



Узнай меня по голосу

Как расширить возможности виртуального ассистента при помощи голосовой биометрии

CONVERSATIONS AI - 2020



Распознавание дикторов

Основные сценарии в голосовых ассистентах



#1: Защита от случайной активации по Wake-Up слову



#3: Голосовая Аутентификация для финансовых транзакций



#5: Персонализация голосовых команд



#2: Идентификация пользователей и персонализация



#4: Разблокировка устройства голосом

Кейс:

Защита от случайной активации WuW

Защита от случайной активации Wake-Up Word

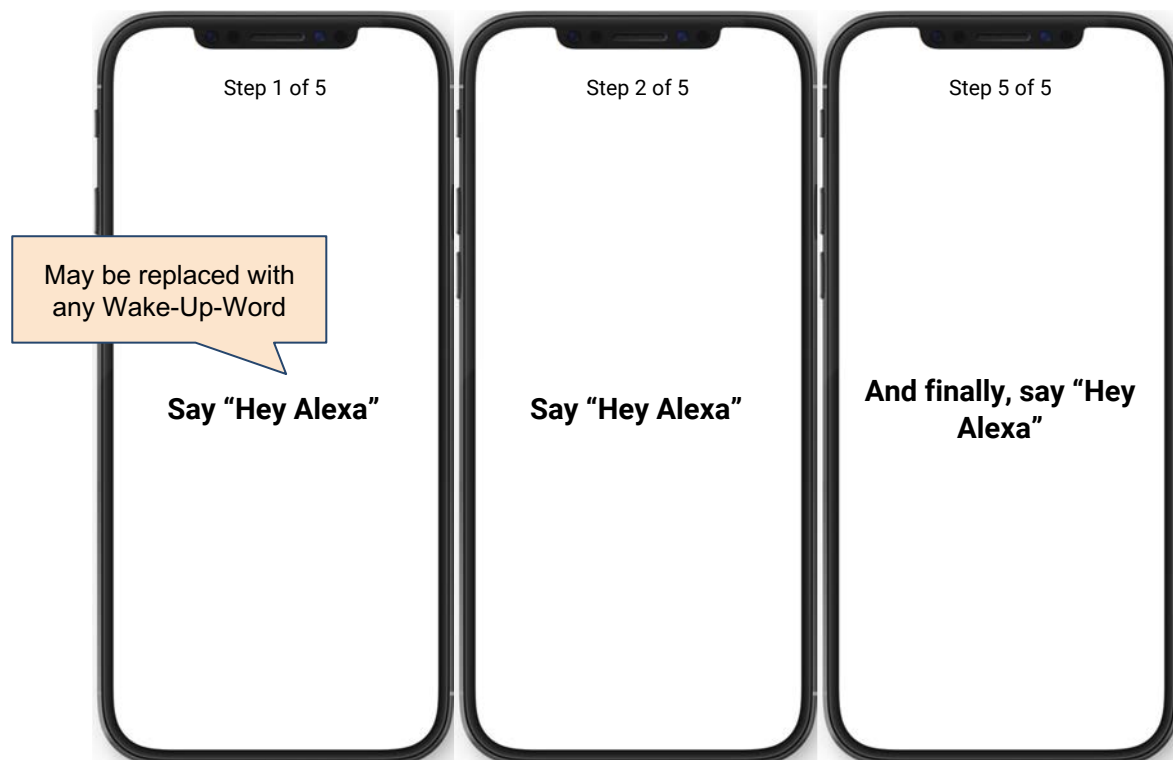


Голосового ассистента легко случайно “разбудить”. Например, фраза “ОК, Гугл” или имя “Алисы” может прозвучать по радио или ТВ

Защита от случайной активации WuW использует распознавание диктора для:

- Снижения потребления CPU (особенно важно для автономных/носимых устройств)
- Улучшение UX за счёт снижения ложных активаций
- Улучшение UX за счёт снижения вероятности пропуска Wake-Up Word (при интеграции с детектором)

Регистрация голоса



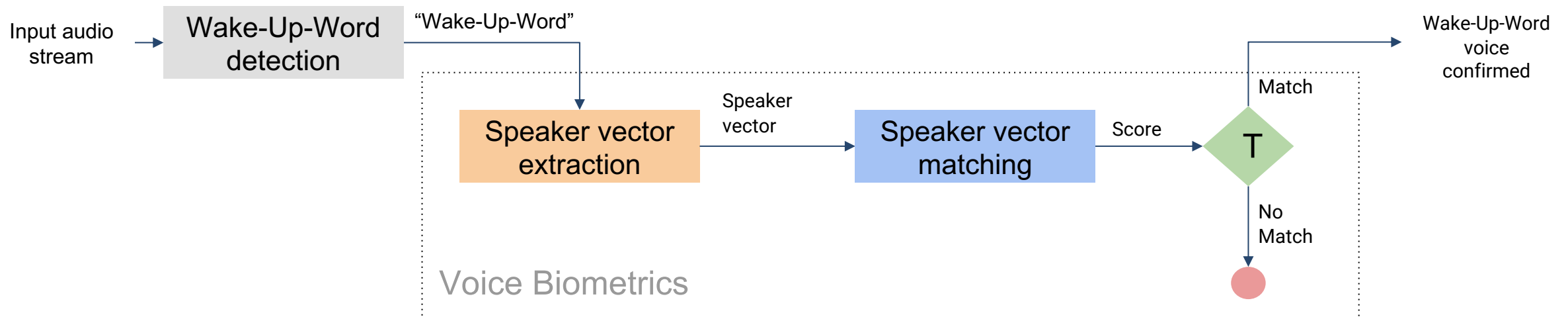
Использование голоса пользователя при стандартной настройке голосового ассистента:

- Стандартный процесс, как у "Siri"
- 5 повторений Wake Up Word
- Собранные команды направляются в биометрический процессор для создания "слепок" голоса

Верификация голоса



Голосовая биометрия работает в пассивном режиме каждый раз, когда детектирован Wake-Up Word.



Технические требования



Основные требования

- Активная регистрация голоса пользователя - 5 повторений **Wake-Up-Word** (может использоваться голос, собранный при настройке голосового ассистента)
- Дистанция < 1 м
- EER < 4% (Сравнимо с Siri ¹)
- **Miss Rate (False Rejection) < 1%**, False Acceptance < 8%

¹<https://machinelearning.apple.com/2018/04/16/personalized-hey-siri.html>

Кейс: Идентификация и Персонализация

Идентификация Пользователей



Голосовые ассистенты часто используются в общих устройствах разными пользователями, например, членами семьи или офиса.

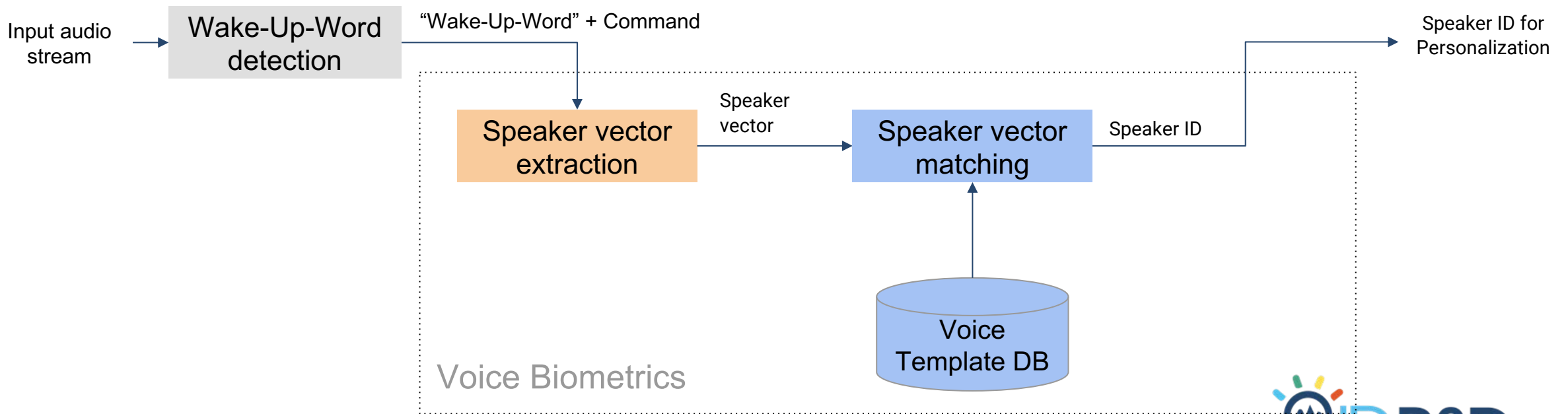
Идентификация пользователей позволяет:

- Уменьшить нагрузку на CPU
- Персонализация ответов — корректная обработка таких запросов, как “включи мою музыку” или “прочитай мой календарь и напоминания”
- Ограничить доступ известным пользователям (например, детям) и неизвестным дикторам (гостям)

Верификация голоса



Голосовая биометрия работает в пассивном режиме каждый раз, когда детектирован Wake-Up Word.



Технические требования



Основные требования

- Активная регистрация голоса пользователя - 5 повторений **Wake-Up-Word** (может использоваться голос, собранный при настройке голосового ассистента)
- Пассивная регистрация предпочтительна (~20 повторений на пользователя)
- Дистанция < 3 м
- Точность Идентификации 99% для списка из 5 человек
- Точность Верификации 94%

Кейс:

Голосовая аутентификация финансовых транзакций

Голосовая Аутентификация Финансовых Транзакций

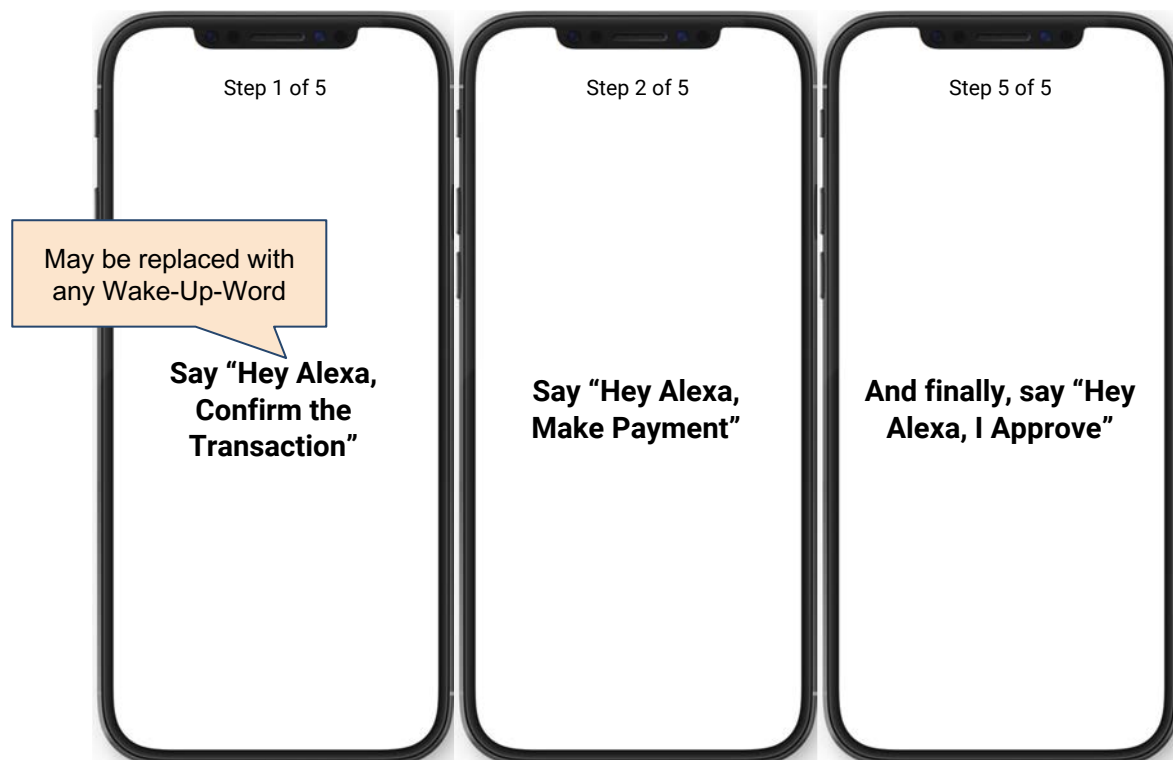


Голосовые ассистенты широко используются для совершения покупок, но насколько это безопасно? В марте 2020 года Google представил возможность подтверждения оплаты голосом.

Голосовая аутентификация может использоваться для улучшения UX и безопасности при совершении:

- Нерисковых финансовых транзакций
- Для подтверждения нефинансовых транзакций

Регистрация голоса



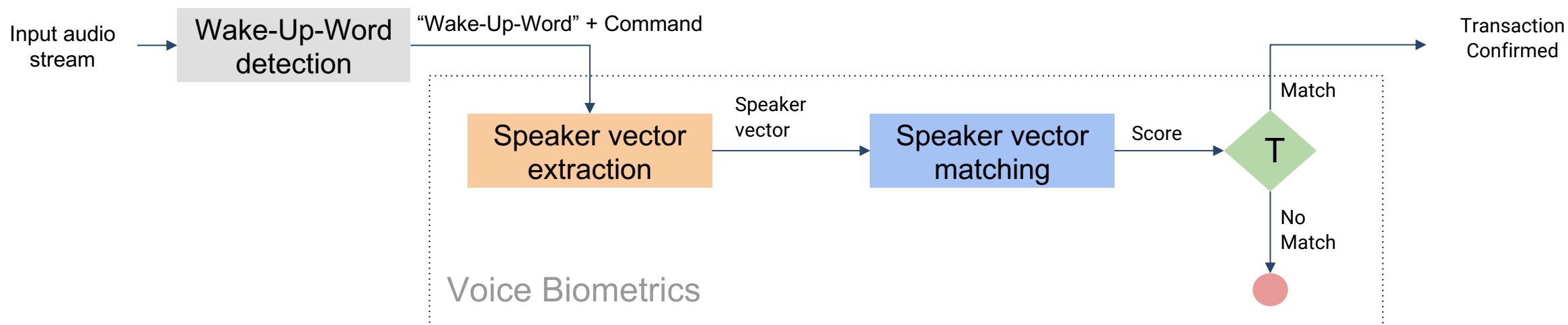
Использование голоса пользователя при стандартной настройке голосового ассистента:

- Стандартный процесс, как у "Siri"
- 5 повторений Wake Up Word и случайных команд
- Собранные команды направляются в биометрический процессор для создания "слепок" голоса

Верификация голоса



Голосовая биометрия работает в пассивном режиме каждый раз, когда детектирован Wake-Up Word.



Технические требования



Основные требования

- Активная регистрация голоса пользователя - 5 повторений Wake-Up-Word + случайная команда (может использоваться голос, собранный при настройке голосового ассистента)
- Активная регистрация на основе Статической Фразы из 3-5 слов
- Дистанция < 1 м
- False Acceptance Rate < 1/10000
- False Rejection Rate < 5-10%

Кейс:

Разблокировка устройства голосом

Разблокировка устройства голосом



Персональные устройства обычно разблокируются при помощи пин-кода или биометрией лица/отпечатков. Однако в ряде случаев такой интерактив неудобен или невозможен. Например, за рулём или домашними делами.

Голосовая разблокировка позволяет:

- Улучшить UX за счёт возможности разблокировки голосом
- Реализовать End-to-End Customer Journey в сценариях с подтверждением транзакций

Разблокировка устройства голосом



← Wake Up Word →

← Команда/Вопрос →

**Текстозависимая
верификация**
Длина фразы **менее 1 сек**

Текстнезависимая верификация
по свободной речи
Длина фразы **1-3 сек**

- Короткий Wake Up Word
- Голосовая команда из 2-6 слов
- Задержка < 1 сек
- Типичный кейс: разблокировка устройства голосом, когда заняты руки (например, за рулём)

Регистрация голоса



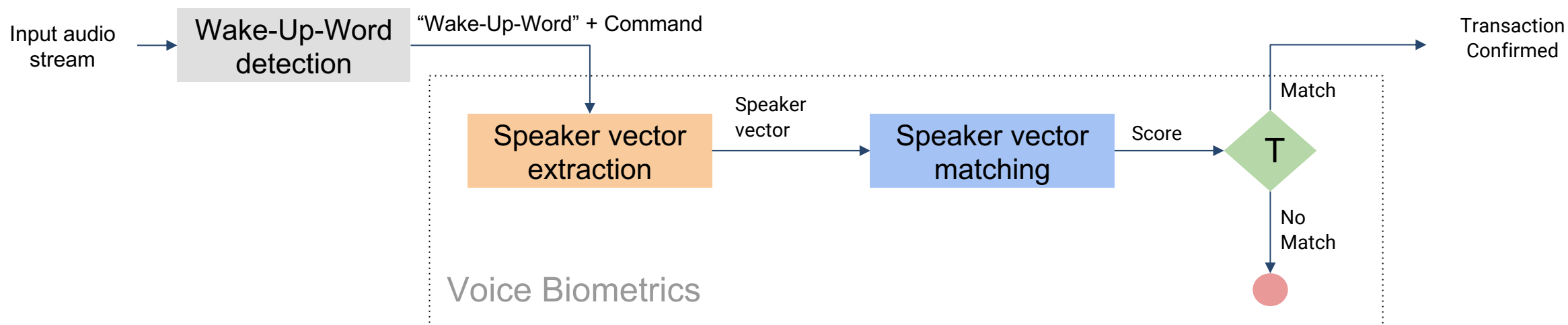
Использование голоса пользователя при стандартной настройке голосового ассистента:

- Стандартный процесс, как у "Siri"
- 5 повторений Wake Up Word и случайных команд
- Собранные команды направляются в биометрический процессор для создания "слепок" голоса

Верификация голоса

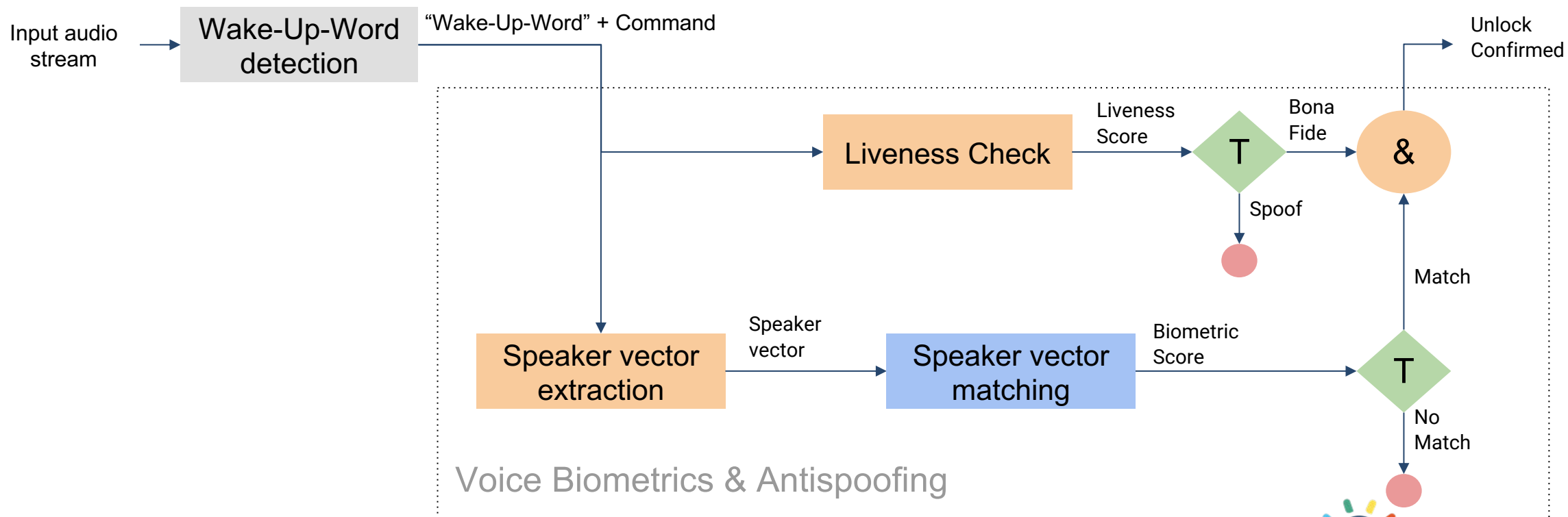


Голосовая биометрия работает в пассивном режиме каждый раз, когда детектирован Wake-Up Word.



... и Анти-Спуфинг!

Голосовая биометрия работает в пассивном режиме каждый раз, когда детектирован Wake-Up Word.



Технические требования

Для использования в Андроиде аутентификация должна соответствовать требованиям, представленным в Android Compatibility Definition Document (CDD). В том числе – защита от спуфинга.

Metric	Android CDD Requirements	Other Market Requirements	Current Status
Spoof Accept Rate (SAR)	<7% (strong), <20% (weak), >20% (convenience)	< 7%	<10%
False Accept Rate (FAR)	1/50k	1/10k	1/10k
False Rejection Rate (FRR)	10%	<5%	<5% for clear cond <15% for SNR>15dB

Кейс:

Персонализация команд для IoT

Персонализация команд для IoT

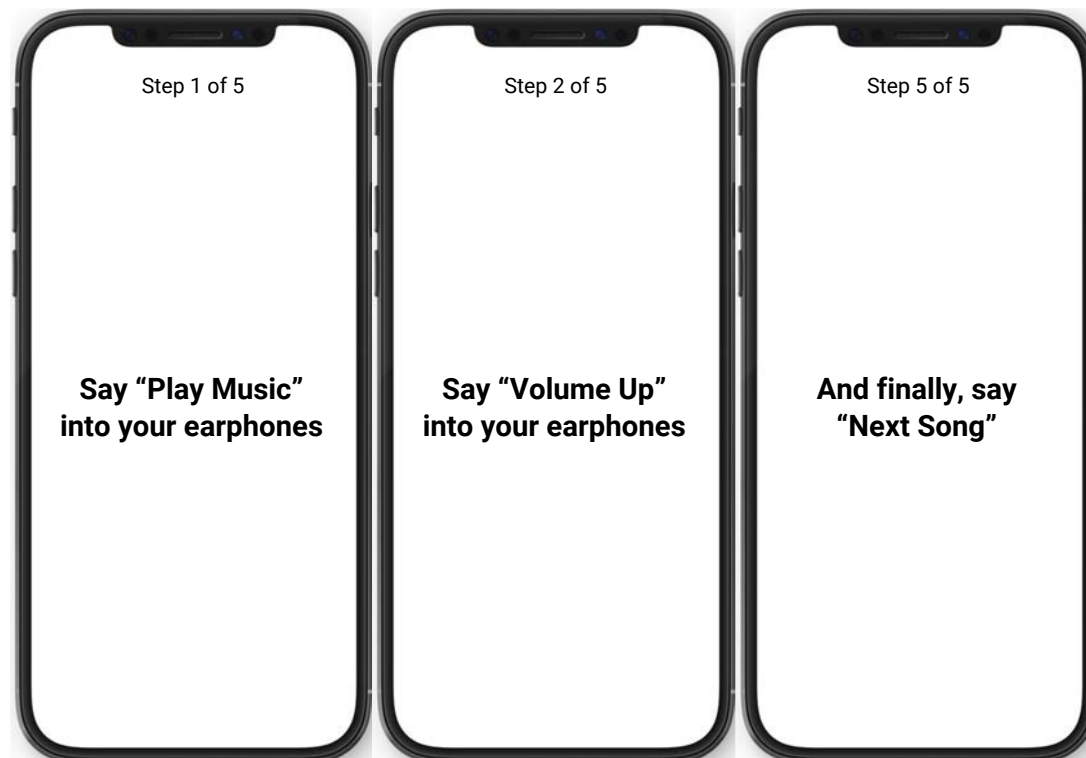


Носимые IoT, такие как наушники или часы, должны отвечать на команды только владельцу. Фразы “Сделай погромче” или “Включи музыку” могут быть произнесены рядом другими людьми или по ТВ.

Персонализация команд использует голосовую биометрию для:

- Улучшение UX за счёт снижения ложных активаций
- Улучшение UX за счёт снижения вероятности пропуска Wake-Up Word (при интеграции с детектором)

Регистрация голоса

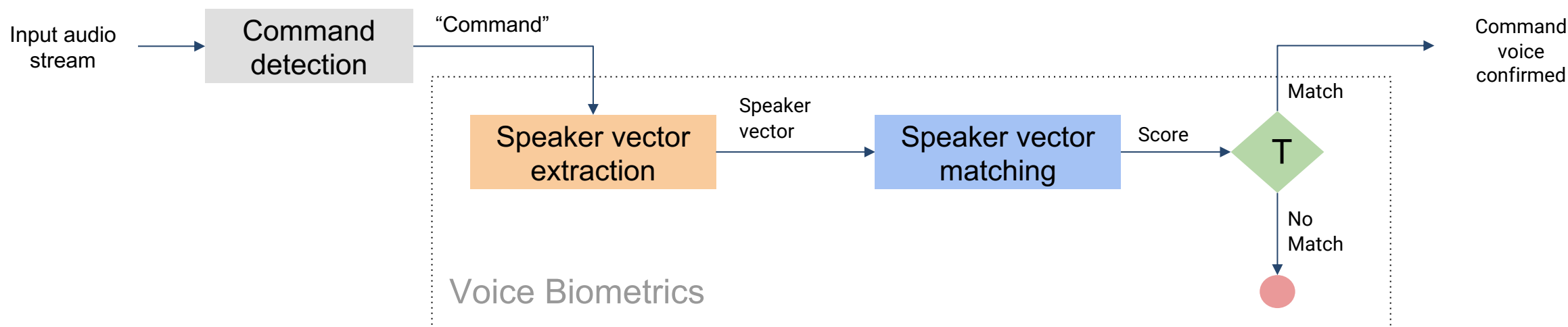


Использование голоса пользователя при стандартной настройке голосового ассистента:

- Стандартный процесс, как у "Siri"
- 5 повторений случайных команд
- Собранные команды направляются в биометрический процессор для создания "слепок" голоса

Верификация голоса

Голосовая биометрия работает в пассивном режиме каждый раз, когда детектирована команда.



Технические требования



Основные требования

- Активная регистрация голоса пользователя - 5 повторений разных **Команд**
(может использоваться голос, собранный при настройке голосового ассистента)
- Дистанция < 0.3 м
- EER < 3%

Спасибо!



www.idrnd.ai

1441 BROADWAY
NEW YORK, NEW YORK 10018

