

# 多模态赋能京音乐创新：基于情感化设计的模式构建与实践研究

吕柯征, 邓逸钰\*

北京印刷学院新媒体学院, 北京

收稿日期: 2025年7月13日; 录用日期: 2025年8月3日; 发布日期: 2025年8月18日

## 摘要

目的: 在全球化与数字技术浪潮下, 京音乐因单一听觉表现形式, 面临传播受限、受众萎缩等问题。本研究以唐纳德·诺曼的情感化设计三层次理论为理论框架, 结合多模态技术特性, 深入探讨赋能京音乐创新模式的构建路径与设计方法, 挖掘其在传承与发展中的创新潜力, 为京音乐的现代化创新实践提供理论支撑与实践指导。方法: 研究基于数字化技术与数字媒体创作深度融合的趋势, 聚焦于北京市智化寺京音乐作为研究对象, 以情感化设计理论为核心方法, 深入挖掘其在京音乐传承创新中的应用路径。结果: 研究证实, 多模态技术与情感化设计三层次理论的融合, 有效拓展了京音乐的呈现维度与艺术表现力, 并且基于游戏引擎创作的AR、VR等实践案例, 在本能层提升用户的视觉与触觉反馈创新体验; 在行为层, 利用沉浸式交互设计提升用户互动体验与操作愉悦感; 在反思层, 借助多模态传播生态与数字科普应用, 传递京音乐的文化价值, 引发深层情感认同。证实多模态技术在情感化设计三层次理论的指导下拓展了京音乐呈现维度, 提升传播效能并革新传承模式。结论: 多模态技术与情感化设计三层次理论的深度融合, 为单一内容呈现的创新创作提供了有效路径。通过满足受众多层次情感需求, 该模式不仅可增强沉浸感、提升感官体验, 还可丰富创新艺术设计与表现形式。未来, 随着数字技术进步, 多模态与人工智能、大数据等技术融合, 将为京音乐创作带来更多可能, 推动其在传承与发展中存续文化活性。

## 关键词

情感化设计, 多模态融合, 京音乐, 沉浸感, AIGC

## Multimodal Empowerment of Jing Music Innovation: Research on Model Construction and Practice Based on Emotional Design

\*通讯作者。

文章引用: 吕柯征, 邓逸钰. 多模态赋能京音乐创新: 基于情感化设计的模式构建与实践研究[J]. 艺术研究快报, 2025, 14(3): 329-337. DOI: 10.12677/ar1.2025.143051

---

## Kezheng Lyu, Yiyu Deng\*

School of New Media, Beijing Institute of Graphic Communication, Beijing

Received: Jul. 13<sup>th</sup>, 2025; accepted: Aug. 3<sup>rd</sup>, 2025; published: Aug. 18<sup>th</sup>, 2025

---

### Abstract

**Objective:** In the context of globalization and the wave of digital technology, Beijing Music faces challenges such as limited dissemination and shrinking audiences due to its single auditory expression form. This study takes Donald Norman's three-level theory of emotional design as the theoretical framework, combines the characteristics of multimodal technology, deeply explores the construction path and design methods for empowering the innovative model of Beijing Music, excavates its innovative potential in inheritance and development, and provides theoretical support and practical guidance for the modern innovative practice of Beijing Music. **Method:** Based on the trend of deep integration of digital technology and digital media creation, the research focuses on the Beijing Music of Zhihua Temple in Beijing as the research object, takes the emotional design theory as the core method, and deeply explores its application path in the inheritance and innovation of Beijing Music. **Result:** The study confirms that the integration of multimodal technology and the three-level theory of emotional design effectively expands the presentation dimensions and artistic expressiveness of Beijing Music. Practice cases such as AR and VR created based on game engines enhance users' innovative experiences of visual and tactile feedback at the instinctive level; at the behavioral level, immersive interactive design is used to enhance users' interactive experience and operational pleasure; at the reflective level, the cultural value of Beijing Music is transmitted through multimodal communication ecology and digital popular science applications, triggering deep emotional identification. It is confirmed that under the guidance of the three-level theory of emotional design, multimodal technology expands the presentation dimensions of Beijing Music, improves communication efficiency, and innovates the inheritance model. **Conclusion:** The deep integration of multimodal technology and the three-level theory of emotional design provides an effective path for the innovative creation of single-content presentation. By meeting the multi-level emotional needs of the audience, this model can not only enhance the sense of immersion and improve the sensory experience, but also enrich the innovative artistic design and expression forms. In the future, with the progress of digital technology, the integration of multimodal technology with artificial intelligence, big data and other technologies will bring more possibilities for the creation of Beijing Music, and promote it to shine in inheritance and development.

### Keywords

Emotional Design, Multimodal Integration, Beijing Music, Immersive Experience, AIGC

---

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 绪论

在全球化与数字技术深度交融的时代背景下, 传统艺术文化的传承发展既迎来机遇, 也面临挑战, 京音乐这一承载五百年历史底蕴的文化遗存, 在当代传播中陷入困境。此时, 情感化设计理论的引入为破局提供了新的思路。

情感化设计理论由唐纳德·诺曼(Donald Arthur Norman)提出, 包含本能层、行为层、反思层三个维度, 强调通过感官刺激、交互体验与情感共鸣的协同, 深度激活用户对文化内容的感知与认同。而京音乐当下受制于单一听觉表现形式, 在多元文化冲击下传播渠道狭窄、年轻受众流失, 正需要借助该理论的多元视角寻求突破: 从本能层看, 多模态技术可整合文本、图像、音频等媒介, 打破单一感官局限; 行为层上, 能依托交互设计构建沉浸式传播场景, 重塑文化传播逻辑; 从情感价值传递角度, 可通过多模态融合技术探索传承创新路径, 优化传播效能与传承模式[1]。借助情感化设计理论, 一方面能构建传统音乐多模态转译模型, 完善非遗数字化研究体系, 另一方面可形成可复制的技术应用范式, 助力京音乐突破传播困境, 让这一古老文化在数字时代重焕活力, 实现文化传承与创新发展的双赢。

## 2. 多模态创新艺术设计: 本能层视角下的视觉与触觉体验升级

在情感化设计的本能层, 人类对信息的感知源于最直接的感官刺激。多模态技术通过视觉与触觉的创新融合, 为京音乐艺术设计带来突破性的感官体验升级, 进而提升用户认知促进, 以直观、强烈的刺激快速吸引用户注意力, 激发其对传统文化的兴趣[2]。

### 2.1. 多模态驱动的视觉风格跨界重构与感官冲击

在情感化设计的本能层, 多模态技术通过整合多元媒介, 深度融合数字艺术与传统文化元素, 为京音乐艺术设计带来视觉风格的跨界重构与多维度感官冲击, 以直观且富有冲击力的方式吸引用户, 革新传统艺术的体验范式[3]。

在视觉创新维度, AIGC 技术成为京音乐创新展示的驱动力, 借助 Midjourney、可灵 AI 等图像生成工具, 联合 Runway、即梦 AI 等辅助工具, 以多模态技术打破传统动画创作边界。从孩童视角构建智化寺虚拟场景, 将京音乐古曲意境转化为生动画面, 同时融入智化寺旋子彩画纹样, 使数字内容与实体建筑遗产风格统一, 赋予古老音乐文化全新视觉生命力(如图 1 和图 2 所示)。在交互体验上, 利用 3dsmax、Maya 等软件对智化殿进行三维建模, 通过 Unity 交互展馆拓展多模态技术在视觉与触觉交互的应用。在“乐器博物馆”探索古代乐器创作, 通过手柄触发音效、展品信息交互等设计, 让观众在视觉观察虚拟场景、动态数据, 进一步实现与虚拟展品的深度互动, 沉浸式感受京音乐的传承与发展(如图 3 和图 4 所示)。



Figure 1. AIGC narrative practice in Beijing music animation

图 1. 京音乐动画中的 AIGC 叙事实践



Figure 2. Scene interaction design of Beijing music animation  
图 2. 京音乐动画的场景交互设计



Figure 3. Interaction design of unified exhibition hall  
图 3. Unity 展馆交互设计



Figure 4. Presentation of the history of Zhihua temple in the virtual  
图 4. 虚拟场景中智化寺历史信息呈现

在虚拟演奏与 Unity 京音乐拼图游戏中, 多模态技术通过压力感应手柄和震动反馈装置实现创新突破(如图 5 所示)。用户操作“云锣”拼图时, 使用手柄依据力度、节拍模拟纸张摩擦、金属共鸣等震动质感, 重音强震、弱音轻颤; 压力感应机制则将按压力度转化为演奏强度, 控制音效变化。这种设计融合触觉反馈与文化科普, 用户操作时触发曲目播放、乐器讲解, 以体感交互实现视、听、触多模态深度融合, 提升文化体验沉浸感。



Figure 5. Design case of the Beijing music jigsaw puzzle game  
图 5. 京音乐拼图游戏的设计案例

多模态技术通过在 AIGC 动画创作、Unity 交互展馆设计、触觉反馈实践等方面的创新应用, 实现了京音乐艺术设计在视觉风格上的跨界重构与多维度感官冲击的提升。这种创新不仅打破了传统艺术形式的局限, 更为传统文化的传承与发展注入新活力, 为用户带来更具吸引力与沉浸感的文化体验。

## 2.2. 多模态交互下的个性化视觉、触觉体验

基于本能层对独特、新奇刺激的偏好, 多模态技术通过数据驱动与交互设计, 实现京音乐艺术体验的个性化定制。在虚拟展馆与数字研学平台中, 系统实时捕捉用户的眼光停留时长、触控轨迹等行为数据, 以此为依据动态生成专属视觉与触觉反馈。

AR 技术作为多模态交互的创新引擎, 为用户解锁了前所未有的自主定制权。用户扫描云锣文创卡片后, 可通过手势滑动调整虚拟云锣模型的表面纹理——从明代斑驳锈迹到现代镜面抛光, 同时选择不同强度的触觉反馈模式: 轻触触发羽毛拂过般的酥麻感, 重击则模拟真实锣槌敲击的反弹力。当用户使用移动设备扫描智化寺文创产品图案, 设备摄像头捕捉图像后, 同时可与虚拟乐器合影实现个性化用户深度交互。触摸锣面触发震动与音效, 长按显示图文并茂的科普卡片与语音讲解, 关联乐器历史文化档案, 实现视觉、听觉、触觉与知识科普的多模态深度融合(如图 6 所示)。上述定制化体验可精准匹配个体对视觉色彩、动态效果及触觉反馈的本能偏好, 让每个用户都能获得独一无二的感官冲击, 极大提升对京音乐文化探索的主动性[4]。



Figure 6. AR personalized experience of Beijing music  
图 6. 京音乐 AR 个性化体验

### 3. 多模态交互革新：行为层驱动下的沉浸体验与操作愉悦升级

#### 3.1. 多模态融合激活多感官联动：基于视觉行为引导的交互闭环构建

依据唐纳德·诺曼的行为层设计理论,人类在交互过程中的视觉感知直接影响行为决策与操作反馈。多模态创作通过视觉元素的动态引导,构建起“视觉输入-行为响应-多感反馈”的闭环体验。在数字艺术展览中,高清投影的动态图像不再是单纯的视觉展示,而是通过色彩渐变、元素位移等视觉线索,隐性引导观众的行为路径。

在京音乐数字化场景中,AIGC技术深度融入视觉行为设计逻辑。将工尺谱字符转化为可交互的动态符号,当用户点击“上、尺、工、凡”字符时,不仅触发对应音符的听觉反馈,字符还会以模拟毛笔书写的轨迹在智化寺动画背景中延展,同时手柄震动强度与笔画粗细同步变化,实现视觉动态引导、听觉即时反馈、触觉精准响应的三位一体[5]。在拼图游戏中,采用“视觉引导线”设计,当用户移动拼图碎片接近正确位置时,碎片边缘会出现高亮光效,手柄同步给予轻微震动提示,降低操作难度的同时提升成就感;而偏离时,光效闪烁频率加快,通过视觉与触觉的双重提示,引导用户调整行为策略(如图7所示)。

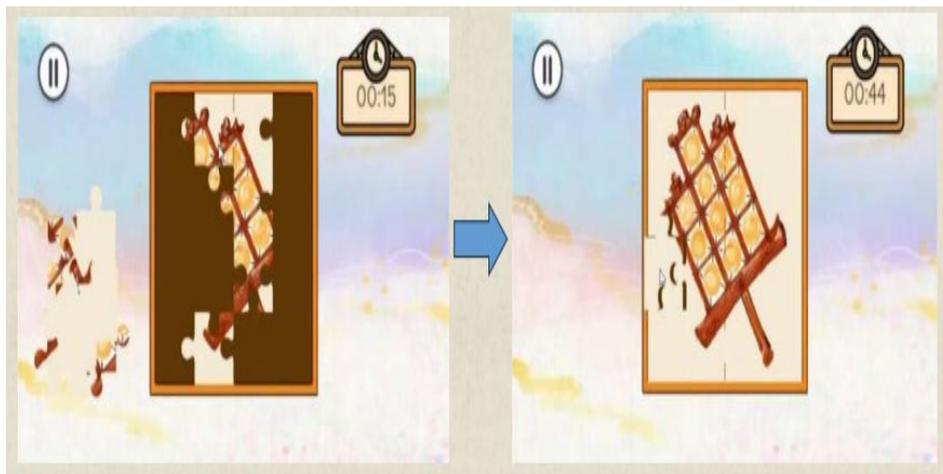


Figure 7. Visual dynamic guidance of Beijing music jigsaw puzzle game  
图7. 京音乐拼图游戏的视觉动态引导

因此,上述基于视觉行为引导的交互闭环可以较好地激活多感官联动,驱动用户沉浸式体验具体操作,这不仅增强了信息传递效率,更通过视觉行为设计建立起用户与数字内容的深度互动,增强了文化记忆的留存度与情感共鸣的深度[6]。

#### 3.2. 多模态适配不同场景下的感官需求：行为层视角的场景化功能拓展

基于行为层设计中“场景-需求-行为”的适配逻辑,多模态技术针对不同传播场景构建动态交互体系:线上短视频通过“视觉焦点+听觉钩子+行为激励”形成点赞-反馈闭环,当用户点击“解锁隐藏曲目”按钮时,点赞数实时进度条与音效升级构成即时奖励,触发“操作-愉悦”神经反馈;移动AR场景中,“智化寺AR导览”APP以建筑构件为视觉锚点,用户双指缩放触发的手柄震动与3D音效拆解,将碎片化学习转化为“探索-成就”行为链条,提升知识获取效率进一步扩大行为层的场景化功能拓展。

通过场景化功能模块与行为适配机制,多模态技术实现从“感官刺激”到“行为沉浸”的质变:视觉引导线降低操作认知负荷,触觉反馈强化行为记忆,听觉钩子激发情感共鸣,三者通过行为层设计形成

协同效应。行为层的即时反馈机制已将技术交互转化为情感化体验。这种“行为触发 - 多感响应 - 情感共鸣”的闭环设计, 不仅让京音乐文化以可操作、可感知的方式融入现代生活, 更通过个性化行为适配, 使每位用户在交互中获得专属的沉浸愉悦感, 实现传统文化传播效率与体验深度的双重突破。

## 4. 多模态数字叙事: 反思层框架下京音乐文化的沉浸式传递与情感共鸣

### 4.1. 多模态叙事空间的文化建构与价值传递

根据唐纳德·诺曼情感化设计三层次理论, 反思层作为人类认知的最高层级, 承载着文化记忆、价值认同与情感共鸣的深层需求。多模态技术通过构建「历史符号可视化-情感线索交互化-文化价值场景化」的叙事链条, 将京音乐 600 余年的历史脉络转化为可感知、可参与的沉浸式体验[7]。AIGC 技术深度挖掘智化寺建筑与京音乐的历史关联, 将工尺谱字符转化为具有叙事性的动态视觉符号, 配合环绕立体声效中渐次增强的编钟音色, 构建起极具历史纵深感的视听叙事。

VR 技术打造的“智化寺数字学院”则进一步突破时空界限, 用户佩戴 VR 设备进入 1:1 还原的明代智化殿, 脚下的压力感应地板随着古乐节奏产生细微震动, 虚拟艺僧的手势动作与工尺谱动态解析同步呈现(如图 8 所示)。当用户触发“历史回溯”功能时, 场景瞬间切换至清代修缮现场, 全息投影投射出匠人修复云锣的真实影像, 手柄震动模拟金属锤击的触感, 使观众在沉浸式体验中直观理解京音乐传承过程中的艰辛与坚守, 引导用户主动思考京音乐作为非遗文化的传承意义, 实现从“观看”到“理解”再到“珍视”的认知升华, 完成文化价值的深度传递。

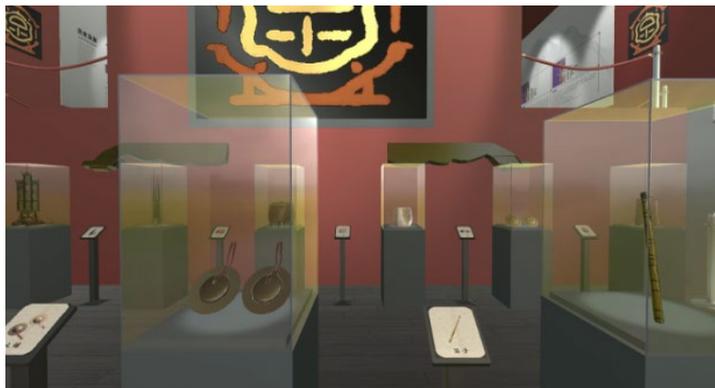


Figure 8. VR narrative immersive experience of Beijing Music  
图 8. 京音乐的 VR 叙事沉浸式体验

### 4.2. 多模态交互中的情感共鸣与认同深化

反思层理论认为, 人类通过深度思考与情感联结, 对体验进行价值判断与意义建构。在多模态交互中, 这一理论被转化为情感化设计的核心逻辑, 通过历史叙事、视觉符号与情感反馈的交织, 推动用户从感官体验升华为文化认同。在数字研学场景中, “京音乐基因解码”互动系统利用生物传感技术捕捉用户心率变化, 当用户成功破译工尺谱古曲密码时, 系统实时生成专属“文化传承者”数字勋章, 勋章设计融合智化寺藻井纹样与现代区块链技术, 既承载历史美学价值, 又赋予数字时代的身份象征意义。

线上线下联动的多模态展演则构建起集体情感共鸣的场域。线下场馆的投影实时同步线上 VR 观众的虚拟形象, 当线上用户通过手柄发送“喝彩”指令时, 线下场馆的灯光系统随之闪烁, 定向音响释放观众欢呼的混响音效, 上述操作源于通过识别与捕捉用户的“情感信号”后, 在多模态情感交互场景下进一步来提升用户的情感理解与传达, 为多模态赋能京音乐创新研究提供新的技术路径[8]。在“云锣合

奏”共创环节, 线上线下参与者通过手势或触屏共同控制虚拟乐器, 生成的乐曲波形实时转化为智能化建筑外墙上的光影艺术, 不仅增强了用户之间的情感联结, 更让参与者体会到自身在传统文化传承中的角色价值, 从情感层面激发对京音乐文化的主观传播意愿, 实现反思层所追求的文化认同与价值内化。

## 5. 多模态叙事革新: 基于情感化设计三层次的传播效能升级

### 5.1. 本能层: 多模态感官融合与多元化表达的数字化吸引力构建

在情感化设计的本能层, 多模态技术通过整合视觉、听觉、触觉等多元感知形式进一步增强用户的情感卷入度, 既打造直击感官的传播内容, 又满足京音乐文化的多元化叙事需求[9]。AIGC 技术基于 Midjourney 等工具生成动态视觉符号, 将工尺谱字符转化为粒子特效与古建筑光影交织的短视频素材, 在各大社交平台上依靠视觉亮点抓取注意力; 同时, 动态数据可视化图谱与专家访谈视频结合, 将 570 年曲目演变等复杂信息转化为直观视觉流, 搭配播客对工尺谱故事的音频演绎, 实现学术性与通俗性平衡。在艺术创作中, AIGC 生成的工尺谱动态符号与 AR 拆解的乐器模型结合, 使抽象音乐理论具象为可交互体验, 并且在创作的过程中, 通过系统性伦理与治理机制确保生成内容既不触及侵权红线, 又坚守文化本真, 将所生成的结果用于文化传播, 建立内容生成系统, 通过技术手段锁定创作边界, 让数字化传播始终扎根于京音乐文化的本质内核, 实现技术创新与文化伦理的双向平衡; Unity 虚拟展馆以 4K 建模呈现云锣细节, 用户缩放时触发的触觉震动与光影变化, 形成“指尖探索 - 视觉反馈”联动, 通过多元感知形式强化数字化传播吸引力。

### 5.2. 行为层: 交互闭环与社交属性驱动的多模态传播

基于行为层“操作 - 反馈”逻辑, 多模态技术通过构建交互闭环与增强社交属性, 推动内容传播裂变。小游戏“京音乐拼图挑战”中, 手柄压力感应将操作力度转化为演奏音效, 生成带个人 ID 的成就海报, 通过“分享解锁曲目”形成“操作 - 分享”闭环; 短视频平台“云锣合奏”功能允许用户触屏生成演奏片段, 自动剪辑为带话题标签的传播视频。线上线下联动的 AR 导览系统中, 用户扫描文创卡片触发虚拟乐师演奏, 手势数据同步至云端排行榜, AR 模型的自由旋转与交互操作支持用户录制创意短视频, 配合“点赞解锁线下彩蛋”机制, 实现从内容消费到社交共创的转化, 强化多模态内容的社交传播力。

### 5.3. 反思层: 文化叙事赋能多模态内容的情感价值沉淀

在反思层, 多模态技术通过历史场景数字化与情感化叙事, 深化用户对京音乐文化的价值认同。VR 沉浸式体验以 1:1 还原的明代智能化殿为场景, 用户触发“历史切片”时, 压力感应地板随古乐震动, 全息投影呈现各朝代乐师影像, 手柄模拟不同时期乐器触感, 助观众理解文化传承脉络, 进一步构建“个人行为 - 文化传承”联结, 生成个性化“文化记忆图谱”, 实现从感官体验到文化认同的反思层传播升级。

## 6. 结语

多模态技术与情感化设计三层次理论的深度融合, 为京音乐这类单一听觉形式的传统艺术提供了系统化的创新路径。该模式通过本能层的感官体验重构、行为层的交互逻辑革新、反思层的文化价值解构, 全面拓展了艺术呈现的维度, 既突破了传统传播的受众局限, 又通过数字技术激活了文化遗产的现代生命力。

从实践价值看, 多模态融合不仅实现了虚拟空间中深度交互场景的构建, 更通过视听触多通道感知刺激与艺术表达边界的拓展, 显著提升了内容传播效能。未来, 随着人工智能、大数据与多模态技术的

协同创新, 京音乐创作有望实现个性化体验定制、跨媒介叙事升级及文化生态重构。持续深化这一融合模式, 不仅能为数字媒体行业注入创新动能, 更将为京音乐这一古老文化遗产注入持久活力, 使其在当代社会以更富感染力的方式传承发展, 在传统与现代的碰撞中产生更具深度和创新性的价值呈现。

## 基金项目

北京印刷学院大创项目《基于 AIGC 技术的京音乐文化交互平台设计与实现——京乐坊》编号: 22150725057。

## 参考文献

- [1] 冯德正. 数字技术与多模态语篇分析[J]. 北京科技大学学报(社会科学版), 2010, 26(4): 12-17.
- [2] 彭智谋, 张瑞英, 黄茜, 等. 基于儿童情感化设计的小学室内环境色彩研究[J]. 建筑科学, 2025(5): 1-8.
- [3] 杨薇. 充分利用数字技术, 做好非物质文化遗产保护传承工作——以智化寺京音乐为例[C]//2019 北京数字博物馆研讨会论文集. 北京: 华夏出版社, 2019: 226-233.
- [4] 孙鑫. 智化寺及所传京音乐的历史考察——兼论其在当代的传承与保护[C]//文化遗产与公众考古(第三辑). 北京: 科学出版社, 2016: 137-179.
- [5] 韩燕玲. 新媒体编辑的多模态叙事能力重构研究[Z]. 嘉峪关市融媒体中心, 2025.
- [6] 裴张龙, 彭纲, 吴佳醜. 虚拟仿真实验情感化设计研究[Z]. 浙江师范大学行知学院设计学院, 浙江师范大学设计与创意学院, 2025.
- [7] 郝壮, 贾营萍. 铁人精神在红色文创产品中的情感化叙事设计研究[Z]. 齐齐哈尔大学美术与艺术设计学院, 2025.
- [8] Lin, D., Li, L., Cao, D. and Li, S. (2017) Multi-Modal Sentiment Feature Learning Based on Sentiment Signal. *Proceedings of the 12th Chinese Conference on Computer Supported Cooperative Work and Social Computing*, Chongqing, 22-23 September 2017, 33-40. <https://doi.org/10.1145/3127404.3127410>
- [9] 李仪玮, 温馨, 黄华秋, 王雨欣. 科技赋能文化传播: 数字艺术展的多模态情感化设计研究[Z]. 中国传媒大学, 2025.