

未来已来-机器人时代触手可及

讯飞开放平台 李繁

IT产业五次浪潮：人机交互引领产业变革



—来源《摩根斯坦利》

万物互联时代



无屏



远场



移动



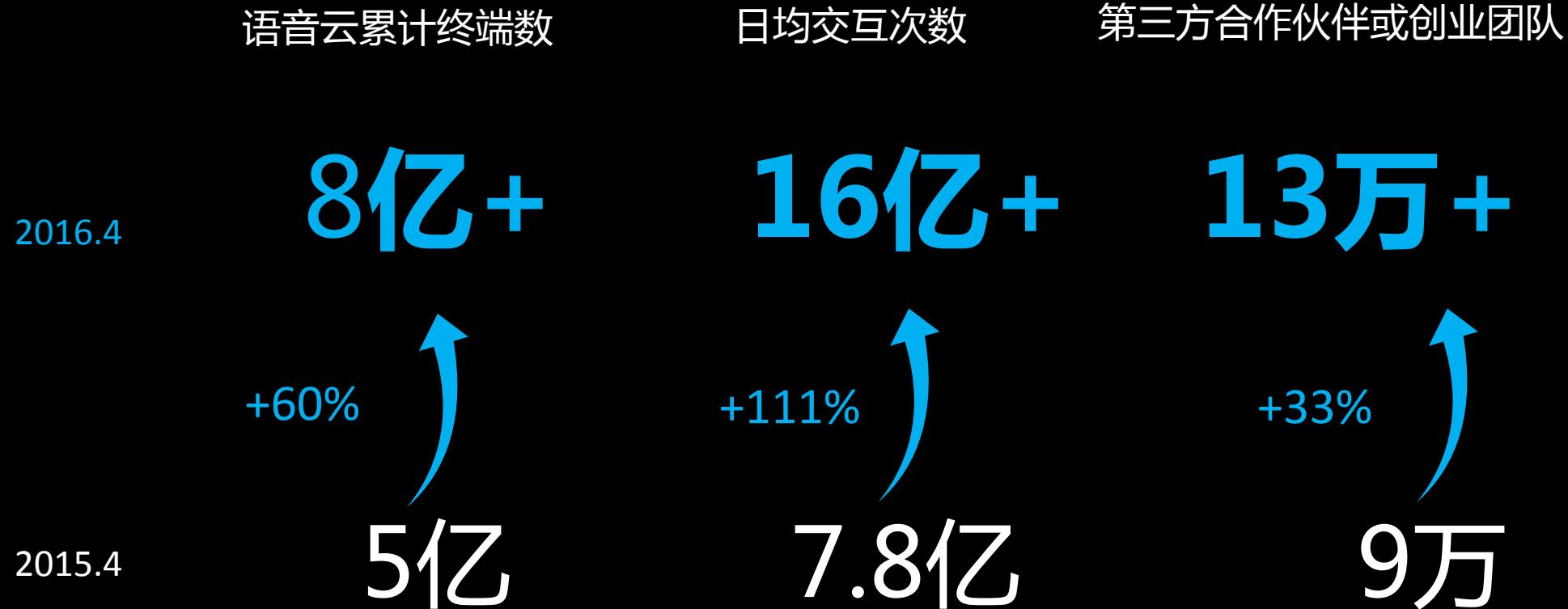
在无屏、移动、远场状态下

以语音为主，键盘、触摸等为辅的人机交互时代正在到来

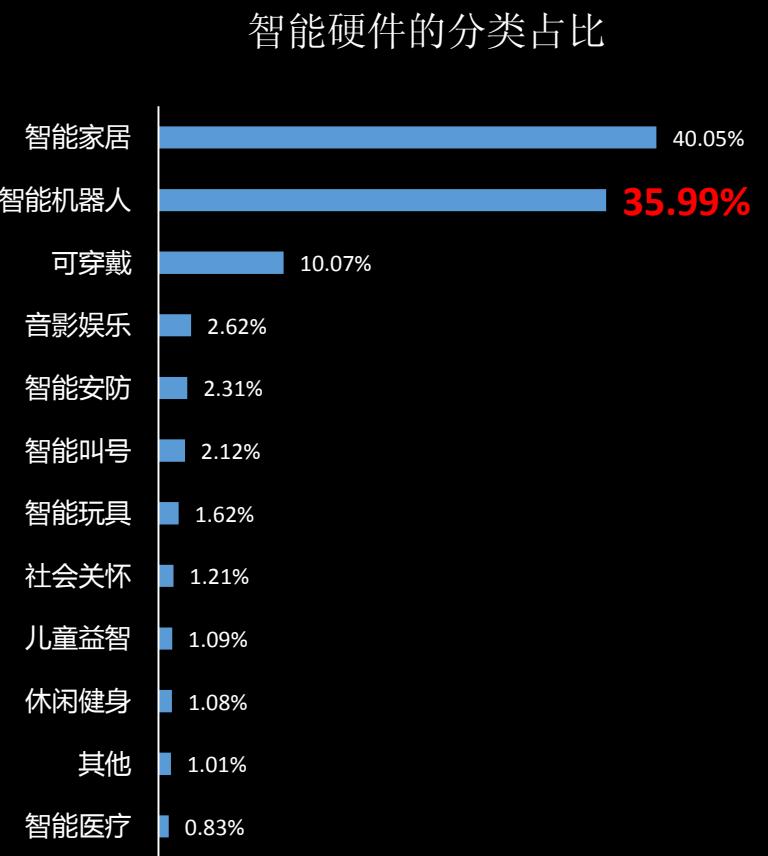
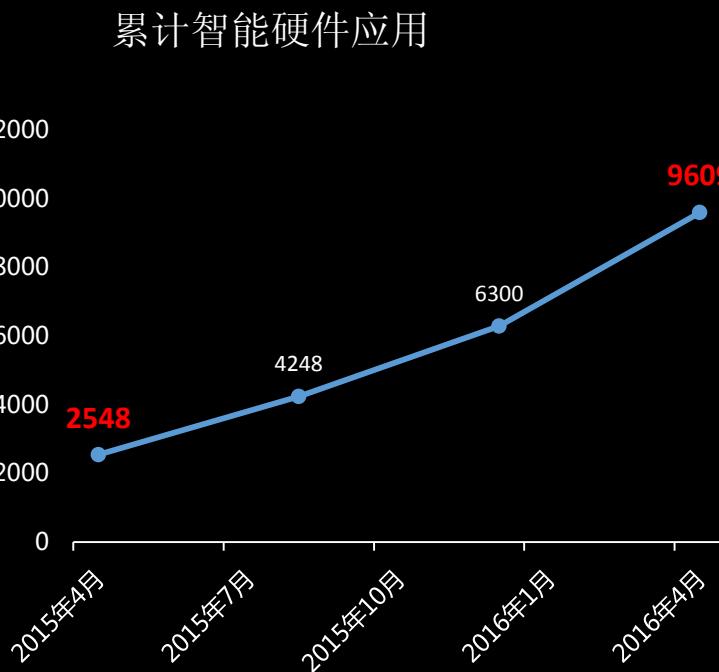
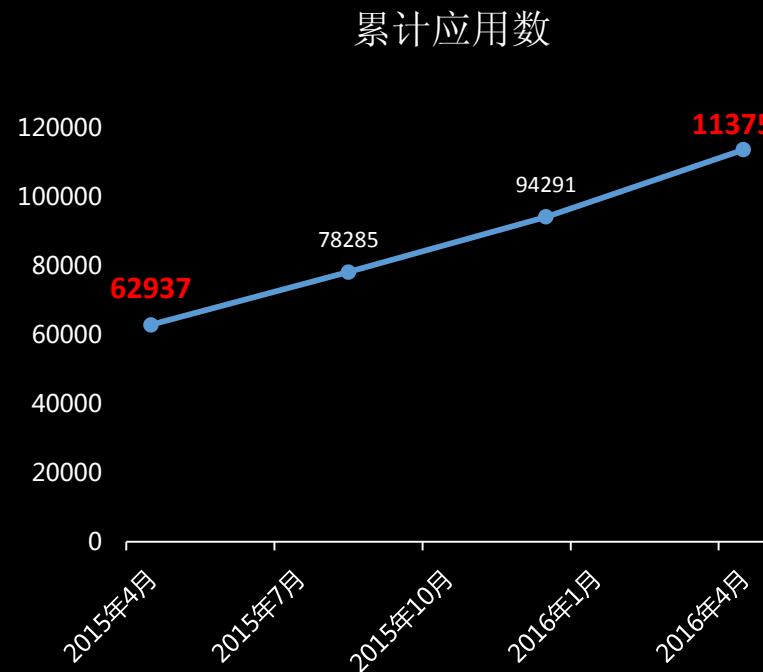
讯飞开放平台开放业界领先的智能人机交互服务



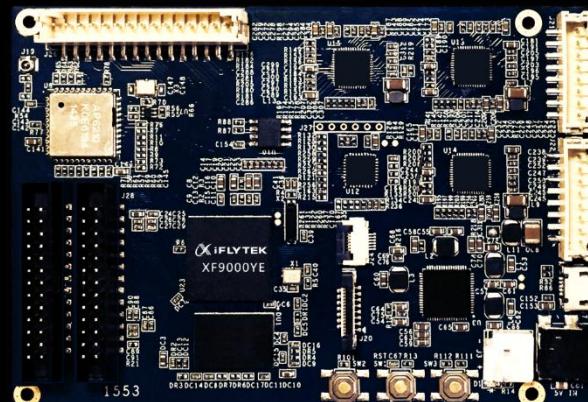
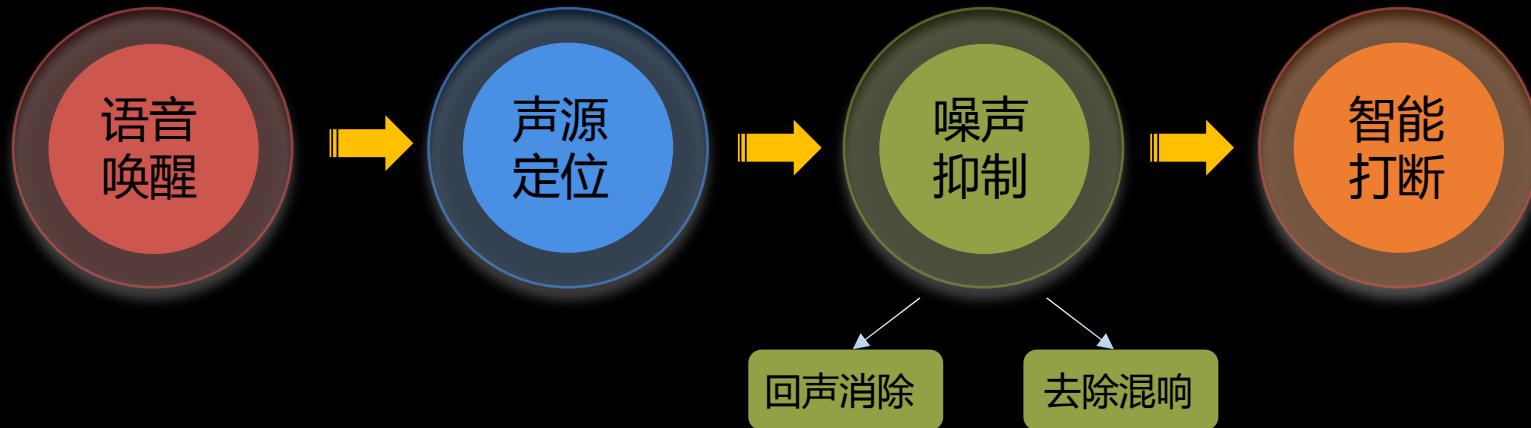
开放平台带动合作伙伴迅猛增长



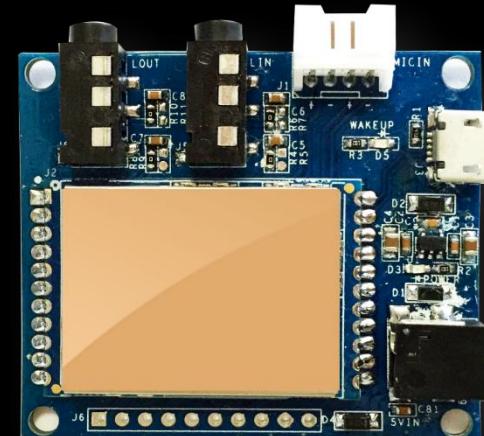
从人工智能开放平台看智能硬件的迅猛增长



麦克风阵列 - 远距离拾音



XFM10411



XFM10211

语音唤醒 – 全面升级

低功耗

高唤醒率

低误唤醒率

小尺寸

95%

10次 / 天

典型
应用



小鱼在家
小鱼帮帮忙

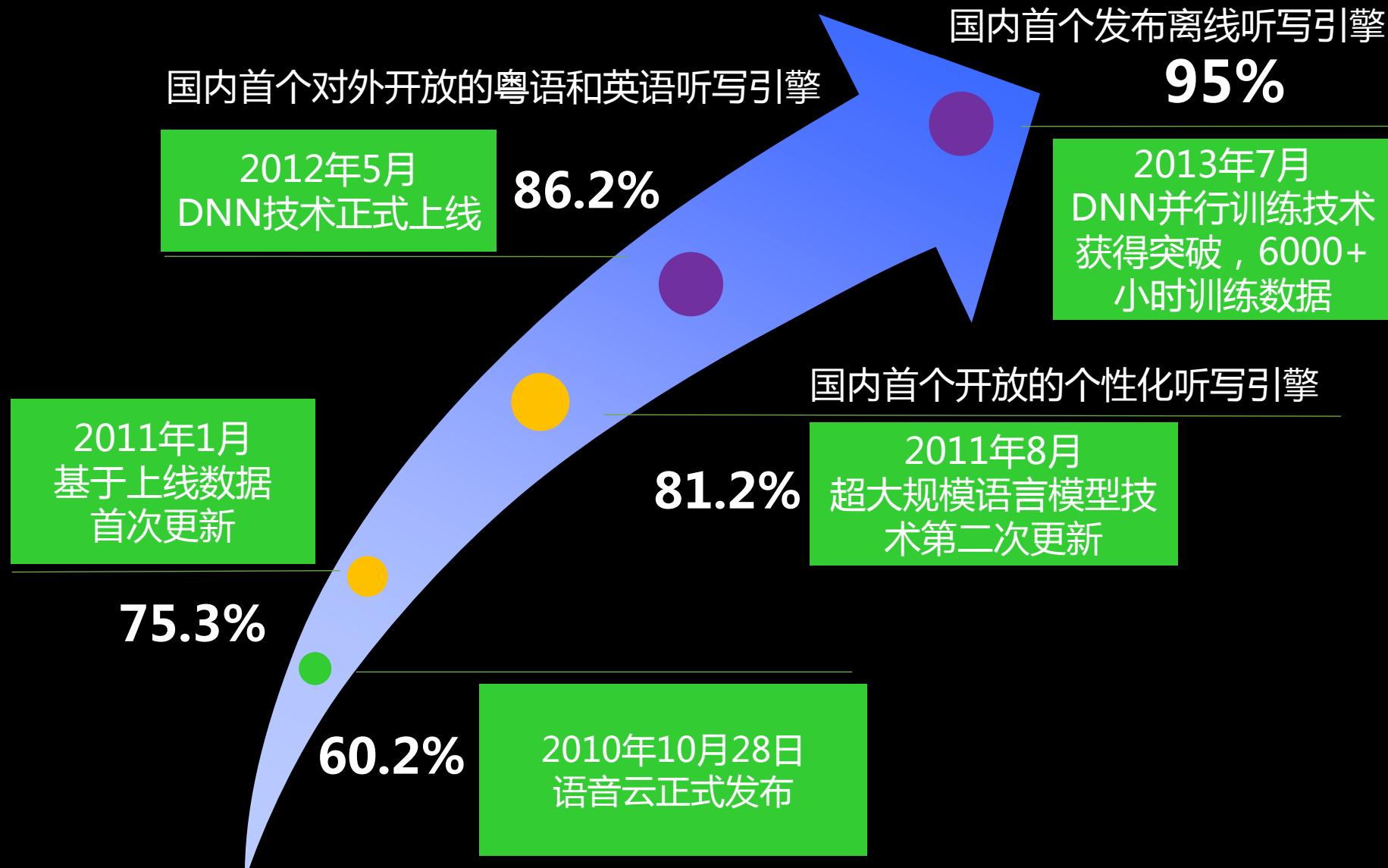


叮咚叮咚



小白小白

语音识别



语音识别 – 多方言识别

普通话识别准确率突破95%，并开放多达十八种方言



语音识别 – 应用级个性化



“琅琊榜”

“海宴的琅琊榜”

“海宴琅琊榜中的麒麟才子”

编写语法
/热词

在线编译

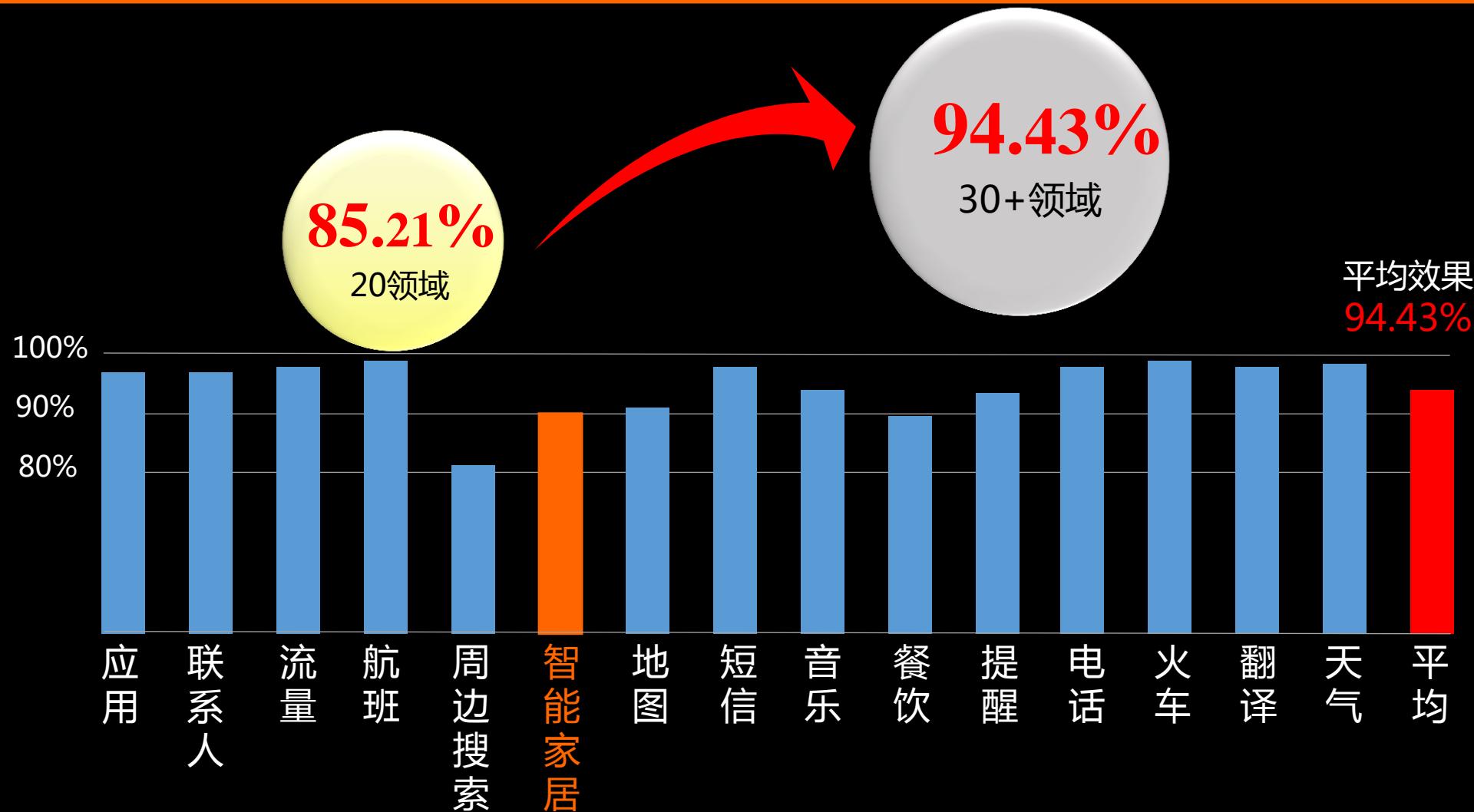
发布上线

语义理解 – 智能家居



语义理解 - 全面覆盖

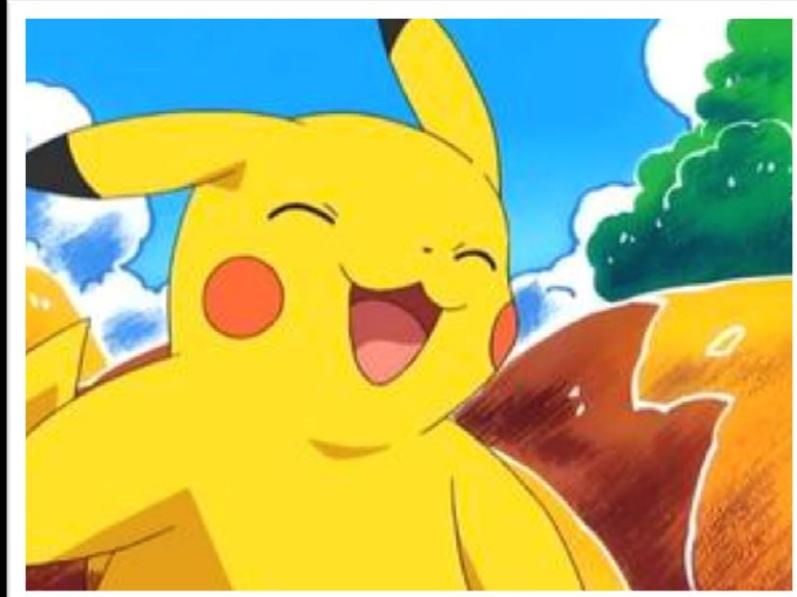
通过语音云海量数据迭代，常用业务语义理解效果持续提升



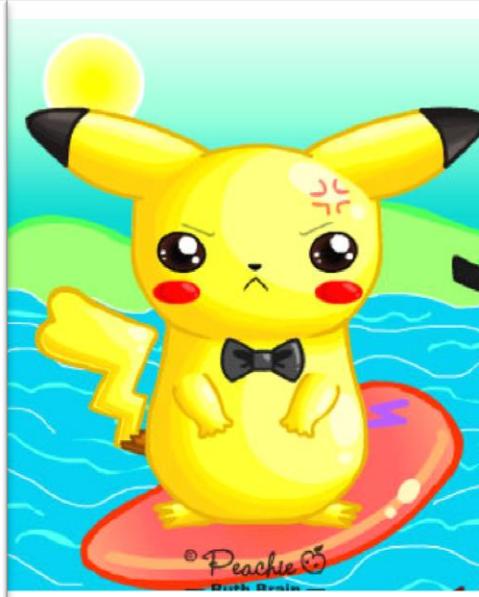
语音合成 – 惟妙惟肖



语音合成 - 情感音库



你好可爱呦



别生气了嘛



光棍节怎么过



身份验证 - 人脸 + 声纹



- 混合验证，多维度降低误判断
- 家庭场景，聚焦鉴定范围，提升准确判断
- 云 + 端一体化，快速稳定



声纹鉴定



家庭场景



人脸验证



麦克风阵列



语音唤醒



身份验证



身份 : 女主



语音合成

```
"text" : 我要看西游记  
"service" : video  
"operation" : play  
"semantic" : {  
    "slot" : {  
        "keywords" : 西游记  
    }  
}
```



语音识别

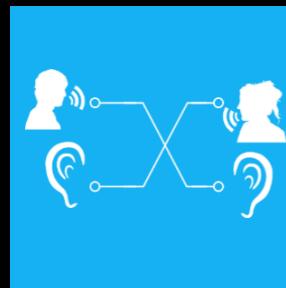
自然人机交互新需求



远场降噪



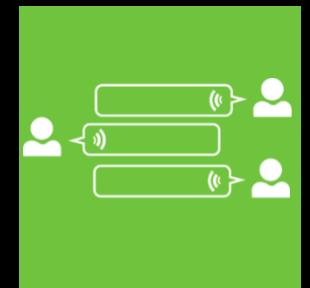
方言识别



全双工



纠错



多轮对话



AIUI

人工智能时代的人机交互界面

AIUI方案特性

- 从单点到方案，进一步降低开发门槛，开发者集成简单便捷
- 麦克风阵列，实现远场拾音
- 语音唤醒自动闭环优化系统
- 多轮对话
- 全双工
- 对话场景自适应
- 全新的通信协议

AIUI方案特性

1 从单点到方案，进一步降低开发门槛，开发者集成简单便捷

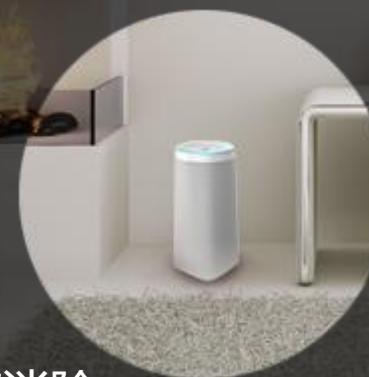


AIUI方案特性

2 麦克风阵列，实现远场拾音

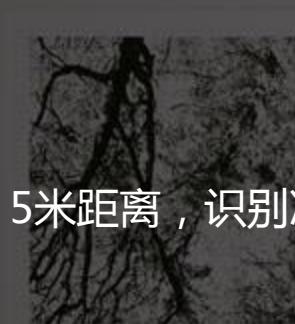
噪声抑制

家居85dB强噪音下，识别准确率依然达到90%



回声消除

喇叭95dB高音量下，唤醒率达到92%



远距离

5米距离，识别准确率达95%



声源定位

360度全范围定位，精度达到5度



AIUI方案特性

3 语音唤醒自动闭环优化系统

芝麻开门))))

说出个性定制唤醒词，
应答听命，直接进入

小尺寸
Sony L39h后台持续侦听，每
隔0.5h唤醒一次，耗电量低于
总电量的1%/h

低功耗
ROM小于40K

高唤醒率
唤醒率总体超过
95%，业界绝对领先

个性化
定制
自定义设置多个唤醒词，
满足个性需求



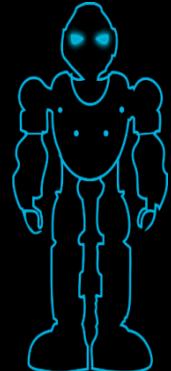
AIUI方案特性

4 多轮对话

- 任务的**复杂**处理为多轮简单问答
- 基于上下文的任务关系，进行任务切换和指代消减
- 基于多轮中的纠错



帮我导航到附近的银行



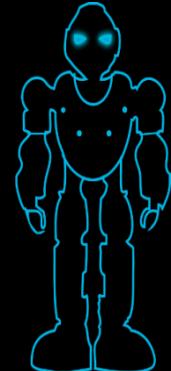
是要去ATM还是有人工
柜台的？



是有人工柜台的



我想听刘德华的歌



我想看他的电影



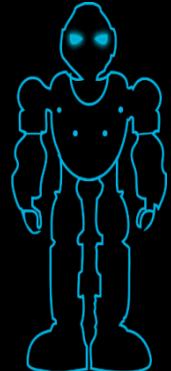
他的富春山居图怎么样？



好



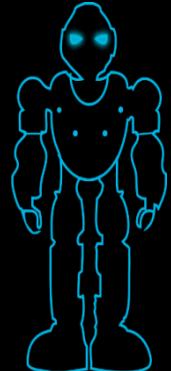
我叫张姗姗



好的，张姗姗



是王字旁的珊，
不是女子旁的姗



不好意思，这次记住你
啦，张珊珊

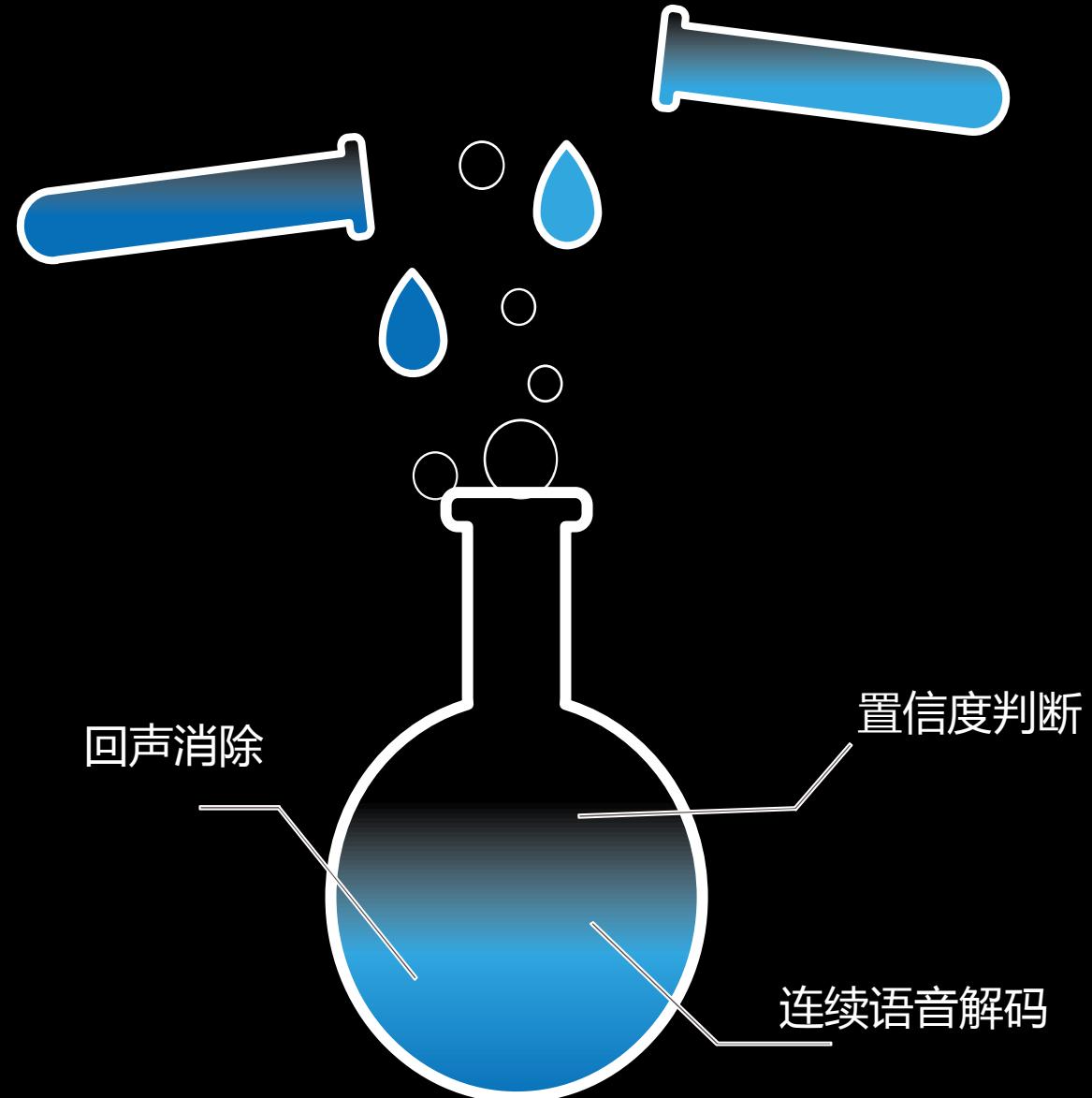
AIUI方案特性

5 全双工

- 机器的反馈和监听，并行同步
- 用户可以在任一时间打断，发出需求
- 机器对用户非相关语音，进行拒识

一问一答
Before

全双工
Now





请进。主上次的酒店吧
添加个乘机人胡小郁。

机器人无反馈，
她知道说话的对象不是自己。

敲上声文理解“刚刚”

持续聆听是弱提问长时追问话打断



行，没问题，那就好，那我还是订国家 那就是12月19号
北京，乘机人是刘大



AIUI方案特性

6 对话场景自适应



- 基于语义的端点监测

支持用户语音交互的短时间停顿，包括语气词的插入，如“嗯”、“啊”

支持智能断句，基于业务场景自动忽略无效语音



AIUI方案特性

7

全新的通信协议

- 少量语音识别容错理解
引入序列标注模型解决集外词

“导航到水墨 **亭**(庭)”

- 大幅降低业务优化的复杂性
技术手段从文法转变为统计模型

基于**CNN**的排序学习的理解模型

- 对话理解的准确率和召回率

对话理解的平均正确率达到**80%+**
重要业务达到**90%+**

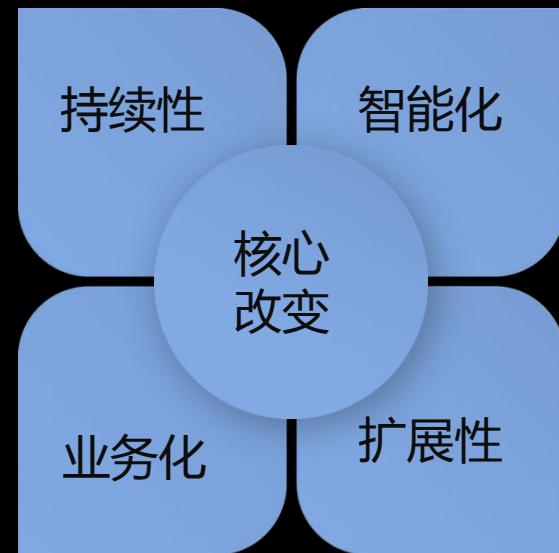
交互传输

鲁棒
稳健



AIUI，核心改变

- 交互过程视为一个持续的会话流，没有明确的分片边界
- 任意时刻的交互处理，都和历史互动相关联
- 交互和业务内容的融合更加紧密
- 进一步的交互会依据业务响应数据和历史互动来决策



- 应用层只定义交互场景边界
- 所需要的服务能力及特性，由交互过程来自主决策
- 从入口传入的音频、图片、文本、动作等信息被视为一致的数据流来收集和汇聚
- 更多的能力服务易于扩展/对接到AIUI框架中来，并能够与原有的输入/输出信息相适配。

AIUI，快速获取



软硬一体化



云端一体化



技术服务一体化

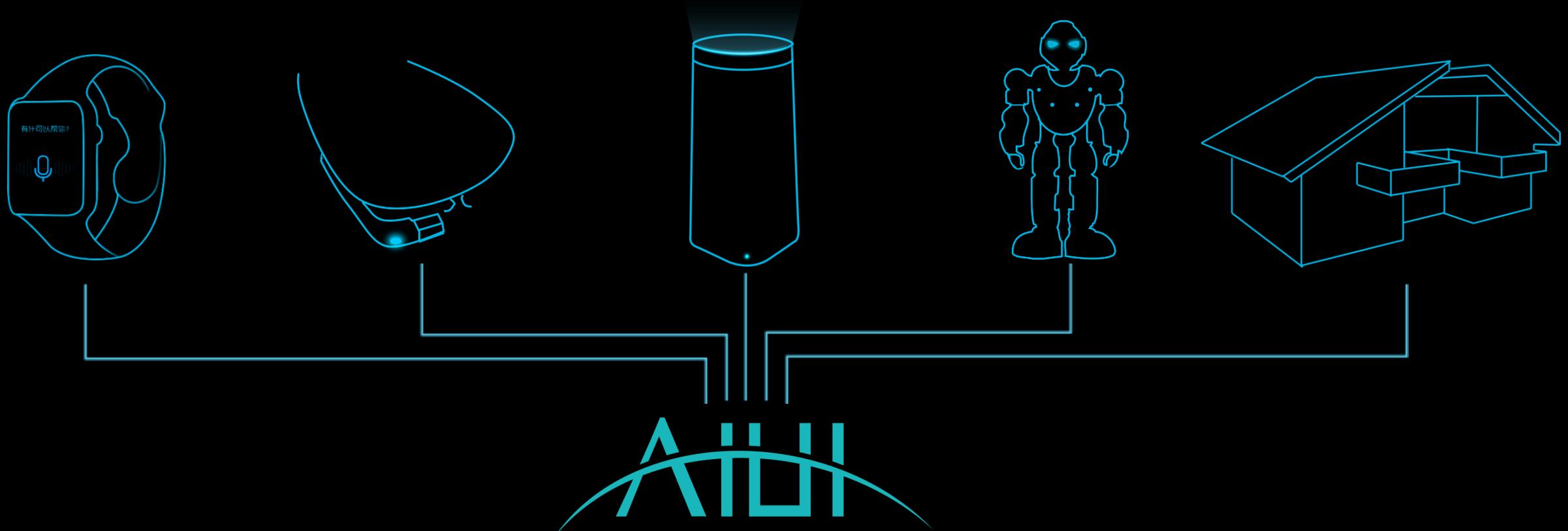
线形麦克风 | 环形麦克风 | 球形麦克风
两麦，四麦 五麦，八麦 十五麦

MSC Mini | MSC | 语音+
ROM : 120k, 通用适配 便捷集成
RAM : 50k

远场全双工 | 生物特征 | 云端服务&
交互 身份验证 差异化定制

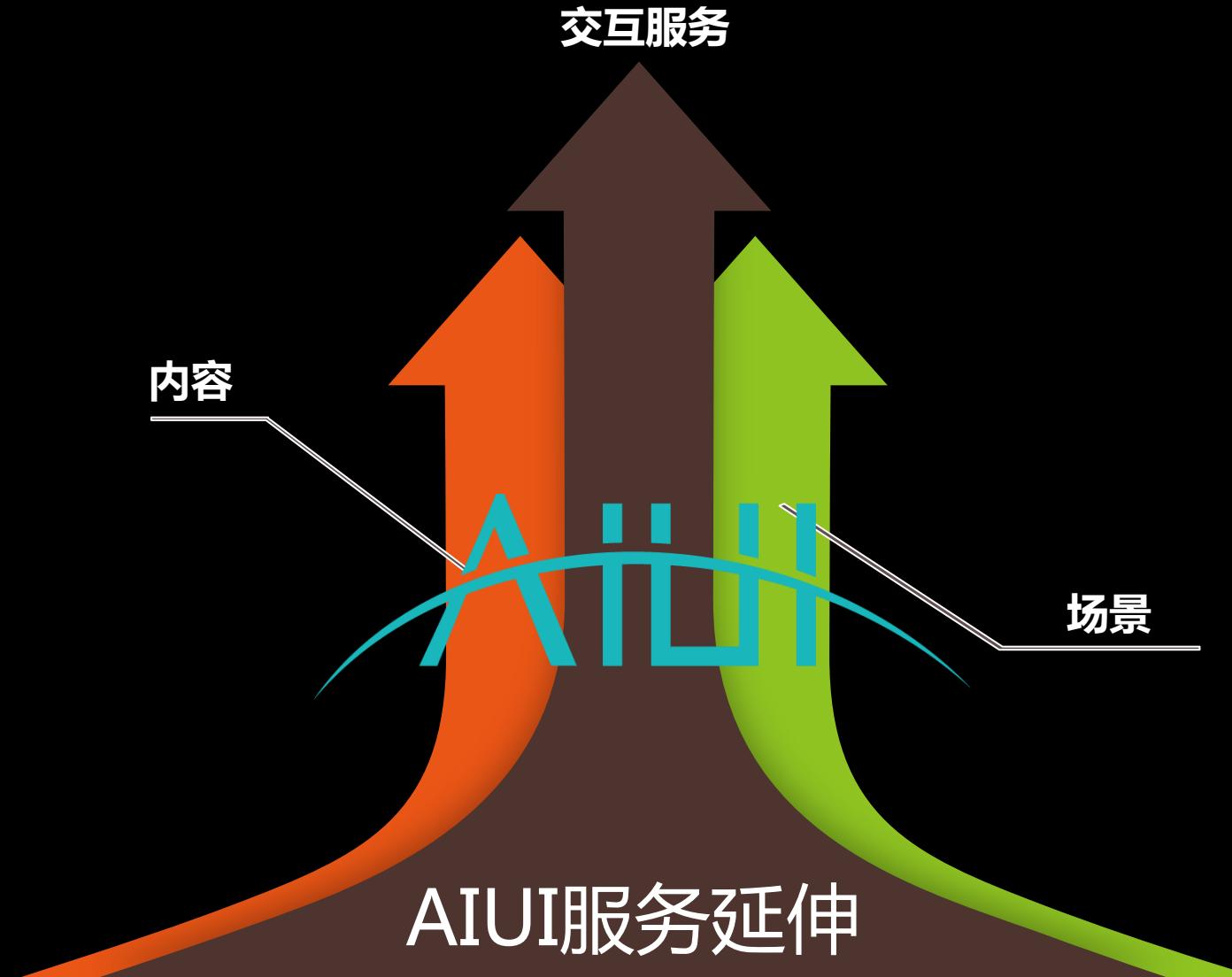
AIUI标准（人机智能交互标准）

AIUI，应用领域



AIUI统一服务平台

AIUI，未来展望



典型合作伙伴



优必选-阿尔法机器人
人形机器人



狗尾草科技-公子小白
情侣社交机器人

典型合作伙伴



小鱼儿科技-小鱼在家
家庭陪伴机器人

新一代儿童机器玩伴

16/32 GB 大内存

红外超声波重力触摸 感应器

1300万超高清 摄像头

自动充电 功能



锐曼智能-小曼机器人
儿童玩伴机器人

THANK

讯飞开放平台

www.xfyun.cn

