

基于第三方点评数据的 科技馆社会评价指标体系构建和实证研究

赵 洋¹ 徐 扬² 王美力¹ 李 丹¹ 杨 洋¹ 周 雷²

(中国科学技术馆, 北京 100101)¹

(北京大学信息管理系, 北京 100871)²

[摘 要] 随着科技馆在公众科学素质提升和科普教育中的作用日益凸显, 如何科学、客观地评价科技馆的社会影响成为重要课题。本研究基于收集的全国 400 余家科技馆第三方点评数据, 利用基于自然语言处理技术的主题建模、层次聚类方法, 构建了一个包含 9 个一级指标、24 个二级指标在内的科技馆社会评价指标体系, 并利用机器学习算法对大众点评、美团、携程等第三方点评平台的 133 371 条评论数据进行实证分析。结果显示, 各种规模科技馆的前 3 个一级指标“儿童娱乐与科学互动”“科学展示与文化体验”“互动学习与互动科技”得分都很高; 各种规模科技馆在交通与环境满意度方面仍有较大提升空间。研究建议, 特大型科技馆应突出科教展示优势, 补足服务短板; 大型科技馆要均衡发展核心功能, 优化基础设施; 中小型科技馆可打造“专题+特色”差异化发展路径, 精细化运营弥补资源不足。

[关键词] 科技馆 社会评价 第三方点评数据 文本情感分析

[中图分类号] N4; G261 **[文献标识码]** A **[DOI]** 10.19293/j.cnki.1673-8357.2026.01.006

科技馆作为提升公众科学素质、推动科学教育和创新的重要场所, 在实践中逐渐形成了覆盖城乡、注重普惠、实用高效的科技馆科普公共服务体系^[1]。科技馆中形式多样、内容丰富的教育活动能够激发公众的科学兴趣, 促进其科学理解和认同, 为不同层次的参观者提供多样化的教育服务^[2]。同时, 通过参与体验型展品为观众提供从实践中学习的途径也是科技馆生存与发展的重要价值所在^[3], 观众体验因而成为一个重要问题, 科技馆必须通过不断提高观众满意度, 扩大社

会吸引力, 才能充分发挥公民科普与教育职能^[4]。2024 年新修订的《中华人民共和国科学技术普及法》明确提出, “国家完善科普工作评估体系”“监测和评估科普事业发展成效”^[5], 以评促建也是推动科技馆事业发展的重要途径, 有利于贯彻和落实中央的科普工作精神与要求^[6]。

科技馆作为科普教育的重要场所, 其社会评价研究始于观众研究, 在 20 世纪末迅速发展。传统的科技馆社会评价主要关注观众对展览内容、展览形式、设施设备、教育活

收稿日期: 2025-07-25

作者简介: 赵洋, 中国科技馆科研管理部主任、研究员, 研究方向: 科技馆理论与实践, E-mail: zhaoyang@cstm.org.cn。徐扬为通讯作者, E-mail: yang.xu@pku.edu.cn。

动、参观体验、安全卫生等方面的反馈，评价手段以问卷调查和专家评估为主，包括利用平衡计分卡、德尔菲法和层次分析法等^[7]。

尽管传统方法具有一定的权威性，但由于样本量小、主观性强等局限，难以全面反映公众的真实感受。随着 Web 2.0 技术发展，第三方点评平台（如大众点评、美团、携程等）成为公众表达体验、分享意见的重要渠道，为科技馆社会评价提供了新的数据来源^[8]。近年来，利用在线评论数据进行社会评价、消费者行为分析和服务质量评估已成为趋势^[9-11]。在科技馆和博物馆领域，已有多项研究进行了相关探索。腊博辰等结合博物馆在线评论、主题分析与服务质量模型，提取关键评价指标，构建了服务质量评价体系^[12]；周瑞等基于携程网和大众点评网的留言文本，分析了公众对博物馆的需求侧重点，包括历史文化、公众教育、参观体验等方面^[13]；赵婕等以清华大学艺术博物馆为例，结合质性与量化研究方法，探讨了影响观众满意度的因素^[14]；王浩等对国家海洋博物馆 1 755 条游客评论进行主题分析，深入了解游客关注点与满意度影响因素^[15]；毛太田等应用主题识别与情感分析方法，研究了民族自治区博物馆的游客在线评论数据^[16]；邵隼等利用情感分析技术，对英国国家美术馆的社交媒体评论进行了研究^[17]。

综上所述，科技馆社会评价研究已从传统的问卷与专家评估法，逐步向结合大数据、文本挖掘与情感分析的新兴方向发展，为更全面、更客观地理解公众需求与改进服务质量提供了有力支撑。然而，现有研究多停留在数据收集和分析层面，尚未构建起系统化、多维度的科技馆社会评价体系，难以切实指导场馆优化运营^[18]。同时，随着大数据时代的到来，既有评价模式逐渐暴露出数据量不足、评价不够全面和时效性不强的问题。本研究整合第三方点评数据，引入机器学习等

大数据技术，构建包含服务质量、展览内容、教育活动、设施设备等核心要素的科技馆社会评价体系，并通过实证分析，验证该体系的可行性与有效性。

1 研究设计

1.1 理论框架

本研究从科技馆服务的价值主体（公众）出发，采集公众评价数据，利用自然语言处理和数据挖掘技术对公众意见进行主题提取和挖掘。以公众的真实需求为基础，对科技馆社会评价开展两阶段主题建模。第一阶段主题建模的主要目标是发现并提取现有研究中观众对于科技馆的体验关注点，比如科技馆的教育服务质量、参观体验满意度、科技创新能力和社会效益等关键领域，并识别出其他在以往研究中未被发现的关键因素，形成一级指标体系。由于评论数据较为庞杂，一次主题建模结果相对粗糙，因此，针对第一阶段的每组评论数据集再运用第二阶段主题建模，旨在分析出每个主题中的细分主题，为科技馆建设提供更加具体的建议。结合广大观众的关切程度（每个主题评论数量在总评论数的占比），赋予相关权重，最终形成一套科技馆社会评价体系。为了验证所建立的评价体系的适用性，实证阶段采用第三方评价数据对全国 400 余家科技馆进行评分，理论框架如图 1 所示。

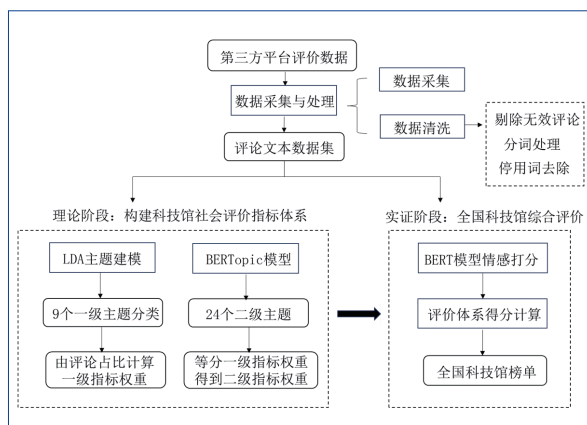


图 1 理论框架示意图

1.2 研究方法

数据采集使用网络爬虫技术，从网页中提取所需信息，包括评论者的评分、评价内容、评论时间等，形成科技馆社会评价体系基础数据集。数据预处理主要包括：（1）剔除无效评论，即去除缺乏有效信息的评论，如长度过短、没有具体描述的评论等；（2）分词处理，即将中文文本切分成有意义的字词，方便后续的文本分析和理解；（3）停用词去除，即去除如“的”“了”“是”等常见的无实际意义的停用词（词语来自哈工大停用词表），以减少噪音对分析结果的影响。

主题建模采用隐狄利克雷分布（Latent Dirichlet Allocation, LDA）主题模型^[17, 19]，

研究在此基础上引入了基于双向变换器模型（Bidirectional Encoder Representations from Transformers, BERT）词向量主题建模技术的BERTopic模型，以进一步挖掘评论中的细粒度主题及其相互关系，从而生成更具解释性的主题^[20]。情感分析旨在从文本中提取用户的情感倾向。虽然大部分第三方评论中含有用户评分，但与用户情感未必一致。如表1所示，尽管用户在评论中表达了某些不满或负面情感，其仍然倾向给予较高的评分，这表明传统的评分系统可能无法准确反映用户的真实感受。为了更准确地评估评论中的情感倾向，本研究采用了BERT模型，重新评估评论数据并更准确地量化每个帖子的情感倾向。

表1 用户评论案例

科技馆名称	评论内容	发布时间	用户评分
东北地区某省级科学技术馆	上大学的时候没有来科技馆，工作了，一定要来圆梦，抽点时间，来开开眼界！（虽然大多数都看不懂），遇到学校学生来学习，混在一群学生中，感叹自己已逝去的青春岁月，学生时代真是美好！转了一圈，有些仪器好像不好使，也可能是我不会用。	2020-04-11 18:32	5

1.3 数据来源

本研究使用的数据来自第三方点评网站，具体包括大众点评、美团和携程。实证阶段涉及科技馆共408家，覆盖国家和地方（省、市、县）各级各类科技馆。其中，特大型科技馆19家、大型科技馆41家、中型科技馆81家、小型科技馆267家。经过数据采集与预处理，本研究自2006年4月至2024年10月共收集133 371条评论数据，包括携程17 161条数据、大众点评70 761条数据、美团45 449条数据（不包括转载大众点评的数据）。

公众评论的主要关切包括孩子、体验、预约、知识、电影、项目、展厅、讲解、科技和交通。在外部条件中，科技馆的交通、地理位置和门票预约比较重要。在内部条件中，科技馆的展厅布局、项目（展品）的科技含量、电影、讲解等是公众的主要关切。

2 科技馆社会评价指标体系

2.1 主题识别

本研究首先通过词云工具对评论数据的宏观特点进行了数据探索，呈现评论数据中的高频关键词（见图2）。由高频关键词可知，

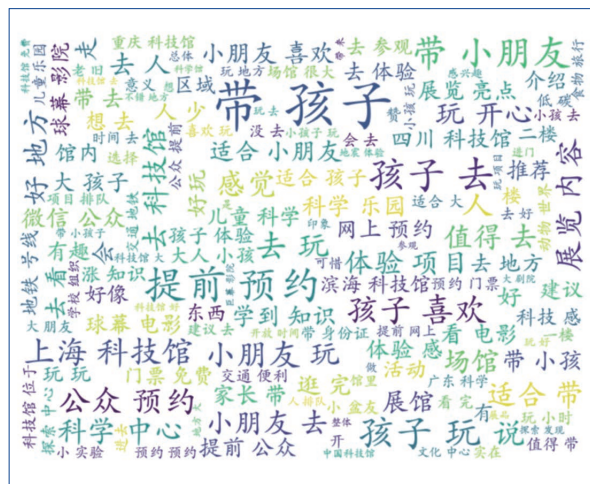


图2 评论高频关键词词云图

其次，本研究对评论数据进行主题建模，根据结果将其概括为 9 个主要主题领域（即一级指标），基本涵盖观众参观科技馆的主要目的及考虑因素。具体包括：儿童娱乐与科学互动、科技馆科学展示与文化体验、多楼层展区与互动科技、预约参观与票务信息、亲子参观与知识体验、参观体验与配套设施、设施维护与服务质量、环境评价与观感体验、交通便利与停车设施。

2.2 指标体系

基于 9 个主要主题领域，研究在主题建模的基础上进一步为其划分了更为具体的子主题，共包括 24 个子主题（二级指标）。综上，本研究建立了包含 9 个一级指标和 24 个二级指标的科技馆社会评价指标体系（见表 2）。一级指标相关权重依据每个主题的评论数在总数中的占比进行计算，如主题 5 “亲子参观与知识体验” 在 133 151 条评论中涉及 54 461 条，其权重为 41.0%。对于二级指标的权重，由于进行细粒度主题提取后，有相当

多的评论包含多个要素，因此对于二级指标采用平分所属一级指标权重的方法进行计算。如主题 3 “互动学习与互动科技” 的权重为 3.0%，则其两个二级指标的权重各为 1.5%。

3 实证研究结果

本研究依照情感分析 1~5 分的机器打分情况，结合各个分项指标的权重，计算出不同规模科技馆的社会评价得分情况。由于不同科技馆所处城市、设施具有较大差别（比如特大型科技馆大多位于一、二线城市，使用的科技手段先进、设备设施先进），观众对于特大型科技馆具有较高期待，评论的侧重也会和中小型科技馆有所不同。因此，本研究根据科技馆规模的不同进行了分类研究，以下是特大型科技馆、大型科技馆、中小型科技馆的具体研究结果。

3.1 特大型科技馆

如表 3 所示，特大型科技馆在“环境评价与观感体验”方面平均得分（3.495 分）明显高于大型科技馆（1.825 分）和中小型科技馆（1.544 分）。这是因为特大型科技馆多为综合性科技馆，展品多、场馆大、展品更新比较及时，所以科学性与视觉冲击会比较强烈。此外，特大型科技馆在核心科学教育指标方面，如“儿童娱乐与科学互动”“科学展示与文化体验”“互动学习与互动科技”，各馆得分均展现出强劲优势。如，华北地区某特大型科技馆 A 在“儿童娱乐与科学互动”指标上获得 4.588 分，“科学展示与文化体验”指标获得 4.676 分，“互动学习与互动科技”指标更是达到 4.692 分，这表明该馆在科普内容设计、互动体验形式上精

表 2 科技馆社会评价指标体系

指标序号	一级指标	权重/%	二级指标
1	儿童娱乐与科学互动	11.0	儿童娱乐与教育价值 视觉与趣味体验 家庭游玩体验 儿童科学启蒙
2	科学展示与文化体验	9.2	启发性科技展示 科技馆的文化与科教地位
3	互动学习与互动科技	3.0	展区中的互动学习 机器人等先进互动科技
4	预约参观与票务信息	11.5	游玩项目安排与提前预约 门票预约与场馆设施
5	亲子参观与知识体验	41.0	高性价比的全家游览体验 亲子互动与知识收获 值得推荐的参观体验
6	参观体验与配套设施	14.2	高峰期参观建议与排队体验 节假日参观感受与氛围 科技馆配套设施
7	设施维护与服务质量	5.6	设施维护 服务质量
8	环境评价与观感体验	1.3	优美环境与优质服务 游客流量与观感体验 自然景观与沉浸式体验
9	交通便利与停车设施	3.2	场馆附近设施 交通便利程度

准契合了公众对科技教育的核心期待，有效实现了寓教于乐的目标；西南地区某特大型科技馆 B 在“儿童娱乐与科学互动”指标上得分为 4.659 分，“科学展示与文化体验”指标得分为 4.684 分，在科技教育的深度与趣味性方面同样表现出色。

然而，在基础设施与服务保障维度，部分馆稍显不足。例如“设施维护与服务质量”指标得分差异显著，西南地区某特大型科技馆 B 仅得 1.645 分，华北地区某特大型科技馆 C 也仅得 1.939 分，侧面反映出部分特大型科技馆存在设施老化、维护不及时，或是工作人员服务意识薄弱、讲解专业性不足等问题。“参观体验与配套设施”得分同样不理想，多

在 4 分以下，如，华北地区某特大型科技馆得分仅为 3.987 分，意味着馆内餐饮、休息区等配套设施在品质、数量上难以满足高峰期大量游客需求，影响整体参观体验。

总分排名靠前的华北地区某特大型科技馆，凭借各指标得分的均衡性脱颖而出，其在“预约参观与票务信息”“亲子参观与知识体验”等指标上均取得较高分，表明该馆不仅在核心业务上表现突出，在服务细节与家庭观众需求满足方面也具备较强竞争力；而排名靠后的场馆，往往在某几个关键指标上存在明显失分项，如西南地区某特大型科技馆在“设施维护与服务质量”的低分，显著影响了整体排名。

表 3 特大型科技馆社会评价得分

	儿童娱乐与科学互动	科学展示与文化体验	互动学习与互动科技	预约参观与票务信息	亲子参观与知识体验	参观体验与配套设施	设施维护与服务质量	环境评价与观感体验	交通便利与停车设施	总分 ^①
华北地区某特大型科技馆 A	4.588	4.676	4.692	4.326	4.600	3.987	3.941	4.616	4.330	4.445
西南地区某特大型科技馆 B	4.659	4.684	4.633	4.651	4.643	4.350	1.645	4.731	4.737	4.444
华北地区某特大型科技馆 C	4.764	4.763	4.710	4.738	4.782	3.919	1.939	1.500	4.626	4.442
东北地区某特大型科技馆 D	4.706	4.770	4.773	4.605	4.790	4.064	1.629	0.000	4.466	4.404
西南地区某特大型科技馆 E	4.646	4.702	4.667	4.480	4.686	3.913	1.472	4.539	4.538	4.362
华南地区某特大型科技馆 F	4.551	4.443	4.609	4.079	4.618	3.694	3.091	4.553	4.324	4.305
华东地区某特大型科技馆 G	4.492	4.665	4.408	4.186	4.578	3.690	2.218	4.524	4.396	4.262
平均分	4.629	4.672	4.642	4.438	4.671	3.945	2.276	3.495	4.488	4.381

3.2 大型科技馆

如表 4 所示，各大型科技馆在不同指标上呈现出明显的得分差异。在总分方面，华北地区某大型科技馆 A 以 4.488 分位居榜首，该馆在诸如“儿童娱乐与科学互动”（4.627 分）、“科学展示与文化体验”（4.712 分）、“亲子参观与知识体验”（4.723 分）等核心教育展示类指标上均取得较高分，这与其独特的由碱厂旧址“冷却塔”元素改造的外观设计密不可分，独特的建筑风格不仅吸引了大量游客，也为科技馆的科学展示与文化体验增添了独特魅力，使其在相关评价维度中脱

颖而出；同时，其在“交通便利与停车设施”指标上获得 4.716 分，反映出该馆地理位置优越，周边交通配套完善，极大提升了观众的整体参观体验。

华中地区某大型科技馆 B 以 4.443 分紧随其后，在“科学展示与文化体验”（4.734 分）、“互动学习与互动科技”（4.718 分）等体现科技馆核心功能的指标上表现突出，说明该馆在科学展示内容与互动科技项目设置上得到了观众的高度认可。然而，该馆在“设施维护与服务质量”方面仅获得 1.738 分，与其他指标差距明显。

①此处“总分”表示加权后的分数，表 4、表 5“总分”同表 3。

从平均得分而言，大型科技馆（4.132分）低于特大型科技馆（4.381分），原因可能是展品更新方面相对逊色，其展品维修可能依靠第三方公司较多，自主维修的能力不足、及时性不够。从整体分项指标来看，大型科技馆在“儿童娱乐与科学互动”“科学展示与文化体验”“亲子参观与知识体验”等核心业务指标上普遍得分较高，平均得分均在4.6分以上。然而，在设施维护与服务质量、

参观体验与配套设施等硬件与服务相关指标上，得分明显偏低，“设施维护与服务质量”平均得分仅为1.884分，“参观体验与配套设施”平均得分2.412分，其中多家科技馆在这两项指标上得分低于2分，如华东地区多家大型科技馆在“参观体验与配套设施”方面得分均在1.3~1.5分区间。此外，在“环境评价与观感体验”指标上，多家华东地区的大型科技馆得分为0分。

表 4 大型科技馆社会评价得分

	儿童娱乐与科学互动	科学展示与文化体验	互动学习与互动科技	预约参观与票务信息	亲子参观与知识体验	参观体验与配套设施	设施维护与服务质量	环境评价与观感体验	交通便利与停车设施	总分
华北地区某大型科技馆 A	4.627	4.712	4.656	4.612	4.723	3.787	3.434	4.453	4.716	4.488
华中地区某大型科技馆 B	4.671	4.734	4.718	4.580	4.680	4.205	1.738	4.708	4.707	4.443
华东地区某大型科技馆 C	4.682	4.678	4.737	4.583	4.637	4.208	1.746	4.286	4.691	4.417
西南地区某大型科技馆 D	4.783	4.648	4.573	3.854	4.818	3.540	2.942	4.802	0.000	4.239
华东地区某大型科技馆 E	4.814	4.660	4.734	4.729	4.865	1.461	1.563	0.000	4.833	4.088
华东地区某大型科技馆 F	4.691	4.703	4.721	4.649	4.711	1.515	1.883	0.000	4.442	4.019
华东地区某大型科技馆 G	4.631	4.456	4.871	4.556	4.721	1.335	1.197	0.000	4.369	3.921
华南地区某大型科技馆 H	4.521	4.596	4.679	4.477	4.700	1.345	1.453	0.000	4.355	3.914
华东地区某大型科技馆 I	4.780	4.680	2.292	4.439	4.742	1.424	1.389	0.000	4.597	3.907
东北地区某大型科技馆 J	4.624	4.742	4.736	4.546	4.528	1.296	1.493	0.000	4.788	3.887
平均分	4.682	4.661	4.472	4.503	4.713	2.412	1.884	1.825	4.150	4.132

3.3 中小型科技馆

如表 5 所示，结合分项指标数据和实践经验分析，中小型科技馆大部分是县级科技馆，人流量较少，服务人员也不多，展品的丰富度较特大型和大型科技馆逊色很多。此

外，中小型科技馆在服务质量、设施维护，以及配套设施等硬件条件方面，明显低于特大型科技馆和大型科技馆。因此，中小型科技馆吸引观众一般可以从举办大型活动、馆校合作、增加教育活动类型等方面入手。

表 5 中小型科技馆社会评价得分

	儿童娱乐与科学互动	科学展示与文化体验	互动学习与互动科技	预约参观与票务信息	亲子参观与知识体验	参观体验与配套设施	设施维护与服务质量	环境评价与观感体验	交通便利与停车设施	总分
华东地区某中型科技馆 A	4.736	4.733	4.760	4.503	4.801	3.630	3.746	4.699	4.554	4.518
华北地区某中型科技馆 B	4.519	4.638	4.594	4.412	4.633	3.887	3.729	4.311	4.516	4.430
华中地区某中型科技馆 C	4.633	4.821	4.632	4.580	4.551	4.437	1.447	0.000	4.896	4.353
西北地区某中型科技馆 D	4.530	4.074	4.447	4.506	4.529	3.769	1.138	0.000	3.955	4.107
东北地区某中型科技馆 E	4.750	4.685	4.574	4.275	4.693	1.466	1.778	0.000	4.817	3.968
华南地区某中型科技馆 F	4.486	4.573	4.760	4.530	4.686	1.389	1.524	0.000	4.639	3.930
华东地区某中型科技馆 G	4.744	4.500	4.531	4.410	4.513	1.384	1.821	0.000	3.925	3.854
华南地区某中型科技馆 H	4.492	4.501	4.705	4.263	4.430	1.362	1.574	4.767	4.133	3.832
东南地区某小型科技馆 I	4.470	4.371	4.333	4.437	4.521	1.234	1.667	0.000	4.308	3.794
华东地区某中型科技馆 J	4.493	4.548	4.438	4.875	4.372	1.103	1.586	1.667	3.938	3.792
平均分	4.585	4.544	4.577	4.479	4.573	2.366	2.001	1.544	4.368	4.058

4 分析与讨论

从现有数据来看,一方面,各规模科技馆的前3个一级指标“儿童娱乐与科学互动”“科学展示与文化体验”“互动学习与互动科技”得分都很高,这说明科技馆的办馆理念有效传递给了观众,家长与孩子都能在科技馆“体验科学、启迪创新”,展现了科技馆与其他科技类博物馆的独特之处。另一方面,就共性问题而言,各规模科技馆在交通与环境满意度方面仍有较大提升空间。尤其是非核心城市的科技馆,常常因地理位置偏远、公共交通不便或高峰期交通拥堵而影响社会评价。在馆址不可改变的前提下,科技馆可加强与城市交通管理部门合作,增设指引标识、设置专属停车区、优化高峰期动线引导;在环境营造方面,可引入艺术装置、主题背景墙、沉浸式声音与光效等空间设计手段,构建科技与美学交融的参观环境,提升整体到馆体验。就具体差异而言,各类科技馆亦有不同,以下是针对特大型科技馆、大型科技馆和中小型科技馆的具体分析与讨论。

4.1 特大型科技馆:突出科教展示优势,补足服务短板

特大型科技馆普遍在教育核心指标上表现突出,充分展现了其在内容设计、展陈科技水平与教育功能上的资源优势,能够满足公众对高水平科普体验的期待。然而,从“设施维护与服务质量”“参观体验与配套设施”“环境评价与观感体验”等方面看,部分场馆得分偏低,反映出有些场馆在日常运营和服务细节方面存在短板。特大型科技馆应积极发挥场地大的优势,将科学性与艺术性相结合,打造地标建筑或热门展览,吸引观众并影响其评价。同时,由于观众人流量大,特大型科技馆在节假日和寒暑假期间流量控制尤为重要。例如,中国科技馆通过控制每

日接待人次总量,高峰时段需要抢票,馆内也实时更新各展区流量示意图,提醒观众做好参观路线规划。综上,特大型科技馆应将部分资源和管理重心从“建设性投入”转向“运行性维护”,强化精细化管理,提升现场服务质量与环境体验,尤其在高峰时段人流疏导、讲解服务、人文关怀等方面加强建设,实现“硬科技”与“软服务”并重。

4.2 大型科技馆:均衡发展核心功能,优化基础设施

大型科技馆整体在科技展示、亲子互动、儿童体验等方面表现较为均衡,尤其是一些科技馆借助旧厂房改造或主题化展览塑造了独特文化符号,增强了公众黏性。但基础设施类指标普遍偏低,显示出当前大型科技馆在规模增长之后,缺乏系统性的基础设施升级计划,且软性服务投入不足。应以“功能性更新”思维为导向,在观众服务能力角度扩展提升,一方面增加各类教育活动,另一方面提升餐饮、休闲、文创等方面的服务水平,构建更友好、舒适的公共空间体验;同时建立定期服务质量评估机制,强化员工培训与志愿者讲解系统,提升整体服务感知度。

4.3 中小型科技馆:打造“专题+特色”差异化发展路径,精细化运营弥补资源不足

中小型科技馆在核心业务指标上表现不弱,尤其部分“专题+特色”鲜明的场馆如企业科技馆、专业类馆(如光学主题馆)在互动体验、科学展示、亲子互动等方面均获得高度评价,体现出其在内容精准化、定位差异化方面的探索成效。但在设施维护、配套服务等硬件方面得分显著低于特大型与大型科技馆,说明受制于资金、场地、运维人员等资源瓶颈,中小型科技馆更易在运营后期出现质量下降、服务跟不上的情况。中小型科技馆应聚焦“轻资产、强内容”的发展思

路, 强化主题策展与教育活动原创性, 通过品牌化、节日活动、企业联动等方式增强社会影响力; 同时, 应探索中小型场馆的“精细运营”机制, 如小程序预约、错峰参观、社会共建、低成本智能导览等创新服务手段, 实现小而精、小而美的长效发展。

5 结语

本研究基于全国 400 余家科技馆的第三方点评数据, 深度融合主题建模、主题层次聚类与情感分析方法, 构建起一套科学、客观的科技馆社会评价指标体系。在模型框架设计上, 从海量评论中抽取出 9 个一级主题,

初步搭建起科技馆社会评价体系架构; 随后通过语义嵌入、降维、聚类等方法, 进一步细化出 24 个二级指标, 实现从宏观到微观的深入剖析; 最后利用情感分析技术, 对每条评论精准打分, 将公众情感倾向量化融入评价体系。在此基础上, 本研究对从大众点评、美团、携程等第三方点评平台收集的 133 371 条评论数据进行实证分析, 对不同规模科技馆进行评估, 验证了指标体系的可行性, 并从数据分析的结果出发, 对各规模的科技馆提供了发展建议。未来研究可以从拓展数据来源角度开展, 整合社交媒体动态、场馆运营日志等多源数据, 构建动态评价体系。

参考文献

- [1] 赵洋, 马宇罡, 苑楠, 等. 中国特色现代科技馆体系建设: 回顾与展望 [J]. 科普研究, 2021, 16(4): 80-86.
- [2] 陈怡, 李佳, 龙美玲, 等. 科技馆教育活动比较研究——以美国纽约科学馆、日本科学未来馆、中国科学技术馆为例 [J]. 科普研究, 2022, 17(4): 71-78.
- [3] 陈善蜀. 科技馆对社会公众实行免费开放的思考 [J]. 科普研究, 2009, 4(3): 30-33.
- [4] 张惠聪, 王珊珊, 樊庆, 等. 基于服务主导逻辑的科技馆观众满意度研究——以中国科技馆的观众满意度调研为例 [J]. 科普研究, 2022, 17(4): 57-64.
- [5] 中华人民共和国科学技术普及法 (最新修订版) [M]. 北京: 科学普及出版社, 2025.
- [6] 佟贺丰, 于洁, 董克. 科技馆综合评价体系构建与实践 [J]. 科普研究, 2020, 15(5): 32-38.
- [7] 宋健, 权赣. 我国县域流动科技馆运营绩效评价初探 [J]. 海峡科学, 2024(1): 168-171.
- [8] 郑建, 宋红娟. 第三方点评对旅游者餐饮消费行为影响研究 [J]. 科技和产业, 2021, 21(11): 92-98.
- [9] 师荣蓉, 张教萌. 中国经济高质量发展的社会评价: 基于微博情感分析的视角 [J]. 统计与决策, 2021, 37(24): 180-184.
- [10] 洪江涛, 陈榴寅, 黄沛. 第三方点评网站对餐饮企业品牌形象与消费者行为的影响研究——以大众点评网为例 [J]. 财贸经济, 2013(10): 108-117.
- [11] 魏欣, 盖泓烨. 基于 LSTM-IPA 模型的冰雪旅游消费服务质量评估研究——以哈尔滨市为例 [J]. 商业经济研究, 2024(21): 177-180.
- [12] 腊博辰, 张健, 齐林. 基于在线评论的博物馆服务质量评价 [J]. 科技和产业, 2024, 24(12): 117-122.
- [13] 周瑞, 李卓. 以“公众关切”为导向的博物馆线上文化传播策略——基于网络用户评价的语料分析 [J]. 中国博物馆, 2024 (1): 62-68.
- [14] 赵婕, 焦越涵. 清华大学艺术博物馆互联网评价与分析 [J]. 科学教育与博物馆, 2022, 8(4): 65-74.
- [15] 王浩, 方俊涛. 基于 LDA 模型对国家海洋博物馆游客在线评论的主题分析 [J]. 科技和产业, 2024, 24(12): 224-230.
- [16] 毛太田, 吴宇豪, 黄文佳. 自治区民族博物馆的游客在线点评内容挖掘与情感分析 [J]. 经济地理, 2023, 43(8): 229-236.
- [17] Shao J, Ying Q, Shu S, et al. Museum Tourism 2.0: Experiences and Satisfaction with Shopping at the National Gallery in London [J]. Sustainability, 2019, 11(24): 7108.
- [18] 徐扬, 周雷, 赵洋, 等. 科技馆社会评价国内外研究综述 [J]. 自然科学博物馆研究, 2025, 10(1): 59-68.
- [19] Blei D M, Ng A Y, Jordan M I. Latent Dirichlet Allocation [J]. Journal of Machine Learning Research, 2003, 3(Jan): 993-1022.
- [20] Grootendorst M. BERTopic: Neural Topic Modeling with a Class-Based TF-IDF Procedure [PP/OL]. arXiv preprint (2022-03-11) [2025-12-11]. <https://doi.org/10.48550/arXiv.2203.05794>.

(编辑 颜 燕 和树美)

science popularization models derived from ‘science popularization+scene’ and ‘scene + science popularization’. First, scenified science popularization reconstructs the production scene of scientific knowledge through technical means, transforming abstract scientific content into perceptible embodied experiences for the audience. Second, popularized scenes embed scientific elements into social spaces such as shopping malls and communities, activating the science popularization potential of non science contexts. Third, scenic popularization achieves the integration of multiple attributes of scenes through functional superposition and spatiotemporal reorganization. Fourth, scene-based science popularization promotes the integration of scientific spirit and daily practice through a context-constructed communication paradigm, thereby providing new ideas and methods for science popularization practice.

Keywords: science popularization; scene; scene theory; science popularization scene

CLC Numbers: N4 **Document Code:** A **DOI:** 10.19293/j.cnki.1673-8357.2026.01.005

Empirical Study and Construction of Social Evaluation Index System for Science and Technology Museums Based on Third-Party Comment Data

Zhao Yang¹ Xu Yang² Wang Meili¹ Li Dan¹ Yang Yang¹ Zhou Lei²

(China Science and Technology Museum, Beijing 100010) ¹

(Department of Information Managemont, Peking University, Beijing 100871) ²

Abstract: With the increasing impact of science and technology museums on enhancing public scientific literacy and science education, how to scientifically and objectively evaluate their social impact has become a critical issue. Based on third-party comments collected from over 400 science and technology museums nationwide, this paper employs natural language processing techniques such as topic modeling and hierarchical clustering to establish a social evaluation index system for science and technology museums. The system includes 9 primary indicators and 24 secondary indicators. Machine learning algorithms are used to conduct empirical analysis on 133 371 valid comments from third-party platforms such as Dianping, Meituan, and Xiecheng. The results show that the three primary indicators—“childrens’ entertainment and ecience interaction” “science exhibitions and cultural experiences, ” and “interactive learning and technology” —score high across all types of science and technology museums. However, there remain significant improvement possibilities in transportation and environmental satisfaction. The study recommends that mega-sized science and technology museums should leverage their strengths in science education exhibitions while addressing service shortcomings; large-sized ones should balance core functions and optimize infrastructure; while medium and small-sized ones can pursue a differentiated development path of “specialization+features” and refine operations to compensate for resource limitations.

Keywords: science and technology museum; social evaluation; third party comment data; text sentiment analysis

CLC Numbers: N4; G261 **Document Code:** A **DOI:** 10.19293/j.cnki.1673-8357.2026.01.006