

Не доверяй нейросети слишком много

Личные данные, ошибки ИИ
и цифровая осторожность



Юрий Тер

Юрий Тер

Не доверяй нейросети слишком много - личные данные, ошибки ИИ и цифровая осторожность

<https://litres.ru/74484703>

SelfPub; 2026

Аннотация

Нейросети помогают решать повседневные и рабочие задачи, но вместе с удобством приносят и новые риски. Пользователи всё чаще отправляют в ИИ личные данные, рабочие документы, переписки, клиентскую информацию и другие чувствительные материалы, а затем без достаточной проверки доверяют полученным ответам.

«Не доверяй нейросети слишком много - личные данные, ошибки ИИ и цифровая осторожность» — практическая книга о безопасной работе с ИИ-сервисами. В ней объясняется, какие данные не стоит бездумно передавать нейросетям, чем опасны ошибки и выдумки ИИ, как проверять ответы модели и какие привычки помогают избежать утечек, дезинформации и неверных решений.

В книге собраны реальные документированные случаи, типовые ошибки пользователей и понятные рекомендации по

цифровой осторожности — для повседневной жизни, учёбы и работы.

Содержание

К читателю	5
Как пользоваться книгой	7
ЧАСТЬ I. ДО КНОПКИ «ОТПРАВИТЬ»	8
Глава 1. Почему чат кажется личным пространством	9
Глава 2. Что происходит после кнопки «Отправить»	12
Глава 3. «Только на минуту»: как удобство расширяет границы доверия	15
Конец ознакомительного фрагмента.	16

Юрий Тер

Не доверяй нейросети слишком много - личные данные, ошибки ИИ и цифровая осторожность

К читателю

Нейросеть редко получает слишком много данных потому, что пользователь решил нарушить правило. Чаще причина буднична: так быстрее. Целый договор вставить проще, чем вырезать один пункт; полный лог — быстрее, чем сделать минимальный пример; всю переписку — проще, чем пересказать ситуацию. Удобство постепенно стирает границу между тем, что мы готовы сообщить внешнему сервису, и тем, что должно остаться внутри устройства, семьи или организации.

Вторая ошибка возникает после ответа. Современная модель умеет писать уверенно, ссылаться на документы, предлагать последовательный план и — всё чаще — выполнять действия через подключённые инструменты. Поэтому старо-

го совета «не верь всему, что написала нейросеть» уже недостаточно. Нужно понимать, что именно можно передавать, что требует независимого источника, какие действия должны оставаться за человеком и какие запреты обязаны быть техническими, а не текстовыми.

Книга построена вокруг одного маршрута: что происходит до отправки данных; как распознать чувствительный контекст; почему профессионалы тоже попадают в ловушку уверенного ответа; что показывают реальные судебные, корпоративные и регуляторные истории; как устроить безопасную работу с агентами; и что делать, если ошибка уже произошла.

Реальные случаи отделены от учебных сценариев. У каждого документированного кейса источник стоит непосредственно после описания. Учебные сценарии помечены как составные: они созданы для тренировки и не выдаются за истории конкретных людей.

Важно. Книга не заменяет юридическую, медицинскую, финансовую или иную профессиональную консультацию. Условия сервисов и законодательство меняются; перед высокорисковым действием проверяйте актуальные первоисточники и правила вашей организации.

Как пользоваться книгой

Если вас прежде всего волнует приватность — начните с частей I–III. Если вы используете ИИ для исследований, права, бизнеса или публикаций — особенно важны части IV–VII. Если подключаете почту, календарь, файлы, код или базы — не пропускайте часть VIII. Если инцидент уже случился, переходите сразу к части IX.

Не нужно запоминать все правила. Цель книги — собрать несколько устойчивых привычек: минимизировать входные данные, отделять черновик от факта и действия, проверять по независимому источнику, ограничивать полномочия и заранее знать, как остановить или отменить процесс.

ЧАСТЬ I. ДО КНОПКИ «ОТПРАВИТЬ»

Первая задача книги — вернуть читателю ощущение реального действия. Сообщение в чат — это не запись в личный блокнот, а передача данных в конкретную систему с конкретными правилами.

Глава 1. Почему чат кажется личным пространством

Чат с нейросетью визуально похож на мессенджер или заметку. Вы пишете фразу, получаете ответ и продолжаете разговор. В интерфейсе не видно серверов, журналов, администраторов, резервного хранения и сторонних компонентов. Психологически это создаёт ощущение закрытого пространства: будто текст существует только между вами и «собеседником».

Но нейросеть — не собеседник в человеческом смысле и не локальный блокнот. Это сервис. У него есть поставщик, инфраструктура, правила хранения и удаления, механизмы безопасности, учётная запись и иногда интеграции с другими системами. Даже если сервис обещает высокую приватность, это всё равно отдельная среда, а не продолжение вашего жёсткого диска.

Особенно легко забыть об этом, когда модель уже помогла десятки раз. После нескольких удачных ответов пользователь начинает расширять контекст: сначала даёт обезличенный абзац, затем целый документ, потом переписку, таблицу и архив. Каждое отдельное действие выглядит небольшим, а через месяц чат превращается в копию рабочего стола.

Полезная привычка — мысленно заменять кнопку «От-

править» на формулировку «Передать этому сервису». Она возвращает реальное значение действия. Если фраза «я передаю внешнему сервису этот паспорт, договор или список клиентов» звучит неуютно, файл, вероятно, нужно сократить, обезличить или обработать в другой среде.

Словом «приватный» часто называют несколько разных вещей. Первая — разговор не публикуется в открытом интернете. Вторая — данные не используются для обучения. Третья — они хранятся ограниченное время. Четвёртая — доступ к ним ограничен внутри организации. Пятая — пользователь может удалить историю. Это пять разных свойств.

Например, чат может быть невидим другим пользователям, но сохраняться у поставщика. Данные могут не идти на обучение модели, но храниться в журнале безопасности. Корпоративный администратор может иметь доступ к метаданным, хотя сам поставщик ограничивает внутренний просмотр. Удаление из интерфейса может не означать мгновенного исчезновения из резервного хранения.

Поэтому вопрос «чат приватный?» почти бесполезен. Полезные вопросы конкретны: кто видит? сколько хранится? для чего используется? можно ли отключить историю? есть ли отдельные условия для корпоративных клиентов? как устроено удаление? где юридически обрабатываются данные?

Если ответы важны для вашей задачи, их нужно получать из документации и договора, а не из рекламной формули-

ровки «ваша приватность для нас важна».

Учебный пример. Менеджер хочет сократить жалобу клиента и машинально вставляет весь тред: фамилию, телефон, номер заказа, домашний адрес и подпись. Для задачи были нужны только три абзаца с сутью претензии. Риск появился не из-за сложной технологии, а из-за ощущения, что чат — продолжение личной переписки.

Ключевой вывод. Интерфейс чата уменьшает ощущение риска, но не превращает внешний сервис в личный блокнот.

Глава 2. Что происходит после кнопки «Отправить»

Упрощённая схема выглядит так: устройство формирует запрос, он передаётся по сети, принимается сервисом, проходит через инфраструктуру обработки и возвращается как ответ. В зависимости от продукта запрос может дополнительно попасть в журналы безопасности, историю диалога, систему аналитики, резервное хранение или инструменты администрирования.

Это не означает, что любой сотрудник поставщика читает ваши чаты и что каждый запрос используется для обучения. Такие обобщения неверны. У разных сервисов и тарифов разные условия. В одних режимах история может сохраняться, в других — нет; в одних данных не используют для обучения, в других пользователь может управлять соответствующей настройкой; корпоративные договоры могут давать дополнительные гарантии.

Главная практическая проблема — пользователь редко знает, в каком именно режиме работает. Он помнит название бренда, но не тип аккаунта, не договор и не настройки. Поэтому важный документ нельзя отправлять на основании воспоминания «кажется, они ничего не хранят». Нужно открыть актуальные условия конкретного продукта.

Даже при хороших правилах остаётся обычный технологический риск. В 2023 году OpenAI публично описала программную ошибку, из-за которой часть пользователей могла видеть заголовки чужих чатов и ограниченные платёжные данные других активных подписчиков. Это был не «сговор модели», а обычный дефект инфраструктуры. Он хорошо показывает, почему чувствительная информация требует минимизации даже у крупного поставщика.

Массовый потребительский сервис оптимизирован под скорость подключения и удобство. Корпоративный продукт обычно добавляет администрирование, разграничение доступа, журналы, единый вход, договорные условия и отдельные настройки обработки. Локальная модель работает в инфраструктуре пользователя или организации и даёт больше контроля над данными, но требует собственных специалистов и собственной защиты.

Ни один вариант не является автоматически безопасным. Корпоративный аккаунт можно настроить небрежно. Локальный сервер можно оставить открытым в интернет. Публичный сервис может иметь строгий режим временных чатов. Поэтому решение должно начинаться не с категории «облако плохо, локально хорошо», а с типа данных и требований к ним.

Для публичного текста достаточно обычного чата. Для внутренних документов компании нужен разрешённый рабочий продукт. Для особо чувствительных данных может по-

надобиться изолированная среда или вообще отказ от передачи содержимого модели. Иногда проще дать нейросети структуру документа и синтетический пример, чем строить дорогую локальную систему ради одной операции.

Режим обработки определяется продуктом, тарифом, настройками и договором, а не одним названием бренда.

Глава 3. «Только на минуту»: как удобство расширяет границы доверия

Многие нарушения начинаются не с осознанного решения передать данные, а с ощущения временности. «Я только проверю формулировку». «Нужно быстро найти ошибку». «Потом удалю чат». В этот момент пользователь оценивает длительность своей работы, а не сам факт передачи.

Конец ознакомительного фрагмента.

Текст предоставлен ООО «Литрес».

Прочитайте эту книгу целиком, [купив полную легальную версию](#) на Литрес.

Безопасно оплатить книгу можно банковской картой Visa, MasterCard, Maestro, со счета мобильного телефона, с платежного терминала, в салоне МТС или Связной, через PayPal, WebMoney, Яндекс.Деньги, QIWI Кошелек, бонусными картами или другим удобным Вам способом.