



中国开放数据指数
CHINA OPEN DATA INDEX

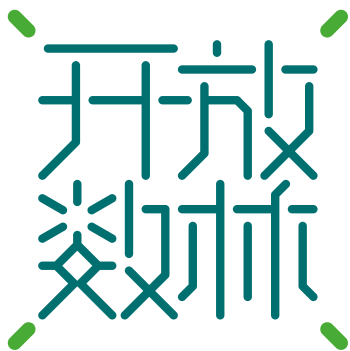
2018
No.03

中国地方政府数据开放报告

联合发布方 复旦大学 / 提升政府治理能力大数据应用技术国家工程实验室 / 国家信息中心数字中国研究院

指数出品方 复旦大学数字与移动治理实验室





中国开放数据指数
CHINA OPEN DATA INDEX

2018

• 联合发布方

复旦大学

提升政府治理能力大数据应用技术国家工程实验室

国家信息中心数字中国研究院

• 指数出品方

复旦大学数字与移动治理实验室

• 合作方

中电科新型智慧城市研究院

复旦发展研究院传播与国家治理研究中心

浪潮集团有限公司

冥睿(上海)信息科技有限公司

开放数据中国

麦明德设计事务所

中国地方政府数据开放报告2018



中国开放数林指数
CHINA OPEN DATA INDEX

2018

目 录 I N D E X

1 背景与趋势	03
2 评估方法	07
2.1 评估指标体系构建	08
2.2 评估范围	10
2.3 数据采集与分析方法	12
2.4 指标计算方法	12
3 主要发现	13
3.1 总体发现	14
3.1.1 上线时间	14
3.1.2 地区分布	14
3.1.3 行政层级	15
3.1.4 平台类型	15
3.2 数据层	15
3.2.1 数据数量	15
3.2.2 数据质量	16
3.2.3 数据标准	18
3.2.4 数据覆盖面	24
3.2.5 数据持续性	31
3.2.6 数据层结论	32
3.3 平台层	33
3.3.1 平台概览	33
3.3.2 平台导引	38
3.3.3 数据获取	43
3.3.4 工具提供	47
3.3.5 利用成果展示	51
3.3.6 互动交流	53
3.3.7 个性化整合	59
3.3.8 平台层结论	60
3.4 准备度	62
3.4.1 法规与政策	62
3.4.2 领导力	69
3.4.3 组织保障	70
3.4.4 准备度层结论	73

4 “开放数林”指数分值与分析	75
4.1 指数分值与排名	76
4.1.1 省级	76
4.1.2 地市级(含副省级)	76
4.2 指数综合分析	78
4.2.1 空间分析	78
4.2.2 长势分析	79
4.2.3 形态分析	80
4.2.4 叶脉分析	83
5 数林建言	87
5.1 及时造林, 后来居上, 提高“开放数林”覆盖率	88
5.2 竞相共生, 枝杈相连, 保护“开放数林”多样性	88
5.3 保持树形, 监测态势, 促进“开放数林”可持续生长	88
5.4 不断修剪, 精耕细作, 满足用户对“开放数林”的向往	88
6 结语	90
7 附录	91

1

背景和趋势 Background



当前，人类正在向智能时代迈进，而数据是智能时代的重要基石和养料。在智能时代，数据作为一种新的基础设施，将和物理基础设施同等重要。在物理世界，政府在大型公共基础设施的建设中发挥了主导作用，在智能时代，政府在构筑公共数据基础设施上的责任仍然义不容辞。

数据是国家的战略性资源，一个国家有大量基础性、关键性的数据掌握在政府手中。这些数据是社会的公共资源，在保障国家秘密、商业秘密和个人隐私的前提下，将政府数据最大限度地开放出来，让社会进行充分融合和利用，合力构筑数据基础设施，有利于释放数据能量，激发创新活力，创造公共价值。

2017年以来，国家层面又提出一系列新的要求，有力推动着数据开放的进程。2月，中央全面深化改革领导小组审议通过了《关于推进公共信息资源开放的若干意见》，要求着力推进重点领域公共信息资源开放，释放经济价值和社会效应。10月，中共十九大报告强调“加快建设创新型国家”，对加快建设网络强国、数字中国、智慧社会等提出新方略。12月，习近平总书记在中共中央政治局第二次集体学习时强调，推动实施国家大数据战略，加快完善数字基础设施，推进数据资源整合和开放共享。

2018年是我国政府数据开放的关键一年。1月，中央网信办、国家发改委、工信部联合印发《公共信息资源开放试点工作方案》，确定在北京、上海、浙江、福建、贵州五地开展公共信息资源开放试点，要求试点地区结合实际制定具体实施方案，明确试点范围，细化任务措施，积极认真开展相关工作，着力提高开放数据质量，促进社会化利用，探索建立制度规范，并于2018年底前完成试点各项任务。根据2015年国务院印发的《促进大数据发展行动纲要》的要求，2018年底前，我国将建成国家政府数据统一开放平台。在此背景下，对我国地方政府的数据开放情况进行评估和分析，对于推进全国范围内政府数据资源开

放与利用具有重要意义。

2017年5月的贵阳数博会上，复旦大学与提升政府治理能力大数据应用技术国家工程实验室联合发布了国内首个政府数据开放评估报告——《中国地方政府数据开放平台报告》，得到了国内政府、学界、企业和媒体的广泛关注。2017年12月，我们又将政府数据开放的评估维度由数据层面扩展到平台层面，发布了第二期《中国地方政府数据开放平台报告》。这一年来，政府数据开放在我国如雨后春笋，迅速推进，纳入报告的地方政府数据开放平台已由去年的19个增加至46个，除在数据层和平台层方面继续提升之外，不少地方还在法规政策、组织管理等方面进行了有益的探索，夯实数据开放的基础层。

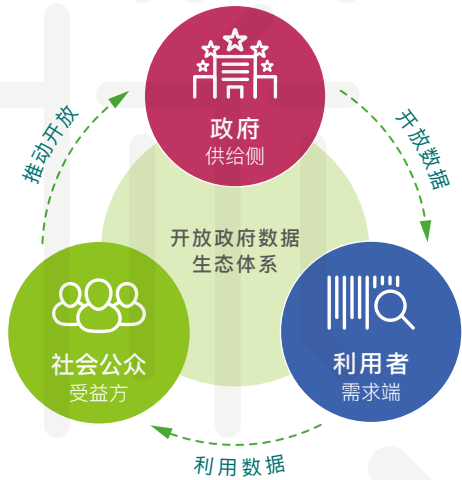


图1 开放数据的动态循环

因此，复旦大学继续联合提升政府治理能力大数据应用技术国家工程实验室、国家信息中心数字中国研究院发布中国地方政府数据开放报告。报告由复旦大学数字与移动治理实验室基于多年研究积累和数据监测分析制作完成，邀请政产学研各界专家共同参与，构建起一个系统、科学、可操作的评估指标体系，并基于该体系对我国已开放数据的地方政府进行综合评价，制作指数报告，提出优化建议。

政府数据从被开放、被利用到产生效益是一个动态循环的过程，政府、数据利用者和社会公众共同构成了一个生态体系。政府部门作为供给侧将数据开放出来，数据利用者作为需求端对数据进行利用，并以其开发的创新应用服务于社会公众，获益的数据利用者和公众又推动政府进一步开放数据。在这个过程中，政府部门是原材料的提供者，数据利用者是加工者，社会公众是最终受益方(如图1)。

◇ 开放数据，蔚然成林



2017 No.01



2017 No.02



2018 No.03

一棵棵开放“数木”的丛然并生、成荫如盖、直至叶茂花开，终将成长为一一片繁盛多样、平衡清朗、持续循环的“开放数林”。

《中国地方政府数据开放报告》将定期发布“开放数林指数”，持续追踪我国地方政府数据开放生态体系的形成与发展，精心测量“数木”们的树形、高度、粗细、材质与价值。

2

评估方法 Methodology

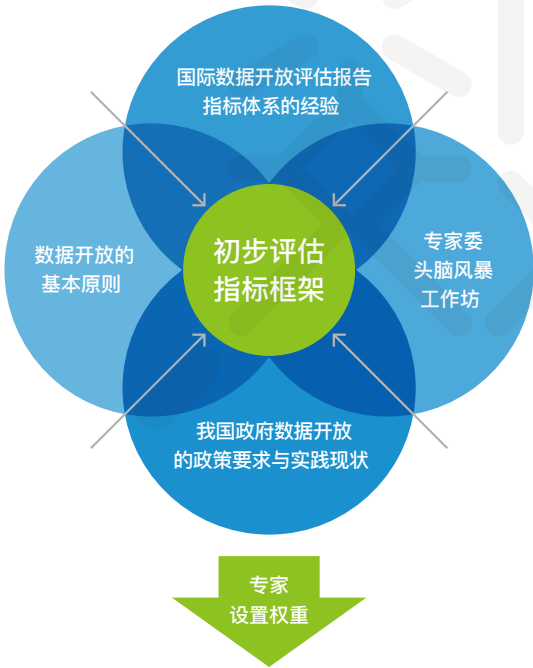


◆ 2.1 评估指标体系构建

• 2.1.1 评估指标与权重的产生过程

报告出品方邀请近四十位我国数据开放领域学、研、产各界的专家与学者，组成“中国开放数林指数评估专家委员会”（以下简称“专家委”，详细成员名单见报告附录），以反映跨界、多学科、第三方、中立的专业视角和实际需求。

基于数据开放的基本原则，借鉴国际数据开放评估报告指标体系的经验，立足我国政府数据开放的政策要求与实践现状，专家委成员围绕开放数林指数评估方法举行了头脑风暴工作坊，通过互动讨论、阐述归类、现场投票等方法得出初步评估指标框架。评估指标框架在经过系统梳理后，专家委成员又通过线上工具对指标的相对重要性进行排序和权重分配。最终，构建起一个系统、科学、可操作的中国地方政府数据开放评估指标体系(如图2)。



中国地方政府数据开放评估指标体系

图2 评估指标体系确立步骤

评估指标体系共包括数据层、平台层、准备度三个维度，每个维度下设有相应的一二三级指标，专家委成员对三个维度和各项一级指标的排序情况如表1：

表1 专家委成员对评估指标相对重要性排序的结果

基本维度	指标名称	排序得分	重要性
数据层	数据质量	1.40	1
	数据标准	2.80	2
	数据可持续性	2.97	3
	数据数量	3.73	4
	数据覆盖面	3.87	5
平台层	数据获取	1.50	1
	平台导引	2.67	2
	工具提供	3.77	3
	利用成果展示	4.17	4
	平台概览	4.70	5
	互动交流	4.73	6
	个性化整合	5.87	7
准备度	法规与政策	1.67	1
	领导力	1.83	2
	组织保障	2.40	3

• 2.1.2 评估指标与权重结果

数据层、平台层、准备度是评估政府数据开放的基本维度。依据专家委成员打分结果，调整后的三个维度的权重分配如表2。

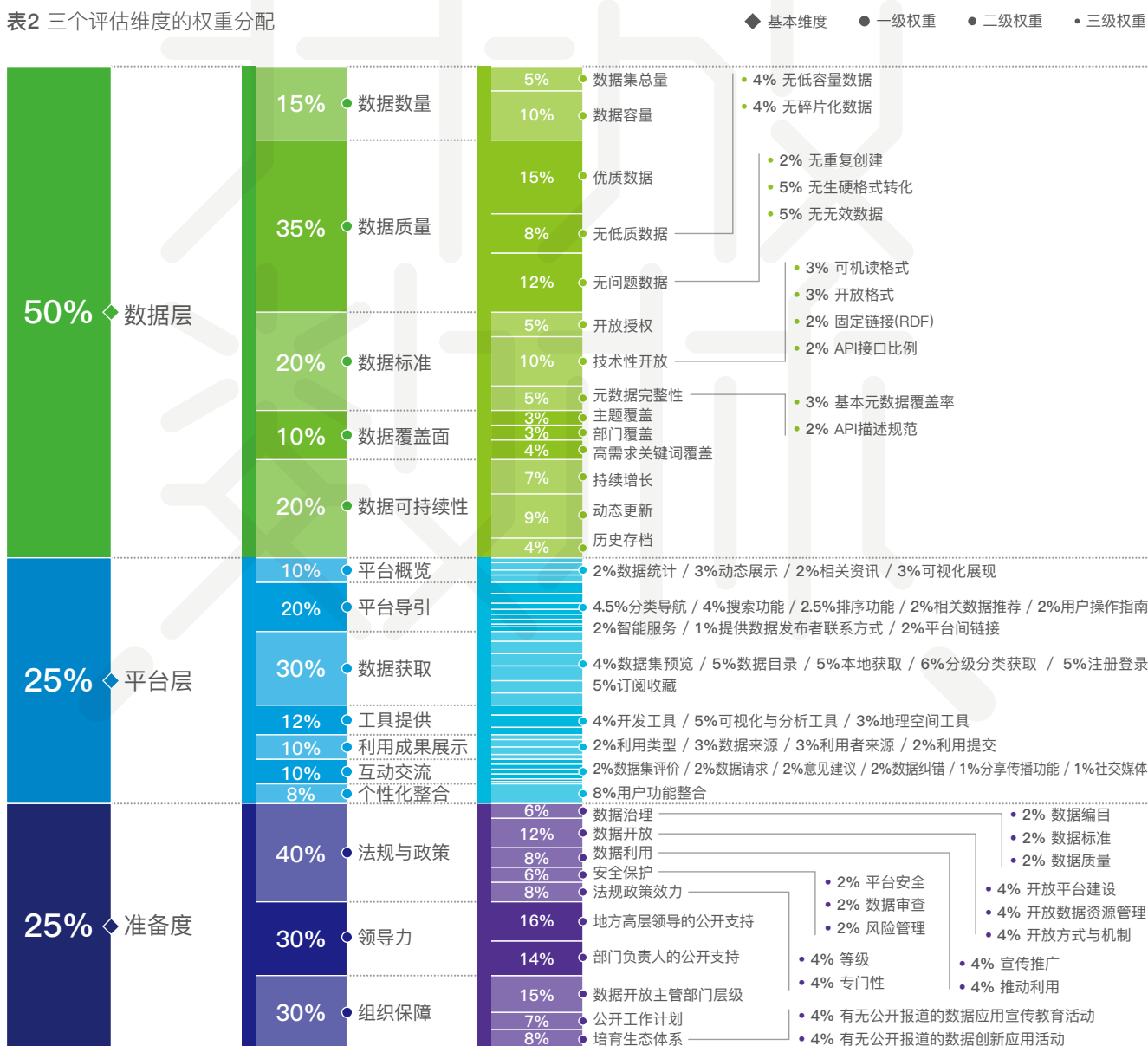
数据层是“里子”，是数据开放的核心。数据层共设置数据数量、数据质量、数据标准、数据可持续性、数据覆盖面五个一级指标，具体评估指标如表2所示。其中，数据质量是今年新设置的指标，被专家委成员一致推选为数据层评估的核心。

平台层是“面子”，是政府开放数据和用户获取数据的载体，是展示应用成果的中心，也是连接数据开放供给侧和需求端的桥梁。平台层共包括平台概览、平台导引、数据获取、

工具提供、利用成果展示、互动交流和个性化整合七个一级指标。这些指标覆盖了数据利用者从在平台上发现数据、获取数据、利用数据，到与政府部门进行互动反馈，再到展示数据利用成果的全过程，具体评估指标如表2所示。其中，个性化整合是今年新增的指标，评估平台是否基于个体用户需求对相关服务进行了订制化、个性化的整合。

准备度是“底子”，是数据开放的基础。准备度是指一项工作的事前准备程度，经专家委成员普遍提议后，今年新设立该维度。在本报告中，准备度着重评估地方政府为实现数据开放所做的基础性工作和现阶段准备情况。包括法规与政策、领导力、组织保障三个一级指标，具体评估指标如表2所示。

表2 三个评估维度的权重分配



◆ 2.2 评估范围

报告根据公开报道，以及使用“数据+开放”、“数据+公开”、“公共+数据”、“政务+数据”、“政府+数据”、“地名+数据”、“地名+政府数据”、“地名+开放数据”等关键词进行搜索，发现了截至2018年4月中旬，我国已上线的政府数据开放平台，并将符合以下条件的地方政府数据开放平台纳入评估范围：

- 1.平台域名中出现gov.cn，作为确定其为政府官方认可的数据开放平台的依据。
- 2.平台形式为“统一专有式”或“统一嵌入式”。“统一专有式”是指开放数据统一汇聚在一个专门的平台上进行开放；“统一嵌入式”是指开放数据统一汇聚为一个栏目版块，嵌入在政府门户网站或政务服务网站上。除浙江、肇庆、阳江、梅州为“统一嵌入式”平台外，其余均为“统一专有式”平台。各个条线部门建设的非集中式开放数据的平台不在本报告

的评估范围内。

- 3.平台所代表的地方政府的行政级别为地级以上。
- 4.平台上确实开放了电子格式的、可通过下载或接口形式获取的、结构化的数据集。有些名为“数据开放”的平台实质上只提供了非结构化的文本内容或跳转到其他相关网页的链接，不存在可通过下载或接口形式获取的、结构化的数据集。这类平台更多属于传统的“信息公开”门户，因而未被纳入本次评估范围，如新疆维吾尔自治区政务数据开放网、四川省人民政府网站上的“开放数据”模块和广东清远市人民政府网的“数据开放”频道等。

基于以上选择标准，被纳入本次报告评估的地方政府数据开放平台共 46个，这些平台符合政府数据开放的基本特征，是我国政府数据开放的先行者。具体平台名称、所属地方政府和平台域名如表3所示。

表3 评估范围 *按行政层级及拼音首字母排序

No.	平台名称	地点	层级	平台域名
1	北京市政务数据资源网	北京市	省级	http://www.bjdata.gov.cn/
2	开放广东	广东省	省级	http://www.gddata.gov.cn/
3	贵州省政府数据开放平台	贵州省	省级	http://www.gzdata.gov.cn/
4	江西省政府数据开放网站	江西省	省级	http://data.jiangxi.gov.cn/
5	开放宁夏	宁夏回族自治区	省级	http://ningxiadata.gov.cn/
6	山东公共数据开放网	山东省	省级	http://data.sd.gov.cn/
7	上海政府数据服务网	上海市	省级	http://www.datashanghai.gov.cn/
8	浙江政务服务网	浙江省	省级	http://data.zjzfw.gov.cn/
9	广州市政府数据统一开放平台	广东省广州市	副省级	http://www.datagz.gov.cn/
10	深圳市政府数据开放平台	广东省深圳市	副省级	http://opendata.sz.gov.cn/
11	哈尔滨市数据开放	黑龙江省哈尔滨市	副省级	http://data.harbin.gov.cn
12	武汉政府公开数据服务网	湖北省武汉市	副省级	http://www.wuhandata.gov.cn
13	济南市公共数据开放网	山东省济南市	副省级	http://www.jndata.gov.cn
14	青岛市公共数据开放网	山东省青岛市	副省级	http://data.qingdao.gov.cn

No.	平台名称	地点	层级	平台域名
15	宁波市政府数据服务网	浙江省宁波市	副省级	http://www.datanb.gov.cn/
16	数据东莞	广东省东莞市	市级	http://dataopen.dg.gov.cn/
17	佛山政府数据开放平台	广东省佛山市	市级	http://www.fsddata.gov.cn/
18	开放惠州	广东省惠州市	市级	http://data.huizhou.gov.cn/
19	开放江门	广东省江门市	市级	http://opendata.jiangmen.gov.cn/
20	梅州市人民政府数据开放平台	广东省梅州市	市级	https://www.meizhou.gov.cn/opendata
21	中国阳江数据开放	广东省阳江市	市级	http://www.yangjiang.gov.cn/sj kf/
22	湛江数据服务网	广东省湛江市	市级	http://data.zhanjiang.gov.cn/
23	肇庆数据开放	广东省肇庆市	市级	http://www.zhaoqing.gov.cn/sj kf/
24	开放中山	广东省中山市	市级	http://zsdata.zs.gov.cn/web/index
25	贵阳市政府数据开放平台	贵州省贵阳市	市级	http://www.gyopendata.gov.cn/
26	荆门市人民政府数据开放模块	湖北省荆门市	市级	http://data.jingmen.gov.cn/app/
27	长沙数据开放	湖南省长沙市	市级	http://data.changsha.gov.cn/
28	苏州市政府数据开放平台	江苏省苏州市	市级	http://www.suzhou.gov.cn/dataOpenWeb/
29	无锡市政府数据服务网	江苏省无锡市	市级	http://etc.wuxi.gov.cn/opendata/
30	扬州市政务数据服务网	江苏省扬州市	市级	http://data.yangzhou.gov.cn/
31	乌海市数据开放平台	内蒙古自治区乌海市	市级	http://whdata.wuhai.gov.cn/odweb/
32	滨州市公共数据开放网	山东省滨州市	市级	http://bzdata.sd.gov.cn/
33	德州市公共数据开放网	山东省德州市	市级	http://dzdata.sd.gov.cn/
34	东营市公共数据开放网	山东省东营市	市级	http://dydata.sd.gov.cn/
35	菏泽市公共数据开放网	山东省菏泽市	市级	http://hzdata.sd.gov.cn/
36	济宁市公共数据开放网	山东省济宁市	市级	http://jindata.sd.gov.cn/
37	莱芜市公共数据开放网	山东省莱芜市	市级	http://lwdata.sd.gov.cn/
38	聊城市公共数据开放网	山东省聊城市	市级	http://lodata.sd.gov.cn/
39	临沂市公共数据开放网	山东省临沂市	市级	http://lydata.sd.gov.cn/
40	日照市公共数据开放网	山东省日照市	市级	http://rzdata.sd.gov.cn/
41	泰安市公共数据开放网	山东省泰安市	市级	http://tadata.sd.gov.cn/
42	威海市公共数据开放网	山东省威海市	市级	http://whdata.sd.gov.cn/
43	潍坊市公共数据开放网	山东省潍坊市	市级	http://wfdata.sd.gov.cn/
44	烟台市公共数据开放网	山东省烟台市	市级	http://ytdata.sd.gov.cn/
45	枣庄市公共数据开放网	山东省枣庄市	市级	http://zzdata.sd.gov.cn/
46	淄博市公共数据开放网	山东省淄博市	市级	http://zbdata.sd.gov.cn/

◆ 2.3 数据采集与分析方法

• 数据层

报告采用网络自动抓取和人工观察方法采集数据。以2018年4月13日至2018年4月18日为数据采集周期。报告主体评估分析部分基于截至2018年4月18日从各地平台上所采集的数据，而在对“动态更新”这一指标的评测中则使用了2018年1月1日至2018年4月18日这一时间段内所采集的数据。报告对采集到的有关各项指标的数据主要使用描述性统计分析、交叉分析、文本分析、空间分析等分析方法。

• 平台层

数据采集时间为2018年4月25日至2018年5月4日。本研究基于上述评估框架，主要采用了人工观察法来采集数据。报告对采集到的各项指标的数据进行了描述性统计分析和交叉分析。

由于武汉平台在数据采集期间内无法访问、宁夏平台需要通过当地电话号码注册才能下载数据等原因，这两地在数据容量、优质数据指标上的数据缺失。

需特别说明的是，山东省除青岛平台较早上线外，全省与省内其它地级城市均在2018年初同时推出数据开放平台。除山东省、济南市、青岛市的平台界面存在一定区别，山东省其他15地市平台完全相同，故报告中将这些地方平台简称为山东其他地市。

• 准备度

准备度的研究主要搜索了相关法律法规、政策文件、新闻报道等数据。搜索方法主要包括以下两种：一是在百度搜索引擎以关键词检索相关法规政策文本、有关领导支持与推广活动的新闻报道以及数据开放主管部门的数据；二是在地方政府门户网站、经（工）信委（局）网站以及数据开放平台上通过人工观察和关键词检索采集数据。数据采集截止时间为2018年5月6日。报告对采集到的数据进行了描述性统计分析。

◆ 2.4 指标计算方法

报告基于各地平台在三个维度的评估指标上的实际表现从低到高按照0-5分共6档分值进行评分，5为最高分，相应数据缺失或完全不符合标准则分值为0。此外，对于部分通过观察法采集数据的指标，根据实际满足评分依据的多少打分（以“开放授权”指标的分值为例，满足4项评分依据得5分，每满足一项得1.25分）。其余指标使用极差归一法将各地统计数据结果换算为0-5之间的数值作为该项得分。

各地平台在数据层、平台层、准备度三个维度上的指数总分等于每个单项指标的分值乘以相应的权重所得到的加权总和。最终，各地开放数林指数等于数据层指数、平台层指数、准备度指数乘以相应权重的加权平均分。

• 各地开放数林指数计算公式如下：

各地开放数林指数

$$= \sum (\text{数据层指标分值} \times \text{权重}) \times 50\% + \sum (\text{平台层指标分值} \times \text{权重}) \times 25\% + \sum (\text{准备度指标分值} \times \text{权重}) \times 25\%$$

3

主要发现 Key Findings



◆ 3.1 总体发现

• 3.1.1 上线时间

自2012年上半年上海市推出全国第一个政府数据开放平台起，截止到2018年上半年，我国已陆续上线40多个符合政府数据开放基本特征的地级市及以上平台(如图3所示)。特别是2017年5月第一期“开放数林指数”发布以来，我国新增了27个地方平台，仅2018年上半年就推出了21个新平台，其中山东省与省内各地市同时上线了政府数据开放平台。

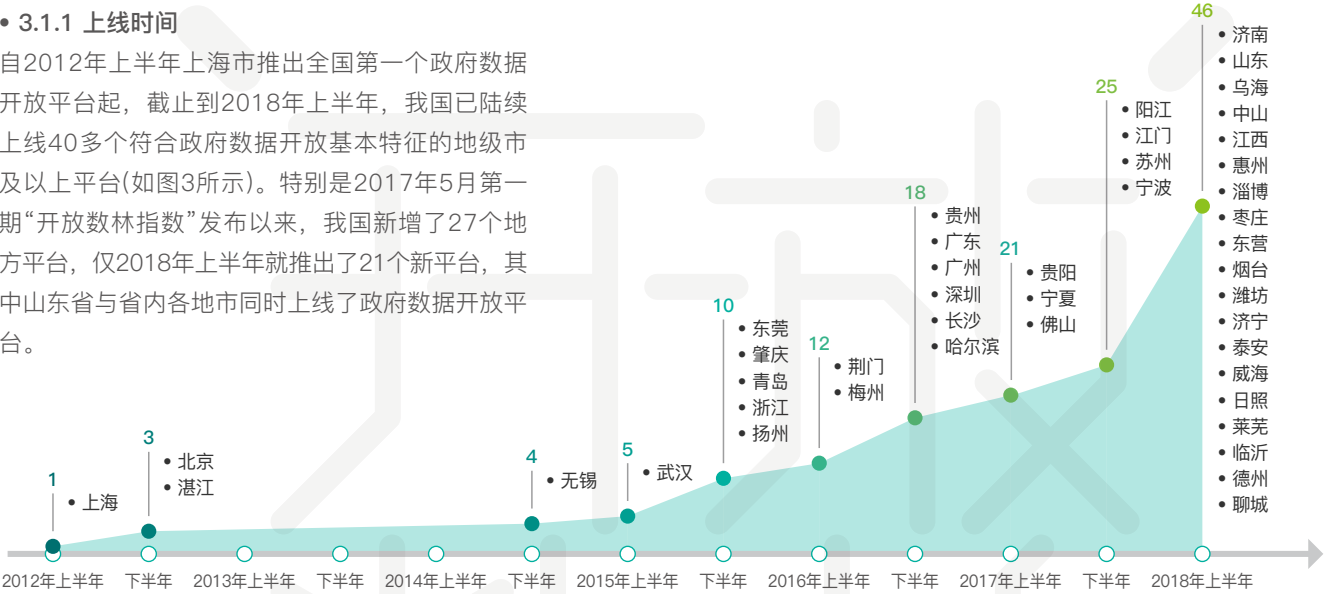


图3 各地平台的上线时间

• 3.1.2 地区分布

截止到2018年4月，我国各地上线的政府数据开放平台的地区分布如图4所示，颜色越深代表平台上线时间越早。从整体上看，我国的政府数据开放平台建设呈现出从东南沿海地区向内陆地区不断扩散的趋势。

目前，全国各地政府数据开放平台呈现“群落式”与“绿洲式”两种分布特征。其中，“群落式”以山东省、广东省和长三角地区最为典型。在这些区域中，地方政府“竞相”开放数据，尤其在广东省和山东省内，地级市政府数据开放平台分布较为集中，形成我国最为密集的省级“开放数林”。相比之下，由于周边地区尚未上线政府数据开放平台，北京、哈尔滨、宁夏等地则成为了所在区域政府数据开放的“绿洲”。

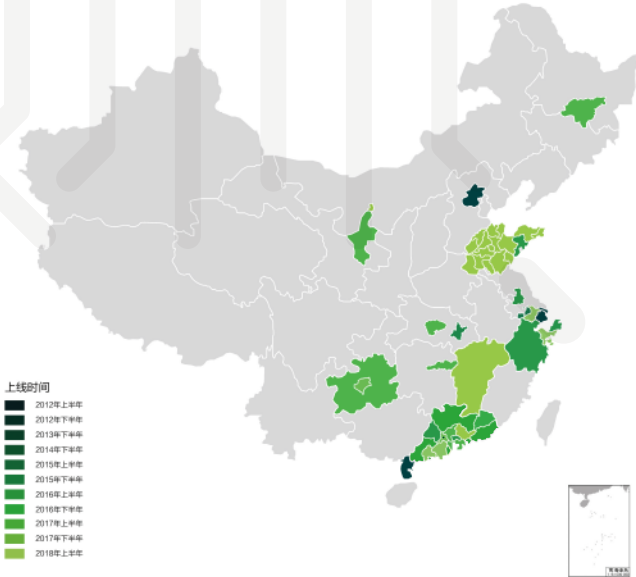


图4 各地平台的地理空间分布

• 3.1.3 行政层级

报告对各地数据开放平台所属的地方政府行政层级进行了统计分析(如图5所示)。目前在已上线平台的各级地方政府中,地市级数量最多(31个),省级行政区和副省级城市数量相当,分别为8个和7个。然而,全国共有31个省级行政区(港澳台除外)、15个副省级城市和334个地级行政区,除副省级城市平台占比接近5成外,我国总体上上线政府数据开放平台的地方所占比例依旧偏低。

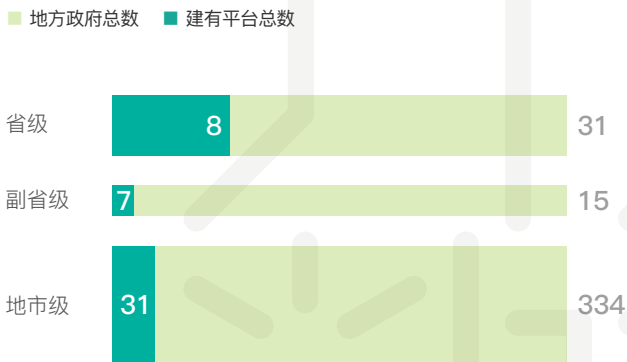


图5 各地平台的行政层级分布

研究还发现,一些区县级地方政府也开设了各具特色的政府数据开放平台。佛山市南海区作为国内最早开放数据的区级政府,其“数说南海”平台在国内政府数据开放领域进行了先行探索,具有示范意义。“凯里市政府数据开放平台”为贵州首个上线的区县级政府数据开放平台,开放传统村落、民族节日、苗族刺绣工艺品等少数民族特色数据,实现了数据开放与民族文化发展的结合。

• 3.1.4 平台类型

从各地政府数据开放平台的类型来看,绝大多数平台以“统一专有”的方式呈现,将数据统一汇聚在一个专门的平台上进行开放。而梅州、肇庆、浙江、阳江、荆门等则为“统一嵌入式”,即将数据统一汇聚为一个栏目板块,嵌入在政府

门户网站或政务服务网站上进行开放。政府数据开放平台与政府门户网站或政务服务网站在建设目的、服务对象和功能设置等方面有明显差异,“统一专有式”的政府数据开放平台更加便于用户集中发现、获取和利用数据,也是目前国际上普遍采用的方式。

◆ 3.2 数据层

• 3.2.1 数据数量

1. 数据集总量

数据集是由数据组成的集合,通常以表格形式出现,每一“列”代表一个特定变量,每一“行”则对应一个样本单位。政府数据开放平台往往以下载或API接口的形式开放数据集。个别平台上出现的以下三类情况,在此次报告中未被视作有效的开放数据集:(1)数据集名称下不存在可直接下载或通过接口获取的数据集;(2)数据集中仅有0-2行数据的低容量数据集(多为一个数据集分拆出的单行数据,或未整合成一个数据集的单行数据)。这类数据的再利用价值很低,适合以信息公开的方式发布,而没有必要作为特定数据集开放,因此不能视作有效数据集;(3)数据集名称下提供的是网页链接,且链接跳转后出现无法通过下载或接口形式获取的文本内容。

同时,报告将各地平台上出现的以下两类情况视作一个有效数据集:(1)同样名称的数据集重复出现;(2)平台上标注的数据集虽名称不同但实际下载后具有相同的名称与内容。

开放的有效数据集总量(含直接下载和API接口开放)最高的10个地方平台如图6所示。在46个政府数据开放平台中,贵阳开放的数据集最多,开放了超过2000个有效数据集与接口;其次是上海、青岛和武汉,开放的有效数据集均超过1000个。

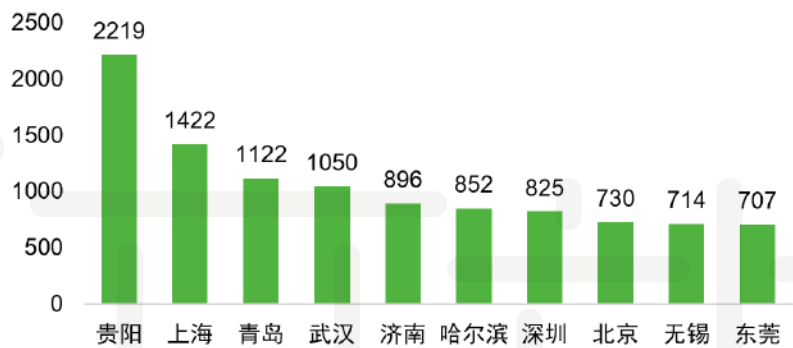


图6 各地平台上的数据集总量(前10名)

2. 数据容量

数据容量是指在各地平台可下载、结构化的数据集中，将字段数（列数）乘以条数（行数）得出的数据总量。作为新增指标，数据容量衡量的是平台上提供的数据集的实际

数据量大小。数据容量排名前10的地方平台如图7所示，其中广州最高，其次是佛山和贵阳，这三地开放的数据容量均超过8000万。有超过三分之一的平台开放的数据容量较小(10万及以下)，平台间数据容量的差距较为明显。

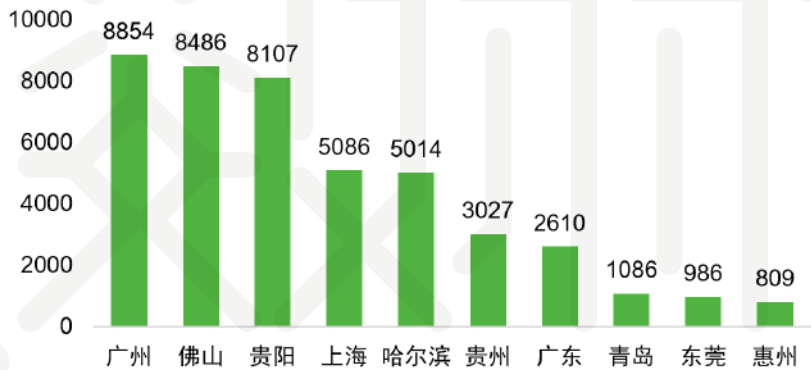


图7 各地平台上的开放数据容量(前10名) 单位：万

• 3.2.2 数据质量

1. 优质数据

优质数据指的是数据量大，社会需求高的数据集。报告对各地平台上所有可下载的数据集按照数据容量进行排序，在数据容量相同的情况下再按照下载量排序，最终选出排名居于前1%的数据集作为优质数据集。报告在46个平台上共发现了146个优质数据集，其分布状况如

图8所示。在46个政府数据开放平台中，17个平台有优质数据集入选，但仍有超半数的地方平台没有优质数据集。其中，拥有优质数据集最多的平台是贵阳，其次是贵州、哈尔滨和上海。

表4是排名前十位的优质数据集名称与其所属平台，这些数据集普遍具有较高的条数、字段数和下载量等，数据内容主要和商事主体、药品等相关。



图8 各地平台的优质数据集数量

表4 前十位优质数据集

No.	数据集名称	所属地方平台	数据容量	条数	字段数	下载量
1	工商登记信息	东莞	46416553	2018111	23	10063
2	商事主体个体年报基本信息	广州	34100000	1705000	20	524
3	商事主体基础信息	佛山	8259768	458876	18	365
4	自然人信息	佛山	6554376	1638594	4	168
5	黑龙江省统一药品信息	哈尔滨	6106135	174461	35	177
6	工程-投标人名单	贵州	4466102	235058	19	--
7	哈尔滨市个体基本信息	哈尔滨	4353930	483770	9	281
8	贵阳市城镇居民医疗保险药品目录	贵阳	4294512	238584	18	841
9	惠州市工商开业登记信息	惠州	3359715	223981	15	80
10	哈尔滨市商事主体个体年报基本信息	哈尔滨	3274398	545733	6	2757

2. 无低质数据

○ 无低容量数据

低容量数据是指条数在两行或两行以内的数据集，其原因可能是数据量本身稀少或是数据经统计归总后颗粒度

过大。这类低容量数据的再利用价值较低。在46个开放数据平台中，近三分之二的平台上存在低容量数据，未发现低容量数据的地方平台有江西、东莞、惠州等。

○ 无碎片化数据

碎片化数据是指按照时间、行政区划、政府部门等被人为分割的数据集，这些数据集进行整合后将更有利于社会的开发利用。报告发现，这一现象目前普遍存在，大部分的地方开放数据平台均存在碎片化数据，未发现碎片化数据的地方平台有东莞、佛山、泰安等。

3. 无问题数据

○ 无重复创建

重复创建是指平台上重复出现标题相同、可下载数据文件相同、且所属主题相同的数据集。在46个开放数据平台中，约三分之一的平台存在重复创建问题，未发现重复创建的地方平台包括贵阳、滨州、中山等。

○ 无生硬格式转化

生硬格式转化是指平台将非结构化的DOC、PDF等文件中的数据通过生硬方式转化成XLS、CSV等机读格式，而数据实质上仍是非结构化的情况。例如，将WORD文件中大段的文字贴到XLS文件中，将DOC格式直接转换成XLS格式等。在此次评估的平台中，有8个地方平台存在上述问题，其他平台尚未发现这一问题。

○ 无无效数据

无效数据是指平台上出现以下三类情况：(1)数据集名称下没有数据可供获取；(2)只提供数据链接，无法获取数据集；

(3)数据集下载打开后，里面实际上并不提供数据。在46个政府数据开放平台中，超过半数存在无效数据，未发现无效数据的地方平台包括贵阳、济南、深圳等。

• 3.2.3 数据标准

1. 开放授权

开放数据应通过数据开放授权协议从法律上保障数据的开放性。目前，各地平台上的数据开放授权通常包含在网站声明、免责条款或服务协议中。报告发现，在46个政府数据开放平台中，共有33个平台配有数据开放授权协议。目前，13个平台缺少专门的数据开放授权协议，其中有6个平台的网站声明与政府门户网站共享，并非专门针对数据开放平台制定，也没有专门涉及数据开放授权的内容；3个平台虽然有独立的免责条款或服务协议，但内容与数据开放授权无关；4个平台完全没有免责条款或服务协议。

开放授权协议的内容应明确授予用户免费获取、不受歧视、自由利用、自由传播与分享“开放数据”的权利。目前有5个地方平台的授权协议全部明确授予了上述四项权利，分别为北京、上海、贵州、广州、贵阳等；大部分地区满足了免费获取和不受歧视两项，这些地区以广东和山东及各自所辖的地级市为主，在开放授权协议的制定上具有趋同性；其余地区在四项指标上未明确提及，或语焉不详(见图9)。

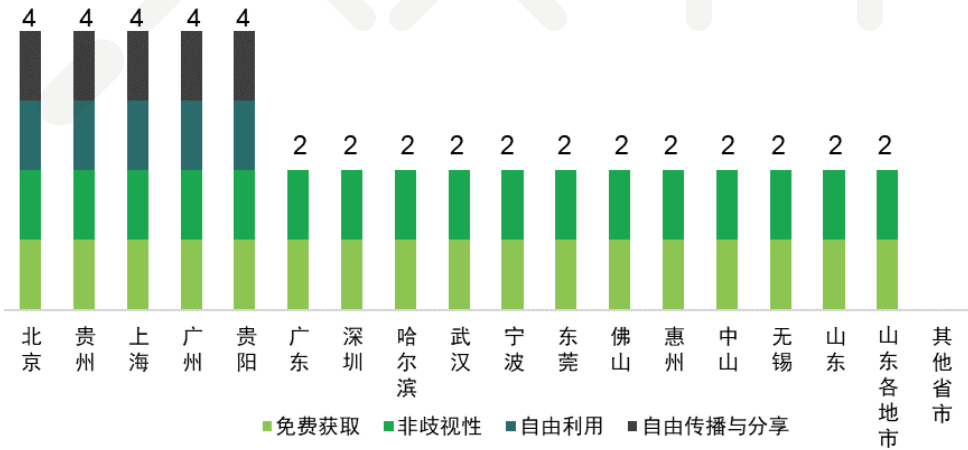


图9 各地平台的开放授权详情

○ 免费获取

免费获取是指平台在开放授权中明确授予用户免费获取和利用“开放数据”的权利。政府数据作为公共资源，原则上应免费向社会开放，除非需要对数据进行额外的增值加工和针对少数用户的个性化加工等。目前，各地平台上的相应条款分为“免费且未设时限”和“现阶段免费”两类。其中，贵州、贵阳、东莞等地的条款中明确指出数据免费且未设时限，用户可永久无偿获取数据平台所提供的所有数据资源。其他大部分地方平台则在服务协议中提到“现阶段免费”，但设置有模糊的期限或限制，如“保留收费权利”等表述(见图10)。

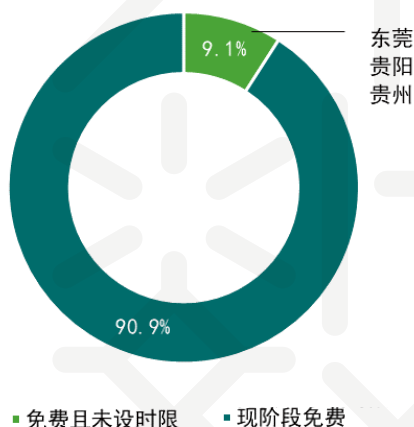


图10 开放数据免费获取比例

○ 非歧视性

开放授权是指平台明确授予任何用户平等访问、获取、使用和分享“开放数据”的权利。目前各地平台上的相应条款均明确保障了数据开放的非歧视性，对任何用户都予以平等的获取和利用权限，如“用户享有数据资源的非排他使用权”、“不受歧视”等表述。

○ 自由利用

开放授权应明确授予用户不受限制地对“开放数据”进行商业和非商业性利用的权利。目前各地平台上相应条款

分为“明确允许自由利用”和“未提及可自由利用”两类。

其中，北京、上海、贵州、广州、贵阳等地明确表示用户可“不受限制的进行商业和非商业性利用”、“享有增值利用的权利”或“可自由利用”，其余28个地方的条款中均未对用户利用数据的权利做出明确说明(见图11)。

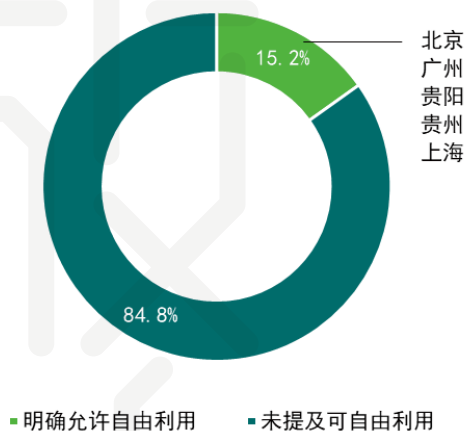


图11 开放数据自由利用比例

○ 自由传播与分享

开放授权应明确授予用户可自由传播和分享“开放数据”的权利。目前各地平台上的相应条款分为“可自由传播”、“未提及可自由传播”、“自由传播受限”三类。北京、广州、贵阳、贵州、上海等地授予用户享有免费传播现有开放数据的权利。为保障用户免费获取数据的权利，上海、广州、贵阳等地还特别说明了用户不得有偿转让获取的数据；东莞、哈尔滨、宁波等城市未对用户自由传播与分享的权限做出明确说明；其他大部分地方在使用条款中要求“用户不得有偿或无偿转让数据资源”，限制有偿转让数据有利于保障用户免费获取数据的权利，但限制用户无偿传播和分享已开放的数据并不符合开放数据的原则，在实际操作中也很难发现和限制(见图12)。

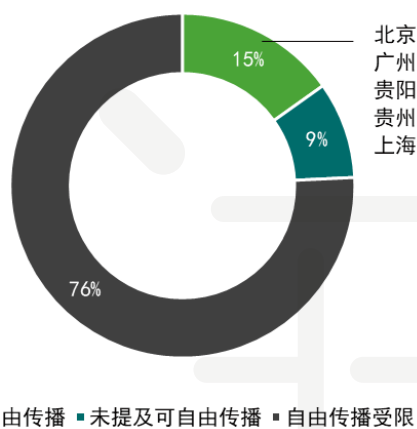


图12 开放数据自由传播与分享比例

2. 技术性开放

关于开放数据的格式标准，万维网发明者Tim Berners-Lee提出了开放数据的五星标准：一星：基于开放授权在网络上开放数据，对数据格式不做要求，可能采用PDF、JPEG等格式；二星：以可机读、结构化格式开放数据(如采用

EXCEL格式而非扫描格式)；三星：在满足二星标准的基础上，以非专属开放格式开放数据（如采用CSV格式代替EXCEL格式)；四星：在满足以上要求的基础上，为每一个数据集设置固定的URL链接，便于使用者发现和链接到数据集的具体位置；五星：在满足以上要求的基础上，数据之间实现关联，以展现数据的背景。本报告基于该标准和其他有关开放数据格式的标准，对各地政府数据开放平台上的数据集的格式标准进行评估。

○ 可机读格式

为方便用户获取和利用数据，数据集应以可机读格式开放，该格式能被计算机自动读取与处理，如XLS、CSV、JSON、XML等格式。图13展示了各地平台上可下载数据集总量与可机读数据集总量对比的前10名。总体来看，各地平台上满足可机读格式的数据集比例较高，贵阳、青岛、上海等38个平台上开放的数据集基本都满足了可机读格式的要求。但也有个别地方平台虽然可下载的数据集总量不小，但其中不少数据集为DOC、PDF、JPG等不可机读格式。

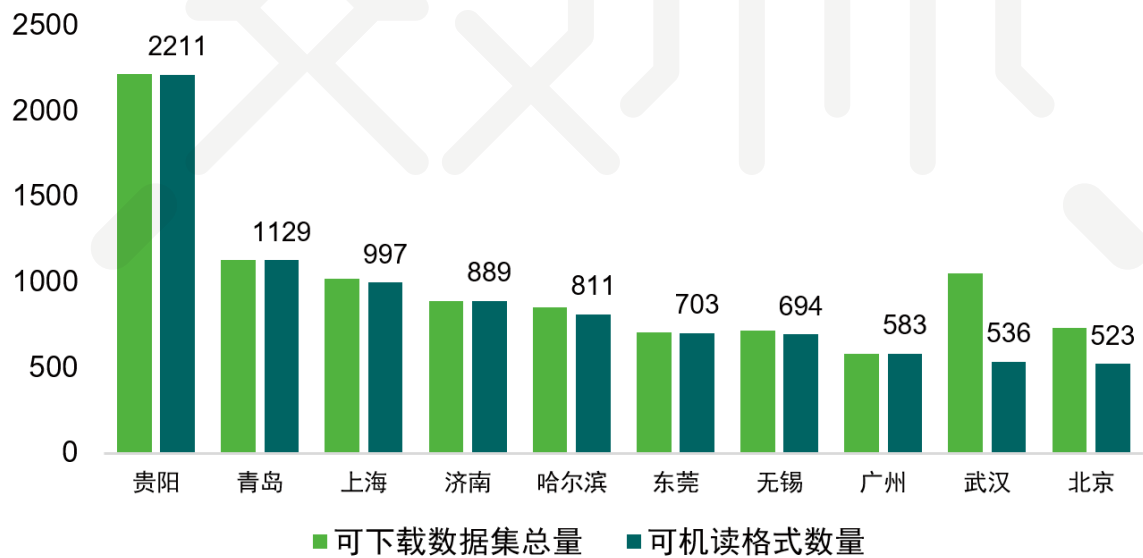


图13 各地平台可下载数据集与可机读格式数据集数量（前10名）

○ 开放格式

开放格式是指可下载数据集应以开放的、非专属的格式提供，任何实体不得在格式上排除他人使用数据的权利，以确保数据无需通过某个特定(特别是收费的)软件或应用程序才能访问。例如CSV是开放格式，而XLS则不是。

图14是各地方平台上可下载数据集总量与开放格式总量对比的前10名，贵阳平台上开放格式的数据集数量最多，且所有数据集均可通过开放格式下载。济南、东莞、广州、山东等23个地方平台上提供的数据集也全部满足开放格式的标准。此外，有部分地方平台没有提供任何开放格式的数据集，而多以XLS等非开放格式为主。

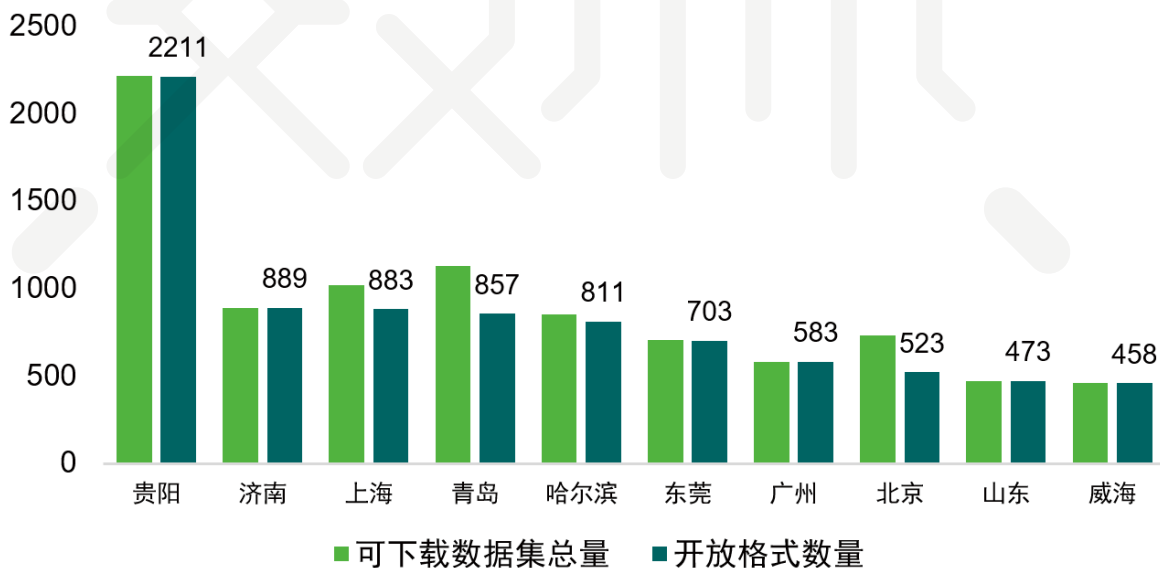


图14 各地平台可下载数据集与开放格式数据集数量(前10名)

○ 固定链接

依据开放数据五星标准，本报告新增了对固定链接格式的评估，即四星标准的要求。固定链接是指为数据集设置固定的URL链接，有助于数据使用者发现和链接到数据集的具体位置。在46个开放平台上，仅有贵阳提供了符合固定链接格式的数据集，共提供216个RDF格式的数据集。

○ API接口比例

除了数据直接下载以外，有些地方平台还提供了应用程序编程接口（即API接口，Application Programming Interface）。与通过直接下载方式获取数据集相比，接口方式使用用户可通过参数实时高效地获取所需数据，满足其开发应用程序的需求，适合用于开放实时性强、规模大的数据。目前，北京、深圳、东莞等16个地方平台为每个数据集提供了接口。然而，仍有近三分之一的平台没有提供或仅提供了少量的API接口，有个别地方平台上的API接口处于无法被调用的状态。

3. 元数据完整性

政府数据开放平台为开放数据集提供元数据有助于数据

利用者清楚地了解数据集的内容与背景，从而更好地获取和利用数据。本报告对各地平台上可下载数据集所提供的元数据进行了分析。

○ 基本元数据覆盖率

综合梳理我国政务信息资源编目指南中关于核心元数据的定义描述、国际开放数据平台上提供的基本元数据条目以及目前我国半数以上的平台已实际提供的元数据条目，报告确定了以下12个条目作为开放数据集基本的元数据条目，包括数据名称、摘要简介、标题关键字、数据主题、数据格式、开放属性、提供单位、发布日期、更新日期、更新频率、数据指标、数据量。其中数据指标是指数据集中每一列字段的具体名称，有助于利用者更好地了解数据结构与内容。

图15反映出以上12个基本元数据条目在46个地方平台的分布情况。目前，所有46个地方平台都已提供了数据集名称，大多数平台提供了摘要简介、数据提供单位、发布日期、数据主题、数据格式等。然而，目前能提供更新频率、数据量、数据指标的地方平台还相对较少。

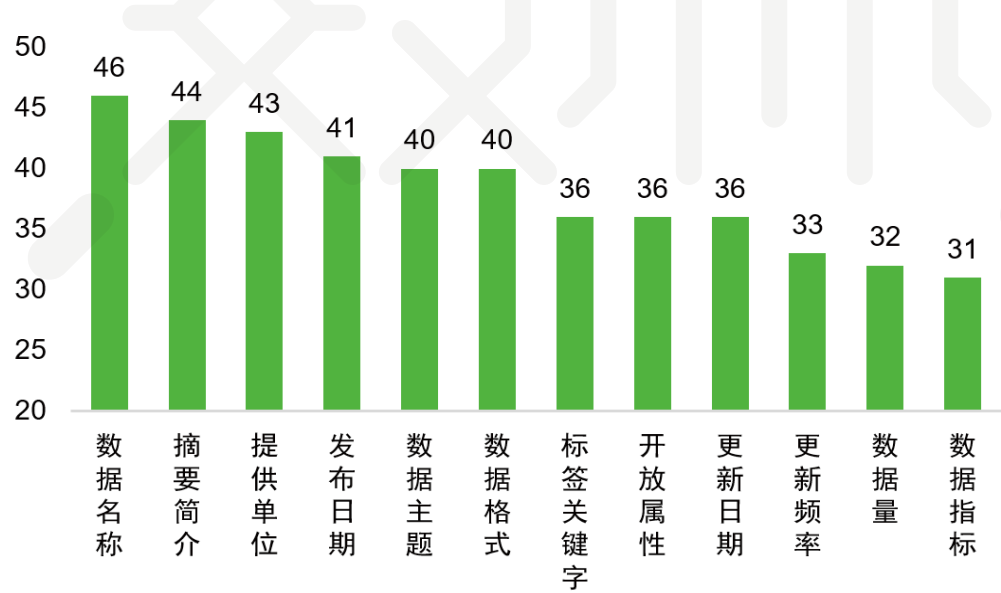


图15 基本元数据条目在各地平台的分布数

目前，佛山、哈尔滨、惠州、江门、宁波、宁夏、深圳、乌海、山东省、山东各地市平台等提供了所有的基本元数据，此外还有3个地方平台覆盖了11个基本元数据条目，但仍有个别地方平台基本元数据覆盖较少。各地平台提供

的基本元数据条目详情可参见表5，无圆点的空格即为该地方缺少的元数据条目，不同颜色的圆点代表不同的元数据条目。

表5 各地平台基本元数据条目详情列表

数据名称	摘要简介	标签/关键字	数据主题	提供单位	数据格式	类型/属性	更新频率	发布日期	更新日期	数据量	数据指标
佛山	●	●	●	●	●	●	●	●	●	●	●
哈尔滨	●	●	●	●	●	●	●	●	●	●	●
惠州	●	●	●	●	●	●	●	●	●	●	●
江门	●	●	●	●	●	●	●	●	●	●	●
宁波	●	●	●	●	●	●	●	●	●	●	●
宁夏	●	●	●	●	●	●	●	●	●	●	●
深圳	●	●	●	●	●	●	●	●	●	●	●
乌海	●	●	●	●	●	●	●	●	●	●	●
山东	●	●	●	●	●	●	●	●	●	●	●
山东各地市	●	●	●	●	●	●	●	●	●	●	●
广东	●	●	●	●	●	●	●	●	●		●
上海	●	●	●	●	●	●	●	●	●		●
广州	●	●	●	●	●	●	●	●	●	●	●
贵阳	●	●	●	●	●	●	●	●	●		
东莞	●	●	●	●	●	●	●	●	●	●	●
武汉	●	●	●	●	●	●	●	●	●	●	
中山	●	●	●	●	●	●	●	●	●	●	
北京	●	●	●	●	●	●		●	●		
肇庆	●	●	●	●	●	●	●	●	●		
湛江	●	●	●	●	●	●		●			
贵州	●	●	●	●	●	●		●	●		
苏州	●	●	●	●	●	●		●			
长沙	●	●	●	●	●	●			●	●	●
荆门	●	●	●	●	●	●		●			
江西	●	●		●	●			●		●	
扬州	●	●		●		●		●			
梅州	●	●		●	●			●	●		
浙江	●	●						●			
阳江	●							●			
无锡	●				●				●		

○ API描述规范

API描述有助于数据利用者清楚地了解API的具体信息及获取方式，从而更好地调用接口并获取数据。本报告从数据资源描述和数据调用说明两方面评估API描述情况。资源描述是指API的基本信息，如名称、简介、提供部门、更新时

间等；数据调用说明指的是API的调用方式、请求地址等信息。如图16所示，在提供API接口的36个地方平台中，33个地方平台均提供了资源描述和数据调用说明，个别平台仅提供一项，也有平台两项均不提供。

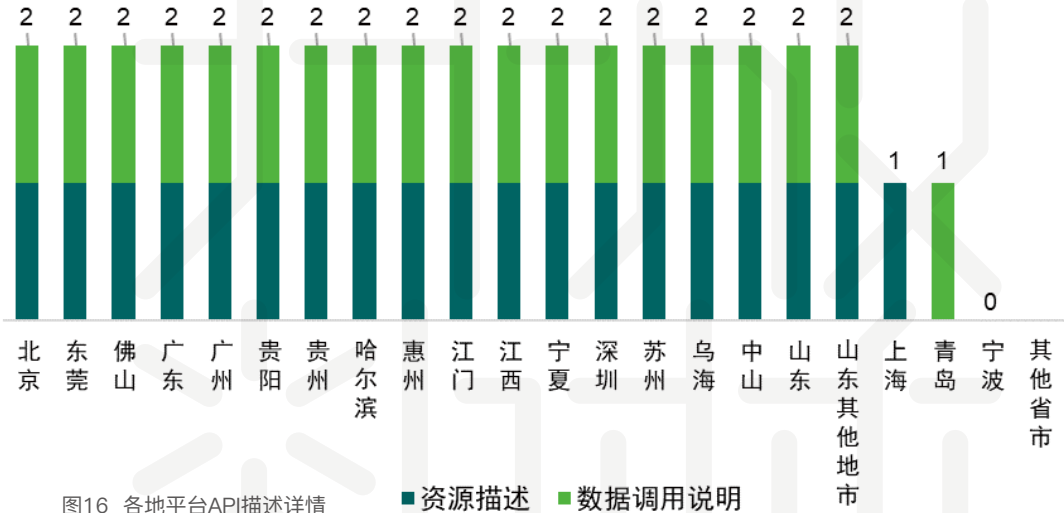


图16 各地平台API描述详情

• 3.2.4 数据覆盖面

1. 主题覆盖率

开放各个主题领域的数据集有利于提高数据开放的广度和覆盖，有利于数据利用者充分获取和整合来自多种领域的数

据，并进行深度挖掘利用。2015年8月国务院印发的《促进大数据发展行动纲要》提出要“优先推动信用、交通、医疗、卫生、就业、社保、地理、文化、教育、科技、资源、农业、环境、安监、金融、质量、统计、气象、海洋、企业登记监管等民生保障服务相关领域的政府数据集向社会开放”。2018年1月，中央网信办、国家发改委、工信部联合印发的《公共信息资源开放试点工作方案》要求“重点开放信用服务、医疗卫生、社保就业、公共安全、城建住房、交通运输、教育文化、科技创新、资源能源、生态环境、

工业农业、商贸流通、财税金融、安全生产、市场监管、社会救助、法律服务、生活服务、气象服务、地理空间、机构团体等领域的公共信息资源”。

基于以上文件的要求，结合国内外数据开放平台上实际开放的重点数据集主题，报告将开放数据主题归纳为经贸工商、交通出行、机构团体、文化休闲、卫生健康、教育科技、社会民生、资源环境、城建住房、公共安全、农业农村、社保就业、财税金融、信用服务共14个大类。

图17体现了各地平台在14个主题下所开放的数据集个数。其中，社会民生、经贸工商、教育科技等主题的数据集开放数量最多，数据开放相对较少的主题为城建住房、社保就业、信用服务等。

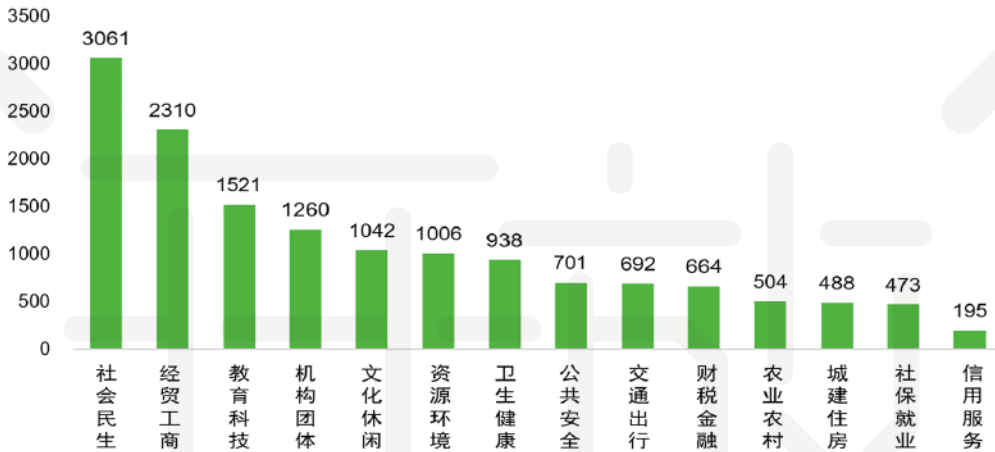


图17 各主题包含的数据集个数

报告进一步分析了各地方平台开放数据集的主题领域覆盖面，图18是覆盖主题个数较多的地方平台。在14个主题领域中，不同地方平台的主题覆盖情况差异明显。其中，广

州和青岛开放的数据集主题覆盖率最高，覆盖了全部的14个主题；其次为贵阳、哈尔滨、深圳、宁夏和宁波等地。

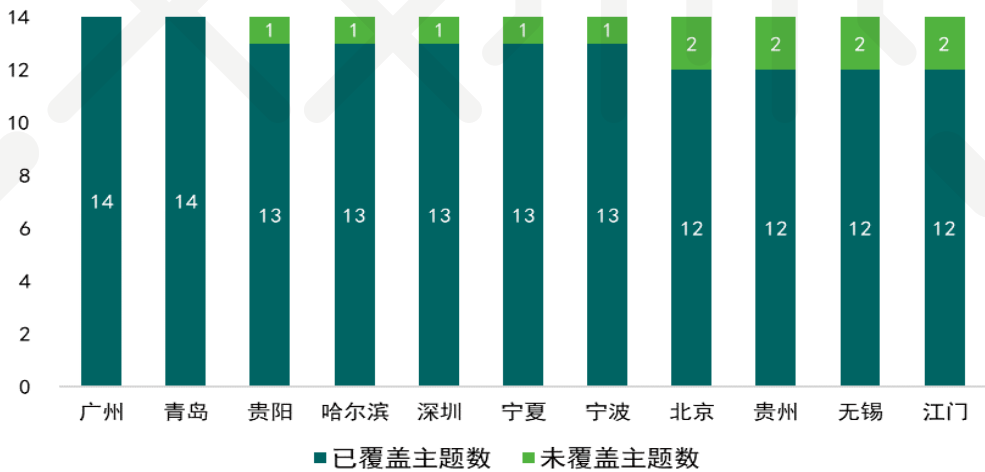


图18 各地平台主题覆盖数(前10名)

2. 部门覆盖率

开放数据集的部门覆盖率反映了一个地方政府的各个部门对于数据开放工作的参与程度，也体现了数据集来源的丰富程度。报告首先梳理了地方平台开放数据集覆盖较多的部门作为“主要数据提供部门”，由于不同地方具有相同职能的部门在名称上存在差异，报告对这些部门的名称做了相应的合并与调整。

图19为各地平台上主要数据提供部门所开放的数据集数量。总体而言，来自统计局的数据集数量最多，远高于其后的

教委、卫计委、交通委、住建委、公安局、人社局等部门。与业务部门开放的数据集相比，统计部门提供的数据多为经过加工归总后的宏观数据，颗粒度较大、数据容量较低，不利于数据被利用和产生价值。如图20所示，统计局数据所占比例较高的地方数据容量也普遍较低。

目前，提供数据集数量相对较少的是档案局、国资委、审计局等部门。图中不同色条的纵向宽度反映了各地平台上不同数据提供部门的数据集数量，如北京提供最多数据的部门为统计局，教育局与卫计委开放较少。

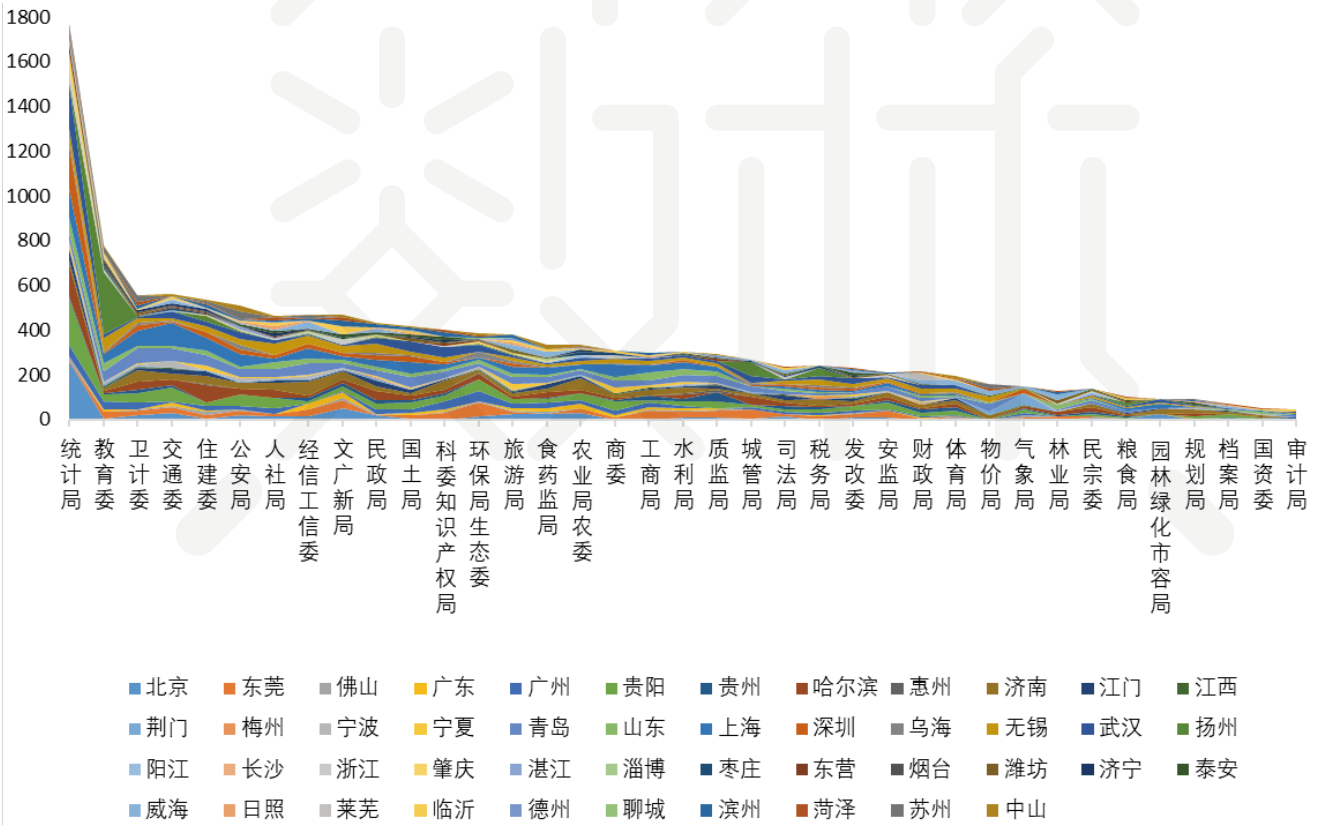


图19 各地平台主要数据提供部门提供的数据集数量分布

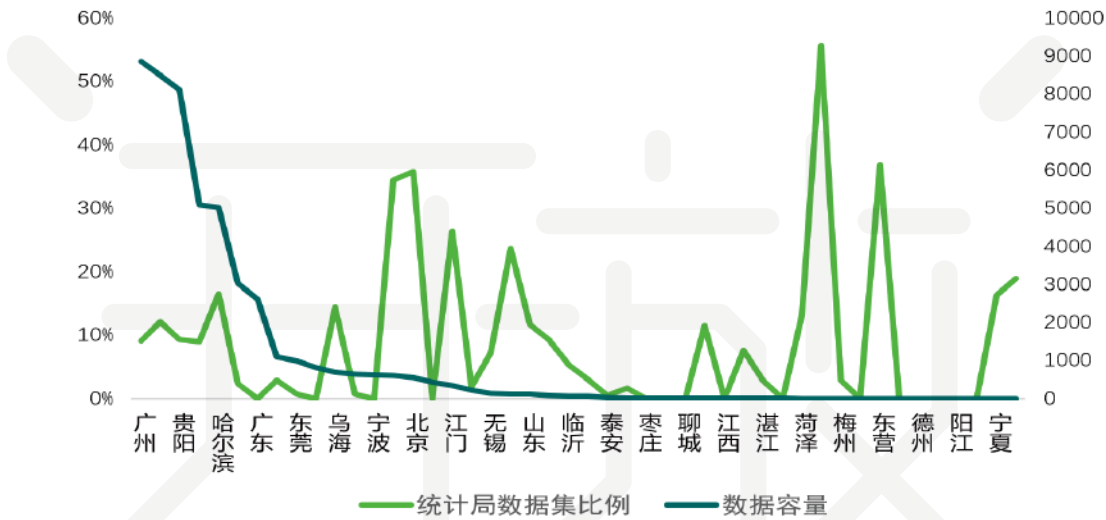


图20 各地数据容量与统计局比例 (数据容量单位：万)

同时，报告还分析了各地平台数据提供部门数量占该地方政府行政职能部门总数（数据来自各地政府门户网站）比例的前10名，以评估该地方数据开放部门的覆盖面。结果如图21所示，贵阳、广州、东莞、济南、宁波和山东的部门

覆盖率达到100%，此后是青岛、哈尔滨、武汉和上海等，部门覆盖率均超过了90%。但仍有部分地方平台的部门覆盖率不到一半，尚有较大提升空间。

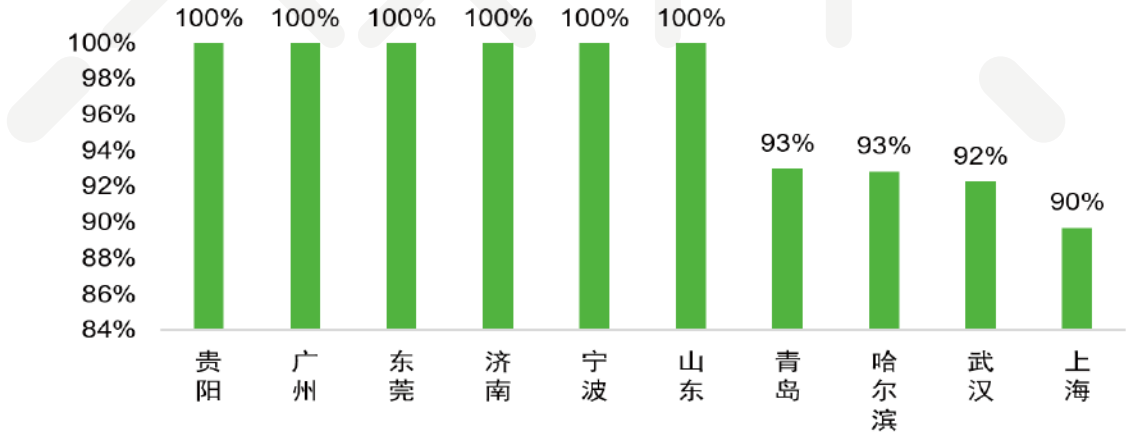


图21 数据提供部门占部门总数的覆盖率 (前10名)

3. 高需求关键词覆盖率

报告对各地平台上可获取、且下载量最高的前20%的数据集名称进行了文本分析，发现了一批出现频次较高的关键词，这些关键词反映了各地开放的高需求数据集的内容及其分布。图22为各地高需求数据集名称中高频出现的描述

性限定词，反映了高需求数据集的主要领域和内容。由图可见，“许可”、“建设”和“生产”出现的频次最高，其后“行政”、“经营”、“审批”、“医疗”、“教育”、“旅游”等词也有较高的出现频次。



图22 各地高需求开放数据集名称中出现频次最高的描述性限定词

报告将各地开放的高需求数据集名称中高频次出现的50个描述性限定词确定为“高需求关键词”，分析各地平台数据集名称在这些高需求关键词上的覆盖率。如图23所示，覆

盖率最高的是贵阳和武汉，达到了100%的覆盖率；其次是上海、青岛、哈尔滨、北京、济南等地，覆盖率也均达到98%；还有少数城市覆盖率较低，不到三成。

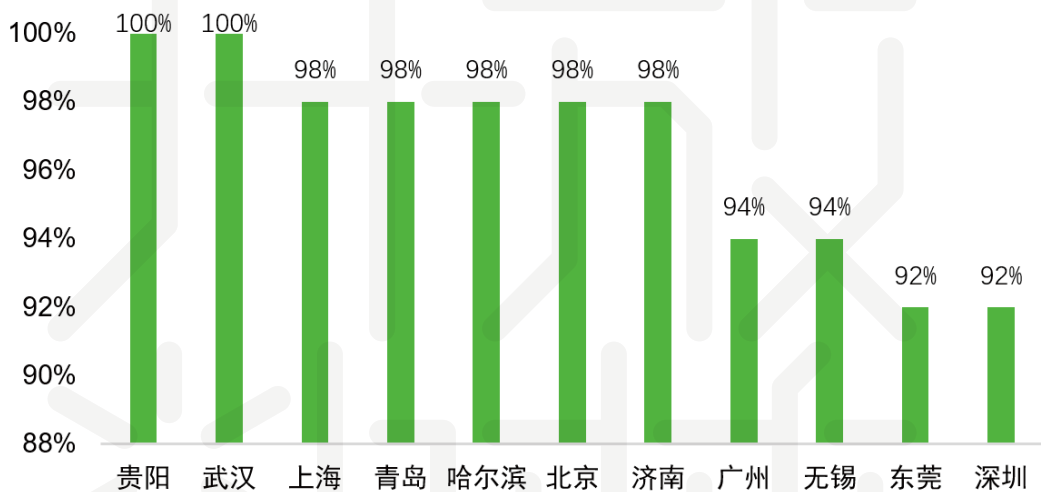


图23 各地开放数据集名称高需求关键词覆盖率（前10名）

覆盖率最高的前20个地方平台详情如图24所示，数字代表覆盖到关键词的数据集数量，所在单元格颜色越深，代表涉及到该关键词的数据集越多，颜色越浅则代表该项数据集越少。

	贵阳	武汉	上海	青岛	哈尔滨	北京	济南	广州	无锡	东莞	深圳	威海	山东	滨州	广东	江门	临沂	贵州	宁波	菏泽
许可	74	9	52	33	41	23	41	17	18	127	85	8	38	11	46	47	11	6	1	4
建设	47	29	28	31	44	6	14	14	10	40	15	23	16	19	20	14	5	6	4	5
生产	68	23	39	32	40	52	14	9	21	28	15	20	30	13	14	28	6	7	5	8
行政	58	18	17	16	24	12	27	23	15	56	16	19	10	3	29	22	8	16	5	6
经营	57	20	20	42	22	30	19	13	14	52	28	16	8	13	14	20	10	6	2	6
审批	19	14	34	12	31	3	10	10	8	86	8	5	10	1	7	6	4	3	5	3
医疗	29	27	22	27	13	18	21	21	8	6	6	8	11	4	4	15	5	9	3	7
教育	50	17	19	18	5	8	14	9	27	10	5	7	27	2	2	3	9	2	15	1
旅游	31	8	21	26	9	12	10	9	11	5	3	12	10	5	10	2	3	3	4	1
工业	32	29	19	17	7	20	8	17	12	5	21	6	12	4	1	6	1	44	2	5
经济	38	34	19	7	15	23	1	10	5	3	15	5	7	1	0	3	0	3	5	1
投资	35	21	22	8	14	14	2	1	16	10	31	3	11	3	9	8	1	1	3	5
卫生	52	14	14	15	13	10	9	7	12	3	11	0	11	4	0	0	1	14	5	1
学校	24	14	7	12	19	7	7	8	10	13	19	7	6	2	2	1	5	1	1	2
交通	19	19	12	9	6	10	8	3	3	6	3	1	5	1	2	2	1	3	10	1
零售	30	13	16	9	6	12	6	5	4	13	8	2	3	0	4	4	0	0	0	4
消费	40	16	25	6	11	26	4	2	7	1	5	1	8	0	0	2	0	0	0	2
安全	34	20	14	19	7	2	19	13	24	17	5	9	11	5	3	5	4	8	6	3
道路	21	9	10	13	10	1	14	0	6	7	3	4	1	1	3	0	1	0	8	0
社区	23	5	9	9	6	24	5	6	8	0	3	2	2	3	0	0	2	0	6	1
资质	20	6	19	10	4	5	13	4	13	21	3	1	11	1	5	11	1	3	1	1
文化	26	11	5	15	4	12	8	5	11	5	4	2	9	5	10	2	5	4	4	4
机动	39	6	5	15	7	12	4	5	3	5	3	0	0	4	4	0	0	2	0	0
运输	21	10	7	10	10	3	10	1	2	12	6	3	6	1	4	8	5	2	2	2
机动车	38	6	5	15	7	12	4	5	3	5	3	0	0	4	4	0	0	2	0	0
建筑	9	14	16	16	40	6	10	5	8	13	18	0	4	1	2	4	1	1	1	0
药品	17	4	9	10	6	8	5	6	10	2	1	6	8	2	10	4	3	6	1	3
职业	25	6	12	7	10	5	8	13	18	3	4	1	9	3	2	2	2	3	3	1
资源	18	9	7	12	7	7	8	13	8	8	3	2	10	1	2	2	1	7	5	0
市场	28	11	5	13	5	7	6	8	9	5	3	3	1	3	7	2	1	1	3	0
培训	20	4	5	9	6	3	12	4	10	6	1	2	7	2	2	2	2	3	4	2
体育	25	12	4	7	7	8	6	7	8	4	11	13	5	0	0	2	6	12	3	1
食品	13	7	13	17	12	12	2	10	6	4	0	21	3	1	4	9	5	0	3	1
建设项目	16	7	8	10	15	1	6	4	1	15	5	4	5	8	3	2	2	1	0	2
医疗机构	10	8	16	13	4	6	13	9	0	5	5	3	2	2	1	11	2	1	1	3
行政许可	10	4	0	3	5	3	8	4	5	18	3	2	0	0	23	9	0	2	0	0
处罚	10	1	8	4	5	5	10	8	1	38	14	4	3	9	4	9	5	0	1	3
企业名单	20	10	4	18	2	7	23	13	27	8	1	5	12	2	6	4	4	36	6	0
用地	17	15	10	4	13	0	1	3	2	12	0	1	1	2	6	4	1	1	0	1
交易	26	7	6	6	2	5	5	2	0	1	8	1	0	2	2	0	0	15	3	0
行政处罚	7	1	8	4	4	5	8	6	1	29	11	3	2	2	4	9	5	0	1	3
生产总值	6	4	17	3	9	13	1	2	4	0	5	2	3	3	0	7	1	0	0	2
监测	26	24	7	13	4	2	6	5	4	3	8	1	4	1	3	0	4	5	3	1
建设工程	5	1	13	6	13	2	4	1	1	11	9	1	2	6	0	6	0	0	1	2
农业	9	11	5	8	6	15	19	9	8	5	6	5	1	3	3	0	7	8	1	3
客运	13	4	5	3	8	2	6	2	1	2	1	1	0	0	2	2	2	1	2	0
普查	3	2	15	0	0	1	0	0	0	0	0	0	0	0	0	0	0	6	0	1
农村	37	25	5	3	7	16	7	3	4	0	0	2	4	0	1	1	0	3	0	0
从业	14	6	6	3	4	4	9	0	6	8	6	3	2	1	2	3	1	1	0	0
施工	5	6	8	7	7	7	5	3	2	16	8	0	4	0	0	6	1	0	0	0

图24 各地开放数据集名称高需求关键词覆盖详情（前20名）

• 3.2.5 数据持续性

1. 持续增长

政府数据开放平台管理运维是一项持续性和常态化工作，平台上线后应持续更新和添加数据集，以满足社会对开放数据日益增长的需求。报告根据各地平台上数据集的创建日期来判断该平台数据集是否持续增长，以季度为时段进行跟踪分析，指标分值取决于该平台数据集保持持续增长的时段的数量。

图25显示的是从平台上线开始，以季度为观测时段的全国各地平台数据集持续增长情况，颜色区域表示该时段有新增数据集，空白区域则表示该时段无新增数据集。上海市平台从2012第四季度发布数据开始，至今数据已保持了15个时段的持续增长，而北京平台持续增长时段为11个。2016年以来，各地密集上线新的平台，这些新上线的平台大都能保持数据集定期增长。

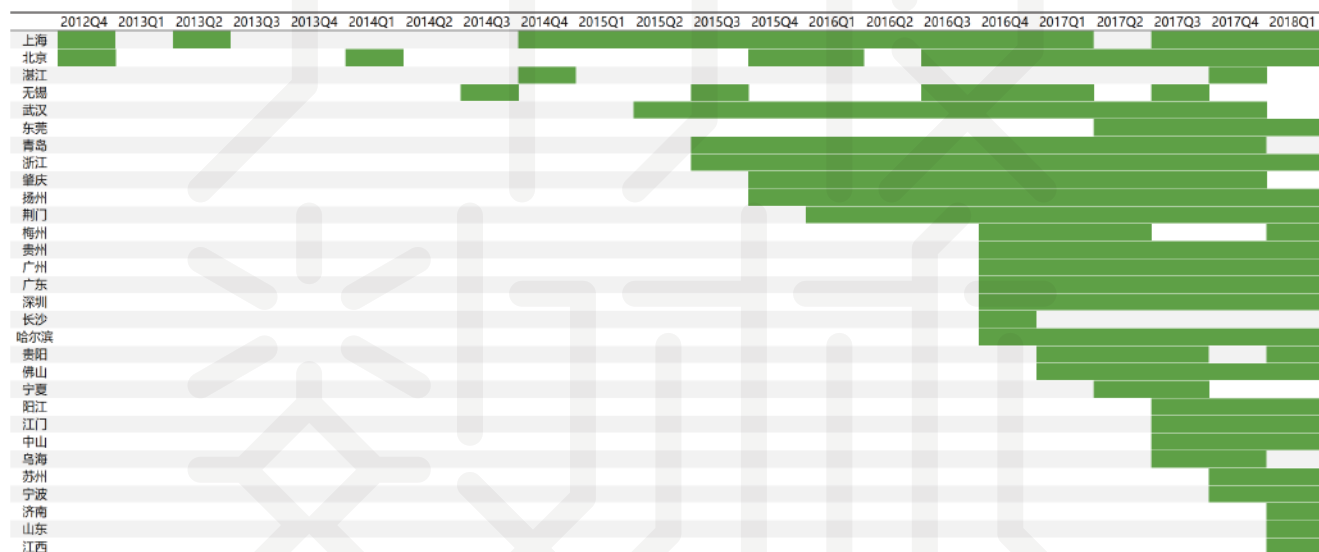


图25 各地平台上线时间与数据集持续增长情况

2. 动态更新

报告跟踪考察地方数据平台在2018年1月至2018年4月期间更新的数据集数量。各地方平台动态更新的情况如图26所示，贵阳平台上该时段内更新的数据集数量最多，超过1000个；其次是威海、贵州、东莞和滨州等，动态更新的数据集超过200个；山东省内各城市平台于2018年4月上线后更新活跃度也较高。然而，在此时间段内，仍有少数地方平台完全没有更新数据。

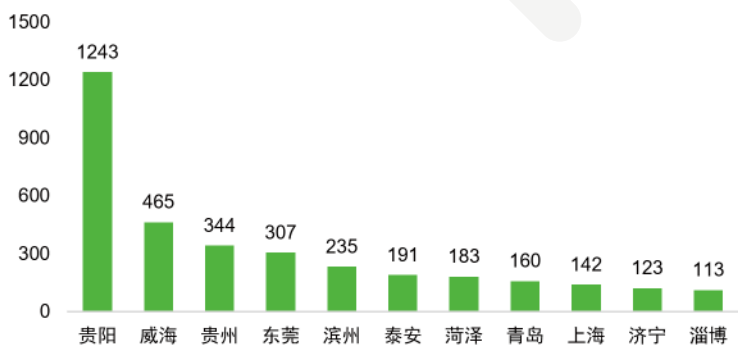


图26 各地平台数据动态更新情况（前10名）

3. 历史存档

历史存档是指平台将历史上不同时间更新的多个批次的数据同时留存在平台上供用户下载，有利于数据利用者按时间线索来获取和利用历史数据。图27反映了数据历史存档的平台分布情况。目前有上海、广东、广州等15个地方平台实现了数据历史存档，较上年在数量和比例上都有较大幅度的提升。

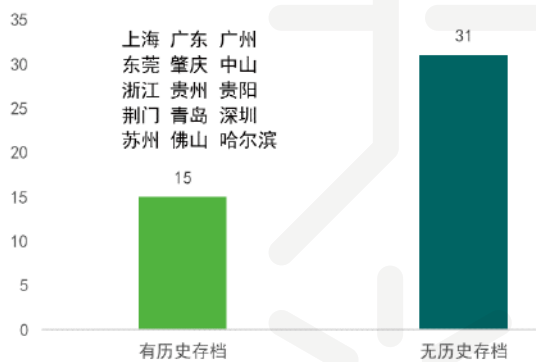


图27 实现数据历史存档的平台分布

• 3.2.6 数据层结论

1. 数据集总量稳步增长，数据容量整体偏低

目前各地平台开放的数据集数量已成规模，贵阳、上海、青岛、武汉等地的数据集总数已突破1000个，其他地区的开放数据集数量相比去年也稳步增长。然而，与此同时，仍有20多个地方的开放数据集总数偏少，不足200个。同时，需要引起关注的是，各地开放的数据集数据容量整体偏低，只有广州、佛山、贵阳等地开放的数据集数据容量较大，利用价值较高，多数地方开放的数据集行数与列数过少，数据容量过低，无法被有效利用。

2. 优质数据初现萌芽，数据质量良莠不齐

各地上线了一批高容量、高需求的优质数据集，尤以贵阳、贵州、哈尔滨、上海等地为代表。然而，大多数地方的数据集质量偏低，体现在多为数据量稀少或大颗粒度的低容

量数据、人为分割的碎片化数据等方面。此外，各地普遍还存在重复创建、生硬格式转化和无效的问题数据等现象。这些低质量数据和问题数据很难被再次利用并产生相应价值。

3. 数据标准逐步规范，元数据完整性参差不齐

在数据的法律性开放上，北京、上海、贵州、广州、贵阳等五个地方平台的授权协议中明确授予了用户免费获取、不受歧视、自由利用、自由传播和分享数据的权利，但许多地方的政府数据开放平台仍未提供明确充分的数据开放授权。

在数据的技术性开放上，各地平台上可机读、非专属、以接口形式提供的数据集比例相比去年稳步增长。贵阳市在全国率先提供了RDF格式的数据集，为开放数据集设置了固定的URL链接，便于使用者发现和链接到数据集的具体位置。然而，全国各地仍然存在很多不符合开放数据格式标准的数据集，如PDF或DOC格式的不可机读数据、XLS格式的非开放格式数据等。

在元数据完整性上，目前大多数地方平台都能提供基本的元数据，这有助于用户理解和应用数据。然而，各地情况参差不齐，在提供发布时间、更新时间、数据指标和数据量等条目上仍有待进一步完善。

4. 数据覆盖面尚不均衡，统计数据数量最多但价值偏低

在主题覆盖率上，目前各地平台上提供最多的数据集主题是社会民生和经贸工商，提供较少的主题是城建住房、社保就业和信用服务。不同地方平台的主题覆盖情况不一，广州和青岛的开放数据集主题覆盖率最高。

在部门覆盖率上，与去年相比，各地平台上开放数据集最多的部门仍是统计局，远高于其后的教委、卫计委、交通委等业务部门。来自统计部门的数据多为二手的、经过加工归总的数据集，其再利用价值低于来自业务部门的、一手、原始的数据集，各地对于开放数据的理解仍存在偏

差。贵阳、广州、东莞、济南、宁波和山东等地的政府部门对数据开放的参与程度最高，几乎所有的部门都开放了数据，但全国仍有接近半数的地方政府其部门参与度不到一半。

在关键词覆盖率上，各地开放的高需求数据集名称中出现“许可”、“建设”和“生产”的频次最高，高需求关键词覆盖率最高的地方是贵阳和武汉，实现了全部覆盖，然而还有少数城市的覆盖率不到三成。

5. 数据可持续性略有进步，后续运维仍是短板

与去年相比，有更多的地方能基本保持新增数据集持续增长与存量数据集动态更新，但仍有不少地方平台出现数据集增长间歇或停滞，而真正实现存量数据集动态更新的比例仍然偏低。此外，仅有不到一半的地方将历史上多个批次的数据留存在平台上供用户获取。总体而言，多数地方平台在上线之后对后续运维与持续更新的重视程度不够。

◆ 3.3 平台层

• 3.3.1 平台概览

政府数据开放平台以醒目方式展示数据统计、数据动态、提供数据开放最新资讯并进行可视化呈现，能够帮助用户了解政府数据开放平台基本概况，为用户进一步发现和获取目标数据集提供方向。

1. 数据统计

数据统计主要是指政府数据开放平台呈现平台总体、分类和个体层面的数据集统计概况。目前，各地政府数据开放平台上提供的数据统计条目种类各异，青岛、中山、江门等平台提供了较为全面的数据统计条目，其中提供条目数量最多的地方为青岛和中山，达到11条。图28展示了排名前10的地方平台的数据统计条目数情况。

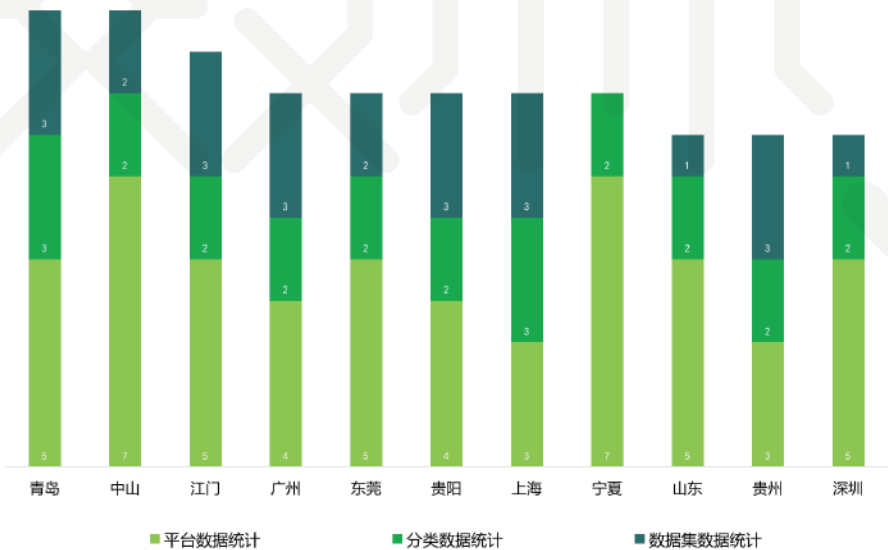


图28 各地平台数据统计条目数 (前10)

2. 动态展示

政府数据开放平台在首页或数据目录页醒目位置展示数据更新和下载动态，可提示用户有关数据集、API接口和APP应用的最新情况。基于国内外政府数据开放平台现状，本报告梳理出了8种动态展示条目，包括最新数据集展示、热门数据集展示(如下载最多、浏览最多等)、好评数据集展示、推荐数据集展示、热门API展示、最新API展示、热门APP展示、最新APP展示等。

图29呈现了各地平台在以上8个基本动态展示条目方面的详细情况。目前，大多数平台都提供了最新数据集展示、热门数据集展示和热门APP展示三项动态展示条目，而好评数据集则无平台提供展示。具体到各省市来看，贵阳提供了最多的动态展示条目7条，宁夏、宁波次之，提供了6条。

地区	最新数据集展示	热门数据集展示	好评数据集展示	推荐数据集展示	热门API展示	最新API展示	热门APP展示	最新APP展示	总计
贵阳	●	●		●	●	●	●	●	7
宁夏	●	●			●	●	●	●	6
宁波	●	●		●	●	●	●		6
哈尔滨	●	●		●	●		●		5
江门	●	●			●	●		●	5
荆门	●	●		●			●		4
惠州	●	●			●		●		4
东莞	●	●		●	●				4
贵州	●	●		●			●		4
青岛	●	●					●	●	4
山东	●	●					●		3
济南	●	●					●		3
山东其他地市	●	●					●		3
北京	●	●					●		3
佛山		●			●		●		3
广州	●	●					●		3
上海	●					●		●	3
深圳	●	●						●	3
武汉	●	●					●		3
扬州	●	●		●					3
浙江	●	●					●		3
中山	●	●					●		3
广东	●							●	2
江西	●	●							2
苏州	●	●							2
梅州	●	●							2
无锡	●						●		2
乌海	●	●							2
湛江		●							1
肇庆	●								1
阳江									0
长沙									0
总计	28	26	0	7	8	5	18	7	

图29 各地平台动态展示条目详情列表

3. 最新相关资讯

政府数据开放平台提供与数据开放相关的最新资讯，包括数据开放活动信息和数据开放政策两类，例如国家和地方最新的政策文件、工作计划、开放清单、数据创新应用大赛消息等,有利于用户了解该领域的最新活动和政

策。目前，有接近半数的省市已在其平台上提供了最新资讯；其中，贵阳、贵州、上海、深圳、青岛五省市既提供了数据开放活动信息，也提供了数据开放政策，并进行及时更新。各地平台最新资讯功能分布详见图30。

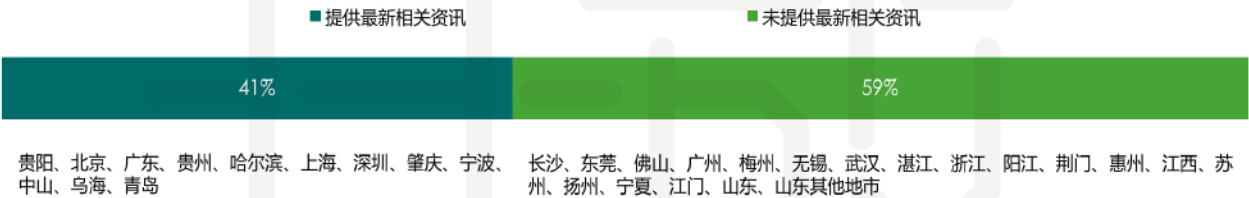


图30 各地平台最新资讯功能分布

4. 可视化展现

对平台数据进行可视化展现，有利于用户更为直观形象地了解平台数据情况。报告发现，目前国内政府数据开放平台主要提供两种可视化展现的方式：分别为展现平台数据之间关联性的数据图谱和展现平台数据开放和访问情况，如数据开放量最多的部门和主题。

目前，能够展现网站数据关联性的政府数据开放平台较少，仅有佛山、广州、贵阳、哈尔滨、宁夏、江门、乌海的平台提供了展现数据关联性的数据图谱。其中佛山、广州、贵阳、哈尔滨、宁夏、江门这六省市除了提供可视化展示外，还提供了数据集链接，即可以通过节点直接进入具体的数据集。



图31 广州市平台的“数据图谱”功能

如图31所示，广州的“数据图谱”对不同领域、部门、主题的数据集之间的关联进行了展示，将鼠标移至某个节点后，页面上会展现该节点所关联的数据集名称等，并以窗口形式展示在页面右侧，使用户对数据之间的关联性有一个基本、宏观和直观的印象。

贵阳的“全局图谱”则对不同领域、部门、主题的数据集之间的关联性进行了展示，将鼠标移至某个节点后，页面会展现该节点的名称、资源数、所属范围、关联节点等（见图32）。



图32 贵阳市平台的“全局图谱”功能

目前，如图33所示，近一半的地方政府数据开放平台提供了基本的基于平台数据开放和访问情况的可视化，即按部门和主题开放数据的可视化。其中，贵阳、贵州、青岛、上海、深圳提供了更为丰富的可视化内容。例如，贵阳在“开放统计”栏目提供了访问次数、访问来源追溯、平台访问历史、关注热度词云、数据开放统计、数据更新统计、资源格式比例、数据需求分类比例、数据评分统计、数据下载和API调用的排名情况等内容（见图34）。深圳市也提供了较为丰富的可视化内容（见图35）。

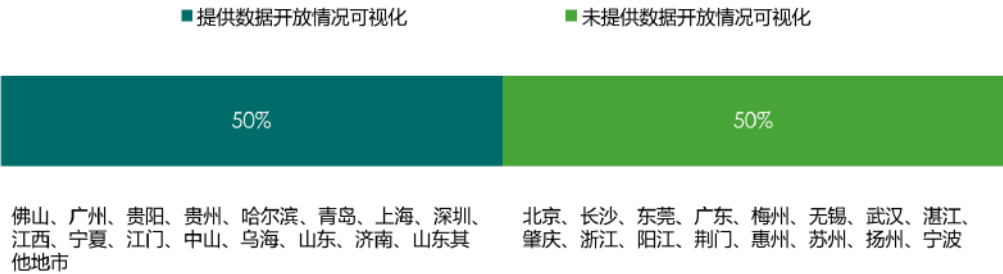


图33 各地平台数据开放情况可视化展现分布



图34 贵阳市平台的“开放指数”栏目



图35 深圳市平台的“开放指数”栏目

3.3.2 平台导引

平台导引是指政府数据开放平台提供了便捷的引导功能，以帮助用户发现目标数据，如分类导航功能、搜索功能、排序功能、用户操作指南、智能交互、提供发布者联系方式和平台间链接等。

1. 分类导航功能

大部分地方政府数据开放平台都提供了多种数据分类导航方式，较为常见的是按主题、机构和标签等分类的导航。除了少数地方平台没有数据导航功能外，其余地方平台都提供了按主题分类的导航方式。同时，大多数地方平台还提供了按部门分类的导航。然而，仅有佛山、广州、青岛、宁夏、济南等五省市的地方平台提供按标签分类的导航。

大部分地方政府数据开放平台都能提供按主题分类的导航功能，差别在于主题命名的方式不同。有的地方平台主要从部门视角进行命名，但名称过于模糊和宽泛(如经济建设、社会发展等)。这种方式虽有利于政府自身对内容进行分类，但对于用户而言则存在一定的理解和辨别困难。有些名称含义不明确，如“红榜”、“黑榜”等，不利于用户发现和获取数据。同时，有些地方平台则主要从用户视角出发进行主题命名，如上海和贵阳，以用户获取服务的场景为基础来进行主题命名。如图36所示，上海的场景分类包括社会保险、婚育、学校教育、培训与就业、就医与保健、交通出行、社区周边生活服务、残疾人、离退休、城市安全、政府办事等。



图36 上海市平台的“场景”分类导航

2. 搜索功能

除个别平台之外，其他地方政府数据开放平台皆提供了搜索功能。但是，少数平台提供的搜索功能无效，即输入了完全匹配的数据集名称后无法搜索到相应的数据集。同时，在提供基本的搜索功能基础上，约有一半的平台提供了高级搜索功能，即可以通过设置多个筛选条件来进行更为精确的搜索。

贵阳市政府数据开放平台提供了较全面的搜索功能，如图37所示，该平台提供元数据目录查询功能，筛选条件包括资源名称、资源关键字、资源摘要、数据领域、主题分类、行业分类、服务分类、数据格式、更新频率、提供部门、发布日期、更新日期等，为用户提供精准的搜索功能。



图37 贵阳市平台的“元数据搜索”功能

3. 排序功能

排序功能，是指用户根据下载次数、更新时间等具体特征，对数据集或APP应用进行排序，以满足自身多样化的查询需求。大部分地方政府数据开放平台设置了按需排序功能，其中尤以上海、山东、广州、贵阳等平台的排序功能为优，可以根据“更新时间”、“数据量”、“访问量”、“下载量”、“评分”等5个标准进行升序或者降序排列。

地搜寻、获取数据，促进数据的融合利用。相关数据推荐目前主要有两种方式：一种是只在单个数据集下面使用文字推荐相关数据集；另一种则不仅使用文字推荐相关数据集，还直接提供具体推荐数据集的链接。

4. 相关数据推荐

在某个数据集下推荐相关联的数据集，能够帮助用户更好

目前，部分平台提供了相关数据推荐功能，详情见图38。东莞、佛山、广州等地提供了相关推荐数据的链接，宁夏、宁波等地使用文字罗列了相关的数据集，而还有超过半数的平台未提供该功能。

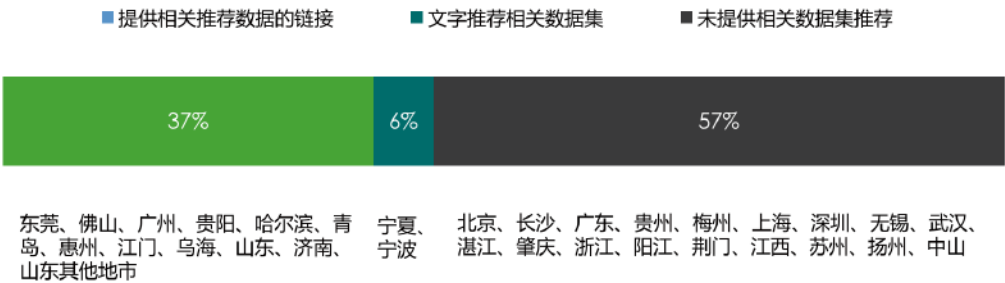


图38 各地平台相关数据集推荐功能分布

5. 用户操作指南

目前，大部分地方平台提供了新用户操作指南，包括开发者文档和用户操作手册等，为新用户使用平台提供指引。其中，上海、青岛和贵阳等地的用户操作指南较为完善，上海和贵阳提供了用户操作指南的下载，青岛还提供了如何调用API服务的指南，如图39所示。然而，多数平台用户操作指南浏览和寻找方式复杂和困难。

6. 智能服务

智能服务是指以人机对话的方式为用户提供指引。智能交互可以方便地为用户提供个性化的指引和服务，节省客服的人力投入，提升服务的效率和效果。目前仅有贵阳平台在此方面进行了尝试，见图40。



图39 青岛市平台的用户操作指南



图40 贵阳市平台的“智能机器人”

7. 数据发布者联系方式

提供数据发布者联系方式有利于用户与数据的发布方直接联系，获得更好的服务和指导。在国际上，数据发布者联系方式的通行标准是数据发布者提供详细的联系方式，包括姓名、电话、电子邮件、地址等，而国内目前尚未有平台提供这些完整的信息，只有少数平台提供了数据发布者的地址，如浙江、贵阳、武汉、湛江等地。图41展示了贵阳提供的数据发布者联系地址。

贵阳市政府数据开放平台

Guiyang Open Government Data Platform

区县目录

首页

数据

应用

网站分析

资讯动态

互动交流

个人中心

退出

数据目录

数据集

API

依申请开放

地图数据

元数据

关联信息

数据图谱

标识符	620000/3900641	名称	贵阳市主要粮油副食品每日监测表			
摘要	提供贵阳市主要粮油副食品日监测价格，包括商品名称、规格、单位、主要农贸市场价格、主要超市价格、环比增幅等。					
关键字	粮油、副食品、价格、监测信息		数据领域	文体休闲 三衣服务		
主题分类	经济管理	行业分类	居民服务、修理和其他服务业		服务分类	服务交付的支撑
开放方式	主动开放	更新频率	实时	发布日期	2017-01-10	
更新日期	2018-05-15	资源格式	CSV,XLS,其他	所属行政区域	市辖区	
数据提供方	市发改委	提供方地址	贵阳市金阳行政中心政府二期B区		数据维护方	市发改委
所属行政事项	暂无	资源状态	激活	语种	中文	
文件大小	80.44M	文件数量	497	浏览量	35339	
下载量	15998	时间范围		空间（地域）范围	市辖区	
来源系统	暂无	在线资源链接	暂无地址			

图41 贵阳市平台的数据发布者联系方式

8. 平台间链接

平台间链接是指不同的平台之间可以相互链接，有利于用户便利地寻找相关的其他平台。目前，山东作为省级平台与其市级平台相链接(如图42)，贵阳作为市级平台与其区县级平台相链接(如图43)，中山作为市级平台提供了省级平台的链接(如图44)。



图42 山东省平台的“平台间链接”功能



图43 贵阳市平台的“平台间链接”功能



图44 中山市平台的“平台间链接”功能

• 3.3.3 数据获取

数据获取主要是指政府数据开放平台提供相关功能便于用户浏览和获取目标数据，如提供数据预览、开放数据目录、本地获取、便利的注册方式、分级分类获取，数据集订或收藏等功能，这是政府数据开放平台的核心功能模块。

1. 数据预览

数据预览是指用户在下载数据集之前，能够对数据集的内容进行预览。如图45所示，北京、东莞、广东等平台提供了基本的数据集预览功能，其中宁波平台要求用户登陆后才可以预览数据集；还有部分平台未提供数据集预览功能。



图45 各地平台数据集预览功能分布

2. 数据目录

数据目录是在平台上提供了开放数据资源目录，帮助用户了解开放数据全貌。目前，只有贵阳提供了能够下载的可

机读格式的开放数据目录(如图46)。

部门名称	标识符	资源名称	资源类别	关键字	资源摘要	一级主题分类	开放类型	更新频率	数据领域	开放方式	发布日期
市金融办	580000/3900575	市金融业运行分析	Excel	金融,运行分析,运行,分析	由市金融办每月提供更新数据	财政	完全开放	按季	财税金融	主动开放	2017-01-10
	580000/3900563	贵阳市上市公司名录	Excel	上市,上市公司,名录	上市公司名录	财政	完全开放	按季	企业服务	主动开放	2017-01-10
	520100/0000392	贵阳市30家“新三板”挂牌企业信息	Excel	新三板,挂牌企业	贵阳市30家“新三板”挂牌企业信息	商业、贸易	完全开放	按年	财税金融	主动开放	2017-01-10
	580000/3900569	贵阳市上市企业信息	Excel	金融,上市,企业信息	该数据由市金融办提供,包含名称、组织机构代码、地址、区镇等属性字段。	财政	完全开放	按年	财税金融	主动开放	2017-01-10
	580000/3900561	贵阳市交易所名录	Excel	交易所,交易,场所,名录	交易所场所名录	财政	完全开放	按年	财税金融	主动开放	2017-01-10
	580000/3900564	贵阳市小额贷款公司名录	Excel	贷款,小额贷款,名录	小额贷款公司名录	财政	完全开放	按年	企业服务 财税金融	主动开放	2017-01-10
	580000/3900565	贵阳市融资性担保机构名录	Excel	融资,担保,名录	融资性担保机构名录	财政	完全开放	按年	财税金融 企业服务	主动开放	2017-01-10
	580000/3900567	贵阳市金融机构信息	Excel	金融,金融机构	该数据由市金融办提供,包含名称、地址等属性字段。	财政	完全开放	按年	财税金融	主动开放	2017-01-10
市商务局	480000/3900538	贵阳市外商合同投资情况	Excel	外商,合同,投资	贵阳市外商合同投资情况数据产品主要发布了贵阳市外商实际投资金额、同比增长等数据信息。	经济管理	完全开放	按月	经济建设 企业服务	主动开放	2017-01-10
	480000/3900539	贵阳市外商实际投资情况	Excel	金融,外商,投资	贵阳市外商实际投资情况数据产品主要发布了贵阳市外商实际投资金额、同比增长等数据信息。	经济管理	完全开放	按月	经济建设 企业服务	主动开放	2017-01-10

图46 贵阳市平台的“数据目录”

3. 本地获取

本地获取是指可下载数据集都能够在平台上直接获取，而无需跳转到其他平台才能下载。目前，大多数地方实现了数据的本地获取，包括北京、长沙、东莞等地；然而，仍有部分平台出现数据需要跳转到外部链接后才能获取的情况(如图47)。

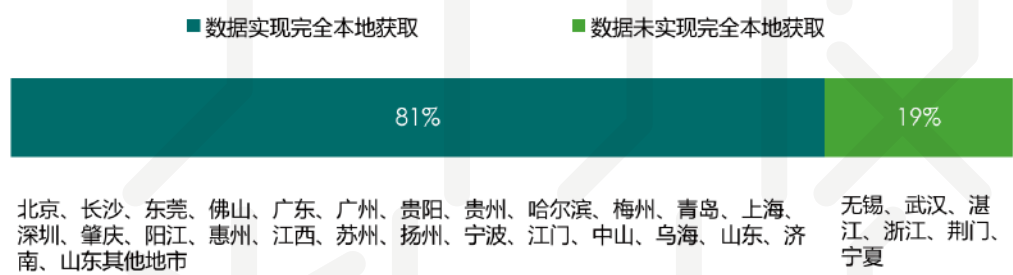


图47 各地平台数据集本地获取情况分布

4. 数据分级分类

对开放数据进行分级分类管理有助于推进数据开放的广度和深度，在保障数据安全的前提下开放更具价值的数据资源。

报告对国内地方政府的数据开放属性进行了梳理分析，发现主要存在以下三类别：(1)普遍开放：用户无需注册或经过简单注册便可获取数据；(2)依申请开放：用户需要向平台提出申请，提供相关信息和理由，经平台审核通过后才可获取数据；(3)有限开放：主要采用数据沙箱模式，可以使用数据获得分析结果但不能获取原始数据，是一种半开放的形式。

目前，大部分平台采取普遍开放的方式，只有上海、贵州、深圳和贵阳等平台上既有普遍开放的数据，也有依申请开

放或有限开放的数据。其中上海平台上开放的数据分为普遍公开、特定公开和依申请开放三种属性。用户获取普遍公开的数据只需凭手机号进行简单注册，获取特定公开的数据集则需要填写身份证信息进行实名认证，而获取依申请公开的数据集则要经过特定的申请审核程序才能获得。贵州的开放数据分为完全开放和依申请开放两种属性，贵州还对具体数据集的内容进行了细分，一个数据集中有些字段是完全开放的，有些字段的数据则需要申请，例如证件号码类的信息就需要申请才能开放。贵阳将平台上的数据分为完全开放和条件开放两类。同时，贵阳还在探索一种如何在有效保障数据安全的前提下开放敏感类数据的方式。这类数据通过“数据孵化平台”有限制地进行开放，构建封闭的数据利用环境，提供相应的数据分析和开发工具，用户可以在平台上分析和利用这类数据，但是不能获取和带走数据(如图48所示)。



图48 贵阳市平台的“交通大数据孵化平台”

5. 注册登录

出于数据安全和跟踪数据利用情况的考量，即使对于普遍开放类的数据，目前大部分地方平台仍然要求用户进行注册登录方可获取数据。

报告经过梳理发现，目前各地平台从易到难主要存在以下几种注册登录方式：(1)无需注册便能下载全部数据；(2)注册时要求提供电子邮箱等非实名信息；(3)注册时要求提供手机号码；(4)注册时要求提供身份证号码等证件信息。此外，还出现了分级分类的注册登录形式，对于获取不同类别的数据有不同程度的注册要求，探索通过精细化管理的方式来实现数据保护和数据开放之间的平衡，然而不同平台对分级分类注册的要求也存在一定差异。

目前，浙江、广东、长沙等地平台无需用户注册登录便能下载全部的数据；青岛、惠州、乌海等地方平台要求用户通过提供电子邮箱等用户非实名信息完成注册；贵阳、苏州、山东等地平台需要用户提供手机号码后完成注册；北京、宁夏的平台要求用户提供身份证号码等信息后才能完成注册。

上海、深圳、宁波已初步实现了分级分类的注册要求，但在获取数据的最低门槛上要求各不相同。例如，深圳设置了简单注册、实名注册和特定审核三种注册级别，第一种需要用户提供手机号等信息方可下载，第二种需要用户实名认证，第三种不仅需要实名认证，还必须通过平台审核后才可下载。宁波将数据公开属性分为完全公开和对注册用户公开，第一类数据无需注册便可下载，第二类数据则需要用户提交手机号等信息进行注册后才可下载。

此外，便捷的第三方登录方式能为用户免去重复繁杂的注册程序。目前浙江、青岛、惠州等九个平台提供了第三方登录方式，可以通过支付宝、微信、微博等第三方账号进行登录，为用户提供了更多选择和便利。

图49展示了各地平台数据下载的具体注册要求和第三方登录情况。

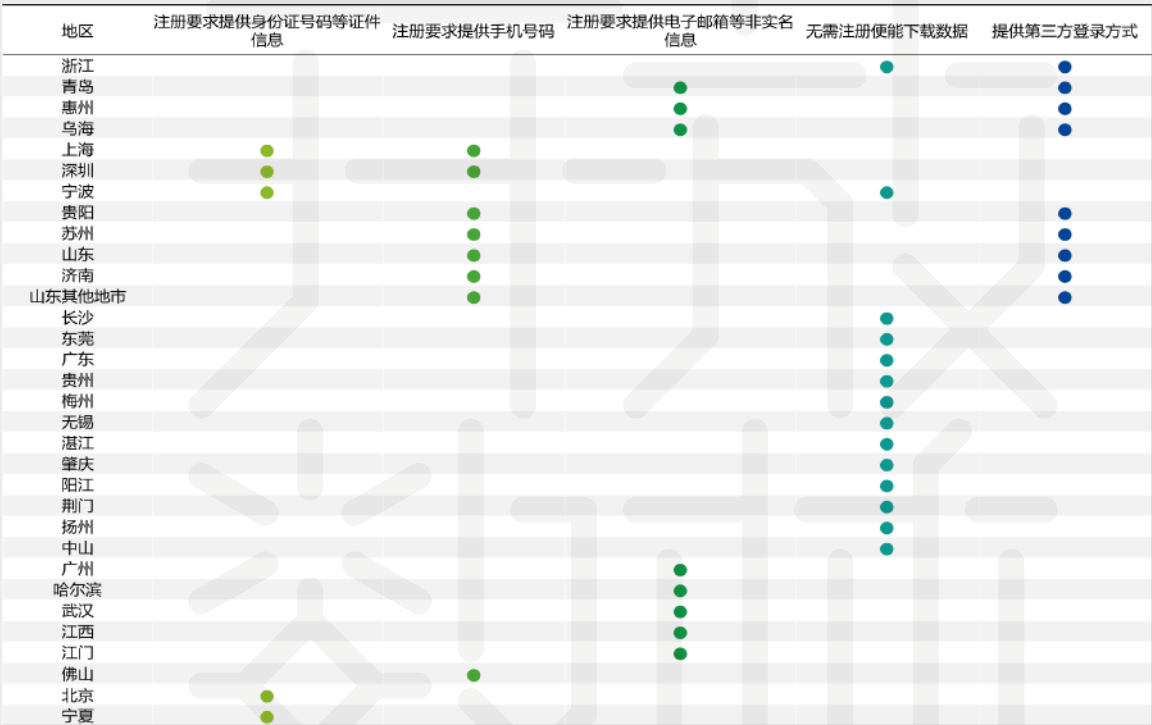


图49 各地平台注册登录方式详情列表

6. 数据集订阅或收藏

数据集订阅或收藏是指用户可以对数据集进行订阅或收藏，之后便可在个人中心中直接发现该数据集而无需另行查找。

目前，北京、佛山、广东等地在具体数据集页面提供了数据集的订阅或收藏功能，但仍有逾四成地区未提供该功能（如图50所示）。

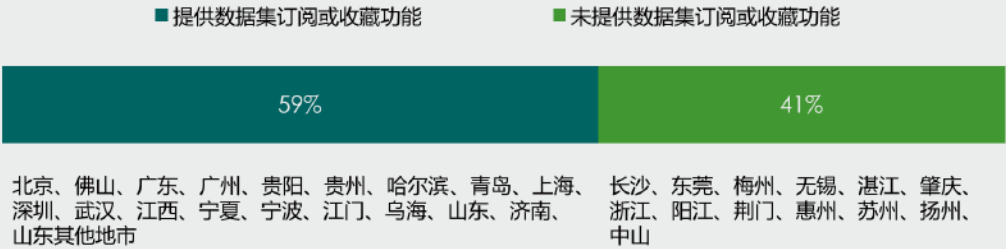


图50 各地平台数据集订阅收藏功能分布

• 3.3.4 工具提供

工具提供是政府数据开放平台提供基本工具，以帮助用户对数据集进行分析 and 开发。政府数据开放平台除了要向用户提供原始数据外，还要提供基本工具帮助用户对数据进行分析 and 开发。目前国内政府数据开放平台提供的工具主要包括三类，分别是：可视化与分析工具、开发工具和地理空间工具。

1. 可视化与分析工具

可视化与分析工具是指平台提供了实现数据集可视化或统计分析的工具。用户可以在数据集页面通过设置不同条件，对数据进行初步的统计分析，还可以对统计结果进行可视化呈现，如饼状图、柱状图和曲线图。

图51、52分别展示了广州市政府数据统一开放平台的数据统计与可视化工具。平台提供这些工具方便用户在下载数据前对数据集有基本的了解，还可帮助非专业的初级数据利用者直接获取所需要的数据分析结果，而无需下载原始数据。



图51 广州市平台的数据统计工具



图52 广州市平台的数据可视化工具

目前，超过三成政府数据开放平台提供了可视化与分析工具，其中大部分实现了分析工具与数据集的联通或嵌入，用户可直接、便捷地调用。图53展示了各地平台数据分析工具提供情况，少部分平台虽然提供了数据分析工具，却

未能与平台内数据集联通，例如贵阳政府数据开放平台提供了功能较为高级的数据分析工具，但需要用户自行上传数据，未能实现与平台内数据集的直接关联。

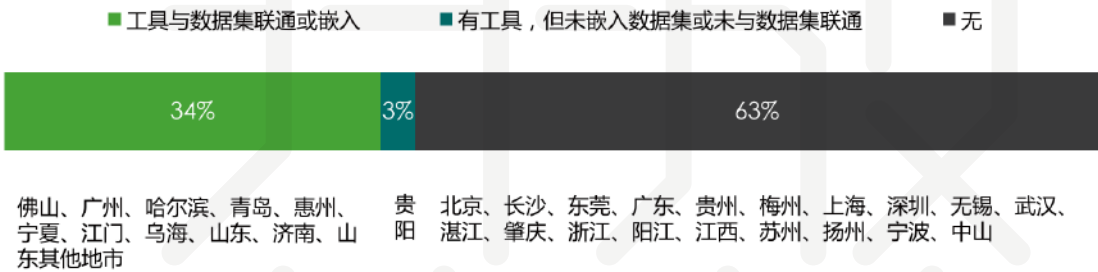


图53 各地平台可视化与分析工具提供情况分布

2. 开发工具

平台提供开发工具可为用户对数据进行深入开发利用提供便利。目前，佛山、广州、贵阳等地均提供了开发工具和服务（见图54）。



图54 各地开发工具提供情况分布

其中，青岛市平台建立了“开发者中心”栏目，建立了“创建应用-申请服务-开发应用-审核发布-管理应用”的全流程管理（见图55）；贵阳市平台则在“开发向导”栏目中形

成了“资源提供-分析工具提供-开发工具提供-推广渠道提供”的多维度管理（见图56）。



图55 青岛市平台“开发者中心”



图56贵阳市平台“开发向导”

开发工具的提供应以便利开发者为目的，但部分地区仅提供了付费下载的跳转链接或止步于对工具的文字介绍，性质更接近于开发工具的推荐者，而不是开发工具的提供者。

3. 地理空间工具

地理空间工具是指平台提供了将政府开放数据和地理信息

进行融合应用的功能。目前，长沙、东莞、佛山等平台提供了地理空间工具，这类地图工具多与高德、百度等企业合作，提供导航功能，或将一些公共部门的位置叠加在地图上提供服务（见图57）。图58展示了山东公共数据开放网上的地图服务功能，用户可在页面上自行检索或根据主题分类查找地理数据集，然后叠加在地图上显示。

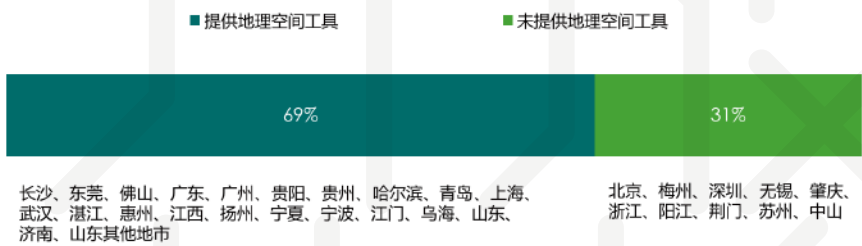


图57 各地平台地理空间工具提供情况分布



图58 山东省平台的开放网地图服务

• 3.3.5 利用成果展示

数据利用成果展示是指政府数据开放平台向社会集中展现用户利用平台上开放的数据后产生的各类数据利用成果，例如APP应用、研究报告、传播产品等，从而形成从数据开放、被利用到产生价值的闭环。展现这些数据利用成果一方面有利于对其进行推广，便于社会公众集中检索和获取，促进数据开放价值的实现；另一方面也可进一步激励数据开发者更好地利用政府开放数据为社会服务，推动开放数据生态体系的可持续发展。

1. 利用成果类型

基于开放的政府数据所产生的利用成果可分为以下三种主

要类型：(1)APP应用，即企事业单位利用开放数据开发的各种服务应用；(2)传播产品，指对开放数据经过分析和可视化处理后，以数据新闻、自媒体推送等方式向社会传播，用易于公众解读的方式展现数据分析结果；(3)研究成果，即高校、研究机构、社会团体等主体利用开放数据形成的研究报告或论文。

目前，大部分地方政府都开设了数据利用频道集中展示数据利用成果。现阶段各地平台展示的主要是APP应用，只有贵阳开设了“数据无限”栏目面向普通公众展示数据传播产品(图59)，而研究报告部分目前还没有平台进行展示。



图59 贵阳市平台的“数据无限”栏目

2. 标明数据来源

数据来源是指政府数据开放平台注明了其展示的数据利用成果所使用的政府开放数据。目前，大部分地方未标明数据来源，只有北京在具体的数据利用成果下用文字标明了所用到的开放数据来源。贵阳在数据利用成果页面，不仅标明了所用的数据来源，还设置了直接进入来源开放数据集的链接（如图60所示）。

3. 利用者来源

政府开放数据的目的在于推动社会对政府数据进行开发利用。然而，目前超过半数的地方平台上展现的数据利用成果（主要是APP应用）基本是由政府部门自身开发，而非数据开放后由市场所开发。目前，佛山、广东、广州等17个地方平台上的数据利用由政府和市场共同开发，只有北京平台上的数据利用基本由市场开发(见图61)。

4. 利用成果提交

政府数据开放平台还应当为开发者开设提交数据利用成果的功能，这既可使开发者主动上报、上传他们利用政府数据所产生的数据利用成果，也有助于政府了解追踪其开放数据的利用情况。

北京、佛山、广东等地方平台为开发者提供了提交数据应用的功能(如图62)，主要包括：在数据应用频道中设置提交数据应用的通道，在用户的个人中心开设应用提交功能，在平台的互动交流频道中设置专栏或在开发者中心栏目中设置提交功能等。其中，贵阳平台还开设了研究报告的提交功能，为日后用户提交利用数据开放产生的研究成果提供了渠道（如图63）。



图60 贵阳市平台上提供的来源数据集链接



图61 各地平台数据利用者来源分布

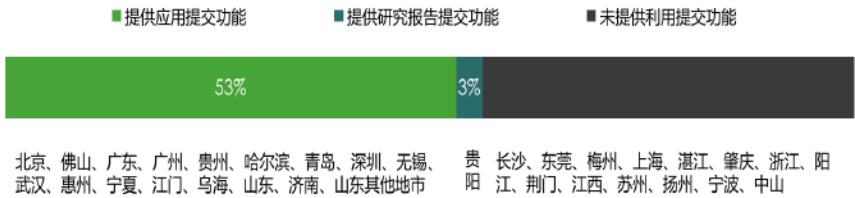


图62 各地平台数据利用成果提交功能分布



图63 贵阳市平台上的研究报告提交功能

• 3.3.6 互动交流

互动交流是指政府数据开放平台提供的功能便于用户与数据提供方进行双向沟通，针对具体数据集提交评价，提出数据请求和向平台提供建议反馈等。

1. 数据集评价功能

目前，上海、荆门、佛山等平台为用户提供了对单个数据集进行评价的功能，图64展示了各地平台数据集评价功能得分前10名。



图64 各地平台数据集评价功能分布(前10名)

数据集评价的主要方式是星级评价和文字评价，上海、贵阳、山东等地同时采用了星级评价和文字评价两种途径，较为具体地展现了用户对于数据集的评分与具体意见。值得注意的是，宁波（如图65）对星级评价进行了进一步细分，引导用户在评价时从完整性、实时性、准确性三个维度进行考量，而上海（如图66）则从准确性、及时性、满意度与可用性四个维度对星级评价进行细化。

此外，只有上海、荆门、贵州、湛江等平台允许用户无需注册便可对数据集直接进行评价，大部分平台只对注册登录的用户开放数据集评价功能。上海、荆门、佛山等10个平台在平台上公开了全部或部分用户对单个数据集的评价。图67展示了山东公共数据开放网的数据集评价情况，用户的评论能够被展示，并被其他用户评论与点赞，增强了平台与用户、用户与用户之间的互动性。



图65 宁波市平台数据集评价



图66 上海市平台数据集评价功能



图67 山东省平台交流互动

2. 数据请求功能

数据请求功能使用户可以在平台上提出其希望开放的数据，开设这一功能有助于数据提供方获知社会的数据需求。目前，青岛、贵阳、深圳等16个地方平台为用户提供了数据

请求功能，但有些地方还需要用户注册登录后才能请求数据。此外，报告还考察了数据请求功能的回复时效和质量。图68展现了排名前9的地方平台数据请求功能的详细情况。

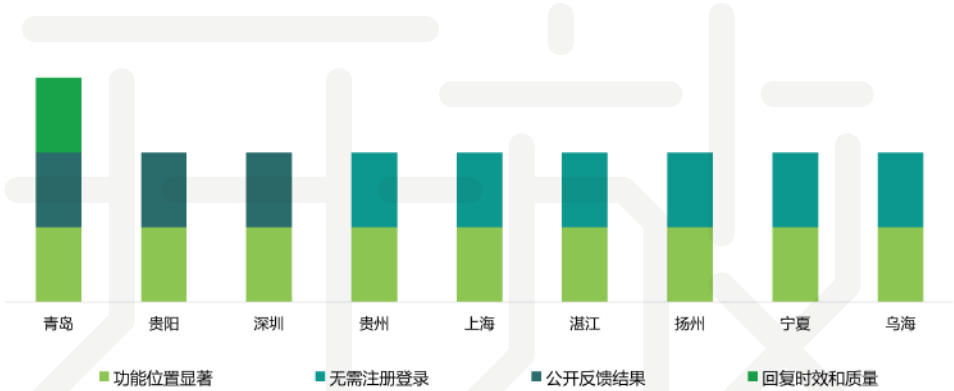


图68 各地平台数据请求功能分布(前9)

青岛对用户的数据请求做出了反馈并予以公开。图69展现了青岛政府的“建议数据”情况。用户能够看到曾经提交的

数据请求的反馈状态，包括“已经开放”、“不予开放”、“暂不开放”三种状态，并能够根据状态进行筛选查看。



图69 青岛市平台建议反馈

3. 建议反馈功能

目前，七成以上的政府数据开放平台都开设了建议反馈功能向用户采集建议，包括问卷调查和填写文字性的咨询建议等方式。图70展现了排名前10的地方平台建议反馈功能的详细情况。北京、贵阳、江西等地向用户公开全部或部

分的建议反馈；贵州、上海、湛江等地允许用户无需注册登录就可对平台建设提出建议。此外，北京、贵阳、江西的回复建议的时效性较强、质量较高。以北京市政务数据资源网为例，平台针对网友的问题及时回应，且回复详尽（如图71）。

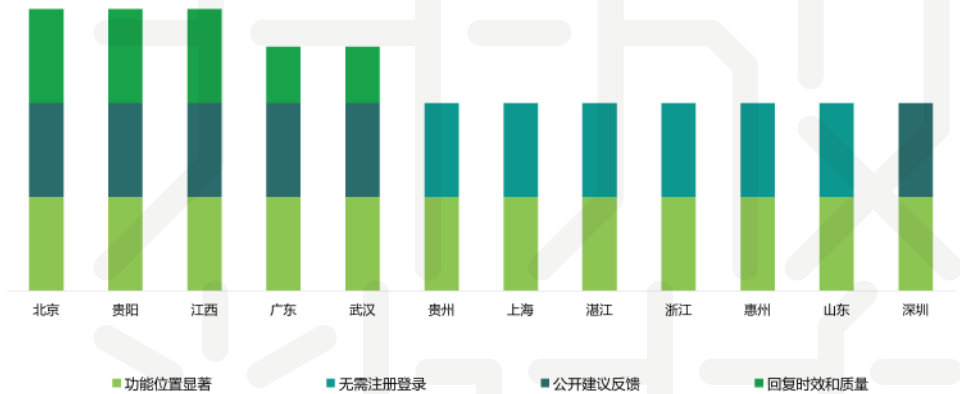


图70 各地平台建议反馈功能分布（前10）

互动交流

问卷调查

历届活动

咨询建议

常见问题

提交APP应用

建议增加的数据类型

用户指南

咨询建议

问题标题

搜索

我要提问

编号	问题	时间	浏览次数
1	有没有行政区划地图（区县级；街道、办事处（乡/镇）级；社区（村）级）	2015-03-25	919
2	这点点数据不能算是资源	2015-03-29	612
3	为什么APP应用这么少呢？	2015-04-11	455

描述：

北京在中国大陆，在开放数据这一领域是先锋，这个网站上也已经发布了很多数据集。但APP却只有个数，想请问这些是北京开放数据应用程序的全部吗？若不是，能否提供一些其他应用程序的链接；若不是，能请教提供的资源不少而应用却很少的原因吗？非常感谢！

回复：

感谢您对北京市政务数据资源网的关注！北京市政务数据资源网欢迎广大用户提交基于本站数据资源开发的APP应用程序，但APP应用的开发和上传完全是出于用户自愿的个人行为，本站并没有强制要求。

北京市政务数据资源网 2015/04/11 13:55:57

图71 北京平台的“咨询建议”

4. 数据纠错功能

数据纠错功能是指用户可以向平台提出其发现的特定数据集存在的错误，有助于数据提供方提高开放数据的准确性。图72展现了排名前10的地方平台数据纠错功能的情况。目前，贵阳、贵州、浙江等地开设了数据纠错功能。贵阳和浙江用户无需进行注册登录就能够进行纠错，其中贵阳

还公开了数据纠错情况与回复。如图73所示，贵阳的数据纠错公开服务位于“互动交流”栏目内，提供纠错内容、用户名称、发布时间、状态、回复时间、回复内容，清晰地向其他数据使用者展示用户数据纠错的内容以及平台的处理情况。

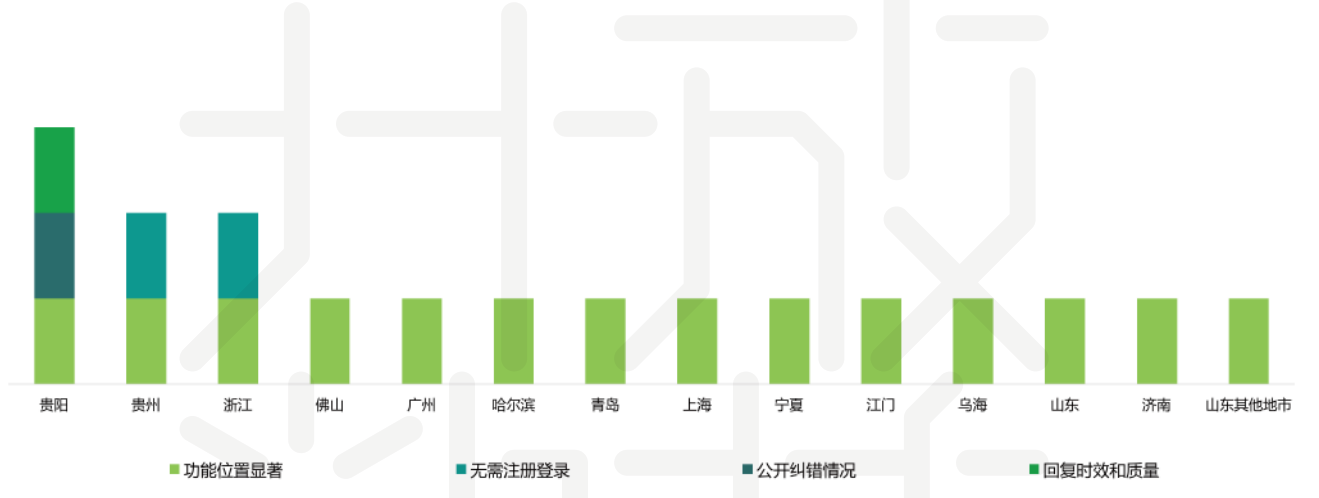


图72 各地平台数据纠错功能得分(前10)

纠错详情

纠错内容:

在关联信息中，区县数据出现重复

网友:

test2017

发布时间:

2018-04-19 16:22:08

状态:

已回复

回复内容:

2018-04-19 17:23:58

回复时间:

重复的数据为各区县相同的数据集名称

图73 贵阳市平台数据纠错结果公开

5. 分享传播功能

在平台上提供分享传播功能，用户可将平台上的信息在微博、微信和QQ等社交媒体平台传播，有助于提升政府数据开放平台的热度和影响力，吸引更多的政府、企业、社会团体和公众参与其中。目前，半数的地方平台提供

了分享传播功能（如图74所示），其中北京、佛山、广州等地方平台在数据集、数据接口与数据应用等页面上，提供了一键分享至微信、微博、QQ的功能。图75展现了广州的分享传播功能。



图74 各地平台分享传播功能分布



图75 广州市平台的分享功能

6. 社交媒体账号开设

社交媒体账号开设是指政府数据开放主管部门设立了专门的社交媒体帐号(如微博、微信等)，向社会宣传推广本地的数据开放工作，并与社会公众通过社交媒体进行直接互动。目前，

仅有贵阳市开设了微信公众号“贵阳政府数据开放”和官方微博账号“贵阳大数据委”，推送与政府数据开放相关的新闻资讯、最新政策、数据更新情况和数据传播产品等。

• 3.3.7 个性化整合

个性化整合是指平台为用户提供个人中心，使用户可以根据自己的个性化需求，集中使用数据获取、工具使用、成果展示、互动反馈等功能。目前，北京、佛山、广东等地均开通了用户个人中心，提供个性化服务（如图76）。



图76 各地平台个性化整合分布

如图77所示，山东的“个人中心”为用户提供了收藏数据和应用、提交建议、参与调查等功能，还可注册成为开发者，与平台的开发应用栏目对接，查看申请的数据接口等信息。



图77 山东省平台的“个人中心”

贵阳市政府数据开放平台开设了自定义场景功能(如图78)。用户可根据自身需求定义场景名称，并在该场景下搜索和组合所需要的数据，同时支持数据下载与可视化分析。然

而，多数平台的“个人中心”还只是实现了功能“整合”，而未提供足够多的个性化服务。



图78 贵阳市平台的自定义场景功能

3.3.8 平台层结论

1. “开门见数”差异悬殊，动态与可视化尤为薄弱

与去年相比，平台概览整体上有了较大进步，多数已向用户展现平台上的数据概况和相关资讯，但地区间仍存在显著差异。青岛和中山的平台已覆盖了所有评测的数据统计条目，但仍有地方只提供了很少的或未能提供数据统计条目。贵阳已提供了七项动态展示条目，但多数平台的动态展示内容还比较单一。接近半数的省市已在其平台上提供了最新相关资讯，但半数平台仍未实现。佛山、广州、贵阳等少数地方平台已能提供较为丰富的可视化功能，但

多数地方仍未提供或提供形式较为初级。

2. 找到数据仍是难题，路径虽多但效果存疑

总体而言，各地方的平台引导功能相较于去年有了较大提升，实现了从无到有的突破。佛山、广州、青岛、宁夏和济南等地能够提供按标签分类的导航方式，东莞、哈尔滨、山东等地提供了相关数据推荐功能，浙江、贵阳、武汉、湛江等地提供了数据发布方的地址，山东和贵阳的平台分别与下级市级和区县级平台建立了链接(如图43)，贵阳已尝试通过智能交互方式提供服务。

然而，各地仍存在一些普遍问题，包括分类导航缺乏用户视角，搜索功能不够精确完整，排序方法简单初级，推荐相关数据却不提供链接，操作指南有而不全，联系数据发布者困难，平台之间缺乏互联互通等。

3. 获得数据门槛高、体验差，“一刀切”多于分级分类

总体上，地方平台在数据获取方面相比去年出现不少亮点，贵阳率先开放了数据目录，上海、贵阳、深圳和贵州开始探索数据分级分类开放，上海、深圳、宁波也已开始实行分级分类的注册要求。

但不少老问题仍然存在：多数地方未提供数据目录和数据集预览，少数平台仍需要用户通过链接跳转到外部平台后才能获取数据，还有部分平台未提供数据订阅或收藏功能。在绝大多数平台上，用户获取数据的方式缺乏便捷性和灵活性，很多地方不分数据类型和用户类型而“一刀切”地采用同样严格的注册登录方式，影响了用户获取和利用数据的体验感和获得感。

4. 工具虽有却不好用，用户体验因人而异

去年以来，越来越多的平台开始向用户提供可视化与分析工具、开发工具和地理空间工具。然而，用户的利用能力高低不一，提供的工具还未能充分考虑不同用户群体的需求差异和使用体验。有三成的地方平台已提供了数据分析工具，但操作难度较高，不便于初级用户使用；许多平台还未提供可视化功能帮助用户解读数据；还有少数平台未能将分析工具与平台内的数据联通，分析工具形同虚设；有些平台上的开发工具实际上只是工具介绍或下载链接；还有部分平台未实现地理空间工具与数据集的叠加应用。

5. 真正的数据利用少，提交利用成果通道不畅

整体而言，各地在数据利用成果展示方面已取得一定进展。贵阳开设了“数据无限”栏目向普通公众展示数据传播产品（图59）。北京在数据利用成果下标明了开放数据来源，贵阳还进一步提供了来源数据集的链接。北京平台上的数据利用已基本来自市场开发。超过半数的地方为用户提供

了提交利用成果的功能，贵阳还开设了提交研究成果的功能。

然而，利用成果展示普遍缺乏广度和深度，尚未充分体现开放数据的价值。各地展现的数据利用成果大多是由政府自身而非市场所开发，这并非政府开放数据供社会开发利用的初心。绝大部分地方的数据利用频道还不够丰富全面，主要展现APP应用，而缺少对数据传播产品和科学研究成果的展现。对于已展现的APP应用多数未标明所使用的开放数据来源，许多地方也未开设通道供数据利用者提交各类利用成果。

6. 互动交流“门难进”、不公开，回复时效性和质量不高

大部分平台都开设了互动交流功能连接开放数据的供给侧与需求端，并在功能上不断优化细化。上海和宁波允许用户从多个维度对数据集进行星级评价。十个地方平台已公开了用户对单个数据集的评价，青岛、贵阳和深圳对用户的数据请求进行了反馈并予以公开。北京、贵阳、江西等地向用户公开了对用户建议的反馈，且回复建议的时效性较强、质量较高。贵阳、贵州、浙江等地开设了数据纠错功能，其中贵阳还公开了数据纠错情况与回复内容。北京、佛山、广州等地方平台提供了将平台信息一键分享至微信、微博、QQ的功能。贵阳市开设的社交媒体账号定期推送与政府数据开放相关的资讯与传播产品。

然而，多数平台在互动交流的功能设置和实际效果上仍存在问题。大多数平台仍要求用户注册登录后才能进行反馈，且注册程序繁琐，增加了公众提出建议和需求的难度。大多数平台回复的时效性差，回复质量低，且未公开用户的数据请求、纠错、评价与意见反馈。

7. 有整合，少个性化

目前，已有近六成的平台通过“个人中心”实现各项服务功能的个性化整合。贵阳开设的自定义场景功能可为用户提供个性化的数据搜索、下载和分析功能。然而，多数平台的个人中心只是实现了“整合”，而缺少个性化服务。

报告中的法规与政策是指对政府数据开放相关环节（如数据治理、数据开放、数据利用、安全保护等）作出规范性要求的法律法规和政策文件，是推进政府数据开放的法治基础和重要依据。

报告对全国46个地方政府共计120项法规及政策的名称进行了文本分析。如词云分析图79所示，目前我国各地对数据开放的规定主要出现在政务公开、大数据、互联网+、信息化、智慧城市、信息共享等相关领域的法规和政策中，多以发展规划、实施意见、实施方案、行动计划、管理办法等形式发布。



图79 法规与政策名称关键词词云图

1. 数据治理

良好的数据治理是数据开放的基础和保障。政府数据治理是将数据视为政府的重要资产，对数据相关事务进行决策和行使权力。当前，全国已有多个地方在相关法规及政策中对政府数据治理提出了明确要求，这对于推进当地政府数据开放起到了重要的支撑作用。

在政府数据治理的多个维度中，与数据开放较为相关的有数据编目、数据标准和数据质量等。数据编目是一项基础性工作，要求在编制过程中对数据的开放属性进行标注，从而提升数据开放的标准化和规范化水平。数据治理不仅能推动政府数据的开放和利用，也能改善数据质量从而提升政府自身的数据资源管理水平。

如图80所示，上海、贵州、山东等6个地方在数据治理方面表现较为突出，各项指标均被纳入其政策规定范畴。

	数据编目	数据标准管理	数据质量管理
上海	●	●	●
贵州	●	●	●
山东	●	●	●
广州	●	●	●
武汉	●	●	●
贵阳	●	●	●
北京	●	⦿	●
广东	⦿	●	●
江西	⦿	●	●
哈尔滨	⦿	●	●
梅州	⦿	●	●
淄博	⦿	●	●
枣庄	⦿	●	●
威海	⦿	●	●
莱芜	⦿	●	●
滨州	⦿	●	●
深圳	⦿	●	●
潍坊	⦿	●	●

注：●表示政策规定较为全面，⦿表示政策规定不充分。

图80 各地政府数据治理政策详情列表

○ 数据编目

对数据资源按照一定的规则 and 标准进行编目，是数据治理的重要组成部分，也是数据开放工作有序进行的重要基础。如图81所示，在本次报告覆盖的46个地方中，有超过80%的地方提出要建立或完善政府数据目录体系。其中，只有上海、贵州、山东等9个地方规定要求明确数据的开放属性。如北京市规定，“政务部门负责编制、维护和管理本部门信息资源目录，明确目录中信息资源的采集途径、更新时限、数据格式以及共享开放属性等”，“根据信息资源的共享开放属性，明确共享信息资源和开放信息资源，作为开展政务信息资源共享开放工作的依据”

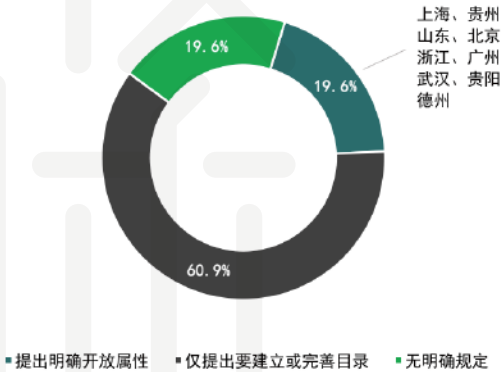


图81 各地数据编目政策情况分布

○ 数据标准

数据的采集、存储、共享、开放等各个环节都需要制定明确的标准作为指导，科学的数据标准是数据资源管理的重要依据。如图82所示，上海、贵州、广东等大部分地方都明确规定要建立或健全包括采集、存储、共享、开放、利用等关键环节在内的标准规范体系。如浙江省提出，“建立标准规范体系。研究制订数据采集标准及分级分类标准、政府数据共享标准、数据交换标准、政府及公共数据开放标准、统计标准，对共享开放的方式、内容、对象、条件等进行规范”²。而有些地方只是笼统地提出要制定数据标准规范，或提出要依照国家或省级标准进行数据资源管理。

² 《浙江省促进大数据发展实施计划》. http://www.zhejiang.gov.cn/art/2016/3/1/art_12460_263726.html

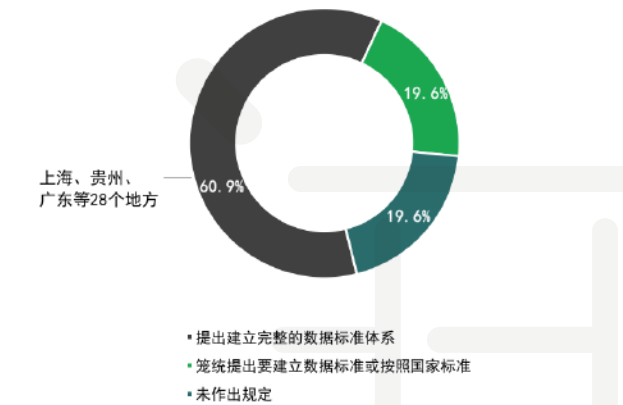


图82 各地数据标准政策情况分布

○ 数据质量

数据质量管理是指在数据采集、加工处理等环节要确保数据的完整性、真实性、准确性、有效性、时效性和有用性等。如图83所示，在报告覆盖的46个地方中，有一半以上的地方都对数据质量提出了要求。其中，北京、上海、广东等23个地方的政策规定中对数据质量的要求较为全面和具体，其余地方则是相关政策要求不完整或未对此作出明确要求。

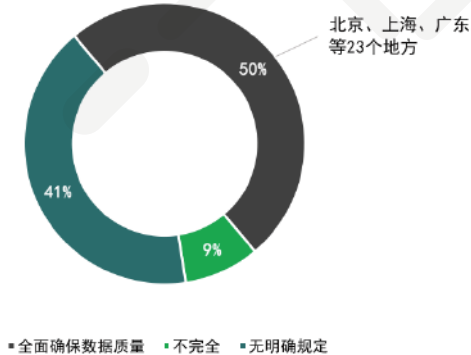


图83 各地数据质量政策情况分布

2. 数据开放

在各地的法规政策中，对开放平台建设、开放数据资源管理和开放方式与过程等数据开放的各个方面提出明确要求，是推进数据开放工作的关键要素。

○ 开放平台建设

政府数据开放平台是开放和获取数据的载体，是展示应用的中心，也是连接开放数据供给侧和用户端的桥梁。如图84所示，除个别地方外，大部分地方的法规政策都要求建设或完善数据开放平台，集中统一向社会开放政务或公共数据资源。有些地方还提到了平台的功能、运维等具体内容。如上海市规定，“建立上海市政府数据服务网，由第三方专业服务机构负责运行维护，统一政府信息资源开放渠道，汇聚各区县、各部门开放的政府信息资源，提供上传、浏览、查询、下载、统计等服务，并强化数据应用，实现应用汇聚、管理等相关功能，为公众和企业提供各类数据服务”³。

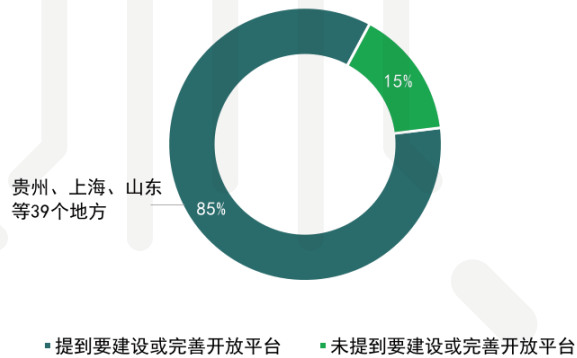


图84 各地开放平台建设政策情况分布

³ 上海市《关于推进政府信息资源向社会开放利用工作的实施意见》. <http://www.shetd.gov.cn/bsxxgk/668391.htm>

○ 开放数据资源管理

开放数据资源管理包括开放数据的范围、目录与清单编制、标准制定以及质量管理等方面，目前贵州、哈尔滨和贵阳等地对上述4项内容都作出了政策规定。

就开放数据的范围而言，目前只有贵州、哈尔滨、武汉等5个地方明确了其数据开放的范围，如哈尔滨规定，“凡不涉及国家秘密、商业秘密和个人隐私以及法律法规明确规定不予开放的政府数据资源，原则上均应以公众可获取、易加工的方式，原始、整体地向社会开放”⁴。其余地方或只提出要确定开放范围的政策要求，或仅确定了重点开放领域，还有少部分地方未对开放范围作出明确政策规定。

在开放数据目录与清单编制方面，目前已有过半数的地方提出要编制政府数据资源开放目录，实行动态清单管理。对于开放数据的标准，大部分地方提到要制定并组织实施工开放的标准，确保标准统一。目前只有北京、广东、贵州等7个地方对开放数据的质量作出相对完整的要求，提出要确保开放数据资源的完整性、准确性、有效性、时效性和可用性；还有少数地方只对“数据质量”的某些方面做出了规定。

○ 开放方式与机制

数据开放的方式与过程包括开放方式、沟通机制、纠错机制、人财物保障和监督考核机制等方面。目前各地在这方面的法规政策建设明显不足，只有贵阳对上述五项内容都作出了政策规定。

根据数据资源的体量、实时性、安全要求等特点并结合实际技术条件，明确数据开放的方式是推动政府数据开放的必然要求。在报告覆盖的46个地方中，目前只有贵州、江西、贵阳等3个地方对数据开放方式做出了清晰的界定。如《贵州省政务数据资源管理暂行办法》中规定，“政务数据

开放应按照无条件开放、依申请开放和依法不予开放三类进行管理：依法不予开放和需要特定相对人申请开放的政务数据，数据目录和限制开放依据应当向社会公布；非涉密但涉及敏感信息的政务数据，经过脱敏清洗后可根据使用条件和适用范围面向社会无条件或依申请开放”⁵。还有部分地方只是明确了数据资源要按照重要性和敏感程度分级分类进行开放，但没有明确具体的开放方式。

沟通机制是指政府要通过各种方式和渠道主动收集并及时回应数据利用主体的需求，体现“以需求为导向”的数据开放原则。目前只有贵州、北京、贵阳、苏州等4个地方对此做出了完整规定。浙江虽提出公共数据开放应当采取听证会、座谈会、征求意见等方式公开听取所涉及公众的意见、建议，但未对政府部门的回应与反馈提出要求。

纠错机制是指当数据利用用户对开放数据有疑义或者发现错误时，应当及时反馈政府部门，相关部门也应及时校核并予以反馈。目前，只有贵州和贵阳对此提出了政策要求，如《贵阳市政府数据共享开放实施办法》规定，“政府数据使用方对政府数据目录和获取的政府数据有疑义或者发现有错误的，应当及时通过共享平台或者开放平台反馈政府数据提供机关予以校核。政府数据提供机关应当自收到校核信息之日起5个工作日内完成校核，进行相应处理，并反馈政府数据使用方。校核期间，有关行政机关办理业务涉及公民、法人或者其他组织已提供合法有效证明材料的，应当照常办理，不得拒绝、推诿或者要求当事人办理数据更正手续”⁶。

人财物保障是指对政府数据开放工作给予必要的机构、人员、资金等方面的支持，是数据开放顺利推进所需的物质基础。在机构与人员方面，各部门应指定专门机构和专人负责本部门公共数据资源开放工作，数据开放主管部门应定期开展公共数据资源开放工作业务培训。如

⁴ 《哈尔滨市推进政府数据向社会开放工作实施方案》. http://www.harbin.gov.cn/art/2017/3/24/art_13791_2106.html

⁵ 《贵州省政务数据资源管理暂行办法》. http://www.gzgov.gov.cn/xxgk/zfxxgkpt/szfxgkml/aztfl/zhzw/201709/t20170919_793971.html

⁶ 《贵阳市政府数据共享开放实施办法》. <http://www.gygov.gov.cn/c9681/20180123/i1459759.html>

贵阳市规定,“市大数据行政主管部门应当定期组织行政机关工作人员开展政府数据共享开放更新维护、安全管理、操作流程、规范使用等专题培训,提高政府数据共享开放工作的业务能力和服务水平”⁷。在资金方面,要为公共信息资源开放工作提供有力经费保障,如上海市提出,“保障资金投入。市和区县有关部门要加大投入力度,将政府信息资源开放工作涉及的信息系统建设改造、信息资源汇集、加工、处理等费用纳入行政机关年度预算,坚持量力而行,尽力而为,按照工作计划,合理安排财政资金,保障相关工作正常进行”⁸。

监督考核是指对数据开放工作的效果进行公开考核,并对相关机构与工作人员的工作行为进行监督。在考核评估方面,上海、贵州、哈尔滨等地都提出要对数据开放工作的情况进行绩效考核,如贵阳市规定,“市人民政府应当制定考核办法,将政府数据共享开放工作纳入年度目标绩效考核,考核结果向社会公布。县级以上人民政府应当定期开展政府数据共享开放工作评估,可以委托第三方开展评估,结果向社会公布。鼓励第三方独立开展政府数据共享开放工作评估”⁹。就效能监督而言,目前浙江、贵州、贵阳等地提出由上级主管部门或其他监察部门对相关机关及工作人员一系列不符合规定的行为进行批评和处分。如贵阳市规定,“违反本条例规定,行政机关及其工作人员有下列行为之一的,由其上级机关或者监察机关责令限期改正,通报批评;逾期不改正的,对直接负责的主管人员和其他直接责任人员依法给予处分:(一)不按照规定建设共享平台、开放平台的;(二)不按照规定采集、更新政府数据的;(三)不按照规定编制、更新目录的;(四)不按照规定汇总、上报目录的;(五)提供不真实、不准确、不完整政府数据的;(六)不按照规定受理、答复、复核或者反馈政府数据共享或者开放需求申请的;(七)要求申请人重复提交能够通过共享平台获取政府数据的;(八)无故不受理或者处理公民、

法人和其他组织投诉举报的;(九)违反本条例规定的其他行为”¹⁰。

3. 数据利用

数据的增值开发与利用是数据开放的目的,在政策法规中要求政府部门采取各种措施推动开放数据的有效利用十分必要。

○ 宣传推广

宣传推广是指政府部门通过加强宣传引导,营造数据利用的良好氛围。如图85所示,在报告覆盖的46个地方中,有约三成的地方(如上海、浙江、广东等)提出要引导企业利用政府数据开展示范应用,积极营造社会广泛参与和利用开放数据的良好氛围。其余地方或仅笼统提出要引导社会主体利用政府数据,或未对此作出明确规定。

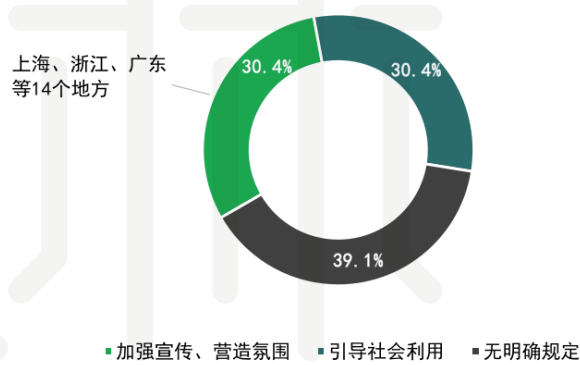


图85 各地宣传推广政策情况分布

○ 推动利用

通过支持数据应用竞赛、提供创业孵化、设立创新基金、减免税收等手段鼓励和支持政府数据的社会化利用有利于提高开放数据的利用效率与效果。如图86所示,目前有70%的地方提出支持企业利用政府开放的数据开发创新应

⁷ 《贵阳市政府数据共享开放实施办法》. <http://www.gygov.gov.cn/c9681/20180123/i1459759.html>
⁸ 上海市《关于推进政府信息资源向社会开放利用工作的实施意见》. <http://www.sheitc.gov.cn/bsxxgk/668391.htm>
⁹ 《贵阳市政府数据共享开放条例》. http://www.gysrd.gov.cn/News_show.aspx?xid=146&lmid=147&nid=7081
¹⁰ 《贵阳市政府数据共享开放条例》. http://www.gysrd.gov.cn/News_show.aspx?xid=146&lmid=147&nid=7081

用，其中上海、浙江、广东等地明确了推动创新利用的具体手段。如上海市规定，“鼓励和支持企业和公众挖掘利用政府和社会的开放数据资源，创新商业模式，开发互联网创新产品。组织开展大数据创业创新开发竞赛，支持大数据创新创业项目，激发创新创业活力。支持传统企业运用大数据推进互联网转型”¹¹。

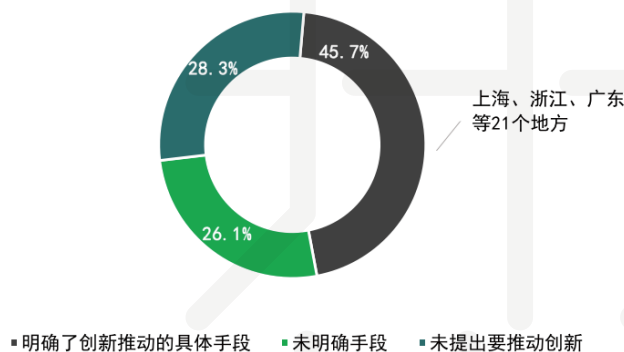


图86 各地推动利用政策情况分布

4. 安全保护

安全保护是数据开放的重要保障，保护是为了更好更可持续的开放。平台是数据开放的主要渠道，平台是否安全可靠直接关系到数据安全；数据在开放前进行安全审查为确保数据安全提供了屏障；对数据开放后利用情况进行追踪和管控则有利于降低安全风险。

○ 平台安全

平台安全管理是指开放平台运营主体根据国家信息安全等级保护的相关要求，通过完善身份认证、访问控制等技术防控措施，定期开展开放平台的系统安全检查，保障平台安全可靠运行。如图87所示，目前只有贵州、浙江、贵阳等3个地方对加强开放数据平台的安全和防御做出了明确规定，如贵州提出要构建平台的安全防护体系和认证体系¹²；

贵阳提出要建立平台政府数据存储和备份机制，防止数据丢失和毁损¹³。

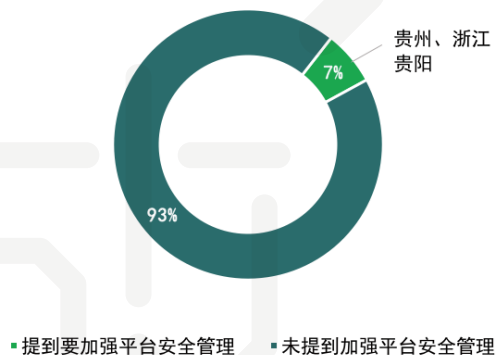


图87 各地平台安全政策情况分布

○ 数据审查

数据资源在开放前要依照相关法律法规进行严格审查，未经安全审查的数据资源不予对外开放。如图88所示，目前只有上海、浙江、贵州等9个地方对此作出了明确的政策要求，如浙江省规定，“公共数据和电子政务主管部门审核公共数据开放时，要对涉及的国家安全、信息风险、社会效益等进行审核”¹⁴。

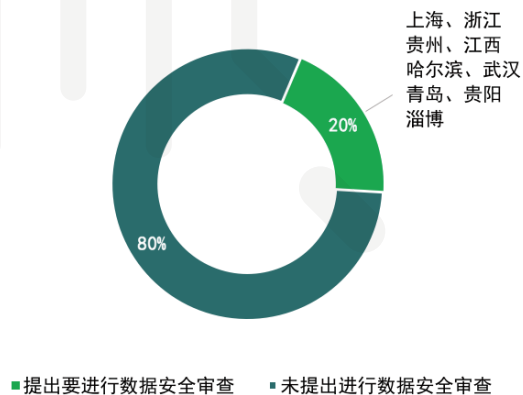


图88 各地数据审查政策情况分布

¹¹ 《上海市大数据发展实施意见》. <http://www.shanghai.gov.cn/nw2/nw2314/nw2319/nw12344/u26aw50056.html>.

¹² 《贵州省政务数据资源管理暂行办法》. http://www.gztyw.gov.cn/zhxw/szfwj/201611/t20161114_1358738.html

¹³ 《贵阳市政府数据资源管理办法》. <http://www.gygov.gov.cn/c1251/20171205/i1415800.html>

¹⁴ 《浙江省公共数据和电子政务管理办法》. http://zfxgk.zj.gov.cn/xxgk/jcms_files/jcms1/web14/site/art/2018/2/1/art_1082_1987774.html

○ 风险管理

风险管理是指要定期对数据开放后的利用情况与多源数据融合利用后的风险进行评估和管理，以确保数据安全。例如，贵州省规定，“实行公共数据共享开放风险评估制度，对通过平台开放的公共数据进行风险审核，发现可能存在风险时，应当及时告知提供单位，提供单位应当及时处理并予以反馈”¹⁵。如图89所示，除贵州、江西、武汉、贵阳外，绝大多数地方都未对此作出明确规定。

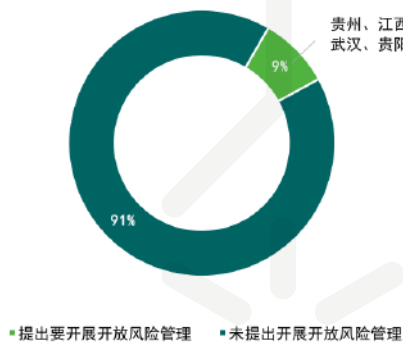


图89 各地风险管理政策情况分布

5. 法规政策效力

法规与政策效力是指法规和政策所具有的约束力，主要表现在效力等级和专门性两方面。本部分的评估范围只包括提及政府数据开放的法规与政策。

○ 等级

法规与政策的效力等级直接决定了其约束力的强弱。报告将各地提及政府数据开放的法规与政策的效力从高到低划分为地方性法规、地方政府规章、规范性文件与其他普通政策四个等级。图90展示了各地制定政策与法规的效力等级，目前只有贵州和贵阳通过地方性法规为当地政府数据开放提供了强有力的立法保障，其中《贵州省大数据发展应用促进条例》是中国首部大数据地方法规。同时，个别

地方(如贵阳和浙江)制定了地方政府规章；贵州、北京、江西等7个地方制定了规范性文件；除个别地方无任何政策外，其余地方均以各种实施意见、发展规划、实施方案等形式的普通政策为政府数据开放进行引导和规范。

地区	地方性法规	地方政府规章	规范性文件	其他普通政策
贵州	●		●	●
贵阳	●	●		●
浙江		●		●
北京			●	●
江西			●	●
山东			●	●
长沙			●	●
威海			●	●
滨州			●	●
其他				●

注：●表示该地制定了该效力等级的政策。

图90 各地法规政策效力等级详情列表

○ 专门性

法规和政策是否专门为政府数据开放而制定也是决定其效力的重要因素。报告以法规或政策标题中是否同时具有“数据/信息资源”与“开放”为标准，来评估某法规或政策是否专门针对政府数据开放。对于不是专门针对数据开放的政策，又根据内容的丰富程度划分为两类情况。

如图91和表6所示，目前只有贵阳、上海、青岛、威海、哈尔滨等5个地方制定了专门性的法规或政策。特别是贵阳出台了三项专门性法规或政策，其中《贵阳市政府数据共享开放条例》是我国首部也是唯一一部专门针对数据开放的地方性法规。同时，大多数地方出台的政策虽然没有专门针对数据开放，但是提及了数据开放、数据利用、安全保护等多方面内容，内容较为丰富的有北京、浙江、广东等地。此外，还有一些地方出台的政策中只涉及了政府数据开放的某些方面，内容较为单薄。

¹⁵ 《贵州省大数据发展应用促进条例》. <http://www.gzrd.gov.cn/dffg/sdgdfxfg/23501.shtml>

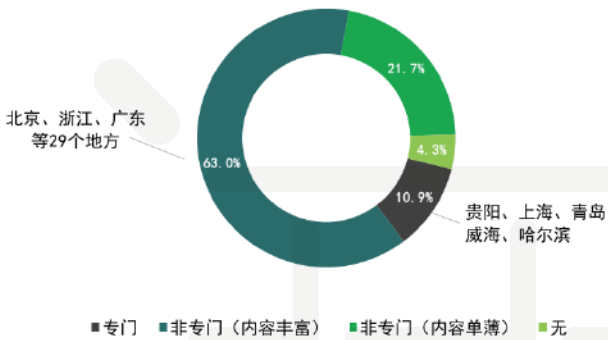


图91 各地法规政策专门性详情列表

表6 各地专门性政策详情列表

地区	法规政策
贵阳	《贵阳市政府数据共享开放条例》 《贵阳市政府数据共享开放实施办法》 《贵阳市人民政府关于贵阳市加快推进政府数据共享开放的实施意见》
上海	《关于推进政府信息资源向社会开放利用工作的实施意见》
青岛	《青岛市人民政府办公厅关于加快推进公共信息资源向社会开放的通知》
威海	《威海市人民政府办公室关于推进政府数据资源向社会开放利用工作的实施意见》
哈尔滨	《哈尔滨市推进政府数据向社会开放工作实施方案》

• 3.4.2 领导力

领导力指领导层对政府数据开放的支持与推动。报告通过公开报道的地方高层领导和部门最高领导对数据开放的表态支持情况来衡量地方政府数据开放的领导力。

1. 地方高层领导的公开支持

政府数据开放的顺利推进离不开强有力的领导推动，地方高层领导更是发挥着巨大的作用。报告将地方高层领导限定于政府(正副)行政首长¹⁶，如图92所示，上海、浙江、贵州、山东等12个地方得到了当地政府最高行政首长在公开场合的发声支持，如在“2015杭州·云栖大会”上，时任浙

江省长李强曾强调了开放数据的巨大价值，并作出了进一步加快政府数据开放的承诺。

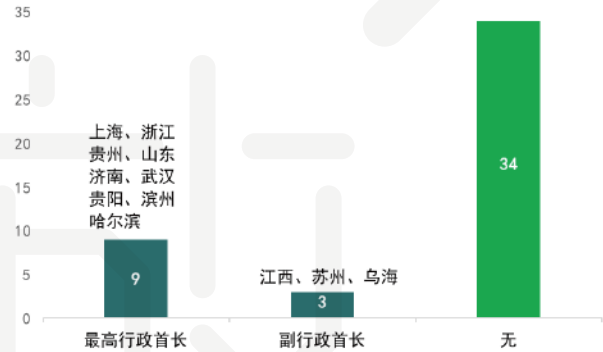


图92 各地地方高层领导支持情况分布

2. 部门负责人的公开支持

各地数据开放的主管部门是具体的推动者、执行者与协调者，部门负责人的意愿与态度直接影响着政府数据开放工作的效率与效果。如图93所示，目前只有贵阳、北京、东营等3个地方主管部门的负责人公开表态支持政府数据开放，如2016年8月，时任北京市经信委主任张伯旭分析了数据开放的价值，承诺到2020年本市公共数据开放单位将超过90%，并且这些数据都将向社会免费开放，数据开放率将超过60%。

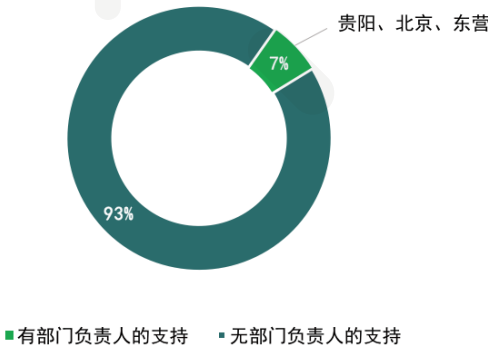


图93 各地部门负责人支持情况分布

¹⁶ 包括（正副）省长、自治区区长、直辖市市长、地级市市长等

• 3.4.3 组织保障

健全的组织保障体系是数据开放长期可持续推进的保障。主管部门是主要的推动者与实施者，是政府数据开放重要的组织基础；公开具体的计划既有助于政府自身厘清阶段性的目标与职责，也能给予有数据需求的组织和个人以明确的预期，并接受社会公众的监督；数据生态体系的培育有助于提高开放数据的利用效率，发挥数据的真正价值。

1. 主管部门层级

负责推进数据开放工作的主管部门所处的相对行政层级及其与数据提供部门之间的关系决定着该部门推动政府数据开放的力度。报告将当前各地开放数据主管部门相对层级分为了一级委办局、委办局下属二级局、政府办公厅或委办局内设处室、事业单位四种情况，其协调力度依次递减。

如图94所示，(1)目前只有贵阳市专门成立了一级委办局“贵阳大数据发展管理委员会”来统筹数据开放，为当地政府数据开放工作的有效推进奠定了坚实的组织基础。此外，贵州省成立的大数据发展管理局虽为省政府直属事业单位，但实际上履行的是行政职能，据公开资料显示，其职能包括研究拟订并组织实施战略、规划和政策措施等，促进政府数据资源开放等¹⁷，而且该局局长同时担任省政府秘书长¹⁸，具有较强的行政协调能力，因此报告也将贵州省大数据发展管理局将归为这一类别；(2)广东、广州、哈尔滨、宁波等四个地方则在经（工）信委下成立了名为“大数据管理局”的二级局；(3)目前，大多数地方是将主管政府数据开放的职能交由政府办公厅、经(工)信委(局)、或网信办的某个内设处室，比如浙江省政府办公厅电子政务处、上海市经信委信息化推进处(大数据发展处)、武汉市网信办信息化发展处等。

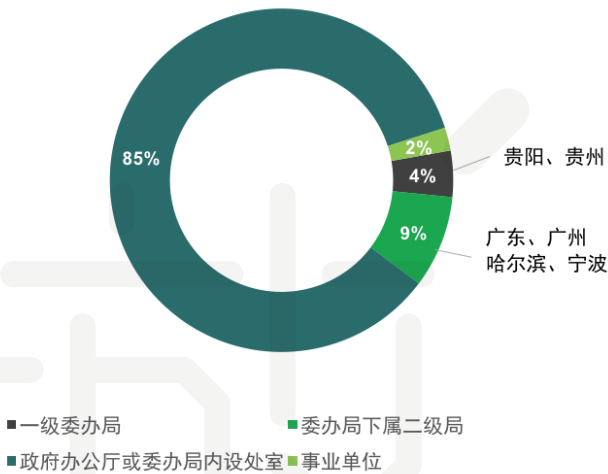


图94 各地主管部门层级情况分布

2. 公开工作计划

工作计划是主管部门制定的、对一定时期的工作预先作出安排和打算的一种公文形式。制定有针对性的、切实可行的开放数据工作计划并向社会公开是推进政府数据开放工作的重要内容。

目前各地政府数据开放的工作计划主要表现为各种工作方案、实施方案、年度工作计划等。如图95所示，根据工作计划的专项性和内容的丰富性，报告将各地公开的工作计划分为专项工作计划、非专项工作计划（内容丰富）、非专项工作计划（内容单薄）三类情况。报告发现，目前只有少部分地方公开了其数据开放工作计划。其中，只有上海制定并公开了专项工作计划。事实上，自2014年起，上海已经连续四年制定《政府数据资源向社会开放年度工作计划及重点领域》，明确了该市政府数据开放的总体思路、重点工作任务、工作要求以及具体分工等，并公开了详细的数据资源开放清单，可操作性强，为当地政府数据开放奠定了有效的管理基础。同时，部分地方将数据开放作为政务公开、政务信息系统整合共享等相关领域工作的一项重要任务，制定了非专项工作计划。其中，广东、山

¹⁷ 贵州省大数据发展管理局主要职责内设机构和人员编制规定。 http://www.gzgov.gov.cn/xxgk/jbxxgk/fgwj/szfwj_8191/qbf_8196/201709/t20170925_823950.html

¹⁸ 来自http://www.gzgov.gov.cn/xxgk/zfxgkpt/szfxgkml/ajcfl/gk/ldjjfjg/201709/t20170928_1051079.html

东、青岛等17个地方的工作计划内容较为丰富，如明确了年度开放目标和主管单位、提出要制定专项行动计划和数据资源开放目录、规范政务数据开放流程、建设或完善开放平台等。此外，个别地方也制定了相关的工作计划，但内容相对单薄。

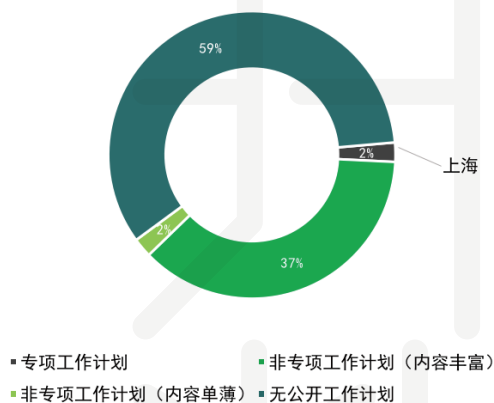


图95 各地公开工作计划情况分布

各种形式的宣传交流活动，多围绕数据创新应用的主题，汇聚来自政府、学界、企业、国内外大数据组织等的相关人士广泛参与，通过政府政策指导、学界理论支撑、企业经验分享与应用展示等，提供了一个良好的宣传交流与应用推广平台，有利于引导和鼓励全社会挖掘数据价值、开发数据应用等。

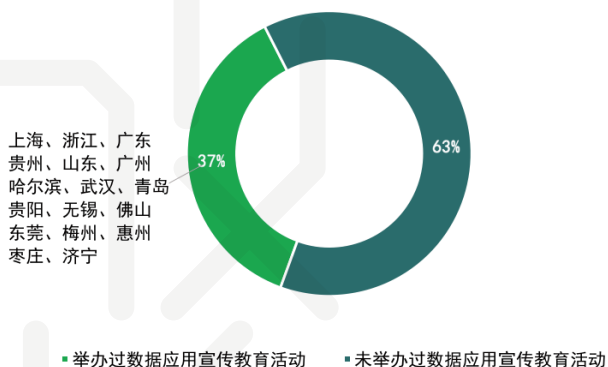


图96 各地数据应用宣传教育活动举办情况分布

3. 培育生态体系

培育生态体系是指政府部门开展了实际工作来营造社会各界广泛参与和利用开放数据的良好氛围，培育良性互动的开放数据生态体系，以有效发挥开放数据的经济、政治和社会价值。本报告重点考察了各地政府部门指导或牵头举办的一系列数据应用宣传交流和创新活动。

○ 公开报道的数据应用宣传教育交流活动

据网上公开报道，目前各地的数据应用宣传交流活动主要以数据应用大会(峰会)和中小型展会为主要形式。如图96所示，目前有17个地方举办并公开报道过相关活动，如贵州、上海、山东等地都举办了大数据高峰论坛。此外，贵阳还通过举办数据开放平台展会的形式加强对政府数据开放的宣传教育和交流。

○ 公开报道的数据创新应用活动

据网上公开报道，目前各地的数据创新应用活动主要以聚焦数据开放和利用的数据应用大赛为主要形式。大赛以数据开放为切入口，通过政府搭台、政企合作等，发动公众参与和协作来共同解决公共问题。这有助于吸引和集聚数据利用人才，培育数据开放利用生态系统，从而激发创新活力、促进产业发展、提升城市服务水平和管理能力等。

据网上公开报道，如表7所示，目前只有北京、上海、广东等9个地方举办过数据开放创新应用大赛。其中，上海市开放数据创新应用大赛(SODA)¹⁹在国内可谓首创，从2015年至2017年已经连续举办了3年，政府开放高质量的数据，吸引了众多参赛者提供大量有特色的创意方案。

¹⁹ 更多内容可关注微信公众号“SODA上海开放数据创新应用大赛”（SODACHallenge）

表7 各地数据创新应用活动举办情况列表²⁰

省/市	活动时间	活动名称
上海	2015 / 2016 / 2017	上海市开放数据创新应用大赛（SODA）
北京	2014	2014年北京市政务数据资源应用创意大赛
广东	2017	2017年广东政务数据创新大赛
贵州	2017	2017数博会最佳大数据挖掘分析（工具）大赛
	2017	2017中国“云上贵州智慧交通大数据应用创新大赛”
哈尔滨	2017	2017哈尔滨“浪潮杯”大数据应用创新大赛
深圳	2017	2017深圳“中国电科杯”城市数据创新大赛
无锡	2016	2016无锡大数据应用创新大赛
	2018	国土资源大数据应用创新大赛
威海	2015	2015威海互联网+数据开放创业创新大赛
贵阳	2016	贵阳交通大数据系列竞赛

²⁰ 每个地方只按级别最高的大赛打分

• 3.4.4 准备度层结论

1. 相关法规与政策“散、弱、空、同”

总体而言，绝大多数地方都已出台了与政府数据开放相关的法规和政策，为数据开放工作奠定了法治基础，提供了政策依据，涉及到数据治理、数据开放、数据利用、安全保护等各个环节。然而，这些法规政策还存在一些普遍问题。

○“散”。目前只有贵阳、上海、哈尔滨、青岛、威海五个地方出台了专门针对政府数据开放的政策，其余各地政策过于分散，普遍缺乏专门性的政策对数据开放的各个方面进行统一规范和引导。

○“弱”。目前只有贵州、贵阳和浙江将政府数据开放纳入具有较高法律效力的地方性法规或地方政府规章规范，少部分地方制定了规范性文件，其余多数地方政策的效力等级较低，约束力不足。

○“空”。当前多数地方的数据开放法规政策多以原则性内容为主，政策规定缺乏指导性与可操作性。

○“同”。目前各地开放数据领域的政策内容存在同质化的现象。尽管后发地区借鉴先发地区的政策文件有利于迅速学习追赶，但不考虑本地实际直接照搬其他地方的政策内容，可能会使本地的数据开放工作缺乏针对性与适用性，甚至流于形式。

2. 缺乏高层领导和部门负责人的公开支持和推动

强有力的领导重视和支持有利于政府数据开放工作的顺利开展与可持续推进。报告发现，尽管已有北京、上海、浙江等十多个地方的高层领导和部门负责人公开表态支持政府数据的开放与社会化利用，但更多地方仍缺乏领导层公开、明确的支持。

3. 主管部门协调力度不够，缺乏专项工作计划，忽视生态体系培育

目前，虽有个别地方在组织保障方面整体表现较为突出，但大部分地方存在不少问题：

○主管部门职能模糊，协调能力弱。一方面，部分地方未在其政府门户网站或政府数据开放平台上明确其主管部门的数据开放职能。另一方面，除极少数地方(贵州和贵阳)专门成立了协调力较强的委办局来主管数据开放和开发利用工作，绝大多数地方只是将数据开放职能交由当地政府的办公厅、经(工)信委或网信办内主管信息化、电子政务的二级处室，其行政级别与其他数据提供部门相同，大都面临协调困难、工作难以推进等问题。

○缺乏专项工作计划。政府数据开放需要分步骤、有计划地专项推进。目前虽然有19个地方制订了公开的数据开放工作计划，但只有上海的工作计划专门针对政府数据开放，其余地方的工作计划均为非专项。

○忽视生态体系培育。培育数据利用生态体系，营造良好的社会氛围是数据开放准备度的重要方面，但目前只有少部分地方政府对培育数据开放生态体系有一定程度的重视，大部分地方的相关意识仍较为薄弱。

4

“开放数林”指数分值与分析 Index and Ranking



◆ 4.1 指数分值与排名

• 4.1.1 省级

省级开放数林指数分值与排名如表8所示，上海排名第一，贵州和山东分列第二、三位。从指数得分来看，上海与贵州处于第一梯队，山东、广东、北京三地位于第二梯队，江西、浙江、宁夏等属于第三梯队。

表8 开放数林指数与排名（省级政府）

地方	数据层指数	数据层排名	平台层指数	平台层排名	准备度指数	准备度排名	开放数林指数	总排名
上海	59.50	1	54.30	2	69.76	1	60.76	1
贵州	55.35	2	54.30	2	68.86	2	58.47	2
山东	40.10	5	58.90	1	59.20	3	49.57	3
广东	44.26	3	42.60	6	45.50	6	44.16	4
北京	41.56	4	43.90	5	43.88	7	42.73	5
江西	28.71	6	39.40	7	47.00	5	35.95	6
浙江	27.57	7	23.80	8	57.44	4	34.09	7
宁夏	22.89	8	44.10	4	25.00	8	28.72	8

• 4.1.2 地市级（含副省级）

地市级（含副省级）开放数林指数分值与排名如表11所示，贵阳排名第一，遥遥领先，处于第一梯队，哈尔滨、青岛、广州、济南处于第二梯队，佛山、滨州、深圳、东营、潍坊等处于第三梯队。进入前十的城市中山东共占据5席，广东占据3席。

表9 开放数林指数与排名(地市级政府, 含副省级)

地方	数据层指数	数据层排名	平台层指数	平台层排名	准备度指数	准备度排名	开放数林指数	总排名
贵阳	85.96	1	89.90	1	89.80	1	87.90	1
哈尔滨	52.34	3	55.30	6	61.38	2	55.34	2
青岛	47.77	5	65.10	2	44.08	5	51.18	3
广州	56.12	2	48.10	24	35.50	11	48.96	4
济南	41.18	8	60.50	3	41.80	7	46.17	5
佛山	51.80	4	55.20	7	21.40	34	45.05	6
滨州	32.69	18	54.60	8	52.96	3	43.23	7
深圳	43.93	7	51.80	23	29.84	18	42.37	8
东营	34.11	15	54.60	8	41.60	8	41.10	9
潍坊	35.23	12	54.60	8	30.60	15	38.92	10
泰安	36.37	9	54.60	8	26.40	24	38.43	11
东莞	44.69	6	28.00	28	36.28	9	38.42	12
淄博	29.80	25	54.60	8	35.72	10	37.48	13
宁波	35.56	11	46.80	26	31.90	14	37.45	14
德州	33.98	16	54.60	8	24.20	30	36.69	15
威海	24.15	35	54.60	8	43.68	6	36.65	16
乌海	29.58	26	56.20	5	30.40	16	36.44	17
江门	29.89	24	57.10	4	26.60	23	35.87	18
枣庄	31.42	20	54.60	8	26.00	25	35.86	19
临沂	33.37	17	54.60	8	21.80	32	35.79	20
莱芜	31.34	21	54.60	8	23.88	31	35.29	21
菏泽	36.24	10	54.60	8	13.80	35	35.22	22
济宁	29.92	23	54.60	8	25.60	26	35.01	23
烟台	30.06	22	54.60	8	24.60	28	34.83	24
日照	34.70	14	54.60	8	13.20	36	34.30	25
惠州	28.85	27	48.10	24	29.20	19	33.75	26
武汉	23.54	36	36.10	27	45.00	4	32.05	27
中山	34.72	13	26.90	29	30.40	16	31.68	28
聊城	27.84	29	54.60	8	13.20	36	30.87	29
苏州	31.69	19	13.10	34	35.48	12	27.99	30
无锡	28.14	28	14.80	32	33.60	13	26.17	31
梅州	26.08	31	13.70	33	28.40	21	23.57	32
长沙	27.27	30	10.70	37	24.60	28	22.46	33
扬州	22.94	37	18.50	30	25.40	27	22.45	34
荆门	24.85	33	11.80	36	27.44	22	22.23	35
湛江	24.64	34	17.80	31	21.80	32	22.22	36
肇庆	24.88	32	9.40	38	29.16	20	22.08	37
阳江	19.27	38	12.70	35	7.80	38	14.76	38

◆ 4.2 指数综合分析

• 4.2.1 空间分析

省级开放数林指数得分的空间分布状况如图97所示，图98则反映了当前地市级（含副省级）开放数林指数的空间分布情况。

从省级开放数林指数空间分布来看（图97），整体表现较好的省份多位于我国东部、西南和华南地区。与去年相比，贵州和广东进步明显，山东虽最新上线，但表现抢眼。

从地市级（含副省级）的空间分布来看（图98），山东省和广东省内的地市已成为两片较为集中的省域“数林”。佛山市南海区在广东省内最早上线数据开放平台，之后省内各地市陆续跟进，主要集中在珠三角地区和南部沿海。同时，广东各地市的数据开放模式较为多元，呈现出一片自由生长的“自然林”。这一模式有利于地方创新，但也同时出现地市级间发展不均衡、参差不齐的问题。

山东省则表现为另一种不同的模式。山东省是首个省内各地市全面推出数据开放平台的省份。除青岛市平台较早上线外，山东省省级平台与下属各地市平台采用统一标准，集中建设，同时上线。各地市数据开放水平普遍较高，省级与地市级平台之间互相连接，形成“群落”。山东省内各地市整体呈现为一片整齐划一的“人工林”。这种模式有利于提升省内各地市政府数据开放的整体水平，推动基础薄弱的地市同步开放，节省各地市建设成本，但也容易出现各地“千城一面”、特色不足、抑制创新等问题。

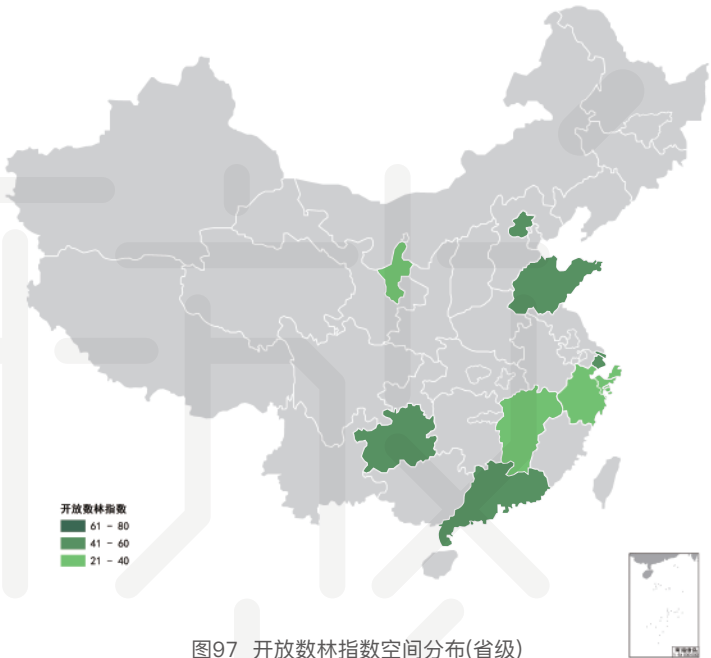


图97 开放数林指数空间分布(省级)

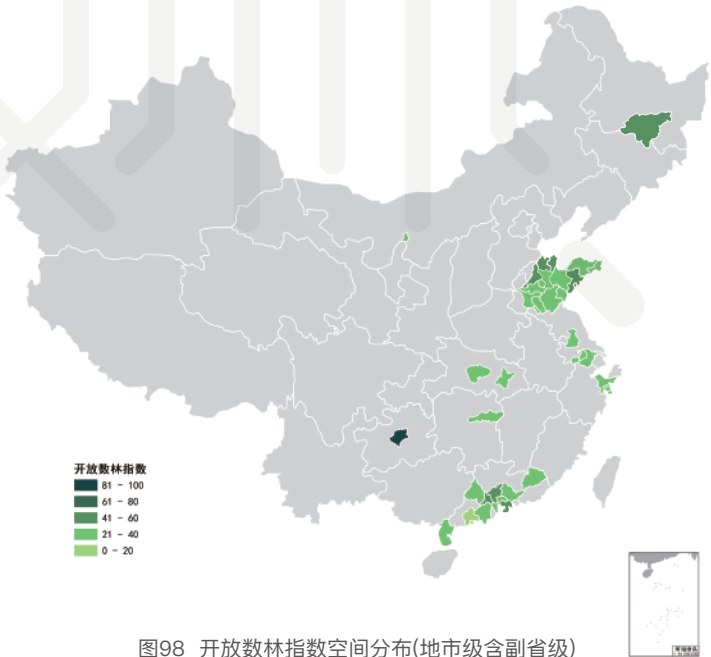


图98 开放数林指数空间分布(地市级含副省级)

• 4.2.2 长势分析

数据层是整体评估框架中权重最高的维度，已连续两年开展评估。报告对去年下半年评估报告中的20个¹地方平台自去年评估以来数据层指数的排名变化情况进行了跟踪分析。图99中的横坐标代表这20个地方2018年的下半年数据层指数分值，纵坐标代表这些地市数据层自去年下半年评估以来的排名变化。数据显示，2017年下半年以来，排名上升幅度最大的地方是贵州，其次为哈尔滨、广东、广州、深圳、佛山和贵阳等地，其余地方的排名均有不同程度的下降。

图中用排名不变(即为0)与指数得分均值作为两条分界线，将各地分为三种类型。位于第一象限(右上方)的地方为“竹林型”，表现为指数较高且上升迅速，包括贵阳、广州、贵州、哈尔滨、佛山、广东、深圳等地；位于第四象限(右下方)的地方为“树林型”，虽排名稳定或小幅下降，但指数仍维持在较高水平，主要包括上海、青岛、东莞、北京等地；而位于第三象限(左下方)的地方则表现为排名下滑且指数处于中下水平，为“灌木丛型”。

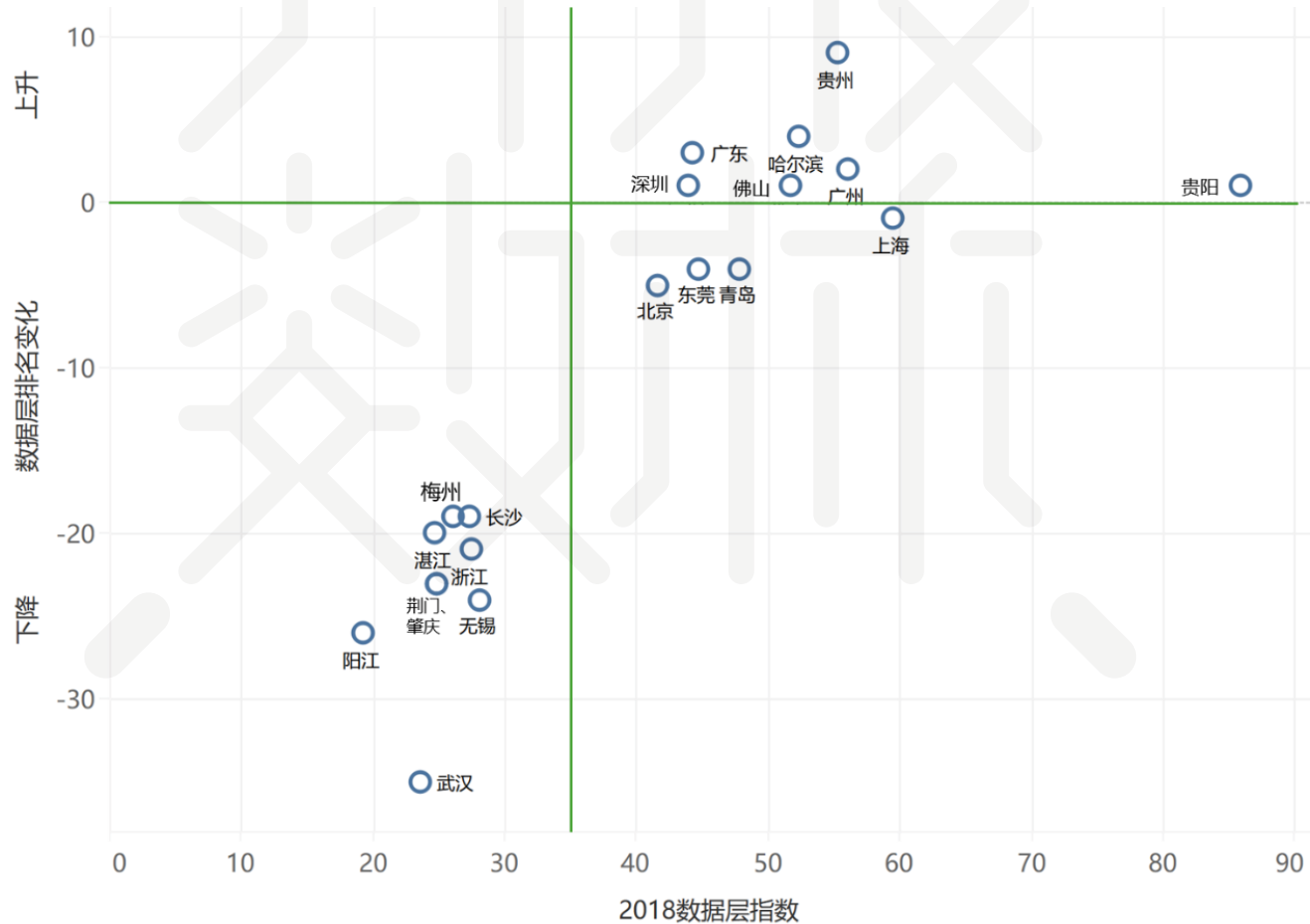


图99 20个地方数据层指数分值与排名变化分析

¹ 去年下半年评估对象共21个，其中佛山南海区为区县级政府，本次未做跟踪，故为20个。

报告发现数据层、平台层和准备度三个维度之间存在一定程度的相关性。如图100所示，数据层指数与平台层指数曲线如藤蔓一样上下“缠绕”着准备度曲线，其中数

据层指数曲线与准备度指数曲线的关系尤为紧密。经统计学检验，数据层指数与准备度指数的相关系数为0.66，平台层与准备度两个指数的相关系数为0.33，都在0.01水平上显著，也就是说，一个地方的准备度指数越高，其数据层和平台层也很有可能表现地更好，准备度的高低会影响数据层与平台层的表现。

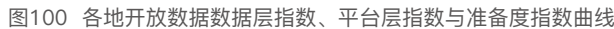


图100 各地开放数据数据层指数、平台层指数与准备度指数曲线

2. 形态特征

以上综合分析表明，准备度是数据层与平台层的基础，准备度犹如树木的根系，根系发达才能有粗壮的枝干与茂密的枝叶。平台层与数据层则相互依赖，平台是树木的枝干，数据是枝干上长出的树叶，没有枝干树叶则无法附着与生长，没有树叶枝干也将成为枯枝。报告根据各地在三个维度上的不同表现，发现了七种“数木”形态。

○ “乔木”型

乔木根系发达、躯干粗壮，枝叶茂密葱茏，该类地方在三个维度上均表现优异，呈现出均衡发展的态势。乔木型地方的代表为贵阳、上海和贵州等地(图101)。

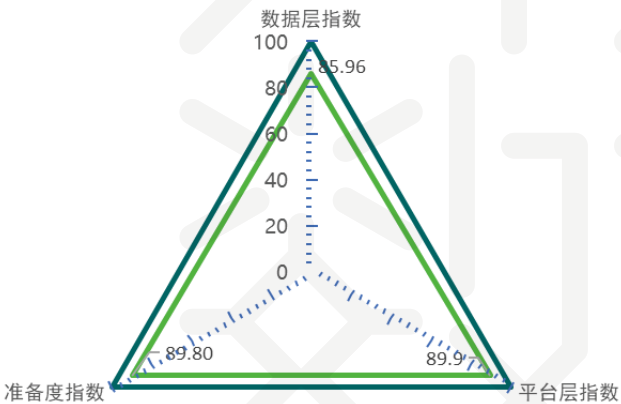


图101 “乔木”型(贵阳)

○ “水培”型

水培植物枝叶繁茂,但缺乏发达的根系。该类地方数据层平台层良好,准备度一般,可能影响其发展后劲(图102)。

○ “仙人掌”型

仙人掌根系发达、枝干粗壮，却缺少茂密的树叶。该类地方准备度与平台表现俱佳，但数据层表现平平，有待发力拿出干货(图103)。

○ “绿萝”型

绿萝叶片繁密，根基较深，但缺少粗壮的茎干，该类地方数据层和准备度表现良好，但在平台层方面尚表现一般，可能影响用户获取和利用数据的体验和效果(图104)。

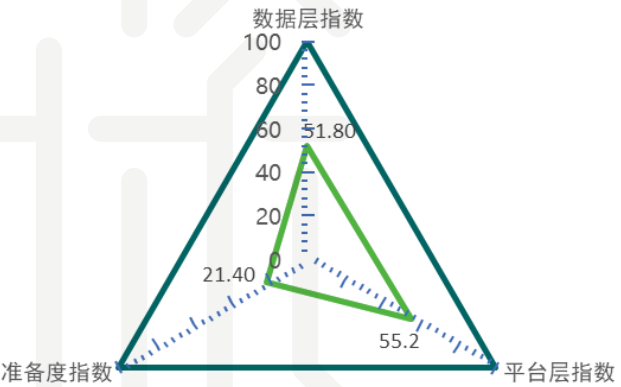


图102 “水培”型(佛山)

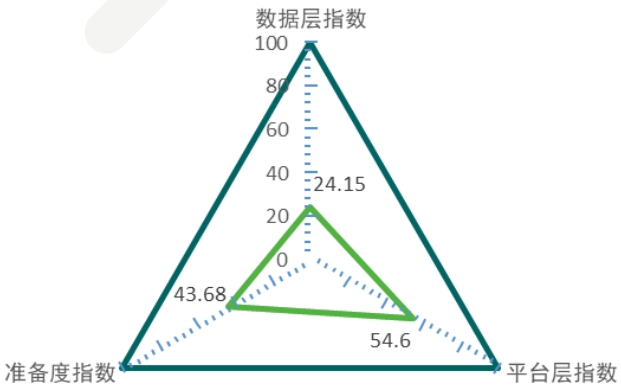


图103 “仙人掌”型(威海)

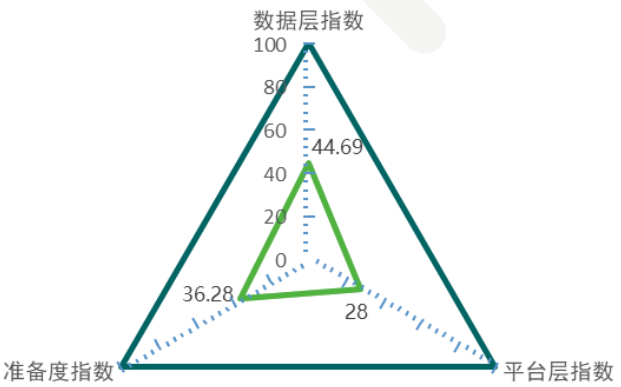


图104 “绿萝”型(东莞)

○ “新栽”型

新栽下的树木虽有躯干，但还缺乏发达的根系与枝叶。该类地方新建了数据开放平台，但在准备度与数据层方面还没有较好表现，还有待夯实基础，做实数据开放(图105)。

○ “树根”型

树根具有发达根系，还未成长出枝叶。此类地方准备度不错，但数据层与平台层比较薄弱，期待早日枝繁叶茂，开花结果(图106)。

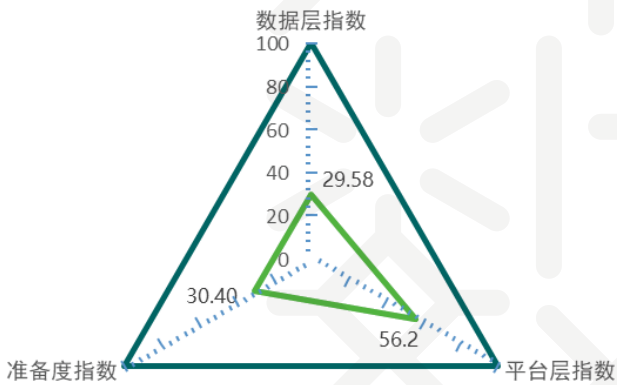


图105 “新栽”型(乌海)

○ “柳条”型

柳条树叶繁密，但还缺少发达的根系和坚实的茎干，该类地方数据层表现较好，但在平台层与准备度方面尚表现平平，可能影响发展后劲和用户体验(如图107)。

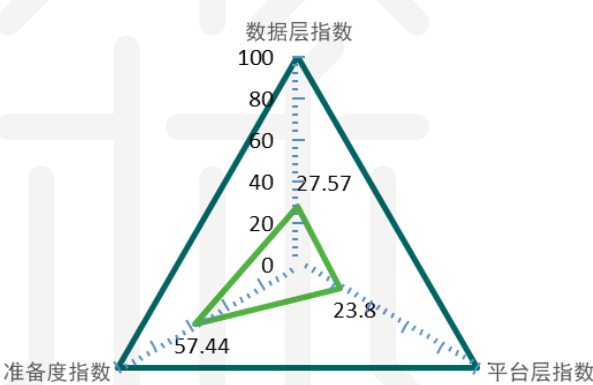


图106 “树根”型(浙江)

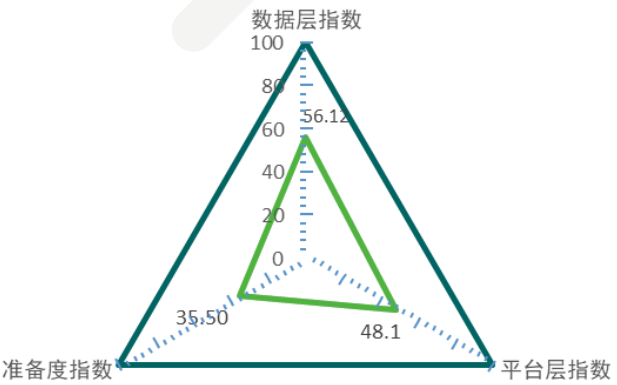


图107 “柳条”型(广州)

• 4.2.4 叶脉分析

各地在数据层指标中得分的均值与标准差如图108所示。均值得分高低代表各地在该项指标上平均达标率的高低，标准差大小则反映了各地在该项指标上表现的差异程度大小。总体上看，各个维度上都缺少达标率高且地区间差异小的指标，数据层中有不少指标整体达标率较高，但是地区间差异显著；平台层和准备度中大部分指标整体达标率不高，且地区间差异明显。

从图108可见，数据层中整体达标率较高且地区间差异适中的指标主要集中在图中右侧偏上的位置，包括无低容量数据和无碎片化数据，说明各地数据中低质量数据的整体比例不高；此外还包括基本元数据覆盖率和部门覆盖率等指标。达标率较低且地区间差异不大的指标主要集中在图中的左上角，包括固定链接、动态更新、开放格式、可机读格式、优质数据和数据集容量等。散落在图中右下角的指标整体达标率虽然较高，但地区间差异较大。

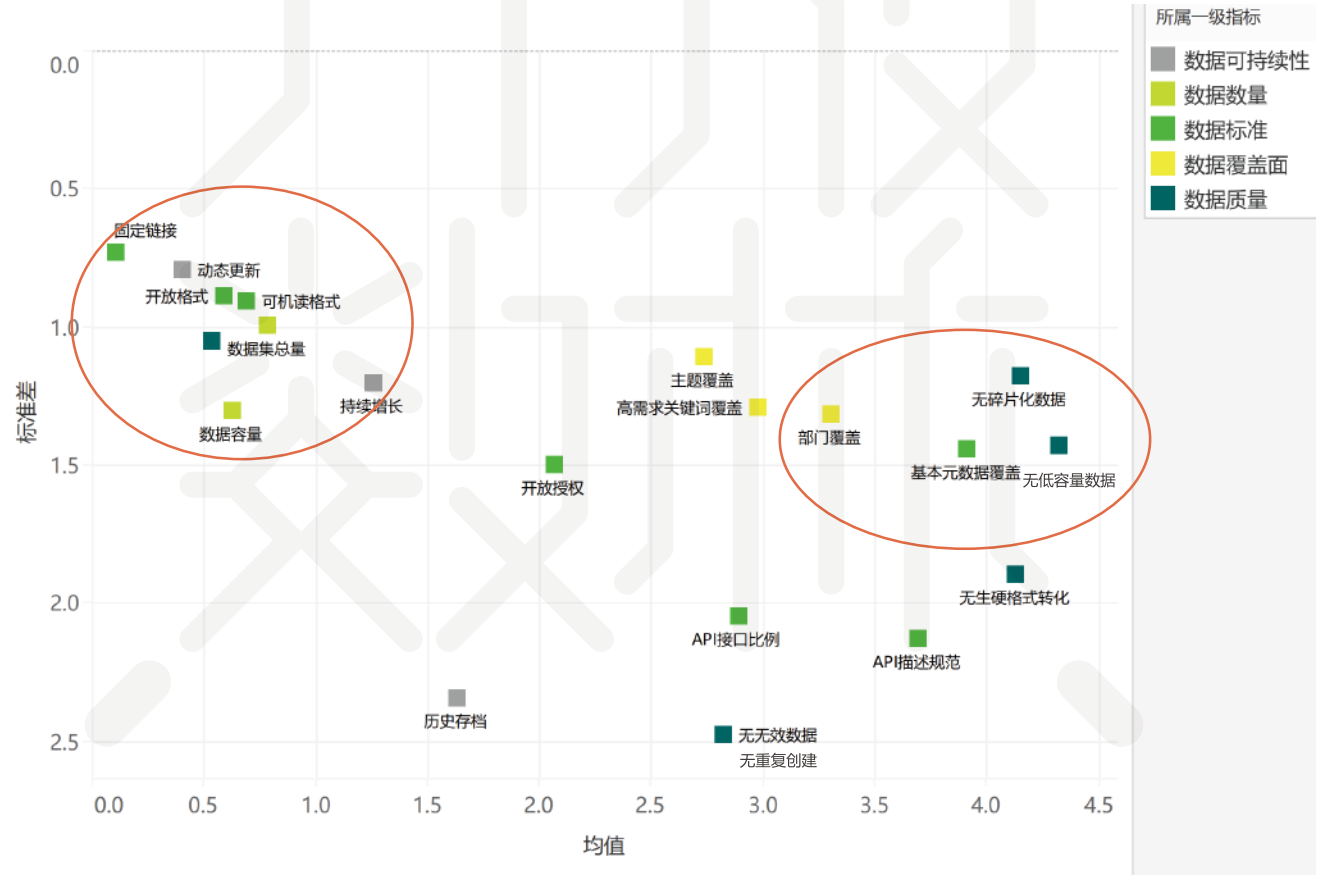


图108 数据层各项指标得分均值与标准差

如图109所示，平台层中整体达标率中上且地方间差异较小的指标主要集中在图中右侧偏上的区域，包括动态展示、分类导航和搜索功能等。整体达标率低且地区间差异不大的指标主要位于图中左侧偏上的区域，包括开放数据目录、智能服务、社交媒体、标明数据来源、数据纠错和数据请求功能等。达标率低且地区间差异较大的指标则位于图中

左侧偏下的区域，具体有分级分类获取、提供数据发布者联系方式、平台间关联三项。达标率高且地方差异较大的指标集中出现在右下角，包括数据本地获取、提供地理空间工具、订阅收藏、用户个性化功能整合和排序功能等指标。

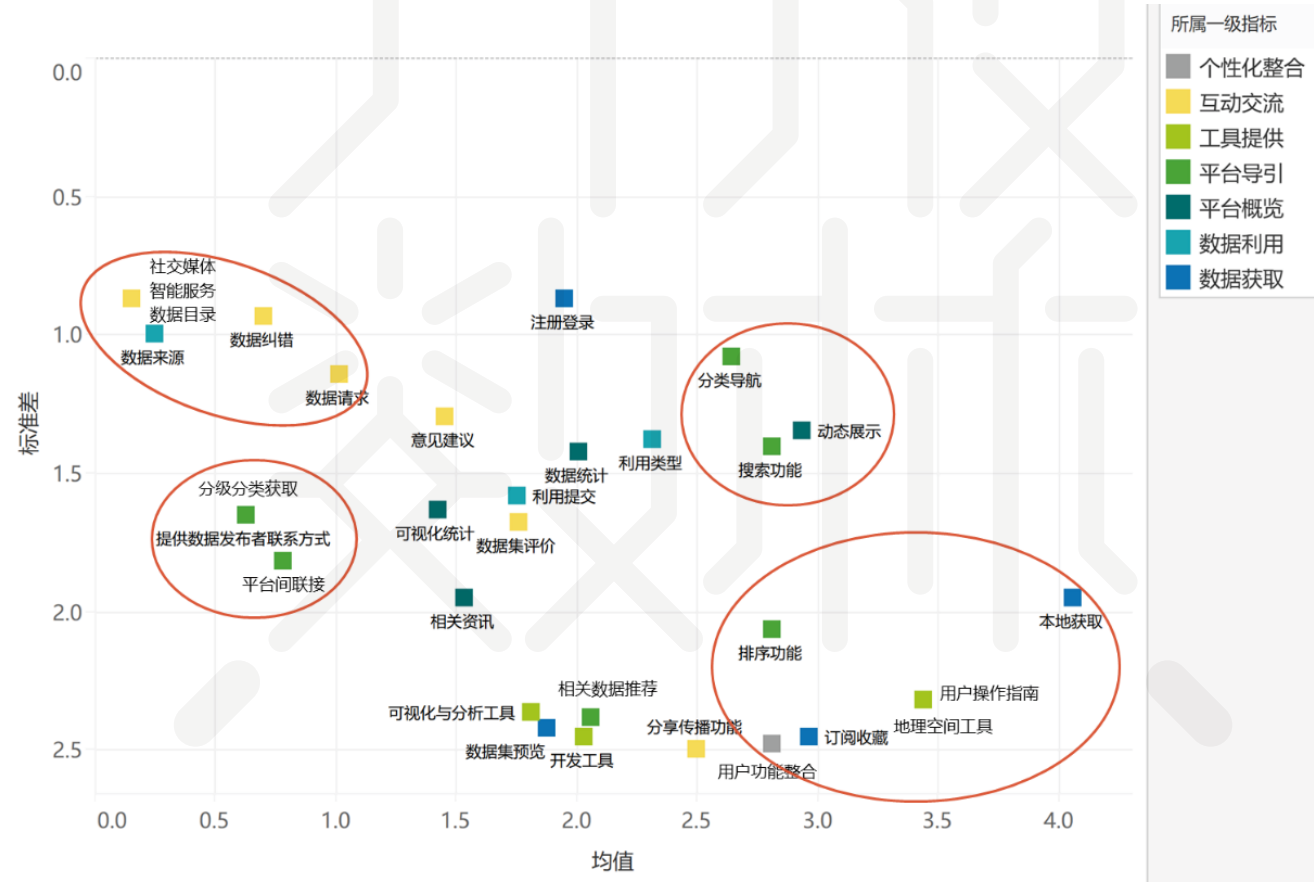


图109 平台层各项指标得分均值与标准差

从图110可见，准备度中达标率中上且地区间差异较小的指标主要是主管部门层级。整体达标率低且地区间差异不大的指标主要位于图中左侧偏中上的区域，包括法规政策中对纠错机制、平台安全、人财物保障等方面的要求和部门领导支持。整体达标率低且地区间差异较大的指标位于图中左下角，包括宣传教育活动、地方高层

领导的支持和法规政策中对数据审查和开放数据质量的要求。散落在图中右下角的指标整体达标率虽然较高但地区间差异较大，包括法规政策中对开放平台建设、数据标准管理、开放数据的目录与清单编制、开放数据标准和推动利用等方面的要求。

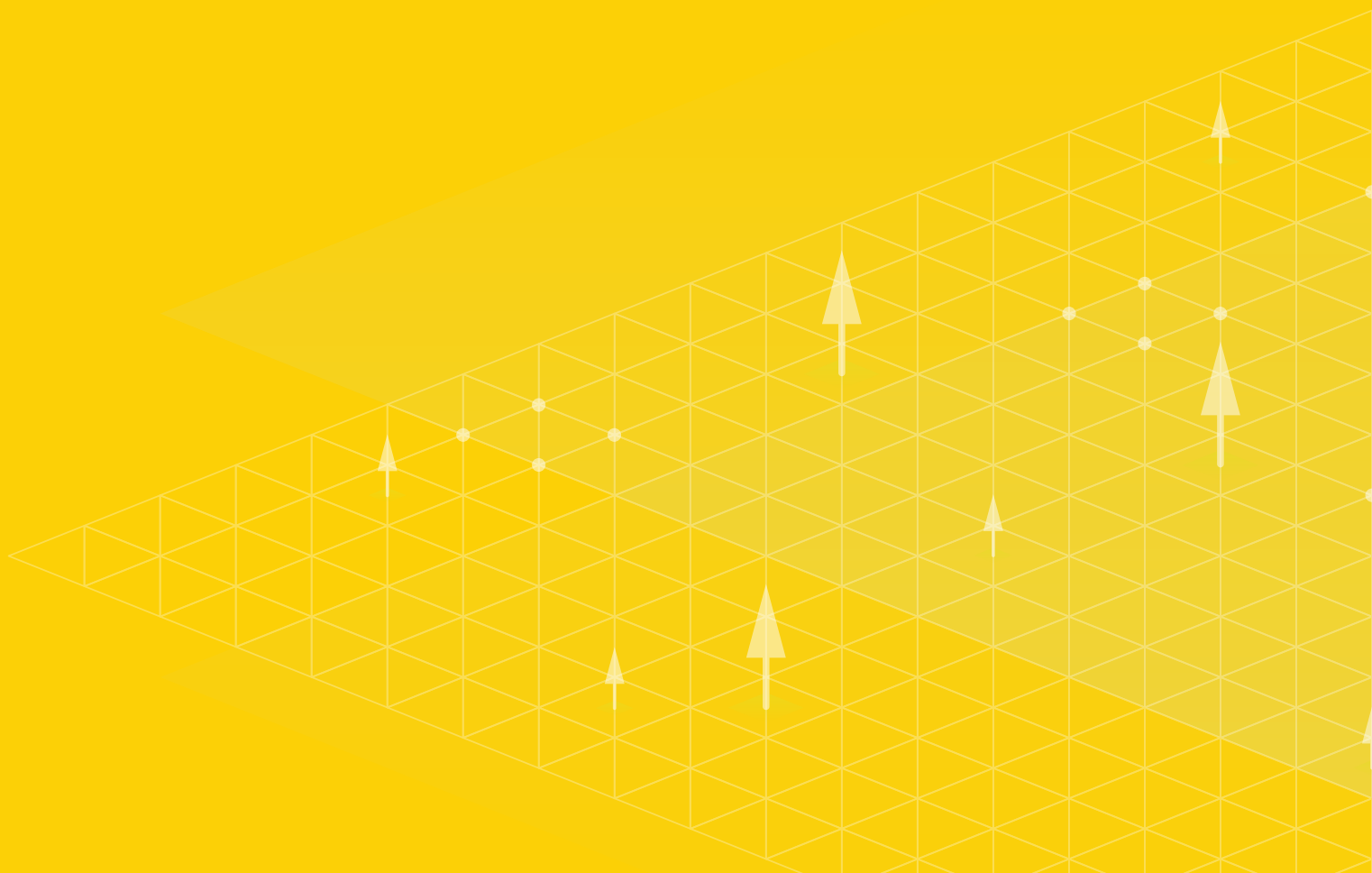


图110 准备度各项指标均值与标准差

5

数林建言

Conclusions



◆ 5.1 及时造林, 后来居上, 提高“开放数林”覆盖率

整体上, 我国“开放数据、蔚然成林”的态势已基本形成, 东、南、西、北各方都生长出一批优良“数木”, 山东、广东已形成两片区域性“数林”。然而, 放眼全国, “绿洲”还形孤影只, “沙漠”仍四野八荒。

与早期探索阶段相比, 当前我国政府数据开放的环境和条件已得到很大改善, 上有国家政策的明确支持, 下有各地实践的成功经验, 天时地利, 俱已齐备, 建议各地政府尽早结束观望, 及时造林。特别是东部沿海地区, 数据基础好, 社会需求高, 却仍存在大片空地, 落后于不少中西部省市。这些地区尤应尽快开放, 植“数”造林, 学习前人经验, 利用后发优势, 力争后“种”居上。已开放地区也要不忘初心, 持续生长, 探索前沿, 突破难点, 防止不进则退。

建议国家和地方层面持续加大数据开放推进力度, 选取有示范性、有利于推动区域扩散的地方作为试点, 不断总结地方经验, 向更大范围复制推广, 稳步提高全国“数林”覆盖率。

◆ 5.2 竞相相生, 枝杈相连, 保护“开放数林”多样性

多元创新、自由生长的“自然林”可能参差不齐, 而集中建设、整齐划一的“人工林”又难免千城一面。自然林, 还是人工林? 能长高长密的就是好林子。开放数据的目标是让政府数据为人民所用, 用户的获得感就是检验林子好坏的最高标准。

建议国家和地方层面着力于制定基本标准, 托起底线水平, 扶植弱小树苗, 解绑优质“数木”。在要求下级平台完成“规定动作”的同时, 鼓励和允许各地根据本地社会需求和发展战略, 开放特色数据, 展现本土应用, 做到既提升数林整体高度, 又避免枪打出头枝, 兼容共性与个性, 保持数林多样性。

各地之间既要学习赶超, 又要协调联动, 不同层级和地域的平台之间相互链接, 关联数据, 便于用户寻找和获取。独木不成林, 只有“数木”竞相相生, 枝杈相连, 才能生长出一片繁盛多样的开放数林。

◆ 5.3 保持树形, 监测态势, 促进“开放数林”可持续生长

准备度是底子, 平台是面子, 数据是里子, 一棵好数木需三者兼备, 缺一不可。只有根系发达, 土壤肥沃, 雨露充足, 勤于浇水施肥, 才能保证数木的茁壮生长; 也只有枝干粗壮, 枝杈舒展, 才能带来成荫如盖, 叶茂花开。因此, 建议各级政府在推动数据开放时一面夯实基础, 营造环境, 一面提升平台体验, 开放优质数据。只有注重根系、枝干、树叶三者平衡生长, 保持树形健康, 才能促进开放数林可持续生长。

在数据开放的过程中, 建立国家和地方层面建立起数据开放监督考核体系, 随时留意区域内生长滞后、质量不佳或畸形发展的“数木”, 及时剪枝修“数”、调控长势。本报告即是学术界从第三方视角所进行的尝试, 将定期发布“开放数林指数”, 持续追踪我国地方政府数据开放生态体系的形成与发展, 精心测量“数木”们的树形、高度、粗细、材质与价值, 助力开放数据蔚然成林。

◆ 5.4 不断修剪, 精耕细作, 满足用户对“开放数林”的向往

政府数据开放是一项长期性、持续性的工作, 开放数林需在多个层面不断修剪, 精耕细作。

准备度是今年的新增维度, 重点关注法规与政策、领导力和组织保障。报告发现, 目前各地法规与政策中关于数据纠错机制、人财物保障、监督考核机制、平台安全、风险管理、沟通机制、开放方式等方面的要求以及部门领导支持、组织创新应用活动等指标的达标情况普遍较低, 建议国家层面及时提出要求, 做出指引。同时, 开展宣传教育活动、地方高层领导支持、法规政策中对数据审查和开放

数据质量的要求等指标则不仅整体达标率低，而且地区间差异较大，需要国家层面加强引导，落后地方加快追赶。报告还发现，各地法规政策中对于开放平台建设、数据标准管理、开放数据标准和推动利用等方面的要求虽然整体达标率较高，但地区间差异较大，落后地方可在这些方面主动学习先进地区经验，尽早补足短板。

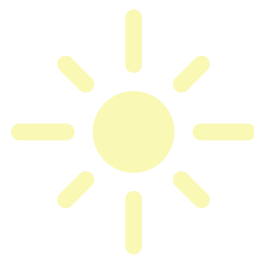
开放数林的繁茂离不开强大的根系、肥沃的土壤和适宜的生长环境。建议上级政府加强制度供给，为下级政府植“数”造林带来更多的“阳光雨露”，营造良好的生长环境，尤其需做好顶层设计，完善制度框架，出台法规政策，探索机制创新，制定标准规范。而对具体正在开展数据开放的地方而言，其政府高层领导需加大支持力度，优化体制架构，组建或指定一支高层级、专业化的“护林团”作为数据开放主管部门，并赋予该部门足够的职权以协调统筹其他业务部门的数据开放，明确各部门在数据开放工作中的职责与分工，并加强人财物保障，夯实“数木”根基，提高“土壤”肥力。作为“护林团”，应负责“数木”的日常管理、开放、利用和保护，既要植树造林，更要浇水施肥。

在数据层方面，数据质量和数据容量是本次评估的重点，从数据集数量，到数据集容量，到数据集质量，要求不断提升。评估发现，各地在开放数据格式、动态更新、优质数据、数据集容量、数据集总量等方面整体水平较低。建议国家层面尽快制定开放数据标准，对开放数据的格式、质量和数量做出明确要求，提升全国整体水平。此外，分级分类获取、提供数据发布者联系方式、平台间关联三项指标在全国范围内不仅整体达标率低，且地区差异显著，建议国家层面予以关注，落后地方尽快改进。此外，在API

描述规范、无问题数据等指标上，全国整体达标率虽然较高，但地区差异也十分明显，落后地区需尽早达标。

本次评估在平台层方面的重点从考察平台功能走向平台体验，更为重视用户的实际感受。报告发现，开放数据目录、标明数据来源、数据纠错功能、数据请求功能等指标在全国整体达标率较低，建议国家主管部门对这些薄弱方面提出明确要求。此外，各地达标率低且地区间差异较大的指标有分级分类获取、提供数据发布者联系方式、平台间链接等三项，这些指标亟需国家和地方层面加强重视，落后地方及早跟进。报告还发现，各地在数据本地获取、提供地理空间工具、订阅收藏、用户个性化功能整合和排序功能等指标上整体达标率较高，但仍存在明显的地区间差异，落后地方需及时弥补。

最后，报告呼吁各地政府部门着力解决好用户日益增长的数据需求和不平衡不充分的政府数据开放之间的矛盾，大力推动数据开放的广度和深度，更好满足用户对开放“数林”的向往。



◎ 结语

大数据发展日新月异。2017年12月，习近平总书记在中共中央政治局第二次集体学习时指出，数据是新的生产要素，是基础性资源和战略性资源，也是重要生产力，要构建以数据为关键要素的数字经济。今年4月，国家主席习近平在博鳌亚洲论坛发言时强调中国开放的大门不会关闭，只会越开越大！2018年是改革开放40周年，改革再出发，开放再扩大，而数据开放正是扩大开放在大数据领域的体现。

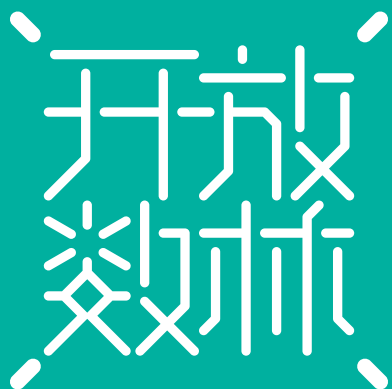
再出发意味着跨越与发展。政府数据开放的最终目的是释放数据能量，创造公共价值，其影响已远远超越数据领域本身。开放政府数据，推动创新发展，是国家治理体系和治理能力现代化的必然要求，也是数字中国建设的重要内容。2018年扩大政府数据开放区域，加快政府数据开放速度，将是大数据领域对改革开放40周年的最好回应。

为此，“开放数林指数”将继续贯彻“创新、协调、绿色、开放、共享”五大发展理念，在不断完善数据层、平台层、准备度评估体系的同时，进一步关注数据利用的产出和效果，打通数据开放生态系统的每个环节，构筑智能时代的公共数据基础设施。

评估专家委员会

安小米	中国人民大学	宋文好	《电子政务》杂志社
陈涛	华中科技大学	王芳	南开大学
陈涛	上海图书馆	王少辉	武汉大学
陈玉梅	暨南大学	王翔	复旦大学
程序	提升政府治理能力大数据应用技术国家工程实验室	王晓斌	浪潮集团有限公司
高冰	中电科新型智慧城市研究院有限公司	夏义堃	武汉大学
高丰	SODA理事会	肖卫兵	上海政法学院
黄璜	北京大学	谢真强	提升政府治理能力大数据应用技术国家工程实验室
贾一苇	国家信息中心	薛梓闻	速溶360
金耀辉	上海交通大学	杨道玲	国家信息中心
李重照	云南财经大学	殷沈琴	复旦大学
刘汪洋	中电科大数据研究院有限公司	张峰	浪潮集团有限公司
刘文静	暨南大学	张会平	电子科技大学
刘新萍	上海理工大学	张楠	清华大学
刘枝	国家信息中心	章燕华	浙江大学
龙怡	上海政法学院	郑磊	复旦大学
马亮	中国人民大学	郑跃平	中山大学
孟天广	清华大学	朱琳	华东理工大学

* 注：评委名单按照姓氏拼音排序



中国开放数林指数 CHINA OPEN DATA INDEX 2018

开放数据 蔚然成林

一棵棵开放“数木”的丛然并生、成荫如盖、直至叶茂花开，终将成长为一片繁盛多样、平衡清朗、持续循环的“开放数林”。

《中国地方政府数据开放报告》将定期发布“开放数林指数”，持续追踪我国地方政府数据开放生态体系的形成与发展，

精心测量“数木”们的树形、高度、粗细、材质与价值。

联合发布方



復旦大學

复旦大学



提升政府治理能力大数据应用技术
国家工程实验室



国家信息中心
数字中国研究院

国家信息中心数字中国研究院

指数出品方



DMG Lab
Fudan University
复旦大学 | 数字与移动治理实验室



扫描二维码关注
复旦大学数字与移动治理实验室
微信公众号

指数出品团队

郑磊、刘新萍、郑跃平、吕文增、朱晓婷、蔡城城、赵丽琼、陈延、吴心琦、袁佳蕾、王翔、戚馨予

合作方



中电科新型智慧城市研究院有限公司
The Institute of New Smart City Research of CEC



复旦发展研究院
治理与国家治理研究中心
Center for Governance
and State Governance Research

inspur 浪潮



冥睿(上海)
信息科技有限公司



开放数据中国

DE'SIGN
表明德设计事务所



中国开放数林指数网 ifopendata.fudan.edu.cn

• 联合发布方

复旦大学
提升政府治理能力大数据应用技术国家工程实验室
国家信息中心数字中国研究院

• 指数出品方

复旦大学数字与移动治理实验室

• 合作方

中电科新型智慧城市研究院
复旦发展研究院传播与国家治理研究中心
浪潮集团有限公司
冥睿（上海）信息科技有限公司
开放数据中国
麦明德设计事务所



DMG Lab
Fudan University
复旦大学 | 数字与移动治理实验室



上海市杨浦区邯郸路220号复旦大学智库楼404室
dmglab@fudan.edu.cn
www.dmg.fudan.edu.cn



关注
DMG
微信公众号