

# 面向跨文化传播的AI多人有声剧生产： 技术赋能、实践挑战与演进路径

陈妍<sup>1</sup>, 李雅箏<sup>1,2</sup>

<sup>1</sup>安徽大学新闻传播学院, 安徽 合肥

<sup>2</sup>安徽新华发行(集团)控股有限公司博士后科研工作站, 安徽 合肥

收稿日期: 2026年4月1日; 录用日期: 2026年5月1日; 发布日期: 2026年5月11日

## 摘要

数字文化产业的蓬勃与听觉经济的升温, 使多人有声剧凭借沉浸式叙事与立体角色塑造, 成为讲述中国故事、勾勒国家形象的数字出版新载体。AI语音合成技术通过自动拆解剧本、同步生成多角色音色并快速适配多语言, 显著提升了优质IP的数字化产能与全球触达效率。AIGC技术红利重塑了创作生态, 并催生出“专业团队 + 大众参与”的协同模式, 为主流价值传播开辟了新路径。但情感细腻度、内容原创力、版权治理及跨文化适配仍是AI语音合成亟待突破的瓶颈。应从四个方面推进: 深化情感语音合成并建设文化专属语音库; 搭建人机互补的内容生产流程; 完善版权保护与行业规范; 针对国际受众制定精准传播与本地化方案, 让技术真正服务于文化传播与国家形象表达。

## 关键词

多人有声剧, AI语音合成技术, 文化生产, 跨文化传播

# AI-Generated Multi-Voice Audio Drama for Cross-Cultural Communication: Technological Enablement, Practical Challenges, and Evolution Pathways

Yan Chen<sup>1</sup>, Yazheng Li<sup>1,2</sup>

<sup>1</sup>School of Journalism and Communication, Anhui University, Hefei Anhui

<sup>2</sup>Postdoctoral Research Station, Anhui Xinhua Distribution (Group) Holding Co., Ltd., Hefei Anhui

Received: April 1, 2026; accepted: May 1, 2026; published: May 11, 2026

文章引用: 陈妍, 李雅箏. 面向跨文化传播的AI多人有声剧生产: 技术赋能、实践挑战与演进路径[J]. 新闻传播科学, 2026, 14(5): 49-57. DOI: 10.12677/jc.2026.145111

## Abstract

The flourishing digital cultural industry and the rising auditory economy have established multi-voice audio drama, with its immersive storytelling and multi-dimensional character building, as a new digital publishing vehicle for narrating Chinese stories and portraying the national image. AI voice synthesis technology significantly enhances the digital production capacity and global reach efficiency of high-quality IPs by automatically deconstructing scripts, synchronously generating multiple character voices, and rapidly adapting to multiple languages. The technological dividends of AIGC are reshaping the creative ecosystem, fostering a collaborative model of “professional teams + public participation”, and opening new pathways for disseminating mainstream values. However, nuanced emotional expression, content originality, copyright governance, and cross-cultural adaptation remain key bottlenecks to be overcome in AI voice synthesis. Progress should be made in four areas: deepening emotional speech synthesis and constructing culturally-specific voice databases; establishing human-AI complementary content production workflows; improving copyright protection and industry standards; and developing targeted communication and localization strategies for international audiences, to ensure that technology truly serves cultural communication and the expression of the national image.

## Keywords

Multi-Voice Audio Drama, AI Voice Synthesis Technology, Cultural Production, Cross-Cultural Communication

Copyright © 2026 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. 引言

在智能语音合成技术不断成熟与数字文化产业快速发展的双重背景下，我们正在经历一场“声音的复兴”。有声内容市场近年来呈现显著增长态势，其中剧情类有声剧作形态迅速崛起。该形态处于有声书与广播剧的交汇地带——既延续了有声书持续叙事的魅力，又融汇了广播剧丰富的声音艺术表达，成为传递文化内容、连接受众情感的重要载体[1]。与早期单播剧形态不同，当前系列化有声剧凭借连贯的叙事结构与灵活的单集时长，更加契合互联网传播特性与当代用户的收听习惯，能够通过立体空间感的塑造给听者带来沉浸式的收听体验[2]。

在文化强国战略与全民阅读政策推动下，有声剧的功能正逐渐由“听觉娱乐”延伸为“文化载体”。人工智能语音合成技术的广泛应用，推动了“专业生产 + 用户共创”生态的形成，UGC 模式的创新促使 AI 多人有声剧作品加速涌现。值得关注的是，在这些创作实践中，人工智能不再是被动的工具或对象，而是成为与人类创作者共同参与创作的主体之一[3]。2024 年，我国有声阅读用户已达 6.06 亿，占数字阅读总用户的 90.45%，这一体量奠定了有声内容承担文化传播功能的底座。当听觉媒介重回主流，如何用技术把“中国故事”讲得更好、让“中国声音”传得更远，已成为绕不开的话题。基于此，本研究以 AI 语音合成技术驱动下的多人有声剧为研究对象，将系统探讨其在文化强国建设中的赋能作用与实践路径，分析技术应用带来的三重契机与四重挑战，并从技术、内容、机制、传播四个维度提出发展路径，以期 AI 语音合成技术在有声剧领域的深入应用提供学理依据，并为文化传播实践提供操作层面的参考。

## 2. 多人有声剧的文化传播价值与传统生产困境

### 2.1. 多人有声剧的艺术特点与文化传播价值

多人有声剧是一种综合性的声音艺术。它把多声部演绎、音效与音乐编织在一起, 让听众仿佛置身故事现场, 不同配音演员的音色和表演风格相互映衬, 角色因而更立体, 叙事也更具情感张力。在保留原文叙事框架的同时, 声音的二度创作为文本注入新的生命力, 使这一形式成为文化传播的活跃通道。

在文化强国建设背景下, 多人有声剧展现出独特的文化传播价值。它能够将传统的文学经典、民间故事、历史传说等文化资源, 转化为更契合现代受众接受习惯的声音产品, 实现文化资源的活态传承与创新表达, 丰富数字出版的内容形态。有声剧《我和我的祖国》以庆祝中华人民共和国成立 75 周年为主题, 紧密围绕“强国复兴有我”的群众性主题宣传教育活动, 精心选取了“超级稻之父”杨守仁、“抗美援朝老兵”张锦华、“工人院士”方文墨三位不同时代、不同领域的典型人物作为原型, 分别讲述他们与祖国同呼吸、共命运的故事, 全面展现了中华人民共和国成立 75 年来所取得的光辉成就和伟大历程[4]。

多人有声剧同时承载着文化普惠的功能。它降低了文字阅读门槛, 让视障者、老年人或识字受限的群体也能借声音进入文化现场; 方言、环境声与民族乐音的嵌入, 又使地域气质和身份认同在听觉层面被重新激活。例如, 水墨动画《雾山五行》的声景被海外观众称为“东方美学的听觉注脚”, 《中国奇谭》的方言对白则唤醒了潜藏的文化记忆。这些实例显示, 多人有声剧在输出中国价值与审美时具备很大的潜力。

### 2.2. 传统生产模式的瓶颈与限制

尽管多人有声剧具有较高的文化价值, 但目前的有声剧制作流程, 从旁白选取到片段剪辑再到后期配音等环节, 高度依赖人工, 费时费力, 导致整体产出效率低, 投入成本高, 制约了有声剧的快速发展[5](表 1)。其生产长期受制于“手工作坊”模式, 高度依赖专业团队的密集协作, 导致生产效率低、成本高昂, 难以适应市场对规模化与即时化内容的需求, 从而限制了文化 IP 的转化效能, 以及在文化出海与国家文化战略中应有的作用。同时, 受困于高昂投入与低效回报: 专业配音演员单集酬金动辄 1~3 万元, 且排期紧张; 人工操作易因情绪不到位、念错台词或断句失误而返工重录。一部多人有声剧从立项到上线常需数月, 内容供给节奏与市场响应速度被大幅拖慢。

Table 1. The production process and challenges of traditional multi-player audio drama

表 1. 传统多人有声剧制作流程与挑战

制作环节	主要内容	持续时间	主要挑战
画本	剧本改编、角色分配	2~4 周	忠于原作并适配听觉叙事
CV 试音	演员选拔、声线确定	7~15 天	匹配角色需求与表演能力
录制	音频采集	3~6 周	协调档期与统一录制条件
对轨	多音轨合成	2~3 周	不同演员的节奏与音质统一
审听	质量审查	1~2 周	纠错并安排补录
后期	音效添加、BGM 融合	2~4 周	艺术表现与技术实现的平衡
上架	平台发布、宣传推广	1~2 周	应对市场竞争与注意力分散挑战

传统模式下的创作生态相对封闭, 创新活力不足。高门槛与资源壁垒把创作主体锁定在少数专业机构与资深团队, 公众难以介入; 题材与风格随之趋同, 难以回应多元市场需求。跨文化传播时, 这种多

人有声剧模式更显捉襟见肘。语言障碍与文化差异叠加,使中文有声剧在国际市场的接受度受限。多语言版本必须另聘外语配音团队,制作成本陡增,国际化步伐被明显拖慢。声音艺术比视觉媒介更易遭遇“文化折扣”,语言差异与声音表达的文化特异性常令听众产生理解偏差,情感共鸣也随之削弱。

### 3. AI 语音合成技术赋能多人有声剧的三重契机

#### 3.1. 生产端提质增效:从“手工作坊”到“智能工厂”

AI 语音合成技术借助自动化生成与智能化处理,把多人有声剧的生产效率和质量同时推上新台阶。在文本处理环节,系统可自动完成剧本拆解、角色标注和情感标记,前期准备的时间被大幅压缩。进入语音生成阶段,预训练 TTS 模型把书面文字直接转化为自然流畅的口语,声音带有人类语调和情感起伏。喜马拉雅平台推出的 AI 智能创作工具“云剪辑”,集合了智能音量、智能配乐、音转文剪辑、AI 分段、智能检测、一键成片等功能,有效降低了音频播客内容的创作门槛。通过“音频转文字剪辑”功能,创作者可以直接通过编辑文字文档来剪辑音频,大大提高了后期制作效率。截至 2023 年初,喜马拉雅平台已通过 AI 技术创作有声书专辑超 37,000 部,其中“单田芳声音重现”系列专辑总播放量超 1 亿次[6]。

AI 语音合成已能在数分钟内为剧本批量生成并即时切换多角色声线,无需再耗费数小时协调多位配音演员,实现了多角色音色的快速生成与灵活切换。借助多角色音色解耦与动态路由技术,用户仅需输入剧本文本,系统即可自动分配不同声线完成对话演绎,将传统需数小时完成的多角色配音工作压缩至分钟级别。目前,面向普通用户的 AI 语音合成工具已大量出现,它们把声音克隆、多语种合成与转换、多情绪角色库及情绪自动匹配等功能整合在同一平台,典型平台代表包括智声云配与 NOIZ 等平台(图 1)。

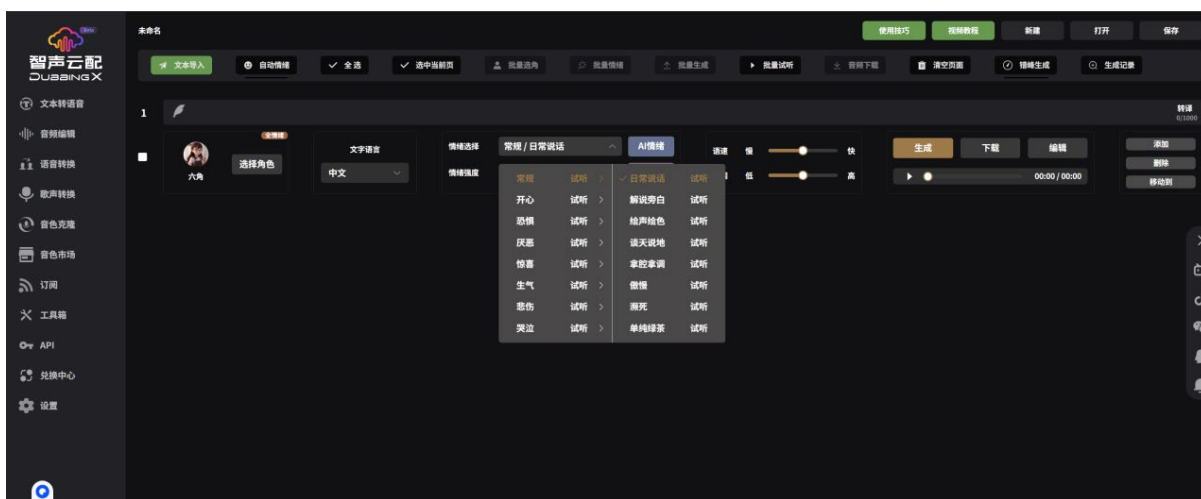


Figure 1. Smart voice cloud with text-to-speech interface

图 1. 智声云配文本转语音功能界面

传统的多轨混音、音效铺陈和音量均衡等都得靠专业工程师手工完成,步骤繁琐。如今, AI 技术可以深度介入音频的后期处理。基于算法, AI 语音合成工具可在无人值守下跑完全程,还能依据脚本情绪起伏实时调节音高、语速与情感强度,细腻度远超传统 TTS,在压缩制作周期的同时,也把音频作品的艺术边界往前推了一步(表 2)。

**Table 2.** The main advantages of 2: AI speech synthesis technology in the application of multiplayer audio drama  
**表 2.** AI 语音合成技术在多人有声剧应用中的主要优势

应用领域	传统模式痛点	AI 技术解决方案
剧本处理	人工分析耗时耗力	自然语言处理自动标注
配音录制	演员档期难协调, 成本高	AI 语音合成, 多声线生成
后期制作	手工对轨繁琐, 易出错	多轨自动混音, 智能对齐
多语言适配	需要重新聘请外语配音	跨语种语音合成与转换
用户参与	专业门槛高, 大众难参与	简易创作工具, 声音克隆

### 3.2. 传播链时空扩展：从“本土传播”到“全球覆盖”

新世代青年在智能算法与多元文化元素新背景下的传播实践，催生出了“转文化传播”(Transcultural Communication)形态。通过跨文化的符号、信息和文本的再生产，体现出“杂糅”和“共生”的新趋势[7]。AI 语音合成技术为多人有声剧的跨文化传播提供了技术可能。通过跨语种语音建模与算法优化，现代 TTS 系统已能够实现高质量、多样化的语音生成，有效打破地域限制，使中华文化能够以更本土化、更亲切的方式传递至全球不同地区。中央广播电视总台推出的《千秋诗颂》多语种版，正是借助 AI 技术将古诗词转化为动画作品，并在欧洲与拉美多国主流媒体播出，触达海外受众近 1 亿，充分体现了技术赋能下文化传播的覆盖广度与“一键出海”的高效性[8]。

技术更新迭代后，AI 语音合成已支持个性化定制：系统可依据不同地区受众的文化偏好与听觉习惯，实时调节语速、语调与音色，输出贴合当地语境的音频。这种文化适配能力有助于提升传播精准性，也更容易唤起海外听众的情感共鸣，使中国故事被顺畅接纳。在传播渠道上，AI 技术推动多人有声剧向多平台、多场景延伸。随着智能硬件和物联网技术普及，有声内容不再局限于音频平台，广泛渗透至车载音频、智能家居、可穿戴设备等多元场景。

### 3.3. 生态多元化：从“专业垄断”到“大众共创”

AI 语音合成技术削弱了专业壁垒，使更多创作者得以进入音频生产领域。集成该技术的工具不仅降低了制作成本，也简化了流程，鼓励公众主动参与文化内容生产。同时，它打破了平台间的隔阂，让多元主体能够协同创新，共同丰富内容生态。借助 AI 辅助平台，听众可自行生成个性化剧集，参与情节设计与改编，逐步形成一个开放、共享的文化生产网络。为激活创作生态，喜马拉雅把“戏精大汇”等年度盛典做成固定仪式，用奖项直接回馈作者。到 2023 年底，站内已沉淀 4.88 亿条音频，覆盖 459 个品类，总时长突破 36 亿分钟，逐步构建出一个丰富而多元的“声音宇宙”。平台还向创作者开放全链路支持，把多人有声剧从专业制作推向大众共创。

## 4. 智能技术用于多人有声剧创作的主要挑战

### 4.1. 技术挑战：“机械凉感”与情感传递的落差

尽管 AI 语音合成在自然度与流畅度上已有明显提升，但在情感渲染和细微语调变化层面仍与真人配音有一定差距。目前由算法生成的声音普遍缺少人类演员特有的情感温度与艺术张力，难以精准传递复杂情绪和文化韵味，这种“机械凉感”容易削弱听众的沉浸体验与情感共鸣，使文化内容的感染力随之下降。

多人有声剧是长叙事，角色情绪需随剧情自然递进，现有 AI 语音合成技术多以句或段为单元，缺少长程记忆，句间段间的情感衔接常显突兀，难以形成连续的情绪弧线，更深层的困难在于模型对文本的文化语境、角色关系及戏剧冲突缺乏语义把握。此外，人工智能在收集人类知识、提取特征、输出知识

的过程中, 人类知识的地方性特征可能被消解[9]。

#### 4.2. 内容同质化: 文化多样性与创新困境

AI 语音合成技术被大规模投入使用, 也可能带来内容趋同的隐忧。因为主流的语音合成模型往往依赖海量且标准化的语料进行训练, 其生成的语音在音色与表达风格上高度一致, 缺少人类配音演员所具有的艺术创造力与细腻的情感层次。这种趋同不仅会稀释文化内容的多样性, 还可能让听众对单一、机械的语音产生审美疲劳, 从而削弱文化传播的广度与深度。

在技术层面, 训练数据的偏差与覆盖范围不足是导致语音同质化的首要原因。商业系统为兼顾通用性, 普遍采用以标准普通话为主、说话人年龄集中在 20~40 岁的中性语料, 结果合成语音在“声音肖像”上呈现系统性偏移, 既难还原特定年龄段的发声特征, 也无法承载方言带来的地域文化特点。在算法层面, 深度学习模型以损失函数最小化为目标的优化机制, 也客观上促成了语音风格的“趋中化”。模型倾向于输出训练数据中的主流模式或平均特征, 导致生成的声音风格容易收敛至一种安全但平庸的“中间值”, 缺乏人类配音演员那种带有个人特质、甚至有些冒险的创造性表演。这种算法“安全区”内的输出, 与艺术创作中对个性与极致的追求形成本质冲突。

#### 4.3. 版权治理: 权责模糊与法律风险

AI 有声剧面临的版权治理困境, 本质上揭示了新技术如何激化了文化生产领域固有的资本、劳动与产权矛盾, 并正在重塑着数字时代的权力与利益分配格局。

首先, 训练数据的权利困境是数字资本进行“声音圈地”与“原始积累”的症候。海量、低成本甚至未经清晰授权的语音数据, 是驱动 AI 模型的核心生产资料。平台与开发者对数据的攫取与利用, 遵循着典型的资本逻辑——将本属于公众或个体的声音资源私有化、资产化, 转化为可产生垄断性利润的“声音资本”。这一过程充满了产权模糊性, 其合规性危机恰恰暴露了在数据资本主义模式下, 法律对新型生产资料的保护与规制严重滞后。

其次, AIGC 作品的权属模糊, 这种权属不确定性或将制约 AIGC 的商业化进程, 包括版权交易主体不明、维权资格缺失及收益分配机制模糊等问题[10]。深度伪造技术的滥用再添伦理与法律风险。语音克隆技术可能被用于模仿公众人物或特定个体声音, 实施诈骗、诽谤或传播虚假信息, 对社会秩序与个人权益构成严重威胁, 如何在鼓励创新的同时实施有效监管, 已成为全球共同的难题。

最后, 多元主体利益分配机制的缺失, 是新兴“产消合一”模式下数字劳动剥削风险的直观反映。在“大众共创”的愿景下, 业余创作者的情感、时间与创意投入被转化为平台的内容资产与流量, 但他们往往难以分享由此产生的核心收益。这种看似自由的参与, 可能掩盖了新型的、更为隐蔽的数字劳动剥削形式。一个公平、透明、可追溯的分配机制, 不仅是为解决纠纷, 更是对参与性文化中劳动者权益的基本保障, 是遏制资本无序扩张、构建良性生态的必要规则。

因此, 版权治理绝非单纯的技术或法律问题, 而是一个深刻的政治经济议题。构建与之相适应的新机制, 意味着需要在数据产权、AI 生成物的法律地位以及多元劳动价值的公平分配上进行制度创新, 这实质上是对文化生产领域新型生产关系的一次重要调整。现阶段, AIGC 产出能否构成作品、权利归谁, 在我国仍处灰色地带。《著作权法》与主流学说均以自然人的独创性智力投入为要件, 且 AI 只能作为工具, 无法成为作者。因此, AI 有声剧的版权归属仍争议不断。用户、开发者、数据提供方都可能主张权利, 而裁判尺度尚未统一这种不确定性可能。

#### 4.4. 跨文化适配: 文化折扣与传播壁垒

在跨文化传播场景中, AI 语音合成技术面临着文化适配的多重挑战。全球市场的多样性决定了不同

国家和地区在文化需求、审美习惯和隐私观念上存在系统性差异[11]。若仅将内容简单翻译并转换为目标语言语音, 往往难以契合当地受众的听觉偏好与文化认知, 甚至可能因文化差异引发误解或抵触情绪, 产生“文化折扣”现象[12], 进而制约多人有声剧等内容的国际传播效果。

目前, 主流的跨语种语音合成系统大多沿用“先翻译文本, 再生成语音”的串行路径。虽然这一路径能在语言符号层面完成转译, 却难以把文化符号系统一并搬过去。文化是一张由价值观、社会规范、思维方式、历史语境和幽默习惯交织而成的意义网, 其深层意涵往往超出字面所能承载。目前, 跨语种语音合成系统普遍采用统一的声学模型进行语音生成, 缺乏对目标文化听觉习惯与审美偏好的适应能力, 无法在表演风格、情感强度与叙述节奏等方面实现针对性调整。以中文独有的“江湖”为例, 直译成“rivers and lakes”后, 原本饱满的文化意象与情感温度在合成语音里被大幅稀释, 厚重的语义背景随之流失。更进一步, 仅靠字面对译与声音替换, 系统对文化专有项几乎束手无策。

## 5. 数智时代多人有声剧赋能跨文化传播的实践路径

### 5.1. 技术优化路径: 深化情感计算与语境感知建模

应突破传统 TTS 架构, 引入更先进的情感计算算法和语境感知技术, 使 AI 系统能够根据文本内容和场景上下文自动调整语音的情感色彩和表达方式, 推进研发深度情感嵌入与上下文感知模型。通过情感编码技术, 在生成过程中引入情感参数, 使模型能够生成特定情感的语音, 如高兴、愤怒、悲伤等。同时, 加强对中文特有语言现象和文化内涵的理解与处理, 如诗词韵律、成语典故、地方俗语等, 使合成语音更能体现中华语言文化的独特魅力。并且采用对抗性领域自适应等迁移学习技术, 让通用语音模型在文化语料上进行精细化微调, 使其学会中文特有的韵律节奏和文化表达习惯。此举是模型产出“有温度”“有文化”声音的基础。

### 5.2. 内容创新路径: 构建人机协同与多元化的创作生态

为抵御内容同质化风险, 必须明确 AI 的工具属性, 将人类的文化洞察与艺术创造力置于核心地位, 构建人机协同的新内容生态。AI 承担海量试错、高效生成、批量处理的“数字劳动力”角色; 而人类创作者则专注于核心创意、审美决策、艺术把关等“价值创造”环节。在创作前期, 创作者利用 AI 进行剧本情绪分析、角色音色海量试听与初步匹配。在中期, AI 快速生成多个表演版本的干音, 供创作者筛选和精修。在后期, 创作者指导 AI 完成多轨混音和效果添加。

内容创新上, 可搭建“AI+ 优质文化资源”的创作框架。一方面, 系统梳理并标注中华优秀传统文化资源, 建成文化语音数据库和知识图谱, 为模型供给多元的训练语料。另一方面, 面向地方戏、民族史诗、传统评书等特定形态, 开发专属语音合成模型, 形成定制化 TTS 系统, 从而维护文化表达的多样与独特。鼓励文化机构、平台与技术方联手, 对既有文化遗产进行体系化的 AI 适配, 把优质 IP 转化为“声音”, 为经典文学、非遗故事设计专属音色与表演风格库。同时支持立足地方文化的原创 AI 有声剧, 借助方言合成技术凸显地域标识, 喜马拉雅“非遗声音博物馆”计划正是此类尝试。

### 5.3. 机制构建路径: 构建合规、可信、可持续的治理框架

针对版权治理困境, 关键在于把制度补全, 为 AI 产出内容搭好版权框架和行业规则。第一步, 训练材料从哪儿来、谁授权, 必须一目了然, 不留模糊地带。接着, 划定 AI 作品的权属与收益怎么分, 让创作者、技术方和声音提供者都能拿到应得份额。与此同时, 嵌入深度伪造识别、数字水印和声纹追溯, 一旦被盗用就能一路追回去。最后, 行业要自立规矩, 把伦理准则和使用标准写清楚, 让 AI 语音合成朝着有益的方向发展。

在法律法规层面,国内已出台一系列相关政策和规定。2022年11月,发布《互联网信息服务深度合成管理规定》,对使用深度合成技术提供互联网信息服务的情形提出了专门性要求。2023年7月,国家网信办联合多部门公布《生成式人工智能服务管理暂行办法》[13],对生成式AI研发及服务作出明确的规定。这些法规为AI语音合成技术的健康发展提供了法律保障,但还需进一步细化相关实施细则和标准。

#### 5.4. 传播增效路径:实现从“技术翻译”到“文化转译”的跨越

要让跨文化传播真正落地,传播环节必须同时解决技术适配与文化解码的关键问题。AI语音合成系统首先要能读懂目标语境,依据文化差异随时调整语调、节奏乃至表演方式。在此基础上,再搭建专门的知识库与规则引擎,把隐含的价值观念、情感结构和文化符号转写成机器可读取的参数,使技术运行与语境需求同频共振。

在国际传播实践中,可以引入人机协同的跨文化创作流程,邀请目标文化背景的专家参与脚本改编与表演指导,以兼顾文化表达的准确性与亲和力。渠道层面,可融合“平台出海”与“生态合作”思路:一方面鼓励国内音频平台推出国际版App,内置低门槛的多语种AI创作与分发工具,形成文化出海的主阵地;另一方面与Netflix Audio、Audible等国际平台以及当地媒体、文化机构开展联合制作、版权共享和内容联运,借助其成熟渠道与本地影响力降低进入壁垒,实现从“走出去”到“融进去”的跨越。

数字内容出海已成为中国文化传播交流的重要载体。中国文化出海需面对复杂的信仰派别和生活文化禁忌[14],多人有声剧的跨文化传播可借鉴此模式,通过深入的目标市场研究和文化分析,制定“一国一策”乃至“千人千面”的更精准的传播策略,增强中国文化的国际传播效果。

## 6. 结语

AI语音合成技术为多人有声剧的生产与传播带来了革命性的变革,通过提升制作效率、拓展传播边界以及促进内容生态的多元化,有力推动了多人有声剧的模式转型。但也需要清醒地认识到,AI语音合成技术在多人有声剧领域的应用仍面临多方面的挑战。具体表现为技术层面的语音自然度不足、创作层面的风格同质化倾向、版权治理机制的不健全以及跨文化传播中的语境适配难题。这些问题不仅源于技术本身的发展阶段,更涉及文化传播的内在规律与伦理边界。亟需从技术优化、内容创新、制度构建与传播策略四个维度系统施策,构建人机协同、健康可持续的行业生态。

随着智能语音合成技术的持续迭代和应用场景的深入拓展,多人有声剧有望成为中华文化国际传播的重要载体和国家文化软实力建设的有力支撑。通过推动技术与艺术的深度融合,AI语音合成将不仅重塑声音内容的生成方式,也将重构文化叙事的传播格局,为文化强国战略注入新动能。在这一进程中,应始终坚持以文化价值引领技术应用,确保人工智能增强而非削弱文化表达的感染力与认同度,使中国故事能以更生动、更亲切、更具深度的方式抵达全球受众。

## 参考文献

- [1] 李雅箏,刘宇星. AIGC 技术赋能数字音频内容生产:应用场景、存在问题与应对策略[J]. 数字出版研究, 2023, 2(3): 13-20.
- [2] 陈企依. 从书写文本到有声剧:莫言小说“有声化”研究[D]:[硕士学位论文]. 杭州:浙江大学, 2024.
- [3] 李雅箏,骆美茹,谢李宁欣. 从工具智能迈向生态智变:AIGC 赋能智慧出版知识生产革新[J]. 出版广角, 2025(7): 3-10.
- [4] 孙革. 传统广播剧向有声剧转化创新路径初探——以有声剧《我和我的祖国》创作传播为例[J]. 北方传媒研究, 2025(4): 26-30.
- [5] 骆迅,杨杰,宋施恩,等. AIGC 在芒果TV有声剧自动化生产中的研究与应用[J]. 广播与电视技术, 2023, 50(11):

---

17-21.

- [6] 黄国斌, 蒋海鸥. AI 赋能出版产业的思路 and 对策[J]. 出版广角, 2023(12): 39-43.
- [7] 史安斌, 杨晨曦. 数智时代新世代青年的跨文化传播: 创变、挑战与愿景[J]. 对外传播, 2025(6): 4-8.
- [8] 刘珊. 新技术浪潮下的智能视听步入深度变革新阶段——2023 年人工智能赋能网络视听产业观察报告[EB/OL]. [https://www.dzwww.com/xinwen/shehuixinwen/202310/t20231021\\_12985193.htm](https://www.dzwww.com/xinwen/shehuixinwen/202310/t20231021_12985193.htm), 2025-09-18.
- [9] 刘大年, 曹月. 知识的幻象: 人工智能与知识变迁[J]. 现代出版, 2024(9): 53-67.
- [10] 李文冰. AIGC 版权治理的兴起: 争议议题、国际实践与启示[J]. 出版发行研究, 2025(6): 5-13.
- [11] 马雪骢. 中国网剧对外传播面临的问题、挑战与路径优化[J]. 济南大学学报(社会科学版), 2025, 35(4): 179-186.
- [12] 陈文泰, 孙仲伯, 李卫东. 权力让渡、话语冲突与空间建构——跨文化传播中多级文化折扣的逻辑理路[J]. 新闻与传播研究, 2024, 31(10): 38-53+127.
- [13] 李婧文, 李雅文. 深度合成技术应用与风险应对[J]. 网络与信息安全学报, 2023, 9(2): 184-190.
- [14] 徐剑, 李晨. 走向“全球南方”: 文化出海与国际传播效能提升[J]. 学术论坛, 2025, 48(1): 138-148.