



中国开放数据指数  
CHINA OPEN DATA INDEX

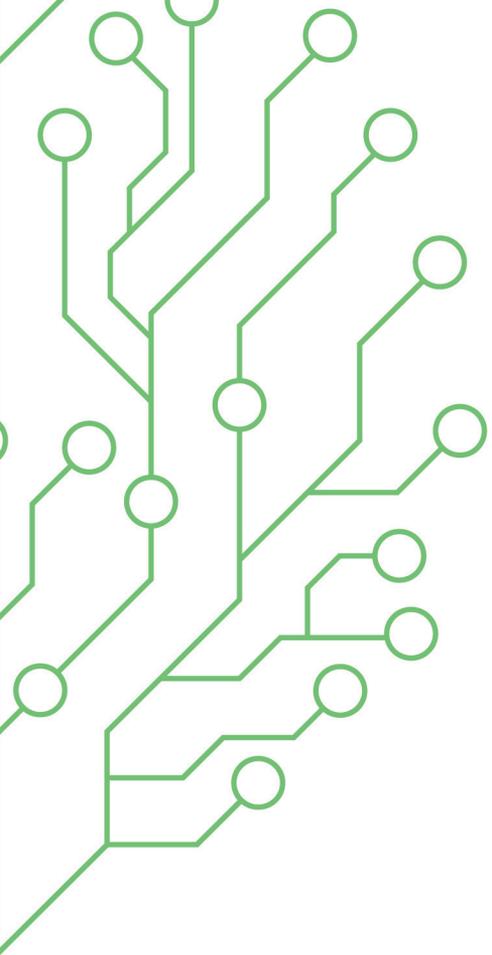
· 2019 ·

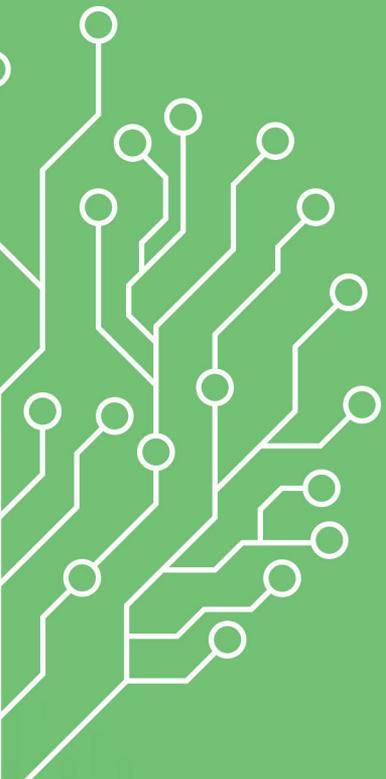


# CONTENTS

总结摘要	EXECUTIVE SUMMARY	002
■ 评估方法	METHODS	006
■ 总体发现与建议	FINDINGS AND RECOMMENDATIONS	018
■ 开放数林指数与排名	INDEX AND RANKING	028
■ 准备度发现	READINESS	032
■ 平台层发现	PLATFORM	050
■ 数据层发现	DATA	070
■ 利用层发现	USE	090
附录	APPENDIX	100

# EXECUTIVE SUMMARY





# 总结摘要



## 总结摘要

### 1. 数林概貌

■截至 2019 年上半年，我国已有 82 个省级、副省级和地级政府上线了数据开放平台。与 2018 年报告同期相比，新增了 36 个地方平台。

■东南沿海地区的省级平台已经逐渐相连成片，并向内陆地区不断扩散。广东和山东省内的各地市都推出了数据开放平台，形成我国最为密集的省级“开放数林”。

■41.93% 的省级行政区、66.67% 的副省级城市和 18.55% 的地级城市推出了数据开放平台，政府数据开放平台已逐渐成为一个地方数字政府建设的“标配”。

■指数分值较高的省份主要集中在东部沿海地区，指数分值较高的副省、地级城市很多都出现在山东和广东两省内；同时，位于中西部和东北部地区的贵州省、贵阳市、成都市、银川市、哈尔滨市也表现抢眼。

### 2. 数叶：数据层

■全国开放数据集总量从 2017 年的 8398 个迅速增长到 2019 年的 62801 个，增幅近七倍。开放数据集的容量与 2018 年报告同期相比，呈现出爆发式增长，一年之内增幅近 20 倍。

■约三成的平台上开放的数据集总量已超过了一个个，约两成的平台上开放的数据容量已超过 1 亿，但仍有超过三成的平台上开放的数据容量仅在 10 万及以下。

■仅有约四成的平台开放了优质数据集，优质 API 接口仍屈指可数，调用难度较高，能调取到的数据容量小、更新频率低。

■近四成的平台缺少专门的数据开放授权协议，只有 16% 的平台在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。

■只有不到一成的平台为每个数据集明确标示了包括有条件开放在内的多种开放类型。

■只有不到一成的平台能在近两年来的每个季度中都持续上线新增数据集，而超过九成的平台曾出现数据集增长中断或停滞的情况。在 2019 年第一季度内，近五成的平台没有更新过数据集。

### 3. 数果：利用层

■有一成多的地方举办过类似开放数据创新利用比赛的利用促进活动，但只有上海市和北京市两个地方连续举办。

■仅在 3.7% 的平台上发现了少量的有效服务应用，主要关于交通出行、教育科技和卫生健康主题。

■超过六成的平台展示的是由政府部门自身开发而不是政府数据开放后被社会开发利用所产生的应用。

■超过九成的平台展示的利用成果未标注数据来源，而即使标注了的数据来源，也普遍存在数据链接无效、数据搜索不到或数据不完整等情况。

#### 4. 数干：平台层

- 只有一成的平台能按用户使用场景和标签分类进行数据导引。
- 不到两成的平台提供了开放数据目录。
- 提供了数据请求功能的地方平台已超过六成，但其中真正能回复用户数据请求的地方平台寥寥无几。
- 开通了意见建议和数据纠错功能的平台已占六成左右，但没有一个地方真正回复用户的意见建议和数据纠错反馈。
- 尚未有平台真正对有条件开放的数据开通了申请功能。
- 标明了利用成果的数据集来源、数据提供部门和利用者名称等信息的平台只占一成。
- 能提供数据发布者联系方式的平台不到一成。
- 超过三成平台的账户注册过程较为繁琐且要求用户提供的个人信息较多。

#### 5. 数根：准备度

- 各地方陆续组建了专门的政府数据管理部门，其中近七成为一级委办局，数据开放工作的组织基础得到加强。
- 尚未有地方出台专门针对数据开放的地方性法规、地方政府规章或规范性文件。
- 只有 16 个地方的党政领导在公开场合讲话要求开放数据。
- 只有广东、山东两省制定了专门针对数据开放的地方标准。
- 只有上海市连续制定和公开了专门针对政府数据开放的年度工作计划。

#### 6. 数林建言

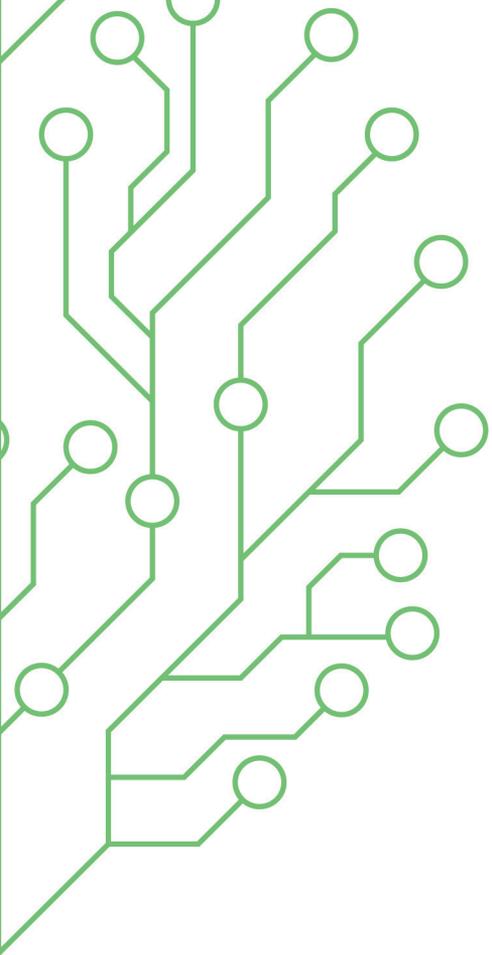
■ 数据是数据开放的核心，没有数叶，就不会有枝繁叶茂的数林。一个地方政府数据开放的水平实际上反映了这个地方的政府数据治理能力，各地应不断提升数据质量，开放更多高价值、高颗粒度、动态更新的优质数据集。对不同安全等级和敏感程度的数据可进行分级分类开放，实现数据开放的精细化管理，在保障数据安全的前提下，尽最大可能将政府数据开放出来。

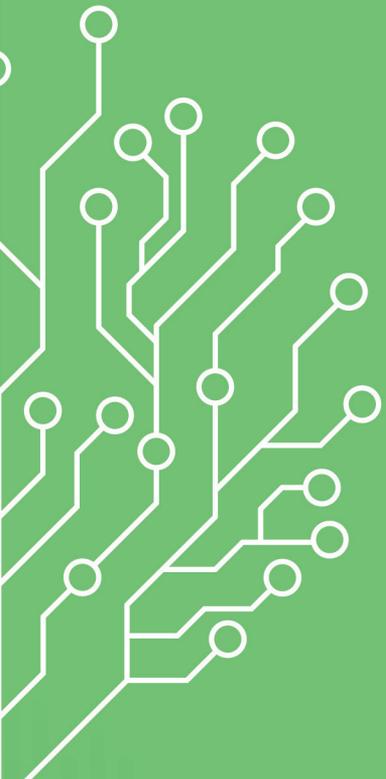
■ 利用成果是数据开放的成果，数木不结果，数据开放的价值无法真正得到释放。目前缺少利用成果的根本原因并不是数据利用者没有需求或缺少能力，而是绝大多数高需求、高价值的政府数据仍没有开放出来。因此，促进利用的关键在于增加数据供给。

■ 平台是连接数据供给侧和利用端的通道、开放和获取数据的载体、传播资讯动态的渠道、展示利用成果的平台。没有数干，数叶中蕴藏的能量无法输送给数果。因此，平台应成为一种持续性提供的服务，而不只是一个一次性开发完成的项目。各地平台的常态化运营服务能力亟待加强，才能及时回应用户提出的意见建议和纠错反馈，快速处理用户提出的数据申请和数据请求，定期审核和上传用户开发的利用成果。

■ 法规、组织、标准方面的准备度是数据开放的基础，失去了数根，数木终将枯萎。各地应加强针对数据开放的法规、政策和标准的制定，地方政府主要领导应更多公开表态支持数据开放。

# METHODS





CHAPTER



评估方法



## 一、评估方法

### 1.1 背景与目的

“中国开放数林指数”由复旦大学数字与移动治理实验室出品，是国内第一个针对地方政府数据开放水平的专业指数，自2017年5月首次发布以来，每年定期对我国地方政府数据开放进行综合评价，精心测量各地“数木”们的树形、高度、粗细、材质与价值，助推我国政府数据开放生态体系的建设与发展。

2018年以来，“中国开放数林指数”为国家互联网信息办公室信息化发展局监测我国公共信息资源开放情况提供数据支撑。2018年4月，“中国开放数林指数”被国家信息中心数字中国研究院发布的数字中国发展指数采纳，作为其评估地区数据开放能力的依据。2019年5月，指数报告被收录为复旦智库报告系列，成为复旦大学每年重点推出的有社会影响力的报告之一。

开放数据，蔚然成林。“开放数林”意喻我国政府数据开放利用的生态系统。一棵棵地方开放“数木”由最初的丛然共生、成荫如盖，直至枝繁叶茂、花开结果，终将成长为一片繁盛多样、枝杈相连、持续循环的中国“开放数林”。



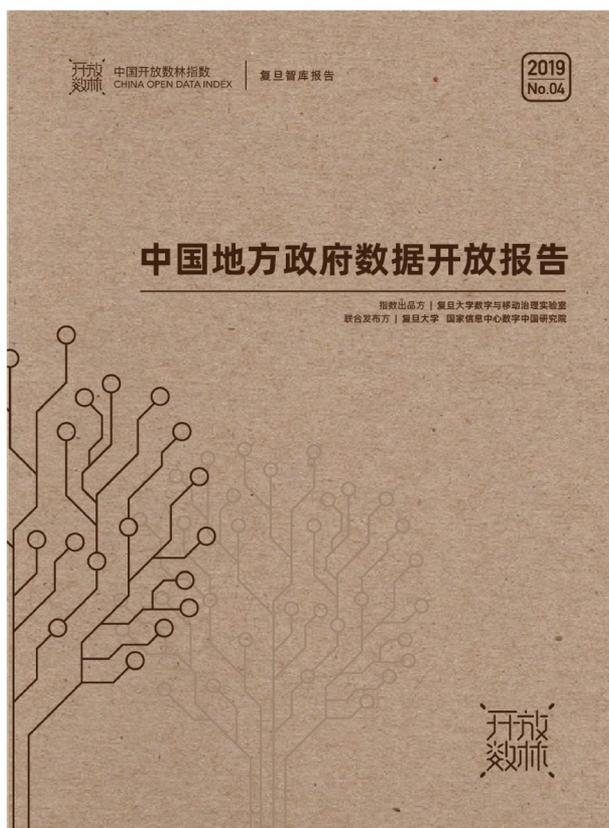
2017 No.01



2017 No.02



2018 No.03



2019 No.04

## 1.2 基本理念和框架



图 1-1 政府数据开放利用生态体系

“中国开放数林指数”对数据开放利用的全过程及其整个生态系统开展评估。政府数据从开放、利用到创造价值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系统。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，数据利用者和社会公众又进一步共同推动政府开放数据。在这个生态系统中，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者（见图 1-1）。

### 1.3 评估指标体系

指数出品方邀请国内外政界、学术界、产业界共七十多位专家共同参与，组成“中国开放数林指数”评估专家委员会（名单附后），体现了跨界、多学科、第三方、中立的专业视角和实际需求。专家委员会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，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，并为每项指标分配了权重（见图 1-2）。

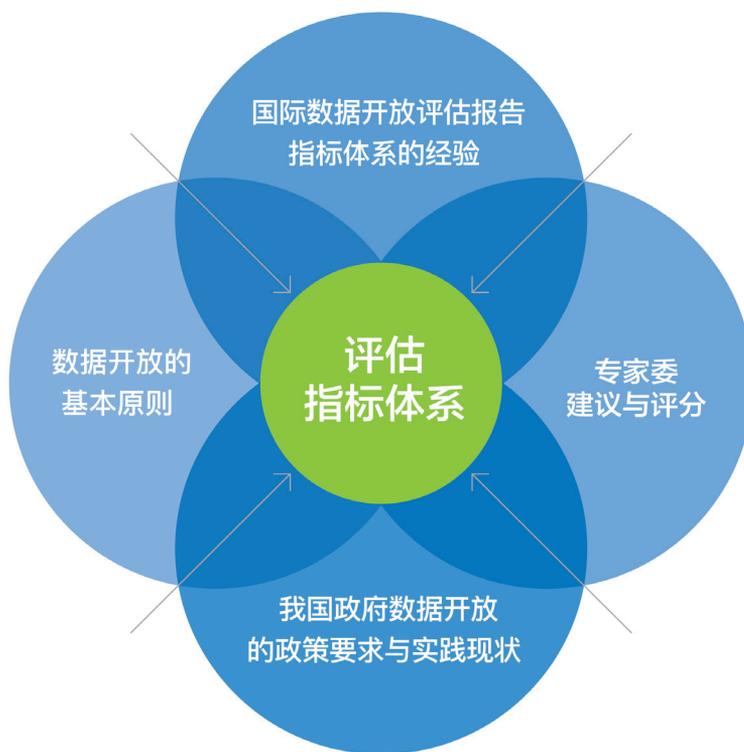


图 1-2 评估指标体系构建方法

评估指标体系共包括准备度、平台层、数据层、利用层四个维度及下属多级指标（见图 1-3）：

准备度是“数根”，是数据开放的基础，包括法规与政策、组织与执行、标准规范等三个一级指标。

平台层是“数干”，是政府开放数据和用户获取数据的载体，是展示利用成果的中心，也是连接数据开放供给侧和需求端的桥梁，包括数据发现、数据获取、工具提供、利用展示、互动反馈、公众传播和账户体验等七个一级指标。

数据层是“数叶”，是数据开放的核心，包括数据数量、数据质量、数据标准、数据覆盖面、数据持续性等五个一级指标。

利用层是“数果”，是数据开放的成果。利用层是今年新增加的评价维度，旨在促进政府数据开放后的社会化利用，包括利用促进、成果产出和数据利用等三个一级指标。

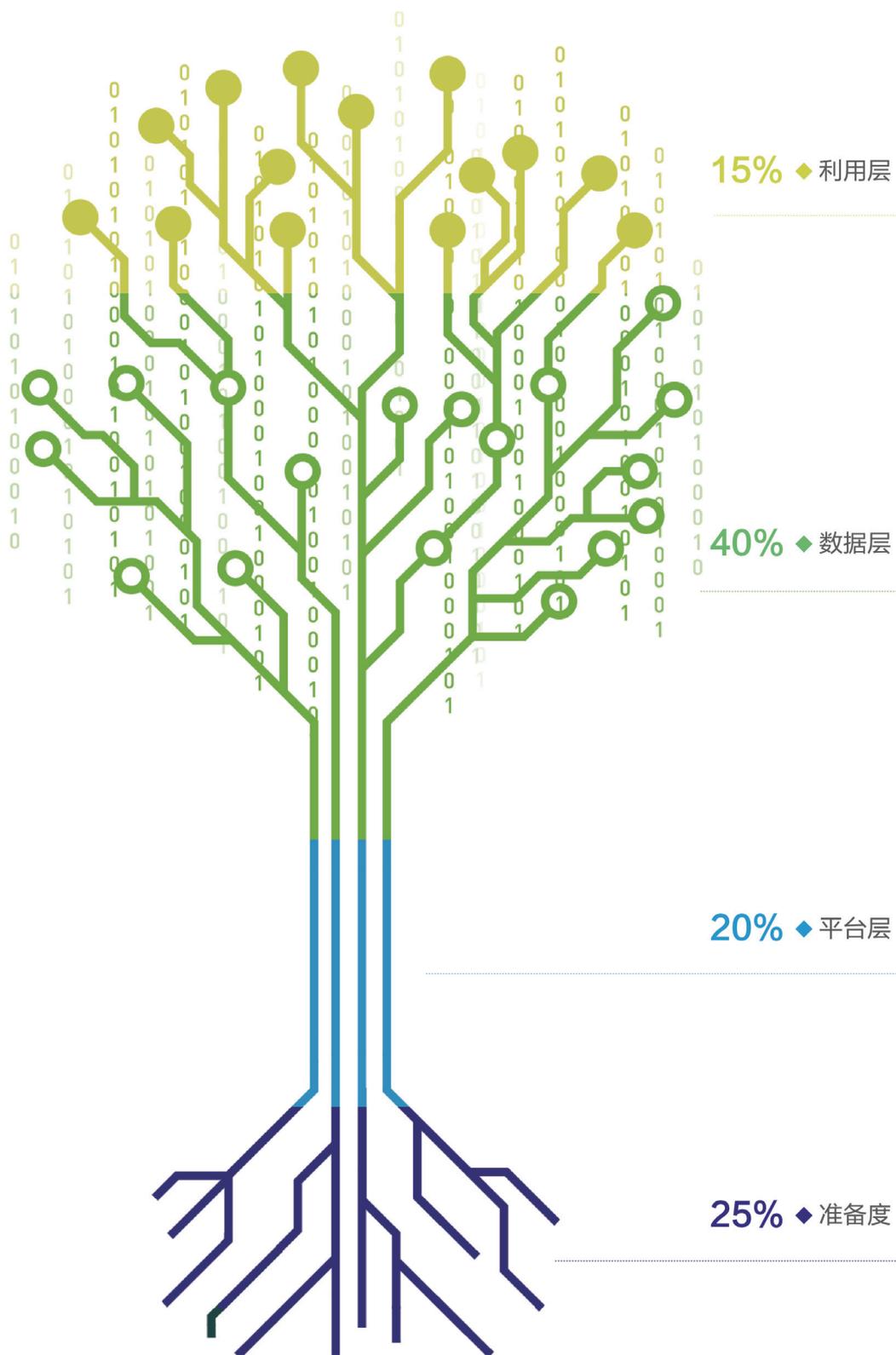
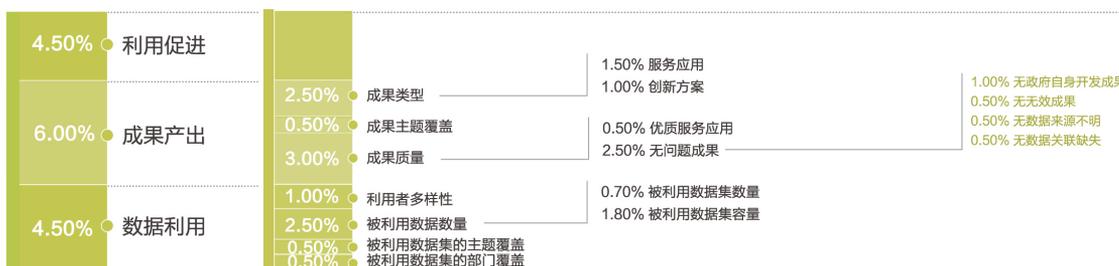
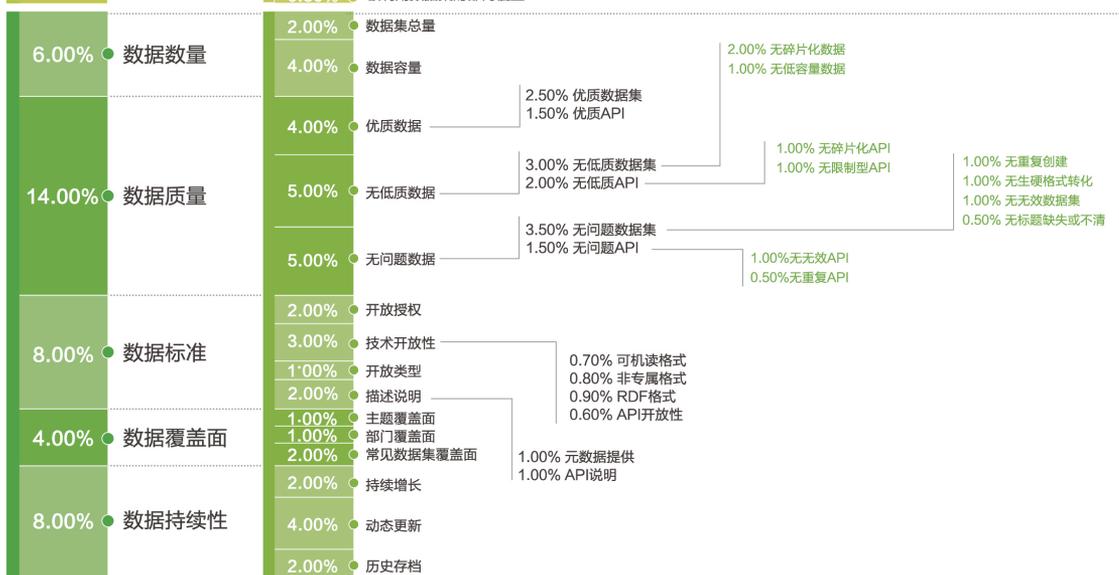


图 1-3 评估指标体系

15%  
利用层



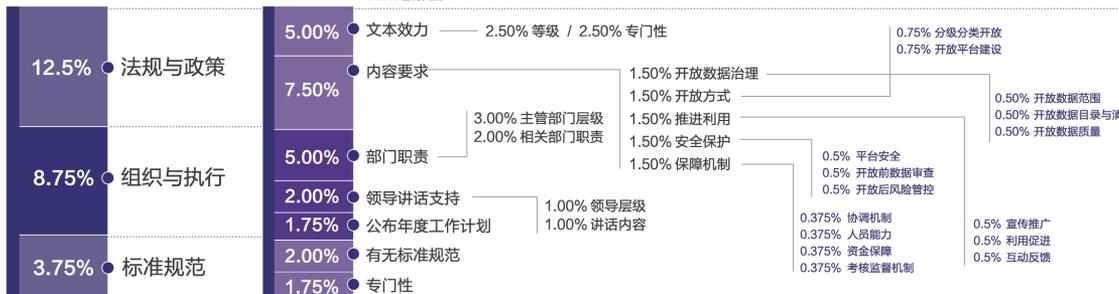
40%  
数据层



20%  
平台层



25%  
准备度



### 1.4 评估范围

指数出品方根据公开报道, 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2019年3月31日, 我国已上线的地方政府数据开放平台, 并从中筛选出符合以下条件的平台:

(1) 原则上平台域名中需出现 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) 平台所代表的地方政府的行政级别为地级及以上。

(3) 平台形式为“集中专有式”或“集中嵌入式”。“集中专有式”是指开放数据集中汇聚在一个专门的平台上进行开放; “集中嵌入式”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。目前, 除常州、阳江、阜阳、雅安等少数地方为“集中嵌入式”平台外, 其余均为“集中专有式”平台。各个条线部门建设的非集中式开放数据的平台不在指数的评估范围内。有些地市既建设了集中专有式的本地数据开放平台, 也在其归属省的省级数据开放平台上开设了本地栏目, 指数出品方以该地市本地平台为评估对象; 对于只在省级平台上开设了本地栏目, 未建设本地平台的地市, 指数出品方以其在省级平台上开设的地方栏目作为评估对象。

(4) 平台上确实开放了电子格式、可通过下载或接口形式获取、结构化的数据集。有些地方上线的网站或栏目虽然名为“数据开放”, 但实质上并未开放任何可通过下载或接口形式获取的、结构化的数据集, 而只是提供一些非结构化的文件或跳转到其他相关网页的链接。这类平台并未真正开放数据, 因此未被纳入本次评估范围。如江苏省人民政府网站上“数据开放”栏目、新疆维吾尔自治区政务数据开放网、山西省人民政府网站上“开放数据”模块和徐州市公共数据开放网等。

指数出品方共发现符合条件的地方平台82个, 并将上线了这些平台的地方作为本次评估的对象。具体地方、平台名称和平台域名如表1-1所示。

**表 1-1 评估范围 (按行政层级及拼音首字母排序)**

序号	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	开放福建	福建省	省级	http://data.fujian.gov.cn
3	广东省人民政府开放广东	广东省	省级	http://www.gddata.gov.cn
4	贵州省政府数据开放平台	贵州省	省级	http://www.gzdata.gov.cn
5	海南省政府数据统一开放平台	海南省	省级	http://data.hainan.gov.cn
6	河南省公共数据开放平台	河南省	省级	http://data.hnzwfw.gov.cn
7	江西省政府数据开放网站	江西省	省级	http://data.jiangxi.gov.cn
8	开放宁夏	宁夏回族自治区	省级	http://ningxiadata.gov.cn
9	山东公共数据开放网	山东省	省级	http://data.sd.gov.cn
10	陕西省公共数据开放平台	陕西省	省级	http://www.sndata.gov.cn
11	上海市政府数据服务网	上海市	省级	http://data.sh.gov.cn
12	天津市信息资源统一开放平台	天津市	省级	https://data.tj.gov.cn
13	浙江数据开放平台	浙江省	省级	http://data.zjzwfw.gov.cn
14	厦门市大数据开放平台	福建省厦门市	副省级	http://www.xmdata.gov.cn
15	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
16	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
17	哈尔滨市政府数据开放平台	黑龙江省哈尔滨市	副省级	http://data.harbin.gov.cn
18	武汉政务公开数据服务网	湖北省武汉市	副省级	http://www.wuhandata.gov.cn
19	南京市人民政府数据开放	江苏省南京市	副省级	http://data.nanjing.gov.cn
20	济南政府数据开放平台	山东省济南市	副省级	http://www.jndata.gov.cn
21	青岛公共数据开放网	山东省青岛市	副省级	http://data.qingdao.gov.cn
22	成都市公共数据开放平台	四川省成都市	副省级	http://www.cddata.gov.cn
23	宁波市政府数据服务网	浙江省宁波市	副省级	http://www.datanb.gov.cn
24	蚌埠市信息资源开放平台	安徽省蚌埠市	地级	http://data.bengbu.gov.cn
25	阜阳市人民政府数据开放栏目	安徽省阜阳市	地级	http://www.fy.gov.cn/openData
26	黄山市人民政府数据开放栏目	安徽省黄山市	地级	http://www.huangshan.gov.cn/DataDevelopment
27	六安市信息资源开放平台	安徽省六安市	地级	http://data.luan.gov.cn:8081/dop

28	马鞍山市人民政府数据开放栏目	安徽省马鞍山市	地级	<a href="http://www.mas.gov.cn/content/column/4697374">http://www.mas.gov.cn/content/column/4697374</a>
29	宣城市人民政府数据开放网	安徽省宣城市	地级	<a href="http://sjkf.xuancheng.gov.cn/open-data-web">http://sjkf.xuancheng.gov.cn/open-data-web</a>
30	广东省政府数据统一开放平台-潮州市	广东省潮州市	地级	<a href="http://www.gddata.gov.cn/index.php/data/ls/Type/0/v/344.html">http://www.gddata.gov.cn/index.php/data/ls/Type/0/v/344.html</a>
31	数据东莞	广东省东莞市	地级	<a href="http://dataopen.dg.gov.cn">http://dataopen.dg.gov.cn</a>
32	佛山市数据开放平台	广东省佛山市	地级	<a href="http://www.foshan-data.cn">http://www.foshan-data.cn</a>
33	广东省政府数据统一开放平台-河源市	广东省河源市	地级	<a href="http://www.gddata.gov.cn/index.php/data/ls/Type/0/v/339.html">http://www.gddata.gov.cn/index.php/data/ls/Type/0/v/339.html</a>
34	开放惠州	广东省惠州市	地级	<a href="http://data.huizhou.gov.cn">http://data.huizhou.gov.cn</a>
35	开放江门	广东省江门市	地级	<a href="http://opendata.jiangmen.gov.cn">http://opendata.jiangmen.gov.cn</a>
36	广东省政府数据统一开放平台-揭阳市	广东省揭阳市	地级	<a href="http://www.gddata.gov.cn/index.php/data/ls/Type/0/v/345.html">http://www.gddata.gov.cn/index.php/data/ls/Type/0/v/345.html</a>
37	茂名市人民政府数据开放栏目	广东省茂名市	地级	<a href="http://www.maoming.gov.cn/zwgk/sj kf">http://www.maoming.gov.cn/zwgk/sj kf</a>
38	广东省政府数据统一开放平台-梅州市	广东省梅州市	地级	<a href="http://www.gddata.gov.cn/index.php/data/ls/Type/0/v/337.html">http://www.gddata.gov.cn/index.php/data/ls/Type/0/v/337.html</a>
39	广东省政府数据统一开放平台-清远市	广东省清远市	地级	<a href="http://www.gddata.gov.cn/index.php/data/ls/Type/0/v/341/t/7.html">http://www.gddata.gov.cn/index.php/data/ls/Type/0/v/341/t/7.html</a>
40	广东省政府数据统一开放平台-汕头市	广东省汕头市	地级	<a href="http://www.gddata.gov.cn/index.php/data/ls/Type/0/v/330.html">http://www.gddata.gov.cn/index.php/data/ls/Type/0/v/330.html</a>
41	广东省政府数据统一开放平台-汕尾市	广东省汕尾市	地级	<a href="http://www.gddata.gov.cn/index.php/data/ls/Type/0/v/338.html">http://www.gddata.gov.cn/index.php/data/ls/Type/0/v/338.html</a>
42	广东省政府数据统一开放平台-韶关市	广东省韶关市	地级	<a href="http://www.gddata.gov.cn/index.php/data/ls/Type/0/v/327.html">http://www.gddata.gov.cn/index.php/data/ls/Type/0/v/327.html</a>
43	阳江市政府网数据开放栏目	广东省阳江市	地级	<a href="http://www.yangjiang.gov.cn/sj kf">http://www.yangjiang.gov.cn/sj kf</a>
44	广东省政府数据统一开放平台-云浮市	广东省云浮市	地级	<a href="http://www.gddata.gov.cn/index.php/data/ls/Type/0/v/346.html">http://www.gddata.gov.cn/index.php/data/ls/Type/0/v/346.html</a>
45	湛江数据服务网	广东省湛江市	地级	<a href="http://data.zhanjiang.gov.cn">http://data.zhanjiang.gov.cn</a>
46	肇庆市人民政府数据栏目	广东省肇庆市	地级	<a href="http://www.zhaoqing.gov.cn/sj kf">http://www.zhaoqing.gov.cn/sj kf</a>
47	中山市政府数据统一开放平台	广东省中山市	地级	<a href="http://zsdata.zs.gov.cn/web/index">http://zsdata.zs.gov.cn/web/index</a>
48	珠海市民生数据开放平台	广东省珠海市	地级	<a href="http://data.zhuhai.gov.cn">http://data.zhuhai.gov.cn</a>
49	贵阳市政府数据开放平台	贵州省贵阳市	地级	<a href="http://www.gyopendata.gov.cn">http://www.gyopendata.gov.cn</a>
50	铜仁市政府数据开放平台	贵州省铜仁市	地级	<a href="http://www.gztrdata.gov.cn">http://www.gztrdata.gov.cn</a>
51	遵义市政府数据开放平台	贵州省遵义市	地级	<a href="http://www.zyopendata.gov.cn/index.html">http://www.zyopendata.gov.cn/index.html</a>
52	佳木斯市公共数据开放网	黑龙江省佳木斯市	地级	<a href="http://data.jms.gov.cn">http://data.jms.gov.cn</a>
53	荆门市人民政府数据开放	湖北省荆门市	地级	<a href="http://data.jingmen.gov.cn">http://data.jingmen.gov.cn</a>
54	常德市人民政府政务数据栏目	湖南省常德市	地级	<a href="http://dataopen.changde.gov.cn">http://dataopen.changde.gov.cn</a>
55	长沙市政府门户网站数据开放平台	湖南省长沙市	地级	<a href="http://www.changsha.gov.cn/data">http://www.changsha.gov.cn/data</a>
56	常州市人民政府数据开放栏目	江苏省常州市	地级	<a href="http://opendata.changzhou.gov.cn">http://opendata.changzhou.gov.cn</a>
57	淮安市公共数据开放网	江苏省淮安市	地级	<a href="http://opendata.huaian.gov.cn/dataopen/showindex/index.do">http://opendata.huaian.gov.cn/dataopen/showindex/index.do</a>
58	连云港市公共数据开放网	江苏省连云港市	地级	<a href="http://data.lyg.gov.cn">http://data.lyg.gov.cn</a>
59	南通市公共数据开放网	江苏省南通市	地级	<a href="http://data.nantong.gov.cn/home/html/index.html">http://data.nantong.gov.cn/home/html/index.html</a>
60	苏州市政府数据开放平台	江苏省苏州市	地级	<a href="http://www.suzhou.gov.cn">http://www.suzhou.gov.cn</a>
61	泰州市政务数据开放平台	江苏省泰州市	地级	<a href="http://opendata.taizhou.gov.cn">http://opendata.taizhou.gov.cn</a>
62	无锡市数据开放平台	江苏省无锡市	地级	<a href="http://etc.wuxi.gov.cn/zfxgk/sj kf/index.shtml">http://etc.wuxi.gov.cn/zfxgk/sj kf/index.shtml</a>
63	扬州市政务数据服务网	江苏省扬州市	地级	<a href="http://data.yangzhou.gov.cn">http://data.yangzhou.gov.cn</a>
64	乌海市数据开放平台	内蒙古自治区乌海市	地级	<a href="http://whdata.wuhai.gov.cn">http://whdata.wuhai.gov.cn</a>
65	石嘴山政府数据开放平台	宁夏回族自治区石嘴山市	地级	<a href="http://szsjsj kf.nxszs.gov.cn">http://szsjsj kf.nxszs.gov.cn</a>
66	银川市城市数据开放平台	宁夏回族自治区银川市	地级	<a href="http://data.yinchuan.gov.cn">data.yinchuan.gov.cn</a>
67	滨州市公共数据开放网	山东省滨州市	地级	<a href="http://bzdata.sd.gov.cn">http://bzdata.sd.gov.cn</a>
68	德州市公共数据开放网	山东省德州市	地级	<a href="http://dzdata.sd.gov.cn">http://dzdata.sd.gov.cn</a>
69	东营市公共数据开放网	山东省东营市	地级	<a href="http://dydata.sd.gov.cn">http://dydata.sd.gov.cn</a>
70	菏泽市公共数据开放网	山东省菏泽市	地级	<a href="http://hzdata.sd.gov.cn">http://hzdata.sd.gov.cn</a>
71	济宁市公共数据开放网	山东省济宁市	地级	<a href="http://jindata.sd.gov.cn">http://jindata.sd.gov.cn</a>
72	聊城市公共数据开放网	山东省聊城市	地级	<a href="http://lcdata.sd.gov.cn">http://lcdata.sd.gov.cn</a>
73	临沂市公共数据开放网	山东省临沂市	地级	<a href="http://lydata.sd.gov.cn">http://lydata.sd.gov.cn</a>
74	日照市公共数据开放网	山东省日照市	地级	<a href="http://rzdata.sd.gov.cn">http://rzdata.sd.gov.cn</a>
75	泰安市公共数据开放网	山东省泰安市	地级	<a href="http://tdata.sd.gov.cn">http://tdata.sd.gov.cn</a>
76	潍坊市公共数据开放网	山东省潍坊市	地级	<a href="http://wfddata.sd.gov.cn">http://wfddata.sd.gov.cn</a>
77	威海市公共数据开放网	山东省威海市	地级	<a href="http://whdata.sd.gov.cn">http://whdata.sd.gov.cn</a>
78	烟台市公共数据开放网	山东省烟台市	地级	<a href="http://ytdata.sd.gov.cn">http://ytdata.sd.gov.cn</a>
79	枣庄市公共数据开放网	山东省枣庄市	地级	<a href="http://zdata.sd.gov.cn">http://zdata.sd.gov.cn</a>
80	淄博市公共数据开放网	山东省淄博市	地级	<a href="http://zbddata.sd.gov.cn">http://zbddata.sd.gov.cn</a>
81	雅安市人民政府数据开放栏目	四川省雅安市	地级	<a href="http://www.yaan.gov.cn/shuju.html">http://www.yaan.gov.cn/shuju.html</a>
82	湖州市公共数据开放平台	浙江省湖州市	地级	<a href="http://data.huzhou.gov.cn/index.html">http://data.huzhou.gov.cn/index.html</a>

### 1.5 数据采集与分析方法

准备度评估主要对相关法律法规、政策、意见、计划、标准规范、新闻报道等资料进行了描述性统计分析和文本分析。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与政策文本、年度工作计划、有关地方党政领导讲话支持的新闻报道以及数据开放主管部门的信息；二是在地方政府门户网站以及政府数据开放平台上通过人工观察和关键词检索采集数据。指数出品方对法律法规与政策的筛选标准为至少有一节专门针对政府数据开放的内容；对标准规范的筛选标准为有关政府数据开放的地方性标准，行业性的标准规范未纳入本次评估范围。数据采集截止时间为2019年4月15日。

平台层评估主要采用人工观察法对各地方政府数据平台上各项功能进行观测并做描述性统计分析，数据采集截止时间为2019年3月31日。同时，指数出品方还对平台的回复情况（包括回复时间和回复质量）进行了评估，回复情况采集截止时间为2019年4月4日。

数据层评估主要通过机器自动抓取各地平台上开放的数据结合人工观察采集相关信息，然后对数据进行了描述性统计分析、交叉分析、文本分析和空间分析。数据采集截止时间为2019年4月15日，对“动态更新”这一指标的评测时段为2019年1月1日至2019年4月15日。对于只在省级平台上开设了本地栏目、未建设本地平台的地市，指数出品方以该地市在省级平台本地栏目下开放的数据作为评估对象。对于开设了下属地市本地栏目的省级平台，指数出品方以其在平台上开放的来自省本地的数据作为评估对象。

利用层评估主要对各地政府数据开放平台上展示的利用成果进行了人工观察和测试，对2017年以来各地开展的开放数据创新利用比赛信息进行了网络检索，并对采集到的数据进行了描述性统计分析。数据采集截止时间为2019年4月4日。

### 1.6 指标计算方法

指数出品方基于各地在各项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，其中5分为最高分，相应数据缺失或完全不符合标准则分值为0。对于连续型统计数值类数据则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。

各地平台在准备度、平台层、数据层、利用层四个维度上的指数总分等于每个单项指标的分值乘以相应权重所得到的加权总和。最终，各地开放数林指数等于准备度指数、平台层指数、数据层指数、利用层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

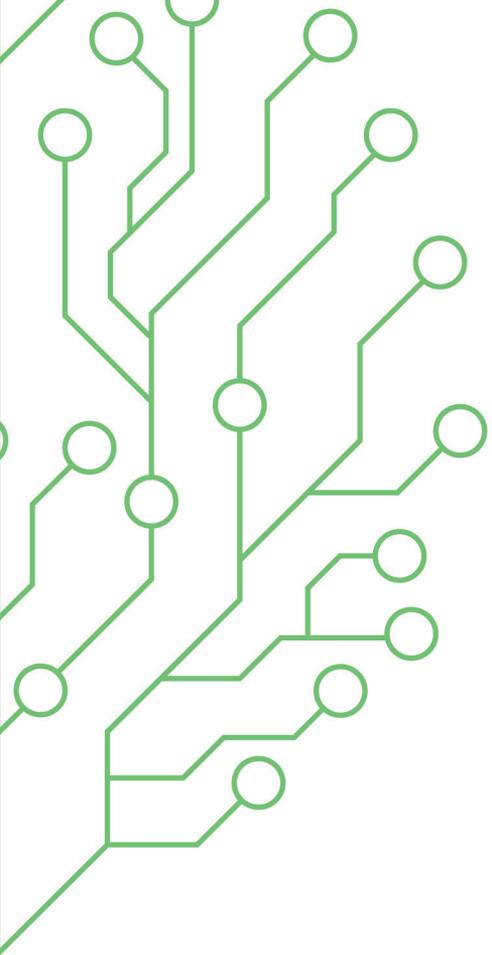
$$= \sum (\text{准备度指标分值} \times \text{权重}) \times 25\% + \sum (\text{平台层指标分值} \times \text{权重}) \times 20\% \\ + \sum (\text{数据层指标分值} \times \text{权重}) \times 40\% + \sum (\text{利用层指标分值} \times \text{权重}) \times 15\%$$

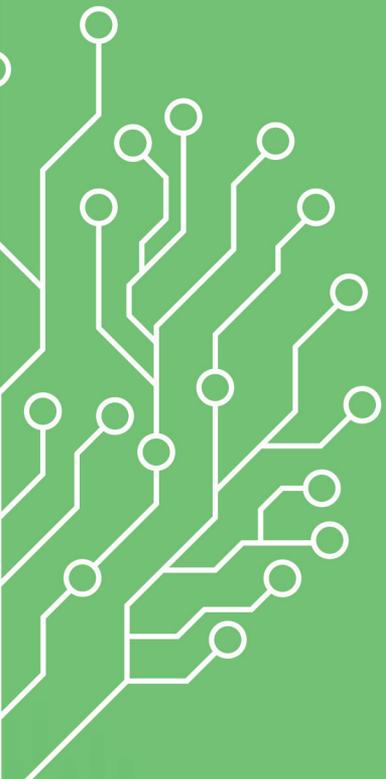


中国开放数据指数  
CHINA OPEN DATA INDEX

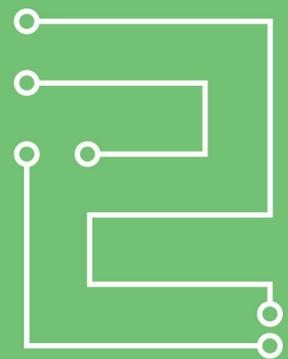
· 2019 ·

# FINDINGS AND RECOMMENDATIONS





CHAPTER



# 总体发现与建议

## 二、总体发现和建议

### 2.1 总体趋势

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，截至 2019 年上半年，我国已有 82 个符合政府数据开放基本特征的地级及以上平台陆续上线（如图 2-1 所示）。与 2018 年报告同期相比，增长速度迅猛，新增了 36 个地方平台。

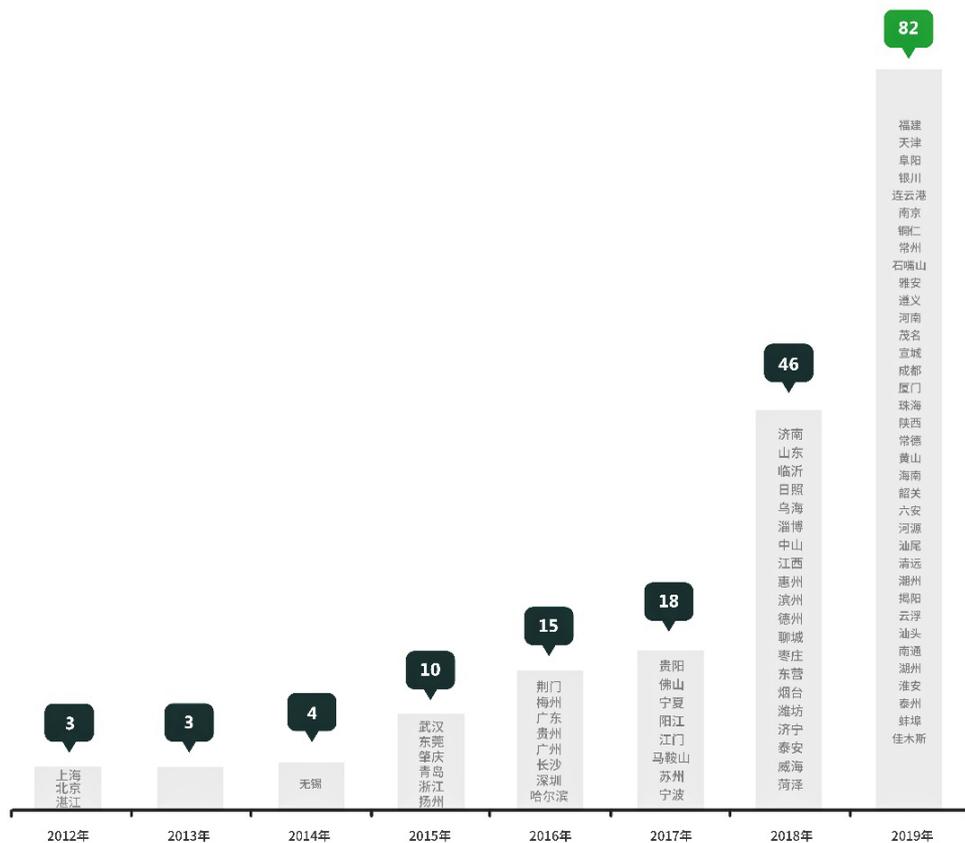
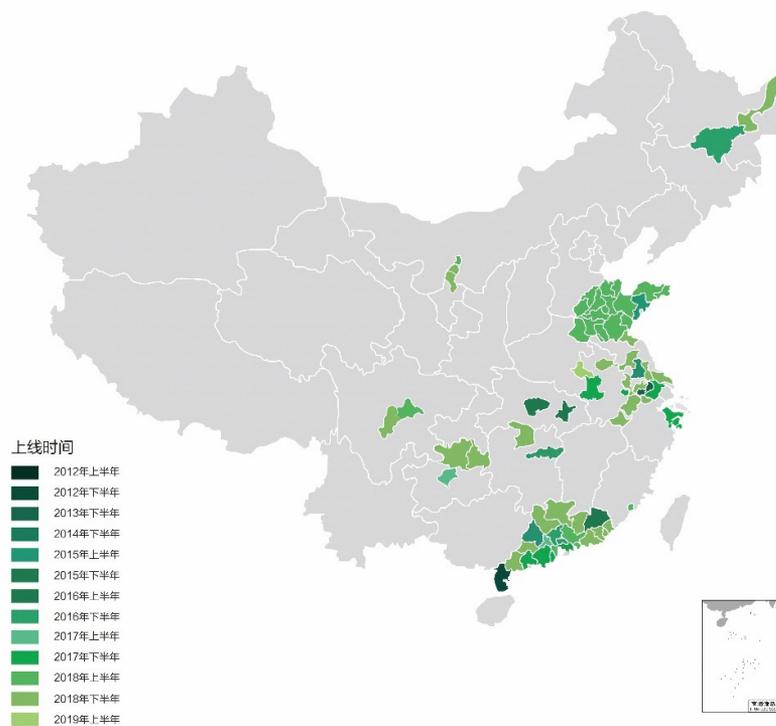
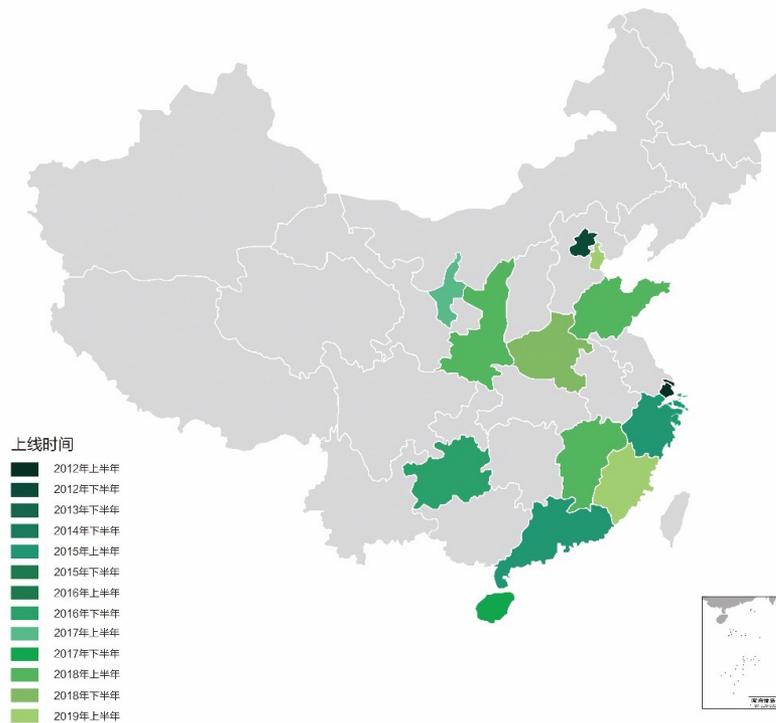


图 2-1 各地平台上线时间

截至 2019 年 3 月底，全国各地上线的省级政府数据开放平台的地区分布如图 2-2 所示，副省级和地级政府数据开放平台的地区分布如图 2-3 所示，颜色越深代表平台上线时间越早。从整体上看，我国地方政府数据开放平台的扩散呈现出从东南沿海地区向内陆地区不断发展的趋势，东南沿海地区的省级平台已经逐渐相连成片。

另外，各地政府数据开放平台还呈现出“群落式”与“绿洲式”两种分布特征。其中，“群落式”特征以山东省、广东省和长三角地区最为典型。在这些区域中，地市平台竞相开放数据，在广东省、山东省，省内各地市都推出了数据开放平台，形成我国最为密集的省级“开放数林”。相比之下，有些地区的周边均未上线政府数据开放平台，使这些地方成为了所在区域政府数据开放的“绿洲”。



报告对全国现有的 82 个政府数据开放平台按行政层级进行了统计（如图 2-4 所示）。其中省级行政区平台（港澳台除外）13 个，占总数的 41.93%；副省级城市平台 10 个，占总数的 66.67%；地级城市平台 59 个，占总数的 18.55%，比例相对偏低。

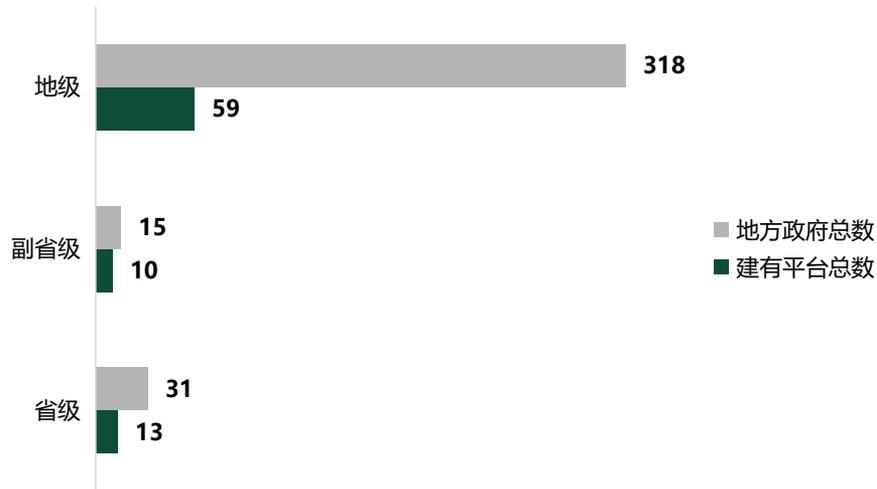


图 2-4 各地平台的行政层级分布

自 2017 年开始，报告连续三年跟踪各地数据开放的情况。期间，全国的开放数据集总量从 2017 年 8398 个迅速增长到目前的 62801 个（如图 2-5 所示）；开放数据集的容量从 2018 年开始呈现出爆发式增长，一年之间增幅近 20 倍（如图 2-6 所示），全国出现了更多高容量、高需求的优质数据集。

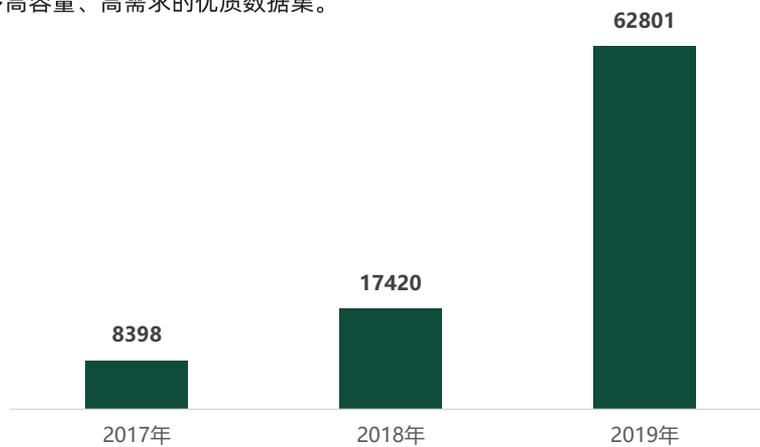


图 2-5 全国开放的数据集总量（单位：个）

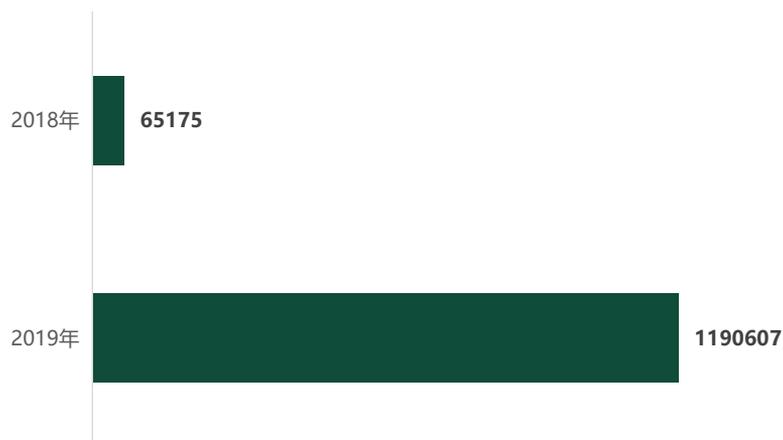


图 2-6 全国开放的总数据容量（单位：万）

2019年各地开放的数据集中满足可机读格式标准的比率为82%（如图2-7所示），与去年同期基本持平。

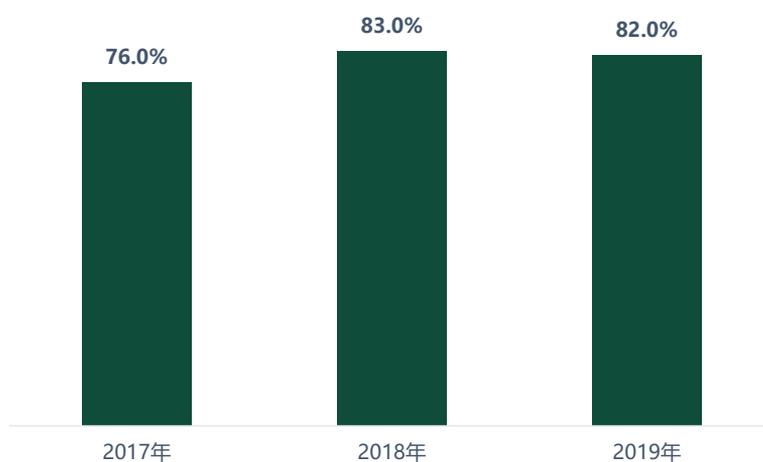
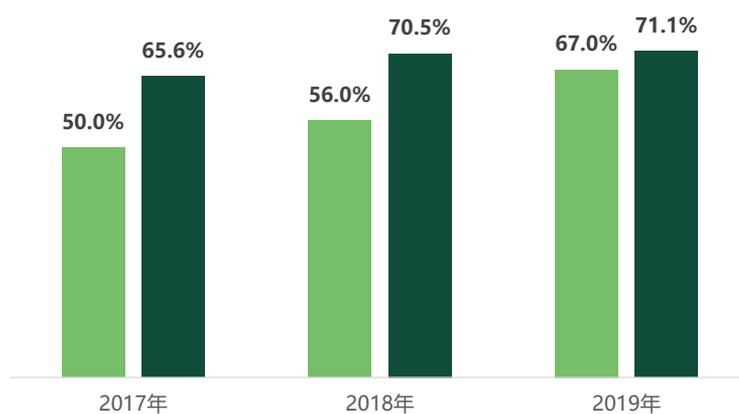


图 2-7 2018-2019 年全国开放的可机读格式数据集比例

全国数据开放的主题覆盖面与部门覆盖面的比例三年来均稳步增长（如图2-8），参与数据开放的政府部门越来越多，开放的数据愈加丰富多样。



■ 部门覆盖 ■ 主题覆盖

图 2-8 2017-2019 年全国开放数据的部门覆盖面与主题覆盖面

值得注意的是，与 2018 年同期相比，今年各地数据定期更新的比例出现了下滑（如图 2-9）。

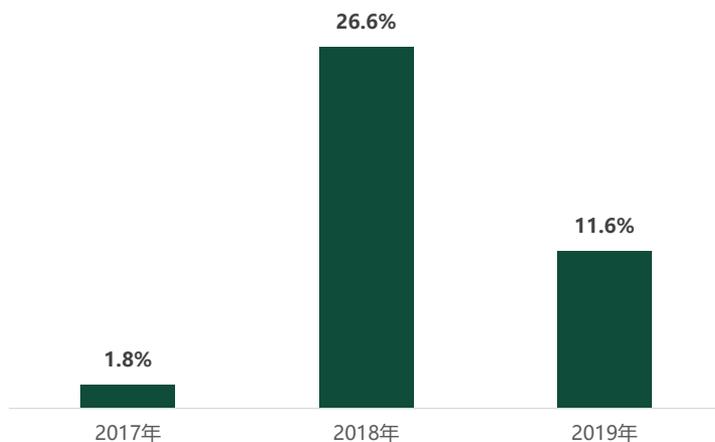


图 2-9 2017-2019 年全国开放数据的更新比例

## 2.2 指数分析

从省级地方指数分值的空间分布情况来看（图 2-10），目前分值高的省份主要集中在东部沿海，西部的贵州省也表现亮眼。与去年相比，福建省、天津市在新上线平台的地方中表现最为突出，浙江省则进步迅猛。

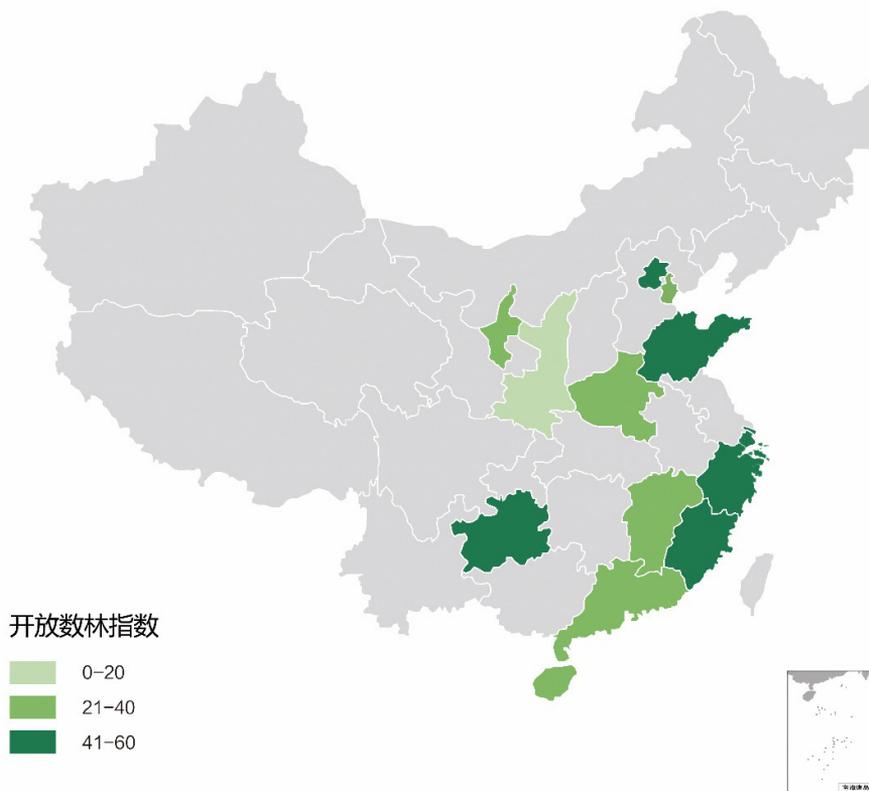


图 2-10 指数空间分布（省级）

从地级（含副省级）的空间分布来看（图 2-11），目前分值较高的地市主要集中在山东省、广东省，但贵阳市、哈尔滨市、成都市、银川市等位于中西部和东北部地区的城市也表现抢眼。

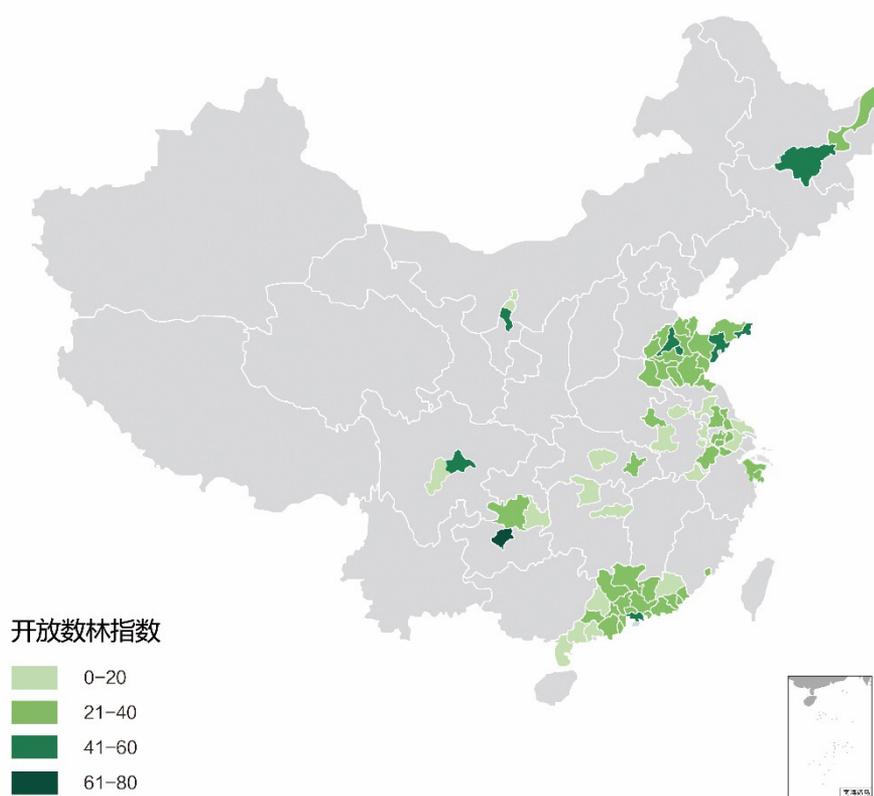


图 2-11 指数空间分布（副省级、地级）

报告对 2018 年下半年评估报告中的 56 个地方半年以来的指数分值和排名变化情况进行了分析。图 2-12 中的横坐标代表这 56 个地方 2019 年上半年的指数分值，纵坐标代表这些地方相比 2018 年下半年的名次变化情况，正数表示名次上升位数，负数表示下降位数。数据显示，2018 年下半年以来，排名上升幅度最大的地方是浙江省、深圳市和烟台市。2019 年新上线平台的地方，由于没有 2018 年的指数分值和名次可做对比，未在图中出现。

图中用排名不变（即为 0）与指数均值作为两条分界线，将各地分为三种类型。右上方的区域为“竹林”，表现为指数分值较高且上升迅速，包括浙江省、深圳市、济南市等地；右下方的区域为“乔木林”，表现为指数分值稳定或小幅下降，但指数分值仍维持在较高水平，主要包括哈尔滨市、成都市、山东省等地；左下方的区域为“灌木丛”，表现为指数分值下滑且指数处于中下水平，目前有大量地方处于该区域。



图 2-12 56 个地方数据层指数分值与排名变化分析

报告发现数据层、平台层、利用层与准备度四个维度之间存在一定程度的相关性。如图 2-13 所示，数据层指数与平台层指数曲线如藤蔓一样上下“缠绕”着准备度指数曲线，其中数据层指数曲线与准备度指数曲线之间的关系尤为紧密。经统计学检验，数据层指数与准备度指数的相关系数为 0.47，平台层指数与准备度指数的相关系数为 0.33，利用层与准备度两个指数的相关系数为 0.49，都在 0.01 水平上显著。也就是说，一个地方的准备度水平越高，其数据层、平台层与利用层的表现也相对更好。

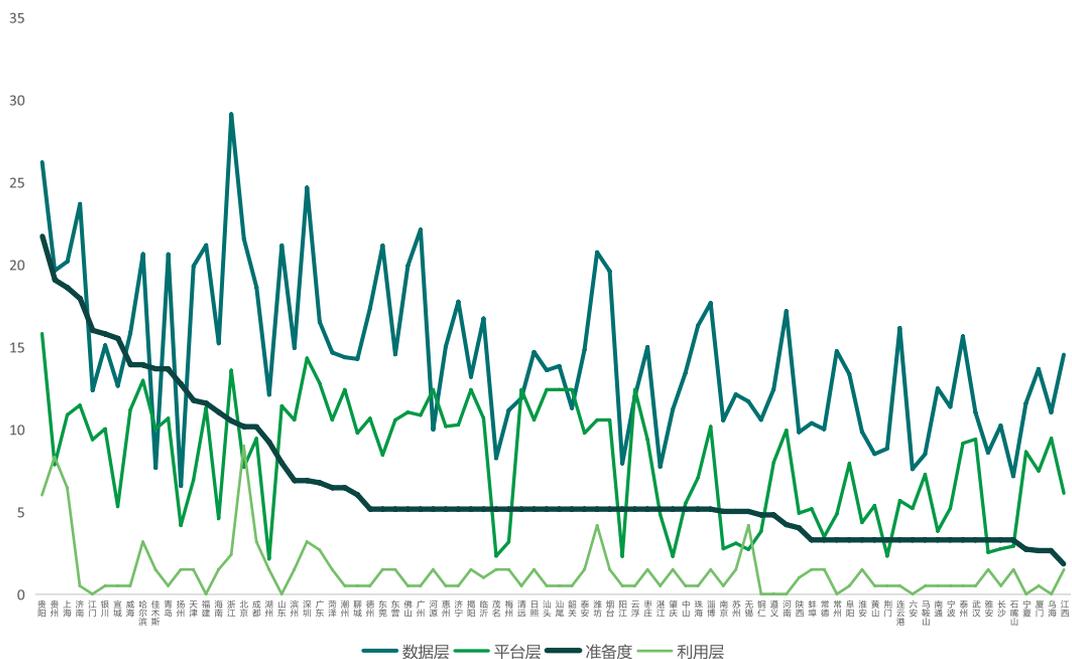


图 2-13 各地数据层、平台层、利用层与准备度指数曲线

### 2.3 主要建议

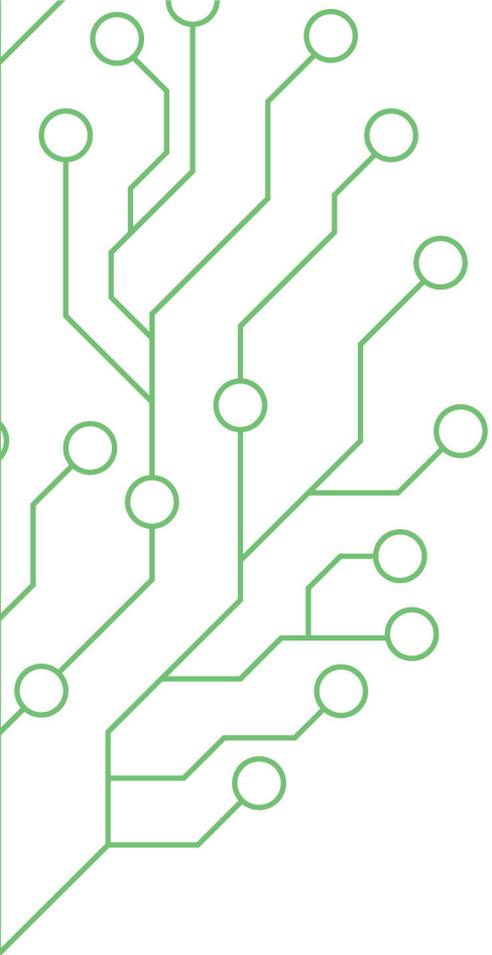
**加强政策供给，提供组织保障。**各地应制定专门针对政府数据开放的地方性法规、地方政府规章或规范性文件，对开放数据治理、开放方式、推动利用、安全保护、保障机制等方面做出明确的规定和要求。地方党政主要领导应更多公开表态支持数据开放，各地应制定和公开专门针对数据开放的标准规范和年度工作计划。

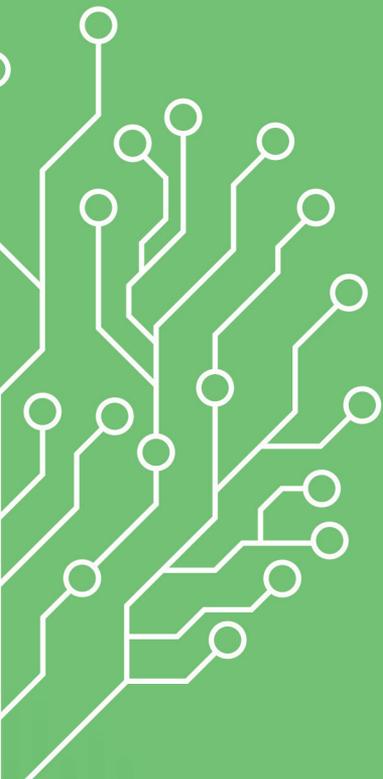
**优化平台服务，提升运维能力。**在数据发现方面，各地平台应改善数据概况和动态的内容及展现方式，提供按用户使用场景和标签分类进行导航的方式，提供按字段名称和多种筛选条件搜索的功能。在数据获取方面，建议平台提供可下载的开放数据目录，对于无条件开放的数据实现分级登陆获取，公开并回复用户对未开放数据的请求，真正开通可申请有条件开放数据的功能。在工具提供方面，平台应注重提高工具的多样性和实用性。在利用展示方面，建议增加提交和展示利用成果的类型，标明利用成果的来源信息。在互动反馈方面，建议平台提供数据发布者联系方式，公开并及时回复用户的意见建议和数据纠错信息，从而在数据提供者和用户之间搭建起真正的供需对接桥梁。在公众传播方面，各地应加大对数据开放工作的宣传力度，探索多渠道、多平台传播。在账户体验方面，建议同时提供面向自然人和法人的账户注册功能，提升账户注册的便捷程度。

**提高数据质量，分级分类开放。**在数据数量方面，各地应不断提升数据集的容量。在数据质量方面，推出更多高容量、高需求的优质数据集，同时减少碎片化和低容量的低质数据集，剔除重复创建、生硬格式转化和无效数据集；建议增加以API接口的形式开放的容量大、更新频率高的数据集，清理碎片化API接口、限制型API接口、无效API接口和重复API接口。在数据标准方面，平台应制定和完善开放授权协议，对用户免费获取、不受歧视获取、自由利用、自由传播与分享开放数据的权利进行明确授权，并且探索对不同类型的数据集分级分类制定开放授权协议。提高RDF格式开放数据的比例，降低申请API接口的难度。平台应推进数据的分级分类开放，对数据集标示不同开放类型，如无条件开放或有条件开放等；提供更为完善的元数据和更加详实的API接口调用说明，帮助用户理解、获取和利用数据。在数据覆盖面方面，各地应鼓励和吸引更多的部门参与到数据开放工作中来，丰富数据内容的多样性，不断提高开放数据集的主题覆盖面和部门覆盖面。数据开放工作刚刚起步或水平不高的地方，可以参考先进地方已经开放的常见数据集，积极稳妥地补齐短板。在数据持续性方面，各地应保障平台的持续性管理和运维，使其成为政府部门的一项常态化工作，以支撑开放数据集的持续增长、动态更新和历史存档。

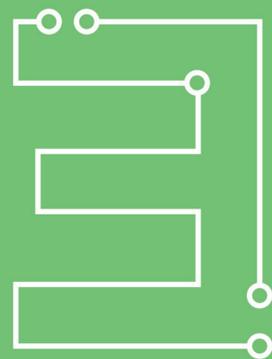
**促进创新利用，营造有利生态。**在开放数据利用促进方面，各地应积极举办并持续开展类似于开放数据创新利用比赛的促进活动，用以提升社会利用政府数据的积极性，营造有利于政府数据开放利用的生态体系。在成果产出方面，平台亟需提升有效、优质利用成果的类型、数量和主题覆盖面，清理政府自身开发的应用，剔除无法获取或正常使用的应用，为利用成果标明数据来源并提供数据集链接。在数据利用方面，地方应鼓励引导多种类型的数据利用者参与政府开放数据的开发利用。开放优质的、高价值的政府数据是推动数据利用的基本条件，各地应把数据供给做实做优，然后在此基础上推动数据利用，提升被利用数据集的数量、容量、主题覆盖面和部门覆盖面。

# INDEX AND RANKING





CHAPTER



# 开放数林指数与排名

### 三、中国开放数林指数与排名

#### 3.1 省级

省级开放数林指数分值与排名如下，上海市排名第一，浙江省和贵州省分列第二、三位，前六名地方获得“数开叶茂”、“数开成荫”、“数开丛生”等奖项，这些地方在准备度、平台层、数据层与利用层方面的综合表现突出。

#### 3.2 地级（含副省级）

地级（含副省级）开放数林指数分值与排名如下，贵阳市排名第一，济南市、哈尔滨市分列第二、第三，前五名地方获得“数开叶茂”、“数开成荫”、“数开丛生”等奖项，这些地方在准备度、平台层、数据层与利用层方面的综合表现突出。

除综合奖项外，今年还新设了单项奖。浙江省凭借其在数据层的优异表现获得“独数一帜”奖；深圳市进步迅速，获得“数飞猛进”奖；福建省、成都市在最新开放数据的地方中综合表现最为优异，获得“新数倍出”奖。

#### ■ 省级

地方	数据层			平台层			准备度			利用层		综合指数	总排名	排名变化	奖项	单项奖
	指数	排名	排名变化	指数	排名	排名变化	指数	排名	排名变化	指数	排名					
上海	20.22	5	↓4	10.91	5	↓2	18.65	2	-	6.50	3	56.28	1	-	数开叶茂	
浙江	29.19	1	↑6	13.63	1	↑9	10.53	6	↓2	2.40	5	55.75	2	↑5	数开成荫	独数一帜
贵州	19.66	7	↓1	7.88	8	↓3	19.13	1	-	8.38	2	55.05	3	-	数开成荫	
北京	21.59	2	-	7.71	9	↓3	10.18	7	↑2	9.04	1	48.52	4	↑1	数开丛生	
福建	21.21	3	/	11.35	4	/	11.60	4	/	0.00	10	44.16	5	/	数开丛生	新数倍出
山东	21.20	4	↑1	11.46	3	↓2	7.95	8	↓5	0.00	10	40.61	6	↓4	数开丛生	
天津	19.94	6	/	6.94	10	/	11.78	3	/	1.50	6	40.16	7	/	十强	
广东	16.54	9	↓6	12.85	2	-	6.75	9	↓3	2.70	4	38.84	8	↓4	十强	
海南	15.23	10	/	4.58	13	/	11.05	5	/	1.50	6	32.36	9	/	十强	
河南	17.20	8	↓4	9.98	6	↓2	4.20	10	↓3	0.00	10	31.38	10	↓4	十强	
江西	14.54	11	↓3	6.13	11	↓2	1.80	13	↓5	1.50	6	23.97	11	↓2		
宁夏	11.58	12	↓3	8.66	7	-	2.70	12	↓2	0.00	10	22.94	12	↓2		
陕西	9.82	13	↓3	4.92	12	↓4	4.00	11	↓6	1.00	9	19.74	13	↓5		

↗ 表示新上线平台

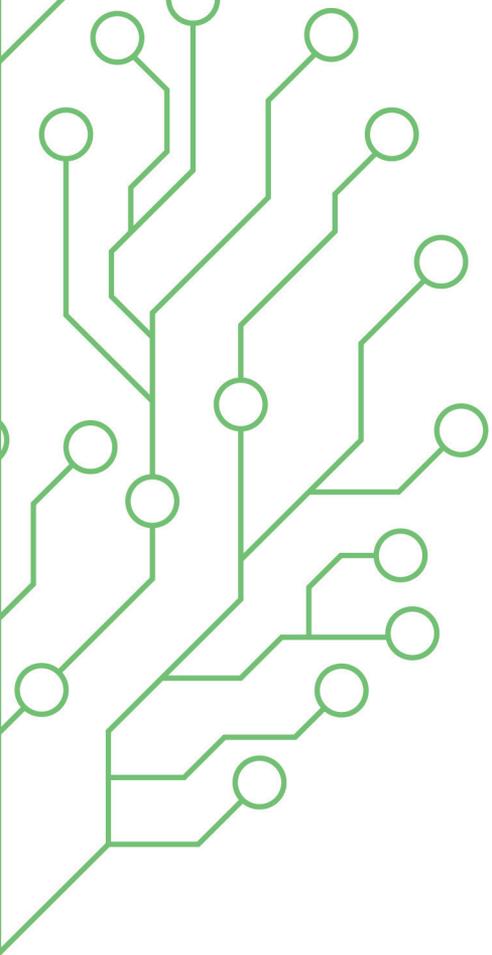
— 表示名次未变动

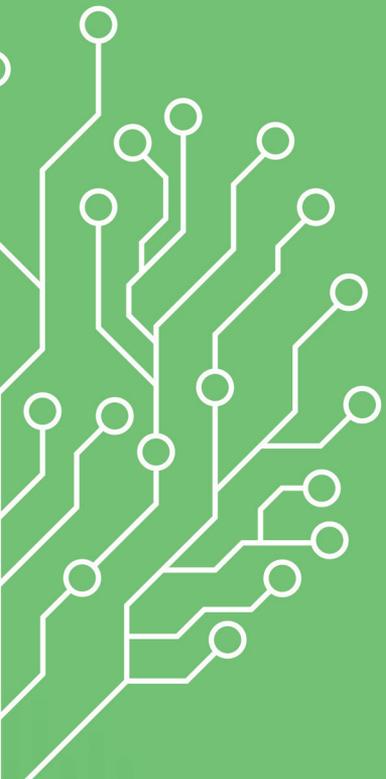
■ 地级（含副省级）

地方	数据层			平台层			准备度			利用层		综合指数	总排名	排名变化	奖项	单项奖
	指数	排名	排名变化	指数	排名	排名变化	指数	排名	排名变化	指数	排名					
贵阳	26.27	1	↑1	15.87	1	-	21.79	1	-	6.04	1	69.97	1	-	数开叶茂	
济南	23.72	3	↑3	11.50	12	↓9	17.98	2	↑14	0.50	31	53.70	2	↑5	数开成荫	
哈尔滨	20.67	7	↓3	13.01	3	↓1	13.93	7	↓5	3.20	4	50.81	3	-	数开成荫	
深圳	24.72	2	↑6	14.37	2	↑25	6.88	13	↑21	3.20	4	49.17	4	↑16	数开丛生	数飞猛进
青岛	20.66	8	↓1	10.70	16	↓11	13.66	9	↓2	0.50	31	45.52	5	-	数开丛生	
威海	15.87	18	↓8	11.22	13	↓5	13.96	6	↑2	0.50	31	41.55	6	↑2	十强	
银川	15.11	20	↑22	10.06	28	↓4	15.83	4	↓1	0.50	31	41.50	7	↑5	十强	
成都	18.64	11	↓2	9.47	33	↓7	10.15	11	↓7	3.20	4	41.46	8	↓1	十强	新数倍出
潍坊	20.78	6	↑10	10.60	19	↓7	5.14	18	↑6	4.20	2	40.72	9	↑7	十强	
广州	22.16	4	↓1	10.89	15	↓8	5.14	18	↑8	0.50	31	38.69	10	↓6	十强	
江门	12.37	40	↓21	9.40	35	↓24	16.05	3	↑9	0.00	64	37.82	11	-	二十强	
烟台	19.63	10	↑21	10.60	19	↓7	5.14	18	↑21	1.50	7	36.87	12	↑17	二十强	
佛山	19.97	9	↓8	11.07	14	↓10	5.14	18	↑18	0.50	31	36.68	13	↓11	二十强	
东莞	21.19	5	-	8.46	38	↓4	5.14	18	↓4	1.50	7	36.29	14	↓2	二十强	
淄博	17.69	13	↓1	10.20	27	↓15	5.14	18	↑3	1.50	7	34.53	15	↓1	二十强	
宣城	12.64	37	↓5	5.32	47	↓16	15.56	5	↑6	0.50	31	34.02	16	↑14	二十强	
滨州	14.94	23	↑3	10.60	19	↓7	6.88	13	↓8	1.50	7	33.92	17	↓8	二十强	
潮州	14.38	29	/	12.45	4	/	6.45	15	/	0.50	31	33.78	18	/	二十强	
济宁	17.78	12	↑16	10.30	25	↓13	5.14	18	↑14	0.50	31	33.72	19	↑5	二十强	
德州	17.38	14	-	10.70	16	↓8	5.14	18	↑22	0.50	31	33.72	19	↑2	二十强	
临沂	16.75	15	↓2	10.70	16	↓4	5.14	18	↑23	1.00	30	33.59	21	↑1	三十强	
菏泽	14.67	27	↓10	10.60	19	↓7	6.45	15	-	1.50	7	33.22	22	↓12	三十强	
佳木斯	7.64	66	/	10.05	29	/	13.71	8	/	1.50	7	32.90	23	/	三十强	
揭阳	13.19	36	/	12.45	4	/	5.14	18	/	1.50	7	32.28	24	/	三十强	
日照	14.71	26	↓11	10.60	19	↓7	5.14	18	↑29	1.50	7	31.95	25	↑3	三十强	
汕尾	13.85	31	/	12.45	4	/	5.14	18	/	0.50	31	31.94	26	/	三十强	
东营	14.56	28	↓10	10.60	19	↓7	5.14	18	↓5	1.50	7	31.80	27	↓15	三十强	
汕头	13.59	33	/	12.45	4	/	5.14	18	/	0.50	31	31.68	28	/	三十强	
泰安	14.84	24	↓13	9.80	30	↓18	5.14	18	↑17	1.50	7	31.28	29	↓12	三十强	
枣庄	15.01	22	↑2	9.40	35	↓23	5.14	18	↑2	1.50	7	31.05	30	↓11	三十强	
惠州	15.04	21	↑4	10.21	26	↓2	5.14	18	↑7	0.50	31	30.89	31	↓8	四十强	
聊城	14.28	30	↓9	9.80	30	↓18	6.01	17	↑5	0.50	31	30.59	32	↓14	四十强	
云浮	12.00	43	/	12.45	4	/	5.14	18	/	0.50	31	30.09	33	/	四十强	
清远	11.89	44	/	12.45	4	/	5.14	18	/	0.50	31	29.98	34	/	四十强	
韶关	11.28	47	/	12.45	4	/	5.14	18	/	0.50	31	29.37	35	/	四十强	
河源	9.98	57	/	12.45	4	/	5.14	18	/	1.50	7	29.07	36	/	四十强	
珠海	16.31	16	↑29	7.08	43	↓14	5.14	18	↑28	0.50	31	29.03	37	↑7	四十强	
泰州	15.66	19	/	9.18	37	/	3.26	51	/	0.50	31	28.60	38	/	四十强	
连云港	16.17	17	/	5.68	44	/	3.26	51	/	0.50	31	25.61	39	/	四十强	
遵义	12.43	39	/	8.02	39	/	4.79	49	/	0.00	64	25.24	40	/	四十强	
阜阳	13.35	35	/	7.96	40	/	3.26	51	/	0.50	31	25.07	41	/	五十强	
扬州	6.54	69	↓26	4.16	54	↓15	12.76	10	↑19	1.50	7	24.96	42	↓2	五十强	
湖州	12.10	42	/	2.12	69	/	9.22	12	/	1.50	7	24.94	43	/	五十强	
中山	13.43	34	↓4	5.50	45	↓10	5.14	18	↑15	0.50	31	24.57	44	↓8	五十强	
厦门	13.67	32	/	7.48	41	/	2.61	68	/	0.50	31	24.26	45	/	五十强	
武汉	11.02	51	↓14	9.43	34	↓1	3.26	51	↓41	0.50	31	24.21	46	↓15	五十强	
无锡	11.69	45	↓16	2.70	63	↓25	5.00	46	↓37	4.20	2	23.59	47	↓15	五十强	
乌海	11.03	50	↓16	9.48	32	↓26	2.61	68	↓37	0.00	64	23.12	48	↓22	五十强	
常州	14.76	25	/	4.86	51	/	3.26	51	/	0.00	64	22.88	49	/	五十强	
苏州	12.12	41	↓19	3.08	59	↓18	5.00	46	↓23	1.50	7	21.70	50	↓15	五十强	
梅州	11.16	49	↓26	3.16	58	↓14	5.14	18	↑12	1.50	7	20.96	51	↓13		
宁波	11.36	46	↓26	5.22	48	↓20	3.26	51	↓34	0.50	31	20.34	52	↓27		
蚌埠	10.37	54	/	5.18	50	/	3.26	51	/	1.50	7	20.31	53	/		
肇庆	11.20	48	↓13	2.28	67	↓21	5.14	18	↑10	1.50	7	20.12	54	↓15		
南通	12.49	38	/	3.82	55	/	3.26	51	/	0.50	31	20.07	55	/		
马鞍山	8.48	62	↓29	7.28	42	↓12	3.26	51	↓24	0.50	31	19.52	56	↓22		
铜仁	10.57	52	↓16	3.80	56	↓25	4.79	49	↓31	0.00	64	19.16	57	↓24		
淮安	9.86	58	/	4.34	53	/	3.26	51	/	1.50	7	18.96	58	/		
南京	10.54	53	/	2.74	61	↓18	5.00	46	↓40	0.50	31	18.78	59	↓22		
常德	10.00	56	/	3.48	57	/	3.26	51	/	1.50	7	18.24	60	/		
湛江	7.71	65	↓27	4.84	52	↓15	5.14	18	↑26	0.50	31	18.19	61	↓19		
黄山	8.49	61	/	5.38	46	/	3.26	51	/	0.50	31	17.63	62	/		
茂名	8.24	63	/	2.30	65	/	5.14	18	/	1.50	7	17.18	63	/		
长沙	10.22	55	↓15	2.74	61	↓16	3.26	51	↓9	0.50	31	16.72	64	↓19		
六安	7.57	67	↓20	5.20	49	↓13	3.26	51	↓32	0.00	64	16.03	65	↓24		
雅安	8.57	60	/	2.52	64	/	3.26	51	/	1.50	7	15.85	66	/		
阳江	7.91	64	↓20	2.28	67	↓20	5.14	18	↑27	0.50	31	15.83	67	↓21		
荆门	8.82	59	↓20	2.30	65	↓25	3.26	51	↓14	0.50	31	14.88	68	↓25		
石嘴山	7.14	68	↓22	2.90	60	↓18	3.26	51	↓8	1.50	7	14.80	69	↓22		

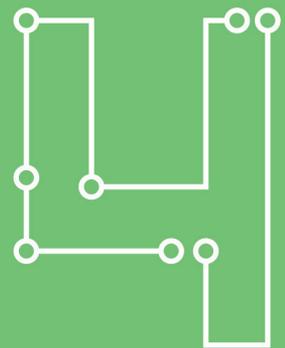
✓ 表示新上线平台    - 表示名次未变动

# READINESS





CHAPTER



# 准备度发现



## 四、准备度

报告从以下三个维度对准备度进行了评测，发现如下：

**在法规与政策方面**，目前尚未有地方出台专门针对数据开放的地方性法规、地方政府规章或规范性文件，已出台的地方法规与政策或是将数据开放与数据共享合并制定，或是将数据开放作为公共数据资源治理或大数据发展促进的一个组成部分来制定，其文本中关于数据开放的内容较为笼统。

从现有地方法规与政策的内容来看：在开放数据治理上，仅有 13% 的地方提出重点开放或优先开放领域，超过 20% 的地方明确要求编制开放数据目录与清单，17% 的地方对开放数据的质量提出了相对全面的要求。在开放方式上，23% 的地方对开放平台的建设作出了要求，但仅有 8.5% 的地方对数据分级分类开放作出了明确要求。在推进利用上，12% 的地方提出要加强宣传推广和促进利用，15.9% 的地方提出要建立与数据利用者之间的沟通机制，但仅有 4.9% 的地方对数据纠错机制提出了要求。在安全保护上，15.9% 的地方明确要求开放前进行数据审查，但仅有 7.3% 的地方对平台安全和开放后的风险管控作出明确规定。在保障机制上，现有法规与政策对协调机制、人员能力、资金保障和考核监督等各项保障机制作出明确要求的占比都未超过 10%。

**在组织与执行方面**，在新一轮机构改革背景下，各地纷纷组建了专门的政府数据管理部门，然而超过八成的地方未对数据开放相关部门作出清晰的职责界定。八成地方仍未发现有主要领导在公开发表的讲话中明确支持数据开放，在已发现的 16 个地方的高层领导公开讲话中，关于数据开放的讲话内容丰富全面的不到 19%。目前只有上海市连续制定和公开了专门针对政府数据开放的年度工作计划。

**在标准规范建设方面**，目前仅有广东、山东两省制定了专门针对数据开放的地方标准，其中山东省制定的标准体系较为全面，对山东省各级政务部门的数据开放、数据脱敏以及数据开放程度评价工作均制定了标准。贵州省虽未专门针对数据开放制定标准，但在政府数据分级分类的标准规范中也涉及到了有关数据开放的内容。

---

综上所述，**报告建议**：在法规与政策方面，各地应制定专门针对政府数据开放的地方性法规、地方政府规章或规范性文件，对开放数据治理、开放方式、推进利用、安全保护、保障机制等方面作出明确的规定和要求。地方党政主要领导应更多公开表态支持数据开放，各地应制定和公开专门针对数据开放的标准规范和年度工作计划。

## 4.1 法规与政策

法规与政策是指对政府数据开放相关环节（如开放数据治理、开放方式、推进利用、安全保护以及保障机制）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。目前，国家层面上尚无独立针对政府数据开放制定的单行法律，报告对地方层面现有相关法规与政策的文本效力和内容要求两个方面进行了分析。

### 4.1.1 文本效力

文本效力是指法律法规和政策文件所具有的约束力，主要表现在等级和专门性两方面。本部分的评估范围包括专门针对政府数据开放的方案意见，以及文本中至少有一节内容专门针对政府数据开放的法规政策。

#### (1) 等级

法律法规与政策的等级直接决定了其约束力的强弱。报告将各地有关政府数据开放的法规与政策的等级从高到低划分为地方性法规、地方政府规章、规范性文件、方案意见四级。图 4-1 展示了制定各个等级的法规与政策在地方的分布情况。目前只有贵州省、天津市和贵阳市通过了地方性法规，为当地政府数据开放提供了有力的制度保障。此外，也有地方以政府令的形式发布了与政府数据开放相关的地方政府规章或制定了规范性文件和专门性方案意见。

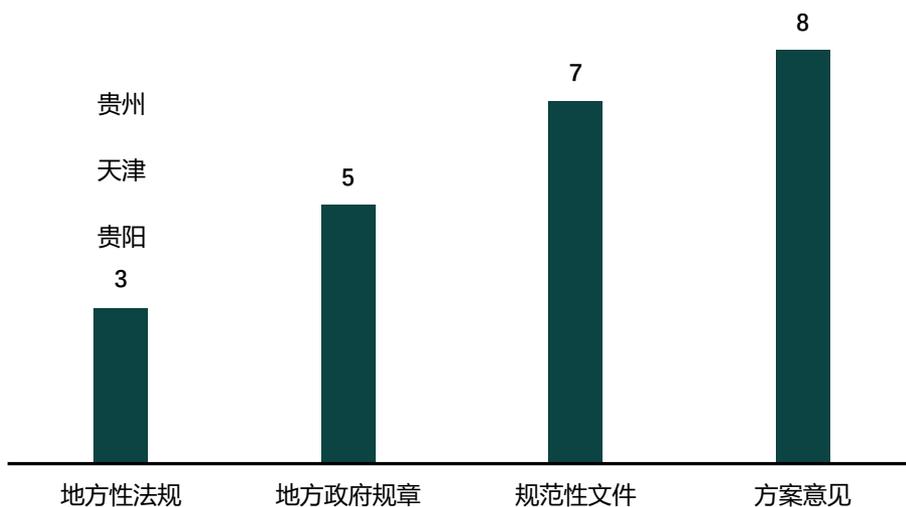


图 4-1 各地法规与政策的效力等级情况分布

#### (2) 专门性

专门性是指法规与政策是否专门为政府数据开放而制定，这也是衡量其效力的重要指标。目前，尚未有地方出台专门针对数据开放的地方性法规、地方政府规章或规范性文件，已出台的法规与政策在内容上主要分为四类：针对公共数据资源共享和开放、针对公共数据资源管理、针对大数据发展应用促进以及针对公共数据和一网通办（或电子政务）管理，整体分布情况如图 4-2 所示。值得一提的是，上海市已于 2019 年 4 月 30 日就《上海市公共数据开放管理办法（草案）》对外征求公众意见，这或将成为国内首部专门针对政府数据开放的地方政府规章。

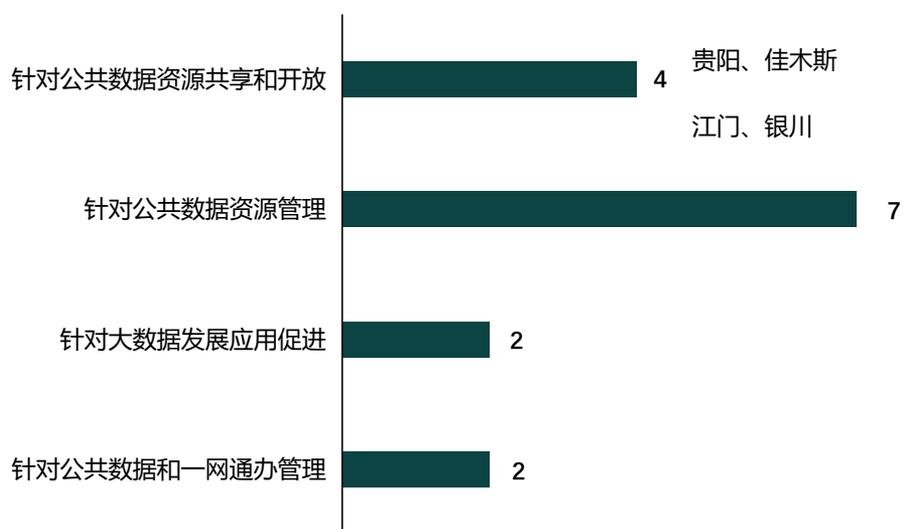


图 4-2 高等级法规与政策专门性情况分布

贵阳市已经出台了四项专门性程度较高的法规与政策，其中《贵阳市政府数据共享开放条例》是我国首部也是目前唯一针对数据共享和开放的地方性法规。有些地方出台的法规与政策只涉及到政府数据开放的某一两个方面，内容较为简略，例如在“政务信息系统项目管理办法”这类文件中，仅对数据开放平台作出简单规定，这类文件未被纳入本报告的评估范围。

表 4-1 中列举了地方政府制定的专门针对数据开放的方案意见，其余地方政府尚未制定专门的规划方案，而只将数据开放列为政务公开、政务信息系统整合共享、“互联网+”政务服务等工作下的分支任务，虽对数据开放作出了一些原则性规定，但内容空泛单薄，这类文件未纳入本报告的评估范围。

表 4-1 各地政府数据开放专门性方案意见列表

地方	专门针对数据开放的方案意见
哈尔滨	《哈尔滨市推进政府数据向社会开放工作实施方案的通知》
济南	《中共济南市委办公厅济南市人民政府办公厅关于推进公共信息资源向社会开放的通知》
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》
上海	《上海市政府办公厅关于推进政府信息资源向社会开放利用工作的实施意见》
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》
宣城	《宣城市人民政府办公室关于宣城市公共信息资源开放工作实施意见》
贵阳	《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》
扬州	《扬州市大数据资源整合共享开放工作实施方案》

### 4.1.2 内容要求

报告对纳入评估范围的法规与政策中的内容进行了分析，内容包括开放数据治理、开放方式、推进利用、安全保护以及保障机制五个方面的要求。如词云分析图 4-3 所示，目前我国各地对数据开放的法规与政策侧重开放数据的标准、质量、社会需求、服务能力、安全、应用、部门分工、管理以及营造社会氛围等方面。



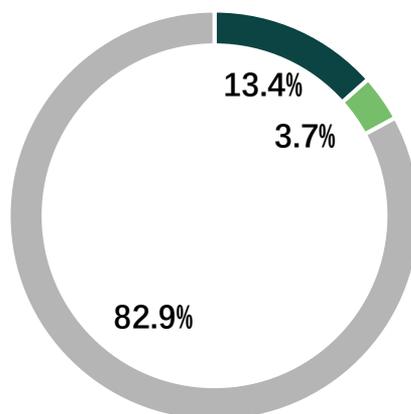
图 4-3 法规与政策内容关键词词云图

#### (1) 开放数据治理

开放数据治理是指对开放数据资源进行管理，良好的数据治理是数据开放的基础和保障。报告对各地现有法规和政策中有关开放数据范围、开放数据目录与清单、开放数据质量的内容要求进行了分析。

### ■ 开放数据范围

目前仅有 11 个地方明确提出了政府数据开放的重点开放或优先开放领域，如天津市规定“市和区人民政府及其有关部门应当采取措施，推进政务数据集中向社会开放，优先推动信用、交通、医疗、就业、社保、教育、环境、气象、企业登记监管等民生保障服务相关领域的政务数据向社会开放”<sup>1</sup>。如图 4-4 所示，还有个别地方只是简要概述了数据开放的范围，大部分地方未对开放范围作出明确政策规定。

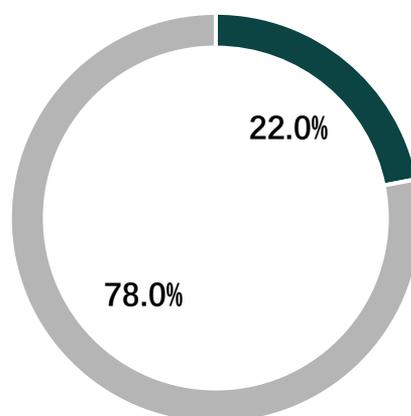


■ 提出重点开放或优先开放领域 ■ 简要概述开放范围 ■ 无相关要求

图 4-4 各地有关开放数据范围要求情况分布

### ■ 开放数据目录与清单

目前有 18 个地方提出编制政府数据资源开放目录、实行动态清单管理、及时维护和更新目录的要求，如图 4-5 所示。



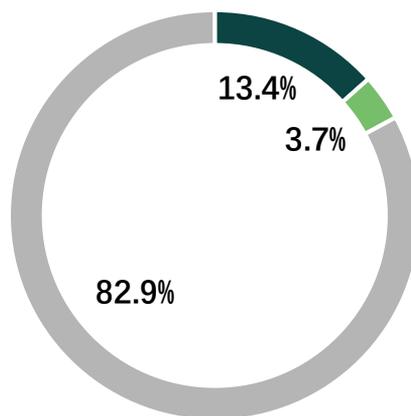
■ 提出编制开放数据目录与清单 ■ 无相关要求

图 4-5 各地有关开放数据目录与清单要求情况分布

<sup>1</sup> 《天津市促进大数据发展应用条例》[EB/OL]. [2019-01-01].  
[http://dsj.guizhou.gov.cn/xwzx/gnpd/201812/t20181219\\_3372603.html](http://dsj.guizhou.gov.cn/xwzx/gnpd/201812/t20181219_3372603.html)

■ 开放数据质量

目前只有 11 个地方对开放数据的质量作出了相对全面的要求，提出要符合开放数据资源的真实性、准确性、完整性、及时性、有效性、原始性和可用性标准。个别地方只对数据质量的某些方面作出要求，大部分未作出明确要求，如图 4-6 所示。



■ 对开放数据质量提出相对全面要求 ■ 提出部分要求 ■ 无相关要求

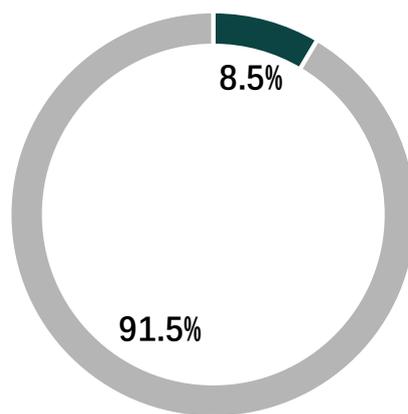
图 4-6 各地有关开放数据质量要求情况分布

(2) 开放方式

开放数据的方式是指推进数据开放工作的手段和渠道。报告对各地现有法规与政策中有关分级分类开放与开放平台建设的要求进行了分析。

■ 分级分类开放

根据数据资源的体量、实时性和安全要求等特性，对数据进行分级分类开放，是深入推进政府数据开放的重要手段。如图 4-7 所示，目前仅有 7 个地方对数据分级分类开放作出明确要求。如贵州省规定：“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”<sup>2</sup>。



■ 要求分级分类开放 ■ 无相关要求

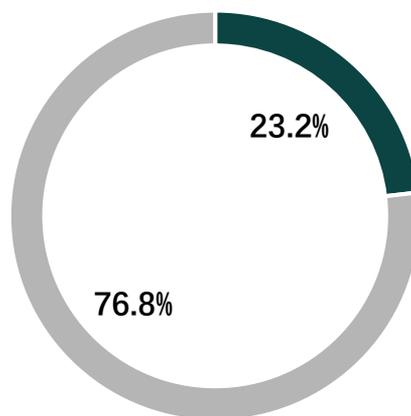
图 4-7 各地数据分级分类开放要求情况分布

<sup>2</sup> 《贵州省政务数据资源管理暂行办法》[EB/OL]. [2016-11-01].

[http://www.gzgov.gov.cn/xxgk/zfxzkgkpt/szfxzkgkml/aztfl/zhzw/201709/t20170919\\_793971.html](http://www.gzgov.gov.cn/xxgk/zfxzkgkpt/szfxzkgkml/aztfl/zhzw/201709/t20170919_793971.html)

### ■ 开放平台建设

政府数据开放平台是政府开放数据和用户获取数据的载体，是展示应用的中心，也是连接开放数据供给侧和用户的桥梁。如图 4-8 所示，有 19 个地方对建设或完善数据开放平台，集中统一向社会开放政府数据资源作出要求，其中有些地方还明确了平台功能与运维等具体要求。如贵阳市规定：“开放平台是发布全市开放的政府数据资源目录和向公民、法人及其他组织开放政府数据的公共基础平台。市大数据行政主管部门应当组织实施开放平台的建设、维护管理工作。市大数据行政主管部门可以依法委托具备相应信息技术服务条件的机构对前款规定平台进行运行维护管理，并依法签订委托合同，明确双方权利义务”<sup>3</sup>。



■ 要求建设数据开放平台 ■ 无相关要求

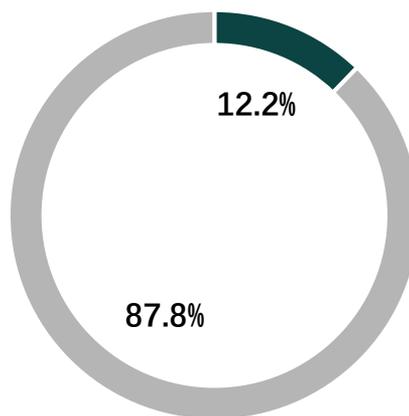
图 4-8 各地有关数据开放平台的要求情况分布

### (3) 推进利用

推进对政府数据的开发与利用是数据开放的根本目的。报告对各地现有政策法规中有关推进数据利用的要求进行了分析，包括宣传推广、利用促进和互动反馈三方面。

#### ■ 宣传推广

宣传推广是指政府部门通过宣传引导，营造推进数据利用的良好社会氛围。如图 4-9 所示，目前仅有 10 个地方明确提出要引导企业利用政府开放数据开展示范应用，积极营造社会广泛参与和利用开放数据的良好氛围。例如，哈尔滨市提出：“各部门要充分利用各类宣传媒体，大力宣传开放政府数据的相关理念和做法，提升社会大众的关注度、参与度和支持度，并借此与公众就相关主题展开切实有效的双向沟通，积极引导社会力量开发利用公共信息资源，促进信息惠民和社会创新”<sup>4</sup>。其他地方未对此作出明确要求。



■ 要求加强宣传推广 ■ 无相关要求

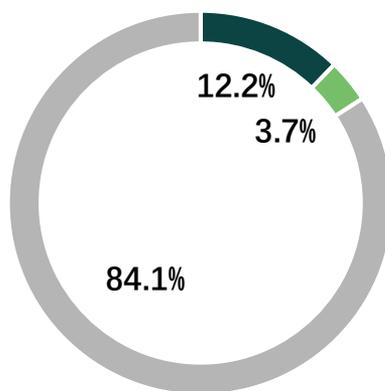
图 4-9 各地宣传推广要求情况分布

<sup>3</sup> 《贵阳市政府数据共享开放实施办法》[EB/OL]. [2018-03-01]. <http://www.guiyang.gov.cn/zfxxgk/fgwj/zfgz/20180123/i1459759.html>

<sup>4</sup> 《哈尔滨市推进政府数据向社会开放工作实施方案的通知》[EB/OL]. [2017-03-10]. [http://www.harbin.gov.cn/art/2017/3/24/art\\_13791\\_2106.html](http://www.harbin.gov.cn/art/2017/3/24/art_13791_2106.html)

■ 利用促进

利用促进是指地方政府采用举办开放数据创新应用竞赛、提供创业孵化服务、设立创新基金、减免税收等举措推进政府数据的社会化利用。如图 4-10 所示，目前有一成多的地方明确提出了利用促进的各项具体措施，如江门市规定：“通过政府购买服务、专项资金扶持和数据应用竞赛等方式，鼓励和支持公民、法人和其他组织利用政府数据创新产品、技术和服务，推动政府数据开放工作，提升政府数据应用水平”<sup>5</sup>。其他地方或仅笼统提出支持企业利用政府开放的数据创新产品、技术和服务，但未提出具体措施，或未对此作出规定。



■ 提出促进利用的具体措施 ■ 未提出具体措施 ■ 无相关要求

图 4-10 各地利用促进要求情况分布

■ 互动反馈

互动反馈是指数据提供者和数据利用者之间的沟通机制和纠错机制。

沟通机制是指政府通过各种方式和渠道主动收集并及时回应数据利用者的数据需求。目前只有 13 个地方对此作出了相关要求，如浙江省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相关公众的意见、建议（如图 4-11）。

纠错机制是指当数据利用者为已开放的数据有疑义或发现错误时，可及时向数据提供部门反馈，相关部门接到反馈后，应及时校核并予以纠正。目前，只有 5 个地方对此作出了相关要求。例如，《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”<sup>6</sup>。

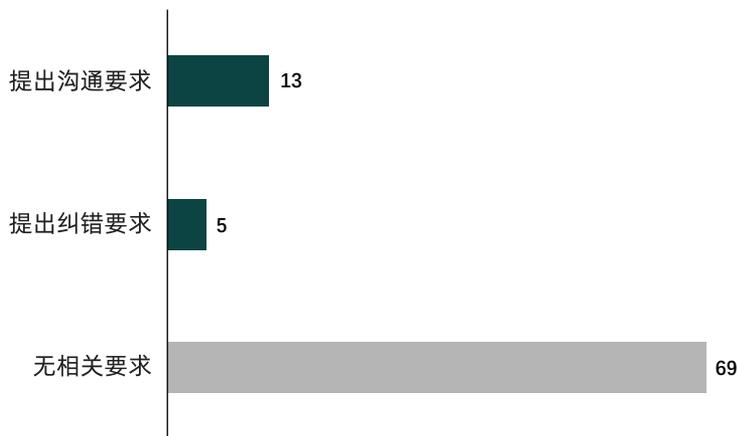


图 4-11 各地互动反馈要求情况分布

<sup>5</sup> 《江门市政务数据资源共享和开放管理暂行办法》[EB/OL]. [2018-06-15]. [http://fzj.jiangmen.gov.cn/zwgk/wgkzl/jcgk/201806/t20180619\\_1495196.html](http://fzj.jiangmen.gov.cn/zwgk/wgkzl/jcgk/201806/t20180619_1495196.html)

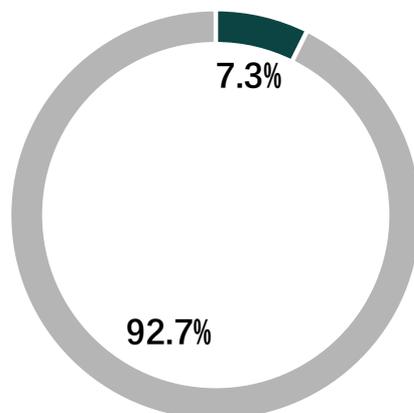
<sup>6</sup> 《贵阳市政府数据共享开放实施办法》[EB/OL]. [2018-01-12]. <http://www.gygov.gov.cn/c9682/20180123/i1459759.html>

#### (4) 安全保护

安全是开放的前提，保护是为了更好、更可持续地开放。报告主要从平台安全、开放前数据审查、开放后风险管控三个方面对各地现有法规和政策中的相关内容进行了分析。

##### ■ 平台安全

平台安全是指数据开放平台根据国家信息安全等级保护等相关要求，保障平台安全运行。如图 4-12 所示，目前只有 6 个地方对数据开放平台安全作出明确规定，如贵州省提出要构建平台的安全防护体系和认证体系<sup>7</sup>。

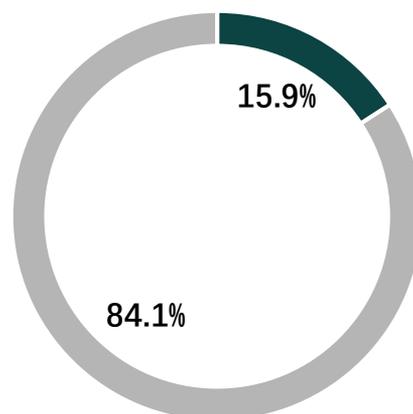


■ 要求保障平台安全 ■ 无相关要求

图 4-12 各地平台安全要求情况分布

##### ■ 开放前数据审查

数据资源在开放前应按照相关法律法规进行安全审查。如图 4-13 所示，目前只有浙江省等 13 个地方对此作出了明确要求，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”<sup>8</sup>。



■ 要求进行安全审查 ■ 无相关要求

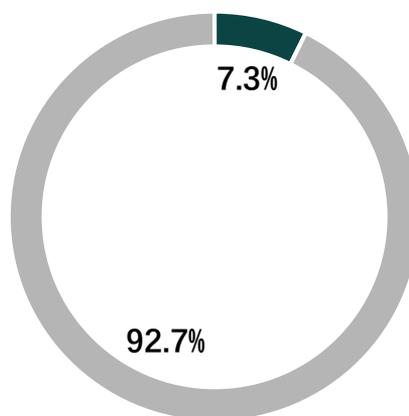
图 4-13 各地开放前数据审查要求情况分布

<sup>7</sup> 《贵州省政务数据资源管理暂行办法》[EB/OL]. [2016-11-01].  
[http://www.gztyw.gov.cn/zhxw/szfwj/201611/t20161114\\_1358738.html](http://www.gztyw.gov.cn/zhxw/szfwj/201611/t20161114_1358738.html)

<sup>8</sup> 《浙江省公共数据和电子政务管理办法》[EB/OL]. [2018-02-01].  
[http://zfxgk.zj.gov.cn/xxgk/jcms\\_files/jcms1/web14/site/art/2018/2/1/art\\_1082\\_1987774.html](http://zfxgk.zj.gov.cn/xxgk/jcms_files/jcms1/web14/site/art/2018/2/1/art_1082_1987774.html)

■ 开放后风险管控

数据在开放后应对其利用情况进行监测评估和风险管控。如图 4-14 所示，目前有 6 个地方对此作出了明确要求。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”<sup>9</sup>。其他地方尚未对此作出明确规定。



■ 要求进行风险管控 ■ 无相关要求

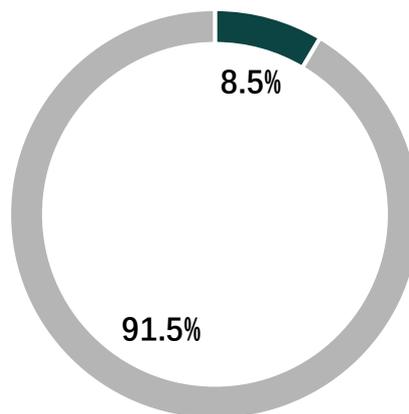
图 4-14 各地开放后风险管控要求情况分布

(5) 保障机制

健全的保障机制是推进数据开放工作的“润滑剂”与“加速器”。报告从协调机制、人员能力、资金保障和考核监督机制四个方面对现有法规与政策中的相关要求进行了分析。

■ 协调机制

协调机制是指建立推进政府数据开放工作的协调机制。如图 4-15 所示，目前仅有 7 个地方作出了有关要求。例如上海市要求建立专门的信息资源开放工作领导协调机制，扬州市要求建立健全全市大数据资源整合共享开放协调机制，定期召开由各地、市级各部门相关人员参加的工作协调会议<sup>10</sup>。



■ 要求建立协调机制 ■ 无相关要求

图 4-15 各地协调机制要求情况分布

<sup>9</sup> 《贵州省大数据发展应用促进条例》[EB/OL]. [2016-03-01]. <http://www.gzrd.gov.cn/dffg/sgdfxfg/23501.shtml>

<sup>10</sup> 《扬州市大数据资源整合共享开放工作实施方案》[EB/OL]. [2018-01-08]. [http://www.yangzhou.gov.cn/yangzhou/szfbgswj/2018-01/08/content\\_28a4c7a3d9374e4ba300930394f0cf45.shtml](http://www.yangzhou.gov.cn/yangzhou/szfbgswj/2018-01/08/content_28a4c7a3d9374e4ba300930394f0cf45.shtml)

### ■人员能力

人员能力是指要求指定专人负责数据开放工作和开展数据开放工作业务培训。相关政策情况分布如图 4-16 所示。个别地方同时提出了以上两项要求，如贵阳市规定：“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”<sup>11</sup>。其余地方或只提到一方面，或未有相关要求。

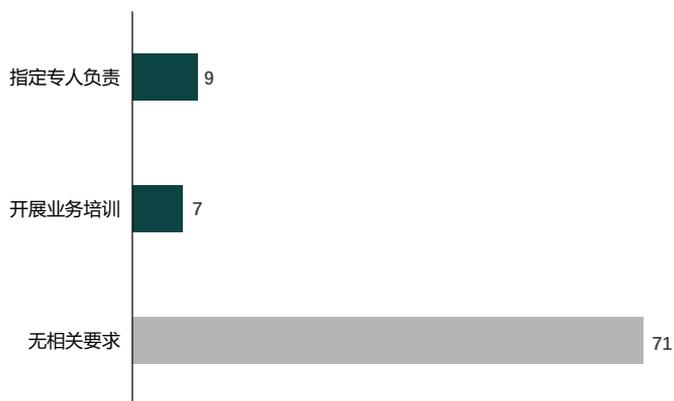
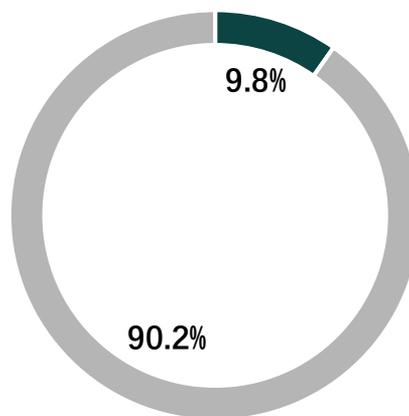


图 4-16 各地人员能力要求情况分布

### ■资金保障

资金保障是指为政府数据开放工作提供有力的经费保障。相关政策情况分布如图 4-17 所示。目前仅有 8 个地方明确要求为数据开放提供资金保障，如上海市要求：“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”<sup>12</sup>。



■要求提供资金保障 ■无相关要求

图 4-17 各地资金保障要求情况分布

<sup>11</sup> 《贵阳市政府数据共享开放实施办法》[EB/OL]. [2018-01-12].  
<http://www.gygov.gov.cn/c9682/20180123/i1459759.html>

<sup>12</sup> 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》[EB/OL]. [2014-05-05].  
<http://sjj.sh.gov.cn/sj2014/zwgk/n386/n604/u1ai22197.html>

■ 考核监督机制

考核监督机制是指对数据开放工作的效果进行评估考核，并对相关机构与工作人员的工作进行监督。目前有 18 个地方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”<sup>13</sup>。

还有 12 个地方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。如成都市规定：“政务部门及其工作人员有下列行为之一的，由公共数据主管部门责令限期整改；拒不整改或情节严重的，由公共数据主管部门提请有关部门对直接负责的主管人员和其他直接责任人员按照相关规定处理。（一）不按规定采集、归集、共享、开放公共数据的；（二）提供不真实、不准确、不完整公共数据的；（三）要求申请人重复提交能够通过大数据中心获取的公共数据的；（四）不按规定编制、更新、汇总、上报目录的；（五）不按规定受理、答复、复核公共数据共享或者开放申请的；（六）不按规定擅自建设政务云平台、业务专网、灾备中心、目录管理系统、共享平台和开放平台的；（七）不按规定开展政务信息系统整合的；（八）不按规定在灾备中心进行备份的；（九）未依法履行公共数据安全职责的；（十）违规使用公共数据并造成损失的；（十一）违反本规定的其他行为”<sup>14</sup>。

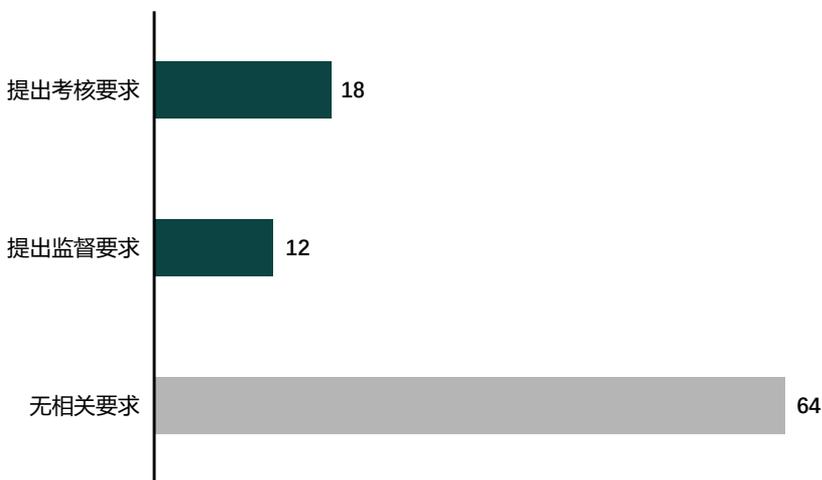


图 4-18 各地考核监督要求情况分布

<sup>13</sup> 《贵阳市政府数据共享开放条例》[EB/OL]. [2018-01-12]. [http://www.gysrd.gov.cn/News\\_show.aspx?xid=146&lmid=147&nid=7082](http://www.gysrd.gov.cn/News_show.aspx?xid=146&lmid=147&nid=7082)

<sup>14</sup> 《成都市公共数据管理应用规定》[EB/OL]. [2018-07-01]. <http://gk.chengdu.gov.cn/govInfoPub/detail.action?id=99573&tn=6>

## 4.2 组织与执行

组织与执行是指开放数据的组织保障与执行推进，包括部门职责、领导讲话支持以及公布年度工作计划三个方面。

### 4.2.1 部门职责

部门职责包括政府数据开放主管部门及相关部门的职责分工。

#### (1) 主管部门层级

数据开放主管部门是工作的执行者与推动者，是保障数据开放工作的组织基础。数据开放工作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对其推动政府数据开放工作的力度具有重要影响。随着机构改革的推进，各地政府纷纷设立了数据管理部门。如表 4-2 所示，报告将当前各地开放数据主管部门的相对层级分为以下几种类型，各地情况分布如图 4-19 所示。

表 4-2 各地数据开放主管部门情况分布

主管部门类型	举 例
一级委办局	如贵阳市大数据发展管理局，作为市政府工作部门
政府办公厅（室）二级局	如广东省政务服务数据管理局，作为省政府办公厅的部门管理机构
政府办公厅（室）挂牌局	如扬州市大数据管理局，在市政府办公室加挂市大数据管理局牌子
政府办公厅（室）内设处室	如乌海市人民政府信息化办公室
党委下设事业单位	如天津市大数据管理中心，由中共天津市委网信办管理
政府办公厅（室）下设事业单位	如佳木斯市信息中心，作为政府办公室所属事业单位
委办局二级局	如成都市大数据管理局，作为经信局二级局
委办局挂牌局	如北京市大数据管理局，在北京市经信局加挂市大数据管理局牌子
委办局内设处室	如上海市经信委信息化推进处
委办局下属事业单位	如江西省信息中心，是省发改委管理的全额拨款事业单位

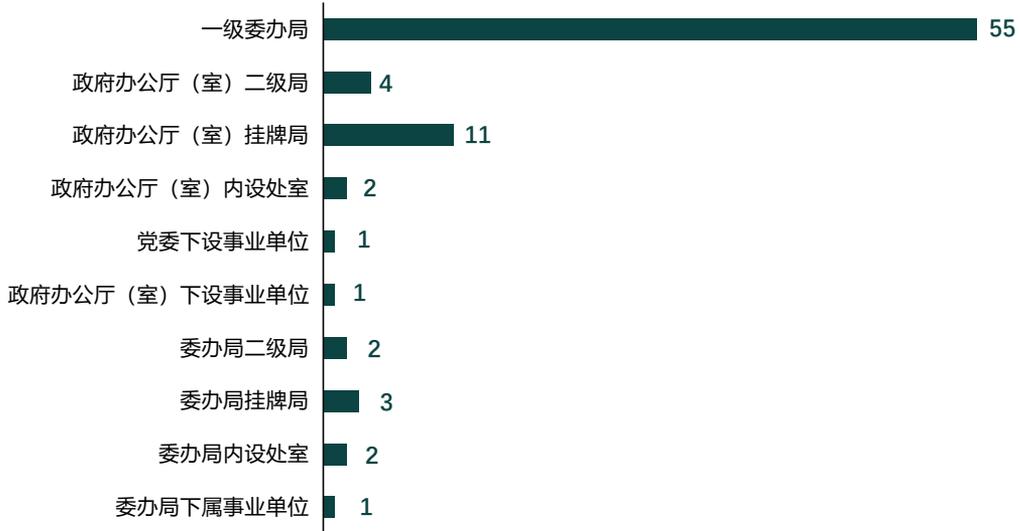


图 4-19 各地数据开放主管部门类型分布

**(2) 相关部门职责**

明确和理顺相关部门的职责，有利于推进部门间的分工协作。各地相关规定的分布情况如图 4-20 所示。目前有 14 个地方在法规和政策中对与数据开放相关部门的职责作出具体明确的规定。如安徽省宣城市规定“市政府办负责政府数据开放的统筹协调、督办检查工作。市信息办负责制定数据开放目录编制指南，制定数据开放相关标准、方法及流程，全市数据开放平台建设，负责市政府数据开放平台日常的运行管理。市财政局负责政府数据开放的资金保障工作。市政府法制办负责政府数据开放工作的法律把关工作。市政府各部门负责编制本部门开放目录、梳理数据并向社会开放数据”<sup>15</sup>。

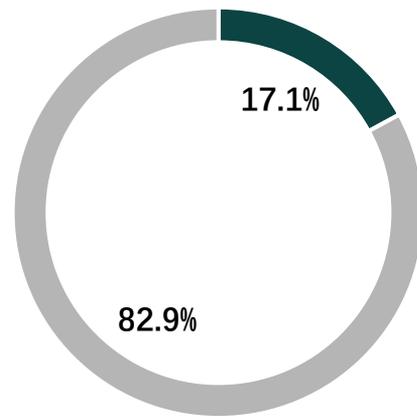


图 4-20 各地相关部门职责要求情况分布

<sup>15</sup> 《宣城市人民政府办公室关于宣城市公共信息资源开放工作实施意见》[EB/OL]. [2018-09-07]. <http://www.xuancheng.gov.cn/openness/detail/content/5bc92f8d20f7fef33a426d6e.html>

### 4.2.2 领导讲话支持

领导讲话支持是指地方党政高层领导在公开讲话中对政府数据开放工作作出明确要求和支 持。报告对 2017 年 1 月 1 日 以来媒体公开报道的领导讲话进行了分析，包括领导的层级和讲话内容。

#### (1) 领导层级

政府数据开放的顺利推进离不开强有力的领导支持，地方高层领导的支持更是作用巨大。报告将地方高层领导限定于地 方党委和政府领导，包括党委正副书记、省长、市长等。目前有 16 个地方党委和政府领导在公开场合讲话要求或支持数据开放， 详情如图 4-21 所示。

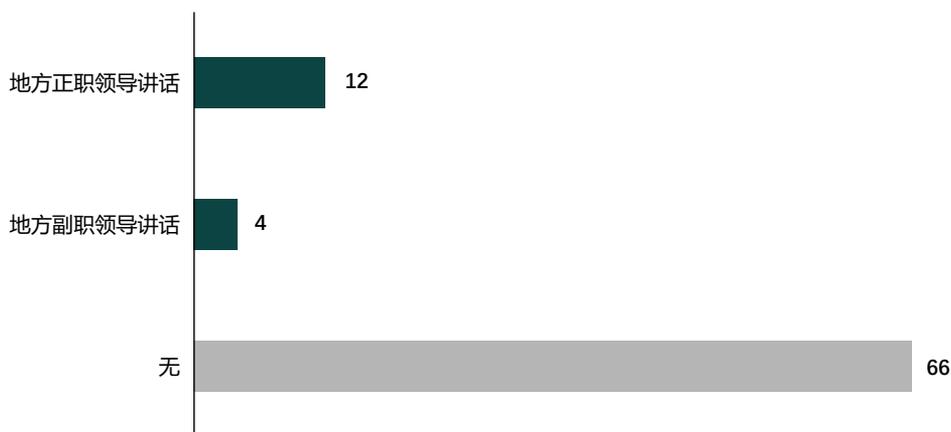


图 4-21 地方高层领导讲话层级分布

#### (2) 讲话内容

各地高层领导的意愿与态度直接影响着数据开放工作的开展与成效，讲话内容分析如图 4-22 所示。有些讲话内容丰富 全面，明确提及数据开放的经验、问题、意义与要求等。例如：上海市委副书记、市长应勇于 2019 年 3 月 7 日在人民大会 堂答记者问时提到上海将根据国家有关法律法规，制定公共数据开放管理规定，实现公共数据进一步开放。在信息和数据时代， 数据是重要的生产资料和资源，公共数据整合开放，既是集中管理、高效便捷应用的需要，也是用好公共数据、发展数字经 济的需要。有的讲话只对当地数据开放的某一方面提出了具体要求，还有些讲话虽然提到了数据开放，但内容较为笼统简略， 例如仅提到要在数据开放上取得新突破等。

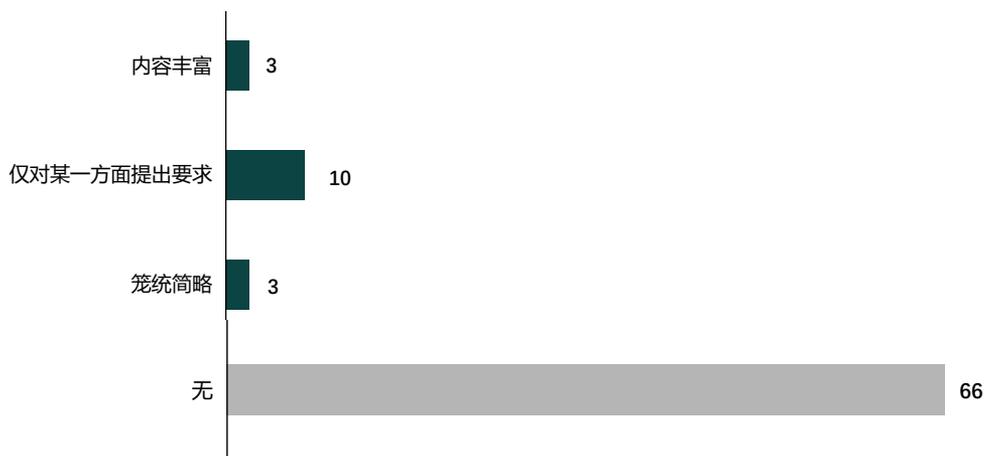


图 4-22 各地高层领导讲话内容丰富程度分析

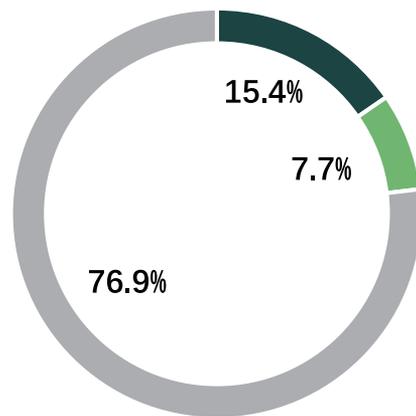
### 4.2.3 公布年度工作计划

制定并向社会公开数据开放年度工作计划既有助于政府部门自身明确阶段性目标与任务，也能给数据利用者明确预期，并接受社会监督。报告发现，目前只有上海市制定并公开了专门针对数据开放的年度工作计划。自 2014 年起，上海连续每年制定和公布《政府数据资源向社会开放年度工作计划及重点领域》，明确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。

### 4.3 标准规范

标准规范是为了在一定的范围内获得最佳秩序，经协商一致制定并由公认机构批准，共同使用的和重复使用的一种规范性文件<sup>16</sup>。制定开放数据标准规范和操作指南，有利于数据开放工作的标准化和规范化实施。

个别省级政府已正式发布并实施了一系列政府数据开放地方标准，整体情况分布如图 4-23 所示。其中山东省的标准体系最为全面，整体分为三个部分，第一部分《基本要求》规定了公共数据开放的原则、总体要求、目录要求、数据要求、安全要求、管理要求和监督评价要求等；第二部分《数据脱敏指南》提供了公共数据开放中数据脱敏的指导和建议，并给出了基本原则、脱敏规划、脱敏流程等方面需考虑的要点信息；第三部分《开放评价指标体系》规定了公共数据开放评价指标的设计原则、指标框架、指标内容以及指标使用要求。



■ 专门针对数据开放 ■ 非专门针对数据开放 ■ 无标准规范

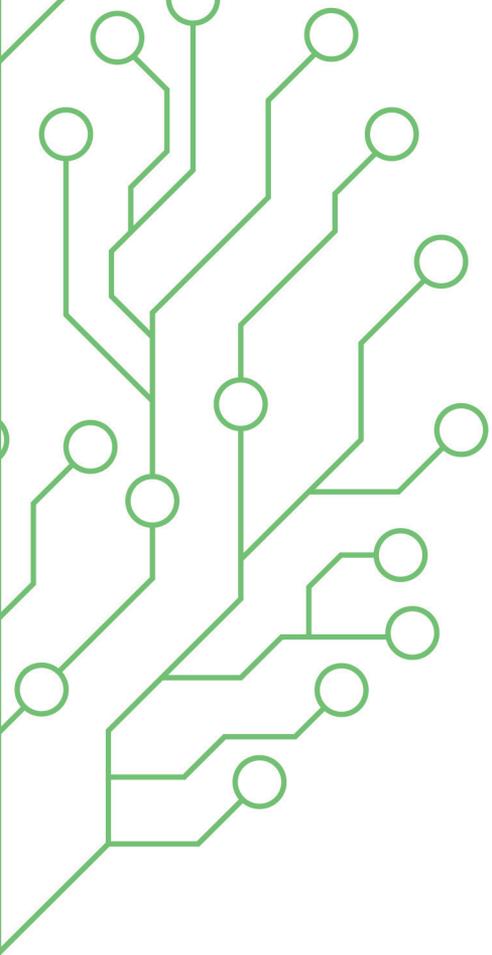
图 4-23 省级标准规范制定情况分布

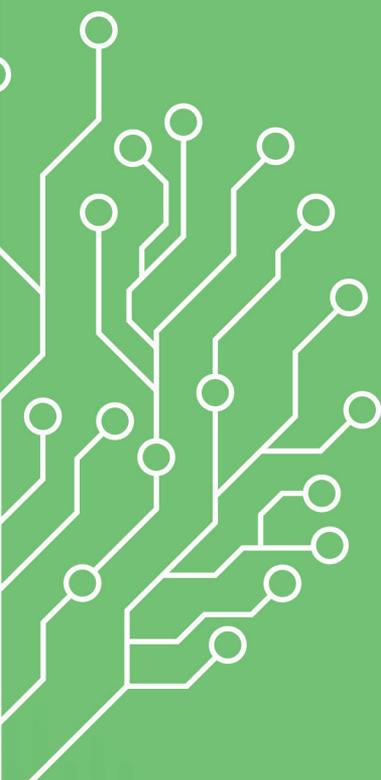
另外，广东省《电子政务数据资源开放数据技术规范》规范了电子政务数据资源开放数据的分类组织方式、元数据、数据格式、版权声明、数据使用策略、数据更新及数据质量要求等，规范了政务数据资源开放数据管理的角色与职责、管理过程、政务数据资源开放内容、数据开放各环节的管理要求等内容。还有一些标准虽然不是专门针对数据开放而制定，但涉及到与数据开放相关的内容，如贵州省的《政府数据 数据分类分级指南》，用于指导政府部门在开放和共享本部门政府数据时，对本部门政府数据进行正确分类，并对分类后的政府数据定级提供参考，以实现政府对政府数据的有效管理，为政府不同类型数据开放和共享策略的制定提供支撑。

<sup>16</sup> ISO/IEC 最新的标准化定义

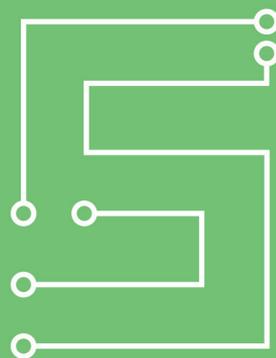
王平 . 国内外标准化理论研究及对比分析报告 [J]. 中国标准化 ,2012(5):39-50.

# PLATFORM





CHAPTER



平台层发现



## 五、平台层

报告从以下七个维度对平台体验进行了评测，主要发现如下：

**在数据发现方面**，各地平台展现的数据概况和动态内容参差不齐，展现了APP总数和API总数的地方平台低于三成，而展现了最新APP和最新API的地方平台只占一成左右。在展现方式上，能通过丰富图表和数据图谱方式展现数据概况和动态的地方平台只占两成左右。在数据导引方式上，超过九成的地方平台提供了按主题和部门进行分类导航的功能和按数据集名称进行搜索的功能，超过七成的地方平台提供了根据数据集更新时间、下载量和浏览量进行排序的功能，有五成多的地方平台提供了相关数据推荐功能。然而，各地平台中能按数据量进行排序的平台只占四成，提供了字段搜索和高级搜索功能的地方平台分别只占三成和两成，提供了用户使用场景和标签分类导航功能的地方平台低于一成。

**在数据获取方面**，提供了数据集预览的地方平台达到七成，但提供了开放数据目录的地方平台不到两成。对于无条件开放数据的获取，超过六成的地方平台设置了“一刀切”的登陆获取门槛。提供了数据请求功能的地方平台虽然已超过六成，但公开了用户数据请求的平台只占2.4%，而其中真正能回复用户数据请求的地方平台更是寥寥无几。目前尚未有地方平台实际开通了申请有条件开放数据的功能。

**在工具提供方面**，提供了地理空间工具的地方平台已超过六成，提供了图表分析工具的地方平台已超过四成，但上述工具的功能与效果还有待进一步完善。

**在利用成果方面**，六成的地方平台提供了能展示和提交服务应用的功能，但能提交其他类型利用成果的地方平台只占16%，能展示其他类型利用成果的平台只占3.7%。然而，标明了利用成果的数据集来源、数据提供部门和利用者名称等信息的地方平台只占一成左右。

**在互动反馈方面**，开通了用户评价功能的地方平台已达七成，但能提供数据发布者联系方式的平台却不到一成。值得关注的是，虽然开通了意见建议和数据纠错功能的地方平台已占六成左右，但其中公开了用户意见建议的地方平台只有三成，公开了用户纠错信息的平台只占1.2%。尤其需要引起重视的是，没有一个地方平台真正能及时回复用户意见建议和数据纠错，互动功能形同虚设。

**在公众传播方面**，提供了数据开放相关资讯和分享传播功能的地方平台比例分别达到了五成和六成，但开设了社交媒体的地方平台有2.4%。

**在账户体验方面**，提供了个性化账户和密码找回功能的地方平台都超过了七成，但同时提供了自然人和法人账户注册功能的地方平台只占三成多。用户注册账户的便捷程度也各不相同，超过三成地方平台的注册过程较为繁琐且要求用户提供的个人信息较多。

---

综上所述，**报告建议**：各地平台在数据发现方面，应改善数据概况和动态的内容及展现方式，提供按用户使用场景和标签分类进行导航的方式，提供按字段名称和多种筛选条件搜索的功能。在数据获取方面，提供可下载的开放数据目录，对于无条件开放的数据实现分级登陆获取，公开并回复用户对未开放数据的请求，真正开通可申请有条件开放数据的功能。在工具提供方面，注重提高工具的多样性和实用性。在利用展示方面，增加提交和展示利用成果的类型，标明利用成果的来源信息。在互动反馈方面，提供数据发布者联系方式，公开并及时回复用户的意见建议和数据纠错信息，从而在数据提供者和用户之间搭建起真正的供需对接桥梁。在公众传播方面，加大对数据开放工作的宣传力度，探索多渠道、多平台传播。在账户体验方面，同时提供面向自然人和法人的账户注册功能，并提升账户注册的便捷程度。

## 5.1 数据发现

数据发现是指政府数据开放平台用醒目的方式展现数据，以帮助用户了解平台上开放数据概况，并引导用户进一步发现和获取目标数据集。

### 5.1.1 数据展现

数据展现是指平台上呈现的开放数据概况、动态以及展现数据的方式，以帮助用户了解数据开放的总体情况。

#### (1) 展现内容

展现内容是指平台在首页醒目位置展现了开放数据概况和数据动态。目前，各地展示的数据概况内容包括数据集总数、按主题分布的数据集情况、按部门分布的数据集情况、数据总量、APP 总数、API 总数和开放数据部门总数。图 5-1 显示了各地平台上展现的数据概况内容的分布情况。各地展示的数据动态内容包括最新数据、最新 API 和最新 APP，图 5-2 显示了各地平台上展现的数据动态内容的分布情况。

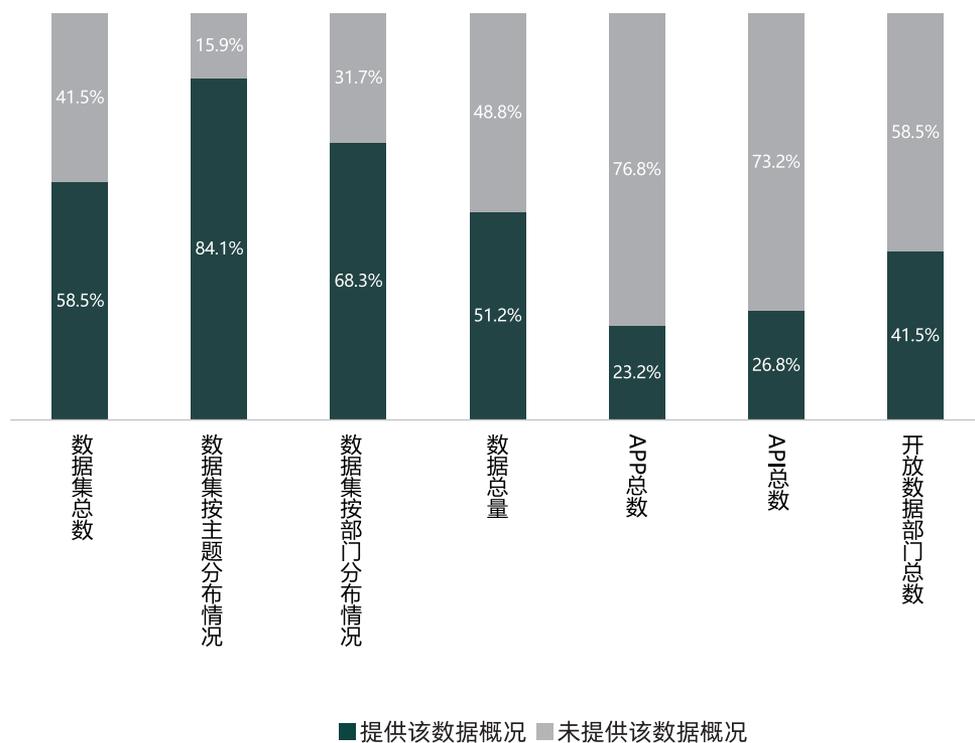


图 5-1 各地平台数据概况展现内容分布

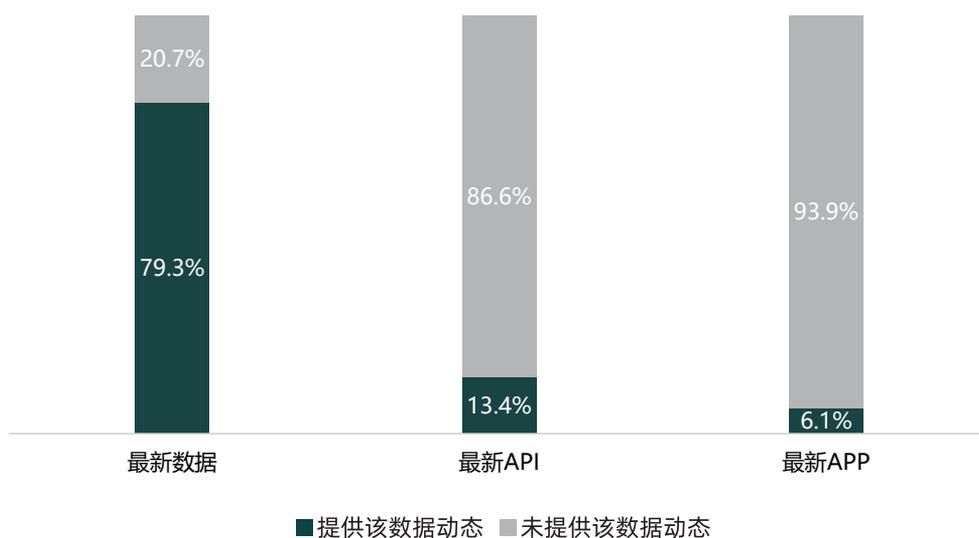


图 5-2 各地平台数据动态展现内容分布

## (2) 展现方式

展现方式是指对平台上开放数据的概况和动态以各种直观、形象的方式进行展现。目前国内政府数据开放平台主要提供了以下四种展现方式：数字罗列、简单图表、丰富图表和数据图谱。以上四种展现方式的分布情况如图 5-3 所示。在展现数据图谱的地方中，深圳和贵阳不仅展现了各数据集之间的关联性，用户还可直接通过点击链接到相应的数据集。

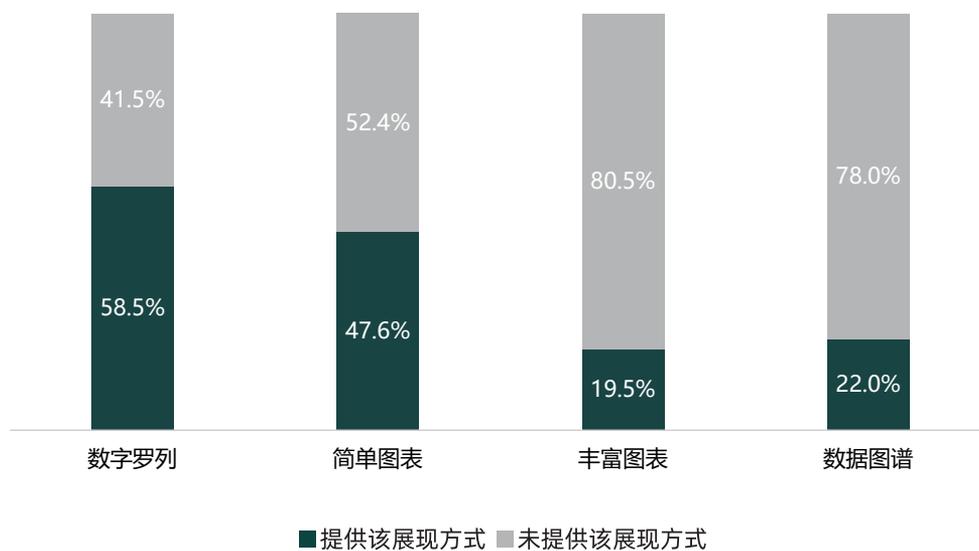


图 5-3 各地数据开放概况和动态展现方式分布

### 5.1.2 数据导引

数据导引是指政府数据开放平台提供了便捷的导引功能,以帮助用户发现所需要的目标数据,如提供分类导航、搜索、排序、相关数据推荐等功能。

#### (1) 分类导航功能

分类导航是指平台提供了按主题、部门、标签和场景等类别进行数据导引的功能。目前各地平台分类导航方式的分布情况如图 5-4 所示。

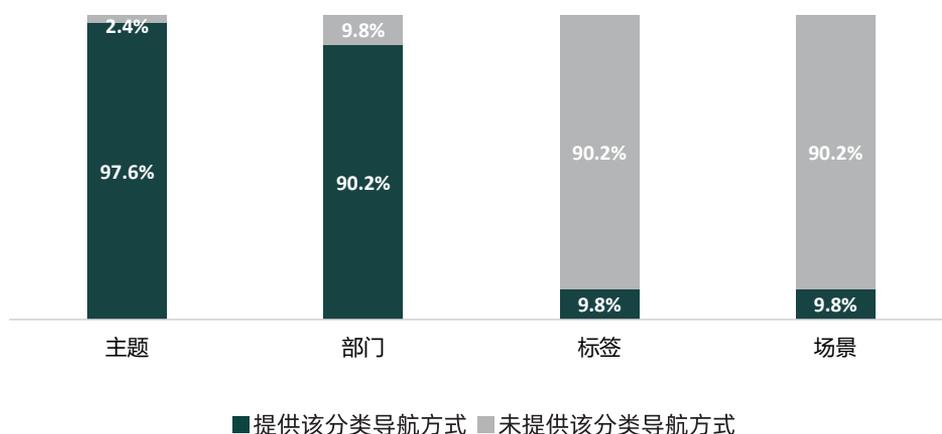


图 5-4 各地平台分类导航方式分布

有些地方平台同时提供了按照主题、行业和领域分类的导航方式,但三者间的区分较为模糊。例如:有的地方平台主题分类中有“教育科技”,领域分类中有“科技教育”,行业分类中有“教育和科技”,这种分类重复且繁杂不明确的导航方式不利于用户获取所需要的数据。另一些地方平台在分类导航中从用户视角出发,以用户的需求场景来进行分类,体验较好。如东莞市平台以用户容易理解的语言命名各种场景,如图 5-5 所示。



图 5-5 东莞市平台的场景分类

### (2) 搜索功能

搜索功能是指平台提供了按数据集名称、字段名称或其他条件搜索数据集的功能。除个别平台外，绝大多数地方平台均提供了搜索功能。目前，主要有以下3种搜索方式：一是名称搜索，即输入数据集全称或部分关键字即可搜索到相应的数据集；二是字段搜索，即除输入数据集名称外，还可通过输入数据集的部分字段名称搜索到相应的数据集；三是高级搜索，即在搜索功能中提供了多种筛选条件。各地平台的搜索方式分布情况如图5-6所示。评估中发现个别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仍无法搜索到相应的数据集。

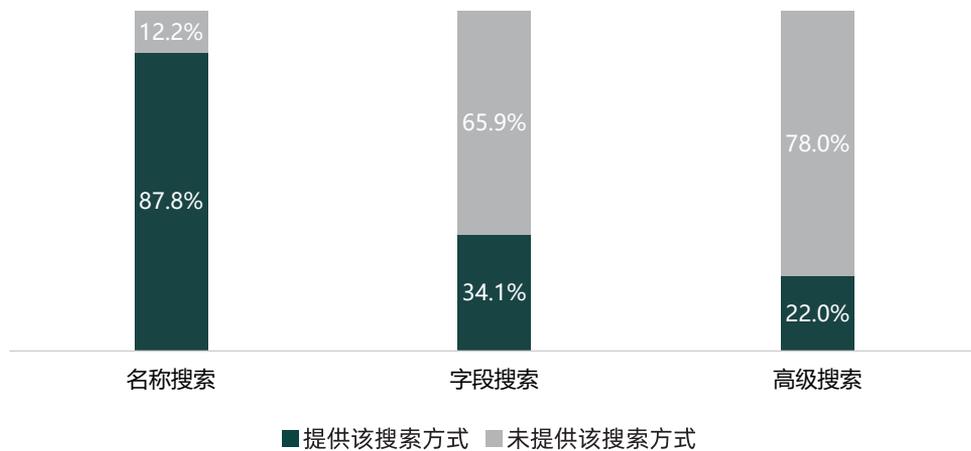


图 5-6 各地平台搜索方式分布

### (3) 排序功能

排序功能是指平台提供了根据某项具体特征或标准对数据集进行排序的功能。目前各地平台分别提供了根据资源名称、数据量、更新时间、下载量、浏览量和用户评分对数据集进行排序的功能。各地平台的排序方式分布情况如图5-7所示。

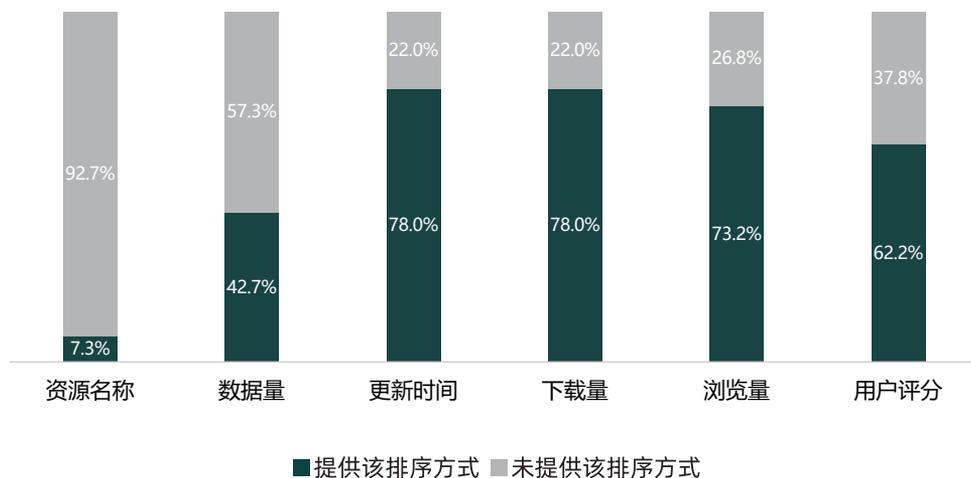
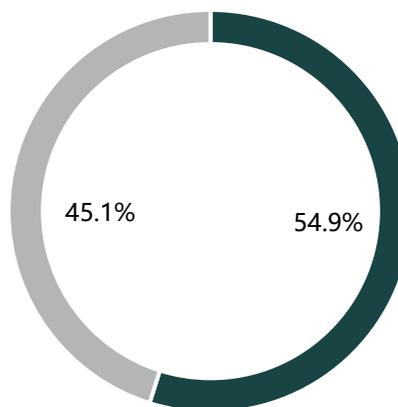


图 5-7 各地平台排序方式分布

#### (4) 相关数据推荐功能

相关数据推荐功能是指在某个数据集下推荐与之相关联的数据集，从而帮助用户更好地搜寻数据，促进数据的融合利用。图 5-8 展示了各地平台提供相关数据推荐功能的分布情况。



■ 提供相关数据推荐 ■ 未提供相关数据推荐

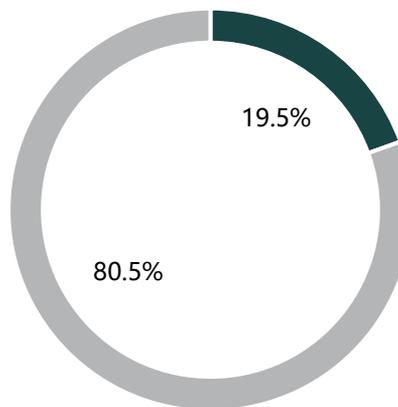
图 5-8 各地平台相关数据推荐提供情况分布

## 5.2 数据获取

数据获取是指平台提供了便于用户浏览和获取目标数据的相关功能，如开放数据目录、数据集预览、无条件开放数据获取、有条件开放数据申请以及未开放数据请求等，这些是政府数据开放平台的核心功能。

### 5.2.1 开放数据目录

开放数据目录是指平台提供了所有已开放数据资源的完整目录，帮助用户了解平台上开放数据的全貌（如图 5-9 所示）。目前贵阳、上海、浙江等 16 个地方平台提供了可下载的开放数据目录，占各地平台总数的 19.5%。

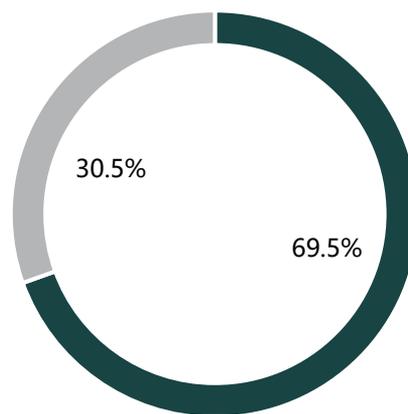


■ 提供数据目录 ■ 未提供数据目录

图 5-9 各地平台开放数据目录提供情况分布

### 5.2.2 数据集预览

数据集预览是指用户在获取数据集之前，能够对数据集进行预览以了解其详细内容。各地平台提供数据集预览功能情况分布如图 5-10 所示。目前各地提供的数据集预览方式主要有两种：一种是基本预览，允许用户简单预览部分数据；另一种是相对高级的预览方式，用户可以通过筛选条件预览相关数据。

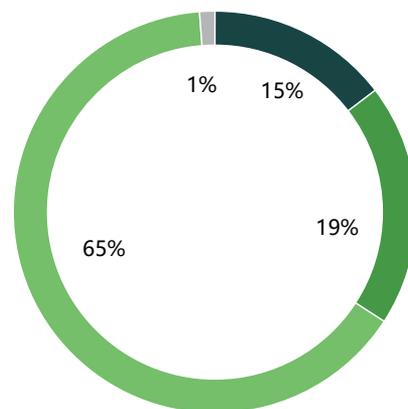


■ 提供数据集预览 ■ 未提供数据集预览

图 5-10 各地平台数据预览提供情况分布

### 5.2.3 无条件开放数据获取

无条件开放数据获取是指用户获取无条件开放类数据集的方式。目前，对于开放属性标注为无条件开放的数据集，各地平台主要提供以下 3 种获取方式：一是直接获取，即所有数据都无需用户登陆便可获取；二是登陆获取，所有的数据都要求用户登陆后方可获取；三是分级获取，即部分数据无需登陆便可获取，部分数据需要登陆才可获取。也有极个别平台存在完全无法获取数据的情况。各地无条件开放数据获取方式分布情况如图 5-11 所示。



■ 分级获取 ■ 直接获取 ■ 登陆获取 ■ 无法获取

图 5-11 各地平台无条件开放数据获取方式分布

### 5.2.4 有条件开放数据申请

有条件开放数据申请是指用户获取有条件开放类数据集的方式，即对于开放属性标注为“有条件开放”、“条件开放”或“依申请开放”等类型的数据集，用户可通过平台向数据提供方提出数据开放申请，并在数据提供方审核同意后获取。这种数据开放方式有利于在保障安全可控的前提下，将具有高价值却又有一定敏感性的数据进行开放和利用。然而，目前尚未有地方政府的数据开放平台实际开通了该功能。

### 5.2.5 未开放数据请求

未开放数据请求是指用户对于尚未开放的数据集，可通过平台提出数据开放请求。开设这一功能有助于数据提供方基于社会的数据需求，不断扩大数据开放的范围。少部分平台还将用户提出的数据请求予以公开。各地平台提供数据请求功能和公开用户请求的分布情况如图 5-12 所示。数据请求回复是指平台对用户提出的数据请求，在规定时效内进行有质量的回复，目前各地平台中仅有贵阳市平台在报告考察时段内回复了数据请求，内容详见图 5-13。

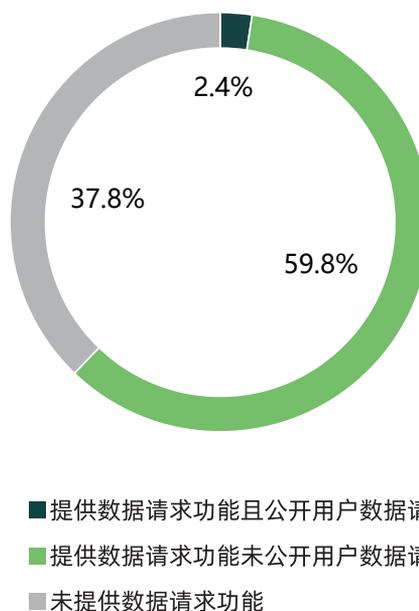


图 5-12 各地未开放数据请求功能提供和请求公开情况分布



图 5-13 贵阳市平台公开的数据请求列表及回复情况

### 5.3 工具提供

工具提供是指政府数据开放平台提供了能够帮助初级用户对数据集进行分析和利用的基本工具。目前国内政府数据开放平台提供的工具主要包括图表分析工具和地理空间工具。

#### 5.3.1 图表分析工具

图表分析工具是指平台提供了数据可视化或统计分析工具，使用户可以在数据集页面通过设置不同条件，对数据进行基本的图表分析或可视化呈现，如饼状图、柱状图和曲线图等。图 5-14 展示了各地平台提供图表分析工具的分布情况，其中提供简单图表分析工具指的是选择条件或图形种类 $\leq 2$  个的情况、提供丰富图表分析工具指的是选择条件或图表种类 $\geq 3$  个的情况。此外，也有少数平台提供的分析工具实际上未能与平台内数据联通。

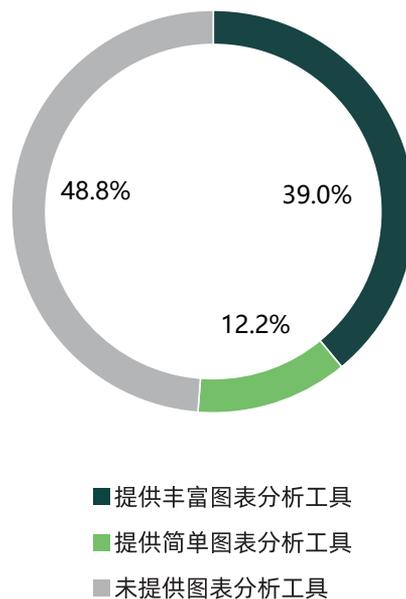


图 5-14 各地平台图表分析工具提供情况分布

#### 5.3.2 地理空间工具

地理空间工具是指平台提供了将位置类数据叠加在地图上进行展现的功能，便于用户直观具象地分析和利用数据。各地平台提供地理空间工具的分布情况如图 5-15 所示。

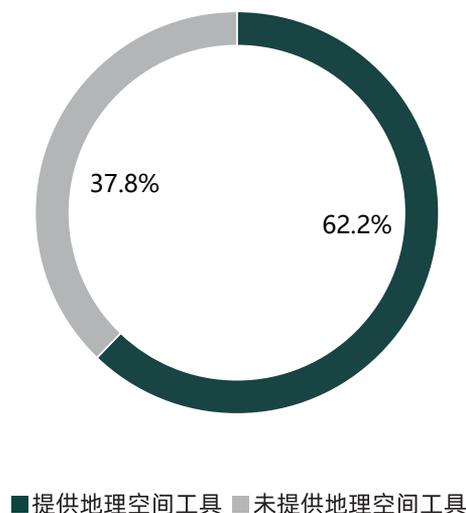


图 5-15 各地平台地理空间工具提供情况分布

### 5.4 利用展示

利用展示是指在平台上集中展现用户利用平台上开放的数据产生的各类利用成果，例如：APP 应用、研究报告、新闻传播产品等。展示利用成果既便于社会公众在平台上集中搜寻和获取成果，也便于向社会推广这些成果，进而激励数据开发者更好地利用政府开放数据，形成从数据开放、数据利用到产生价值的闭环。利用展示主要考察利用成果展示和利用成果提交两个功能。

#### 5.4.1 利用成果展示

利用成果展示主要包括利用成果类型和利用成果来源两部分。

##### ■ 利用成果类型

目前，各地平台展示的数据利用成果主要分为以下四种类型：（1）服务应用，即企事业单位利用开放数据开发的各种服务应用，如：APP 应用、小程序应用或网站应用；（2）传播产品，即将开放数据经过分析或可视化处理后，以公众易于理解的数据新闻、自媒体文章等形式向社会传播；（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产生的研究报告、学术论文等成果；（4）创新方案，即各地开放数据创新利用比赛中参赛者产生的创新方案。

现阶段各地平台上展示の利用成果类型的分布情况如图 5-16 所示。六成的地方平台展示了服务应用，贵阳和哈尔滨平台面向普通公众展示了数据传播产品，深圳在平台上开设了展示研究报告的栏目，北京和上海展示了创新方案。

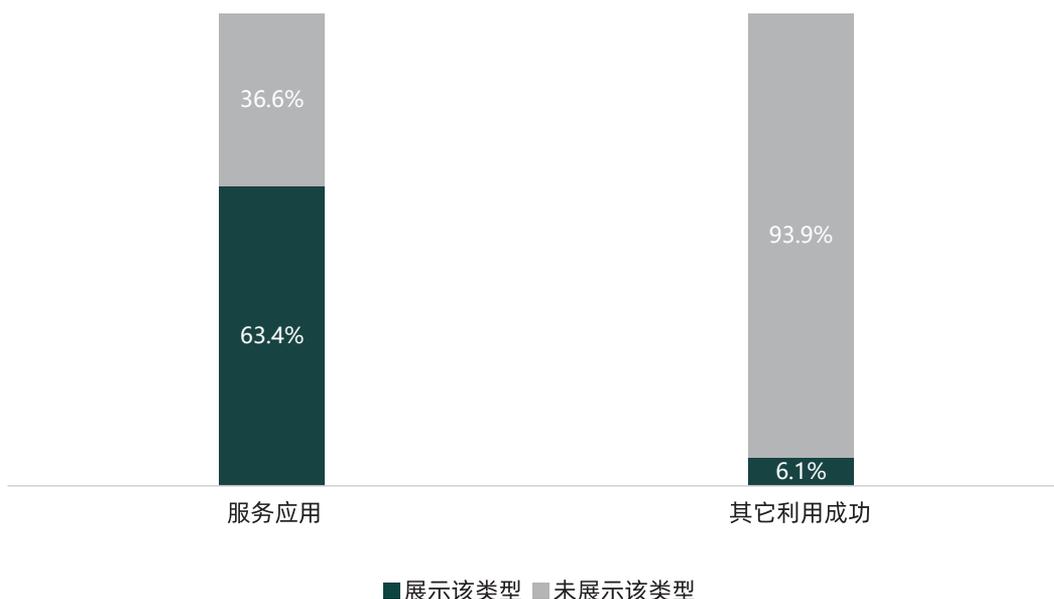


图 5-16 各地平台利用成果类型展示情况分布

### ■ 利用成果来源

利用成果来源是指平台上展示了数据利用成果的来源信息，从而在开放数据和利用成果之间建立关联。目前各地平台上展现的利用成果来源信息主要包括以下三类：（1）数据集来源；（2）数据提供部门；（3）利用者名称。图 5-17 显示了各地平台展示利用成果来源信息的分布情况。

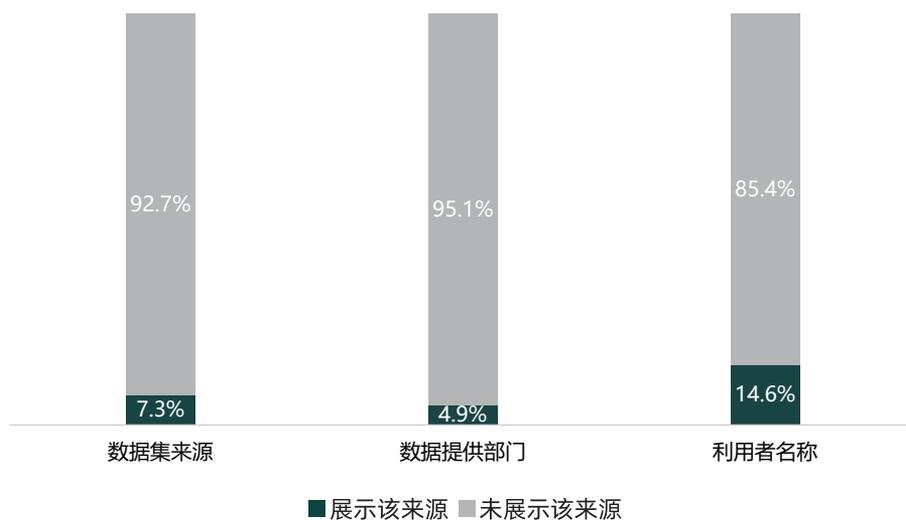


图 5-17 各地平台利用成果来源信息展示情况分布

### 5.4.2 利用成果提交

利用成果提交是指政府数据开放平台为利用者开设了提交数据利用成果的功能，从而既便于利用者主动上传其利用政府开放数据所产生的各类数据利用成果，也有助于政府追踪其开放数据的利用情况。目前各地平台上提供利用成果功能的分布情况如图 5-18 所示。图 5-19 展示了深圳市平台提供的多种利用成果提交功能。

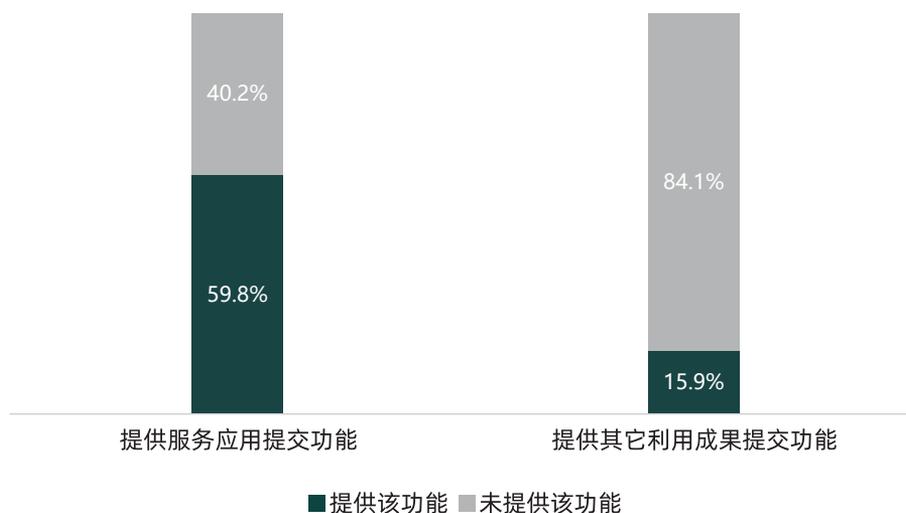


图 5-18 各地平台利用成果提交功能提供情况分布



图 5-19 深圳市平台上的“应用成果提交”功能

### 5.5 互动反馈

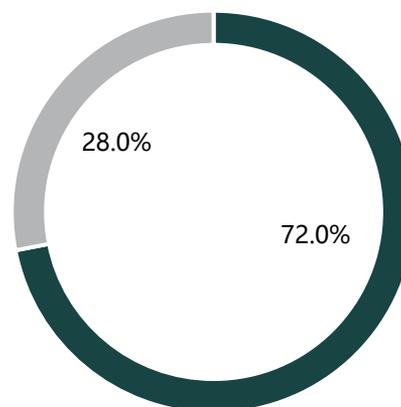
互动反馈是指政府数据开放平台提供了便于用户与平台或数据提供方进行沟通反馈的功能，包括获得客服方式、数据集评价、意见建议、数据纠错等。

#### 5.5.1 客服方式

目前的平台客服方式主要有人工客服和智能客服两种。

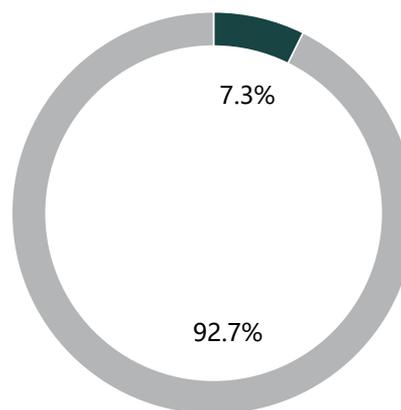
##### ■人工客服

人工客服指用户能够通过平台上提供的联系方式与平台方或数据发布方联系，以获得相关信息和服务。平台提供的联系方式包括电话、邮箱、地址等。各地提供平台方和数据发布方联系方式的分布情况如图 5-20 和图 5-21 所示。目前，仅有浙江平台同时提供了数据发布者的具体地址、联系电话或邮箱，如图 5-22 所示。



■提供平台联系方式 ■未提供平台联系方式

图 5-20 各地平台方联系方式提供情况分布



■提供联系方式 ■未提供联系方式

图 5-21 各地平台数据发布者联系方式提供情况分布

浙江·数据开放		注册	服务条款	快速登录
首页	开放数据	接口服务	应用超市	地图服务
开发者中心	开放指数	互动交流	政策动态	输入搜索内容
				高级检索
首页 > 数据集 > 医保医疗服务、诊疗项目查询信息				
医保医疗服务、诊疗项目查询信息 ☆☆☆☆☆				
★ 收藏    📡 订阅    🗑 纠错				
摘要:	查询医保医疗服务、诊疗项目的名称、项目编码、限定支付条件等信息			
资源代码:	30701600003029/299			
标 签:	查询、信息、医疗、项目、医保、服务、诊疗	更新周期:	不定期	
资源格式:	电子表格-xlsx	资源提供方:	省人社保厅	
提供方地址:	杭州市省府路8号省府大院2号楼	联系方式:	0571-87057390	
数据领域:	社保就业	访问/下载次数:	834/76	
系统名称:	无	评分/评价次数:	0/0	
行业分类:	暂无	数据预览:	<input type="button" value="预览"/> <input type="button" value="可械化"/>	
是否向社会开放:	是	数据接口:	<input type="button" value="API"/>	

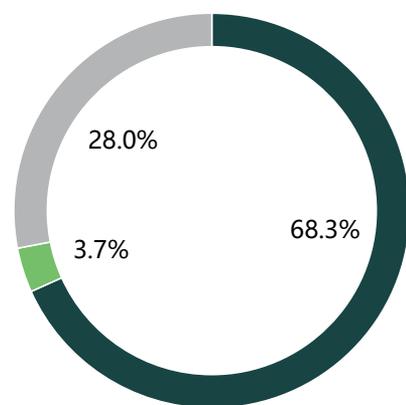
图 5-22 浙江省平台提供的数据发布者的联系方式

### ■ 智能客服

智能客服是指平台提供了人机智能交互的方式来为用户提供个性化的指引和服务，从而在保障服务体验的前提下节省客服的人力投入。目前仅有贵阳市和佳木斯市平台提供了智能化的服务方式。

### 5.5.2 数据集评价

数据集评价是指用户可对单个数据集进行星级评分和文字评价的功能。目前大多数平台允许用户登录后对单个数据集进行星级评价和文字评价，并直接公开用户的星级评价结果，文字评价则须经平台审核后才能公开。此外，部分平台还允许已公开的文字评价被其他用户评论和点赞，加强了平台与用户、用户与用户之间的互动。图 5-23 展示了各地平台提供数据集评价功能和公开评价内容的分布情况。



- 提供评价功能且公开用户评价
- 提供评价功能未公开用户评价
- 未提供评价功能

图 5-23 各地平台数据集评价功能提供和评价公开情况分布

### 5.5.3 意见建议

意见建议功能是指平台通过问卷调查和提交文字内容等方式向用户采集意见或建议，以获得用户反馈。各地平台提供意见建议功能和公开意见建议的分布情况如图 5-24 所示。在评估时段内，未有任何平台回复评估组提出的用于测试的意见建议。

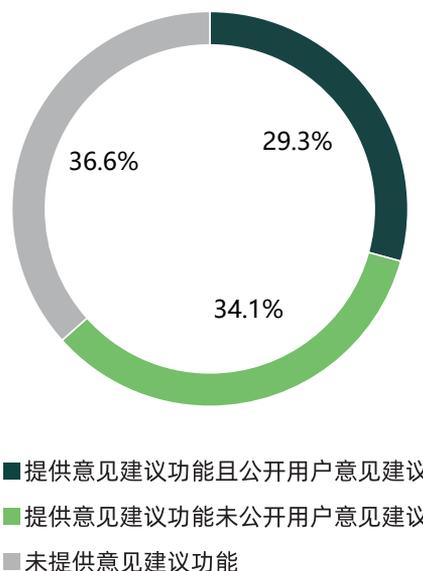


图 5-24 各地平台意见建议功能提供和意见建议公开情况分布

### 5.5.4 数据纠错

数据纠错功能是指用户可向平台反馈其发现平台中数据集存在的问题或错误，有助于数据提供方不断提高开放数据的质量。各地平台提供数据纠错功能和公开数据纠错信息的分布情况如图 5-25 所示。目前，仅有贵阳市平台公开了数据纠错信息和回复。但在评估时段内，未有任何平台回应评估组提出的用于测试的数据纠错请求。

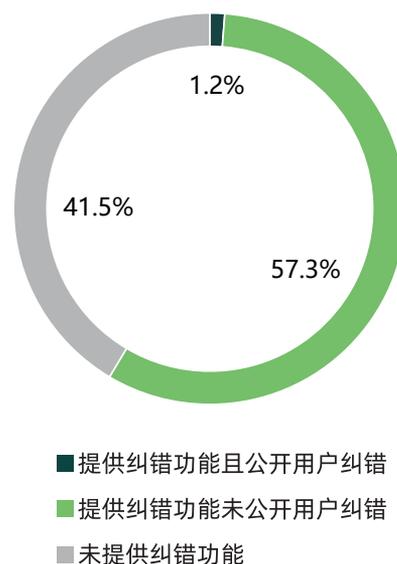


图 5-25 各地平台数据纠错功能提供和纠错信息公开情况分布

## 5.6 公众传播

公众传播是指政府通过开放平台或社交媒体等渠道向社会进行传播，帮助公众了解有关数据开放的基本知识、最新政策、开放动态、利用成果、利用活动等资讯。目前各地运用各种公众传播方式的分布情况如图 5-26 所示，主要包括资讯提供、分享功能和社交媒体等。

### 5.6.1 资讯提供

资讯提供是指通过开放平台提供与数据开放相关的资讯，有利于用户及时了解数据开放相关的知识、政策与动态，提高对数据开放的知晓和认知程度。

### 5.6.2 分享功能

分享功能是指在平台上提供信息分享传播功能，使用户可将平台上的信息在微博、微信和 QQ 等社交媒体平台进行转发传播，从而提升政府数据开放平台的热度和影响力，吸引社会公众访问平台和利用数据。

### 5.6.3 社交媒体

社交媒体是指开设专门的社交媒体账号（如微博、微信公众号等），有助于向社会公众宣传推介本地的数据开放工作，并与社会公众通过社交媒体开展互动。目前，仅有贵阳市和济南市开设了微信公众号（“贵阳政府数据开放”和“开放济南”），推送与政府数据开放相关的资讯。

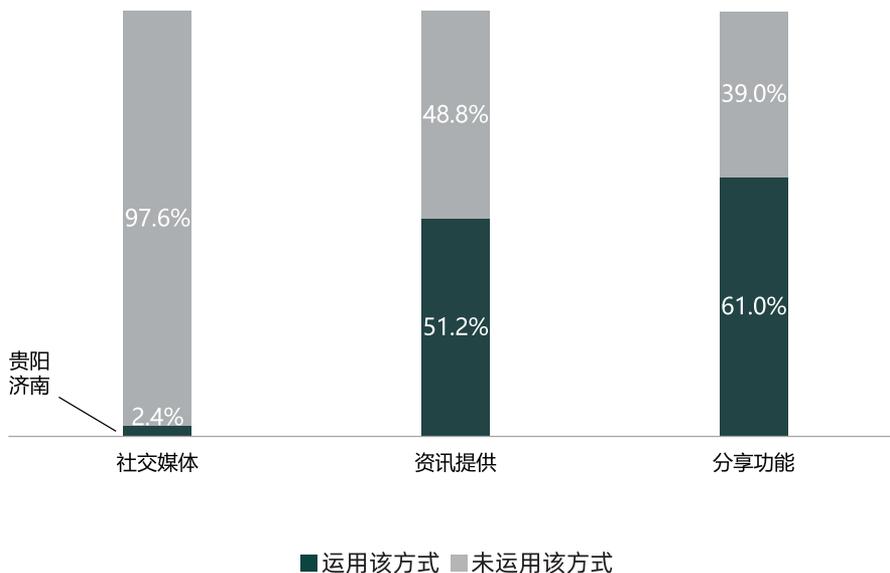


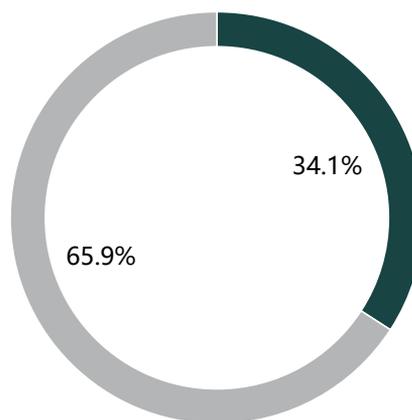
图 5-26 各地平台公众传播方式情况分布

## 5.7 账户体验

账户体验是指平台为用户在开设和使用账户时提供便捷的个性化服务。账户体验包括账户分类、个性化账户、账户注册便捷程度、密码找回等。

### 5.7.1 账户分类

账户分类是指平台不仅为自然人，还为法人和非法人组织提供了注册功能。各地平台账户分类的分布情况如图 5-27 所示。

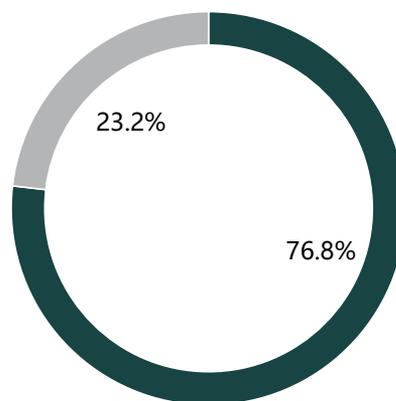


■ 提供账户分类 ■ 未提供账户分类

图 5-27 各地平台账户分类提供情况分布

### 5.7.2 个性化账户

个性化账户是指平台为用户提供个性化的账户功能，使用户可以根据其需求来使用数据发现、数据获取、工具使用、成果展示、互动反馈等功能。各地提供个性化账户的分布情况如图 5-28 所示。



■ 提供个性化账户 ■ 未提供个性化账户

图 5-28 各地个性化账户提供情况分布

### 5.7.3 账户注册便捷程度

用户注册账户的便捷程度会影响用户使用平台的意愿和体验。目前各地平台要求的注册方式主要有 5 种，按便捷程度由低到高分别为：（1）无须提交任何信息即可完成注册；（2）可通过第三方（如微信、支付宝、微博等）完成注册；（3）必须提供电子邮箱地址才能完成注册；（4）必须提供手机号才能完成注册；（5）必须提供身份证号才能完成注册。此外，还有极个别平台无法完成注册。各地平台要求的注册方式的分布情况如图 5-29 所示。

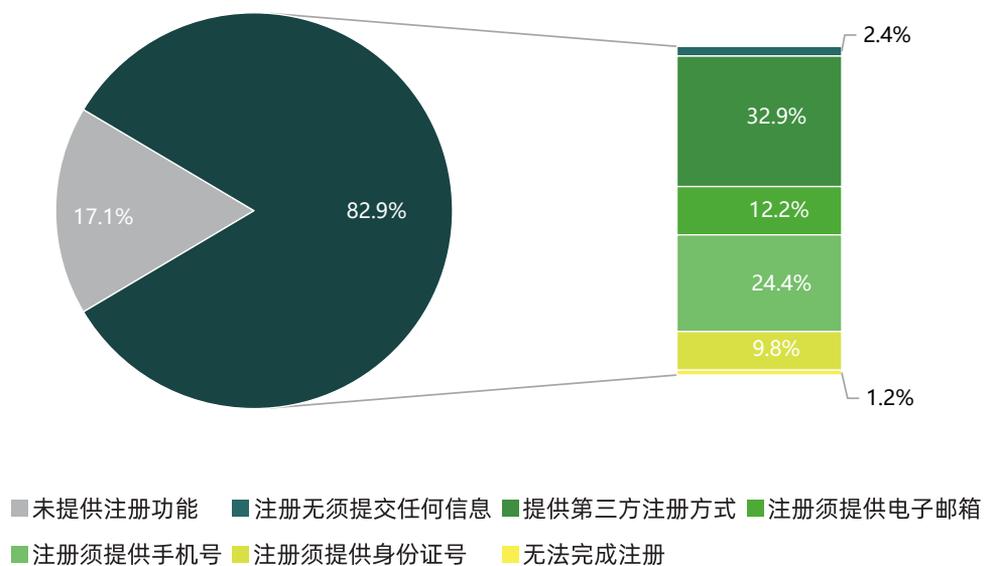


图 5-29 各地平台注册要求情况分布

### 5.7.4 密码找回

密码找回是指用户可以通过注册时使用的邮箱、手机号或身份信息修改或找回账户密码。密码找回功能可方便用户维护自己的个人账户。各地平台提供密码找回功能的分布情况如图 5-30 所示。

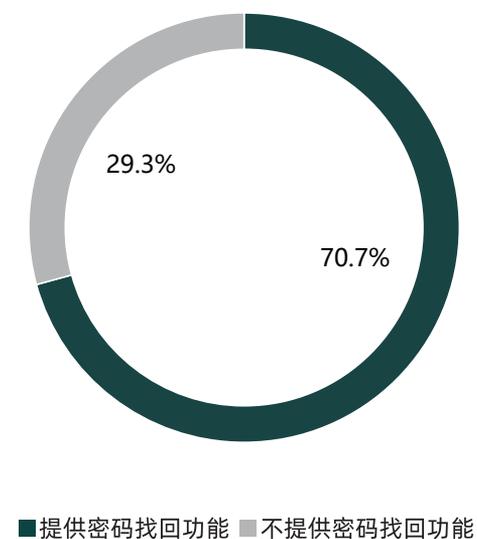


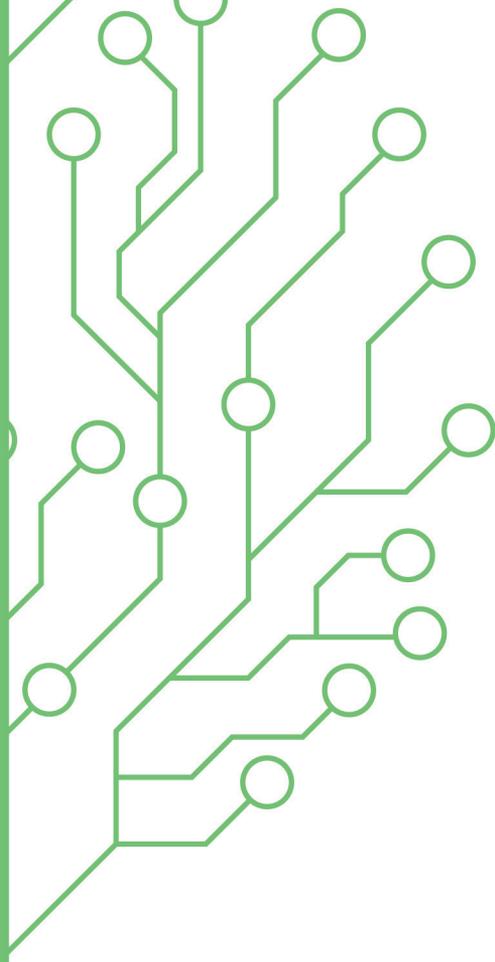
图 5-30 各地平台密码找回功能提供情况分布

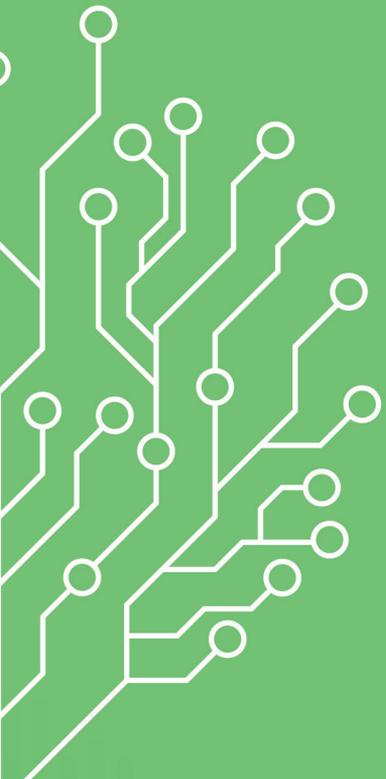


中国开放数据指数  
CHINA OPEN DATA INDEX

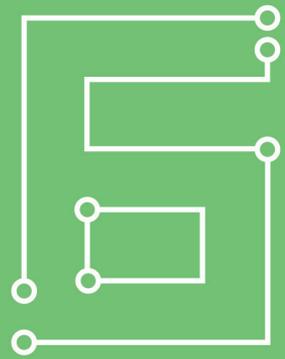
· 2019 ·

DATA





CHAPTER



数据层发现

## 六、数据层

报告从以下五个维度对数据层进行了评测，发现如下：

**在数据数量方面**，目前各地平台开放的数据集总量已成一定规模，但各地之间差异明显，全国约三成的地方平台上开放的数据集总量已超过了一千个，但也有约四成的地方平台上开放的数据集总量不到两百个。在开放数据容量方面，各地平台差异更为显著，约两成的地方平台上开放的数据容量已超过 1 亿，而超过三成的地方平台上开放的数据容量仅在 10 万及以下。

**在数据质量方面**，虽然优质数据集在容量上与去年同期相比有显著提升，但全国仅有约四成的地方平台开放了优质数据集。不到七成的地方平台上存在碎片化和低容量数据集。虽然各地平台上重复创建、生硬格式转化和无效数据集的情况数量有所下降，但目前国内依旧只有三成的地方平台不存在以上问题。各地平台上提供的 API 接口质量普遍堪忧，优质 API 接口屈指可数，近五成地方平台上存在碎片化 API 接口，几乎所有地方平台上都存在不必要的限制型 API 接口，还有八成的地方平台上存在无效 API 接口，近半数地方平台上存在重复 API 接口。能够调取数据的 API 接口大都存在数据容量小和更新频率低的问题。

**在数据标准方面**，只有 16% 的地方平台在提供的授权协议中明确授予用户免费获取、不受歧视、自由利用、自由传播和分享数据的权利，近四成的地方平台缺少专门的数据开放授权协议。各地平台上约八成的数据集已符合可机读格式的标准，超过六成的数据集已符合非专属格式的标准，但全国仅有浙江、深圳和贵阳三个地方平台提供了符合 RDF 格式的数据。API 接口调用难度较低的平台和为数据集标注了多种开放类型的平台都只占一成。超四成的地方平台已提供所有的基本元数据，近六成的地方平台能为 API 接口提供基本的调用说明，但只有不到一成的地方平台提供了详细的操作指南。

**在数据覆盖面方面**，不到一成的地方平台开放的数据集覆盖了所有重要主题，而近两成的地方平台开放的数据集覆盖的主题不到一半；近两成的地方实现了所有政府部门都开放了数据集，而超过三成的地方参与开放数据的部门覆盖面不到一半。只有 16% 的地方平台上开放了所有的常见数据集，而近五成的地方平台上只开放了不到一半的常见数据集。

**在数据持续性方面**，各地平台对数据持续运营的重视程度逐年提高，但仍然只有不到一成的地方平台能在近两年来的每个季度都持续上线新增数据集，而超过七成的地方平台曾出现数据集增长中断或停滞的情况。能实现数据动态更新的地方平台仍然偏少，在 2019 年第一季度内，近五成的地方平台没有更新数据集。约四成的地方平台将多个批次的历史数据留存在平台上供用户获取。

---

综上所述，**报告建议**：在数据数量方面，各地应不断提升数据集的容量。在数据质量方面，推出更多高容量、高需求的优质数据集，同时减少碎片化和低容量的低质数据集，剔除重复创建、生硬格式转化和无效数据集；增加以 API 接口的形式开放的容量大、更新频率高的数据集，清理碎片化 API 接口、限制型 API 接口、无效 API 接口和重复 API 接口。在数据标准方面，制定和完善开放授权协议，对用户免费获取、不受歧视获取、自由利用、自由传播与分享开放数据的权利进行明确授权，探索对不同类型的数据集分级分类制定开放授权协议；提高 RDF 格式开放数据的比例，降低申请 API 接口的难度；推进数据的分级分类开放，对数据集标示不同开放类型，如无条件开放或有条件开放等；提供更为完善的元数据和更加详实的 API 接口调用说明，帮助用户理解、获取和利用数据。在数据覆盖面方面，鼓励和吸引更多的部门参与到数据开放工作中来，丰富数据内容的多样性，不断提高开放数据集的主题覆盖面和部门覆盖面；数据开放工作刚刚起步或水平不高的地方，可以参考先进地方已经开放的常见数据集，积极稳妥地补齐短板。在数据持续性方面，保障平台的持续性管理和运维，使其成为政府部门的一项常态化工作，以支撑开放数据集的持续增长、动态更新和历史存档。

## 6.1 数据数量

### 6.1.1 数据集总量

数据集是指由数据组成的集合，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。政府数据开放平台通常以下载或 API 接口的方式提供数据集，数据集总量统计的是平台上可通过下载或 API 接口获取的数据集总数。目前有 26 个地方平台开放的数据集总量超过 1000 个。开放数据集总量（含直接下载和 API 接口开放）最多的十个地方平台如图 6-1 所示。

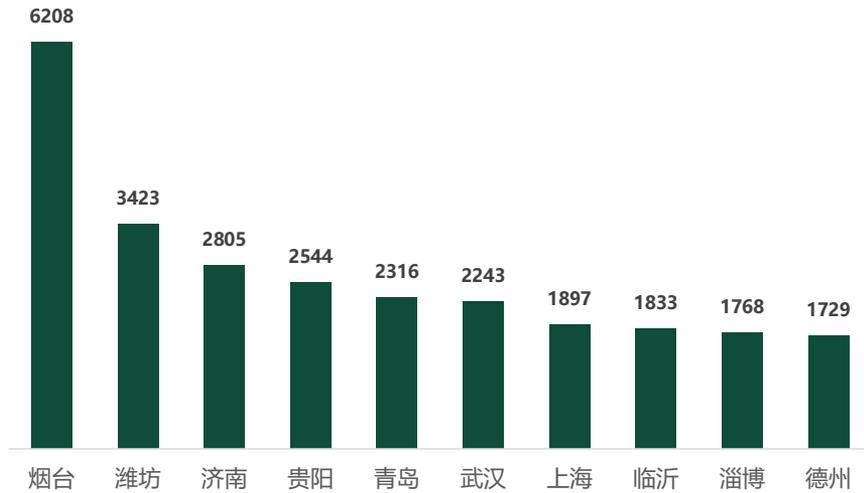


图 6-1 数据集总量（前十名）（单位：个）

### 6.1.2 数据容量

数据容量是指将一个地方平台中可下载的、结构化的、各个时间批次的数据集的字段数（列数）乘以条数（行数）后得出的数据总量。数据容量体现的是平台上开放的可下载数据集的数据量大小。数据容量排名前十的地方平台如图 6-2 所示。目前，各地平台间数据容量差异显著，约 1/5 的地方平台上开放的数据容量超过 1 亿，而近 1/3 的地方平台上开放的数据容量仍在 10 万及以下。图 6-3 展示了各地平台中单个可下载数据集平均容量排名前十的地方。

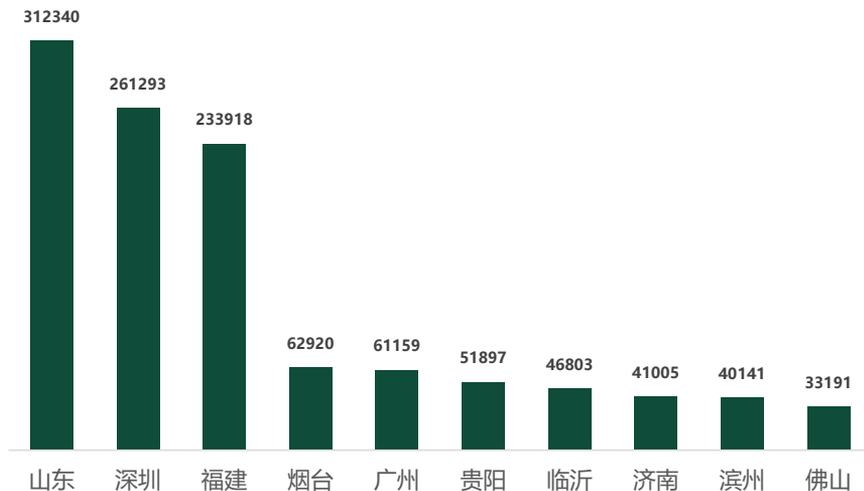


图 6-2 可下载数据集容量（前十名）（单位：万）

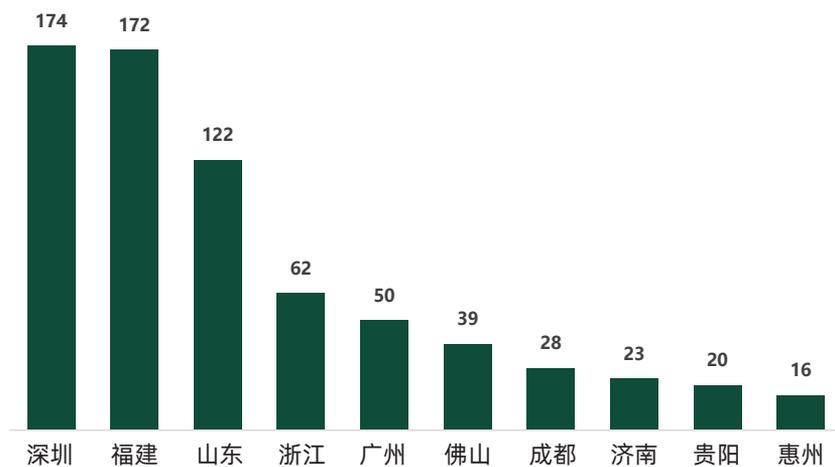


图 6-3 单个可下载数据集平均容量 (前十名) 单位: 万

## 6.2 数据质量

### 6.2.1 优质数据

#### (1) 优质数据集

数据容量大、社会需求高的优质数据集是开放数据的重点。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 (共 392 个)。绝大部分的优秀数据集集中出现在 33 个平台上, 其他超过半数的地方平台没有出现优质数据集。拥有优质数据集最多的十个地方平台如图 6-4 所示。

表 6-1 是排在前十位的优质数据集列表, 这些数据集普遍具有较高的条数、字段数和下载量, 内容主要涉及商事主体、司法、气象、食品和药品等。

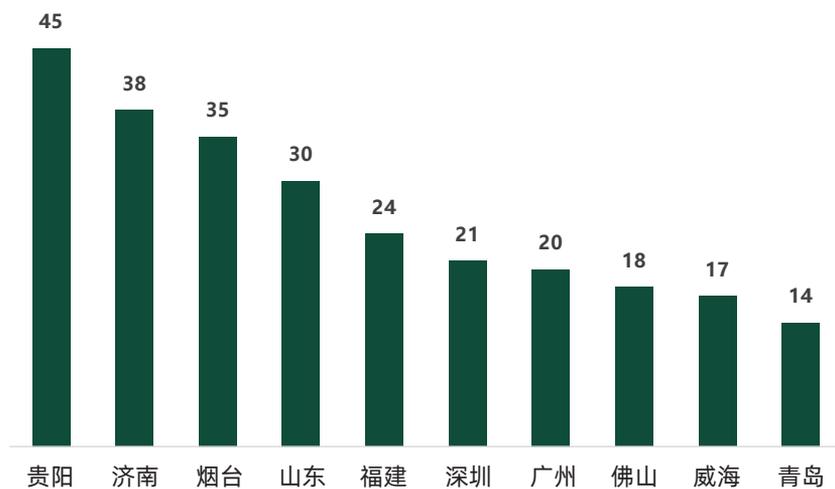


图 6-4 优质数据集数量 (前 10 名)

表 6-1 前十位优质数据集列表

序号	数据集名称	地方	数据容量	条数	字段数	下载量
1	商事主体个体年报基本信息	广州	162962240	8148112	20	74
2	双定户超定额信息	佛山	132320110	10178470	14	46
3	司法判决信息	成都	99678650	19935730	15	147
4	区域气象站观测实况	广州	68062798	4003694	17	189
5	工商登记信息	东莞	51980529	2260023	23	15651
6	威海市肉菜追溯信息	威海	42779242	2516426	17	164
7	热带气旋数据	深圳	28523550	1296525	22	316
8	企业注册登记信息	青岛	23896110	796537	30	147
9	贵阳市城镇居民医疗保险药品目录	贵阳	15335748	851986	18	1398
10	济南市企业变更信息	济南	13540910	2708182	5	424

## (2) 优质 API

优质 API 是指地方平台上提供的必要、可调取真实数据、至少每天动态更新、数据集容量较大的 API 接口。其中必要是指该数据集仅适合 API 接口获取，不适合通过下载方式获取，例如高频率动态更新的、大容量的数据。本次评估中发现各地平台上同时符合以上四点特征的优质 API 接口较少，能够调取数据的 API 接口普遍存在数据容量小、更新频率低等问题。表 6-2 为各地平台上发现的优质 API 接口列表。

表 6-2 优质 API 接口列表

API名称	地方	数据容量	调用次数
企业基础信息	浙江	62492751	333
工商登记信息	东莞	51980529	未显示
养老服务补贴给付信息	浙江	1803760	85
贵阳市空气日报	贵阳	40048	459

## 6.2.2 无低质数据

### (1) 无低质数据集

#### ■ 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、批次等被人为分割的数据集，不利于数据利用者进行融合获取和利用。目前约七成的地方平台上存在碎片化数据。

#### ■ 无低容量数据

低容量数据是指条数在三行或三行以内的数据集，存在这类数据的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 82 个开放数据平台中，近七成的地方平台上存在低容量数据。

### (2) 无低质 API

#### ■ 无碎片化 API

碎片化 API 是指通过 API 接口提供的数据集按照时间、行政区划、政府部门、批次等被人为分割，部分地方平台的 API 接口还按页面和数据量被分割，这些碎片化的 API 接口都不利于数据利用者整体性获取和利用。目前近半数的地方平台上存在碎片化 API 接口。

### ■无限制型 API

API 接口便于用户获取数据容量大、更新频率高的数据集。部分地方平台为数据集提供了 API 接口，但是这些数据集本身容量低、相对静态，对用户来说通过直接下载的方式获取这类数据可能更为方便。还有些地方平台对于数据容量较小、更新频率低的数据集仅提供了通过 API 接口获取一种方式，而未提供下载获取的方式，这类 API 接口就不仅未给用户带来方便，反而还限制了用户获取数据的方式，提高了获取数据的门槛。目前，几乎所有地方平台上都存在这种限制型 API 接口的情况。

## 6.2.3 无问题数据

### (1) 无问题数据集

#### ■无重复创建

重复创建是指平台上出现标题相同、可下载数据文件内容相同、所属主题也相同的数据集。在 82 个地方平台中，约四分之一存在重复创建问题。

#### ■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 等文件生硬地转化成 XLS、CSV、XML 等可机读格式。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 XML 格式等。目前有 9 个地方平台存在上述问题。

#### ■无无效数据集

无效数据集包括以下三类情况：数据集名称下没有可获取的数据集；只提供数据链接，却无法获取数据集；数据集内实际上不存在数据。在 82 个地方平台中，目前超过六成的地方存在无效数据集。

#### ■无标题缺失或不清

标题缺失或表达不清是指平台上开放的数据集标题不全，或是标题名称表达不清、不知所云，且未在数据简介中做出说明，使用户无法查询、理解和利用。例如：数据集标题为“公示”、“文档”，却未说明是关于什么内容的公示或文档；又如：数据集名称和简介都叫“出租车人车绑定 1 比 4”，却未提供任何解释说明。目前约五分之一的地方平台存在类似情况。

### (2) 无问题 API

#### ■无无效 API

无效 API 是指存在于平台之上，但却无法调取的 API，主要包括以下三种情况：一是提交 API 接口调用申请后未收到任何反馈；二是 API 接口调用说明不清，无法调用；三是调用申请审核通过且按照调用说明正确操作后，仍无法获取数据。目前有五分之四的地方平台上存在无效 API 接口。

#### ■无重复 API

重复 API 接口是指平台上出现标题相同或极其相似，实际调取到的数据集完全相同、所属主题也相同的 API 接口。目前有接近半数的地方平台上存在重复 API 接口。

### 6.3 数据标准

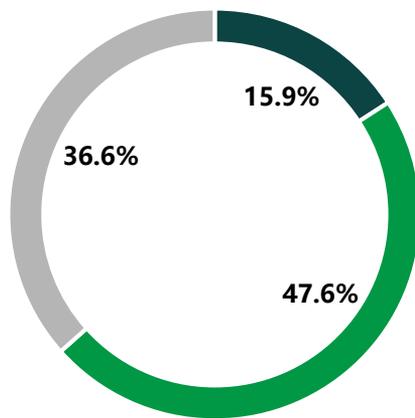
#### 6.3.1 开放授权

开放授权协议是以契约的形式确定数据开放主体和数据使用主体之间关于数据利用的权利义务关系，有利于保障和规范数据的合理使用。目前，各地平台上的开放授权协议通常包含在网站声明、免责条款或服务协议中。在 82 个地方平台中，共有 52 个制定了开放授权协议，其余 30 个缺少专门的开放授权协议。

开放授权协议应对用户免费获取、不受歧视获取、自由利用、自由传播与分享开放数据的权利做出明确授权。目前有 13 个地方平台的协议明确授予了用户上述全部四项权利；大部分地方平台授予了用户免费获取和不受歧视两项权利，其余地方平台未明确授权或语焉不详（见图 6-5）。目前还未出现对不同类型的数据集分级分类制定开放授权协议的地方平台。

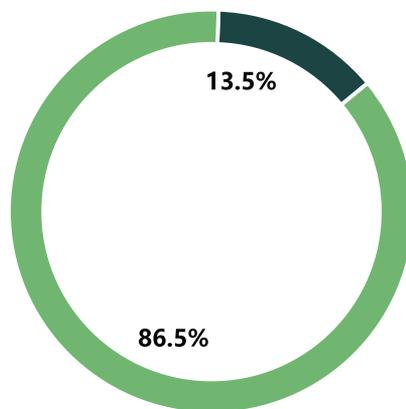
#### （1）免费获取

免费获取是指平台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或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。“免费且未设时限”授予用户永久无偿获取平台上开放数据资源的权利，其他大部分地方平台则在服务协议中采用了“现阶段免费”的表述，并设置了例如“保留收费权利”这样的条件（见图 6-6）。



■ 授予 4 项权利 ■ 授予 2 项权利 ■ 未明确授权

图 6-5 各地开放授权情况分布



■ 免费且未设时限 ■ 现阶段免费

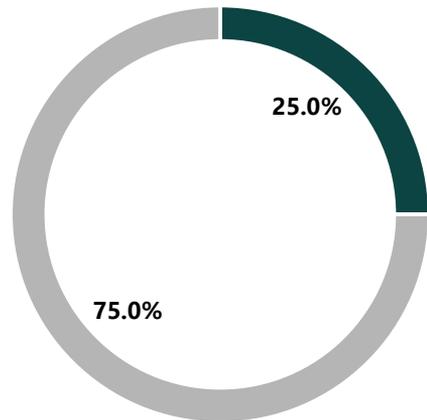
图 6-6 各地免费获取授权情况分布

### (2) 非歧视性

非歧视性是指平台授予各类用户平等访问、获取、使用和分享开放数据的权利。目前除宁夏回族自治区在注册过程中只赋予拥有本地手机号的非用户获取数据的权利外，各地平台上用户均享有平等获取和利用数据的权利，文字上，主要通过“用户享有数据资源的非排他使用权”、“不受歧视”等样式来表述这种权利。

### (3) 自由利用

自由利用是指平台授予用户对开放的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。部分地方平台授予了用户可“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”的权利，其余 39 个地方的条款未对用户利用数据的权利做出明确说明（见图 6-7）。

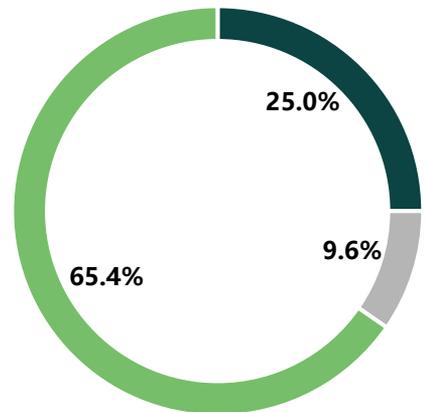


■ 明确允许自由利用 ■ 未提及可自由利用

图 6-7 各地自由利用授权情况分布

### (4) 自由传播与分享

开放授权是指平台授予用户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。部分地方平台授予了用户免费传播开放数据的权利；少数地方平台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”，尽管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（见图 6-8）。



■ 可自由传播 ■ 自由传播受限 ■ 未提及可自由传播

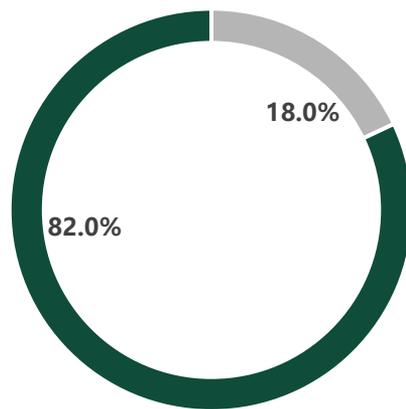
图 6-8 各地自由传播与分享授权情况分布

### 6.3.2 技术开放性

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，采用万维网联盟（W3C）的开放标准（如 RDF 和 SPARQL）来识别数据集中的事物或主体，每个事物采用统一资源标识符（URI）来进行标识；五星：在满足以上要求的基础上，数据集提供更多的与其他数据集数据之间的关联链接，以便使用者在使用数据的同时发现更多的相关信息。

#### （1）可机读格式

为方便用户获取和利用数据，开放数据集应为可机读格式，该格式能被计算机自动读取与处理，如 XLS、CSV、JSON、XML 等格式。总体上，目前各地平台上开放的可下载数据集中超过八成成为可机读格式（见图 6-9），超过半数的地方平台上开放的数据集基本满足了可机读格式的要求，但仍有部分地方平台上存在 DOC、PDF、JPG 等不可机读格式的数据集。

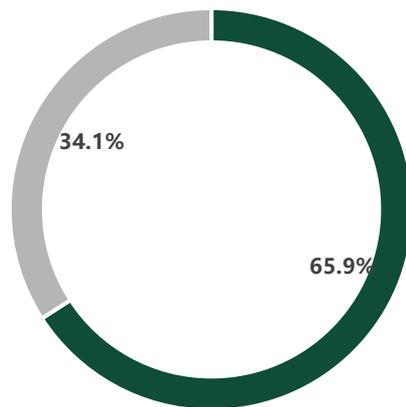


■ 可机读格式 ■ 不可机读格式

图 6-9 可机读格式分布情况

#### （2）非专属格式

非专属格式是指可下载数据集应以开放的、非专属的格式提供，任何数据提供主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 是非专属格式，XLS 则不是。目前在各地平台上开放的可下载数据集中已有约 66% 符合了非专属格式的标准（见图 6-10）。



■ 专属格式 ■ 非专属格式

图 6-10 非专属格式分布情况

### (3) RDF 格式

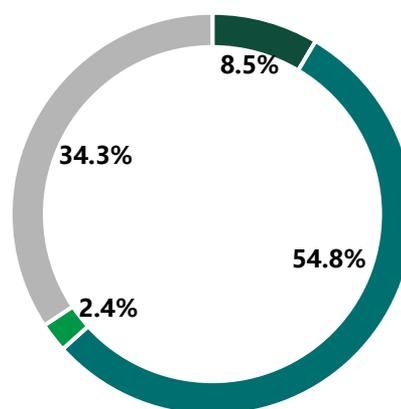
RDF 格式（资源描述框架）作为 W3C 的推荐标准，使用了“主体 - 谓词 - 客体”三元组形式来描述数据资源和它们之间的关系，使数据资源更易于被机器阅读和理解，有助于数据利用者发现和链接相关数据资源。目前，全国仅有浙江、深圳、贵阳三地提供了符合 RDF 格式的数据集。

### (4) API 开放性

API 开放性是指申请与调用 API 接口的难易程度。繁琐的申请流程与高技术门槛不利于数据的获取和利用。目前仅有一成的地方平台上调用 API 接口的难度较低，而在近 35% 的地方平台上，API 接口调用难度较高。

#### 6.3.3 开放类型

开放类型是指平台对数据集明确标示了不同的开放类型，包括无条件开放、有条件开放和不予开放等，明确标示有利于推进数据的分级分类开放。目前只有不到一成的地方平台为每个数据集明确标示了包括有条件开放在内的多种开放类型，多数地方平台只提供了无条件开放一种类型，各地平台提供的开放类型分布情况如图 6-11 所示。



■ 提供多种类型 ■ 只提供无条件开放 ■ 分类混乱 ■ 未分类

图 6-11 各地开放类型提供情况分布

### 6.3.4 描述说明

#### (1) 元数据提供

在平台上提供数据集的元数据有助于用户更清晰地了解数据资源的内容与背景，从而更好地获取和利用数据。通过梳理我国政务信息资源编目指南中关于核心元数据的说明，国际开放数据平台上提供的元数据以及目前我国半数以上的地方平台已实际提供的元数据，报告确定了以下十二项内容作为开放数据集的基本元数据，分别是：数据名称、提供方、数据摘要、数据主题、关键字、数据格式、数据项信息、开放属性、更新周期、发布日期、更新日期、数据量。

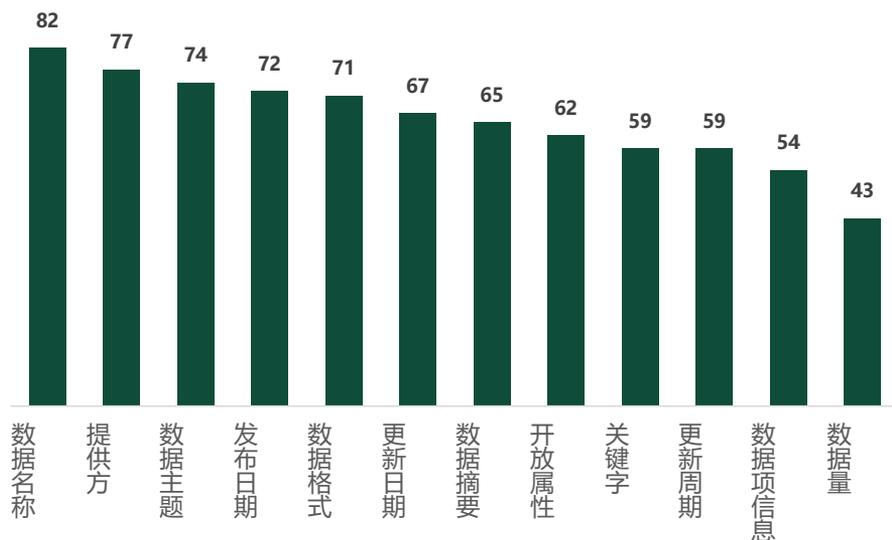
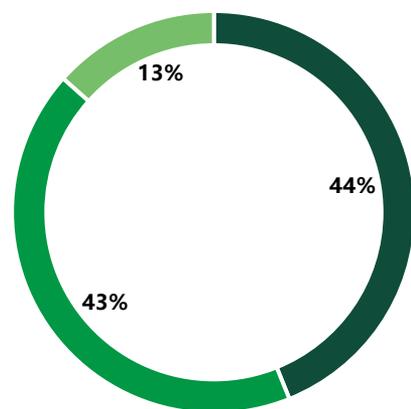


图 6-12 基本元数据在各地平台上的出现次数

图 6-12 展示了十二项基本元数据在 82 个地方平台上的出现次数。图 6-13 展示了各地平台基本元数据的提供情况，仅有 44% 的地方平台提供了所有的基本元数据。



■ 全提供 ■ 提供过半 ■ 提供未过半

图 6-13 各地基本元数据提供情况分布

(2) API 说明

在平台上提供 API 说明有助于用户了解 API 接口的调用方式。API 接口说明包括基本调用说明与使用操作指南两种形式。基本调用说明是指提供了调用接口的说明，包括请求地址、请求参数与返回参数等信息。使用操作指南是指提供了详细的 API 接口使用方法与流程。如图 6-14 所示，在 82 个地方平台中，有 46 个提供了基本调用说明，6 个提供了使用操作指南，还有 30 个地方平台未提供任何说明。

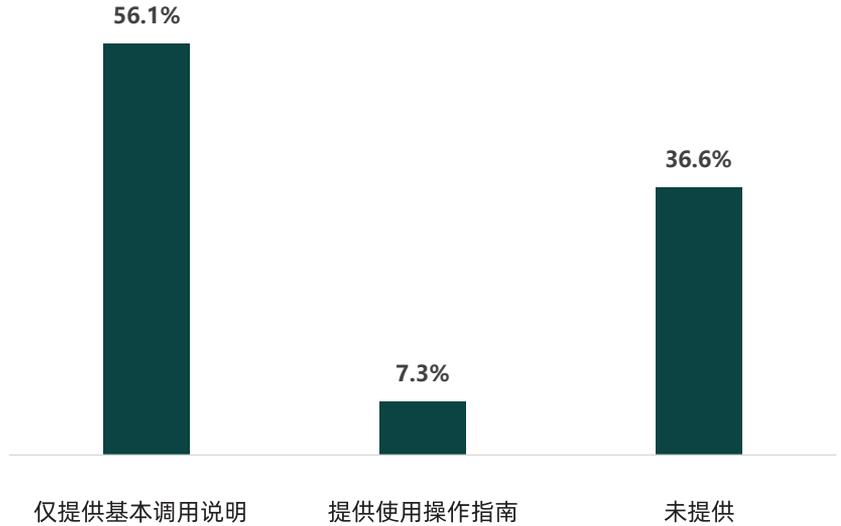


图 6-14 各地 API 说明提供情况分布

6.4 数据覆盖面

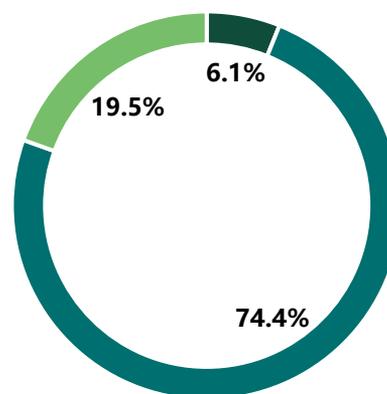
6.4.1 主题覆盖面

开放多个主题领域的数据集有利于提高数据开放的广度和覆盖面，也有利于数据利用者获取和整合来自多个领域的数据进行融合利用。根据 2015 年 8 月国务院印发的《促进大数据发展行动纲要》和 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》中的相关要求，结合国内外数据开放平台上划分的数据集主题，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、信用服务共十四个大类。图 6-15 展现了各地开放的数据集主题分布情况。



图 6-15 开放数据集主题分布

报告进一步分析了各地平台上开放的数据集对各个主题的覆盖情况（见图 6-16），目前只有五个地方平台覆盖了全部十四个主题。

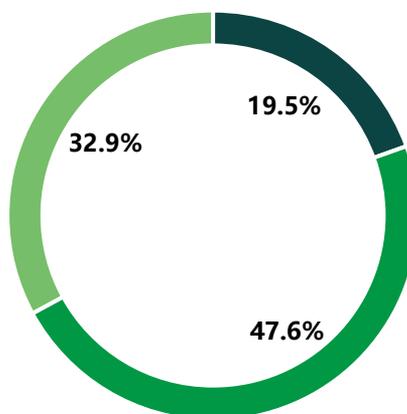


■ 全覆盖 ■ 覆盖过半 ■ 覆盖未过半

图 6-16 各地主题覆盖情况分布

#### 6.4.2 部门覆盖面

开放数据的部门覆盖面反映了各个政府部门开放数据的参与程度，也体现了数据集来源的丰富程度。由于不同地方相同职能的政府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 6-17 为各地平台上数据提供部门数量占该地方政府部门总数（数据来自各地政府门户网站）的比例。目前，全国只有不到 20% 的地方实现了数据提供部门全覆盖。



■ 全覆盖 ■ 覆盖过半 ■ 覆盖未过半

图 6-17 各地数据提供部门覆盖情况分布

各地平台上开放的数据集的部门分布情况如图 6-18。目前，来自统计部门的数据集数量最多，数量远大于其后的市场监管、教育、住建等部门。然而，开放数据集总量中统计部门数据集占比较高的地方，其数据容量普遍较低（如图 6-19 所示）。与业务部门开放的数据集相比，统计部门提供的数据集多为经过加工归总后的宏观分析结果，颗粒度较大、数据容量较低、利用价值也不高，不应成为数据开放的主力部门。

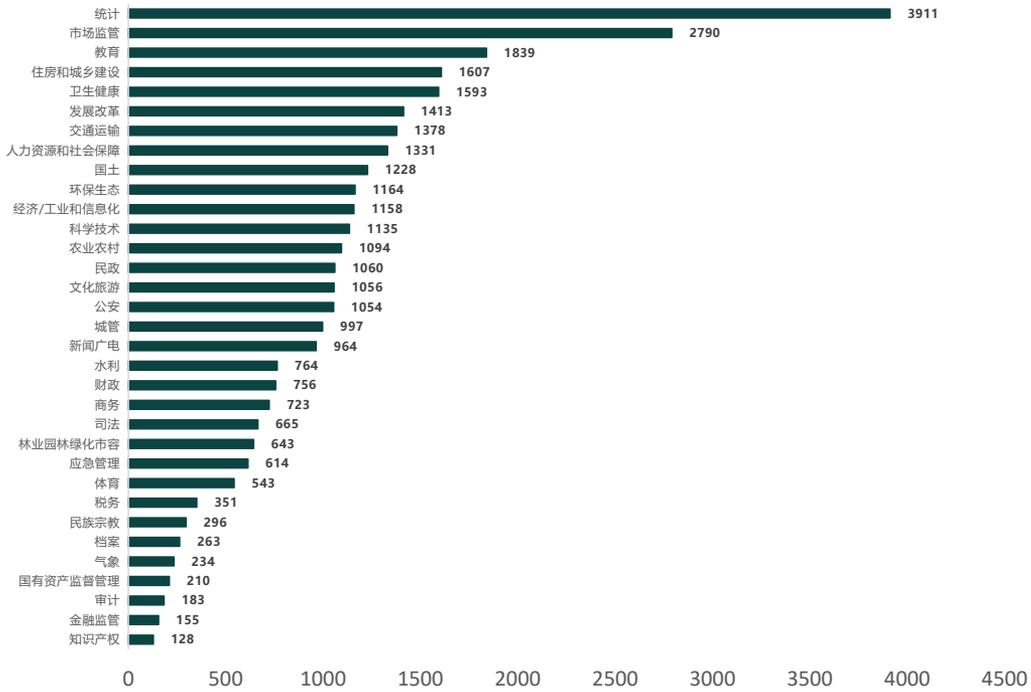


图 6-18 开放数据集部门分布

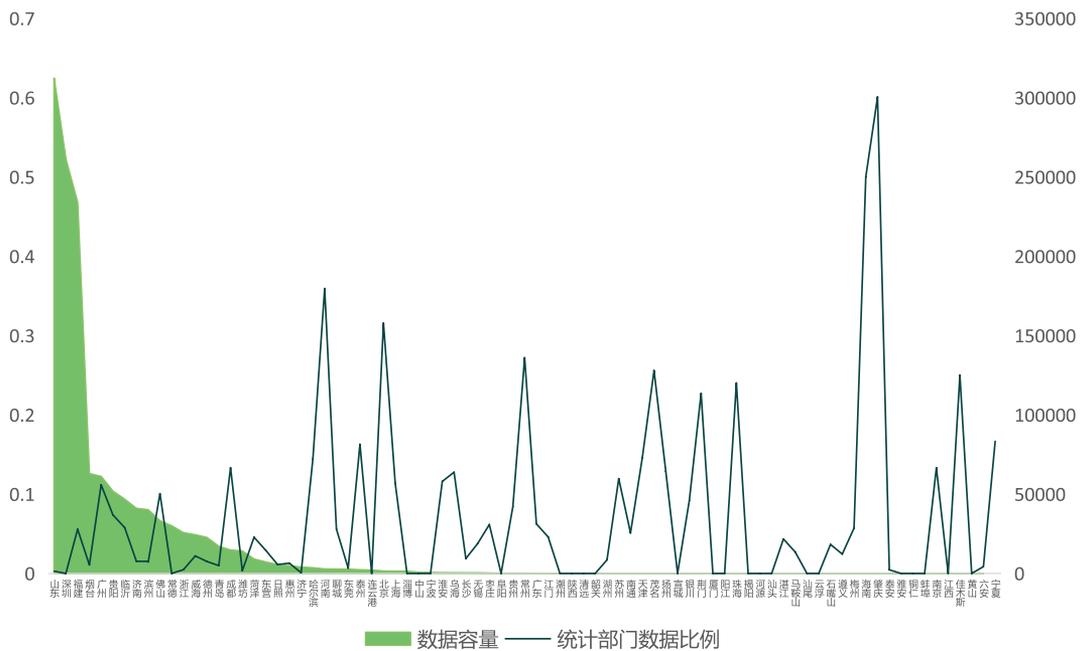


图 6-19 各地数据容量与统计部门比例比较 (数据容量单位: 万)



报告对这十四类常见数据集在 82 个地方平台上的开放情况进行了分析（如图 6-21 所示），其中行政许可处罚类是地方平台开放最多的数据，共有 58 个地方平台开放了此类数据。在地方平台开放的数据集中出现常见数据集的情况如图 6-22 所示。目前，只有 15% 左右的地方平台开放了所有常见数据集，近 50% 的地方平台上只开放了不到一半的常见数据集。



图 6-21 常见数据集各地开放情况分布

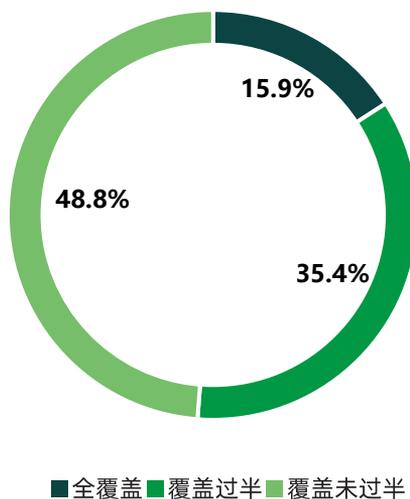


图 6-22 各地常见数据集覆盖情况分布

## 6.5 数据持续性

### 6.5.1 持续增长

平台上线后应持续更新和增加数据集，以满足社会对开放数据日益增长的需求。报告通过跟踪各地平台从 2017 第二季度到 2019 年第一季度这一时间段内每季度新增数据集的情况后发现，只有 8.5% 的地方平台在近八个季度连续发布了新的数据集，保持数据集持续增长，而其他地方存在新增数据集中断或停滞的情况（如图 6-23 所示）。

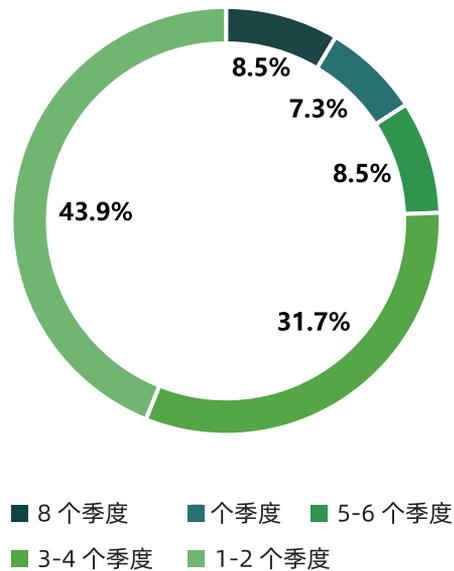


图 6-23 各地数据集持续增长的季度数

### 6.5.2 动态更新

报告监测了地方平台在 2019 年 1 月至 2019 年 3 月期间更新的数据集数量，约 46% 的地方平台在该时段内没有出现更新数据（如图 6-24 所示）。数据动态更新前十名的地方如图 6-25 所示。

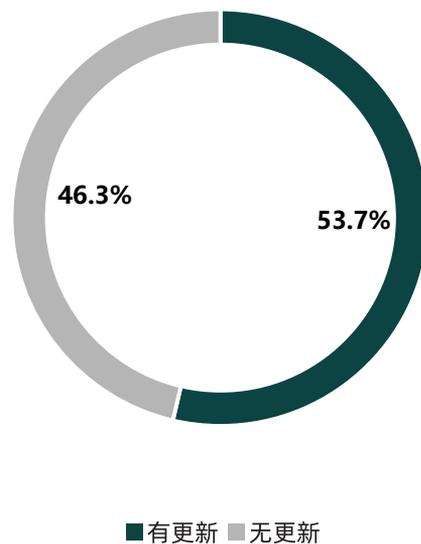


图 6-24 各地发现数据更新情况分布

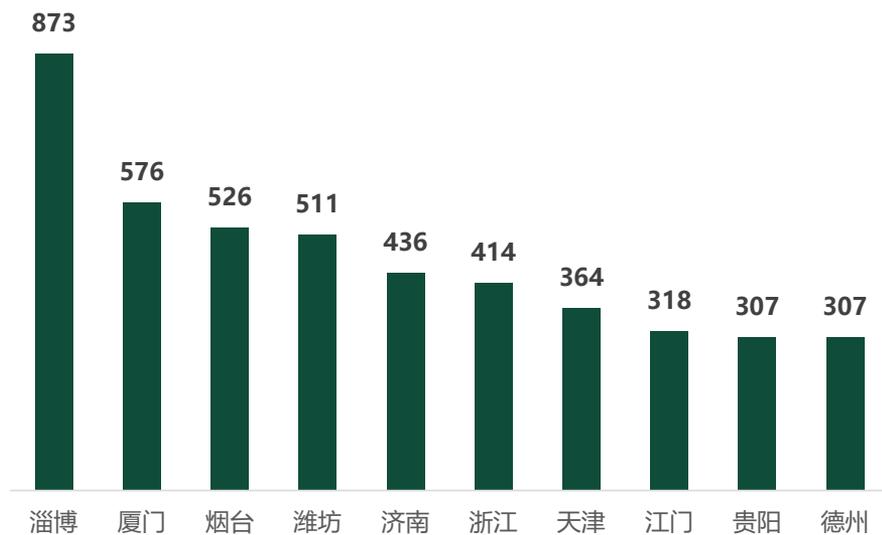


图 6-25 各地平台数据动态更新情况（前十名）

### 6.5.3 历史存档

历史存档是指平台留存了历史上不同时间发布的多个批次的数据，使用户可获取和利用。图 6-26 展示了各地提供历史存档的分布情况，约 41.5% 的地方平台提供了历史存档。

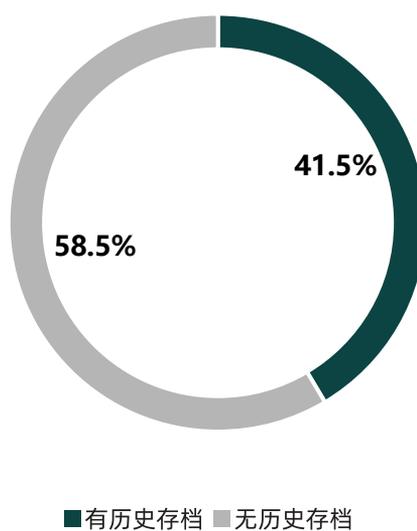
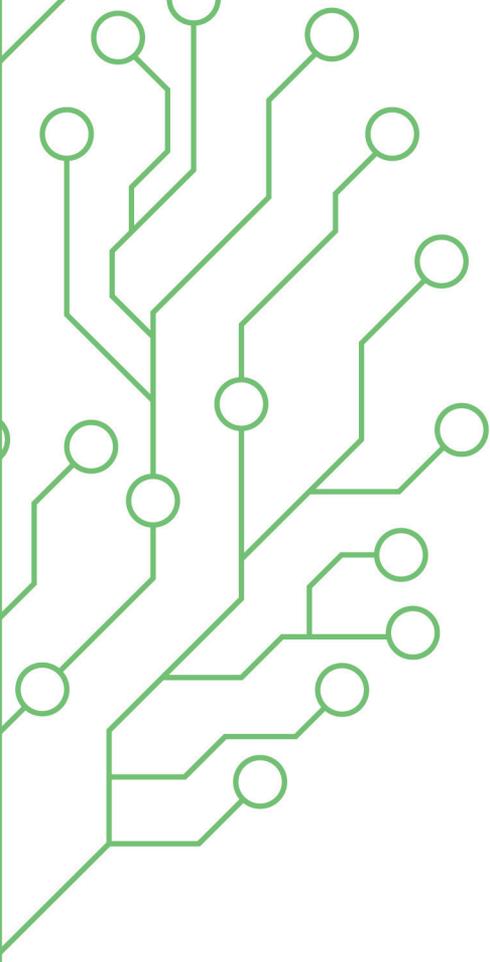
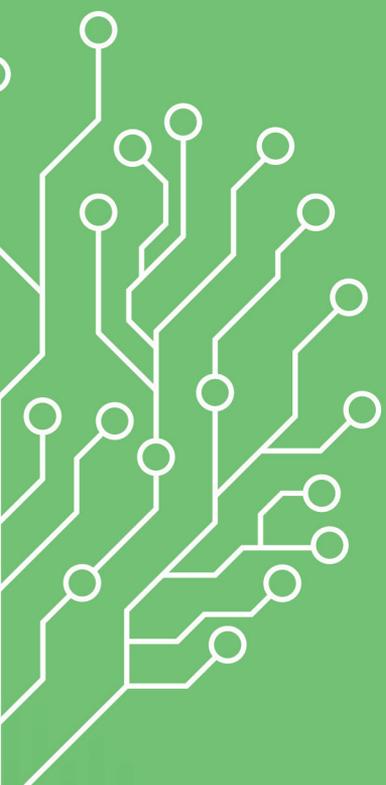


图 6-26 各地提供历史存档情况分布

USE





CHAPTER



利用层发现



## 七、利用层

报告从以下三个维度对开放政府数据利用情况进行了评估，主要发现如下：

**在利用促进方面**，目前仅有一成多的地方举办过开放数据利用促进活动，主要是以开放数据创新利用比赛的形式开展，其余近九成的地方还未举办过类似活动。然而，这些比赛大多未能持续，只有上海和北京两个地方做到了连续举办。

**在成果产出方面**，利用成果类型单一，数量极少，目前仅在 3.7% 的地方平台上发现了少量的有效服务应用，仅有 2.4% 的地方平台展示了部分开放数据创新利用比赛的创新方案。利用成果的主题覆盖面较小，仅在交通出行、教育科技和卫生健康主题下发现了有效应用。利用成果的质量也普遍堪忧，优质应用屈指可数。超过六成的地方平台展示的是由政府部门自身开发而不是政府数据开放后被社会开发利用所产生的应用；近两成的地方平台展示的应用无法获取或无法正常使用；超过九成的地方平台展示的利用成果未标注数据来源；而在标注了数据来源的平台中，又有六成的地方平台上出现了来源数据集链接无效、搜索不到数据或数据不完整的情况。

**在数据利用方面**，开放政府数据的利用者类型较为单一，在发现了有效利用成果的地方平台中，仅出现了企业和个人两种类型的利用者，尚未发现有高校、研究机构和社会团体等其他利用者。超过九成的地方平台上开放的数据未被利用，仅在 3.7% 的地方平台上出现了被利用数据集，它们分布在交通出行、卫生健康、信用服务、经贸工商、教育科技、机构团体、公共安全和农业农村 8 个主题，交通运输、统计、教育、市场监管、人力资源与社会保障、农业农村、公安和卫生健康 8 类政府部门中。

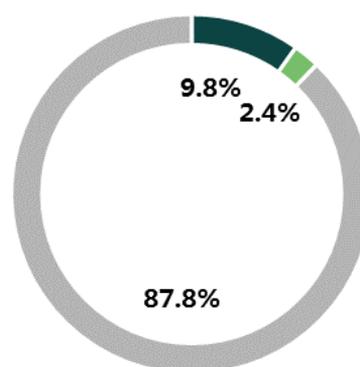
---

综上所述，**报告建议**：在开放数据利用促进方面，各地应积极举办并持续开展类似于开放数据创新利用比赛的促进活动，用以提升社会利用政府数据的积极性，营造有利于政府数据开放利用的生态体系。在成果产出方面，亟需提升有效、优质利用成果的类型、数量和主题覆盖面，清理各地平台上政府自身开发的应用，剔除无法获取或正常使用的应用，并为利用成果标明数据来源并提供数据集链接。在数据利用方面，鼓励引导多种类型的数据利用者来参与政府开放数据的开发利用。开放优质的、高价值的政府数据是推动数据利用的基本条件，各地应把数据供给做实做优，然后在此基础上推动数据利用，提升被利用数据集的数量、容量、主题覆盖面和部门覆盖面。

## 7.1 利用促进

利用促进指政府为了推动开放数据的社会化利用而组织的各类活动。组织利用促进活动有利于提升社会利用政府数据的积极性和创造力，也有助于政府部门更好地了解社会的数据需求，开放更多的优质数据，从而通过促进数据提供者与数据利用者之间的双向互动，营造有利于政府数据开放利用的生态体系和社会氛围。

组织开放数据创新利用比赛是一种常见的利用促进活动。各地近两年举办开放数据创新利用比赛的分布情况如图 7-1 所示。例如：2018 年上海 SODA 开放数据创新应用大赛开放了来自政府部门和公共企事业单位的数据供参赛团队进行创新应用比赛，这些数据有的已在上海市的平台上开放，有的则是为大赛专门提供。



- 举办开放数据创新利用比赛
- 持续举办开放数据创新利用比赛
- 未举办开放数据创新利用比赛

图 7-1 各地举办开放数据创新利用比赛情况分布

## 7.2 成果产出

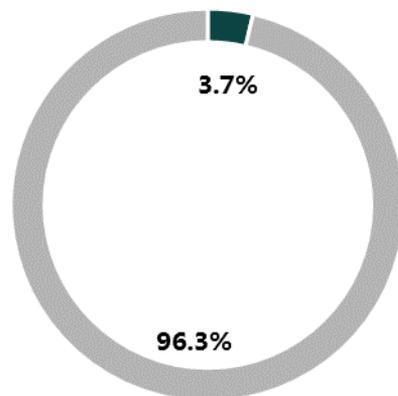
成果产出是指社会利用开放数据所产生的利用成果的类型、主题覆盖和质量。政府数据开放的根本目的是推动政府数据的社会化利用及再利用，并产生各类优质利用成果。

### 7.2.1 成果类型

成果类型是指平台上展示的服务应用、创新方案、传播产品和研究成果等类型的利用成果。目前，各地平台上展示的利用成果只有服务应用和创新方案两类。

### (1) 服务应用

有效的服务应用是指平台上展示的同时满足以下标准的应用：由数据利用者而非政府部门自身开发利用所产生；确实利用了开放的政府数据；能正常获取；能正常使用。各地平台上展示的有效服务应用的分布情况如图 7-2 所示。目前，仅在北京、贵州和贵阳 3 个地方平台上发现了有效的服务应用，例如贵阳市平台上展示的“慧停车”APP（见图 7-3）。



■ 有效服务应用 ■ 无有效服务应用  
图 7-2 各地平台有效服务应用情况分布

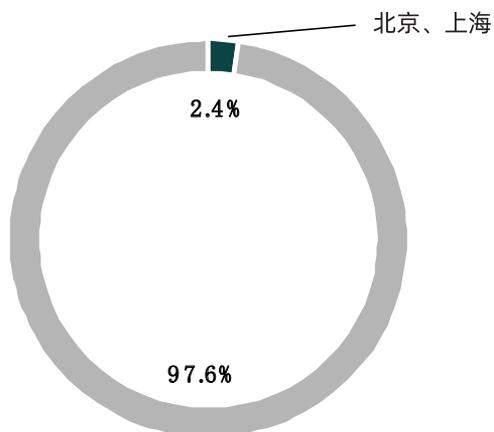
#### 应用详情



图 7-3 贵阳市平台上的“慧停车”APP

### (2) 创新方案

创新方案是指平台上展示的开放数据创新利用比赛参赛者产生的创新方案。展示这类创新方案有利于创新方案的落地，也有利于启发其他数据利用者的思维，推动更多创新利用成果的产生。目前，仅有北京和上海在平台上展示了开放数据创新利用比赛中形成的创意方案（见图 7-4）。上海市平台在其“典型应用”模块所展示的创新方案样例如图 7-5 所示。



■ 展示创新方案 ■ 未展示创新方案  
图 7-4 各地平台展示的创新方案情况分布



图 7-5 上海市平台展示的 SODA 大赛创新方案样例

### 7.2.2 成果主题覆盖

成果主题覆盖是指平台上有效应用的主题覆盖面。目前，各地平台上发现的有效应用的主题覆盖了交通出行、教育科技和卫生健康，有效应用覆盖的主题分布情况如图 7-6 所示。

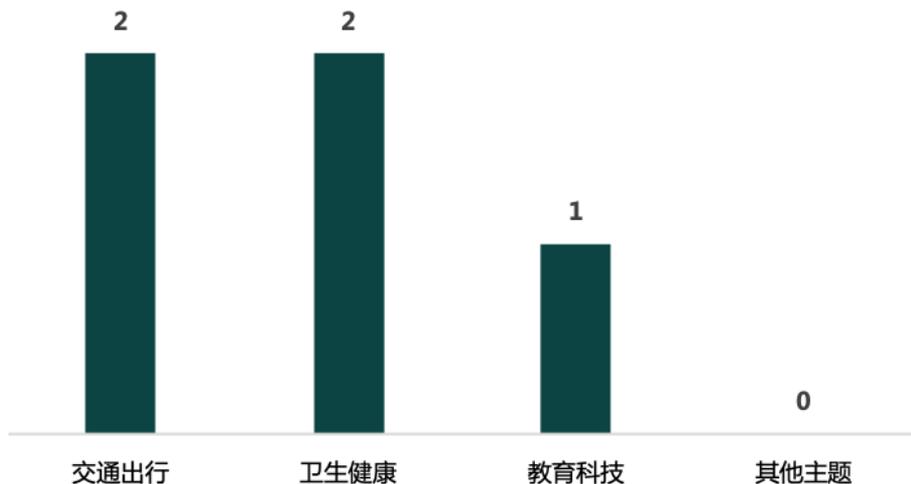


图 7-6 各地有效应用主题覆盖情况分布

### 7.2.3 成果质量

成果质量是指平台上展示的开放数据利用成果的质量。报告目前主要从优质服务应用和无问题成果两个方面进行评估。

#### (1) 优质服务应用

优质服务应用需要同时满足以下 4 个条件：符合有效服务应用的标准；利用的核心数据集为政府开放数据；利用了大容量的政府开放数据；除利用政府开放数据外，还融合利用了其他来源的数据。这类优质服务应用具有产生较大社会经济价值的潜力。目前，贵州省平台上的货车帮应用、贵阳市平台上的慧停车和贵医云等应用基本符合优质服务应用的标准，其他地区尚未发现优质服务应用。

#### (2) 无问题成果

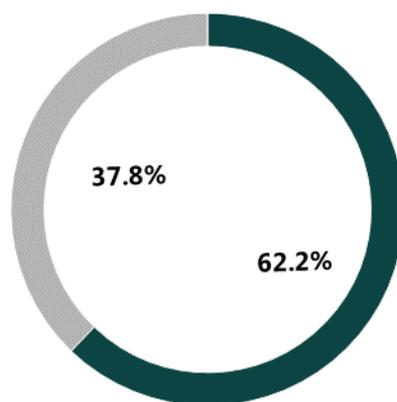
问题成果包括政府自身开发的成果、无效成果、数据来源不明成果和数据关联缺失的成果。

#### ■ 无政府自身开发成果

政府自身开发成果是指平台上展示的是由政府部门自身或委托第三方开发的政务服务类应用，这类成果不是政府向社会开放数据后被社会开发利用所产生的。各地平台上政府自身开发成果的分布情况如图 7-7 所示。

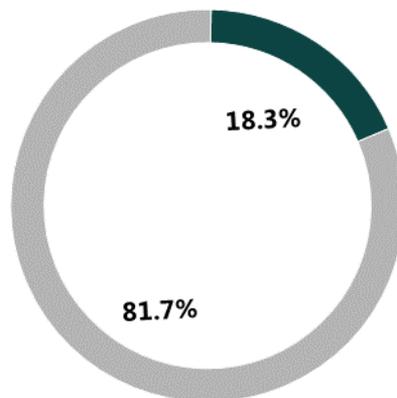
#### ■ 无无效成果

无效成果是指平台上展示了无法搜索到、能搜索到但无法下载或能下载但无法正常使用的应用。各地平台上展示的无效成果的分布情况如图 7-8 所示。



■ 有政府自身开发成果 ■ 无政府自身开发成果

图 7-7 各地平台政府自身开发成果情况分布

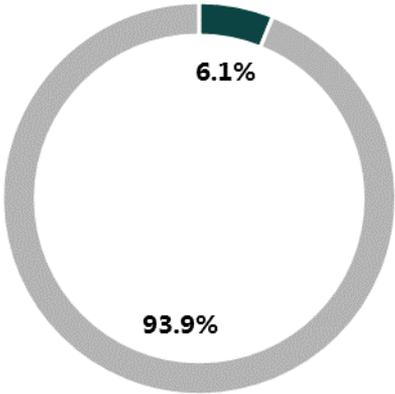


■ 有无效成果 ■ 无无效成果

图 7-8 各地平台展示的无效成果情况分布

■ 无数据来源不明

数据来源不明是指平台上展示的利用成果没有标示数据来源，因而无法辨别这些成果是否真正利用了政府开放数据以及利用了哪些开放数据。目前，各地平台上标示利用成果数据来源的分布情况如图 7-9 所示。

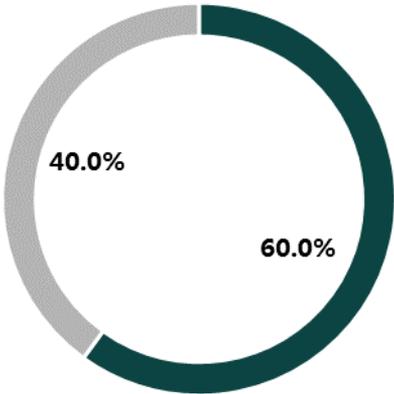


■ 标注数据来源 ■ 未标注数据来源

图 7-9 各地平台标示成果数据来源情况分布

■ 无数据关联缺失

数据关联缺失是指利用成果虽然标注了所利用的开放数据集，但这些数据集在平台上链接不到、搜索不到或数据不完整，无法追溯这些利用成果是否真正利用了开放数据。目前，在标注了数据来源的地方平台中，数据关联缺失的分布情况如图 7-10 所示。



■ 有数据关联缺失 ■ 无数据关联缺失

图 7-10 各地平台数据关联缺失情况分布

### 7.3 数据利用

数据利用是指数据利用者为开放数据的利用方式和过程。本报告主要从数据利用者多样性、被利用数据数量、被利用数据集的主题覆盖和被利用数据集的部门覆盖四个方面评估数据利用。

#### 7.3.1 利用者多样性

利用者多样性是指平台上开放政府数据利用者类型的多样性。开放政府数据的利用者可包括企业、高校、研究机构、社会团体和个人等。目前，在发现了有效利用成果的地方平台上，仅出现了企业和个人两种类型的利用者。

#### 7.3.2 被利用数据数量

被利用数据数量是指平台上有效应用所使用的数据数量。报告主要从被利用数据集数量和被利用数据集容量两个方面进行评估。

##### (1) 被利用数据集数量

被利用数据集数量是指平台上有效应用所使用的开放数据集的数量。各地平台上被利用数据集数量的分布情况如图 7-11 所示。

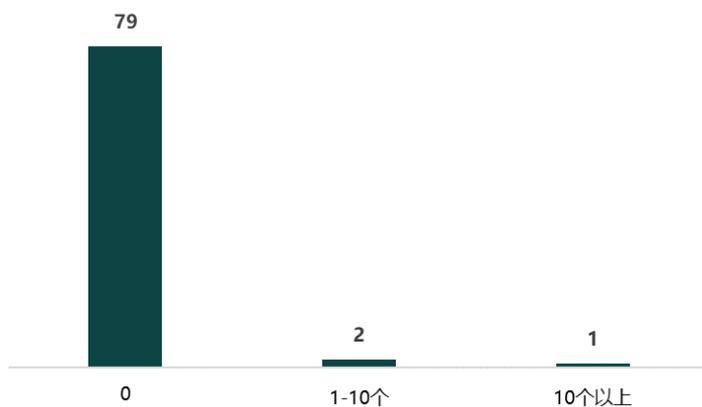


图 7-11 各地平台被利用数据集数量分布

##### (2) 被利用数据集容量

被利用数据集容量是指平台上有效应用所使用的开放数据集的容量，反映了被利用数据集的实际数据量大小。各地平台上被利用数据集容量的分布情况如图 7-12 所示。

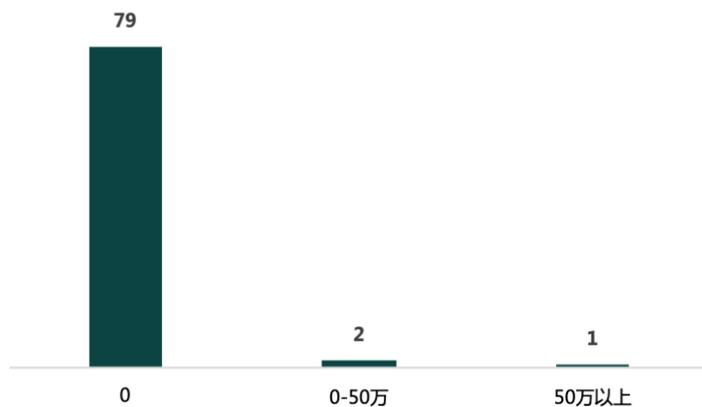


图 7-12 各地平台被利用数据集容量分布

### 7.3.3 被利用数据集的主题覆盖

被利用数据集的主题覆盖是指平台上有效应用所使用的开放数据集涵盖的主题，即这些应用利用了哪些主题下的开放数据。目前，各地平台上被利用的数据集覆盖了交通出行、卫生健康、信用服务、经贸工商、教育科技、机构团体、公共安全和农业农村这 8 个主题。各地平台被利用数据集所覆盖的主题分布情况如图 7-13 所示。

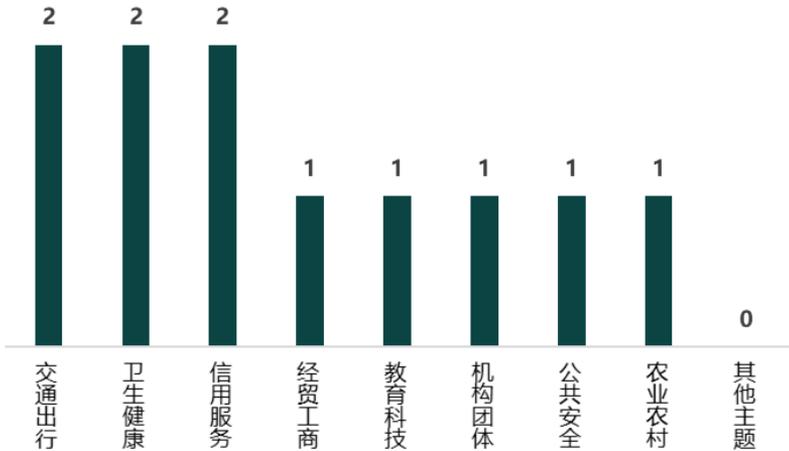


图 7-13 各地平台被利用数据集的主题覆盖情况分布

### 7.3.4 被利用数据集的部门覆盖

被利用数据集的部门覆盖是指平台上有效应用所使用的开放数据集覆盖到的部门，即这些应用利用了来自哪些政府部门的开放数据。目前，各地平台上被利用的数据集来自交通运输、统计、教育、市场监管、人力资源与社会保障、农业农村、公安和卫生健康这 8 类政府部门。各地平台被利用数据集覆盖的部门分布情况如图 7-14 所示。

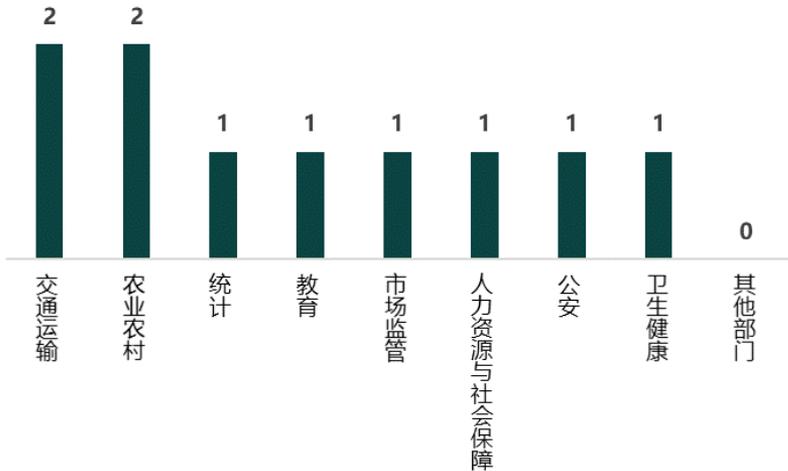
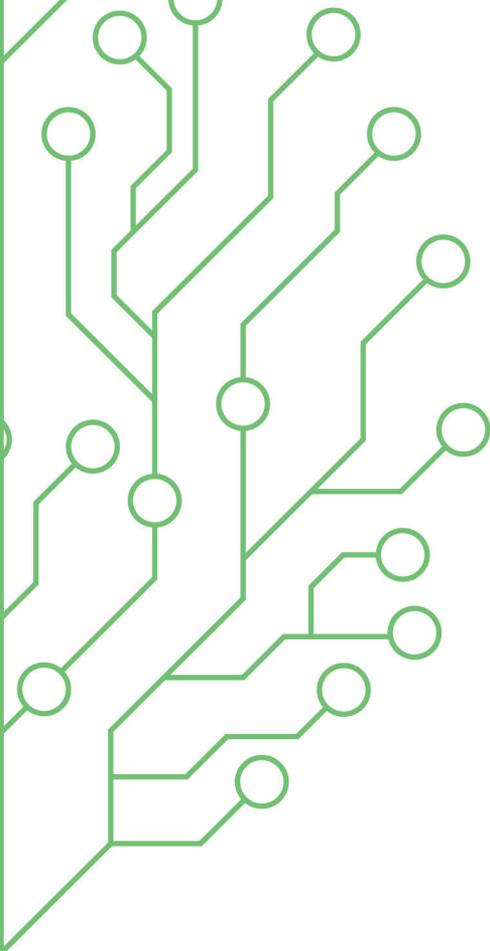
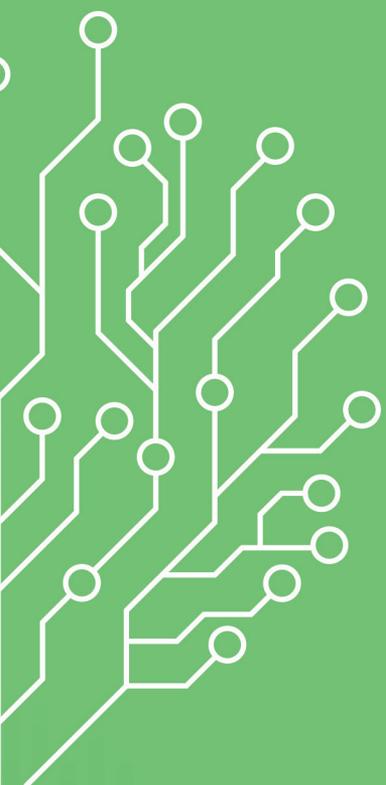


图 7-14 各地平台被利用数据集的部门覆盖情况分布

# APPENDIX





# 附录



## 评估专家评委会 METHODOLOGY COMMITTEE

Amparo Ballivian	世界银行
安小米	中国人民大学
常江	上海市司法局
陈德柱	影子数据
陈涛	上海图书馆
陈涛	华中科技大学
陈晓建	优刻得
陈瑶	上海对外经贸大学
陈玉梅	暨南大学
邓崧	云南大学
樊博	上海交通大学
范佳佳	上海社会科学院信息研究所
方小敏	虎牙直播
付熙雯	西北大学
高冰	中电科新型智慧城市研究院有限公司
高丰	开放数据中国
巩福	深圳太极云软技术有限公司
郭伟民	联合国经社部
何渊	上海交通大学
胡广伟	南京大学
胡业飞	天津大学
黄萃	浙江大学
黄璜	北京大学
黄志敏	数可视
惠志斌	上海社科院
贾西贝	深圳市智慧城市大数据研究院
金耀辉	上海交通大学
李伯一	埃克塞特大学
李雯	城市数据派
李宵声	欧莱雅中国
李晓方	南京理工大学
李重照	云南财经大学
梁春鼎	广州零点有数数据科技有限公司
刘冬梅	交通部公路院
刘红波	华南理工大学
刘汪洋	中电科大数据研究院有限公司
刘文静	暨南大学

刘新萍	上海理工大学
刘勇	九次方大数据
刘枝	国家信息中心
龙怡	上海政法学院
陆小敏	上海市信息中心
马亮	中国人民大学
孟天广	清华大学
孟晔	美团城市创新实验室
宋华琳	南开大学
宋文好	电子政务杂志
孙荣	同济大学
孙宗锋	山东大学
谭海波	湖南大学
汤志伟	电子科技大学
王法硕	华东师范大学
王芳	南开大学
王强	百度
王少辉	武汉大学
王翔	复旦大学
王晓斌	浪潮集团
吴昊	第一财经数据科技
夏义堃	武汉大学
肖卫兵	上海政法学院
徐亚敏	世界经济论坛
薛梓闻	速溶360
阎冠和	北京市经济和信息化委员会
杨道玲	国家信息中心
殷沈琴	复旦大学
袁莉	四川大学
张峰	浪潮集团
张会平	电子科技大学
张楠	清华大学
章燕华	浙江大学
郑跃平	中山大学
朱琳	华东理工大学
庄国波	南京邮电大学

注：评委名单按照姓氏拼音排序

## 指数出品方



## 联合发布方



## 指数出品团队

郑磊、刘新萍、吕文增、周业光、袁佳蕾、温祖卿、韩笑、  
王翔、戚馨予、付熙雯、刘利亚、董煜、方金传、孙一蕾、陈欢

## 合作单位



冥睿（上海）信息科技有限公司



中山大学数字治理研究中心



复旦发展研究院  
FUDAN DEVELOPMENT INSTITUTE



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



中国开放数据指数  
CHINA OPEN DATA INDEX

· 2019 ·

