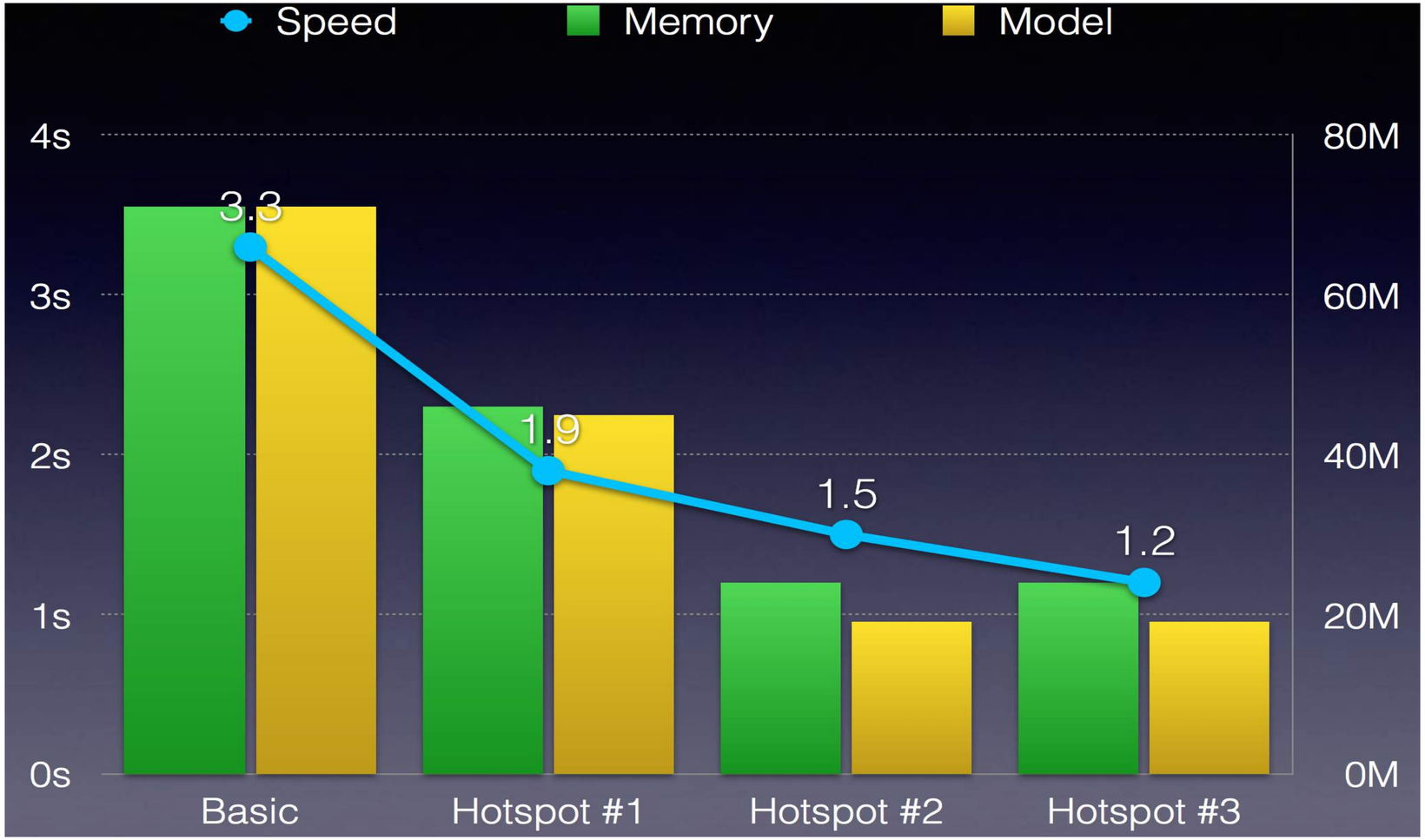
kika


Speed
Memory
Model
4s
3s
2s
1s
0s
80M
60M
40M
20M
0M
3.3
1.9
1.5
1.2
Basic
Hotspot #1
Hotspot #2
Hotspot #3

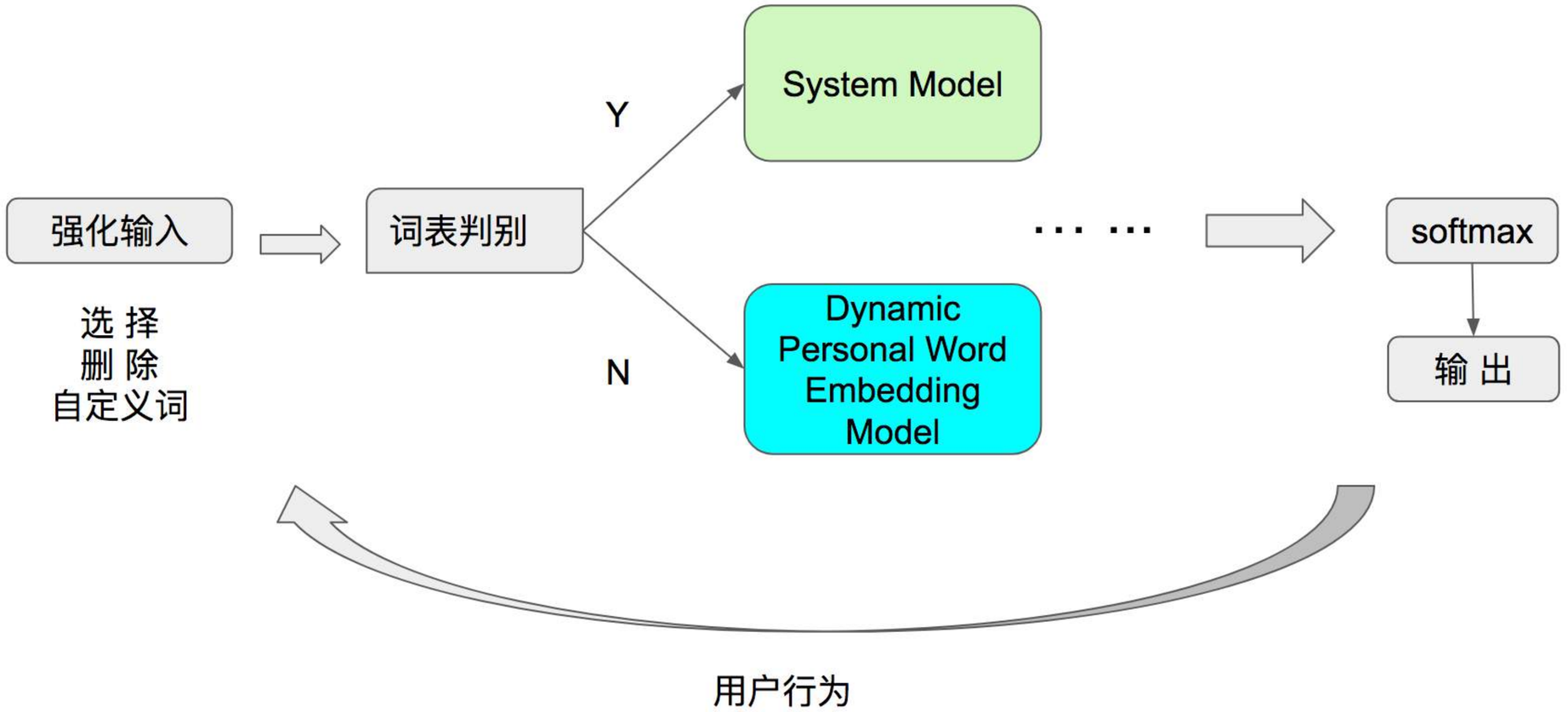
kika

Client - 个性化强化学习

- 关键问题
  - 高效增量计算
  - 个性词频
  - 个性词表

kika

Client - 强化学习模型基本结构



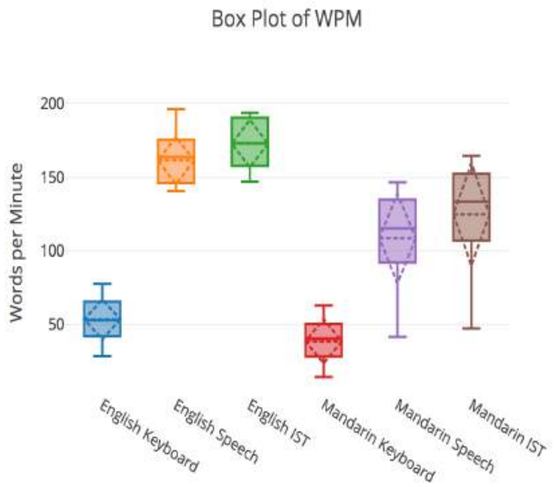
kika

TABLE OF CONTENTES

kika

Why Voice

- Voice Input 10X Typing Input
- ASR Is not enough

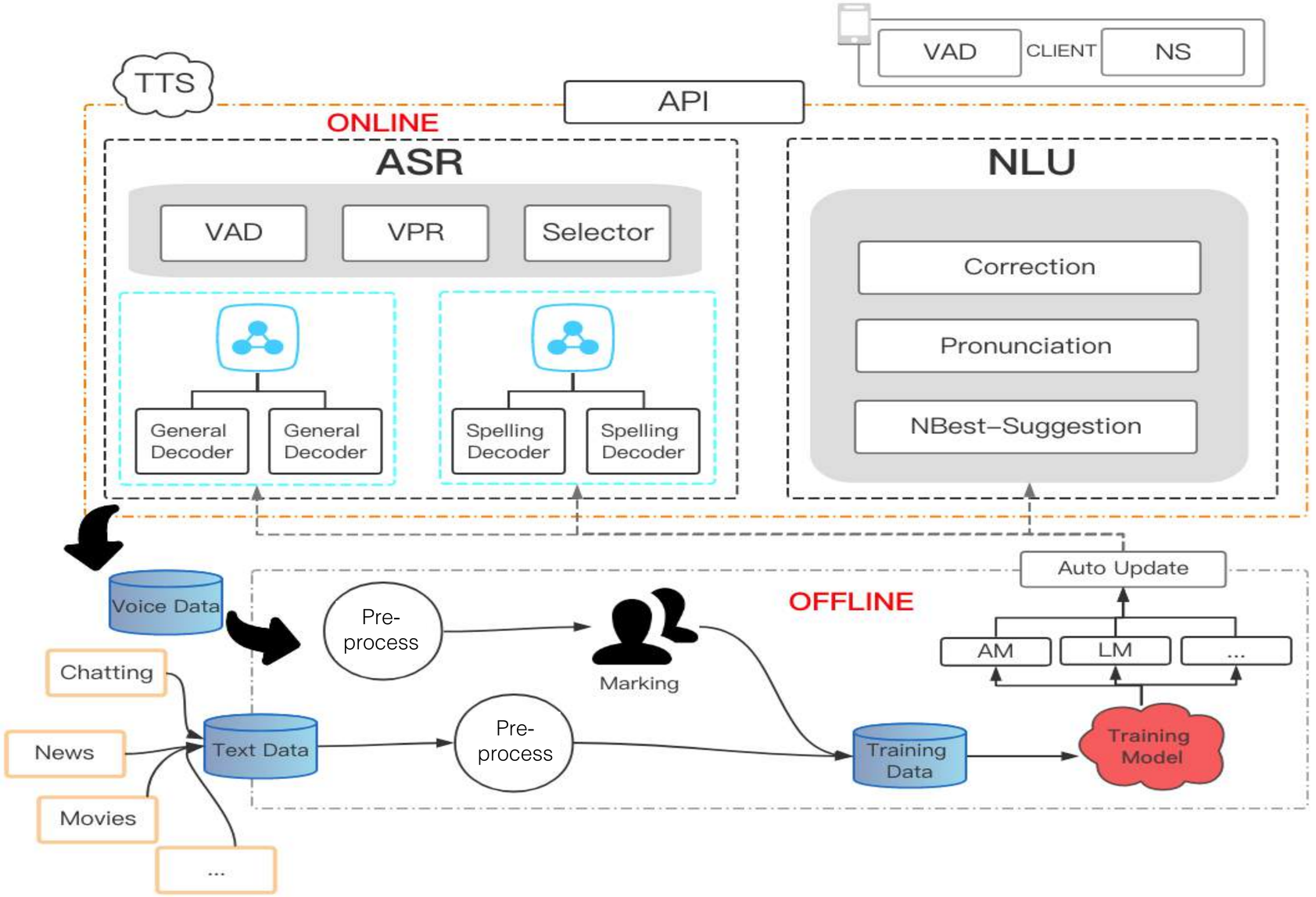


kika

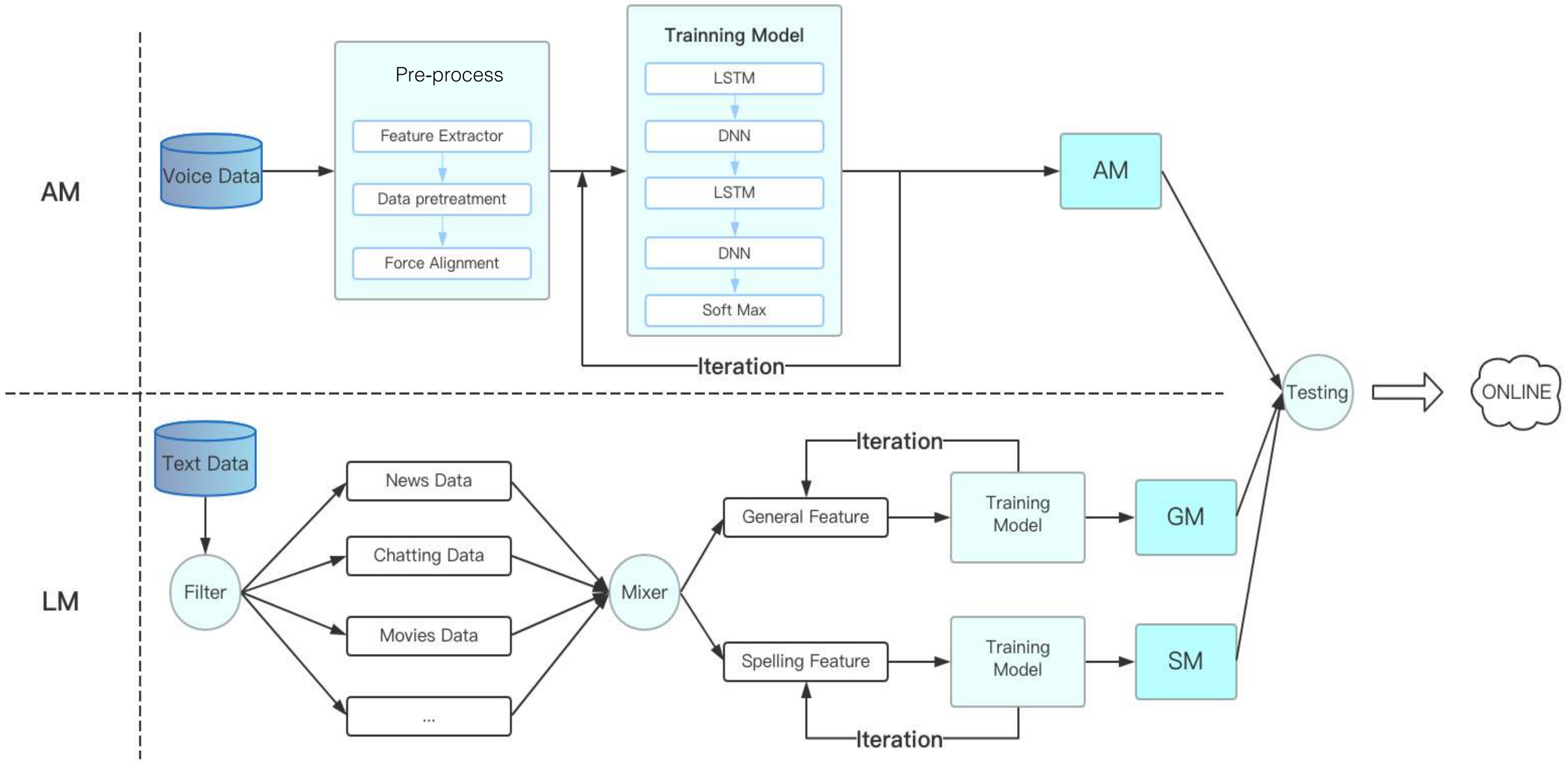
Why Voice Only

- What's Voice Only
- Key Technology
  - 语音识别
  - 语音纠错
  - 语音编辑

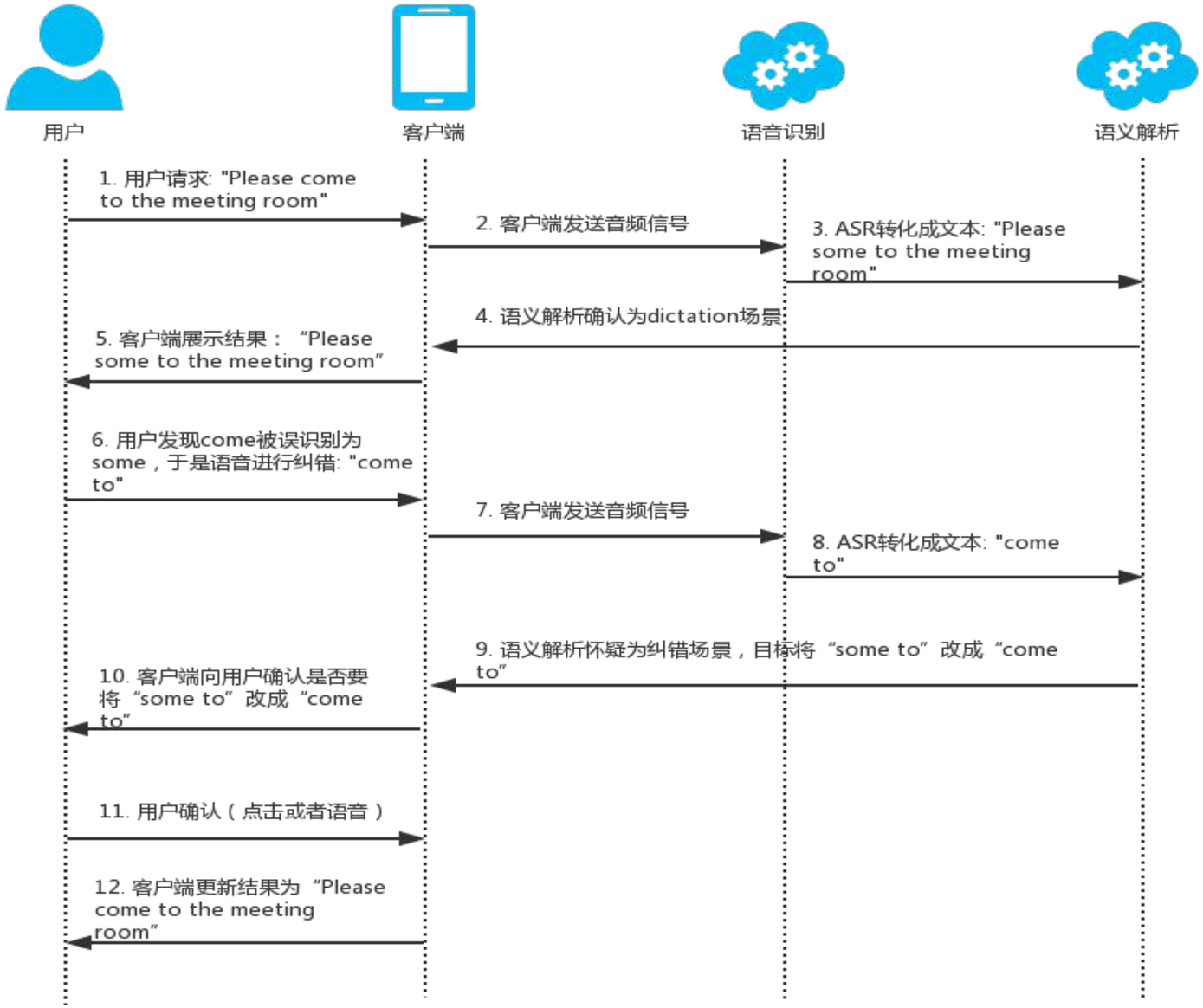
kika


VAD
CLIENT
NS
TTS
API
ONLINE
ASR
VAD
VPR
Selector
General Decoder
General Decoder
Spelling Decoder
Spelling Decoder
NLU
Correction
Pronunciation
NBest–Suggestion
Auto Update
OFFLINE
Voice Data
Pre-process
Marking
AM
LM
...
Chatting
News
Movies
...
Text Data
Pre-process
Training Data
Training Model

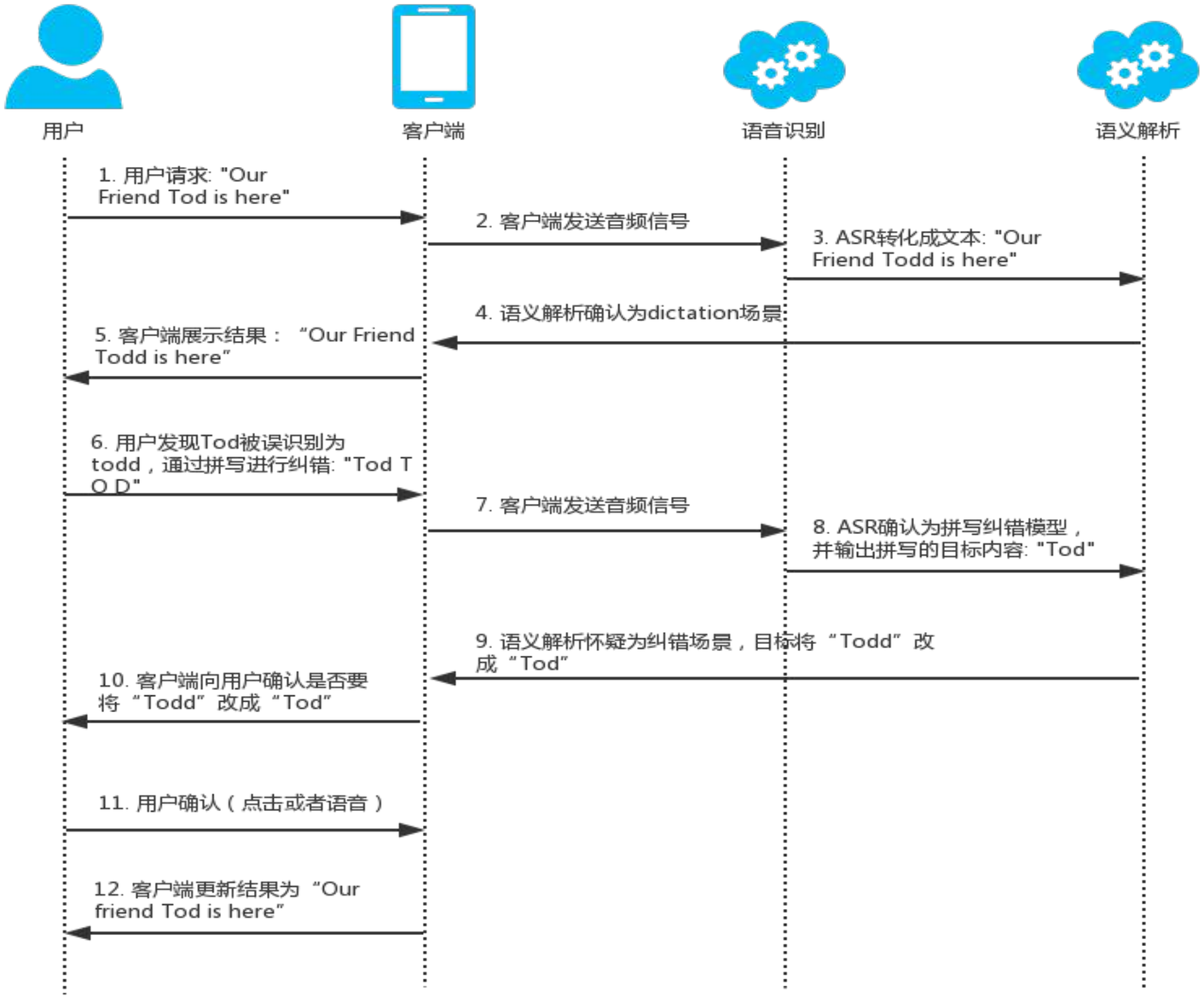
kika


AM
Voice Data
Pre-process
Feature Extractor
Data pretreatment
Force Alignment
Trainning Model
LSTM
DNN
LSTM
DNN
Soft Max
Iteration
AM
Testing
ONLINE
LM
Text Data
Filter
News Data
Chatting Data
Movies Data
...
Mixer
Iteration
General Feature
Training Model
GM
Spelling Feature
Training Model
SM
Iteration

kika


用户
客户端
语音识别
语义解析
1. 用户请求: "Please come to the meeting room"
2. 客户端发送音频信号
3. ASR转化成文本: "Please some to the meeting room"
4. 语义解析确认为dictation场景
5. 客户端展示结果： “Please some to the meeting room”
6. 用户发现come被误识别为some，于是语音进行纠错: "come to"
7. 客户端发送音频信号
8. ASR转化成文本: "come to"
9. 语义解析怀疑为纠错场景，目标将“some to”改成“come to”
10. 客户端向用户确认是否要将“some to”改成“come to”
11. 用户确认（点击或者语音）
12. 客户端更新结果为“Please come to the meeting room”

kika

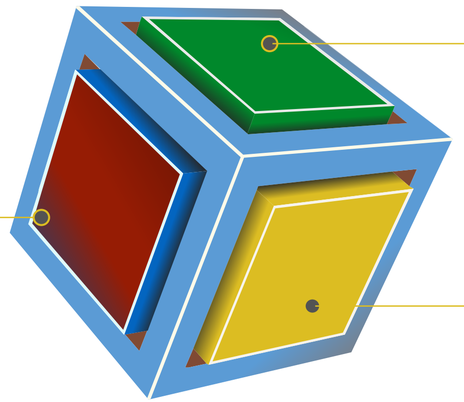

用户
客户端
语音识别
语义解析
1. 用户请求: "Our Friend Tod is here"
2. 客户端发送音频信号
3. ASR转化成文本: "Our Friend Todd is here"
4. 语义解析确认为dictation场景
5. 客户端展示结果： “Our Friend Todd is here”
6. 用户发现Tod被误识别为todd，通过拼写进行纠错: "Tod T O D"
7. 客户端发送音频信号
8. ASR确认为拼写纠错模型，并输出拼写的目标内容: "Tod"
9. 语义解析怀疑为纠错场景，目标将 “Todd” 改成 “Tod”
10. 客户端向用户确认是否要将 “Todd” 改成 “Tod”
11. 用户确认（点击或者语音）
12. 客户端更新结果为 “Our friend Tod is here”

kika

“3A”引擎

数据

4亿用户的语音和文本数据
多场景上万小时音视频数据



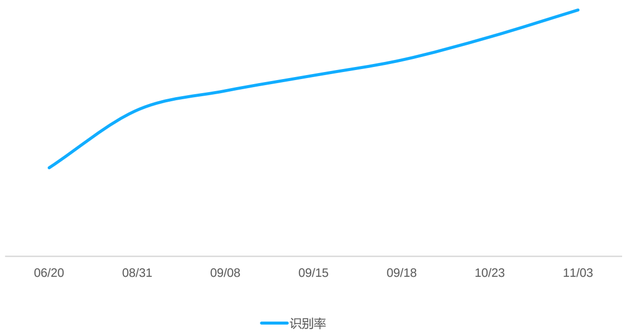
技术

深度学习引擎

场景

基于对话沟通延展开的多种场景，包括语音输入，商务会议，医生问诊，车载沟通

语音识别准确率提升



精准语义理解能力

在语义模糊时，根据上下文正确理解用户意图

- Dictation Intent
- Correction Intent
- Spelling Intent
- …

表达增强

基于精确的语义理解，智能增强用户表达

用户：do you want to have dinner
at 8:00?

AI 识别： do you want go at 8:00?

kika

TABLE OF CONTENTES

kikaGO

CES INNOVATION AWARDS 2018 HONOREE

CES 2018 Innovation Award Winner

- In-Vehicle Audio/Video
- Tech for a Better World
- Software Mobile Apps
- Smart Home

kika

AI Lab

- Language Model Adaptation
- Unsupervised Acoustic Model
- RNN Model Adaptation on Mobiles

kika

TABLE OF CONTENTES

kika

Alps

- Forward Prediction VS. Backward Correction
- Performance optimization on Android Go
- Sentence-level Paraphrasing

kika

Appalachian

- From ASR Dictation to Application Adaptation
- Semi-Supervised Acoustic Model training
- Oral voice to formal text transformation

kika

Thanks!