

ПРАКТИЧЕСКОЕ РУКОВОДСТВО · ИИ-АГЕНТЫ

СИСТЕМА ИИ-АГЕНТОВ: НАСТРОЙКА И КОНТРОЛЬ

Распределение задач, прав доступа и ответственности за проверку. На примере моей рабочей системы с Claude Code и Codex.

5

шагов · практические примеры
· проверка результата

Анастасия Тарасова

РЕЗУЛЬТАТ НАСТРОЙКИ

Два агента с разделёнными функциями выполнения и проверки. Окончательные решения принимаете вы.

Чат

Подготавливает ответы и рекомендации.
Перенос данных и выполнение команд
остаются за пользователем.

Агент

Самостоятельно работает с файлами,
пишет код и запускает проверки. Примеры
— **Claude Code** от Anthropic и **Codex**
от OpenAI.

Работа агента с файлами и кодом требует **распределения полномочий и контроля вносимых изменений**.

ОСНОВНЫЕ ТЕРМИНЫ

CLAUDE.md	Файл с правилами проекта для Claude Code. У Codex такой файл — AGENTS.md.
-----------	---

Субагент	Агент для выполнения определённой функции. Например, проверяющий с доступом только для чтения.
----------	--

Навык	Описанная процедура выполнения повторяемой задачи. Например, порядок развёртывания обновления.
-------	--

Хук	Автоматическая проверка с возможностью блокировки операции.
-----	---

Коннектор	Подключение внешнего сервиса. В моей работе — Notion для управления задачами.
-----------	---

ПРИМЕР ИЗ МОЕЙ ПРАКТИКИ

Я внедряла эти механизмы для решения конкретных задач. Хуки добавила после ошибки агента, из-за которой откат стал невозможен. Навык развёртывания — для повторного использования единой процедуры обновления.

РАСПРЕДЕЛИТЕ РОЛИ

ДЕЙСТВИЯ

Зафиксируйте распределение ролей в заметках или таблице до запуска агентов.

ФОРМУЛИРОВКА ЗАДАНИЯ

Для каждой роли укажите задачи, ограничения и ответственного за проверку.

ОЖИДАЕМЫЙ РЕЗУЛЬТАТ

Для каждой задачи назначены исполнитель и отдельный проверяющий.

ПРОВЕРКА РЕЗУЛЬТАТА

Проверку выполняет участник, не отвечающий за выполнение задачи.

ПРИМЕР ИЗ МОЕЙ ПРАКТИКИ

Я — принятие решений и утверждение развёртывания.

Claude в чате — подготовка заданий агентам и проверка предлагаемых команд.

Codex — разработка и тестирование в отдельной копии проекта.

Claude Code — проверка, слияние изменений и развёртывание.

ЗАФИКСИРУЙТЕ ПРАВИЛА ПРОЕКТА

ДЕЙСТВИЯ

Создайте в папке проекта файлы: CLAUDE.md для Claude Code, AGENTS.md для Codex. Черновик CLAUDE.md создаст команда `/init`.

ФОРМУЛИРОВКА ЗАДАНИЯ

Кратко опишите проект, порядок работы и действия, требующие предварительного согласования.

ОЖИДАЕМЫЙ РЕЗУЛЬТАТ

Агент учитывает правила в начале каждой работы.

ПРОВЕРКА РЕЗУЛЬТАТА

«Перечисли действующие инструкции проекта». Текстовые инструкции не гарантируют соблюдение ограничений. Дополнительно проверьте права доступа по шагу 4.

ПРИМЕР ИЗ МОЕЙ ПРАКТИКИ

Слияние изменений и развёртывание — только после моего утверждения.

Если подтверждение не найдено, сообщите об этом. Не добавляйте неподтверждённые сведения.

НАЧНИТЕ С АНАЛИЗА БЕЗ ИЗМЕНЕНИЯ ФАЙЛОВ

ДЕЙСТВИЯ

Откройте агента в папке проекта.

ФОРМУЛИРОВКА ЗАДАНИЯ

«Проведи анализ без изменения файлов. Укажи причину проблемы и подтверждающие сведения. Если причина не установлена, сообщи об этом».

ОЖИДАЕМЫЙ РЕЗУЛЬТАТ

Отчёт с причиной проблемы и её обоснованием.

ПРОВЕРКА РЕЗУЛЬТАТА

Сопоставьте выводы агента с результатами самостоятельной проверки.

ПРИМЕР ИЗ МОЕЙ ПРАКТИКИ

Агент определил ИИ-сервис по названию параметра настройки, хотя в конфигурации находился ключ другого сервиса. После этого я добавила роль проверяющего: каждый вывод должен содержать ссылку на файл и строку.

НАСТРОЙТЕ И ПРОВЕРЬТЕ ПРАВА ДОСТУПА

ДЕЙСТВИЯ

Создайте файл субагента в каталоге `.claude/agents/`. В поле `tools` укажите доступные инструменты.

ФОРМУЛИРОВКА ЗАДАНИЯ

Для проверяющего разрешите только инструменты чтения и поиска.

ОЖИДАЕМЫЙ РЕЗУЛЬТАТ

Субагент отображается в списке при вводе `@`.

ПРОВЕРКА РЕЗУЛЬТАТА

Попросите изменить файл. После запроса на изменение проверьте фактическое состояние файла.

ПРИМЕР ИЗ МОЕЙ ПРАКТИКИ

`name: code-auditor`

`description: Проверяет код. Только чтение.`

`tools: Read, Grep, Glob`

При проверке субагент отказался изменять файл, однако право записи оставалось доступным из-за дополнительного параметра конфигурации.

РАЗДЕЛИТЕ ВЫПОЛНЕНИЕ И ПРОВЕРКУ

ДЕЙСТВИЯ

Добавьте правило в файлы обоих агентов.

ФОРМУЛИРОВКА ЗАДАНИЯ

«Codex разрабатывает код в отдельной копии проекта. Claude Code проверяет результат и выполняет развёртывание. Обновление сайта — только после моего утверждения».

ОЖИДАЕМЫЙ РЕЗУЛЬТАТ

Каждый результат проходит отдельную проверку.

ПРОВЕРКА РЕЗУЛЬТАТА

В отчёте указано, кто выполнил задачу и кто проверил результат.

ПРИМЕР ИЗ МОЕЙ ПРАКТИКИ

При первоначальной проверке Claude Code не сообщил об ошибке. Codex обнаружил, что строковое значение "false" интерпретировалось как подтверждение согласия. В результате согласие могло считаться полученным без действия пользователя.

Уверенность ответа **не подтверждает его корректность.**

ПРОВЕРЯЙТЕ ФАКТИЧЕСКОЕ ИСПОЛЬЗОВАНИЕ НАСТРОЕК

ФОРМУЛИРОВКА ЗАДАНИЯ

«Перечисли компоненты, которые использовались за последнюю неделю. Отдельно укажи неиспользуемые».

ДЕЙСТВИЯ

Исключите неиспользуемые компоненты. Если необходимый механизм не запускается, установите причину.

ПРИМЕР ИЗ МОЕЙ ПРАКТИКИ

- Субагенты были настроены, однако основной агент не использовал их и выполнял задачи самостоятельно.
- Запланированный еженедельный обзор был выполнен один раз.
- Я подготовила конфигурацию команды из 11 агентов, но не запускала её из-за ограничений сервера на выполнение длительных задач.

Добавляйте субагентов для решения конкретных задач, требующих отдельной роли.

ЧЕК-ЛИСТ

- ☐ Для каждой задачи назначены исполнитель и отдельный проверяющий.
 - ☐ Правила проекта зафиксированы в файлах агентов.
 - ☐ Первая задача выполнена без изменения файлов.
 - ☐ Ограничения прав проверены на практике.
 - ☐ Развёртывание выполняется только после вашего утверждения.
 - ☐ Использование настроенных компонентов проверяется еженедельно.
-

ПЕРВЫЙ ШАГ

Определите роли по шаблону из шага 1. Если агент один, выполняйте проверку самостоятельно.

В этом гайде — основы для самостоятельного старта.

Если хотите глубже разобраться в работе с ИИ-агентами и настроить их под задачи своего проекта, свяжитесь со мной — обсудим подходящий формат работы.

Мой сайт и контакты → anastasia-tarasova.ru