

Главная идея: остаточное векторное квантование (RVQ)

Нейрокодек учится сжатию по данным, а не по правилам. RVQ округляет звук слоями: первый берёт грубое значение, следующий - остаток ошибки, и так далее - как сдача рублями, потом гривенниками, потом копейками. Убрать нижние слои - и битрейт падает сразу, поэтому одна модель даёт много битрейтов.

Кодеки кратко

Кодек	Битрейт	Качество (равное)	Цена / статус 2026
Opus (2012)	6-510 кбит/с	эталон для сравнения	микросекунды; основа WebRTC
SoundStream	3-18 кбит/с	3 кбит/с примерно Opus 12	исследовательский эталон (RVQ)
Lyra v2	3.2/6/9.2	9.2 кбит/с примерно Opus 14	0.57 мс/кадр на телефоне; в проде
EnCodec	1.5-24	3 кбит/с > Opus 12; +музыка	дорого; топливо для аудио-ИИ
Mimi	~1.1 кбит/с	токенайзер для речевых LLM	внутри голосовых ИИ-моделей

Цифры, которые решают

- Нейрокодеки дают качество Opus при 3-4 раза меньшем битрейте - но сеть работает на каждом кадре, груза CPU и батарею куда сильнее Opus.
- 50 дикторов x 12 кбит/с Opus = 600 кбит/с; при 3 кбит/с нейро = 150 кбит/с. Экономия в 4 раза - реальна, где мало полосы, но вычисления не бесплатны.
- Lyra v2 - исключение, что работает на телефонах: 0.57 мс на кодирование+декодирование кадра 20 мс на Pixel 6 Pro, ~35x реального времени, задержка 20 мс.
- Главное применение в 2026 - голосовой ИИ: кодек превращает звук в токены для LLM. Mimi делит их на семантические (что сказано) + акустические (как звучит).
- Opus остаётся выбором для живых звонков; следующий традиционный кодек (AOMedia OAC, начат в 2026) - преемник Opus, а не нейросетевой.

Прежде чем брать нейрокодек - барьеры

- ☐ Звонки в реальном времени? По умолчанию Opus - почти бесплатное декодирование, без лицензии, основа WebRTC.
- ☐ Нейрокодек только там, где полоса - жёсткий предел И есть бюджет вычислений на клиента.
- ☐ Проверьте худшее устройство клиента: CPU на кадр, расход батареи и добавленную задержку.
- ☐ Строите голосовой ИИ (ассистент, дубляж, TTS на устройстве)? Нейрокодек - фундамент, не опция.
- ☐ «Обгоняет Opus на низком битрейте» - это лабораторный результат; перепроверьте на своих устройствах.
- ☐ Следите за гранью кодек/генератор: восстановить голос из 1 кбит/с - почти то же, что клонировать его.