

1 — Три плоскости: держите эти задачи отдельно

Медиа	Несёт звук + видео. WebRTC + SFU. Самый жёсткий бюджет.
AI	Все ML-модели. Живут в одном из трёх домов (см. ниже).
Control	Комнаты, участники, сигналинг, кто говорит. Свободный бюджет.

2 — Где работает каждая AI-фича (по нужному ей сигналу)

НА УСТРОЙСТВЕ	Размытие фона, шумоподавление, AR/бьюти-фильтры. (отправитель) Нужны сырые кадры + звук. \$0 на участника. Лучшая приватность.
ОБЛАЧНЫЙ АГЕНТ	Живые субтитры, перевод, заметки встречи, lip-sync аватара. (как участник) Подписка на потоки, публикация результатов. Цена за минуту.
СТОРОНА SFU	Только модерация контента. Централизованно, клиент не обойдёт. (копия потока) Копию потока — в воркер; SFU никогда не декодирует встроено.

3 — Два параллельных бюджета (с арифметикой)

ИНТЕРАКТ. МЕДИА одна сторона, цель ~200 мс (говорят ли друг через друга)
Захват 10 + AI устр-ва 35 + Кодир. 10 + Сеть 120 + Вывод 25 = 200 мс
-> если сеть 120 мс, на всё остальное остаётся ~80 мс. Держите модели малыми.

ГОЛОСОВОЙ ЦИКЛ слышать-отвечать, цель ~800 мс (отзывчив ли ассистент)
Смена реплики 200 + STT 100 + первый токен LLM 300 + первый чанк TTS 150
+ сеть 50 = 800 мс. Потокowo каждую стадию; детектор реплики, не таймер тишины.

4 — Порядок сборки: каждый шаг выдаёт работающий продукт

Шаг 1 Обычный звонок — WebRTC + SFU. Надёжные звук/видео + переподключение.
Шаг 2 Очистка на устройстве — размытие фона + шумодав. Лучшее качество/усилие.
Шаг 3 AI-агент — живые субтитры, затем перевод, затем заметки (тот же транскрипт).
Шаг 4 Специализация — модерация, аватары, ассистенты + уведомление AI Act.

5 — Чек-лист перед запуском (инженерный контекст, не юрсовет)

- ☐ Ровно ОДИН шумодав в цепочке — отключите встроенный в браузер, если ставите свой.
- ☐ Никакого аудио-AI после Opus-энкодера; шумодав на сыром звуке в AudioWorklet, до кодирования.
- ☐ SFU никогда не транскодирует; модерация на копии потока, а не встроено на сервере.
- ☐ Письменный бюджет задержки на устройство; измеряйте каждую модель на реальных телефонах.
- ☐ Раскройте AI-сгенерированный или изменённый голос/лицо участникам — статья 50 EU AI Act.
- ☐ Берите медиа-плоскость и фреймворк агента; стройте только свои фичи и правила.