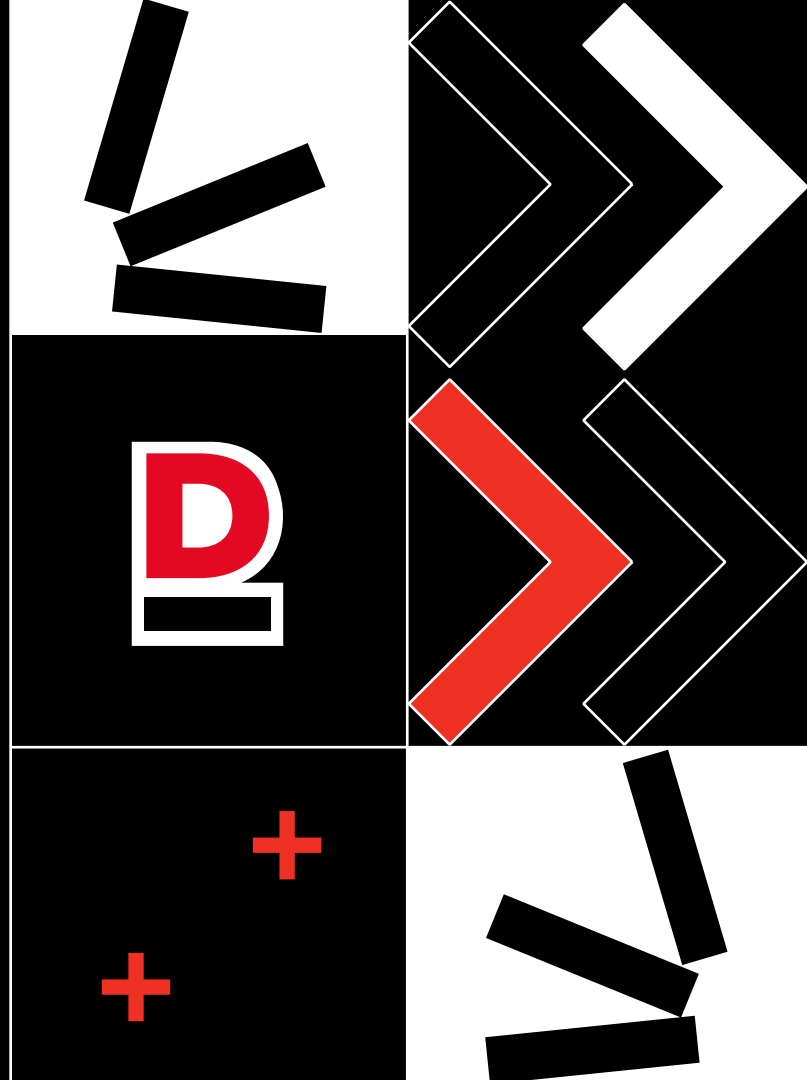


# Альфа-Помощник: Голос



Скидан Сергей  
Альфа Банк



# История голоса



**2018**

Понимает  
только  
кнопочное  
тональное меню

**2019**

Понимает вопрос  
и осуществляет  
маршрутизацию  
на оператора

**2020**

Понимает  
базовые вопросы  
и отвечает  
предзаписаны  
ответами

**2021**

Ведет диалог  
с клиентом,  
все ответы  
предзаписаны

**2022**

Внедрение  
активных операций  
за счет интеграций  
с внутренними  
системами

**2023**

Анализ профиля  
клиента: продукты,  
проблемы,  
обращаемость  
и т.д.

**2024+**

Потоковый  
синтез, качество  
распознавания,  
интонация,  
ударения

# Какой объем обрабатываем?

**A**

**~90%**

решенных  
запросов в 1 раза

**50%**

закрытых запросов

**3 млн**

звонков в месяц

**Входящая  
линия**



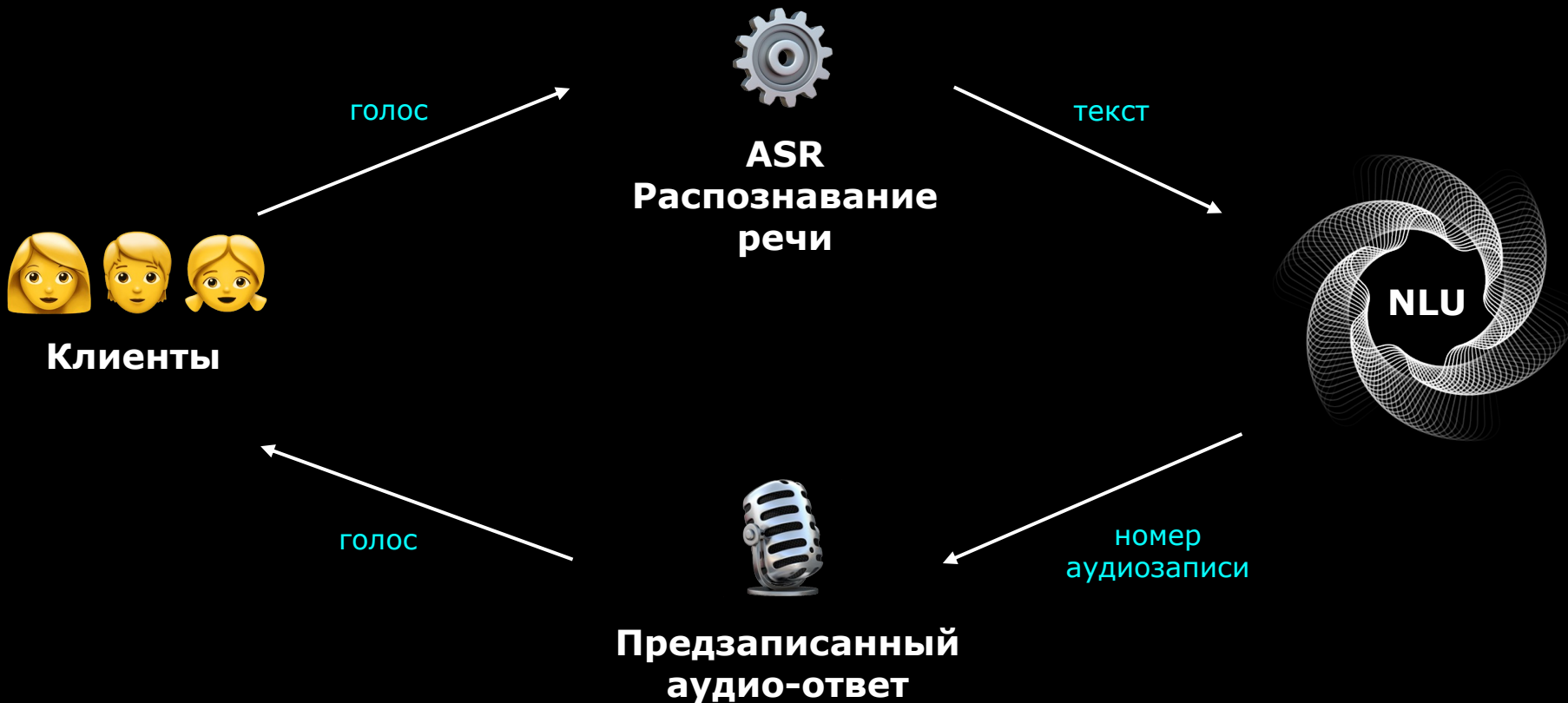
**Как мы добились  
таких результатов?**



A

**STAGE 1.**

# Единая платформа

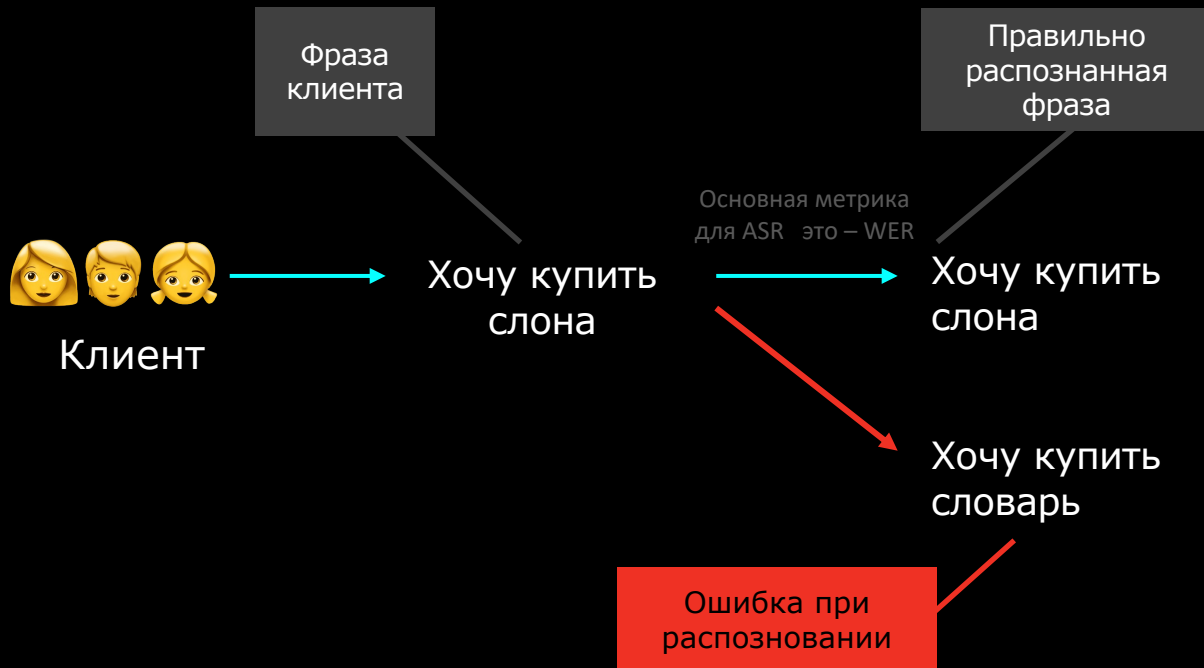


# Что такое ASR?



Клиент → ASR → NLU → Аудио-ответ

Automatic Speech Recognition (Технология автоматического распознавания речи)



**WER (word error rate)** – качество распознавания. Измеряется в %

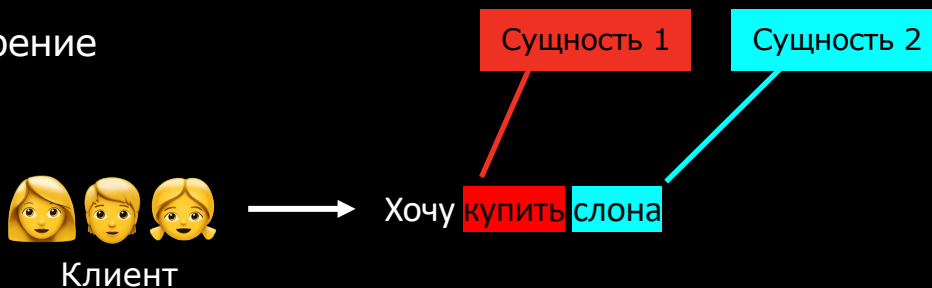
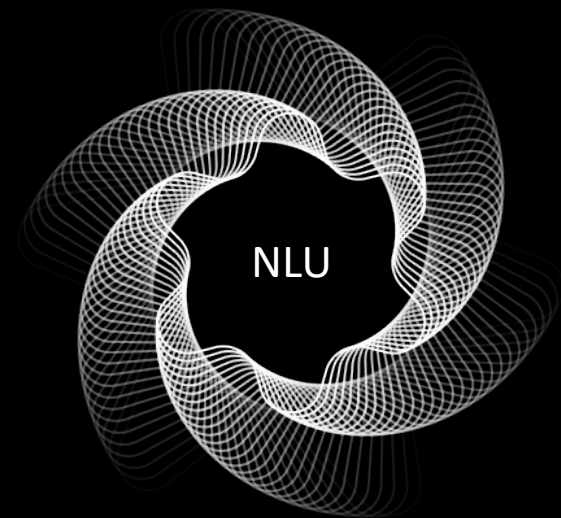
Высокий WER – 👎  
Низкий WER – 👍

Средний wer по рынку 4%

# Для чего нужен NLU?

Клиент → ASR → NLU → Аудио-ответ

Извлекает смысл из текста, а точнее намерение



**Сущность** – это единица информации, которую система хочет извлечь из текста

**Намерение** = сущность + сущность + ...  
(купить слона)

Намерению «**купить слона**»  
соответствует определенный ответ



# Аудио ответ

Клиент -> ASR -> NLU -> Аудио-ответ

Воспроизведение .wav аудиозаписи

A



Запись диктором  
Правильные ударения  
и интонация



A

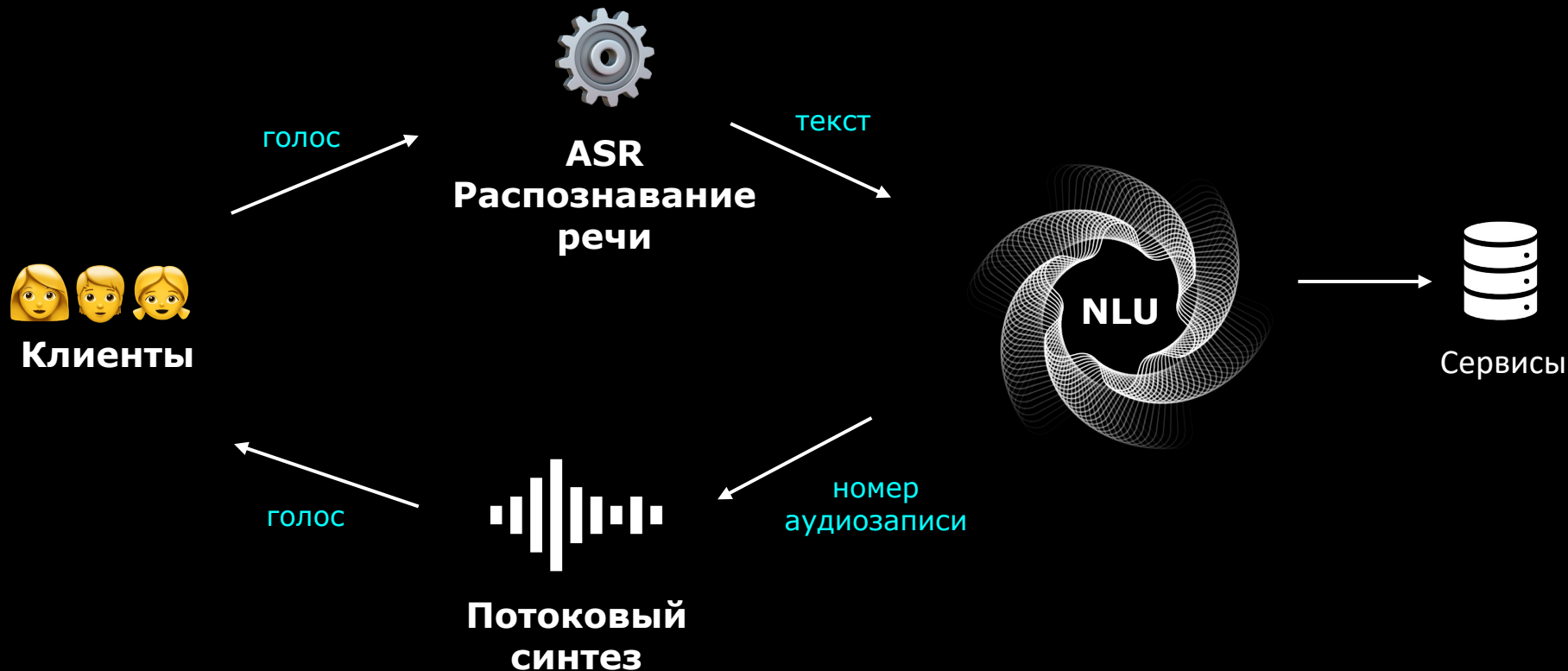
**STAGE 2.**

# Единая платформа



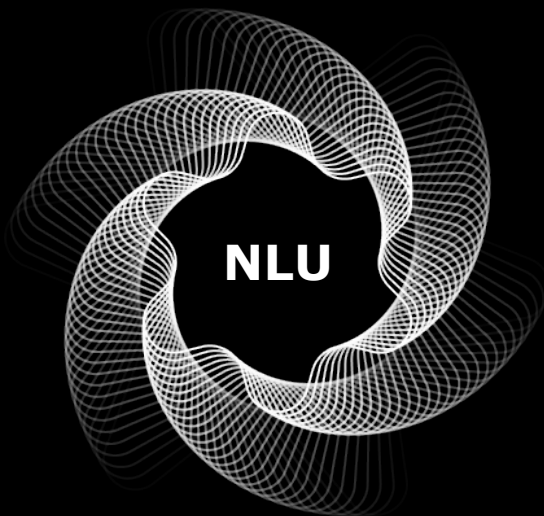
# Единая платформа

A



# Зачем интегрироваться с сервисами? A

Клиент -> ASR -> NLU -> Сервисы -> TTS



Обмен  
данными,  
активные  
операции



Профиль клиента



История взаимодействия



Статистика использования



История обращений



Сервисные данные

# ПОТОКОВЫЙ СИНТЕЗ



Клиент -> ASR -> NLU -> Сервисы -> Потокóвый синтез

TTS - Text - To - Speech



**Быстрое внедрение  
изменений**



**Отсутствие диктора,  
нет предзаписи**



**Быстрая разработка  
и настройка**



**Детальная  
настройка речи**



A

**Что дальше?**

**P**



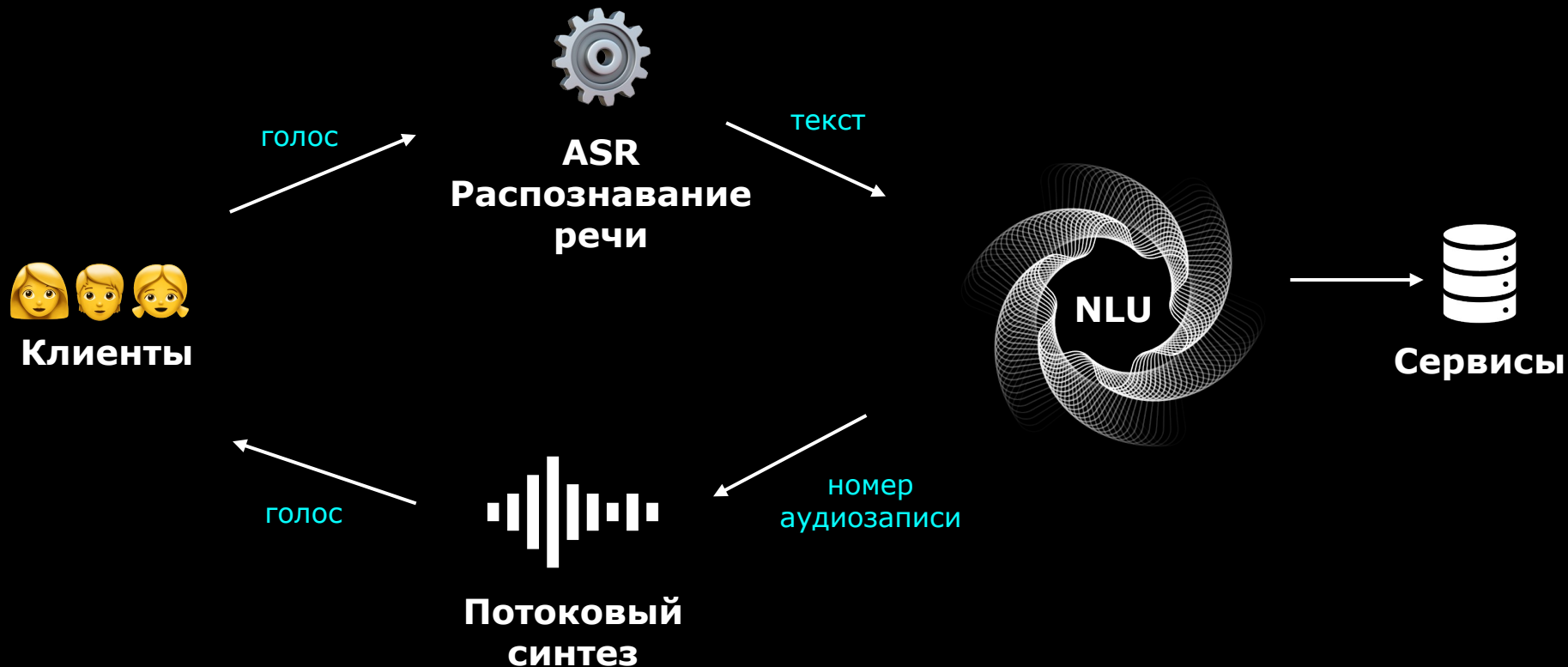
**A**

**LLM Модели**



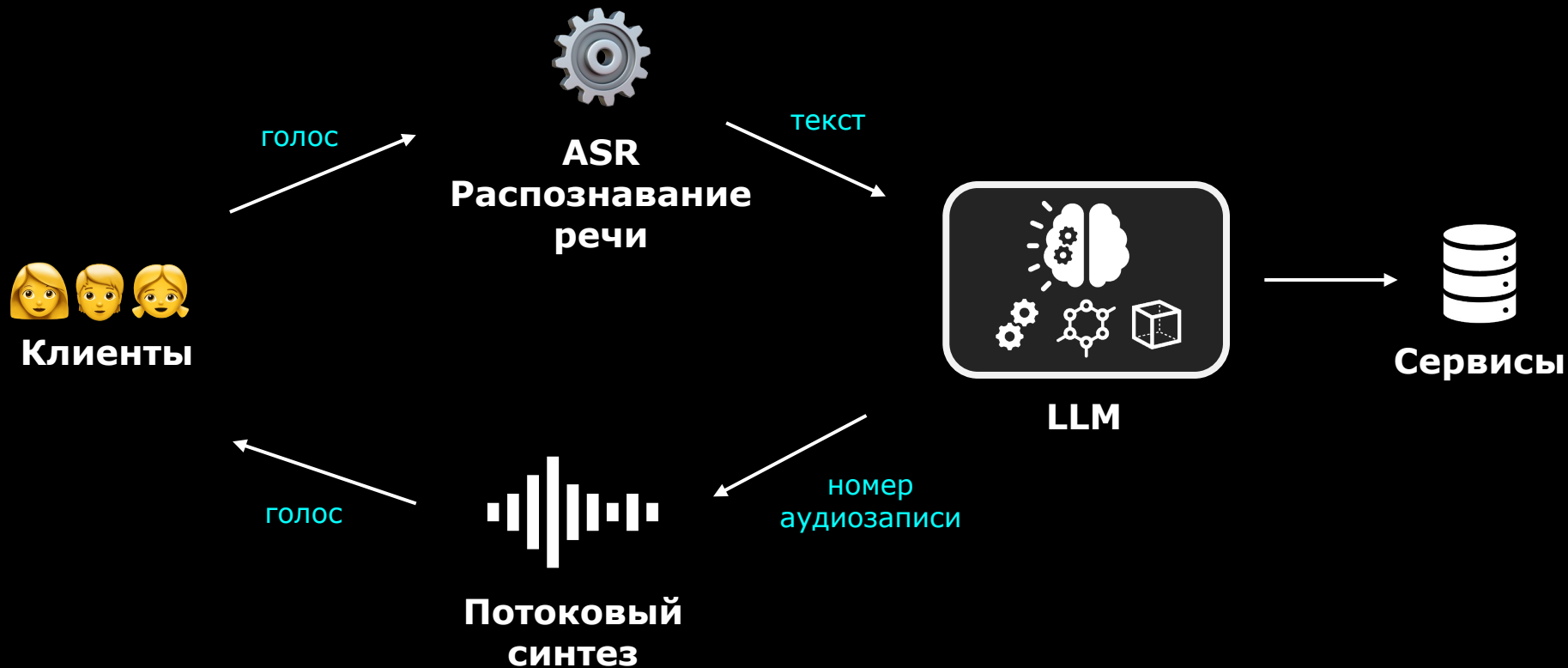
# Единая платформа

**A**

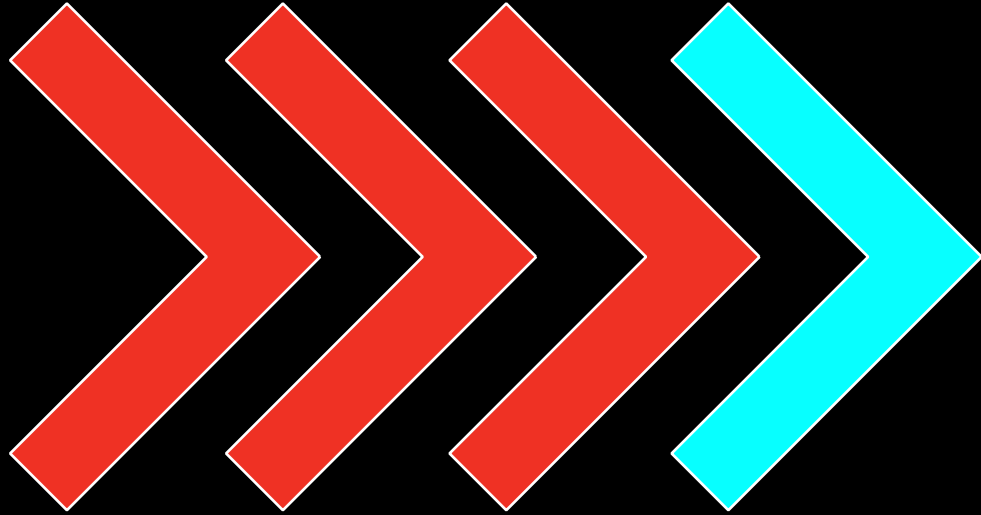


# Единая платформа

A



**В**



**A**

**Вопросы?**