

Использование данных вендорами и провайдерами LLM

Анализ безопасности и конфиденциальности

OpenAI ChatGPT / API

ChatGPT (веб версия)

- Хранение данных: ⚠ Бессрочно (судебное решение)
- Обучение на данных: ⚠ По умолчанию
- Корпоративные гарантии: ☒ ZDR доступен

OpenAI API

- Хранение данных: ⚠ Бессрочно (судебное решение)
- Обучение на данных: ☒ Нет
- Корпоративные гарантии: ☒ ZDR доступен

Enterprise/Team: Не используют данные для обучения по умолчанию, доступен ZDR

Инциденты 2025: Судебное решение о нарушении авторских прав при обучении GPT-4

Anthropic Claude / API

Claude (веб версия)

- Хранение данных: 30 дней после удаления
- Обучение на данных: ☒ Нет
- Корпоративные гарантии: ☒ ZDR доступен

Claude API

- Хранение данных: ☒ Не хранит
- Обучение на данных: ☒ Нет
- Корпоративные гарантии: ☒ ZDR по умолчанию

Лидер по безопасности: ZDR по умолчанию в API, не обучается на пользовательских данных

Инциденты 2025: Утечка данных в 2024 году (исправлена)

Google Gemini / Workspace

Gemini

- Хранение данных: 18 месяцев
- Обучение на данных: ⚠ По умолчанию
- Корпоративные гарантии: ☒ Workspace защищен

Google Workspace

- Хранение данных: Можно выбрать регион хранения
- Обучение на данных: ☒ Нет
- Корпоративные гарантии: ☒ Полная защита

Google Workspace: Защита данных от обучения и просмотра человеком, интеграция с Google экосистемой

Инциденты 2025: Не зафиксировано

GitHub Copilot

Individual план

- Хранение данных: 28 дней
- Обучение на данных: ⚠ По умолчанию
- Корпоративные гарантии: ✗ Нет

Business план

- Хранение данных: Можно выбрать регион хранения
- Обучение на данных: ☒ Нет
- Корпоративные гарантии: ☒ Business защищен

Business план: Не используют данные для обучения, корпоративные гарантии безопасности

Инциденты 2025: Утечки репозитория, судебные решения по авторским правам

GigaChat / API

GigaChat (веб)

- Хранение данных: ✗ Данные хранятся
- Обучение на данных: ⚠ Данные могут использоваться для обучения
- Корпоративные гарантии: ⚠ Размытые формулировки

GigaChat API

- Хранение данных: ⚠ Косвенно из документации
- Обучение на данных: ⚠ Прямо не указано
- Корпоративные гарантии: ⚠ Пункт идентичен с соглашением для физических лиц

Из соглашения: Контент используется исключительно для целей предоставления Сервиса

Инциденты 2025: Не зафиксировано

YandexGPT / API

YandexGPT (веб)

- Хранение данных: Весь период
- Обучение на данных: **✗** Да
- Корпоративные гарантии: **⚠** В рамках соглашений

YandexGPT API

- Хранение данных: **⚠** Опционально
- Обучение на данных: **⚠** Только при логировании
- Корпоративные гарантии: **☑** Отключение логирования

API с выключенным логированием: Stateless inference, не сохраняет данные запросов

Инциденты 2025: Утечка кода в 2023 году

DeepSeek

Политики данных

- Хранение данных: Не указано
- Обучение на данных: ✗ Да
- Корпоративные гарантии: ✗ Нет гарантий
- Передача третьим лицам: Возможна

Особенности

- Открытые веса: Модели доступны для скачивания
- Локальный запуск: Возможность разместить в своем контуре
- API: Стандартные политики

Преимущество: Открытые веса моделей позволяют локальное развертывание без передачи данных

Инциденты 2025: Утечка 1M записей из ClickHouse

Агрегаторы: OpenRouter, Cursor, Windsurf

OpenRouter

- Хранение: Минимальное
- Обучение: ☒ Нет
- Передача: ☐ Провайдерам

Преимущество: Доступ к множеству моделей

Cursor

- Хранение: ☒ Privacy Mode
- Обучение: ☒ Privacy Mode
- Передача: ☒ Privacy Mode

Преимущество: Режим приватности блокирует сбор и использование данных

Windsurf

- Хранение: ☒ ZDR по умолчанию
- Обучение: ☒ Нет
- Передача: ☒ Отдельные ZDR с провайдерами

Преимущество: Zero Data Retention по умолчанию

Инциденты 2025: Зафиксированные инциденты не обнаружены

Выводы для личного использования



ТОП-3 для максимальной конфиденциальности:



1. Anthropic Claude API с ZDR

Лидер по безопасности с Zero Data Retention по умолчанию в API



2. Cursor Privacy Mode

Режим приватности с отдельной инфраструктурой для приватных запросов и отключением обучения



3. Windsurf Enterprise Hybrid

Данные не используются, отдельный ZDR с вендорами, доступны все AI возможности

Рекомендации по использованию:

Используйте API вместо веб-чатов для максимальной защиты

Проверяйте настройки конфиденциальности перед использованием

Отключайте историю чатов, где это возможно

Избегайте передачи личных данных в промптах

Выводы для корпоративного использования

Для российских компаний:

YandexGPT API — с выключенным логированием

On-premise во внутреннем контуре — полный контроль

GigaChat корпоративный — с уточнением формулировок

Ключевые критерии выбора для корпоративного использования:

Юридические гарантии неиспользования данных для обучения

Соответствие требованиям российского законодательства

Для компаний вне российской юрисдикции:

OpenAI Enterprise/API с ZDR — максимальная защита данных с юридическими гарантиями

Anthropic Claude API — ZDR по умолчанию, заявлена высокая безопасность

Google Workspace — полная защита корпоративных данных



habrahabr.ru/company/pt



pt@ptsecurity.com



ptsecurity.com

Спасибо