Построение голосового интерфейса мобильного приложения

Докладчики:

Карлов Сергей, «Аурига»

Соколов Дмитрий, «Аурига»

Elite Software R&D Services

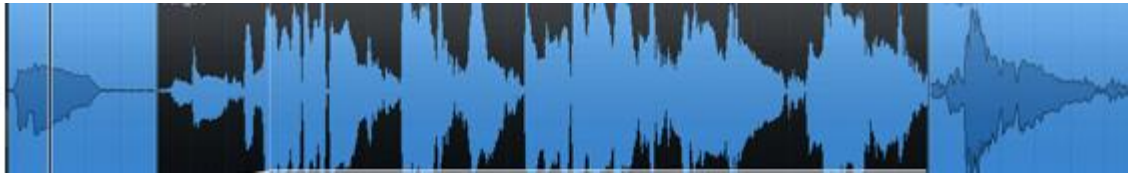
Since 1990

Почему голосовой интерфейс

- Быстрый ввод данных
- Диалог в режиме беседы
- Естественно для мобильных устройств
- Альтернатива для людей с ограниченными возможностями

Что под «капотом» ?

- Речь – цифровой сигнал (аудио-поток)



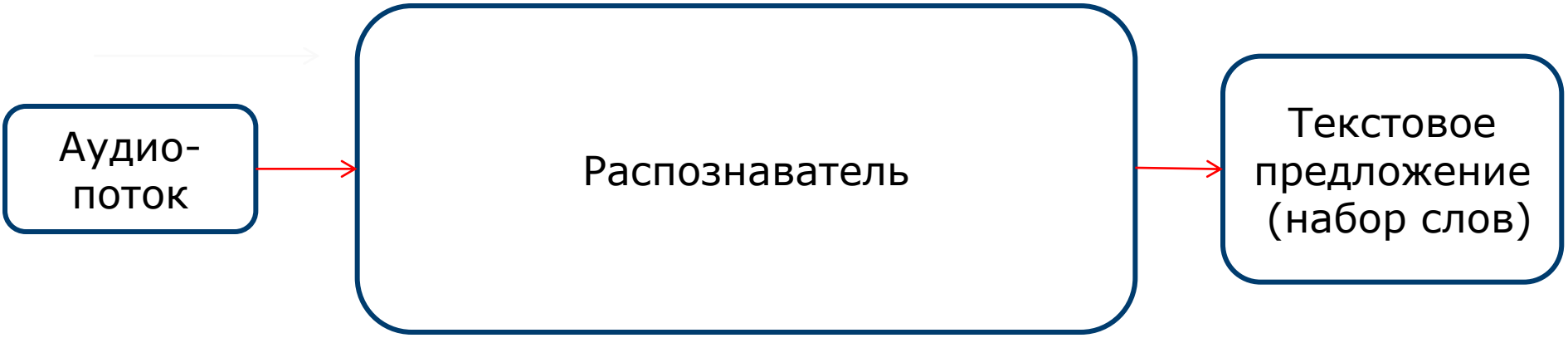
фонемы

Этапы:

- Фильтрация
- Нормализация
- Выделение фрагментов (фонемы)

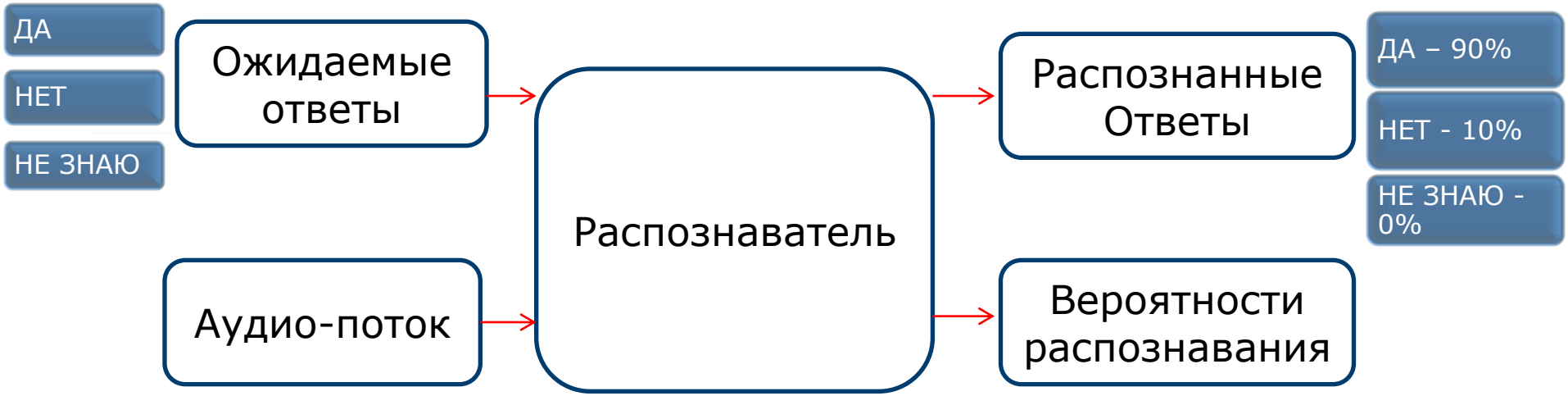
Возможные решения

- Вариант «открытые грамматики»



Возможные решения

- Вариант «закрытые грамматики»



Применение голосовых технологий в медицине

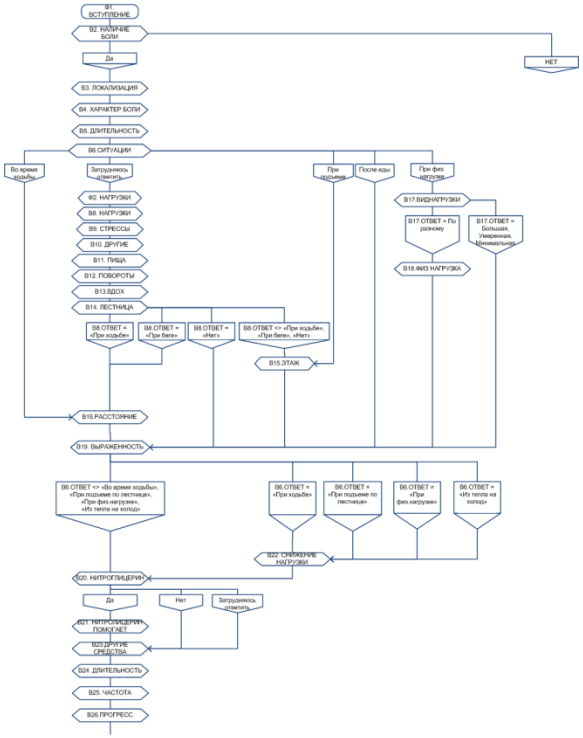
- Анкетирование пациентов

- Предварительные рекомендации

Модель данных для диалога

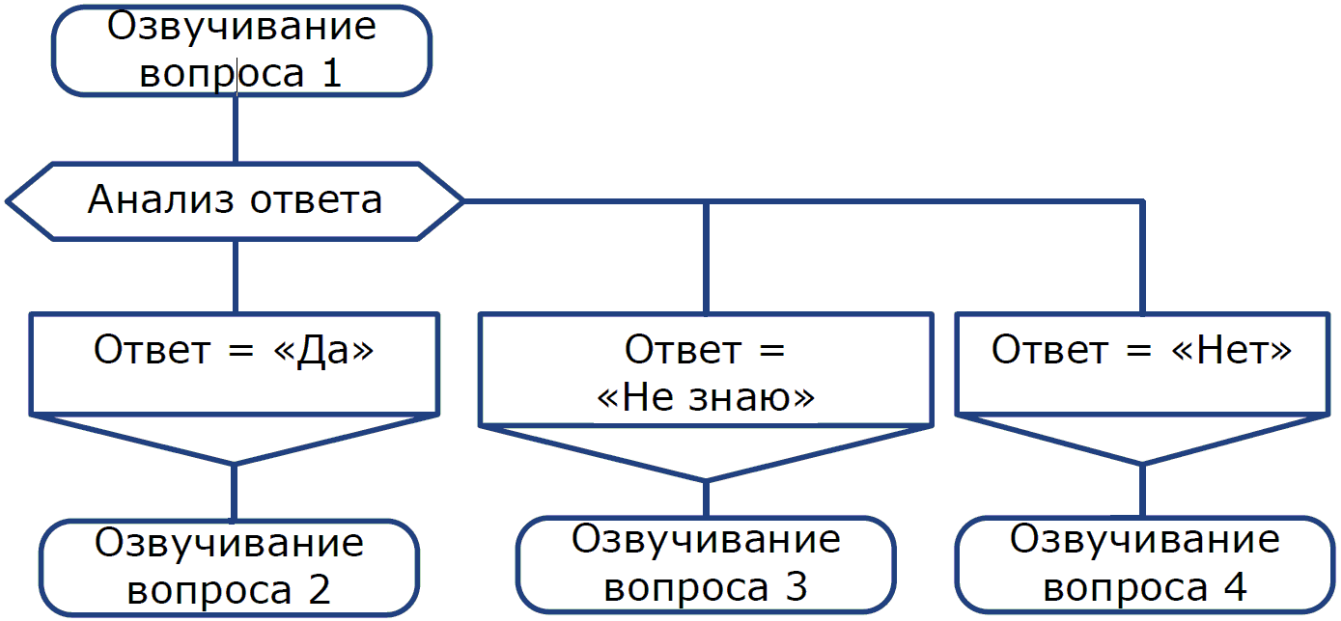
- Мед. анкета

- Сценарий



Модель данных для диалога

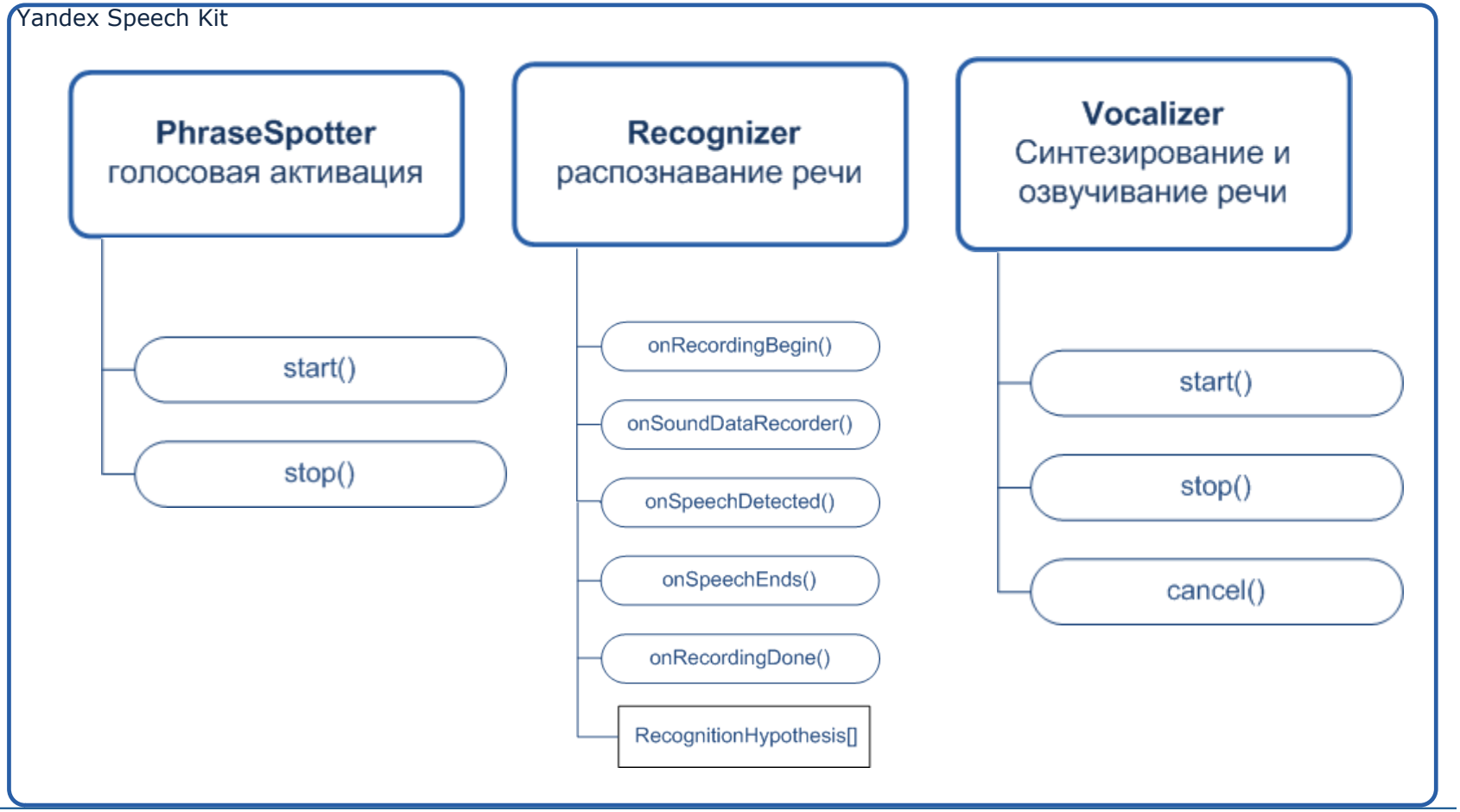
- NoSql
- Json
- J2V8



Yandex Speech Kit («открытые грамматики»)

- Легко встраивается в мобильные приложения
- Online и Offline версии
- Поддерживает несколько языков
- Стандартные языковые модели
- Выбор голоса для синтеза речи
- Понятная документация по компонентам библиотеки

Yandex Speech Kit Компоненты



Решения по модели «Закрытые грамматики»

Осталось «за кадром»:

- Протокол VXML (Voice eXtensible Markup Language)
- Протокол MRCP