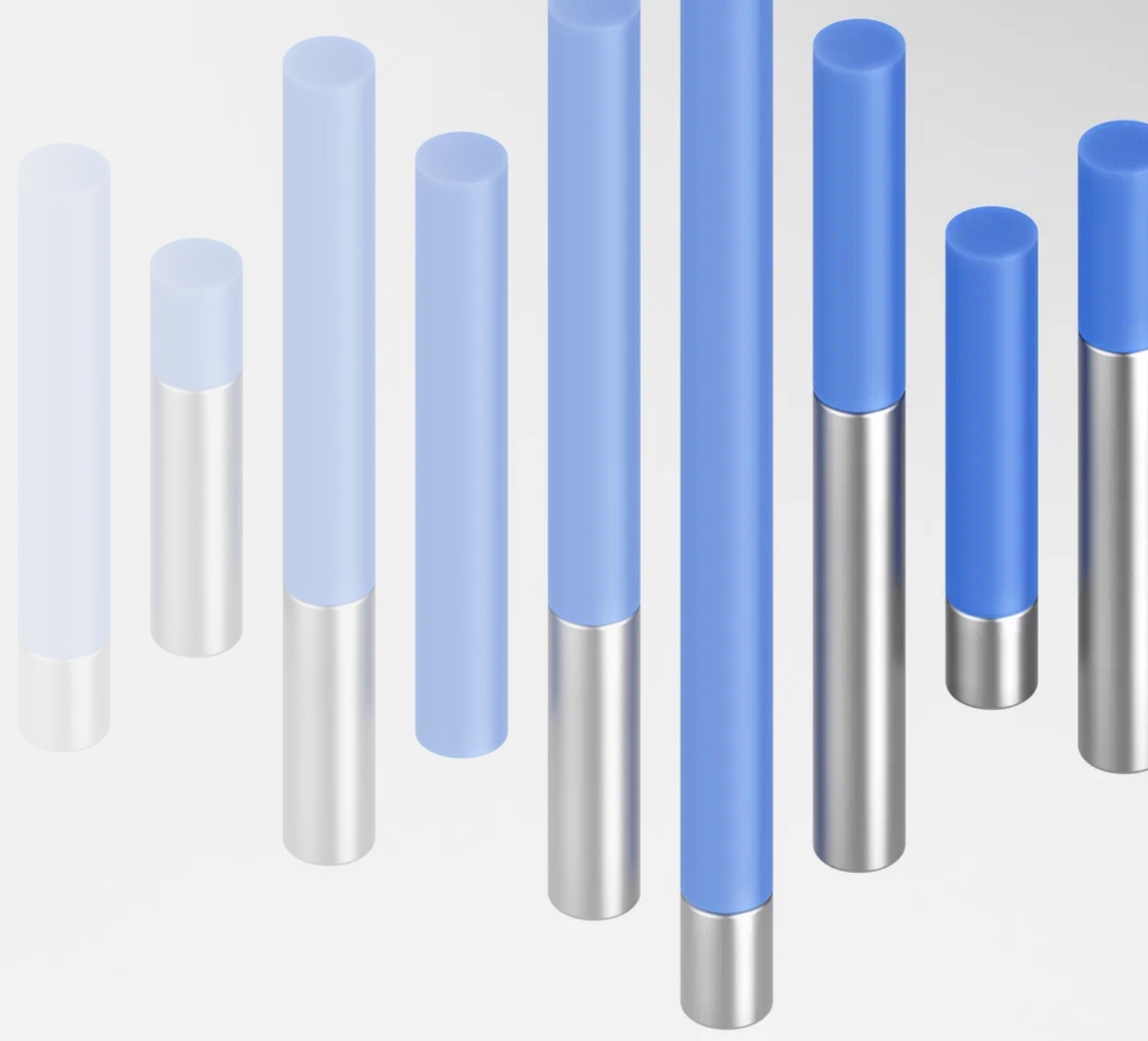


Принципы синтеза речи

Синтез речи — это технология, которая позволяет переводить текст в речь. С её помощью создают голосовых помощников, автоматизируют телефонные коммуникации, начитывают аудиокниги и не только.



У Яндекса есть собственная технология синтеза речи. Её применяют в Алисе, Браузере, Картах, Навигаторе и других продуктах Яндекса. Она лежит в основе приложения Разговор, которое помогает людям с нарушениями слуха и речи общаться с окружающими; её использует виртуальный рассказчик в сервисе Яндекс Книги. Нашей технологией синтеза речи могут воспользоваться и другие компании — с помощью сервиса **Yandex SpeechKit** на облачной платформе **Yandex Cloud**.



Технология синтеза речи обучается на аудиозаписях голосов людей. Чтобы получить записи, приглашают дикторов, которые начитывают тексты перед микрофоном. Дикторами могут быть как профессиональные актёры озвучивания, так и обычные люди.

Яндекс развивает речевые технологии более десяти лет и много сотрудничает с дикторами и звукозаписывающими студиями. На базе нашего опыта мы сформулировали принципы, которых, на наш взгляд, необходимо придерживаться в работе над синтезом. Они нужны, чтобы наша технология оставалась безопасной для всех, кто так или иначе с ней соприкасается — дикторов, которые помогают ей обрести голос, партнёров, которые внедряют её у себя, и пользователей, которые с ней взаимодействуют.

Мы публикуем этот внутренний документ, поскольку считаем, что наши принципы могут быть полезны в масштабах всей индустрии речевых технологий.

✓ **Мы не синтезируем речь людей без их разрешения**

Когда мы создаём модель синтеза — такую, с помощью которой можно озвучивать произвольные тексты, — то всегда сперва получаем согласие на использование голоса у человека, которому он принадлежит, или у его законных представителей. При этом мы оставляем за собой право не запрашивать согласие при решении прочих задач, связанных с обработкой аудиозаписей голоса. Это может быть, например, повышение разборчивости речи, очистка записи от шумов или перевод сказанного на другой язык.

✓ **Мы информируем дикторов о том, как будут использоваться их голоса**

При сотрудничестве с дикторами мы максимально подробно и открыто объясняем им, что такое синтез речи и как будут использоваться сделанные в студии записи. Важно, чтобы дикторы понимали: впоследствии их голосами будут озвучены другие тексты — не те, которые они начитывают в микрофон. Мнение диктора может не совпадать с позицией, которая высказана в озвученном его голосом тексте.

✓ **Мы храним и обрабатываем данные таким образом, чтобы они всегда оставались в безопасности**

Мы очень бережно обращаемся с данными, которые используются для обучения модели и в процессе синтеза речи:

- Данные, которые мы обрабатываем на своей стороне, например записи голосов дикторов, хранятся в дата-центрах Яндекса, отвечающих самым строгим стандартам в сфере информационной и физической безопасности.
- Доступ к данным имеет ограниченное количество сотрудников — только те, кому они непосредственно нужны в работе. Такие сотрудники проходят дополнительный инструктаж.
- Когда пользователи синтезируют речь через API **Yandex SpeechKit**, мы не имеем доступа ни к исходному тексту, ни к результатам синтеза.
- Если партнёр синтезирует речь с помощью нашей технологии на собственном сервере, мы также не имеем доступа ни к текстам, ни к результатам синтеза. Это особенно важно для компаний, которые работают с чувствительными данными.

✓ **Мы сохраняем за собой право ограничивать сценарии использования нашей технологии синтеза речи**

Наша технология синтеза речи доступна сторонним пользователям. У нас нет возможности контролировать, как именно они её применяют в каждом из случаев, но мы оставляем за собой право отозвать доступ, если нам станет известно о злоупотреблениях. Например, с помощью **Yandex SpeechKit** мы не разрешаем озвучивать некоторые виды контента, в том числе противоправный контент и контент, где есть оскорбления или дискриминация, — это оговорено в [правилах использования](#).