



第九届中国系统架构师大会
SYSTEM ARCHITECT CONFERENCE CHINA 2017

语音识别技术回顾 及应用



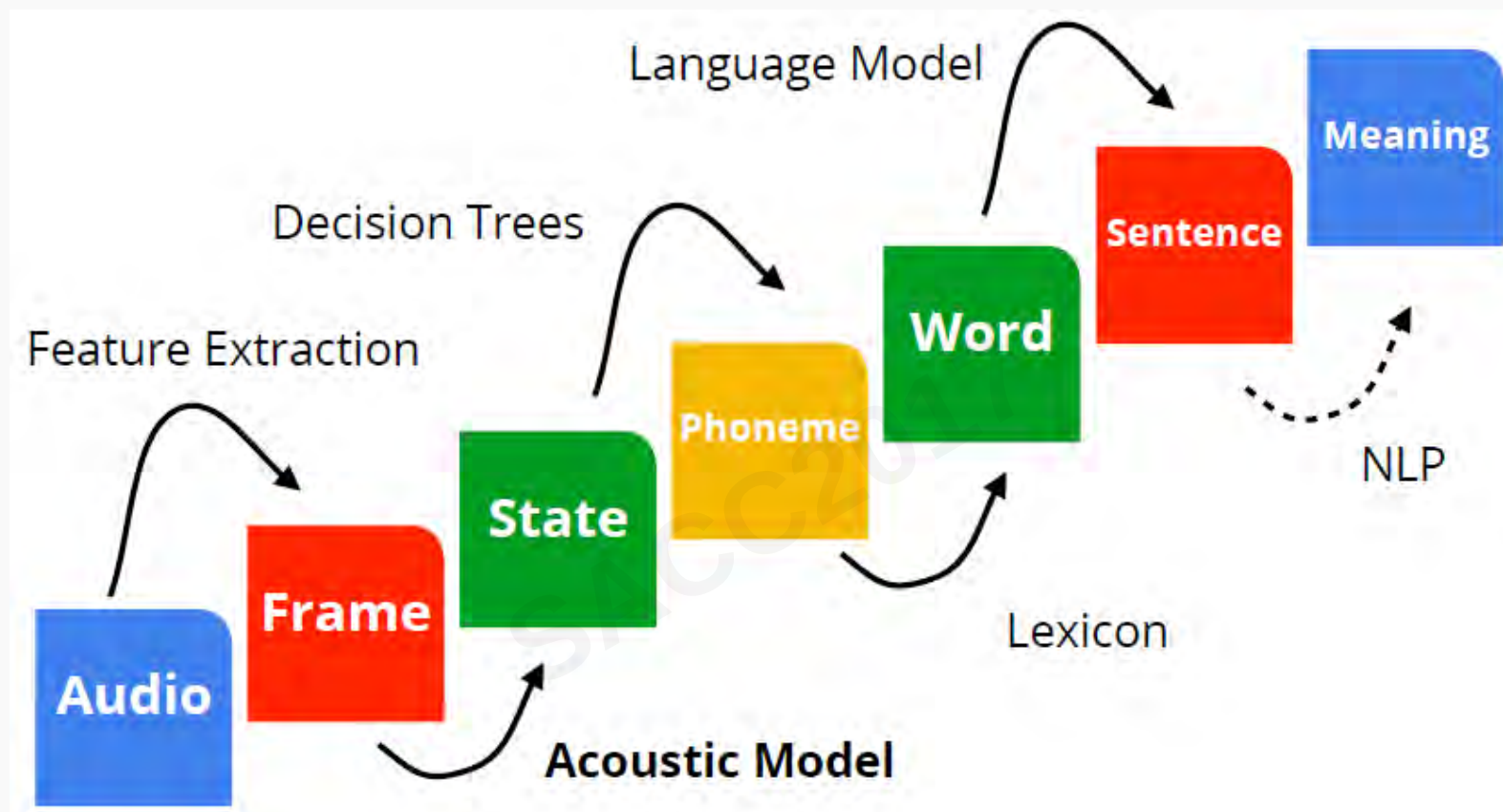
Ali A.I. Labs
阿里人工智能实验室

空海

人如何理解语音

1) 语音交互技术

2) 语音识别相关应用



人如何进行语音识别

如何听见声音

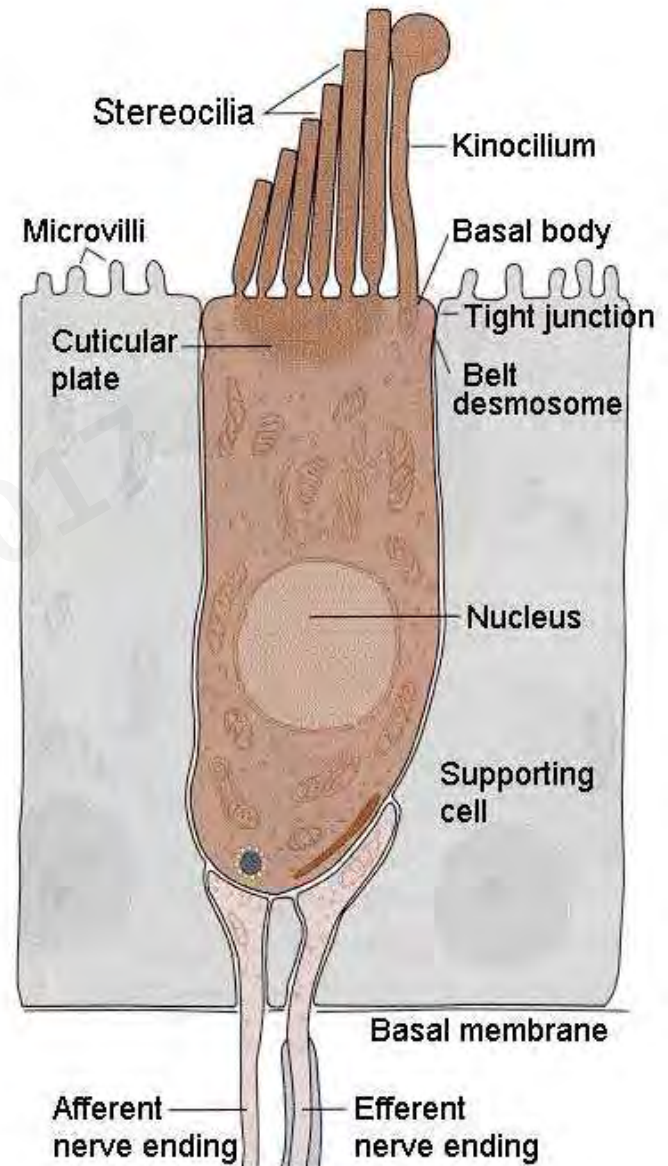
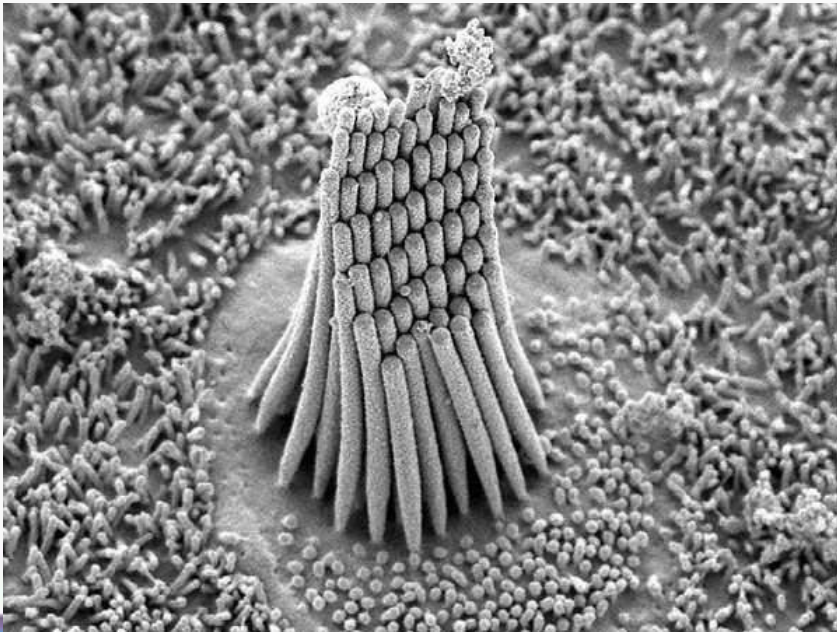
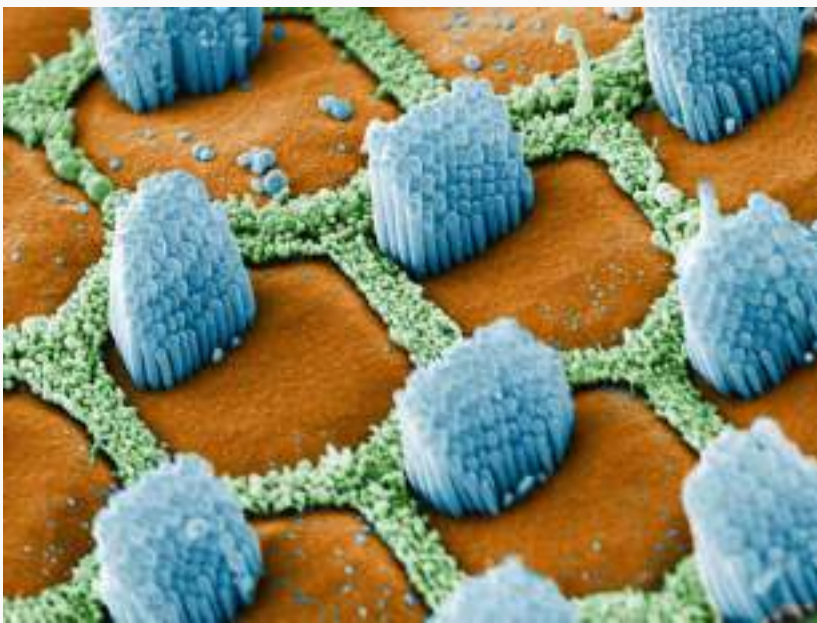
如何进行记忆理解

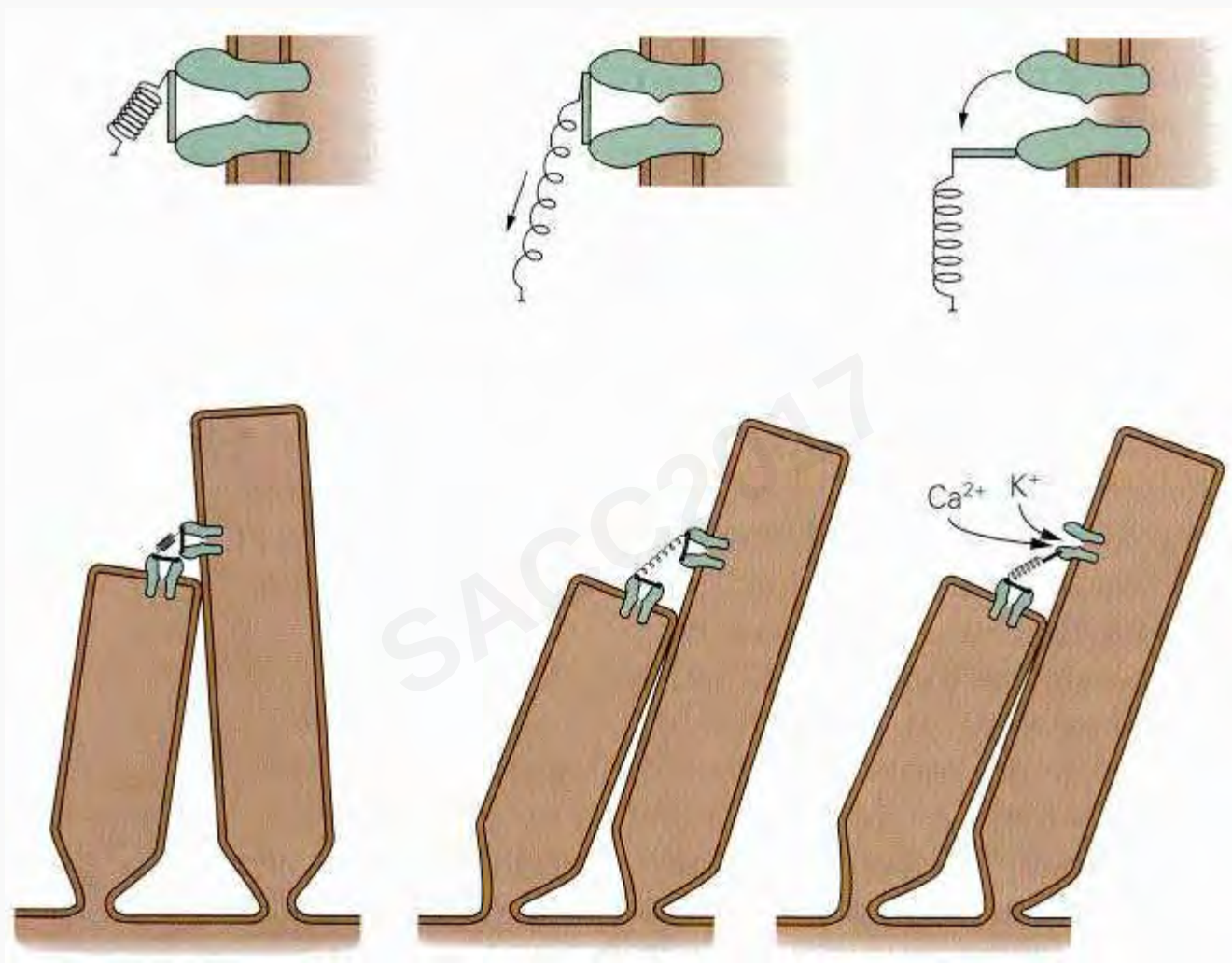
SACC2017

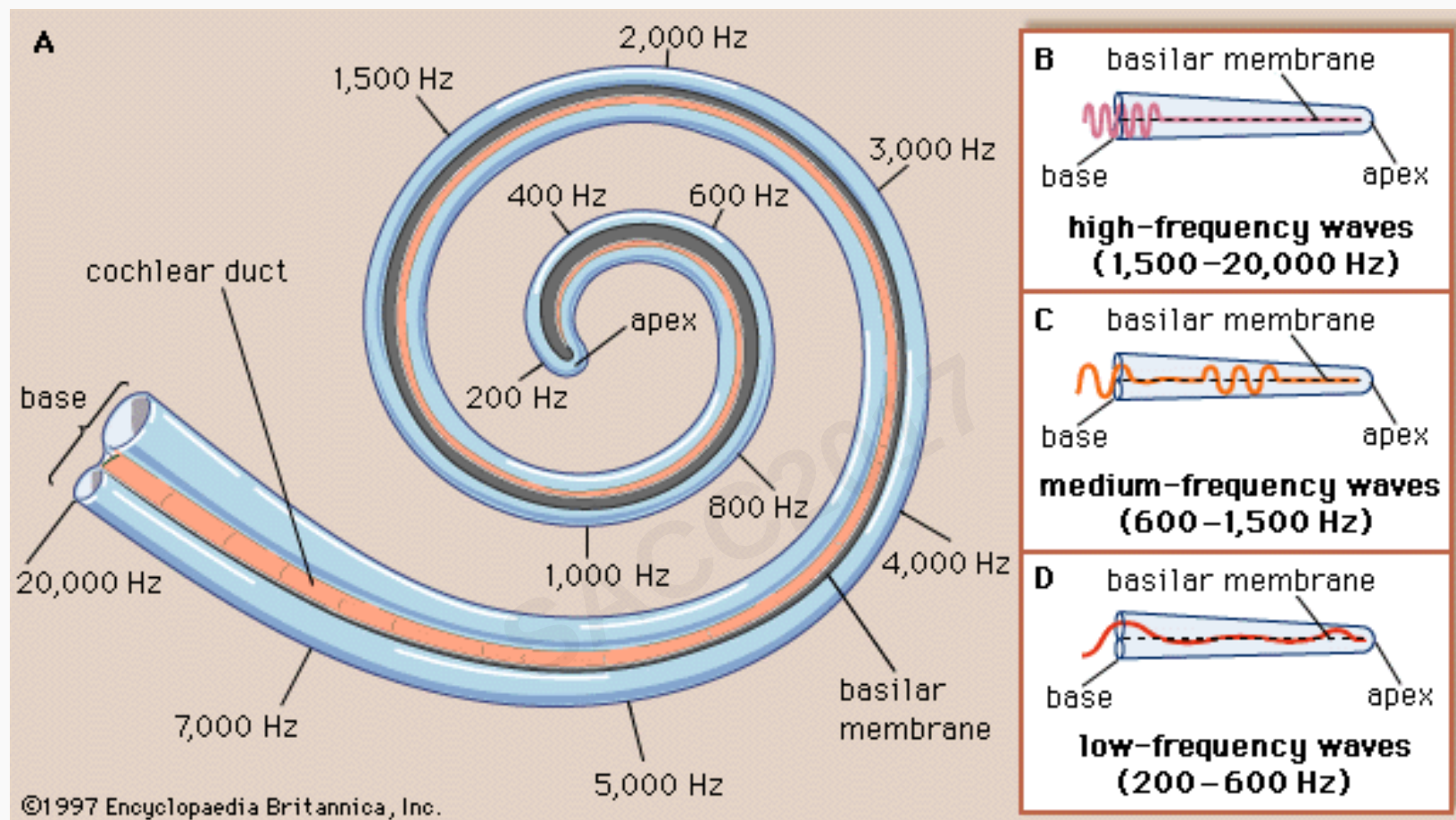
耳朵的构造



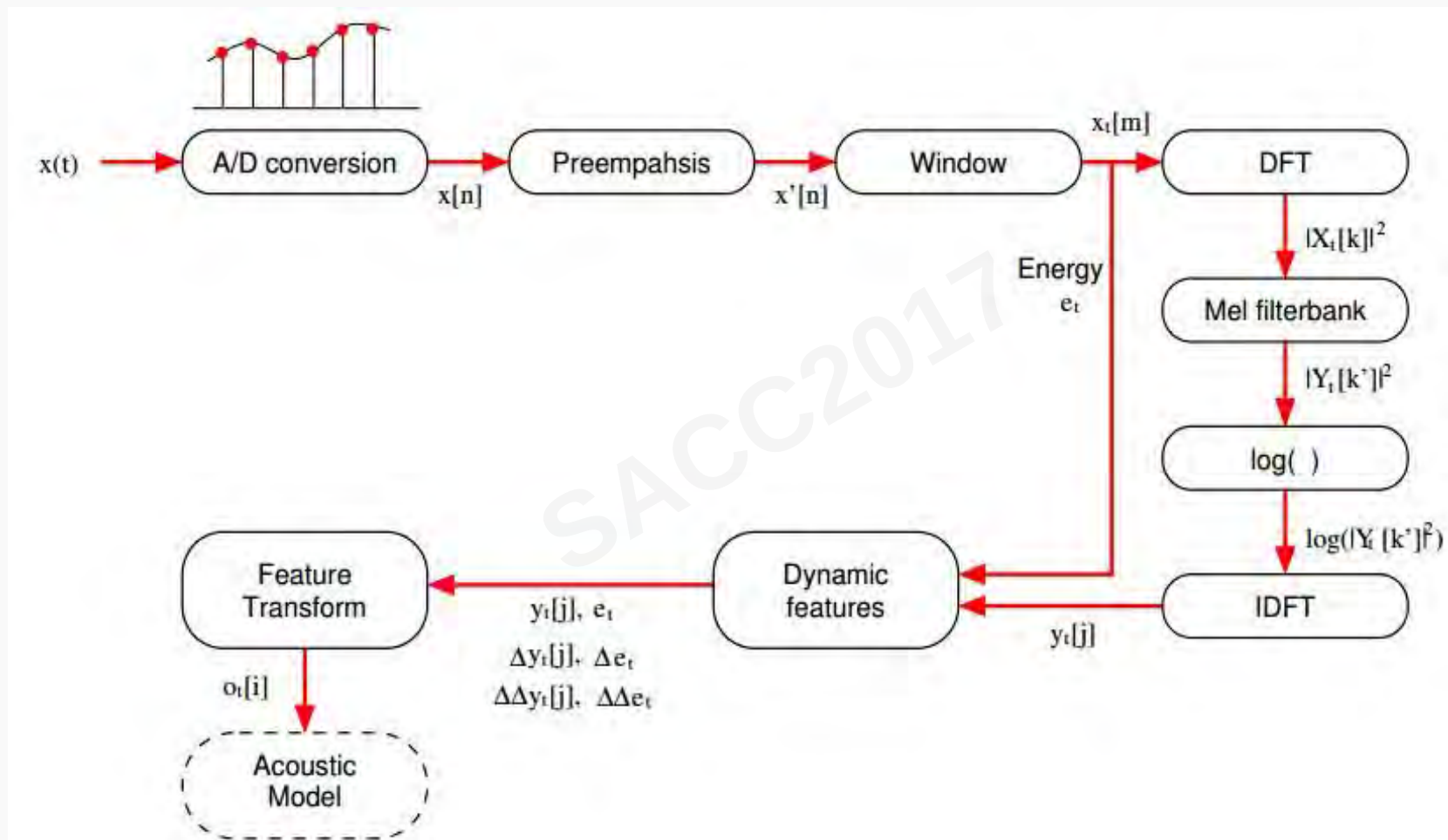
听觉毛细胞(hair cell)



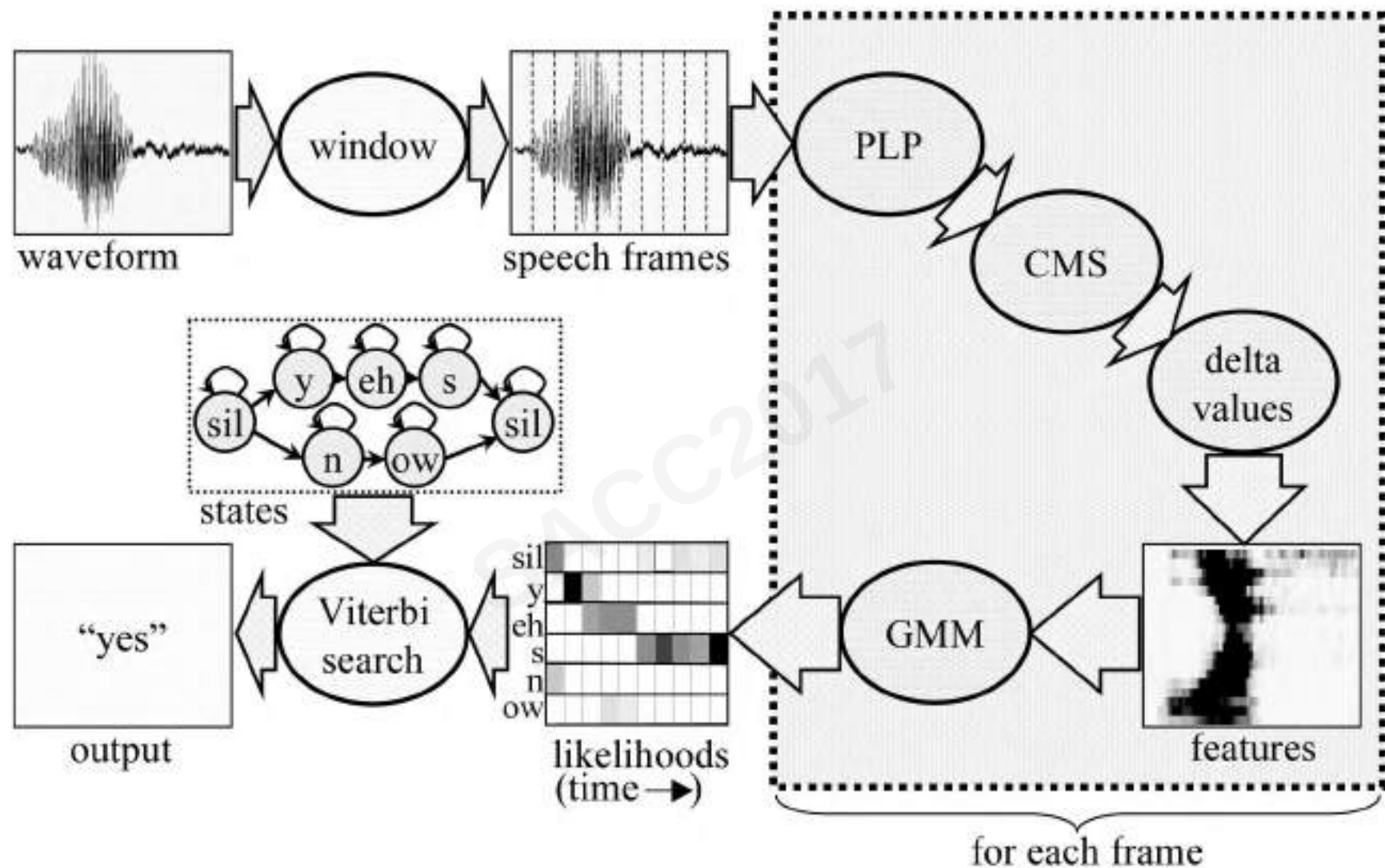




机器如何感知数据



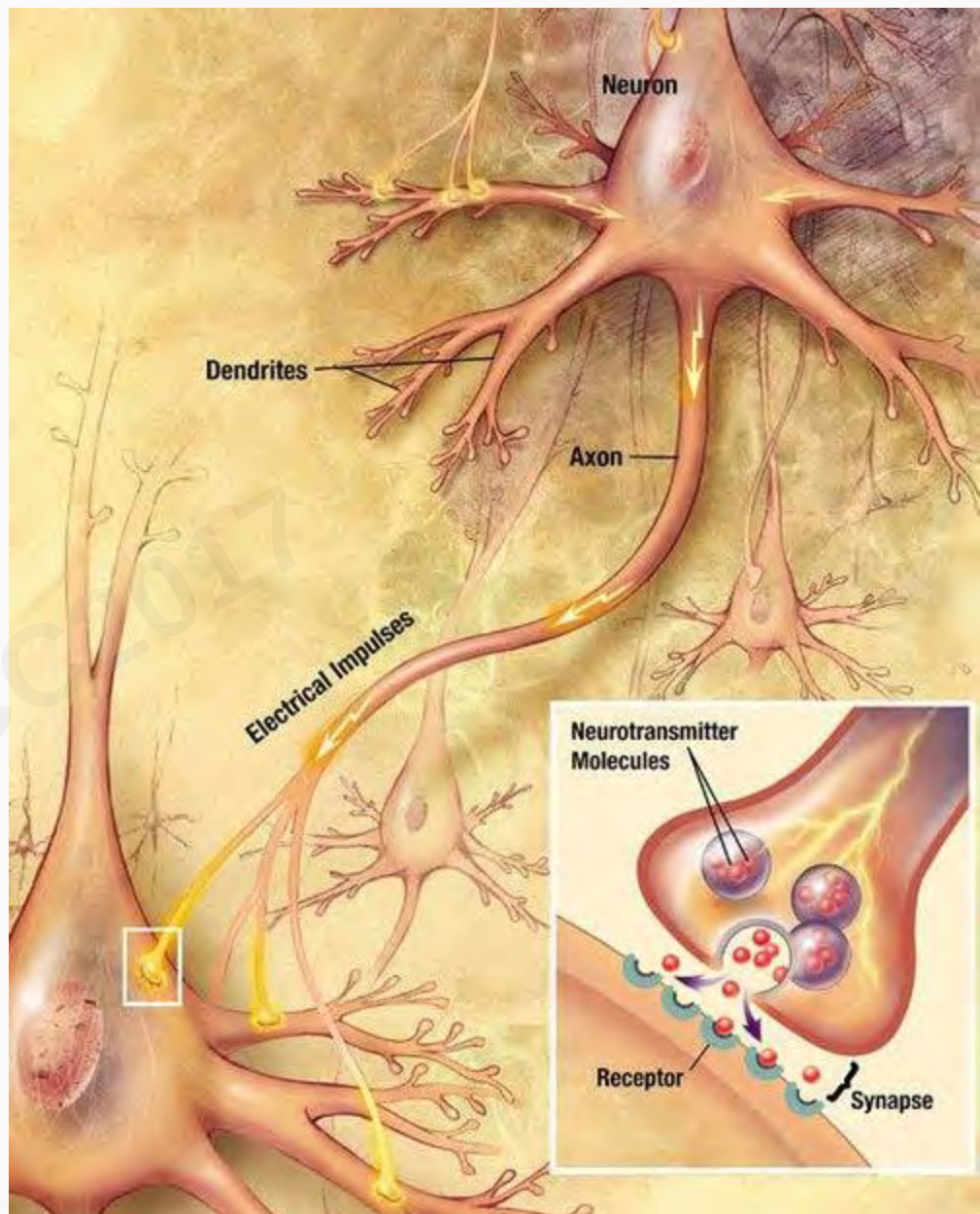
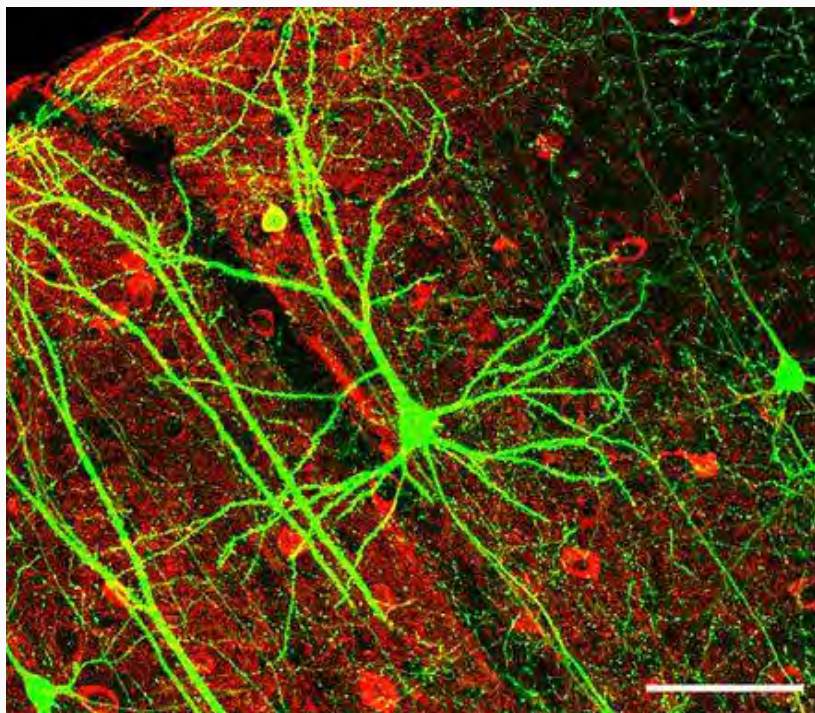
语音识别总体流程

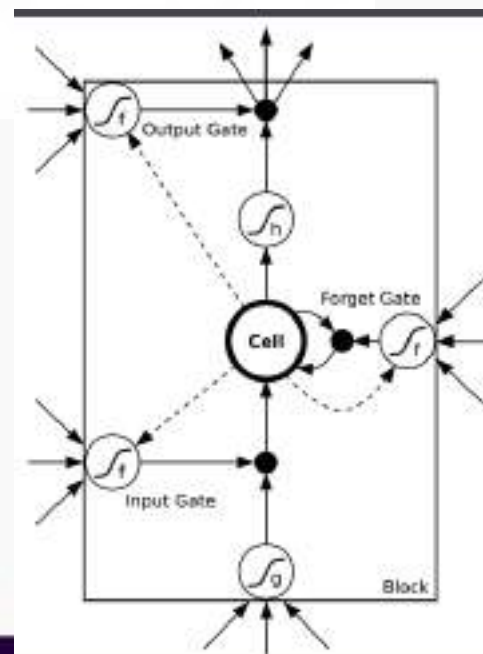
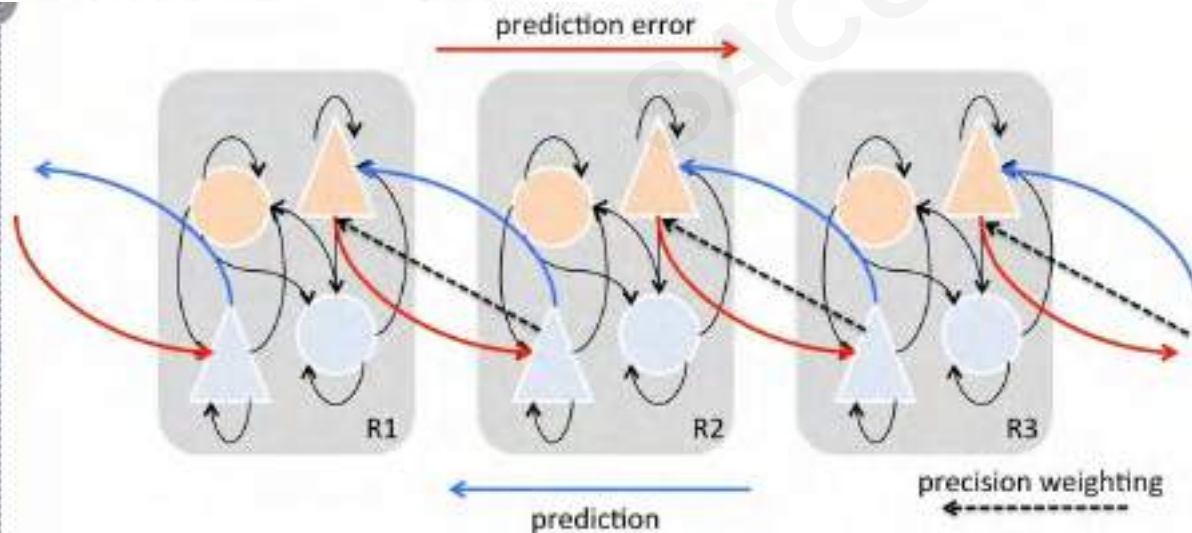
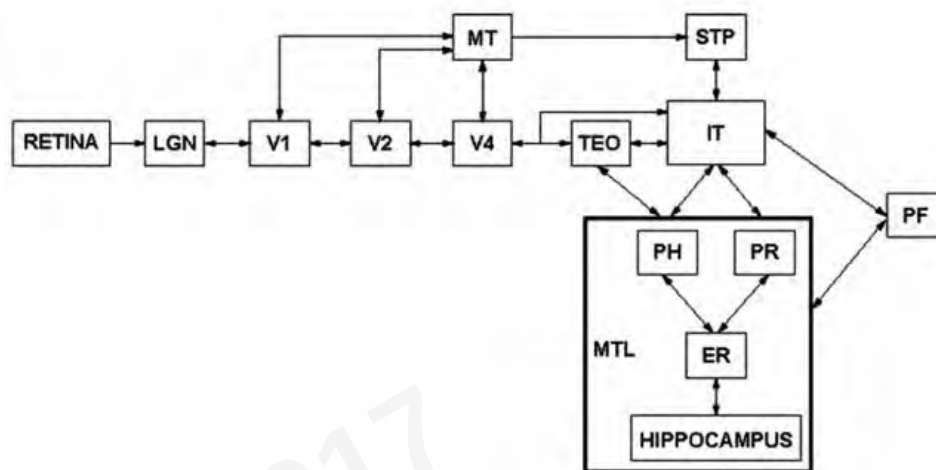
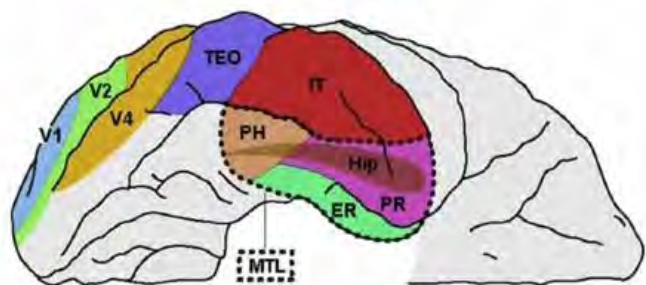
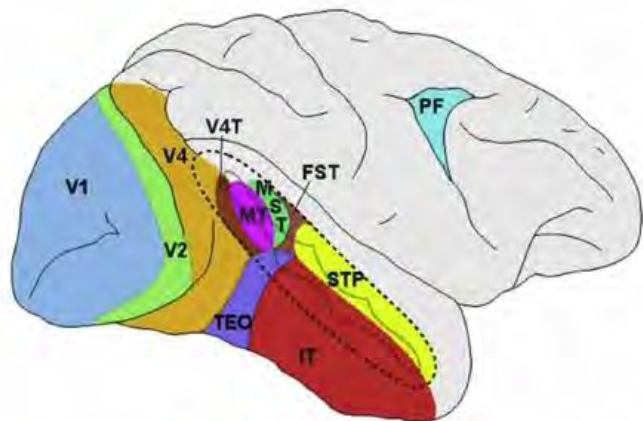


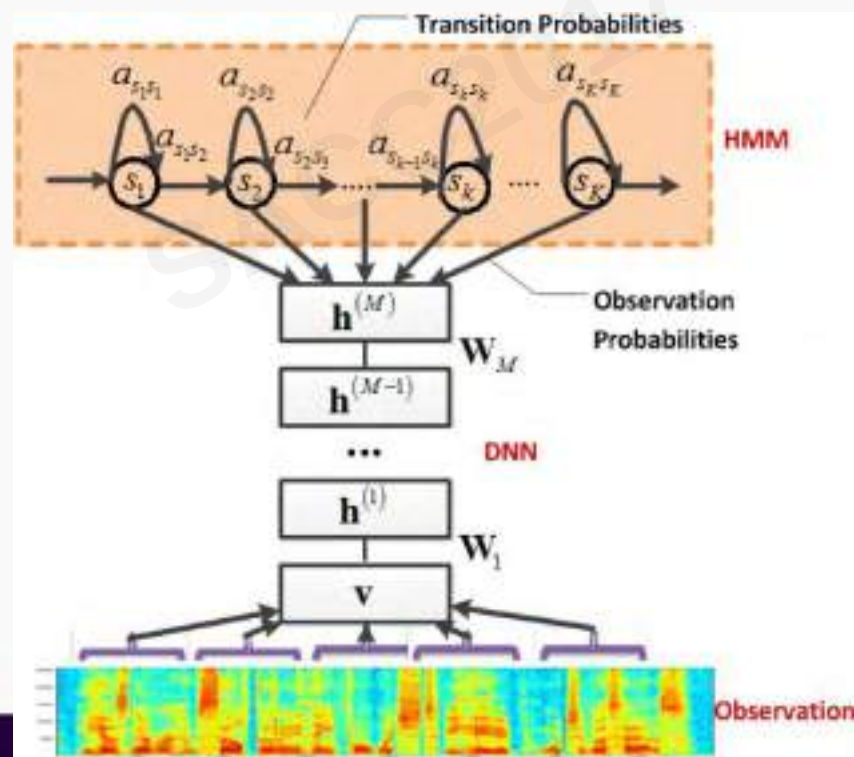
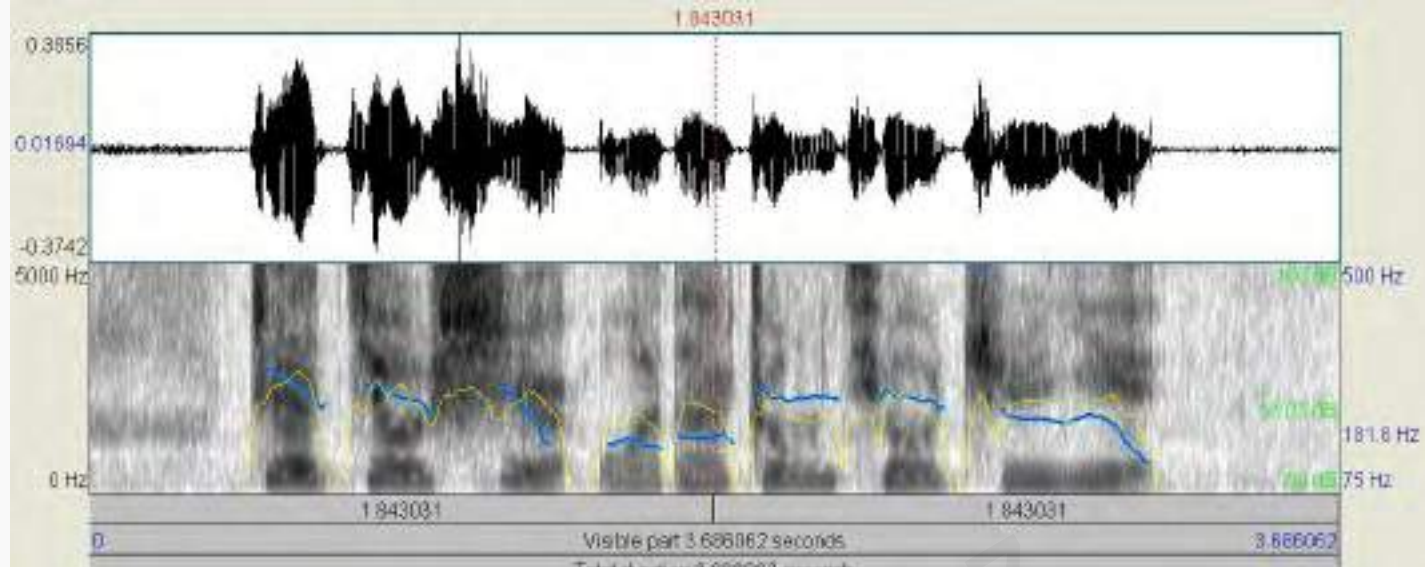
人如何进行 记忆和推理

90亿个神经元

3 million 年才能才能数完







语音识别相关技术

数据 算法 实现

SACC2017

低成本数据方案： 阿里众包团队

关键：标注人+质检

Creating Speech and Language Data With Amazon's Mechanical Turk

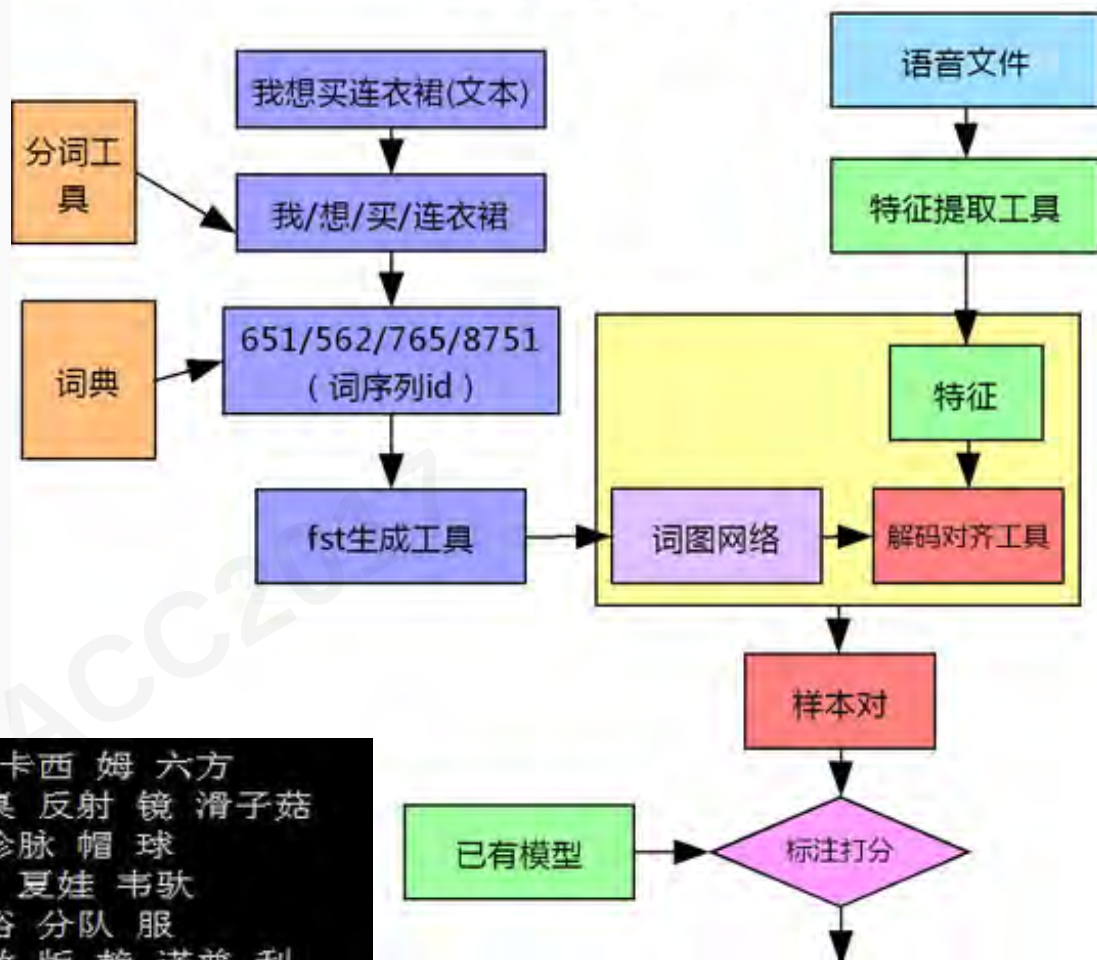
Chris Callison-Burch and Mark Dredze
Human Language Technology Center of Excellence
& Center for Language and Speech Processing
Johns Hopkins University
ccb, mdredze@cs.jhu.edu

Abstract

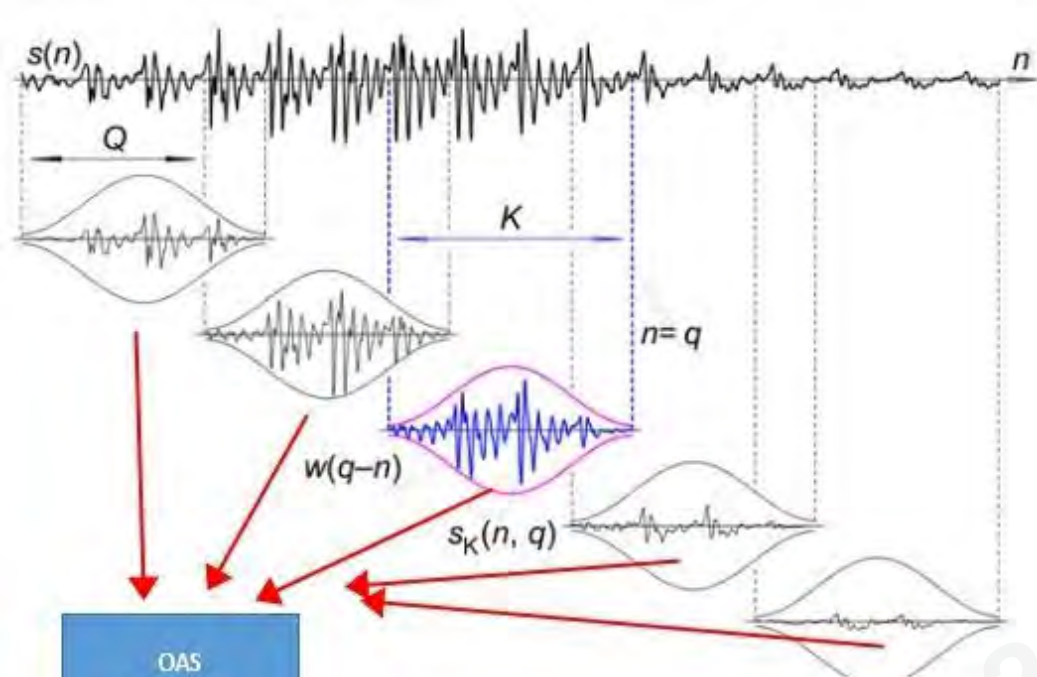
In this paper we give an introduction to using Amazon's Mechanical Turk crowdsourcing platform for the purpose of collecting data for human language technologies. We survey the papers published in the NAACL-2010 Workshop. 24 researchers participated in the workshop's shared task to create data for speech and language applications with \$100.

Turk to spend on an annotation task of their choosing. They were required to write a short paper describing their experience, and to distribute the data that they created. They were encouraged to address the following questions: How did you convey the task in terms that were simple enough for non-experts to understand? Were non-experts as good as experts? What did you do to ensure quality? How quickly did the data get annotated? What is the cost per label? Researchers submitted a 1 page proposal

在线质检流程



RDDIC_PART100000001 选票 小鸡 快跑 卡西姆 六方
RDDIC_PART100000002 锅 苏泊尔 食堂 桌 反射 镜 滑子菇
RDDIC_PART100000003 紫 牡丹 天津市 诊脉 帽 球
RDDIC_PART100000004 竹叉 德鲁伊 亚当 夏娃 韦驮
RDDIC_PART100000005 峡谷 珍珠 托 粟裕 分队 服
RDDIC_PART100000006 录像 盒 献忠 高教版 赖 诺普 利
RDDIC_PART100000007 雷 哈 逗留 超博 缩 管
RDDIC_PART100000008 糖化酶 菲 纹 张翰 音王 子
RDDIC_PART100000009 五金 包 黄梅 县 陌上 纳斯达克
RDDIC_PART100000010 鹅毛 扇 呼 朋 唤 友 切断 新娘 款



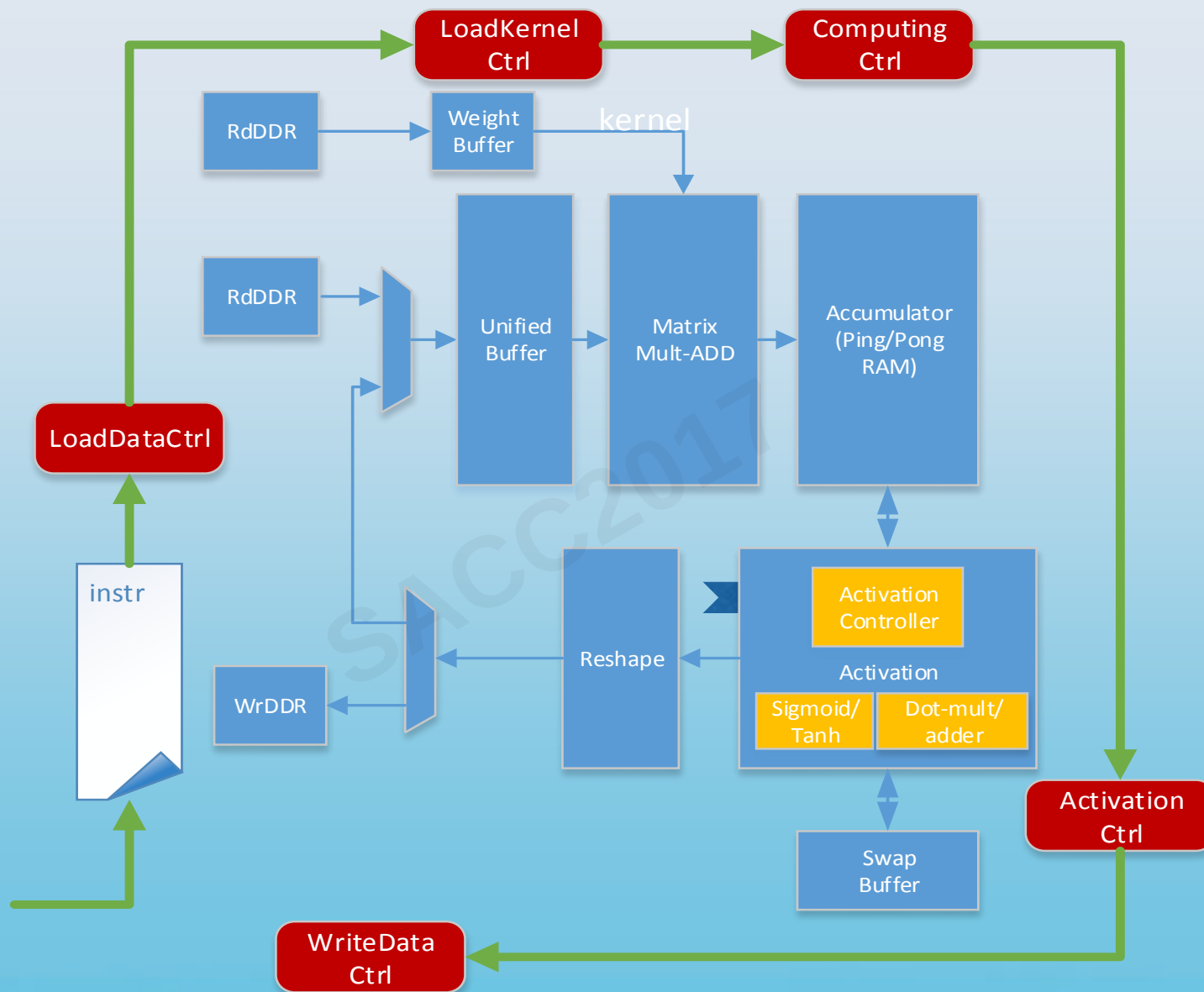
OAS
特征服务

CPU, FPGA, GPU

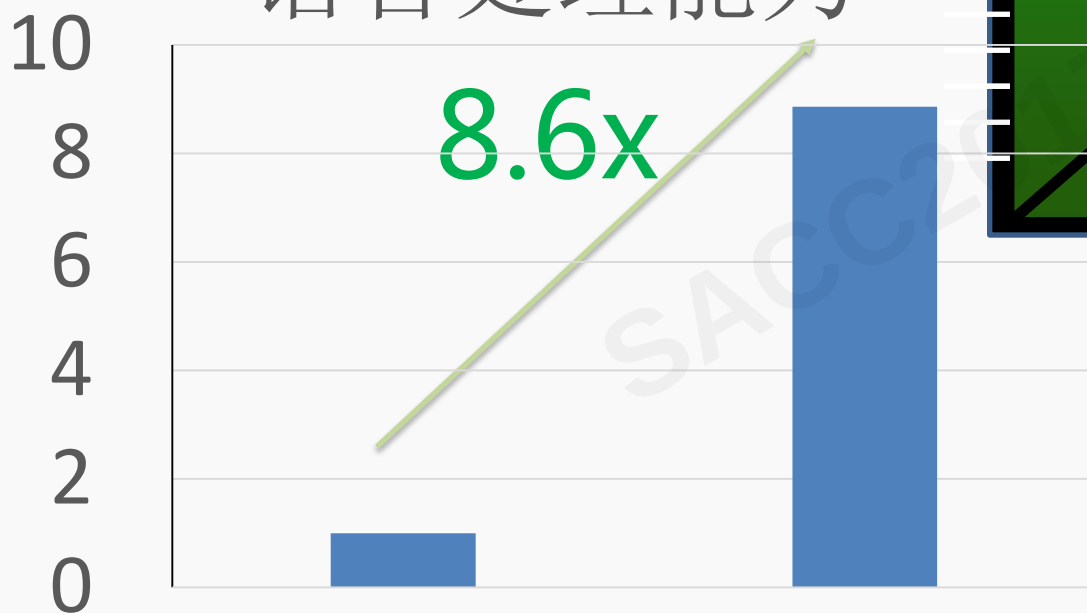
OASDNN
计算服务

OAS speech
Merge 服务



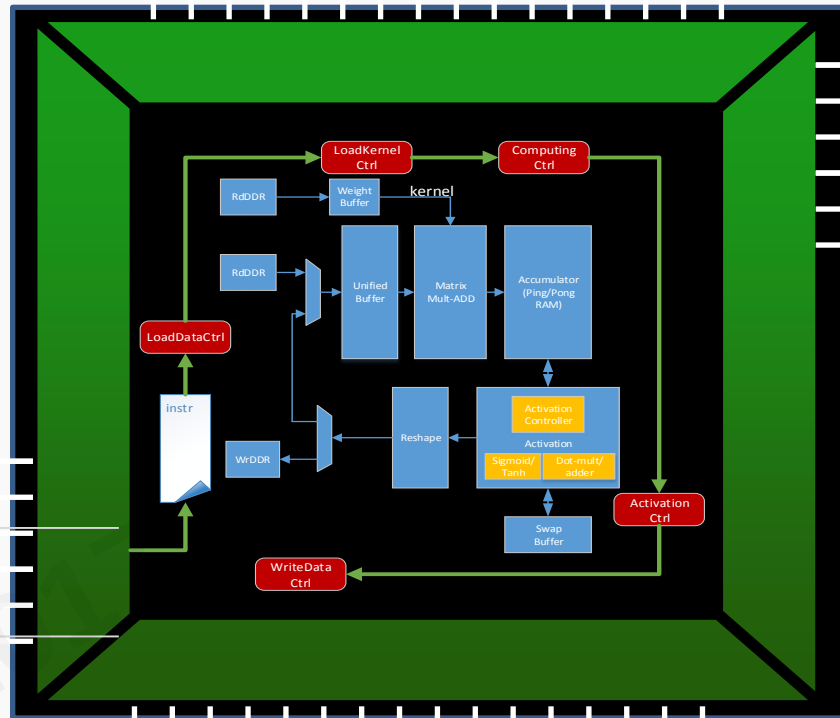


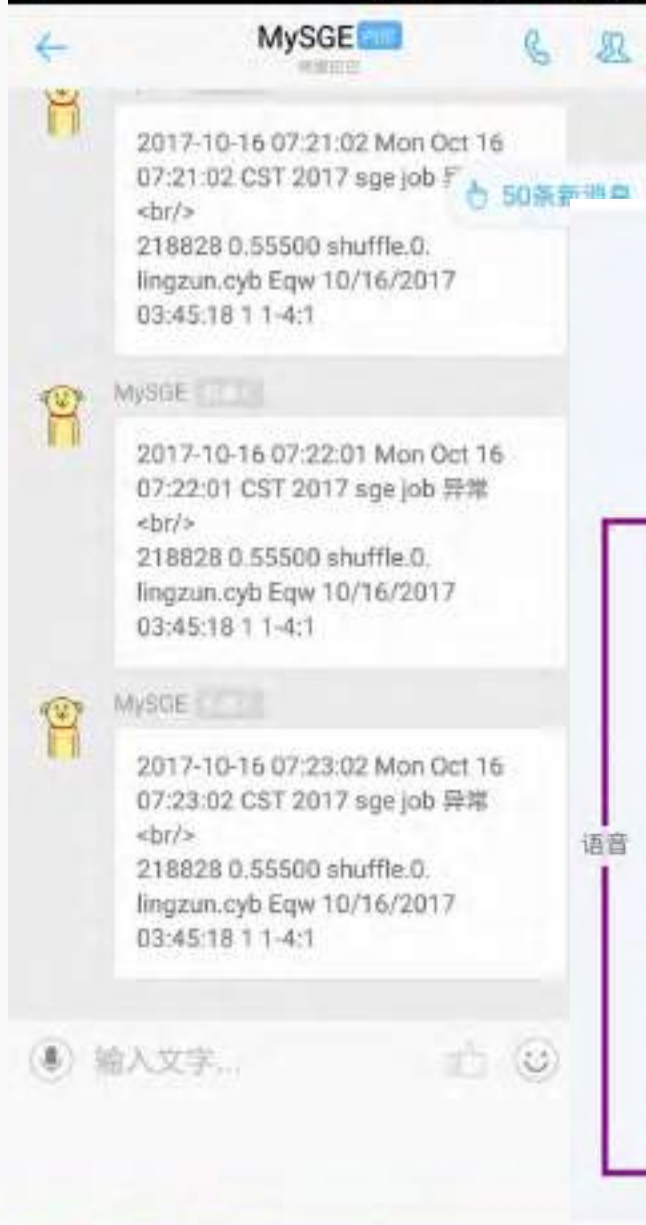
语音处理能力



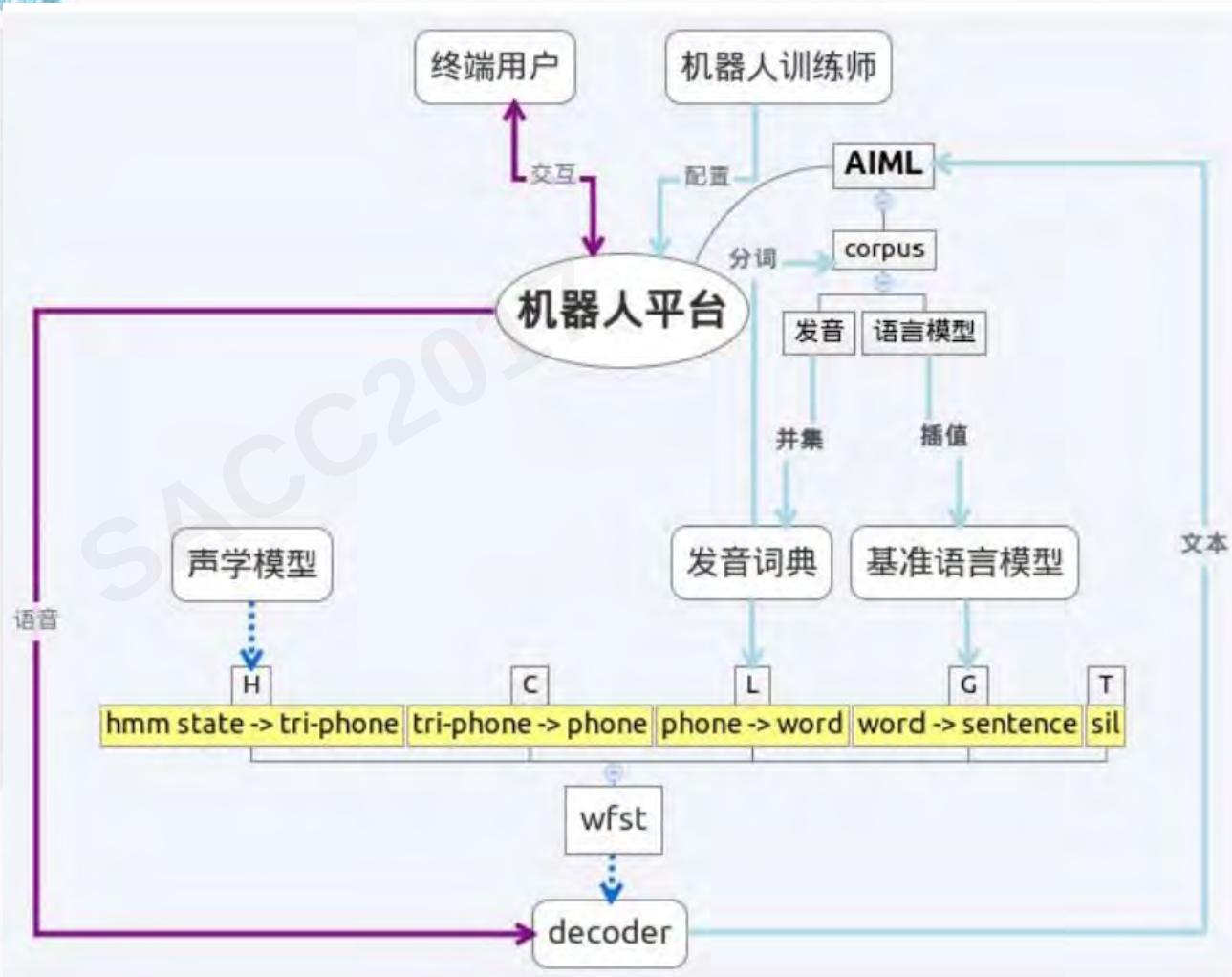
High End 2
Sockets Server

AliFPGA





机器人自动值守训练



云智未来^{9th}

第九届中国系统架构师大会
SYSTEM ARCHITECT CONFERENCE CHINA 2017

语音识别相关应用



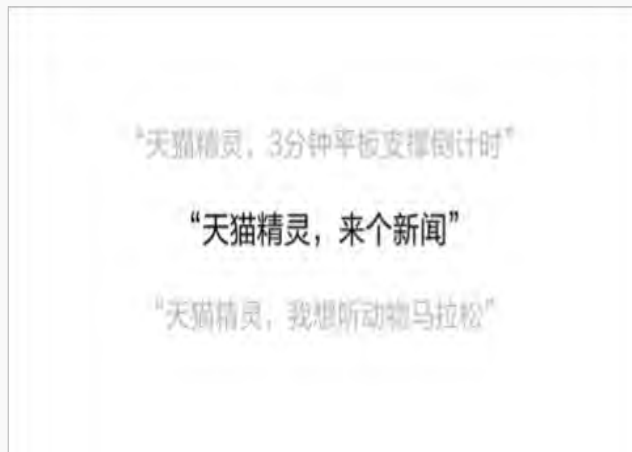




从有趣

到实用

满足不同年
龄不同职业
用户



睡前故事、儿歌、算术、
成语解释、诗词歌赋

儿童教育

购物、叫外卖、手机充值
等声纹支付轻松消费

外卖充值购物

健身、音乐、脱口秀、广
播、笑话、玩游戏

休闲娱乐

天气、星座运势、名人查
询、你想知道的

百科查询

空调、窗帘、灯、加湿器、
电视等上千种智能家电

智能家居控制

闹钟、日程提醒、抛硬币、
报菜谱、表白、每日鼓励

其他黑科技

A photograph of a woman and a young girl. The woman is in the foreground, smiling broadly with her hands covering her eyes. The girl is behind her, also smiling. The background is softly blurred, suggesting an indoor setting with natural light.

声纹识别

——记得你的声音



你的小秘书

效率工具

订日程 / 备忘提醒 /
定闹钟

查天气 / 找手机

购物助手

话费充值 / 日常生活
用品购买

健身教练

倒计时 / 卡路里查
询

信息查询

听新闻 / 查星座运
势

查价格 / 查快递 /
查股票

你孩子的小老师

唐诗宋词 / 成语解
释
听儿歌 / 讲故事
算算术 / 十万个为
什么





你的娱乐和家 电管家

家居控制 / 场景控制 笑话 / 毒鸡汤
 场景化设置 / 全屋控制 音乐库 / 音频 / 有
 魔盒控制 / 智能电视控声书
 制 海量视频 / 儿童内
 容

超强云端功能服务

更多功能
持续开放
中...



智能家居



找手机



音乐



笑话



魔盒控制



幼儿儿歌



娱乐节目



脱口秀



有声书



语音购物



查天气



叫外卖



娱乐新闻



设置闹钟



星座运势



幼儿故事



手机充值



备忘录



日程提醒



商业财经



时间提醒

“天猫精灵”

“我在，你说”



计算 Computing

算法 Algorithm

数据 Data

FPGA 云端硬件定制
化逻辑电路加速
前端云端联合优化

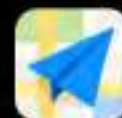
全球领先的语言特征表达
KDD 收录全球首创的混合神经网络
conv-RNN 的通用推理混合框架
深度优化的强化学习策略
个性化产生对话策略

AliGenie 认识的实体超过 1 亿+
中文语言标注标准 CTB
用户画像 / 语境和上下文
同义词 / 反义词 / 上下位 / 口语化表达

AliGenie 合作伙伴



阿里云IoT





AliGenie
inside



西溪天堂



喜来登



悦榕庄



西轩



悦椿



布鲁克

内容服务商

Content Provider

内容快速接入

Fast Access System

虾米音乐

喜马拉雅FM

微博

YOUKU

Radio.cn

中信书院

新华智云

babytree

从心EVA

爱奇艺

腾讯视频

优酷

应用提供商

Service Provider

专人的定制化接入

In-person Service

盒马

肯德基

饿了么

淘票票

高德地图
amap.com

淘宝网
taobao.com

支付宝
ALIPAY

阿里通信
ALITELECOM

中国南方航空
CHINA SOUTHERN

K
keep

飞猪

T.MALL 天猫

双十一全球狂欢节

天猫精灵

99元开启人工智能

领券499减400

全年仅限一次

限量抢购



券后价 ¥99

立即预定 >

预售时间：10月20日00:00 - 11月10日22:00



智能家居设备商

Smart Home

设备快速接入

Fast Device System

美的 Midea

PHILIPS

SIEMENS

AUX 奥克斯

海尔 Haier

LifeSmart

ECOVACS

Honeywell

美的 Midea

阿里通信

Galanz 格兰仕

美的 Midea

SACC
2017

云智未来^{9th}

ITPUB

ChinaUnix

ITPUB

硬件方案

Hardware Solution

+

支持芯片

Chip Solution

+

技术服务

Technique Support



汽车

CAR



家居

HOME



随身设备

PORTABLE DEVICE



案例

会说话的托马斯火车

TALKING TOYS

THANKS

The background is a dark, deep blue gradient. It features several dynamic, glowing trails of small, bright blue particles or dust. These trails originate from a bright, white-to-yellow light source located near the center-left, just below the word 'THANKS'. From this point, the particle trails extend outwards in various directions, creating a sense of motion and depth. The overall aesthetic is futuristic and ethereal.