



# Нейросети для создания музыки, видео, голоса и перевода устной речи

Из курса Учебника Т—Ж

«Как облегчить жизнь с помощью нейросетей»

Генерация текстов и картинок с нейросетями — только вершина айсберга. В этой инструкции советуем новые сервисы и показываем, на что они способны.

1. [Пишем музыку](#)
2. [Имитируем пение](#)
3. [Переводим речь на видео с синхронизацией мимики](#)
4. [Создаем видео](#)

## Пишем музыку

**Что это.** Нейросеть создает музыку и песни по промпту.

**Как это работает.** Принцип похож на работу с Midjourney — промпты составляются по одним и тем же правилам. Например, вы пишете в запросе что-то вроде «мелодичный хип-хоп со скрипкой и клавишными, 100 bpm, мрачная атмосфера», а нейросеть выдает вам результат. Подробно о технологии и способах ее применения можно прочитать [в статье Тинькофф Журнала](#).

**Какие сервисы попробовать:**

1. [Stable Audio](#). Инструмент от разработчиков популярной нейросети для создания картинок Stable Diffusion. Можно генерировать мелодии с указанием жанра, инструментов, атмосферы, ритма. Трек ограничен 45 секундами, с платным аккаунтом можно расширить до полутора минут.

[Stable Audio — Generate music with AI](#)

2. [Suno Chirp](#). Эта нейросеть генерирует целые треки. Сначала необходимо придумать несколько строчек песни — или доверить это ChatGPT. А музыка уже подстроится под текст — можно выбрать разве что жанр и настроение. Своего сайта или приложения у Suno Chirp нет, генерация проходит в «Дискорде».

[Everything is Fake — AI Rap Metal Song V2 \(Suno Chirp + Runway Gen-2\)](#)

## Имитируем пение

**Что это.** Нейросеть генерирует песни в исполнении разных людей.

**Как это работает.** Самый простой способ — загрузить трек и выбрать голос, который должен его спеть. Минус у подхода один: галерея голосов ограничена, и там может не быть такого, который вам нужен. Более сложный способ — самому сделать модель нужного голоса и загрузить ее в нейросеть. Проблема в том, что для этого надо самому собрать набор данных для обучения алгоритма или найти слепок нужного голоса на форуме энтузиастов.

В итоге можно получить неожиданный результат. Например, недавно [энтузиасты взяли голоса рэпера Drake и солиста The Weeknd](#) и загрузили их дуэт на стриминговые платформы. Трек стал вирусным — например, до удаления на Spotify его успели прослушать больше 600 000 раз.

### Какие сервисы попробовать:

1. [Covers AI](#). Работает очень просто: нужно загрузить трек или просто оставить на него ссылку, а затем выбрать нужный голос. Все остальное нейросеть сделает сама. Проблема этого онлайн-сервиса и подобных в одном: на них постоянно жалуются правообладатели песен, поэтому там мало голосов известных исполнителей. Их заменяют, например, Губкой Бобом или персонажами «Маппет-шоу».

[Linkin Park — Pokémon Theme Song \(AI Cover\)](#)

2. So-vits-svc. Достаточно сложная программа, зато универсальная. Вы можете найти в сети «слепок» любого голоса, загрузить его и сделать кавер. Можно генерировать каверы у себя на компьютере

или воспользоваться нейросетью в среде Google Colab. Как именно это сделать, мы рассказывали [в гайде Тинькофф Журнала](#).

[Freddie Mercury sings Toxicity by System of a Down \(AI Cover\)](#)

## Переводим речь на видео с синхронизацией мимики

**Что это.** Нейросеть переводит видео, сохраняет оригинальные голоса и адаптирует мимику под новый язык.

**Как это работает.** Нужно загрузить файл в нейросеть — и все! На выходе вы получите то же видео с оригинальным голосом, но только на другом языке с оптимизированной под новые слова мимикой.

### Какие сервисы попробовать:

1. [HeyGen](#). Умеет все то, о чем мы говорили выше. Бывают ошибки, но в целом очень похоже переносит мимику и голос. Главное — чтобы на видео говорил только один человек, при диалоге два разных голоса синтезируются в один усредненный. Есть недостаток: умеет переводить с русского языка на иностранные, но не наоборот. И есть ограничения по длительности — не более пары минут.

[Мем «Мы не знаем что это такое, если бы мы знали, что это» на английском \(heygen\)](#)

[Никита Литвенко — «Борщик с капусткой» на английском \(heygen\)](#)

2. [ElevenLabs](#). Главный плюс — умеет переводить с иностранных языков на русский. Но есть и минус — мимика не сохраняется, что очень сильно бьет по качеству роликов. Сам перевод тоже неидеальный.

[One Piece — русский дубляж от нейросети AI by ElevenLabs \(Japanese to Russian\)](#)

# Создаем видео

**Что это.** Нейросеть создает видео по промпту.

**Как это работает.** Такие сервисы за последний год сильно прогрессировали. Весной 2023 нейросети генерировали очень абстрактный видеоряд. Он напоминал какой-то сон, который начинается нормально, а потом переходит во что-то странное и невозможное. Взгляните на рекламу бургеров: [Burger Blast Ad 1995 — AI Generated Commercial](#)

А вот что произошло через полгода. С помощью той же нейросети сообщество создает качественные короткометражки, которые сложно отличить от сделанных человеком: [Genesis — Official Trailer \(Midjourney + Runway\)](#)

**Какие сервисы попробовать.** Мы рекомендуем попробовать [Runway Gen-2](#): сейчас это самый мощный сервис. Нейросеть создает видео по текстовому описанию и анимирует статичные изображения, поэтому его часто используют в связке с Midjourney. Есть редактор, чтобы быстро смонтировать сгенерированные сцены. Есть бесплатные попытки — 105 секунд генерации, что равно примерно 26 коротким роликам. Их достаточно, чтобы попробовать нейросеть, но для полноценных работ лучше приобрести подписку.

[CYBER CHASER | cinematic AI video | runway gen-2 | midjourney](#)

