

国家治理进程中的政府大数据开发利用*

王芳 陈锋

[摘要] 大数据对于推进国家治理现代化具有重要意义。本文探讨了国家治理现代化进程中政府大数据开放与利用的若干问题。首先辨析了大数据、政府开放数据的含义及相互关系。然后探讨了大数据对于国家治理的意义，提出大数据既是国家治理的环境和工具，也是国家治理的对象。在此基础上，分析了政府大数据开放与利用过程中的深层利益关系以及我国的现状和存在的问题，并提出几条对策建议，包括依法治理、民主协商管理、风险评估、成本效益评价与科学利用数据分析结果等。

[关键词] 国家治理 政府开放数据 大数据 公共管理

[中图分类号] D630

Openness and Exploitation Of Government Big Data in the Course of State Governance

Wang Fang Chen Feng

[Abstract] Big data has important significance in promoting the modernization of state governance. This paper discussed a few problems related to the openness and exploitation of government big data in the course of the modernization of state governance. It firstly analyzed the concept of big data, open government data and their relationship. Then it explored the influence of big data on state governance and put forward that big data is the environment, tool and object of state governance. On this basis, it analyzed the deep interest relationship in the openness and utilization of government big data and present situation and problem existing in China. At the end a few policy suggestions were provided, including governance according to law, management through democratic consultation, risk evaluation, cost-effectiveness assessment and scientific application of the results of data analysis.

[Keywords] State governance, Open government data, Big data, Public administration

近年来，大数据引起了企业界、科研机构和政府部门的高度关注。一些企业利用大数据技术实施精准营销或提高内部管理效能^[1]。目前，美国、英国等国家已开始利用大数据提高政府决策水平或改进公共服务，在教育、医疗、公共交通等领域已初见成效^{[2][3]}。中国共产党的十八届三中全会提出“推进国家治理体系和国家治理能力现代化”，使得我国政府大数据利用面临新的政策环境。大数据利用对推进国家治理体系和国家治理能力现代化将会发挥何种作用？政府如何更好地利用大数据进行国家治理？针对这一问题目前已出现了一些探讨，但为了更好地回答这些问题，还需要进一步的分析研究。本文拟在对相关实践与研究成果进行回顾的基础上，分析国家治理进程中的政府大数据利用问题，探讨当前我国需要解决的问题并提出相应的对策建议。

* 基金项目：国家社会科学基金重大项目“我国网络社会治理研究”（14ZDA063）。

作者：王芳，南开大学商学院网络社会治理研究中心主任，信息资源管理系教授、博士生导师；陈锋，南开大学商学院信息资源管理系博士研究生。

一、政府大数据与开放数据

(一) 大数据

目前，随着互联网应用的深化，尤其是物联网和移动互联网的广泛覆盖，产生了大量分散异质的数据。信息技术的发展使得实时采集和分析利用这些数据成为可能。近年来关于大数据的定义有很多，其中得到普遍认可的是“无法在可容忍的时间内用传统 IT 技术和软硬件工具对其进行采集、存储、管理和分析的数据集合”^[4]。人们对于大数据基本特征的认识也在不断深化^[5]，从“3V”^[6]扩展到了“7V”，包括：Volume（大量）、Variety（多样）、Velocity（快速生成）、Viscosity（黏性）、Veracity（准确性）、Volatility（时效性）和 Variability（易变性）等^[7]。作为一种观察世界的新方法，大数据将世界的本质看成是数据的集合，通过发掘混杂数据间的关系来认识世界^{[8][9]}，并在此基础上为用户提供把握需求、精准营销、个性化服务以及高效的决策支持。与企业、科研机构相比，政府作为社会资源的主要管理者和分配主体，掌握的数据类型更多，地域范围更广，时间跨度更大。自 1993 年实施“三金工程”以来，我国信息基础设施逐步完善，电子政务建设成绩显著^[10]。与西方国家相比，我国政府直接承担的公共管理事务更多，各级政府部门也积累了更多的公共数据资源。

根据数据产生的来源，可以将政府数据划分为五种类型：一是政府各部门内部行政管理过程中所产生的数据，如内部公文和人事数据，大部分与公共事务没有直接关系^[11]。二是政府在进行社会管理和提供公共服务的过程中实时产生的数据，这类数据种类多，规模大，不断更新，广泛存在于税务、工商、海关、教育、交通、卫生、公安、民政等承担公共管理与服务职能的部门，如工商企业登记、税务缴纳、公民户籍、道路视频等信息。三是由政府专门的职能机构所采集的社会管理数据，如政府统计部门定期采集的宏观经济、社会、人口方面的数据^[12]，以及气象、遥感、测绘、环境监测等数据。四是由政府通过业务外包或政府采购方式所获得的数据，如由财政资金资助的专项调查报告、科学研究成果数据等。五是互联网上的公开数据，如上市公司报表、互联网百科全书、社交媒体数据等，这类数据不仅可以为个人或企业所利用，也可以为政府所收集、整理和利用。在这五类数据中，除了第一类外，其它都具有大数据的多项特征，可以看作是政府所拥有的大数据。

根据数据管理机构的职能，可以将政府数据划分为三种类型：一是政府各部门直接管理的数据，如文书档案以及工商、海关、税务、公安、公共交通、文化等部门拥有的企业、个人及社会组织登记信息。随着电子政务的深化和智慧城市的推进，这一类数据的数量正在不断增加。二是由公共事业机构或国有企业管理、政府在必要时有权调取的数据，如医院、供水、供电、教育、图书馆等公共事业部门所存储的数据；三是由企业收集和管理的数据库，为了国家安全、查处腐败或侦破刑事案件，政府部门或司法机构在法律许可的条件下可以调取的数据，例如企业会计信息、公民通话记录等，但使用目的受到严格限制。

根据信息公开的程度，可以将政府数据划分为三种类型：依据《中华人民共和国政府信息公开条例》对社会完全公开的信息，例如政府政策法规、政府采购信息等；不完全公开的信息，主要是政府各部门内部管理使用的信息，例如税收征管信息、居民户籍信息；保密信息，涉及国家安全、商业秘密或个人隐私的数据，例如公民的通话记录和网络访问记录等。

(二) 政府开放数据

开放数据是移动互联时代政府信息公开的一种理念和实践。美国政府的研究人员将开放

数据定义为“公众可获取的，能够被用户完整观测和使用的数据”^[13]。开放数据是政府大数据得以开发利用的前提和基础^[14]。为了实现数据的开放和共享利用，需要依照一定的规则对数据进行加工和组织^[15]，将不同来源的数据进行语义关联处理，从而有益于人机交互。可操作性和标准化是关联数据的关键要素^[16]，一般采用 RDF（Resource Description Framework, 资源描述框架）数据模型，利用 URI(统一资源标识符)命名数据实体，通过 HTTP（超文本传输协议）表达并读取数据。美国和英国的政府信息发布都应用了关联数据的技术^[17]。可以说，开放数据使得政府数据可用，而关联数据强调数据发布和组织的规范，使得政府数据易用。

目前，世界许多国家都在积极推动政府数据的开放。美国、英国在推动开放数据、引导社会利用方面一直走在前列^{[18][19][20]}。2011年9月，在联合国大会上，巴西、印度尼西亚、墨西哥、挪威、菲律宾、南非、英国、美国等八个国家联合签署了《开放数据声明》，成立开放政府合作伙伴联盟(OGP, Open Government Partnership)。截至2014年2月24日，开放政府合作伙伴已由最初的8个成员国发展为63个成员国^[21]。目前我国中央政府的开放数据工作尚处于论证阶段，数据开放较少，部分地方政府在开放数据、引导社会力量参与方面已经取得了一定的进展。2014年上海率先实行政府数据资源向社会开放，通过政府数据服务网(www.datashanghai.gov.cn)提供数据查询、浏览、下载等功能，出自28个市级政府部门的190项数据内容成为重点开放的对象。总体来看，我国的政府开放数据进程和美国、英国等国家相比还有一定的距离，三国相关的立法与政策情况比较详见表1。

表1：美、英、中三国政府开放数据进程比较

国家地区	法律法规	行政命令或行动计划	数据开放日期	开放的资源
美国 [17][18]	1967年《信息自由法》(FOIA)生效；1976年国会通过《阳光下的政府法》(Government in the Sunshine Act of 1976)；1974年《隐私权法》(The Right of Privacy)；1988年《电脑匹配和隐私权法》(The Computer Matching and Privacy Protection Act)；1996年国会通过《电子信息自由法》；2014年颁布了《数字问责与透明法案》(Digital Accountability and Transparency Act)	2009年《透明和开放的政府备忘录》(Memorandum on Transparency and Open Government)；2009年《开放政府令》(Open Government Directive)；2013年《美国开放政府全国行动计划》(U.S. Open Government National Action Plan)	2009年 data.gov 上线；	截至2015年5月18日有132038个数据集；包含农业、商业、制造、海洋、能源、科学研究、地方政府、气候、消费者、生态、医疗、财政、公共安全等主题；多种数据格式及数据利用工具；汇集联邦政府、各州、大学等数据。
英国[17] [19]	2005年《信息自由法》(Freedom of Information Act)生效	2012年《开放数据白皮书》(open data white paper)；2011年《开放政府合作伙伴2013-2015英国国家行动方案》；2013年11月英国发布《八国集团开放数据宪章2013年英国行动计划》	2010年1月 data.gov.uk 上线；	截至2015年5月18日共有24753个数据集；
中国	2007年颁布《中华人民共和国政府信息公开条例》(行政法规)	2014年《当前政府信息公开重点工作安排》；国务院办公厅关于加强政府网站信息内容建设的意见 国办发〔2014〕	2013年7月北京市政务数据资源网上线；2014年上海市政务数据服务网上线。	截至2015年5月18日，上海市政务数据服务网共有508项数据。

（三）政府大数据与开放数据之间的关系

大数据与开放数据是当前两个备受关注的概念，二者既相互联系，又有一定的区别，需要从内涵和外延上加以辨析。

首先，从内涵上来看，大数据是数据存在的一种形态，强调数据的规模；而开放数据则是数据提供公共利用的一种方式，强调数据的无限制使用。二者并不在同一个分类框架之下。

其次，从数据范围上来看，政府开放数据是政府大数据的组成部分。政府大数据是政府拥有的各类数据，既包括互联网上的政府开放数据，也包括大量涉及个人隐私和商业机密的非公开数据。而开放数据的范围相对较窄，主要涉及具有公共利用价值的信息，不包括涉及国家安全、个人隐私和商业机密的数据。尽管开放数据可能包括少量由个人或企业收集、制作的数据集，但主要来源是政府数据。

第三，从数据存在形式上来看，政府开放数据追求以标准化格式与关联数据的方式向公众提供开放利用，是经过一定的组织整理而成的数据集，有时还附有支持数据利用的应用软件；而政府大数据则存在于不同的部门或系统中，具有多种格式，既包括经过加工整理的数据集，也包括实时产生、采集的各类数据。

第四，从利用目标来看，开放数据旨在促进政府的开放与透明，将具有公共利用价值的信息及时向社会公众开放，目的是方便公众使用、促进企业创新、改善公共服务；而政府大数据则关系到政府信息资源的积累保存、整合集成和开发利用，政府大数据挖掘和分析的主要目标是挖掘政府数据资源的价值，提高政府的公共决策水平，改善政府的公共服务质量。

第五，从开发利用的主体来看，开放数据的主要开发利用主体是社会公众、企业、社会组织和政府机构，目的是实现信息的增值利用。例如北京龙信公司基于国家经济户籍库，通过数据分析发现1990年至2011年我国财政收入与企业注册资本之间的关系呈高度线性相关，其相关系数高达0.987，而斜率为0.148，表明放开企业注册可以增加政府财政收入^[22]。而对于政府大数据而言，除了开放数据之外，大量的政府数据的开发利用以政府为主导，主要目的是为政府提供决策支持。虽然部分具体业务可以通过服务外包方式委托给企业、科研机构或社会组织承担，但政府依然是数据开发利用的主体。

第六，从开发利用方式来看，对于政府开放数据而言，既可以以数据集的形式直接提供浏览和下载，也可以对公众开放API（应用程序接口），提供查询接口，鼓励社会进行二次开发，还可以直接提供基于政府数据的应用软件，例如适时查找公厕、停车位的APP等。对于政府大数据而言，首先要实现信息的采集、清洗、集成，在此基础上才能实现信息的挖掘分析或开放共享。

可以说，政府开放数据是政府大数据的组成部分，是政府大数据开发利用的目标和方式；而政府大数据则是政府开放数据的基础和来源。

二、大数据开发利用对国家治理的意义

国家治理（state governance）现代化是指运用现代化治理的理念、方法和技术工具，对国家的政治、经济、社会、文化进行有效管理的过程，包括经济治理、社会治理、环境治理、文化治理以及各类组织治理，等等。不同于以单极治理为特征的国家统治，国家治理强调治理主体的多元化和治理工具的多样性。何增科认为，国家治理的主体是政权所有者、管理者和利益相关者等多元主体，强调合作管理，其客体是社会公共事务，其目的是增进公共利益和维护公共秩序^[23]。政府大数据是政府所拥有的重要信息资产，对其进行有效的开发利用

将会提高政府的决策水平，增强政府的社会管理和公共服务能力，对于推进国家治理体系现代化具有重要价值。可以说，大数据既是国家治理的环境和工具，同时也是国家治理的对象和结果。

（一）大数据是当前国家治理的环境

目前，互联网的应用已经渗透到经济、社会、文化生活的各个层面，并产生了大量的数字化数据。2011年，全世界产生和复用的信息约有1.8ZB^[24]，而到了2013年已经达到4ZB^[25]，两年时间内翻了一倍多。这些数据不仅作为静态信息资产存在，而且反映了经济、社会、政府组织动态运行的特征、质量和存在的问题。可以说，在移动互联时代，国家治理处于一个全新的大数据环境之中，对国家治理主体的行为准则、策略选择、相互关系以及治理效果的评价都提出了新的挑战。胡洪彬认为大数据的利用有利于扩展国家治理的主体范围，治理过程可以更透明^[26]。刘叶婷、唐斯斯认为大数据对政府治理理念的影响是开放与包容，并以“智能化”重塑治理范式^[27]。孙粤文认为大数据可以从产业转型升级、公民政治参与和政治发展的新通道、社会管理、国家安全等方面提升国家治理能力^[28]。透彻理解大数据环境的复杂性，充分把握大数据环境多变、透明、互联互通的特点，有效应对大数据环境带来的意识形态斗争、信息安全、文化渗透等风险，是当前国家治理必须要考虑的问题。

（二）大数据是国家治理的工具

国家治理能力现代化要求治理手段与时俱进，符合时代的要求和趋势，体现出科学化、信息化与结果的有效性^[29]。大数据的开发利用可以为治理主体提供更加快速准确的决策信息，帮助治理主体及时了解国内外的局势、公众的需求、政策效果反馈信息以及存在的问题，从而成为国家治理的重要工具。麦肯锡公司2011年的研究报告发现，大数据技术可以从提高透明度、发现需求、提高政策的针对性、辅助决策和引导管理部门内外创新等五个方面改善公共管理^[30]³⁷⁻⁹⁷。美国白宫行政办公室经过3个月的调研于2014年5月发布了一份近80页的研究报告，论述了大数据和开放数据对美国政府和个人隐私的影响，指出大数据在医疗卫生、教育和国土安全等公共服务领域的重要价值^[31]。新加坡利用智能交通综合信息管理平台预测交通流量和流速的准确率达到85%^[32]。马慧民和周曦民认为大数据利用可以在科学决策、民主发展、限权反腐、危机预测以及简政放权等国家治理的各方面起到巨大的推动作用^[33]。Talking Data利用上海移动设备的大比例抽样数据进行分析，结果发现加班指数最高的公共部门依次是城市规划部门、人民政府和公安系统以及税务、财政和发改部门，相比较而言，市级部门的公务员平均加班比率最低，通勤距离也较短，而镇级加班率最高，通勤距离达到了9.5公里^[34]。这些数据为政府行政管理提供了科学准确的依据。

（三）大数据是国家治理的对象

国家治理的目标是保障国家安全、促进经济繁荣、增进公共利益、实现社会公平正义和维护公共秩序。国家治理的对象涉及到各类组织、社会单元、生态环境、公共事务以及政治、经济或社会活动。大数据既是各类组织和社会单元活动的记录，也是特定公共事务或社会活动的表征，因而成为不可忽视的治理对象。由于政府大数据产生于不同的部门、机构和系统中，存在着部门分割、元数据标准不统一、数据质量参差不齐、共享困难等种种问题，因此需要进行科学治理。目前我国实施政府数据治理统筹治理的典型是广东省佛山市南海区。南

海实行先统筹后开放的策略,通过政府数据资源目录与交换平台,建设主题数据库和跨部门综合管理服务平台实现数据共享、业务协同,并为决策提供支持。截至2014年底,数据资源目录平台和数据资源服务平台已经实现了工商、质检、劳动、社保、民政、卫生、财政、公安、教育等66个单位的数据注册、查询、共享和交换。数据资源目录平台注册数据表共802个,数据指标达6930项,数据资源服务平台平均每月跨部门数据交换批次达600多次,交换总量超过4000万条^[35]。南海数据统筹局结合网络安全协调委员会、专家咨询委员会与数据治理委员会对数据进行统筹治理。借助于有效的政府数据治理,南海市在公共服务、社会网格化治理、物联网监管、重点人群社会关系、舆情管理、综合治税等方面取得了显著的成效。

三、政府大数据开发利用过程中的利益关系分析

大数据在为国家治理带来机遇的同时,也提出了新的挑战,如增加了治理复杂性和信息安全风险^{[26]91},对数据的连续性管理和个人信息保护提出了更高的要求^[36],等等。为了更好地解决这些问题,需要对其背后深层的利益关系进行分析。

一是价值取向的争议。开放与保密是一对矛盾。就政府开放数据而言,持积极态度的学者一般认为政府数据开放可以更好地方便公众利用,促进产业创新,增加政府透明度,从而增加数据价值;而持谨慎态度的学者则认为数据开放涉及到个人隐私、企业机密和国家信息安全,开放数据运动隐含数据权利的概念,同时造成不公平和低效率^[37]。这两种观点的冲突实质上是一个价值取向的问题:在市场效率与社会公平、个人价值与社会价值之间,哪个应当优先考虑?只有首先确定价值取向,才能细化制度建设^[38]。就当前我国的情况而言,应当加快个人隐私保护、国家信息安全立法的进程,制定大数据交易的法规与实施细则,使得政府数据的开放与大数据的开发利用有法可依,健康发展。

二是权益分配问题。大数据的开发利用涉及多个利益主体,不解决权益分配问题,就难以实现跨部门的数据共享。例如个人数据的利益相关者包括个人、个人数据收集者、个人数据处理者、个人数据应用者及监督者5种角色,涉及个人、数据企业、政府、媒体、第三方隐私保护组织等利益主体^[39]。数据权利涉及所有权、控制权、使用权,还包括交易权、查询权和收益分享等权利。已有一些案例反映了关于数据权利的争议。2013年北京市的高考成绩还未公布,网上已有人出售2013年高考生信息,包括手机号、监护人、学校、年级、年龄等,2.1万条信息售价800元^[40]。教育部门是否可以把高考成绩出售给信息服务公司,医院能否出售病人病历?王芳认为,政府信息资源的所有权属于社会公众,公民有权利获取和使用政府信息,并享受部分收益权,作为管理者,政府在信息公开时应当保护所有权人的利益,比如隐私或国家安全^[41]。

三是效益评价与成本考量。在给定的社会经济发展水平与技术条件下,信息公开与共享的收益成本比大于垄断状态^[42]。政府数据的开放与共享涉及到数据的采集、存储、整理、发布、挖掘和分析等整个过程,需要行政机关付出人力和物力等成本。但是,在数据拥有部门和数据使用部门之间的成本分担、收益分配以及权责归属却是难以清晰界定的问题。例如由公安部门管理的户籍信息是许多业务部门需要的数据,但是与其他部门或社会公众共享户籍数据不仅会给公安部门带来额外的工作量,而且存在个人信息泄露的巨大风险。在政府数据的管理开放过程中,所投入的成本容易量化,但是其经济社会效益与潜在风险却难以量化计算。为此,需要明确界定政府数据部门的职能、人员编制及财政预算,建立政府大数据利用的成本效益评价标准。

四是数据管理变革与风险防范。随着电子政务的深化,政府信息管理部门积累了业务、

财务、人事等多种数据，但是对信息部门的内控管理并没有引起足够的重视，数据泄露的风险不断增加。2014年12月底，铁路客户服务中心12306网站大量用户数据泄露并在网上传播售卖，包括13万条账号、密码、手机、身份证号、邮箱地址等个人私密信息^[43]。知名漏洞响应平台曝光江苏、陕西等全国至少19个省份的社保系统存在漏洞，数千万用户的社保信息遭遇泄露危机^[44]。武汉上百个新建、在建的楼盘约100万条个人信息被通过网络转让出售，包括业主姓名、电话、身份证号、住址及房屋面积等隐私^[45]。为了防范数据开发利用过程中的风险，一些地区在数据管理体系上进行了新的探索，例如广东省佛山市南海区成立了数据统筹局来统筹管理各部门的数据，在管理过程中通过数据治理委员会来决定数据的开放层次和物理存放地点。当然，数据的集中管理也存在着一定的安全隐患，比如，技术人员将有机会接触更多敏感数据，增加了数据泄露风险，美国斯诺登事件表明技术人员可能会泄露敏感数据。

四、政府大数据开发利用的现状、问题和对策

大数据作为国家治理的工具，具有透明、科学、民主等优点。政府在开发利用大数据方面具有独特的优势：作为数据拥有者，政府掌握的数据更多更广；作为数据提供者，政府可以向社会免费开放数据，还可以通过税收减免、资金补贴等方式调动社会各界参与数据的开发利用；作为数据的利用者，电子政务、网络问政、智慧城市的发展都为政府大数据的开发利用提供了巨大的需求空间。目前，世界许多国家在政府数据开放和开发利用方面取得了进展，但同时也存在不少问题。

（一）当前政府大数据开发利用取得的进展

从规划层面来看，目前美国、英国、澳大利亚等国都制定了国家层面的大数据战略和规划^[46]，通过立法、制定规划和政策支持等形式加强对大数据的研究和利用，引导社会组织参与利用政府大数据，在社会治安、天气预报与应用、交通管理、卫生保健等方面有较多应用案例^[47]。2012年3月，奥巴马政府在白宫网站发布了《大数据研究和发展倡议》^[48]。2012年5月，美国数字政府战略发布，提出以信息和客户为中心，改变联邦政府工作方式，尽可能开放数据，为美国民众提供更优质的公共服务^[49]。欧盟将数据信息基础设施作为 Horizon 2020 计划的优先领域之一，并专门征集针对大数据的研究项目，例如预算为5000万欧元的 FP7Call 8 项目^[50]。

目前我国尚未出台独立的大数据法律或规划，一些部委和省市开展了大数据研究和产业规划的前期工作。2013年5月国家发改委和中科院启动了“基础研究大数据服务平台应用示范”项目，着手研究大数据处理与分析的关键技术，提供在线基础研究大数据的集成共享与知识发现服务，并在天文、材料领域开展应用示范^[51]。2013年科技部将大数据列入973基础研究计划；2013年度国家自然科学基金指南中，管理学部、信息学部和数理学部都将大数据相关课题列入其中。与此同时，一些地方政府纷纷出台大数据相关产业的发展规划，例如上海市2013年发布了《上海推进大数据研究与发展三年行动计划》，同年重庆市制定了《重庆市大数据行动计划》。这些地方规划以市场需求为主导，把大数据作为有发展潜力的产业进行培育，重视大数据利用的基础设施建设、技术研究和产业扶植。同时，也提出政府应该在民生服务（医疗等）、城市管理（交通、治安等）等方面发挥大数据应用的示范作用。在具体的举措上，都提出探索制定地方性法规制度、加强组织协调和推进数据共享。

从实践案例来看，美国以开放数据为先导，鼓励各类社会主体参与利用政府数据，公众不仅利用政府数据改善公共服务，也可以对政府实施监督，例如对白宫访客记录的分析就是对公务接待的监督^{[52][221-239]}。目前我国政府大数据主要用于决策支持和舆情管理，如传染

病预警预报和卫生舆情监测预警^[53]。从实施主体来看，目前单个政府部门实施较多，跨部门协作实施较少，社会主体参与政府数据利用也较少。上海市交通管理部门通过交通综合信息平台全面、实时整合、处理全市道路交通、公共交通、对外交通领域车流、客流、交通设施等多源异构基础信息数据资源，交管部门可以根据实时路况采取相应的处理措施^[54]。审计机关在工作中可以接触财政、企业、工商管理、金融机构、水利等不同行业和部门，采集的数据范围广，可以开展跨机构的大数据开发利用。例如利用某市地税局税收征管数据、社会保障中心的失业保险数据、劳动就业服务局的失业人员登记数据和财政局的财政资金审核拨付数据进行整理分析，结果发现市保险中心少征失业保险费 1291 万元^[55]。人民网舆情监测平台通过对网页中新闻的跟帖评论和网络社交软件的数据进行分析，适时监测网络舆情动态^[56]。

（二）当前我国政府大数据开发利用存在的问题

总体来讲，目前我国政府数据的开放和大数据开发利用尚处于起步阶段，还存在一些问题：

第一，政策法规与技术标准的建设较为滞后。政府大数据的开发利用涉及到数据采集、处理、保存、挖掘和分析利用等多个环节，每一个环节都需要有相应的制度法规和技术标准作为依据。目前我国除了《政府信息公开条例》及其实施细则之外，尚未出台针对政府数据开放和开发利用的政策法规，在一定程度上制约了政府数据的开放和开发利用步伐。

第二，政府数据的统筹管理才刚刚起步，部门间的数据共享利用还存在很大的障碍。统筹管理可以促进数据共享，提高行政审批效率。目前政府正在大力推进行政审批改革，2013 年 3 月出台了《国务院机构改革和职能转变方案》。2015 年 5 月，改革委《关于 2015 年深化经济体制改革重点工作意见》要求“再取消和下放一批行政审批事项，全部取消非行政许可审批，规范行政审批行为，推广网上并联审批等新模式；……实施企业投资项目网上并联核准制度，加快建立健全投资项目纵横联动协同监管机制等”。天津市滨海新区政府已经率先实现了“一颗印章管审批”的改革，取得了初步的成效。但是，网上并联审批需要各审批机构协同工作，这就需要对分散于不同部门和机构的数据进行统筹管理。目前除了广东南海之外，其它地方的数据统筹尚未正式启动。

第三，政府数据开放进程缓慢。目前北京、上海的政府数据开放已取得了一定进展，但是开放的种类、数量还比较有限，主要集中于公共交通、医疗卫生、舆情监控等领域，在满足用户实际需求方面还有较大差距，而中央部委和其它地区的政府数据开放基本上尚未启动。

第四，政府数据的质量和可用性有待提高。为了使服务对象能够更加方便快捷地使用政府数据，需要采用统一的元数据标准和方便用户下载浏览的数据格式，提高数据的真实性、准确性和完整性。

第五，政府大数据开发利用的技术平台建设尚不够完善。与商业企业相比，目前政府开放数据网站以及应用软件的建设都还不够完善丰富，需要进一步加强。

（三）我国政府大数据开发利用的对策

政府大数据的开发利用应该符合国家治理体系现代化的根本目标，坚持法治化、民主化、市场化与科学化导向。针对当前我国政府大数据开发利用过程中存在的问题，可以从以下方面探索解决办法：

一是加快立法进程，在立法、规划、管理机构调整等多方面展开布置，明确数据公开和使用的范围，明确政府、企业、社会团体、个人在数据收集、存储、利用过程中的权利和义务，实现依法对数据进行采集、管理、开放与利用。美国在政府开放数据进程历经数十年时间，先后通过了《信息自由法》、《电子信息自由法》、《数据质量法》、《开放政府法》等法案^[51]，值得我国借鉴。

二是用民主协商的方式解决价值观争议和权益分配问题。政府大数据开发利用面临复杂

的利益和风险问题,需要通过民主协商方式细致协调各方利益,寻求有效的解决方案。对一些有争议的问题进行深入分析,准确定位,制定具体明确的解决方案,兼顾数据开发利用的公平与效率。

三是对政府开放数据实施风险评估。政府数据的开放利用存在着一定的风险,涉及个人隐私与企业秘密泄露以及危害国家安全等问题。因此,需要针对不同类型政府数据的开放与利用进行风险评估,制定合理的风险防范策略,完善相关的法律法规,保证数据的安全利用。

四是采用经济分析方法进行大数据开发利用的成本效益评价,运用市场化的方法对成本与收益进行合理分配。没有科学合理的效益评价,政府数据开放和利用就难以有政策依据和稳定的财政保障。

五是科学利用大数据分析结果。科学合理地使用大数据是国家治理能力现代化的要求。大数据分析结果通常只是揭示了相关关系,而相关关系并不一定就是因果关系。大数据关于样本等于总体的假设通常是不符合事实的,因而统计学中的抽样误差问题依然存在^[57]。“数据驱动的科学”鼓励人们进行被动的数据收集,而不是主动的实验与测试,也诱使人们搞“数据欺瞒”(hornswoggling,指缺乏统计学依据的数据欺诈)^[58]。对大数据结果的偏信和轻率接受,可能导致错误的决策。因此大数据的开发利用应结合正确的理论和其它科学分析方法,实现科学管理与利用,避免滥用和误用。

参考文献

-
- [1] C.L. PHILIP CHEN ,CHUN-YANG ZHANG. Data-intensive Applications, Challenges, Techniques And Technologies: A Survey On Big Data[J]. *Information Sciences*, 2014, 275(11): 314-347.
 - [2]王茜. 英国大数据战略分析[J]. 全球科技经济瞭望, 2013, (8): 24-27.
 - [3]Executive Office of the President. Big Data Across the Federal Government [R/OL]. 2012-3-29 [2015-3-2].
https://www.whitehouse.gov/sites/default/files/microsites/ostp/big_data_fact_sheet_final_1.pdf
 - [4][30]The Mckinsey Global Institute . Big Data: The Next Frontier For Innovation, Competition, And Productivity[EB/OL]. 麦肯锡公司:
http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation, 2011
 - [5] GANDOMI A, HAIDER M. Beyond The Hype: Big Data Concepts, Methods, And Analytics[J]. *International Journal of Information Management*, 2015, 35(2): 137-144.
 - [6]李国杰,程学旗. 大数据研究: 未来科技及经济社会发展的重大战略领域——大数据的研究现状与科学思考[J]. 战略与决策研究. 2012(6):648.
 - [7]IBM Center for The Business of Government. Realizing the Promise of Big Data: Implementing Big Data Projects[R/OL]. [2015-03-24].
http://www.businessofgovernment.org/sites/default/files/Realizing%20the%20Promise%20of%20Big%20Data_0.pdf
 - [8]张引,陈敏,廖小飞. 大数据应用的现状与展望[J]. 计算机研究与发展. 2013(50):219-233.
 - [9]维克托·迈尔-舍恩伯格,肯尼思·库克耶. 大数据时代:生活、工作与思维的大变革[M]. 杭州:浙江人民出版社, 2013: 1-60.
 - [10]王芳,赖茂生. 我国电子政务发展现状与对策研究[J]. 电子政务, 2009(8):51:57.
 - [11]王芳. 政府信息共享障碍及一个微观解释[J]. 情报科学, 2006, 24(2):194-199, 221.

-
- [12]徐继华,冯启娜,陈贞汝.智慧政府:大数据治国时代的来临[M].北京:中信出版社,2014:20-21.
- [13] M. BURWELL SYLVIA, VANROEKEL STEVEN, PARK TODD, etc.. Open Data Policy--Managing Information as an Asset[R/OL]. 2013-5-9 [2015-4-15].
<https://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf>.
- [14]Executive Office of the President. Big DATA :Seizing opportunities, Preserving Values[R/OL]. 2014-5-1 [2015-3-2].
https://www.whitehouse.gov/sites/default/files/docs/big_data_privacy_report_may_1_2014.pdf
- [15]谭健. 开放数据及其应用现状[J]. 图书与情报, 2011, (4): 42-47.
- [16]FLORIAN BAUER, MARTIN-KALTENBÖCK. Linked open data: the essentials. a quick start guide for decision makers[R]. Vienna, Austria: monochrom, 2012:30-36
- [17]刘炜. 关联数据: 概念、技术及应用展望[J]. 大学图书馆学报, 2011, (2): 5-12.
- [18]吴曼. 开放数据在英、美政府中的应用及启示[J]. 图书与情报. 2012, (1) :127-130.
- [19]Making open and machine readable the new default for government information [R/OL]. <http://www.whitehouse.gov/the-press-office/2013/05/09/executive-order-making-open-and-machine-readable-new-default-government>
- [20]Cabinet office and paymaster general. Open Data White Paper [R/OL]. 2013-6-27[2015-6-1] <http://data.gov.uk/library/open-data-white-paper>
- [21]Cabinet Office. Open Government Partnership UK National Action Plan 2013 to 2015 [EB/OL]. :<https://www.gov.uk/government/consultations/open-government-partnership-uk-national-action-plan-2013-to-2015>
- [22]刘维涛. 大数据: 政府治理“如虎添翼” [N]. 人民日报, 2014-05-2 (20)
LIU WEITAO. Big Data: An Excellent Assistant Of Government Governance[N]. *People's Daily*, 2014-05-2 (20)
- [23]何增科. 国家治理及其现代化探微[J]. 国家行政学院学报. 2014(4):11-14.
- [24] Gantz J, Reinsel D. Extracting Value from Chaos, IDC. 2011-6-30
[2015/5/19] :<http://www.emc.com/about/news/press/2011/20110628-01.htm>
- [25] Daily Infographic . “2016:the Year of the Zettabyte.”, 2013-3-23[2015-4-3]<http://visual.ly/2016-year-zettabyte>
- [26]胡洪彬. 大数据时代国家治理能力建设的双重境遇与破解之道[J]. 社会主义研究. 2014(4):90-95.
- [27]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) .
- [28]孙粤文. 思维与技术——大数据时代的国家治理能力现代化[J]. 行政领导, 2015 (2): 19-23.
- [29]郑慧,何君安. 试论国家治理体系和国家治理能力现代化[J]. 新视野. 2014(3):38-43.
- [31] Executive Office of the President. Big data: seizing opportunities, preserving values [R/OL]. www.whitehouse.gov/BigData.
- [32]黄小勇 王鉴. 政府大数据重在服务和人本关怀[N/OL]. 学习时报. 2015-5-11[2015-5-20].
<http://www.studytimes.cn/shtml/xxsb/20150511/11135.shtml>.
- [33]马慧民,周曦民. 大数据拓展国家治理空间[J]. 社会观察, 2014 (8): 36-38.
- [34]团支书. 大数据告诉你,哪些公务员最辛苦[EB/OL]. 微信公众号”城市数据团”.
<http://dy.qq.com/article.htm?id=20150519A0021R00>.
- [35]赵越 黄艳姿. 这半年,数据统筹局做了什么? [N/OL]. 南方日报, 2014-12-5. <http://news.163.com/14/1205/05/ACM9BGSM00014AED.html>.
- [36]安小米. 现代国家治理的云端思维——信息治理能力与政府转型的多重挑战[J]. 人民

论坛·学术前沿, 2015, (2).

- [37] 吴冬琪. 论财产权视野下当代数据权利的保护[D].: 复旦大学, 2013: 6-17.
- [38] 刘文静, 陈耿. 政府信息共享制度建设的价值取向[J]. 电子政务, 2014, (10): 25.
- [39] 王忠, 殷建立. 大数据环境下个人数据隐私治理机制研究--基于利益相关者视角[J]. 技术经济与管理研究, 2014, (8): 71-74.
- [40] 于海波. 高考考生信息网上叫卖 2 万条信息仅售 800 元. 北京晚报, 2013-06-14, http://www.chinadaily.com.cn/dfpd/dfjyzc/2013-06-14/content_9312923.html.
- [41] 王芳. 政府信息资源的经济学特征及其产权界定. 图书情报工作, 2005, 49(5): 50-55.
- [42] 王芳. 政府信息公开与共享的成本收益分析[J]. 南开管理评论, 2004, 8(5): 74-81.
- [43] 魏兵. 大数据时代, 谁能阻止信息泄露[EB/OL]. 中国经济网, 2015-03-03. http://www.ce.cn/xwzx/gnsz/gdxw/201503/03/t20150303_4715669.shtml
- [44] 周蕊, 邓中豪, 徐博. 多地社保系统现信息安全漏洞追踪[EB/OL]. 新华网. 2015-04-23[2015-5-2]. http://news.xinhuanet.com/local/2015-04/23/c_1115071365.htm.
- [45] 冯国栋 李劲峰. 武汉百万业主个人信息泄露 9 名嫌疑人被刑拘[EB/OL]. 2015-05-25. <http://house.people.com.cn/n/2015/0525/c164220-27053148.html>.
- [46] 刘兰, 闫永君. 澳大利亚公共服务大数据战略研究[J]. 图书馆学研究, 2014, (5): 47-51.
- [47] WULLIANALLUR RAGHUPATHI, VIJU RAGHUPATHI. Big Data Analytics In Healthcare: Promise And Potential[J]. *Health Information Science and Systems*. 2014, (2): 3.
- [48] Tom Kalil. Big Data is a Big Deal[EB/OL]. <http://www.whitehouse.gov/blog/2012/03/29/big-data-big-deal>.
- [49] Office of the Press Secretary. Digital Government Strategy [EB/OL]. U.S. department of state. [2015-5-3]. <http://www.state.gov/digitalstrategy/>
- [50] 李国杰, 程学旗. 大数据研究: 未来科技及经济社会发展的重大战略领域——大数据的研究现状与科学思考[J]. 中国科学院院刊. 2012(6): 649.
- [51] 国家发改委“基础研究大数据服务平台应用示范”项目启动会在网络中心举行[EB/OL]. 中国科学院信息化工作网, 2013-5-17[2014-7-2]. http://www.cnict.ac.cn/xw/rdxx/201305/t20130517_3841618.html
- [52] 涂子培. 大数据: 正在到来的数据革命[M]. 桂林: 广西师范大学出版社, 2012: 1-331
- [53] 马家奇. 公共卫生大数据应用[J]. 中国卫生信息管理杂志, 2014, (2): 174-177.
- [54] 上海市交通综合信息平台点亮智慧城市 [EB/OL]. 2013-04-16[2015-5-2], 中国交通通信信息中心, <http://info.cttic.cn/information/detail-53.html>.
- [55] 彭华彰等, 社会保障审计理论与实务[M]. 北京: 中国时代经济出版社, 2007: 169-182.
- [56] 唐涛. 基于大数据的网络舆情分析方法研究[J]. 现代情报. 2014, 34(03): 3-6.
- [57] HARFORD TIM. Big Data Are We Making A Big Mistake[EB/OL]. Financial Times. 2014-3-28 [2015-4-28]. <http://www.ft.com/cms/s/2/21a6e7d8-b479-11e3-a09a-00144feabdc0.html#axzz2xZaSBx6o>.
- [58] Frické, M. Big Data and its Epistemology[J]. *Journal of the American Society for Information Science and Technology*, 2015, 66(4): 651-661.