

快手直播的AI进化论： 技术落地与用户体验升级

演讲人：吴翔宇

快手 / 算法专家

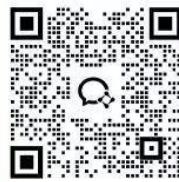
AiCon
全球人工智能开发与应用大会

目录

- 01 秀场直播的AI落地趋势
- 02 理解能力的突破
- 03 用户触点的升级
- 04 互动智能体的落地
- 05 直播AI的下一步

极客邦科技 2026 年会议规划

促进软件开发及相关领域知识与创新的传播



参会咨询



查看会议

北京

1200人

QCon

全球软件开发大会

会议时间：4月16-18日

- Agentic Engineering
- AgentOps
- 下一代模型架构与推理优化
- AI 原生基础设施
- 知识工程实践
- AI 安全

深圳

1000人

AiCon

全球人工智能开发与应用大会

会议时间：8月21-22日

- Agentic AI
- 轻量化与高效推理
- 多模态应用
- AI + IoT 场景实践
- AI 工业化落地

北京

1000人

AiCon

全球人工智能开发与应用大会

会议时间：12月18-19日

- 大模型架构创新
- 多模态 AI 产业融合
- 具身智能
- AI for Science
- 大模型安全

4月

6月

8月

10月

12月

AiCon

全球人工智能开发与应用大会

会议时间：6月26-27日

- AI Infra 系统工程
- 多 Agent 协作与实践
- 多模态融合
- 模型训练与推理创新
- 数据平台与特征服务

上海

1000人

QCon

全球软件开发大会

会议时间：10月22-24日

- AI Agent
- Vibe Coding
- 智能可观测
- 推理基建
- 模型攻防
- AI x 创造力

上海

1200人

01 秀场直播的AI化趋势

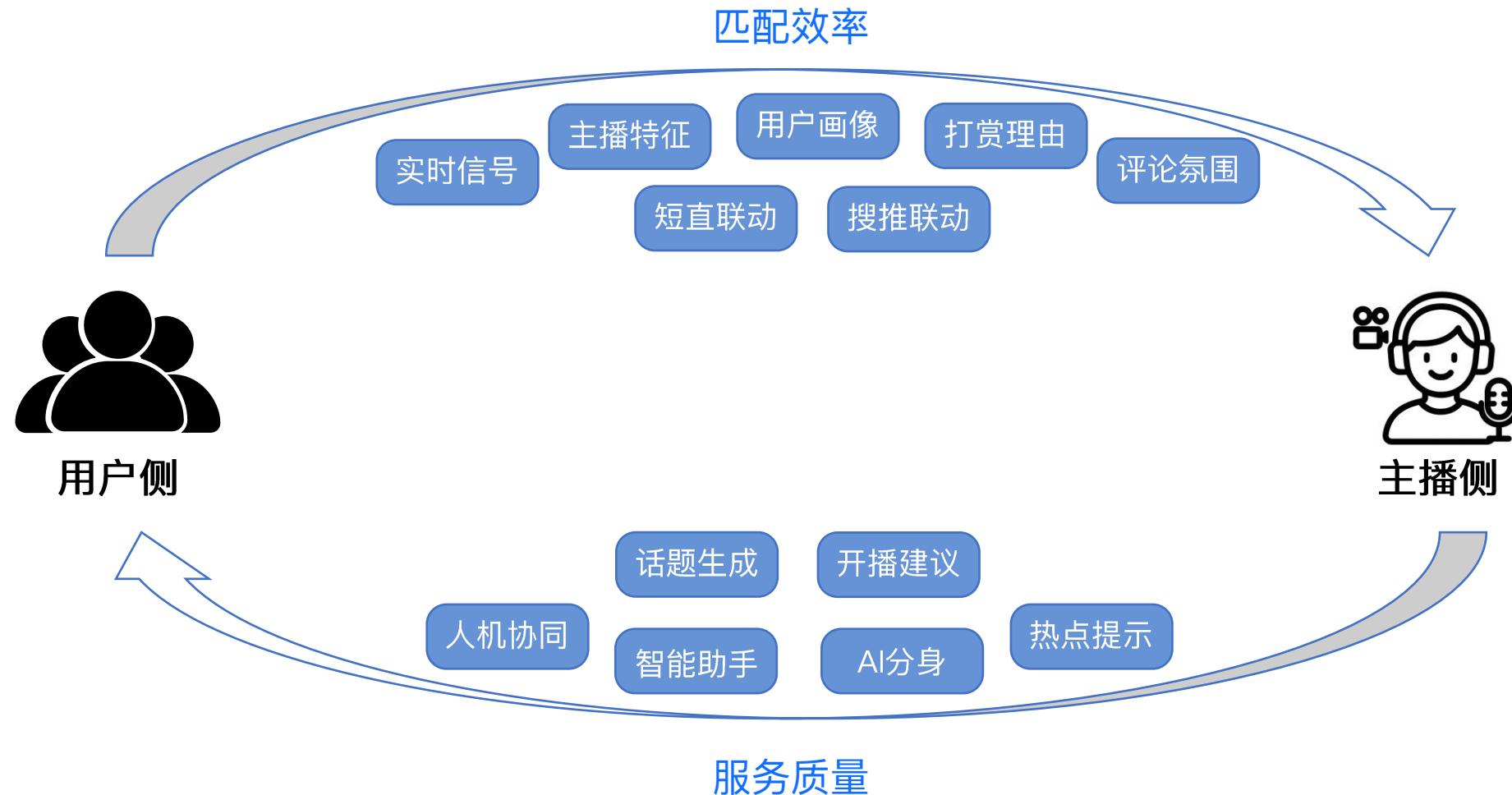
直播行业发展历程：从娱乐产业到社会基础设施



直播业务的本质，是UA关系的建立



AI之于秀场直播：匹配效率与服务质量的双重提升

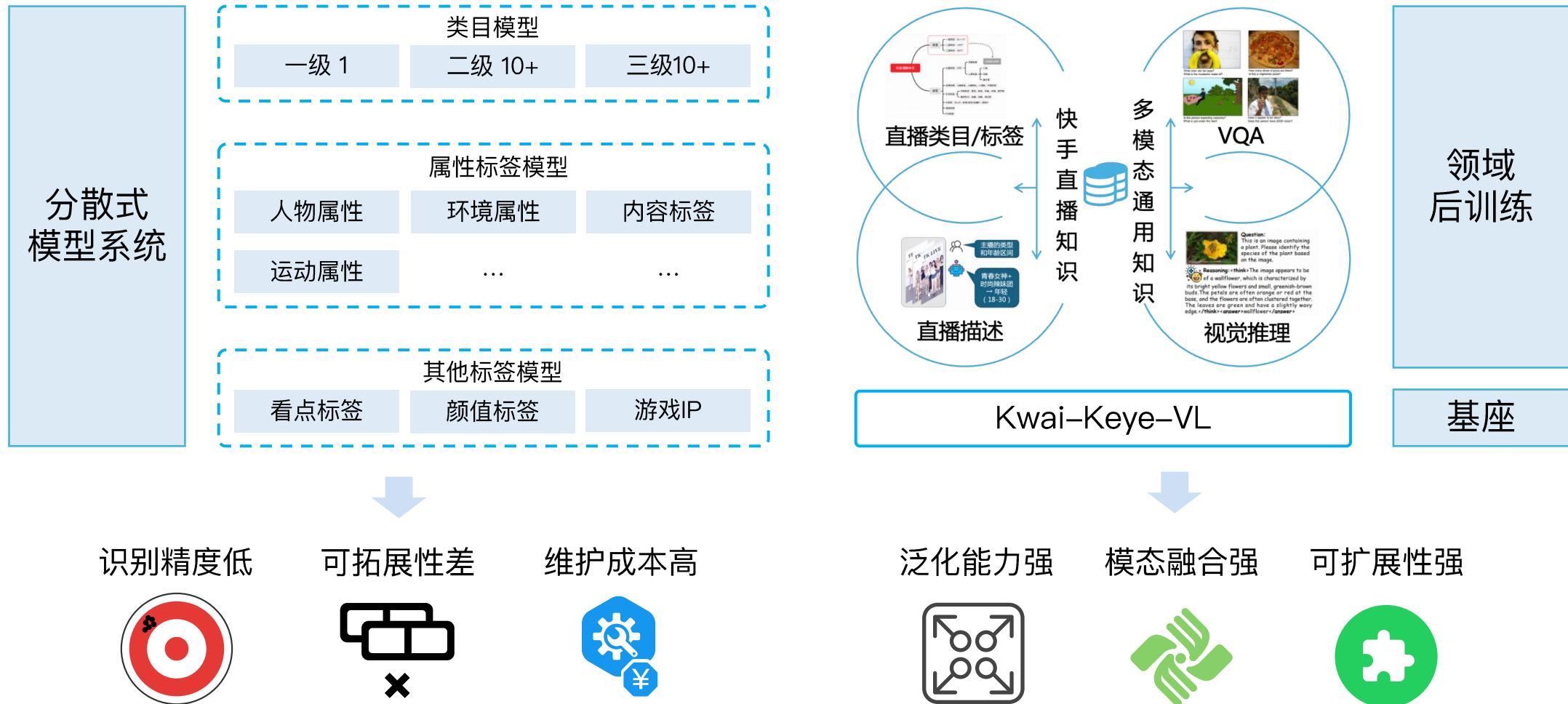


02 理解能力的突破

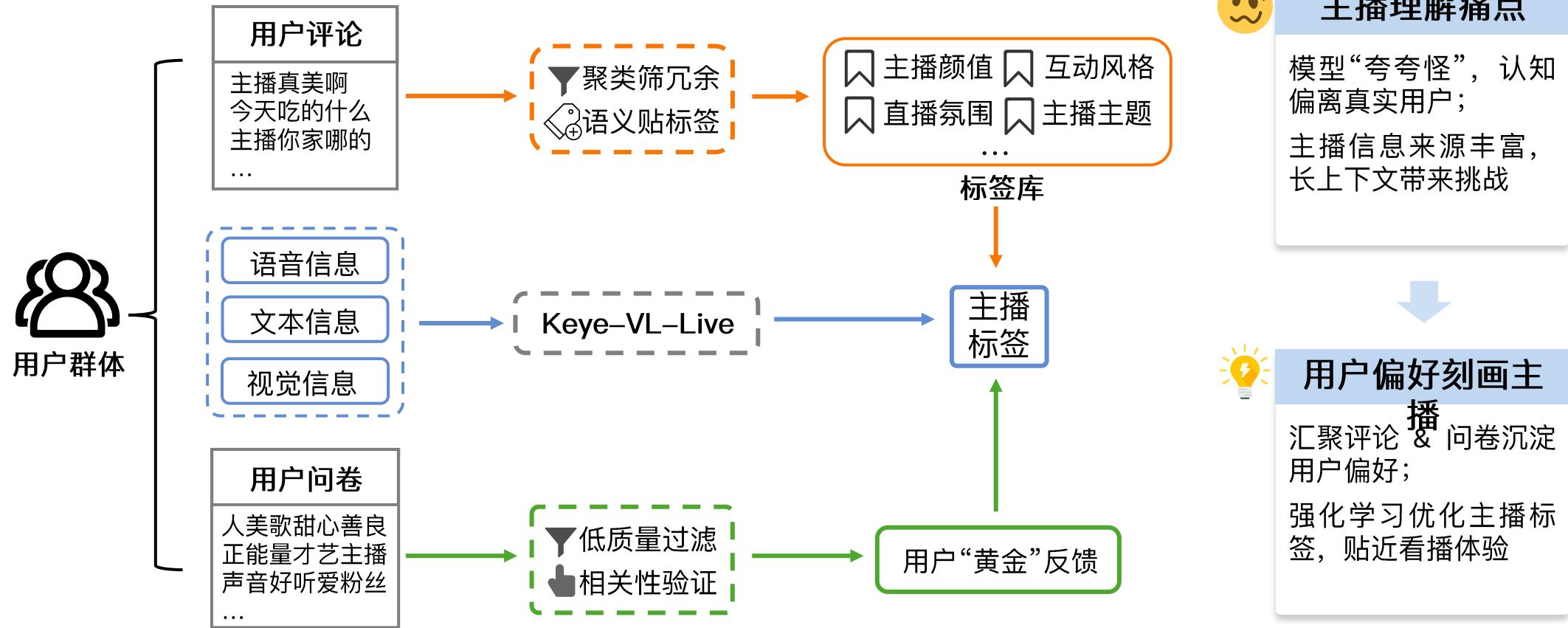
理解主播、认识用户，挖掘有潜力的主播-用户兴趣联系



构建更懂快手直播的多模态直播基座

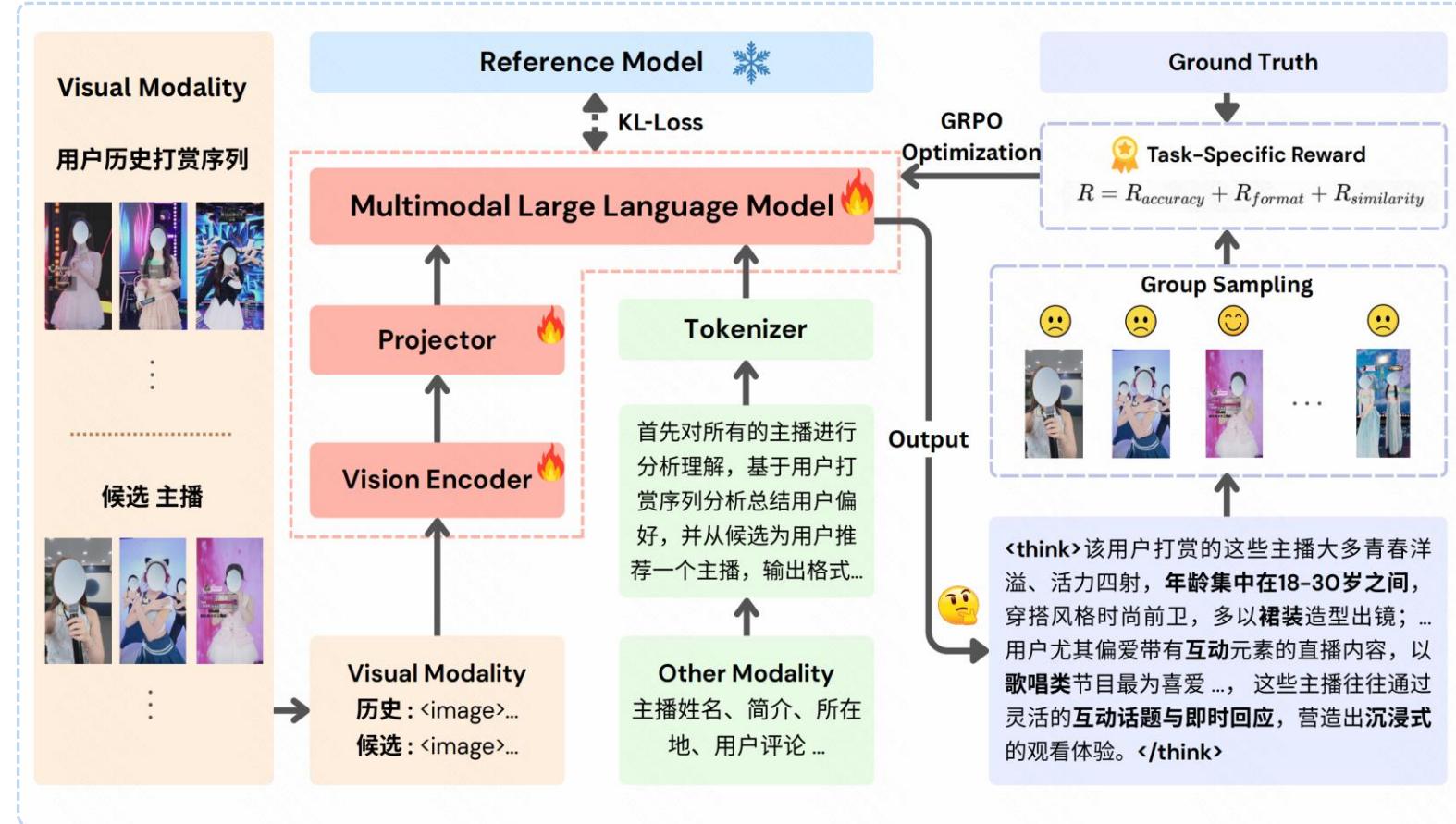


对齐用户主动表达，从用户视角理解主播





推演用户看播偏好，从用户行为理解UA关系



从单边理解迈向
双向建模

1. 统一建模用户多模态偏好与主播特征；
2. 使用打赏等行为信号引导模型优化；
3. 推荐结果同时输出偏好画像与推荐理由



直播场景下的生成式推荐落地



用户打赏序列



用户偏好分析

从用户的打赏行为与直播内容特征综合来看，用户更倾向于为高互动、强陪伴感的直播体验付费。相较于单向输出型内容，用户对“被看见、被回应”的互动链路更敏感——当直播间能够提供更明确的参与感、存在感与即时反馈（例如被点名、被提及、被互动承接）时，更容易触发停留与打赏行为。

在内容形态上，用户明显偏好联动感更强的互动形式，尤其是同框互动、对话推进、节奏更活跃的直播场景。这类形式天然更容易制造情绪起伏与互动节点，用户的注意力与投入度也更容易被持续拉住，从而形成更稳定的付费动机。

整体审美与观看选择上，用户更青睐亲和、轻松、生活化的内容氛围：画面与表达更贴近日常，沟通方式更像“聊天陪伴”而非“表演展示”。与此同时，用户对社交/情感向主题的直播更敏感，更愿意为其中的情绪价值与关系感买单——本质上是在为“互动质量”和“陪伴体验”的综合感受付费。

个性化偏好标签

双人互动PK; 青春女神; 室内明亮互动; 情感交友向



用户兴趣
主播预测



理解能力升级，全面带动匹配效率提升

召回

深度优化主播描述与标签精度，
进行白盒化召回，
提升推荐系统多样性与可解释性。

营收

通过用户看播偏好的精准刻画，
推理潜力主播，
兼顾推荐准确度与体验新鲜感。

渗透

大模型驱动
理解能力升级

看播

关注

打赏

排序

将用户与主播标签引入
精排模型，首次将理解信号
以side-info的形式在排序中。

冷启动

对主播类别、质量等信息进行
全面刻画，指导流量分配，
加速好主播成才路径。

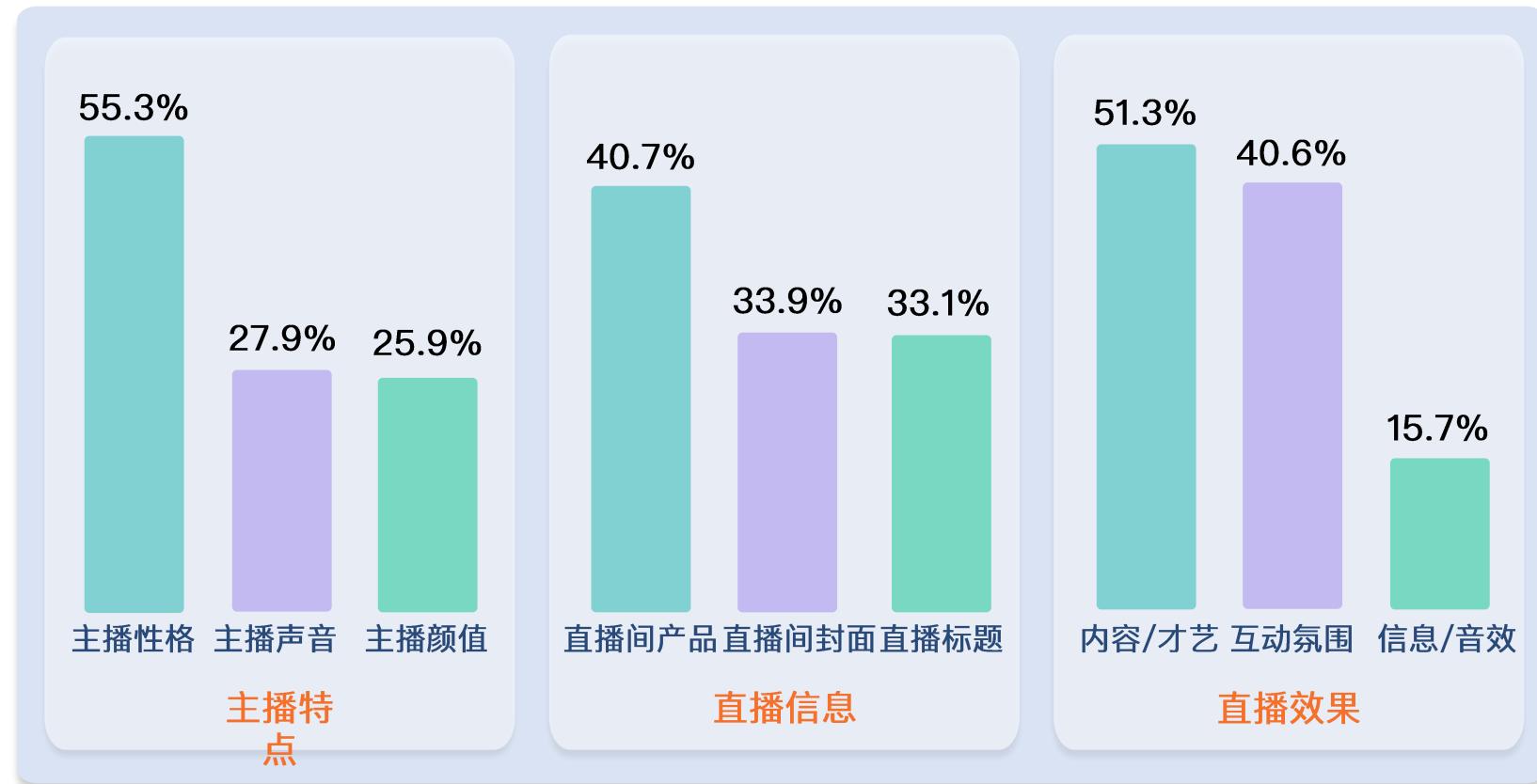
03 用户触点的升级

从「惊鸿一瞥」到「一见如故」，提升每一次匹配成功的效率

用户存在明确的了解主播的需求



当用户第一次刷到某个新主播时，TA们是怎么认识&筛选主播的？



用户原声

- 看是否开朗活泼，合眼缘就留意
- 看重颜值 / 气质，合心意就点进直播间
- 看商品 & 价格，少则直接划走
- 被好看封面 /“下饭”背景吸引，点进看
- 聊天有趣的直播间，会主动关注

降低用户认识主播的成本

面对每一刷都在变化的主播新面孔，用户了解主播的信息触点不足，导致选择成本越来越高。

直播间外 强化主播特质外显

让用户一眼看出这是什么样的主播

- 主播无法随时展现自己最好的一面，用户触达时刻难以始终最优
- 把主播核心亮点浓缩成间断文字，先让用户愿意点进来

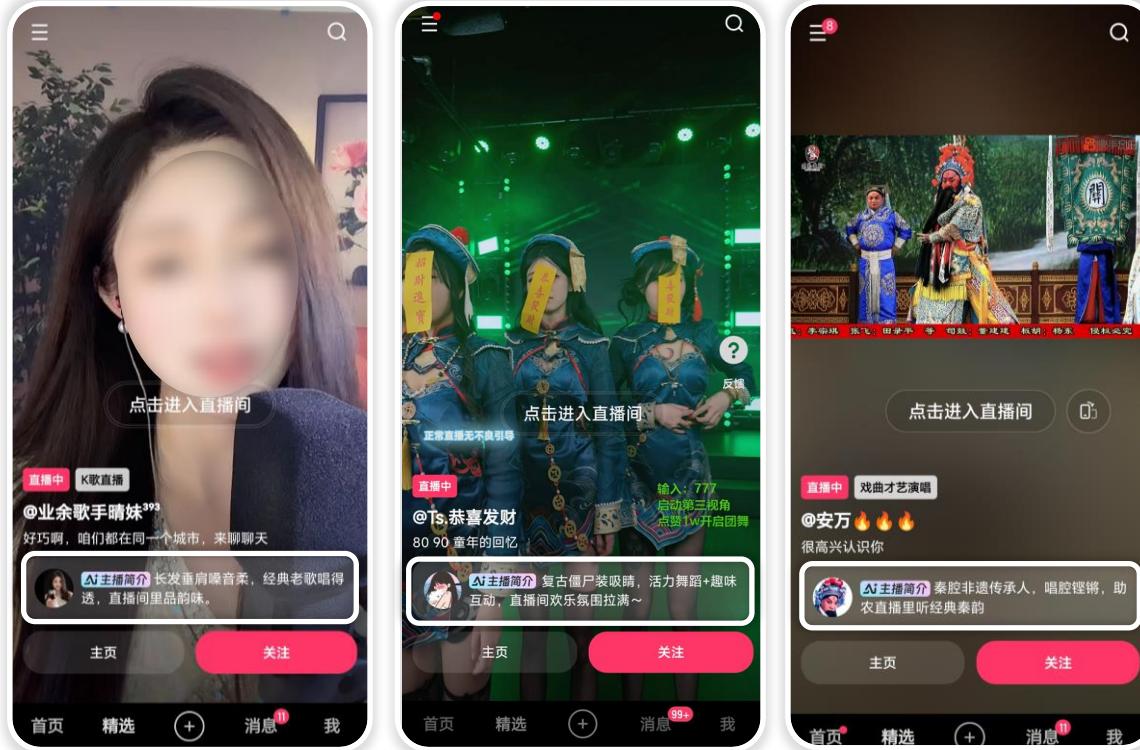


直播间内 实时内容滚动展现

让用户快速看懂主播在干什么

- 用户可能不了解上下文，看不进去，不知道主播在干什么
- 把直播精华和直播预告进行滚动播放，让用户看得懂，看的进去

AI主播简介：提升用户筛选效率



主播AI简介的落地逻辑



多源数据智能分析

整合主播历史直播内容、直播间行为、个人简介、直播预约信息等全站多源数据，构建“直播内容 + 主播画像”的基础信息底座。

大模型深度总结提炼

依托大模型强大的语义理解与概括能力，自动生成精炼、准确、具高吸引力的主播与内容解读文案，突出主播差异化特点与核心卖点。

核心阵地前置外显与引导

- 触点前置：**在简易直播间等首触点优先展示解读文案，缩短用户认知路径。
- 降低流失：**用户快速看懂内容与主播，缓解“认知前流失”，提升点击与停留。

AI小快播报：拉近用户看播距离



实时根据直播内容做出反应



内容梗概型



还原本场直播的“内容时间线”与核心节点

氛围鼓动型



提炼直播间情绪高点，激发用户参与感



浏览深度



观看时长



互动次数

04 互动智能体的落地

让开播更简单，让互动更自然，完成情绪价值到直播打赏的转化

主播们面临的共同挑战与各自困境

共同挑战：单人秀场模式的天然天花板

- 秀场直播依赖主播一人同时输出内容与互动，形成情绪与消费价值
- 单人模式天然限制营收上限，且用户体验高度依赖主播当下状态，难以持续放大



头部主播

- 知道如何呈现内容
- 无法一一服务所有用户
- 需要 AI 拓宽服务半径

中小主播

- 服务特色尚未形成
- 难以承接和服务好每位观众
- 需要 AI 补齐基础服务能力

AI 主播互动智能体：补齐服务能力 & 放大服务半径

互动助手：主播服务半径内的价值挖掘



播前准备

精准入场 × 赢在起跑线

- 话术模板库：个性化欢迎语、感谢话术
- 标题优化器：爆款标题生成、关键词植入
- 定向邀请：VIP专属邀请、纪念日提醒



播中辅助

实时识别用户 × 智能互动引导

- 用户标签：新老客识别、价值评级
- 破冰话题：昵称解读、地域互动、兴趣挖掘
- 熟人互动：历史回顾、动态关怀、专属记忆



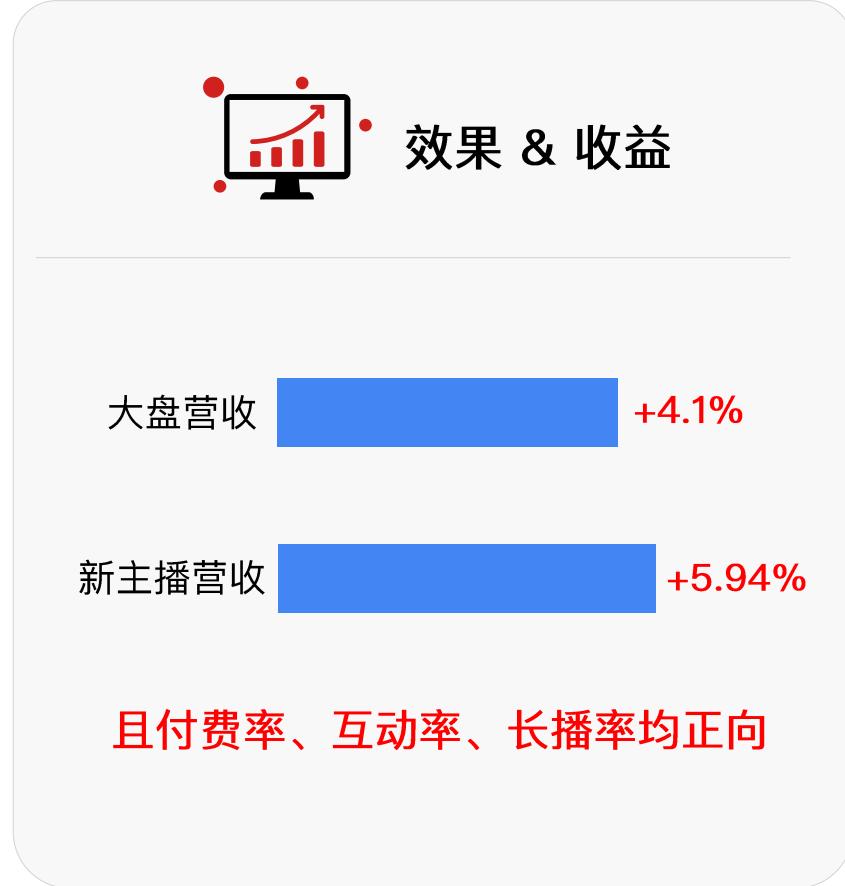
播后维护

深度关系维护 × 潜力用户挖掘

- 感谢私信：榜单答谢、个性化表达、情感连接
- 潜客发现：新用户筛选、价值评估、培养策略
- 需求洞察：弹幕分析、偏好识别、内容优化



协助主播更好地完成每一场直播



■ AI分身：主播服务半径外的价值扩散

核心能力

- ⌚ 复刻主播的灵魂：学习并复制有效的互动话术与节奏
- ⟳ 承接内容做延伸：对用户评论进行反馈，延续用户兴趣话题
- 🎤 引导行动促转化：根据用户关注、打赏等行为触发感谢等互动行为

关键挑战

- ❗ 语料臃肿复杂，特征稀疏。
- ❗ 垂类数据稀缺，模型直出效果差。
- ❗ 实时高并发，badcase率需极低。

核心模块



ASR清洗Agent

将语音转为干净文本



语料清洗Agent

自动化清洗和标注数据



多Agent协作

多智能体协同完成任务



模型微调

用垂类数据优化模型



RAG能力

外挂知识库增强生成



目标

完成从语音到策略的闭环，
奠定超越基础



为用户提供更沉浸式的看播体验



效果 & 收益

营收



付费UA对



评论UA对



04 直播AI的下一步展望

直播AI下一步：交互式主播助手



直播表现回顾



关键指标

随时提取直播数据，精准分析观看数、互动频率，考核直播表现指标。



亮点复盘

总结直播亮点与不足，生成简明的表现汇报。



优化建议

针对表现瓶颈提出针对性的改进建议，提升直播效果。



个性化内容建议



脚本创意

结合现有定制语料方案，引导观众互动氛围与互动。



实时话题推荐

推荐互动话题与当前热点，让直播成为共享信息的窗口。



直播技巧提示

跟随直播间实时节奏，进行生成互动专属提示卡，提升直播节目效果。

直播AI下一步：直播伴侣与对话式推荐



✓ 从被动接受到主动表达

泛意图理解

通过多样化兴趣和需求，提供广泛的推荐，帮助用户发现新的直播体验。

推荐策略：基于行为分析和兴趣拓展，提供多元化内容推荐。

窄意图理解

根据用户明确的需求，精准匹配相关内容，快速满足个性化需求。

推荐策略：通过关键词匹配、过滤条件和个性化推荐，提供精准内容。

✓ 让用户看懂每一次推荐

逻辑透明

通过推荐的历史数据或用户行为，增加用户对推荐系统的信任感。

例：推送时附加信息：“最近常看户外直播，系统为您推荐该内容。”

场景解释

通过结合热点提升推荐的相关性和即时性，让用户感受到推荐的实用性。

例：“北京下雪了，快来看看雪景现场直播吧。”

极客邦科技 2026 年会议规划

促进软件开发及相关领域知识与创新的传播



参会咨询



查看会议

北京

1200人

QCon

全球软件开发大会

会议时间：4月16-18日

- Agentic Engineering
- AgentOps
- 下一代模型架构与推理优化
- AI 原生基础设施
- 知识工程实践
- AI 安全

深圳

1000人

AiCon

全球人工智能开发与应用大会

会议时间：8月21-22日

- Agentic AI
- 轻量化与高效推理
- 多模态应用
- AI + IoT 场景实践
- AI 工业化落地

北京

1000人

AiCon

全球人工智能开发与应用大会

会议时间：12月18-19日

- 大模型架构创新
- 多模态 AI 产业融合
- 具身智能
- AI for Science
- 大模型安全

4月

6月

8月

10月

12月

AiCon

全球人工智能开发与应用大会

会议时间：6月26-27日

- AI Infra 系统工程
- 多 Agent 协作与实践
- 多模态融合
- 模型训练与推理创新
- 数据平台与特征服务

上海

1000人

QCon

全球软件开发大会

会议时间：10月22-24日

- AI Agent
- Vibe Coding
- 智能可观测
- 推理基建
- 模型攻防
- AI x 创造力

上海

1200人

THANKS

探索 AI 应用边界

Explore the limits of AI applications

AiCon

全球人工智能开发与应用大会

