

# 人工智能驱动的非物质文化遗产资源库建设研究

明 慧

(北京联合大学应用文理学院, 北京, 100191)

**摘要:** 非物质文化遗产(非遗)是民族文化传承的核心载体, 资源库则是利用现代科技对非遗进行数字化保护与活化的关键基础设施。当前传统非遗资源库受限于技术架构, 普遍在数据处理、知识挖掘与应用活化方面存在显著局限, 难以适配数字时代需求。本文系统探讨AI技术在非遗资源库建设的整体框架、技术适配与功能开发, 分析出当前人工智能驱动下的非遗资源库建设存在的现实挑战与相对应策。研究旨在明确AI驱动非遗资源库建设的核心逻辑与个性化应用场景, 为解决传统资源库痛点, 推动非遗资源从“沉睡数据”转化为文化创新与教育传播的活性素材, 为非遗资源数字化保护提供理论支撑与实践参考。

**关键词:** 非物质文化遗产; 资源库; 人工智能; 数字化保护; 活态传承

中图分类号: G270.7 文献标识码: A

## 一、引言

非物质文化遗产(以下简称“非遗”)作为民族文化的重要组成部分, 承载着历史记忆、价值理念与生活智慧, 其活态传承与系统性保护已成为文化建设的核心议题之一。进入数字时代, 非遗数字化转型逐步推进, 传统非遗资源库作为数字化保护的核心载体, 在非遗资源存储利用、公开共享等方面发挥了重要作用。然而, 面对海量增长的非遗数据与日益多元的传承需求, 传统资源库在数据处理效率、知识挖掘深度与应用活化能力上的局限逐渐凸显, 难以满足非遗保护对于活化利用的转型需求。人工智能技术凭借其在数据智能处理、知识关联挖掘、个性化服务推送等方面的独特优势, 为非遗资源库的升级重构提供了新的技术路径。本文围绕人工智能驱动非遗资源库的建设, 从现实瓶颈与适配基础出发, 系统探讨了资源库的整体架构、实施路径及应对挑战的策略, 旨在为构建兼具技术先进性与文化适配性的非遗资源库提供理论参考与实践指引。

## 二、非遗资源库建设的现实瓶颈与AI适配基础

在数字化转型的背景下, 传统非遗资源库虽已完成初步数据积累, 但在应用于实践场景时逐渐显现其功能局限。人工智能技术的快速发展为突破当前困境提供了全新路径, 其技术特性与非遗资源库的功能需求形成高度契合, 为后续整体架构设计与实施路径规划奠定理论基础。

### (一) 传统非遗资源库建设的核心瓶颈

传统非遗资源库的建设多以资源的数字化存储为目标, 在技术应用与功能设计上仍停留在信息管理层面, 难以适配当下非遗资源的复杂性与传承需求的多元性, 具体表现如下:

#### 1. 数据结构碎片化与标准缺失

其一, 数据处理低效, 呈现碎片化特征。非遗资源类型复杂多样, 涵盖文字资料、口述音频、技艺影像、实物模型等多种形态, 不同类型资源的格式标准不统一。传统资源库主要

依赖由人工操作完成的数据采集、分类与标注，不仅消耗大量人力成本，效率极低，还易因主观判断差异导致数据标注偏差。同时，不同部门与地区间的共享机制尚未完善，各区域建设的资源库多为独立运行，在建立资源库时参照的数据标准缺乏协同，形成了“信息孤岛”的局面，且跨区域、跨类型的非遗数据整合难度极大，无法形成完整、系统的数据体系。

### 2. 知识关联薄弱与语义挖掘不足

其二，知识关联薄弱，信息挖掘表层化。传统资源库的功能开发集中于数据保存与简单检索，但仅能实现对单一资源的定位查询，难以挖掘非遗资源背后的深层关联。在实际运用中，无法自动梳理某一非遗技艺与传承人的代际脉络，或是与其地域民生间的共生关系，也无法将不同流派技艺的技法差异与演变轨迹进行对比。在信息挖掘方面，仅处于表层化的呈现方式，使得非遗资源中的文化内涵难以被有效提取，无法为学术研究提供深层数据，也难以满足信息网络社会中对于知识关联的需求。

### 3. 功能单一与用户体验不足

其三，功能应用单一，用户体验感不足。传统非遗资源库的应用功能设计多停留在以展示为主的单向输出层面，缺乏交互性设计，用户通过资源库只能浏览信息，缺乏对用户主动参与行为的设计与支撑。这种应用模式将资源库的功能局限于“保护端”的存储需求，未能将非遗资源的活态内涵转化为可感知、可参与的体验，无法向“传承端”的应用需求延伸，进而造成非遗资源与现代社会场景的适配性不足，难以实现大幅度的功能升级，制约了非遗资源在大众传播与当代传承中的实际价值发挥<sup>[1]</sup>。

## （二）AI 赋能非遗资源库的适配基础

人工智能技术的成熟与发展，在技术特性、数据基础与政策环境上均与非遗资源库建设形成高度适配，构成了人工智能赋能的坚实基础。

从技术适配角度，人工智能关键技术与非遗资源库的功能需求高度契合。计算机视觉技术能够对非遗技艺影像进行智能分析，自动识别并标注技艺动作的关键帧，替代人工完成高效精准的数据处理；自然语言处理技术可对非遗口述史、文献资料进行语义分析，提取其中的文化信息与情感表达，实现从文字转写到知识提取的升级；知识图谱技术能够整合分散的非遗数据，构建传承人、技艺、项目、民俗文化、传承地等元素间的关联网络；机器学习技术则可基于用户行为数据，提供非遗资源的个性化推荐服务，提升应用体验。

从数据基础角度，国内非遗数字化积累已为人工智能技术的应用提供了充分条件。近年来，国家级、省级及地方部门已建成多个非遗专题数据库，累计完成数百万条非遗数据的数字化采集与存储，涵盖传统技艺、民俗、戏曲、音乐等多个类别。同时，部分科研机构与文化单位已开始构建非遗专属数据集，对数据进行标准化标注，为人工智能大模型的训练提供了高质量的数据支撑。

从政策环境角度，国家战略导向为人工智能赋能非遗资源库提供了制度保障。在《“十四五”文化发展规划》明确提出，“推动人工智能、大数据等新技术在文化遗产保护领域的

应用，构建数字化、智能化的文化遗产保护体系”，多地政府也出台专项政策，支持“AI+非遗”融合创新项目，鼓励科研机构、文化单位与科技企业协同合作。国家政策的导向指引不仅为人工智能驱动非遗资源库建设提供了发展方向，也为跨部门协作、资源整合与资金投入提供了保障，减少了项目推进的制约条件。

### 三、AI 驱动非遗资源库的总体架构与技术路径

人工智能驱动的非遗资源库架构设计以融合先进技术、释放资源价值为导向，围绕非遗资源全流程智能流转展开，从原始数据输入到智能处理、安全存储，再到多元服务输出，构建以层级协同、功能创新且兼顾差异化需求为逻辑的系统框架。

#### （一）系统框架设计

在人工智能驱动下的非遗资源库，构建以数据智能流转、知识深度挖掘、服务精准赋能为核心的多层次框架，遵循“数据输入-智能处理-存储管理-应用输出”流通过程的逻辑链条，自上而下分为数据层、处理层、存储层与应用层，各层级间既独立承担其核心功能，又将高效协同，形成完整连贯的全周期智能体系。

数据层是资源库的基础支撑，负责收集各类非遗数据资源，由原始数据与标准化数据两部分组成。原始数据包括非遗传承人访谈音频、技艺操作视频、民俗活动影像、文献资料扫描件等未加工资源；标准化数据是经过初步处理后的结构化数据，如传承人的基本信息、非遗项目分类表等。数据层通过多源数据接入接口，实现对政府部门、文化机构、科研单位及传承人个人数据的整合，同时建立数据接入标准，确保数据格式的统一性与规范性，为后续处理提供高质量的数据源。

处理层承担数据转化与知识挖掘的关键功能，通过部署人工智能算法模型，对数据层输入的原始数据进行智能处理，将非结构化数据转化为可分析关联的智能数据。通过计算机视觉、自然语言处理、知识图谱等技术，完成非遗数据的智能标注、语义提取与关联构建，形成包含技艺特征、文化内涵、关联关系的非遗知识体系。同时，该层级还应具备模型迭代能力，根据新接入的数据与用户反馈，能够持续优化算法精度，提升处理效率。

存储层结合数据的种类与重要性进行分类，采用分级管理模式，对于核心数据采用分布式本地存储，确保数据安全性与访问稳定性；对于普通数据采用云存储模式，降低存储成本，提升远程访问效率。建立常态化运维机制，定期对核心数据进行备份，设置安全自检功能，确保系统的隐私性与安全性，实现非遗数据与知识体系的长期存储与高效管理。

应用层是资源库与用户交互的主要渠道，面向不同用户群体提供差异化的智能服务，分为不同类型的服务模块。面向非遗保护工作者，提供数据的具体数据与资料，并开发自主上传、采集辅助、影像修复等工具，便于其高效完成非遗数据的补充完善与动态更新；面向科研人员，提供知识关联查询、技艺特征分析等功能，为其提供非遗知识体系的深度挖掘与量化分析支持，辅助开展相关学术课题；面向大众用户，提供个性化资源推荐、虚拟体验等服

务，通过互动体验增强对非遗文化的认知与认同<sup>[2]</sup>。

## （二）关键技术适配

人工智能驱动的非遗资源库需根据不同层级的功能需求，适配相应技术，确保技术应用与业务场景深度融合，避免与实际需求脱节。

### 1. 多模态数据融合

该技术聚焦数据层面，针对非遗数据文本、音频、视频、图像、三维模型的多模态特征，解决数据格式不统一、关联断裂、质量参差的问题。在数据整合部分，对单一模态数据，通过自适应算法实现各类形态档案的标准化格式转换；对跨模态数据，依托语义锚点与时间戳匹配技术建立关联，避免模态间数据孤立。在预处理部分，按模态特性设计质量控制机制，文本数据可以进行语法纠错与冗余过滤剔除错误表述与重复内容；音频数据通过降噪去除背景噪音；视频数据通过画面稳定性处理修正抖动画面；三维模型优化模型精度与展示效果。同时，通过多模态数据一致性校验算法，检查跨模态数据的信息匹配度，并标记冲突数据并生成修正建议，为后续智能处理提供格式统一、关联清晰、质量达标的 data 源。

### 2. 智能标注与特征提取

该技术作用于针对数据价值转化的智能处理体系，将非结构化多模态数据转化为可分析、可关联的结构化知识，同时保留非遗的人文属性。针对不同模态数据设计差异化处理逻辑，文本数据采用实体识别与语义分析技术，提取“传承人姓名、传承地、技艺流派”等核心实体，挖掘文本中的语义关联；视频数据标注技艺关键帧，分析动作轨迹，生成可量化的动作特征向量；图像数据从纹理、色彩、形状维度提取实物特征；三维模型通过捕捉实物的空间结构特征。此外，引入人工与模型的双重审核机制，非遗专家对自动标注结果进行校验，修正偏差，并将修正数据反馈至模型进行增量训练，持续提升标注与特征提取的精度，确保数据处理既高效又贴合非遗实际<sup>[3]</sup>。

### 3. 知识图谱与语义索引

该技术衔接智能处理与存储，实现知识系统化与检索高效化，兼顾数据间的关联与更新。在知识关联构建环节，通过构建非遗知识的基础框架，定义传承人、非遗项目、技艺等核心实体类型，分析其传承关系、所属关系、关联关系等实体间关系类型，形成结构化的知识图谱。在智能索引构建环节，采用语义索引与特征索引的双重结构：语义索引基于自然语言理解模型将非遗知识转化为语义向量，支持基于语义相似度的检索；特征索引则将视觉、音频、文本特征向量构建倒排索引，支持基于特征匹配的检索。此外，索引部分需具备实时更新能力，当知识图谱新增或修改信息时，可自动同步更新，确保检索结果的时效性。

### 4. 生成式 AI 与虚拟仿真

该技术聚焦用户的差异性需求，需从用户视角对资源库的精准性与体验感进行设计。在个性化服务环节，设置协同过滤模块，通过分析用户浏览记录等历史行为数据挖掘用户兴趣偏好，构建用户兴趣模型；内容特征模块基于非遗资源的特征数据构建资源特征模型；混合

推荐模型通过对两个模块进行融合，生成个性化资源列表。在交互体验环节，采用生成式人工智能与虚拟仿真技术，针对非遗传承人数字形象，采用基于生成对抗网络（GAN）的数字人生成技术，结合动作捕捉与语音合成技术，开展“虚拟讲解”等展示场景，通过场景渲染技术构建非遗技艺的应用场景，提升体验的沉浸感。同时，资源库需具备跨终端适配能力，针对移动端、VR设备等不同终端优化交互逻辑，确保不同场景下的交互便捷性。

### （三）模块功能开发

模块功能开发需基于关键技术适配成果，以四层体系的核心任务为出发点，实现技术成果向实际功能的转化，确保各模块逻辑独立且无内容重复，同时贴合非遗资源库建设的实际需求。每个模块需设计清晰的功能边界与数据接口，确保与其他模块的协同高效。

#### 1. 数据采集与接入

数据采集模块围绕数据来源拓展与质量前置控制展开，功能设计需兼顾便捷性与规范性。在多端采集功能上，可以开发移动端APP与小程序，提供专业采集工具，包含视频拍摄的帧率、分辨率调节、音频降噪处理、图像标注辅助等功能，同时支持离线采集，满足不同场景的采集条件。在跨系统对接功能上，开发标准化接口，支持与国家级非遗数据库、地方文化部门系统的数据同步，接口设计包含数据请求、格式转换、校验反馈三个环节，设置数据请求认证机制确保安全性，格式转换可以自动将外部系统数据转换为资源库标准格式，校验反馈则实时返回数据质量检测结果，便于外部系统调整数据报送内容。在数据初筛功能上，开发可视化质量检测界面，支持批量上传数据的质量分析，对不合格数据进行警报与提示，同时提供数据修复建议，用户可根据建议选择修复或舍弃数据，确保进入智能处理环节的数据符合基础质量要求<sup>[4]</sup>。

#### 2. 智能数据处理

智能处理体系聚焦数据价值的深度挖掘，其功能设计包含数据处理的标注、审核、特征提取等多方面内容。在智能标注功能上，开发自动化标注与人工审核一体化界面，自动化标注模块支持批量导入数据，通过调用数据智能标注技术自动生成标注结果；人工审核模块则提供标注结果列表，支持单条或批量审核，审核人员可修改错误标注、补充遗漏标注，同时支持标注规则自定义，标注规则修改后系统自动更新标注模型，提升后续标注精度。在特征分析功能上，开发特征提取与管理界面，特征提取模块支持按数据类型选择提取维度，提取完成后生成标准化特征库，支持按类型、特征相似度进行检索；特征管理模块则支持特征向量的查看、编辑与删除，同时提供特征相似度分析工具，可对比两个或多个数据的特征差异，生成可视化对比报告<sup>[5]</sup>。

#### 3. 资源存储与检索

资源库存储以数据安全存储与高效调用为双重目标，设置分级存储机制，便于根据资源的等级与使用频率进行分类与调用。在分级存储功能上，开发存储策略配置界面，用户可根据数据重要性与使用频率设置存储策略，如核心数据采用本地加密存储并定期离线备份，高

频使用数据采用云端存储并开启缓存加速。在安全管理功能上，开发权限管理与安全监控界面，权限管理根据级别设置不同角色，拥有不同的数据访问权限，同时支持角色权限的自定义配置；安全监控界面实时展示数据访问日志，通过异常检测算法识别可疑操作，触发预警机制，防范数据泄露风险。在多维度检索功能上，开发综合检索界面，提供多种智能检索方式，语义检索支持自然语言输入，特征检索支持上传图像或音频进行相似匹配，关联检索支持基于某一数据查询关联信息，提供检索结果导出功能，满足用户数据使用需求。

#### 4. 多端应用服务

资源库的应用与提供服务是其建立的根本目的，面向不同用户的差异需求提供个性化场景服务，针对保护者，开发保护工具集，包含技艺影像修复、数据采集进度统计、濒危技艺预警，支持数据导出与报告生成，便于保护者开展工作复盘。针对研究者，开发研究辅助工具，包含技艺流派对比、知识图谱分析、文献语义检索，同时可以提供 AI 辅助使用，降低操作难度。针对公众，开发互动体验功能，包含个性化资源推荐、虚拟体验、活动互动，设计注重趣味性与传播性，推动非遗资源走向大众生活<sup>[6]</sup>。

### 四、AI 驱动非遗资源库建设的挑战与对策

#### （一）现实挑战

##### 1. 数据安全风险

AI 驱动的非遗资源库在数据全生命周期中面临多重安全隐患，除基本的信息泄露外，较为特殊的是会直接威胁非遗数据与传承人知识产权方面的权益。非遗数据包含大量敏感信息，既有传承人个人隐私，也有独家技艺内容，如未公开的技艺秘方、核心操作流程。在存储阶段，若防护措施不足，可能遭遇攻击或内部人员违规泄露。在模型训练阶段，部分数据需上传至第三方平台，存在被擅自用于商业用途的风险。此外，AI 技术的“黑箱特性”使得数据使用轨迹难以全程监控，一旦发生安全事件，难以快速定位问题源头，进一步加剧了数据安全管控的难度<sup>[7]</sup>。

##### 2. 人才培养缺口

AI 与非遗的跨领域融合特性，使得资源库建设面临复合型人才与实操型人才双重短缺的困境。一方面，核心人才供给不足，现有非遗保护人员多具备深厚的文化素养，但缺乏对先进技术认知，难以直接参与到资源库技术架构设计工作中；但技术人员对非遗的活态传承特征、文化内涵理解浅显，导致开发功能脱离实际需求。另一方面，基层实操人才匮乏，县域及偏远地区的非遗保护单位人员有限，且多未接受过系统培训，面对智能采集时的复杂参数设置、数据审核流程时，操作不熟练，导致采集效率低、数据质量差。

##### 3. 数据共享壁垒

非遗数据的跨主体、跨区域共享面临多重阻碍，难以形成规模化的数据体系支撑 AI 模型训练。非遗数据多由政府部门、文化机构、传承人等多主体共同产生，数据权属界定模糊，

缺乏明确的权属划分标准，各主体因担心权益受损，不愿主动开放数据。数据标准也并不统一，不同地区、不同部门建设的非遗数据库，在数据格式、标注规则上存在差异，在进行跨系统数据整合时需大量人工调整，成本极高。同时，各单位多基于自身需求独立运营数据，缺乏统一的跨部门共享平台与机制，即便部分单位愿意共享，也因缺乏便捷的对接渠道与激励措施，难以实现常态化数据流通<sup>[8]</sup>。

## （二）应对策略

### 1. 筑牢数据安全防护体系

针对数据全生命周期安全风险，构建集技术、制度与监测三位一体的防护机制。技术层面，采用端到端加密技术保障数据采集与传输安全，核心敏感数据采用本地加密存储，搭配定期离线备份，降低云端存储风险，避免原始数据外泄。制度层面，制定与非遗数据安全相关的管理规范，明确数据采集、存储、使用的权限划分，并建立数据使用审批流程，任何单位或个人使用数据需提交申请并说明用途，经审核通过后才可获取。监督层面，搭建数据安全监控平台，实时监测数据访问、传输、使用轨迹，通过AI算法识别异常操作，触发自动预警，同时保证定期开展数据安全审计，确保防护措施落地生效。

### 2. 加大专项人才培养力度

从高校、科研机构、企业等单位共同打造多层次人才培养体系，重点培养兼具人文素养与技术能力的复合型人才，填补现有缺口。推动高校与科技企业、文化机构联合成立“非遗创新实验室”，组织学生参与到实际的资源库项目中，协助设计技艺智能模型，邀请企业技术专家、非遗传承人开设讲座，实现理论与实践深度融合。对于现有岗位工作人员，制订AI工具操作手册与短视频教程，聚焦核心操作步骤，避免复杂技术术语。同时，开展线下实操培训，邀请专家上门指导，针对偏远地区提供线上一对一辅导，并建立“基层人才帮扶机制，由核心团队定期对接基层单位，帮助解决实际操作难题<sup>[9]</sup>。

### 3. 建立数据共享协同机制

从权属界定、标准统一、平台搭建三方面入手，打破数据共享壁垒。首先，明确数据权属划分，由文化部门牵头制定相关认定范围，根据数据产生主体确定权属，明确数据使用的收益分配规则，保障各主体权益。其次，统一数据标准，由政府部门牵头，成立跨部门的工作小组，共同协商制定涵盖数据格式、分类维度、标注规则的统一标准，确保数据可直接跨系统整合。搭建统一共享平台，全国非遗数据共享中心，支持各单位通过标准化接口接入数据，平台设置数据检索、申请、授权等功能，同时建立激励机制，对主动开放数据的单位给予政策倾斜或资金补贴，对使用数据产生的研究成果、文创产品，在平台进行优先展示，激发各主体共享积极性。

## 五、结语

人工智能驱动的非遗资源库已突破传统模式局限，有效衔接非遗保护与传承间的需求，

为其数字化传承筑牢基础。展望未来，随着技术迭代，资源库将进一步提升非遗技艺数字还原度与体验感，同时打破领域壁垒，联动文旅、教育等产业推动非遗场景化应用。建设非遗资源库不仅提升了非遗保护与传承的效能，更助力非遗文化融入现代社会场景，为守护民族文化记忆、推动文化高质量发展提供了有力支撑。

### 参考文献

- [1] 赵跃.新时期档案机构参与非遗保护的反思与再定位[J].档案学通讯,2020,(02):40-48.DOI:10.16113/j.cnki.daxtx.2020.02.007.
- [2] 倪晓春,张蓉.关于非物质文化遗产档案数字资源库建设的思考[J].档案学通讯,2017,(02):53-57.DOI:10.16113/j.cnki.daxtx.2017.02.013.
- [3] 赖雅琳.生成式AI赋能非遗传承人特色信息库建设路径[J].图书馆工作与研究,2024,(12):53-60.DOI:10.16384/j.cnki.lwas.2024.12.004.
- [4] 胡晓庆.信息生命周期理论视角下的档案数据治理策略研究[J].山西档案,2020,(06):58-61.
- [5] 国家级非物质文化遗产代表性传承人抢救性记录工程操作指南[EB/OL].[2024-08-07].<https://wenku.so.com/d/1e7fc2ab339a0177672f2e72043604fd>.
- [6] 赵跃,周耀林.国际非物质文化遗产数字化保护研究综述[J].图书馆,2017,(08):59-68.
- [7] 温昊臻.数智技术赋能的非物质文化遗产档案全生命周期管理策略[J].山西档案,2025,(03):158-160.
- [8] 刘明真,李元旭.数字文化战略视域下文化遗产数字化传承：机理、挑战与路径[J].档案学研究,2024,(05):107-116.DOI:10.16065/j.cnki.issn1002-1620.2024.05.013.
- [9] 刘春辉.大数据背景下非遗档案保护对策[J].山西档案,2025,(04):182-184.

## Study on the Construction of AI-Driven Intangible Cultural Heritage Resource Databases

Ming Hui

(College of Applied Arts and Sciences of Beijing Union University, Beijing, 100191)

**Abstract:** Intangible Cultural Heritage (ICH) is the core carrier of national cultural inheritance, and resource databases serve as the key infrastructure for the digital protection and activation of ICH using modern technology. Currently, traditional ICH resource databases are limited by their technical architecture, generally exhibiting significant shortcomings in data processing, knowledge mining, and application activation—making it difficult to meet the needs of the digital age. This study systematically explores the overall framework, technical adaptation, and functional development of AI technology in the construction of ICH resource databases, and analyzes the practical challenges and corresponding countermeasures in the current AI-driven construction of ICH resource databases. The research aims to clarify the core logic and personalized application scenarios of AI-driven ICH resource database construction, address the pain points of traditional resource databases, transform ICH resources from "dormant data" into active materials for cultural innovation and educational communication, and provide theoretical support and practical references for the digital protection of ICH resources.

**Keywords:** Intangible Cultural Heritage; Resource Database; Artificial Intelligence; Digital Protection; Living Inheritance