

政府旅游数据开放的评价指标体系构建与组态分析: 基于21个省级行政区的数据

胡安安^{1,2}, 郭梦珂², 黄丽华^{2,3}, 黄荔桐¹

1. 复旦大学旅游学系, 上海 200433;
2. 复旦大学电子商务研究中心, 上海 200433;
3. 复旦大学管理学院, 上海 200433

摘要

在信息时代和智慧旅游背景下, 政府主动对外开放内部的旅游数据, 有利于充分释放旅游数据资源的商业价值与社会价值。使用变异系数法与熵权法进行组合赋权, 构建了政府旅游数据开放评价指标体系, 计算21个省级行政区的具体得分, 并采用模糊集定性比较方法进行高评价值与低评价值条件组态分析。结果表明: 指标体系包含4项一级指标、17项二级指标与51项三级指标, 可以对当前省级行政区的旅游数据开放绩效进行有效评价; 利用层、数据层是权重较高的一级指标, 法律政策效率与内容、平台关系等是权重较高的二级指标; 高评价值组态为综合发展型、数据辅助应用型, 低评价值组态为政策与利用短板型、平台与数据问题型。研究结果可以为旅游数据开放评价相关学术研究与管理实践提供参考。

关键词

旅游数据; 数据资源; 数据开放; 信息构建; 评价体系; 组态分析

中图分类号: F590

文献标志码: A

doi: 10.11959/j.issn.2096-0271.2023070

Research on the evaluation index system of government tourism data opening: based on the data of 21 provincial administrative regions

HU Anan^{1,2}, GUO Mengke², HUANG Lihua^{2,3}, HUANG Litong¹

1. The Department of Tourism, Fudan University, Shanghai 200433, China
2. E-Commerce Research Center of Fudan University, Fudan University, Shanghai 200433, China
3. School of Management, Fudan University, Shanghai 200433, China

Abstract

In the context of the information age and smart tourism, the government's initiative to open up internal tourism data is

conducive to releasing the commercial and social value of tourism data resources. This paper used the coefficient of variation method and the entropy weight method for combined weighting, constructed an evaluation index system for government tourism data opening, calculated the scores of 21 provincial administrative regions, and used fsQCA to compare high and low evaluation values with configuration analysis. The results showed that the index system included 4 first-level indicators, 17 second-level indicators and 51 third-level indicators, which effectively evaluated the performance of tourism data opening in provincial administrative regions; the utilization layer and data layer were the first-level indicators with higher weights, legal policy efficiency, content, platform relationship, etc. were second-level indicators with higher weights; high evaluation values were configured as comprehensive development type and data-assisted application type, and low evaluation values were configured as insufficient policy & utilization type and platform & data problem type. The research results are helpful to provide a reference prototype for the research and practice of tourism data opening evaluation.

Key words

travel data, data resource, data openness, information construction, evaluation system, configuration analysis

0 引言

政府数据开放(open government data)是电子政务的一个子集,指政府主动向社会公众免费开放透明、可重用、标准化、持续更新的数据^[1],包括公共行政部门直接拥有的数据,如财政预算、人口普查等数据,以及其间接拥有的数据,如旅游、气候、污染等数据^[2]。政府数据开放的典型实现方式是将数据集发布到中央数据门户,为数据消费者提供数据服务的“一站式平台”,提高信息透明度、政府公信力、公众参与度和数据重用度^[3-4]。

国内针对政府数据开放评价的研究多以省级行政区为研究对象开展。省级行政区是中国地方政府数据开放工作最主要的推动力,也是制定行业发展规划、汇集行业数据资源的核心主体,将其作为研究特定行业的政府数据开放评价切入点,具备可操作性,在医疗卫生、交通等领域已有研究成果。

中国旅游行业同样迈进创造数据、获取数据、运用数据的“数据时代”,旅游数据开放工作愈发重要。然而,目前国内外针对旅游数据开放的研究仍较为薄弱,缺少

聚焦。国务院2015年8月印发的《促进大数据发展行动纲要》提出优先推动信用、交通等民生保障服务相关领域的政府数据集向社会开放;中共中央办公厅、国务院办公厅2016年2月印发的《关于全面推进政务公开工作的意见》同样提出优先推动民生保障、公共服务和市场监管等领域的政府数据向社会有序开放。但目前,旅游领域数据开放尚未形成突破。能否针对旅游行业特征与数据需求,制定一套适用于旅游领域的政府数据开放评估体系?各省级行政区的旅游数据开放工作实际进展如何?有哪些可落地的对策建议?探索以上问题,道阻且长。

1 文献综述

1.1 数据开放评价相关研究

数据开放旨在推动政府提供完整、公开、免费的公共数据^[5]。在数据开放政策逐步落地、开放数据得到应用与推广的同时,数据开放的评价体系也在不断完善。2007年12月,开放政府数据的八大基本原则被提出:完整的、原始的、及时的、可

公开获取的、可机器读取的、非歧视获取的、非专属的、免于授权许可的。学术界也给出了具体评价要求。Evans等^[6]指出政府数据开放需要解决数据格式异常、技术术语滥用、用户指南缺失等问题。Janssen等^[7]描述了组织与个人在政策、法律、文化领域对不同开放数据的需求,提出防止侵犯用户隐私的基本要求。Attard等^[8]提出了政府数据开放生命周期模型,区分了用户从获取数据到自主使用数据的递进阶段。Victoria等^[9]建议政府数据发布平台明确区分开放数据和普通信息,以便公众找到结构化、细颗粒度、高价值的开放数据集。现阶段关于政府数据开放评价体系的研究仍在起步阶段,核心还是围绕着完整度、准确性、可机读性等几项指标展开。

在旅游行业,虽然相关研究较少,但数据开放工作的重要性不容小觑。旅游市场的体量较大,诸如客流量、出入境游客占比、消费额等众多数据难以由市场各方实现准确统计;同时,旅游活动对数据的时效性、季节性的要求较高,较之其他行业,旅游行业存在着较大的数据开放缺口。

1.2 旅游数据开放与服务相关研究

在旅游数据开放的研究中,国内外学者大多着眼于数据的最终应用场景,如Eleonora等^[10]指出开放旅游数据有利于地方旅游企业预测游客需求,优化旅游广告,生成餐饮、住宿、交通的精准营销方案。李萌^[11]提倡通过深度挖掘旅游大数据,让旅游公共服务实现以需定产,从“被动管理”转为“主动管理”。与此相比,聚焦旅游数据开放指标建构、旅游数据资源采集、旅游数据开放门户网站效果等方面的研究成果较少,侧面反映出旅游数据开放指标体系构建仍有较大的探索空间。

在对旅游政务网站服务的评价中,目前代表方式有3类:一是针对网站建设工作的测评,主要集中在网站内容和结构的评价,这种模式可以得到量化参数,侧重在技术层面^[12];二是基于服务水平的评价方式,容易施行且具备权威性,但在各地旅游业态不同的情境下,较难得出统一的评估结论;三是基于公众满意度层面的测评,可以反映用户对旅游网站信息的需求和建议,但在评估标准方面不够直接和清晰。

在已有研究的基础上,业界和学界主要将“网站建设”“政务水平”“公众满意度”三者设为核心评估标准,不仅可以反映某一地区旅游电子政务平台的建设水平,而且只需加以修改就可以推出数据开放功能,呈现较为系统的数据开放模型结构。

2 理论基础

2.1 信息构建理论

信息构建理论(information architecture)由美国学者Wurman提出,包括信息环境、信息内容、信息用户三大维度。其一,信息环境由信息开发商、政府政策、技术基础设施、业务流程、质量准则、标准化协议等构成。其二,信息内容是获取和处理信息的过程中对应的模块,如数据库、文献、应用程序等。信息内容既包括内容质量、数据结构、目标主题覆盖率等,也涉及内容处理效率、更新速度、收录数量等。其三,信息用户是信息需求的提出者,基本目的是获取和利用信息。从信息构建理论出发,由专家团队建立一致的评估标准和准则,进行启发式评估,使信息具备可理解性、可用性、易获取性,有利于用户从海量、复杂的信息集合中快速识别、利用信息。

2.2 “开放数林”评价框架

“开放数林”评价框架全称为“中国开放数林指数”，是复旦大学数字与移动治理实验室在2017年提出并持续更新的一套指标体系。该评价框架以树木的生长部位来类比、解析政府数据开放工作的各个维度，如将数据开放的准备基础比喻为树根，将数据利用比喻为树果。前者侧重评价政策的完备性，后者更关注数据利用成果的产出率与有效性^[13]。在数据采集方面，框架提出了“政务+数据”“地名+政府数据”等关键词检索方式，为研究者获取评价数据提供了范式^[14]。“开放数林”评价框架自首次发布以来，每隔半年进行一次更新，基于专家团队意见更新评价指标，动态适应行政管理领域的发展方向。

“开放数林”评价框架被广泛用于省级、市级行政区政府数据开放工作的绩效评估中。在评价对象的类型划分上，“开放数林”评价框架将省域归为一个整体，将直辖市、副省级市、地级市归为城市，主要原因在于省域政府在数据治理上对下辖地有着统筹管理、规范引导的作用，而城市是人口集中的密集“聚落”。在具体评分操作上，“开放数林”评价框架通过样本和指标对比、区域趋势分析等方式反映开放数据发展态势，从多个维度对政府数据开放绩效进行评价。

3 研究方法

3.1 变异系数法

变异系数法是一种基于统计学的客观赋权法，其指标权重由变异系数，即标准差与平均数的比值决定。根据变异系

数法的性质，统计学意义上变化差别越小的指标被赋予的权重越小；反之，统计学意义上变化差别越大的指标的权重将越大。

3.2 熵权法

熵权法是一种基于指标信息量大小以确定指标客观权重的赋权方法，指标值的离散程度越大，信息熵越小，说明该指标提供的有效信息量越大，在综合评价中相应权重越高。熵权法赋权有两大优势：一是避免人为因素对指标权重的过度干扰；二是通过计算熵值衡量信息量多寡，确保所建立的指标能反映绝大部分的原始信息。

3.3 模糊集定性比较分析

模糊集定性比较分析(fuzzy set qualitative comparative analysis, fsQCA)与传统的回归分析方法不同，不囿于单个条件变量的“独特效应”，而是采用集合概念处理变量间的关系，遵循布尔代数的运算原理，分析条件变量的不同组合对结果变量的影响。组态是指条件变量的组合。在处理复杂因果问题上，模糊集定性比较分析可以还原特定结果出现的条件组态，探究组态间的等效性，并对导致或不导致给定结果的组态进行单独分析。

4 政府旅游数据开放的评价指标体系设计

4.1 具体指标设计

本文提出的各项指标不同于现有数据

开放评价体系下的一般性概念,而是结合数据开放视角与旅游行业视角的复合型概念,即:一方面考察各省级行政区在基础、通用的数据开放环节的建树;另一方面考察其对旅游数据开放这一核心点的支持程度,关注其数据开放措施对旅游行业发展、旅游数据应用场景挖掘的有利性。

基于信息构建理论和“开放数林”评价框架,依据信息的可理解性、可用性、易获取性原则,本文确立了准备度、平台层、数据层、利用层4个一级指标。在对一级指标进行细分的基础上,进一步设计了17个二级指标,用于衡量政府在旅游数据开放某一个特定的工作方向上的综合表现。三级指标则对应政府旅游数据开放中涉及的各项基础工作,是评价指标体系中最具体、颗粒度最细的指标。通过拆解二级指标代表的工作类型,并结合数据开放评价研究、旅游行业特点进行设计,最终得到51个三级指标。

为突出旅游行业特色,三级指标的设立遵循5项原则。第一,旅游相关应用场景。该指标是指政府在法律、政策中支持的与旅游相关的开放数据应用场景数目。第二,数据资源目录。该指标是指政府数据开放平台提供旅游数据资源目录的水平。基础目录应具有分类、排序等功能,高级目录则应将专题嵌入导航按钮。比如平台能单独设立“旅游住宿”“智慧旅游”等目录,方便旅游数据需求者直接定位到网站上的开放数据集,降低信息检索成本。第三,优质数据集。参考学界现有标准,本文将优质数据集定义为“数据开放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,排名居于前10%的数据集”。第四,无质量问题。该指标是指可下载或通过接口获取、数据量大于1、字段数和记录数基本涵盖了标题的主要内容,且不存

在内容或格式错误等问题的数据集数量。第五,旅游相关关键数据集。关键数据集反映了具体数据项对于特定行业研究的价值,应至少包含一类以下数据项。

- 人口统计学特征类:如客源地、年龄等。
- 企业绩效类:如满意度、网站打开率等。
- 特色专题类:如老年旅游、康养旅居、旅游小镇等。
- 时间类:如日期、时刻等。
- 空间类:如经度、纬度等。

4.2 评价框架与规则设计

在三级指标评分机制层面,本文重点解决“各指标的具体表现如何转化为数值型得分”的问题,步骤如下。

- 指标类型划分:设置“数值型”与“等级型”两类指标,前者通过公式计算或用特定数据直接折算。
- 指标级别划分:“等级型”指标有5个级别,得分对应1到5分,并设置0级“无表现”,即“从未开展任何相关项目”“未制定任何计划或规范”的无活动状态。
- 关键实践设计:确定旅游数据开放达到某一级别所要求的关键实践,用定量或定性形式表述,遵循科学性(由低层次发展到高层次,且均可实现)、系统性(在相同级别上水平相当)、可操作性(可直接衡量)。

已有研究中,针对数值型指标的评分规则包括是否满足特定条件、特定事件的发生频数或频率、表述与规章的一致程度、定额目标的达标比率等^[15-16]。通常在以下3种条件存在时采用数值型指标:①指标具有定量特征,且数值可通过直接的观测、度量或计算得到;②根据定义,指标不适于拆分为多个等级;③样本在指标的表

现差异过大,难以形成统一等级标准。据此,本文将存在上述条件的23个三级指标设为数值型指标,结合“开放数林”评价框架设定计算式。

对于余下的28个等级型指标,评分标准设定如下。

- 准备度三级指标。政策规划类指标的关键实践从级别1到级别5可概括为:只制定个别相关性较低的政策、已制定一些相关政策、政策得以实施、政策得到推广、政策得以评估并不断优化^[17]。准备度指标大多围绕法规、政策的制定与领导组织展开,故本文以此为基础设定评分标准。

- 平台层三级指标。对于数据开放平台功能有关的关键实践,从级别1到级别5的进阶则可被看作某项功能从雏形到基础再到优化,不断被赋予高级化体验的过程。除C17“旅游开放数据目录”更适合根据旅游在平台目录内容中的独立性来判断,其他12个功能型指标都可采用这种功能优化的视角,将基础功能设为较低等级,将增值型或个性化功能设为较高等级。

- 数据层三级指标。在数据质量管理相关研究中,等级的“进阶”主要体现在与特定标准的对标程度上。指标C38可以根据“开放数据五星标准”进行评分标准设定。指标C36“数据持续性”和C37“开放协议”暂无通用的规范标准,因此分别根据数据更新的频率和开放协议内容的完备性设定评分标准。

- 利用层三级指标。指标C43“数据竞赛举办”可以通过竞赛与旅游的关联度来设定评分标准,指标C44“赋能活动引导”和C51“服务应用质量”则分别根据数据利用者权限的高低和服务应用在产品设计中从基础到升级的路径设定。

据此,政府旅游数据开放评价指标体系见表1,接下来将对其进行指标赋权,得到完整体系。

4.3 数据采集

本文于2021年11月1日先采用“数据+开放”“数据+公开”“政务+数据”“省级行政区名称+政府数据”等关键词检索省级行政区政府数据开放平台,并依据国际通用的数据门户命名规则“data.gov”进行域名筛选。经测试,平台均为统一专有式门户或统一嵌入式门户,符合数据开放平台标准,但部分平台存在无法访问或未开放旅游数据集的情况。为确保筛查结果客观、准确,对于存在问题的网站,本文于2021年11月8日、11月22日、12月6日再次访问排查,尝试切换操作系统、网络状态、浏览器后重新访问,对始终无法访问或显示未开放旅游数据集的省级行政区予以剔除,最终获得21个省级行政区(北京市、福建省、广东省、广西壮族自治区、贵州省、海南省、河南省、湖南省、吉林省、江苏省、江西省、辽宁省、宁夏回族自治区、山东省、陕西省、上海市、四川省、天津市、香港特别行政区、浙江省、重庆市)作为研究对象。

根据各类指标原始数据的特性,本文采用针对性的数据采集方法,如下。

- (1)对准备度相关指标采用文本检索法综合平台政策模块检索、专业数据库检索(“北大法宝”“国家法律法规数据库”等)及通用搜索引擎检索(必应、百度等)3种模式,对相关法律法规、政策、授权协议、使用声明等内容进行搜索与提取。根据两项筛选条件剔除无效信息:①发布机构应为省级行政区政府;②文本内容至少在1处体现“开放数据”字样,不能只有其相关主题或子主题。

- (2)对平台层、利用层相关指标(大多只能人工识别,无法自动抓取)采用人工观察法和人工测试法,并记录判断依据

表1 政府旅游数据开放评价指标体系

一级 指标	二级 指标	三级 指标	等级度型三级指标在各成熟度级别对应的关键实践 (或数值型三级指标的计算公式)				
			I级	II级	III级	IV级	V级
A1 准 备 度	B1 法律 政策 效力 与 内 容	C1数据开 放要求	简述数据 开放范围	要求开放旅游相关 领域(如交通)数据	要求开放 旅游数据	要求优先开放 旅游数据	编制旅游数据 目录,完成旅 游数据分类
		C2数据利 用要求	简述禁 止事项	规范产品开发	要求签订 利用协议	制定利用追踪机制	建设开发利 用平台
		C3全生命周 期安全管理	数据分类 分级开放	指定安全主管部门	建立应急 处置机制	建立风险评估机制	建立监测预 警机制
		C4保障机制	简述禁 止事项	保障数据隐私	保障成果权益	开设维权通道	建立申诉反 馈机制
		C5数据基 础设施	【数值】该省份法律、政策中要求设立的服务于数据开放的数据基础设施 (如存储、服务器、网络、数据库、大数据系统等)类型数				
	C6旅游相关应 用场景支持	【数值】该省份法律、政策中支持的旅游相关应用场景数					
	B2 标准 规范	C7标准规 范等级	规划、意 见、指南	部门规范性文件 (各部门制定的)	规范性文件 (市政府等 制定的)	地方政府规章	地方性法规
		C8数据标 准规范	规范数 据质量	规范数据格式	规范元数 据管理	制定数据校核机制	制定定期维 护机制
		C9平台标 准规范	要求设 立平台	规范平台数据(要求 平台开放旅游数据)	规范平台功能	规范平台整 合与互通	制定平台评 估指标
	C10流程标 准规范	【数值】该省份法律、政策中已制定规范的数据开放流程数					
B3 组 织 与 领 导	C11统筹管 理机制	简述主管 部门职责	设立专项小组	设立专门机构	要求统筹管理	制定协调机制	
	C12跨域协作	【数值】该省份数据开放跨领域合作对象类型数					
	C13领导重视	【数值】该省份与数据开放有关的领导公开讲话次数及会议举办次数					
	C14年度工 作计划	制定单 项计划	制定多项计划	计划内容 可量化	形成明细清单	建立绩效监督 或公示机制	
A2 平 台 层	B4 平台 关系	C15省域整体性	【数值】该省份数据开放平台已整合的下一行政级别的数据 开放平台数/该省份的下一行政级别地区总数				
		C16同业互通性	【数值】该省份数据开放平台已链接的其他省份、其他行业及其他类型的数据开放平台数				
B5 发 现 预 览	C17旅游开 放数据目录	有目录, 但没有任 何旅游相 关目录	有文旅局目录	有文化休闲等 旅游相关目录	有带有旅游字 样的目录但与其 他行业合并,如 “旅游交通”	有为旅游单独 设立的目录, 如智慧旅游	
		C18搜索功能	提供搜 索功能	搜索结果排序	高级搜索		
		C19数据集 预览功能	展示数 据项	预览数据详情	数据可视 化分析	推荐相关数据	提供数据图谱

续表

一级指标	二级指标	三级指标	等级度型三级指标在各成熟度级别对应的关键实践 (或数值型三级指标的计算公式)				
			I级	II级	III级	IV级	V级
A2 平台层	B6 数据获取	C20无条件开放数据获取	有该功能,但访问受限	有该功能,但需要申请	登录账号即可获取	无需登录账号即可获取	无需登录账号且可批量获取
		C21有条件开放数据申请	有该功能,但访问受限	要求明确使用用途	要求选择使用时间	要求填写安全保障措施	要求阅读许可说明或签署申请协议
		C22未开放数据请求	有该功能,但访问受限	要求描述数据需求	要求选择数据格式	要求选择更新频率	提供请求示例或填写指导
	B7 成果提交展示	C23利用成果提交功能	有该功能	支持提交多种形式利用成果	展示相关协议	提供开发者文档	介绍或提供开发工具
		C24利用成果展示	有该功能	可分类	可排序	可搜索	可下载
	B8 互动反馈	C25数据发布者联系方式	有发布者,但没有具体到部门	告知部门,但没有联系方式	提供联系方式(邮箱、地址)	有效链接	实时对接入口
		C26用户评价	电话、邮箱向平台反馈	表单反馈	数据集评分评价	用户评价公开	用户评价答复
		C27意见建议	电话、邮箱向平台反馈	表单反馈	专门版块反馈	意见建议公开	意见建议答复
		C28数据纠错	电话、邮箱向平台反馈	表单反馈	专门版块反馈	纠错公开	纠错答复
	B9 用户体验	C29权益申诉	电话、邮箱向平台反馈	表单反馈	专门版块反馈	申诉公开	申诉答复
C30账号互通性		【数值】该省份数据开放平台提供的互通账号类型数					
C31个性化服务		【数值】该省份数据开放平台提供的个性化服务类型数					
A3 数据层	B10 数据数量	C32数据集总数	【数值】旅游开放数据集总数				
		C33单个数据集平均容量	【数值】所有旅游开放数据集的行数*列数/旅游开放数据集总数				
	B11 数据质量	C34优质数据集	【数值】优质旅游开放数据集个数/旅游开放数据集总数				
		C35无质量问题	【数值】无质量问题旅游开放数据集个数/旅游开放数据集总数				
		C36数据持续性	显示发布时间、更新时间	持续发布新数据集	持续更新现有数据集	承诺更新频率,但更新频率包含“自定义”或“不定时”	承诺更新频率,且更新频率不包含“自定义”或“不定时”

续表

一级 指标	二级 指标	三级 指标	等级度型三级指标在各成熟度级别对应的关键实践 (或数值型三级指标的计算公式)				
			I级	II级	III级	IV级	V级
A3 数 据 层	B12 数 据 范 围	C37开放协议	明确违 约条款	规范数据用途	展示数据 利用主体权 利及义务	展示数据责任主 体权利及义务	包含被申请数 据集的信息
		C38开放格式	使用任 意格式	使用结构化数据 (如xls格式)	使用开放格式 的结构化数据 (如csv、xml 或json格式)	使用可关联数据 (如url链接)	使用关联数据 rdf格式(除提供 url链接外还可 链接其他数据)
		C39描述说明	【数值】数据描述说明中包含的元数据类型数				
	B13 开 放 范 围	C40部门覆盖	【数值】旅游开放数据集的数源部门类型数				
C41旅游相关常 见数据集覆盖		【数值】常用旅游数据集个数					
A4 利 用 层	B14 利 用 促 进	C42旅游相关关 键数据集覆盖	【数值】关键旅游数据集个数				
		C43数据竞 赛举办	赛题或赛 事数据集, 包含少量 旅游元素	开放旅游相 关数据集	开设旅游 相关赛道	开设旅游专题赛道	举办旅游专 题赛事
	C44赋能活 动引导	优化开 发环境	表彰优质成果	开拓应用场景	授权成果交易	授权外部运营	
	B15 利 用 多 样 性	C45旅游相关数 据利用者多样性	【数值】利用者类型数				
C46旅游相关成 果形式多样性		【数值】利用成果类型数					
B16 成 果 数 量	C47旅游相 关成果数量	【数值】展示在旅游主题目录下、名称带有旅游、标签含 有旅游或主要功能涉及旅游的利用成果个数					
	C48旅游相关 有效成果数量	【数值】能够被正常访问、下载、使用,不是政府自身开发,标示数据来源且数 据关联有效(数据来源确实可以搜索、链接到)的旅游相关有效成果个数					
	C49旅游相关 成果有效率	【数值】旅游相关有效成果数量/旅游相关成果数量					
B17 成 果 质 量	C50旅游相关 优质成果数	【数值】在界面设计、功能设计、数据利用或用户反馈 上表现优秀的旅游相关有效成果个数					
	C51服务应 用质量	界面清 晰美观	功能设计完备, 富有特色	大量使用旅游 开放数据集	融合关联数 据、外部数据	用户反馈良 好(下载量、 好评量)	

具体操作上,在进入每个数据开放平台的相应页面后,通过人工观察网页内容并手动测试相应功能。以河南省在指标“未开放数据请求”上的评分为例,其相关页面虽然没有填写说明,但提供了数据格式选项栏,根据该项指标的评分标准,达到3级水平(3分)。

(3)对数据层相关指标,采用网络爬虫方式完成数据抓取和统计分析

以2021年12月6日为时间节点,对模型框架各项指标涉及的数据资源进行采集,采用八爪鱼采集器及Python脚本收集原始数据,并通过微词云、Excel对原始数据进行清洗处理,得到可用于评分的数据。由于21个样本的数据开放平台均采用了相似的网站结构布局,因此本文确定了一套可复用的网络爬虫逻辑,如图1所示。

4.4 权重计算

4.4.1 变异系数法权重计算

由于各项指标具有不同的量纲和数量级,本文采用Min-Max标准化方法进行数据标准化处理,即设 X_{\min} 和 X_{\max} 分别是属性A的最小值和最大值,将A的一个原始值 X 通过线性变换映射到[0,1]区间中得到可供比较与加权的纯数值 x 。鉴于本文指标均为正向指标,不存在负向指标,标准化计算式为:

$$x = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}} \quad (1)$$

本文提出的评价指标体系具有三级层次,在变异系数法计算过程中,需要先确立

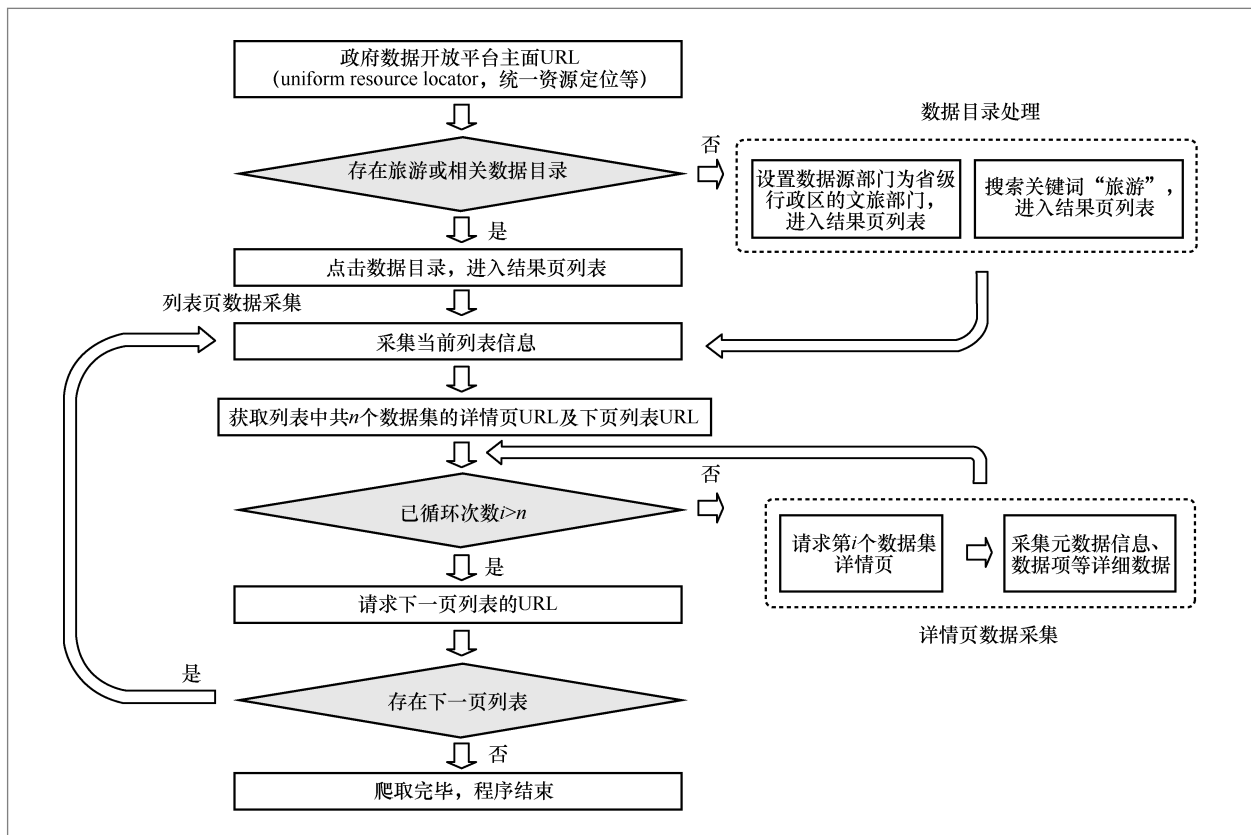


图1 本文采用的网络爬虫核心逻辑

三级指标的变异系数权重。假设二级指标B共有 p 项三级指标,先计算各项三级指标的均值 \bar{x}_j 和标准差 S_j :

$$\begin{cases} \bar{x}_j = \frac{1}{n} \sum_{i=1}^n x_{ij} \\ S_j = \sqrt{\frac{\sum_{i=1}^n (x_{ij} - \bar{x}_j)^2}{n-1}} \end{cases}, j=1,2,\dots,p \quad (2)$$

而后,计算各项三级指标的变异系数 v_j :

$$v_j = \frac{S_j}{\bar{x}_j} \quad (3)$$

对变异系数进行数据标准化处理,进而得到各项三级指标的变异系数权重 w_j :

$$w_j = \frac{v_j}{\sum_{i=1}^p v_i} \quad (4)$$

经过计算可得到二级指标B包含的 p 项三级指标的变异系数权重 W :

$$W = \{w_1, w_2, \dots, w_p\} \quad (5)$$

在应用变异系数法确立多级指标权重的研究中,有学者将较高一级指标的权重直接设定为其包含的所有下一级指标的变异系数权重之和,该方法运算简便,但由此得到的权重将受下一级指标的数量干扰^[18-19]。鉴于此,可以结合较低一级指标的变异系数权重,先计算出上一级指标数据,进而得到上一级指标的变异系数权重^[20]。

本文以三级指标的变异系数权重为基础,确立二级指标的变异系数权重。对于样本 T ,其二级指标B的原始数据 y 等于B包含的各项三级指标变异系数权重 w 与三级指标的标准化数据 x 的乘积的总和:

$$y = \sum_{i=1}^p w_i x_i \quad (6)$$

将二级指标B的原始数据进行标准化处理后,代入式(2)至式(5)的变异系

数权重计算式,可得到二级指标B的变异系数权重。同理,以二级指标的变异系数权重为基础,可确立一级指标的变异系数权重。

4.4.2 熵权法权重计算

熵权法确定各级指标权重的推导过程如下所示。

第一步:将各个指标的数据进行标准化处理,计算式同式(1)。

第二步:求各个指标的信息熵。根据信息论的定义,一组数据的信息熵 e_j 的计算式为:

$$e_j = -\frac{1}{\ln n} \sum_{i=1}^n p_{ij} \ln p_{ij} \quad (7)$$

其中, $p_{ij} = \frac{x_{ij}}{\sum_{i=1}^n x_{ij}}$, $i=1, \dots, n$; $j=1, \dots, m$; n

为样本数; m 为评价指标数。

由于 \ln 的函数自变量大于零,为了消除零值影响,定义当 $p_{ij}=0$ 时, $p_{ij} \ln p_{ij}=0$,在计算软件中,将 p_{ij} 向右平移极小值0.0001,使运算成立。由熵的极值性可知,各样本的 p_{ij} 的值越接近,则熵值越大,当 p_{ij} 的值完全相等时熵达到最大,为 $e_{\max}=1$ 。

第三步,计算第 j 个指标的熵权 w'_j :

$$w'_j = \frac{d_j}{\sum_{j=1}^m d_j} \quad (8)$$

其中, $d_j = 1 - e_j$,为差异系数。

最终,得到总计 m 个指标的熵权 W' :

$$W' = (w'_1, w'_2, \dots, w'_m) \quad (9)$$

参考将熵权法应用于多层指标体系的研究,先确定各三级指标的权重,然后将各项三级指标的标准化数据根据熵权进行

加权算术平均,求得各二级指标原始数据,再利用熵权法确定各二级指标的权重。同理,由各二级指标权重可得到各一级指标权重。

假设通过计算得出二级指标B下的 p 个三级指标的权重分别为 w_1' 、 w_2' 、 \dots 、 w_p' ,三级指标的标准化数据分别为 x_1 、 x_2 、 \dots 、 x_p ,则二级指标B的原始数据指数 y 计算式为:

$$y = \sum_{i=1}^p w_i' x_i \quad (10)$$

同理,以二级指标的熵权为基础,可确立一级指标的熵权。

4.4.3 组合权重计算

组合权重通常采用线性加权组合法或乘法组合法来确定^[21]。线性加权组合法通过将某一指标由多种赋权方法下得到权数进行加权汇总,产生某一指标的组合权重,计算式为:

$$\begin{cases} w_j^* = \sum_{i=1}^k \alpha_i w_{ij} \\ \text{s.t.} \sum_{i=1}^k \alpha_i = 1 \end{cases} \quad (11)$$

其中, w_j^* 为第 j 个指标的线性加权组合权重, α_i 为第 i 种方法的权系数, w_{ij} 为第 i 种方法得出的第 j 个指标的权重, k 为方法数^[22]。本文综合采用变异系数法与熵权法两种赋权方法,因此线性加权组合式可简化为:

$$w_j^* = \alpha w_j + (1 - \alpha) w_j' \quad (12)$$

其中, w_j 第 j 个指标的变异系数权重, w_j' 为第 j 个指标的熵权, α 和 $1 - \alpha$ 为两种方法所占权重的比例。考虑到变异系数法和熵权法所依赖的原始数据的统一性、在评价体系中地位的相似性,参考相关组合赋权研究的参数设定^[23], α 取0.5,以获得相对客观的整体评价结果。

乘法组合法将多种赋权方法得出的某一指标的权数相乘,然后进行归一化处理得到组合权数,计算式为:

$$w_j^{\wedge} = \frac{\prod_{k=1}^q w_j(k)}{\sum_{j=1}^m \prod_{k=1}^q w_j(k)} \quad (13)$$

其中, w_j^{\wedge} 为第 j 个指标的乘法组合法组合权重, $w_j(k)$ 为第 k 种方法得出的第 j 个指标的权重, q 为方法数, m 为指标数。在本文情境下,乘法组合法计算式可进一步被简化为:

$$w_j^{\wedge} = \frac{w_j w_j'}{\sum_{j=1}^m w_j w_j'} \quad (14)$$

其中, w_j 表示第 j 个指标的变异系数权重, w_j' 表示第 j 个指标的熵权, m 为指标总数。

根据式(12)、式(14)分别计算各项指标,得到组合权重。4种方法计算权重的结果如图2所示。一方面,线性加权组合权重值介于变异系数法与熵权法之间,具有明显的“调节效应”,避免赋权依据单一导致的权重偏移问题。另一方面,基于乘法组合法组合权重显示出“倍增效应”,即原始权重偏大的指标的组合权重倍增,反之则权重倍减,由此存在两个问题:一是两种原始权重乘积的涵义很难给出明确解释^[24],二是部分指标权重被过度放缩为极端值,而这些极端值在旅游数据开放领域缺乏有力的理论依据。

综上,鉴于线性加权法的组合权重结果较乘法组合法的调节性、解释性与合理性更强,本文将线性加权法的组合权重结果作为最终权重,见表2,各项指标的“权重”列表示该指标在其上一层级指标下所占的权重。将各省级行政区数据根据指标体系进行逐级加权,可以得到其总得分。

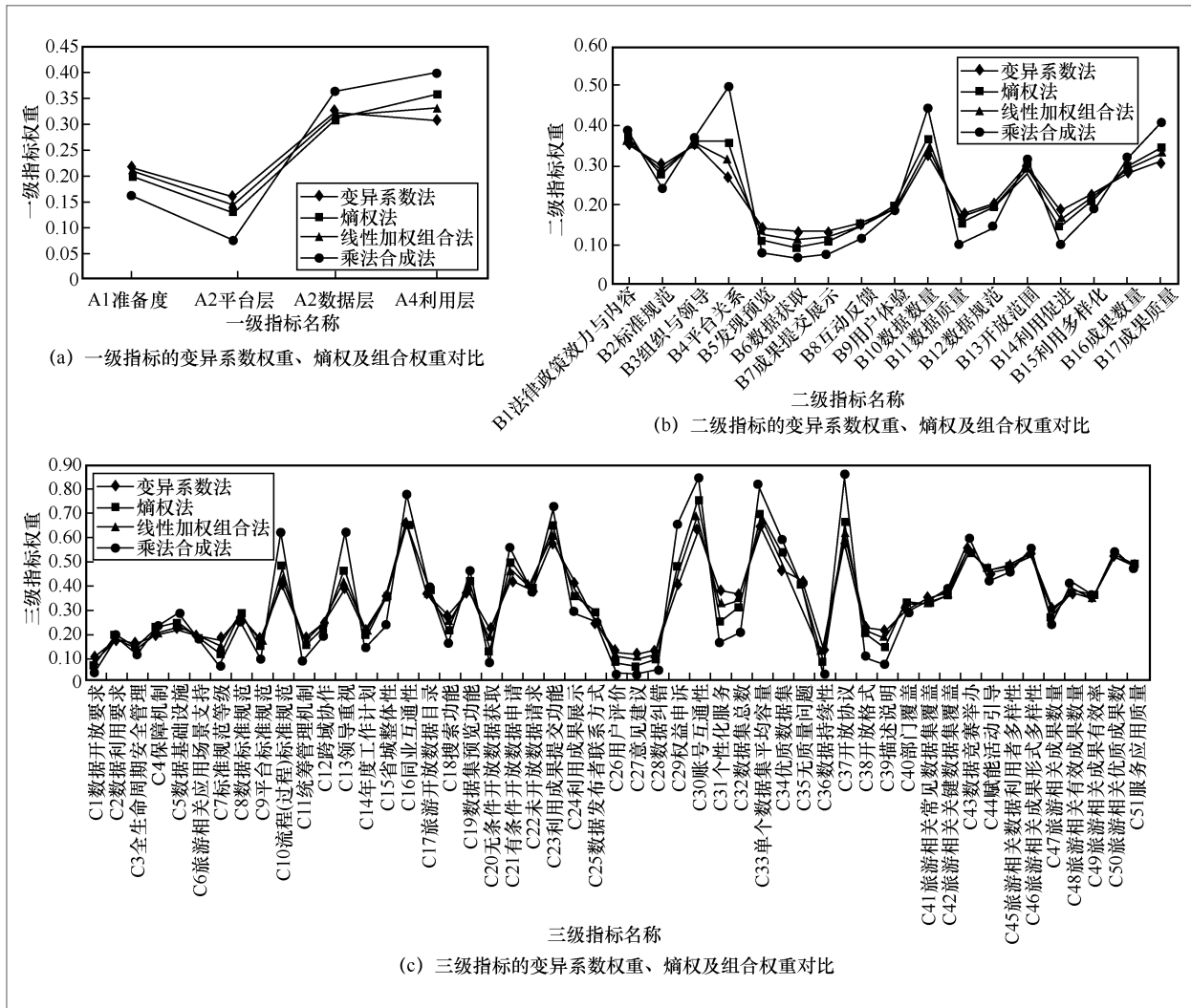


图2 各级指标的变异系数权重、熵权及组合权重对比

对于省级行政区样本 T , 其总得分 Y_T 的计算式为:

$$Y_T = \sum_{i=1}^p \sum_{j=1}^{q_i} \sum_{k=1}^{r_{ij}} w_i w_{ij} w_{ijk} x_{ijk} \quad (15)$$

其中, w_i 为第 i 项一级指标的权重, w_{ij} 为第 i 项一级指标下的第 j 项二级指标的权重, w_{ijk} 、 x_{ijk} 分别为第 i 项一级指标下的第 j 项二级指标下的第 k 项三级指标的权重与标准化数据。 p 为指标体系的一级指标数, q_i 为第 i 项一级指标下的二级指标数, r_{ij} 为第 i 项一级指标下的第 j 项二级指标下的三

级指数。

4.5 省级行政区旅游数据开放水平的评价与分析

根据式(15), 各省级行政区旅游数据开放水平的总得分由各一级指标得分与一级指标权重的乘积累加得到, 具体得分见表3。其中一级指标A1准备度、A2平台层、A3数据层、A4利用层的权重 w_1 、 w_2 、 w_3 、 w_4 根据表2分别取0.2090、0.1422、0.3167、0.3321。

表2 政府旅游数据开放评价指标体系的最终权重

一级指标	权重	二级指标	权重	三级指标	权重		
A1准备度	0.2090	B1 法律政策效力与内容	0.3623	C1数据开放要求	0.0769		
				C2数据利用要求	0.1788		
				C3全生命周期安全管理	0.1362		
				C4保障机制	0.2031		
				C5数据基础设施	0.2265		
				C6旅游相关应用场景支持	0.1783		
		B2 标准规范	0.2833	C7标准规范等级	0.1363		
				C8数据标准规范	0.2745		
				C9平台标准规范	0.1547		
				C10流程标准规范	0.4345		
				C11统筹管理机制	0.1528		
				C12跨域协作	0.2296		
				C13领导重视	0.4239		
A2平台层	0.1422	B3 组织与领导	0.3544	C14年度工作计划	0.1937		
				C15省域整体性	0.3470		
				C16同业互通性	0.6530		
				B4 平台关系	0.3136	C17旅游开放数据目录	0.3672
						C18搜索功能	0.2357
				B5 发现预览	0.1225	C19数据集预览功能	0.3971
						C20无条件开放数据获取	0.1665
						C21有条件开放数据申请	0.4608
				B6 数据获取	0.1094	C22未开放数据请求	0.3728
						C23利用成果提交功能	0.6135
				B7 成果提交展示	0.1182	C24利用成果展示	0.3865
						C25数据发布者联系方式	0.2689
				B8 互动反馈	0.1473	C26用户评价	0.0958
C27意见建议	0.0856						
A3数据层	0.3167	B9 用户体验	0.1889	C28数据纠错	0.1079		
				C29权益申诉	0.4417		
		B10 数据数量	0.3478	C30账号互通性	0.6958		
				C31个性化服务	0.3042		
		B11 数据质量	0.1636	C32数据集总数	0.3329		
				C33单个数据集平均容量	0.6671		
		B12 数据规范	0.1971	C34优质数据集	0.4949		
				C35无质量问题	0.4025		
		B13 开放范围	0.2916	C36数据持续性	0.1026		
				C37开放协议	0.6195		
		C38开放格式	0.2075				
		C39描述说明	0.1729				
		C40部门覆盖	0.3108				
C41旅游相关常见数据集覆盖	0.3313						
C42旅游相关关键数据集覆盖	0.3579						

续表

一级指标	权重	二级指标	权重	三级指标	权重
A4利用层	0.3321	B14 利用促进	0.1609	C43数据竞赛举办	0.5463
				C44赋能活动引导	0.4537
		B15 利用多样性	0.2214	C45旅游相关数据利用者多样性	0.4753
				C46旅游相关成果形式多样性	0.5247
				C47旅游相关成果数量	0.2815
		B16 成果数量	0.2889	C48旅游相关有效成果数量	0.3722
				C49旅游相关成果有效率	0.3463
				C50旅游相关优质成果数	0.5180
		B17 成果质量	0.3289	C51服务应用质量	0.4820

由表3可得到以下结论。

● 排名前三的省级行政区依次为山东省、广西壮族自治区、浙江省；其中，山东省在准备度等多个一级指标维度均表现优异。此外，处于第一梯队的省级行政区还

包括广东省、四川省等。

● 处于第二梯队的省级行政区包括贵州省、香港特别行政区、江苏省、重庆市等，其主要特点是在某个一级指标层面表现优秀，在其他一级指标水平上则处于筹

表3 各省级行政区的总得分 (N=21)

省级行政区	A1准备度得分 $\times w_1$	A2平台层得分 $\times w_2$	A3数据层得分 $\times w_3$	A4利用层得分 $\times w_4$	总分	排名
山东省	0.1272	0.0853	0.2099	0.2320	0.6545	1
广西壮族自治区	0.1242	0.0679	0.0787	0.2290	0.4997	2
浙江省	0.0877	0.0762	0.0525	0.2226	0.4390	3
广东省	0.0855	0.0891	0.0740	0.1603	0.4089	4
四川省	0.0905	0.0760	0.1868	0.0437	0.3970	5
上海市	0.1027	0.0716	0.0506	0.1302	0.3551	6
福建省	0.0385	0.0622	0.0381	0.1662	0.3050	7
贵州省	0.0998	0.0788	0.0472	0.0286	0.2544	8
香港特别行政区	0.0131	0.0291	0.0264	0.1739	0.2426	9
江苏省	0.0962	0.0604	0.0193	0.0640	0.2399	10
重庆市	0.1136	0.0508	0.0141	0.0145	0.1930	11
北京市	0.0149	0.0483	0.0722	0.0572	0.1926	12
海南省	0.1087	0.0316	0.0153	0.0194	0.1750	13
天津市	0.0915	0.0464	0.0146	0.0204	0.1728	14
宁夏回族自治区	0.0297	0.0491	0.0627	0.0000	0.1415	15
河南省	0.0120	0.0464	0.0196	0.0227	0.1007	16
辽宁省	0.0220	0.0413	0.0218	0.0000	0.0851	17
江西省	0.0329	0.0188	0.0143	0.0058	0.0719	18
吉林省	0.0266	0.0083	0.0022	0.0145	0.0516	19
湖南省	0.0129	0.0152	0.0144	0.0000	0.0424	20
陕西省	0.0069	0.0191	0.0005	0.0000	0.0266	21

资料来源：作者自制

备或发展阶段。

● 处于第三梯队的省级行政区包括陕西省、湖南省、吉林省等, 主要存在数据利用率低等短板。

对4个一级指标进行数据标准化处理, 计算结果见表4, 可知: ①平台层表现较好, 得分平均值达到0.5292, 说明大部分省级平台已完成基础功能配置, 运行流畅; 但同时数据层表现欠佳, 平均值仅为0.2329。参考交通领域数据粒度的划分方法^[25], 归纳数据层在不同粒度的问题实例, 得到表5。数据开放起步早的省份由于旅游数据存量较大, 数据层的矛盾更多体现在多个数据集的关系上, 要确保“同源不同格式”的数据集内容一致, 并采用“数据集名称+更新时间”等防重名命名规则; 由于旅游数据的生成经常会涉及地址信息录入、复杂坐标转换、文本转码等易错流程, 因此需要加强从数据集到数据单元的清洗、筛查, 减少格式、元数据及数据行列项值错误。②准备度、利用层的标准差偏大, 分别为0.3533、0.3588, 各省份表现参差不齐。在法律政策效力与内容层面, 得分较高的山东、贵州等省份的数据管理机

表4 一级指标、二级指标经数据标准化处理后得分平均值与标准差 (N=21)

一级指标	得分平均值	得分标准差	二级指标	得分平均值	得分标准差
准备度	0.4718	0.3533	B1法律政策效力与内容	0.3784	0.3051
			B2标准规范	0.4484	0.2969
			B3组织与领导	0.4172	0.3346
平台层	0.5292	0.2926	B4平台关系	0.2571	0.3078
			B5发现预览	0.4667	0.2859
			B6数据获取	0.4922	0.2838
			B7成果提交展示	0.5630	0.3313
			B8互动反馈	0.4404	0.3173
数据层	0.2329	0.2568	B9用户体验	0.3681	0.3203
			B10数据数量	0.1416	0.2396
			B11数据质量	0.2563	0.2445
			B12数据规范	0.2672	0.2830
			B13开放范围	0.1634	0.2460
利用层	0.3295	0.3588	B14利用促进	0.3448	0.3038
			B15利用多样性	0.3293	0.3636
			B16成果数量	0.2630	0.3654
			B17成果质量	0.2088	0.3191

表5 数据层不同粒度对应的问题实例 (N=21)

数据层问题	粒度	实例 (省级行政区)
D1同一数据集的不同格式文件的内容不一致	多个数据集 (多个文件)	“摄像头设置点” XLS与CSV文件数据不一致 (上海)
D2多个数据集名称重复	多个数据集 (多个文件)	历年更新的“潮州市A级旅游景区名录”的数据集名称重复 (广东)
D3格式不符合开放标准	数据集 (文件)	“密云区红色旅游信息”等为DOC格式 (北京)、文件“公司名称”等为XLS格式 (河南)
D4更新频率与承诺频率不一致	数据集 (文件)	“浙江省旅游业主要指标信息”未按照承诺频率每月更新 (浙江)
D5元数据缺失	数据集 (文件)	“贫困县乡村旅游”“三资企业、外商投资、旅游情况”目录简介缺失 (湖南)
D6数据被人工编辑过, 非原始数据	记录 (行)	“蓟州区天津市级文物保护单位”第21行至第41行 (天津)
D7字段内数据格式	字段 (列)	“年度规范旅游市场处罚情况”的“行政处罚依据”字段数据有的以顿号分隔, 有的没有 (天津)
D8出现乱码	数据单元	“博物馆入场人次”个别行出现特殊符号乱码 (香港)

制较为健全,但尚未有专门针对旅游行业细化数据目录的编制要求。

对17个二级指标进行数据标准化处理,计算结果见表4,可知:①数据数量的总体平均分最低,为0.1416,说明大部分省份并未实现大数量、高容量的旅游数据供给,难以满足数据使用主体的核心诉求;②利用多样性和成果数量是各省级行政区表现差异最大的两个二级指标,标准差分别达到0.3636、0.3654,二者同属于利用层面,反映了旅游数据利用水平不平衡,开发者创作动力不足的问题。

在三级指标方面,有几项指标需要重视。

第一,单个数据集平均容量。数据容量是数据集行数与列数的乘积。单个数据集平均容量越大,则数据集字段越充实,颗粒度越细,数据利用的便利性越强。将该指标得分最高的10个省级行政区按平均容量降序排列,可知:①部分省份拥有一些大容量的优质数据集,但平均容量较低,还存在着大量空数据集和低质量的无效数据集;②平均容量排名靠前的省份大多存在单体数据容量在千万量级的超大数据集;③超大容量数据集主要生成于景区相关场景,如景区预约、人流量、饱和指数、门票等。景区是最重要的旅游活动场域,与其他数据元素交集广泛,因此数据积累量大,分析价值也较高。

第二,部门覆盖。山东省涉及旅游开放数据的源头部门有19类,包括交通运输局、城市管理局、公安局、卫生健康局等非旅游部门,得分最高;北京市覆盖了16个数源部门,涉及园林绿化局、统计局、气象局等非旅游部门,得分次高;其余省级行政区的数源部门均少于10个,有11个省级行政区的数源部门仅有文旅部门一家,未体现“食住行游购娱”的旅游特色。

第三,旅游相关常见数据集覆盖。对21个平台上旅游数据集进行词频分析,前

10名高频词依次为“景区”“旅行社”“文化遗产”“星级”“酒店”“服务指南”“接待”“博物馆”“智慧旅游”“游客”,这些是旅游数据开放工作中最成熟的板块。统计各省份的名称、摘要包含上述高频词的数据集个数,即可测算指标“常见数据集覆盖”。经统计,常见数据集个数排名前5的省级行政区为:山东省(437个)、四川省(133个)、广东省(127个)、贵州省(116个)、上海市(50个)。四川省提供了27个智慧旅游数据集,反映了其大力发展“旅游+新技术”的态势。广东省提供了34个文化遗产数据集和25个旅行社数据集,在两类数据集总量中占比分别达到44.16%和16.78%,体现了侧重旅游业务供需两端的特色。贵州省是博物馆数据集的主要提供者,上海市则未体现出明显的类别倾向。

5 政府旅游数据开放高评价价值与低评价价值的模糊集定性比较分析

5.1 变量选取与校准

在条件变量的数量上,模糊集定性比较分析的条件变量数量一般为4~7个^[26-28]。在条件变量与结果变量的选择上^[29-30],将一级指标准备度、平台层、数据层、利用层作为条件变量,将评价价值总分作为结果变量,进行政府旅游数据开放的高评价价值与低评价价值的条件组态分析。

对变量得分进行校准,首先设定3个校准锚点:完全隶属点、交叉点、完全不隶属点,然后将所有数据转化至[0,1]区间内,以判定赋值后案例的取向。参考现有研究,选择25分位作为完全不隶属点,50分位作为交叉点,75分位作为完全隶属点^[31]。将校准后数值为0.5(即处于交叉点)的数据修改为0.501^[32-33]。

5.2 必要条件检验

首先,检验必要条件的一致性(consistency)。必要条件检验的一致性指标用于测量单项条件的重要程度,由一致性可知结果的模糊集合在多大程度上是条件模糊集合的子集。必要条件一致性的计算式为:

$$\text{consistency}(X_i \leq Y_i) = \frac{\sum_i \min(X_i, Y_i)}{\sum_i X_i} \quad (16)$$

其中, X_i 为条件组合中的隶属分数, Y_i 为高评价价值结果中的隶属分数,一致性的取值范围为[0,1]。若某条件变量的一致性达到0.9以上,则说明该变量为结果变量的必要条件。

随后,对必要条件的覆盖率(coverage)进行检验。覆盖率代表前因变量对于结果变量的解释力度。必要条件覆盖率的计算式为:

$$\text{coverage}(X_i \leq Y_i) = \frac{\sum_i \min(X_i, Y_i)}{\sum_i Y_i} \quad (17)$$

其中, X_i 、 Y_i 的含义同式(16),覆盖率取值范围为[0,1]。

经计算,与政府旅游数据开放高评价价值的一致性最高的条件变量为平台层,一致性为0.8916,但是包括其在内的各条件变量的一致性水平和覆盖率均未高于0.9,均不单独构成导致政府旅游数据开放高评价价值的原因。与政府旅游数据开放低评价价值的一致性最高的条件变量为“~利用层”(符号“~”表示该变量的相反值,即不存在),一致性为0.9030,说明其是单独导致政府旅游数据开放低评价价值的必要条件。

5.3 条件组态分析与稳健性检验

真值表是变量的各种可能取值及其对应的目标函数值的集合。通过操作fsQCA程序计算得到政府旅游数据开放高评价价值与低评价价值组态的初始真值表,再对其进行简化处理,以剔除矛盾组合。首先,设置最低可覆盖样本量为1,原始一致性阈值为0.8,剔除样本量与原始一致性不合格的组态。其次,设置一致性阈值为0.5,剔除组态呈现显著不一致性的样本,避免同时子集关系。最后,将更新后的真值表进行标准化分析,得到简单解、复杂解和中间解3种解状态下的条件组态。其中,简单解将无样本对应的逻辑余项组态完全纳入运算;复杂解完全不纳入运算;中间解将部分有意义组态纳入运算,并提出结果变量的核心条件,优于简单解和复杂解。一般研究通常取中间解,并结合简单解来确定中间解的各条件变量的重要程度。根据中间解析出结果,当前中国政府旅游数据开放高评价价值与低评价价值都主要表现为2种组态,见表6。

解的覆盖率和解的一致性检验所有组态整体解释力的指标,前者为“中间解的样本占有所有符合结果变量的样本的比例”,作用类似传统定量研究的拟合优度;后者为“中间解的样本中符合结果变量的样本的占比”,用于表示中间解的所有组态在多大程度上与结果相符。原始覆盖率、唯一覆盖率、一致性用于衡量各个组态自身的解释力。其中,原始覆盖率表示结果在多大程度上可以被某组态解释;唯一覆盖率表示结果在多大程度上仅可以被某组态解释;一致性表示某组态在多大程度上与结果相符^[34-35]。上述指标的取值范围均为[0,1],越接近1表明解释力度越好,各指标之间不存在直接的加减运算关系。

表6 政府旅游数据开放高评价值与低评价值的组态结果 (N=21)

条件变量	高评价值组态1	高评价值组态2	低评价值组态1	低评价值组态2
	综合发展型	数据辅助应用型	政策与利用短板型	平台与数据问题型
准备度	●	⊗	⊗	
平台层	●		●	⊗
数据层		●	●	⊗
利用层	●	●	⊗	⊗
原始覆盖率	0.5917	0.3121	0.1756	0.7217
唯一覆盖率	0.4206	0.1409	0.0825	0.6286
一致性	0.9984	0.9287	0.8775	0.9800
解的覆盖率		0.7326		0.8042
解的一致性		0.9671		0.9535

注：●与●表示该条件存在，⊗与⊗表示该条件不存在。●与⊗为核心条件，●与⊗为边缘条件。空白表示条件可以存在也可不存在。

由表6可知，高评价值条件组态解的覆盖率为0.7326，低评价值条件组态解的覆盖率为0.8042，表示政府旅游数据开放上处于高评价值的样本中有73.26%可以被条件组态解释，处于低评价值的样本中有80.42%可以被条件组态解释。同时，0.8是解的一致性的可接受参考值，而本文得到的高评价值条件组态解的一致性达到0.9671，低评价值条件组态解的一致性为0.9535，均高于可接受参考值，说明本文的模糊集定性比较分析结果较好，具有较高的可信度与解释力。根据组态的条件结构与各个条件变量的重要性强度，将高评价值组态1、组态2分别命名为综合发展型、数据辅助应用型，将低评价值组态1、组态2分别命名为政策与利用短板型、平台与数据问题型。

稳健性检验参照相关研究，采用“将一致性阈值由0.8提升至0.85”与“将最低案例频数从1提升至2”的两种检验模式^[36-38]，结果见表7。对比表6与表7，高评价值组态1、组态2在调整一致性检验阈值后结果无变

化；调整最小案例频数后仅分别新增了1个核心条件和1个边缘条件，整体变动较小。低评价值组态1、组态2在调整一致性检验阈值后结果无变化；调整最小案例频数后，只筛去了仅有1个案例的组态1，组态2的条件变量关系和内部拟合参数无显著变化。此外，调整后的中间解的一致性分别为0.9671、0.9791、0.9535、0.9800，仍有较强的解释力。综上所述，原分析结果具有稳健性，可以接受。

5.4 结果解释与典型组态分析

5.4.1 综合发展型

在模糊集运算中，“*”表示“逻辑与”，“~”表示“逻辑非”。综合发展型是旅游数据开放高评价值的第1种条件组态，其条件结构为“准备度*平台层*利用层”，各个维度综合发展是该组态最突出的特征。在构成该组态的所有条件变量中，利用层是核心条件，准备度和平台

表7 政府旅游数据开放高价值与低评价价值的组态结果稳健性检验(N=21)

条件变量	调整一致性检验阈值				调整最小案例频数			
	高评价值组态1	高评价值组态2	低评价值组态1	低评价值组态2	高评价值组态1	高评价值组态2	低评价值组态1	低评价值组态2
准备度	●	⊗	⊗		●	⊗	/	
平台层	●		●	⊗	●	⊗	/	⊗
数据层		●	●	⊗	●	●	/	⊗
利用层	●	●	⊗	⊗	●	●	/	⊗
原始覆盖率	0.5917	0.3121	0.1756	0.7217	0.5066	0.1622	/	0.7217
唯一覆盖率	0.4206	0.1409	0.0825	0.6286	0.4399	0.0955	/	0.7217
一致性	0.9984	0.9287	0.8775	0.9800	1.0000	0.9265	/	0.9800
解的覆盖率	0.7326		0.8042		0.6021		0.7217	
解的一致性	0.9671		0.9535		0.9791		0.9800	

注：●与●表示该条件存在，⊗与⊗表示该条件不存在。●与⊗为核心条件，●与⊗为边缘条件。空白表示条件可以存在也可不存在。

层是辅助条件。与数据辅助应用型组态相比，综合发展型组态的一致性较高，为0.9984。山东省、四川省、广东省是综合发展型组态的重要实践者。

5.4.2 数据辅助应用型

数据辅助应用型是旅游数据开放高评价价值的第2种条件组态，其条件结构为“~准备度*数据层*利用层”。其中，强大的赋能利用能力是该组态的核心条件，数据层建设的存在与准备度建设的缺失则为边缘条件，平台层表现对该组态不产生影响。31.21%的高评价价值实践可以用此组态解释，14.09%的高评价价值实践仅可被此组态解释。福建省、北京市和香港特别行政区是数据辅助应用型组态的典型代表。

5.4.3 政策与利用短板型

政策与利用短板型是旅游数据开放

低评价价值的第1种条件组态，由4个条件变量构成，其中核心条件变量为“~准备度”“~利用层”，边缘条件变量为平台层、数据层。政策与利用短板型的原始覆盖率、唯一覆盖率较低，分别为0.1756、0.0825，这是由于该组态仅可解释宁夏回族自治区这一样本；同时，该组态的一致性较高，为0.8775。

5.4.4 平台与数据问题型

平台与数据问题型是旅游数据开放低评价价值的第2种条件组态，其条件结构为“~平台层*~数据层*~利用层”。其中，“~平台层”“~数据层”为核心条件，“~利用层”为边缘条件。该组态的原始覆盖率为0.7217，可以解释72.17%的低水平样本，说明该组态是旅游数据开放低水平的重要成因，其矫正方案可以作为表现薄弱的省份“扭转战局”的路径。陕西省、湖南省、江西省、吉林省、海南省等8个省级行政区属于该组态。

6 结论与展望

6.1 研究结论

第一,构建政府旅游数据开放评价指标体系框架并实现组合赋权,形成独特的旅游数据指标权重。本文构建了具有4项一级指标、17项二级指标、51项三级指标的政府旅游数据开放评价指标框架,上级指标为下级指标的凝练概括,下级指标为上级指标的具体构成,结合了等级型与数值型两类评分标准类型,且两者配比大致平衡。

赋权结果显示:一级指标层面,A4利用层权重最高,A3数据层次高,A2平台层权重最低,符合“目标导向”原则与产业发展实际;二级指标层面,B1法律政策效率与内容、B4平台关系、B10数据数量等基础类、旅游关联类工作被赋予较高权重,且整体权重配比较为均衡;三级指标层面,高权重指标包括反映通用开放要求的C33单个数据集平均容量、C37开放协议等指标,以及体现旅游行业特色,专用于旅游数据开放工作评价的C16同业互通性、C42关键数据集覆盖、C50优质成果数等指标,可以充分体现数据要素的价值。

综合评分显示:总体得分层面,得分最高的3个省级行政区依次为山东省、广西壮族自治区、浙江省;处于第二梯队的有贵州省、香港特别行政区等;有较大进步空间的是陕西省、湖南省等。一级指标层面,各省级行政区在数据层整体表现欠佳,准备度、利用层表现差异偏大。二级指标层面,数据数量的总体平均分最低,利用多样性和成果数量分异较大。三级指标层面,以景区为数据源生成大容量数据集、

加强部门场景联动等具体实践值得关注。

第二,对21个省级行政区的旅游数据开放水平进行组态分析,聚类高评价组态和低评价组态的典型特征,给出具体优化建议。本文基于模糊集定性比较分析,对应高评价和低评价旅游数据开放分别得到2种条件组态。

高评价组态结果显示:首先,“综合发展型”组态的覆盖率、一致性较高,其经验适用于在数据开放基础工作的某个或多个方面已筹备完善的情形,山东省、四川省、广东省等是典型代表。该组态可以从细化数据规范条款、提高旅游数据集在平台资源位的展示率等方面寻求突破。其次,“数据辅助应用型”组态以数据建设和市场促活机制为跳板,提升数据赋能利用能力,福建省、北京市和香港特别行政区是典型代表。该组态可以尝试依靠数据时效与数据质量支撑应用开发,引进道路运营、手机信令等跨领域关联数据,或调整应用的前端设计、互动机制,提高旅游数据应用成果的访问率、留存率、转化率。

低评价组态结果显示:首先,“政策与利用短板型”组态的痛点在于“平台层与数据层建设初见成效,但准备度、利用层严重拉低了整体开放水平”,典型代表为宁夏回族自治区。其次,“平台与数据问题型”组态多处在数字经济发展的中级或初级阶段,表现为旅游数据开放平台与数据集建设存在不足,典型代表为陕西省和海南省。两种组态亟须采取针对性的优化策略,如出台数据开放地方性法规、更新可用性高的数据描述信息、整改平台基础功能错误等,消除关键症结,实现逆势突围。

6.2 管理启示

第一,出台旅游专域数据开放法规与政策,构建多主体运营管理机制。法规层

面,以《中华人民共和国数据安全法》《中华人民共和国个人信息保护法》等法律及省级数据共享开放条例为框架,制定涉及旅游数据开放基本要求、脱敏指南、中台建设、更新频率的地方标准^[39]。基本政策的制定应吸纳多元主体意见^[40],使旅游/非旅游行政部门、旅游研究机构、旅游企业、公民群策群力,与国家相关战略协同呼应。具体政策应清晰界定、重点描述,内容范围类政策应回应“旅游领域开放哪些数据”“以何方式开放数据”;问责机制类政策应规定“有哪些责任主体”“违规行为如何追责”;人才培养类政策应确定“旅游-数据”复合专业人才的培养原则、评价方法与认证体系。分工协作体系的组织可以参考贵阳市的经验,设立地方大数据职能部门与领导小组,统筹旅游开放数据资源融通,将旅游大数据产业专家咨询委、同业协会等纳入管理体系,形成对必要设施、流量垄断的监管规制。

第二,强化省内平台垂直互通性,上线旅游数据分析功能。垂直互通是指省级数据开放平台与下属的市、区级平台以及旅游同业平台之间功能、数据的实际互通。在旅游数据开放领域,强化省内平台垂直互通可以从“平台许可制度的标准化”着手。目前,江苏等多个省份已自主提出平台许可制度。以此为基础,各省级行政区数据开放平台可以制作协议模版,避免数据再利用的不兼容性。此外,可提供数据偏好追踪、个性化数据分析功能,向用户推荐与其高频点击标签相关联的数据集与接口,或将平台的地图分析功能与旅游空间地理数据整合,打造三维影像、历史影像、地图叠加数据集信息等高级分析功能。其三,优化“用户帮助”界面设计,重构热点问题详解、人工智能客服等服务界面,嵌入旅游数据分析功能的演示教程。

第三,引入“基础库-行业库-整合库”多渠道采集模式,提升高价值旅游数据集占比。各省级行政区可以考虑引入多渠道数据采集模式。首先,梳理空间地理数据、法人及人口数据、宏观经济等基础信息数据库中使用率高、接入难度小的旅游类数据,得到粗粒度数据;其次,检索历年《中国文化和旅游统计年鉴》《入境游客抽样调查资料》等与旅游相关的专业化统计资源,得到细粒度数据;最后,将部分已过著作权保护期的旅游课题资料使用光学字符识别技术转化成文本型数据,与旅游资讯网站合作,导入具备公共许可证的社交场景数据、游客评论数据,并从公开咨询报告中汲取数据主题,扩充到政府旅游开放数据库。此外,还需提升高价值数据集在旅游开放数据集中的占比。首先,开展数据评估试点,以用户(特别是中小旅游企业)的实际收益为价值取向,设置动态更新等准入门槛,构建数据评估维度;其次,加强旅游关键数据集建设,参照中共中央、国务院发布的《关于构建更加完善的要素市场化配置体制机制的意见》,构建旅游社会信用、经济运行、应急指挥等事关产业战略的旅游关键数据集,提供专题资源服务;最后,优质数据集应强调数据容量与数据质量的有机统一,关注景区服务、旅游教学与科研等关键命题及旅游沉浸式体验、夜景经济等新兴领域的大容量数据,配置必要数据字段。

第四,孵化旅游相关有效成果,丰富数据赋能与利用形式。省级行政区可以尝试开放数据下载链接,邀请终端用户测试,评估旅游数据开放的实际效用,淘汰低质量数据产品。旅游数据主管部门需要改变“以奖促利”的短效奖励机制,引导旅游数据开放成果增设订阅费、交叉补贴等市场化收益形式,提升商业性客户对旅游数据服务的支付意愿,促使开发者愿意在

旅游数据事业中“投资自我”。在数据赋能利用形式上,基础较薄弱的省级行政区可以从3个方面入手:一是设置数据通行证、数据科学短期合作课题、开源项目托管等新型成果通道;二是唤起数据服务提供者、基础设施提供者、代理运营者等能够发挥场景识别与中介作用的数据利用生态“中间生产者”;三是举办旅游开放数据创新应用赛事,向旅游电商平台、旅行社、酒店等商业主体征集行业痛点作为命题。数据赋能基础雄厚的省级行政区则可以基于关联数据,重点捕捉多元的数据应用场景,开展旅游数据的高阶应用。

6.3 未来展望

受限于研究能力与资源,本文存在一定局限性。首先,原始数据的准确性与全面性待加强。旅游类数据分布较分散,有些数据集可能不以“旅游”命名,或未添加旅游相关标签,但对旅游产业的发展影响重大,如安防数据等,本文暂未实现这类数据采集。未来研究可以通过预设更多检索关键词、扩大检索来源库等方式改进。其次,评价体系的可复制性尚待验证。本文属于静态横截面研究,未来应该增加样本,引入面板数据或者时间序列数据,或针对特定行政区的开展长期案例跟踪,重点考察评价体系的可复制性,使评价研究与实践工作结合更加紧密。最后,本文评价体系主要依据客观实践设立,为旅游数据开放工作提供参考雏形,后续有待广泛邀请产业界与学术界专家进行主观指标赋权与评分,优化评价体系。

参考文献:

- 架、现状与建议[J]. 电子政务, 2015(7): 8-16.
- ZHENG L, GAO F. Research on China open government data platform: framework, status and suggestions[J]. E-Government, 2015(7): 8-16.
- [2] SANDOVAL-ALMAZAN R, GIL-GARCIA J R. Toward an integrative assessment of open government: proposing conceptual lenses and practical components[J]. Journal of Organizational Computing and Electronic Commerce, 2016, 26(1/2): 170-192.
- [3] WIRTZ B W, PIEHLER R, THOMAS M J, et al. Resistance of public personnel to open government: a cognitive theory view of implementation barriers towards open government data[J]. Public Management Review, 2016, 18(9): 1335-1364.
- [4] HELLBERG A S, HEDSTRÖM K. The story of the sixth myth of open data and open government[J]. Transforming Government: People, Process and Policy, 2015, 9(1): 35-51.
- [5] YOUNG M M. Implementation of digital-era governance: the case of open data in US cities[J]. Public Administration Review, 2020, 80(2): 305-315.
- [6] EVANS A M, CAMPOS A. Open government initiatives: challenges of citizen participation[J]. Journal of Policy Analysis and Management, 2013, 32(1): 172-185.
- [7] JANSSEN M, VAN DEN HOVEN J. Big and open linked data (BOLD) in government: a challenge to transparency and privacy?[J]. Government Information Quarterly, 2015, 32(4): 363-368.
- [8] ATTARD J, ORLANDI F, SCERRI S, et al. A systematic review of open government data initiatives[J]. Government Information Quarterly, 2015, 32(4): 399-418.
- [9] VICTORIA W, DAVID S. Exploring the extent of openness of open government data - a critique of open government datasets in the UK[J]. Government Information Quarterly, 2020, 37(1): 101405.
- [10] ELEONORA P, CONSTANTINOS P,
- [1] 郑磊, 高丰. 中国开放政府数据平台研究: 框

- NIKOLAOS S. “You will like it!” using open data to predict tourists’ response to a tourist attraction[J]. *Tourism Management*, 2017, 60: 430-438.
- [11] 李萌. 基于智慧旅游的旅游公共服务机制创新[J]. *中国行政管理*, 2014(6): 64-68.
LI M. The innovation of tourism public service mechanism based on smart tourism system[J]. *Chinese Public Administration*, 2014(6): 64-68.
- [12] 李莉. 基于层次分析法的旅游政务网站评价研究[J]. *重庆第二师范学院学报*, 2014(6): 44-48, 167.
LI L. Evaluation research on tourism e-government web based on analytic hierarchy process method[J]. *Journal of Chongqing University of Education*, 2014(6): 44-48, 167.
- [13] 夏姚璜, 邢文明. 开放政府数据评估框架下的数据质量调查与启示: 基于《中国地方政府数据开放报告(2018)》[J]. *情报理论与实践*, 2019, 42(8): 44-49, 66.
XIA Y H, XING W M. The investigation and enlightenment of data quality based on the open government data evaluation framework: a study of “the local government data open report in China (2018)”[J]. *Information Studies (Theory & Application)*, 2019, 42(8): 44-49, 66.
- [14] 郑磊, 韩笑, 朱晓婷. 地方政府数据开放平台研究: 功能与体验[J]. *电子政务*, 2019(9): 12-22.
ZHENG L, HAN X X, ZHU X T. Research on local government data open platform: functions and experiences[J]. *E-Government*, 2019(9): 12-22.
- [15] 林鹏, 李丽红, 郝生跃. 众创空间发展水平评价研究: 基于能力成熟度模型[J]. *科技管理研究*, 2020, 40(17): 61-67.
LIN P, LI L H, HAO S Y. Evaluation on the development level of maker space: based on the capability maturity model[J]. *Science and Technology Management Research*, 2020, 40(17): 61-67.
- [16] 陈媛媛. 欧洲国家开放数据成熟度研究及发展建议[J]. *图书馆学研究*, 2021(2): 67-72.
CHEN Y Y. Research on the maturity of open data in European countries and development suggestions[J]. *Research on Library Science*, 2021(2): 67-72.
- [17] 李君, 邱君降, 窦克勤, 等. 基于成熟度视角的工业互联网平台评价研究[J]. *科技管理研究*, 2019, 39(2): 43-47.
LI J, QIU J J, DOU K Q, et al. Assessment research on industrial Internet platform from the maturity aspect[J]. *Science and Technology Management Research*, 2019, 39(2): 43-47.
- [18] 宋彦蓉, 张宝元. 基于地区现代化评价的客观赋权法比较[J]. *统计与决策*, 2015(11): 82-86.
SONG Y R, ZHANG B Y. Comparison of objective weighting methods based on regional modernization evaluation[J]. *Statistics and Decision*, 2015(11): 82-86.
- [19] 耿松涛, 杨晶晶, 严荣. 自贸区(港)建设背景下海南会展业发展评价及政策选择[J]. *经济地理*, 2020, 40(11): 140-148.
GENG S T, YANG J J, YAN R. Development evaluation and policy choice of mice industry in Hainan under the background of construction of free trade zone[J]. *Economic Geography*, 2020, 40(11): 140-148.
- [20] 姚爽, 黄玮强, 张展, 等. 区域金融生态环境评价: 基于修正变异系数法[J]. *技术经济*, 2015, 34(10): 61-67.
YAO S, HUANG W Q, ZHANG Z, et al. Evaluation on regional financial ecology environment: based on adjusted variation coefficient method[J]. *Technology Economics*, 2015, 34(10): 61-67.
- [21] 韩亚娟, 杨玉琪. 基于组合赋权法和未确知测度模型的发动机健康状况评估[J]. *运筹与管理*, 2020, 29(4): 204-211.
HAN Y J, YANG Y Q. Health status evaluation of aero-engines based on combination weighting method and unascertained measure model[J]. *Operations Research and Management Science*, 2020, 29(4): 204-211.
- [22] 杨宇. 多指标综合评价中赋权方法评析[J]. *统计与决策*, 2006(13): 17-19.
YANG Y. Evaluation of weighting

- methods in multi-index comprehensive evaluation[J]. *Statistics and Decision*, 2006(13): 17-19.
- [23] 查奇芬, 董洁, 陈祖功. 主客观组合赋权法在科技成果转化评价分析中的应用[J]. *统计与决策*, 2009(16): 152-154.
ZHA Q F, DONG J, CHEN Z G. Application of subjective and objective combination weighting method in evaluation and analysis of scientific and technological achievements transformation[J]. *Statistics and Decision*, 2009(16): 152-154.
- [24] 李刚, 李建平, 孙晓蕾, 等. 主客观权重的组合方式及其合理性研究[J]. *管理评论*, 2017, 29(12): 17-26, 61.
LI G, LI J P, SUN X L, et al. Research on a combined method of subjective-objective weighing and the its rationality[J]. *Management Review*, 2017, 29(12): 17-26, 61.
- [25] 龙莎. 基于数据治理的开放交通数据发布流程优化及质量控制[D]. 大连: 大连海事大学, 2020.
LONG S. Optimization and quality control of open traffic data release process based on data governance[D]. Dalian: Dalian Maritime University, 2020.
- [26] 梁乙凯, 戚桂杰. 基于模糊集定性比较分析的政府开放数据使用影响因素研究[J]. *情报杂志*, 2019, 38(3): 180-186.
LIANG Y K, QI G J. The determinants of open government data usage based on fuzzy-set qualitative comparative analysis[J]. *Journal of Intelligence*, 2019, 38(3): 180-186.
- [27] 王法硕. 我国地方政府数据开放绩效的影响因素: 基于定性比较分析的研究[J]. *情报理论与实践*, 2019, 42(8): 38-43.
WANG F S. Influencing factors of local government data open performance in China: a research based on qualitative comparative analysis[J]. *Information Studies (Theory & Application)*, 2019, 42(8): 38-43.
- [28] 赵玉攀. 组织资源对开放政府数据实施的影响: 基于回归和模糊集定性比较分析[J]. *情报杂志*, 2020, 39(10): 158-164, 99.
ZHAO Y P. The influence of organizational resources on the implementation of open government data—based on the qualitative comparative analysis of regression and fuzzy set[J]. *Journal of Intelligence*, 2020, 39(10): 158-164, 99.
- [29] 汤志伟, 王研. TOE框架下政府数据开放平台利用水平的组态分析[J]. *情报杂志*, 2020, 39(6): 187-195.
TANG Z W, WANG Y. Configuration analysis of utilization level of government data open platform under the framework of TOE[J]. *Journal of Intelligence*, 2020, 39(6): 187-195.
- [30] 刘淑妍, 王湖葩. TOE框架下地方政府数据开放制度绩效评价与路径生成研究: 基于20省数据的模糊集定性比较分析[J]. *中国行政管理*, 2021(9): 34-41.
LIU S Y, WANG H P. Research on performance evaluation and path generation of local government data open system under TOE framework: qualitative comparative analysis of fuzzy sets based on data from 20 provinces[J]. *Chinese Public Administration*, 2021(9): 34-41.
- [31] 张琳, 黄美玲, 杨毅. 场景视域下特色小镇创生发展的文化介入路径研究: 基于川渝33个样本的模糊集定性比较分析[J]. *福建论坛(人文社会科学版)*, 2020(10): 88-97.
ZHANG L, HUANG M L, YANG Y. A Study on the Path of Cultural Intervention in the Creation and Development of Characteristic Towns from the Perspective of Scenes—based on the qualitative comparative analysis of fuzzy sets of 33 samples from Sichuan and Chongqing[J]. *Fujian Tribune*, 2020(10): 88-97.
- [32] FISS P C. Building better causal theories: a fuzzy set approach to typologies in organization research[J]. *Academy of Management Journal*, 2011, 54(2): 393-420.
- [33] CAMPBELL J T, SIRMON D G, SCHIJVEN M. Fuzzy logic and the market: a configurational approach to investor

- perceptions of acquisition announcements[J]. Academy of Management Journal, 2016, 59(1): 163-187.
- [34] GRECKHAMER T, MISANGYI V F, FISS P C. Chapter 3 the two QCAs: from a small-N to a large-N set theoretic approach[M]//Research in the sociology of organizations. Bingley: Emerald Group Publishing Limited, 2013: 49-75.
- [35] GARCIA-CASTRO R, AGUILERA R V. Family involvement in business and financial performance: a set-theoretic cross-national inquiry[J]. Journal of Family Business Strategy, 2014, 5(1): 85-96.
- [36] 张正荣, 肖文丽. 我国自由贸易试验区的设立与推广模式: 基于fsQCA的省际比较分析[J]. 浙江理工大学学报(社会科学版), 2022, 48(1): 13-20.
- ZHANG Z R, XIAO W L. On the establishment and promotion mode of pilot free trade zones in China: an inter-provincial comparative analysis based on fsQCA[J]. Journal of Zhejiang Sci-Tech University (Social Sciences Edition), 2022, 48(1): 13-20.
- [37] 郝政, 何刚, 王新媛, 等. 创业生态系统组态效应对乡村产业振兴质量的影响路径: 基于模糊集定性比较分析[J]. 科学学与科学技术管理, 2022, 43(1): 57-75.
- HAO Z, HE G, WANG X Y, et al. Effect route of entrepreneurial ecosystem on rural industry revitalization: a research based on fuzzy set qualitative comparative analysis[J]. Science of Science and Management of S & T (Monthly), 2022, 43(1): 57-75.
- [38] 张明, 杜运周. 组织与管理研究中QCA方法的应用: 定位、策略和方向[J]. 管理学报, 2019, 16(9): 1312-1323.
- ZHANG M, DU Y Z. Qualitative comparative analysis(QCA)in management and organization research: position, tactics, and directions[J]. Chinese Journal of Management, 2019, 16(9): 1312-1323.
- [39] 袁莉, 李姗蔓, 赵婧. 政策文本视角下地方政府数据开放动态能力的识别与培育研究[J]. 情报理论与实践, 2022, 45(10): 96-102.
- YUAN L, LI S M, ZHAO J. Research on identification and cultivation of local open government data's dynamic capability from the perspective of policy text[J]. Information Studies (Theory & Application), 2022, 45(10): 96-102.
- [40] 毛太田, 赵绮雨, 朱名勋. 基于协同理论的政府开放数据共建共享研究[J]. 图书馆学研究, 2020(11): 28-32, 51.
- MAO T T, ZHAO Q Y, ZHU M X. Research on Co-construction and sharing of government open data based on synergy theory[J]. Research on Library Science, 2020(11): 28-32, 51.

作者简介



胡安安(1980-),男,博士,复旦大学旅游学系副教授,复旦大学电子商务研究中心研究员,主要研究方向为老年人信息技术应用、旅游数据开放与治理。



郭梦珂(1991-),女,复旦大学电子商务研究研究中心研究员,主要研究方向为数据要素治理。



黄丽华(1965-),女,博士,复旦大学管理学院信息管理与商业智能系教授、博士生导师,复旦大学电子商务研究中心主任,大数据流通与交易技术国家工程实验室常务副主任,国务院学位委员会管理科学与工程学科评议组成员,上海市信息化专家委员会成员,主要研究方向为数据资产与数据要素治理、大数据流通与交易。



黄荔桐(1996-),女,复旦大学旅游学系硕士生,主要研究方向为旅游数据开放。

收稿日期: 2023-05-15

基金项目: 国家社会科学基金项目(No.21&ZD119); 国家自然科学基金项目(No.72074052)

Foundation Items: The National Social Science Foundation of China (No.21&ZD119), The National Natural Science Foundation of China (No.72074052)