

Fair use и приватная автоматизация

Версия: 1.0

Дата изменения: 24 августа 2026 года

Правила для личных агентов, дайджестов и автоматизированного доступа.

Допустимые сценарии

Допустимы личный поиск, локальный агент, приватный дайджест, внутренний некоммерческий прототип и исследование качества данных при сохранении ссылки на первоисточник.

Техническая добросовестность

Используйте API вместо scraping сайта, кэшируйте ответы минимум 15 минут, применяйте backoff, соблюдайте Retry-After и не распараллеливайте массовые запросы. Не обходите robots.txt, CAPTCHA, авторизацию и ограничения источников.

Контроль человека

Автоматизация не должна самостоятельно отправлять отклики, сообщения работодателям, соглашаться с условиями или раскрывать персональные данные. Такие действия выполняются только после явного подтверждения человека для каждого адресата.

Публичное и коммерческое повторное использование

Массовая перепубликация, создание конкурирующего публичного агрегатора, продажа набора данных или высокая нагрузка требуют отдельного разрешения и самостоятельной проверки лицензий источников.

Статус проекта

Публичный сайт — пример работы open-source библиотеки, а не коммерческая услуга scraping-as-a-service. Проект не продаёт доступ к данным, не принимает заказы на обход защиты и не обещает доступность конкретного источника.

Private automation означает запуск пользователем для личного поиска, локального исследования или внутреннего некоммерческого прототипа. Пользователь сам предоставляет инфраструктуру, конфигурацию и законные права доступа.

Уважение технических и правовых ограничений

Не используйте проект для обхода CAPTCHA, авторизации, paywall, блокировки аккаунта или IP, если владелец ресурса не дал разрешение либо такое действие не допускается

законом. Остановите сбор при запрете, жалобе, ошибках 401/403/429, требовании CAPTCHA или очевидной деградации ресурса; используйте официальный API, минимальную частоту и кэш.

Минимизация и хранение

Собирайте только поля, необходимые для личного поиска вакансий; не создавайте досье на людей, не обогащайте контакты и не храните исходные страницы дольше необходимого. Не публикуйте полный корпус или существенную часть базы источника. Сохраняйте атрибуцию и прямую ссылку.