

Практический чек-лист для команды

Чек-лист вынесен в отдельный PDF для сохранения и использования при расследовании сбоев tool calls.

До запуска теста

- Сохранить model ID, endpoint, версию сервера и SDK.
- Зафиксировать параметры генерации, режим стриминга и хэш доступного набора tools.
- Измерить полный размер prompt, включая системные инструкции, схемы, историю и tool results.
- Разделить сценарии, где текст допустим, и сценарии с обязательным вызовом инструмента.

При разборе неудачного хода

- Сопоставить raw output с итоговым полем tool_calls.
- Проверить сборку потокового ответа и работу парсера.
- Проверить имя инструмента по белому списку.
- Разобрать arguments как JSON и применить схему.
- Выполнить бизнес-валидацию прав, объекта операции и допустимого статуса.
- Зафиксировать факт запуска функции и подтверждённый outcome.

При восстановлении

- Для предполагаемой ошибки стриминга проверить контролируемый повтор без стриминга.
- Для невалидных аргументов вернуть модели структурированную ошибку валидации.
- Для отсутствия обязательного tool call использовать ограниченный repair-ретрай или fallback.
- Не повторять автоматически операцию, отклонённую политикой или правами.
- Для действий с побочным эффектом применять idempotency key, журнал операции и проверку фактического результата.