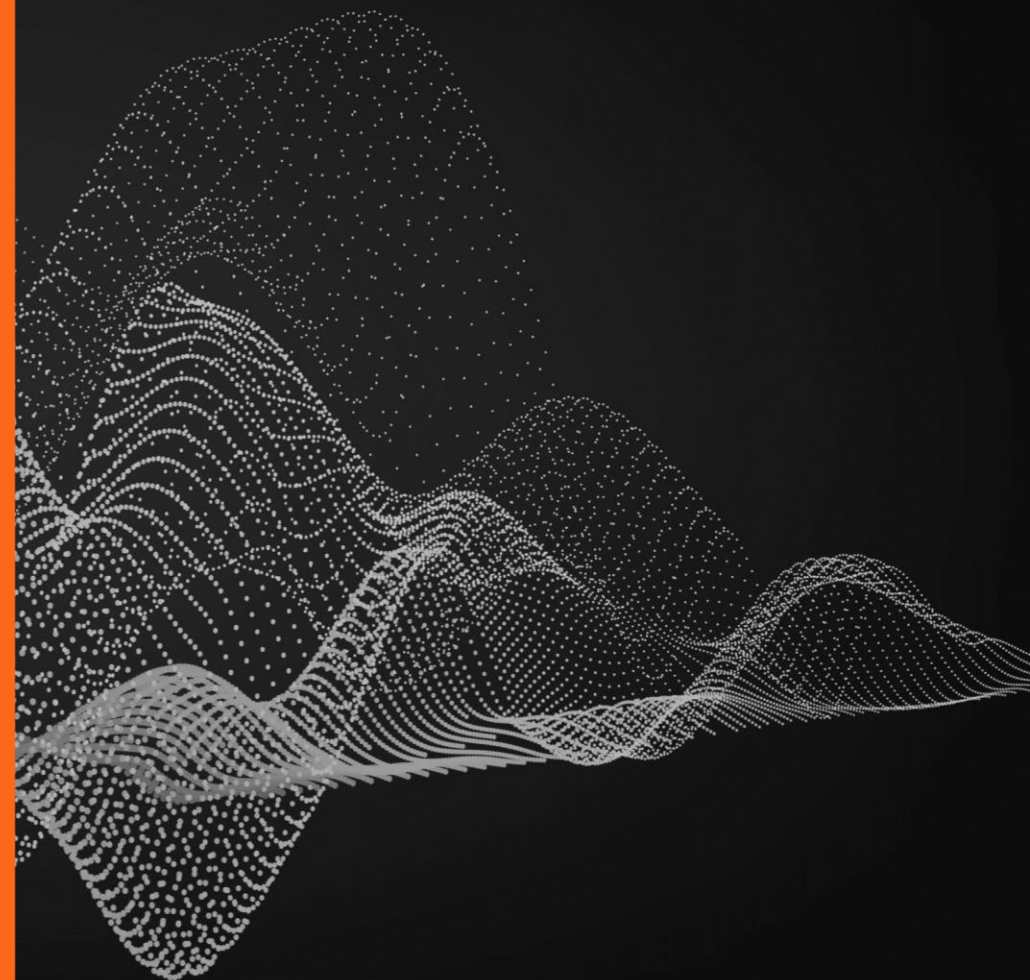


Синтез и распознавание речи Naumen Speech AI

Возможности, преимущества,
use-кейсы



NAUMEN развивает российскую экосистему продуктов для автоматизации дистанционного обслуживания

NAUMEN

500+

реализованных проектов автоматизации
контакт-центров

> 7

лет опыта работы с речевыми
технологиями

с 2001

аккумулируем экспертизу в организации работы
контактных центров

100+

профессионалов в команде для внедрения решений в
клиентский сервис

Naumen Speech AI – система автоматического **распознавания и синтеза речи**, которая помогает создавать ИИ-сервисы в крупных компаниях



Голосовые ассистенты для обслуживания клиентов и сотрудников



Голосовой интерфейс для управления процессами и навигации



Речевая аналитика для контроля работы операторов



Голосовой бот для исходящих обзвонков и информирования



Замена IVR голосовым классификатором для распределения входящих звонков



Озвучка контента и книг и другой информации

NAUMEN предоставляет полный комплекс технологий для создания голосовых ботов



Naumen Speech AI

Распознавание и синтез речи



Naumen Erudite

Платформа для создания диалоговых роботов



Naumen Contact Center

Коммуникационная платформа

-
- Высокое качество и скорость распознавания речи ботом
 - Живой и естественный синтез речи без задержек
 - Бесшовная работа в интерфейсе для транскрибирования аудиозаписей диалогов, которые используются в обучении ботов
 - Широкие возможности кастомизации распознавания для более высокого качества работы бота

Naumen Speech AI может работать в связке с Naumen Contact Center для распознавания во время обработки звонка



Naumen Speech AI

Распознавание и синтез речи



Naumen Contact Center

Коммуникационная платформа

-
- Данные для внесения в систему и персонализации предложений
 - Данные для речевой аналитики
 - Бот-суфлер и бот-помощник для операторов, который распознает звонок real-time



NAUMEN

Распознавание речи

Распознавание речи от NAUMEN

NAUMEN

Распознавание NAUMEN (ASR) позволяет автоматически преобразовать текст в речь в режиме онлайн и офлайн

Потоковый режим

real-time распознавание речи непосредственно с микрофона с возможностью получения промежуточных результатов распознавания и детектированием окончания речи

Распознавание коротких и длинных аудио

для транскрибирования как одноканальных, так и двухканальных аудиозаписей



Почему распознавание Naumen отлично подходит для создания голосовых ботов?

NAUMEN

01

Высокая скорость и работа без задержек

Низкий показатель latency и высокая скорость обработки потока речи

02

Универсальное распознавание для любых отраслей и специфики

Покажет достаточно хороший результат на любых данных без дообучения

03

Заточено под работу в звонках и устойчиво к шуму

Адаптировано к качеству потока в телефонном канале и корректно работает в шумных условиях

04

Легко масштабируется

Решение работает как на GPU, так и на CPU и легко синхронизируется с ростом объемов бизнеса

Как боты на базе технологий NAUMEN делают диалог с клиентом естественным и комфортным?

NAUMEN

- Распознавание **молчания и тишины** со стороны клиента, которое позволяет боту корректно на это реагировать
- Бот умеет отличать **естественные паузы в речи** от тишины
- Бот **игнорирует сторонние шумы** со стороны клиента и хорошо распознает невнятную речь

Примеры работы с зашумленной речью



noise_case_1.wav

Выбрать все

00:01

нужен терморезистор 117



noise_case_2.wav

Выбрать все

00:01

связаться с бобриком богданом петровичем

Как боты на базе технологий NAUMEN делают диалог с клиентом естественным и комфортным?

NAUMEN

- **Активное слушание**, в рамках которого бот подтверждает, что слышит клиента
- Интеллектуальная **работа с перебиваниями**: бот понимает, в каких случаях клиент пытается его перебить, а в каких - реакция не требуется



Настройки распознавания
Вендор
Naumen
Название открытой грамматики
default_extractor
Слова-исключения для перебивания
угу
ага
так
допустим
ла
Открыть расширенные настройки



Как достичь качественного распознавания вариативной речи?

NAUMEN

Распознавание, которое требует дообучения для любой задачи

- Работает на базе акустических и языковых моделей с открытыми и закрытыми грамматиками
- Обязательная кастомизация перед стартом
- Затраты труда и денег для дообучения перед стартом
- Это повторяется для каждой новой задачи бота, а значит, снова нужны деньги и время



Создание универсальной – генерализованной - модели

- Принятие решений опирается на ML
- Может распознавать любые данные сразу и без тюнинга
- Способ работы более приближенный к человеческому восприятию речи

Naumen ASR – универсальное распознавание которое работает с достаточным качеством в разных отраслях независимо от специфики

NAUMEN



финансы



энергетика
и ЖКХ



телеком



ритейл и
HoReCa



медицина



логистика и
перевозки

**Такое решение обеспечит быстрый старт работы с голосовыми ботами,
так как модель ASR не требует обязательного дообучения**

Высокая точность и качество распознавания

NAUMEN

WER

(Word Error Rate) для оценки точности распознавания понимается как % неверных слов

**Чем меньше WER –
тем лучше**

YouTube



Финансы



Что мы делаем, если точности все же не хватает для качественного решения узкой задачи?

NAUMEN

Большой
объем
данных

+

Распространенные
типовые данные

Пример задачи: точное распознавание трек-номеров, фамилий, адресов или телефонов

Решение: тюнингуюем саму модель распознавания, работая с **языковой моделью и акустической моделью**

Конечное
число единиц
данных

+

Данные
индивидуальны
для компании

Пример задачи: точное распознавание названий банковских продуктов, тарифов телеком-операторов или брендов товаров в ритейле

Решение: тюнингуюем с пост-обработку результатов распознавания и учитываем контекст путем работы **со словарем и проводим рескоринг гипотез на основе доменных эвристик**



NAUMEN

Синтез речи

Синтез речи от NAUMEN

NAUMEN

Синтез NAUMEN (TTS) позволяет автоматически генерировать речь на основе текста на русском языке

Живое естественное звучание

Речь звучит реалистично и приближена к человеческой. Нейронные сети, используемые в обучении, позволяют сделать звучание плавным, сохранять паузы и интонации



Анна



Павел



Ева



Ксения



Собственная библиотека голосов

Три женских голоса и один мужской на выбор в соответствии с задачами и профилем бизнеса



Создать индивидуальный голос бренда – просто и быстро



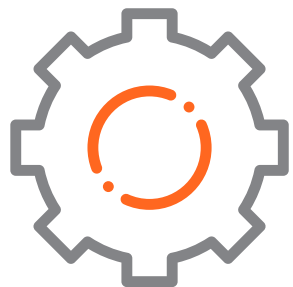
Для создания собственного голоса достаточно 5-6 часов дикторских записей, сделанных на хорошем микрофоне



Возможность трансфера уникального голоса с на другие языки без дополнительных дикторских записей на этом языке

Так звучит один из голосов, который создали мы на 6 часах записей

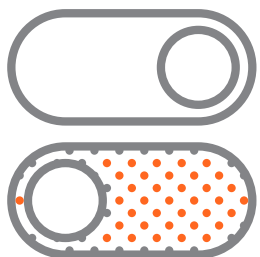




**Изменение
скорости речи**



Дополнительная кастомизация
голоса для естественного
звучания



**Изменение
тональности речи**



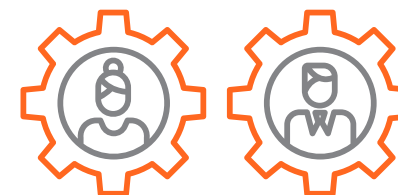
Инструмент для гибкого
сочетания предзаписи и
синтеза, который помогает
создавать бесшовные переходы

Как мы работаем с динамическими данными в сценариях: примеры

NAUMEN



Числа и суммы



**Фамилии, имена
и отчества**



Где уже работает Naumen Speech AI?

NAUMEN

chipdip.

70 000

входящих звонков
ежемесячно обрабатывает
голосовой помощник,
используя распознавание
Naumen Speech AI

28%

составляет уровень
автоматизации на проекте



Уральский Банк
реконструкции и развития

210 000

входящих звонков ежемесячно
принимает и классифицирует
голосовой помощник,
применяя распознавание
Naumen Speech AI



Новосибирскэнергосбыт

45 000

входящих звонков от физлиц
и юрлиц ежемесячно
обрабатывает умный
помощник используя
распознавание Naumen
Speech AI



NAUMEN

Остались вопросы?

Ждем на стенде NAUMEN

Хотите потестировать синтез и распознавание речи?

Пишите на sales@naumen.ru или звоните +7 (495) 145 90 45