

未来已来

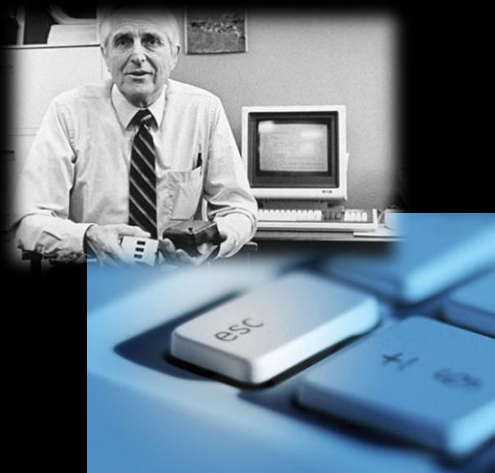
机器人时代触手可及

讯飞开放平台

李繁

人机交互的发展历程

PC时代



键盘+鼠标

移动互联时代



多点触摸，
手写，手势

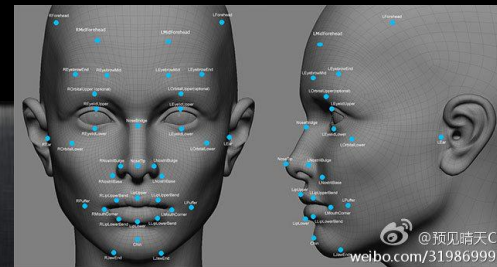


语音

+

视觉

智能生活时代

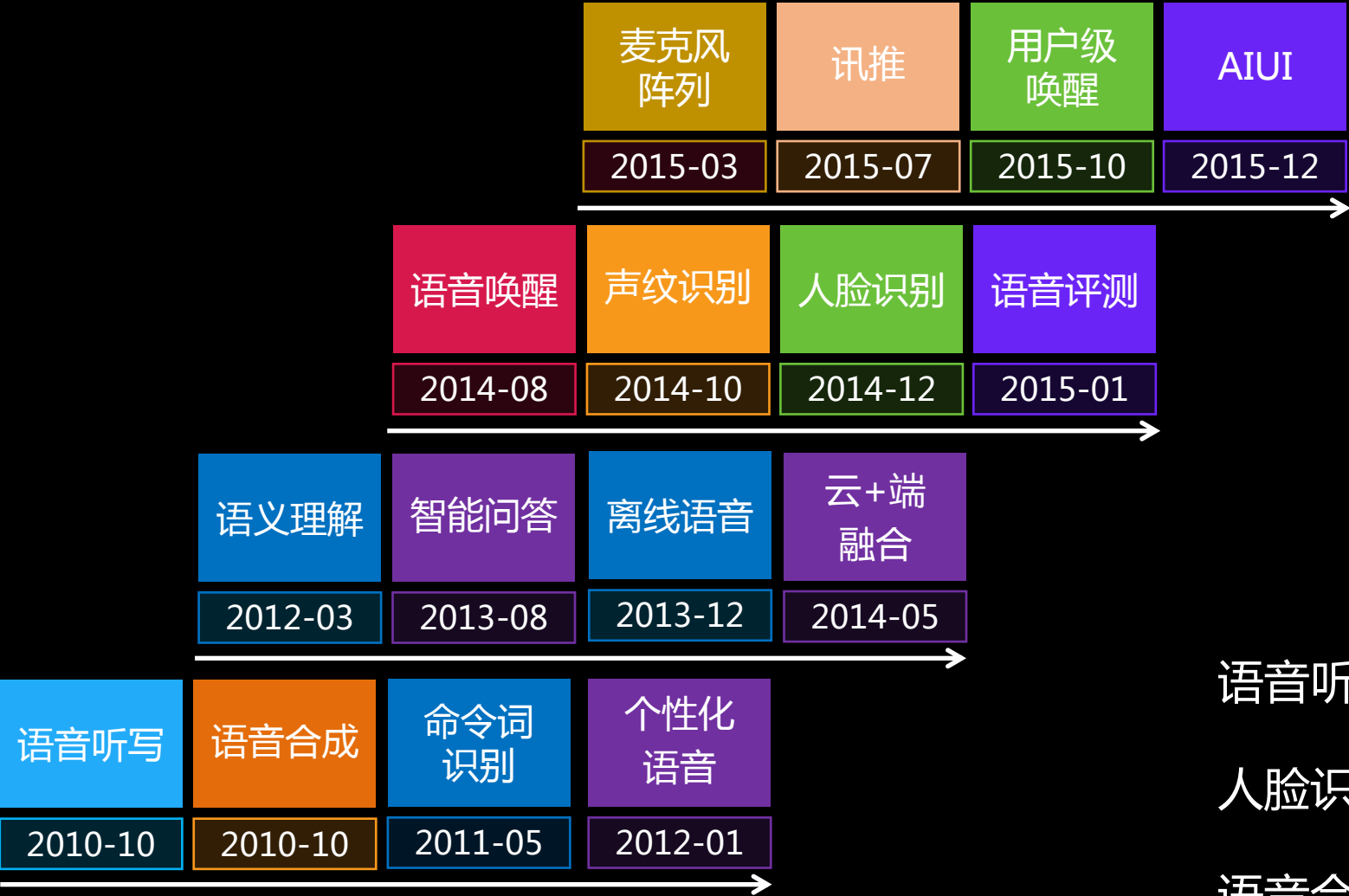


交互的发展史就是走向自然交互的发展过程

从以机器为中心的人机交互，走向以人为中心的自然交互

万物互联浪潮下，语音为主，
键盘、触摸、图像、手势为辅的人机交互时代已经到来

讯飞开放平台开放业界领先的智能人机交互服务

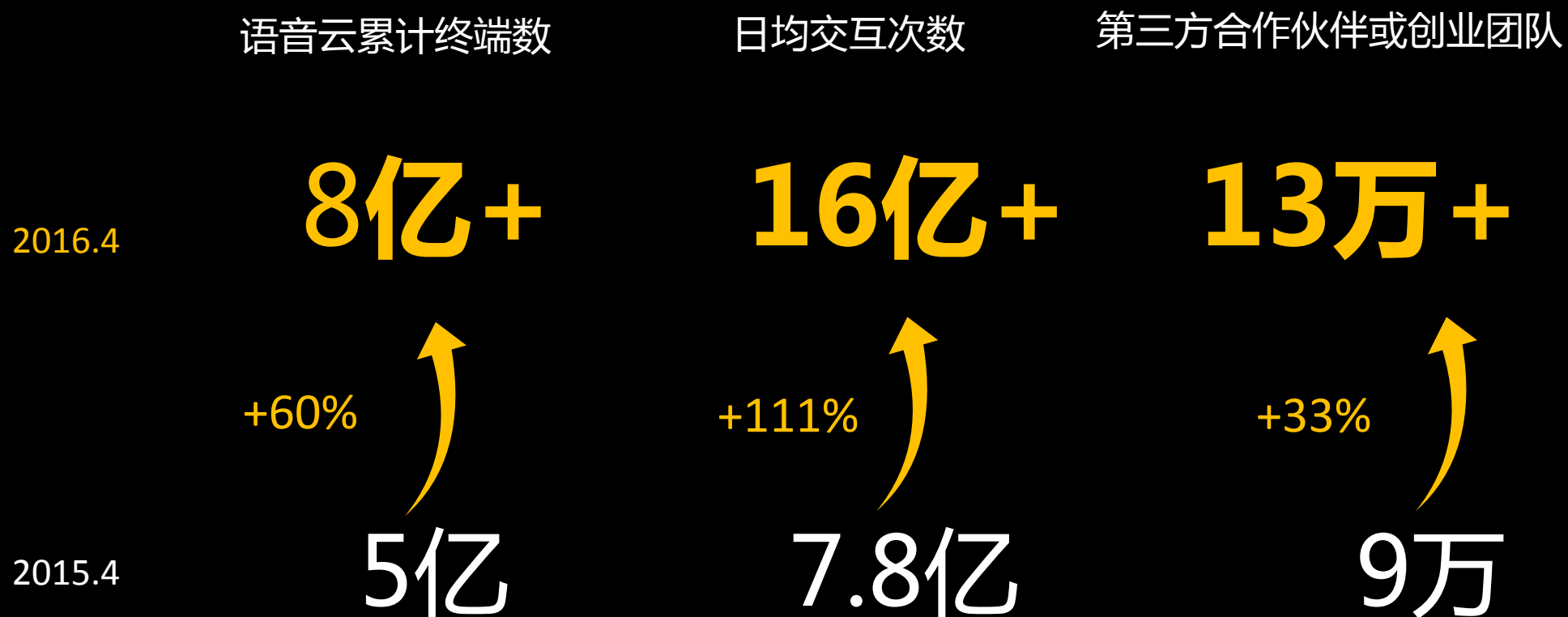


语音听写 95%

人脸识别 99%

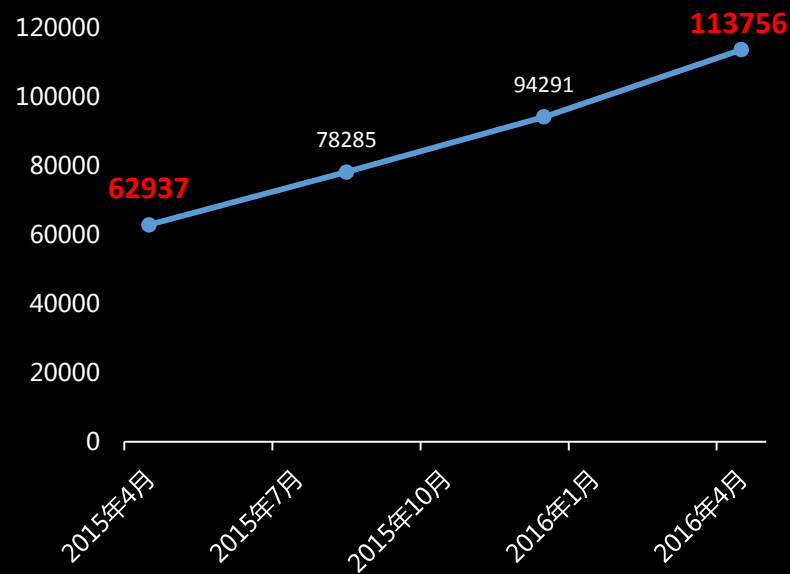
语音合成 自然度超过普通人

开放平台带动合作伙伴迅猛增长

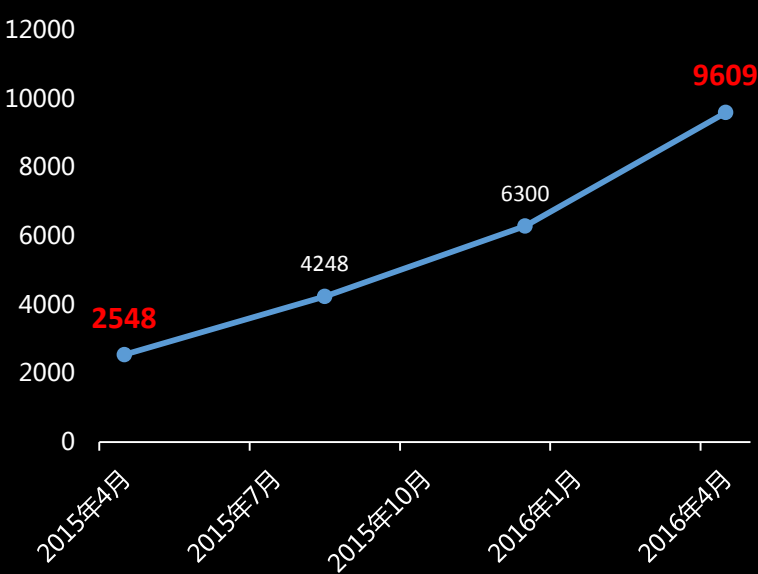


从人工智能开放平台看智能硬件的迅猛增长

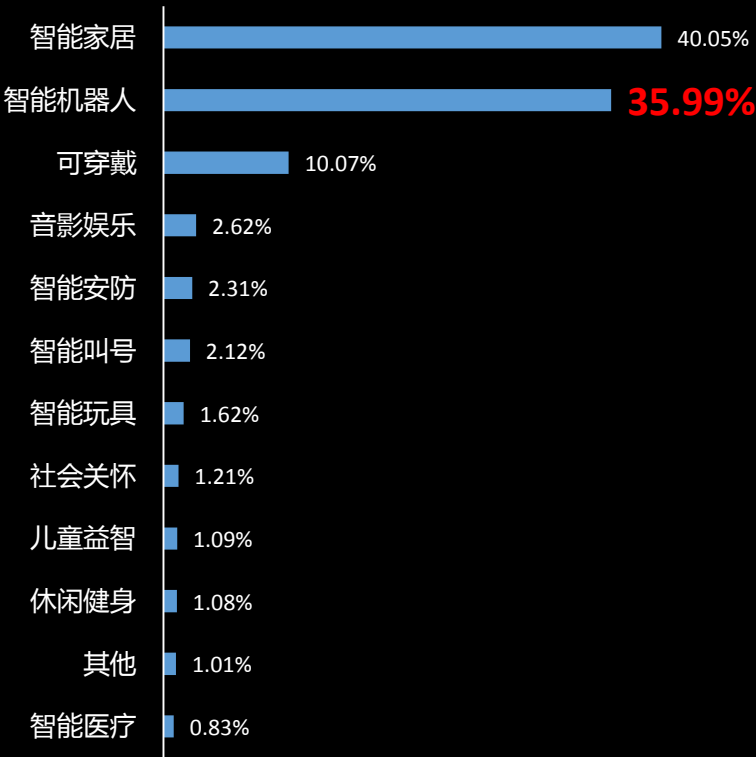
累计应用数



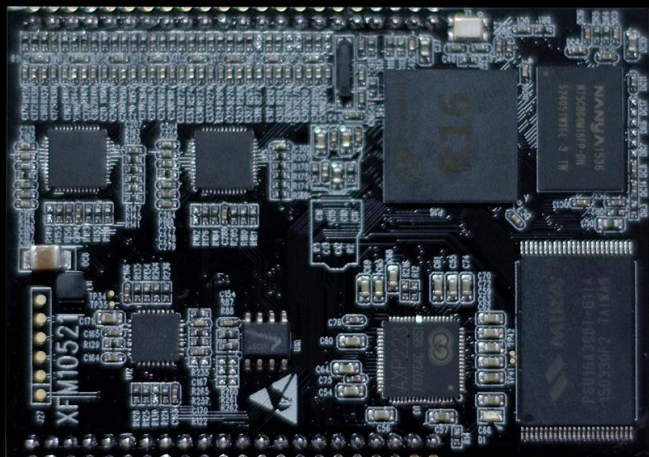
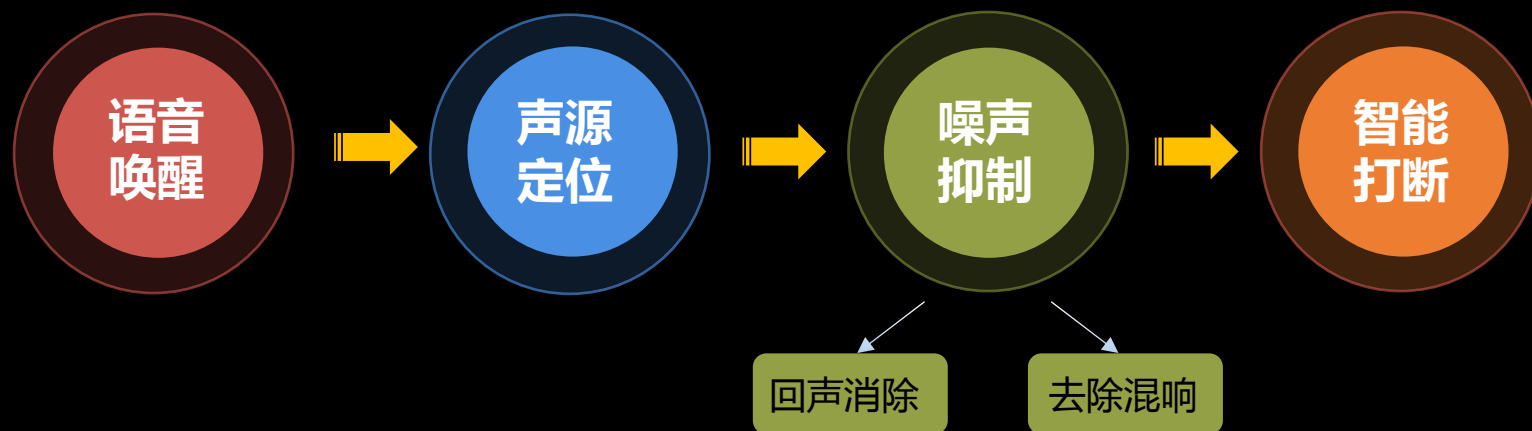
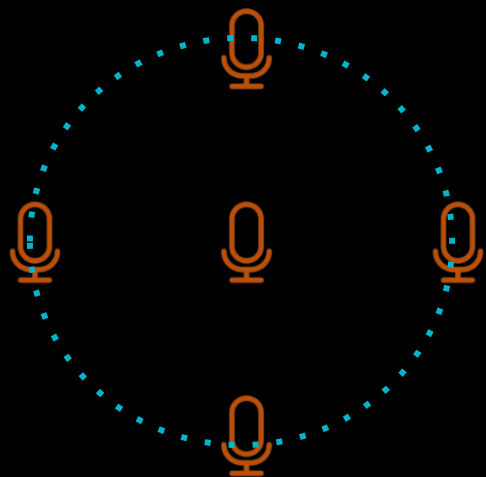
累计智能硬件应用



智能硬件的分类占比



麦克风阵列 – 远距离拾音



- ✓ 拾音距离**5米**
- ✓ 国际领先的回声消除技术，消除量可达**50dB**
- ✓ 稳态噪声抑制
- ✓ 动态噪声抑制
- ✓ **360°**声源定位
- ✓ **±10°**声源定位精度
- ✓ **5m**距离语音听写识别率突破**90%**

语音唤醒 – 全面升级



典型
应用



小鱼在家

小鱼帮帮忙



叮咚叮咚



小白小白

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

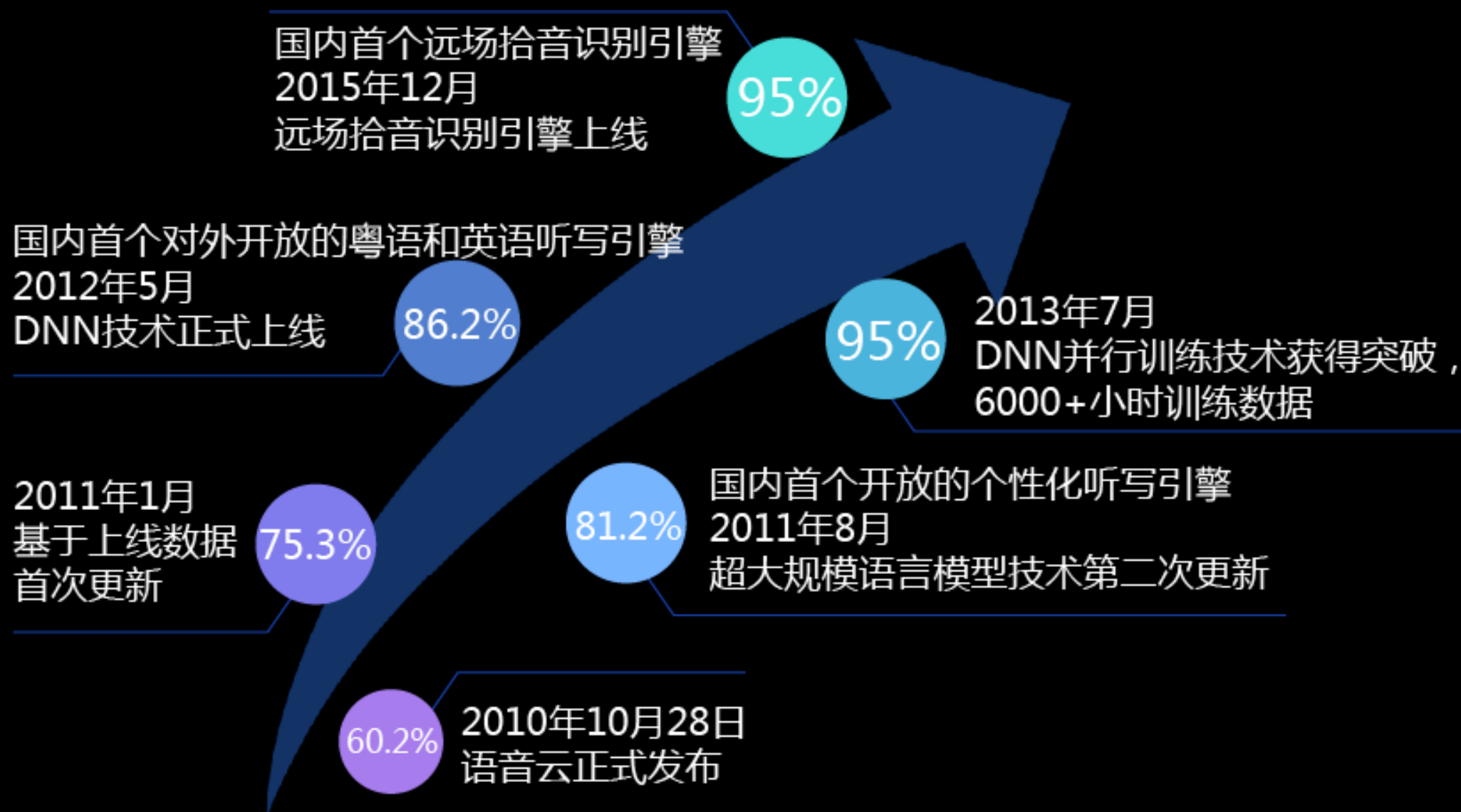


家庭场景



人脸验证

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



“芈月传”

“芈八子”

“芈姝嫁给秦王嬴驷”

编写语法
/热词



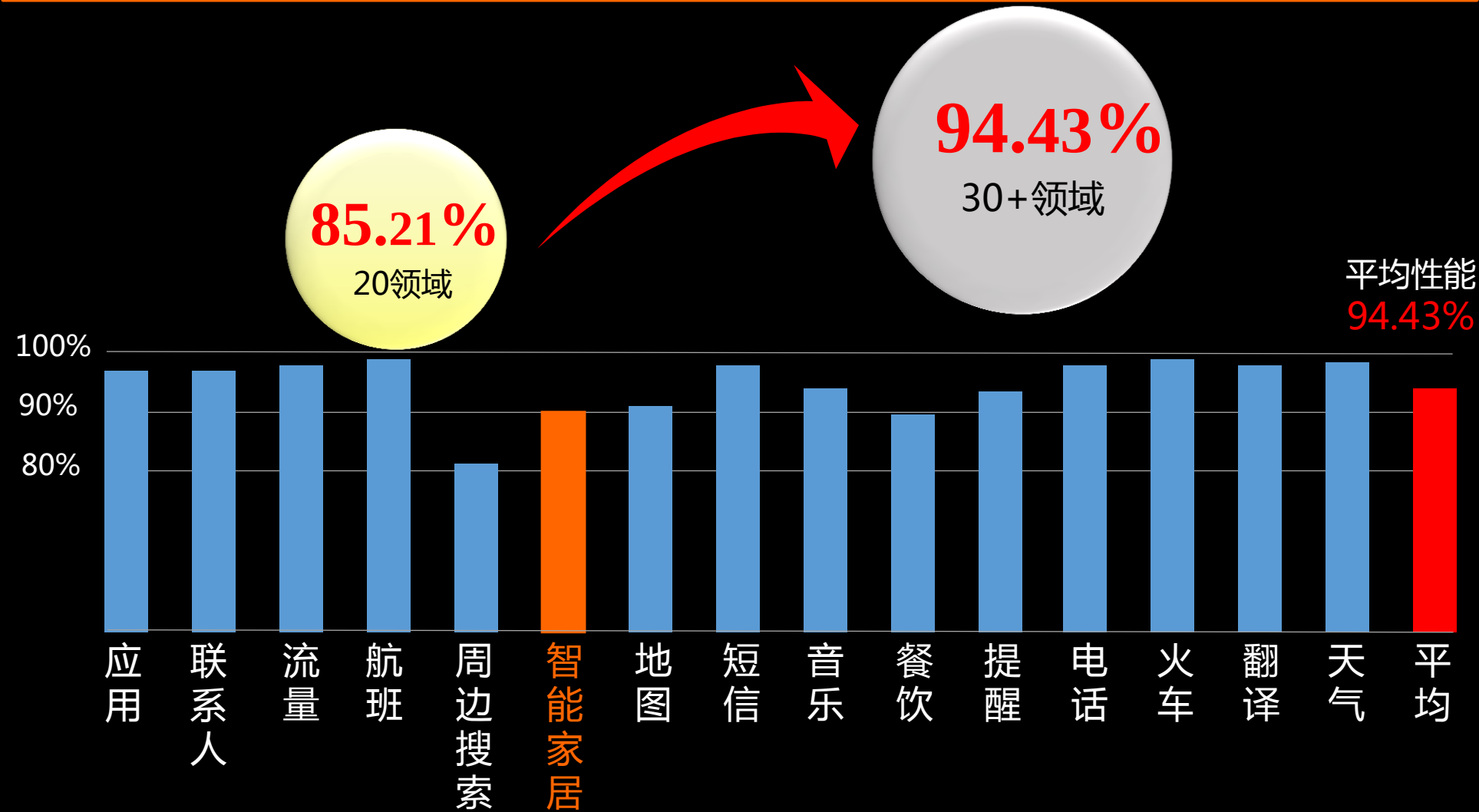
在线编译



发布上线

语义理解 - 全面覆盖

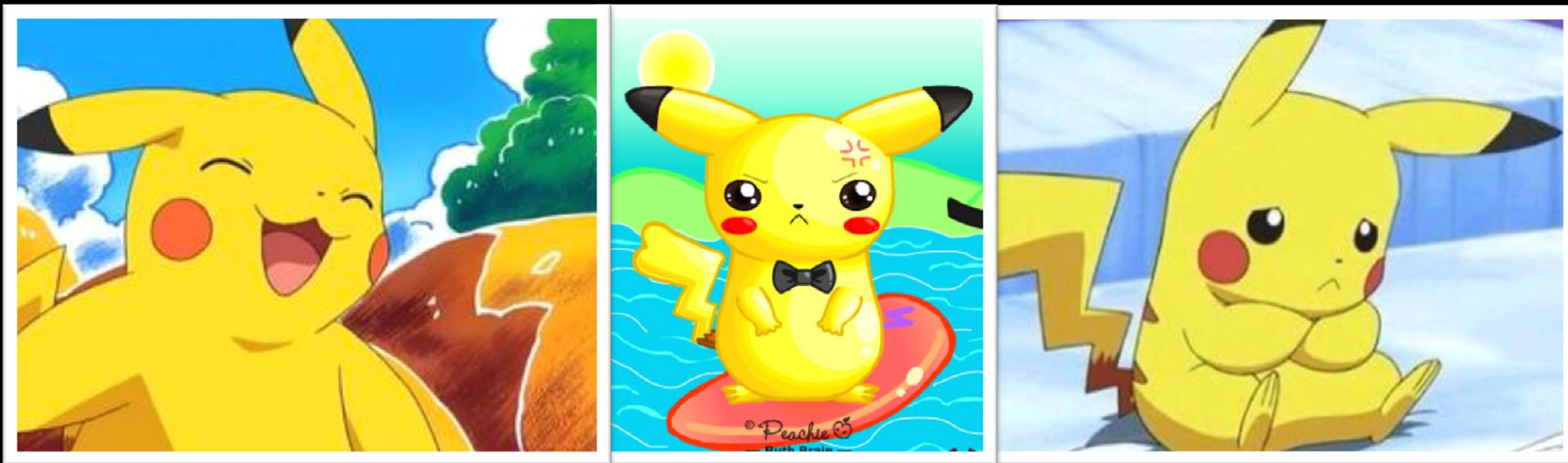
通过语音云海量数据迭代，常用业务语义理解效果持续提升



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





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



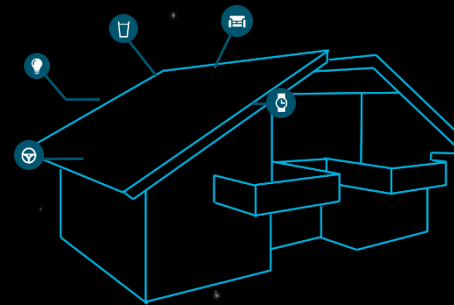
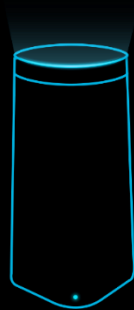
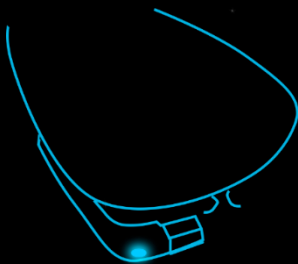
纠错



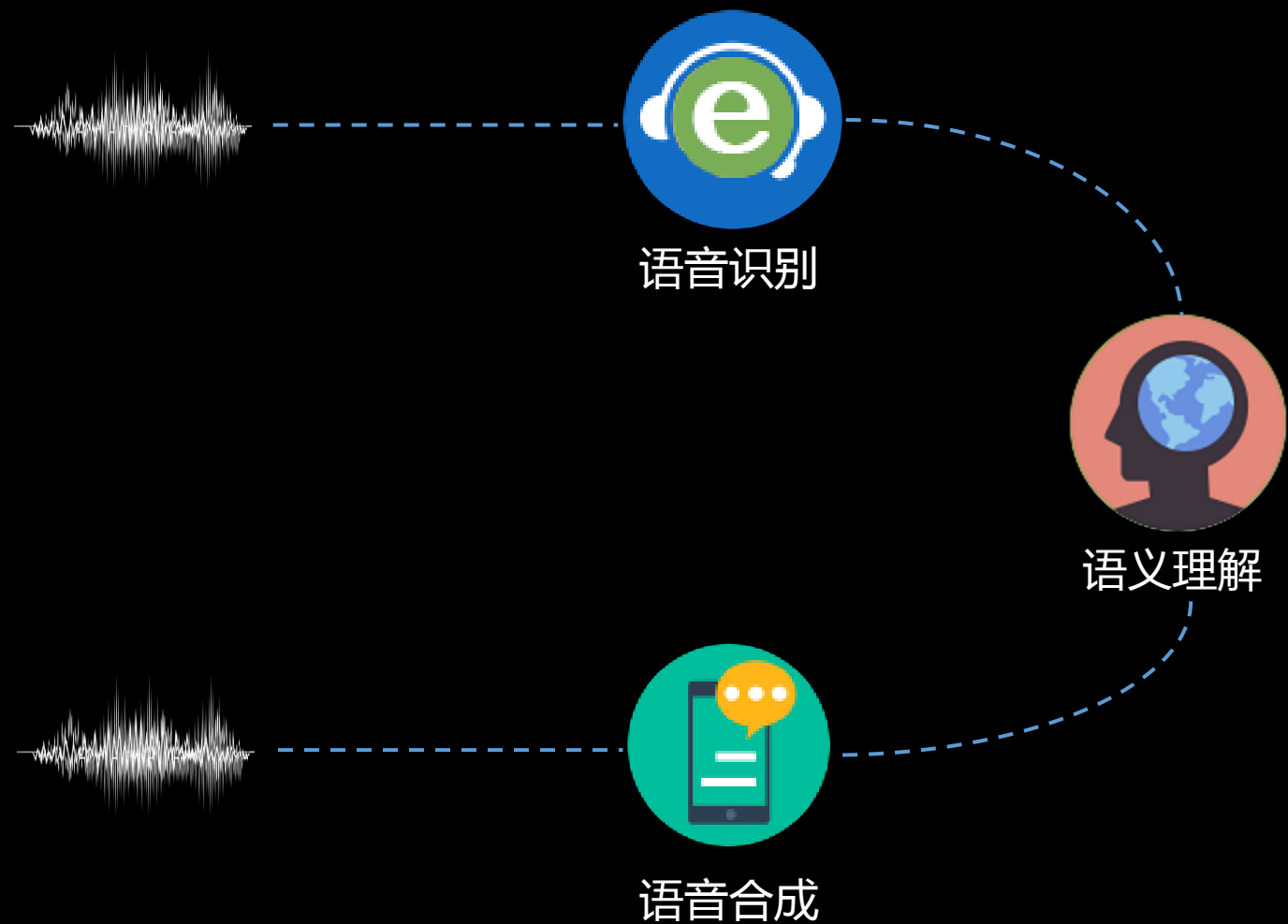
多轮对话



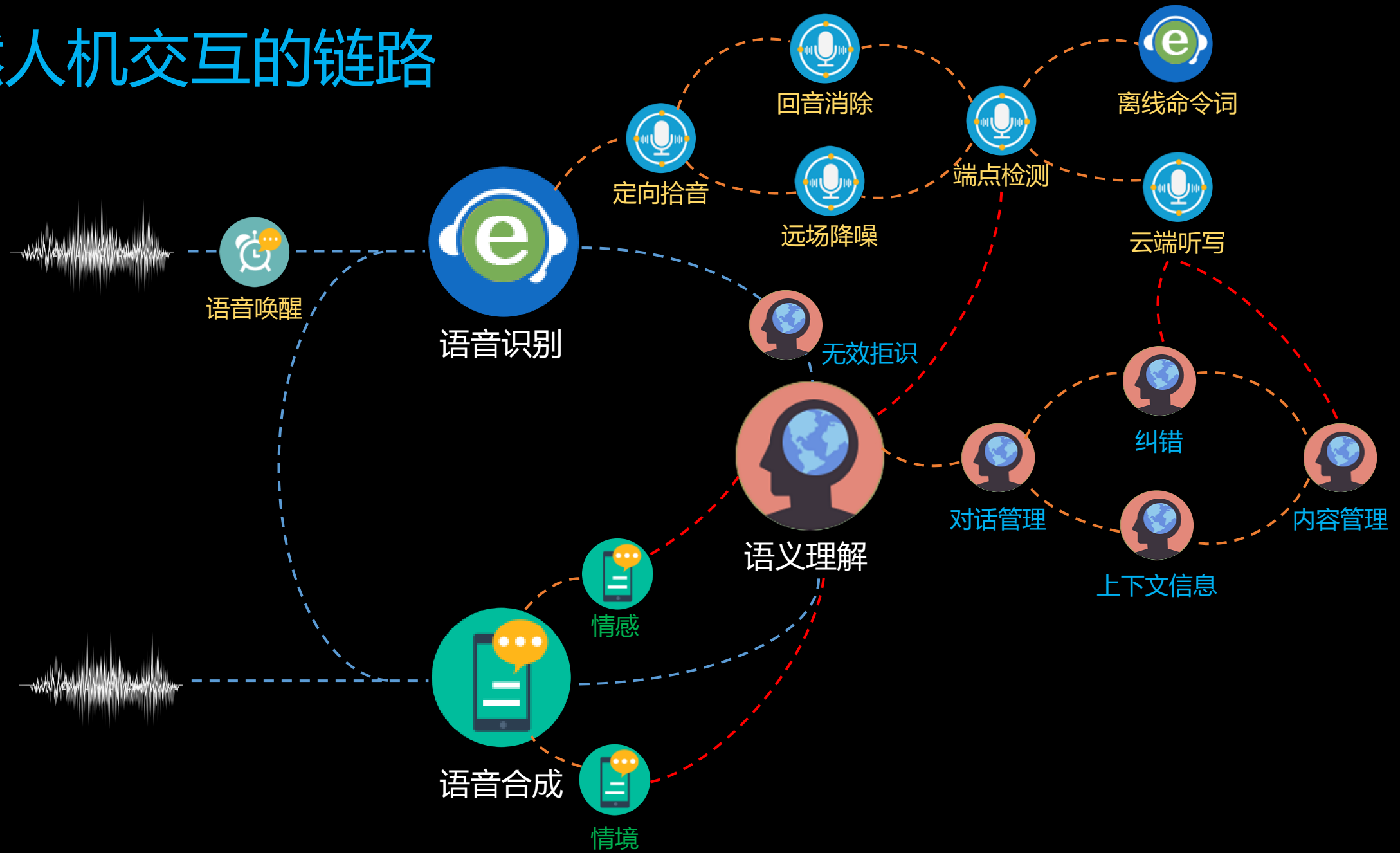
人工智能时代的人机交互界面



传统人机语音交互的链路环节



自然人机交互的链路



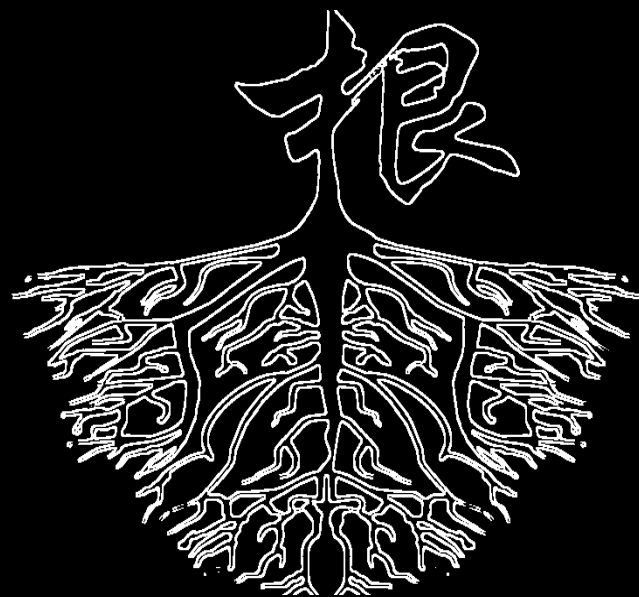
AIUI人机交互的核心改变

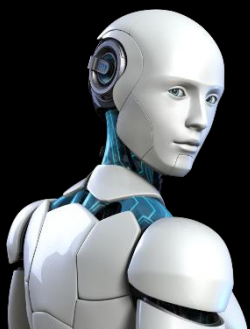
交互全链路贯穿

全链路综合优化，从单点突破的思路转变成整体演进的思路，综合优化交互效果。

紧密融合业务场景

结合实际业务需求，让人机交互是为了解决实际问题而交互





麦克风阵列



语音唤醒



在线命令词识别



对话闲聊



在线语音合成



智件云



人脸识别



声纹识别



离线命令词识别



语义理解



离线语音合成



内容服务



灵活能力搭配

开放服务扩展



统一接口交互

AIUI 人机智能交互解决方案



软硬一体化

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



云端一体化

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



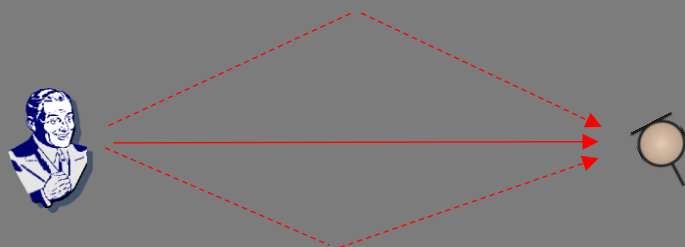
技术服务一体化

远场全双工交互 | 生物特征身份验证 | 云端服务&差异化定制

AIUI标准（人机智能交互标准）

AIUI特性 – 远场精准识别

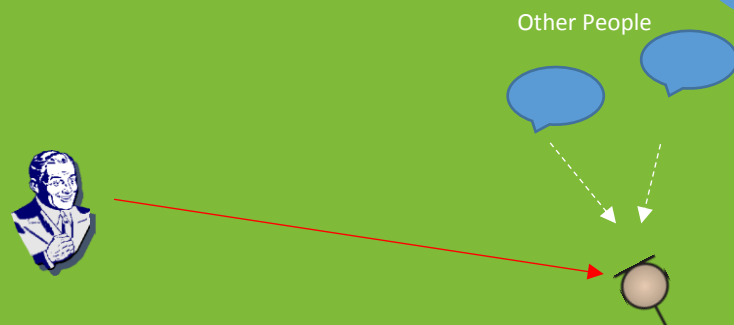
混响：Reverberation



背景噪声：Background Noise



人声干扰：Interference

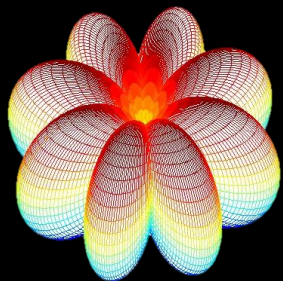


Noise

回声：Echo



AIUI特性 - 远场精准识别



麦克风阵列硬件

远场定向拾音
降噪
回音消除



方言口音适配

方言口音普通话自适应
近二十种方言精准识别



远场识别引擎

适配优化远场拾音特征

一诺千金



语义上下文自修正

结合上下文动态修正 应用场景热词优化

AIUI特性 - 全双工持续交互

唤醒+识别

叮咚叮咚,查杭州的天气
叮咚叮咚,我要听歌
叮咚叮咚,...



循环识别

叮咚叮咚,查杭州的天气
嗯,那个...
(播放爱你)一万年

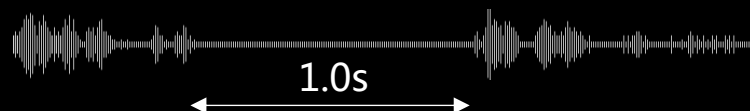


全双工交互

- 持续录音,连续识别
- 拒识 - 过滤无效语音
嗯、啊、哦等语气词
非业务场景的闲聊
- 基于语义的智能断句

我想听

周杰伦的歌



- 随时可打断,支持主动式交互



AIUI特性 – 上下文对话



我要到去兑换外币

附近有交行和建行，你
去哪个？



交行吧

温馨提示：办理外汇业
务需要带身份证。交行
地址为：*****



那里附近有什么好吃的？

附近有很多餐厅，其中
有你最爱吃的老乡鸡



换个口味吧。

为你推荐肥东老母鸡，
是否需要介绍特色菜？



不需要，你退下吧。



与业务深度结合

上下文理解

理解是要去银行

基于内容的提问

多个银行可以办理此业务

多对话场景管理

银行兑换外币场景替换成餐饮场景

跨场景信息共享

地理位置的跨场景使用

长时记忆

用户最近的就餐历史

AIUI特性 - 个性化可扩展的交互

产品特性定制

唤醒词定制

发音人定制



巴布,你好.我要听故事.

[童声]小主人,下面给你播放
灰太狼的故事



还是看蝙蝠侠大战超人的电影吧

这个需要付费收看哦,是否
联系爸爸确认?

业务对话场景定制



好

外部交互扩展



关灯

[启动视频通话,推送确认消息]

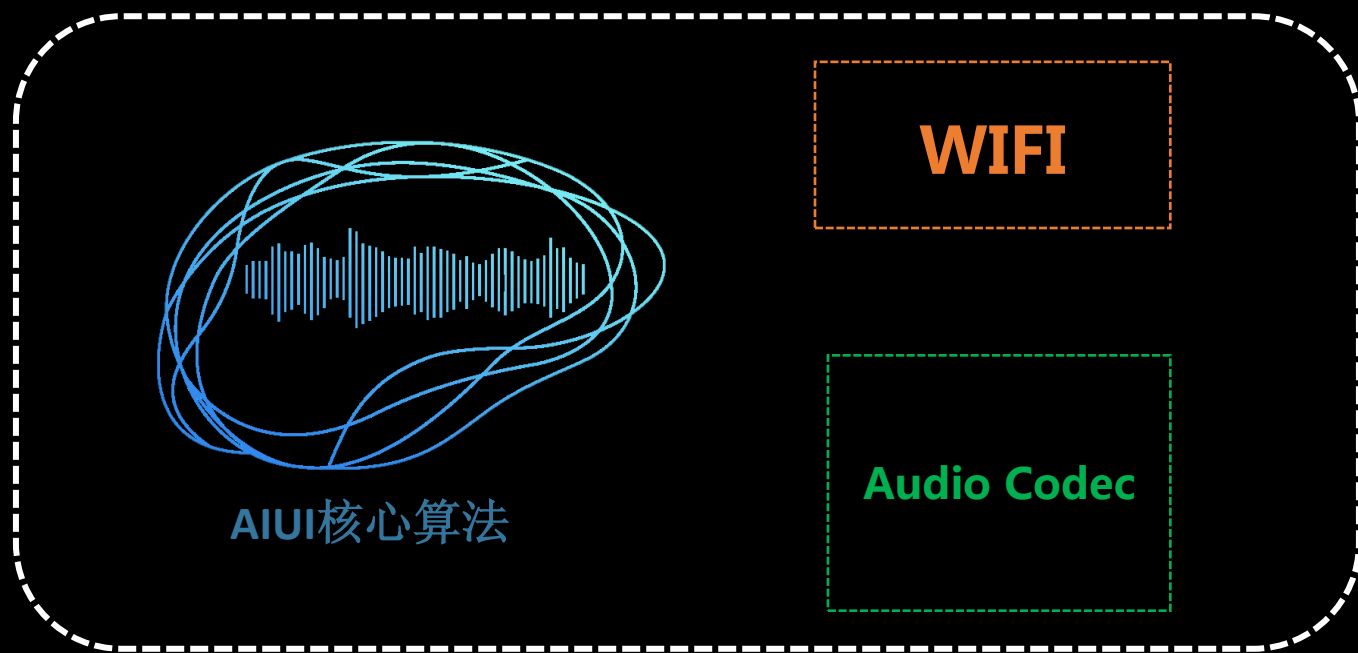
准备在客厅电视上播放蝙蝠
侠大战超人,关上客厅灯吗?

用户个性化交互

基于身份的个性化回答

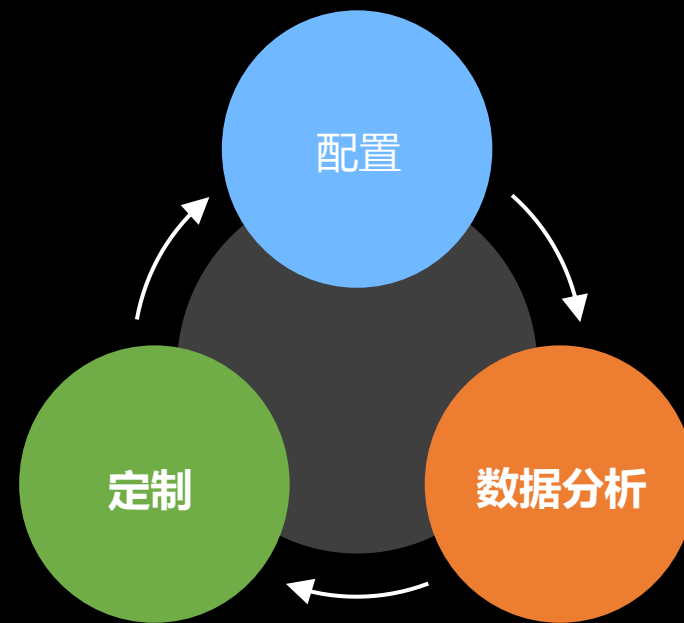
基于用户环境的个性化
提示

AIUI特性 – 简单易用的产品集成



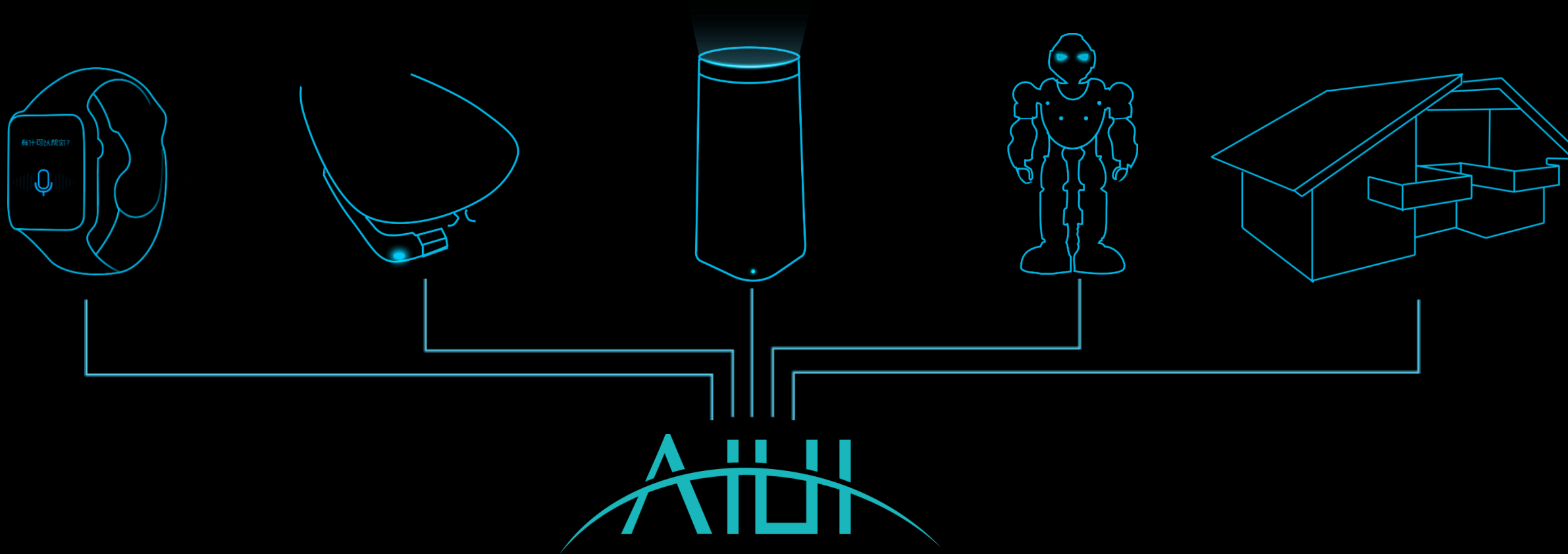
产品中植入软硬一体的AIUI模块
忽略交互能力及流程调度
直接处理业务交互状态及结果

```
文本: "关灯"  
服务: 家居控制  
对象: 灯  
操作: 关闭  
参数: {  
    位置: 客厅  
}
```



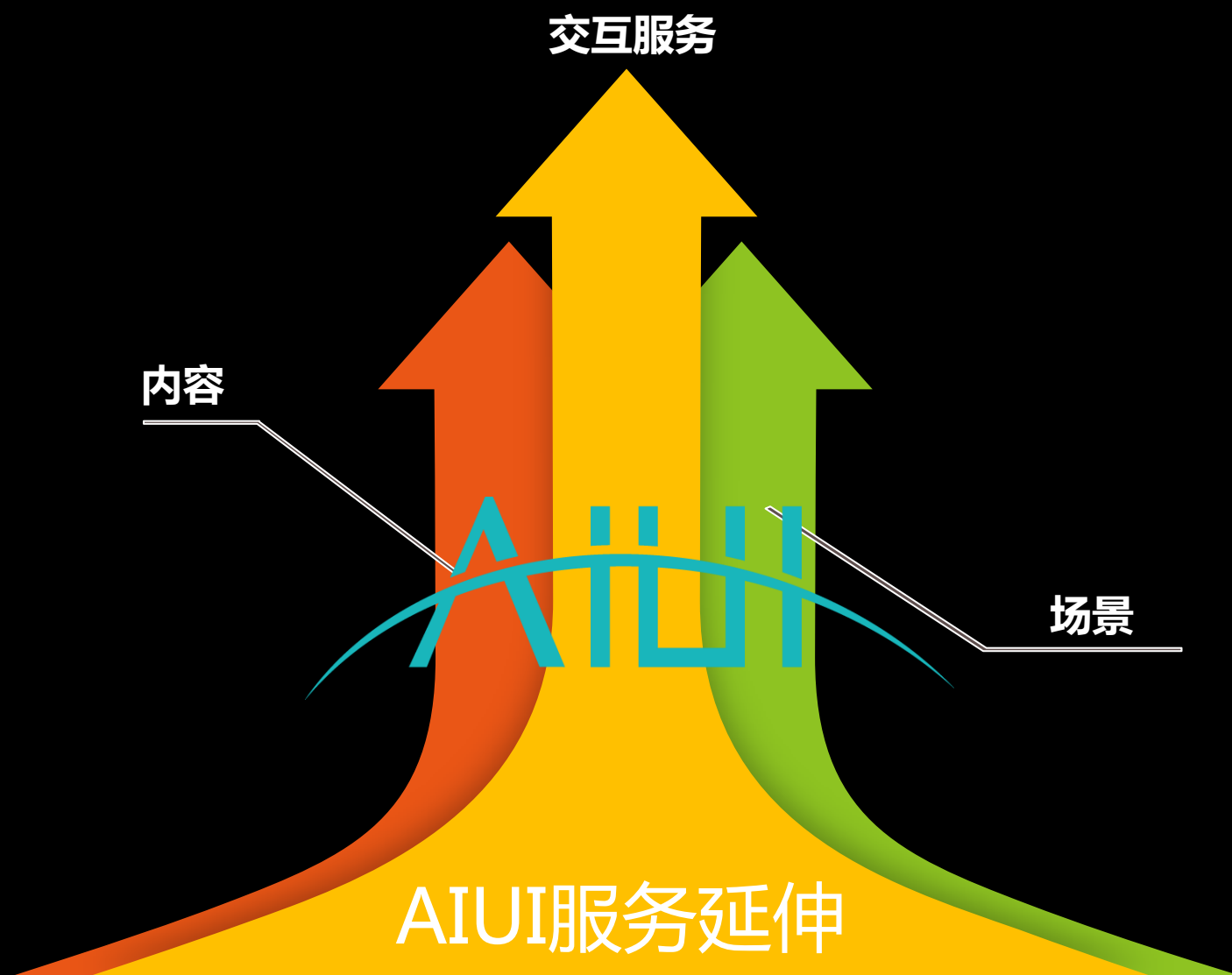
云端配置产品特性,定制产品功能

AIUI – 面向智能硬件的交互方案平台



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

让老人和孩子0门槛享受互联网带来的便利



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

