

## **Hermes Agent: автономный ИИ-ассистент на вашем сервере**

---

# **Hermes Agent: Ваш автономный AI-ассистент 24/7**

### Ключевые тезисы:

- Hermes — автономный AI-ассистент с открытым исходным кодом, работающий на вашем сервере.
- Он не просто отвечает на вопросы, а **учится, запоминает и создаёт навыки (skills)** для автоматизации задач.
- Общение происходит через мессенджеры (Telegram, Discord), что делает его похожим на личного помощника.
- Установка возможна **без единой команды в терминале** с помощью AI (OpenAI Codex).
- Стоимость содержания составляет примерно **10-20 долларов в месяц** (сервер + модель AI).

## **Отличия от ChatGPT/Claude и OpenClaw**

- **ChatGPT/Claude:** Инструменты для диалога и решения задач «здесь и сейчас». Требуют вашего присутствия.
- **Hermes:** **Автономный агент**, работающий постоянно. Он учится на ваших задачах, создаёт и оптимизирует навыки для их повторного выполнения.
- **Claude Code:** Инструмент для разработки («кабина пилота»), с которым вы работаете напрямую.
- **OpenClaw:** Изначально оркестратор, теперь также агент. Имеет проблемы со стабильностью и безопасностью, тратит значительно больше токенов на старте (30-60к против 6-8к у Hermes).

# Пять столпов (слоёв) Hermes

Hermes построен на пяти ключевых компонентах, которые делают его уникальным.

## Память (6 слоёв)

1. **Текущий контекст:** Рабочий стол агента (сообщения, результаты команд).
2. **Компрессия контекста:** Автоматическое сжатие длинных диалогов в конспект с двумя уровнями защиты (на 50% и 85% заполнения).
3. **Постоянная память:** Два небольших файла (`user.md` — портрет пользователя, `memory.md` — факты о среде). Агент использует её молча, учитывая ваши предпочтения.
4. **Архив прошлых сессий:** Все диалоги сохраняются в SQLite базе для последующего поиска.
5. **Навыки (Skills):** Про это отдельно ниже.
6. **Внешняя память:** На выбор — платный **Honcho** (моделирует пользователя) или бесплатный локальный **Holographic Memory** (ловит противоречия в фактах).


## Навыки (Skills)

- *Процедурная память агента* — не что он знает, а **как правильно сделать**.
- Создаются автоматически после выполнения сложных задач (5+ вызовов инструментов).
- **87 навыков** идут из коробки (работа с файлами, API, браузером, генерация изображений/видео и т.д.).
- Навыки **самообновляются** при изменении инструментов или процессов.
- **Hermes Curator** автоматически архивирует неиспользуемые навыки (через 30/90 дней), чтобы они не засоряли контекст.

## Душа (Soul)

- Файл `soul.md` определяет личность агента: тон общения, реакцию на ошибки.
- Можно выбрать готовую или написать свою. Личность эволюционирует на основе вашей обратной связи.

## 17 **Расписание (Cron)**

- Задачи на естественном языке (например, «каждое утро в 10:00 присылай дайджест новостей»).
- Запускаются в изолированных сессиях. Есть полезные флаги ( `content_from` , `work_dir` , `no_agent` ).
- Можно назначать разные модели для разных задач (мощная — для аналитики, дешёвая — для рутины).
-  **Ограничение:** задачи по расписанию не могут рекурсивно создавать новые.

## **Самообучение**

- Агент улучшается с каждым разговором: фиксирует ошибки, определяет успешный путь и обновляет память и навыки.
- В следующий раз ту же задачу выполнит быстрее и качественнее.

## **Практические сценарии применения**

- **Новостной дайджест:** Автоматический сбор и фильтрация новостей из заданных источников по расписанию.
- **Мониторинг сервера:** Уведомления о перегрузке CPU, заполнении диска, падении сайта с предложениями по решению.
- **Маркетинговый анализ:** Автоматический сбор данных о конкурентах и формирование отчёта с последующим созданием навыка.
- **Личный ассистент 24/7:** Напоминания, помощь в написании текстов, подбор информации под ваши цели.
- **Автоматизация рутины:** Например, оптимизация сайта (сжатие медиафайлов, ленивая загрузка) по одной команде.

## **Установка и настройка (способ без терминала)**

Цель: Установить Hermes на VPS-сервер одной командой через **OpenAI Codex**.

Что нужно подготовить:

1. **VPS-сервер** (например, Timeweb: Ubuntu, 2 ядра, 4 ГБ ОЗУ, 20 ГБ SSD).

2. **Telegram-бот:** Создать через `@BotFather`, сохранить его токен.
3. **Ваш User ID:** Узнать через бота `@userinfobot`.
4. **API-ключ модели AI** (рекомендуется OpenRouter или Severel для оплаты в рублях). Обязательно установить лимит расходов.
5. **Файл `credentials.txt`** в папке проекта с данными: IP сервера, пароль root, токен бота, User ID, API-ключ.

### Процесс:

1. Установите **OpenAI Codex** на компьютер.
2. В Codex откройте папку с проектом (где лежит `credentials.txt`).
3. Вставьте специальный промпт (предоставляется автором), который инструктирует Codex прочитать файл и выполнить всю установку.
4. Нажмите Enter. Codex самостоятельно подключится к серверу, установит Hermes, все зависимости, запустит мастера настройки и подключит Telegram-бота.

## Основные команды и советы

- `mes` / `mes-tui` — запуск чата в терминале (tui — красивее).
- `mes-setup` — мастер настройки (провайдер, модель, мессенджер).
- `mes-update` — обновление Hermes (рекомендуется раз в неделю).
- `mes config set <key> <value>` — безопасное хранение секретов (никогда не вводите ключи в чат!).
- `/model` — смена модели прямо в чате.
- `/new` — новый разговор (контекст очищается, память сохраняется).

### Важно о моделях:

- **Не экономьте на основной модели** для агентских задач. Дешёвая модель будет ошибаться и заикливаться.
- **Рекомендация:** `DeepSeek-V4-Pro` для сложных задач + `DeepSeek-V4-Flash` для рутины и кронов.
- Кэшированные запросы у DeepSeek — **бесплатные**, что даёт огромную экономию при круглосуточной работе.



## Продвинутые возможности

- **Мультиагентная система:** Запуск нескольких независимых агентов на одном сервере (для клиентов, мониторинга, аналитики).
- **Субагенты:** Параллельное выполнение частей одной задачи (например, анализ 5 конкурентов одновременно).
- **Канбан-доска:** Визуальное управление мультиагентным workflow (исследователь, писатель, рецензент).
- **Интеграции (CLA):** Подключение тысяч внешних сервисов (Gmail, Notion, YouTube) одной командой через Compose.io.
- **Голосовые сообщения:** Транскрибация и озвучка работают из коробки в Telegram.
- **Hyperframes:** Генерация нативного HTML-видео через AI.
- **Официальный дашборд:** Просмотр сессий, аналитики расходов токенов, расписаний.



## Критически важные правила безопасности

1. **Бэкапы на GitHub** — это обязанность. Настройте ежедневный автоматический бэкап в 3:00 ночи.
2. **Никогда не используйте подписку Claude Pro/Max через API** — рискуете получить бан аккаунта.
3. **Anthropic (Claude) сканирует Git-коммиты и системные промты** на слова «mes» или «орен». Это может привести к блокировке или неожиданным списаниям.
4. **Hermes может врать о выполненных задачах.** Всегда проверяйте результат.
5. **Секреты (API-ключи, пароли) храните только через `mes config set`**. Никогда не вводите их прямо в чат.
6. **Принцип минимальных привилегий:** Создавайте для агентов отдельные аккаунты, почты и API-ключи с ограниченными правами.
7. **Еженедельный аудит безопасности:** Настройте крон-задачу для проверки здоровья сервера и подозрительной активности.

## Итоговая стоимость и экономия

- VPS-сервер: ~\$5-10 в месяц.
- AI-модель (DeepSeek-V4-Pro + Flash): ~\$6 в месяц.
- Итого: ~\$10-20/мес за автономного ассистента 24/7.
- Для сравнения: Claude Sonnet при аналогичной нагрузке обойдётся ~\$100/мес.

Способы экономии:

1