

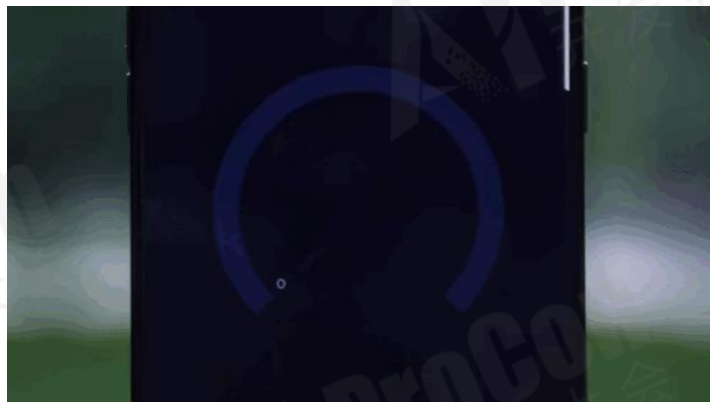
# 5G 驱动 AIoT 多模态智能技术与应用

云知声董事长/CTO 梁家恩

**5G 变革：不止于快！**



## 5G 变革：大带宽——刷爆全网的5G测速



平均下载速率在700Mbps左右  
差不多是4G的十倍



手机软件，**下载速度比解压速度还快**



下载音乐，只能看到“**未下载**”和“**已完成**”两个画面



在线看一部超高清的电影，**进度条随便拖**

## 5G 变革：除了快，还有什么？

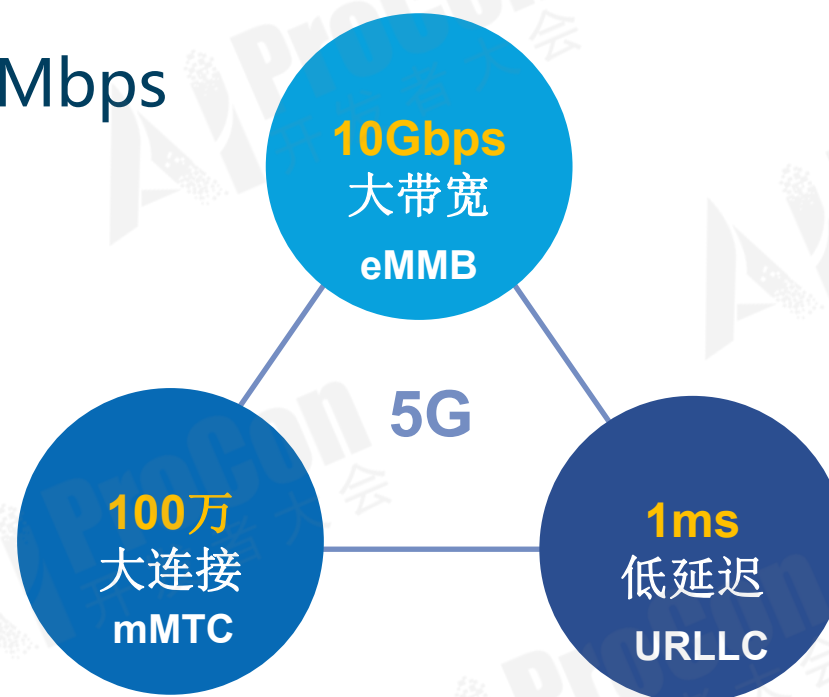
**5G** 2020s, AIoT / 超高清 / VR/AR, 10Gbps

**4G** 2010s, 移动 / 视频, 100Mbps

**3G** 2000s, 网络 / 图片, 2Mbps

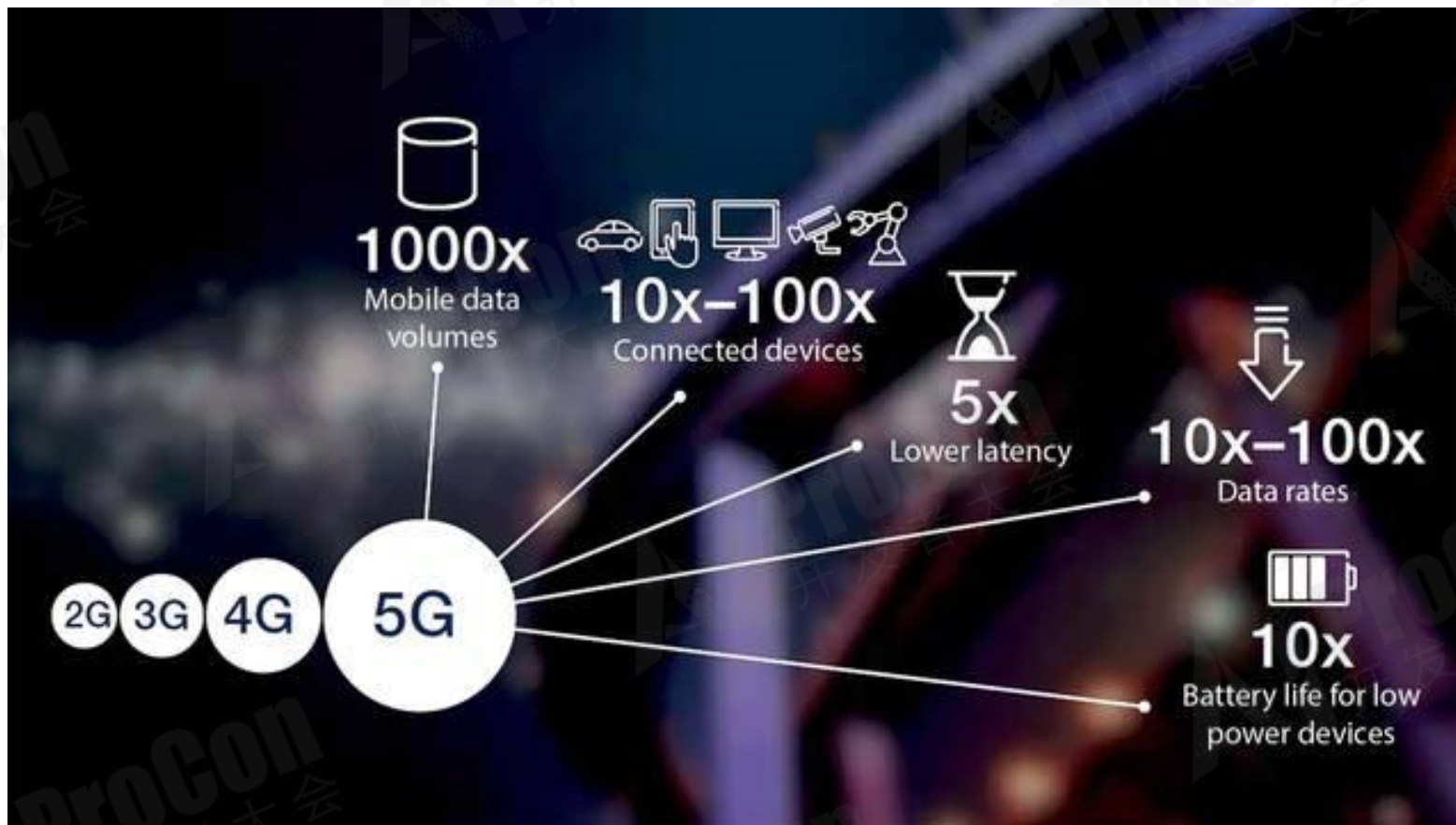
**2G** 1990s, 数字短信, 100kbps

**1G** 1980s, 模拟语音, 2.4kbps





## 5G 变革：低延迟 / 大连接——万物互联（IoT）



### 从移动互联到万物互联

从 4G 时代 10 亿级手机到 5G 时代百亿、千亿级物联网终端，不仅是数量级上的增长，更将带动应用、生活与生产方式的变革。



## 5G 机遇：AIoT 时代到来——应用 + 数据 + 智能



从移动应用到智能产业应用



数据规模与维度爆炸性增长

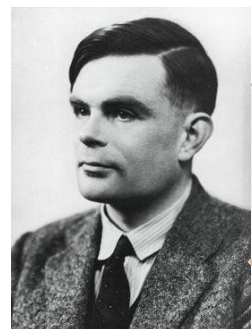
# AI 技术演进与多模态趋势



# 第四次工业革命：从信息化到智能化

“计算 + 存储 + 通信” 技术发展  
Moore' s & Gilder' s Law

互联网产业应用蓬勃发展  
Metcalfe' s Law



图灵机@1936

ENIAC  
@1946



ARPANet &  
Unix @1969

个人电脑(PC)  
1970s



阿波罗登月  
@1969

移动电话 / 1G  
1980s

专家系统  
统计学习  
1980s

WWW / 2G  
1990s

支持向量机  
1990s

PC互联 / 3G  
2000s

移动互联 / 4G  
2010s

万物智联 / 5G  
2020s

A: 移动应用  
B: 大数据  
C: 超级计算

D: 深度学习



第一次 AI 寒冬  
~ 1974

第二次 AI 寒冬  
~ 1987

AI 技术探索：两起两落

深度学习  
@2006

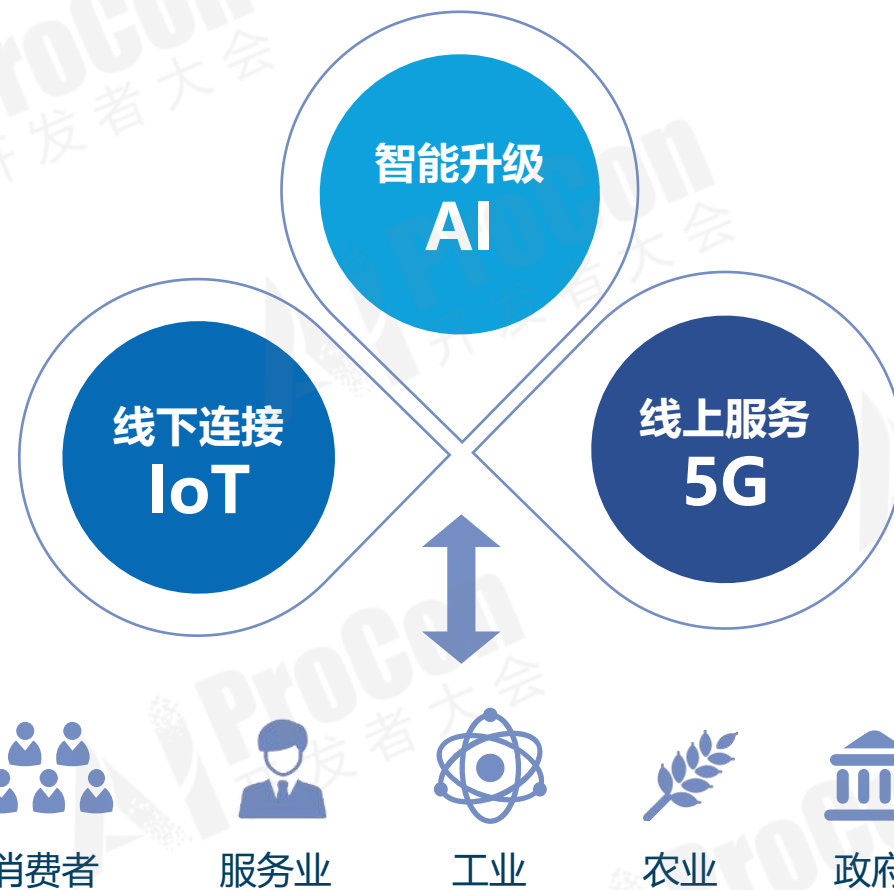
ImageNet  
@2012

AlphaGo  
@2016

第三次 AI 浪潮

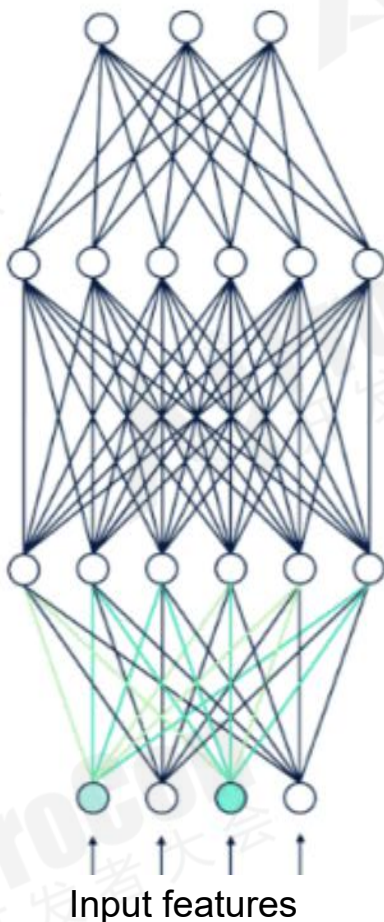


# AIoT 变革：交互智能与产业智能



# 深度学习演进：从模式匹配到序列影射

Output Probabilities



大数据与超级计算加持下的“AI 暴力美学”

**DNN:** 深度神经网络

**CNN:** 卷积神经网络

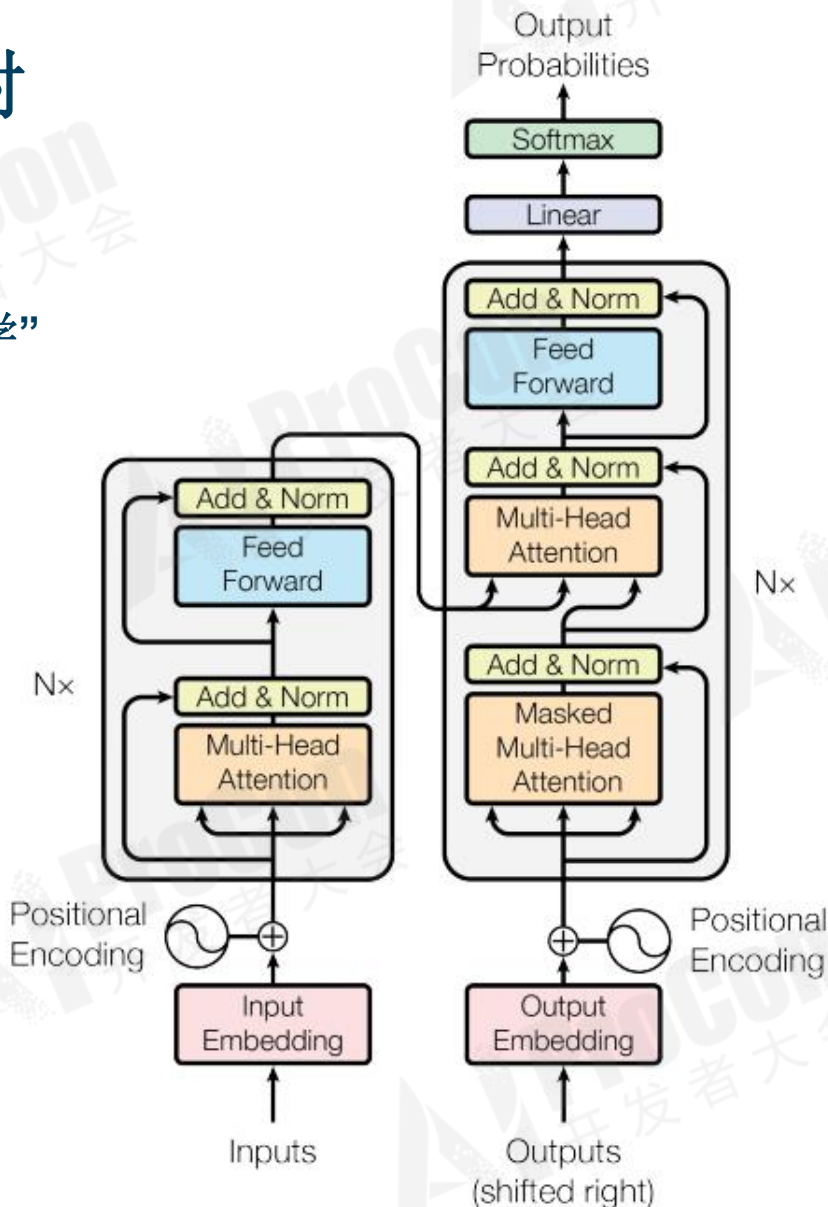
**RNN:** 回归神经网络

**LSTM:** 长短时记忆机制

**Attention:** 注意力机制

**Transformer:** 序列编解码模型

是表征与映射，非认知和推理。

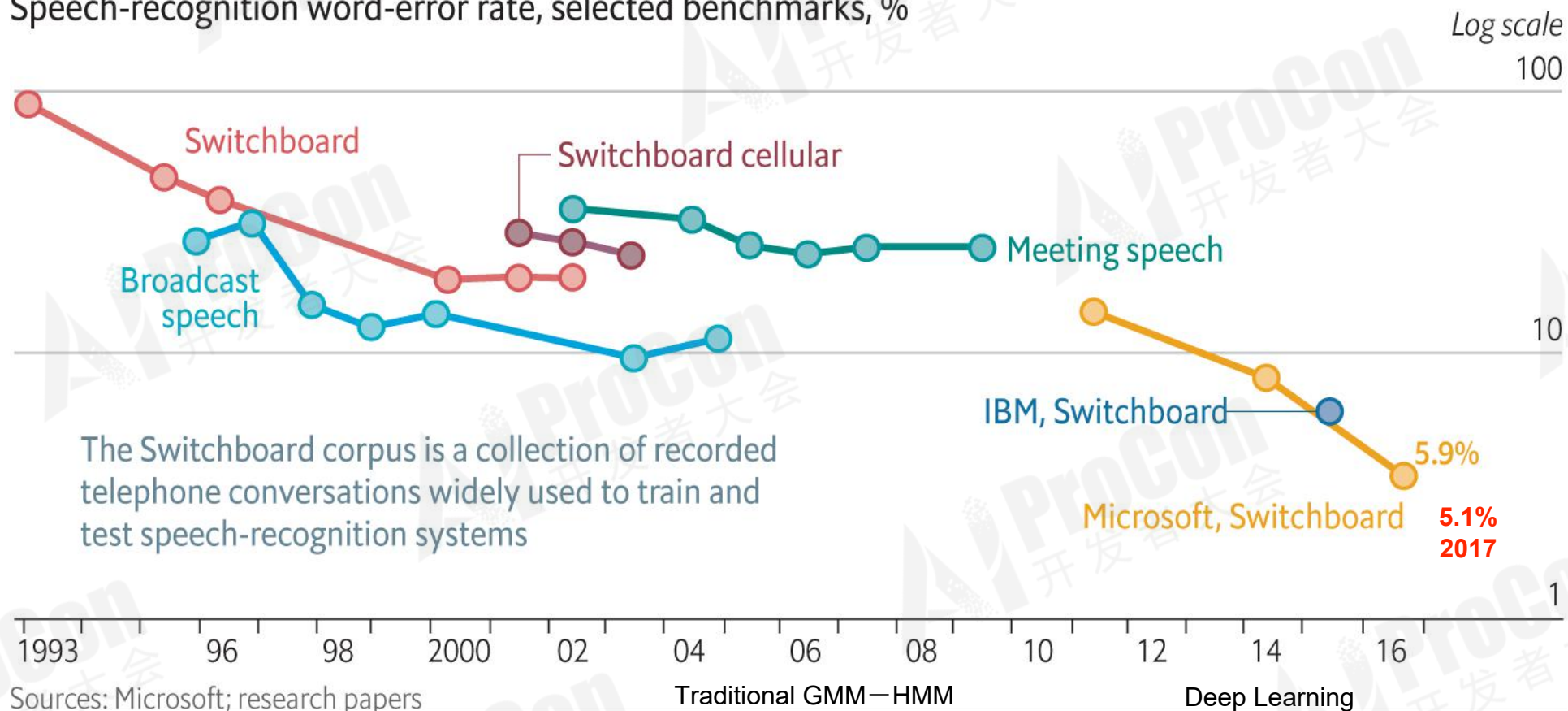




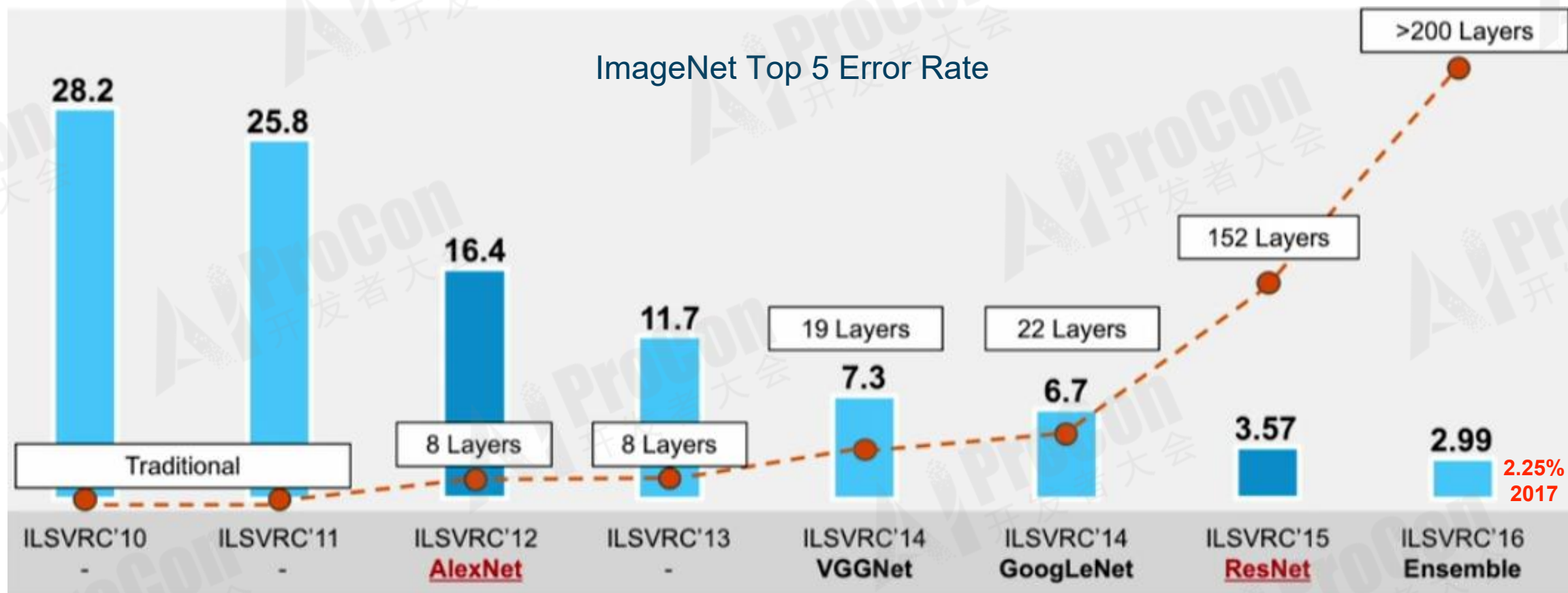
# 深度学习进展：语音识别

## Loud and clear

Speech-recognition word-error rate, selected benchmarks, %



## 深度学习进展：图像识别





## 深度学习应用：AI 感知与生成

输入序列**表征**，及到输出序列的**映射**：

- 语音识别：Speech to Text
- 图像识别：Image to Label / Description
- 语音合成：Text to Speech
- 图像生成：Description to Image
- 机器翻译：Source to Target Language
- 游戏博弈：Input to Reaction
- .....

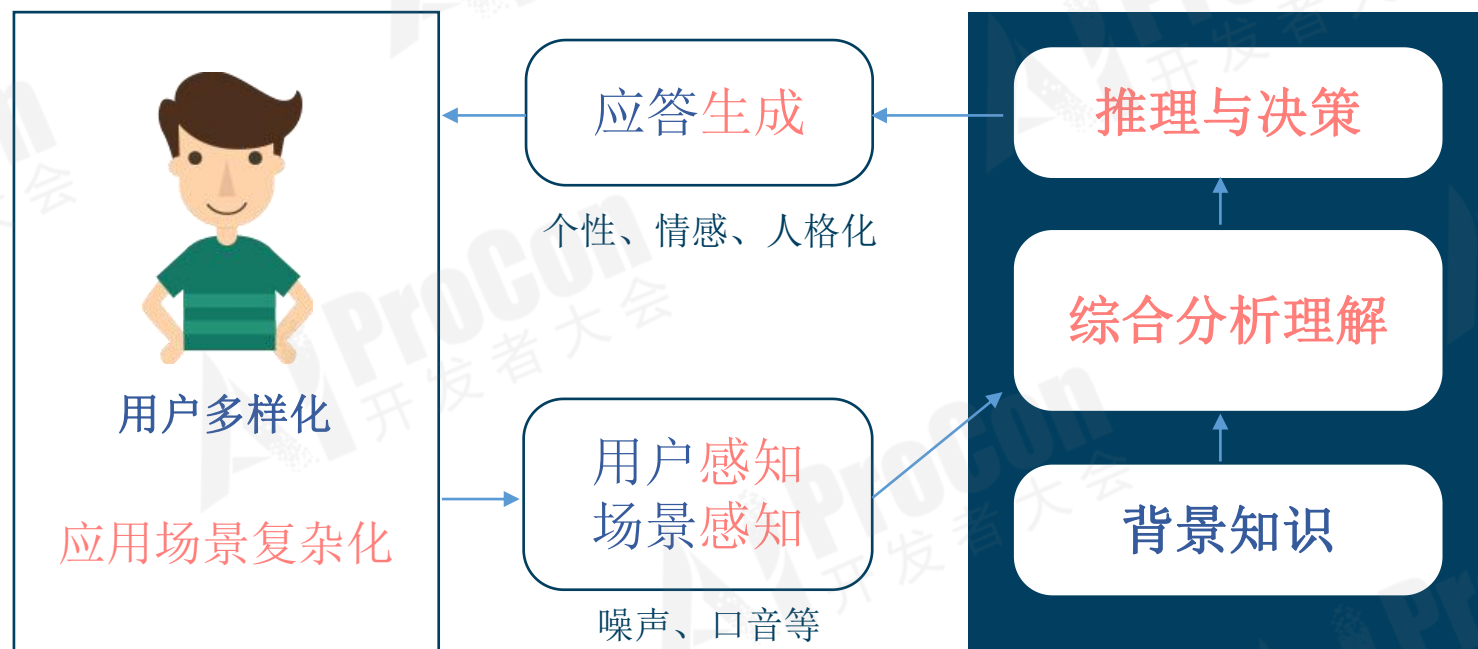
### 优势

- 大计算，简单粗暴有效
- 大数据，精准度提升显著
- 上手快，专家知识依赖少

### 局限

- 数据需求大，规模 & 质量
- 黑盒子，可控性、可解释性差
- 异常处理、攻击防御挑战大
- 非映射：理解、推理、决策等

# 深度学习屡次“超越人类”，为何 AI 系统还那么傻？



## 准确理解与应答的挑战:

- 多模态综合感知：用户+场景
- 背景知识：常识+专业
- 心智模式：用户画像+机器角色
- 目标管理：理解、推理与决策
- 澄清与纠错策略
- 个性化应答
- .....

AI 在“耳聪目明”方面进步显著，在“全局统筹、灵活机变”方面明显不足。

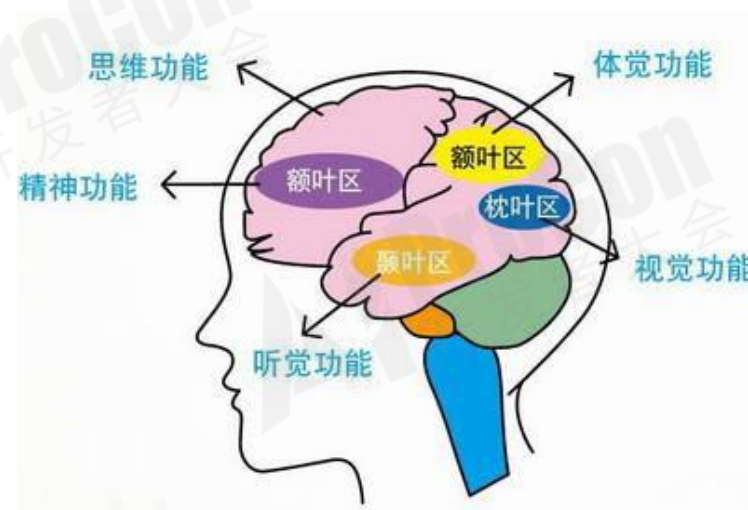


## 智能挑战：AI 如何能更聪明？



### 垂直化应用，分解问题复杂度：

- 明确垂直应用目标，设计可计算架构
- 针对应用，设计相关知识库、策略和流程
- 采集真实数据，设计和优化应用系统
- 用户数据和知识库迭代，滚动优化性能
- 行业智能技术架构提炼和演进，新应用拓展



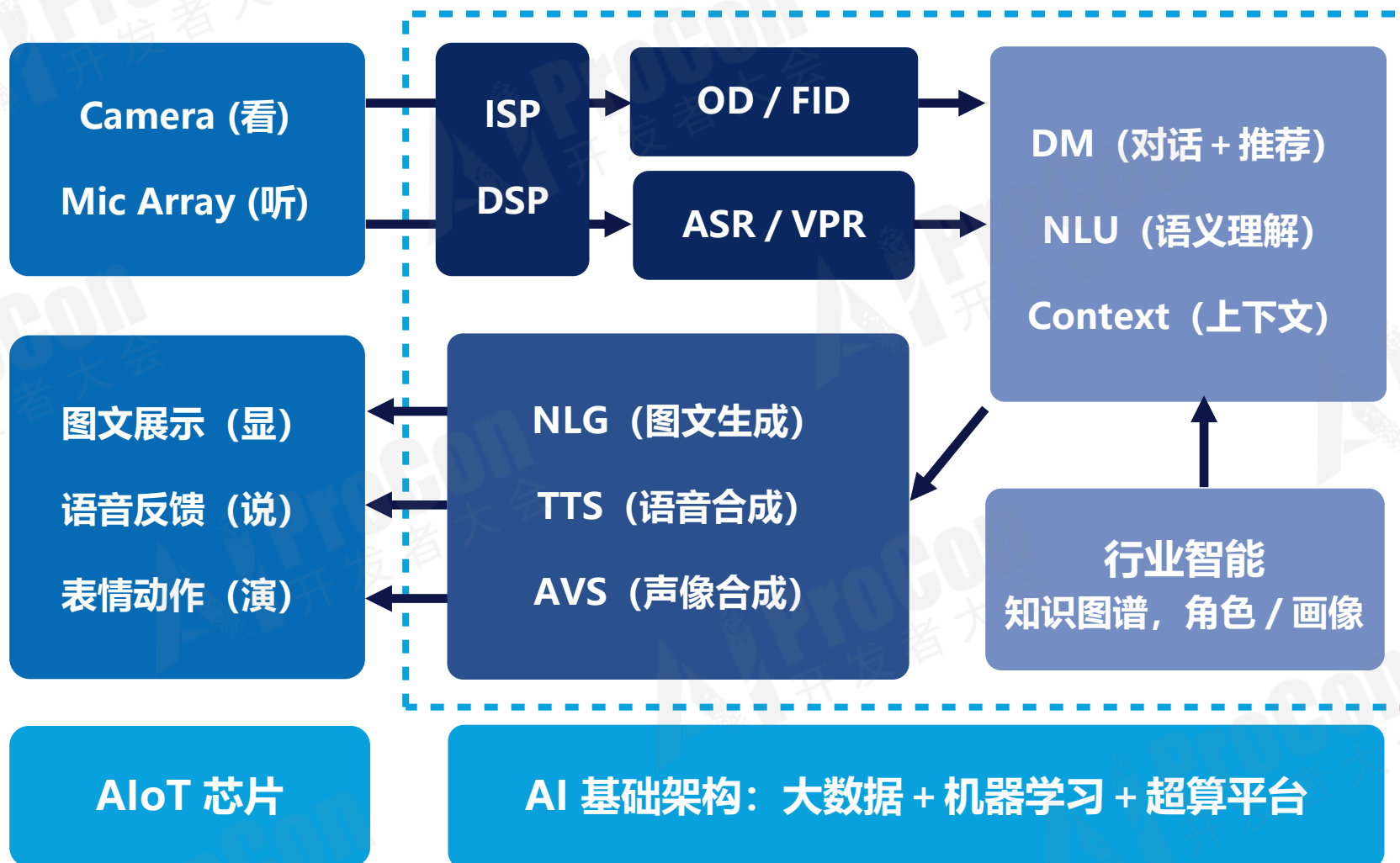
### 多模态融合，提升精准度和灵活性：

- 音视频等多模态场景分析
- 场景实体和事件检测
- 用户语音、表情、动作等多模态融合感知
- 多模态（维度）用户画像
- 声像融合多模态生成

# 云知声多模态基础：感知、认知、生成（全栈 AI 技术）



多模态、拟人化  
智能系统





# AIoT 芯片：多模态边缘计算

# AIoT 规模化增长，边缘计算芯片成为关键

## 交互智能化

语音/图像/手势等多模态智能交互



## 应用场景化

家居/汽车/办公/零售/  
医疗/金融.....

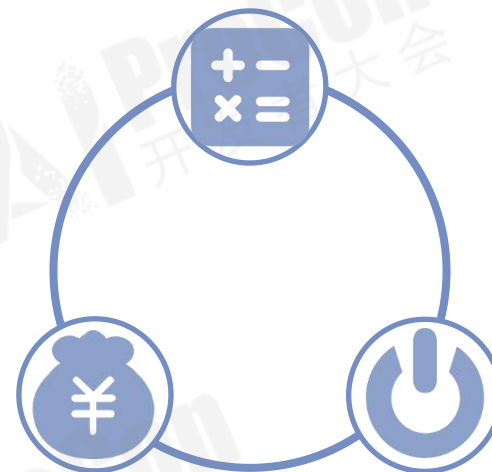
## 数据规模化

2018年全球总数据量为2ZB，机器  
产生占比达56%，首次超过人类



## 边缘算力需求

规模巨大、复杂度高，存储和计算挑战大  
智能交互标准化，保证随时可用、隐私安全



## 低成本需求

在保障边缘算力前提下  
降低成本

## 低功耗需求

低功耗不仅涉及芯片制造工艺，  
还涉及场景、应用、算法

AIoT 时代，数百亿级的智能终端及其数据规模，对边缘计算芯片设计提出了极高要求。



# 云知声 AIoT 产业实践闭环

## 认知技术

知识、理解、分析、决策等

## 感知 + 生成

信号、语音、图像、文本等

## Atlas 基础架构

大规模机器学习超算平台

行业 Turnkey 方案

“云端芯” 产品体系



## 创新业务

城市 / 酒店 / 金融等

## 智慧教育

SaaS 服务 / KAR 玩具等

## 智慧医疗

病历录入 / 质控 / 导医等

## 智慧生活

家居 / 车载 / 中控等

## 理论原创 + 技术演进

论文、刷榜、炫技  
(0 => 1 => 10)

原理 + 工具



## 工程优化 + 产品化

场景、需求、边界  
(10 => 30 => 60)

体验 + 规模化



## 产业变革 + 商业化

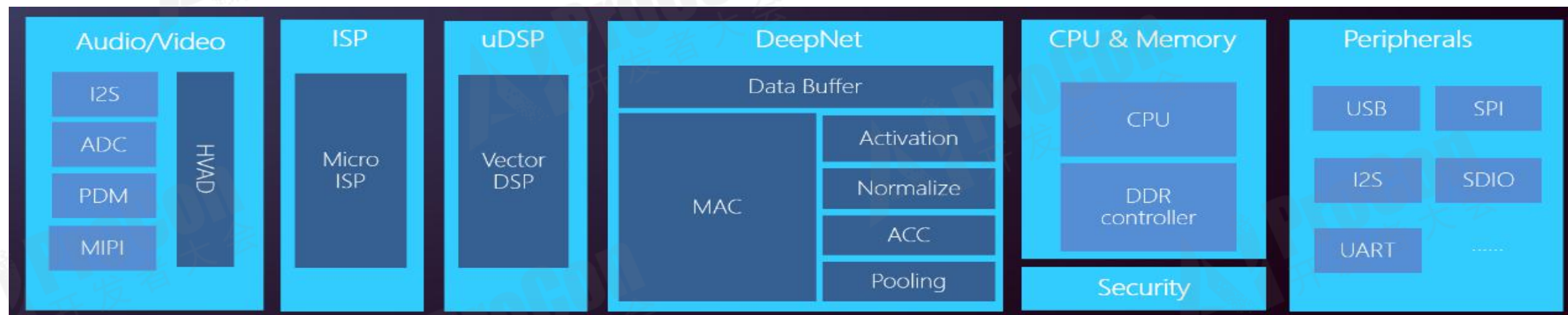
渠道、运营、模式  
(60 => 100)

数据智能

# 云知声 AIoT 边缘计算芯片



## 云知声 AIoT 芯片架构



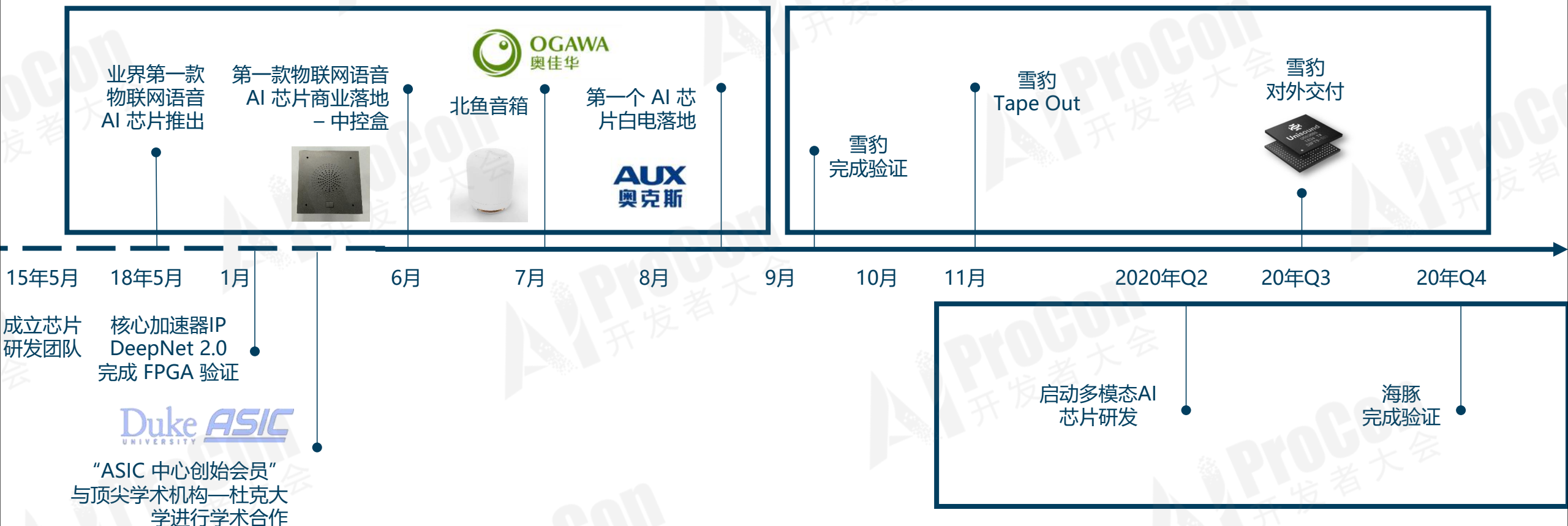
# 云知声 AIoT 芯片之路：从语音到多模态

雨燕

业界第一款语音 AI 芯片正式商业落地

雪豹

面向智慧出行的语音AI芯片



海豚 —— 智慧城市 多模态AI 芯片快速推进



# 云知声 AIoT 产业落地成果



家居终端



车载终端



儿童教育机器人



行业服务机器人



智慧医疗



口语评测

# 云知声 AIoT 产业应用拓展

## 智慧城市

智慧家居、智慧医院、智慧地产等



厦门市人民政府



華通集團  
HUATONG GROUP

## 智慧社区

全屋智控、语音管家、智能安防、智能能耗等



世茂集团



平安好医生

## 中控智能

设备与设备、服务与服务连接

## 单体智能

智慧生活/智慧服务



电视



空调



机器人



后视镜



远程教育



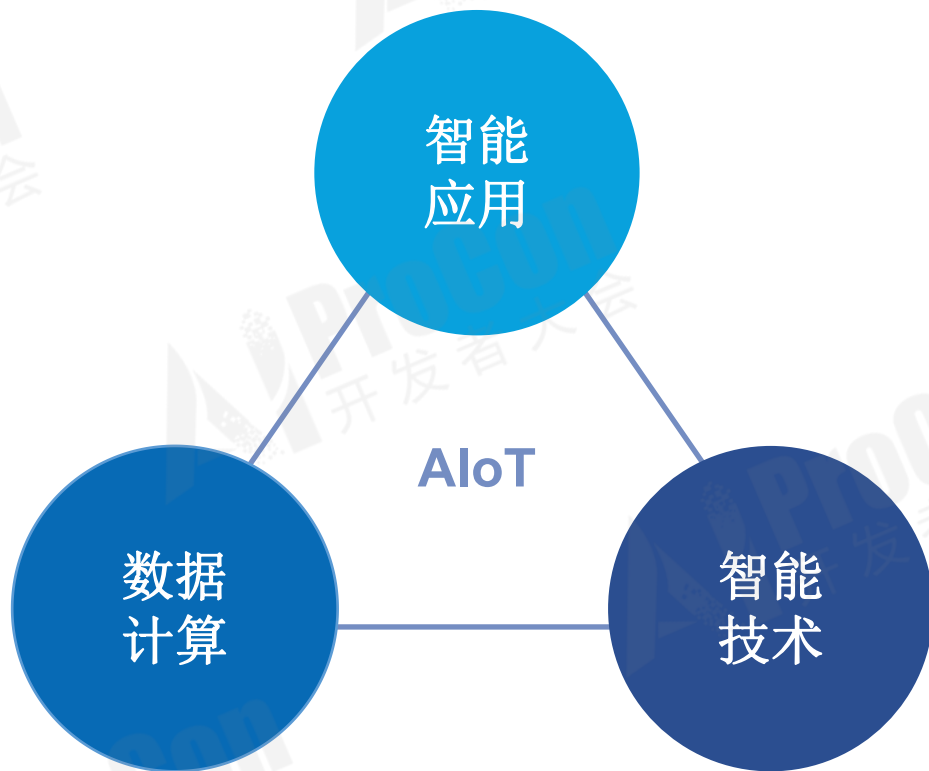
在线教育



病历录入



## AIoT 产业升级展望：空间巨大、任重道远



### AIoT 产业升级关键挑战：

- 交互创新：多模态智能化，回归自然
- 应用创新：服务与生产重构，Thinking in AI
- 用户互动：活跃度，心智与行为模式改变
- 边缘计算：多模态，高性能、低功耗、低成本
- 数据迭代：模型、画像与知识库高效更新
- 技术突破：高效学习与认知能力



感谢聆听!



Just for smart life!