

1 — Пять семейств open-weights (сначала лицензия + VRAM)

Семейство	Версия	Лицензия	Мин. VRAM	Сильная сторона
Wan (Alibaba)	2.2 (open)	Apache 2.0	~8 GB (5B)	топ открытый VBench 84.7%
HunyuanVideo (Tencent)	1.5	Community (огранич.)	~14 GB (FP8)	топ-качество T2V+I2V
CogVideoX (Zhipu)	1.5 / 5B	Apache (2B) / своя (5B)	~8 GB (5B)	проще всего запустить
Mochi 1 (Genmo)	Mochi 1	Apache 2.0	~20-60 GB	свободная; сильное движение
LTX-Video / LTX-2	LTX-2	Apache 2.0	consumer-DC	real-time; 4K + звук

2 — Два барьера (по порядку, до любого бенчмарка)

БАРЬЕР 1 ЛИЦЕНЗИЯ -> Разрешает ли ПРОДАВАТЬ вывод, на моём масштабе, на моём рынке?

БАРЬЕР 2 VRAM -> ВЛЕЗАЕТ ли в GPU, который я могу позволить? (quantization ужмёт)

Качество — РАЗРЕШЕНИЕ ничьей среди прошедших оба барьера, а не фильтр.

3 — Ловушки лицензий, проверить до старта

- ☐ Apache 2.0 = чистая коммерция: продавать, дообучать, релизить, без лимита пользователей. Подходят CogVideoX-2B, Mochi 1, Wan 2.1/2.2, LTX-2.
- ☐ HunyuanVideo идёт под Tencent Community License — ограничение на коммерцию; прочтите пороги до любого платного продукта.
- ☐ CogVideoX-5B — своя лицензия (НЕ Apache, как у 2B). Wan 2.5+ перешёл на API-first — сверяйте лицензию ИМЕННО той версии, что релизите.

4 — Считаем счёт: self-host vs аренда

- ☐ Стоимость self-hosted клипа = (минуты рендера / 60) x часовая ставка GPU. H100 ~ \$2-2.70/час; A100 sub-\$1/час spot.
- ☐ Полная стоимость = аренда GPU + простой + амортизированные 20-40ч настройки + множитель операций 3-5х. НЕ «модель x электричество».
- ☐ Окупаемость севернее ~5 000 клипов/мес (GPU загружены). Ниже — платный API дешевле И куда меньше работы.

5 — Три неценовые причины для self-hosting

- ☐ Резидентность данных: чувствительный контент (телемед, surveillance, приватные загрузки) не покидает ваши машины.
- ☐ Объём: выше окупаемости при высокой загрузке GPU стоимость своей инфраструктуры за клип падает ниже платы API.
- ☐ Лицензия + fine-tune: только свободная открытая модель даёт дообучить на своих данных и релизить результат в продукт на продажу.