Чек-лист запуска аватара в реальном времени

1 — Цикл: пять стадий гонятся с часами

```
Mic → ASR → LLM → TTS → синтез аватара → WebRTC → зритель. Стадии последовательны.
Идите сквозь них потоком (LLM/TTS/аватар рано); не ждите завершения каждой.
```

2 — Одно число: задержка переключения хода

```
Цельтесь в воспринимаемую паузу ~200 мс; держите её под ~1,5 с, иначе как робот.
До первого кадра: 150 ASR + 300 LLM + 150 TTS + 200 аватар + 100 WebRTC ≈ 900 мс.
```

3 — Архитектура: аватар входит в звонок

```
Шлите аватару только ЗВУК агента; пусть аватар публикует видео прямо в комнату.
Это убирает обратный рейс и двойное кодирование (наивный WebSocket платит дважды).
Прерывания и сигнал завершения — по RPC / каналу данных, за миллисекунды.
```

4 — Купить или собрать (2026 — перепроверьте перед выбором)

```
Купить (быстро): Tavus Phoenix-4 (sub-600 мс), HeyGen LiveAvatar (1–2 с), Simli, bitHuman.
Self-host (лицо/данные/объём): MuseTalk + LivePortrait (MIT), NVIDIA ACE на своих GPU.
Решают задержка, контроль над лицом и поминутный объём против аренды GPU.
```

5 — Раскрытие: закройте это ДО запуска в ЕС

```
EU AI Act, статья 50: синтетическое лицо — дипфейк. Раскрытие + машиночитаемая метка.
Прозрачность с 2 августа 2026. Живое видео: постоянная метка «AI» + вступление.
```

6 — Чек-лист перед запуском

- [ ] Измерили реальную воспринимаемую паузу хода end-to-end, а не задержку одной стадии.
- [ ] Рендерер аватара входит в звонок как участник; агент шлёт только звук, без рейса видео.
- [ ] Прерывания отменяют кадры в полёте и переключают в позу слушания за ~200 мс.
- [ ] Выбрали купить/собрать по задержке, контролю над лицом и объёму минут стриминга.
- [ ] Зафиксировали голосовой путь (потоковый TTS или speech-to-speech) и тест под прерыванием.
- [ ] Добавили раскрытие по статье 50 EU AI Act: постоянная метка «AI» + вступительное раскрытие.
- [ ] Подтвердили согласие на любое клонированное лицо или голос до выхода в эфир.