基于生态位的我国省级公共数据开放平台资源利用状况研究*

李成熙[1,2,3]　吴新年[1,2,3]
（1. 中国科学院西北生态环境资源研究院，兰州 730000；2. 甘肃省知识计算与决策智能重点实验室，兰州 730000；3. 中国科学院大学经济与管理学院，北京 100190）

摘要：公共数据蕴涵着巨大价值，但数据鸿沟阻碍了公共数据进一步的开发利用。深入分析不同省级公共数据开放平台的数据供给现状和用户行为特征及差异，对厘清平台资源与用户需求的匹配关系，实现公共数据价值具有重要意义。从生态学研究视角，引入生态位理论与方法，分析公共数据资源类型，测度公共数据用户生态位宽度和生态位重叠指数，探讨不同省级公共数据开放平台的公共数据资源供给和用户群体下载行为特征与差异，分析省级公共数据开放平台的资源供需关系。研究发现：资源总量与用户群体下载行为之间呈现出非线性关系，资源多样性和需求多样性对用户数据下载行为多样性和均匀性具有显著影响；省域间用户需求偏好差异显著，下载行为多样性和均匀性高的用户需求多样，且平台数据供给效果较好，能较好地满足用户数据需求，而下载行为多样性和均匀性低的用户更关注少数领域公共数据利用，平台数据供给效果欠佳；资源多样性、平台开放倾向共同影响用户下载行为的相似性。

关键词：公共数据；数据开放；生态位；资源多样性；价值共创
中图分类号：F49；G203　DOI：10.3772/j.issn.1673-2286.2025.04.002
引文格式：李成熙，吴新年. 基于生态位的我国省级公共数据开放平台资源利用状况研究[J]. 数字图书馆论坛，2025，21（4）：11-20.

2024年9月，《中共中央办公厅 国务院办公厅关于加快公共数据资源开发利用的意见》发布，标志着中央层面对公共数据资源开发利用的系统性部署正式启动，涵盖扩大供给、规范授权运营、推动数据产业健康发展等17项具体举措，凸显了公共数据开发利用的紧迫性和重要性。公共数据开放作为公共数据开发利用的关键基础环节，发挥着不可或缺的作用。尽管我国政府数据资源占全国数据资源比例超过3/4，但相较于美国，无论是开放规模还是用户实际可利用的规模均存在显著差距[1]，我国公共数据开放仍存在供给不充分等问题[2]。我国已上线的24个省级公共数据开放平台（统计不含直辖市和港澳台）数据资源量持续增长，但面对用户群体较高的数据需求，平台开放程度仍显不足[3]。相关研究表明，政府数据开放平台数据资源利用效率受开放规模制约处于较低水平[4]，这说明当前各平台在资源配置方面存在问题，资源供给与用户需求之间匹配性不足仍是亟待解决的难题。公共数据开放平台作为数据用户获取公共数据的主要途径之一，扮演着匹配供需的枢纽角色。用户在平台上的注册、访问、获取和互动等行为数据反映了用户的数据需求偏好，对这类数据进行深入分析，有助于平台了解用户需求特征，更好地推进公共数据开放。

收稿日期：2025-03-04
*本研究得到国家社会科学基金项目“演化视角下新兴技术形成机制与识别方法研究”（编号：20BTQ094）、甘肃省知识产权计划项目“甘肃省科研院所专利转化平台建设”（编号：22ZSCQ054）资助。

1 文献回顾

目前，我国关于公共数据开放平台资源利用状况的研究成果较少，主要基于用户在公共数据开放平台上生成的多种行为数据，包括访问量[5]、下载量[6-7]、数据集满意度[8]、典型应用数量[9]以及用户在线评论文本[10]等。其中，下载量数据是衡量平台供给方公共数据资源与用户需求匹配性的重要指标。储节旺等[11]从数据下载量（数据集下载率）和应用成果APP数量两个维度评价了合肥市政府数据开放平台的公共数据资源利用效果，得出其处于较低水平；赵龙文等[9]则以用户浏览量、典型应用成果数量、下载量位列前25%的数据资源平均浏览量和平均下载量为核心指标，评价了部分省级和副省级公共数据开放平台的资源利用与创新应用效能。这些研究证明了基于下载量指标来评价公共数据开放利用效果、分析用户数据需求特征的研究方法具有科学性和有效性，但现有研究未能从生态位视角综合探讨不同平台的公共数据资源供给状况，并揭示平台资源多样性与用户需求偏好的关系。

生态位理论作为生态学领域的一个核心概念，由生物学家Grinnell于1917年提出，是生物在生态环境中占据的特定位置[12]，具体指一个物种在特定环境中的生活方式、资源利用能力以及与其他个体相互作用过程中形成的相对地位和作用[13]。在公共数据开放研究领域，学者们从生态学视角讨论了公共数据的生态系统建设[14]、安全保障[15]、价值创造[16]、利用瓶颈[17]以及生态链[18-19]等。这些研究表明，运用生态学视角和理论解决数据开放利用问题已成为研究趋势，生态位理论也逐渐被应用到公共数据开放的研究中。赵佩玉等[20]基于生态位理论，构建了政府开放数据生态链的结构模型，提出了节点演进模式和链式演进模式。Ruijer等[21]则创新性地将战略生态位用于政府开放数据研究，通过干预实验刺激了政府开放数据的使用。但上述研究并未利用生态位测度方法进一步计算分析公共数据开放平台资源多样性特征和用户数据偏好特征，且多聚焦于单一平台分析或少量平台的对比分析，缺乏对不同省级公共数据开放平台用户群体数据需求偏好差异化特征的对比分析。

针对已有研究的不足，引入生态位理论及方法，研究我国各省级公共数据开放平台的资源供给和用户下载情况，旨在更好地理解平台资源利用情况，并为区域间公共数据价值共创提供有益参考。

2 理论基础及适用性阐释

公共数据开放生态系统是一个由多主体参与的复杂生态网络，其核心在于数据的产生、流通、利用与治理。在这个生态系统中，不同省级平台的用户群体类似于物种，公共数据类似于自然资源，公共数据开放平台则可被视为物种获取自然资源的栖息地，用户数据行为则类似于物种在生态系统中的资源利用和生态位占据状态。公共数据用户和公共数据资源就如同自然生态系统中的物种和自然资源，因空间分布差异会呈现出不一样的资源利用行为和资源多样性，并随着需求演变和技术进步而不断变化。

生态位理论中的生态位宽度和生态位重叠概念及方法常用于分析物种对资源的利用情况。生态位宽度用于反映物种或种群对环境的适应状况或对资源的利用程度：如果一个物种所利用的资源占整体可利用资源的比例较少，则该物种的生态位较窄；若该物种能够充分利用其所能利用的资源，则其生态位较宽。生态位宽度的计算方法主要包括Shannon指数、Levins指数和Simpson指数。其中，Shannon生态位宽度指数基于信息熵原理，能很好地反映公共数据用户利用的不同类型数据的多样性和均匀性。在公共数据开放环境中，用户对各类数据的利用情况较为复杂，Shannon指数可以反映用户对不同类型数据资源的利用情况，更全面地描述用户生态位宽度，与研究用户如何广泛且均衡地利用不同公共数据资源的目的相契合。因此，将生态位宽度概念及方法引入公共数据开放研究，可以将公共数据用户生态位宽度理解为用户在各省级平台中对公共数据的需求的多样性和广度，以此评估用户对公共数据开放平台数据资源的利用情况。

生态位重叠指数常用于揭示物种或种群间的竞争关系，能映射出不同用户对同领域公共数据资源利用行为的相似性。生态位重叠的计算方法以Pianka生态位重叠指数为主，该指数在测度不同物种或种群资源利用行为相似性方面具有优势，能准确地量化不同省级公共数据开放平台的用户群体在下载行为和需求偏好方面的相似性。

鉴于此，借鉴生态位理论，通过分析不同省级平台的公共数据资源组成及多样性指数，测度不同区域公共数据用户生态位宽度和生态位重叠指数，不仅可以为理解和分析用户对公共数据的需求偏好以及供给主

体优化数据资源配置提供有力的理论支撑，也可以为公共数据开放政策的制定提供科学依据，从而更好地满足用户需求，促进公共数据价值挖掘和公共数据治理整体效能的提升。

3　数据来源与研究方法

基于生态位理论，从平台、用户和行为3个维度出发，确定研究样本及变量，为系统考察不同省级公共数据开放平台的公共数据资源供给特征及用户群体数据需求特征奠定基础。进一步地，对不同省级公共数据开放平台的公共数据领域分类进行标准化处理，采用Shannon-Wiener指数量化平台资源多样性，该指数能同时衡量公共数据开放平台的资源多样性和分布均匀性；通过用户下载率计算Shannon生态位宽度指数，该指数基于信息熵原理，能很好地反映公共数据用户利用的不同类型数据的多样性和均匀性；利用Pianka指数测度不同省级平台用户群体间的生态位重叠指数，该指数可以反映出不同用户群对的数据下载行为相似性。

3.1　样本选择

通过应用生态位理论，将生态学中的“物种—资源”关系转化为“平台、用户—数据”关系，并从平台、用户和行为3个层级逐渐深入，以确定研究对象及范围。

首先，以省级平台为样本。在复旦大学数字与移动治理实验室发布《中国地方公共数据开放利用报告（省域）》[3]的24个省级公共数据开放平台中，通过人工检索和复核，排除存在网址无法访问、数据难以统计、下载量统计为0和统计数据不清晰等问题的省级平台，最终筛选出15个平台样本（见表1），为全面分析不同经济发展水平地区公共数据开放平台的发展水平及用户行为特征提供了可靠的样本基础。其次，以注册用户为对象。调研发现部分平台的用户行为统计数据存在局限，如未能提供用户注册总量数据，无法按用户类别（如政府、企业、个人等）细分平台使用情况，无法获取个人行为数据。基于上述数据获取的困难，将研究对象限定为各省级公共数据开放平台的所有注册用户。最后，以下载量为变量。平台要求用户在注册并完成实名认证后下载公共数据资源，这一措施在一定程度上筛选了用户群体。这些用户的下载行为数据具备数量庞大、可靠性强及结构清晰等特点，不仅直接反映了用户对特定数据资源的需求，有助于识别用户群体关注的重点领域，也是衡量公共数据开放平台供给成效的重要指标。鉴于此，将用户下载量作为研究变量。

表1　研究选定的15个省级公共数据开放平台

序号	省域	平台名称	网址
1	安徽	安徽省公共数据开放平台	http://data.ahzwfw.gov.cn:8000/dataopen-web/index
2	福建	福建省公共数据资源统一开放平台	https://data.fujian.gov.cn/#/oportal/center
3	广西	广西壮族自治区公共数据开放平台	https://data.gxzf.gov.cn/portal/index
4	贵州	贵州省政府数据开放平台	https://data.guizhou.gov.cn/home
5	海南	海南省政府数据统一开放平台	https://data.hainan.gov.cn//index.htm
6	黑龙江	黑龙江省政务数据开放平台	http://116.182.12.53:8001/oportal/index
7	湖北	湖北省公共数据开放平台	https://data.hubei.gov.cn/#/portal/index
8	江苏	江苏省公共数据开放平台	http://data.jszwfw.gov.cn:8118/extranet/openportal/pages/default/index.html
9	江西	江西省公共数据开放平台	http://data.jiangxi.gov.cn/
10	辽宁	辽宁省公共数据开放平台	https://data.ln.gov.cn/oportal/index
11	内蒙古	内蒙古自治区公共信息资源开放平台	http://open.nmgdata.org.cn/
12	山东	山东公共数据开放网	https://data.sd.gov.cn/
13	山西	山西省公共数据开放网站	http://data.shanxi.gov.cn/
14	四川	四川公共数据开放网	https://scdata.net.cn/oportal/index#
15	浙江	浙江·数据开放	https://data.zjzwfw.gov.cn/jdop_front/index.do

3.2 数据采集与处理

综合采用八爪鱼采集器和人工统计方法，对15个省级公共数据开放平台进行数据采集（时间范围为2024年10月20—30日）。统计各平台的“数据目录/数据集”栏目中的目录总量和用户下载总量。在采集过程中存在平台分类与统计结果不匹配等情况，如福建平台的领域分类中未显示“交通执法”选项，但数据采集结果中却存在“交通执法”主题领域，且相关下载统计数据不完整。考虑到类似情况及数据量较少，剔除无法实现匹配的数据，以确保最终数据分析的准确性和可靠性。

同时，由于各公共数据开放平台的分类标准缺乏统一性，为确保后续分析的可操作性和结果的可比性，对15个平台的分类方式进行比较分析，发现浙江、福建、贵州、广西、山东、湖北和海南等7个平台采用了教育文化、生活服务、医疗卫生等22个领域的标准化划分。因此，以这22个领域为统一分类标准，对其他平台的数据进行分类标准化处理，将相似领域的统计数据归并到这22类中，以保证最终数据的可比性和一致性。

3.3 资源多样性计算

公共数据开放平台的公共数据资源供给直接影响用户群体的数据下载行为。为有效分析不同省级平台的资源供给特征，首先，根据22个领域类别对公共数据资源组成进行统计分析，并计算各省级平台某一领域公共数据资源量占该平台公共数据资源总量的比例；然后，利用Shannon-Wiener指数[22]计算不同公共数据开放平台的资源多样性指数，见式（1）。

$$H=-\sum_{j=1}^{r}D_{ij}\ln D_{ij}\quad(H>0) \tag{1}$$

式中：H表示公共数据开放平台资源多样性指数；D_{ij}表示第i个平台中第j类资源数量占该平台资源总量的比例；r表示样方总数。

利用式（2）计算出15个省级平台中22类公共数据资源的下载率，以反映用户群体对不同领域资源的利用程度。

$$R=\frac{C}{P}\times 100\% \tag{2}$$

式中：R表示用户下载率；C表示用户下载量；P表示资源总量。

3.4 用户生态位测度

基于生态位理论中生态位宽度和生态位重叠等核心概念，进行用户生态位测度。首先，利用式（3）测度用户生态位宽度，取值范围是0~1，值越大表示资源利用率及均匀性越高。其次，利用式（4）测度用户生态位重叠，取值范围是0~1。通过用户生态位宽度和生态位重叠的测度分析，可进一步识别出用户使用数据资源的多样性和均匀性，以及用户群对下载行为相似性的区域性差异，为优化公共数据开放策略（如资源倾斜、分类服务）提供科学依据。

$$W=-\sum_{j=1}^{r}P_{mj}\ln P_{mj} \tag{3}$$

$$O_{mn}=\frac{\sum_{j=1}^{r}P_{mj}P_{nj}}{\sqrt{\sum_{j=1}^{r}P_{mj}^{2}\sum_{j=1}^{r}P_{nj}^{2}}} \tag{4}$$

式中：W为用户生态位宽度；O_{mn}为用户群体m与用户群体n的生态位重叠指数；P_{mj}为用户群体m利用的第j类资源量；P_{nj}为用户群体n利用的第j类资源量。

4 结果与分析

4.1 公共数据资源组成及多样性

15个省级平台的数据目录总量与分类如表2所示。整体上看，福建、贵州、广西、黑龙江和四川等省级平台在公共数据开放总量方面位居前列，表明这些省级平台在资源生态位中占据优势地位，具有较高的资源供给能力。其中，福建数据目录数量最多，高达44 557个，生活服务类公共数据规模尤为突出，体现了该平台在生活服务领域具有较高的资源供给能力。相比之下，山西、辽宁、内蒙古、江西等省级平台的资源总量较少，供给能力弱，资源生态位较窄，在资源供给的多样性上有所不足，这可能制约了用户群体的需求满足和创新应用。同时，部分平台存在某些领域分类中数据量为0的现象，这意味着这些平台在相应领域的资源生态位缺失，可能导致公共数据开放生态系统的资源不完整，无法满足用户数据需求。

表2　公共数据开放平台数据目录总量与分类

单位：个

分类	安徽	福建	广西	贵州	海南	黑龙江	湖北	江苏	江西	辽宁	内蒙古	山东	山西	四川	浙江
教育文化	440	414	926	1 304	27	137	127	33	60	47	13	139	30	334	151
生活服务	57	20 295	642	1 450	77	186	28	7	124	0	61	63	46	520	25
地理空间	1	1 459	54	199	9	20	0	0	0	2	2	14	2	79	15
社保就业	28	373	419	548	3	1 056	3	83	0	7	10	65	0	1 106	5
公共安全	13	257	112	459	0	0	3	0	23	22	21	16	1	202	16
安全生产	19	136	279	503	7	207	33	10	0	0	7	0	0	50	12
医疗卫生	7	292	620	713	4	126	55	12	16	13	17	45	31	617	52
市场监管	5	364	1 142	632	446	241	30	57	0	35	23	58	12	564	95
社会救助	3	398	533	0	1	1	0	0	0	20	8	13	0	106	2
交通运输	39	407	400	753	11	78	10	26	10	2	9	47	9	184	33
气象服务	3	243	101	392	0	7	0	0	0	1	0	22	1	48	5
城建住房	49	1 299	610	1 217	32	0	9	11	25	0	8	9	3	200	36
商贸流通	157	349	132	358	9	0	18	0	0	0	6	22	32	72	39
财税金融	46	1 370	394	896	9	127	27	3	65	0	22	60	10	214	53
法律服务	2	190	215	507	10	0	1	0	0	0	16	0	0	56	9
机构团体	10	929	515	988	19	17	13	0	15	0	19	67	39	82	25
工业农业	44	5 757	767	1 184	80	78	15	27	0	22	15	54	13	260	65
生态环境	18	680	541	885	156	95	64	17	71	11	12	11	14	290	70
资源能源	31	284	224	602	0	109	79	0	0	17	11	158	0	98	17
科技创新	134	0	212	354	26	39	55	0	0	5	11	0	0	42	32
信用服务	0	4 048	231	58	0	247	1	0	4	11	11	21	1	181	45
其他	194	5 013	929	5 143	49	3 112	0	416	33	42	0	0	0	0	105
合计	1 300	44 557	9 998	19 145	975	5 883	571	702	446	257	302	884	244	5 305	907

广西的资源多样性指数最高，内蒙古和浙江分别排在第2、3位，上述省级平台占据了较高的资源生态位。这不仅体现在资源数量上，更体现在资源类型的多样性和分布均匀性上。高生态位的平台能满足更广泛的用户需求，促进数据价值的多元化实现。相比之下，位于中值区（资源多样性指数为2.10~2.65）的省级平台虽然数据资源多样性相对较低，但仍保持了一定的资源均匀性，表明其在公共数据开放生态系统中形成了相对稳定的次级资源生态位，具备基本的资源供给功能。而处在低值区（资源多样性指数低于2.10）的平台（如江苏平台）的公共数据资源类型多样性、均匀性都比较低，这种低生态位可能导致用户需求满足度低、创新应用受限等问题，制约区域公共数据开放生态系统的良性发展。

已有研究认为，公共数据开放总量与数据价值实现呈正相关关系[23]，即开放总量越多，用户可选择资源越多，数据价值实现效率越高。然而，分析15个省级公共数据开放平台发现，资源总量与资源多样性之间存在显著差异：广西平台的资源总量排名并非最高的，但资源多样性指数排在第1位；福建平台虽然资源总量最多，但资源多样性指数排在倒数第4位。进一步分析发现，同时拥有高资源总量和高资源多样性指数的平台（如广西平台），资源数量多，资源类型多样且分布均衡，因此形成了高资源生态位，这种均衡的资源生态位特征能支持更广泛的用户需求满足，促进公共数据开放。相反，高资源总量、低资源多样性指数的平台（如福建平台）存在资源供给与用户需求都集中于少数领域的问题（如福建平台中生活服务类数据占比超过45%），导致资源分布单一，资源生态位受限。在当前公共数据开放平台的资源建设中，部分平台的资源总量和资源多样性并不完全呈同步增长趋势。资源总量所决定的资源生态位“广度”并非平台满足用户需求的唯一决定性因素，资源多样性所构建的资源生态位“丰度”同样至关重要。只有实现资源总量与资源多样性的均衡发展，即达到资源生态位“广度”与“丰度”的协调统一，才能更好地实现公共数据开放，满足用户需求，充分发挥公共数据价值，并服务于社会经济发展。

4.2 公共数据用户下载率分析

根据式（2）计算各省级平台中各领域的资源下载率。数据统计结果表明，生活服务、工业农业、市场监管、教育文化、其他类数据资源的下载率排名前5，说明当前我国省级公共数据开放平台的用户数据需求主要集中在上述领域，在生活服务和工业农业领域的需求尤为突出。

不同省级平台的用户需求存在差异。福建平台用户对生活服务类公共数据的需求最多，下载率为37 114.68%，是贵州平台同类数据资源下载率（687.89%）的53倍。贵州平台用户更关注市场监管类公共数据，下载率为13 945.36%。山东和浙江平台中教育文化类公共数据的下载率最高，分别为4 318.44%和2 378.83%。这种现象不仅反映了公共数据开放生态系统中不同省级平台用户群体的生态位分化，也凸显了省域间公共数据资源多样性、资源开放总量等因素的差异对公共数据用户生态位的塑造作用。另外，江苏、辽宁和江西等平台中其他类数据资源的下载率最高，分别为1 109.54%、1 058.37%和887.22%，这说明了这些省级平台的公共数据领域分类不完善。部分平台分类过于细致，部分则过于笼统，这种分类体系的失衡也在一定程度上影响了公共数据用户生态位的有效形成，易导致公共数据用户生态位宽度受限。因此，如何合理有效地将公共数据资源分类，成为更好地满足用户数据需求的关键问题之一。从资源配置与需求匹配看，公共数据开放与用户需求间仍存在失衡。研究发现，福建、黑龙江、浙江等7个省域的平台中，各领域资源的用户下载率与资源目录总量呈正相关关系，而贵州、广西、四川等8个省域的平台则未呈现这一趋势。可见，在资源开放与用户需求方面，近半数省级平台仍存在一定的匹配失衡问题，主要体现为平台自主开放的数据未能满足用户需求（如安徽平台开放量最多的教育文化类公共数据下载率为0），而用户主动申请开放的数据又难以获批。这种供需失衡限制了用户生态位宽度的拓展，导致用户需求集中于少数领域，生态位功能未能充分发挥。

4.3 公共数据用户生态位宽度

用户生态位宽度用于衡量公共数据用户群体使用公共数据资源的多样性和均匀性。用户生态位宽度指数如图1所示。由图1可知，15个省级公共数据开放平台的用户生态位宽度在0.77以上，生态位宽度指数范围为0.77~2.89，平均宽度为2.03。其中：用户生态位宽度大于2.03的平台有8个，浙江平台的用户生态位宽度最大，其次为广西、内蒙古和山东平台；用户生态位宽度小于2.03的平台有7个，安徽平台的用户生态位宽度最小。用户生态位宽度较高的平台在多个数据资源领域中有较高的下载率，用户需求多样化，平台数据供给效果较好；用户生态位宽度较低的平台则更集中于少数领域的公共数据利用，平台数据供给效果欠佳。

由图1可知，不同省级平台的用户生态位宽度计算结果与资源总量存在显著差异。广西的资源总量（9 998个）远高于浙江（907个），但广西的用户生态

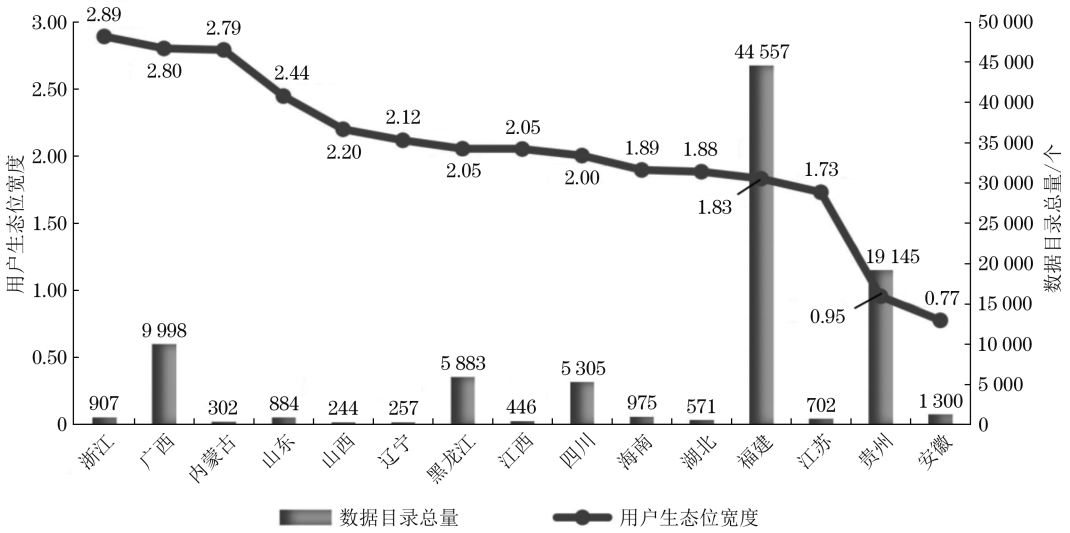

图1 公共数据用户生态位宽度指数与数据目录总量

位宽度指数（2.80）却略低于浙江（2.89）；福建的资源总量（44 557个）最高，但用户生态位宽度指数（1.83）仅排在第12位；安徽的资源总量（1 300个）虽排名第6，但其用户生态位宽度指数最低（0.77）。这些结果表明，在公共数据开放生态系统中，资源多样性和需求多样性对用户下载数据类型的多样性和均匀性有显著影响。具体而言，数据供给与用户需求的良好匹配，以及用户行为的多样化和高活跃度，会不断提升用户下载数据类型的多样性和均匀性；用户需求的集中性和行为的单一性，甚至可能存在的对特定领域资源的过度依赖，都会降低用户下载数据类型的多样性和均匀性。平台只有在资源供给上实现总量与质量的平衡，兼顾资源的多样性和均匀性，同时精准匹配用户多样化的数据需求，才能有效提升用户下载数据类型的多样性和均匀性，提升公共数据的利用效率和价值。

4.4　公共数据用户生态位重叠

资源多样性和用户需求偏好共同影响不同平台之间的用户下载行为相似性。在120对用户群对（两两组合的省级平台用户群体）中，生态位重叠指数在0.01~0.89范围内波动。从聚类结果来看，低生态位重叠区中的安徽、贵州平台的用户群体与其他省级平台用户群体在资源利用上相似性较低，安徽与其他省级平台用户仅有轻微重叠，生态位重叠指数均在0.27及以下。相比之下，贵州与海南（0.70）、贵州与广西（0.52）的生态位重叠指数较高。中生态位重叠区中的生态位重叠指数分布不太均匀，最高达到0.82（浙江与内蒙古用户群对），最低则低至0.26（福建与山东用户群对）。高生态位重叠区中的黑龙江、辽宁、江苏、广西以及江西两两之间的生态位重叠指数都较高，在0.69~0.89范围内。这些计算结果表明，用户生态位重叠指数的高低直接反映了不同用户群体在公共数据利用模式和需求偏好上的相似程度。具体而言，黑龙江与江苏用户群对的生态位重叠指数最高（0.89），生态位重叠现象显著，说明两省级平台中的用户群体在公共数据下载利用行为上相似性很高：①两个平台的资源多样性指数接近，资源领域分布具有较高相似性；②两个平台的用户下载领域相似性较高，用户需求偏好较为一致。生态位重叠指数最低为0.01，涉及安徽与贵州、安徽与山东用户群对，说明两省级平台中用户群体资源利用行为的相似性非常低：①平台间的资源多样性指数差距明显，且资源总量差距过大，资源匮乏在一定程度上会导致低重叠度现象；②用户群对的下载领域相似性较低，用户需求偏好差异显著。由此可以得出，资源多样性和用户需求偏好对用户生态位重叠指数存在显著影响。

不同平台的开放倾向和用户需求偏好共同影响两个平台用户的下载行为相似性。用户下载行为相似性较高的平台之间的平台开放倾向、用户需求偏好都呈现出显著的相似性，即两个平台的各数源部门的数据供给态势和供给效果相似性较高，且用户对这些领域公共数据的需求偏好也高度相似，这为区域间公共数据协同价值共创提供了潜在可能性。而在用户下载行为相似性较低的平台之间，用户需求偏好存在显著差异，平台开放倾向也各具特色，这也为平台差异化布局和协同发展提供了依据。分析用户生态位重叠测度结果，有助于公共数据开放平台及特定领域的数源部门深入分析用户群对的数据行为，识别出用户的需求热点、需求偏好及行为模式，进而调整数据开放策略、优化审核及服务流程、创新开放模式，以更好地满足用户数据需求。

需要注意的是，高重叠指数仅表征需求结构相似性，关于两省级平台之间公共数据资源是否具备可替代性，须结合数据内容具体特征和用户需求进行分析。以学术研究需求为例，通过对各省级平台用户数据需求公示的调研分析发现，本硕博学生群体因学术研究需要而提交数据申请的现象较为普遍。其中，部分研究需求具有显著的地域依赖性，如青藏高原独特的地理数据在生态学、气候学等领域的研究中具有不可替代的科学价值。相比之下，对于通用性较强的研究需求（如经济模型验证），当用户无法从某一平台获取目标数据时，可尝试从其他平台申请同类数据作为替代。此外，用户生态位重叠指数较低的现象表明，两省级平台用户群体在数据需求和下载行为上存在显著差异，这为公共数据开放平台的差异化布局与协同发展提供了重要依据。

5　公共数据开放平台的优化策略

5.1　以用户需求为导向，提升公共数据供给质量

评估公共数据开放平台的数据供给状况时，应综

合考虑资源总量和资源多样性指数，以更全面地反映资源的数量、多样性和均匀性，单纯依靠数据开放总量难以准确揭示不同省级平台在资源组成上的差异。公共数据分类是用户查询、下载、利用数据的关键[24]，但目前各平台在公共数据分类上存在标准不统一、分类精细度不足的问题，给用户查询、下载和利用数据造成不便。因此，平台应进一步优化数据分类体系，提升分类的精细化程度。此外，用户是公共数据开放的重要参与者，公共数据开放平台应重视细化用户分类，并深入分析不同类别用户的需求及数据行为，为提升公共数据供给质量奠定坚实基础。

随着数据开放事业不断发展，各平台积极响应用户的公共数据需求，我国公共数据开放从“政府主导需求”向“用户需求导向”[25]有序转变。然而，单纯增加数据开放量并不能保证用户需求的满足和公共数据的有效利用。因此，公共数据供给主体应以用户需求为导向，提升对高价值数据的识别和确定能力[26]，最大限度地实现公共数据价值。具体而言，公共数据开放平台可以采取两个策略。①建立一套涵盖资源供给、用户需求和数据利用效率等方面的综合评估指标体系，定期评估公共数据开放平台的运行效果，并根据评估结果不断改进供给方案。②基于用户数据需求及数据行为模式，构建公共数据价值评估模型和数据应用追踪反馈机制。综合考虑数据特性、技术、政策等多种因素，优先开放真正有价值且易于利用的数据，定期进行评估和更新，跟踪开放数据的应用情况和社会经济效益，并根据用户反馈动态调整开放策略。

5.2 优化公共数据开放平台生态，拓展用户生态位宽度

资源总量与用户生态位宽度之间呈现非线性关系。虽然更大规模的资源可能蕴含更高的生态位宽度拓展潜力，但若缺乏多样性且未能实现与用户需求的高度匹配，则可能导致资源利用率下降，反而限制生态位宽度的拓展。目前，公共数据用户与公共数据开放平台之间主要呈现互利共生的生态关系，但调研发现部分平台的数据需求响应明显滞后，存在最新数据申请通过时间停留在2022年和平台长时间无反馈等情况。

为提升用户体验和平台服务效能，引入用户生态位宽度指标进行公共数据供需关系评价具有重要实践价值。基于评价结果，不同类型的公共数据开放平台可采取差异化的数据供给策略。①针对资源总量较大或资源多样性较高的平台。一方面，应开展数据资源的深度挖掘与用户精细化运营。深入剖析平台现有数据资源，结合用户生态位宽度及需求偏好，构建多维度的用户画像体系，为不同类别的用户提供个性化数据服务，满足用户的差异化需求。另一方面，建立高效的公共数据开放需求受理反馈机制。组建负责处理用户需求的专业团队，确保及时捕捉用户需求变化，并快速调整数据供给和服务策略。②针对资源总量较小或资源多样性较低的平台。首先，实施对标学习、协同发展策略。加强与用户生态位宽度较高的省级平台的合作，积极借鉴其先进的数据管理经验、高效的平台运营模式以及丰富的用户服务策略。其次，实施精准化服务策略。精准定位核心用户群体，深入分析其生态位宽度及行为特征，集中特色数据资源，优化数据资源供给。可通过建立重点用户服务档案、设置专属服务通道等方式，提供精准化、高质量的数据服务，逐步提升公共数据开放平台的服务能力。

5.3 加强区域数据协同，促进公共数据价值共创

生态位重叠指数高的用户群对的数据下载行为更为相似，平台在多个数据领域的数据供给效果相似；生态位重叠指数低的用户群对，其数据下载行为相似性较低，平台在多个领域的数据供给效果差异较大。目前，公共数据开放平台与本地企业间的省域内数据协同模式和川渝一体化数字资源枢纽底座建设等省域间协同模式，都促进了公共数据价值共创，激发数据价值的协同效应。可见，多主体协同的公共数据价值共创已经成为公共数据价值实现的重要路径之一。

平台开放倾向受到各省域的经济发展水平、城市数字化水平和数源部门的开放意愿及能力等多重因素的综合影响，各省级平台在这些方面的差异为相互学习合作提供了契机。具体而言，用户生态位重叠指数高的平台应积极搭建合作桥梁，如建立区域公共数据联盟或合作平台、达成战略项目合作，挖掘隐藏在高相似性的用户数据行为背后的数据需求趋势和区域性关联，优化资源配置，促进区域协同的公共数据价值共创，实现互利共赢。用户生态位重叠指数低的平台可加

强资源互补，拓展合作空间，通过数据共享机制、跨领域项目合作等方式促进公共数据价值最大化。

参考文献

[1] 构建数据产权、突出收益分配、强化安全治理，助力数字经济和实体经济深度融合：基于对“数据二十条”的解读[EB/OL]. [2024-10-31]. https://www.ndrc.gov.cn/xxgk/jd/jd/202303/t20230317_1351341.html.

[2] 郑磊，刘新萍. 我国公共数据开放利用的现状、体系与能力建设研究[J]. 经济纵横，2024（1）：86-92.

[3] 复旦大学数字与移动治理实验室. 中国地方公共数据开放利用报告 省域（2024）[R]. 上海：复旦大学数字与移动治理实验室，2024.

[4] 杜佳璘，袁永仪，胡广伟，等. 政府数据开放平台数据资源利用效率评估[J]. 图书馆学研究，2024（9）：24-34.

[5] 徐慧娜，郑磊. 面向用户利用的开放政府数据平台：纽约与上海比较研究[J]. 电子政务，2015（7）：37-45.

[6] 曾倩. 政府数据开放平台的用户需求预测模型研究[D]. 成都：成都理工大学，2019.

[7] 王瑞云，贾君枝. 基于用户适用度的开放数据质量提升研究[J]. 数字图书馆论坛，2018（12）：18-26.

[8] 张子良，马海群. 我国政府数据开放平台利用效果比较研究[J]. 数字图书馆论坛，2016（6）：8-15.

[9] 赵龙文，洪逸飞，赵雪琦. 政府开放数据资源利用与创新应用效能评价研究[J]. 现代情报，2023，43（7）：133-144.

[10] 李冠，赵毅. 基于在线评论的政府数据开放平台用户增量需求研究[J]. 数字图书馆论坛，2022（12）：37-46.

[11] 储节旺，张林轩，宫雨晨，等. 合肥市政府数据开放平台建设及发展路径研究[J]. 数字图书馆论坛，2021（2）：11-18.

[12] 曹凑贵. 生态学概论[M]. 北京：高等教育出版社，2002.

[13] GRINNELL J. The niche-relationships of the California thrasher[J]. The Auk，1917，34（4）：427-433.

[14] 王今，邹纯龙，马海群. 公共数据生态系统测度及优化路径研究[J]. 情报理论与实践，2022，45（10）：19-26.

[15] 张哲，王英，马海群. 数字生态视角下公共数据安全保障体系研究[J]. 现代情报，2025，45（2）：106-115.

[16] 郑磊. 开放政府数据的价值创造机理：生态系统的视角[J]. 电子政务，2015（7）：2-7.

[17] 汤志伟，罗意. 开放政府数据生态系统中的数据利用：瓶颈与组态[J]. 数字图书馆论坛，2024，20（6）：12-22.

[18] 张晓娟，莫富传，冯翠翠. 政府开放数据生态链数据流转机制研究[J]. 现代情报，2023，43（1）：91-99.

[19] 赵需要，郭义钊，姬祥飞，等. 政府开放数据生态链上数据要素价值分析及评估模型构建：基于“数据势能”的方法[J]. 情报理论与实践，2022，45（12）：50-59.

[20] 赵佩玉，赵需要. 政府开放数据生态链演进模式及其实现路径研究：基于生态位理论[J]. 西安石油大学学报（社会科学版），2023，32（2）：27-35.

[21] RUIJER E，MEIJER A. Open government data as an innovation process：lessons from a living lab experiment[J]. Public Performance & Management Review，2020，43（3）：613-635.

[22] SHANNON C E，WEAVER W. The mathematical theory of communication[M]. Urbana：University of Illinois Press，1949.

[23] 陈艳利，蒋琪. 数据生产要素视角下开放公共数据与企业创新：基于建立公共数据开放平台的准自然实验[J]. 经济管理，2024，46（1）：25-46.

[24] 时祖光. 公共数据分类分级的基本内涵、形成逻辑与实现路径[J]. 图书馆，2024（10）：15-21.

[25] 迪莉娅. 论我国政府数据深度开放的产生背景、特点与发展策略[J]. 情报资料工作，2024，45（4）：96-103.

[26] 宋烁. 公共数据开放中用户需求导向的高价值数据认定[J]. 求索，2024（5）：176-185.

作者简介

李成熙，男，博士研究生，研究方向：数据治理、公共数据开发利用。

吴新年，男，博士，研究员，博士生导师，通信作者，研究方向：情报理论与方法、知识管理与知识服务，E-mail：wuxn@lzb.ac.cn。

Resource Utilization Status of Provincial Public Data Open Platforms in China Based on Ecological Niche

LI ChengXi[1,2,3] WU XinNian[1,2,3]
(1. Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou 730000, P. R. China; 2. Gansu Key Laboratory of Knowledge Computing and Decision Intelligence, Lanzhou 730000, P. R. China; 3. School of Economics and Management, University of Chinese Academy of Sciences, Beijing 100190, P. R. China)

Abstract: Public data contain great value, but the data gap hinders the further development and utilization of public data. An in-depth analysis of the current data supply situation and user behavior characteristics and differences of different provincial public data open platforms is of great significance for clarifying the matching relationship between platform resources and user demands and realizing the value of public data. This study, from the perspective of ecological research, introduces niche theory and methods, analyzes the types of public data resources, measures the niche width and niche overlap index of public data users, explores the characteristics and differences of public data resource supply and user group download behavior of different provincial public data open platforms, and analyzes the resource supply and demand relationship of provincial public data open platforms. The research finds that there is a nonlinear relationship between the total amount of resources and the download behavior of user groups. Resource diversity and demand diversity have a significant impact on the diversity and uniformity of user data download behavior. There are significant differences in user demand preferences among provinces. Users with higher diversity and uniformity of download behaviors have more diverse demands, and the platform's data supply effect is better, which can better meet users' data demands. However, users with lower diversity and uniformity of download behaviors are more concentrated in the utilization of public data in a few fields, and the platform's data supply effect is poor. Resource diversity and the tendency of platform openness jointly affect the similarity of users' download behaviors.

Keywords: Public Data; Data Openness; Ecological Niche; Resource Diversity; Value Co-Creation

（责任编辑：管清潆）