

ИИ-сборщик данных с сайтов

Вытаскивает данные с сайтов без ручного копирования: тексты страниц, цены, таблицы, каталоги. Может обойти целый сайт и собрать всё в структурированную таблицу.

6 800+ звёзд у MCP-версии

Нужен ключ Firecrawl

500 запросов бесплатно

1 ЧТО УМЕЕТ

- Контент любой страницы в чистом виде
- Обход целого сайта или раздела
- Каталог или база из веб-страниц в Excel
- Сбор цен и позиций с сайтов конкурентов
- Структурированные данные: название, цена, описание
- Скриншоты страниц

Под капотом: скилл на базе Firecrawl (у официального MCP-сервера Firecrawl 6 800+ звёзд на GitHub). Firecrawl обходит защиты и отдаёт чистый контент.

2 УСТАНОВКА, ОДИН РАЗ

- Зарегистрируйся на firecrawl.dev и возьми API-ключ: бесплатный тариф даёт 500 запросов.
- Скачай архив `firecrawl-skill.zip` из этой папки, распакуй и положи папку в `~/.claude/skills/`.
- Поставь библиотеку: `pip install firecrawl python-dotenv`
- Добавь ключ в файл `~/.env` строкой `FIRECRAWL_API_KEY=твой-ключ`: скилл читает его оттуда. Готово.

Совет. Бесплатного тарифа (500 запросов) хватает для регулярного сбора небольших объёмов. Следи за остатком в личном кабинете Firecrawl, чтобы не упереться в лимит посреди задачи.

3 КАК ПОЛЬЗОВАТЬСЯ

Специальных команд нет. Пишешь задачу обычными словами в Claude Code, сотрудник включается сам, когда задача по его части. Примеры фраз:

“Собери все тарифы с этих 5 сайтов в таблицу

“Вытащи контент этой страницы в markdown

“Обойди раздел блога [ссылка] и собери заголовки со ссылками

“Собери каталог курсов с этого сайта: название, цена, формат

4 КАКОЙ РЕЗУЛЬТАТ ПОЛУЧИШЬ

Чистые структурированные данные из веба: таблица, список или текст без рекламы и мусора. Часы ручного копирования превращаются в один запрос.

ПРОМПТ

Собери с этих пяти сайтов школ по нейросетям (ссылки прилагаю) данные в одну таблицу: название курса, цена, длительность, формат, есть ли рассрочка. На выходе Excel с колонкой «источник» и датой сбора.

6 ЧАСТЫЕ ВОПРОСЫ

Закончились бесплатные запросы? Сотрудник предупредит и остановится до трат. Дальше или платный тариф Firecrawl, или ждать обновления лимита.

Сайт не отдаёт данные? Некоторые сайты закрыты жёстко. Скажи об ошибке, сотрудник попробует другой способ или честно скажет, что сайт недоступен.

Это легально? Открытые страницы собирать можно. Персональные данные и закрытые разделы: нет. Уважай правила сайтов.