

中国地方政府环境数据开放的现状、问题与对策： 基于国内部分省市开放数据平台的分析*

刘新萍**^① 肖鑫^① 黄奕奕^②

①上海理工大学管理学院 上海 200093

②复旦大学国际关系与公共事务学院 上海 200433

摘要：环境数据开放是指将政府环境治理过程中采集和存储的大量数据开放给社会，以扩大环境数据的增值利用水平，提升公民对环境相关议题的知晓程度和环保参与程度。研究对国内部分省市的开放数据平台上的环境数据资源进行了数据层和平台层两方面的评估，数据层主要包括数据整体开放情况、数据可用性、数据可访问性、数据再利用情况等指标，平台层主要包括用户体验、数据应用展示、数据可视化、互动交流等指标，从中发现当前环境数据开放现状与问题，并为未来环境数据开放提出对策与建议。

关键词：政府数据；环境数据；数据开放；数据开发

DOI：10.16582/j.cnki.dzzw.2017.09.004

一、引言

近年来，政府数据开放成为各级政府关注的焦点和前沿话题。数据开放将政府在治理过程中采集和存储的数据资源开放给社会，对促进数据的增值利用、推动产业升级和经济转型具有重要意义。而环境问题一直是百姓关心的热点问题，也是实现经济转型增长不可绕开的难题。在环境领域，公民获取的信息越多，参与政府环境管理事务的可能性就越大^[1]，因此，环境数据的开放对扩大公民对环境相关议题的知晓程度，促进和加强公民参与环境管理事务的能力，方便公众对政府环境监管提出批评建议，具有重要意义。^[2]

当前，国外政府在环境数据（或生态数据）开放中已经进行了一定探索，但成效甚微。有学者指出，生态环境类的数据由于其科学性和技术性，大部分高价值信

息大数据掌握在科研机构及生态学家手里，并没有贡献增值。^[3]因此，有学者鼓励生态学、环境学研究者积极分享数据，呼应“大科学”（Big Science）倡议，聚合学界力量，探讨生态环境问题的解决新道，甚至有学者提出生态领域有必要发展一种透明数据交换和聚合文化^[4]。

2015年8月，国务院印发的《促进大数据发展行动纲要》中明确提出要优先推动包括环境数据在内的民生保障服务相关领域的政府数据集向社会开放，这一部署也进一步推动了环境数据开放的进程。当前，地方政府已经开始了环境数据开放的探索，在已经开通的地方政府开放数据平台上，均列有资源环境类或生态文明类主题。然而，尽管各地政府的环境部门在数据开放方面已经做了不少实践和努力，但总体来看，中国的环境数据开放工作仍处于起步阶段。本研究选取了7个地方政

*基金项目：2014年度国家自然科学基金项目“大数据背景下开放政府数据的因素与机理研究：系统动力学建模与政策仿真”（编号：71473048）。

**通讯作者

收稿日期：2017-08-20

刘新萍 肖鑫 黄奕奕·中国地方政府环境数据开放现状、问题与对策:基于国内部分省市开放数据平台的分析

府,研究其在数据开放平台上的环境数据开放现状,分析当前存在的问题,并为政府未来的环境数据开放提出建议。

二、研究方法 & 评估维度设计

(一) 研究方法 & 样本选择

本研究选取国内目前数据开放相对较早且比较成熟的平台作为研究对象,评估其开放数据平台的环境数据主题下数据开放现状,分析总结当前中国环境数据开放平台中的整体现状与问题,并提出可行对策,为其他地方政府环境数据开放作出参考。为保证样本的代表性,本研究综合参考地域分布、经济发展水平和行政级别等指标,最终选取北京市、上海市、浙江省、武汉市、青岛市、贵阳市、佛山市南海区等七个省、市或区县政务数据开放相对较早且比较成熟的平台作为研究对象(参见表1)。

从表1可以看出,当前中国地方政府的环境数据开放将环境、资源、生态等数据列在同一主题目录中,但实际上环境、资源与生态的侧重点存在差异。党的十八

大报告把生态文明建设纳入中国特色社会主义建设“五位一体”的总体布局,开始努力走向社会主义生态文明新时代,全面推进资源节约、加大自然生态系统和环境保护力度成为生态文明建设的重点议题。因此,在数据资源的筛选中,为突出表现环境数据的特征,最终将资源数据从资源环境目录中剔除,筛选出标准环境数据作为分析对象。

(二) 评估维度界定

在对环境数据开放的评估维度构建中,本研究参考了郑磊等人的政府开放数据评估框架的研究成果,他们从“基础、数据、平台、使用、效果”等五个层面构建了中国政府数据开放评估的三级指标体系^[5];同时也参照了OECD的开放政府数据指数(OGD)^[6]评估框架,该框架从数据可用性(data availability)、可获取性(data accessibility)和数据再利用(re-use)三个维度进行评估。最终本研究数据层的评估包括:

——数据整体开放情况,主要包括数据总量、数据资源内容、数据提供机构、数据简介与描述等;

——数据可用性方面,主要评估数据分类、数据发

表1 评估样本选择

研究对象	行政级别	开放数据平台名称	环境类数据的数据领域名称	环境数据目录网址
北京市	省级	北京市政务数据资源网	环境与资源保护	http://www.bjdata.gov.cn/zyml/azt/hjzyzbh/index.htm
上海市	省级	上海市政务数据服务网	资源环境	http://www.datashanghai.gov.cn/query!queryDataByField.action?dataField=2
浙江省	省级	浙江政务服务网	资源环境	http://data.zjzwfw.gov.cn/toCate?catecode=SJLY003
武汉市	市级	武汉市政务公开数据服务网	能源环境	http://www.wuhandata.gov.cn/whdata/resources_listPage.action
青岛市	市级	青岛市政务数据开放网	资源环境	http://data.qingdao.gov.cn/data/gov/index.htm
贵阳市	市级	贵阳政务数据开放平台	生态文明	http://www.datagy.cn/topics.jsp?topics_id=5
佛山市南海区	区县级	数说南海	环境与资源保护	http://data.nanhai.gov.cn/cms/sites/sjzy/folder.jsp

刘新萍 肖鑫 黄奕奕·中国地方政府环境数据开放现状、问题与对策：基于国内部分省市开放数据平台的分析

布与更新时间、数据承诺更新频率、数据格式与可机读情况等；

——数据开放授权方面，主要考察数据是否可免费获取、不受歧视、自由利用、自由传播和分享等；

——数据利用情况，主要考察数据访问与下载量等内容。

在平台层，主要考察互动交流、数据应用展示、数据可视化等内容。

本研究的主要分析维度如图1所示。

（三）数据采集与分析

关于评估框架中数据层评估涉及到的数据，本研究以2017年2月13日为时间节点，下载了各地方政府开放数据平台上的资源环境或生态文明主题下的数据，并对环境数据进行了筛选，基于前述评估框架对涉及的相关数据资源进行评估分析。同时，本研究也进一步对各平台的授权协议、使用声明、免责声明等进行文本分析，研究各平台对开放数据的授权、收费、使用者权利义务

等方面的具体规定。此外，本研究还对各平台是否可以评论或打分、是否有可供下载的客户端应用等功能进行了观察与统计分析，研究各平台的用户体验以及与数据使用主体的互动交流情况。

三、研究发现：政府环境数据开放现状分析

基于前述评估框架，本研究对环境数据开放的评估将从数据层与平台层两个维度分别展开。

（一）数据层

1.数据整体开放情况

(1)数据总量

截至2017年2月13日，研究样本所覆盖的各地开放政府环境数据集共279个，其中环境数据开放最多的是武汉（85个），其次是上海（50个），详见表2。从表2中可以看出，当前在各地政府数据开放网站上资源环境主题下的数据中，以环境类数据为主，占75.8%；资源类数据相对较少。

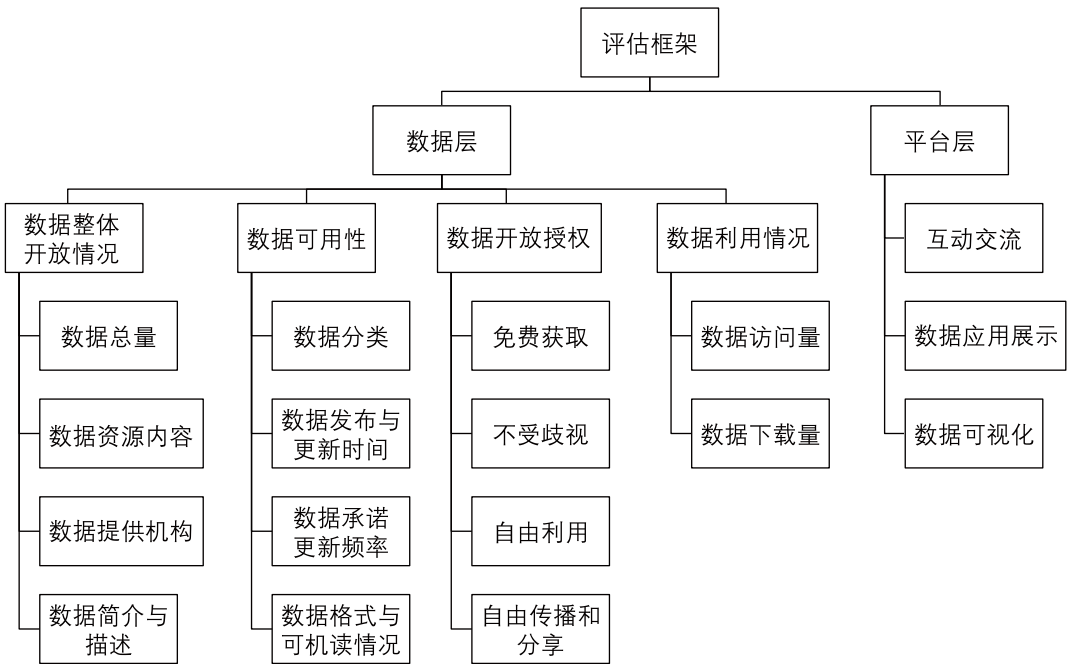


图1 中国地方政府环境数据开放的总体评估框架

刘新萍 肖鑫 黄奕奕·中国地方政府环境数据开放现状、问题与对策：基于国内部分省市开放数据平台的分析

表2 环境数据开放总量

地方	数据总数	环境类	资源类	环境类数据占比
北京市	25	24	1	96.0%
上海市	77	50	27	64.9%
浙江省	30	26	4	86.7%
武汉市	119	85	34	71.4%
青岛市	37	33	4	89.2%
贵阳市	62	43	19	69.4%
佛山市南海区	18	18	0	100.0%
合计	368	279	89	75.8%

(2)数据资源内容

数据资源内容主要考察当前环境数据主要开放了哪些数据资源，本研究对各平台提供的数据资源的关键词进行了词云分析，其中青岛市数据资源开放时未提供关键词，故本分析只显示其余6个地区的环境数据开放情况（参见图2）。

从图2中可以看出，当前环境数据资源主要开放了信息、环境、单位、许可、监测、企业、许可证、质量、重点、保护、空气、危险、环保、污染等数据内容，从关键词可以看出，当前所开放的环境数据资源多为许可类、监测类信息，而不是原始数据，这与数据开

放应尽量开放原始数据的原则相违背。

(3)数据提供机构

数据提供机构主要考察当前主要有哪些部门参与环境数据的采集、整理、更新与开放工作，以确认环境数据开放的责任主体。在开放的数据资源中清晰地标明责任主体有利于在用户发现数据存在数据质量、数据更新、数据权属纠纷等问题，或者用户需要更多相关数据资源时，可以更容易找到相应权责部门。鉴于各地环境部门的名称与职能存在差异，本研究将其进行了归并，其中，出现频次较低的少数部门列入“其他”（参见表3）。



图2 数据资源关键词词云分析

刘新萍 肖鑫 黄奕奕 · 中国地方政府环境数据开放现状、问题与对策：基于国内部分省市开放数据平台的分析

表3 数据提供部门汇总

部门类别	包含的部门	数据资源数量
环境保护部门	北京市环境保护局、上海市环境保护局、浙江省环境保护厅、武汉市环境保护局、青岛市环保局、贵阳市生态委、南海区环境保护局	150
水务、水利管理部门	上海市水务局、浙江省水利局、武汉市水务局、青岛市水利局、贵阳市水务局	57
园林绿化部门	北京市园林绿化局、浙江省林业厅、武汉市园林和林业局	14
气象局	浙江省气象局、贵阳市气象局、南海区气象局	8
食品药品监督管理局	浙江省食品药品监督管理局	8
安全生产监督管理局	北京市安全生产监督管理局、南海区安全生产监督管理局	4
海洋、渔业管理部门	浙江省海事局、青岛市海洋与渔业局	3
国土资源管理部门	上海市规划和国土资源管理局、贵阳市国土局	3
文物局	北京市文物局	3
其他	上海市工商管理局, 贵阳市工信委, 上海市公安局、南海区环境运输和城市管理局, 武汉市武昌区、蔡甸区、东湖生态旅游风景区管委会、硚口区、江夏区、城市管理委员会、洪山区, 武汉经济技术开发区管委会, 武汉市	24
合计	274	

注：环境类数据资源一共279条，其中上海有4条数据应用、武汉有1条数据资源未标明提供机构，故仅统计了274条数据资源。需说明的是，环境数据资源的提供机构以各部门主动提供的信息为准，本研究不对未主动提供机构信息的数据资源另做专门判断。

由此可见，环境数据的提供主体涉及多个部门，在环境数据开放的责任主体中除环境保护部门外，水务水利部门、园林绿化部门、气象局等也发挥着非常关键的作用。

(4)数据简介与描述

数据简介是对数据资源内容、数据项等的简要描述。除青岛市以外，其他6省市均对其开放的环境数据资源提供了数据简介，但各地的描述内容存在差异。北京市、上海市、浙江省、贵阳市的多数数据资源包括对数据资源内容的简要介绍和主要的数据项，少数数据资源仅介绍了数据资源简介或仅有主要数据项；武汉市的数据资源简介包括提供机构和数据资源的简要介绍，其中有少数数据资源也介绍了内含的主要数据项；佛山市南海区的数据资源简介仅包括提供机构和数据项。

2.数据可用性

数据可用性主要与数据是否清晰分类、及时发布、及时更新，数据格式是否符合可机读要求等组成，以保证数据质量，确保数据完整、准确。只有完整准确的数据才能够被数据利用主体进行开发使用。

(1)数据分类

清晰的数据分类有助于数据利用主体准确、快速地找到对应的数据资源，同时也便于数据利用主体综合不同主题、不同地区的数据资源，开发跨地区的数据应用产品。就数据分类而言，上海市和贵阳市分别对其开放的环境数据资源进行了详细分类。上海市的政务数据资源区分了国家主题分类和部门主题分类，其中国家主题分类主要包括环境保护、治理，农业、水利，气象、水文、测绘、地震，综合类，城乡建设、环境保护，公安，水利等类别。部门主题分类包括业务类、规划计

刘新萍 肖鑫 黄奕奕·中国地方政府环境数据开放的现状、问题与对策：基于国内部分省市开放数据平台的分析

划、机构职能等。贵阳市的政务数据分类除了包括国家主题分类（和上海市相同的分类标准）外，还包括了行业分类和服务分类。其中，行业分类包括水利、环境和公共设施管理业，居民服务、修理和其他服务业，电力、热力、燃气及水生产和供应业，文化、体育和娱乐业，科学研究和技术服务业，公共管理、社会保障和社会组织等；在服务分类方面包括惠民服务、服务交付的支撑、政府资源管理等。其余地区均未对数据分类做详细规定，但部分地区对数据资源稍作了区分，如北京市的环境数据资源分为自然资源环境、环境污染与防治、特定区域环境保护三类；佛山市南海区将其环境数据资源分为环境污染与防治、环境质量两类。

同时，需要特别说明的是，青岛市所开放的环境数据资源还增加了相关数据的引导以及标签功能，用户可以通过标签链接到同一类别或者相关的数据，由此提高了对数据资源的利用水平，实现环境数据资源与其他数据资源的交叉利用。

(2)数据发布与更新时间

数据发布与更新时间是描述数据资源的重要指标，有利于数据利用主体对数据资源有更清晰的了解。在各地的开放数据平台上，关于数据发布和更新时间的表述存在差异，本研究用原始数据创建时间（即该数据的产生时间）、数据首次发布时间（即在各自平台上传的时

间）和数据最新更新时间来进行区分。研究发现，七个省市区均标明了数据资源的首次发布日期；佛山市南海区和武汉市的部分数据资源也标明了原始数据发布时间；上海市、青岛市、武汉市的部分数据资源还标明了数据的最新更新时间（参见表4）。这些描述有利于数据利用主体评估该数据资源是否为所需数据资源，亦可根据数据是否具有时效性来甄选数据。

(3)数据承诺更新频率

数据承诺更新频率是衡量数据时效性的重要标准。在研究对象中，除北京、武汉、佛山市南海区未对数据更新频率做明确说明外，其余各平台的环境数据存在不定期、按需要（或依申请）、每年、每半年、每季度、每月、每周、每日、实时等更新频率。通过对数据资源的统计分析发现，仅有38.2%的数据可以每季度、每月、每周、每日或者实时更新，其余61.8%的数据更新频率在每半年及以上更新或按需要更新（参见表5）。由此可见，当前在开放数据平台上已经开放的数据多为静态数据，数据更新不及时，严重限制了数据利用主体对环境数据资源的开发利用。

(4)数据格式与可机读情况

为方便数据利用主体对已经开放的环境数据资源进行开发利用，各平台上的环境数据资源应当以可机读的格式进行开放。当前各平台上已经开放的环境数据资

表4 关于数据资源创建、发布与最新更新时间的描述

地方	原始数据发布时间	数据首次发布日期	数据最新更新时间
北京市		√	
上海市		√	√
浙江省		√	
武汉市	√	√	√
青岛市		√	√
贵阳市		√	
佛山市南海区	√	√	

注：武汉市的环境数据资源，部分数据含有首次发布日期和原始数据发布时间；部分数据有首次发布日期和数据最新更新时间；部分数据仅有首次发布日期。

刘新萍 肖鑫 黄奕奕·中国地方政府环境数据开放现状、问题与对策：基于国内部分省市开放数据平台的分析

表5 数据更新频率分布

	未明确规定	不定期	按需要 (或依申请)	每年	每半年	每季度	每月	每周	每日	实时
上海市	4	0	4	17	0	2	4	0	4	15
浙江省	12	8	0	1	0	1	0	1	3	0
青岛市	0	0	0	14	2	2	3	0	0	12
贵阳市	0	0	0	32	0	1	0	0	0	10
合计	16	8	4	64	2	6	7	1	7	37
94 (61.8%)						58 (38.2%)				

源的数据格式包括CSV、XML、JSON、XLS、API、在线链接、网页、空间数据、TXT、RAR、ZIP、PDF、JPG、DOC等格式。由表6可以看出，从各地提供的数据资源的格式上来看，多数地方对同一数据资源提供多种数据格式，以方便用户选择使用，仅有浙江省、青岛市的数据资源格式相对单一；同时，从整体上来看，XLS、CSV、XML、JSON等可机读格式的数据资源比例较高，并且已经有相当比例的数据提供了API接口。然而，诸如PDF、JPG、DOC等不可机读格式的数据也占有一定比例。可机读格式是数据利用主体进行数据开发利用的重要保障，各数据平台应提高可机读格式数据的比例，提高数据的利用程度。

3. 数据开放授权情况

数据开放授权情况是考察数据利用主体能否对各数据资源免费获取、不受歧视使用、自由利用、自由传播和分享。在各数据开放网站平台上，通常以用户使用协

议、使用条款、免责声明等形式对上述进行说明，包括北京市政务数据资源网、上海市政务数据服务网的《使用条款》、浙江政务服务网的《网站声明》、武汉市政府公开数据服务网、青岛政务网的《免责声明》和《青岛市政府数据开放网站数据下载使用许可》、《贵阳大数据开放服务平台用户注册协议》以及《南海区政务数据资源使用协议》等，这些条款对规范当前环境数据的开发利用具有重要的指导意义。

(1) 免费获取

免费获取是指用户是否可免费获取和利用已开放的数据资源。通过对各地的开放数据授权协议的研究发现，北京市、武汉市、贵阳市指出“现阶段”用户有权免费获取数据资源；而佛山市南海区指出，现阶段“认证用户”可以免费获取本网站所公开的政务数据资源；上海市、青岛市指出目前提供的服务均为免费，但保留收费预览和收费下载的权利。

表6 各地数据资源格式分布情况

	资源 总数	CSV	XML	JSON	XLS	API	链接或 网页	空间 数据	TXT	RAR	ZIP	PDF	JPG	DOC
北京市	24	10	0	0	14	24	0	6	0	0	5	0	0	0
上海市	50	0	1	0	24	14	6	0	0	2	0	1	0	1
浙江省	26	0	0	0	0	0	26	0	0	0	0	0	0	0
武汉市	85	0	0	0	55	0	83	11	0	0	0	8	1	0
青岛市	33	0	0	0	33	0	0	0	0	0	0	0	0	0
贵阳市	43	6	6	6	43	0	0	0	0	0	0	0	0	0
佛山市南海区	18	15	15	15	15	0	2	0	15	0	0	0	0	0
总数	279	31	22	21	184	38	115	17	15	2	5	9	0	1

刘新萍 肖鑫 黄奕奕·中国地方政府环境数据开放的现状、问题与对策:基于国内部分省市开放数据平台的分析

(2)非歧视性

非歧视性是指在开放数据平台上所发布的数据是否能够实现人人均可无条件地平等访问、获取、使用、分享。在各地开放数据的开放授权协议中,已经有部分城市提到数据使用的非歧视性,如北京市、武汉市、贵阳市、佛山市南海区均指出用户“享有数据资源的非排他使用权”。在实践中,上海市的大部分环境类数据集为普遍公开,有3项为特定公开,有2项为依申请公开,需要用户申请且满足条件才予以下载使用。由此可见,各平台基本保障了数据资源使用的非歧视性,赋予数据利用主体相对均等的获取和利用权限。

(3)自由利用

在自由利用方面,各平台均缺乏对用户的自由利用权限的保障。尽管部分地方没有限制数据利用情况,但是在界定用户权利时,均要求用户在使用其网站的政务数据资源产生增值利用中,应将应用情况进行备案,如北京市、武汉市、贵阳市;或清楚地说明数据来源及数据下载日期,如上海市、佛山市南海区。青岛市则明确用户可以基于开放数据网站提供的数据“进行商业或非商业性的开发活动”;但也有平台甚至限制对数据的商业利用,如浙江提到,“任何媒体、互联网站和商业机构只有在获得浙江政务服务网或其他相关权利人的授权之后才能使用,不得用于任何商业用途或创造与内容相关的派生产品”,这一规定限制了数据的增值利用水平,与开放数据的初衷不符。

(4)自由传播和分享

自由传播和分享有利于提升数据的用户知晓情况,让更多用户了解并使用数据资源,实现不同行业和技术交叉开发,产生更多的数据增值利用产品。北京市、武汉市、贵阳市、佛山市南海区指出,用户不得有偿或无偿转让在本网站中获取的各种数据资源。青岛市明确

指出,用户可以在法律、法规允许的范围内对开放数据网站提供的数据下载、复制与传播,但不得有偿转让从本网站获取的数据。

4.数据利用情况

鉴于对数据利用情况很难进行有效跟踪和评估,本研究主要以数据访问量和下载量两个指标来评估当前对已经开放的环境数据资源的利用情况。

数据访问量(或浏览量)可以展现用户对环境数据资源的关注程度,上海市、武汉市、青岛市、贵阳市等地提供了数据访问量指标。为衡量各城市环境数据资源的整体关注程度,本研究引入平均访问量,即环境数据资源总访问量除以该地区的数据资源条数。

数据下载量是反映数据是否被使用的重要指标,可引导用户根据下载量多少对数据资源进行筛选利用。研究发现,北京市、上海市、武汉市、青岛市、贵阳市、佛山市南海区等六个地方政府提供了数据资源的下载量指标。本研究用平均下载量来衡量数据资源的下载情况,即环境数据资源总下载量除以该省市数据资源条数。

由表7可以看出,当前数据资源的平均访问量参差不齐,这种现象一方面与数据资源平台的上线时间及数据资源的发布时间有关,同时也间接反映了当前各平台所开放的数据资源的质量与水平存在较大差异。在下载

表7 环境数据资源的平均访问量和下载量

城市	数据资源总数	平均访问量	平均下载量
北京市	24	---	89.3
上海市	50	1543.6	240.3
浙江省	26	---	---
武汉市	85	75.2	2.93
青岛市	33	108	1.1
贵阳市	43	27.3	12.1
佛山市南海区	18	---	92.1

量方面,当前开放数据资源的整体下载量较低,可见数据利用主体对数据的下载和使用情况较差。

(二) 平台层

平台功能的完善有利于提升数据利用主体运用开放数据平台获取环境数据资源。在对样本省市区平台层的评估中,主要涉及互动交流、数据应用展示、数据可视化等方面。

1. 互动交流

互动交流是指政府部门与数据利用主体之间关于数据需求、开放数据建议咨询与意见反馈、数据应用提交等方面的互动。通过对各平台分析可以发现,上海市平台可以在线提交或发送电子邮件来提出意见建议;上海市、青岛市的数据资源平台提供五星评价功能;北京市、青岛市、佛山市南海区提供评论功能,并可查看相关评论;青岛市可以将数据资源分享至QQ空间、微博、人人网、微信等平台,同时具有订阅功能,可以订阅相关数据;武汉市还有满意度评价,在每项等级后面对应描述评判标准。在数据请求方面,当前各平台均未公开用户的数据请求信息。

2. 数据应用展示

开放数据平台除提供原始数据下载之外,还是数据应用展示的重要平台,如上海提供了“上海防汛”(Android版)、“阿拉自来水”(关于上海水质监控的应用,iOS版)和上海空气质量(Android版和iOS版)等移动应用。其余省市区均没有在环境类主题下展示相关数据应用。

3. 数据可视化

数据可视化可以方便用户在下载数据之前对环境数据资源的内容有更加清晰的了解,方便用户下载并使用数据。在当前各平台所提供的环境数据资源中,上海市、青岛市和佛山市南海区的部分数据资源提供了在线

预览功能;北京市的环境数据资源除提供数据在线预览之外,还提供折线图、柱状图等多种格式预览,同时可以在线保存相应图片,以满足用户基本的数据分析需求。

四、中国环境数据开放的问题分析

通过以上分析可以发现,尽管各地政府均已开始了环境数据开放的探索,并开放了一定数量的环境数据,但环境数据开放工作仍存在不少问题。

第一,从总体上来看,已开放的环境数据资源数量少,各地均仅有十几条到几十条,这与庞大的环境数据总量形成明显反差。同时,在主题分类上,当前环境数据多数与资源或能源数据、生态文明混为一谈,限制了数据利用主体对环境数据的搜索与利用。在数据分类上,多数地方政府尚未建立数据资源间的链接,不便于用户查询并使用相关数据资源。

第二,从已经开放的环境数据资源内容来看,当前主要开放了信息、环境、单位、许可、监测、企业、许可证、质量、重点、保护、空气、危险、环保、污染等数据内容,这些信息多数为加工归总后的统计数据,不符合开放数据应当开放完整的、原始的数据集的基本原则。

第三,当前开放的环境数据资源缺乏数据规范,导致数据质量参差不齐、数据分类模糊、更新频率低、数据格式混杂。当前,各地政府均基于工作的方便开放了一定数据内容的数据集,但是由于缺乏统一的标准,导致各地政府对环境数据资源开放中的表现差异很大,也限制了数据利用主体进行跨地域的数据应用产品的开发。从数据更新频率上来看,当前已经发布的数据多为静态数据,更新速度慢,导致数据的可用性差;而从数据格式上来看,尽管当前已经开放的数据多为可机读格

刘新萍 肖鑫 黄奕奕·中国地方政府环境数据开放的现状、问题与对策：基于国内部分省市开放数据平台的分析

式,但是开放格式的数据占比仍然较低,且各地选用的可机读格式也差别很大。

第四,从数据授权协议上来看,当前的授权协议很难对数据利用主体提供有效的保障。尽管各地通过开放授权协议对数据利用主体的权限进行了界定,但是对数据免费获取、自由利用、自由传播或分享的相关规定的用词模糊不清,难以保障数据利用主体的合法权益。

第五,从数据利用情况来看,由于当前已经开放的环境数据资源数据量少、质量整体堪忧,因此当前数据利用主体浏览、下载环境数据资源的整体情况较差,更无从谈起基于开放的环境数据资源进行环境数据应用产品的开发或其他增值利用。

第六,从平台功能上来看,当前的开放数据平台功能相对简单,且建设中缺乏统一标准,在平台建设中也未考虑到环境数据资源的特殊性,比如环境治理的区域性特征,跨区域的环境数据资源难以在当前各地市的开放数据平台上进行开放,导致环境数据资源受现有平台功能的限制而阻碍其数据开放工作。

五、中国环境数据开放的对策分析

针对以上问题,环境数据开放工作除依赖于宏观层面开放数据相关法规政策、标准规范及管理制度的完善之外,还需要各环境数据相关部门转变观念意识、提升环境数据资源质量、规范开放授权协议和平台建设,同时也需要社会各界更多地关注、利用环境数据资源,并督促相关政府部门开放更多的环境数据资源。

第一,应从总体上改变环境部门的观念意识和内部技术能力。在环境数据工作中,应建立全局性的生态环保大数据发展观,以“开放、共享”精神推进生态环境的大数据发展^[7]。各环境数据相关部门应提高环境数据开放的意识,将更多的环境数据资源开放出来,供社会

增值利用。同时,各环境数据相关部门应提高内部技术水平,甄别潜在的数据开放的风险,对敏感或涉密的环境数据资源进行加工处理后进行开放。

第二,在数据内容方面,应尽可能开放更多原始的数据,而非加工统计类数据。在数据主题分类中,应遵照党的十八大报告关于生态文明建设中的有关论述,将环境、资源、生态类数据加以区分,更好地甄别环境数据,便于数据利用主体对环境数据的有效运用。

第三,应建立开放数据的规范,从数据分类、数据更新、数据格式等方面进行明确规定,提升环境数据的质量。地方政府环境数据相关部门应在国家环保部门的指导下,加强对本部门的环境数据资源的梳理、清洗与加工,促进更多高价值的环境数据资源对外开放。

第四,在开放授权协议方面,鉴于环境数据的敏感程度、涉密情况与其他类别的数据资源存在差异,相关政府部门应根据数据资源的不同类别细化开放授权协议条款,以更好地推进环境数据资源被免费获取,并自由利用、自由传播或分享。

第五,从社会对环境数据的利用来看,社会各界应加强对环境数据资源的关注程度和利用水平。一方面,利用当前的环境数据资源开发更多的环境数据产品;另一方面,允许民众表达需求是推进数据开放的最好方式^[8],数据利用主体应主动向政府环境部门请求更多环境数据资源,让政府部门了解社会需求,开放更多数据。政府层面也应该有目的地培育环境数据利用的生态圈,充分利用各类民间社会组织进行环境数据的增值开发与利用。

第六,从平台建设上来看,一方面要规范平台建设标准,将各地平台进行统一规范并建设;另一方面,环境问题需跨部门、跨地域的合作,建议建立专门的环境数据开放平台,与各地开放数据平台上的环境数据互联

互通,以提高环境数据的整体数量和质量,提升对环境数据的利用效率和水平,切实让环境数据开放服务于环境治理及环境质量的提升。

六、结论

环境数据开放有利于提升政府的环境治理能力,基于开放的环境数据,可以实现政府、企业、环保组织、社会公众、专家学者与媒体等不同行为主体的跨界合作,提升民间对环境事务的认知程度和参与能力,更好地预测环境污染演变趋势,为制定基于数据驱动的、科学的、精细化的环境治理政策具有重要意义。当前,环境数据开放乃至整个政府数据开放工作尚处于探索与起步阶段,应着手从数据开放的法规政策、标准规范与管理等入手,制定环境数据开放的针对性战略与规范,提升社会各界对环境数据的关注程度和利用水平,推动环境数据开放工作的进展。此外,环境数据开放应以环境保护、环境质量提升,甚至整个社会经济的全面发展作为终极目标,实现环境保护、社会治理与经济增长的融合发展。

参考文献:

- [1]其春. 大数据驱动政府环境监管转型及建设途径探讨[J]. 太原理工大学学报:社会科学版, 2015(4): 55-59.
- [2]迪莉娅. 大数据环境下政府环境数据开放研究[J]. 绿叶, 2013(9): 22-23.
- [3]Hampton S E, Strasser C A, Tewksbury J, et al. Big Data and the Future of Ecology[J]. Frontiers in Ecology & the Environment, 2013, 11(3): 156-162.
- [4]Jones M B, Schildharer M P, Reichman O J, et al. The New Bioinformatics: Integrating Ecological Data from the Gene to the Biosphere[J]. Annual Review of Ecology Evolution & Systematics, 2006, 37: 519-544.

- [4]郑磊, 关文雯. 开放政府数据评估框架、指标与方法研究[J]. 图书情报工作, 2016, 60(18): 43-55.
- [5]Ubaldi B. Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives[C/OL]. OECD Working Papers on Public Governance, No. 22, 2013, OECD, Paris. [2017-08-20]. <http://dx.doi.org/10.1787/5k46bj4f03s7-en>.
- [6]Parr C S. Open Sourcing Ecological Data[J]. Bio Science, 2007(4): 309-310.
- [7]程春明, 李蔚, 宋旭. 生态环境大数据建设的思考[J]. 中国环境管理, 2015(6): 12-13.

作者简介:

刘新萍(1986—),女,博士,上海理工大学管理学院讲师,复旦大学数字与移动治理实验室研究员,研究方向为电子治理、跨部门协同、开放政府数据。

肖鑫(1993—),男,上海理工大学管理学院硕士研究生。

黄奕奕(1993—),女,复旦大学国际关系与公共事务学院本科生。