

1 — Конвейер: два правила, пять стадий (воронка)

Правило 1	Сначала дешёвая проверка — каждая стадия решает или эскалирует; дорогие не видят массу.
Правило 2	Каждое решение — в журнал с причиной; нелегальное — отчёт + сохранение, не удаление.
Стадии	Приём → S0 hash → S1 классификаторы → S2 VLM → S3 человек, над governance-этажом.
Ключ	Модерируйте выбранные кадры + аудиотранскрипт, а НЕ целые видео поминутно.

2 — Строить или купить (2026): арендуйте модели, партнёртесь по hash

BUY/HOST	классификаторы (Rekognition/Hive/Azure/NudeNet/CLIP) + VLM (Gemini/Claude/GPT/LLaVA/Qwen-VL)
PARTNER	hash известно-плохого (PhotoDNA, NCMEC/IWF, PDQ+TMK, Lantern) — никогда не строить самим
BUILD	маршрутизация + пороги + хранилище решений + audit + отчётность + гуманная проверка = продукт
ABSTRACT	классификатор и VLM за одним интерфейсом — смена вендора это правка конфига.

3 — Поле инструментов 2026 (сверьте цены перед запуском)

КЛАССИФИКАТОРЫ	Azure ~\$1,50/1k изобр (~\$0,0015) Rekognition ~\$0,10/мин Google Hive open
VLM	frontier Gemini/Claude/GPT для сложного open LLaVA/Qwen-VL для рутинного контекста
HASH	PhotoDNA + NCMEC/IWF + Meta PDQ/TMK + Tech Coalition Lantern/VHIP (784k+ видео)
ВНИМАНИЕ	hash ловит ТОЛЬКО известное; AI-CSAM без прежнего hash → классификаторы + VLM важны.

4 — Модель затрат (за день при 100 000 видео)

НАИВНО	200 000 видеомин/день × \$0,10/мин = \$20 000/день ~ \$7,3 млн/год (и слишком медленно).
ВОРОНКА	hash \$50 + классиф. \$1 800 + VLM(~7%) \$140 + люди(~1%) \$250 = ~\$2 240/день ~ \$0,8 млн/год.
ИТОГ	~9× дешевле И точнее — дорогие стадии видят лишь то, что им действительно нужно.
РЫЧАГ	настраивайте пороги Стадии 1 (долю, что она очищает), а не модель — там счёт.

5 — Порядок сборки: каждый майлстоун защищает пользователей

- 1 Hash-матч + обязательная отчётность (S0 + юр. этаж) — дешевле всего и требуется законом, первым.
- 2 Дешёвые классификаторы, консервативные пороги (S1) — щедро эскалировать, ужесточать с данными.
- 3 Консоль + очередь людей (S3) ДО VLM — с фидами благополучия, встроенными с первого дня.
- 4 VLM-стадия контекста (S2) сжимает очередь. 5 Масштаб + измерение + состязательная петля.

6 — Чек-лист governance (инженерный контекст, не юридическая консультация)

- ☐ Каждый кусок несёт источник + таймкод с приёма, чтобы каждое решение логировалось с причиной.
- ☐ На совпадении S0: БЛОК + ОТЧЁТ (NCMEC CyberTipline) + СОХРАНЕНИЕ — не тихое удаление (18 U.S.C. 2258A).
- ☐ Пороги ПО КЛАССУ ВРЕДА — тяжелейшее детерминированно (hash + человек), а не тюнингуемой оценкой.
- ☐ DSA изложение причин + прозрачность; защита детей (Ст. 28); EU AI Act Ст. 50 метит AI-медиа (авг 2026).
- ☐ Благополучие проверяющих как требование: размытые превью, лимиты за смену, ротация, достойная оплата.
- ☐ Отложенный eval-набор с известными ответами прогоняется непрерывно; новое подтверждённое — органам hash.