

腾讯云智能语音行业落地 探索与实践

倪捷

腾讯云高级产品经理

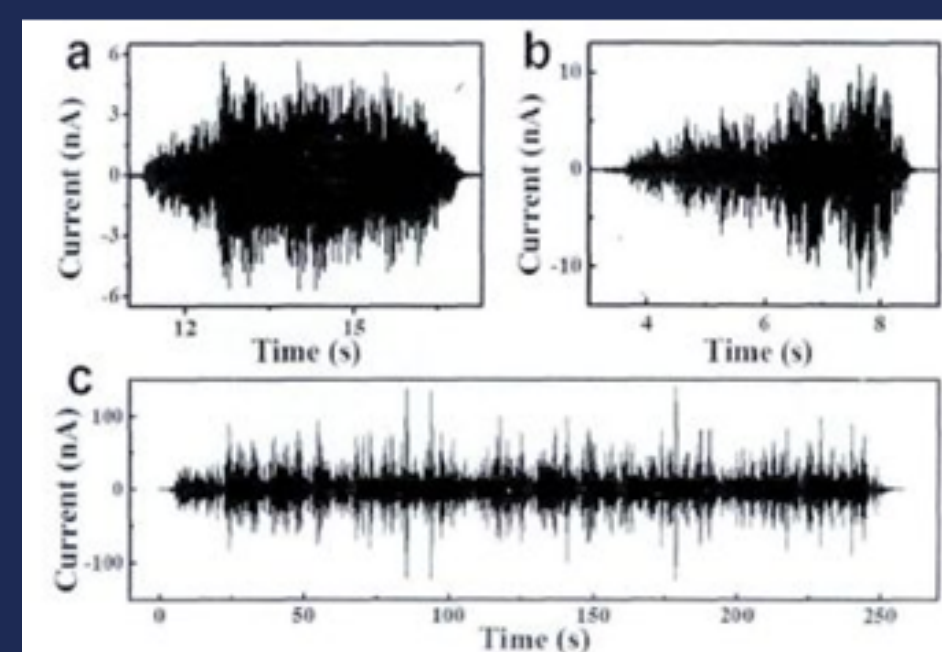
目录

- 智能语音技术发展情况及应用中面临的挑战
- 典型落地场景下如何优化语音技术的体验
- 腾讯云智能语音落地案例分享

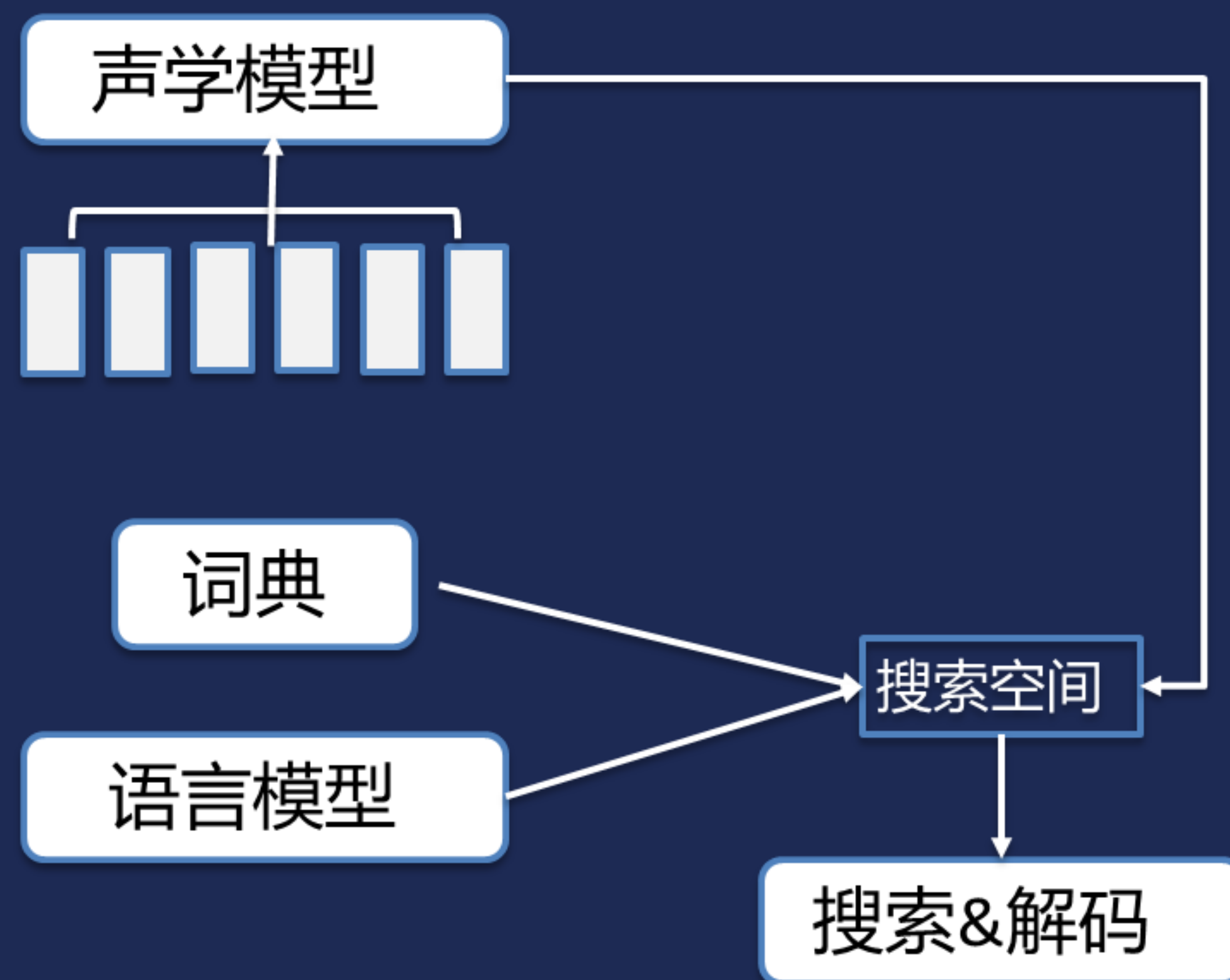
人类系统性学习中文的入门方法



语音识别基础原理 (ASR)



特征提取



业界语音识别水平

- 背景安静
- 近场
- 普通话标准
- 朗读

- 口语化
- 轻度口音

- 背景嘈杂
- 远离拾音器
- 严重口音

语音识别面临的挑战



- ✓口语化：不同情绪下语速、声调不同，吞字、结巴会造成音频质量参差不齐。
- ✓高质量收音：在多人、声音嘈杂的场景下，如何“听清楚”。



- ✓噪声影响：例如车载环境下会有回声以及各种噪声源。
- ✓远场识别：麦克风和声源距离过远时收音识别效果较差。



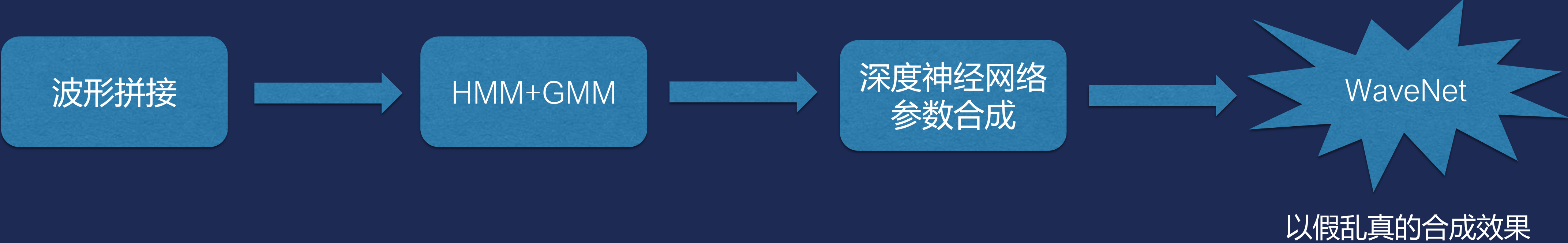
- ✓专业领域识别：导航、办公、旅游、美食等不同场景有语言模型。
- ✓方言口音：全球有上千种语言，中国的方言种类也非常多。

语音合成 (TTS)

- 把文字转化成拟人化的声音
- 完成人机语音交互的闭环
- 满足多种场景对不同声音的需求



语音合成技术演进



	MOS
传统声码器（World等）	4.2+
WaveNet	4.4+
真人录音	4.5+

样例1

样例2

语音合成面临的挑战



- ✓拟真度：发音准确性，流畅度，抑扬顿挫等一系列因素决定合成效果与真人差距
- ✓主观性：对声音好坏的判断有一定个人主观性，难有客观的标准
- ✓声音的适应性：有的声音适合读物朗读，有的声音适合做语音助手
- ✓多音字，特殊读法



- ✓语音定制化：声音也是品牌的代表，追求品牌的公司不愿意共享声音
- ✓录音时长：合成要求一定量的高质量录音，会带来录音时间及其它成本

目录

- 智能语音技术发展情况及应用中面临的挑战
- 典型落地场景下如何优化语音技术的体验
- 腾讯云智能语音落地案例分享

语音输入



通用引擎是基础，场景化引擎进一步提高准确率

录音转写（人与人交互）

需求：

- 判断服务的质量，甚至责任判定

痛点：

- 人工处理速度慢，抽检覆盖率低
- 标准不一：无法避免个人主观性影响
- 成本高：需要大量人力，且需要定期培训、抽查

难点：

- 人与人交互口语化程度高
- 背景噪声不可控



语音质检

提高录音转写准确率的方法

选择正确的引擎参数

- ✓ 采样率：8K or 16K
- ✓ 单声道 or 双声道

尽可能提升录音质量

- ✓ 降低背景噪音
- ✓ 减少无关声音的输入
- ✓ 使说话人不要远离麦克风
- ✓ 不要采用大压缩比的音频格式

定制化提升

- ✓ 语料训练
- ✓ 热词

客服机器人



公众号、服务号



电话客服



网页客服



低门槛和7X24
带来海量用户请求



80%的客户问题重复



机器人客服

基于人工智能的客服系统逐渐变为企业的竞争力

新时代数字化服务的要求

- 更好的人机交互体验
- 更多样的人机交互渠道
- 更丰富的智能化功能

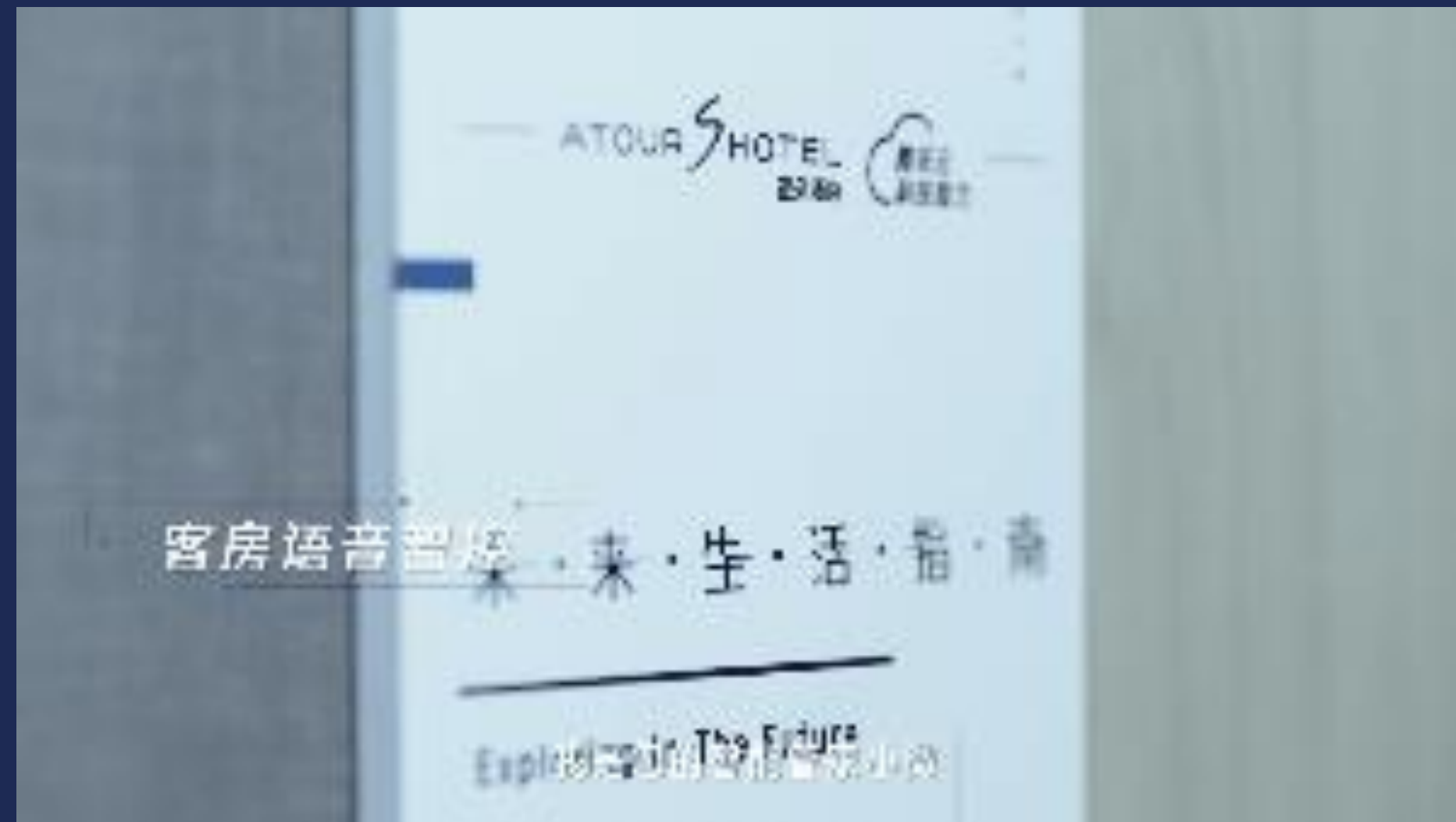


语音是人类最自然的沟通方式，语音技术与机器人能力的结合会带来全新的服务体验

目录

- 智能语音技术发展情况及应用中面临的挑战
- 典型落地场景下如何优化语音技术的体验
- 腾讯云智能语音落地案例分享

智慧酒店——亚朵酒店

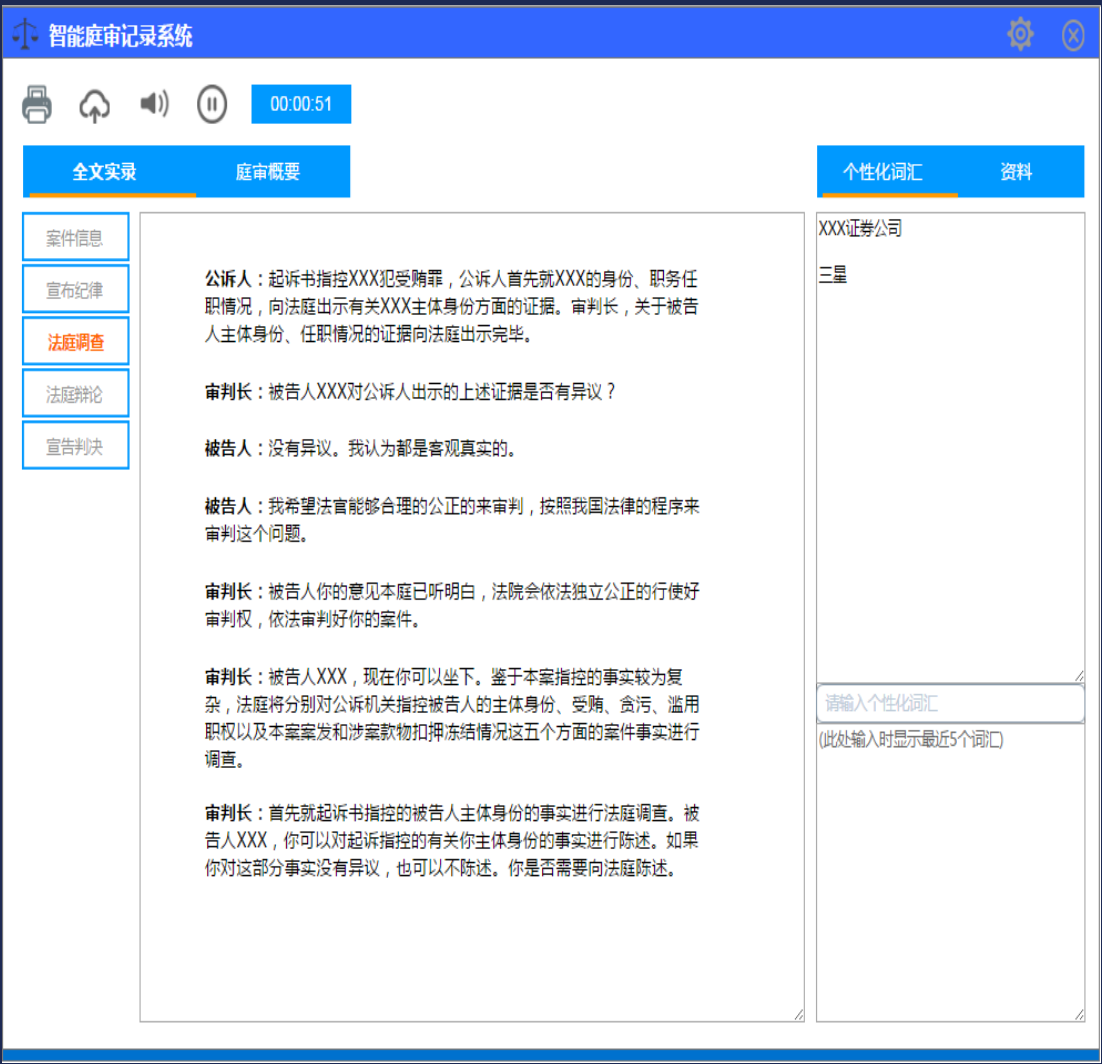
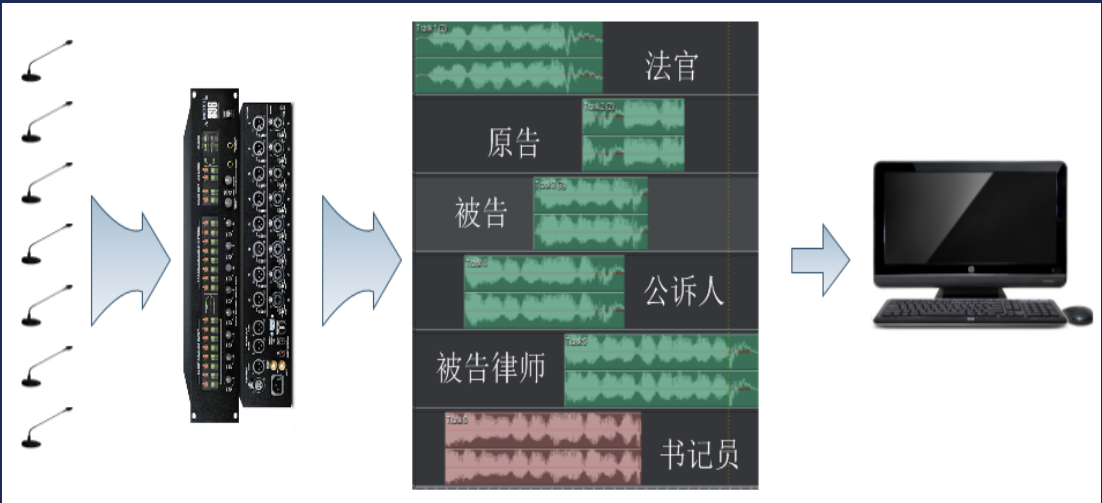


- 只需要接入相关设备，他可以听得懂、看得见。一个专门适合懒人的黑科技，只需要张开嘴轻声呼唤"Hi，小微"。
- 窗帘、音箱、灯光、空调都在你的掌握之中。
- 除此之外，它还可以查询天气、交通路线、新闻资讯、查询机票等。

其它案例



音频内容审核



法庭记录



同传



Thanks For Watching

对话系统101

