

2026年

音乐大模型行业研究：从“基础生成”到“成品级输出”，AI音乐开启商业化新浪潮

AI Music Large Model Industry

音楽大規模モデル業界

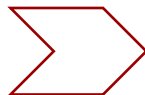
概览标签：音乐大模型、多模态融合生成、Suno、Mureka

报告提供的任何内容（包括但不限于数据、文字、图表、图像等）均系头豹研究院独有的高度机密性文件（在报告中另行标明出处者除外）。未经头豹研究院事先书面许可，任何人不得以任何方式擅自复制、再造、传播、出版、引用、改编、汇编本报告内容，若有违反上述约定的行为发生，头豹研究院保留采取法律措施，追究相关人员责任的权利。头豹研究院开展的所有商业活动均使用“头豹研究院”或“头豹”的商号、商标，头豹研究院无任何前述名称之外的其他分支机构，也未授权或聘用其他任何第三方代表头豹研究院开展商业活动。

调研总结：我们主要回答了哪些问题？

01：音乐大模型核心架构及发展现状如何？

- 音乐大模型核心技术架构采用“三层架构”模式，即数据层、模型层、应用层。2025年以来，音乐大模型技术主要实现以下突破：一是模型稳定性显著增强，AI音乐作品达商用发布标准；二是上线参考生成能力，升级为可复用、可迭代的专业创作系统；三是打造出即时响应的演奏级应用新模式。



03：音乐大模型核心应用场景有哪些？市场规模有多大？

- 目前，C端用户主要借助模型生成短视频BGM、产出参考旋律、制作广告Demo、创作活动主题歌/知识记忆歌等。在B端，音乐大模型主要应用于：（1）为AI音乐、AI MV等AI创作平台提供各类功能；（2）为影视制作公司快速生成符合剧情氛围的音乐；（3）为游戏开发企业生成背景音乐或特效音乐。
- 2025年，全球音乐大模型市场规模约30.8亿元，预计到2030年，随着B、C端持续渗透，该市场有望增至171.4亿元，2025-2030年CAGR达40.9%。



02：主流音乐大模型有哪些？各模型有什么优势？

- 目前，主流音乐大模型包括：以Suno、Udio、Stability Audio为代表的海外厂商；以昆仑万维、MiniMax、趣玩科技等为代表的中国厂商。从生成能力看，Suno和Stability Audio功能相对全面；Udio音质质感顶尖；Mureka在分轨导出方面表现突出；天谱乐风格覆盖广、MV生成链路较成熟；MiniMax Music对多元歌曲结构变体支持度高。从商业化现状看，针对个人用户，除MiniMax Music外，其它模型均采用订阅制，价格集中在\$10-30/月，针对企业/专业开发者，国产模型提供标准化API服务，海外仅有Stability Audio提供定制模型服务，Suno等尚未推出官方API接口。



目录

CONTENTS

◆ 音乐大模型行业发展概述	-----	4
• 音乐大模型核心架构		
• 音乐大模型创作模式		
• 音乐大模型发展历程		
• 音乐大模型发展现状		
◆ 海内外主流音乐大模型及其对比	-----	10
• 音乐大模型市场参与者		
• 海外主流音乐大模型		
• 中国主流音乐大模型		
• 海内外主流音乐大模型对比		
◆ 音乐大模型市场应用洞察	-----	23
• 音乐大模型核心应用场景		
• 音乐大模型市场规模		
• 音乐大模型发展趋势		
◆ 联系我们	-----	29
◆ 方法论与法律声明	-----	30

Chapter 1

音乐大模型行业 发展概述

音乐大模型核心架构

音乐大模型核心技术架构采用“三层架构”模式，即数据层（包括数据采集/预处理/增强）、模型层（主流模型采用“多模态融合Transformer”架构）、应用层（包括音乐理解/生成/编辑/交互类应用）

音乐大模型核心架构

数据层

01 数据采集

主要包括：①**音频数据**：涵盖不同曲风/语种/年代的歌曲/纯音乐/音效等，是模型**核心训练数据**；②**文本数据**：歌词/音乐描述/曲风标签/情感标签等，用于实现**音频与文本对齐**；③**乐谱数据**：用于**模型理解音乐旋律/和声结构**；④**跨模态辅助数据**：图像/视频等，用于**多模态音乐生成与个性推荐**。

02 数据预处理

核心任务：①**音频降噪**：通过信号处理技术去除音频中环境/设备噪声；②**格式标准化**：将不同格式音频数据转换为统一格式，统一采样率和声道数；③**特征提取**：将音频数据转换为模型可识别的特征表示；④**数据清洗**：剔除低质量、重复、违规数据，确保数据集纯度与合规性。

03 数据增强

为解决数据集规模不足、多样性不够问题，采用数据增强技术，**对原始数据进行变换，生成更多训练数据，提升模型的泛化能力**。常用方法包括：音频变速、音调调整、时域翻转、频谱增强等，**部分模型采用“跨模态数据增强”技术**，通过文本生成音频、音频生成乐谱等方式，进一步扩大数据集规模。

模型层

- Transformer架构凭借强大的长序列建模能力与并行计算效率，成为音乐大模型核心架构。当前，主流音乐大模型**采用“多模态融合Transformer架构”**，实现音频、文本、乐谱等不同模态对齐与融合，核心原理为：**通过不同模态编码器，将不同模态的数据转换为统一的特征向量，再通过跨模态注意力机制，实现不同模态特征的交互融合，使模型能够理解不同模态之间的关联。**

关键技术1：自监督学习与预训练策略

自监督学习是音乐大模型训练核心技术，核心是利用**数据自身内在结构与规律，构建监督信号，让模型自主学习通用音乐特征**。常见的预训练策略包括**单模态预训练与多模态预训练**，常见自监督学习任务包括**音频掩码重建、文本音频匹配、旋律预测等**。

关键技术2：思维链推理与可控性生成

思维链技术的引入提升音乐生成的**结构完整性与可控性**，核心原理为：**通过“生成——批判——优化”循环机制，逐步完善音乐作品；可控性生成核心目标是让用户精准控制音乐生成结果，满足个性化需求**，包括**文本驱动控制、参数调节控制、交互编辑控制等**。

关键技术3：音频合成与音质优化

音频合成负责将模型生成的音乐特征转换为可聆听的音频信号，其性能直接决定音乐生成的音质与自然度，当前主要包括**基于波形合成与基于频谱合成**；**音质优化是提升用户体验的保障**，主要包括**提升采样率与比特深度、优化发音准确率、降低生成延迟等**。

应用层

音乐理解类任务

核心是让模型“听懂”音乐，包括**曲风识别、情感分析、歌词识别、乐谱识别、音频分离等**，主要用于**音乐检索、版权监测、音频编辑等场景**。

音乐生成类任务

核心是让模型“创作”音乐，包括**旋律生成、编曲生成、歌词生成、完整歌曲生成、多模态音乐生成**，主要用于**专业创作辅助、广告/游戏配乐等场景**。

音乐编辑类任务

核心是让模型“优化”音乐，包括**音乐旋律修改、和声调整、混音优化、风格转换等**，主要用于**专业音乐创作、大众音乐编辑等场景**。

音乐交互类任务

核心是实现人与音乐的**智能化交互**，包括**个性化推荐、实时伴奏、人声合成、音乐问答等**，主要用于**音乐APP、智能音箱、在线K歌等场景**。

音乐大模型创作模式

音乐大模型的创作模式主要分为文本驱动（适合零基础创作）、歌词驱动（适合旋律创作）、图/视频驱动（适合影视、游戏、广告等快速配乐）及音乐驱动（适合风格迁移、翻唱、改编）四类

音乐大模型创作模式

模式一：文本驱动 01

路径详情

- 用户输入一段文本描述（如主题、风格、情绪、乐器、场景等），模型根据提示词生成完整音乐作品。如Suno。

路径优势

- **当前最主流、最基础的创作路径**，降低创作门槛、无需专业知识即可快速生成音乐；支撑个性化定制，满足不同风格和情感需求；可快速迭代优化。

模式二：歌词驱动 02

路径详情

- 用户提供歌词文本，模型根据歌词内容生成匹配的旋律、节奏和编曲，部分模型支撑歌词结构标记，确保音乐与歌词段落对应。如Song Generation。

路径优势

- 同时兼顾歌词与旋律的协调性，能生成更完整的作品；支撑风格和情感表达，增强音乐表现力；**适合旋律创作、说唱等音乐形式**。

模式三：图/视频驱动 03

路径详情

- 支持图片、视频等多模态作为输入，模型自动分析输入内容的色彩、氛围、节奏等，生成匹配的音乐。如天谱乐。

路径优势

- 实现模态创作，拓展创作灵感来源；能更精准捕捉用户意图，生成与输入内容情感和风格契合的音乐；**适用于影视、游戏、广告等场景的快速配乐需求**。

模式四：音乐驱动 04

路径详情

- 用户提供参考音乐，模型通过学习生成具备相似风格的歌曲，或基于自动解析，根据用户指令进行续写、扩写、风格转换。如Mureka。

路径优势

- 帮助用户快速探索特定风格的音乐创作；可生成创新性的变奏作品，避免内容同质化；**适用于音乐风格研究、翻唱、改编、音乐扩展等场景**。

音乐大模型发展历程

2016年以来，音乐大模型的发展经历“单点功能AI化→多任务融合→多模态协同创作”三阶段。目前，多模态交互融合正重塑音乐创作边界，具备完整歌曲创作、实时交互的新一代音乐大模型成为主流

AI音乐发展历程

2016-2020年，单点功能AI化

- 代表：2016年，谷歌推出基于机器学习的艺术生成研究平台——Magenta（开源），旨在通过机器学习模型自动生成音乐、图像等创意内容，例如旋律/和弦；2018年，AI作曲工具AIVA（专注利用AI为电影、广告等制作配乐）基于使用强化学习技术的深度学习算法，发布首张中国音乐专辑——《艾媧》；2019年，Amper Music推出供企业内容创作者使用的端到端AI作曲平台——Amper ScoreTM，客户可将Amper的API与企业自身创作平台、发行平台及其它音乐创作软件结合（期间QQ音乐与Amper达成API合作）。
- 技术：基于CNN的音频分离技术实现人声与伴奏分离；基于机器学习的曲风识别、情感分析技术，可对风格与情感进行初步判断；基于RNN等的旋律生成技术，能识别简单单声部旋律。
- 特点：泛化能力弱，无法跨任务协同工作，生成的音乐内容缺乏创新性与完整性。

2020-2024年，多任务融合

- 代表：2020年，OpenAI发布AI音乐生成模型——Jukebox，它采用Transformer架构和VQ-VAE技术，能够生成多风格、多声部的音乐片段，奠定Transformer文生音乐的基本架构；2023年1月，谷歌发布MusicLM大模型，可直接将文字、图像自动生成曲风多样的音乐，同年6月，谷歌发布基于Transformer架构的AI音乐生成模型，可以根据文本提升生成新的音乐片段，并选择性与现有旋律对齐，其由单级变压器语言模型和高效令牌交错模式组成，可生成高质量音乐样本，消除层级或采样等多个模型级联的需求。
- 技术：预训练技术的应用解决传统模型标注数据依赖问题，提升泛化能力；实现旋律生成、编曲、歌词匹配等多任务协同优化，生成的音乐由单声部旋律升级为多声部、多乐器的完整片段。
- 特点：核心为“预训练+微调”，依托大规模未标注数据进行训练，实现多任务协同处理。

2025年-至今，多模态协同创作

- 代表：2025年，Suno V4.5突破音轨分离局限，允许创作者将AI生成的主要音轨单独导入数字音频工作站，进行混音、均衡、母带处理等专业操作，首次支持最多十二轨分轨音频导出；Udio允许用户通过指定主题、流派等，从简单文本提示中创作音乐；Mureka V9以音乐思维链技术为核心，将大模型推理范式深度融入音乐生成逻辑，将音乐创作拆解为“理解意图——结构建模——分环节生成——精准调整”的可控流程；天谱乐多模态音乐生成大模型，支持文本、图片和视频等多种输入方式生成音乐。
- 技术：思维链技术与多模态交互融合重塑音乐创作边界；多模态预训练架构实现不同模态数据对齐与融合，实现“跨模态理解与生成”。
- 特点：具备完整歌曲创作、个性化编曲、实时交互等核心能力，广泛应用于专业创作、大众娱乐、产业服务等多个领域。

音乐大模型发展现状（1/2）

2025年以来，音乐大模型迎来以下突破：一是模型稳定性显著增强，AI音乐作品达商用发布标准；二是上线参考生成能力，升级为可复用、可迭代的专业创作系统；三是打造出即时响应的演奏级应用新模式

音乐大模型技术突破

1 创作从“基础生成”→“可发布标准”→“一键MV生成分发”

□ 基础生成→可发布标准（生成稳定性提升）：“生成完整歌曲”已成为音乐大模型基础能力，AI音乐创作的价值重心开始转向“审美目标设定→多版本筛选→结构与声线一致性控制→后期制作整合”的系统性能力。例如：Suno灵感激发功能，可通过分析用户音乐列表，生成符合用户审美的音乐。

①音质质感显著改善（录音棚级音质），拼接感、噪声与不稳定片段减少，支持自动调整混音等，使得输出内容更易进入后期制作与发布流程

②作品结构完整性明显提升，支持多种歌曲结构变体和段落式风格设定，主歌——副歌——桥段的段落组织更自然，段落对比与情绪推进更具可预期性

③提示词遵循率持续提升，对情绪、配器、跨流派等指令的响应更稳定，实现更强的提示词遵循能力与更丰富的表达能力

端到端生成模型愈发稳定成熟，使得“音乐生成”转为常规环节，为后续局部重绘、改编扩展等创作环节的模式变革奠定技术基础。

□ 可发布标准→一键MV生成分发：SunoMV原生支持Suno链接转为AI MV，整个流程从粘贴链接到导出成品，通常不超5min；Tune智能体支持“智能解读需求→用户选择方案→MV生成→一键分享”全流程。

2 从“文本驱动”到“参照式创作”，一次性输出转向可复用迭代

□ 2025年以来，主流音乐模型引入参考生成功能，以音频/音乐作为提示或约束，有效弥补自然语言描述音乐的局限性，实现文本驱动到音频参照两种创作模式深度融合。通过音频参照、时间线编辑、版本管理等，创作者可精准定义音乐风格与整体结构，推动AI音创作从单次灵感输出，转向工程化、可迭代的生产模式。例如：Suno推出Personas、Covers、Extend等工具，强化全流程可控创作链路；Udio新增Styles、Sessions、Voice功能，搭建起“参照创作→精细化编辑→多版本迭代”完整工作流。可以看出，AI音乐生成已突破单一结果输出形态，演变为一套可复用、可持续迭代的专业创作系统。

3 “实时交互”能力逐渐形成相对稳定的产品化路径

□ 2025年4月，谷歌DeepMind推出LyriaRealTime，基于Lyria2音乐生成模型所打造的创新功能，它支持用户实时控制音乐生成，如随时调整音乐风格、节奏和情感，且在响应用户控制生成音乐流的同时，实现“即兴演奏”，推动AI音乐从传统“离线渲染”模式升级为“即时响应的演奏级场景”。

音乐大模型发展现状 (2/2)

技术持续迭代的同时，音乐大模型行业也逐步迈入版权授权新阶段。Suno、Udio等均通过与版权方深度合作，实现训练数据合规、商用权限清晰、内容版权权责分明

音乐大模型版权问题

2024年，全球三大唱片公司环球音乐、索尼音乐、华纳音乐起诉Suno和Udio，控诉其利用唱片公司内容来训练生成音乐的AI系统，大规模侵犯版权。

音乐大模型版权问题

音乐大模型训练数据是否合法获取？

大模型生成的音乐是否构成“衍生作品”？是否侵权？

大模型生成的音乐若属于原创，是否需要授权？

用户行为与平台责任边界在哪？

训练数据合规化

- ❑ 与版权拥有方合作推出全新授权平台：Suno与华纳音乐达成合作协议，推出全新授权AI模型，使用经过正式授权的音乐数据进行训练，所有用户统一迁移至授权体系；Udio与环球音乐达成合作推出全新授权平台，环球音乐将提供部分曲库数据用于模型训练，并共享未来商业化收益。

商业使用权明确划分或在受控环境中商用

- ❑ 商业化权限向付费用户倾斜：Suno明确划分免费和付费账户权限边界，通过免费计划生成的作品无法下载，仅支持在线播放、分享，无法变现；付费用户在生成内容时自动获得商业使用许可，可用于流媒体发行等商业场景，**Suno不从中分成收益。**
- ❑ 仅可在特定平台播放、分享音乐：Udio免费版即拥有商业使用权，但关闭下载权限，仅支持一键分享至X、Facebook等平台。

内容版权归属明确

- ❑ 对于Suno全新平台，付费用户只获得商业使用权，但通常不被视为歌曲所有者，Suno对输出内容本身负有最终责任。

Chapter 2

海内外主流音乐
大模型及其对比

音乐大模型市场参与者

音乐大模型市场参与者主要包括：以Suno、Amper、OpenAI、Google、Udio、Stability Audio为代表的海外厂商；以昆仑万维、字节、MiniMax、趣玩科技、阶跃星辰、网易云音乐等为代表的中国厂商

音乐大模型市场参与者图谱



海外主流音乐大模型——Suno

Suno大模型发展历程

2024年3月，V3：端到端生成

- 端到端生成完整歌曲（最长2min，可续写，生成内容包含伴奏+主唱+和声+歌词+歌曲结构）
- 支持“一句话生成”、“给定歌词→按歌词填曲+演唱”，可选风格、语言、节奏、情绪等（自定义模式）
- 生成质量：频响/动态范围接近人类混音、人声音色自然清晰、歌曲结构清晰

2024年6月，V3.5：时长扩容

- 生成歌曲时长最长延至4min，提供最多2min的歌曲扩展；同时改进歌曲结构（包括引入、发展、高潮和结尾等，使得生成音乐更流畅自然，情感表达更丰富细腻）和声音流畅度等；增强的元标签支持与和弦识别，能够更精准识别和弦；提供wav格式音频下载权限（具备更高音质）

2024年11月，V4：音质升级

- 音质全面提升，允许用户直接将旧版本作品升级；ReMi AI歌词助手协作用户创作更清晰的歌词；新增封面设计功能；Covers功能支持一键生成歌曲翻唱内容；人声一致性技术助力打造独特、稳定的声音特征；Personas功能支持曲风复用；允许用户上传照片/视频（多模态生成），生成一段30s的音乐

2025年5月，V4.5：时长扩容+编排

- 支持长达8min的单曲创作；扩展流派和流派混搭；人声具有增强的音域和情感深度；音质具有更平衡饱满的混音，减少音频衰减和闪烁效果；对情绪、氛围等的描述理解更准确，并上线提示增强助手功能；引入编排工具（如替换、扩展、重制、翻唱等），支持段落式风格设定，创作具备编排感与层次感

2025年7月，V4.5+：风格扩容

- 人声替换功能：用户上传任意伴奏音频或使用内置乐器伴奏，通过输入歌词即可生成完整歌曲，同时允许用户将自己的歌声/哼唱转化为完整音乐作品，极大降低创作门槛
- 灵感激发功能：通过分析播放列表的音乐风格、情绪、元素等，生成符合用户审美的音乐作品
- 更丰富多元的音乐风格

2025年10月，V5：音质升级

- 录音棚级音质：根据歌词内容自动注入情感起伏，支持多声部合成与和声编排；音频动态响度与声场空间感显著提升，音乐具备立体层次；用户精确指定节奏、分段结构、乐器编排，告别“随机生成”不可控性
- 母带翻唱处理：用户导入现有作品，由AI重新编曲并控制风格偏离程度
- 商业版权开放，创作者可直接变现

2026年4月，V5.5：个性化定制

- 声音克隆Voices：捕捉用户的声音，并在生成的歌曲中使用
- 定制模型Custom models：基于用户拥有版权的歌曲，构建专属模型，最多可支持创建3个私有模型变体，用于不同创作方向
- 精准定义用户音乐偏好：捕捉个人音乐口味，自动优化创作，减少每次从头重建相同上下文的时间

- 从V3到V5.5，Suno逐步突破时长限制（从2min到8min），持续升级音质（已达录音棚级别）、优化歌曲结构与入声表现，再到推出歌词助手、翻唱、多模态生成、编排工具等创作辅助功能，后期开放商业版权，向入声替换、声音克隆、专属定制模型等个性化、商业化方向深化，整体呈现从基础生成能力到专业制作工具、再到高度个性化定制的发展脉络。

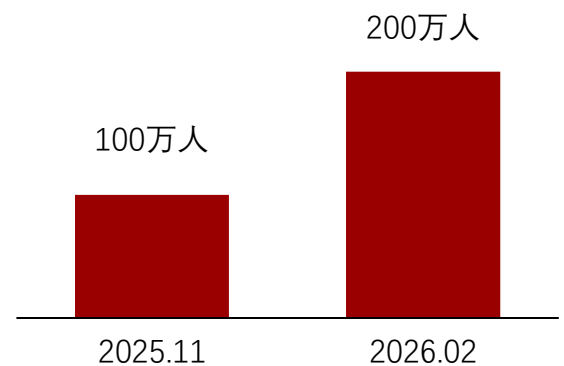
海外主流音乐大模型——Suno（接上页）

Suno正加速从专业AI音乐生成工具，朝着个性化应用与商业化落地纵深发展。平台依托订阅制盈利模式快速崛起，目前付费用户突破200万，年化经常性收入约3亿美元，付费用户周留存率高达78%

Suno商业化现状

版本	定价	权益
免费版	\$0/月	使用V4.5+及之前模型；每日赠送50积分（一首音乐消耗5积分）但不累积；不可商用；仅可使用标准功能；最长支持上传8min音频；生成队列为共享模式，最多同时处理2个任务；不能额外购买积分； 适合初次体验用户
Pro版	\$10/月（年付享8折）	使用V5.5模型及其功能；包含2,500积分，每月刷新；新歌的商业使用权；标准功能+专业版功能（角色+高级编辑）；最长支持上传30min音频；优先生成，最多可同时处理10个任务；新功能优先体验；可额外购买积分； 适合小型工作室和个人创作者
Premier版	\$30/月（年付享8折）	使用Studio（用户创意空间，集成多轨编辑与AI分轨生成）；使用V5.5模型及其功能；包含10,000积分，每月刷新；新歌的商业使用权；标准功能+专业版功能；优先生成，最多可同时处理10个任务；新功能优先体验；可额外购买积分； 适合专业音乐制作机构及高频次商用场景

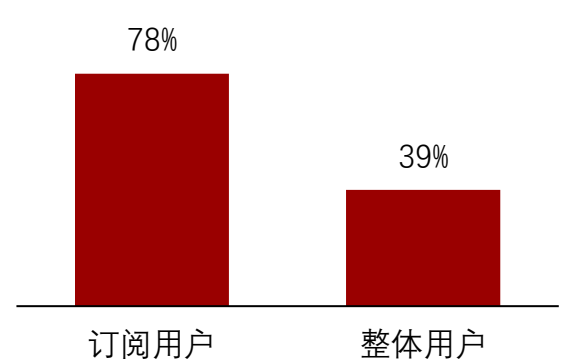
- 商业化方面，Suno采用订阅制收费，提供免费版（适合新用户体验）、Pro版（适合小型工作室和个人创作者）和Premier版（适合专业制作机构及高频次商用场景）。截至2026年2月，Suno付费用户已由2025年11月的100万增至200万，实现翻倍，年化经常性收入已达3亿美元。从用户粘性看，平台整体用户周留存率为39%，订阅用户周留存率达78%，约25%的订阅用户在30天后仍保持活跃。



Suno付费用户数从2025年的约100万增至2026年2月的200万



Suno年度经常性收入达3亿美元，2026.02



Suno订阅用户和整体用户周留存率对比，2026.02



Suno约25%的订阅用户在30天后仍保持活跃，2026.02

海外主流音乐大模型——Udio

Udio主打高保真音质与精细化编辑能力，适合对音质和细节有较高要求的专业用户。平台采用订阅制为主的盈利模式，定价对标Suno；受版权合规约束，目前用户仅可在受控环境定制、播放、分享音乐

Udio发展现状

□ Udio是一款用于音乐创作和共享的AI工具，主打高保真音质与精细化编辑能力，适合对音质和细节有较高要求的专业用户。2024年4月，Udio发布V1版本，支持多风格多语言创作、音轨时长扩展和定制（默认生成时长32s，但允许用户基于母样扩展音乐长度）以及音乐剪辑。2024年5月，Udio推出udio-130模型，可直接生成130s音乐片段。2024年7月，Udio V1.5发布，可直接生成48kHz立体声轨道，音乐清晰度、连贯性、乐器分离度、瞬态响应得到极大提升，同时新增全球语言生成、分轨下载（将混音曲目拆分为4个独立音轨）、音频到音频（用户上传曲目进行混音创作）、调性控制等功能，同时支持歌词视频（自动生成的歌曲配备AI封面、动态高亮字幕，然后合成视频）在社交平台分享。2025年后，Udio平台陆续新增Styles（允许用户用一两个音频片段/现有的Udio歌曲作为风格参考生成新歌）、Sessions（精准编辑和扩展任何曲目）、Voices（使用Voices库中的声音或用户之前在Udio创作歌曲的声音作为声音来源创作，暂不可上传用户自己的声音作为来源）等功能。

Udio新版本	定价	权益
免费版	\$0/月	每天10积分（24h更新），每月100积分（不累积）；最多可同时生成4首歌曲；创作32s歌曲；创作130s歌曲（每天最多3次）；扩展歌曲；混音歌曲；使用音乐参考；生成AI封面；含商业版权
标准版	\$10/月 (年付享8折)	每月2,400积分（不累积）；最多同时生成6首歌曲；创作32s/130s歌曲（无限制）；用Udio的声音库创作歌曲；支持用户上传自己的音频文件和使用自己的声音；扩展歌曲；混音歌曲；使用音乐参考；使用高级艺术家风格；混合风格；简化样式；编辑音乐；生成AI封面；上传封面；调整歌曲访问权限；含商业版权
尊享版	\$30/月 (年付享8折)	每月6,000积分（不累积），最多同时生成10首歌曲，其它权益通标准版

□ 2024年6月，Udio和Suno遭环球音乐、索尼音乐、华纳音乐三大唱片公司联合起诉，被指控其AI系统在未经授权的情况下，使用受版权保护的音乐作品进行训练，导致生成的歌曲与原创作品高度相似。2025年10月，环球音乐与Udio达成和解，并计划于2026年联合推出一款新平台，该平台技术训练数据全部来自获得授权的音乐，确保创作过程中合规。与此同时，平台将提供订阅服务，通过创新功能提升用户体验，用户可在一个受控环境（无法自由下载WAV/MP3，只能在平台内播放和分享；每次创作产生的收益将汇入版税池，分配给授权艺术家；无法基于现有艺术家训练自定义声音，但可混合风格）中定制、播放、分享音乐，这种模式满足用户个性化音乐需求的同时，又保护音乐艺术家的权益。

海外主流音乐大模型——Stability Audio

Stability Audio通过“完全授权数据+企业法律保障”建立差异化定位，服务覆盖设备端生成、完整音乐创作及企业级部署不同场景。平台面对个人用户提供订阅服务，面对中大型企业客户，提供定制模型服务

Stability Audio发展现状

2023年9月，Stability Audio 1.0推出，基于音频素材平台AudioSparx提供的80万首歌曲、音效及乐器采样完成训练；2024年4月，**Stability Audio 2.0推出**，可商用、支持**44.1kHz无损音质、最长3min完整曲目创作**；2024年6月，**Stability.ai开源文生音频模型Stability Audio Open**，用户通过文本就能生成最长**47s**、不同风格的**44.1kHz**音效，同时支持用户根据自定义音频数据对模型进行微调。

2025年9月，Stability Audio 2.5发布（公司首个面向企业级应用场景开发的型号），新模型在速度和输出质量上明显提升，**生成时长可达3min**，可生成**多声部**作品，同时提升提示词遵循率，能**更有效响应情绪描述词和跨流派语言**。除文生音频和音频到音频外，新模型还支持**音频修复**。此外，还新增**定制音频功能**，只需微调模型即可将品牌音频嵌入定制生成的 workflow 中。

2026年5月，Stability.AI推出Stability Audio 3.0（新增局部编辑、因果续写等功能），**该系列共有4款模型**，覆盖设备端生成、完整音乐创作及企业级部署不同场景，其中三款开源，面向开发者、音乐平台以及创作者社区开放下载与二次开发，**所有模型均基于“完全授权”数据训练，开放模型允许商业化使用**。Small SFX与Small版均可以0.44s的推理速度生成最长2min音频，前者主要针对音效生成，面向手机、电脑等设备；后者则支持完整音乐创作，能实现设备端完整音乐生成；Medium版最长可生成约6min20s的完整音乐；**Large版暂不开放权重**，主要面向音乐平台与高并发商业应用，仅可通过Stability.AI API、合作伙伴fal.ai或企业授权形式部署。

服务	定价	权益
免费版	\$0	每月使用V2.5生成10次，V3.0生成6次；最长6min；每月上传30s音频，累计6min； 无商业版权（不可以商业音乐发行、在社媒等平台分享）
Pro版	\$11.99/月	每月使用V2.5生成250次，V3.0生成166次；最长6min；每月上传6min音频，累计30min； 有创作者版权（可发行、可用于商业上面、可分享、可用于月活低于10万的商业产品，不可用于电影/电视/广告/应用与游戏）
Studio版	\$29.99/月	每月使用V2.5生成675次，V3.0生成450次；最长6min；每月上传6min音频，累计60min； 有创作者版权
Max版	\$89.99/月	每月使用V2.5生成2,250次，V3.0生成1,500次；最长6min；每月上传6min音频，累计90min； 有创作者版权
企业版	——	针对年收入超100万美元的客户，提供定制训练模型服务（企业法律保障），拥有企业版权（可用于月活超10万的商业商品、电影/电商/广告、应用与游戏）

中国主流音乐大模型——Mureka (1/2)

技术方面，Mureka音乐大模型以MusiCoT音乐推理技术为核心，逐步强化其编曲能力、音乐结构连贯性、音质质感、表达意图理解与创作可控性，同时实现响应速度、系统稳定性的显著提升

昆仑万维音乐大模型发展路径

2024.04, SkyMusic

- 乐器、人声、旋律、音量、音符一体化端到端音乐生成
- 能根据用户输入的歌词风格灵活生成对应音乐风格
- 人声合成技术强，可高度模拟人声，且具备参考音乐生成与方言歌曲生成能力
- 允许用户细致控制歌曲情绪节奏

2025.03, Mureka V6

- 自研ICL技术，声场更开阔，强化质感和混音设计；训练含合成及标注数据、链式思维数据库等
- 支持多语言生成多风格；支持风格控制、旋律录制（根据用户录制旋律生成音乐伴奏）；支持纯音乐生成；支持文本提示、音色音频参考等生成个性化音乐

2025.07-08, Mureka V7-7.5

- V7整体音质较V6提升一倍，良品率明显提升，音乐创新性更强；V7.5优化自动语音识别技术，通过智能划分乐句、确定自然换气与停顿位置，显著提示生成人声的段落清晰度与结构真实感
- 支持10种语言；新增歌曲参考、自定义歌手、音频编辑等功能

2026.01, Mureka V8

- 在音乐性、编曲完成度、人声表达、质感等多个关键维度同步提升，能生成具备清晰主副歌对比、时长达标的音乐成品
- 在人声音色、性别匹配、唱法选择上更贴合提示词，演唱更自信，整体表现更具“主唱存在感”；编曲层次更丰富；空间层次更立体

最新功能

- **单轨生成**：支持为已有音频生成配套单轨，支持人声、伴奏、制定乐器三种模式
- **歌曲续写**：新增向前续写，与原有向后续写共同覆盖两个方向
- **局部编辑**：选取任意片段重写歌词、调整风格或替换乐器，前后衔接自动处理
- **分轨**：底层模型全面升级，最多支持导出12条独立音轨，新增MIDI导出
- **Remix（内测）**：在保留原曲旋律与结构的基础上，自由改变风格或重写歌词

2024.08, Mureka

- 昆仑万维正式推出AI音乐商用创作平台Mureka

2025年3月, Mureka O1

- 首款音乐推理模型：引入业界首创MusiCoT技术，利用思维链方法在生成音乐前预生成整体音乐结构，大幅提升音乐结构连贯性和乐器编排精准度
- 支持多语言多风格；支持音色克隆，生成个性化专属作品
- 开放API服务与模型微调

2025.11, Mureka V7.6及O2

- 在音乐性、编曲能力、音质质感和Prompt贴合度等多个维度实现显著提升
- V7.6的响应速度、系统稳定性和推理效率大幅增强，在MusiCoT分层编曲建模基础上，强化对段落关系、乐器互动与情绪走向的建模能力，情绪推进与结构变化更自然；重构声场建模机制，通过更精细的频率动态处理与声部定位，使音质更自然清晰

2026.03, Mureka V9

- 在MusiCoT技术体系持续迭代上，进一步强化模型对音乐结构、歌词段落、表达意图、声音呈现（混音音质）之间关系的协同建模；与V8相比，在可控性、音质、效率、表达性等方面明显提升
- 最长支持5.5min，可单独对某一段落进行二次生成；支持将生成的歌曲分离为独立音轨，可单独调整某一音轨的音量、音色、效果；支持多版本并行生成

中国主流音乐大模型——Mureka (2/2)

商业化层面，Mureka采取ToB和ToC并行策略：企业端主打标准化API服务，采用充值付费制，各项功能单独定价，目前客户数超8,000家；C端推行会员订阅制，支持月付、年付两种方式，年付性价比更高

Mureka商业模式

会员服务	定价	权益
免费版本	¥ 0	注册时免费赠送50金币：V7.6或V7.5-all为10金币/首、V8或V9为12金币/首
基础会员	¥ 88/月 (¥ 0.5/首)	最多生成180首歌曲；使用V7.6、O2、V8、V9等高级模型；将歌曲混音成任意风格；为视频和图片定制配乐；下载MP3及视频； 新生成歌曲可商用 ；根据参考音乐创作歌曲；优先队列，最多同时生成10首歌曲；音频扩展； 编辑选区实现歌词/旋律精调
	¥ 398/年 (¥ 0.33/首)	最多可生成1,200首歌曲；其它权益同上
尊享会员	¥ 168/月 (¥ 0.4/首)	最多可生成440首歌曲；使用V7.6、O2、V8、V9等高级模型；将歌曲分离为最多12音轨；将歌曲混音成任意风格；使用Mureka Studio；为视频和图片定制配乐；下载WAV、MP3、伴奏和分轨音频、视频；新生成歌曲可商用；优先队列，最多同时生成10首歌曲；音频扩展；编辑选区实现歌词/旋律精调；任何声音可成为用户AI音乐的歌手
	¥ 598/年 (¥ 0.25/首)	最多可生成2,400首歌曲；其它权益同上

API服务内容：【1】音乐生成：歌曲/背景音乐/配乐/歌词/单轨生成、音色克隆、分轨音频、歌曲续写、局部编辑、音乐理解（每项功能单价不同）；【2】模型微调（自定义模型）：用户只需提供200首具有一致特征的曲目，API即可为其微调专属模型，此后生成所有歌曲均遵循该风格，适合拥有大型歌曲数据集、并需要批量生成相似风格音乐的企业；【3】内容服务：对于音乐内容有需求，但不想投入开发资源接入API的客户提供成品音乐内容。

Studio API定价	权益
入门版（¥ 200）	12个月有效；1个并发；每个账户仅限购买一次
基础版（¥ 10,000）	12个月有效；5个并发
标准版（¥ 30,000）	12个月有效；15个并发
商业版（¥ 50,000）	12个月有效；25个并发

□ 在商业化方面，针对个人用户，Mureka采用会员订阅制，分为基础会员和尊享会员，月费分别为¥ 88和¥ 168，年费分别为¥ 398和¥ 598。自Mureka O1和V6发布，截至2025年11月，平台新增注册用户近700万，用户遍布全球超100个国家/地区。针对企业，Mureka开放平台提供API服务，截至2026年1月，已服务全球超8,000家客户，是全球率先推出官方AI音乐API开放服务的平台服务商。未来，Mureka坚持ToB和ToC并行策略，ToC的APP可能采用类似短剧的免费广告模式，以吸引海量用户，计划在3-5年内，海外月活达3-5亿。

中国主流音乐大模型——天谱乐（1/2）

作为多模态AI音乐生成模型，天谱乐通过分层Codec表达体系，完成从高层语义到细粒声学的结构化生成。同时推出Tunee智能体，以对话式实现“需求理解→音乐视频MV创作”全流程，大幅降低创作门槛

天谱乐发展现状

2024.05: Tempolor 1.0

基于语言模型与离散音乐音频语义表征进行建模，可生成**48kHz、双声道**音频。该版本重点解决从**文本描述到音乐语义结构建模问题**，能较稳定生成具有明确旋律走向、节奏框架与段落组织的音乐内容。



2024.09: Tempolor 2.0

该版本基于掩码式自回归生成模型与连续音乐音频声学特征，支持生成**44.1kHz、双声道**音乐音频，其在1.0的基础上，进一步强化歌曲级别的结构建模能力，**更注重曲式组织、段落衔接与前后呼应**，使生成结果在旋律发展、情绪递进和主题呈现上更自然真实。



2025.01: Tempolor 3.5

基于扩散模型与连续引用音频表征，支持生成**44.1kHz、双声道**音乐。该版本从偏重语义骨架建模的生成范式，进一步演进至以**连续声学表征为核心的高保真音乐生成范式**，显著提升声学纹理、空间层次、动态表现及整体听感的自然度和完成度。同时，该版本通过引入ControlNet等可控生成能力，进一步拓展模型在**旋律控制、歌词编辑、局部重绘**等场景下的可编辑性和可控制性，为更精细化的音乐创作与交互式编辑提供技术基础。



2026.03: Tempolor 4.6

对bottleneck进行重新开发，通过对音乐性、音乐语义和音乐音频声学Codec的分层和渐进式表达体系，支持生成**48kHz、双声道**音乐。该版本将音乐生成拆解为不同层级的表示学习与生成任务，实现音乐结构化生成：**高层负责音乐性与结构组织、中层负责语义与表达、底层负责声学细节与高保真还原**。在可控性方面，模型还进一步支持**精准Remix改写、细粒度音频编辑**等功能。

2025年9月，AI音乐智能体Tunee上线：从需求理解到一键生成MV，对话式创作全流程

创作流程（创作→制作→分发全链路闭环）：**用户通过自然语言提出创作需求→Tunee智能解读需求，快速生成多套不同风格、旋律的创作方案→用户选择对应方案或持续对话调整方案→Tunee一键生成MV，画面与音乐节奏自动匹配→生成作品可直接发布至抖音、TikTok等短视频平台**。例如：用户可直接输入“创作一首适配XX场景的XX风格歌曲”。

Tunee MV Agent主要功能：（1）**对话式创作**：用自然语言描述画面，直接驱动MV生成，减少复杂参数设置；（2）**音乐理解**：上传音乐后，系统智能分析节奏、情绪和关键时刻，辅助生成创作方向；（3）**@角色上镜**：在对话框中@指定角色，最多支持5个角色同时参与，精准控制角色在镜头中的位置；（4）**分镜级调整**：支持调整剧本、分镜和单个镜头；（5）**画布直接编辑**：在画布中预览、修改和替换素材。

□ 作为多模态AI音乐生成模型，天谱乐支持文本生曲（最长可生成3.5min歌曲/音乐）、图片生曲（上传图片，生成30s音乐视频）以及视频生曲（上传10-60s视频，生成音乐视频）。为进一步降低AI音乐制作门槛，2025年9月，音乐智能体Tunee上线，实现“需求理解→音乐视频MV生成→分发”全链路闭环。截至2025年12月，Tunee产品页面浏览量与日活均突破100万。

中国主流音乐大模型——天谱乐 (2/2)

商业化层面，ToC端，天谱乐采用订阅制，各版本商用授权范围存在差异，目前已有超5,600万用户使用过该模型；针对企业客户/专业开发者，天谱乐提供标准化API服务，采用充值制，各模型/功能定价不同

天谱乐商业化现状

Tempolor	定价	权益	Tunee	定价	权益
Lite版	\$9.79/月 (年付\$5.59/月)	最长生成4min曲目；每月1,000首；最多保存4个声音克隆；每天支持10次下载免版权音乐库音乐；授权范围含社媒平台和播客	免费版	\$0	每天100积分，约15首歌曲；音轨分离（2轨）；音乐视频方面，1个并发视频项目
Pro版	\$13.99/月 (年付\$7.99/月)	最长生成4min曲目；每月3,000首；支持无限分割曲目为人声和器乐主干；最多保存10个声音克隆；无限次下载免版权音乐库音乐；授权范围含社媒平台、播客、商业视频广告等客户工作、付费广告	Basic版	\$18/月（年付\$15/月）	每月1,800积分，约250首歌曲； 版权证明 ；每月领取1首历史歌曲版权证明；5个并发项目；音轨分离（6轨）；音乐视频方面，约8个MV和3个并发视频项目
Ultra版	\$29.99/月 (年付\$15.99/月)	最长生成4min曲目； 无限生成轨道 ；无限分割曲目为人声和器乐主干；最多保存50个声音克隆；无限次下载免版权音乐库音乐；授权范围含社媒平台、播客、商业视频广告等客户工作、付费广告、 电视广播等传统媒体、音乐发行	Plus版本	\$38/月（年付\$32/月）	每月4,000积分，约500首歌曲；版权证明；每月领取1首历史歌曲版权证明；5个并发项目；音轨分离（6轨）；音乐视频方面，图片模型15%折扣、视频模型5%折扣、约15个MV、3个并发视频项目、 移除官方水印、商业用途
<p>❑ ToC端，Tempolor和Tunee均采用订阅制收费，Tempolor定价为每月\$9.79-\$29.99，最长支持生成4min的曲目，各版本商业授权范围不同；Tunee定价为每月\$0-\$89，其中免费版不可商用，Basic版仅生成音乐可商用，生成MV不支持商用，Plus版和Pro版均支持MV商用。针对企业客户与专业开发者，Tempolor开放平台提供标准化API服务，采用充值制，价格为¥1/100积分，各项功能定价不同，如使用Tempolor v4.6生成一首歌曲为¥0.35（最长支持5min人声音乐）、使用Tempolor v3生成一首歌曲为¥0.2（最长支持2min人声音乐）。截至2026年1月，已有超5,600万用户使用过天谱乐大模型，累计创作音乐2,300万首。</p>			Pro版	\$89/月（年付\$75/月）	每月9,200积分，约1,300首歌曲；版权证明； 每月领取3首历史歌曲版权证明 ；10个并发项目；音轨分离（6轨）；音乐视频方面，图片模型30%折扣、视频模型15%折扣、约40个MV、 6个并发视频项目、移除官方水印、商业用途

中国主流音乐大模型——MiniMax Music

技术方面，Music 2.6最长支持4min歌曲生成，音包延迟降至20s以内，新增翻唱与歌词优化器功能；商业化方面，ToC端采取积分充值制，定价为1积分/首，为企业和专业开发者提供API服务，定价为¥1/首

MiniMax Music发展现状

2025.09: Music 1.5

- 最长4min，可对歌曲风格、情绪、场景进行自定义和强控制，如歌词编排
- 对人声唱腔发音技巧进行建模，可生成不同声线、唱腔的音色
- 对乐器进行细粒度建模，使得编曲丰富、乐器层次清晰、演奏技巧富于变化
- 结构清晰，前奏、主歌、副歌等段落间区别明显

2026.01: Music 2.5

- 段落级精准控制：开放全段落标签控制，精准支持包括前奏、桥段等在内的14+种结构变体，能够承载高复杂度音乐创作
- 人声/风格/混音的物理级保真：对华语流行音乐进行深度优化、旋律设计贴合流行趋势、咬字清晰与中英文无缝衔接；优化人声合成，真实自然；针对不同风格，自动调整混音策略

2026.04: Music 2.6

- 最长4min，音包延迟降至20s以内、指令控制全面增强、中低频声学表现系统性优化
- 新增AI翻唱：将任何现有歌曲转变为新风格，同时保留原有旋律；新增歌词优化器：可自动根据提示词写歌词，用户无需自己提供任何歌词
- 最高44.1kHz采样率，导出MP3、WAV或PCM等格式
- 开源3个Music Skill：minimax-music-agent具备完整音乐生成能力、minimax-music-playlist生成整份定制歌单、buddy-sings构建专属声音身份，以角色第一人称为用户即兴创作

- ❑ 2026年4月，MiniMax发布最新音乐生成模型MiniMax Music 2.6，支持最长生成4min歌曲，音包延迟降至20s以内，新增AI翻唱和歌词优化器等功能，同时开源3个Music Skill，降低音乐生成门槛。商业化层面，针对个人用户，MiniMax采用充值制，1积分可生成一首完整歌曲，积分充值后永久有效。针对企业客户和专业开发者，提供标准API服务，按量计费，Music 2.5、Music 2.5+、Music 2.6均为¥1/首。

来源：公司官网、头豹研究院

版本	定价	权益
Basic版	\$9.9	120积分
Advanced版	\$29.9	375积分
Pro版	\$59.9	788积分
Premium版	\$89.9	1,266积分

Music1.5每首歌1积分，Music 2.5和2.6每首歌4积分，积分充值后永久有效，所有版本生成音乐均可商用

海内外主流音乐大模型对比 (1/2)

从生成能力看，Suno和Stability Audio功能相对全面；Udio音质质感顶尖；Mureka在分轨导出方面表现突出；天谱乐风格覆盖广、MV生成链路较成熟；MiniMax Music对多元歌曲结构变体支持度高

海内外主流音乐大模型生成能力对比

维度	Suno	Udio	Stability Audio	Mureka	天谱乐	MiniMax Music
最长支持时长	8min	130s	6min	5.5min	3.5min	4min
音乐参考	✓ (人声/伴奏/音乐)	✓ (现有音乐)	✓	✓	✓	✗
支持风格	10+	11+	—	—	25+	—
Studio	✓	✗	✓	✓	✗	✗
结构变体	6	—	—	—	—	14+
采样率	48kHz	48kHz	44.1kHz	—	48kHz	44.1kHz
优化歌词	✓	✓	✓	✓	✓	✓
音乐拓展	✓	✓	✓	✓	✗	✓
AI翻唱	✓	✗	✓	✓	✗	✓
Remix (混音)	✓ (可控制偏离程度)	✓	✓ (可控制偏离程度)	✓ (内测阶段)	✓ (精准改写)	✗
声音克隆	✓	✓	—	✓	✓	✗
局部编辑	✓	✓	✓	✓	✓	✗
一键MV生成	✓	✗	✓	✗	✓	✗
分轨导出	12	4	2	12	6	多轨
自定义模型	✓	✗	✓	✓	✗	✗

海内外主流音乐大模型对比 (2/2)

从商业化现状看，针对个人用户，除MiniMax Music外，其它模型均采用订阅制，价格集中在\$10-30/月，针对企业/专业开发者，国产模型提供标准API服务，海外仅有Stability Audio提供定制模型服务

海内外主流音乐大模型价格及商业化现状对比

维度	Suno	Udio	Stability Audio	Mureka	天谱乐	MiniMax Music
商业化模式						
个人用户（按月）	订阅制（\$0/10/30）	订阅制（\$0/10/30）	订阅制（\$0/11.99/29.99/89.99）	订阅制（¥0/88/168）	订阅制（\$9.79/13.99/29.99）	充值制（充值后，积分永久有效）
企业或专业开发者	暂无官方API	—	定制模型服务	标准API服务（V8、V9歌曲生成¥0.33/首）	标准API服务（V4.6歌曲生成¥0.35/首；V3为¥0.2/首）	标准API服务（Music 2.5、Music 2.5+、Music 2.6均为¥1/首）
免费版商业使用权	✘	✓	✘	✘	✘	无免费版（测试阶段除外）
充值后，历史歌曲商业使用权	✘（除平台给予特殊作品许可）	✓	✘	✘	✓（可领取历史歌曲版权证明）	—
生成歌曲自由下载	✓（付费用户）	✘	✓	✓	✓	✓
版权问题解决方案	与华纳音乐达成合作推出授权AI模型	与环球音乐达成合作推出授权平台	开源，授权数据训练+企业定制模型	—	—	—
商业化现状	截至2026年2月，付费用户达200万，ARR达3亿美元	2026年4月，网站访问量约为140万	—	2025.03-2025.11，新增注册用户数700万，服务企业/专业开发者客户数超8,000家	截至2026年1月，超5,600万用户使用过	—

Chapter 3

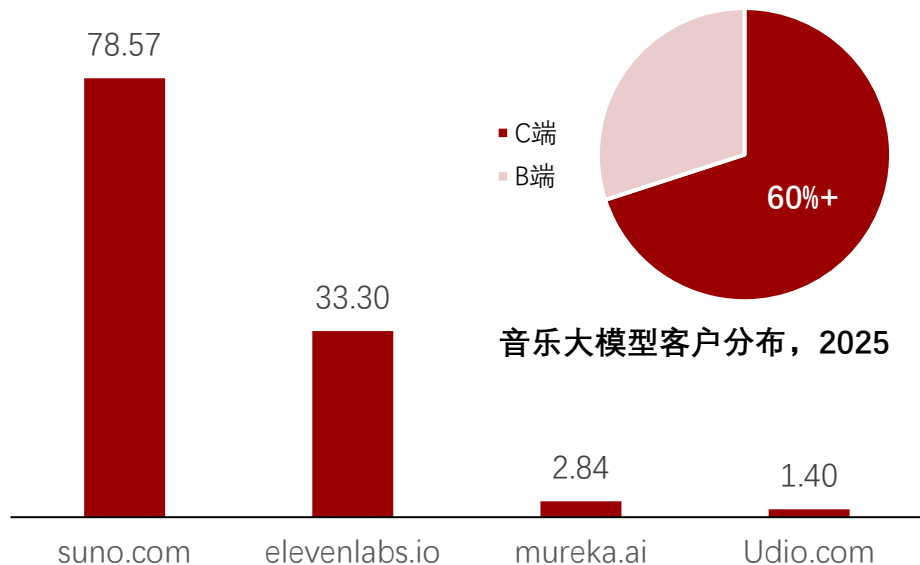
音乐大模型市场 应用洞察

音乐大模型核心应用场景 (1/2)

目前，C端用户是音乐大模型核心消费群体，其主要借助模型生成短视频BGM、产出参考旋律/部分核心段落、制作广告Demo、创作活动主题歌/知识记忆歌/趣味儿歌等

音乐大模型核心用户

➤ 部分主流音乐大模型网站访问量（单位：百万），2026.04



短视频BGM：交付以原创感背景音乐为主，核心用户包括短视频团队、自媒体用户、电商卖家。

广告Demo：交付以口号歌、广告旋律草稿为主，核心用户包括品牌方、创意代理商等。



创意草稿：交付以歌曲方向、hook、旋律参考为主，核心用户为音乐制作人、编曲师。

活动主题歌：交付以活动歌曲、片头片尾音乐为主，核心用户为企业活动策划人员、社区运营人员。

教学与娱乐：交付以知识记忆歌、趣味儿歌为主，核心用户包括音乐教育从业者、亲子内容创作者。

- 目前，音乐大模型用户群体以C端个人为核心，该类用户整体占比达60%以上。其中Suno作为最热门的C端音乐大模型（Suno暂不提供官方API服务），2025年平台日均生成音乐数可达700万首，2026年4月单月访问量达7,857万，用户活跃度保持较高水平。结合实际使用需求与应用场景，Suno用户可以分为5类，不同群体的使用方向各有侧重。（1）短视频团队、自媒体创作者、电商卖家，主要借助模型快速打造具备原创属性的背景音乐，视频剪辑、店铺宣传等；（2）专业音乐制作人和编曲师将其作为创作辅助工具，用来敲定歌曲整体风格、产出参考旋律与核心Hook段落，提升创作效率；（3）品牌方、创意代理商多用于制作品牌口号曲、广告旋律初稿，缩短营销音乐的创作周期；（4）企业活动策划人员、社区运营者依托平台定制活动主题曲、栏目片头片尾曲，满足线下活动与线上栏目配乐需求；（5）音乐教育从业者、亲子内容创作者等主要创作知识记忆歌曲、趣味儿歌等作品。

音乐大模型核心应用场景（2/2）

在B端，音乐大模型主要应用场景为：（1）为AI音乐、AI MV等AI创作平台提供各类功能；（2）为影视制作公司快速生成符合剧情氛围的音乐；（3）为游戏开发企业生成背景音乐或特效音乐

音乐大模型B端客户主要分布

AI创作平台

- **AI音乐创作平台**：例如①中国快歌作为自助写歌平台，其技术栈同时集成海内外多家模型，其海外版音乐大模型合作方为Mureka；②北美AI音乐创作平台Sondo将Mureka声音克隆、Remix等功能接入产品后，两者合作体量增长超70%。
- **AI MV生成平台**：某头部海外音乐MV平台已接入谷歌Lyria、昆仑万维Mureka等模型。

影视制作公司

- 影视制作需要为电影、电视剧、纪录片等制作背景音乐、主题曲、场景配乐等，音乐大模型可快速生成符合剧情氛围、风格需求的音乐，降低配乐成本和时间。例如：MiniMax旗下音乐大模型主要面向B端客户，目前公司与恒星引力达成战略合作，发布面向东方幻想IP的AI创作体系，开放影视AI内容孵化平台。

游戏开发企业

- 游戏中的角色动作、场景切换、战斗环节等需要大量不同风格的音乐，音乐大模型能按需生成动态音乐，支持游戏音乐的可变性和交互性。例如：据中国音数协游戏工委调研，在中国22家具有代表性的游戏企业中，AI在游戏研发场景的应用率达86.4%，在合成游戏背景音乐和特效音乐环节的渗透率为52.6%。

其它行业客户

- **短视频内容平台**：短视频创作者需要大量短时、风格多样的背景音乐，音乐大模型可满足其快速生成个性化音乐的需求，帮助创作者提升内容质量。
- **广告及营销机构**：为品牌广告、产品宣传视频等制作专属音乐，通过音乐大模型快速生成符合品牌调性、目标受众喜好的音乐，提升广告吸引力和传播效果。

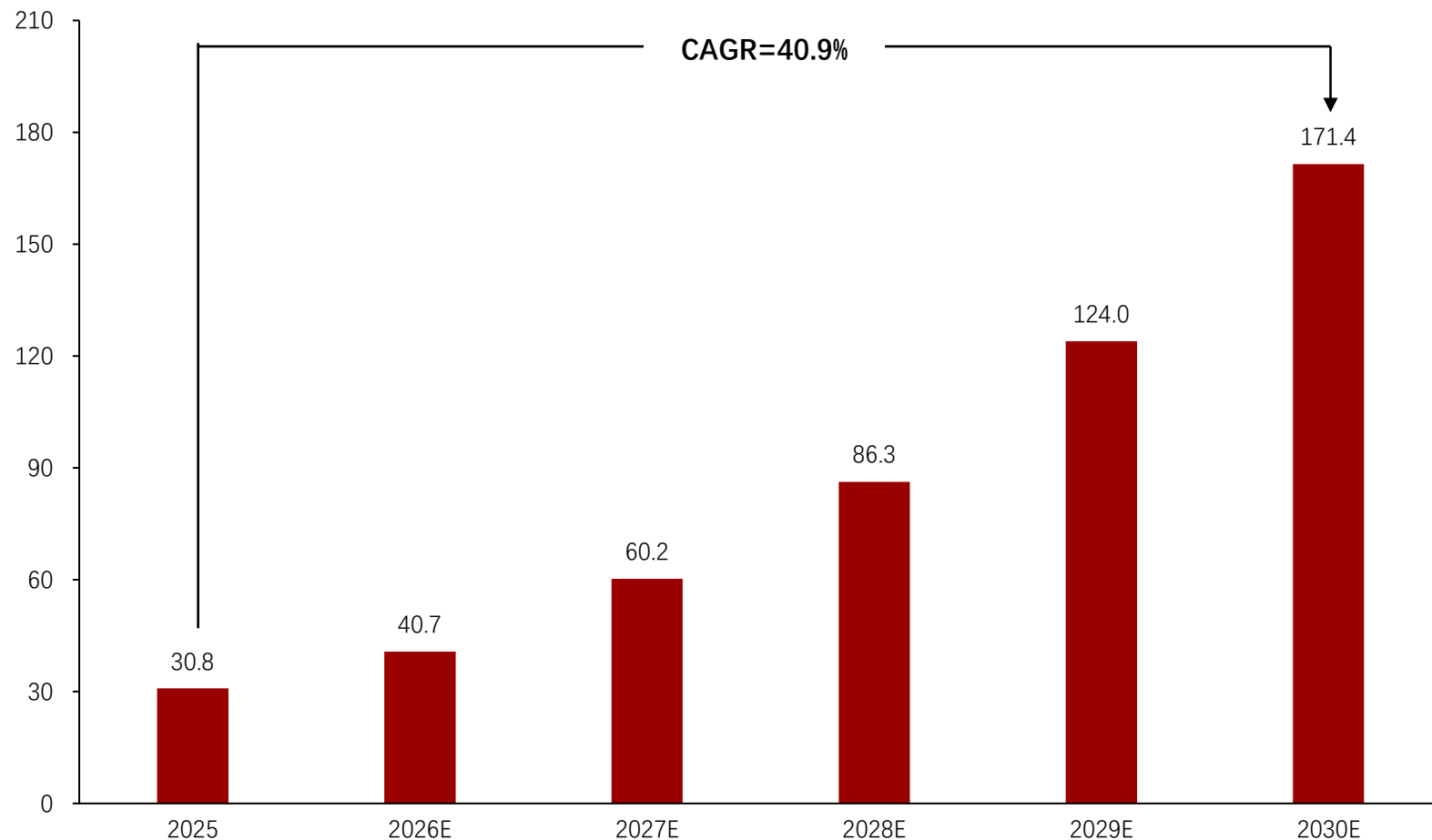


音乐大模型市场规模

2025年，全球音乐大模型市场规模约30.8亿元，预计到2030年，随着B、C端持续渗透，该市场有望增至171.4亿元，2025-2030年CAGR达40.9%

音乐大模型市场规模（按营收计）

➤ 全球音乐大模型市场规模（单位：亿元），2025-2030E



来源：头豹研究院

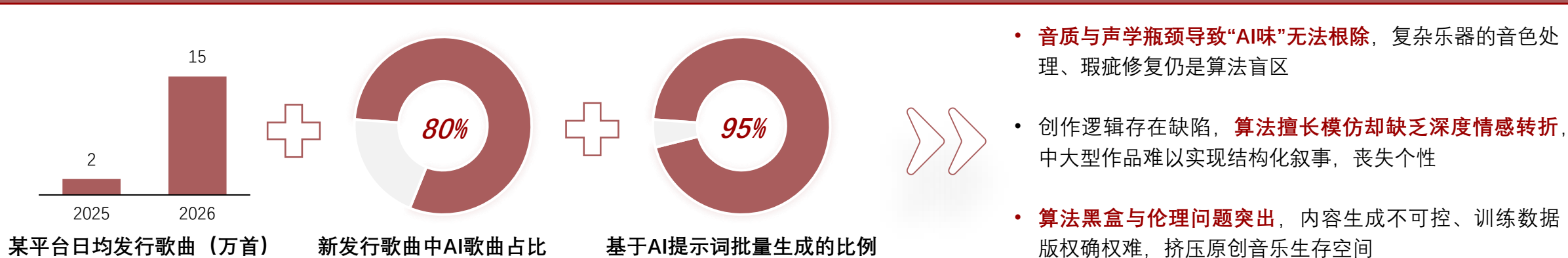
- ❑ 2025年，全球音乐大模型市场规模约为30.8亿元，预计2030年将增至171.4亿元，2025-2030年期间CAGR约为40.9%。这一增长主要得益于：
- ❑ **C端付费用户持续增长**。AI音乐应用使用门槛极低，普通用户仅凭文字描述就能创作专属音乐，轻松满足大众的创作意愿与表达需求。同时，音轨分离、专业工作站、自定义模型等进阶功能的落地，也让音乐大模型逐步获得半专业及职业音乐人的青睐，进一步推动用户规模扩张。据调研，50%的音乐消费者对AI音乐应用高度感兴趣。
- ❑ **政策支持与效率革命成为核心驱动力，推动音乐大模型不断渗透B端领域**。凭借降本增效的突出优势，AI音乐快速渗透至商业领域，已成为影视配乐、广告音乐、游戏音效制作的重要工具。未来，随着《关于深入实施“人工智能+”行动的意见》、《北京市促进“人工智能+视听”产业高质量发展行动方案（2025-2029年）》等政策出台，将持续推动音乐大模型与各产业深度融合。

音乐大模型发展趋势（1/2）

当前AI音乐生成虽在平台歌曲发行中占比快速提升，但存在音质、创作逻辑与版权伦理等问题，正通过人机协同创作与动态生成服务向开放、可交互的创作模式演进

音乐大模型发展趋势一

当前AI音乐生成存在问题



解决方式：人机协同创作+动态AI音乐生成服务

一次性
输出



局部精修重绘

多轨分轨生成

MIDI

分段生成

时间线编辑

参考创作

AI虚拟人声

版本管理

定制模型

.....



循环交
互输出

- 2025年以来，音乐大模型在音轨分离、局部生成重构、参考创作、定制模型等功能上迭代升级，AI音乐生成从算法“一键生成”被重塑为一种开放、动态且高度协同的机制，AI从执行者转向执行+协作者，参与到创作决策与音乐结构建构全过程。
- AI驱动的个性化音乐推荐引擎已成为流媒体平台核心竞争力，推动**动态生成式音乐服务**逐渐崛起。即结合实时数据和实时场景，生成符合用户需求、独一无二的音乐。例如：Mubert可根据用户场景适配需求（工作、运动、睡眠等）或实时情绪响应，生成无限时长、无缝衔接的个性化音乐流。

音乐大模型发展趋势 (2/2)

AI音乐创作重构产业价值链，各方正依托授权训练、版权分成、联合创作、确权等合规合作模式，形成技术输出工具、内容注入创意、多方利益共享的良性生态

音乐大模型发展趋势二

海外AI音乐合规实践

01

授权训练合作模式

AI音乐开发者与版权方主要合作模式

版权方将音乐作品和录音授权给AI开发者用于模型训练、优化与迭代，AI音乐开发者以一次性费用、按量计费或收益分成方式作为对价。例如：Stable Audio采用合作伙伴AudioSparx的授权数据集进行训练。

02

版权分成合作模式

AI音乐开发者与创作者主要合作模式

AI音乐开发者开放AI创作工具，创作者完成音乐制作，双方约定商用收益分成比例，由版权方统一结算分配。AI音乐开发者可通过分成获得持续收益，不依赖单次授权费用。例如：AI平台Loudly与数千个独立音乐人合作。

03

联合创作模式

AI音乐开发者与创作者深度合作形态

AI音乐开发者提供AI技术支持、创作者提供核心创意，双方组建项目组共同参与创作、制作、发行全流程，版权共有、收益按约定分配、风险共担。例如：华纳音乐与Suno合作。

04

确权赋能合作模式

版权清晰的保障机制

在协议中明确作品版权归属，AI音乐开发者依托技术为创作者提供版权注册、权属证明、侵权监测等支持，通过搭建区块链存证平台，对接全球版权机构，全面保障创作者版权权益。例如：Deezer上线AI音乐检测+标注系统。

≥80%

头部AI音乐平台至少与一家主流唱片公司/版权方达成授权合作

≥65%

独立音乐人使用AI辅助创作，且通过合作模式实现AI作品商用

- AI音乐创作正重构音乐产业原有价值体系。当前，AI音乐平台、音乐企业、创作者及版权机构逐步搭建起全链路协作体系，形成“技术输出工具、内容注入创意、多方收益共享”的良性生态。行业主流合作分为授权训练、版权分成、联合创作、确权赋能四大类型，均遵循版权合规、利益共享、风险共担的核心原则。调研数据显示，全球超80%的头部AI音乐平台已与主流唱片公司或版权机构开展授权合作；65%的独立音乐人借助AI工具辅助创作，并依托各类合作模式完成作品商业化变现。

头豹业务合作

全球视野 · 本土洞察 · 研究数据 · 可信知识网络



行业数据API

开放原创报告与研究数据接口，
支持企业知识库、系统平台及AI
应用高效接入和调用



KNIT解决方案

构建企业可信内容体系，提升品
牌在AI搜索与问答中的可见度、
准确性与转化效果



报告会员账号

可阅读全部原创报告和百万数据，
提供PC及移动端，方便触达平台
内容



定制报告/白皮书

对产业及细分行业进行现状
梳理和趋势洞察，输出全局
观深度研究报告



商业尽调

面向投资并购和商业决策，
评估标的公司的商业前景、
价值及风险



招股书引用

研究覆盖国民经济19+核心产
业，内容可授权引用至上市
文件、年报



报告作者

- 袁栩聪 | 首席分析师
- 付淑芳 | 行业分析师



service@leadleo.com



业务咨询

- 客服电话：400-072-5588
- 官方网站：www.leadleo.com



深圳办公室

广东省深圳市南山区粤海街道华润
置地大厦E座4105室
邮编：518057



上海办公室

上海市静安区南京西路1717号会
德丰国际广场2701室
邮编：200040



南京办公室

江苏省南京市栖霞区经济开发区兴
智科技园B栋401
邮编：210046

方法论

- ◆ 头豹研究院布局中国市场，深入研究19大行业，532个垂直行业的市场变化，已经积累了近100万行业研究样本，完成近10,000多个独立的研究咨询项目。
- ◆ 研究院依托中国活跃的经济环境，研究内容覆盖整个行业的发展周期，伴随着行业中企业的创立，发展，扩张，到企业走向上市及上市后的成熟期，研究院的各行业研究员探索和评估行业中多变的产业模式，企业的商业模式和运营模式，以专业的视野解读行业的沿革。
- ◆ 研究院融合传统与新型的研究方法，采用自主研发的算法，结合行业交叉的大数据，以多元化的调研方法，挖掘定量数据背后的逻辑，分析定性内容背后的观点，客观和真实地阐述行业的现状，前瞻性地预测行业未来的发展趋势，在研究院的每一份研究报告中，完整地呈现行业的过去，现在和未来。
- ◆ 研究院密切关注行业发展最新动向，报告内容及数据会随着行业发展、技术革新、竞争格局变化、政策法规颁布、市场调研深入，保持不断更新与优化。
- ◆ 研究院秉承匠心研究，砥砺前行的宗旨，从战略的角度分析行业，从执行的层面阅读行业，为每一个行业的报告阅读者提供值得品鉴的研究报告。

法律声明

- ◆ 本报告著作权归头豹所有，未经书面许可，任何机构或个人不得以任何形式翻版、复刻、发表或引用。若征得头豹同意进行引用、刊发的，需在允许的范围内使用，并注明出处为“头豹研究院”，且不得对本报告进行任何有悖原意的引用、删节或修改。
- ◆ 本报告分析师具有专业研究能力，保证报告数据均来自合法合规渠道，观点产出及数据分析基于分析师对行业的客观理解，本报告不受任何第三方授意或影响。
- ◆ 本报告所涉及的观点或信息仅供参考，不构成任何证券或基金投资建议。本报告仅在相关法律许可的情况下发放，并仅为提供信息而发放，概不构成任何广告或证券研究报告。在法律许可的情况下，头豹可能会为报告中提及的企业提供或争取提供投融资或咨询等相关服务。
- ◆ 本报告的部分信息来源于公开资料，头豹对该等信息的准确性、完整性或可靠性不做任何保证。本报告所载的资料、意见及推测仅反映头豹于发布本报告当日的判断，过往报告中的描述不应作为日后的表现依据。在不同时期，头豹可发出与本报告所载资料、意见及推测不一致的报告或文章。头豹均不保证本报告所含信息保持在最新状态。同时，头豹对本报告所含信息可在不发出通知的情形下做出修改，读者应当自行关注相应的更新或修改。任何机构或个人应对其利用本报告的数据、分析、研究、部分或者全部内容所进行的一切活动负责并承担该等活动所导致的任何损失或伤害。

