

妙思宇宙

——基于生成式 AI 技术的文博领域多模态交互界面设计

白妮妮 巴晨旭 韩秋硕 刘畅

西安工程大学 陕西 西安 710048

【摘要】：在新一轮科技革命和产业变革加速突破的形势下，为了顺应我国推动文化高质量发展、建设社会主义文化强国的重要举措，在进行充分的市场调研后，研发一款名为“妙思宇宙”的文博领域的智能 APP 设计，app 以多模态自然交互^[1]、深度个性化服务及活跃的智能体社区为核心体验支柱，并以抽象的智能体形象作为主要交互载体，全面支持用户通过图像输入、语音指令及文字交流进行双向互动^[2]。

【关键词】：文博领域；UI 设计；双向互动；智能体形象

DOI:10.12417/3041-0630.26.10.038

2024 年 10 月 28 日，习近平总书记在中共中央政治局第十七次集体学习讲话中提出：“要探索文化和科技融合的有效机制，用互联网思维和信息技术改进文化创作生产流程，推动“硬件”和“软件”全面升级，实现文化建设数字化赋能、信息化转型。”2026 年《中共中央关于制定国民经济和社会发展第十五个五年规划的建议》提出：“推进文化和科技融合，推动文化建设数智化赋能、信息化转型，发展新型文化业态。”这是在新一轮科技革命和产业变革加速突破的形势下，推动文化高质量发展、建设社会主义文化强国的重要举措。文化与科技相生互补，文化为科技注入价值导向，科技为文化注入新鲜血液，因此要大力发展文化与科技的融合产品，把焦点关注与二者融合的视角上，不能只单一的注重其中某一个板块。

在进行充分的市场调研后，梳理调查获取的结果，剖析文博领域的应用，发现当下传统文博平台界面普遍存在诸多短板：信息架构与导航层级杂乱，核心信息被冗余内容掩盖，标签设计缺乏直观性与场景化入口；交互模式以单向图文展示为主，缺少手势缩放、3D 旋转、AR 叠加等沉浸式操作，用户参与感薄弱；视觉呈现形式单一，整体以静态页面为主，滚动渐显、点击微动效等贴合内容的动态交互缺失，难以和用户建立情感联结，同时页面视觉负荷较高，长时间浏览易产生疲劳，文物高清细节、专项标注展示不足；产品个性化服务能力欠缺，未搭建用户画像、偏好记忆体系，无法针对学生、研究者、普通游客等不同群体推送定制化展品内容，难以满足多元使用需求。

针对上述行业与文博应用领域所面临的诸多痛点与挑战，我们创新性地采用了混合用户界面^[3](HUI)与视觉问答(VQA)^[4]技术相结合的综合解决方案，旨在全面提升用户在文博场景中进行信息交互的整体体验感与操作效率。我将对全文进行精细化改写，重构语句逻辑、替换专业表述、剔除重复词汇，同

时完整保留核心技术逻辑、理论出处与项目内核，保证学术严谨度与语句流畅度。

为革新文博场景数字化体验、升级展馆交互模式，本团队自主研发“妙思宇宙”文博生成式交互测试系统。该系统依托多模态交互融合技术，重构博物馆线下场景与线上导览应用的人机交互逻辑，有效简化用户游览、查询、体验全流程，全方位优化大众文博观赏与数字化游览体验。1991 年，Feiner 与 Shamash 两位学者首次提出混合用户界面(HUI)概念，该理论的核心思路是整合各类异构终端设备，融合多元互补的人机交互范式，构建一体化、复合型的交互体系。从架构维度划分，混合用户界面体系可归纳为三类核心交互形态。第一类为智能交互融合形态，依托智能算法实现图形交互界面(GUI)与自然语言交互体系的深度联动，是当下数字化交互的主流形式。该形态构建“语义意图识别+智能任务落地”的闭环运行模式，可适配文博场景中全场景、全流程的用户交互与功能使用需求。第二类为虚实协同交互形态，通过 AR、VR 沉浸式终端与二维平面设备的联动协作，兼顾交互的精准性与场景的沉浸感。既能够满足用户对文物细节精准查看、精细操作的需求，又可为用户搭建沉浸式文博空间场景，实现虚实融合的游览体验。第三类为高效极简交互形态，结合图形可视化界面与指令输入模式，针对性适配文博专业研究场景，可大幅降低高频重复操作的工作成本，有效提升文物研究、资料整理、数据核验等专业工作的运行效率。视觉问答技术作为跨领域前沿技术，整合计算机视觉识别与自然语言语义解析两大核心能力，通过图像特征精准抓取、文字语义深度拆解、跨模态信息联动推演等多层技术逻辑，实现图文交互式咨询服务。用户可结合文物实景画面搭配文字提问，快速获取文物年代溯源、传统制作工艺、历史文化内涵、馆藏价值等精细化、多元化的文博信息。

目前，该跨模态技术已落地应用于智慧展馆导览、文博学

术深耕研究、数字展厅沉浸式搭建、文物数字化修复辅助等多个核心文博领域，为文博文化资源的大众化传播、数字化普及，以及传统文博交互模式的迭代升级筑牢技术根基。基于上述混合界面交互理论与跨模态视觉问答核心技术，本团队完成“妙思宇宙”文博专属大模型视觉展示展板设计。该设计聚焦文博领域多模态交互创新落地，将复合型交互架构与视觉问答功能深度融合，为文博数字化交互体系的创新实践、场景落地与模式探索提供全新思路与可行路径。

“妙思宇宙”这一品牌精准定位于16至45岁之间的文化探索者、潮流追随者及科技爱好者群体，致力于通过先进的生成式人工智能技术，构建一个融合多种模态的文博创新应用平台。该平台以多模态自然交互、深度个性化服务及活跃的智能体社区为核心体验支柱，并以抽象的智能体形象作为主要交互载体，全面支持用户通过图像输入、语音指令及文字交流进行双向互动。产品的核心理念可概括为两大组成部分：一是“MUSE 行星智能体”，二是“VERSE 妙思宇宙”，二者共同构成了完整的“MUSEVERSE”生态体系。MUSE 行星智能体能够与用户进行流畅自然的语言互动，借助语音对话、图像识别与智能问答等多种形式，为用户提供包括文博资料查询、个性化观展路线规划、实时增强现实导览辅助等在内的多项实用功能；而 VERSE 妙思宇宙则是一个集文化、潮流与科技于一体的综合性赛博社交平台，有机整合官方发布内容（GGC）、用户原创内容（UGC）以及人工智能生成内容（AIGC），从而打造出一个丰富多元、充满活力的线上文化互动空间。

在展品识别与内容解析的深度层面，该软件系统依托先进的多模态大模型作为核心技术基座，实现了对展品图像的全方位、多维度智能识别。系统能够精准提取展品的主体轮廓、画面中包含的文字信息以及独特的艺术风格等关键视觉特征，并进一步结合强大的语言模型能力，自动生成蕴含丰富历史发展脉络与文化背景内涵的深度解读文本，从而全方位、多层次地满足观众在观展过程中获取知识、深化理解的核心需求。妙思宇宙产品架构与核心功能优化改写

在整体架构搭建层面，本平台搭建了层级分明、逻辑闭环的可视化信息结构。平台首页展示界面、智能交互咨询界面、文物展品详情界面、桌面快捷功能组件以及虚实融合导览功能等全部功能板块，均经过系统化视觉排布与精细化交互打磨。在简化用户操作路径、提升整体使用效率的基础上，实现界面视觉风格的统一协调，兼顾实用性与美学观感。

产品整体视觉体系采用现代极简设计范式，搭配专属原创智能IP视觉形象，大幅精简界面冗余元素，降低大众用户的上手学习成本与视觉认知压力。同时，平台通过精简多余操作步骤、过滤无效信息推送的设计思路，适配各年龄层次、不同使用习惯和审美偏好的用户群体，具备极强的大众适配性与普

惠性。整套平台拥有成熟完善、高效闭环的运行逻辑，突破性搭建了语音交互+实景图像上传的双向融合交互模式，彻底革新传统博物馆单一的咨询体验。在语音交互场景中，参观者可直接以口述方式提出观展疑问，依托高精度语音辨识技术，平台可实时完成语音转译与用户深层需求拆解，适配老人、孩童等全年龄段人群的无障碍使用需求，兼容各类用户的操作习惯。在图像交互场景中，用户可现场拍摄馆藏文物并上传至平台系统，依托成熟的机器视觉识别算法，精准抓取文物整体形制、纹样肌理、色彩特征等核心细节，快速对接海量馆藏资源数据库，完成高精度信息比对匹配。两种交互渠道获取的用户需求数据，将统一汇入平台核心智能处理单元进行统筹运算。该核心单元可同步融合解析图像视觉特征与用户文字、语音的语义诉求，联动标准化文博知识库完成智能检索整合，最终输出涵盖文物基础参数、历史渊源、文化寓意、工艺特点的全方位、体系化科普内容，彻底解决了传统人工讲解、固定展板导览信息零散、内容片面、不成体系的痛点问题。与此同时，平台搭载毫秒级实时运算响应机制，极大缩短数据解析、资源检索与内容输出的耗时，实现用户操作即刻反馈，保障全程观展交互流畅不间断，带来连贯自然的使用体验。

除此之外，MUSEVERSE 妙思宇宙平台深度融合虚实仿真展示技术^[5]，打造沉浸式、场景化的新型文博观赏模式。用户可通过移动端设备或AR穿戴设备，激活文物专属动态交互特效，既可还原古代器物的实际使用场景与应用方式，也可让静态书画展品实现动态叙事，以拟人化形式演绎背后的历史故事。依托虚拟场景复刻技术，平台还可高度还原文物诞生的古代工坊、历史场景与时代环境，让使用者沉浸式置身于历史时空之中，直观感受不同时代的社会风貌与人文底蕴，有效增强传统文化传播的感染力、代入感与记忆深度，实现文博文化的沉浸式传播与创新性传承。

MUSEVERSE 平台具备实时、高效的多模态对话能力，经过持续的算法优化，其响应速度显著提升，能够迅速理解并解析用户提出的各类问题，及时给出准确反馈，帮助用户实现高效的信息检索与获取。平台依托于强大的预训练语言模型、高精度的语音识别技术以及高度拟人化的语音合成系统，可完成流畅自然、贴近真人语感的语音交流体验。妙思宇宙在设计上也会充分考虑用户群体的多样性，能够为不同语言背景、文化习惯的参观者提供量身定制的导览服务，并且能够最大限度地兼顾国内外游客的差异化需求，应用不单单只针对于国内用户，使用者还可以根据个人喜好，自主调整智能体的性格特质与对话风格，能获得高度定制化和贴心陪伴式的智能游览体验，跳出导游式的传统逛文博。

本项目意在构建“高校技术研发+博物馆资源供给”共生发展模式，为中小型博物馆的数字化转型探索并提供一条切实

可行的实施路径。在此想法与实操下，高校科研团队针对当前文博行业存在的具体技术瓶颈与共性难题，开展深入专项技术研究，并将前沿的学术研究成果高效地落地转化为可实际部署、功能完善的商用软件产品；与此同时，博物馆作为重要的文化机构，不仅为这些前沿技术的迭代升级与功能验证提供真实和丰富的应用场景，还贡献出海量的、多模态的馆藏文物数据资源，从而在技术与文化资源之间形成了深度耦合、相互促进、双向赋能的“双螺旋”协同发展体系。为确保合作高效、规范，双方需要共同建立标准化的协作与数据管理机制：博物馆方面系统性地将馆藏文物的文字描述、高清图像资料、相关学术研究成果等数字化资源，上传至统一的共享数据平台；高校研发团队在此基础上，进行统一的专业数据清洗、规范化与

标准化处理工作，进而构建起结构清晰、标准统一的文物专题数据库。

MUSEVERSE 平台拥有完整、清晰的信息架构与交互设计，在软件具体页面呈现上——包括首页、智能对话页面、文物详情页面、桌面便捷小组件以及增强现实（AR）导览模块等——其各组件页面布局、视觉排版以及各类操作按钮的交互逻辑，均经过精心的用户体验设计，旨在确保用户能够便捷地获取丰富的参观内容与信息服务的同时，持续优化和提升妙思宇宙整体的使用流畅度与满意度。此外，平台将抽象智能体设计理念与文博浏览相结合，不仅为整个用户界面注入了一种契合数字时代的极简主义美学风格，提升了视觉上的高级感与科技感。

参考文献：

- [1] 夏子硕.面向智能实验的多模态自然交互算法设计及其应用研究[D].济南大学,2024.
- [2] 杨永标,李逸驰,陈霜,等.双向互动服务平台的关键技术与总体构架[J].广东电力,2015,28(8):28-32.
- [3] 蔡登旭,王欢.基于文旅行业大模型的移动终端垂域 VQA 应用 HUI 设计[J].电脑知识与技术,2025,21(4):152-154.
- [4] 包希港,周春来,肖克晶,等.视觉问答研究综述[J].软件学报,2021,32(8):2522-2544.
- [5] 王明政.基于虚拟仿真技术的展示艺术设计对策[J].鞋类工艺与设计,2024,4(12):104-106.