

# 人工智能在有声书制作和传播中的应用研究

熊梦欣

北京联合大学应用文理学院新闻与传播系，北京

收稿日期：2026年6月1日；录用日期：2026年6月23日；发布日期：2026年7月6日

## 摘要

人工智能技术不断赋能各行各业，音频领域的有声书产业也正积极探索与AI的深度融合之路。基于参与观察法和案例分析法分析人工智能技术在有声书产业制作与传播中的应用、影响、局限性及优化建议，为听觉领域智能转型提供参考，发现在有声书生产创作方面，人工智能的介入降低了使用门槛，提高了生产效率。在传播分发方面，人工智能深化了服务的精准化、个性化与互动性。但是，人工智能技术在有声书的应用中不可避免地存在同质化内容泛滥、声音情感化不足、版权归属模糊等问题。在人工智能与有声书的结合运用上，需要审慎把握，促进智能音频行业的良性发展。

## 关键词

人工智能技术，有声书，AIGC，生产传播，音频

# Application Research of Artificial Intelligence in Audiobook Production and Dissemination

Mengxin Xiong

Department of Journalism and Communication, College of Applied Arts and Sciences, Beijing Union University, Beijing

Received: June 1, 2026; accepted: June 23, 2026; published: July 6, 2026

## Abstract

Artificial intelligence technology continues to empower all walks of life, and the audiobook industry in the audio field is also actively exploring in-depth integration with AI. Adopting participant observation and case analysis methods, this paper analyzes the application, impacts, limitations and suggestions of artificial intelligence technology in the production and dissemination of the audiobook industry, so as to provide references for the intelligent transformation of the auditory field. The study

**finds that in the creation and production of audiobooks, the intervention of artificial intelligence has lowered the entry threshold and improved production efficiency. In terms of dissemination and distribution, artificial intelligence has realized more refined, personalized and interactive services. Nevertheless, the application of artificial intelligence in audiobooks inevitably faces problems such as the proliferation of homogeneous content, insufficient emotional expression of synthetic voices, and ambiguous copyright ownership. The integration of artificial intelligence and audiobooks needs to be rationally regulated to promote the sound development of the intelligent audio industry.**

## Keywords

**Artificial Intelligence Technology, Audiobook, AIGC, Production and Dissemination, Audio**

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. 人工智能技术赋能有声书生产创作

### 1.1. AI 对有声书文本的生成与润色

移动音频平台与数字阅读平台都是以内容为主导的平台，有声书的内容主要来源于文学作品的合法授权和人工改编。以 GPT、DeepSeek 为代表的生成式 AI 的引入，能够对有声书的内容进行智能化处理，进行文本生成文本的操作。在有声书生产流程的前期筹备阶段，平台与出版社或作者签约之后，编辑团队需要将书面文本优化为适宜有声书形式的文本，进行书籍的改编。AI 具有语义重构能力，自然语言处理技术 (NLP) 能够让计算机理解并处理人类语言，从中提取出有用的信息，帮助人类更高效地处理各种任务。使用 NLP 技术优化文本，AI 能够识别书籍中的人物对话、时空环境，进而拆分长句，补充口语化表达、标注情绪与角色标签，这大大加快了前期筹备阶段的文本处理效率，使书面文字更好地适应有声书的听觉体验。

### 1.2. AI 语音合成助力有声书生产降本增效

文语转换(text to speech, 以下简称 TTS)技术是将计算机产生的或外部输入的文字信息转变为可以输出的听得懂的流利口语，是 AI 有声书发展的关键技术[1]。传统有声书依赖专业主播录制，不仅成本高昂，而且声音风格受限于主播个人特质，难以满足不同题材的适配需求。而 TTS 技术能够以较低的成本，将任意文本快速转化为音频，解决一些真人主播可能受限于生理机能、档期协调的掣肘。2025 年 5 月，亚马逊旗下的有声书 Audible 宣布大规模启用 AI 配音技术，计划将能提供 100 多种人工智能生成的英语和其他语言语音，Audible 对希望利用这项技术的出版商提供两种选择：一种由 Audible 管理生产；一种是“自助服务”，即出版商在 Audible 的人工智能技术的帮助下制作自己的有声书[2]。当前，有声书单纯的语音合成功能正朝着情感化、音色定制化、多语种化等多功能方向延展。有声书平台购买版权后，可利用 TTS 技术将优质文学作品直接转变为有声书，为有声书生产提质增效。

作为行业领导者，喜马拉雅 FM 很早就将 TTS 技术作为其内容生态建设的重要一环。截至 2023 年底，喜马拉雅平台上 AIGC 内容的总时长已超 2.4 亿分钟，其中，有超过 10 部 AIGC 有声专辑的播放量破亿[3]。这一数据充分说明了 AIGC 技术在有声书领域的巨大影响力。比如，通过模仿真人主播音色，喜马拉雅 FM 的四位 AI 音色——喜小道、喜小迪、喜小玖和苏小刀能够独立完成有声书的内容制作。与专业人员配音成本相比，AI 实现了有声内容的快速迭代，极大地提高了平台内容库的丰富度，并有效降低了制作周期和费用，并提升了听户的内容消费体验，为有声书的生产提供“AIGC + UGC + PGC + PUGC”

的模式可能性。

行业内正在向人工智能的情感表达方面努力，早期，用户对 AI 有声书的听感评价大多为“机器人朗读”、“缺乏感情”，如今 AI 有声书的声音媲美真人，用户的听感更加逼真和情感化。情感 TTS 是一种能够生成带有特定情感或情绪的语音合成技术，它不仅能够将文本转换为语音，还能根据上下文或指定情感标签调整语调、语速和音量等韵律特征，使输出语音更自然、富有表现力。喜马拉雅 FM 已推出多款 AI 合成主播，并将其应用于海量网络文学、知识类和新闻类有声书的制作，极大地提升了用户听觉体验，打破了以往机械感强烈的瓶颈。

### 1.3. AI 后期制作降低用户创作门槛

有声书后期制作环节是音频从录制完成到可供传播的关键步骤，在 AI 辅助出现之前，有声书的后期制作高度依赖于人工编辑，比如人工逐字逐句地听审校对、耗时的章节划分等，这使得编辑在内容生产中扮演了严格的把关人角色，有声书的上线周期与质量受制作人员经验影响较大。在 AI 技术的推动下，智能创作软件为有声书制作的流程提速，为创作者降低了内容生产的门槛。

MiniMax 是全球领先的通用人工智能科技公司，自主研发了一系列多模态通用大模型，能够处理超长文本，理解、生成并整合成音频。在 MiniMax 语音合成中，具有“说书”的功能，可以根据语言、口音、性别、年龄来选择适合的 AI 音色进行生成。喜马拉雅 FM 也为创作者提供了智能创作工具，“音剪”是喜马拉雅 FM 推出的一站式 AI 音频创作平台，提供多情感多模态的 AI 辅助工具，强大的在线剪辑能力，辅助用户轻松创作出优秀的音频作品。“音剪”集合了多项 AI 音频创作功能，包括 AI 音频剪辑、AI 配乐、文章转语音、AI 小说、AI 分段、智能检测以及一键成片等。2025 年 7 月，全新音频智能创作平台“音剪 AI”上线，上新了多种功能，包括 50 种 AI 音色，多人制作实时同步，喜马拉雅签约主播免费试用。喜马拉雅平台整合了录音、剪辑、混音等制作流程 AI 技术深度渗透到音剪 App 中，将故事脚本、书籍粘贴进入软件，能够选择“新闻腔”“温柔音”等专业音色进行朗读，并能够生成有声小说，自动拆章节、识别角色，搭配不同 AI 音乐，能够帮助内容创作者快速打造高品质有声书。AI 可贯穿有声书的录制、后期等全过程，相比于人工创作过程的时间，AI 的创作时间提效超过 50 倍；而由 AI 制作人优先参与的 AI 精制作，相比于人工创作提效超过 3 倍[4]。

AI 音频剪辑工具还能够提升音质稳定性，识别噪声、口水音等问题，对创作者的录制进行调整，这使得即使是业余主播在非专业环境下录制的作品，也能够“一键优化”。从而提升有声书作品质量，满足平台要求，有效降低了内容创作的门槛，吸引更多业余用户投入有声书内容创作，丰富平台作品。AI 已经深度渗透到有声书制作的流水线后端，目前已有超过数十万名创作者从中受益，为创作者提供了丰富的创作选择，使得无论新手还是资深主播都能轻松利用这些工具创作出优质的音频作品。同时，自动化剪辑工具将专业的后期制作人员从简单枯燥工作中解放出来，转向更加需要创造力的工作当中。

媒介可供性理论最早由吉布森提出，用以描述环境为行为主体提供的行为可能性。潘忠党在 2017 年将媒介可供性理论用于传播学，并提出“媒介可供性”可以分为生产可供性、社交可供性和移动可供性三个维度。在人工智能技术的加持下，有声书的生产可供性体现在降低门槛、减少成本、提高效率方面，AI 技术拓展了更多用户参与有声书的生产制作，传统由专业主播和团队主导的生产模式逐渐向平台化、协同化、智能化生产模式转变，推动有声书产业生产关系发生重构。

## 2. 人工智能技术助推有声书传播分发

### 2.1. 算法匹配的精准推荐和个性化服务

卡茨提出的使用与满足理论认为，受众基于特定需求主动选择媒介内容，而人工智能技术的介入使

得用户主动寻找的行为变为平台精准预判。平台借助算法对海量内容进行筛选、排序和分发，不仅影响用户接触何种内容，也会在一定程度上影响内容的曝光与传播。基于大数据与深度学习算法，能够对用户的基础属性、收听历史、搜索行为等日常数据进行挖掘，构建精准的用户画像，为用户提供精准化和个性化定制服务，实现内容与用户需求的毫秒级匹配。AI 所具备的技术优势和显著特点，为有声书实现精准化服务提供了技术支撑。

具体而言，如荔枝平台从可量化的指标入手，综合运用点击量、播放量、完播率、点赞、评论、转发数等正负反馈行为作为特征，进行深度学习训练并调整和优化模型，以提供更符合用户体验及期待的精准传播内容与服务[5]。以喜马拉雅 App 为例，基于平台的用户行为数据，如收听时长、点赞、收藏、评论、跳过行为、搜索记录等，通过协同过滤算法与深度学习模型，构建了精细化的用户画像，理解用户需求，从而推荐用户感兴趣有声书作品。例如用户连续收听 3 部“职场成长”类有声书后，AI 会自动推荐同题材的作品。这种功能能够提升用户的收听体验，降低用户的搜索成本，实现用户碎片化收听的最优化，提高用户粘性和满意度。

个性化推荐的口号是“为你量身定制”，背后依赖的是庞大的数据收集与算法逻辑。网易云音乐、Spotify 等流媒体平台的个性化推荐功能也是市场竞争和吸引用户的核心。2025 年 2 月，喜马拉雅 App 与 DeepSeek 合作，接入 Deepseek-R1 模型，在 App 首页的推荐页上新了“DeepSeek 猜你想听”，根据用户收听习惯提供个性化服务。除此之外，喜马拉雅 FM、番茄畅听、讯飞有声等平台在“听书”功能中提供多种 AI 音色，用户能够自行选择符合自己需求的个性化音色收听书籍。“AI 换声·爸妈分身”功能是喜马拉雅自研的一种个性化定制算法，它使用少量录音数据就能模拟用户音色，并将其替换为有声书的音频。在喜马拉雅 App 的“我的”页，点击“全部服务”，选择“儿童服务”下的“AI 换声”，可以定制我的专属音色，父母可以定制属于自己的 AI 声音给孩子讲故事、读名著等。对于用户来说，算法精准推荐的核心是最大化满足用户的需求。不仅能够满足用户的收听动机，也能挖掘用户的潜在需求，从而提升用户体验，增强了用户对于平台的黏性，将平台打造成个人专属内容电台。

## 2.2. AI 助手驱动的智能交互体验

人工智能基于人机协同实现人类福祉，目前 AI 在社会中的渗透度和接受度不断提高，各种 AI 软件层出不穷，多个平台推出了专属自己平台的 AI 智能助手。这种智能化交互设计通过技术创新构建双向互动的体验，依托自动语音识别(ASR)、自然语言理解(NLU)、自然语言生成(NLG)，识别用户需求并快速响应。

在实际应用中，这种智能化交互设计深度覆盖用户收听全流程。喜马拉雅 FM 内置的 AI 助手——小雅，支持深度思考、创建自己的 AI 频道、实时通话、每日资讯功能、切换 AI 助手音色、全语音回答等功能。喜马拉雅 FM 还推出了全新功能“AI 听中间”，直击长音频内容消费痛点，AI 助手为听众总结全文有声书、内容提炼、大纲生成或背景追溯，用户可以在听书时随时唤起助手，通过文字或语音发起提问，用户能够在不中断收听、不切换 App 的情况下，与 AI 助手进行交流。在海外市场，亚马逊旗下的有声书平台 Audible 针对《傲慢与偏见》等经典作品测试了“Ask a Question”功能，允许听众在不暂停播放的情况下，询问关于 19 世纪英国婚姻观等历史背景问题。这种“边听边问”的模式不仅没有破坏沉浸感，还能通过 AI 助手进行即时的知识补偿，消解用户收听门槛。

在传统有声书传播模式中，受众被动接受，而智能化交互助手的介入，改变了这种局面，使有声书从播放工具转变为智能书友，实现知识的深度渗透。这种功能设计增添了听众“听书”的趣味性和知识性，提高了用户的收听效率。与 AI 助手进行对话互动，这种拟人化的反馈也能够为用户提供社交和情感需求，让听众收听有声书的体验更加充满活力，增强听众对平台的好感。

### 2.3. 多终端场景化的智能分发

人工智能技术进一步提升了满足用户需求的能力，影响着用户的媒介使用习惯和消费行为模式。人工智能技术通过算法结合地理位置和时间数据，可以在适当的场景和时间自动推荐适合的有声书内容，精准适配听书环境，进一步提高读者听书的便利性和满意度[6]，推动用户形成碎片化、伴随式的消费习惯。随着物联网与智能终端的发展，人工智能使有声书的传播不再局限于智能手机终端，而逐步渗透进车载系统、智能家居、穿戴设备等场景中。跨终端同步技术基于云端数据存储，实现收听进度、收藏列表、定制模板在手机、平板、智能音箱、车载系统等终端的无缝同步。

在实际应用中，喜马拉雅 FM 推出了小雅智能音箱，理想汽车、特斯拉汽车等车载终端也能够同步播放有声书平台作品，搭配用户出行场景，提高用户通勤幸福感和有声书市场占有率。AI 多终端场景化适配功能能够打破单一终端局限，拓宽传播渠道，触达更多潜在用户，提升用户使用时长。同时，开展用户付费和商业化新空间，拓展有声书产业的盈利边界。

## 3. 人工智能在有声书制作与传播中的局限性及发展建议

### 3.1. 规模化生产导致同质化内容泛滥

人工智能技术赋予了有声书低成本、高效率的生产能力，在释放了内容生产力的同时，由于算法逻辑的统一与创作门槛的降低，催生了严重的内容同质化问题。许多音频平台利用 AI 批量将网络小说转化为音频，生产的网络作品类型高度集中于市场热度高、变现能力强的领域，如悬疑推理、都市言情、武侠玄幻题材。这些有声书往往套用相似的语音合成模型。当成千部有声书呈现出高度一致的语调、停顿和节奏时，就失去了有声书的艺术独特性。

平台为降低传播风险，保障商业利益，倾向于使用 AI 生产符合大众审美，易获点击的作品。同质化的有声书泛滥会造成听众审美疲劳，用户失去收听兴趣，流失现有用户。并且用户会因内容套路化、听感相同，而选择回避有声书，同质化已经成为影响用户留存的关键因素。还会掩盖优质原创内容的生存空间，阻碍有声书的创新生产，最终制约有声书的发展。

### 3.2. 声音的情感表达仍存在缺陷

传统影视作品大多通过画面与声音的结合来酝酿情感，当视觉画面成为用户关注的焦点时，作为辅助元素的声音，其情感饱满程度的要求就相对较低。而有声书完全依赖声音来营造戏剧冲突，通过声音叙述故事情节、描绘多样化的角色性格以及展现不同情境下的情感起伏[7]。有声书本质上是一种声音表演，需要主播根据情节进行深度的二次创作。尽管情感 TTS 技术已经取得显著进步，AI 有声书从机械音到自然化大幅跨越，但在情感传递的精准度、音色的自然与流畅层面，仍然与真人主播存在着明显的差距。以喜马拉雅推出的 AI 系列作品《单田芳声音重现》专辑为例，不少听众在评价时直言，合成语音只复刻了单田芳先生的音色，却没能还原其独特的节奏与神韵。

人工智能技术难以真正理解人类情感的复杂内核，AI 的情感表达更多依赖于算法对文本标签的匹配，而不是像真人一样，能够对剧情、角色、语境深度共情。情感 TTS 技术的核心逻辑是通过 NLP 算法提取文本中的情感关键词，并匹配对应的语调、语速、音量参数，但文学作品中的情感表达往往并非直白呈现，而是蕴含在细节描写、对话转折、场景铺垫之中。AI 难以复制真人主播的情感张力，真人主播的声音承载着个人的音色特质、表达习惯与情感理解，产生的一些不完美反而使有声书更具温度，这种特色正是有声书的核心魅力。AI 的情感表达遵循算法预设，长期收听难免会让用户感到刻板和冰冷。所以，在有声书演播中，AI 会表现出缺点，难以引起听众的情感共鸣。

### 3.3. 版权归属问题模糊不明

当一部有声书由 AI 自动生成时，其著作权主体应该归属于软件开发者、指令输入者，还是属于公有领域？目前法律界对此没有统一的界定。例如，2023 年 1 月，苹果公司悄然在 Apple Books 应用中上线系列由人工智能技术生成、播讲的有声读物，后遭到由美国音频业务巨头 Findaway 公司联合业内诸多播音主播发起的谴责，认为苹果公司收集他们的播音样本来培训人工智能，可能涉嫌版权侵权[8]。人工智能技术运用于产业的快速发展，加上现行法律规范与行业标准的滞后性，使得有声书版权归属问题成为制约产业发展的痛点。从版权主体、改编边界、声音克隆几个方面，都可能引发法律纠纷。版权问题直接关系到创作者和平台企业的利益，并且这种版权不明还会引发公众对 AI 产品的信任危机，从而阻碍作品的商业化运营。

### 3.4. 有声书智能化应用的具体优化策略

#### 3.4.1. 推动有声书内容精品化与差异化发展

内容始终是有声书发展的核心竞争资源。无论技术如何迭代发展，用户消费的根本在于内容。针对大量同质化内容泛滥的情况，平台可以适当提高内容准入标准，减少对套路化内容的批量生产。平台在提升内容生产效率的同时，要进一步优化推荐机制，避免单纯以流量逻辑作为内容标准，加大对原创内容和优质 IP 的扶持力度。随着有声书市场的扩大，用户的需求呈现出多元化、个性化、圈层化的发展趋势，越来越多的用户开始关注垂直细分领域内容。平台可以建立差异化创作激励机制，对于非热门的精品内容进行倾斜，设置专属推荐、创作激励等措施，鼓励创作者深耕内容创作。有声书平台可以利用人工智能的数据分析能力，挖掘用户需求，鼓励小众和特色内容创新生产。

#### 3.4.2. 技术升级和人机协同模式

技术优化是破解有声书智能应用缺陷的基础，人工智能语音合成技术不断进步，能够实现自然流畅的声音生成，但仍达不到真人演播的效果。未来，应该进一步加强情感计算、自然语言处理等技术的升级，提升人工智能对文本内容、人物关系和叙事情境的理解能力。有声书作为信息传播与艺术表达的作品，其价值不仅体现在语言传递方面，也体现在主播传递的情感能力方面。因此，我认为有声书行业的发展应该推动人机协同发展，人工智能完成标准化环节，专业主播再对其情绪等方面优化调整。同时，可以探索真人主播和人工智能联合训练的模式，通过高质量声音样本提升模型表现能力，增强人工智能声音的真实感和感染力。

#### 3.4.3. 完善版权管理与行业规范

未来应加快相关法律法规建设，明确人工智能生成有声书的版权主体和责任归属，规范声音数据采集、模型训练与商业化应用流程。同时，加强对主播声纹、声音数据以及数字声音资产的保护，建立人工智能声音授权和收益分配机制，保障创作者合法权益。此外，行业组织和平台企业也应加强自律，推动形成统一的技术标准与伦理规范，为人工智能技术在有声书领域的创新应用营造规范有序的发展环境。

## 4. 结语

人工智能技术的日益成熟，为有声书的制作与传播过程注入发展动能，通过探讨 AI 技术在生产端的降本增效、传播端的精准触达，虽然 AI 凭借其强大的算力与学习能力，让大规模、个性化、场景化的内容供给成为可能。但是不可否认的是内容同质化、声音情感缺失以及法律版权归属等问题，制约着有声书行业高质量发展。人工智能与有声书产业如何深度融合仍是音频行业需要重点关注的问题，推动内容的精品化和差异化建设，积极探索人机协同生产模式，技术持续向情感化、拟人化深度发展，并且行业

---

亟须建立健全相关法律法规与伦理准则，为技术创新划定底线。AI 时代下的听觉文化，要实现文化和商业双重价值，为用户提供更加优质的体验。

## 参考文献

- [1] 李武, 谢泽杭, 杨飞. AI 有声书: 价值优势与未来进路[J]. 科技与出版, 2023(6): 41-47.
- [2] 新京报书评周刊. 2025 图书行业, 学着与 AI 共舞[EB/OL].  
<https://baijiahao.baidu.com/s?id=1852454620521274561&wfr=spider&for=pc>, 2025-12-25.
- [3] 半岛都市报. AIGC 赋能“声工厂”助推有声阅读持续繁荣[EB/OL].  
<https://yule.bandao.cn/a/1713861758698255.html>, 2024-04-23.
- [4] 中国商报. 喜马拉雅举办 2024-2025 戏精大汇: 持续扶持激励创作者[EB/OL].  
<https://www.zgswcn.com/news.html?aid=228328>, 2024-11-06.
- [5] 余苗, 赵文聪. 人工智能发展与有声阅读平台创新探析[J]. 传媒, 2022(6): 54-56.
- [6] 郭小坤. 人工智能技术助力有声阅读服务精准化研究——以喜马拉雅 APP 为例[J]. 现代视听, 2024(7): 42-45.
- [7] 赵远杰, 宋霄钰, 秦振宇. 从文本到声音: AI 语音合成技术对有声书产业的影响[J]. 传播与版权, 2025(4): 37-40.
- [8] 程辉, 刘松强, 李武. 作为新质生产力的技术创新驱动: AI 2.0 时代有声出版产业链建设现状、挑战及对策[J]. 出版发行研究, 2024(6): 30+46-53.