

10. Не передавайте код, ключи и личные данные «AI-посредникам» неизвестного происхождения, особенно когда AI автоматически выполняет команды

Выгодность исходной версии: обычная

Доказательность В · исходная версия

для свободы

немного денег

мимоходом

немного силы воли

Частично Китай

Аналог в России не найден

Искали в официальных российских источниках (проверено 08.10.2026), прямого аналога не нашли. Это не значит, что его нет. Китайские условия на Россию не переносите.

Оценки выше относятся к исходной версии; отдельная оценка российского варианта не проводилась.

Простыми словами

Китай — исходная версия. Посредник стоит между вами и поставщиком модели. Он видит и может менять все, что вы отправляете, и все ответы, которые получаете. В одном исследовании проверили более 400 посредников: 9 вставляли в ответы вредоносный код, а 17 использовали ключи облачных сервисов, которые исследователи специально туда поместили. Если AI автоматически выполняет команды, одной измененной команды достаточно, чтобы компьютер оказался под чужим контролем.

Затраты

Китай — исходная версия. Прямое подключение к официальному сервису обычно дороже, чем через посредника, и пополнять баланс не так удобно. Сложность в том, что посредники дешевы, а через одну точку входа можно обращаться к моделям нескольких компаний

Выгода

Китай — исходная версия. Министерство государственной безопасности (国家安全部) в июне 2026 года опубликовало предупреждение о рисках. Некоторые посредники сохраняют присланные пользователями данные на своих серверах.

Некоторые без разрешения удерживают их и перепродают другим поставщикам больших моделей для обучения. Некоторые скрывают в своих сервисах бэкдоры, внедряют вредоносный код на устройства пользователей, крадут ключи учетных записей и учетные данные для доступа к облачным сервисам, даже устанавливают программы удаленного управления. В предупреждении рекомендуется подключаться напрямую к официальным сервисам или пользоваться платформами с надлежащей авторизацией, а не платформами неизвестного происхождения, без разрешения на ведение деятельности и без гарантий безопасности. Перед использованием нужно обезличить личные данные и материалы проектов, надежно хранить ключи и регулярно менять их. При неожиданных списаниях, блокировке учетной записи без причины или аномалиях с данными нужно немедленно прекратить использование, сменить ключи, проверить устройство на вирусы и удалить их, сохранить доказательства. В 2026 году исследователи из Калифорнийского университета в Санта-Барбаре и других организаций провели измерительное исследование. Они купили доступ к 28 платным посредникам на китайской торговой площадке (淘宝), на китайской площадке продажи поддержанных товаров (闲鱼) и в зарубежных интернет-магазинах, а также собрали в открытых сообществах 400 бесплатных посредников. Оказалось, что 1 платный и 8 бесплатных посредников внедряли вредоносный код в возвращаемые вызовы инструментов. Вызов инструмента — это команда, которую AI дает вашему компьютеру, например для установки программного пакета. Еще 2 выбирали момент для атаки. В статье приведен пример: атаковали только сеансы с полностью автоматическим выполнением команд. 17 использовали ключи облачных сервисов Amazon, которые исследователи специально туда поместили. 1 вывел деньги из Ethereum-кошелька исследователей. В кошельке было совсем мало денег, потери составили менее 50 долларов США. Показанная в статье подмена очень незаметна: в команде установки requests заменяют на похожее по написанию requeusts, и с первого взгляда этого не видно (по всему Китаю)

Примечания

Китай — исходная версия. Уровень В присвоен потому, что исследование только одно, и это препринт, то есть версия, выложенная до официальной публикации. Большинство проверенных посредников были бесплатными, платных было только 28, поэтому нельзя вычислить, «какая доля посредников имеет проблемы». Некоторые СМИ написали, что в том случае украли 500000 долларов США, но в самой статье указано менее 50 долларов США. Масштаб выгоды оценен как средний, исходя из последствий: размер потерь от кражи ключей и кода может

сильно различаться, а в исследовании нет суммы, которую можно было бы взять за основу. Если вы уже пользовались посредником неизвестного происхождения, отзовите использованные ключи у поставщика модели и создайте новые. Заодно замените ключи облачных сервисов и репозиторий кода. При работе через посредника не позволяйте AI выполнять команды полностью автоматически: сначала просмотрите команду и только потом разрешайте выполнение. Сторона, управляющая посредническим сервисом, рискует столкнуться с уголовной ответственностью в Китае, см. раздел 11, пункт 19 (AI-посредники). Кому это выгодно — вам самим.

Источники

Китай — исходная версия — источники

1. 国家安全部 (2026). 「AI中转站」, 风险要防范.
www.szzg.gov.cn/2026/xwzx/szcx/202606/t20260608_5331487.htm (数字中国建设峰会官网转载国家安全部微信公众号)
2. Liu H, Shou C, Wen H, Chen Y, Fang RJ, Feng Y. (2026). Your Agent Is Mine: Measuring Malicious Intermediary Attacks on the LLM Supply Chain. arXiv:2604.08407.
arxiv.org/abs/2604.08407