

政府数据开放平台建设效率评估及其启示^{*}

韩磊^{1,2} 胡广伟^{1,2}

(1. 南京大学信息管理学院, 南京 210023; 2. 南京大学政务数据资源研究所, 南京 210023)

摘要: 2012年上海政府数据服务网上线以来, 截至目前我国已建成40多个政府数据开放平台。对政府数据开放平台的建设效率展开评估, 能够揭示政府数据开放的投入产出转化效果, 促进政府数据开放平台的良性发展。研究以我国39个典型的政府数据开放平台为例, 从投入产出视角出发, 基于数据和平台两个维度构建评估指标体系, 运用数据包络分析 (Data Envelopment Analysis, DEA) 方法评估平台的综合效率、纯技术效率和规模效率。研究发现, 我国政府数据开放平台的整体建设效率较高, DEA综合有效平台占比59%, 对非DEA有效平台来说, 其在用户数据利用方面的产出亟待提升。

关键词: 政府数据开放; 效率评估; 数据包络分析(DEA); 投入产出分析

中图分类号: G203

DOI: 10.3772/j.issn.1673-2286.2018.09.008

大数据是国家基础性战略资源, 被喻为21世纪的“石油”。中央到地方的各级政府部门在履行其行政职能、参与社会公共治理的过程中产生和积淀了大量的数据资源^[1]。2009年奥巴马政府签署《开放透明政府备忘录》、建立全球首个国家数据开放门户网站Data.gov以来^[2], 全球多个国家和地区都相继展开政府数据开放的探索和实践。我国在“十三五”规划中也明确指出, 要实施国家大数据战略, 推进数据资源开放共享。2012年上海市建立我国首个政府数据开放门户网站至今, 已有40多个省市逐渐建立起相应的政府数据开放平台^[3], 且国家层面的数据开放平台预计于2018年底建成^[4]。

政府数据开放的快速发展也带来一些不可避免的问题, 如开放水平仍然较低, 开放数据利用率低下等^[5]。鉴于此, 国内外都展开了一系列针对政府数据开放的评估研究和实践^[6-8], 以期评估政府数据开放平台的建设现状和存在的问题, 进而提出相应的建议来促进其良性发展。目前国内外的学者多从建设成效的视角对政府数据开放进行评估, 且国外的研究(如开放数据晴雨表等)多关注国家和地区层面, 城市层面的评估较少^[9]; 国内针对政府数据开放评估的研究实践主要针对省市层

面, 具有不同行政级别的省市级政府一般在一个统一的框架下进行比较^[10], 并没有作详细区分。但是处于不同层级的地方政府其数据开放体量往往不同, 如省级政府所开放的数据一般包括所辖地市级政府开放的数据, 故而前者的数据体量大于后者, 在数据开放的投入方面也往往要大于后者, 因此单从建设成效的角度进行评估并不能准确地体现其发展效率的高低。

政府对数据开放的投入和产出体现在多个方面, 且人力、资金等投入往往难以直接量化, 而产出方面对于开放数据的利用及其所产生的社会价值和经济价值更是无法直接衡量。因此学界对于政府数据开放的投入和产出的研究一般从最直观的数据开放平台及其开放的数据入手^[11]。本研究在借鉴前人研究基础上, 从平台和数据两个维度构建政府数据开放的投入和产出评估指标体系。在涉及具有多投入和多产出的同类型决策单元(政府数据开放平台)效率评估问题中, 数据包络分析(Data Envelopment Analysis, DEA)方法能够避免主观因素, 且能简化运算, 具有成熟的理论基础和广泛的应用^[12]。因此, 本研究以我国39个典型的政府数据开放平台为例, 运用DEA方法对其建设效率进行实证

^{*}本研究得到国家自然科学基金面上项目“电子政务服务价值共创机制及实现模式实证研究”(编号: 71573117)和江苏省“六大人才高峰”项目“政务大数据资源开发技术与实现方法研究”(编号: 2015-XXRJ-001)资助。

分析,以期丰富政府数据开放平台建设效率评估方面的研究实践,并对非DEA有效的平台提出改进意见和建议。

1 研究现状

1.1 政府数据开放“数据”维度评估

数据是政府数据开放评估的核心,当前学者针对数据维度的评估研究多集中于数据质量、数据类别、关键数据集等方面^[13]。Vetrò等^[14]构建了开放数据的质量评估框架,从数据集和数据项两个层面对数据的完整性、可理解性等7个维度进行评估;Veljković等^[15]从基础数据集、开放度、透明度等5个视角切入,以美国政府数据开放门户网站(Data.gov)的数据集为例进行评估分析。我国学者赵继娣等^[16]从类目完备性、更新及时性、可读性、可访问性4个角度对上海市政府数据服务平台的数据内容进行评估;谭必勇等^[17]从内容层面(可用性、全面性、易用性、及时性)和结构层面(安全性、开放性、可读性)出发,对我国10个有代表性的省市开放数据平台的数据质量进行研究。

综合来看,学者针对数据维度的评估研究多注重数据格式标准、数据质量等“质”的评估,对数据集数量等“量”的评估却较少重视^[13]。本研究在构建数据维度的政府数据开放效率评估指标体系时,将二者结合起来,在考察数据完整性等数据质量的基础上,引入数据集数量这一指标来综合衡量政府在数据层面的投入大小。

1.2 政府数据开放“平台”维度评估

平台作为政府数据开放的载体,是连接政府与数据用户的桥梁,对于提高用户数据获取体验和利用数据的有效性至关重要。由于当前政府数据开放平台大多以网站为载体,因此学术界对开放数据平台的评估多结合网站评估进行。如马海群等^[18]从网站可达性、网页可用性、服务舒适性等角度评估政府数据开放网站服务质量,并基于结构方程模型开展实证研究;邹纯龙等^[19]从网站规模、网站性能、网站体验、网站影响等5个维度出发构建评价指标体系,以美国20个政府开放数据网站为例,基于BP神经网络算法研究总体评价效果。

当前学者针对政府数据开放平台维度的评估已有不少探索和尝试,且研究方法趋于多样化,除专家打分法外也引入了利用Web工具进行客观评估的方法^[20]。本研究在前人研究的基础上,采用Maxamine Web Analysist软件并结合站长之家工具来获取政府数据开放网站的客观评价数据。

1.3 政府数据开放效率评估

目前国内外针对政府数据开放的研究多从建设成效的维度进行评估^[21],而从投入、产出角度评价政府数据开放效率的研究较少。马海群等^[11]构建以数据量、数据内容、连通率等为输入指标和以下载量、访问量等为输出指标的评估体系,运用DEA方法对我国8个省市的政府开放数据网站进行评价分析,其中对数据内容的分析主要考察界面的友好程度,采用专家打分的方式获得指标数据,没有涉及对元数据的完整性等数据质量维度的评估;且由于我国政府数据开放平台的建设总体起步较晚^[22],该评估涉及的对象还较少。因此本研究在构建政府数据开放平台的投入产出评估指标体系时,结合已有学者针对数据维度的评估研究增加了评价指标,并将评估对象扩大到我国现已建成的所有省市政府数据开放平台。

2 评估指标选择与数据获取

2.1 政府数据开放平台建设效率的DEA评估指标体系

在运用DEA方法进行评价时,如果被评价对象较多,而输入、输出指标较少,则会出现绝大多数被评价对象非DEA有效的结果^[23],因此本研究在借鉴前人研究基础上,结合各省市政府数据开放平台的实际情况,从数据和平台两个维度构建了具有9个投入指标和4个产出指标的DEA评估指标体系(见表1)。

2.2 样本选取与数据来源

本研究以截至2018年6月我国各省市已建成的政府数据开放平台为样本,由于各个平台给出的数据描述信息不尽相同,如部分平台并没有数据集的访问量、下载量等信息(广东省阳江市),综合考虑本研究构建的

表1 政府数据开放平台建设效率的DEA评估指标体系

逻辑维度		指标名称	指标内容
投入 指标	平台	网页总数 ^[11] （个）	网页总数是网站内部的所有网页数，是衡量网站大小和规模的重要指标之一，网页总数越多说明网站整体内容方面投入越大
		站内链接数 ^[11] （条）	站内链接数是网站指向自身的网页链接总数，用来衡量网站内部的层次性和结构合理性，站内链接数越多说明网站整体结构方面投入越大
		网站连通率 ^[11] （%）	网站连通率是指从网站主页出发，通过各个层级的链接所能到达的网页数与网页总数之比，能够反映网站导航设计的合理性和所有页面的可达性，网站连通率越高说明网站整体页面设计方面投入越大
		下载速度 ^[11] （毫秒）	下载速度是指用户访问数据开放网站时的平均响应时间，一般访问时间越短用户越能获得好的体验，也间接说明网站整体性能优化方面投入越多
	数据	数据集总量（个）	数据集总量是指数据开放平台已经开放的所有数据资源的总量
		数据领域覆盖度 ^[3] （%）	数据领域覆盖越广说明投入越大
		数据可机读格式占比 ^[24] （%）	可机读格式是指计算机可直接读取的CSV、XLS（X）、XML、JSON等格式，而DOC、PDF等不是可机读格式，数据可机读格式越高说明数据预处理方面的投入越大
		承诺高频更新数据集占比 ^[3] （%）	高频更新数据集是指平台标识的实时或者每天更新的数据集，数据更新频率越快说明投入越大
		元数据完整性 ^[25] （%）	元数据是对开放数据资源的描述，元数据越完整，用户能获得的信息就越多，说明数据资源维护方面的投入越多
产出 指标	平台	网站外链占比 ^[11] （%）	网站外链占比指链接总数比重越高说明其投入越大，这是从链接维度来衡量网站的外部影响力
		外部网站指向链接的数量 ^[11] （个）	与网站外链的区别在于该指标表示指向该网站的外部网站总数，从网站维度来衡量其外部影响力
	数据	数据集平均访问量（次）	数据集访问量越多说明用户关注度越高
		数据集平均下载量（次）	数据集下载量越多说明用户利用率越高

DEA指标体系和各省市数据开放平台的实际情况,最终选取了39个典型的政府数据开放平台作为评价对象,评估对象总数符合数据包络分析中决策单元个数不少于输入输出指标总数2倍的要求。

2.3 DEA指标描述性统计分析

对于开放数据平台层面的投入和产出指标,本研究综合利用Maxamine Web Analyst分析软件和站长之家工具进行数据获取,下载速度指标采用站长之家工具获取,其他指标来源于Maxamine Web Analyst产生的结果报告。开放数据内容层面的数据是通过自编Python爬虫工具进行获取,以上数据均集中于2018年6月4—6日获取。

投入指标中,从网页总数、站内链接数及数据集总量来看(见图1),最大值和最小值在数值上的差距较大,且中位数均低于均值,说明半数以上的平台连平均水平都达不到,建设起步较早的贵州省贵阳市数据集开

放总量达到2 754个,而广东省湛江市只开放46个数据集,可见政府数据开放平台在网站维护和数据集开放上的投入力度存在较大差异。

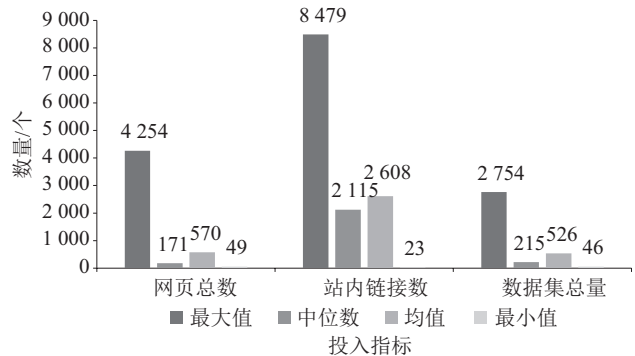


图1 DEA投入指标统计分析 (1)

从元数据完整性、数据领域覆盖度、数据可机读格式占比指标来看(见图2),三者的均值和中位数均处于较高水平,说明政府在这些方面的投入总体较大;但从承诺高频更新数据集占比来看,其均值仅为8.23%,这可能与我国政府数据开放平台总体还处于起步建设阶

段的实际情况有关,由此可以看出数据集的更新频率还有较大的提升空间。

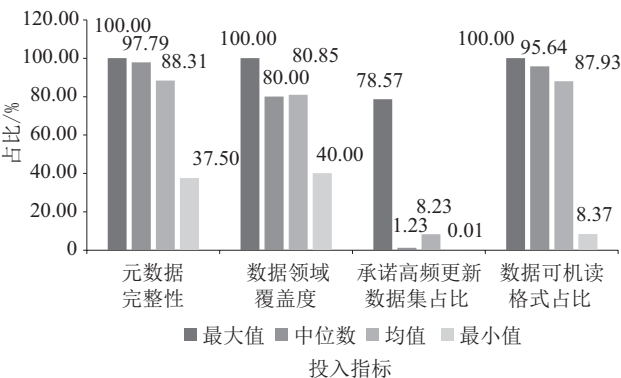


图2 DEA投入指标统计分析 (2)

表2 DEA投入和产出指标统计分析

项 目	网站连通率/%	下载速度/毫秒	网站外链占比/%	外部网站指向链接的数量/个	数据集平均访问量/次	数据集平均下载量/次
最大值	100.00	2 021	67.25	2 091	465.566 2	3 175.440 5
中位数	100.00	338	14.95	314	54.080 4	13.024 7
均值	96.33	420	17.67	386	55.994 4	331.492 2
最小值	47.83	11	0.04	3	0.009 2	0.647 1

3 模型求解与效率评估

3.1 DEA评估模型的建立与求解

由于所抓取的数据量纲不同,需要进行标准化处理以消除量纲和数量级的影响,如下载速度指标通过响应时间来测量,需要对该指标进行倒数化处理。最终对上述所有输入、输出指标进行0~1标准化后,利用Deap2.1软件,采用基于投入视角的规模收益可变DEA模型对政府数据开放平台的建设效率展开分析,最终结果见表3。

39个政府数据开放平台的整体综合效率较高,平均值为0.907,纯技术效率的平均值要高于规模效率,且16个(占比41%)平台的规模报酬处于递增阶段,说明政府数据开放平台的资源配置水平有待提高。结合表3中的结果数据和DEA有效性的一般分类方法,可将39个政府数据开放平台划分为3个层次。

第1层次: DEA有效(综合效率=1)。该层次的政府数据开放平台既包括北京市、上海市、广东省广州市、江苏省无锡市等发展起步较早的平台,也包括江西省、宁夏回族自治区、山东省德州市等运营上线较晚的平台,共计23个(占比59%),表明这些省市及自治区从数

从网站连通率和下载速度指标来看(见表2),二者衡量的是网站的易用性,网站连通率的平均值达到96.33%,说明网站内页面的可达性较高;下载速度用响应时间来衡量,响应时间越短则下载速度越快,该指标的中位数小于平均值,说明半数以上的平台下载速度优于平均水平。

产出指标中,从网站外链占比、外部网站指向链接的数量及数据集平均访问量来看(见表2),三者的中位数均低于均值,但相差不大;而从数据集平均下载量来看,其中位数要远低于平均水平,说明开放数据用户的下载量整体处于较低水平,有待进一步提升。

据和平台两个维度衡量其投入产出配置时效率上达到了相对最优。以广东省5个城市的DEA有效平台为例,将其标准化后的投入和产出累加,得到如图3所示的变化趋势,对于DEA有效单元来说,一般其投入越多产出就越多。由于DEA方法给出的相对有效性结论只是针对被评价群体,且当下我国政府数据开放平台的建设总体还处于起步阶段,因此对于投入相对较少但是效率较高的省市来说,其在平台建设上的投入仍应继续保持或增加。

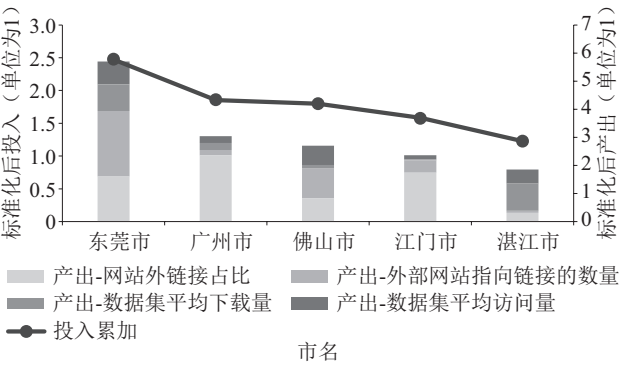


图3 DEA有效单元投入产出分析 (以广东省5市为例)

第2层次: 弱DEA有效(综合效率<1, 纯技术效率=1)。该层次的政府数据开放平台包括广东省、山东

表3 DEA评估模型求解结果

平台名称	综合效率	纯技术效率	规模效率	规模报酬	平台排名
北京市政务数据资源网	1.000	1.000	1.000	--	1
开放广东	0.820	1.000	0.820	irs	29
江西省政府数据开放网站	1.000	1.000	1.000	--	1
开放宁夏	1.000	1.000	1.000	--	1
上海政府数据服务网	1.000	1.000	1.000	--	1
广州市政府数据统一开放平台	1.000	1.000	1.000	--	1
哈尔滨市数据开放	1.000	1.000	1.000	--	1
武汉政府公开数据服务网	1.000	1.000	1.000	--	1
济南市公共数据开放网	1.000	1.000	1.000	--	1
青岛市公共数据开放网	0.930	1.000	0.930	irs	26
宁波市政府数据服务网	1.000	1.000	1.000	--	1
数据东莞	1.000	1.000	1.000	--	1
佛山政府数据开放平台	1.000	1.000	1.000	--	1
开放惠州	0.799	0.908	0.879	irs	32
开放江门	1.000	1.000	1.000	--	1
湛江数据服务网	1.000	1.000	1.000	--	1
开放中山	0.379	1.000	0.379	irs	39
贵阳市政府数据开放平台	0.752	0.755	0.996	irs	33
荆门市人民政府数据开放模块	1.000	1.000	1.000	--	1
长沙数据开放	1.000	1.000	1.000	--	1
苏州市政府数据开放平台	1.000	1.000	1.000	--	1
无锡市政府数据服务网	1.000	1.000	1.000	--	1
扬州市政务数据服务网	1.000	1.000	1.000	--	1
乌海市数据开放平台	0.811	0.947	0.857	irs	30
滨州市公共数据开放网	0.607	0.953	0.637	irs	38
德州市公共数据开放网	1.000	1.000	1.000	--	1
东营市公共数据开放网	0.736	0.976	0.755	irs	34
菏泽市公共数据开放网	0.947	1.000	0.947	irs	25
济宁市公共数据开放网	0.800	0.983	0.814	irs	31
莱芜市公共数据开放网	1.000	1.000	1.000	--	1
聊城市公共数据开放网	1.000	1.000	1.000	--	1
临沂市公共数据开放网	0.911	1.000	0.911	irs	27
日照市公共数据开放网	0.711	0.968	0.735	irs	35
泰安市公共数据开放网	0.608	0.932	0.653	irs	37
威海市公共数据开放网	0.700	1.000	0.700	irs	36
潍坊市公共数据开放网	0.949	1.000	0.949	irs	24
烟台市公共数据开放网	1.000	1.000	1.000	--	1
枣庄市公共数据开放网	1.000	1.000	1.000	--	1
淄博市公共数据开放网	0.908	0.990	0.917	irs	28
均值	0.907	0.985	0.920	-	-

注：“irs”和“--”分别表示规模报酬递增和规模报酬不变。

省青岛市等7个（占比18%）省市，这些平台的建设处于纯技术效率有效且规模报酬递增的阶段，因此成比例

地增加投入可以显著提升总体的建设效率。以广东省“开放广东”数据开放平台为例，2018年数据集总量相

比2017年增速达到400%，而处于DEA有效层次的北京市和上海市的增速则相对要小许多（见表4），由此也说明广东省数据开放平台处于高速发展阶段，其综合效率仍然有提升的空间。

表4 2017—2018年政府数据开放平台数据集总量变化

数据开放平台	2017年数据集总量 ^[3] /个	2018年数据集总量/个	数据集总量同比增长率/%
开放广东	110	550	400
北京市政务数据资源网	483	754	56
上海政府数据服务网	1 120	1 556	39

第3层次：非DEA有效（综合效率<1，纯技术效率<1）。贵州省贵阳市、山东省淄博市、广东省惠州市等9个（占比23%）省市的政府数据开放平台建设效率属于非DEA有效层次，从纯技术效率的结果来看，该层次的平台需要优化平台的建设投入和资源配置；从规模效率的结果来看，这些平台处于规模报酬递增阶段，因此可以适当增加各项指标的投入力度，进而提高其建设效率。以贵阳市政府数据开放平台为例，其建设起步

较早，且数据集开放总量居于39个平台中的第1位，但是从建设效率的角度来说，贵阳市政府数据开放平台虽投入较多，但是在产出方面相对不足，因此其纯技术效率和综合效率均小于1，故政府今后应合理地优化数据开放上的资源投入配比，加大数据开放宣传力度，提升用户利用率，从而优化平台的建设效率。

3.2 非DEA有效平台的效率改进

利用Deap2.1软件对9个非DEA有效的政府数据开放平台展开冗余度分析，从产出指标来看（见表5），数据集的平均访问量、下载量都严重不足，如访问量的效率改进平均值达到749%，说明这些平台在用户关注和利用上亟需提高，除贵州省贵阳市的政府数据开放平台建设较早外，其他省市平台在数据集下载量和访问量上都有提升的空间，应该加大宣传力度，确保所开放的数据能够被用户发现和使用，进而才有可能产生更大的社会效益和经济效益。就网站影响力来说，这些平台还有一定的改进空间，对应的省市政府应当采取适当措施提升宣传效果，使得平台更容易被用户发现。

表5 非DEA有效政府数据开放平台的产出效率改进

平台名称	网站外链接量	外部网站指向链接的数量	数据集平均下载量	数据集平均访问量
淄博市公共数据开放网	21.72	0.00	63.91	607.14
乌海市数据开放平台	329.22	0.00	260.00	188.41
济宁市公共数据开放网	46.59	0.00	0.00	326.92
开放惠州	8.55	0.00	917.16	402.44
贵阳市政府数据开放平台	192.58	0.00	0.00	0.00
东营市公共数据开放网	6.90	2.94	94.67	86.96
日照市公共数据开放网	92.92	12.77	45.23	807.69
泰安市公共数据开放网	160.99	20.38	151.96	2 250.00
滨州市公共数据开放网	149.96	25.76	28.57	2 071.43
均值	112.16	6.87	173.50	749.00

从投入指标来看（见表6），非DEA有效政府数据开放平台在数据集总量、更新频率、数据领域覆盖度及站内链接数等指标上均存在相对较大的投入冗余，表明在这些指标上的投入相比DEA有效的决策单元来说没有转化成对应的产出增加。以贵州省贵阳市为例，其数据集总量上的投入冗余值达到71.10%，说明平台在数据量方面的投入已经相对比较可观，后续应该注重提升开放数据的利用率，进而达到相对更优的投入产

出配置和更高的建设效率。

4 结论与展望

本研究以我国39个典型的政府数据开放平台为例，基于数据、平台两个维度构建投入产出的评估指标体系，并运用DEA方法分析平台的建设效率。研究发现，平台的整体建设效率相对较高，且规模报酬均处于

表6 非DEA有效政府数据开放平台的投入效率改进

%

平台名称	网页总数	站内链接数	网站连通率	下载速度	数据集总量	承诺高频更新数据集占比	数据领域覆盖度	元数据完整性	数据可机读格式占比
淄博市公共数据开放网	0.00	-8.98	0.00	0.00	-11.68	0.00	0.00	-0.71	0.00
乌海市数据开放平台	0.00	-62.48	0.00	0.00	-26.17	-84.08	-36.40	-1.32	0.00
济宁市公共数据开放网	0.00	-2.43	0.00	0.00	-13.67	-95.57	0.00	0.00	0.00
开放惠州	0.00	0.00	0.00	-55.28	-2.88	-73.70	0.00	-6.16	0.00
贵阳市政府数据开放平台	-68.50	-54.30	-2.52	0.00	-71.10	-53.76	-37.28	-14.04	0.00
东营市公共数据开放网	0.00	-6.13	0.00	0.00	-43.98	-97.62	-14.25	-1.20	-2.52
日照市公共数据开放网	0.00	-22.53	0.00	0.00	0.00	-27.86	-21.71	-1.13	0.00
泰安市公共数据开放网	-14.62	-45.26	0.00	0.00	-31.07	-31.85	0.00	-2.66	0.00
滨州市公共数据开放网	0.00	0.00	0.00	0.00	-26.29	-90.26	0.00	0.00	-9.60
均值	-9.24	-22.46	-0.28	-6.14	-25.20	-61.63	-12.18	-3.02	-1.35

不变或者递增的阶段,说明这些平台整体上投入产出的配置较为合理,但相对效率并不等于建设成效。结合我国政府数据开放平台建设普遍起步较晚的实际情况,政府应该继续保持或加大投入力度,以期促进数据开放平台建设的良性发展,获得更大的社会效益和经济效益。对于非DEA有效的平台,本研究发现其投入方面存在一定的冗余,但更多地是由于产出方面的不足导致的,整体来说开放数据的利用率还亟待提升,相关政府应该加大数据开放的宣传力度,让更多的用户发现和利用数据。

在大数据时代下,世界各国政府都在加速推进本国政府数据开放的进程,图书馆也被赋予了新的历史使命。公共图书馆在数据开放的过程中也应广泛借鉴政府数据开放平台的建设经验,在逐步增加开放数据量的同时,也应保证开放数据能够广泛地被大众利用,确保在开放服务上投入的资源能够转化为相应的产出,从而为公众提高更好的信息服务及不断提高图书馆自身的影响力。

由于我国各省市政府数据开放平台投入使用的时长不尽相同,未来的研究应该限定一个时间周期,如以一年内的投入和产出为例进行分析,从而使得政府数据开放平台的建设效率评估结果更具有时效性,反馈结果也有助于政府及时作出资源投入等方面的调整。此外,本研究运用DEA方法对政府数据开放平台进行效率评估时,并没有考虑各个评价指标所占的权重,后续的研究应该结合专家打分等方法赋予指标相应的权重,以区分不同指标之间的相对重要性。另外,由于DEA评价的是相对效率,适用于同类型决策单元之间的相互比较^[26],但是不能区分水平都较低且相近的情

况^[27],因此未来的研究还应考虑在研究方法上予以改进。最后,随着政府数据开放平台的不断发展和完善,其提供的功能和开放的数据将会日趋增加,本研究所构建的评估指标体系也应结合实际动态更新。

参考文献

- [1] 郑磊. 开放政府数据研究: 概念辨析、关键因素及其互动关系 [J]. 中国行政管理, 2015 (11): 13-18.
- [2] GODA S. Open data and open government [J]. Informacios Tarsadalom, 2011, 11 (1): 181-187.
- [3] 2017中国地方政府数据开放报告 [EB/OL]. [2018-06-28]. <http://www.dmg.fudan.edu.cn/wp-content/uploads/中国地方政府数据开放平台报告.pdf>.
- [4] 国务院关于印发促进大数据发展行动纲要的通知 [EB/OL]. (2015-08-31) [2018-06-28]. http://www.gov.cn/zhengce/content/2015-09/05/content_10137.htm.
- [5] 黄思棉, 张燕华. 当前中国政府数据开放平台建设存在的问题与对策研究——以北京、上海政府数据开放网站为例 [J]. 中国管理信息化, 2015 (14): 175-177.
- [6] 王迪. 开放政府数据评估体系比较研究 [J]. 情报资料工作, 2017 (5): 27-33.
- [7] 赵玉攀. 国内开放政府数据研究的系统评述 [J]. 现代情报, 2018 (2): 164-170.
- [8] 莫祖英. 国外政府开放数据及其质量研究述评 [J]. 情报资料工作, 2018 (2): 24-28.
- [9] 郑跃平, 刘美岑. 开放数据评估的现状存在问题——基于国外开放数据评估的对比和分析 [J]. 电子政务, 2016 (8): 84-93.
- [10] 郑磊, 熊久阳. 中国地方政府开放数据研究: 技术与法律特

- 性[J]. 公共行政评论, 2017(1): 53-73.
- [11] 马海群, 王今. 基于DEA的政府开放数据网站效率评价[J]. 数字图书馆论坛, 2016(6): 2-7.
- [12] 李美娟, 陈国宏. 数据包络分析法(DEA)的研究与应用[J]. 中国工程科学, 2003(6): 88-94.
- [13] 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(18): 43-55.
- [14] VETRÒ A, CANOVAL, TORCHIANO M, et al. Open data quality measurement framework: Definition and application to open government data[J]. Government Information Quarterly, 2016, 33(2): 325-337.
- [15] VELJKOVIĆ N, BOGDANOVIĆ-DINIĆ S, STOIMENOV L. Benchmarking open government: an open data perspective[J]. Government Information Quarterly, 2014, 31(2): 278-290.
- [16] 赵继娣, 张罕仑. 地方政府数据开放成效评价研究——以上海市为例[J]. 电子政务, 2017(9): 11-21.
- [17] 谭必勇, 陈艳. 我国开放政府数据平台数据质量研究——以十省、市为研究对象[J]. 情报杂志, 2017(11): 99-105.
- [18] 马海群, 唐守利. 基于结构方程的政府开放数据网站服务质量评价研究[J]. 现代情报, 2016(9): 10-15.
- [19] 邹纯龙, 马海群. 基于神经网络的政府开放数据网站评价研究——以美国20个政府开放数据网站为例[J]. 现代情报, 2016(9): 16-21.
- [20] 韦忻伶, 安小米, 李雪梅, 等. 开放政府数据评估体系述评: 特点分析[J]. 图书情报工作, 2017(18): 119-127.
- [21] SUSHA I, ZUIDERWIJK A, JANSSEN M, et al. Benchmarks for evaluating the progress of open data adoption: usage, limitations, and lessons learned[J]. Social Science Computer Review, 2015, 33(5SI): 613-630.
- [22] 范佳佳. 中外开放数据辅助政府治理现状比较研究[J]. 情报科学, 2017(11): 38-45.
- [23] 李娜, 许晓东. 基于数据包络分析的省级电子政务绩效评估研究[J]. 现代情报, 2013(8): 73-78.
- [24] 郑磊, 高丰. 中国开放政府数据平台研究: 框架、现状与建议[J]. 电子政务, 2015(7): 8-16.
- [25] KUBLER S, ROBERT J, NEUMAIER S, et al. Comparison of metadata quality in open data portals using the analytic hierarchy process[J]. Government Information Quarterly, 2018, 35(1): 13-29.
- [26] 袁群. 数据包络分析法应用研究综述[J]. 经济研究导刊, 2009(19): 201-203.
- [27] 闫现洋, 余小萍, 吴江. 基于DEA方法的高校图书馆网站效率评价研究——基于6所部属师范高校的实证分析[J]. 图书情报工作, 2011(23): 68-71.

作者简介

韩磊, 男, 1995年生, 硕士, 研究方向: 信息管理、电子政务。

胡广伟, 男, 1975年生, 教授, 博士生导师, 通信作者, 研究方向: 信息资源管理、电子政务、政务大数据, E-mail: hugw@nju.edu.cn。

Efficiency Evaluation on Open Government Data Portals Based on DEA

HAN Lei^{1,2} HU GuangWei^{1,2}

(1. School of Information Management, Nanjing University, Nanjing 210023, China;

2. Government Data Resources Institution of Nanjing University, Nanjing 210023, China)

Abstract: Since the Open Government Data Portal of Shanghai came into service in 2012, more than 40 open government data portals in China has been online. Efficiency evaluation on open government data portals will reveal the effectiveness of input-output transformation as well as promote healthy development of these portals. This study is based on China's 39 open government data portals on provincial and city level. We established the input-output evaluation system based on the portal and data dimensions. The data envelopment analysis method (DEA) is used to analyze the technical efficiency, pure technical efficiency and scale efficiency of these portals. The results showed that generally the efficiency of open government data portals is relatively high, and the DEA effective portals accounts for 59%. But the output of non-DEA effective portals in the aspect of user data utilization needs to be improved.

Keywords: Open Government Data; Efficiency Evaluation; Data Envelopment Analysis; Input-output Analysis

(收稿日期: 2018-09-02)