府逐渐认识到政府数据本身具有可利用的特征,可以从这些数据中获得商业价值和社会价值。过去10年来,在一些专注于开发公共部门信息市场的游说团体和开放政府和信息的自由活动家的倡导和宣传下,开放数据运动在全球有了相当大的进展,该运动要求政府以开放许可、可重复使用的格式提供原始政府数据。美国、英国等建立了开放政府数据平台(如data.gov、data.gov.uk),开放了非常多的政府数据集。开放数据运动强调了对政府开放数据的要求:“开放的数据必须是原始数据,而不是汇总或者被修改的数据”“开放数据可以自由使用、再利用和再分配,如果需要许可,最多只要满足属性和共享许可的要求”“数据可以通过网络批量下载”。但是现在开放的政府数据有很多并不是原始数据,而是衍生或者统计汇总的数据,只提供链接到查询接口或者进行交互式访问,而不提供批量数据集下载,大多数政府开放数据都只允许在非商用许可条款下使用,并对下游数据用户提出限制性要求^[21]。

如今,企业开发的很多产品和服务包含对政府数据的增值利用。在绝大多数情况下,政府是这些数据的唯一提供者,并且在很多情况下,政府也向用户提供同类的增值服务和产品,与企业进行竞争。在向用户提供服务和产品方面,政府与企业相比,

存在服务质量不足、价格过高等问题。

政府部门出于多种原因(如法定义务、公共任务、政府活动)收集并生产数据和信息,数据和信息涉及经济和商业、法律、政治、文化、环境、交通运输、科学、气象、水文、地理、卫生、社交等方面,范围十分广泛。这些数据和信息大部分可以供政府、企业和民众使用,作为信息来源、决策工具和新产品新服务开发的输入。这些数据和信息形成了政府大数据,除了一部分可以直接出售给消费者和企业外,还被政府用来开发自己的产品和服务,然后与企业竞争,政府数据增值再利用过程如图3所示。

政府数据增值利用涉及产权问题、服务质量问题、竞争问题。政府倾向于对政府数据拥有全部产权或者部分产权,一方面是为了从产权中获益,另一方面是为了鼓励产权保护下的创新。企业则认为政府数据的产权许可模式增加了企业利用政府数据的成本,妨碍了企业使用政府数据创新。在为用户提供高质量的服务方面,政府

部门可能存在动力不足的现象,主要是因为政府提供的服务往往属于垄断地位,不像企业存在很多的竞争对手。如果企业使用政府数据,相对于政府本身使用政府数据,可能开发出更有竞争力的高质量的产品和服务。但是如果政府不提供这些数据给企业,企业就无法开发新的产品和服务。有些政府数据(如地图数据)由于具有高附加值,只有少数机构或者企业能获得,形成垄断市场,破坏了竞争。

政府数据开放和政府数据再利用成为政府数据大数据治理的新挑战。若不建立适合政府数据开放和再利用的政策框架,将无法获得预期的效益。因此政府大数据治理的新目标就是建立和调整新的法律框架、监管框架,以各种方式监督并确保政府数据开放和再利用遵守相应的治理框架。

5 政府大数据治理的新机制

科学、顺畅、高效的政府大数据治理机制是实现上述目标的重要保证,翟云提出了大数据治理的框架结构^[22],本文在该框架的基础上进行了扩展,提出政府大数据治理机制框架,如图4所示,主要包括六大机制。

(1) 决策机制

大数据治理的对象是指为了保证大数据治理的责、权、利的一致而设计的规则、制度和流程。政府大数据治理决策机制的核心要点是建立与政府大数据应用相关的规则和程序,同时构建相应的组织结构,并确保决策权力机构及其对应的决策内容和范围与政府大数据治理的范畴相匹配^[23]。

(2) 政府数据开放共享机制

政府数据开放共享机制是指政府数据通过协同作用,由部门内部数据共享逐渐实现数据透明开放。它包含2层含义:第

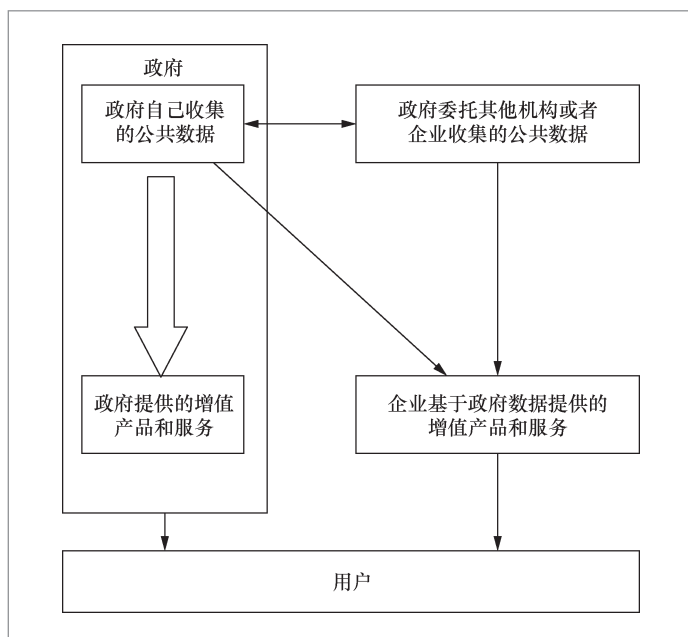


图3 政府数据增值再利用过程

一,为高效地实现数据共享和开放,需要建立适用法律,划清数据开放程度,思考数据价值实现等问题;第二,在原始数据价值实现层面,改变既往的政务信息公开方式,推动政府数据开放,实现政府部门数据资产的增值利用。自21世纪以来,大数据在信息公开、数据开放、大数据汇聚和综合利用这4个环节表现出其强势之处,同时还为挖掘数据资产潜在价值、提升机构协同治理水平以及促进大数据市场化开发奠定了重要基础^[22]。

(3) 政府数据资产可持续运营机制

政府数据资产运营是指集中精力进行大数据整理、开放、分析,结合实际业务优化操作流程,利用数据资产获取经济效益和社会价值。根据信息流理论,政府数据资产流包括数据的获取、存储、共享、开

放、传输、加工、利用等过程。供应商与用户之间发生数据交换,使数据复用,即数据在由供应商和用户组成的闭环价值生态链中流动,不断复用和增加价值,这就是政府数据可持续资产运作的基本形式。政府内部共享数据、对外开放数据、与其他政府机构共享数据3种方式构成了一种数据供应形式。政府数据通过数据存储、数据加工、数据产品开发和数据交易平台在供应商和用户之间流通,在整个过程中保持并增加价值^[24]。

(4) 政府公信机制

政府公信机制是指通过信息公开、信息共享,实现政府决策、执行、监督和有序融合等信息以及数据的公开,以便政府大数据能够在政府机构之间交换共享,让政府大数据支撑决策,允许政府数据被社会

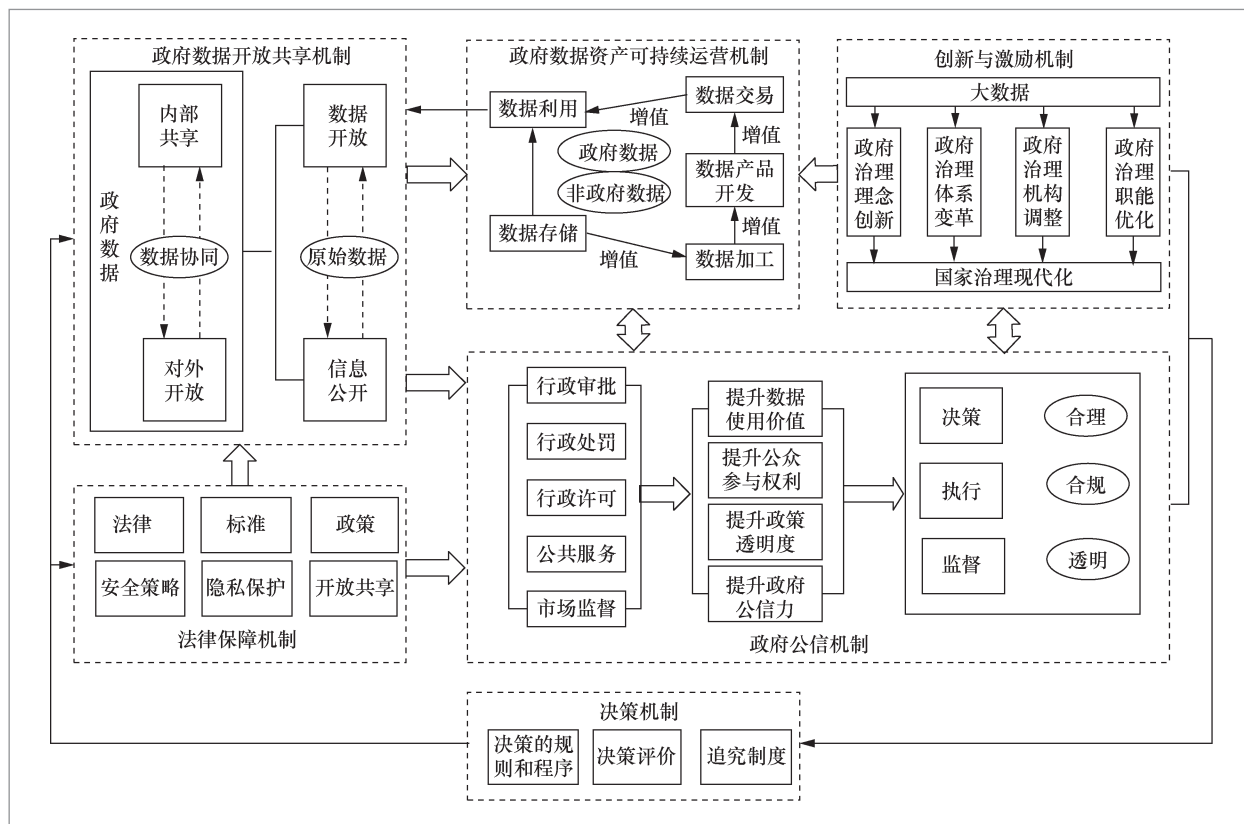


图4 政府大数据治理机制框架

化利用,同时促进新一轮政府改革,并不断提高政府的决策能力、执行力和透明程度,从而达到提升政府公信力的目标。实现这一目标的要点有3个:一是加大政府内部以及与市场相关的信息的公开和数据的开放;二是提升数据利用价值,鼓励群众积极参与开放数据的使用,增强群众对政府管理的主人翁意识;三是提升政策透明度,便于广大群众看见与理解,并更加实在地感受到参与政府数据治理带来的便利^[221]。

(5) 创新与激励机制

大数据已成为国家重要的创新突破方向之一。具体而言,这些改革和创新主要涉及理念创新、体系变革、机构调整和职能优化4个方面。建立符合人民标准、满足大众需求的服务型政府,明确各级机构职权划分,精简部门体系和行政管理流程,实现权力下放,以最大程度地激发市场参与者和参与者的活力,实现中国经济的可持续发展与社会进步^[251]。

(6) 法律保障机制

任何制度的有效实施都离不开完备的法律保障。在政府大数据治理过程中,建立健全、完善的法律保障机制需要明确资产权益归属、探讨开放制度以及保护个人隐私。数据保护以及个人隐私权已成为大数据时代每个公民的新权利,滥用和侵犯这些权利必将带来新的道德危机。依法加强对滥用数据和侵犯个人隐私行为的管理和处罚,也是大数据治理的内在需求。建立大数据使用方面的相关法律和政策以及相对应的标准规范势在必行^[221]。

6 政府大数据治理是政府管理的新形态

大数据时代背景下,政府对大数据的

有效治理可以催生许多新的政务、新的服务、新的形式,实现政府管理形式的创新,打破政府部门之间的数据孤岛。政府的数据治理能力将得到显著提升,不仅可以政府行为的受益者扩展到公众,还可以扩大公共服务的范围,提高公共服务的质量,更好地实现政策的预期管理。政府数据要发挥出社会效益,首先需要解决政府内部利用存在的问题。政府数据的内部利用可以分为3个阶段。第一阶段是跨单位使用(内部使用),即政府数据在政府内部跨不同单位使用,如市级政府不同单位共享数据。政府内部数据跨不同单位的使用,多数是在内部开放自建数据库,授权部门内部的其他单位使用。第二阶段是跨层级纵向使用,即政府数据同一领域内跨不同部门使用,包括一些垂直管理的领域,如司法、公安、监察、银行、住房和城乡建设等领域。同一领域内部不同部门之间的数据共享交换使用,除了技术上需要实现不同系统间的数据共享交换外,还需要解决不同部门由职责不同、利益关联等导致的管理问题。第三阶段是跨部门横向使用,即政府数据跨不同部门使用。跨部门使用主要是指部门之间进行信息交换和共享时,相互之间开放数据接口,签署合作备忘录。这种方式可以减少不同部门之间的数据信息不对称问题。跨部门使用可以看作一定程度上的相互协商、有限共享。除了政府数据的内部使用,还可以提供各种数据的外部服务,如政府数据开放、“互联网+”政府服务、新型公共服务、大数据应用等。政府大数据可以与企业合作提供协同数据公共服务,创新外部服务治理新形态。大数据的高效利用已成为实现政府治理体系和治理能力现代化的一条重要途径。政府必须为管辖范围内的大数据制定新的治理目标,找到新的理论支撑,发展新的体制机制,创新政府治理形态,积极回

应大数据时代提出的新要求。

政府大数据治理就是一种依托信息资源的政府管理新形态,离开信息资源,政府治理就无从谈起。大数据是一种数据形态的信息资源,在政府治理的全过程和全部职能活动中,都需要由大数据提供支撑。在政府治理过程中,从始至终充满着各种选择、判断、决断等具有决策性质的活动,这些活动是影响治理状态和结果的关键性、决定性因素。大数据对决策活动的支撑方式、支撑效果,与政府治理的基本特征及其所反映的客观需求具有高度的契合性,大数据可以比其他类型的信息资源更好、更充分地支撑政府治理。

政府管理的新形态基本特征包含以下五大方面。

(1) 政府治理的元主体共治特征。政府治理的相关决策活动既表现出活动种类的多样化,又呈现出治理过程的非官方、非传统正规和开放性特色,这表现在对决策信息的需求上,非海量、非多样、单一渠道、非初始性、非封闭性的政府数据不能满足包括党政机关在内的相关社会组织和公民个人、群体参与决策的需求,不能达到通过协调整合更多的社会力量有效处置更多、更复杂的社会事务的协同共治要求。传统意义上,政府管理基本上单纯依靠正规渠道的随机调查统计数据信息进行决策,但这种决策方式在当今的大数据时代背景下无法持续使用。

(2) 政府治理以服务为主要职能,以服务提供为职能实现的主要方式,政府大数据治理的决策活动不再仅仅是决策者的“主动”行为,而是需要及时响应服务对象不断发展变化的服务诉求,因此决策者必须更多地“被动”地进行各种决策。这样的决策活动需要更充分的数据信息储备,若存在数量不足、种类不齐、内容不够丰富、时效性不强、来源渠道单一、反映的关联

关系简单等数据处理过程中的缺陷,决策活动有效进行的基础条件将被严重削弱,进而严重影响决策活动的质量和效果。服务型政府治理中的“服务”和“服务提供”使治理主体不再像传统政府管理主体那样处于强势地位,若服务对象有合理的决策需要和要求,政府就必须及时反应,化消极和被动为积极主动。因此,在以服务为主要职能、以服务提供为职能实现的主要方式的情况下,决策活动就一定要用大数据。

(3) 政府治理以法律政策的颁布施行等间接手段作为主要和基本管理手段,政府大数据治理决策活动的重点就是在制定和施行有效处置多种多样的复杂社会关系的准则的同时,做出合理、科学的判定和选择。法律政策是处理各种复杂社会关系的产物和依据,因此,决策不可避免地要涉及高度复杂的社会关联。这时以反映包括因果关系在内的多种多样关联关系见长的大数据,理所当然地要比传统政府管理决策中主要依靠传统统计调查获得的随机抽样调查数据的信息质量更好、针对性更强,使用大数据协助完成决策活动有利于获得更高水平的决策效果。

(4) 政府治理在政府职能转变以及职能实现方式转变中的特殊要求,使行政审批等常规性职能活动必须完成从单纯的事前审批到事前审批、事中审批、事后审批相结合的转变。这种转变使政府监管方式发生根本性调整,传统的依靠申请人提供证据信息,政府部门进行审核验证,最终决定“许可”或者“不许可”的方式,必须让位给减少事前审批、强化事中和事后监督管理和服务的新方式。这种新方式要求政府部门必须有强大的信息捕获和及时处理能力,而大数据则可以利用自己海量处理和深度关联分析判断的长处,在这其间大有作为。例如,政府部门可以利用互联网上相关论坛和自媒体的数据以及上市公

司公开披露的数据,发现必须制止的行为及违法违规的“坏人坏事”,及时而有确凿证据支持地进行处罚并采取有效的预防措施,杜绝或大幅度减少违法违规行为。

(5) 政府治理强调综合应用多种管理工具的特征,此种特征时常导致决策活动超高频次发生,其中对不同管理工具的适用性和施行效果的预测准确程度成为决定决策活动质量和效果的关键。大数据拥有不局限于因果关系的各种关联关系的特点以及特有的低成本特性,使得大数据管理工具在选择应用决策活动中进行精准的预计和测度有了前所未有的优势条件,甚至可以说,在缺乏大数据的情况下,管理工具选择应用方面的盲目性是难以避免的客观事实,如今大数据的高效应用无疑将改变此前的面貌。

7 结束语

政府大数据治理是政府管理在大数据、“互联网+”时代背景下的新命题。一些国家已经启动了政府大数据应用项目,可提升公共服务的效率和效能、提高政策决策水平、增加民众参与政府治理的深度和广度。政府部门的大多数大数据应用和计划才刚刚开始或者正在计划实施,这意味着政府大数据应用仍处于早期阶段。政府大数据治理的概念、理论、机制、框架等还需要随着政府大数据应用的深入进行补充和修正,其中还有若干问题需要进一步探索,如政府数据确权、政府数据资产化管理、政府数据商业利用以及政府数据定价机制等。

参考文献:

- [1] SUN Z H. Managerial perspectives on intelligent big data analytics[M]. Pennsylvania: Information Science Reference, 2019.
- [2] Information Resources Management Association. Big data: concepts, methodologies, tools and applications[M]. Pennsylvania: Information Science Reference, 2016.
- [3] DE VRUEH R L A, DE VLIETGER J S B, CROMMELIN D J A. Public-private partnerships as drivers of innovation in healthcare[J]. *Frontiers in Medicine*, 2019,114(6): 20-35.
- [4] European Commission. Digital agenda for Europe: engagement for growth and jobs[M]. Luxembourg: Publications Office of the European Union, 2014.
- [5] BROEKEMA P C, BOONSTRA A J, VICTORIA C C, et al. Dome: towards the ASTRON & IBM center for exascale technology[C]//The 2012 Workshop on High-Performance Computing for Astronomy Data, June 18, 2012, Delft, The Netherlands. New York: ACM Press, 2012: 1-4.
- [6] CAROLE S. Horizon scanning: gathering research evidence to inform decision making[M]. Ottawa: The Conference Board of Canada, 2008: 1-8.
- [7] AGGARWAL A. Managing big data integration in the public sector[M]. PA: Information Science Reference, 2016: 280-320.
- [8] MASARU K. Challenge for Info-plosion[C]//The 10th International Conference on Discovery Science, October 1-7, 2007, Sendai, Japan. Berlin: Springer, 2007 :1-8.
- [9] 李国杰. 对大数据的再认识[J]. *大数据*, 2015, 1(1): 8-16.
- LI G J. Further understanding of big data[J]. *Big Data Research*, 2015,1(1): 8-16.
- [10] KIM G H, TRIMI S, CHUNG J H. Big-data applications in the government sector[J]. *Communications of the ACM*, 2014, 57(3): 78-85.
- [11] SANTISO C. Good governance and

- aid effectiveness: the world bank and conditionality[J]. The Georgetown Public Policy Review, 2001, 7(1): 1-22.
- [12] PANIAN Z. Some practical experiences in data governance[J]. World Academy of Science Engineering & Technology, 2010: 939-946.
- [13] AL-BADIA, TARHINIA, KHAN A. Exploring big data governance frameworks[C]// The 9th International Conference on Emerging Ubiquitous Systems and Pervasive Networks, November 5, 2018, Leuven, Belgium. Leiden: Elsevier, 2018: 271-277.
- [14] 韩兆柱, 翟文康. 西方公共治理前沿理论述评[J]. 甘肃行政学院学报, 2016(4): 23-39.
HAN Z Z, ZHAI W K. A review of western public governance theory[J]. Journal of Gansu Administration Institute, 2016(4): 23-39.
- [15] MARGETTS H, DUNLEAVY P. The second wave of digital-era governance: a quasi-paradigm for government on the Web[J]. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 2013, 371(1987): 20120382.
- [16] 孙宝文, 王天梅. IT治理模式及影响因素: 一个基于中国电子政务实施的多案例研究[J]. 中国软科学, 2012(6): 156-171.
SUN B W, WANG T M. IT governance patterns and their influencing factors: a multiple cases study on E-government implementation in China[J]. China Soft Science, 2012(6): 156-171.
- [17] THOMPSON N, RAVINDRAN R, SALVATORE N. Government data does not mean data governance: lessons learned from a public sector application audit[J]. Government Information Quarterly, 2015, 32(3): 316-322.
- [18] 夏义堃. 政府数据治理的国际经验与启示[J]. 信息资源管理学报, 2018(3): 64-72, 101.
XIA Y K. International experiences and implications on government data governance[J]. Journal of Information Resources Management, 2018(3): 64-72, 101.
- [19] 宋懿, 安小米, 马广惠. 美英澳政府大数据治理能力研究——基于大数据政策的内容分析[J]. 情报资料工作, 2018(1): 12-20.
SONG Y, AN X M, MA G H. A study on big data governance capability of the United States, UK and Australia: based on the content analysis of big data policy[J]. Information and Documentation Services, 2018(1): 12-20.
- [20] 范灵俊, 洪学海, 黄晔, 等. 政府大数据治理的挑战及对策[J]. 大数据, 2016, 2(3): 27-38.
FAN L J, HONG X H, HUANG C, et al. Challenge and countermeasure of governing government big data[J]. Big Data Research, 2016, 2(3): 27-38.
- [21] 杨孟辉. 开放政府数据: 概念、实践和评价[M]. 北京: 清华大学出版社, 2017.
YANG M H. Open government data: concept, practice and review[M]. Beijing: Tsinghua Press, 2017.
- [22] 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 innovation and development path of big data governance model in China[J]. E-Government, 2018(8): 12-26.
- [23] 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.
ZHEGN D Q, HUANG L H, ZHANG C H, et al. Concept and reference architecture of big data governance[J]. R & D Management, 2017, 29(4): 65-72.
- [24] 范佳佳. 政府数据资产可持续运营评估理论模型研究[J]. 上海商学院学报, 2018, 19(3): 21-27.
FAN J J. Research on the theoretical model for the sustainable operation assessment of government data assets[J]. Journal of Shanghai Business School, 2018, 19(3): 21-27.
- [25] 孟天广, 张小劲. 大数据驱动与政府治理能力提升——理论框架与模式创新[J]. 北京航空航天大学学报(社会科学版), 2018, 31(1): 18-25.

MEGN T G, ZHANG X J. Big data driven and government capacity building: theoretical framework and innovative

models[J]. Journal of Beijing University of Aeronautics and Astronautics (Social Sciences Edition), 2018, 31(1): 18-25.

作者简介



杨孟辉 (1969-), 男, 博士, 中国人民大学信息资源管理学院教授、博士生导师, 主要研究方向为情报学、信息计量。主持国家自然科学基金面上项目1项、重点项目子课题1项, 参与“863”计划重点项目1项、面上项目1项。近年来在《通信学报》《电子学报》《电子与信息学报》《清华大学学报》《北京邮电大学学报》《高技术通讯》等国内重要期刊, 以及国际重要期刊和国际会议上发表论文50余篇。曾获中国电子学会科学技术奖一等奖等奖项。



杜小勇 (1963-), 男, 博士, 中国人民大学信息学院教授、学术委员会主任、博士生导师, 中国人民大学校长助理, 数据工程与知识工程教育部重点实验室主任。兼任教育部科学技术委员会信息学部委员, 国家重点研发计划“云计算与大数据”专家组成员, 中国计算机学会理事、教育工作委员会主任、大数据专家委员会主任, 《大数据》副主编、全国信息技术标准化技术委员会大数据标准工作组副组长等。先后获得国家科技进步奖二等奖、北京市科技进步奖一等奖、教育部科技进步奖一等奖、中国计算机学会科学技术奖一等奖等奖项。

收稿日期: 2020-02-01

基金项目: 国家自然科学基金资助项目 (No.U1711263)

Foundation Item: The National Natural Science Foundation of China (No.U1711263)