Рабочий лист: строить или покупать ИИ-функции

Пройдите этот список для каждой ИИ-функции до подписания договоров. Выбор определяют четыре фактора, рейтинг моделей вторичен. Пересчитайте все варианты в затраты на одного учащегося в месяц.

А. Функция и вариант

- [ ] Мы определили ИИ-функцию и данные учащихся, с которыми она работает
- [ ] Мы рассмотрели все варианты: купить продукт, вызвать облачный API, развернуть открытую модель
- [ ] Новую функцию запускаем на облачном API, чтобы проверить её до оптимизации

Б. Четыре фактора

- [ ] Объём: оценены учащиеся, минуты или сообщения в месяц
- [ ] Точность: определено, насколько функции позволено ошибаться
- [ ] Чувствительность данных: решено, могут ли данные учащихся покидать наш контур
- [ ] Зависимость от поставщика: оценено, насколько болезненно будет потом сменить провайдера

В. Затраты на учащегося в месяц

- [ ] Места, минуты, кредиты и токены пересчитаны в одну единицу: ₽ на учащегося в месяц
- [ ] Ставки взяты из актуального прайса выбранной модели, а не из чужих примеров
- [ ] Сравнили с затратами на свою модель: GPU + время ML/инфраструктурного инженера
- [ ] Свою модель выбираем только за точкой окупаемости при стабильном объёме

Г. Данные, зависимость и трекинг

- [ ] Подписано поручение на обработку (ч. 3 ст. 6 152-ФЗ); запрет обучать модели на наших данных
- [ ] Базы с ПДн находятся в России (ч. 5 ст. 18); подрядчики и сроки хранения проверены
- [ ] Согласие оформлено отдельным документом; за несовершеннолетних его дают законные представители
- [ ] Если ИИ выставляет итоговую оценку, решение принимает человек (ст. 16 152-ФЗ)
- [ ] Для госзаказчика проверен статус ПО в реестре российских программ (п. 21 ПП 1678)
- [ ] Есть тонкий слой абстракции; функция отправляет xAPI-утверждения в нашу аналитику