

1 - Две задачи под одним названием (решите это первым)

LIP-SYNC: старт с реального видео, рисуем ТОЛЬКО рот. Дешевле, быстрее, безопаснее на вид.
АВАТАР: старт с ОДНОГО фото, выдумываем всю голову. Мощнее, тяжелее, рискованнее.

2 - Пять опенсорсных моделей

Модель	Задача	Real-time?	Для чего
Wav2Lip	Lip-sync	Да	Любое лицо, дёшево
MuseTalk	Lip-sync	Да	Real-time рот на видео (~30 fps)
LatentSync	Lip-sync	Нет	Лучшее качество lip-sync (диффузия)
SadTalker	Аватар	Нет	Аватар из одного фото (3D-каркас)
ЕМО / OmniHuman	Аватар	Нет	Выразительные офлайн-аватары (диффузия)

3 - Три коммерческие платформы

Платформа	Под что	Цена входа 2026
Tavus	Живые AI-видеоагенты (CVI, < 600 мс)	Free, далее ~\$59/мес
HeyGen	Сценарные многоязычные ведущие + API	Free, далее ~\$29/мес
Synthesia	Управляемые корп. обучающие библиотеки	~\$29/мес (~\$22 годовых)

Цены - прайс-диапазоны 2026, часто меняются; перепроверяйте при интеграции.

4 - Граница реального времени (почему диффузия не стримит вживую)

30 fps -> 1 с / 30 = 33 мс на кадр. Каждый кадр должен уложиться в 33 мс, иначе дёрганья.
Диффузия в 25 шагов: 33 / 25 = 1,3 мс на шаг. Ни одна стоковая диффузия так не успевает.
=> Real-time = только одношаговые методы. Вывод класса ЕМО рендерится ОФЛАЙН, потом играется.

5 - Чек-лист согласия и раскрытия (до запуска)

- ☐ Храните, КТО согласился, ровно на КАКОЕ использование, КОГДА - со сроком и путём отзыва.
☐ Привязывайте каждый сгенерированный ассет к записи о согласии, его разрешившей (аудит).
☐ Получите записанное согласие до клонирования голоса или лица - даже если вендор не требует.
☐ Добавьте видимое раскрытие «сгенерировано AI» - ст. 50 EU AI Act обязательна со 2 авг 2026.
☐ Следите за NO FAKES Act (США, предложен 2025): федеральное право на цифровую реплику близко.