

Чек-лист: агент-ресёрч

Семь шагов. Пропустил хоть один — агент вернёт красивый отчёт, которому нельзя верить.

1 Одна задача, одно предложение

Напиши, что должно стать правдой после работы агента.

«Каждое утро в файле лежат пять свежих тем с рабочей ссылкой и готовым углом подачи.»

Не помещается в предложение — это две задачи. Раздели.

2 Формат выхода — до промпта, не после

Реши, что именно агент вернёт, и опиши структуру.

Один файл. Пять блоков. В каждом: **что произошло + ссылка · угол подачи · почему важно нашей аудитории.**

Без этого агент отдаст стену текста, и разбирать её будешь руками.

3 Критерии приёмки — бинарные, 3–5 штук

Каждый критерий должен **уметь провалиться**. «Хорошо написано» — не критерий. «Ровно пять тем» — критерий.

- ровно 5 тем;
- у каждой рабочая ссылка на первоисточник, не выдуманная;
- у каждой угол подачи, а не пересказ;
- у каждой строка «почему важно»;
- ничего старше суток.

Критерий, который нельзя провалить, не существует.

4 Модель и потолок расхода

- модель — самая дешёвая из подходящих, дорогая берётся как исключение;
- **одна оценка, без цикла переделок** — иначе агент крутится и жжёт лимит;
- потолок вызовов и потолок расхода — числом, а не «постараюсь».

5

Запреты — в промпт, прямым текстом

- не выдумывать факты, цифры и ссылки;
- не выдавать старое за новое;
- **ничего никуда не публиковать** — только черновик в файл;
- не выходить за названные папки и адреса.

Запрет, которого нет в промпте, будет нарушен.

6

Кто проверит результат — назвать до запуска

Одна команда, которая даёт «прошло» или «не прошло». Например: поиск по файлу, который обязан ничего не найти.

Проверку запускает не агент, а ты. Его слова о собственной проверке не принимаются никогда — ни «ок», ни «найдено 0», ни числа.

Не можешь придумать такую команду — задача поставлена плохо. Переформулируй до запуска.

7

Живой прогон и мутация

- **Живой прогон:** настоящий запуск на настоящих данных. Не мок, не «команда не упала».
- **Мутация:** сломай нарочно — проверка **обязана** упасть. Починил — обязана пройти.

Не упала на сломанном — проверки нет, есть её видимость.

Три ошибки, которые стоили дороже всего

Ошибка	Как выглядит	Чем лечится
Агент выдумал числа	«замерить не смог» — и рядом три правдоподобные цифры	доказательство файлом, а не текстом отчёта
Агент проверил не всё	«найдено 0», а на деле два десятка вхождений	точный список папок в задании, проверка своим скриптом
Проверка не умела падать	зелено на сломанном коде	мутационный прогон обязателен

Короткая версия — на один экран

1. Задача в одно предложение

2. Формат выхода

3. Бинарные критерии
4. Дешёвая модель, одна итерация, потолок

5. Запреты прямым текстом

6. Приёмочная команда — запускаешь ты

7. Живой прогон + мутация