

基于大数据的花山岩画数字化传播效果与公众认知研究

申航硕

桂林信息科技学院, 广西桂林, 541004;

摘要: 在数字经济助力文化遗产传承的背景下, 花山岩画的数字化传播已成为重要手段, 但其实际效果与公众认知状况尚缺乏实证研究。本研究基于大数据技术, 旨在系统评估花山岩画数字化传播的效果, 并深入分析公众认知特征。通过采集社交媒体、新闻门户及视频平台的公开数据, 综合运用情感分析、主题建模与社会网络分析等方法进行挖掘。研究结果表明, 当前传播已形成广泛覆盖, 但公众情感以中性为主, 认知多聚焦于“旅游”与“遗产”标签, 对文化内涵的理解尚显不足。未来传播需从广度转向深度, 通过精准内容策略深化公众价值认知, 以提升活态传承效能。本研究为文化遗产传播效果的量化评估提供了数据支持与分析范式。

关键词: 花山岩画; 数字化传播; 大数据分析; 公众认知; 情感分析

DOI: 10.64216/3080-1516.25.04.056

1 引言

1.1 研究背景与问题的提出

在数字经济时代, 国家文化数字化战略的施行, 给文化遗产保护与传承带来了新的发展机会。花山岩画是世界文化遗产, 有着重要的历史与艺术价值, 由于其地理位置偏远, 文化内涵不易解读, 给传承与推广带来了巨大挑战。近年来, 相关机构投入大量资源对数字化传播开展实践, 但实际效果以及公众的认知如何, 还缺少系统性评估和深入分析。现有的研究大多只围绕在定性描述方面, 缺少在大数据技术的量化评估与深度洞察方面开展研究, 这就成了本研究的问题所在。通过引入大数据分析的方法, 从多个维度对花山岩画的数字化传播成效进行系统评估, 深入探究公众认知的结构与特征, 弥补现有研究的欠缺之处。

1.2 研究目的与意义

本研究采用大数据分析的方法, 对花山岩画数字化传播的效果加以系统评估, 从多方面展现公众的认知结构与情感态度。其一, 对传播的广度以及公众的参与程度进行量化评估。其二, 要识别出公众讨论中的核心议题及其情感导向。其三, 依据研究结果提出可操作的传播优化策略。本研究从理论上能够丰富文化遗产数字化传播效果评估理论, 推动数据驱动研究范式在该领域的应用。在实践中, 花山岩画及相关文化遗产管理机构能从研究成果里得到决策参考, 实现传播效能和文化传承力的提升。

1.3 研究范围与限制

本研究重点探究中文互联网生态环境和花山岩画有关的公众讨论, 以及内容传播情况, 数据的来源主要有主流社交平台和旅游社区平台。需指出的是, 本研究有一定局限性, 其一, 数据收集难以涵盖所有网络平台, 可能有样本遗漏。再者, 文本分析方法在识别语境和反讽时, 会有一定误差。最后, 研究结论主要体现的是经常使用互联网的活跃公众群体的认知情况, 无法完全涵盖线下公众或者非网络用户的观点。

2 文献综述

2.1 国内外研究现状

国内外学者在文化遗产数字化传播方面已经取得了一定进展。早期的研究主要集中在数字存档、虚拟博物馆建设这些方面, 然而近年来, 研究更多地关注沉浸式体验以及社交媒体传播。在对花山岩画的深入研究里, 学者们大多是凭借考古学、保护技术以及旅游开发等方面为契机, 其数字化传播的效果与公众的认知, 还没有系统地加以探讨。从方法论的角度来讲, 传统研究多用问卷调查和访谈, 在大数据分析方面还处于起步阶段, 有着较大的发展前景。虽然有部分研究开始尝试用社交媒体数据来做分析, 但大多集中在单一平台或者小规模样本上, 很少有综合考察多平台、长时段的情况, 这也为本研究提供了进一步创新的机会。

2.2 主要理论框架与关键概念

本研究把传播效果理论、建构主义学习理论以及文化表征理论当作主要的理论框架。这些理论对理解公众怎样接收、解释以及重构文化遗产信息有着重要的视角。

在此基础上,本研究对核心概念予以明确界定,即数字化传播效果包含认知、情感以及行为这三个层面。公众认知着重阐述受众对文化符号的理解及意义构建。大数据分析,就是利用计算技术从海量数据里提取有价值模式的方法。

2.3 现有研究的争议与不足

当下研究存在不少争议,主要是技术导向跟人文导向之间的张力,也就是怎样平衡数字化手段和文化内涵传播的关系。另外,现有的研究大多只是进行定性分析,缺少大规模定量数据的支撑。宏观策略方面有较多讨论,但对公众微观认知的深入分析却十分欠缺。研究方法相对传统,缺乏有效利用新兴数据科学技术。总的来讲,现有的研究对花山岩画数字化传播的传播效果,没有系统性的评估,无法基于海量数据深入地描绘公众的认知图谱。本研究力求填补这一空白,通过引入大数据分析的方法,在理论方面推动文化遗产传播研究范式创新,在实践上给相关机构提供精准、可靠的决策依据,进而从方法论与应用价值上作出创新贡献。

3 研究方法

3.1 研究设计与方法论

本研究运用数据驱动的探索性研究方法,系统性地搜集并分析多源网络数据,旨在客观展现花山岩画数字化传播的效果及公众的认知特点。研究按照“数据采集→清洗处理→多维度分析→结果解读”的技术流程来开展,实现对研究的科学性和文化遗产研究的人文属性同时兼顾。按照传播效果理论和公众认知逻辑,构建一个“传播广度-参与深度-认知结构-情感倾向”四维分析框架,以此保证研究能涵盖数字化传播的关键维度。

3.2 数据收集方法

研究的数据来自新浪微博、微信公众号、抖音、哔哩哔哩,还有携程、马蜂窝等旅游社区平台,时间从2016年(花山岩画申遗成功)到2023年。使用Python Scrapy框架进行数据采集,以“花山岩画”、“左江岩画”、“世界遗产花山”等为核心关键词,共获取有效文本数据12.8万条。在数据收集时,严格遵循各平台的数据使用协议,对个人信息也做了匿名化处理。为保证数据全面且具有代表性,研究还对不同平台制定了不同的采集策略,像在微博上重点抓取原创帖文和热门评论,在旅游

社区就聚焦游记和评分数据。

3.3 数据分析方法

首先要数据清洗并进行预处理,像去重、去除无关内容这些操作都是要做。其次运用Jieba分词工具对中文进行分词,再进行去停用词处理。接着借助停用词表(像“的”“了”之类的无意义词汇)来去除冗余信息,从而构建出标准化分析语料。最后借助多维度分析技术对传播效果予以评估:借助时间序列分析声量的变化,运用描述性统计分析各平台的公众参与度。采用基于BERT预训练的情感分类算法对情感倾向进行分析,把公众情感划分成积极、中性、消极这三类,进而计算各类情感的占比分布。采用社会网络分析法对认知结构加以分析,把高频关键词当作节点,共现关系当作边构建语义网络,借助网络密度、中心度等指标来解析公众认知的结构特点。

3.4 研究的效度与信度

为确保研究质量,采用多种方法来保证效度与信度,借助多平台数据交叉验证来提升结构效度,邀请2位文化遗产研究领域的专家对LDA主题模型的命名结果予以人工校验,以保证内容效度。随机抽取500条文本数据,由2名研究人员进行独立情感标注,计算科恩卡帕系数(Cohen's Kappa=0.82),表明编码一致性达到较高水平。另外,要把数据采集参数、预处理流程以及分析算法设置完整地记录下来,然后形成研究方法说明书,保证其他研究者能重复验证研究结果。

4 数据分析与结果

4.1 传播效果数据分析结果

数据分析表明,花山岩画相关讨论声量有明显波动,年均发文量为2.1万条。声量峰值和重要时间节点密切相关,2016年申遗成功后单日声量达1,842条,2021年世界遗产日期间声量增长了3.2倍。各平台的数据显示,微博的占比最高,达到38.7%,其次是抖音,占27.3%,旅游社区占比为18.5%。在用户参与度上,短视频平台的互动率明显比其他平台高,抖音平台平均点赞数是文字平台的4.3倍,说明视觉化传播形式更易激起公众互动意愿。

4.2 公众情感分析结果

情感分析结果表明,公众情感分布以中性为主(占

比 62.3%)，积极情感占比 31.5%，消极情感占比 6.2%。从不同平台情况，旅游社区积极情感占比最高，达 45.8%，微博平台里情感占比最大的是中性情感，占比为 71.2%。进一步分析后发现，正向情感较多，像“震撼”“值得”等词汇出现频率颇高。消极情感主要体现在交通不便、解说信息不足以及门票价格偏高等实际问题上。

4.3 公众认知主题挖掘结果

采用 LDA 主题模型，识别并确定了五个核心讨论主题。其中，旅游体验主题占比最高，达到了 24.7%，该主题主要聚焦于交通方式、门票价格以及观光路线等方面。文化遗产价值主题占比达 21.3%，其涉及世界遗产的状况、历史价值等。18.9%的文化符号主题，其中关注岩画图案、人像、符号意义等。保护与开发主题占比 17.2%，包含保护措施、旅游开发等方面。民俗文化主题占比达 17.9%，并且会探讨有关壮族文化、传统习俗之类的。主题分布显示，公众讨论还是以实用性和体验性内容为主，对文化深层价值的关注比较有限。

4.4 公众认知语义网络分析

语义网络分析表明，公众认知有着明显的中心 - 边缘结构。“花山岩画”是核心节点，它和“世界遗产”、“广西旅游”等概念联系较为紧密。要知道，“祭祀仪式”“铜鼓文化”这些深层次文化概念大都存在于网络的边缘地带，跟核心节点的联系也相对较弱。网络密度达 0.36，这表明公众的认知结构比较松散，还没形成系统化的文化认知体系。

5 讨论

5.1 对研究发现的综合解读

本研究指出，花山岩画数字化传播里有“高声量与浅认知并存”的情况。一方面，传播声量在重要时间节点有明显提升，这表明数字化传播在扩大影响力上有所成效。另一方面，公众对旅游体验等表层内容讨论较多，对文化内涵的认识不够系统，缺少关注点。这种现象可能与下面因素有关，首先就是传播内容同质化程度颇深，缺乏深度文化解读。其次，平台算法所偏好的内容较为轻松易懂。最后，文化符号跟现代生活的连接不够，公众难以产生深度共鸣。

5.2 与现有理论的对话

研究发现与传播效果理论存在多方面的契合，公众

对花山岩画的认知更倾向于“使用与满足”理论所提及的工具性利用，而非深度卷入。这与建构主义理论相契合，公众认知是传播内容与自身经验共同构建起来的。另外，研究结果也证实了文化表征理论，表明文化符号在传播时被简化、重构，原有的文化内涵在一定程度上被旅游消费话语所取代，这反映出商业逻辑对文化传播的影响。这一发现让我们明白，在数字化传播里，得更看重文化价值的精准传递，别让过度商业化造成意义的流失。

5.3 实践启示与对策建议

根据研究成果，对于如何改进提出以下优化建议：在内容方面，要从简单宣传转向深度对话，开发故事化且解读性强的文化内容。从技术方面来讲，能结合 AR、AIGC 这类技术，创新文化体验的形式，把理解门槛降低。在策略方面，建议搞分众化传播，给不同受众设计有区别的内容。另外，要建立长久的监测机制，定期对传播效果加以评估，以便于及时对传播策略做出调整。

5.4 研究局限性

本研究存在局限性，其一，数据来源仅限中文互联网平台，没有涵盖所有传播渠道。再者，文本分析方法在处理图像、视频等多模态内容时能力有限。最后，研究结论主要体现线上活跃群体的认知情况。这些局限给未来研究指出了方向，像开展多平台对比研究、开发多模态分析技术之类的方式。在日后的研究里，要进一步拓宽数据范围，添加更多分析维度，借此更好地把握花山岩画数字化传播效果与公众认知特征。

6 结论与展望

6.1 主要研究结论

本研究运用大数据分析手段，对花山岩画数字化传播的效果予以系统评估，深入剖析公众认知特征，首先指出，数字化传播在扩大声量上成效显著，声量峰值和重要时间节点紧密相连，可公众参与以浅层互动为主，深度参与明显欠缺。其次，公众情感态度以中性为主，占比达 62.3%，积极情感多与旅游体验有关，消极情感多是设施与服务方面的问题。最重要的是，公众的认知有明显的“标签化”与“浅层化”特点，讨论的主题主要集中在旅游体验（占比 24.7%）和遗产身份（占比 21.3%）上，对文化内涵、艺术价值等深层内容的关注明显不够，语义网络分析表明，文化核心概念在认知网络里

处于边缘位置。

6.2 研究的创新性与贡献

本研究的创新之处主要体现在方法论与应用价值这两方面，方法论上，创新地把大数据分析范式引入文化遗产传播研究领域，构建起了包含声量分析、情感计算、主题建模和社会网络分析的多维研究框架，给文化遗产传播效果评估提供了可量化、能重复的分析模型。就研究价值而言，这突破了传统定性研究的局限，以数据驱动的方式，精准地揭示出公众认知的深层结构，发现了“高声量与浅认知并存”这一核心问题，给花山岩画及相关文化遗产的传播实践提供了实证依据，并给出了精准的优化方向。

6.3 未来研究方向

根据本研究的发现与局限，往后的研究能从下面几个方向深入拓展。首先，要开展多模态数据的融合分析，把图像、视频等多媒体内容的分析方法引入进来，便于更全面地掌握数字化传播的整体情况。其次，拓展研究的边界，把分析范围拓展到国际社交媒体平台，探究花山岩画在国际视野里的传播效果与认知特点。接着，运用纵向追踪的研究方法，长期观察公众认知的变化轨迹，以揭示认知形成的动态机制。最后，要推动交叉学科合作，把计算科学和文化遗产研究深度融合起来，开发专门用于文化遗产传播分析的计算工具与理论模型，让研

究方法体系更丰富。

参考文献

- [1] 蒋楠, 张菁. 数字人文视角下文化遗产信息的演进、呈现与管理[J]. 同济大学学报(社会科学版), 2023, 34(06): 82-93.
- [2] 胡昊天, 邓三鸿, 张逸勤, 等. 数字人文视角下的非物质文化遗产文本自动分词及应用研究[J]. 图书馆杂志, 2022, 41(08): 76-83. DOI: 10.13663/j.cnki.lj.2022.08.010.
- [3] 黄晓瑜, 徐祥伍. 元宇宙语境下花山岩画文化资源数字化传播路径与展望[J]. 边疆经济与文化, 2024, (02): 85-88.
- [4] 申楠. 非遗数字化传播战略SWOT模型研究[J]. 同济大学学报(社会科学版), 2023, 34(01): 58-67.
- [5] 魏建中. 数字技术赋能非物质文化遗产保护路径创新——基于区块链的视角[J]. 数字技术与应用, 2024, 42(01): 104-106.

作者简介: 申航硕(1996.01), 女, 汉族, 辽宁省丹东市, 硕士研究生/助教, 桂林信息科技学院, 研究方向: 艺术设计。

基金项目: 2024年度桂林信息科技学院人才引进科研启动基金项目“数字经济助推广西花山岩画活态传承研究”(项目编号: XJ2024003)。