



中国开放数林指数
CHINA OPEN DATA INDEX

复旦智库报告

2021

2021 粤港澳大湾区 政府数据开放报告

报告制作发布方 | 复旦大学数字与移动治理实验室
联合发布方 | 粤港澳大湾区大数据研究院



Guangdong-Hong Kong-Macao
Greater Bay Area





中国开放数林指数
CHINA OPEN DATA INDEX

· Since 2017 ·

开放数据 蔚然成林

CONTENTS

引言	001
■ 大湾区数林概貌	003
■ 大湾区数林指标	005
■ 大湾区数林指数	008
■ 大湾区数林标杆	010
■ 数林建言	035
■ 评估方法	037
附录	040

引言

粤港澳大湾区包括香港特别行政区、澳门特别行政区和广东省广州市、深圳市、珠海市、佛山市、惠州市、东莞市、中山市、江门市、肇庆市，是我国开放程度最高、经济活力最强的区域之一，在国家发展大局中具有重要战略地位。2019年2月18日，中共中央、国务院印发《粤港澳大湾区发展规划纲要》，明确“粤港澳大湾区不仅要建成充满活力的世界级城市群、国际科技创新中心、‘一带一路’建设的重要支撑、内地与港澳深度合作示范区，还要打造成宜居宜业宜游的优质生活圈，成为高质量发展的典范。”

2020年10月11日，中共中央办公厅、国务院办公厅发布《深圳建设中国特色社会主义先行示范区综合改革试点实施方案（2020-2025年）》，提出要“试点推进政府数据开放共享。支持建设粤港澳大湾区数据平台”。2021年9月22日，广东省人民政府通过《广东省公共数据管理办法》，要求“公共数据应当依法有序开放”。

粤港澳大湾区是我国政府数据开放的先行地区，进一步提升粤港澳大湾区的政府数据开放水平，不仅有利于促进数据要素在粤港澳大湾区的流动和利用，还有利于充分发挥粤港澳综合优势，提升粤港澳大湾区在国家经济发展和对外开放中的支撑引领作用。在此背景下，本报告对粤港澳大湾区政府数据开放的现状和水平进行了研究和评价。

《粤港澳大湾区政府数据开放报告》是“中国开放数林指数”发布的区域性系列报告之一。“中国开放数林指数”是我国首个专注于评估政府数据开放水平的专业指数，由复旦大学数字与移动治理实验室制作出品。

“开放数林”意喻我国政府数据开放利用的生态体系，“开放数林指数”自2017年5月首次发布以来，定期对我国地方政府数据开放水平进行综合评价，精心测量各地“开放数木”的繁茂程度和果实价值，助推我国政府数据开放生态体系的培育与发展。

开放数据，蔚然成林，粤港澳大湾区的每一棵开放“数木”不仅枝繁叶密、花开结果，更能根系相通、枝叶相连，终将成长为一片茂盛多样，互联互通的粤港澳大湾区“开放数林”。

一、大湾区数林概貌

香港于 2011 年上线了政府数据开放平台“资料一线通”。自 2015 年以来，粤港澳大湾区上线的地方政府数据开放平台数量持续快速增长（如图 1 所示），目前，粤港澳大湾区的两个特别行政区与九个城市都已上线了政府数据开放平台。各地平台上线时间的地理空间分布如图 2 所示，具体上线时间如表 1 所示。

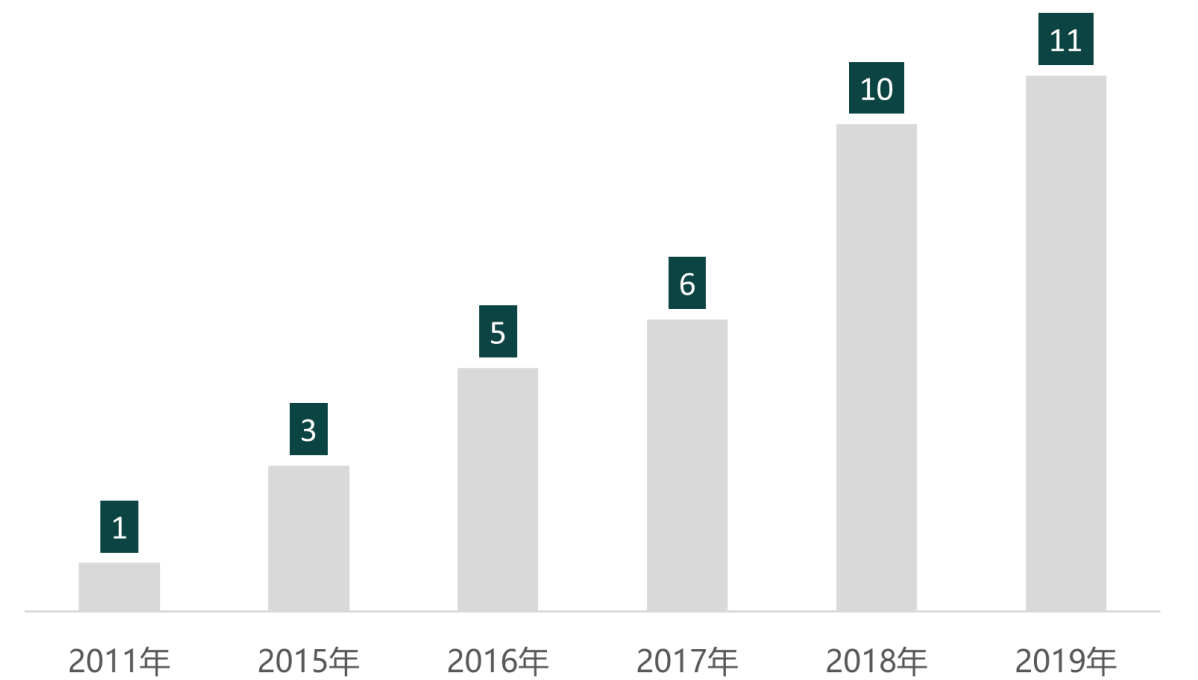


图 1 粤港澳大湾区政府数据开放平台数量

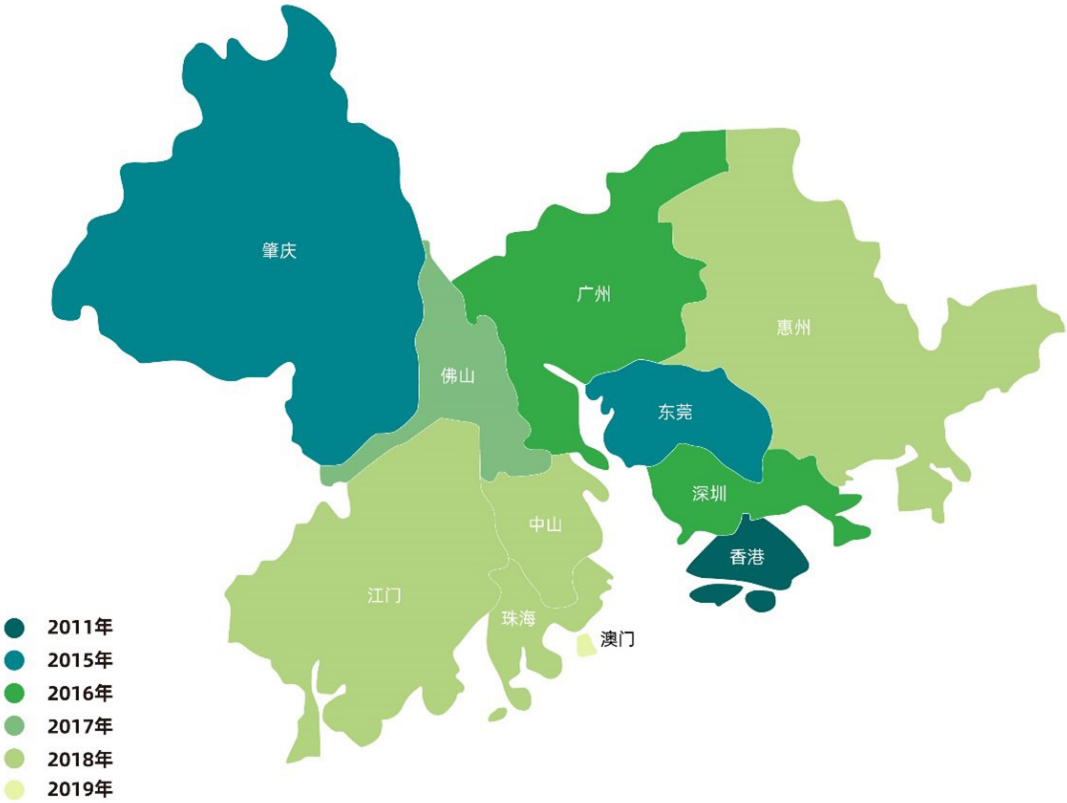


图 2 粤港澳大湾区政府数据开放平台上线时间的空间分布

表 1 粤港澳大湾区平台上线时间一览表（按拼音首字母排序）

上线时间	地方
2019 年	澳门
2018 年	惠州、江门、中山、珠海
2017 年	佛山
2016 年	广州、深圳
2015 年	东莞、肇庆
2011 年	香港

二、大湾区数林指标

2.1 评估指标体系

《粤港澳大湾区政府数据开放报告》在“中国开放数林指数”评估指标体系的基础上，立足国家政策方向，根据粤港澳大湾区的区域特点和创新实践，对部分指标进行了调整，如准备度层面增加了“数据集开放计划”指标，数据层增加了对数据描述说明中数据字典指标。最终确定的粤港澳大湾区开放数林指数评估指标体系包括准备度、平台层、数据层、利用层四个维度及下属多级指标（如图 3 所示）：

准备度是“数根”，是数据开放的基础，包括法规政策效力与内容、组织与领导等两个一级指标。

平台层是“数干”，是数据开放的枢纽，包括发现预览、数据获取、成果提交展示、互动反馈、个性化体验等五个一级指标。

数据层是“数叶”，是数据开放的核心，包括数据数量、数据质量、数据规范、开放范围等四个一级指标。

利用层是“数果”，是数据开放的成效，包括利用促进、利用多样性、成果数量、成果质量等四个一级指标。

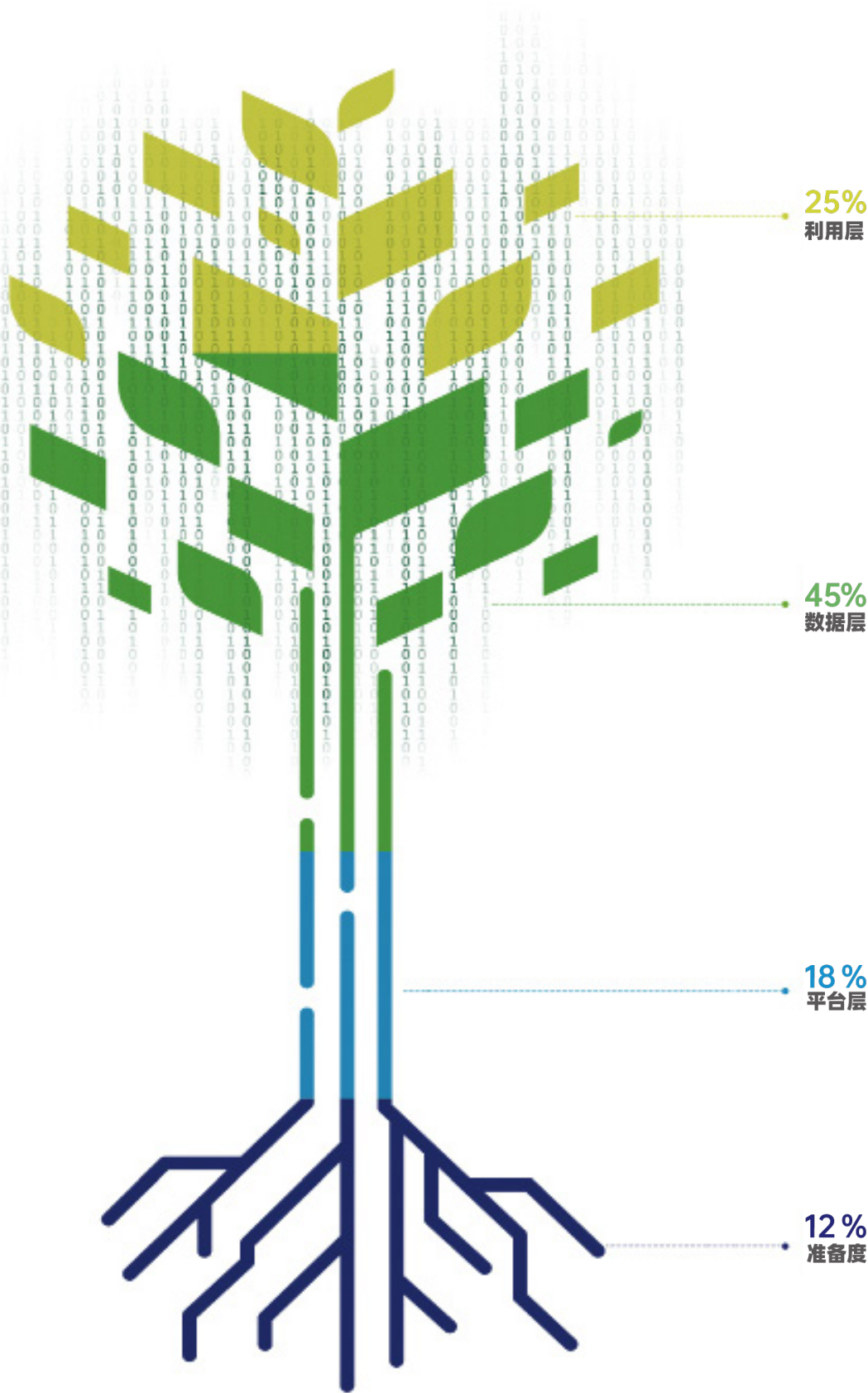


图 3 粤港澳大湾区开放数林指数评估指标体系

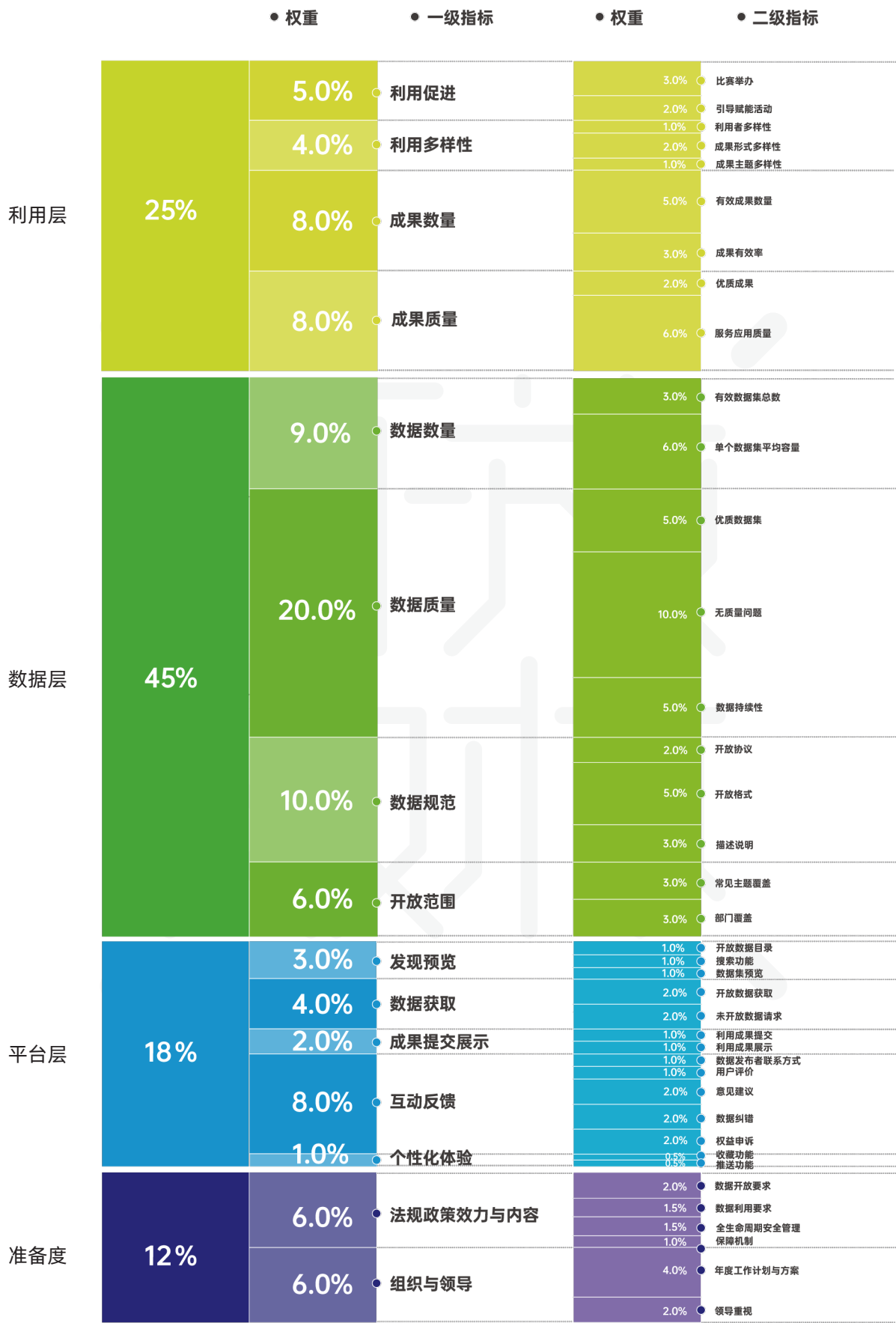


图 3 粤港澳大湾区开放数林指数评估指标体系

三、大湾区数林指数

截至 2021 年 6 月，粤港澳大湾区各地开放数林指数等级如表 2 所示，各地指数综合得分可分为五个等级。深圳、香港开放数据综合等级最高，为 A+ 等级；广州、东莞紧随其后，为 A 等级；其次是江门、澳门和中山等地。在单项维度上，香港在准备度和利用层上表现最优，进入 A+ 等级；深圳在平台层和利用层上处于领先地位，进入 A+ 等级。

■ 表 2 粤港澳大湾区各地开放数林指数等级

地 方	准备度等级	平台层等级	数据层等级	利用层等级	综合等级	综合指数区间
深 圳	A	A+	A	A+	A+	55以上
香 港	A+	B+	A	A+	A+	
广 州	C	A+	B+	B+	A	40 ~ 45
东 莞	C	B	A+	B	A	
江 门	C+	B+	B+	B+	B+	30 ~ 40
澳 门	B+	B	B+	B	B+	
中 山	B	C+	B+	B	B+	
惠 州	C	B	B+	C+	B	25 ~ 30
肇 庆	C	A	B	B	B	
佛 山	C+	A	B	C+	B	
珠 海	C	C+	C+	B	C+	15 ~ 20

大湾区开放数林指数综合等级的空间分布如图 4 所示，颜色越深代表综合等级越高，综合等级较高的地方主要集中在香港、深圳、东莞、广州这一条绿色“林带”上。

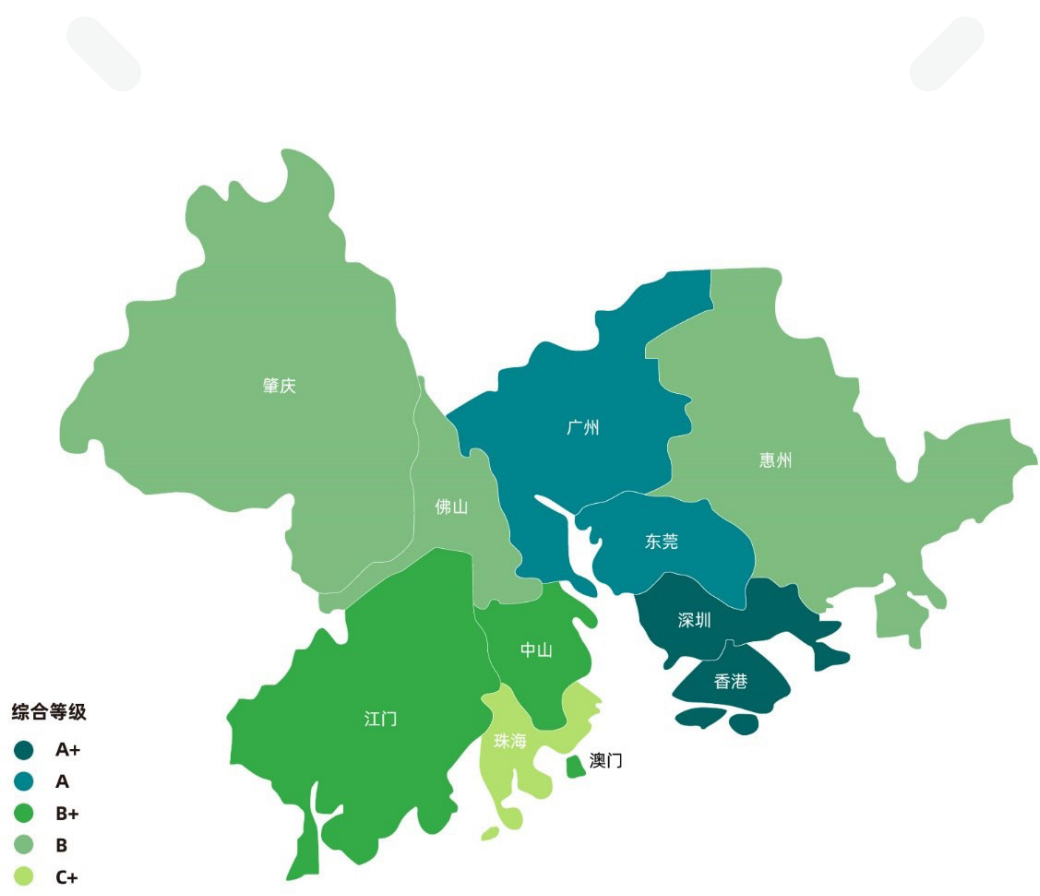


图 4 粤港澳大湾区开放数林指数综合等级的空间分布

四、大湾区数林标杆

本章将重点推介粤港澳大湾区政府数据开放的标杆案例，以供各地进行对标分析。

4.1 准备度

4.1.1 法规政策

深圳制定并公布了地方性法规《深圳经济特区数据条例》，如图 5 所示，其中专门针对公共数据开放作出了要求，如图 6 所示。



图 5 《深圳经济特区数据条例》

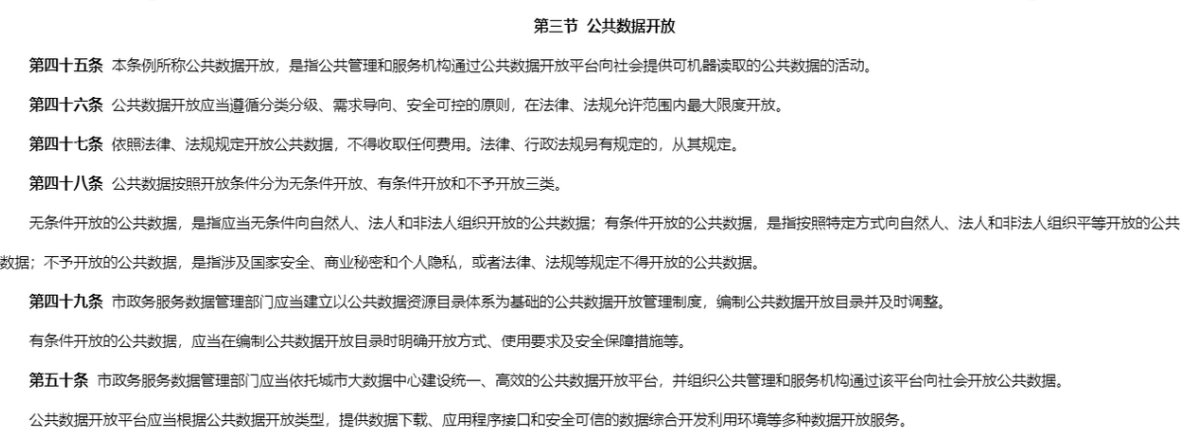


图 6 《深圳经济特区数据条例》中关于“数据开放”的内容条款

4.1.2 组织与领导

香港的年度开放数据计划由特别行政区政府的各部门分别制定，并在香港资料一线通（DATA.GOV.HK）上发布，如图 7 所示。同时，用户还可以在互联网上方便快捷地搜索到各部门的年度开放数据计划，如图 8 所示。



图 7 香港 " 资料一线通 " 首页

香港海关 年度开放数据计划

2020年12月24日 香港海关每年会发放年度开放数据计划,列出未来三年及已于「资料一线通」网站 (data.gov.hk (在新窗口开启))开放的数据。我们乐意听取你对计划的意见,...
www.customs.gov.hk/sc/about_us... 百度快照

香港金融管理局 年度开放数据计划

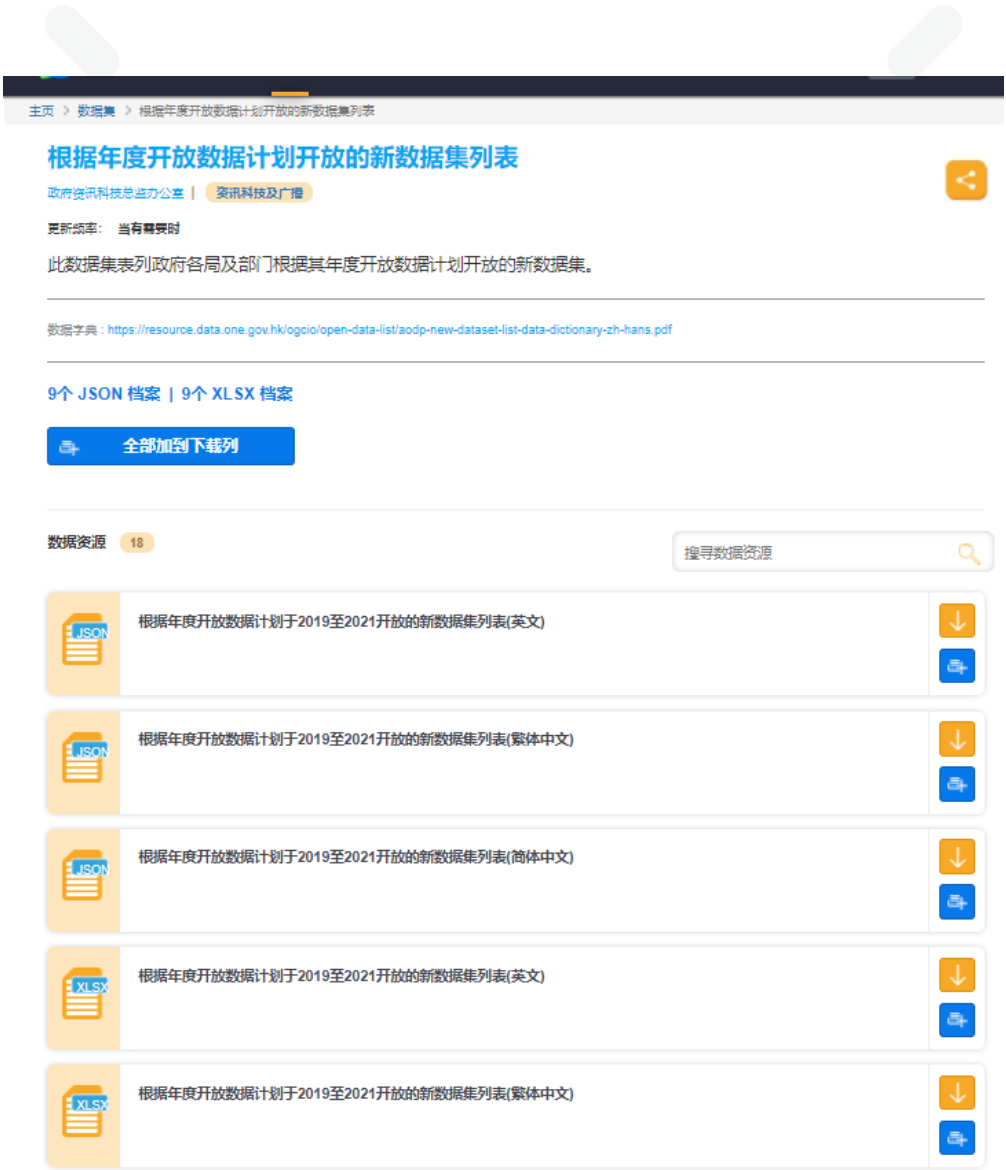
2020年12月24日 香港金融管理局每年会发放年度开放数据计划,列出未来三年及已于金管局网站(hkma.gov.hk)开放的数据。我们乐意听取你对计划的意见,并会参照你的需要制...
www.hkma.gov.hk/gb_chi/other-i... 百度快照

年度开放数据计划 香港天文台(HKO) 关于香港天文台

2020年12月22日 年度开放数据计划香港天文台每年会发放年度开放数据计划,列出未来三年及已于天文台网站及「资料一线通」网站(data.gov.hk)开放的数据。
www.hko.gov.hk/sc/abouthko/ann... 百度快照

图 8 “香港年度开放数据计划”网络搜索截图

在数据集开放计划中，香港在资料一线通平台公布了计划开放的数据集列表，如图 9 所示。例如，香港房屋委员会（房委会）/ 房屋署 在 2021 至 2023 年开放数据计划中列明了已开放的数据集和计划开放的数据集，如图 10 所示。



■ 图 9 香港年度开放计划 - 数据集列表

香港房屋委员会(房委会) / 房屋署
2021至2023年的开放数据计划

A. 在 2021 年发放的部门数据集

#	数据类型 / 数据集名称	目标发放日期	更新频率	备注
1	出售绿表置居计划单位(绿置居)2019 的成交记录	02/2021	一次性	单位成交记录数据将以机读的 JavaScript对象表示法 (JSON) 格式及应用程序编程接口提供。
2	出售居者有其屋计划单位(居屋)2020 的成交记录	另行公布	当资料齐备时 发放	目标发放日期将依据居屋2020选楼期结束而定。 单位成交记录数据将以JSON格式及应用程序编程接口提供。
3	绿置居2020/21的成交记录	另行公布	当资料齐备时 发放	目标发放日期将依据绿置居 2020/21选楼期结束而定。 单位成交记录数据将以JSON格式及应用程序编程接口提供。

B. 在 2022 年及 2023 年发放的部门数据集

#	数据类型 / 数据集名称	目标发放日期	更新频率	备注
1	居屋2021的成交记录	另行公布	当资料齐备时 发放	目标发放日期将依据居屋2021的选楼时间表而定。 单位成交记录数据将以JSON格式及应用程序编程接口提供。
2	居屋2022的成交记录	另行公布	当资料齐备时 发放	目标发放日期将依据居屋2022的选楼时间表而定。 单位成交记录数据将以JSON格式及应用程序编程接口提供。
3	绿置居2022的成交记录	另行公布	当资料齐备时 发放	目标发放日期将依据绿置居2022的选楼时间表而定。 单位成交记录数据将以JSON格式及应用程序编程接口提供。

图 10 香港房屋委员会 (房委会) / 房屋署 2021 至 2023 年开放数据计划

4.2 平台层

4.2.1 高级搜索功能

深圳、肇庆、佛山平台提供了有多个筛选项的高级搜索功能，并同时支持对数据集和利用成果的搜索，如图 11 所示。



图 11 深圳平台的高级搜索功能

4.2.2 可按时间段下载历史数据的功能

香港平台提供了可按时间段下载历史数据的功能，用户可根据特定时间段下载数据集的各个历史版本，如图 12 所示。



图 12 香港平台的按时间段下载历史数据功能

4.2.3 未开放数据请求

深圳、广州平台提供了对未开放数据的请求功能，并对外公开了用户的数据请求和平台的回复，如图 13 所示。此外，深圳平台还对用户的数据请求进行了及时有效的回复，如图 14 所示。



■ 图 13 广州平台公开的未开放数据请求和平台回复



■ 图 14 深圳平台对未开放数据请求的回复

4.2.4 意见建议

香港平台在意见建议提交页面特别提醒用户不要提供任何可识别个人身份的信息，以保护用户的自身信息，如图 15 所示。

意见内容

类型*:

选择类型

提供者*:

选择数据提供机构

数据集名称*:

输入数据集名称

网址:

输入网址

意见*:

输入意见

剩下3000字

关于你的资料

你是*:

☐ 公众人士

☐ 开发人员

☐ 学术界

你的电子邮件*:

输入你的电子邮件

再输入一次电子邮件

☐ 进行人机身份验证


reCAPTCHA
隐私权 - 使用条款

当你提交此表格时，即表示你同意将此表格中所提供的任何资料转发给香港特别行政区政府辖下的局和部门。填写表格时请务必注意不要包含任何可识别个人身份的讯息，例如个人姓名、电话号码、地址、身份证号码等。

提交

重新输入

■ 图 15 香港平台在意见建议提交页面对用户个人信息保护的提醒

广州平台对用户的意见建议进行了及时有效的回复（如图 16 所示），并对外公开了用户的意见建议和平台的回复（如图 17 所示）。



图 16 广州平台对用户意见建议的回复



图 17 广州平台公开的用户意见建议和平台回复内容

4.2.5 数据纠错

深圳平台对用户的数据纠错进行了及时有效的回复（如图 18 所示），并对外公开了用户的纠错和平台的回复（如图 19 所示）。



图 18 深圳平台对用户数据纠错的回复



图 19 深圳平台公开的用户纠错和平台回复内容

4.3 数据层

4.3.1 数据数量

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。截至 2021 上半年，深圳、香港和广州开放的有效数据集的总数最高，均开放了超过一千个数据集（如图 20 所示），其次是中山、东莞、佛山、珠海等地。

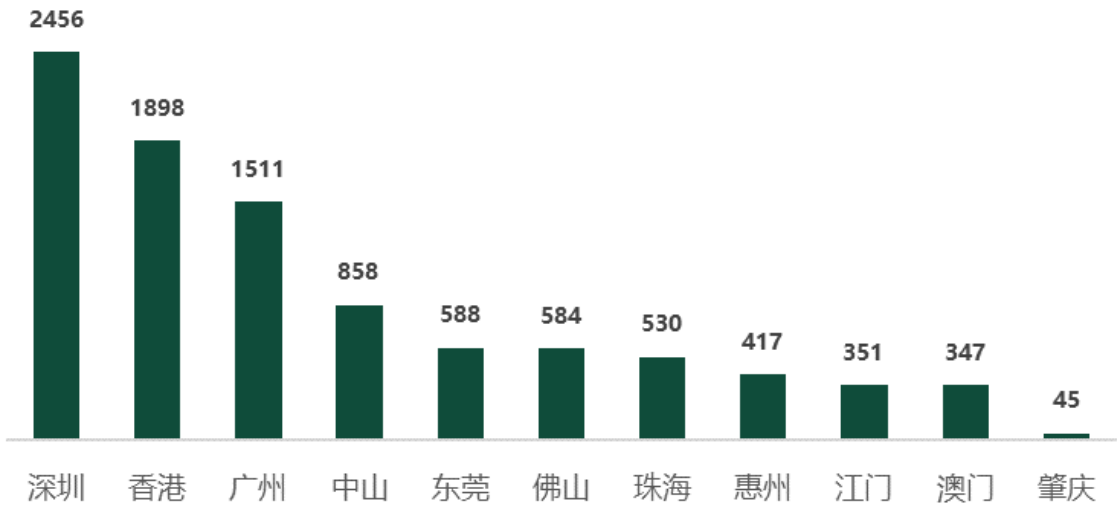


图 20 粤港澳大湾区各地平台有效数据集数量（单位：个）

数据容量是指将一个地方平台中可下载的、结构化的、各个时间批次发布的数据集的字段数（列数）乘以条数（行数）后得出的数据总量。截至 2021 上半年，东莞平台开放的有效数据集的容量最高，达到 2.6 亿，其次是深圳和香港平台，总容量均超过了 5 千万（如图 21 所示）。

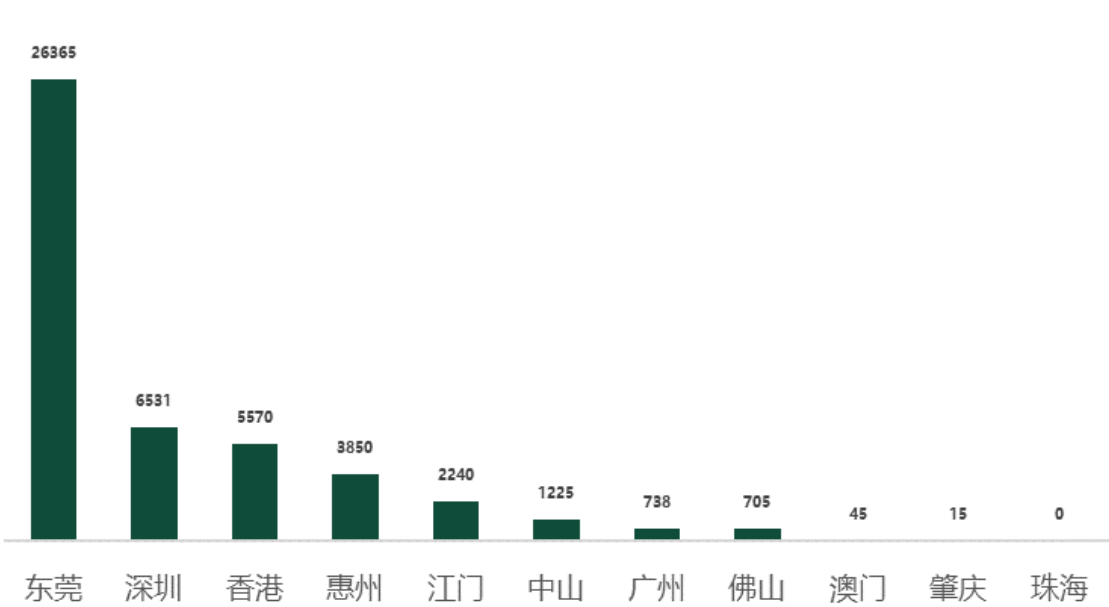


图 21 粤港澳大湾区各地平台有效数据集容量（单位：万）

报告进一步分析了地方开放的单个数据集平均容量，单个数据集平均容量最高的是东莞，平均容量达到 31 万，其次是惠州、江门、香港、深圳等地（如图 22 所示）。

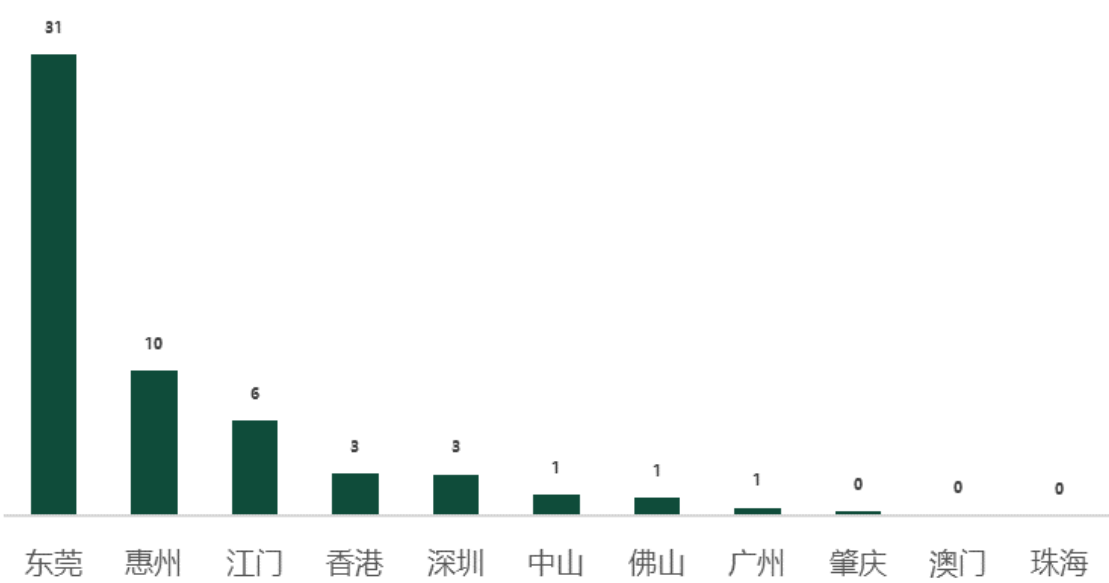


图 22 粤港澳大湾区各地平台单个数据集平均容量（单位：万）

4.3.2 数据质量

数据容量大、社会需求高的优质数据集是开放数据的重点。报告对粤港澳大湾区各地平台上所有可下载的数据集（共约 8610 个）按照数据容量进行排序，最终选出排名居于前 1% 的数据集作为高容量数据集（共约 86 个）。高容量数据集数量的地区分布如图 23 所示，所有的高容量数据集集中在 5 个地方，其中东莞的高容量数据集最多，其次是香港、惠州等地。

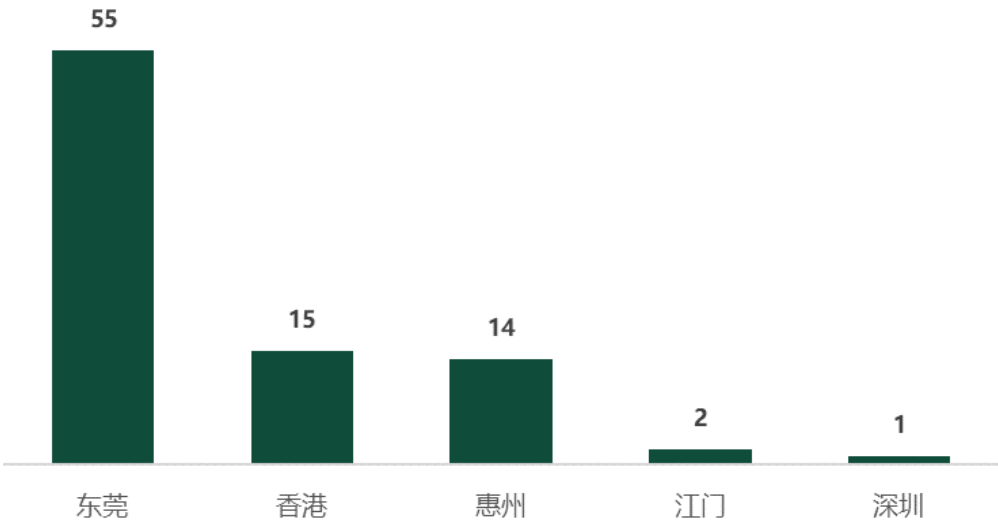


图 23 粤港澳大湾区高容量数据集的地区分布

表 3 是大湾区排在前十位的高容量数据集列表，这些数据集普遍具有较高的条数和字段数，内容涉及面较广，主要涉及工程项目、商事主体、行政许可、交通等方面。表 4 是香港平台提供的前五位高容量数据集，主要涉及地理、交通、环保、选举、研究数据等方面。

表 3 粤港澳大湾区前十位高容量数据集

序号	数据集名称	地方	数据容量	条数	字段数
1	涉企信息_工程竣工验收备案信息	东莞	35125800	2341720	15
2	就业人员信息	东莞	34371848	2148241	16
3	工商登记信息	东莞	32197058	1694582	19
4	深圳公交线路轨迹表	深圳	28401471	4057353	7
5	推送市交易中心信息 - 专业分	东莞	24203544	3025443	8
6	工商行政许可公示信息	东莞	23410785	1377105	17
7	工程施工许可信息	东莞	18239623	868553	21
8	涉企信息_企业不良行为信息	东莞	10671794	762271	14
9	惠州市工商开业登记信息	惠州	9727450	648497	15
10	航空照片飞行索引	香港	7205072	257324	28

表 4 香港平台的前五位高容量数据集

序号	数据集名称	提供部门	数据容量	条数	字段数
1	航空照片飞行索引	地政总署	7205072	257324	28
2	交通运输资料月报	运输署	3222000	537000	6
3	公共政策研究资助计划研究项目的实证数据 (项目编号2013.A6.012.13A) — 我们能否摆脱 香港烟雾弥漫的日子：形成机理与控制策略？	政策创新 与统筹办事处	2476283	72348	34
4	过去海水水质数据	环境保护署	2395458	82602	29
5	2020年选民登记数字	选举事务处	2348352	391392	6

在“无问题数据”这项指标下，高缺失数据集是指数据集中有 60% 以上的空缺数据；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，不利于数据的融合利用；低容量数据是指因数据量本身稀少或颗粒度过大等原因造成的数据条数在三行或三行以内的数据集，其再利用价值较低；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。总体上，存在数据质量问题较少的地方是香港、澳门、东莞、中山、江门等地。

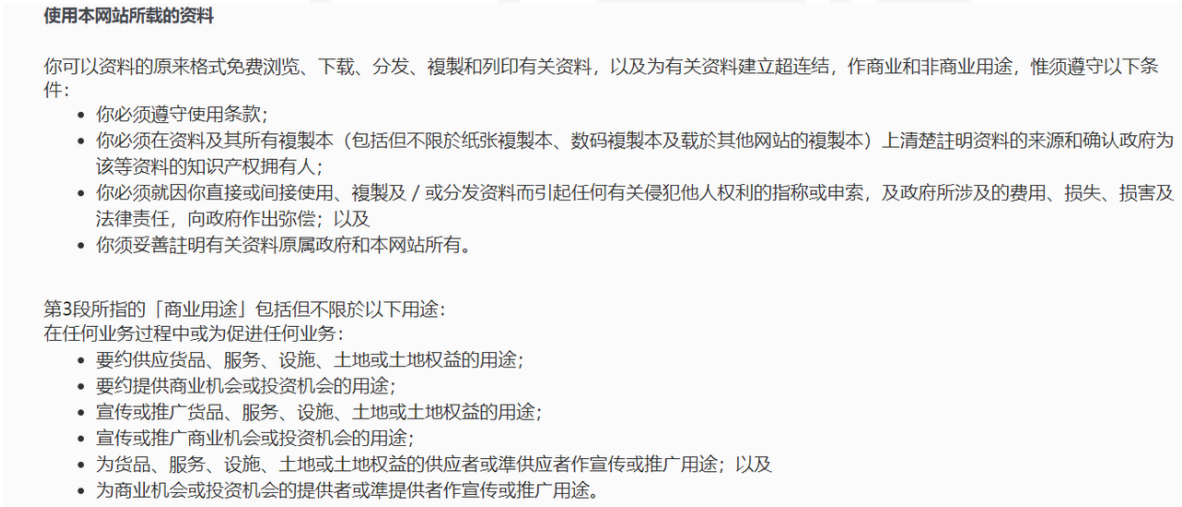
在开放数据持续性上，澳门、中山、东莞能对平台上已开放数据集保持常态化更新；香港平台为每个数据集留存了可按日下载的历史数据（如图 24 所示）。



■ 图 24 香港平台提供的历史数据存档与获取

4.3.3 数据规范

开放授权协议是指以契约的形式确定数据开放主体和数据使用者之间有关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。香港平台在数据使用条款中，对用户须遵守的条件作出了详细的要求，如图 25 所示。



■ 图 25 香港平台提供的使用条款及条件

深圳和江门为平台上开放的所有数据集均提供了可机读格式、非专属格式和 RDF 格式。深圳平台提供了较为规范 RDF 格式（如图 26 所示）。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，每一条数据都是一个唯一的资源 URI，有助于数据利用者发现和链接相关数据资源。

```

<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:szmeta="http://opendata.sz.gov.cn/rdf/1.0#"
  <rdf:Description rdf:about="http://opendata.sz.gov.cn/data/dataSet/toDataDetails/29200_01300931#data3529">
    <szmeta:ZZJGDM>763481754</szmeta:ZZJGDM>
    <szmeta:JYCS>深圳市福田区福华路金田新村2栋305房</szmeta:JYCS>
    <szmeta:JYFW>一般经营范围：日用品、护肤品、化妆品、家居护理用品、厨房用品的销售；信息咨询（不含限制项目）。；许可经营范围：</szmeta:JYFW>
    <szmeta:TYSHXYDM>91440300763481754M</szmeta:TYSHXYDM>
    <szmeta:BZ>（未换照）#预赋码。（因2014、2015年度未年报且未纳税长期停业未经营个人独资企业，市局于2017年06月15日拟吊销）(因2014、2015年度未年:
    <szmeta:QYMC>深圳市安惠贸易商行</szmeta:QYMC>
    <szmeta:HZRQ>2004-05-31</szmeta:HZRQ>
    <szmeta:CLRQ>2004-05-31</szmeta:CLRQ>
    <szmeta:DJJGDH>4403</szmeta:DJJGDH>
    <szmeta:RECORDID>005AA23AD3304ADC935FD138C32EE149</szmeta:RECORDID>
    <szmeta:YYZT>2</szmeta:YYZT>
    <szmeta:ZCH>4403046007324</szmeta:ZCH>
    <szmeta:BIZHONG>156</szmeta:BIZHONG>
    <szmeta:ZCZB>1.000000</szmeta:ZCZB>
  </rdf:Description>
  <rdf:Description rdf:about="http://opendata.sz.gov.cn/data/dataSet/toDataDetails/29200_01300931#data6158">
    <szmeta:HZRQ>2018-04-19</szmeta:HZRQ>
    <szmeta:TYSHXYDM>92440300MA5JTQD3X9</szmeta:TYSHXYDM>
    <szmeta:JYCS>深圳市龙岗区坂田街道南坑村民营路13号</szmeta:JYCS>
    <szmeta:ZCZB>10.000000</szmeta:ZCZB>
    <szmeta:DJJGDH>4403</szmeta:DJJGDH>
    <szmeta:CLRQ>2013-10-29</szmeta:CLRQ>
    <szmeta:QYMC>深圳市龙岗区坂田湘味香人家餐厅</szmeta:QYMC>
    <szmeta:ZCH>440307811140322</szmeta:ZCH>
    <szmeta:BIZHONG>156</szmeta:BIZHONG>
    <szmeta:RECORDID>3df89a65e24c427396da48e93b0112b0</szmeta:RECORDID>
    <szmeta:ZZJGDM>MA5JTQD3X</szmeta:ZZJGDM>
    <szmeta:JYFW>一般经营范围：；许可经营范围：中餐制售</szmeta:JYFW>
    <szmeta:BZ>（未换照）#预赋码。</szmeta:BZ>
    <szmeta:YYZT>4</szmeta:YYZT>
  </rdf:Description>
  <rdf:Description rdf:about="http://opendata.sz.gov.cn/data/dataSet/toDataDetails/29200_01300931#data8688">
    <szmeta:ZZJGDM> </szmeta:ZZJGDM>
    <szmeta:ZCH>440304804295931</szmeta:ZCH>
    <szmeta:CLRQ>2002-05-22</szmeta:CLRQ>
    <szmeta:TYSHXYDM> </szmeta:TYSHXYDM>
    <szmeta:BIZHONG>156</szmeta:BIZHONG>
    <szmeta:JYFW>一般经营范围：电子元器件；许可经营范围：</szmeta:JYFW>
    <szmeta:DJJGDH>4403</szmeta:DJJGDH>
    <szmeta:QYMC>深圳市福田区新群豪仪器仪表工具行</szmeta:QYMC>
    <szmeta:YYZT>4</szmeta:YYZT>
  </rdf:Description>

```

■ 图 26 深圳平台为数据集提供了较为规范的 RDF 格式

香港平台为每个数据集的字段提供了详细的数据字典。如图 27 所示，在香港平台上的开放数据字典中，提供了数据集介绍、数据集内可获取文件的名称、格式和标注，并为每个字段提供名称、数据类别、内容长度、度量单位、备注等信息。

		endTime	纪录完结时间 YYYY-MM-DD'T'hh:mm:ssZ 例子: 2020-09-01T08:19:00+08:00
rainfall	雨量	data	unit: 量度单位
			place: 数据位置
			max ¹ : 最高雨量纪录
			min ¹ : 最低雨量纪录
			Main: 分区站是否维修中 (TRUE/FALSE)
		startTime	纪录开始时间 YYYY-MM-DD'T'hh:mm:ssZ 例子: 2020-09-01T08:19:00+08:00
		endTime	纪录完结时间 YYYY-MM-DD'T'hh:mm:ssZ 例子: 2020-09-01T08:19:00+08:00
icon	图示	以数组表示 天气图示列表: https://www.hko.gov.hk/textonly/v2/explain/wxicon_sc.htm	
iconUpdateTime	图示更新时间	YYYY-MM-DD'T'hh:mm:ssZ 例子: 2020-09-01T08:19:00+08:00	
uvindex	紫外线指数	data	place: 位置
			value: 值
			desc: 强度
			message ¹ : 信息
		recordDesc	资料详情
updateTime	更新时间	YYYY-MM-DD'T'hh:mm:ssZ 例子: 2020-09-01T08:19:00+08:00	

污水处理量数据

此数据集提供 4 间污水处理厂(昂船洲、小蚝湾、赤柱及马湾)之每月污水处理量的数据(GeoJSON 格式)。

数据名称	数据类型	数据描述
月份	日期	污水流量月份 (格式 MM/YYYY)
污水每月总流量 (立方米/月)	数字	4 间污水处理厂（昂船洲、小蚝湾、赤柱及马湾）的每月污水流量总额

排放水水质 (每月几何平均值)

此数据集提供有关昂船洲污水处理厂和马湾污水处理厂之排放水水质的每月几何平均值数据(GeoJSON 格式)。

数据名称	数据类型	数据描述
月份	日期	大肠杆菌计量几何平均值月份 (格式 MM/YYYY)
几何平均值	数字	以几何平均算法表达一个月内每日平均大肠杆菌计量 (个/ 100 毫升)

排放水水质

此数据集提供有关昂船洲污水处理厂和马湾污水处理厂之排放水水质的数据(GeoJSON 格式)。

数据名称	数据类型	数据描述
日期	日期	随意取集排放水样本日期 (格式 DD/MM/YYYY)
随意取集排放水样本中大肠杆菌 (个/100 毫升)	数字	随意取集排放水样本中大肠杆菌总数 (每天于昂船洲污水处理厂 15A 号室和每星期三马湾污水处理厂取集放水样本)

■ 图 27 香港平台为每个数据集的字段提供数据字典

4.4 利用层

4.4.1 利用促进

深圳举办了 2021 全球开放数据应用创新大赛（如图 28、图 29 所示），推动数据跨域跨界融合，参赛者可利用香港、澳门及珠三角其他八个城市开放的来自政府、企业、社会等多种渠道的数据。



■ 图 28 2021 深圳全球开放数据应用创新大赛



■ 图 29 2021 深圳全球开放数据应用创新大赛利用大湾区多地开放的数据

4.4.2 利用多样性

香港平台上展示的利用成果主题覆盖面较广（如图 30 所示），涵盖城建住房、交通出行、卫生健康、社会民生、财税金融等领域。



■ 图 30 香港平台上展示的成果覆盖了多个领域

4.4.3 成果数量

香港平台上展示的利用开放数据产生的有效成果数量最多（如图 31 所示），达到了 22 个。



图 31 香港平台上展示的有效成果

4.4.4 成果质量

香港与深圳平台上展示了相对较多的优质利用成果。如：香港平台上的 Citymapper（如图 32 所示）综合利用了公交路线、巴士到站等数据，为市民提供便捷的交通出行引导服务。深圳平台上的小区罗盘(如图 33 所示)融合利用了交通、文旅等多种主题的开放数据，为市民提供选房指南；



■ 图 32 香港平台展示的 Citymapper 服务应用



■ 图 33 深圳平台展示的“小区罗盘”小程序

五、数林建言

基于报告总体发现和各地先进经验，提出以下优化建议：

5.1 数根：政策供给与组织保障

通过制定政策法规与文件促进和规范政府数据开放，对数据开放要求、数据利用要求、数据开放全生命周期安全管理、保障机制等方面作出规定。

制定和公开政府数据开放的总体计划与方案；地方高层领导加大对政府数据开放的重视和支持力度。

5.2 数千：持续运营与改善体验

提高数据发现和数据获取的便利度，提供和完善高级搜索和数据集预览功能，降低开放数据获取门槛并提供按时间段下载历史数据的功能，加强对未开放数据请求的回复及公开。

增强平台对开放数据利用的支撑作用，完善利用成果提交功能，展示多种利用成果类型及来源信息。

提升互动反馈的时效和质量，提供数据发布者联系方式，公开用户对数据集及利用成果的评价，加强对意见建议和数据纠错的回复及公开，提供权益申诉功能，并提醒用户在互动中注意保护自身的个人信息。

注重平台用户的体验及获得感，完善收藏和推送功能。

5.3 数叶：容量提升与描述说明

提升有效数据集的数量与容量，开放高容量、高需求的优质数据集。

减少高缺失、碎片化、低容量的数据集，保持开放数据集的不断增长和动态更新，并提供历史数据。

为单个数据集提供差异化的开放授权协议，在协议中详尽说明责任义务。

提高开放数据集可机读格式、非专属格式、RDF 格式的比例，为开放数据集提供丰富的元数据信息和详细的数据字典说明。

5.4 数果：数据利用与生态培育

促进数据在开放之后的开发利用，包括综合性的开放数据利用比赛和条线性的引导赋能活动，举办跨区域的开放数据创新利用大赛，在粤港澳大湾区营造有利于政府数据开放利用的良好生态。

提高有效利用成果的数量和质量，清理与数据开放无关的、由政府自身开发的、无法获取或无法正常使用的成果，为展示的利用成果标明其所利用的开放数据集并提供链接。

提高开放数据利用者的多样性，提升利用成果形式与主题领域的多样性。

六、评估方法

6.1 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 2021 年 6 月粤港澳大湾区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，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。

本次评估中，指数出品方共发现符合条件的地方平台 11 个，以此作为本次评估的对象。具体地方、平台名称和平台链接如表 5 所示。

表 5 评估范围（按拼音首字母排序）

序号	地方	平台名称	类别	平台链接
1	澳门特别行政区	澳门特别行政区政府数据开放平台	特别行政区	https://data.gov.mo/
2	香港特别行政区	资料一线通	特别行政区	https://data.gov.hk/sc/
3	广东省广州市	广州市政府数据统一开放平台	副省级	http://data.gz.gov.cn/
4	广东省深圳市	深圳市政府数据开放平台	副省级	https://opendata.sz.gov.cn/
5	广东省东莞市	数据东莞	地级	http://dataopen.dg.gov.cn/
6	广东省佛山市	开放广东-佛山市	地级	https://gddata.gd.gov.cn/data/dataSet/toDataSet/dept/38
7	广东省惠州市	开放惠州	地级	http://data.huizhou.gov.cn/
8	广东省江门市	开放江门	地级	http://data.jiangmen.gov.cn/
9	广东省肇庆市	开放广东-肇庆市	地级	https://gddata.gd.gov.cn/opdata/base/collect?deptCode=518&t=1634143844073
10	广东省中山市	中山市政府数据统一开放平台	地级	http://zsdata.zs.gov.cn/web/index
11	广东省珠海市	珠海市民生数据开放平台	地级	http://data.zhuhai.gov.cn/#/

6.2 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、总体计划与方案、领导讲话的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：一是在搜索引擎以关键词检索相关法律法规与政策文本、总体计划与方案、有关地方党政领导讲话支持的新闻报道；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。评估范围数据采集的截止时间为 2021 年 6 月。

平台层评估主要采用人工观察法对各地方政府数据开放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 2021 年 5 月。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 2021 年 5 月。

数据层评估主要通过机器自动抓取各地方政府数据开放平台上开放的数据，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 2021 年 5 月，对“动态更新”这一指标的评测时段为 2021 年 1 月至 2021 年 5 月。

利用层评估主要对各地方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 2019 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 2021 年 5 月。

此外，为确保数据采集准确全面，对部分指标采用报告出品方自主采集和向地方征集相结合的方式，从各地征集到的公开资料经验证后也纳入评估数据。

6.3 指数计算方法

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 分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：

粤港澳大湾区开放数林指数

$$\begin{aligned} &= \sum (\text{准备度指标分值} \times \text{权重}) \times 12\% + \sum (\text{平台层指标分值} \times \text{权重}) \times 18\% \\ &+ \sum (\text{数据层指标分值} \times \text{权重}) \times 45\% + \sum (\text{利用层指标分值} \times \text{权重}) \times 25\% \end{aligned}$$

指数制作方



复旦大学
DMG 实验室

指数制作团队

郑磊、刘新萍、张忻璐、吕文增、华蕊、张宏、侯铖铖

观察员

蔡伊南、常媛媛、董万怡梦、段晓耀、冯妍、管莹、胡海琛、黄思雅、嵇江夏、姜怀舒、蒋佳钰、景宇欣、李佳欣、李慧瑜、李行行、卢一鸣、莫涛、王芮、王艳梅、王野然、尉苇、吴海珍、肖姝阳、徐丹丹、叶晴琳、于丹、余芊蓉、朱钰谦

报告校对组

纪昌秀、王翔、毛天慧、韩笑、马李滨

联合发布方

复旦大学数字与移动治理实验室
粤港澳大湾区大数据研究院

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
晴禾（南京）文化有限公司
汇纳科技股份有限公司



中国开放数林指数网 ifopendata.fudan.edu.cn

联合发布方:

复旦大学数字与移动治理实验室
粤港澳大湾区大数据研究院

合作方:

冥睿（上海）信息科技有限公司
中山大学数字治理研究中心
晴禾（南京）文化有限公司
汇纳科技股份有限公司



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
复旦DMG
公众号