

推进开放政府数据： 提升政府部门大数据能力^{*}

樊 博

内容提要 开放政府数据能够助力政府实现公共治理变革。近年来,我国各级政府均在进行开放政府数据工作,但目前存在开放数据量少、质量不高、再利用性低等问题,这些问题与缺乏政府部门大数据能力有关,因此,探索如何提升政府部门大数据能力极为必要。通过分析我国目前开放政府数据的现状与美国的案例,本文认为,可以通过探索首席信息官制度、加强数据脱敏等工作、推动法制建设等方式来加强政府部门大数据能力,从而将我国开放政府数据工作落于实处,充分发挥政府大数据的巨大价值。

关键词 开放政府 大数据能力 治理
DOI:10.16091/j.cnki.cn32-1308/c.2018.02.002

引 言

大数据时代的来临,为创新政府治理、提高政府服务效率提供了有利条件。近年来,我国大力推行大数据发展战略,其中,开放政府数据是重点工作之一。开放政府数据拥有诸多好处,一是能增进公众与政府的互动,转变公众被动沟通为主动参与决策;二是能借助外界力量充分挖掘政府数据,实现增值利用;三是推动“互联网+政务”工作,创新政府治理模式,提高政府服务效率。因此,推进开放政府数据,创建开放数据平台,成为各级政府主动适应新形势、创新政府治理的重大举措。自2012年来,我国已有近20多个省市建设开放数据平台,部分省市增设大数据管理机构,也有省市探索大数据相关立法,如贵阳市人

大常委会于2017年1月通过《贵阳市政府数据共享开放条例》。可见,开放政府数据工作正在全国范围内稳步进行。

然而,目前我国的开放政府数据工作仍处于初期阶段,存在一些问题:数据层面上,目前开放的数据数量和质量有限,未能充分发挥政府数据的价值;虽然各级政府创建了开放数据平台,但这些平台之间并不互通,“数据孤岛”现象依旧存在,并向“平台孤岛”转变。而法制层面上,法律制度保障缺乏,数据监管和开放推进方面均存有隐患。其实,这些开放政府数据的不足与我国政府部门大数据能力缺失密切相关。所谓政府部门大数据能力,涉及开放数据周期中的整合、监管、开放这些关键环节,要求政府能够共享整合各部门数据,在保证个人隐私、商业机密、国家安全的

* 本文系国家重点研发计划重点专项(2016YEF0122300)、国家自然科学基金项目(71371122)和国家社科基金重大项目(14ZDB152)的阶段性成果。

情况下,无障碍地将政府数据高质量地开放给社会各界进行利用。政府部门大数据能力是一个综合能力,涉及政府数据共享、监管、开放等环节,涵盖整个数据开放的周期。如若政府部门缺乏大数据能力,各部门的数据会依旧处于割裂状态,开放政府数据也因缺乏有效监管,数据的数量、质量、安全性、可供再利用程度等都难以得到保障,从而充分挖掘政府数据将成为一纸空谈。

因此,要想突破当前的困境,急需推进政府部门大数据能力建设。本文在阐述政府数据开放瓶颈的基础上,引入美国开放政府数据案例进行分析,提出需要提升政府的大数据能力。这样能够帮助落实部门间的数据共享,并对政府大数据进行较好的监管,实现隐私与安全的平衡。有助于走出当前政府数据开放的困境,推动高质量的政府数据开放工作,从而能够加快转变政府职能,促进社会经济结构转型,激发大众创业和万众创新的热情和动力。

开放政府数据的兴起与研究

将政府的数据开放给公众这一概念并不陌生,在美国其演变可以追溯到100多年前。无论以纸质版或是电子版的形式,美国政府一直在大范围搜集与公开政府数据,这种透明与公开反映出美国政府获取与传播政府信息的核心原则(Jaeger & Bertot & Shuler, 2010)。基于此原则,美国奥巴马政府在2009年于全球首先推行开放政府运动,此后,开放政府数据成为世界各国政府在大数据浪潮中的关注焦点。

政府数据通常包含多种数据集,如预算支出、人口、普查、地理等数据,同时也包括一些通过下属分支机构等间接得到的数据,如公共交通、医疗、教育等数据。开放政府数据的过程中,典型的推进方法就是收集整理这些数据集,通过数据脱敏等处理后,在开放政府数据门户网站上进行开放,如美国(data.gov)、新加坡(data.gov.sg)等(Attard et al, 2015)。虽然一些政府建立了数据门户网站,但并不意味着这些数据是真正透明、可获取的,如有的政府仅通过PDF来公开数据,这给再利用数据带来了困难(Arcelus, 2012)。可见,搭建开放平台仅起到促进政府与外界沟通的桥梁作用,开放政府数据更重要的在于政府数据

是否真正处于开放和可利用的状态。

在开放政府模型方面,奥巴马政府于2009年提出,开放政府应包含透明、参与、合作三个维度。^①一些学者基于此进行深层次的研究,提出改进模型。如有学者利用系统动力学模型来展现整个数据开放过程,从政府数据公开、政府数据适合再利用、开放政府信息语境化、对数据进行管理维护以减少冲突、外部开放者如何利用政府数据创造价值这五个方面对此模型进行了解读。^②从这些复杂的模型可以发现,开放政府数据工作是一个综合过程,包含很多关键环节,需要政府相应地具备全方位的大数据能力,才能够将开放政府数据工作落到实处。

聚焦国内,我国开放政府数据平台运行的情况、如何促进我国开放政府数据处于良好的生态等话题成为国内学者的研究热点(徐慧娜、郑磊, 2015; 丁念、夏义堃, 2015; 樊博、陈璐, 2017)。如郑磊、高峰(2015)通过创建基础、数据、平台三大层面的评估框架对我国多地的开放数据平台进行实证研究,发现地方政府开放的政府数据存在数量少、质量不高、更新频率低、可利用性不高等问题。能够看出,我国的开放政府数据工作刚刚起步,在整个开放数据周期中各方面的工作均存有不足,导致了大部分政府数据并未处于开放和可利用的状态。究其原因,是政府部门大数据能力的缺失,以致于散落于政府各个部门的数据不能进行有效的共享整合,也缺乏对政府数据的安全监管,开放政府数据工作面临很大的瓶颈。因此,探讨政府部门大数据能力的现状及建设路径极为必要。

政府部门大数据能力:构成与我国现状

政府部门大数据能力是一个综合能力,贯穿整个开放数据的生命周期。从政府后台产生、管理、提供数据的动态过程来看,涉及数据共享、数据监管、数据开放和数据再利用四个关键环节。

数据共享主要涉及政府部门内的数据活动,要想聚集整合各部门的数据,部门间需具备数据共享能力。只有在充分共享整合的基础上,才能够产生多数量、高质量的政府数据(李明, 2014)。在搜集、保存和维护政府大数据的过程中,政府部门应调动其数据监管能力,对已有数据集进行清

洗、脱敏等关键工作,为开放数据做好准备(Bertot,2013)。而在开放数据时,政府部门不仅需要拥有开放数据能力,也需要具备数据再利用能力,在提供大量高质量数据的同时,注重数据的时效性和可获取性等,确保数据“有用”(Dawes,2010),从而降低外界对于数据处理难度,提高数据的再利用水平。总体而言,我国政府部门大数据能力存在以下不足:

一是数据共享能力不足,“信息孤岛”现象犹存,并逐步转向“平台孤岛”。在我国电子政务发展的过程中,由于“部门本位”思想的存在,缺乏与政府其他部门实现信息系统的互联互通,信息共享意识不强,从而形成了大量的“信息孤岛”。更为严重的是,“平台孤岛”正在逐步形成。一些发达的省市,如上海,由于拥有较为雄厚的财政、技术、产业等资源支持,一些部门以搭建平台为名义立项,不断建设、充实本部门的信息平台。事实上,这些信息平台上的所有数据和相关系统仅供本部门使用,且部门较少共享信息给其他相关部门。于是,在“条强块弱”的局面下,这种平台建设不但难以消弭“信息孤岛”,反而极大地强化了已存在的“信息孤岛”,并且导致“信息孤岛”向着“平台孤岛”转变。

二是数据监管能力不足,政务大数据基础建设及数据安全问题亟待解决。从国际经验来看,数据开放工作需要信息化部门与专业的大数据团队统一建设政务大数据仓库,在此基础上,由首席数据官带领大数据团队依据法律分析现有数据集,在保障个人隐私、商业机密和国家安全的基础上,对整个政府的政务大数据进行分析、脱敏之后确定开放数据集的范围。然而,目前我国多数省市既缺乏统一数据仓库,也缺乏具有大数据分析能力的团队进行统一数据资源建设、监督和运营。而法律制度层面,对于政务数据资源开放的范围、隐私界限等无明确规定,数据资源的编目、数据审查、权限设定、数据公布、数据更新等环节也缺乏法律法规指导,较难实现数据隐私与安全的平衡。

三是数据开放能力不足,开放的数据数量和可获取性均有待提高。依据《中国地方政府数据开放平台报告》,上海市政府在开放数据工作中位列全国首位,代表着国内顶尖的开放政府数据水平。^③然而,相较于国外发达城市如纽约,仍有

一定的距离。首先,从开放的数据来看,未能做到非保密即公开,存在部门自由决定是否开放的模糊地带。其次,上海市政府不同领域内的开放水平差距较大。例如,截至2017年底,上海经济建设领域开放数据集有277个,而资源环境领域开放数据集仅有62个。此外,开放平台上的数据并非完全的开放,存在需要通过审核才能下载数据的情况,公众与企业的数据获取方面存有障碍。

四是数据再利用能力不足,开放数据的质量与格式难以满足社会需要。当前国内各省市开放平台上的数据多为静态数据,动态数据中也有未按原始设定标准进行更新的情况。这直接导致开放的政府数据难以及时、准确地反映现状,存在着时效性问题。此外,在数据格式方面,很多省市的开放数据多是RAR/XLS/XLSX这三类可读格式,甚至有的平台上呈现的信息为PDF格式,更有存在网页文本信息公示的情况。大多数平台提供的数据格式中,缺少可机读格式和开放API接口,这给数据使用者获取数据、分析数据造成极大的困难。

加强政府部门大数据能力: 美国推进开放政府数据的经验

美国是最早实施电子政务的国家,也是最早推进开放政府数据的国家。总体而言,其注重通过加强相关法律、体制建设,促进政府各部门数据资源的有效整合,使得监管、开放、再利用数据有法可依,政府部门大数据能力进而得以提升。具体而言,在法律方面,美国先后发布了《信息自由法》《开放政府指令》以及《数字政府战略》三项法规政策。其中,2009年发布的《开放政府指令》确立了透明、参与和协作的开放政府三原则。在此基础上,全球第一个国家层面的开放数据网站Data.gov应运而生。而2012年出台的《数字政府战略》则对开放数据提出了更加具体的要求,如要求政府数据总是处于“开放和机器可读”状态,从而方便公众可随时访问高质量的政府数据。聚焦美国地方层面,地方政府也在积极探索法规、体制建设,来提升部门大数据能力,保障开放政府数据各环节能够成功运转。以纽约为例,纽约市政府于2012年通过了《开放数据法案》(Open Data Law, Local Law 11 of 2012)^④,并陆续出台了一些更有针对性的法案,分别对数据存储、数据目录、

地理空间数据标准、公众需求、及时更新提出要求。^⑤

在体制方面,联邦政府对电子政务建设的配套机制工作也进行了积极探索,如在2012年正式设立了首席信息官(CIO)制度,随后联邦及各州又陆续设立的首席技术官(CTO)、首席数据官(CDO)和首席分析官(CAO)等来推动相关配套建设。例如,纽约市在承接联邦政府、州政府首席信息官基础上,在信息技术与通讯局设立本市首席信息官(Agency CIO),由该局局长兼任,依托信息技术与通讯局开展工作;同时在纽约市政府各部门设立部门首席信息官(Bureau CIO),由市首席信息官任命,实行垂直管理,并于2013年在市长办公室下设市长数据分析办公室,成立8人的专家团队支持其工作,办公室主任为首席分析官(Chief Analytics Officer, CAO)。职责上,首席分析官向纽约市长首席政策与战略规划顾问汇报工作,并与分管运行和分管经济发展的副市长,以及信息技术与通讯局协同工作。此外,分管运行的副市长还负责建立政府分析师委员会,并任主席;委员会成员由首席分析官、信息技术与通讯局局长以及运行办公室的其他被指派代表构成。政府分析师委员会主要负责纽约市整体层面的分析战略规划。

在完善立法、体制架构的基础上,纽约市首席信息官、首席分析官基于自身核心职能,在其他部门特别是管理和预算委员会、信息技术与通讯局、财务和采购部门等支持下,对纽约市政府的前台数据开放、中台数据共享以及后台数据分析等方面不断进行革新和优化。DataBridge是首席分析官带领纽约市长数据分析办公室与信息技术与通讯局协作,以纽约市刑事司法系统牵头建立的DataShare数据交换系统为媒介,利用纽约市311数据库备用容量来存储数据而建立起的统一数据系统。该系统正逐步由实体的数据仓库形态向着基于云的数据仓库形态转变。首席分析官会带领团队向各部门提供政府大数据相关培训支持,从而确保各部门均拥有向DataBridge传送和获取数据的技术,部门人员也能够顺利进行相关操作,从而综合提升部门的大数据监管能力。此外,依托DataBridge,首席信息官和市长数据分析办公室也会为其他政府部门提供咨询分析服务,帮助科学

决策。

在提升数据共享能力方面,除了在法规上明确要求各部门进行数据共享之外,首席分析官还牵头进行了DEEP项目的建设,构建起机构间蜘蛛网状的数据交换系统。通过该系统,纽约市各部门可以以一致的格式,定期甚至实时进行数据交换。当前,DEEP项目已帮助实现了30多个政府部门、外部供应商间200余项数据的自动交换。从而使政府部门的数据共享能力得到极大提升,政府数据的整合变得更加有效率。在数据开放和再利用方面,主要由首席分析官带领其团队,基于美国《开放数据法案》以及各部门提交的数据开放计划,对政府数据进行筛选和“脱敏”后决定开放数据的内容,从而保障个人隐私、商业机密和国家安全。此外,首席分析官及其团队也会定期对各部门开放数据工作进行审查,每年至少审查三个市政部门。具体而言,审查内容主要是核实部门数据开放情况是否与部门开放数据计划书一致,并就现实情况撰写相应报告,对外界进行公示。为了确保政府开放的数据能够满足各方面需求,并实现隐私与安全之间的平衡,首席分析官及其团队也会结合审查情况及各部门反馈,对开放数据的范围进行不断调整。为了加强公众企业与政府的沟通,首席分析官带领市长数据分析办公室通过扮演“联络人”的角色,主要负责收集与分析外界使用者在开放政府数据平台上反馈的相关问题;并会在每年的特定时间段发放问卷,以收集外界对于政府其他未开放数据集的需求提议,从而依据外界需求帮助确定数据开放的优先级。结合外界反馈与政府各部门提交的开放计划,数据分析团队不断分析与完善政府数据开放政策,有选择有步骤地扩大数据开放内容。同时,数据分析团队也与一些学术机构和高校进行合作,对政府数据开放中的不足进行改进。由此可见,纽约市政府具备着极强的数据开放和再利用能力。

美国经验于提升我国政府部门 大数据能力的启示

从美国经验来看,美国政府主要是从国家立法、管理体制、人员素质三个方面来提升政府部门大数据能力。由于国情与政府体制的不同,美国经验并不能完全照搬,需要结合我国国情,寻找提

升政府部门大数据能力的方法。

在立法方面,美国联邦及地方政府通过颁布相关政府信息安全法以及开放政府数据相关法案,对开放数据工作进行规范,从而保障信息安全;并可减少因各环节标准不一而导致的重复作业,提高政务效率;也能通过法律的规范性和强制性,极大地推动落实横向部门间数据共享工作及对外开放数据工作。我国长期积极推进法治建设,2017年我国施行了《中华人民共和国网络安全法》,贵州颁布了我国首个大数据地方性法规《贵州省大数据发展应用促进条例》,这些均能反映出国家对电子政务相关法律建设的重视。然而,纵观全局,我国政府部门信息共享仍存在推进困境,开放数据工作也存在数据量少与质量不高等瓶颈,这些问题显现出我国目前在数据共享、开放等诸多环节仍缺少相应规范,在相关法律建设方面仍存有不足。因此,美国通过法律的强制力来落实和规范开放数据工作的思路,值得我国借鉴。

在部门管理体制方面,美国在联邦政府及地方政府均设立了首席信息官制度,是政府信息化建设的抓手。首席信息官由于拥有部门信息资源处置权和业务流程决定权,能够通过监督与协调,成功推进各部门信息数据的共享与整合。此外,美国的一些地方政府也成立了专门的政府大数据团队,牵头建设开放数据的基础硬件设施,并推进数据监管与计划落实等工作,保证开放数据工作的正常运转。我国的电子政务管理体制与美国不同,现阶段,管理体制存在条块分割的困境,电子政务建设项目也存有多头治理的难题,无法做到统一规划、建设各部门的电子政务项目,导致了重复建设、信息系统不一等问题(樊博,2008)。因此,参考美国顶层设计的经验,我国政府相关信息化建设需要站在全局的角度,以需求为导向结合国情进行顶层设计。而部门管理体制上,急需一个权威的能够实现体制变革的部门,结合地方现状,整体推进、协调各部门的相关开放政府数据工作。

在人员素质方面,美国政府的首席信息官、大数据团队都具备极强的信息能力,并且会定期对各部门成员提供培训支持。从开放数据的生命周期来看,其数据收集、共享、整合、开放等环节,都会涉及一些基层公务员的操作,而开放数据中所

涉及的部分决策工作,也需高一级的公务人员在拥有特定知识背景的情况下,做出科学的分析与决策。我国政府也需要考虑相应提升部门人员的数据能力,掌握大数据时代所需的一些信息知识与技术,来迎接时代的挑战。

因此,本文认为提升我国政府部门大数据能力、推进开放政府数据工作需要重视以下几点:

一是尽快出台开放政府数据的相关法规,为数据开放提供法制保障,从而综合提升政府部门大数据能力。需要人大立法来完善相关数据共享的法律,来突破当前管理体制的困境;需要加强数据安全法制建设,如个人数据保护法、商业机密保护法、安全保密条例等,保证各系统和网络能安全可靠地运行;同时,也需要加强规范来确定开放数据的目录、开放标准、数据获取技术、开放申请流程与数据开放接口等,并对数据的时效性和可获取性提出规范,尽可能保证数据的可利用性。进而推动政府在法律框架内做到非保密即公开,在保障安全与隐私的情况下最大程度地开放政府数据。

二是逐步探索适合中国国情的信息化推进体制,加强政府部门数据共享、数据监管、数据开放能力。可以探索建设符合我国国情的首席信息官制度,或者成立大数据管理团队,对电子政务按需求进行顶层规划,牵头推动相关硬件设施如数据交换系统建设。并在数据共享方面,综合协调各部门的共享应用,推进数据整合;在开放数据清单方面,结合专家意见和各部门的计划,并不断修改完善开放数据清单;在数据开放前的数据分析与脱敏工作方面,提升数据监管流程的科学性和准确性,防止开放数据后隐私泄露和数据攻击等问题出现。

三是提升部门人员的相关数据能力,推动政府数字信息化进程,并逐步完善开放数据的绩效考核和监督制度。公务人员的信息素质不高会造成工作效率低下的问题,且目前我国政府很多信息仍以纸质文档而非电子数据形式存在,给开放数据工作造成很大难度。除了需要培训基层公务员的相关技能外,也需要制定数字化存储标准与流程,规范政府信息电子化备案,提升存储数据的质量,为开放数据工作奠定基础。另一方面,通过对于开放数据工作的内部绩效考核与外部监督,来引起部门重视相关工作。

此外,除了加强政府部门内部的建设外,还可推进数据开放的互动管理,提升企业、个人、科研院所及其他社会组织等多元主体的参与,汲取大众的智慧。可以通过数据开发竞赛、座谈会等形式鼓励社会主体积极参与数据的开发利用,并在此过程中不断反馈开放数据的有关问题,从而提升政府部门的数据开放和数据再利用能力。☞

- ①Transparency and Open Government, <https://www.whitehouse.gov/the-press-office/transparency-and-open-government>.
- ②The dynamics of opening government data Center for Technology in Government. <http://www.ctg.albany.edu/publications/reports/opendata>.
- ③2017 中国地方政府数据开放指数排名: 上海、青岛位列前二, 北京跌出前五, <http://www.thebigdata.cn/YeJieDongTai/35148.html>.
- ④Read the Open Data Law, <https://www1.nyc.gov/site/doitt/initiatives/open-data-law.page>.
- ⑤NYC Open Data Laws and Reports, <https://opendata.cityofnewyork.us/open-data-law/>.

参考文献

1. Arcelus, Josune “Framework for useful transparency websites for citizens”, *International Conference on Theory and Practice of Electronic Governance ACM*, (2012): 83-86.
2. Attard, Judie, et al. “A systematic review of open government data initiatives”, *Government Information Quarterly*, 32. 4 (2015): 399-418.

3. Bertot J. C., Choi H., “Big data and e-government: issues, policies, and recommendations”, *Information Polity*, (2013): 1-10.
4. Dawes, S. S., “Stewardship and usefulness: policy principles for information-based transparency”, *Government Information Quarterly*, (2010): 377-383
5. Jaeger, Paul T., J. C. Bertot, and J. A. Shuler, “The Federal Depository Library Program (FDLP), Academic Libraries, and Access to Government Information”, *Journal of Academic Librarianship*, 36. 6(2010): 469-478.
6. 樊博, 陈璐 《政府部门的大数据能力研究——基于组织层面的视角》, 《公共行政评论》2017 年第 1 期。
7. 樊博 《跨部门政府信息资源共享的推进体制、机制和方法》, 《上海交通大学学报》(哲学社会科学版) 2008 年第 2 期。
8. 才世杰, 夏义堃 《发达国家开放政府数据战略的比较分析》, 《电子政务》2015 年第 7 期。
9. 李明 《大数据技术与公共安全信息共享能力》, 《电子政务》2014 年第 6 期。
10. 徐慧娜, 郑磊 《面向用户利用的开放政府数据平台: 纽约与上海比较研究》, 《电子政务》2015 年第 7 期。
11. 郑磊, 高丰 《中国开放政府数据平台研究: 框架、现状与建议》, 《电子政务》2015 年第 7 期。

作者简介: 樊博, 教育部“长江学者奖励计划”首批青年学者, 上海交通大学国际与公共事务学院教授、博士生导师、副院长, 中国城市治理研究院应急研究中心主任。上海, 200240

(责任编辑: 王 婷)